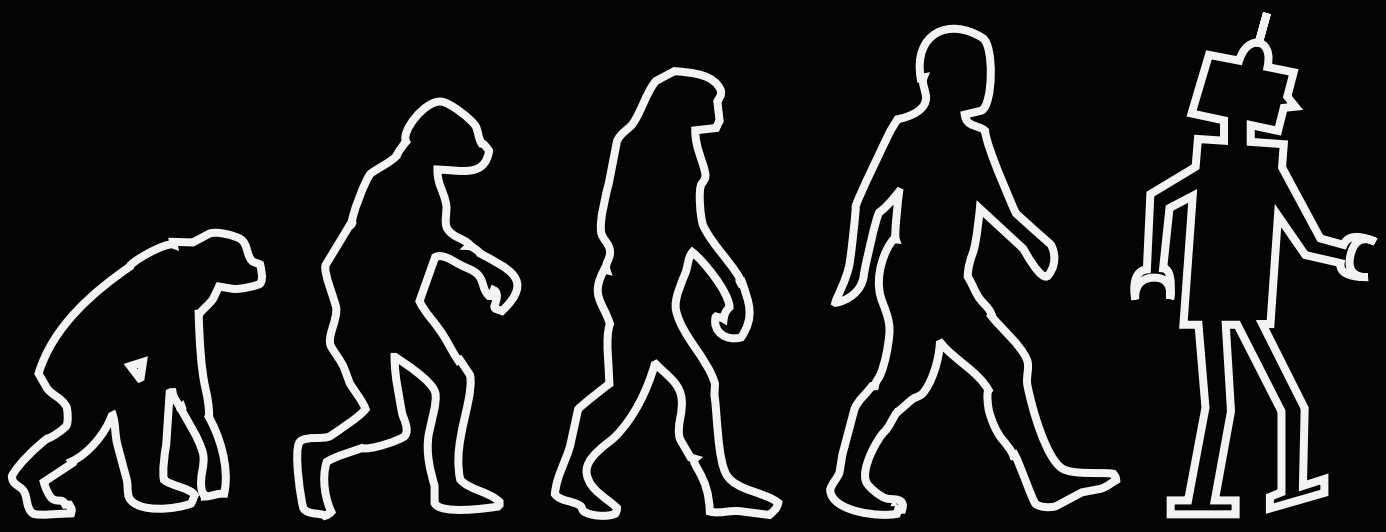
BNAIC2016

PROCEEDINGS

OF THE 28TH BENELUX CONFERENCE ON ARTIFICIAL INTELLIGENCE

NOVEMBER 10-11 2016
AMSTERDAM



BNAIC 2016
Benelux Conference on Artificial Intelligence

PROCEEDINGS OF THE TWENTY-EIGHT

BENELUX CONFERENCE ON ARTIFICIAL INTELLIGENCE

Amsterdam, November 10-11, 2016

General Chairs: Tibor Bosse and Bert Bredeweg

Student Program (MSc & BSc): Arnoud Visser

Interactive Demos: Tom Kenter

Contact and administration: Mojca Lovrenčak

Sponsoring: Natalie van der Wal

Finances: Mark Hoogendoorn

Cover design: Tom Kenter

Website: Adnan Manzoor

Preface

This book contains the proceedings of the 28^{th} edition of the annual Benelux Conference on Artificial Intelligence (BNAIC 2016). BNAIC 2016 was jointly organized by the University of Amsterdam and the Vrije Universiteit Amsterdam, under the auspices of the Benelux Association for Artificial Intelligence (BNVKI) and the Dutch Research School for Information and Knowledge Systems (SIKS).

Held yearly, the objective of BNAIC is to promote and disseminate recent research developments in Artificial Intelligence within Belgium, Luxembourg and the Netherlands. However, it does not exclude contributions from countries outside the Benelux. As in previous years, BNAIC 2016 welcomed four types of contributions, namely A) regular papers, B) compressed contributions, C) demonstration abstracts, and D) thesis abstracts.

We received 93 submissions, consisting of 24 regular papers, 47 short papers, 11 demonstration abstracts and 11 thesis abstracts. After a thorough review phase by the Program Committee, the conference chairs made the final acceptance decisions. The overall acceptance rate was 88% (63% for regular papers, 100% for compressed contributions and demonstration abstracts, and 91% for thesis abstracts).

In addition to the regular research presentations, posters and demonstrations, we were happy to include several other elements in the program of BNAIC 2016, among which keynote presentations by Marc Cavazza (University of Kent), Frank van Harmelen (Vrije Universiteit Amsterdam), Hado van Hasselt (Google DeepMind), and Manuela Veloso (Carnegie Mellon University), a Research meets Business session, a panel discussion on Social Robots, with contributions by Elly Konijn (Vrije Universiteit Amsterdam), Ben Kröse (University of Amsterdam & Amsterdam University of Applied Sciences), Mark Neerincx (TNO & Delft University of Technology), and Peter Novitzky (University of Twente), a special FACt (FACulty focusing on the FACts of AI) session with presentations by Bart de Boer (Vrije Universiteit Brussels), Catholijn Jonker (Delft University of Technology), and Leon van der Torre (University of Luxembourg), and a special session on open access publishing with contributions by Rinke Hoekstra (Vrije Universiteit Amsterdam), Maarten Fröhlich (IOS Press), Bernard Aleva (Elsevier), and Hilde van Wijngaarden (University of Amsterdam/Amsterdam University of Applied Sciences).

To conclude, we want to express our gratitude to all people who made this conference possible: in addition to all invited speakers mentioned above, many thanks to all organizing and program committee members for their hard work in assuring the high quality of this conference. Moreover, we wish to thank all student volunteers, administrative and secretarial assistants, and of course our sponsors. We also gratefully acknowledge help from the BNVKI and from previous organizers. And last, but certainly not least, we cordially thank all the authors who made important contributions to the conference. Without their efforts, this conference could not have taken place.

November 7, 2016
Amsterdam

Chairs:
Bert Bredeweg
Tibor Bosse

Program Committee

Stylianos Asteriadis	University of Maastricht
Reyhan Aydogan	Delft University of Technology
Floris Bex	Utrecht University
Michael Biehl	University of Groningen
Mauro Birattari	IRIDIA, Université Libre de Bruxelles
Peter Bloem	University of Amsterdam
Sander Bohte	Centrum Wiskunde & Informatica
Peter Bosman	Centrum Wiskunde & Informatica
Tibor Bosse	Vrije Universiteit Amsterdam
Bruno Bouzy	Paris Descartes University
Frances Brazier	Delft University of Technology
Bert Bredeweg	University of Amsterdam
Tristan Cazenave	LAMSADE Université Paris Dauphine
Tom Claassen	Radboud University Nijmegen
Fokie Cnossen	University of Groningen
Tom Croonenborghs	University of Leuven
Walter Daelemans	University of Antwerp
Gregoire Danoy	University of Luxembourg
Mehdi Dastani	Utrecht University
Victor de Boer	Vrije Universiteit Amsterdam
Patrick De Causmaecker	University of Leuven
Martine De Cock	University of Washington Tacoma
Mathijs De Weerdt	Delft University of Technology
Benoît Depaire	Hasselt University
Frank Dignum	Utrecht University
Virginia Dignum	TU Delft
Jason Farquhar	Radboud University Nijmegen
Bart Goethals	University of Antwerp
Pascal Gribomont	University of Lige
Perry Groot	Radboud University Nijmegen
Franc Grootjen	Radboud University
Marc Gyssens	Universiteit Hasselt
Frank Harmsen	Ernst & Young Advisory
Tom Heskes	Radboud University Nijmegen
Koen Hindriks	Delft University of Technology
Rinke Hoekstra	University of Amsterdam/Vrije Universiteit Amsterdam
Arjen Hommersom	University of Nijmegen
Bert Kappen	Radboud University
Maurits Kaptein	Eindhoven University of Technology
Uzay Kaymak	Eindhoven University of Technology
Tom Kenter	University of Amsterdam
Walter Kosters	LIACS, Leiden University
Johan Kwisthout	Radboud University Nijmegen
Tom Lenaerts	Universite Libre de Bruxelles
Marco Loog	Delft University of Technology
Elena Marchiori	Radboud University

John-Jules Meyer	Utrecht University
Peter Novák	Science & Technology Corporation B.V.
Mykola Pechenizkiy	Eindhoven University of Technology
Aske Plaat	Leiden University
Eric Postma	TiCC, Tilburg University
Henry Prakken	University of Utrecht & University of Groningen
Jan Ramon	INRIA
Nico Roos	Universiteit Maastricht
Makiko Sadakata	University of Amsterdam
Stefan Schlobach	Vrije Universiteit Amsterdam
Pierre-Yves Schobbens	University of Namur
Johannes Scholtes	University of Maastricht
Martijn Schut	AMC, Universiteit van Amsterdam
Evgueni Smirnov	Maastricht University
Matthijs T. J. Spaan	Delft University of Technology
Jennifer Spenader	University of Groningen
Ida Sprinkhuizen-Kuyper	Radboud University Nijmegen
Pieter H.M. Spronck	Tilburg University
Thomas Stützle	Université Libre de Bruxelles (ULB)
Johan Suykens	University of Leuven
Niels Taatgen	University of Groningen
Annette ten Teije	Vrije Universiteit Amsterdam
Dirk Thierens	Universiteit Utrecht
Jan Treur	Vrije Universiteit Amsterdam
Karl Tuyls	University of Liverpool
Antal van Den Bosch	Radboud University Nijmegen
Egon L. van den Broek	Utrecht University
Jaap van Den Herik	Leiden University
Wil van der Aalst	Eindhoven University of Technology
Peter van der Putten	LIACS, Leiden University & Pegasystems
Leon van der Torre	University of Luxembourg
Natalie Van Der Wal	Vrije Universiteit Amsterdam
Tom van Engers	University of Amsterdam
Tim van Erven	Leiden University
Frank Van Harmelen	Vrije Universiteit Amsterdam
Sietse van Netten	University of Groningen
Martijn Van Otterlo	(former) Radboud University Nijmegen
M. Birna van Riemsdijk	Delft University of Technology
Peter van Rosmalen	Open University of the Netherlands
Maarten Van Someren	University of Amsterdam
Marieke van Vugt	University of Groningen
Joost Vennekens	University of Leuven
Katja Verbeeck	Odisee
Bart Verheij	University of Groningen
Arnoud Visser	Universiteit van Amsterdam
Louis Vuurpijl	Radboud University Nijmegen
Willem Waegeman	Ghent University
Martijn Warnier	Delft University of Technology

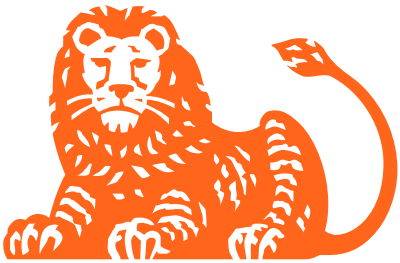
Gerhard Weiss	University Maastricht
Max Welling	University of Amsterdam
Marco Wiering	University of Groningen
Floris Wiesman	AMC, Universiteit van Amsterdam
Jef Wijsen	University of Mons
Mark H. M. Winands	Maastricht University
Radboud Winkels	University of Amsterdam
Cees Witteveen	Delft University of Technology
Yingqian Zhang	Eindhoven University of Technology

Additional Reviewers

Franz, Robin
Merhej, Elie
Voulis, Nina
Wolf, Ben
Ye, Qing Chuan

BNAIC 2016 IS SPONSORED BY

ING

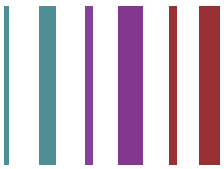


NWO

Netherlands Organisation for Scientific Research

textkernel

Semantic Recruitment Technology



NETWORK INSTITUTE

HAL 24K

Data Intelligence Labs

SNN | ADAPTIVE INTELLIGENCE

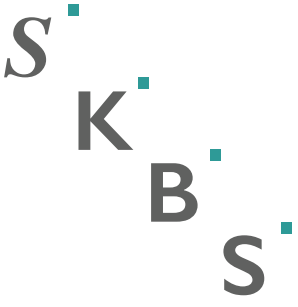


Table of Contents

Regular papers

Compressed contributions

Thesis abstracts

Demonstration abstracts

Textual Inference with Tree-structured LSTMs[1]

Kolawole J. Adebayo[a] Luigi Di Caro[a] Livio Robaldo[b] Guido Boella[a]

[a]University of Turin - {*kolawole.adebayo, luigi.dicaro, guido.boella*}*@unito.it*
[b]University of Luxembourg - *livio.robaldo@uni.lu*

Abstract

Textual Inference is a research trend in Natural Language Processing (NLP) that has recently received a lot of attention by the scientific community. Textual Entailment (TE) is a specific task in Textual Inference that aims at determining whether a hypothesis is entailed by a text. Usually tackled by machine learning techniques employing features which represent similarity between texts, the recent availability of more training data presupposes that Neural Networks that are able to learn latent feature from data for generalized prediction could be employed. This paper employs the *Child-Sum Tree*-LSTM for solving the challenging problem of textual entailment. Our approach is simple and able to generalize well without excessive parameter optimization. Evaluation done on SNLI, SICK and other TE datasets shows the competitiveness of our approach.

1 Introduction

Textual Inference stands at the heart of many NLP tasks. Presently, machines seem to be a bit far from reproducing human capability in reasoning and making semantic deductions from natural languages, e.g., spoken or written text. This is due to the phenomenon of variability and ambiguity in natural languages, since we have different ways of expressing similar ideas[2] [11]. The challenge is for machines to overcome these limitations for it to identify semantic connections such as similarity, entailment and paraphrasing etc., in a body of text. By Textual Inference, we mean the ability of Machines to recognize and quantify similarity between two text portions [1, 18], extract summary or paraphrases from a given text and most importantly, being able to infer the type of semantic connection between two text. The latter is generally referred to as Recognizing Textual Entailment (RTE), where, given two text fragments e.g., a text *T* and an hypothesis *H*, the machine's ability to determine whether T entails H is put to test [3]. This paper focuses on the Textual Entailment task.

Since Dagan et al [12] conceived the task of Recognizing Textual Entailment, it has continued to receive interest from a lot of researchers thus leading to projects and conferences dedicated to it [4, 21]. It has also inspired interests in similar NLP applications such as Factoid Question answering and information extraction [14].

TE tasks can be grouped into two i.e., the binary and multi-class categories. The former requires a Yes/No answer while the latter could be an n-way prediction[3], e.g., a 3-way task where label could be either of Entailed, Neutral or Contradict. The latter is a bit complicated than the former e.g., S_2 below contradicts S_1 while S_3 and S_4 are neutral and entailment with respect to S_1 respectively.
$\mathbf{S}_1$: This church choir sings to the masses as they sing joyous songs from the book at a church.

[1]Kolawole J. Adebayo has received funding from the Erasmus Mundus Joint International Doctoral (Ph.D.) programme in Law, Science and Technology. Luigi Di Caro and Guido Boella have received funding from the European Union's H2020 research and innovation programme under the grant agreement No 690974 for the project "MIREL: MIning and REasoning with Legal texts". Livio Robaldo has received funding from the European Union's H2020 research and innovation programme under the grant agreement No 661007 for the project "ProLeMAS: PROcessing LEgal language in normative Multi-Agent Systems".

[2]E.g., synonymy, polysemy etc.

[3]n is in order of 3 and above. The PASCAL RTE 1-3 are examples of binary class, RTE 5-7 are examples of 3-way task and the Semeval track [13] is an example of a 5-way task

$\mathbf{S}_2$:A choir singing at a baseball game. (Contradiction)
$\mathbf{S}_3$: The church has cracks in the ceiling. (Neutral)
$\mathbf{S}_4$: The church is filled with song. (Entailment)

Most of the reported systems approached the problem as a classification task. They use hand crafted features fed into some machine learning algorithms and often relying on some knowledge resources e.g., WordNet, as well as syntactic and semantic resources, e.g., co-reference resolution, named entity recognizer, parts of speech-tagger and dependency parsing libraries. The features employed typify similarity between the text and hypothesis. The assumption is that common named entities, dependencies such as subject-verb-object-predicate agreement, presence of negation words as well as information retrieval (IR) based intuition such as TFIDF and cosine similarity might be good features for identifying entailment [17]. Other researchers approached the task simply as IR-based task without Machine learning. Reported systems in this category use the word alignment between the text and hypothesis, word overlap, word-word similarity and synonyms substitution using WordNet, surface string similarity e.g. levenhstein distance as well as other syntactic and semantic pointers. However, handcrafting features is usually time-consuming. Moreover, it is often uncertain the combination of features that might work best and ablation test is often done to weigh the features. Nevertheless, these systems rarely or slightly outperform simple baselines relying on just surface string similarity and word-overlap approaches [3].

Recently, Neural Networks such as Convolutional Neural Networks (CNN) [19], Recurrent Neural Networks (RNN) [22] and Long Short-Term Memory (LSTM) [16] have shown to possess the ability to autonomously identify latent features provided there is sufficient amount of data to learn from.

Moreover, it has been successfully applied to several NLP tasks while producing state of the art results, e.g., paraphrase detection [27], Question answering [14], text classification [30], and semantic relatedness [28]. In particular, the authors in [26] using attention-based LSTM have reported state of the art results in textual entailment.

This study builds on these works and show that a more simplified LSTM model which uses fixed raw embeddings learned from even the training data achieves similar result on some well known textual entailment datasets, even without any constraint on size of training data available as well as requiring little or no excessive hyper-parameter fine-tuning.

2 Background

The task of recognizing textual entailment (RTE) aims at making machines to mimic human inference capability, i.e., given a text T and some world knowledge H which could be a ground truth or partial truth (hypothesis), a human reading both the text and the hypothesis can recognize whether the meaning of the hypothesis can be directly inferred from the text [12], the goal is to make automated systems replicate this capability. As earlier pointed out, several authors used machine learning approaches with engineered features [15, 24].

The introduction of SNLI[4], a large dataset of 570K human generated text-hypothesis pairs by Bowman et al.,[9] and the encouraging results obtained from their LSTM-RNN neural networks model, there has been renewed interest in applying deep learning to textual entailment. Moreover, their model outperformed lexical classifiers which use word overlap, bigram and other features. The authors evaluated their models on SICK[5] and SNLI.

The authors in [26] proposed a word-by-word neural attention mechanism based on LSTM. The idea is to have two LSTMs, one reasoning over the sequence of tokens in the text while the other is reasoning on the hypothesis sequence. The second LSTM is conditioned by the first one as its memory is initialized by the output (i.e., the last cell state) of the first LSTM. This is different from the tree-structured LSTM networks proposed in [28] which being order insensitive, is able to capture semantics between two sentences. Notwithstanding, the authors evaluated their model on SNLI and obtained an accuracy of 83.7 and 83.5 on the development and test set respectively.

Baudis et al., [2] also reproduced an attention based model similar to the LSTM-based question answering model in [29]. Also, as in the former, the idea is to attend preferentially to token sequences in the sentence when building its representation. They proposed a RNN model, a CNN model as well as a hybrid RNN-CNN. The RNN captures the long-term dependencies and contextual representation of

[4] http://nlp.stanford.edu/projects/snli
[5] http://clic.cimec.unitn.it/composes/sick.html

words before being fed to the CNN. They report an accuracy of 0.829 and 0.774 respectively on SNLI train and test set respectively. Since this model also rely on embedding sequences, importance is also placed on word order. As pointed out in [28], order-insensitive models capture the semantics of natural language without recourse to syntactic structure differences. We therefore proposed an augmented tree-structure LSTM network which builds sentence representation from constituent subphrases of a text, but takes into account more compositional features for better generalization.

3 Methods

We describe the general LSTM architecture. Specifically, this work employs the *Child-Sum Tree*-LSTMs proposed by [28]. We describe our modified version of the algorithm in this work.

Long Short-Term Memory Networks

A characteristic of deep networks is their ability to autonomously learn semantic representation from text without recourse to time-consuming feature construction. Recurrent Neural Networks (RNNs) have connections that have loops, adding feedback and memory to the networks over time. This memory allows this type of network to learn and generalize across sequences of inputs rather than individual patterns. LSTM Networks [16] are a special type of RNNs and are trained using backpropagation through time, thus overcoming the vanishing gradient problem. LSTM networks have memory blocks that are connected into layers, the block contains gates that manage the blocks state and output. These gates are the *input* gates which decides the values from the input to update the memory state, the *forget* gates which decides what information to discard from the unit and the *output* gates which decides what to output based on input and the memory of the unit. LSTMs are thus able to memorize information over a long period of time since this information are stored in a recurrent hidden vector which is dependent on the immediate previous hidden vector. A unit operates upon an input sequence and each gate within a unit uses the sigmoid activation function to control whether they are triggered or not, making the change of state and addition of information flowing through the unit conditional. We follow the definition and notation of LSTM in [28].

At each time step *t*, let an LSTM unit be a collection of vectors in $\mathbb{R}^d$ where *d* is the memory dimension: an *input gate* i_t, a *forget gate* f_t, an *output gate* o_t, a *memory cell* c_t and a *hidden state* h_t. The state of any gate can either be open or closed, represented as [0,1]. The LSTM transition can be represented with the following equations (x_t is the an input vector at time step *t*, σ represents sigmoid activation function and $\odot$ the elementwise multiplication. The u_t is a tanh layer which creates a vector of new candidate values that could be added to the state) :

$$
\begin{aligned}
i_t &= \sigma\Big(W^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)}\Big),\\
f_t &= \sigma\Big(W^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)}\Big),\\
o_t &= \sigma\Big(W^{(o)}x_t + U^{(o)}h_{t-1} + b^{(o)}\Big),\\
u_t &= \tanh\Big(W^{(u)}x_t + U^{(u)}h_{t-1} + b^{(u)}\Big),\\
c_t &= i_t \odot u_t + f_t \odot c_{t-1},\\
h_t &= o_t \odot \tanh c_t
\end{aligned}
\tag{1}
$$

Tree-Structured LSTMs

Tree-LSTMs are specialized LSTMs that adopt the tree-structure topology, i.e., at any given time step *t* the LSTM is able to compose its states from input vector and hidden states of its child-nodes simultaneously. This is unlike the standard LSTM that assumes a single child per unit since the gating vectors and memory cell updates are dependent on the states of all child-nodes while also maintaining a forget

state for each child. A well-known variant of this structure, i.e., the *Child-Sum Tree*, has been proposed by Tai [28]. The *Child-Sum Tree* LSTMs transition is represented by the following equation, where C_j is the set of children in a node j and k $\in C_j$.

$$\widehat{h_j} = \sum_{k \in C_{(j)}} h_k \tag{2}$$

$$i_j = \sigma\left(W^{(i)}x_j + U^{(i)}\widehat{h_j} + b^{(i)}\right) \tag{3}$$

$$f_j k = \sigma\left(W^{(f)}x_j + U^{(f)}\widehat{h_k} + b^{(f)}\right) \tag{4}$$

$$o_j = \sigma\left(W^{(o)}x_j + U^{(o)}\widehat{h_j} + b^{(o)}\right) \tag{5}$$

$$u_j = \tanh\left(W^{(u)}x_j + U^{(u)}\widehat{h_j} + b^{(u)}\right) \tag{6}$$

$$c_j = i_j \odot u_j + \sum_{k \in C_{(j)}} f_{jk} \odot c_k \tag{7}$$

$$h_j = o_j \odot \tanh c_j \tag{8}$$

Tree-LSTM for Textual Entailment

We use the *Child-Sum Tree*-LSTM to generate sentence representation for both the text and hypothesis by committing a tree to each of text and hypothesis. Assuming both the text and hypothesis are sentences, each with constituent words that are well connected. The structural connection between the words form a deep branching graph, with elements and their dependencies (in case of dependency parsing) where each connection in principle unites a superior term and an inferior term. The inferior term defers to its superior, thus distinguishing between semantically useful words like nouns and verbs to say, a determinative word. With constituency parsing, a phrase-like one-to-one correspondence between the words is observed. The *Child-Sum Tree*-LSTM works better for dependency parse tree representation of a sentence, where each child is a node in the representation. For each node, the LSTM unit takes as input the vectors of its superior word to which it is dependent. In the case of constituency parsing, an LSTM unit takes as input the exact vector of the node.

Considering our n-way classification, given a set of classes Y whose label cardinality corresponds to n, i.e., y_1, y_2, y_3...y_n are the given labels. First, we obtain a representation r_j = (h_{txt}, h_{hyp}) from both the text and hypothesis using the dependency tree representation. Recall that an inferior node takes the fixed raw vectors of its superior while a superior node takes the sum of its vectors and that of all its dependents. At each superior node of the text and hypothesis tree, we encode the entailment relationship as the distance and angle between their element-wise summed vectors and angle of their vectors product. We use the * operator to denote vectors product. We predict the label $\hat{y}_j$, given the input representation r_j observed which encodes the entailment relationship at nodes in the subtree rooted at j. The classifier takes the hidden state h_j at the node as input. The process is represented by the equations below:

$$h_\times = h_{txt} \odot h_{hyp} \tag{9}$$

$$h_+ = h_{txt} + h_{hyp} \tag{10}$$

$$h_\diagdown = |h^*_{txt} \odot h^*_{hyp}| \tag{11}$$

$$h_s = \sigma\left(W^{(\times)}h_\times + W^{(\diagdown)}h_\diagdown + W^{(+)}h_+ + b^{(h)}\right) \tag{12}$$

$$\widehat{p}\theta(y|\{r\}_j) = softmax\left(W^{(p)}h_s + b^{(p)}\right) \tag{13}$$

$$\widehat{y_j} = \arg\max_y \widehat{p}\theta(y|r_j) \tag{14}$$

4 Evaluation

RTE PASCAL challenge [12] is an important avenue for researchers to submit TE systems for public evaluation. We evaluated our system on the PASCAL RTE3 dataset which consists of 800 sentence pairs both for development and test set. The RTE3 dataset has only two classes, i.e., the entailment relation can either be true or false. The SEMEVAL track offering similarity and entailment task also makes use of the SICK dataset[6]. SICK consists of 10000 sentence pairs annotated for use in both sentence similarity and 3-way entailment task. Finally we evaluated our system on the SNLI corpus [9] which till date is the biggest manually constructed entailment dataset publicly available. However, we only use half of the training data in our experiment. Nevertheless, the result obtained shows that our model is able to generalize quite well.

In the context of our ongoing work in the legal domain[7] [5, 8, 7], we explored the option of evaluating our models on a dataset deeply rooted in legal nuances. The three datasets cited above contain sentences that are domain independent and thus have no technical jargons. Our goal is to see how our model would perform within the complex legal domain. Legal texts seem intuitive as they have some peculiarities which set them apart from day-to-day natural language since they employ legislative terms. For instance, a sentence could reference another sentence (e.g., an article) without any explicit link to its text from inside the quoting text. Also, sentences could be long with several clausal dependencies, that is notwithstanding of its inter and intra-sentential anaphora resolution complexity. We opined that a system that is able to achieve good result in this scenario would generalize well given other domain dependent texts.

We used the COLIEE dataset[8] which is a Question-Answering legal corpus made available in the context of COLIEE Legal Information Extraction/Retrieval challenge. Task 2 of the challenge addresses TE, such that, given a sentence and a Query, a system identifies if there is an entailment or not. We provide our evaluation on the 2015 training and test set.

Experiment

We implemented our adaptable Tree-Structured LSTM as proposed by [28]. We obtained dependency tree of both text and hypothesis using Stanford dependency parser [10]. Instead of encoding both text and hypothesis as one-hot encoding representation of the token sequences, we used trained word embedding vectors with fixed weight throughout the experiments. We used 300-dimensional Glove vectors [25]. However, we also generated our own 300-dimensional word embeddings using the popular Word2Vec algorithm[9] [23] on a minimal set of text including SNLI, RTE2, RTE3 and SICK datasets. For the experiment on COLIEE corpus, we included its text in those used to generate the first Word2Vec word embeddings, using the same parameter for training, we obtained a separate embeddings to use for this particular run. An oddity is that even with the embeddings generated from a rather small text, the obtained vectors still capture the semantics of the sentences. We therefore compare the result obtained when we use Glove as well as our trained embeddings.

We used the Keras[10] Deep Learning library to construct two models which only differ based on their configuration depth. Model1 is quite basic, it has a single hidden layer with 300 LSTM units. Model2 mimics the feed-forward MLP networks in that we maintain a stack of layers to increase the depth and scale-up the performance. The first hidden layer size = 300, the second hidden layer size = 200, the third hidden layer size = 150. Depending on the cardinality of classification in each dataset, the sigmoid output layer predicts the right label based on the class distribution. We observed that applying a heavy dropout between 0.4 and 0.5 in the hidden layers lowers the performance. Hence, a moderate dropout between 0.1 and 0.2 was used depending on the performance on the validation data.

As earlier pointed out, our goal is to avoid excessive parameter fine-tuning specifically for each dataset. Because of this, we maintain a uniform batch-size of 25. We used ADAM, a stochastic optimizer with learning rate set at 0.01 and a decay value of 1e-4 as well as momentum of 0.9. Typically, we set the

[6] `http://clic.cimec.unitn.it/composes/sick.html`

[7] Specifically, the MIREL project: `http://www.mirelproject.eu`, which is drawn from our past project EUCases [6].

[8] `http://webdocs.cs.ualberta.ca/~miyoung2/COLIEE2016`

[9] No stopword removal or stemming was done, all terms were lemmatized. We used the skip-gram representation with context window of 10, dimension of 300 and min-count of 5 using the gensim package.

[10] `https://github.com/fchollet/keras`

number of training epochs to 250. However, throughout the experiments, we track the model for over-fitting by monitoring the peak accuracy on the validation data as well as the epoch where the model does not seem to learn sufficiently again and the accuracy begins to drop or does not increase anymore over a period of time. Usually, we pick this epoch for our test runs.

Tables 1, 2 and 3 display the results obtained on the datasets used. For table 1, we used the results from Rocktaschel et al.,[26], Baudis et al., [2] and Bowman et al., [9] as the baseline systems on SNLI and SICK respectively. For PASCAL-RTE3, we did not include any baseline system since there is no recent work which use similar deep learning approach on that dataset. For the result on COLIEE dataset in table 3, we include as our baseline, the best and the baseline results as reported by the authors in [20]. Note that we only use a randomly sampled half of the SNLI data for training due to computation time for our experiments were conducted on the CPU and not the GPU. We also give a comparison of the performances of the models with Glove vectors and our generated Word2Vec embeddings from the training data. It turns out that there is no clear distinction between the results obtained in this respect. In fact, instances exist where we obtained higher accuracy with our trained embeddings. We also observed that the depth/configuration complexity of the network influences accuracy by some order of magnitude since we obtained marginally improved accuracy with *model2* that has more hidden LSTM layers stacked.

Model	SNLI Train	SNLI Test	SICK Train	SICK Test
Model1 (Glove)	84.76	78.30	85.10	76.00
Model1 (w2vec)	84.40	78.00	-	-
Model2 (Glove)	85.30	79.40	88.00	80.10
Model2 (w2vec)	85.90	78.30	-	-
Neural Attention[26]	85.30	83.50	-	-
attn1511 [2]	82.90	77.40	85.80	76.70
LSTM-RNN [9]	84.80	77.60	99.90	80.80

Table 1: Evaluation on SNLI and SICK datasets

Model	Train	Test
Model1 (Glove)	95.76	86.90
Model1 (w2vec)	94.80	82.24
Model2 (Glove)	95.30	87.20
Model2 (w2vec)	93.20	87.60

Table 2: Evaluation on PASCAL-RTE3 dataset

Model	Train	Test
Model1 (w2vec)	76.00	64.19
Model2 (w2vec)	81.33	70.10
Baseline [20]	-	55.87
Best [20]	-	63.87

Table 3: Evaluation on COLIEE dataset

Discussion and Conclusions

In this paper, we have presented a *Child-Sum Tree*-LSTM model for solving the challenging problem of textual entailment. We showed that our approach performs competitive with other state of the art deep learning systems applied to textual entailment. We reported our evaluations on SNLI, SICK, and relatively small PASCAL-RTE3 datasets. We evaluated our models on a domain specific Question-Answering corpus with background in the field of Law. Results are shown in Tables 1, 2, and 3.

The table 1 reveals slight over-fitting in our models since it did not generalize well when compared to the baselines, i.e., we seem to obtain better accuracy on the training data than the authors in [26] while their system generalize better on unseen data. However, observing deeply, we see only slight variations between train and test accuracy. Probably, the model is not able to minimize error or learn better. For the SNLI dataset, it is possible that there is an imbalance in class distribution when we randomly shuffled the data in order to pick one-half of the original training data used in our experiments. Class imbalance can skew the learning curve, thus leading to improper generalization. Note that for SICK, we split the data into train, validation and test set at the ration 60:20:20 respectively. Likewise, COLIEE has no official validation set, we therefore split the test set into validation and test in the ratio 50:50.

The seemingly poor result obtained on the COLIEE corpus is quite noticeable. Two issues might be connected to this. First, the general consensus is that with deep neural networks, more data gives

a better generalization. In fact, a lexicalized classifier with carefully handcrafted features or a simple n-gram based classifier might generalize better with little training data in comparison to a very complex deep network architecture with little data to learn from. The COLIEE corpus with less than 500 training sentence pairs is some order of magnitude smaller than the PASCAL-RTE3 corpus which is generally agreed to be insufficient for training neural networks[11]. Secondly, we believe that the result on PASCAL-RTE3 is not totally bad, given that the training corpus is of similar size. Another factor may be considered. For instance, we believe that the technicality of the text and the fact that it is domain specific pose some issues. First, the embedding was not obtained from the training of a large legal text from which the semantics of legal jargons are best captured. Furthermore, compared to SNLI and SICK, both the text and hypothesis in COLIEE are unusually long, which as pointed out earlier, is a clear characteristic of legal text. However, we posit that more training data and the inclusion of more legal data for generating word embedding might improve results. Nevertheless, we report improved result to the baselines from [20].

Compared to all the benchmarked baselines, our approach is simple and required little or no hyperparameter fine-tuning. Obtaining a better generalization should be possible by optimizing various parameters, e.g., a grid search optimization of some parameters might lead to a more accurate model. Also, introducing more layers in our model might achieve better generalization even though it can also lead to model memorizing features without any significant improvement in the learning curve.

References

[1] E. Agirrea, C. Baneab, D. Cerd, M. Diabe, A. Gonzalez-Agirrea, R. Mihalceab, G. Rigaua, J. Wiebef, and B. Donostia. Semeval-2016 task 1: Semantic textual similarity, monolingual and cross-lingual evaluation. *Proceedings of SemEval*, pages 497–511, 2016.

[2] Petr Baudiš and Jan Šedivỳ. Sentence pair scoring: Towards unified framework for text comprehension. *arXiv preprint arXiv:1603.06127*, 2016.

[3] Luisa Bentivogli, Peter Clark, Ido Dagan, Hoa Dang, and Danilo Giampiccolo. The seventh pascal recognizing textual entailment challenge. *Proceedings of TAC*, 2011, 2011.

[4] Luisa Bentivogli, Ido Dagan, Hoa Trang Dang, Danilo Giampiccolo, and Bernardo Magnini. The fifth pascal recognizing textual entailment challenge. *Proceedings of TAC*, 9:14–24, 2009.

[5] G. Boella, L. Di Caro, L. Humphreys, L. Robaldo, R. Rossi, and L. van der Torre. Eunomos, a legal document and knowledge management system for the web to provide relevant, reliable and up-to-date information on the law. *Artificial Intelligence and Law*, to appear, 2016.

[6] Guido Boella, Luigi Di Caro, Michele Graziadei, Loredana Cupi, Carlo Emilio Salaroglio, Llio Humphreys, Hristo Konstantinov, Kornel Marko, Livio Robaldo, Claudio Ruffini, Kiril Simov, Andrea Violato, and Veli Stroetmann. Linking legal open data: Breaking the accessibility and language barrier in european legislation and case law. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law*, ICAIL '15, New York, NY, USA, 2015. ACM.

[7] Guido Boella, Luigi Di Caro, Daniele Rispoli, and Livio Robaldo. A system for classifying multi-label text into eurovoc. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, ICAIL '13, pages 239–240, New York, NY, USA, 2013. ACM.

[8] Guido Boella, Luigi Di Caro, and Livio Robaldo. *Semantic Relation Extraction from Legislative Text Using Generalized Syntactic Dependencies and Support Vector Machines*, pages 218–225. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[9] Samuel R Bowman, Gabor Angeli, Christopher Potts, and Christopher D Manning. A large annotated corpus for learning natural language inference. *arXiv preprint arXiv:1508.05326*, 2015.

[10] Danqi Chen and Christopher D Manning. A fast and accurate dependency parser using neural networks. In *EMNLP*, pages 740–750, 2014.

[11] e.g., compared to 570,000 sentence pairs from SNLI

[11] Ido Dagan, Bill Dolan, Bernardo Magnini, and Dan Roth. Recognizing textual entailment: Rational, evaluation and approaches–erratum. *Natural Language Engineering*, 16(01):105–105, 2010.

[12] Ido Dagan, Oren Glickman, and Bernardo Magnini. The pascal recognising textual entailment challenge. In *Machine learning challenges. evaluating predictive uncertainty, visual object classification, and recognising tectual entailment*, pages 177–190. Springer, 2006.

[13] Myroslava O Dzikovska, Rodney D Nielsen, Chris Brew, Claudia Leacock, Danilo Giampiccolo, Luisa Bentivogli, Peter Clark, Ido Dagan, and Hoa T Dang. Semeval-2013 task 7: The joint student response analysis and 8th recognizing textual entailment challenge. Technical report, DTIC, 2013.

[14] Minwei Feng, Bing Xiang, Michael R Glass, Lidan Wang, and Bowen Zhou. Applying deep learning to answer selection: A study and an open task. In *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, pages 813–820. IEEE, 2015.

[15] M. Gaona, A. Gelbukh, and S. Bandyopadhyay. Recognizing textual entailment using a machine learning approach. In *Mexican International Conference on Artificial Intelligence*. Springer, 2010.

[16] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8), 1997.

[17] S. Jimenez, G. Duenas, J. Baquero, A. Gelbukh, J. Bátiz, and A. Mendizábal. Unal-nlp: Combining soft cardinality features for semantic textual similarity, relatedness and entailment. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 732–742, 2014.

[18] Adebayo Kolawole John, Luigi Di Caro, and Guido Boella. Normas at semeval-2016 task 1: Semsim: A multi-feature approach to semantic text similarity. *Proceedings of SemEval*, 2016.

[19] N. Kalchbrenner, E. Grefenstette, and P. Blunsom. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*, 2014.

[20] M.Y. Kim, Y. Xu, and R. Goebel. A convolutional neural network in legal question answering.

[21] Bernardo Magnini, Roberto Zanoli, Ido Dagan, Kathrin Eichler, Günter Neumann, Tae-Gil Noh, Sebastian Pado, Asher Stern, and Omer Levy. The excitement open platform for textual inferences. In *ACL (System Demonstrations)*, pages 43–48, 2014.

[22] LR Medsker and LC Jain. Recurrent neural networks. *Design and Applications*, 2001.

[23] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[24] Shachar Mirkin, Ido Dagan, and Eyal Shnarch. Evaluating the inferential utility of lexical-semantic resources. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 558–566. Association for Computational Linguistics, 2009.

[25] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In *EMNLP*, volume 14, pages 1532–43, 2014.

[26] Tim Rocktäschel, Edward Grefenstette, Karl Moritz Hermann, Tomáš Kočiskỳ, and Phil Blunsom. Reasoning about entailment with neural attention. *arXiv preprint arXiv:1509.06664*, 2015.

[27] R. Socher, E. Huang, J. Pennin, C. Manning, and A. Ng. Dynamic pooling and unfolding recursive autoencoders for paraphrase detection. In *Advances in Neural Information Processing Systems*, pages 801–809, 2011.

[28] Kai Sheng Tai, Richard Socher, and Christopher D Manning. Improved semantic representations from tree-structured long short-term memory networks. *arXiv preprint arXiv:1503.00075*, 2015.

[29] Ming Tan, Bing Xiang, and Bowen Zhou. Lstm-based deep learning models for non-factoid answer selection. *arXiv preprint arXiv:1511.04108*, 2015.

[30] Xiang Zhang, Junbo Zhao, and Yann LeCun. Character-level convolutional networks for text classification. In *Advances in Neural Information Processing Systems*, pages 649–657, 2015.

Towards legal compliance by correlating Standards and Laws with a semi-automated methodology[1]

Cesare Bartolini[a] Gabriele Lenzini[a] Livio Robaldo[a]

[a] *University of Luxembourg,*
Interdisciplinary Centre for Security, Reliability and Trust (SnT)
{*cesare.bartolini,gabriele.lenzini,livio.robaldo*}*@uni.lu*

Abstract

Since legal regulations do not generally provide clear parameters to determine when their requirements are met, achieving legal compliance is not trivial. If there were a clear correspondence between the provisions of a specific standard and the regulation's requirements, one could implement the standard to claim a presumption of compliance. However, finding those correspondences is a complex process; additionally, correlations may be overridden in time, for instance, because newer court decisions change the interpretation of certain provisions. To help solve this problem, we present a framework that supports legal experts in recognizing correlations between provisions in a standard and requirements in a given law. The framework relies on state-of-the-art Natural Language Semantics techniques to process the linguistic terms of the two documents, and maintains a knowledge base of the logic representations of the terms, together with their defeasible correlations, both formal and substantive. An application of the framework is shown by comparing a provision of the European General Data Protection Regulation against the ISO/IEC 27018:2014 standard.

1 Introduction

Generally, achieving legal compliance is not a trivial task for enterprises. Laws and regulations generally do not specify what measures should be implemented to match the requirements that they state. For instance, in the European Union (EU), where there is the need to support a single market while letting each country establish its own legal framework, legislation normally defines the safety/security of products and systems only at a very abstract level, leaving to the Member States the choice on how to demonstrate compliance[2].

In this situation, implementing standards is a viable way for enterprises to create a "presumption of conformity with the specific legal provision they address" [6]. Implementing a standard *per se* does not guarantee legal compliance, but it can provide a significant clue of conformity with regulations. When widely adopted and subject to repeated audits by conformity-assessing bodies, a standard offers an argument of compliance. Of course, such an argument is a presumption, giving a plaintiff the possibility to demonstrate that the organization failed to comply with the legal framework. Still, this *inversion of the burden of proof* is often preferable to a personalized solution [6]. However, the problem of establishing and assessing this presumption remains, and it may become a serious issue when new laws reshape the legal landscape for businesses: while awaiting for the establishment of feasible solutions, an enterprise faces the risk of liability for not being compliant and may, on that ground, be sanctioned.

[1] Livio Robaldo has received funding from the European Union's H2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 661007. Cesare Bartolini has received funding from the European Union's H2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 690974.

[2] Regulation (EU) 1025/2012, Article 2.1

A domain that will need to be thoroughly addressed from this point of view is data protection, because the recent adoption of the General Data Protection Regulation (GDPR)[3], which will be applied from 25 May 2018[4], will pose significant challenges for undertakings in terms of ensuring compliance with it, as it brings major changes to the regulatory framework for personal data protection in Europe. On the other hand, the ISO/IEC 27018:2014 standard, concerning public clouds acting as personal data processors, can be regarded as a building block [6] that helps data-processing organizations comply with the principle of accountability and with their obligations resulting from data protection laws.

We propose a software framework that aids in determining the formal and substantive correlations between the provisions in a standard and in a law. The framework's core is a logic-based methodology to represent, in a machine-processable format, (*a*) the relevant syntactic concepts in the provisions, and (*b*) the relevant correlations between them. In this paper, we describe this logic-based methodology and exemplify how it works using provisions from the ISO 27018 standard and the GDPR.

The framework depends on two auxiliary functional blocks (see Section 3): (*i*) a *logic knowledge base* that can be used, corrected and extended by legal experts, that stores the machine-processable logic correlations; (*ii*) a *set of Natural Language Semantics (NLS) and Natural Language Processing (NLP) techniques* that allow a user to browse a XML representation of the documents and to search and retrieve the words, terms, and sentences that have been found relevant for correlation. The NLS and NLP techniques will help users show the established correlations within the knowledge base. Any expert user can contribute to reinforce, correct, justify, and expand the correlations. Due to differences in legal interpretation, some of the correlations could be in contradiction. Interpretations from authoritative sources (such as high courts) will eventually sort contradictions out.

Technically, the different correlations are expressed in a deontic and defeasible logic for legal semantics called Reified Input/Output Logic (see Section 2). The selection of relevant terms and the definition of correlations, requiring human reading, processing and decision-making, is therefore semi-automatic. It will be supported by the extensible knowledge base; the process of browsing, aided by the tools and techniques from the NLS and NLP domains, add efficiency and precision.

2 Related work

Reified Input/Output Logic is the defeasible logic that we propose as the formal language to express correlations. It was designed as an attempt to extend the state-of-the-art research in legal informatics, by investigating the *logical architecture* of the provisions, which are available in natural language only.

Reified Input/Output logic is a logical framework [15] for representing the meaning of provisions. Contrary to the large majority of its competitors (*e.g.*, [8, 10, 7]), Reified Input/Output logic integrates modern insights from the NLS literature. Specifically, it merges Input/Output logic [11], a well-known formalism in Deontic Logic, with a first-order logic for NLS grounded on the concept of *reification* [9]. Reification [4] is a concept that allows to move from standard notations in first-order logic such as "$(give\ a\ b\ c)$", asserting that "a" gives "b" to "c", to another notation in first-order logic "$(give\ e\ a\ b\ c)$", where "e" is the *reification* of the giving action. "e" is a first-order logic term denoting the giving event of "b" by "a" to "c". Thanks to reification, the logic is able to express a wide range of phenomena in NLS such as named entities, quantification, anaphora, causality, modality, time, and others. In particular, the simplified version of Reified Input/Output logic that we use in our example is an extension of first-order logic that distinguishes three kinds of implication: "$\rightarrow$", "$\rightsquigarrow$", and "$\Rightarrow$".

The implication "$\rightarrow$" is the standard trust-value implication of first-order logic, whereas "$\rightsquigarrow$" is its *defeasible* [14] version. Defeasible here means that an implication "$\Phi \rightsquigarrow \Psi$" holds by default unless overridden by "stronger" implications. When instantiated properly, this notion of stronger implications resolves the potential contradictions emerging because of the non-monotonic nature of the defeasible reasoning. Reified Input/Output logic also includes other mechanisms to deal with unresolvable conflicts[5]. A possible solution to deal with this type of conflicts is to leave the conflict open until more evidence will allow the reasoner to take a decision. This is, in short, what better fits a situation with conflicts due to multiple legal interpretations. The third implication, "$\Rightarrow$", is a deontic implication. Taken a

[3] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[4] GDPR, Article 99.2.

[5] See `plato.stanford.edu/entries/logic-nonmonotonic/#sec-2-2`.

formula Φ, referring to a set of pre-conditions, and another formula Ψ, referring to a set of actions, the meaning of "$\Phi \Rightarrow \Psi$" is "given the pre-condition Φ, actions Ψ are obligatory", that is, the actions must be undertaken if the pre-conditions hold. Note that, according to the interpretation of "$\Rightarrow$", the formula "$\Phi \Rightarrow \Psi \wedge \Phi \wedge \neg\Psi$" is not inconsistent: it only means obligation Ψ has been violated.

The framework in [15] further distinguishes the formulæ belonging to the assertive contextual statements (ABox), *i.e.*, the formulæ denoted by the provisions, from those belonging to the terminological declarative statements (TBox), *i.e.*, the definitions, axioms, and constraints on the predicates used in the ABox formulæ. Formulæ] in the ABox are in the form "$\forall_{x_1} \ldots \forall x_n[\Phi(x_1 \ldots x_n) \Rightarrow \Psi(x_1 \ldots x_n)]$", where arguments "$\Phi(x_1 \ldots x_n)$" and "$\Psi(x_1 \ldots x_n)$" are conjunctions in standard first-order logic. Formulæ in the TBox can be any formula in standard first-order logic augmented with the operator "$\rightsquigarrow$".

The GDPR is the most recent step in the evolution of data protection rules in the European Union. The Regulation, which will apply from May 2018, replaces the current Directive 95/46/EC[6], which introduced a minimum level of protection but required implementation by means of Member State laws. As such, the Directive produced a heterogeneous legislation throughout the European Union, although with a common base of protection. On the other hand, the GDPR, being a Regulation, will be directly applicable in all Member States, and thus create a homogeneous data protection framework in the EU.

The purposes of the GDPR, according to its promoter [13], are to align data protection rules with the most recent developments in data-processing technologies, while still providing a legislation that is flexible enough not to become outdated over the course of a few years. In addition, the rights of the data subject are strengthened by burdening the data controller with new obligations, and enforcing such obligations with heavy penalties. Controllers and processors are required to be compliant with the Regulation, which sets up fines as high as four percent of the total annual worldwide turnover of a company in some cases of infringements of its provisions.

The interpretation of the Regulation will be provided mainly by doctrine and jurisprudence. The latter is not available yet: since the Regulation was recently published and won't be applicable for the next two years, no decisions based on its provisions exist yet. When it will be applied, relevant decisions on its interpretations are expected to be issued by the Data Protection Authoritys (DPAs) of Member States, by national courts, and by the Court of Justice of the European Union (CJEU). On the other hand, some doctrinal analysis from legal experts already exists (*e.g.*, [5]), and significant literature will be published in the upcoming months.

3 Methodology

The framework we propose offers a computer-aided methodology to analyze standards to make an argument of compliance with respect to a specific piece of legal text, and is schematically summarized in Figure 1. Users (who may be lawyers, regulators, auditors, or other legal experts) access a digital and annotated XML representation of the normative texts (laws and standards). While browsing a document and selecting the relevant concepts, NLP and NLS tools help traverse the rest of the documents, find related terms, and recall previous correlations between them. Correlations from different sources have different degrees of importance, which need to be tracked using specific metadata. The framework implements a collaborative strategy to evaluate the stored correlations. The user's decisions are stored in the knowledge base, after being appropriately represented in a logic for legal semantics.

The framework, and in particular its knowledge base, does not pretend to be complete. Rather, it provides the expert user with an updated knowledge that helps him take autonomous and informed decisions, both when confirming the correlations the tool suggests and when choosing to define new correlations. The knowledge base is designed to support defeasible reasoning, *i.e.*, to tolerate (apparent) inconsistencies of different interpretations of terms, by overriding general assertions into more contextually-specific ones. Conflicts are especially frequent in legal interpretation, but they can generally be solved considering that the interpretation by higher-instance courts, such as the CJEU, prevails over lower ones. In order to cope with interpretations of different legal weights, which may supersede one another, the logic formalism that the framework embeds is defeasible: correlations can be updated, modified, rewritten and weighed. If conflicts do remain, the framework still embeds strategies that help the user take a decision.

[6]Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

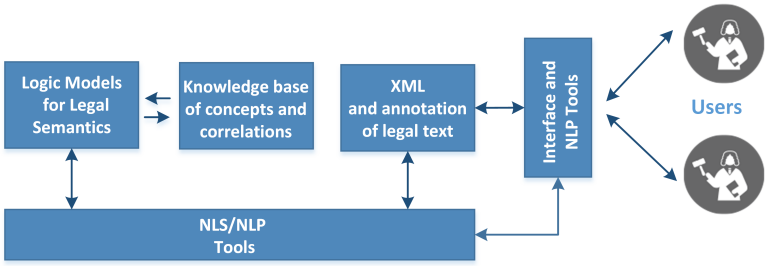


Figure 1: The framework at a glance.

The methodology we propose is highly interdisciplinary, involving a strict interaction between law and computer science. Its purpose is to establish correlations between the provisions in the standards and those in the law. As pointed out in Section 1, correlations can be further divided into two different categories: *formal* and *substantive* correlations.

Formal correlations entail a mere textual overlap between concepts. For example, formal correlations would allow us to observe that both the GDPR and the ISO 27018 standard use the term "notify" (see Section 4). On the other hand, substantive correlations are more complex and entail the analysis of the actual meaning of terms. To assert a correlation of this kind, requirements must be met in a concrete way. Following the previous example, to assert a correlation between a provision of the GDPR and one of the standard concerning notification, it is necessary to verify the exact meaning of the term "notify" in the two texts.

Implications that define a match are *defeasible* and can be overridden, for instance, because of a new decision by a court or DPA, or simply because of the evolution of the interpretation in doctrinal analysis. This way, the output of the methodology can be easily kept up to date by introducing new correlations as they are developed by legal actors.

The methodology follows three steps to build the correlations, involving both a legal and a technical approach. The legal approach is focused on the interpretation of the provisions of laws (the GDPR in our example) and standards (ISO 27018), whereas the technical approach consists of modelling those provisions into an ontology, and expressing the interpretation by means of logical formulæ.

Step 1: analysis of the provisions. The provisions of the law and of the security standard are analyzed by legal experts, who provide an interpretation of the terms used in the provisions and compare them in search of semantic correlations. There is no need for this interpretation to be final, as more interpretations can be added later, and old interpretations can be overridden by newer ones, but this start requires a significant manual activity.

To support the execution of this step, legal experts are assisted by external NLP procedures that suggest (semi-automatically and during the browsing of the documents) previous translations and correlations on the basis of the information currently stored in the knowledge base. Ultimately, it is the legal expert who must decide whether and how correlations need to be overridden. In that case, new correlations are added and annotated with the source which contributed to define it (*e.g.*, Court of Justice of the European Union, *Dapreco and Copreda Corp.*, C-XYZ/16). Implications by more authoritative sources override defeasible implications by less authoritative ones.

Step 2: creation of legal ontologies. The legal interpretations are mapped onto legal ontologies of the law and the security standard. Legal ontologies [2] model the legal concepts, parties and stakeholders affected by the law, the duties and rights of each stakeholder, and the sanctions for violating the duties. As per ontologies in general, legal ontologies are expressed in a knowledge representation language. For this work, we chose the popular abstract language OWL, which can be serialized using various XML notations. For example, the OWL representation of the data protection ontology will contain concepts such as "controller", "data subject", "personal data", "processing" and so on.

A preliminary version of a legal ontology for the GDPR has been defined already [1]. Albeit partial and based on an older version of the GDPR, it was designed to express the duties of the controller. As such, it can be used to find the correspondences between the requirements expressed in the GDPR and in security standards, until an improved ontology is built.

Step 3: generation of logic formulæ. The third and final step of the methodology consists of generating the logical formulæ representing the set of provisions in the law and the set of provisions in the security standard, as well as the implications between them. These formulæ are expressed in Reified Input/Output logic [15]. An example is shown in the next section.

Associating textual provisions to logical formulæ amounts to converting ambiguous and vague terms into non-ambiguous items (predicates and terms). Words in the provisions are represented via predicates reflecting their vagueness. For example, the word "notify", included in the sample provisions used in Section 4, will be represented via the homonym predicate "notify". These "vague" predicates may be defined by adding implications and further constraints (axioms). Those implications will be *defeasible*, so that they can account for different legal interpretations. Predicates are associated with classes of the ontologies developed in Step 2 or with standard general-purpose ontologies/repositories belonging to the NLS literature, *e.g.*, Verbnet [16].

4 Generation of Logic Formulæ: example

We exemplify step 3 of our methodology. This step, which lies at the core of the methodology, is the most technical, and more innovative than steps 1 and 2 which instead rely upon existing techniques. We use a provision from the GDPR and an article of the ISO 27018 security standard:

(*a*) GDPR, Article 33.2: *The processor shall notify the controller without undue delay after becoming aware of a personal data breach.*

(*b*) ISO 27018, Article 9.1: *The public cloud PII processor should promptly notify the relevant cloud service customer in the event of any unauthorized access to PII.*

The formalization of the two provisions in the simplified version of Reified Input/Output logic we use in this paper is rather intuitive. As explained above in Section 3, the formulæ will include predicates reflecting the vagueness of the terms occurring in the sentences. Thus, for instance, the verb "notify", which occurs in both provisions, is formalized into an homonym predicate "notify".

On the other hand, the provisions in the ISO 27018 use the term "PII" (Personally Identifiable Information) while the ones in the GDPR use the term "personal data". Although it sounds rather obvious to consider the two terms as synonyms (as suggested in ISO 27018, Article 0.1), and thus associate them with the same predicate, our methodology keeps them distinct, *i.e.*, it formalizes them via two different predicates "personalData" and "PII". An additional axiom is then added to the TBox, in order to correlate the two predicates:

$$\forall_x[\,(\mathsf{PII}\ x)\ \rightsquigarrow\ (\mathsf{personalData}\ x)\,] \tag{1}$$

The axiom correlates the two predicates via a defeasible implication: all that is considered PII is also by default considered personal data. Note that (1) allows the knowledge base to include instances of PII which are *not* personal data. Those would be *exceptions* to the default rule in (1), and it would be necessary to add further higher-priority axioms in order to consistently account for them.

In the light of this, the GDPR provision in (*a*) is formalized as follows:

$$\begin{aligned}\forall_e\forall_x\forall_y\forall_z\forall_{e_p}\forall_{e_c}\forall_{e_b}[\,(\,&(\mathsf{dataProcessor}\ x)\ \wedge\ (\mathsf{dataController}\ y)\ \wedge\ (\mathsf{personalData}\ z)\ \wedge\\ &(\mathsf{process}\ e_p\ x\ z)\ \wedge\ (\mathsf{control}\ e_c\ y\ z)\ \wedge\ (\mathsf{dataBreach}\ e_b\ z)\,)\\ &\Rightarrow\ \exists_{e_n}[\,(\mathsf{notify}\ e_n\ x\ y\ e_b)\ \wedge\ (\mathsf{nonDelayed}\ e_n)\,]\,]\end{aligned} \tag{2}$$

In (2), "e_p", "e_c", and "e_b" are variables referring to three events. "e_p" is an event of processing the personal data "z" (patient) performed by the data processor "x" (agent). "e_c" is an event of controlling the personal data "z" (patient) performed by the data controller "y" (agent). Finally, "e_b" is an event of data breach of the personal data "z" (patient). Note that the agent of the data breach has been omitted; that could be unknown, but this is irrelevant for the obligation in (2), which concerns the data processor.

(2) states that for any data breach of personal information "y", it is mandatory for the data processor "x" who processes "y" to notify the data controller "z" who controls "y". Such a notifying event "e_n" must be carried out without undue delay; this is formalized via a unary predicate "nonDelayed", which is assumed to be true if the event in its argument has been performed without undue delay.

In (2), "notify" and "nonDelayed" are predicates whose meaning is subject to different legal interpretations. Recalling the difference between *formal* and *substantive* compliance outlined in Section 1, we note that the formalization in (2) only enforces formal compliance. The formula in (2) simply requires the data controller to notify data breaches without undue delay, but it does not specify *how* notifications should be performed for being legitimate.

For instance, the data processor could require the data controller to acknowledge the notification, in order to make sure it was received. Similarly, the processor could be required to avoid sending notifications of data breaches via standard paper mail, in that the time needed by the postal service to deliver the mail could be considered as an undue delay. It is up to judicial authorities to establish the substantive compliance of the obligation in (2).

Of course, we do not have the authority to decide whether (2) is performed in the proper way. In our work, we only aim at providing a methodology to keep track of all legal interpretations of the provisions. From a formal point of view, the knowledge base must be enriched with axioms defining the conditions under which the predicates in the formula are true. Note that different authorities could establish different (conflicting) interpretations of these predicates; this is why *defeasible* implications are needed to model their substantive meaning.

Specifically, the TBox of the ontology will include defeasible axioms that define when, by default, a data processor properly notifies a data controller. For instance, by assuming that email with electronic signature and registered high-priority paper mail are both proper and prompt means to notify the data controller, the following two (defeasible) axioms are added to the TBOX.

$$\begin{aligned}&\forall_x\forall_y\forall_{e_1}\forall_{e_2}[\,(\mathsf{sendEmailWithES}\ e_1\ x\ y\ e_2)\ \rightsquigarrow\ (\,(\mathsf{notify}\ e_n\ x\ y\ e_2)\ \wedge\ (\mathsf{nonDelayed}\ e_n)\,),\\&\forall_x\forall_y\forall_{e_1}\forall_{e_2}[\,(\mathsf{sendRegHPMail}\ e_1\ x\ y\ e_2)\ \rightsquigarrow\ (\,(\mathsf{notify}\ e_n\ x\ y\ e_2)\ \wedge\ (\mathsf{nonDelayed}\ e_n)\,)\,]\end{aligned}\tag{3}$$

In case a judicial authority later decides, for example, that emails with electronic signature are no longer proper and prompt means for notifying data breaches, an additional higher-priority axiom will be added to the TBox in order to override the first axiom in (3). Axioms may be associated to time stamps (although this is not shown in (3)), so that the new axiom will override the old one only for notifying actions performed after a certain date, *i.e.*, since the new decision was issued. On the other hand, the old axiom will still assert that emails with electronic signature are proper and prompt means for all notifications performed before that date.

Formula (4) models the ISO 27018 provision in (*b*):

$$\begin{aligned}\forall_e\forall_x\forall_y\forall_z\forall_{e_p}\forall_{e_c}\forall_{e_b}[\,(\,&(\mathsf{PIIProcessor}\ x)\ \wedge\ (\mathsf{PIIController}\ y)\ \wedge\ (\mathsf{PII}\ z)\,\wedge\\&(\mathsf{process}\ e_p\ x\ z)\ \wedge\ (\mathsf{control}\ e_c\ y\ z)\ \wedge\ (\mathsf{unauthorizedAccess}\ e_a\ z))\\&\Rightarrow\ \exists_{e_n}[\,(\mathsf{notify}\ e_n\ x\ y\ e_a)\ \wedge\ (\mathsf{promptly}\ e_n)\,]\,]\end{aligned}\tag{4}$$

As it was done for formalizing (*a*) into (2), the formula introduces predicates that reflect the vague terms used in the text. With an important exception: we formalized "the relevant cloud service customer" via the predicate "PIIController". The reason is that ISO 27018 includes a constitutive provision that defines the cloud service customer[7].

Therefore, the cloud service customer *is* the PII controller, *i.e.*, the organization handling the data of the PII principals. In case the PII principals handle their data themselves, without any "broker" doing that on their behalf, they *are* the PII controller, *i.e.*, the data cloud service customer, that should receive the notification of unauthorized accesses from the PII processor.

The final ingredient needed to (4) and (2) are axioms relating to the predicates occurring in both, similar to the axiom in (1), which state that PII is by default considered as personal data:

$$\begin{aligned}&\forall_x[\,(\mathsf{PIIProcessor}\ x)\ \rightsquigarrow\ (\mathsf{dataProcessor}\ x)\,],\ \forall_x[\,(\mathsf{PIIController}\ x)\ \rightsquigarrow\ (\mathsf{dataController}\ x)\,],\\&\forall_x\forall_z[\,(\mathsf{unauthorizedAccess}\ x\ z)\ \rightsquigarrow\ (\mathsf{dataBreach}\ x\ z)\,],\\&\forall_x[\,(\mathsf{promptly}\ x)\ \rightsquigarrow\ (\mathsf{nonDelayed}\ x)\,]\end{aligned}\tag{5}$$

Axioms (5) are quite intuitive: for instance, the first one states that any entity that is taken to be a PII processor, with respect to the ISO 27018 standard, is also, by default, taken to be a data processor with respect to the GDPR.

[7] ISO 27018, Article 0.1: "The cloud service customer, who has the contractual relationship with the public cloud PII processor, can range from a natural person, a 'PII principal', processing his or her own PII in the cloud, to an organization, a 'PII controller', processing PII relating to many PII principals".

It is easy to verify that every tuple of variables "e", "x", "y", "z", "e_p", "e_c", and "e_b" that satisfies formula (4) also satisfies formula (2) by default.

Again, such default correlations may be rewritten. For example, a judicial authority could later decide that although a notifying action may be considered as "prompt" with respect to the ISO 27018 standard, it cannot be considered as being "without undue delay" with respect to the GDPR. In such a case, the last implication in (5) needs to be overridden by higher-priority axioms: notifications would be henceforth considered "without undue delay" with respect to the GDPR only if they are "prompt" with respect to the ISO 27018 standard *and* they sport the extra features decided by the judicial authority.

Other instances of data breaches may be encompassed as well. For instance, ISO 27018 also imposes the prompt notification of "unauthorized access to processing equipment or facilities resulting in loss, disclosure or alteration of PII". If we formalize such an unauthorized event "e" in terms of a predicate "$(\mathsf{lossDisclosureOrAlteration}\ e\ z)$", the corresponding formula is obtained by inserting this predicate in place of "$(\mathsf{unauthorizedAccess}\ e_a\ z)$" in formula (4). The correlation with respect to the GDPR provision formalized in (2) is achieved by adding the following axiom to the TBox:

$$\forall_x \forall_z [\, (\mathsf{lossDisclosureOrAlteration}\, x\, z) \rightsquigarrow (\mathsf{dataBreach}\, x\, z)\,] \tag{6}$$

5 Discussion and Conclusion

In this paper we address the problem of legal compliance, *i.e.*, establishing how to be conforming to regulations and laws. The problem is not one of easy solution for at least two reasons. First, unless specific practices of compliance are directly mentioned in a law, what to do to achieve compliance is often unclear for an enterprise. Second, it is becoming hard to cope with a growing landscape of laws and regulations. To cope with such an overload, companies often resort to standards, which offer to the subject implementing them (not full compliance, but) a presumption of compliance. The question of establishing such a presumption thus reduces the problem to finding formal and substantive correlations between the terms used to describe the practices in the standards (*e.g.*, sending a digitally signed email to the PII controller) and those used in the provisions of the law (*e.g.*, duty of notification to the data controller).

This paper introduces a semi-automated methodology to help practitioners find and establish such correlations. The paper exemplifies the methodology in the context of data protection regulations (in particular, the GDPR) and security standards (ISO 27018), two domains which display significant overlaps. By following the illustrated methodology, one can build a machine-processable *knowledge base* of logic formulæ that model and store relevant concepts from a law and a standard, together with their possible formal correlations. The knowledge base, which will be collaboratively accessible, will have its records updated and labelled considering the outcomes of specific auditing processes or decision of the courts; in time, it will embed substantive correlations on which a legal argument for compliance can be relied upon. It should be pointed out that the logic into which we translate the correlations is defeasible, allowing them be overridden. The methodology herein is currently a work in progress and not fully implemented yet. The complete methodology requires the definition of a detailed taxonomy of concepts extracted from the law and the security standard. In the future, we envision significant developments of the research presented in this paper. In addition to the technical skills needed to manage the knowledge base, the methodology would greatly benefit from a close interaction with legal authorities. In a small country like Luxembourg, this could be easily achieved.

Several technical challenges related to building and updating the knowledge base are raised. The translations from natural language to logical formulæ must be uniform for excerpts of text that are similar to each other. To achieve this, we must overcome the limitations of a manual translation, which would be time-consuming and error-prone. For this reason, our work must rely on NLP technologies. However, even at the best of their performances, current NLP algorithms are still unable to automatically carry out the translation with a reasonable level of accuracy, so we advocate a *semi-automatic* translation of the provisions. Similar approaches are applied to translations in general, where translators are helped by collaborative tools such as the "SDL Trados Studio"[8], which suggests, via pattern-recognition text-similarity NLP techniques [12, 3], how to translate a sentence on the basis of the translations of similar sentences that the translators have previously stored in the tool. Inspired by that approach, we will

[8] http://www.translationzone.com/products/trados-studio/

develop an enhanced text editor to assist the manual translation of provisions into formulæ. For each provision, the editor will display the translations of similar provisions found via NLP procedures applied to the provisions already stored in the knowledge base. Finally, we are aware that the knowledge base must be consistent, *i.e.*, without contradictions, even after having applied the defeasibility measures. To check for consistency, we plan to store formulæ using a XML-based data model, and employ/extend reasoners to monitor the consistency of the knowledge base, whenever new formulæ are added to it.

References

[1] C. Bartolini, R. Muthuri, and C. Santos. Using Ontologies to Model Data Protection Requirements in Workflows. In *Proc. of the 9th Int. Work. on Juris-informatics (JURISIN)*, pages 27–40, Nov. 2015. Extended version to be published in LNAI book.

[2] R. Benjamins, B. Selic, and A. Gangemi, editors. *Law and the Semantic Web: Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications*, volume 3369 of *LNAI*. Springer-Verlag Berlin Heidelberg, 2005.

[3] G. Boella, L. Di Caro, A. Ruggeri, and L. Robaldo. Learning from syntax generalizations for automatic semantic annotation. *The J. of Intelligent Information Systems*, 43(2):231–246, 2014.

[4] D. Davidson. The logical form of action sentences. In N. Rescher, editor, *The Logic of Decision and Action*. Univ. of Pittsburgh Press, 1967.

[5] P. De Hert and V. Papakonstantinou. The proposed data protection Regulation replacing Directive 95/46/ec: A sound system for the protection of individuals. *Computer Law & Security Review*, 28(2):130–142, April 2012.

[6] P. De Hert, V. Papakonstantinou, and I. Kamara. The cloud computing standard ISO/IEC 27018 through the lens of the EU legislation on data protection. *Computer Law & Security Review*, 32(1):16–30, February 2016.

[7] G. Governatori, F. Olivieri, A. Rotolo, and S. Scannapieco. Computing strong and weak permissions in defeasible logic. *Journal of Philosophical Logic*, 42(6):799–829, 2013.

[8] J. Hansen. Prioritized conditional imperatives: problems and a new proposal. *Autonomous Agents and Multi-Agent Systems*, 17(1):11–35, 2008.

[9] J. R. Hobbs. *Deep Lexical Semantics*, volume 4919 of *LNCS*, pages 183–193. Springer-Verlag Berlin Heidelberg, 2008.

[10] J. Horty. *Reasons as Defaults*. Oxford University Press, 2012.

[11] D. Makinson and L. W. N. van der Torre. Input/output logics. *Journal of Philosophical Logic*, 29(4):383–408, 2000.

[12] R. Mihalcea, C. Corley, and C. Strapparava. Corpus-based and knowledge-based measures of text semantic similarity. In *Proc. of the 21st National Conference on Artificial Intelligence - Volume 1*, AAAI'06, pages 775–780. AAAI Press, 2006.

[13] V. Reding. The upcoming data protection reform for the European Union. *International Data Privacy Law*, 1(1):3–5, February 2011.

[14] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[15] L. Robaldo, L. Humphreys, L. Sun, L. Cupi, C. Santos, and R. Muthuri. Combining input/output logic and reification for representing real-world obligations. In *Post-proc. of the 9th Int. Work. on Juris-informatics. Lecture Notes in Artificial Intelligence*, 2016.

[16] Karin Kipper Schuler. *Verbnet: A Broad-coverage, Comprehensive Verb Lexicon*. PhD thesis, Philadelphia, PA, USA, 2005. AAI3179808.

Agents for Mobile Radio Tomography

K. Joost Batenburg [a b] Leon Helwerda [c] Walter A. Kosters [c]
Tim van der Meij [c]

[a] *Mathematical Institute, Universiteit Leiden*
[b] *CWI, Amsterdam*
[c] *Leiden Institute of Advanced Computer Science (LIACS), Universiteit Leiden*

Abstract

Mobile radio tomography uses moving agents that perform wireless signal strength measurements in order to obtain a reconstruction of objects inside an area of interest. We propose a toolchain to facilitate the planning, data collection and reconstruction process. The main components are the automated planning of missions for the agents, and the dynamic tomographic reconstruction. Preliminary experiments show that the approach is feasible and results in smooth images that clearly depict objects at the expected locations, when using missions that sufficiently cover the area of interest.

1 Introduction

Radio tomography is a technique for measuring the signal strength of low-frequency radio waves which are exchanged between *sensors* around an area, and reconstructing information about objects in that area. We send a signal between a source and target sensor of a bidirectional *link*. The signal passes through objects that attenuate it, which results in a detectably weaker signal at the receiving end. This phenomenon makes it possible to determine where objects are located. The typical setup for radio tomography is illustrated in Figure 1 (left), in which the sensors are situated on the boundaries in an evenly distributed manner. Gray lines represent unobstructed links and red lines indicate links attenuated by the object.

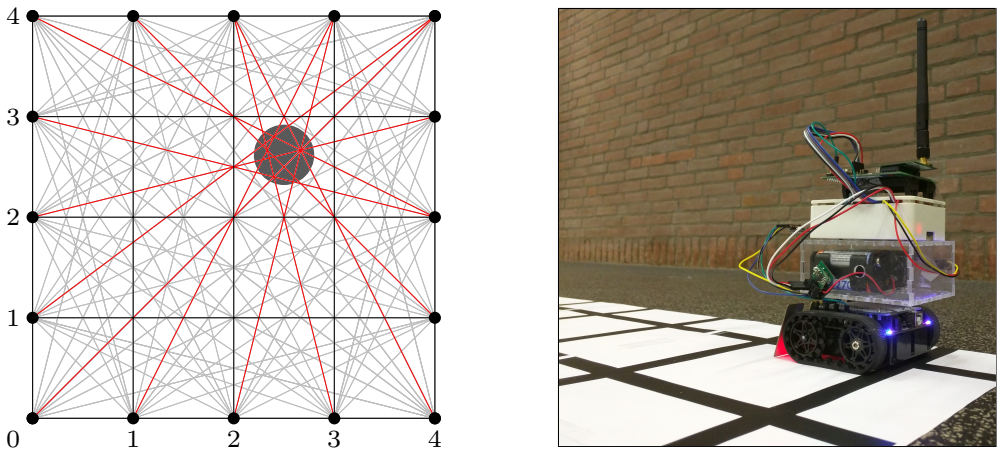


Figure 1: Network of sensors for radio tomography, and the physical realization of a vehicle.

There are benefits of radio tomography over other detection techniques. We can see through walls, smoke or other uninteresting obstacles. The technique is non-intrusive and does not require objects to carry sensor devices themselves. Furthermore, it is less privacy-invasive than optical cameras, because the possible level of detail is limited so that a detected object can only be pictured as an unidentifiable blob.

A static sensor network with a large number of affordable sensors, placed around an area of interest, can be used to reconstruct and visualize a smooth image in real time [12]. This discrete setup is in contrast to the more traditional continuous approach from, e.g., [9]. The drawbacks of such a static network are that it requires a large number of sensors, and there is no way to resolve gaps in the sensor coverage or to react to information obtained through the reconstruction.

One way to resolve these issues is to move the sensors around using *agents*, which are realized as autonomous *vehicles* as pictured in Figure 1 (right). We position them along a *grid* which defines the discrete sensor positions and resulting *pixels* in the reconstructed image. In comparison to the static setup, we require fewer sensors and less prior knowledge about the area. We may adapt the coverage dynamically, for example by zooming in on a part of the area. We name this concept *mobile radio tomography*, which includes both the agent-based measurement collection and the dynamic reconstruction approach.

In this paper, we present our toolchain for mobile radio tomography using intelligent agents. In Section 2 we describe the key challenges for mobile radio tomography and the components in our toolchain that address them. We then cover two such challenges in greater detail: (i) planning the paths of the agents in Section 3, and (ii) reconstructing an image from the measurements in Section 4. Results for real-world experiments are presented in Section 5, followed by conclusions and further research in Section 6. This paper is based on two master's theses on the subject of mobile radio tomography [8, 10].

2 Toolchain

Compared to existing radio tomographic imaging techniques that use statically positioned sensors, imaging using dynamically positioned agents leads to several new challenges. In particular, routes must be planned for each agent to obtain isotropic sampling of the network, short total scanning time and collision-free movement. Images must be reconstructed from the measurements in real time, requiring algorithms and models that work with highly limited and noisy data. Communication between the agents must be interleaved with data acquisition using a robust protocol. To deal with these challenges, we develop an open-source, component-based toolchain. The toolchain is written in Python, with low-level hardware components written in C. The diagram in Figure 2 shows the components in the toolchain.

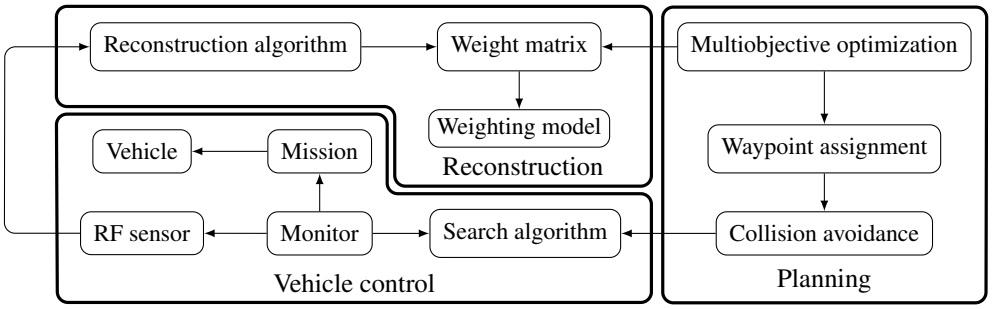


Figure 2: Diagram of components in the toolchain.

The planning components are responsible for automatic mission generation, which we discuss in Section 3. Creating a set of links to measure may be done manually or using an evolutionary multiobjective algorithm that uses a weight matrix to ensure that the links cover the entire network. Next, a waypoint assignment algorithm distributes the sensor positions for each link over the vehicles. Collision avoidance, using a path graph search algorithm, is applied to prevent the vehicles from clashing with each other.

Execution of the mission is taken care of by the vehicle control components. The monitor oversees the process and tracks auxiliary sensors on the vehicles, such as distance sensors for obstacle detection. It makes the RF (radio frequency) sensor perform the signal strength measurements and it may use the search algorithm for collision avoidance during a mission. The mission, consisting of the waypoints for the vehicle, instructs the vehicle to control its motors for moving to the next waypoint.

The reconstruction components convert the signal strength measurements to a two-dimensional image. We study this process in Section 4. The reconstruction algorithm outputs a reconstruction of the area of interest, which can be visualized. The weight matrix determines which pixels are intersected by a link, and a weighting model describes how the contents of pixels contribute to a measured signal strength.

3 Missions

We instruct the autonomous vehicles to travel to locations around the area of interest, that two-by-two correspond to the positions where sensor measurements are performed. The vehicles then execute this *mission*, consisting of *waypoints* that denote which locations they should visit in order. They should choose a short and safe path that does not conflict with any other concurrent route.

We wish to plan the mission algorithmically, rather than assigning the waypoints by hand. The vehicles should visit *sensor positions* such that they can perform a measurement together. This problem is related to other vehicle routing problems [7] with synchronization constraints [5]. We propose a two-stage algorithm, the first part described in textual format, and the second part also in pseudocode.

Assuming that we already know which links we want to measure for collecting tomographic data, we can distribute the positions of the sensors required for each link among the vehicles. Thus, our assignment algorithm is given as input a set $P = \{(p_{1,1}, p_{1,2}), (p_{2,1}, p_{2,2}), \ldots, (p_{\omega,1}, p_{\omega,2})\}$ with ω location pairs of *coordinate tuples* (two-dimensional vectors), and a set $V = \{v_1, v_2, \ldots, v_\eta\}$ of $\eta \geq 2$ vehicles, which are initially located at the predefined coordinate tuples $S_1, S_2, \ldots, S_\eta$. Now define U as the pairwise unique permutations of the vehicles, e.g., with two vehicles, this is $U = \{(v_1, v_2), (v_2, v_1)\}$.

Our greedy assignment algorithm then works as follows: for each vehicle pair $\vartheta = (v_a, v_b) \in U$ and each sensor pair $\rho = (p_{c,1}, p_{c,2}) \in P$, determine the distances $d_1(\vartheta, \rho) = \|S_a - p_{c,1}\|_1$ and $d_2(\vartheta, \rho) = \|S_b - p_{c,2}\|_1$. We use the L^1 norm $\|\cdot\|_1$ because we move only in cardinal directions on a grid; in other applications we may use the L^2 norm $\|\cdot\|_2$. Next, take the maximum of the distances, and finally select the overall minimal pair combination, i.e., solve the following optimization problem:

$$\underset{(\vartheta,\rho)\in U\times P}{\arg\min} \Big(\max(d_1(\vartheta,\rho), d_2(\vartheta,\rho))\Big) \tag{1}$$

The selected positions are then assigned to the chosen vehicle pair, and removed from P. Additionally, S_a becomes the first position and S_b becomes the second sensor position. The greedy algorithm then continues with the next step, until P is empty, thus providing a complete assignment for each vehicle.

Secondly, we design a straightforward collision avoidance algorithm that searches for routes between waypoints that do not *conflict* by crossing any concurrent route of another vehicle; see Algorithm 1. The algorithm is kept simple in order to incorporate it into an evolutionary algorithm (see [11] for more involved methods). We use a search algorithm to find a *safe route* that crosses no other routes. Only when a vehicle performs a measurement involving some other vehicle, their prior routes no longer conflict.

Algorithm 1 The collision avoidance algorithm.

let $V = \{v_1, v_2, \ldots, v_\eta\}$, and $W_1, W_2, \ldots, W_\eta$ be sets, in which $W_i = \{v_i\}$ for $i = 1, 2, \ldots, \eta$
initialize the graph G of possible positions and connections in the area
remove incoming edges of nodes in G that enter forbidden areas, and those of $S_1, S_2, \ldots, S_\eta$
initialize empty sequences of routes $R_1, \ldots, R_\eta$
procedure AVOID($S_1, S_2, \ldots, S_\eta, v_p, v_q, N_p$)
    for all $v_i \in V \setminus W_p$ **do**
        remove the edges for nodes in R_i from G
    end for
    let $R^* \leftarrow$ SEARCH(G, S_p, N_p) ▷ find a safe path R^* in G from S_p to N_p
    append R^* to R_p, reinsert the edges for S_p to G and remove incoming edges for the N_p node
    $S_p \leftarrow N_p$ and $W_p \leftarrow W_p \cup \{v_q\}$
    for all $v_i \in V$ **do**
        if $v_i \notin W_p$ **then**
            reinsert the edges for nodes in R_i to G
        end if
        if $v_i \neq v_p \wedge W_i = V$ **then**
            clear the sequence R_i
            $W_i \leftarrow \{v_i\}$
        end if
    end for
end procedure

Let v_p be the vehicle that we currently assign the position N_p to, and v_q the vehicle that will visit the other sensor position. We also initialize sets $W_1, W_2, \ldots, W_\eta$, where each W_i indicates with which other vehicles the given vehicle v_i has recently performed a measurement. We assume that the search algorithm is given as input a graph G, start point S_p and end point N_p, and outputs a route of *intermediate* points R^*, or an empty sequence if there is no safe path. We can use the collision avoidance algorithm every time the waypoint assignment algorithm assigns a position to a vehicle, so twice per step. Thus, we detect problematic situations as they occur, which are either solved via detours (although the vehicle might also search for a faster safe path during the actual mission), or by rejecting the complete assignment.

In order to supply sensor positions to the waypoint assignment and collision avoidance algorithms, we use an evolutionary multiobjective algorithm [6] to iteratively generate sets of positions that converge toward a theoretical optimal assignment. We keep a population $(X_1, X_2, \ldots, X_\mu)$ of multiple *individuals*, each of which contains variables that encode the positions in adequate form. After a random initialization, the algorithm performs iterations in which it selects a random individual X_i and slightly mutates it to form a new individual [2].

In our case, the variables encode coordinates for positions, of which $m^{(i)}$ are correctly placed such that they intersect the network. From this, we can deduce other information, such as a weight matrix $A^{(i)}$ (containing link influence on pixels, see Section 4) for each individual X_i. Then the algorithm removes an individual that is infeasible according to domains or the *constraints* in Equations 2 and 3 such as a minimum number of valid links ζ, wrapped into a combined feasibility value in Equation 4:

$$(2)\quad Q_1^{(i)} : \exists j : \forall k : A_{j,k}^{(i)} \neq 0 \qquad (3)\quad Q_2^{(i)} : m^{(i)} \geq \zeta \qquad (4)\quad f_i = \begin{cases} 0 & \text{if } \neg Q_1^{(i)} \vee \neg Q_2^{(i)} \\ 1 & \text{if } Q_1^{(i)} \wedge Q_2^{(i)} \end{cases}$$

If all constraints and domain restrictions are met, the multiobjective algorithm uses a selection procedure based on the *objectives*. We remove an individual if its objective values are strictly higher than those of one *dominating* individual in the population. If no individuals are dominated, we remove the one with the least area around it when placed in a plotted function, the *Pareto front*. The objectives to minimize are the two functions (the algorithm favors two over more than two objectives) in Equations 5 and 6:

$$(5)\quad g_1(X_i) = -\sum_{j=1}^{m^{(i)}} \sum_{k=1}^{n} A_{j,k}^{(i)} \qquad (6)\quad g_2(X_i) = \delta \cdot \left(\sum_{j=1}^{m^{(i)}} \left\| p_{j,1}^{(i)} - p_{j,2}^{(i)} \right\|_2 \right) + (1 - \delta) \cdot T^{(i)}$$

For the selection step of the evolutionary algorithm, we use the reconstruction, waypoint assignment and collision avoidance algorithms to check that a new individual adheres to the constraints and to calculate the objective values. Aside from the weight matrix $A^{(i)}$ for one individual X_i, we calculate the pairwise L^2 norms between sensor positions, and $T^{(i)}$, the sum of all minimized distances of Equation 1, weighted by the factor δ. These algorithms generate missions that provide sufficient network coverage.

4 Reconstruction

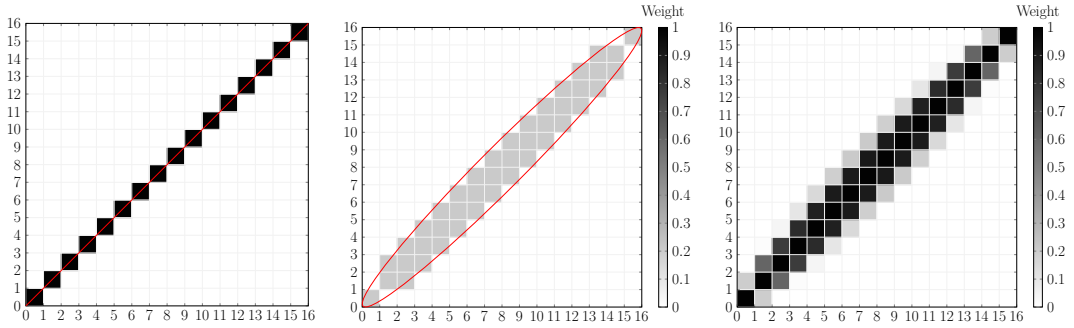
The reconstruction phase converts a sequence of signal strength measurements to a two-dimensional image of $m \times n$ pixels that may be visualized. Let $M = \{(s_1, t_1, r_1), \ldots, (s_k, t_k, r_k)\}$ be the input, in which s_i and t_i are pairs of integers indicating the x and y coordinates on the grid for the source and target sensor i, respectively, r_i is the *received signal strength indicator* (*RSSI*) and k is the total number of measurements. Performing the conversion is done by expressing the problem algebraically as $A\vec{x} = \vec{b}$, in which $\vec{b}$ is a column vector of RSSI values $[r_1, r_2, \ldots, r_k]^T$, $\vec{x}$ is a column vector of $m \cdot n$ pixel values (in row-major order) and A is a weight matrix that describes how the RSSI values are to be distributed over the pixels that are intersected by the link, according to a weighting model.

In general, signal strength measurements contain a large amount of noise due to *multipath interference*. The wireless sensors send signals in all directions, and thus more signals than those traveling in the line-of-sight path may reach the target sensor, causing interference. This phenomenon is especially problematic in indoor environments due to reflection of signals. Techniques exist to include an estimate of the contribution of noise in the model [12], but we suppress noise outside the model using calibration measurements and regularization algorithms. The difficulty lies in the fact that the reconstruction problem is an ill-posed inverse problem, which is primarily caused by the unstable nature of the measurements.

The *weight matrix* A defines the mapping between the input $\vec{b}$ and the output $\vec{x}$. If and only if the contents of a pixel with index i attenuate a link with index ℓ, then the weight $w_{\ell,i}$ in row ℓ and column i is nonzero. Given a link ℓ, a *weighting model* determines which pixels have an influence on the link and are thus assigned nonzero weights, which may be normalized using the link length d_ℓ to favor shorter links [13]. The variable $d_{\ell,i}$ is the sum of distances from the center of pixel i to the endpoints of link ℓ. The *line model* in Equation 7 assumes that the signal strength is determined by objects on the line-of-sight path, as shown in Figure 3a; the *ellipse model* in Equation 8 is based on the definition of Fresnel zones:

$$(7)\quad w_{\ell,i} = \begin{cases} 1 & \text{if link } \ell \\ & \text{intersects pixel } i \\ 0 & \text{otherwise} \end{cases} \qquad (8)\quad w_{\ell,i} = \begin{cases} 1/\sqrt{d_\ell} & \text{if } d_{\ell,i} < d_\ell + \lambda \\ 0 & \text{otherwise} \end{cases} \qquad (9)\quad w_{\ell,i} = e^{-\frac{(d_{\ell,i}-d_\ell)^2}{2\sigma^2}}$$

Fresnel zones, used to describe path loss in communication theory, are ellipsoidal regions with focal points at the endpoints of the link and a minor axis diameter λ. Only pixels inside this region are assigned a nonzero weight, as depicted in Figure 3b. Moreover, we introduce a new *Gaussian model* in Equation 9 which assumes that the distribution of noise conforms to a Gaussian distribution. The log-distance path loss model is a signal propagation model that describes this as well [1]. We use a Gaussian function that assigns most weight to pixels on the line-of-sight path and less weight to pixels farther away from this path, based on the parameter σ, as indicated in Figure 3c.



(a) Line model (b) Ellipse model (c) Gaussian model

Figure 3: Illustration of weight assignment for a link from (0, 0) to (16, 16) by the weighting models.

Due to the ill-posed nature of the problem, in general there exists no exact solution for $A\vec{x} = \vec{b}$ because A is not invertible. Instead, we attempt to find a solution $\vec{x}_{min}$ that minimizes the error using *least squares approximation* [3] as defined in Equation 10:

$$\vec{x}_{min} = \arg\min_{\vec{x}} \left(\left\| A\,\vec{x} - \vec{b} \right\|_2^2 + R(\vec{x}) \right) \tag{10}$$

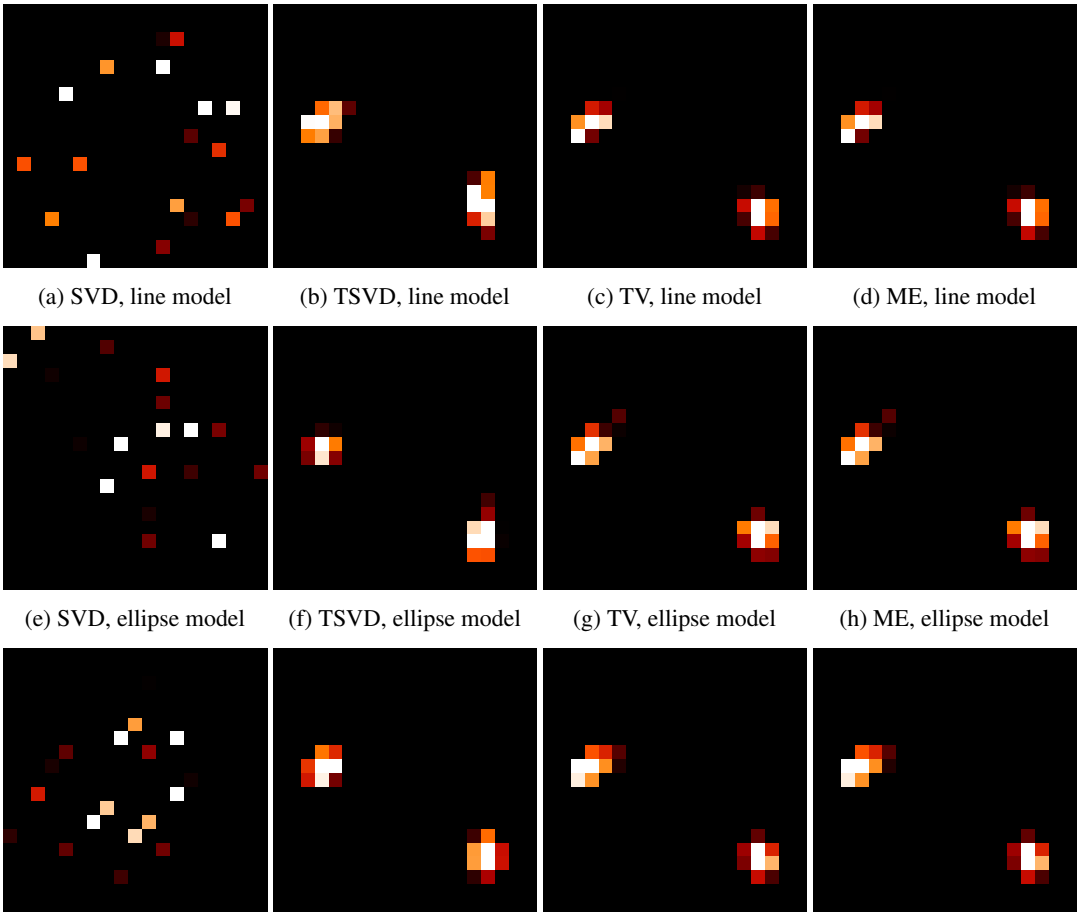
The *singular value decomposition* (*SVD*) of A may be used to solve this and is defined as $A = U\Sigma V^T$, in which U and V are orthogonal matrices and Σ is a diagonal matrix with singular values [13]. Singular value decomposition is exact and does not apply any regularization, so the regularization term $R(\vec{x}) = 0$. *Truncated singular value decomposition* (*TSVD*) is a regularization method that only keeps the τ largest singular values in the SVD and is defined as $A = U_\tau \Sigma_\tau V_\tau^T$ [13]. Small singular values have low significance for the solution and become erratic when taking the reciprocals for Σ.

Iterative regularization methods incorporate desired characteristics of the reconstructed images. *Total variation minimization* (*TV*, see [13]) enforces that the reconstructed images are smooth, i.e., that the differences between neighboring pixels are as small as possible, by penalizing slow changes. The gradient $\nabla\vec{x}$ of $\vec{x}$ is a measure of the variability of the solution. The regularization term in Equation 10 is set to $R(\vec{x}) = \alpha \sum_{i=0}^{\xi-1} \sqrt{(\nabla\vec{x})_i^2 + \beta}$, in which ξ is the number of elements in $\nabla\vec{x}$. The parameter α indicates the importance of a smooth solution and leads to a trade-off as a high value indicates more noise suppression, but less correspondence to the actual measurements. The term is not squared, so we need an optimization algorithm for the minimization. The parameter β is a small value that prevents discontinuity in the derivative when $\vec{x} = 0$, as that generally needs to be supplied to optimization algorithms.

Finally, we consider another variability measure. *Maximum entropy minimization* (*ME*) smoothens the solution by minimizing its entropy. The *Shannon entropy* is defined as $H = -\sum_{i=0}^{\gamma-1} q_i \log_2(q_i)$, in which γ is the number of unique gray levels in the solution and q_i is the probability that gray level i occurs in the solution. Low entropy indicates a low variation in gray levels (which we observe as noise). While this regularization technique is well-known [4], we have found no previous work about its application to radio tomographic imaging. The regularization term in Equation 10 is set to $R(\vec{x}) = \alpha H$. We calculate a numerical approximation of the derivative. The described reconstruction methods and weighting models allow us to obtain a clear image of the area.

5 Experiments

To study the effectiveness of our approach, we perform a series of experiments. Two vehicles drive around on the boundaries of a 20×20 grid in an otherwise empty experiment room. We use hand-made missions that apply common patterns used in tomography, such as fan beams. With this setup we create a dataset with two persons standing in the middle of the left side and in the bottom right corner of the network, and a dataset with one person standing in the top right corner of the network. A separate dataset is used for calibration. The first experiment compares all combinations of regularization methods and weighting models to determine which pair creates the most accurate reconstructions. We use the dataset with the two persons, so both must be clearly visible. The outcome of this experiment is presented in Figure 4, in which darker pixels indicate low attenuation and brighter pixels indicate high attenuation.



(a) SVD, line model (b) TSVD, line model (c) TV, line model (d) ME, line model

(e) SVD, ellipse model (f) TSVD, ellipse model (g) TV, ellipse model (h) ME, ellipse model

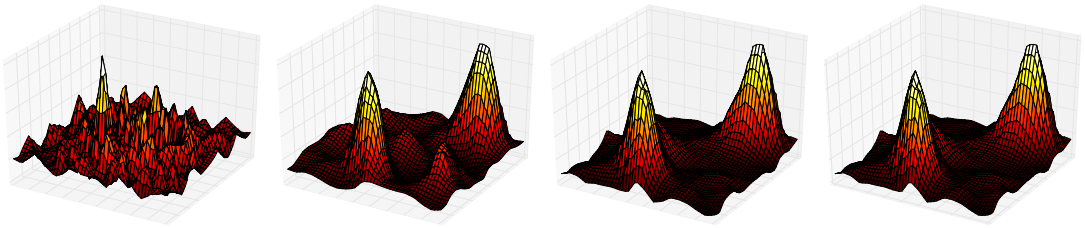
(i) SVD, Gaussian model (j) TSVD, Gaussian model (k) TV, Gaussian model (l) ME, Gaussian model

Figure 4: Reconstructions combining regularization methods and weighting models (two persons dataset).

SVD indeed leads to major instabilities because of the lack of regularization. Noise is amplified to an extent that the images do not provide any information about the positions of the persons. TSVD

provides more stable results that clearly show the positions of the two persons. The ellipse model and the new Gaussian model yield similar clear results, whereas the use of the line model leads to slightly more noise compared to the former two. Even though TV and ME use different variation measures, the reconstructions are visually the same.

Besides providing a clear indication of where the persons are located inside the network, it is important that the reconstructed images are smooth. The second experiment studies the smoothening effects of the regularization methods using 3D surface plots of the raw grayscale images, i.e., without any additional coloring steps applied. The only difference between the experiment runs are the regularization method, so any other parameters remain the same, such as the Gaussian weighting model and the dataset with the two persons that we use. The results for this experiment are shown in Figure 5.



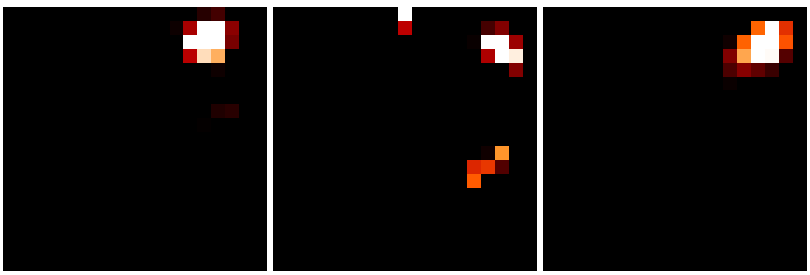
(a) SVD (b) TSVD (c) TV (d) ME

Figure 5: 3D surface plots of the reconstructions for each regularization method (two persons dataset).

The ideal surface plot consists of a flat surface with two spikes exactly at the positions of the persons. The surface plot for SVD is highly irregular, which leads to noticeable noise in the image as there is a high variance in pixel values. In contrast, the surface plot for TSVD is smooth and the two spikes are clearly distinguishable. However, there are still small instabilities. The surface plots for TV and ME are, again, practically the same and have even fewer instabilities.

We now create a mission using the evolutionary multiobjective algorithm from Section 3 and compare it to a hand-made mission. The algorithm places sensors for at least 320 and up to 400 valid measurements. Other parameters are tuned such that they result in lower objective values. At 7000 iterations, we end the run, and pick a solution that optimizes both objectives. The collision avoidance algorithm determines that this assignment is safe.

In Figure 6, we show the images resulting from the tomographic reconstruction of the dataset with one person. The reconstruction, which is run in real time during the collection of measurements, uses TV and the Gaussian model. We can see how long it takes for each mission to provide a smooth and correct result in terms of quality and realism. In Figure 6a, we are around halfway through the planned mission, with 204 out of 382 measurements collected. The reconstructed image clearly shows the person standing in the top right corner, so we end the mission here. Figure 6b shows the reconstructed image obtained from the handmade mission after 413 out of 800 measurements. Although this mission is executed faster due to fewer erratic movements, it is not stable enough to clearly show one person at this point. Figure 6c shows the end result, where the hand-made mission does provide an acceptable reconstructed image.



(a) Algorithmically planned mission, partial result (b) Hand-made mission, partial result (c) Hand-made mission, final result

Figure 6: Reconstructions for algorithmically planned and hand-made missions (one person dataset).

6 Conclusions and further research

We propose a mobile radio tomography toolchain that collects wireless signal strength measurements using dynamic agents, which are autonomous vehicles that move around with sensors. We plan missions, which consist of the locations that the agents must visit and in what order. Novel algorithms provide us with generated missions, guaranteeing that two sensors are at the right locations to perform a measurement. The algorithms avoid conflicts between the routes and provide an optimized coverage of the network.

The measurements are then passed to the reconstruction algorithms to create a visualization of the area of interest that corresponds to the patterns in the data as best as possible. Regularization methods suppress noise in the measurements and increase the smoothness of the resulting image. We introduce a new Gaussian weighting model and apply maximum entropy minimization to the problem of radio tomographic imaging. Results from our preliminary experiments show that the mobile radio tomography approach is effective, i.e., it is able to provide smooth reconstructed images in a relatively short amount of time using algorithmically planned missions.

There is much potential for further research. One interesting topic is to replace the agents, that currently operate on the ground using small-scale robotic rover cars, with drones that fly at different altitudes. This leads to 3D reconstruction, e.g., by performing a reconstruction at different altitudes and combining the images, which are slices of the 3D model. The reconstruction algorithms could also be modified to allow performing measurements anywhere in the 3D space, although this makes the problem more difficult to solve in real time. Finally, improvements can be made to the evolutionary multiobjective algorithm, such as altering the objectives or using predetermined patterns that we encode in the variables. It is also not entirely clear yet how the network coverage, or the lack thereof, influences the quality of the reconstruction. The greedy waypoint algorithm could factor in the time that certain actions take, such as turning around. Altering missions dynamically helps making adaptive scanning a viable approach.

References

[1] J. B. Andersen, T. S. Rappaport, and S. Yoshida. Propagation measurements and models for wireless communications channels. *IEEE Communications Magazine* 33 (1995), pp. 42–49.

[2] T. Bäck. Evolutionary algorithms in theory and practice: Evolution strategies, evolutionary programming, genetic algorithms. Oxford University Press, 1996.

[3] A. Björck. Numerical methods for least squares problems. SIAM, 1996.

[4] A. C. Bovik. Handbook of image and video processing. Academic Press, 2005.

[5] D. Bredström and M. Rönnqvist. Combined vehicle routing and scheduling with temporal precedence and synchronization constraints. *European Journal of Operational Research* 191.1 (2008), pp. 19–31.

[6] M. Emmerich, N. Beume, and B. Naujoks. An EMO algorithm using the hypervolume measure as selection criterion. In: *Proceedings of the Third International Conference on Evolutionary Multi-Criterion Optimization*. LNCS 3410. Springer, 2005, pp. 62–76.

[7] B. L. Golden, S. Raghavan, and E. A. Wasil. The vehicle routing problem: Latest advances and new challenges. Springer, 2008.

[8] L. Helwerda. Mobile radio tomography: Autonomous vehicle planning for dynamic sensor positions. Master's thesis. LIACS, Universiteit Leiden, 2016.

[9] A. C. Kak and M. Slaney. Principles of computerized tomographic imaging. SIAM, 1999.

[10] T. van der Meij. Mobile radio tomography: Reconstructing and visualizing objects in wireless networks with dynamically positioned sensors. Master's thesis. LIACS, Universiteit Leiden, 2016.

[11] G. Sharon, R. Stern, A. Felner, and N. R. Sturtevant. Conflict-based search for optimal multi-agent pathfinding. *Artificial Intelligence* 219 (2015), pp. 40–66.

[12] J. Wilson and N. Patwari. Radio tomographic imaging with wireless networks. *IEEE Transactions on Mobile Computing* 9 (2010), pp. 621–632.

[13] J. Wilson, N. Patwari, and F. Guevara Vasquez. Regularization methods for radio tomographic imaging. In: *Virginia Tech Symposium on Wireless Personal Communications*. 2009.

Aspects of the Cooperative Card Game Hanabi

Mark J.H. van den Bergh [ab] Walter A. Kosters [a] Flora M. Spieksma [b]

[a] *Leiden Institute of Advanced Computer Science, Universiteit Leiden*
[b] *Mathematical Institute, Universiteit Leiden*

Abstract

We examine the cooperative card game HANABI. Players can only see the cards of the other players, but not their own. Using hints partial information can be revealed. We show some combinatorial properties, and develop AI (Artificial Intelligence) players that use rule-based and Monte Carlo methods.

1 Introduction

The game of HANABI, meaning "fire flower" or "fireworks" in Japanese, is a cooperative card game which requires the players to combine efforts in order to achieve the highest possible score. Designed by Antoine Bauza in 2011 and published by R & R Games [7] (see Figure 1), among others, it is designated as a game in the categories "cooperative play" and "hand management" by the game analysis site BOARDGAMEGEEK [3]. The goal of the game is simple: play out several sequences of cards in the right order. The catch, however, is that the players can only see the cards in other player's hands and not their own; information has to be gathered by a system of hints that reveal partial information.

Figure 1: HANABI, as sold by R & R Games [7].

A game in the named categories which is somewhat similar in gameplay is THE GAME designed by Steffen Benndorf [3]. Though leading to interesting combinatorial and strategic questions, there is not much to be found in the literature on this kind of games. More is known about another game which bears similarity to Hanabi in view of its dealing with hints: BRIDGE; some issues here touch upon research as in [5]. For instance, even *not* giving a hint might reveal some information. Moreover, it is possible to use artificial hints having non-natural meaning, like conventions in BRIDGE. Yet another game which gives rise to theoretical problems comparable to those inspired by Hanabi, because of its similarity in nature, is SOLITAIRE; as an example, NP-completeness results for HANABI can be found in [1].

This paper examines some interesting properties of the game, but is far from being complete. In fact, we show complicated mathematical behavior for a one-player version without hints, and provide

an exploratory examination of some artificial players (cf. [6, 2]), including rule-based and Monte Carlo versions.

We start with a comprehensive explanation of the rules of HANABI in Section 2. We then deal with two main questions. In Section 3, we consider the playability of a game of HANABI: given a certain start configuration, is it possible to obtain a maximal score when playing perfectly? We will address some theoretical issues. Next, in Section 4, we look for a good strategy for any arbitrary game. Among others, we consider Monte Carlo methods which seem promising. In Section 5, we conclude with a summary of the given results as well as some interesting open questions.

2 Game Rules

The classic game of HANABI is played with a stack of $N = 50$ cards. Every card has one out of $C = 5$ colors — blue (B), red (R), green (G), yellow (Y) or white (W) — and a value between 1 and $k = 5$. In the classic game, for every color there are three 1s, two 2s, 3s and 4s and one 5, hence fifty cards in total. At the start of the game, the stack is shuffled and a hand of cards is dealt to each player: $R = 5$ cards are given to every player if $P = 2$ or $P = 3$ persons are playing, and $R = 4$ cards are dealt in games with $P = 4$ or $P = 5$ players. Now, every player picks up their hand in such a way that the other players can see them, but they themselves cannot. The rest of the cards forms the face-down stack.

The goal of the game is to create C stacks of cards going from 1 through k on the table, one of each color. To do so, players take turns, choosing exactly one of the three following actions every turn:

- give a hint,
- discard a card,
- play a card.

At the start of the game, a pool of $H = 8$ hint tokens is available to the players. To give a hint to another player, one token must be removed from this pool. If there are no tokens left, this action cannot be chosen. Giving a hint is done by either pointing out all cards (perhaps zero) of a certain color or all cards of a certain value in the hand of one other player. This is explained most easily by considering an example hand like $(\mathrm{R3}, \mathrm{W1}, \mathrm{B1}, \mathrm{R4}, \mathrm{B5})$. A hint may be expended to point out the position of the W1 and B1, telling the player that *these two cards are 1s*. One might also point out the R3 and R4 by telling that *these cards are red*. Also, pointing out B5 by saying *this is a 5* is fine. However, one cannot point out B5 by telling that *this card is blue*, because B1 must then also be pointed at. One is allowed to tell the player that *there are no green cards in the hand*, as this hint effectively points out all (zero) green cards.

The second possible action is to discard a card: the active player takes a card from his/her hand (without first looking at it) and announces that it will be discarded. It is then put face-up in the discard pile, which may be viewed by all players at any point in the game. Once in the discard pile, a card will never re-enter the game. A new card is now taken from the face-down stack to replenish the hand and as an added bonus, a hint token is added to the available hints pool. Therefore, if there are already H hints available, this action cannot be chosen.

Finally, one may choose to try to play a card much in the same way as discarding a card. However, the players now look whether the card can be added to one of the stacks on the table. This is again best clarified by an example. Suppose the following stacks are on the table:

R1	G1	W1
R2		W2
		W3

It is now possible to play a R3, which can be appended to the leftmost stack. Similarly, a G2 or W4 would be fine. It is also possible to play a Y1 or B1, starting a new stack. However, one may for example not play a R4 (as a R3 is first needed) or R2 (as this is already on the table). Moreover, another R1, G1 or W1 cannot be played as there may only be one stack of every color on the table at any time. If a card is successfully played, it is put in the fitting position on the table. If a card turns out not to be playable, it is moved to the discard pile and the players score one error. In either case, the active player draws a new card from the stack, but no hint token is added.

In order to show that there is more to a hint than just its plain meaning, consider the following example. If a player has only one 2, the G2, a hint pointing at this card might in this situation be considered containing the message "come on, play this card".

The game ends when one of three conditions is met. First, if $E = 3$ errors are made by trying to play a card, the game ends with the lowest possible score of 0. Second, if the C-th stack is completed by playing a card, i.e., if there is a sequence of 1 through k of every color on the table, the game ends with the highest possible score of $C \cdot k$ (for the classic game: 25). Finally, when the last card of the face-down stack is drawn, every player gets one more turn, including the player that drew the last card. The achieved score is then determined by adding the highest number of every stack on the table. In the previous example, a score of $2 + 1 + 3 = 6$ points would be obtained. Apart from just aiming at the highest score, one might also focus on the 0–1 target of obtaining the maximal score.

Naturally, the rules for classic HANABI can easily be generalized and extended. One can for example alter the parameters mentioned above, see Section 3. A more formal approach can be found in [2].

3 Playability

In this section, we will address the first main question: given an initial configuration of a game of HANABI, is it possible to obtain the maximum score if playing perfectly? An initial configuration of a game of HANABI (or simply a game of HANABI) is called *playable* if the maximum score can be achieved. We first turn to a theoretical approach involving combinatorics; a perhaps more practical approach using dynamic programming is presented in [1] and also in [2].

As in [1], we consider the simplified situation where the players can also see their own cards (making the hints system and errors obsolete). We give a result for the one-player version, with $R = 1$, called SINGLEHANABI: a player must immediately play or discard the newly received card. We consider only one color, but the number of cards of each value can be arbitrary.

Fix an integer $k \geq 1$. Let $x = (x_1, x_2, \ldots, x_k)$ be a vector with k non-negative integers $x_1, x_2, \ldots, x_k$, and let $F(x) = F(x_1, x_2, \ldots, x_k)$ denote the number of ordered sequences of length $N = N(x) = x_1 + x_2 + \ldots + x_k$ with x_1 occurrences of the integer 1, x_2 occurrences of the integer 2, …, x_k occurrences of the integer k, *without* a subsequence 1–2–…–k: the so-called *bad* sequences; the others are called *good*. Note that a subsequence is not necessarily consecutive; e.g., the sequence 1–3–2–3 has 1–3–2 and 1–2–3 as subsequence, but not 2–3–1. So $F(x_1, x_2, \ldots, x_k)$ is the number of unplayable SINGLEHANABI games, having x_i cards of value i ($1 \leq i \leq k$): there is no subsequence 1–2–…–k that would allow the player, who has "no memory", to immediately play these k "cards" in order.

If for some i with $1 \leq i \leq k$ we have $x_i = 0$, then we know that $F(x) = \binom{N}{x_1, x_2, \ldots, x_k, N - \sum_{j=1}^{k} x_j}$, since all sequences are bad in that case.

We also note that F is symmetric in its arguments. Indeed, we construct a bijection between good sequences with interchanged numbers of, e.g., 1s and 2s: $x_1 \leftrightarrow x_2$. To do this, view every good sequence in parts: the part up to but not including the first 1, the part between this 1 and the next 2, etc. Now, the required bijection is given by swapping the first and second part and changing all 1s in the parts to 2s and vice versa. We clarify this by an example, in which the numbers which divide the parts are shown in bold:

$$24\mathbf{1}31341313\mathbf{2}1\mathbf{3}32\mathbf{4}21421 \leftrightarrow 32342323\mathbf{1}14\mathbf{2}2\mathbf{3}31\mathbf{4}12412$$

Note that the resulting sequence contains the proper amount of every number. Moreover, applying the proposed bijection twice results in the original sequence, which shows that it is indeed bijective.

Inspired by a discussion by anonymous contributors at STACKEXCHANGE [9], our main result in this section is:

Theorem We have

$$F(x) = \sum_{\substack{y \prec x \\ |y| \leq k-2}} a(N, y)$$

Here we denote $y \prec x$ if the ordered sequence y with k non-negative integers satisfies $y_i \leq x_i$ for all i with $1 \leq i \leq k$. Furthermore, $|y|$ denotes the number of non-zero elements in y. We put $a(n, y) = (-1)^{|y|} \binom{n}{y}^{*} (k - |y| - 1)^{n - s(y)}$. The multinomial coefficient $\binom{n}{y}^{*}$ is defined as follows: $\binom{n}{y}^{*} = \binom{n}{y_1 \ominus 1, y_2 \ominus 1, \ldots, y_k \ominus 1, n - s(y)} = n!/((y_1 \ominus 1)!(y_2 \ominus 1)! \ldots (y_k \ominus 1)!(n - s(y))!)$ (the bottom last term

in the multinomial coefficient, here $n - s(y)$, is often omitted by convention) with $s(y) = \sum_{i=1}^{k}(y_i \ominus 1)$, where we used $t \ominus 1 = \max(t-1, 0)$.

The equation can also be written as

$$F(x) = \sum_{\ell=0}^{k-2}(-1)^{\ell} \sum_{\substack{y \prec x \\ |y|=\ell}} \binom{N}{y}^{*} (k-\ell-1)^{N-s(y)}$$

For example, with $k = 3$ and $\ell = |y| = 1$, in the computation of $F(2,3,1)$ we encounter sequences $(1,0,0)$, $(2,0,0)$, $(0,1,0)$, $(0,2,0)$, $(0,3,0)$ and $(0,1,0)$; and, e.g., $\binom{6}{(0,2,0)}^{*} = \binom{6}{0,1,0,5} = \binom{6}{1,5} = \binom{6}{1} = 6$. The term with $\ell = |y| = 0$ equals $(k-1)^N$. Note that $F(x_1, x_2)$ evaluates to 1, as expected. By the way, $F(x_1)$ is 0, if $x_1 > 0$.

Proof The theorem can be proven through the obvious recurrence (for $x_1 > 0$)

$$\begin{aligned} F(x_1, x_2, \ldots, x_k) &= \binom{N-1}{x_1-1} F(x_2, \ldots, x_k) + \\ &\quad + F(x_1, x_2-1, \ldots, x_k) + \ldots + F(x_1, x_2, \ldots, x_k - 1) \end{aligned}$$

where we interpret a term as 0 if one of its arguments is negative. The respective terms count bad sequences that start with a 1, with a 2, …, with a k.

We first note that $F(1, \overbrace{0, \ldots, 0}^{k-1 \text{ times}})$ equals 0 if $k = 1$, and equals 1 if $k > 1$. And $F(0, 0, \ldots, 0) = 1$. This is the basis for an inductive proof, with respect to $\langle k, N \rangle$.

Using the symmetry of F in its arguments, we may assume that $x_1 > 0$. If $x_2 > 0$ we compute (analogous for the other terms):

$$F(x_1, x_2-1, \ldots, x_k) = \sum_{\substack{y \prec x;\, y_2 < x_2 \\ |y| \le k-2}} a(N-1, y) = \sum_{\substack{y \prec x;\, y_2 = 0 \\ |y| \le k-2}} a(N-1, y) + \sum_{\substack{y \prec x;\, y_2 \neq 0 \\ |y| \le k-2}} a(N, y) \frac{y_2 - 1}{N}$$

Now we look at a fixed y, and combine all contributions of the $k-1$ terms. If $y_1 = 0$ we arrive at $a(N-1, y)(k - |y| - 1) + a(N, y)s(y)/N = a(N, y)$. However, if y has non-zero y_1, we have to be more careful. We then still arrive at $a(N, y)$, but now with an additional term $a(N-1, y) - a(N, y)(y_1 - 1)/N$. If we let y_1 increase from 1 to x_1 (where we keep the other elements from y unchanged), these terms telescope to $a(N-1, y')$ with $y'_1 = x_1$ and $y'_i = y_i$ $(1 < i \le k)$. We rewrite

$$\binom{N-1}{x_1-1} F(x_2, \ldots, x_k) = -\sum_{\substack{y \prec x;\, y_1 = x_1 \\ |y| \le k-2}} a(N-1, y)$$

which exactly cancels the remaining terms.

If we happen to have $x_2 = 0$, the argument above remains valid. Indeed,

$$\sum_{\substack{y \prec x \\ |y| \le k-2}} a(N-1, y) = \sum_{\substack{y \prec x;\, y_1 < x_1 \\ |y| \le k-2}} a(N-1, y) + \sum_{\substack{y \prec x;\, y_1 = x_1 \\ |y| \le k-2}} a(N-1, y)$$

is equal to 0, since the first term from the right hand side in that case equals

$$F(x_1 - 1, 0, x_3, \ldots, x_k) = \binom{N-1}{x_1 - 1, x_3, \ldots, x_k}$$

whereas the second equals

$$-\binom{N-1}{x_1-1} F(0, x_3, \ldots, x_k) = -\binom{N-1}{x_1-1}\binom{N-x_1}{x_3, \ldots, x_k} = -\binom{N-1}{x_1 - 1, x_3, \ldots, x_k}$$

thereby completing the proof. □

One consequence is that

$$F(\overbrace{1,1,\ldots,1}^{k \text{ times}}) = k! - 1 = \sum_{\ell=0}^{k-2} (-1)^\ell \binom{k}{\ell} (k-\ell-1)^k$$

which happens to be a special case of a formula from [8]. In this same category, another special case is

$$F(\overbrace{4,4,\ldots,4}^{13 \text{ times}}) = \sum_{\ell=0}^{11} (-1)^\ell \binom{13}{\ell} \sum_{i_1=0}^{3} \cdots \sum_{i_\ell=0}^{3} \binom{52}{i_1,\ldots,i_\ell} (12-\ell)^{52-i_1-\ldots-i_\ell}$$

which evaluates to 91973270026324484565579418350194489655582912237019 (a prime number approximately equal to $9 \cdot 10^{49}$) using a straightforward Python program, that takes a few seconds. Therefore, the probability that a shuffled deck of standard playing cards contains a full increasing subsequence (from ace to king, disregarding suits) turns out to be equal to $1 - F(4,4,\ldots,4)/(52!/4!^{13}) \approx 0.000554$, which is a folklore result.

4 Strategies

We will now consider the second main question: what is a good strategy? In this section, a *strategy* is a means to determine which action to take in any given state of the game. We will look at games with $P = 3$ players in which every player has a hand size of $R = 5$ as in the classic game rules in Section 2. We call a card *useful* if it can be appended to one of the stacks on the table at the current moment, and otherwise *useless* — the latter term perhaps being somewhat preliminary.

In [6], strategies are considered in which the players try to estimate their hand by analyzing actions of other players. This analysis is for the two-player version, and relies heavily on online learning. The author claims an average result of 15.85 points. Here, we will try two other types of strategies, and explore their potential. The first one implements several rules of thumb that tend to come up quickly in human play: a rule-based strategy. The second is an implementation of a basic Monte Carlo strategy. Note that it is hard to determine the maximal score that can be achieved for any given game, see also [1].

For the first strategy, every player acts according to the following preset rules:

1. If there is a card in my hand of which I am "certain enough" that it can be played, I play it.
2. Otherwise, if there is a card in my hand of which I am "certain enough" that it is useless, I discard it.
3. Otherwise, if there is a hint token available, I give a hint.
4. Otherwise, I discard a card.

In this framework, there are several parameters to be determined. First, we may choose the definition of "certain enough" in steps 1 and 2: we let $\omega_\mathrm{p}, \omega_\mathrm{d} \in [0,1]$ be the thresholds above which a player knowing that the probability of a card being useful, resp. useless, exceeds the threshold, it is played, resp. discarded, in step 1, resp. 2. Moreover, one can choose whether or not to take any risk to end the game on three errors by prohibiting to play cards which are not certainly useful after having made two errors; this is referred to as "safe play".

In step 3, we can follow different guidelines which determine the hint to be given. We consider four of them: #1) random; #2) giving a hint that gives information on the largest number of cards; #3) giving a hint on the next useful card in sight or on the largest number of cards if no useful card is seen; or #4) giving a hint on the next useful card or on the next useless card if no useful card is available or otherwise on the largest number of cards.

Finally, in step 4, we choose from four different rules by which the card to be discarded is chosen. There are: #1) random; #2) discarding the card of which it is most certain that it is useless; #3) discarding the card which has been stored in hand the longest; or #4) discarding the card of which it is most certain that it is not absolutely necessary to complete all stacks.

In addition to these choices, we also explore whether it is profitable to sometimes swap the order of steps 3 and 4. We let $\omega_\mathrm{h} \in [0,1]$ be the probability that this is done during a turn.

To test the various settings of the parameters, we let three players using the same strategy play 10,000 different starting configurations. Every configuration is played ten times in order to account for the randomness in the strategies. A test run of this kind takes approximately one minute on a computer with an Intel i7 2.3GHz core and 6 GB of RAM.

First, we take $\omega_{\mathrm{d}} = \omega_{\mathrm{h}} = 1.0$, and vary ω_{p} in $\{0.5, 0.6, 0.7\}$. The average scores for all sixteen combinations of hint and discard rules are shown in Table 1. It is apparent that the third hint rule, giving a hint on the next useful card or otherwise on the most cards, is dominant for this setting of the parameters. Surprisingly, discard rules #1 and #2 are about as effective, meaning that discarding randomly is competitive with discarding the card of which we most think it is useless.

		ω_{p}	Hint rule			
			#1	#2	#3	#4
Discard rule	#1	0.7	6.3	13.5	14.5	11.9
		0.6	7.1	13.9	15.3	12.7
		0.5	7.3	13.8	15.2	12.5
	#2	0.7	6.6	13.8	14.5	12.0
		0.6	7.4	14.4	**15.4**	12.7
		0.5	7.5	14.2	15.3	12.6
	#3	0.7	5.8	13.1	14.1	11.5
		0.6	6.6	13.5	14.9	12.3
		0.5	6.9	13.4	14.8	12.3
	#4	0.7	5.8	13.0	13.9	11.5
		0.6	6.7	13.7	14.8	12.3
		0.5	6.8	13.5	14.7	12.2

Table 1: Scores obtained with different rules and varying ω_{p}, for $\omega_{\mathrm{d}} = \omega_{\mathrm{h}} = 1.0$, playing safely.

Now, we vary the other parameters. For a full overview of the results, see [2]. It turns out that for the combination of hint rule #3 and discard rule #2, the best results can be obtained. In Figure 2 the results for varying ω-thresholds can be found, playing safely: no more risk is taken when playing cards after two errors have been made. We see (also in Table 1) that the best average score obtained is 15.4 when taking $\omega_{\mathrm{p}} = 0.6$, $\omega_{\mathrm{d}} = \omega_{\mathrm{h}} = 1.0$ and playing safely. Apparently, it is profitable to try and play a card once we are 60% sure that it will be correct (unless we have already made two errors in which case we require certainty), and only discard a card if it is certain that it is useless and give a hint if possible otherwise.

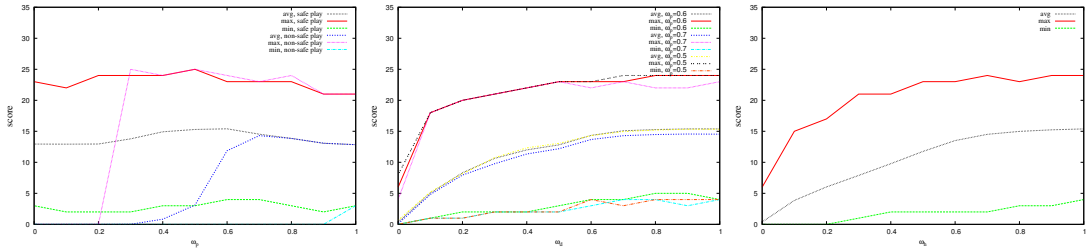


Figure 2: Scores obtained using hint rule #3 and discard rule #2, for varying ω-thresholds.

For the Monte Carlo strategy, the basis is well-known. In every turn, we try every action after which the game is played out by random players many times. Each of these random games is evaluated in some way, after which we choose to do the action which led to the best score.

In the implementation for HANABI, special care has to be taken in at least two situations. First, note that by trying to play a card and seeing how this turns out, a player could illegitimately obtain information on this card. Therefore, when trying to play a card in the Monte Carlo phase, the hand of the active player is shuffled through the deck and a new hand is dealt which is consistent with all hint information obtained so far. This way, the factual information on the cards is stored without allowing the agent to cheat. However, information on the exact hints that were given and the time at which these were given

is lost, which possibly results in the agent not picking up implied hints, e.g., a hint pointing out a card actually meaning that it can successfully be played. This is in sharp contrast with the inner working of the strategies as described in [6]. Still, much progress could be booked here by better judging or even learning the probability distribution of the cards in the hand of the active player, incorporating the hints provided so far.

Second, a truly random player will end the game on three errors with high probability. To circumvent this, we choose to not end the game after three errors have been made in the play-out phase. Moreover, we prohibit the random player from playing a card of which it is certain that it cannot be played based on the hints received so far. If the random player knows that none of the cards in his/her hand can be played, he/she will always randomly discard a card or give a hint. If he/she knows that only some of his/her cards are useless, he/she may also randomly try to play one of the other cards.

Contrary to the standard implementation of Monte Carlo, we do not evaluate each of the play-outs on the final score obtained. Instead, we register for the next D turns the amount of new points obtained as well as the amount of errors made with respect to the current turn in each play-out. For each new point we administer a $+1$ and for the k-th error we administer a score of $-k$. We then compute a weighted sum over the scores for each of the D turns, counting later turns with (linear) higher weight, after which we pick the action with the highest average score among the play-outs. Note that taking the value $D = 1$ would result in greedy play.

Furthermore, it turns out to be profitable to not let the random player choose an action uniformly at random (indeed, such a player performs hardly better than a random player). If the random player decides to play a card, in the style of MCTS [4], he/she favors playing a card on which he/she has more information over playing a card on which he/she has less information. It seems a challenge not to incorporate too much game knowledge, certainly if it can be avoided.

In order to test the performance of the Monte Carlo player, we let three players using this strategy play 100 different starting configurations. A test run of this kind takes approximately thirty minutes on a computer with an Intel i7 2.3GHz core and 6 GB of RAM, when using 500 play-outs.

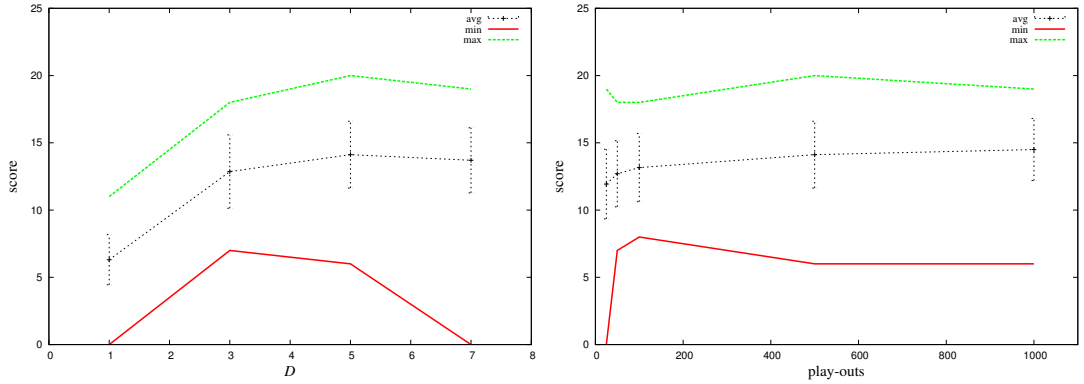


Figure 3: Scores obtained using 500 play-outs and varying D, and using $D = 5$ and varying number of play-outs, both with standard deviation.

In the experiments, the results of which can be seen in Figure 3, we have varied the amount of turns D that the Monte Carlo player takes into account and the number of play-outs allowed per different action. Surprisingly, the value of D taken, when not being too small, does not fundamentally affect the obtained score. Furthermore, while raising the amount of allowed play-outs does seems to somewhat improve the average score obtained, the differences are relatively small.

In fact, we note that it is not unlikely that these differences are to be contributed for a great part to random chance. Indeed, the standard deviation in the scores obtained seems to be quite high, with scores as low as 6 (or sometimes even 0) and as high as 20 being observed with an average around 14 points; the highest average score obtained is 14.5, with 1000 play-outs. We also see this large range in the scores obtained by the rule-based strategy discussed earlier in this section. It appears likely that the result of a game is heavily dependent on the starting configuration, and it could be worthwhile for the algorithm(s) to somehow adapt to this. As discussed in the previous section and in [1], it is sometimes impossible to

score a perfect game with 25 points, in which case it may already be difficult to obtain a final score of more than 15. An easy example is a game where many 1s and/or 2s arrives late in the deck.

A provisional conclusion could be that, with the current settings, the rule-based players slightly outperform the Monte Carlo players, considering the one-point difference between the average scores obtained. The results from [6] for the two-player version, using online learning, even seem to be a little better than for the three-player version, although it could be that the two-player version allows for higher final scores.

5 Conclusions and Further Research

In this paper we have examined some interesting aspects of the cooperative card game HANABI, that has the special property that players can only see the cards in the hands of the other players — but not in their own. Hints provide partial information about the game states.

Even simplified versions of the game, like SINGLEHANABI, give rise to complicated combinatorial questions and corresponding formulas. Furthermore, we have shown that different game playing strategies offer promising results. In particular, Monte Carlo techniques require little game knowledge, but are capable of delivering high quality competitive players. One issue is the problem of the amount of information that is used during the play-outs. The rule-based players perform a little bit better, at the cost of incorporating more game knowledge.

As further research we first mention the quest for a proof of the result from Section 3 using the principle of inclusion and exclusion, perhaps also providing links to more general theorems. It is also of interest to find (substantial) classes of games that are unplayable. For the Monte Carlo players, we like to combine the presented method with techniques as mentioned in [4], like UCT, meanwhile somehow learning knowledge. And finally, there is much potential in research regarding the hints system.

Acknowledgements

The authors would like to thank Anne Hommelberg and Hendrik Jan Hoogeboom, as well as the anonymous contributors from [9].

References

[1] J.-F. Baffier, M.-K. Chiu, Y. Diez, M. Korman, V. Mitsou, A. van Renssen, M. Roeloffzen and Y. Uno, Hanabi is NP-complete, even for cheaters who look at their cards, in: Proceedings of 8th International Conference on Fun with Algorithms (FUN 2016), Leibniz International Proceedings in Informatics (LIPIcs) 49, pp. 4:1–4:17, 2016.

[2] M.J.H. van den Bergh, Hanabi, a cooperative game of fireworks, Bachelor thesis, Leiden University, www.math.leidenuniv.nl/scripties/BSC-vandenBergh.pdf, 2015. 1899848

[3] BoardGameGeek, website, www.boardgamegeek.com [retrieved 24.8.2016].

[4] C. Browne, E. Powley, D. Whitehouse, S. Lucas, P.I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis and S. Colton, A survey of Monte Carlo Tree Search methods, IEEE Transactions on Computational Intelligence and AI in Games 4, 1–43, 2012.

[5] H. van Ditmarsch and B. Kooi, One hundred prisoners and a light bulb, Springer, 2015.

[6] H. Osawa, Solving Hanabi: Estimating hands by opponent's actions in cooperative game with incomplete information, in: Proceedings of the Workshop at the Twenty-Ninth AAAI Conference on Artificial Intelligence: Computer Poker and Imperfect Information, pp. 37–43, 2015.

[7] R & R Games, website, www.rnrgames.com [retrieved 24.8.2016].

[8] S. Ruiz, An algebraic identity leading to Wilson's theorem, The Mathematical Gazette 80, 579–582, 1996.

[9] StackExchange, website, http://math.stackexchange.com/questions/1215764/sequences-that-contain-subsequence-1-2-3 [retrieved 24.8.2016].

A Domain-Independent Method for Entity Resolution by determining Textual Similarities with a Support Vector Machine

Kerstin Borggrewe Jan C. Scholtes

Department of Data Science and Knowledge Engineering, Maastricht University

Abstract

As digital text collections continue to increase in size and number, their importance to scientific research in fields such as medicine, humanities and law continues to grow as well. These text collections originate from various different sources, using different terminology, name variations and various abbreviations for the same semantic categories. This causes problems when semantic information is extracted or when these collections are used for community detection, clustering or other analysis. Mapping different textual variations to their corresponding real-world entities is called Entity Resolution. An important part of this process is entity matching, which calculates a measure of textual similarities between the entities. In this research several well-known similarity measures are combined in a voting algorithm, using a Support Vector Machine (SVM). The quality of the matching is evaluated against several tagged data-collections in order to create a domain independent solution for the text matching problem. As will be shown, this SVM approach outperforms all individual text-matching methods. To evaluate the method in a practical application, it is used to resolve entities before a typical clustering task is applied. The clustering task is evaluated on a subset of the real world Enron email data-set. As can be seen, the clustering and subsequent visualization of the clusters benefits significantly from the reduction of number of entities from the raw textual occurrences to their significant lower number of real-world occurrences.

1 Introduction

The performance of many data-science algorithms applied on large text collections decreases due to the occurrence of morphologically and syntactically different variations of semantically identical entities. This harms research in ontology mapping [11] and extraction of Resource Description Framework (RDF) for the semantic web [6, 10]. To use identified entities effectively, they have to be normalized. This process is called entity name matching. Entity name matching can be considered to be a classification problem: a string pair is a match when describing the same real world object. When not, it is labeled as a non-match. A match can be defined by measuring textual similarities. To be able to compare strings, various string distance metrics have been defined over the years. However, certain string distance metrics seem to work better on certain types of entities. For instance, Jaro-Winkler works well on company names, but not very well on personal names. Term Frequency – Inverse Document Frequency (TF-IDF) based methods seem to work better on personal names, but less well on company names. This is because different string matching algorithms take different characteristics of strings in consideration when calculating the distances. This behavior justifies research which combines different string distance metrics in a voting algorithm. In this research we will evaluate the most common individual string distance metrics on several published and pre-tagged data collections. Hereafter, we will train and evaluate the performance of a Support Vector Machine (SVM)-based voting algorithm. In order to show the real-world benefit of this method, we will first normalize entities derived from the ENRON data set and then cluster and visualize the differences between the original and the normalized data. This will result in a significant lower number of duplicate references to real-world entities and improves the quality of the clustering and the visualizations.

2 Related Work

The task of entity resolution is about determining which entity names are co-referent and represent the same real-world entity. To build a machine-learning-based entity-resolution engine, the common approach is to define a set of features used by a classification algorithm to separate matches from non-matches. Soon *et al.* [12] considered a combination of sentence structure based features and string similarities with more advanced high-level features. However, they show that a set of basic syntactic and lexical features performs as good as the combination of these with more advanced high-level features. From this work we can conclude that just concatenating more features does not necessarily improve the performance. In William Cohen *et al.* [2], several string distance metrics have been tested, measured and compared. No domain knowledge was taken into account. According to their research, the best performing individual approach was a hybrid distance metric on words named Soft-TF-IDF which we will discuss in more detail later. Gravano et al. [8] extend this approach by using n-grams as tokens instead of words. In order to combine different methods, they trained adaptive distances across multiple fields and domains of records. Bilenko and Mooney [1] trained a Support Vector Machine (SVM, [3]) using the numeric scores of the best performing distances per domain. In our research, we do not provide any information on the domain. All we provide is the entity name. We let the SVM decide when which (non-linear combination of) string distance measures is best to use. This will result in a general applicable text-matching solution for a variety of different entity types, without the need to define what type of semantic category or form an entity is.

3 String Distance Metrics

There are several types of string distances, that each measure different notions of (dis)similarity. In what follows, we will provide a short introduction to the most used methods.

As we seek to develop a generic and domain independent approach, we do not look at rules-based approaches that take specific domain knowledge into consideration. The only assumption we make is that the language is English and thus make use of an English Phonex method to measure phonetic aspects of string similarity.

Edit-based techniques compare two strings based on individual characters and are able to detect the most common typing and spelling errors. Examples of such metrics are the Hamming distance, the Jaro and the Jaro-Winkler metrics. Other variations are the Levenshtein and Damerau distances, which both measure the minimal required number of edits to transform one string into the other. Levenshtein uses unweighted edits, Damerau weights edits. Edit-based distances are very sensitive to word order. This causes problems when matching entities that consist of multiple tokens such as person names or more complex company or organizational names.

Therefore, when using entities that consist of multiple tokens, a token-based distance which is insensitive to word order is a better choice. Token-based methods split the strings into short sub-strings. When an entity consists of multiple tokens, their similarity can be measured with similarity functions like Jaccard or the Cosine distance. The last one can be combined with another widely used token-based weighting scheme: TF-IDF. In this case, each term of a string's token set is assigned a TF-IDF weight. As a result, the most distinguishing tokens in a collection will be prevailed. With the Cosine similarity, token sets can be ranked to measure the similarity between the strings represented by the TF-IDF weighting vectors.

For string matching, a hybrid approach that is often used, is called Soft-TF-IDF. First, the string's tokens are separated and assigned with TF-IDF weights. Then all tokens of the one string S are compared with the tokens of the other string T by using the inner similarity function $sim(w, v)$, where $w \epsilon S$ and $v \epsilon T$. Nearly matching tokens or exact matches are exploited by considering token pairs which have a similarity value above a certain threshold (see token pair set $CLOSE$). If there are several tokens $v_i \epsilon T$ that are exact or near matches to a token $w \epsilon S$, only the $\max_{sim(w,v_i)>0}$value pair is considered. The mapping of a token pair to a similarity value is described as the relation $SIM(w, v)$. When applying the Cosine similarity to $\vec{a}$, the token weight vector of string S, and $\vec{b}$, T's TF-IDF vector, the product of weights for non-identical, nearly matching tokens is multiplied by the similarity value, which is less

than one.

$$
\begin{aligned}
sim(w,v) &= similarity\, between\, w \epsilon S\, and\, v \epsilon T \\
CLOSE &= \{\,(w,v)\,|\,w \epsilon S \wedge v \epsilon T \wedge sim(w,v) > threshold\,\} \\
SIM(w,v) &= \begin{cases} sim(w,v) & if\ sim(w,v) = \max_C sim(k,j), \\ & C = \{\,(k,j)\,|\,k = w \wedge j \epsilon T \wedge (k,j) \epsilon CLOSE\,\} \\ 0 & otherwise \end{cases} \\
a_w, b_v &= weight\, of\, token\, w \epsilon S,\ weigh\, of\, token\, v \epsilon T \\
N &= number\, of\, all\, tokens\, in\, the\, collection \\
soft.TF-IDF.sim(\overrightarrow{a}, \overrightarrow{b}) &= \frac{\sum_{(w,v) \epsilon CLOSE} a_w \cdot b_v \cdot SIM(w,v)}{\sqrt{\sum_{i=1}^{|N|} a_i^2} \cdot \sqrt{\sum_{i=1}^{|N|} b_i^2}}
\end{aligned}
$$

4 Corpora used

SecondString Data Sets

There are no widely accepted benchmark data-sets for evaluating entity-name matching-algorithms, but there are a few. For instance, the data sets provided by the SecondString project of Cohen [13], include different domains. There are a number of problems with these data sets that we had to overcome: the data sets are relatively small compared to real world data sets. We also found that the automatically labeled data-sets included significant inconsistencies. For instance: ten to fifteen percent of the records included in the *business* data set and even more in the *park* data set are labeled to be a non-match when they clearly should actually match. Other data sets used were manually created and thus create a more reliable ground truth. In addition, most of the data sets are unbalanced.

In order to overcome these problems, we created a balanced data set by combining records of the *ucdPeopleMatch*, *censusTextSegmented*, *business*, *parks*, *vaUniv* and *animal* data sets, leaving out the obvious wrong records from the *business* and the *park* data set. This data set represents different domains, measures different characteristics of the strings and also contains sufficient records. We will call this set the *Multi-Domain* set. To measure the performance on some balanced single domain data sets, the individual *animal*, *vaUniv*, *business* and *censusTextSegmented* were used.

Enron

A widely known real world data set is the Enron corpus. It is a large email collection from employees of the energy company Enron. The Federal Energy Regulatory Commission published the emails of 158 Enron employees as part of the Enron investigations. Several versions of the Enron corpus exists. We used the EDRM Enron Email Data Set [4]. It includes 168 PST files, which are 40.7 GB of digital information in total and 767.806 PST extracted messages. From the Enron set a few thousand records are sampled. Out of these 65 person and 40 company names are received. It was made sure that several duplicates were included. Also a set containing 500 person and 500 company names set is created.

5 Experimental Approach

First, the best string distance metrics are determined for each of the single domain sets: *ucdPeopleMatch*, *censusTextSegmented*, *business*, *parks*, *vaUniv* and *animal* data sets and for the *Multi-Domain* set. Next, the voting algorithm is implemented (python Scikit learn package) by using a Support Vector Machine (SVM). The SVM input consisted of the output values of best performing individual string matching methods are used (see table 1):

Data Sets	Selected Distances for SVM Input
multi-domain data set	Jaro-Winkler, n-gram, abbreviation handling, Soft-TF-IDF, Phonex, hybrid
single domain data sets	Soft-TF-IDF, hybrid, Jaro-Winkler

Table 1: used string distances for the SVM's input

An SVM is more effective when the attribute values have the same range, so the feature weighting is normalized to values between zero and one. Because the combination of string distances isn't necessarily a linear problem, a SVM with a Radial Basis Function (RBF) kernel was used as well as a linear SVM to compare the performance differences between them. Grid search was used to find out the best hyper-parameters for the SVM algorithm. 3-fold-cross validation was used (unfortunately some of the the data sets were too small for the more commonly used 10-fold cross validation). The data sets were split randomly in 30% and 70%. To be able to compare the performance of the voting SVM to the individual sting distance metrics, the matching confidence-value returned by the SVM's was used as the final distance measure. So, evaluation is done by comparing distance-based string-pair rankings. The next step is to use the algorithm and the trained SVM for a more practical application. An example of such a practical applications is normalizing a data collection of entity names before clustering or visualization. In order to test the model, a typical cluster algorithm is implemented which uses the confidence of the voting SVM's match class as cluster distance. The clustering using the combined string distance metrics from the SVM should perform better than clustering based on individual string distances.

6 Results

6.1 Voting Algorithm versus Individual Distances applied to a Single-Domain Matching Task

The results of the best individual string matching metrics can be seen in the most left columns of table 2.Next, the SVM was trained with the *Multi-Domain* data set and tested against the the individual *animal*, *vaUniv*, *business* and *censusTextSegmented* domain sets. The outcome of these experiments were then compared to the results of the best individual string distance metrics. These results can be found in the right columns of table 2. The effort of training with the Multi-Domain data set and then matching it against the single domain data sets did not perform better than each of the individual string matching metrics. So, when it is possible to use domain dependencies, this will obviously lead to better results than using domain independent data.

	Best String Matching Distance		SVM Trained with Multi-Domain Set and Tested against Single Domain Sets		
Name of Best Matching String Distance	*Avg Precision*	*Max F1*	*Avg Precision*	*Max F1*	*data set*
Soft-TF-IDF combined with longest common sequence	0.9832682	0.9303905	0.97987205	0.9255319	*vaUniv*
Hybrid approach combined with n-grams	0.9994171	0.9917628	0.999335	0.986711	*animal*
Soft-TF-IDF combined with n-grams	0.8068306	0.69916224	0.76029325	0.66703725	*business*
Jaro-Winkler	0.9801477	0.9355323	0.9817391	0.9432836	*census-TextSegmented*

Table 2: Performance of the SVM trained with the multi-domain set, evaluated against the Single Domain data set, compared to the best performing individual string matching distances

6.2 Voting Algorithm versus Individual Distances applied to a Multi-Domain Matching Task

As it is not always possible or practical to use domain dependencies, the SVM was also trained with - and tested against the *Multi-Domain* data set. In this case, we have also used the best performing individual string distance metrics against the *Multi-Domain* data set. As an additional variation, we used both a linear SVM as well as a Radial Basis Function (RBF) kernel-based SVM. The RBF-based SVM achieves a higher accuracy (88.47%) than the linear one (87.19%), so it was decided to use this one for the final experiments. As can be seen in table 3, the SVM with the RBF kernel outperformed all individual string matching distances for the Multi-Domain data set. .

String Distances Metrics	**Avg P**	**Max F1**
Soft-TF-IDF N-Gram	0.9116784	0.8351758
Hybrid N-Gram	0.91953653	0.8532527
N-Gram	0.92979455	0.8541368
SVM with RBF Kernel	0.9511016	0.875549

Table 3: Results of the string distances and the SVM, which are parameterized based on the training set, applied to the evaluation set

6.3 Combining Clustering with the Voting Algorithm

To evaluate the effect of the normalization on a clustering task, the DBSCAN and Affinity Propagation [7, 5] cluster algorithms were used on the original *SecondString* data sets *ucdPeopleMatch*, *censusTextSegmented*, *vaUniv*, *business* and *animal*. The results were then evaluated by computing the adjusted Rand indices provided by the scikit-learn python package [9]. A Rand index ranges from *minus one* to *plus one*. The value *plus one* represents a perfect clustering, negative values indicate a bad performance and zero is returned when cluster labels seem to be assigned randomly. By comparing the adjusted rand indexes of both DBSCAN and Affinity Propagation, it can be seen that DBSCAN performs better (see table 4). For this reason, we decided to use the DBSCAN clustering for the remainder of the experiments.

	animal	*business*	*vaUniv*	*censusTextSegmented*	*ucdPeopleMatch*	Avg
DBSCAN	0.7576	0.0001	0.5740	0.5341	0.7594	0.5250
Aff. Prop.	0.5555	0.0228	0.6246	0.4108	0.7373	0.4702

Table 4: Evaluation of the clustering: adjusted rand indexes per data set and cluster algorithm

The *business* data set is noticeably bad clustered. This is caused by the high label error rate in the original *business* data set (roughly fifteen percent of the records are labeled wrong). To determine that the clustering benefits from the SVM based data normalization, a DBSCAN using the SVM normalization and a DBSCAN using the N-Gram normalization are compared.The N-Gram based normalization was selected as it was the best performing individual string distance metric on the Multi-Domain data set. When applying the DBSCAN algorithm the eps parameter has to be set. Eps defines the maximum distance of neighboring samples: the lower the eps value the more similar the samples have to be to be called neighbors. Based on the density of a neighborhood DBSCAN creates clusters. For each eps value a clustering can be received from DBSCAN. A clustering can be evaluated like a ranking: string pairs that are labeled to be in the same cluster are identified as matches and the other ones as non-matches. Ranking the matches higher as non-matches makes it possible to compute the average precision and a maximum F1 score. The adjusted rand index is also computed. The DBSCAN-SVM performance's

optimum is reached with $eps = 0.01$, $AvgP = 0.7896$, $MaxF1 = 0.7721$ and $ARI = 0.45$ (see table 5). These values are never reached by the DBSCAN-N-Gram, so the SVM based method clearly outperforms the N-Gram one.

eps	AvgP	MaxF1	ARI	*eps*	AvgP	MaxF1	ARI
0.005	0.74777764	0.7130333	0.421424061982	0.005	0.577478	0.6666667	0.296266634279
0.01	0.7895747	0.77214485	0.45003736691	0.01	0.577478	0.6666667	0.296266634279
0.1	0.70621276	0.8349033	0.0197151111155	0.1	0.6735914	0.6666667	0.341255272681
0.15	0.60691273	0.82164145	0.00542474609333	0.15	0.7030967	0.6714801	0.231759674647

Table 5: Evaluation of the clustering with DBSCAN using the SVM's confidence as distance (on the left) and using the N-Gram distance (on the right)

6.4 Applying the Clustering to a Real World Data Set

Figure 1: Visualization of the Enron subset before (on the left) and after (on the right) normalization

In this experiment, the SVM-DBSCAN combination is used on the real-world Enron data set, just to show the results. No additional training was performed. These experiments were done with the SVM trained with the *Multi-Domain* data set. As no ground truth is available for the strings present in the Enron set, we cannot provide a detailed quantitative evaluation of the results. This experiment is therefor provided for demonstrative purposes only. The DBSCAN, which uses the SVM based distance combination, grouped the records of the Enron subset using the eps value of 0.005. The subset only includes two of the five trained domains, so the eps parameter has to be adjusted. The subset and the record's distances are visualized in figure 1 (see also figures 2 and 3 for more details). The nodes represent the records, which are connected to each other when the similarity is higher than or equal 0.5 and a node's size and color is related to the number of connections to other ones. The thickness of connecting edges is determined by the similarity value of the two connected records. It is noticeable that there are less notes than before but still clusters and relations recognizable. There are eleven records grouped wrong by the clustering when counting the ones that are labeled as wrong by at least fifty percent of the labeling people. The subset consists of 105 records and after normalizing it is reduced to 66 groups or entities. The accuracy averaged over the eighteen people is 0.8624. This is a satisfying result taking into account that the SVM based distance isn't trained on the Enron data set.

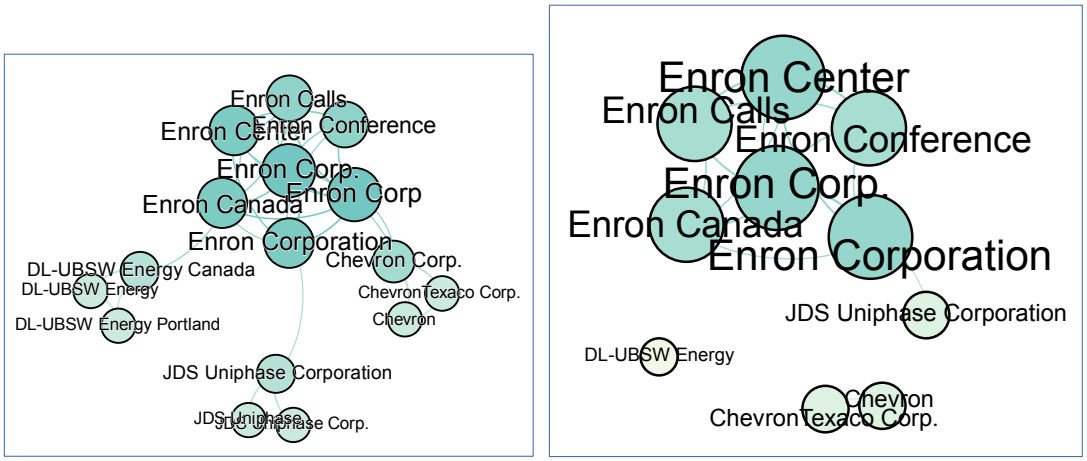


Figure 2: Enron Corporation, Chevron and DL-UBSW Energy

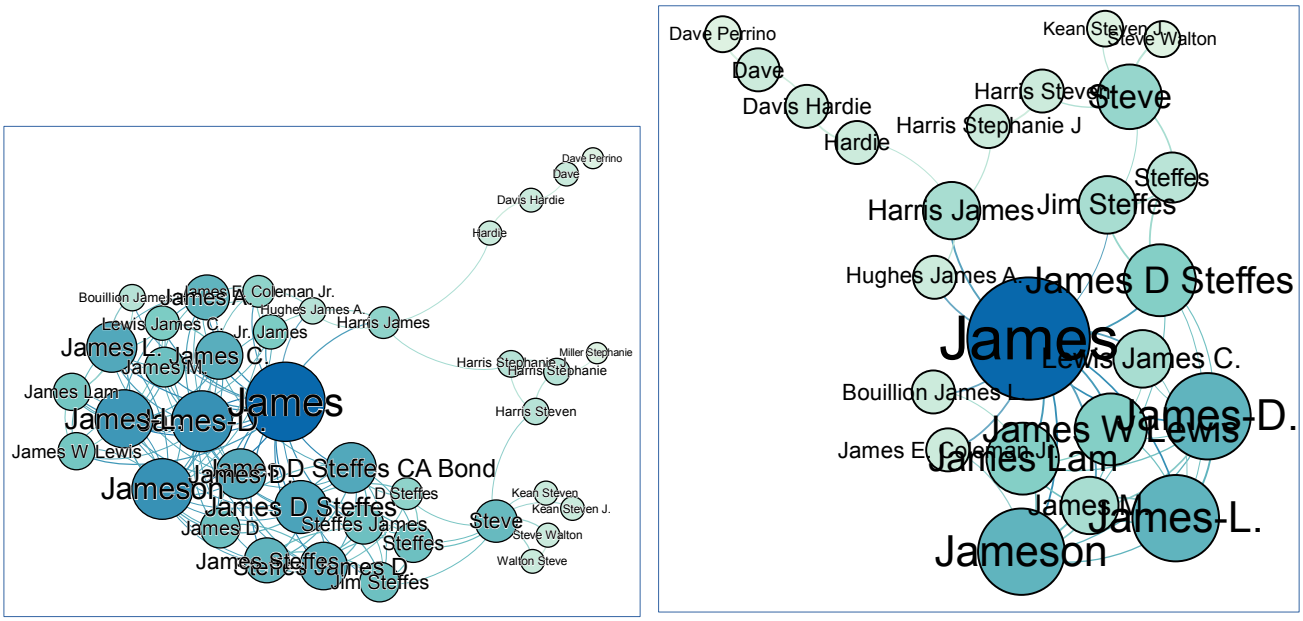


Figure 3: James

7 Conclusion and Further Research

In this research, it is shown that a combination of string distance metrics, which detect different notions of textual similarities, performs better than each of the individual string distance metrics for domain independent data sets. This method is a useful tool for domain-independent entity-resolution and reduces the need for user interaction. In order to show the benefit of the normalization on a clustering task, the SVM-based combined-distance approach can be combined with a clustering algorithm. It has been successfully applied to an entity-clustering task using the DBSCAN algorithm. An SVM-DBSCAN combination outperforms a tested DBSCAN-N-Gram combination on the *Multi-Domain* data set. TheN-gram distance is the best performing individual string distance on the *Multi-Domain* data set. This shows the advantage of the SVM-based distance approach when applied to a clustering task. Investigating the applicability of the combined distance approach to a real world data set and its scalability, it is applied to a subset of the Enron data set. Comparing the results of the clustering algorithm by human judges, it is shown that the SVM distance combination benefits entity clustering on real world data, even by using

the SVM trained with the *Multi-Domain* data set. Training it on annotated real-world data will probably improve the results a bit more.

During the exploration of the cluster algorithm's scalability the number of received clusters is highly reduced. It introduces the question, whether the algorithm's parameters can be optimized when a high number of records has to be normalized. An additional point of interest is the customization of the distance function used by the cluster algorithms: it would be interesting to investigate whether a hierarchical cluster algorithm combined with the trained distance achieves better results.

References

[1] Mikhail Bilenko and Raymond J. Mooney. Learning to combine trained distance metrics for duplicate detection in databases. Technical Report AI 02-296, Artificial Intelligence Lab, University of Texas at Austin, 2002.

[2] William W. Cohen, Pradeep Ravikumar, and Stephen E. Fienberg. A comparison of string distance metrics for name-matching tasks. In *Proceedings of the IJCAI-2003 Workshop on Information Integration on the Web*. American Association for Artificial Intelligence, 2003.

[3] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[4] EDRM. New edrm enron email data set. `http://www.edrm.net/resources/data-sets/edrm-enron-email-data-set`. Last visited Mar. 4, 2016.

[5] M. Ester, H. P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pages 226–231, 1996.

[6] Christiaan Fluit, Marta Sabou, and Frank Van Harmelen. Ontology-based information visualization: Towards semantic web applications. In *Visualizing the semantic web*, pages 45–58. Springer, 2006.

[7] Brendan J Frey and Delbert Dueck. Clustering by passing messages between data points. *Science*, 315(5814):972–976, 2007.

[8] Luis Gravano, Panagiotis G. Ipeirotis, Nick Koudas, and Divesh Srivastava. Text joins in an rdbms for web data integration. In *Proceedings of the 12th international conference on World Wide Web*, pages 90–101. ACM, 2003.

[9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[10] Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabec, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard. Building event-centric knowledge graphs from news. *Web Semantics: Science, Services and Agents on the World Wide Web*, 2016.

[11] Frederik Christiaan Schadd. *Ontology Mapping with Auxiliary Resources*. PhD thesis, SIKS, the Dutch Research School for Information and Knowledge Systems, 2015.

[12] Wee Meng Soon, Hwee Tou Ng, and Daniel Chung Yong Lim. A machine learning approach to coreference resolution of noun phrases. *Computational linguistics*, 27(4):521–544, 2001.

[13] Stephen Fienberg Kathryn Rivard William W. Cohen, Pradeep Ravikumar. Secondstring project page. `http://secondstring.sourceforge.net/`, 2010. Last visited Jun. 29, 2015.

Grades of Responsibility

Estelle Doriot [a] Jan Broersen [a]

[a] *Departement of Philosophy and Religious Studies, Utrecht University, The Netherlands*

Abstract

In this paper, we give a formal definition of grades of responsibility using stit logic and probabilities. We formalize the notion of responsibility based on probability increase. Then, we compare our framework with similar notions developed by Braham and van Hees.

1 Introduction

The question of responsibility for the outcome of one's choices is now relevant to the study of intelligent systems as those are increasingly integrated to our daily life. The framework we present here reflects the notion of agent-responsibility as defined by [8] which says that an individual is agent-responsible for an outcome to the extent that it suitably reflects the exercise of his agency. This notion is more general than moral responsibility as it only requires that the agent has the ability to make an autonomous choice. For example, we may consider that a child is not morally responsible for a crime he committed because he is not morally aware; but he would be agent responsible. Thus, there is no deontic component in our definition of responsibility, but the addition of a deontic operator could be a future research topic. To be more precise, we are interested in representing grades of responsibility as we believe that responsibility is not an all or nothing notion [8]. How to represent responsibility in relation to agents, causal relations and probabilities is still an open question. We formalize the notion of responsibility based on probability increase.

To that end, we extend stit theory, which is a formal theory of agency well established in philosophy [1]. More recently, stit theory has been used for the specification of multi-agent systems [6]. Stit models the agents' choices; however, the final outcome is not known by the agents before the execution of the action because it is the result of the combined choices and actions of all the participants. Thus an agent's action might not be successful due to other agents – or the environment – counteracting him. The responsibility of an agent depends on his beliefs about the outcome of his choice. That is why we use probabilities to represent the subjective beliefs of an agent about the combined choices of the other agents. The responsibility of one agent's choice for a particular outcome is defined as the increased probability of this outcome happening due to the agent's choice, which measures the contribution of the agent to the outcome. This definition of responsibility is subjective due to the use of subjective probabilities. We suppose that each agent consciously makes a choice and knows what choice he makes.

After giving the syntax and the semantics of an object language that represents grades of responsibility, we compare it with the notions of responsibility and degree of causation defined by Braham and van Hees in [3] and [4]. Even though theses notions are not similar to our definition of grades of responsibility, they are the closest we have found in the literature.

2 Syntax

We introduce here the syntax of our logic, which is an extension of the syntax of the probabilistic XSTIT logic defined in [5]. A new responsibility operator has been added to the existing operators. Moreover, there are differences in the semantics: we use a $Choices$ function instead of an h-effectivity function, and the definition of subjective probability varies as well.

Definition 2.1 Given a countable set of propositions P, a finite set $\mathcal{A}$ of agent names and a set of real numbers $C \subseteq [0,1]$, the formal language $\mathcal{L}$ is:

$$\varphi \quad ::= \quad p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi \mid X\varphi \mid [\alpha\ \texttt{xstit}^{\geq c}]\varphi \mid [\alpha\ \texttt{resp}^{=c}]\varphi$$

where $p \in P$, $\alpha \in \mathcal{A}$ and $c \in C$.

The modal operators have the following interpretation: $\Box\varphi$ expresses 'historical necessity'. We abbreviate $\neg\Box\neg\varphi$ by $\Diamond\varphi$; $X\varphi$ has a standard interpretation as the transition to a next moment; $[\alpha\ \texttt{xstit}^{\geq c}]\varphi$ stands for 'agent α sees to it that φ in the next moment with a probability of at least c' and $[\alpha\ \texttt{resp}^{=c}]\varphi$ stands for 'agent α has a responsibility of c for φ in the next moment'.

3 Semantics

The underlying structure of the frame we use is a branching-time structure made up of moments and histories going through these moments.

Definition 3.1 (History) Given M a non-empty set of *moments*, H is a non-empty set of possible system *histories* isomorphic with $\ldots, m_{-2}, m_{-1}, m_0, m_1, m_2, \ldots$ with $m_x \in M$ for $x \in \mathbb{Z}$. $H_m = \{h \in H \mid m \in h\}$ is the set of histories going through m. We define two functions $succ$ and $prec$ by: $m' = succ(m,h)$ iff m' succeeds m on the history h, and $m = prec(m',h)$ iff m precedes m' on the history h. We have the following constraints on the set H:
(One successor) $\forall m, h$, if $m \in h$ then $\exists! m' = succ(m,h)$
(Definition predecessor) $m = prec(m',h)$ iff $m' = succ(m,h)$
(One predecessor) $\forall m, h$, if $m \in h$ then $\exists! m' = prec(m,h)$
(One past) $\forall m, \forall h, h' \in H_m$, $prec(m,h) = prec(m,h')$

Condition **(One successor)** states that there is one and only one successor of a moment m along a history h going through m; condition **(definition predecessor)** is the definition of $prec$ as the converse of $succ$; condition **(One predecessor)** states that there is one and only one predecessor of a moment m along a history h going through m; and condition **(One past)** expresses the fact that there is one past.

Definition 3.2 (Choices) $Choices$: $M \times \mathcal{A} \to \wp(\wp(H))$ is the *choices function* yielding for an agent α in a moment m the set $Choices(m,\alpha)$ of subsets of H_m containing the agent's choices. We have the following constraints on $Choices(m,\alpha)$:
(No empty choice) $\forall K \in Choices(m,\alpha),\ K \neq \emptyset$
(No absence of choice) $\forall h \in H_m,\ \exists K \in Choices(m,\alpha)$ such that $h \in K$
(No choice between undivided histories) $\forall h, h'$, if
$succ(m,h) = succ(m,h')$ then $\forall K \in Choices(m,\alpha)$, if $h \in K$ then $h' \in K$
(Independence of agency) $\forall \alpha_i \in \mathcal{A}, \forall K_i \in Choices(m,\alpha_i), \bigcap_{\alpha_i \in \mathcal{A}} K_i \neq \emptyset$
The choice made by agent α in a moment m relative to the history h is given by:

$$Choice(m,\alpha,h) = \left\{ \bigcup_i K_i \mid K_i \in Choices(m,\alpha) \text{ and } h \in K_i \right\}$$

We thus define a choice K as a set of histories; and $Choices(m,\alpha)$ is the set of all the choices available to agent α at m. It is equivalent to an effectivity function. Condition **(No empty choice)** states that a choice cannot be empty; condition **(No absence of choice)** states that every history is accessible to any agent through a choice. The following conditions are the well-known stit condition 'no choice between undivided histories' and the 'independence of agency' which states that the choice of one agent cannot limit the choices the other agents make simultaneously. The set $Choices(m,\alpha)$ is not a partition of H_m because we omit the fact that the choices in $Choices(m,\alpha)$ are mutually disjoint. We do not include this condition because it is not modally expressible and will thus have no impact on our logic.

The definition of $Choices$ can be extended to groups of agents as the intersection of the choices of all the agents in this group.

Definition 3.3 (Group choices) $Choices_G$: $M \times \wp(\mathcal{A}) \rightarrow \wp(\wp(H))$ is the *group choice function* yielding for a group of agents A in a moment m the set of subsets of H_m containing the group's combined choices: $K \in Choices_G(m, A)$ iff $\exists(K_{\alpha_1}, \ldots, K_{\alpha_k}) \in \times_{\alpha_i \in A} Choices(m, \alpha_i)$ where $K = \cap K_{\alpha_i}$.

The choice made by the group of agents A in a moment m along the history h is given by:

$$Choice_G(m, A, h) = \left\{ \bigcup_i K_i \mid K_i \in Choices_G(m, A) \text{ and } h \in K_i \right\}$$

We introduce the state of an agent α at a moment m, which is the combined choices of all the other agents at that same moment m.

Definition 3.4 (States) The *states* of an agent α at a moment m is defined by:

$$States(m, \alpha) = Choices_G(m, \mathcal{A} \setminus \{\alpha\})$$

For each agent, the expectation is a subjective probability distribution over the states of this agent at a moment. We choose to define the expectation as a probability distribution over states (which are the combined choices of all the other agents) rather than on the individual choices of the other agents. This allows for cases where the choices of the agents are correlated.

Definition 3.5 (Expectation) The *expectation* function B: $M \times \mathcal{A} \times \wp(H) \rightarrow C$ is a subjective probability function such that $B(m, \alpha, K)$ expresses agent α's expectation that he will be in a state K in a moment m. We apply the following constraints:

1. $B(m, \alpha, K) \geq 0$ if $K \in States(m, \alpha)$
2. $B(m, \alpha, K) = 0$ otherwise
3. $\sum_{K \in States(m,\alpha)} B(m, \alpha, K) = 1$

Conditions **1.** and **2.** express that only states can be assigned a non-zero expectation. Condition **3.** is a standard probability condition stating that the sum of the expectation of one agent over the possible states adds up to 1. As stated in the definition of the syntax, the set C is a subset of $[0, 1]$. One might argue that the probability of a state should not be zero (and that the inequality in condition **1.** should be strict) because it would represent an impossible situation. However, we are using subjective probabilities which reflects the beliefs of the agents. This means that a zero probability would represent a situation that the agent did not know was possible.

We now have all the elements necessary to define the probabilistic XSTIT-frame: a set of moments, a set of histories, a choice function and an expectation function.

Definition 3.6 (Probabilistic XSTIT-frame) A *probabilistic XSTIT-frame* is a tuple $\mathcal{F} = \langle M, H, Choices, B \rangle$ such that:

1. M is a non-empty set of moments
2. H is a non-empty set of histories
3. $Choices$: $M \times \mathcal{A} \rightarrow \wp(\wp(H))$ is a choice function
4. B: $M \times \mathcal{A} \times \wp(H) \rightarrow C$ is an expectation function

Definition 3.7 (Probabilistic XSTIT-model) A probabilistic XSTIT-frameframe is extended to a *model* $\mathcal{M} = \langle M, H, Choices, B, V \rangle$ by adding a valuation V of atomic propositions V: $P \rightarrow \wp(M)$ assigning to each atomic proposition the set of moments relative to which they are true.

Figure 1 visualizes a probabilistic XSTIT-model with two agents, as given by definition 3.7. The cells represents the moments and the dashed lines going through them are the histories, grouped in bundles. In each game form, the columns correspond to the choices of the first agent α_1 and the rows to the choices of the second agent α_2. The numbers present alongside the rows are the probabilities assigned by agent α_1 to the actions of agent α_2 and, symmetrically, the numbers alongside the columns are the probabilities assigned by agent α_2 to the actions of agent α_1.

We have defined the frames and models but we still need a few additional definitions to reach the notion of responsibility. First of all, we are looking at the result of the agent's actions: when does an action leads to φ? The 'possible next φ-states' function gives the set of states that, given the agent's choice, will ensure that φ is true at the next moment.

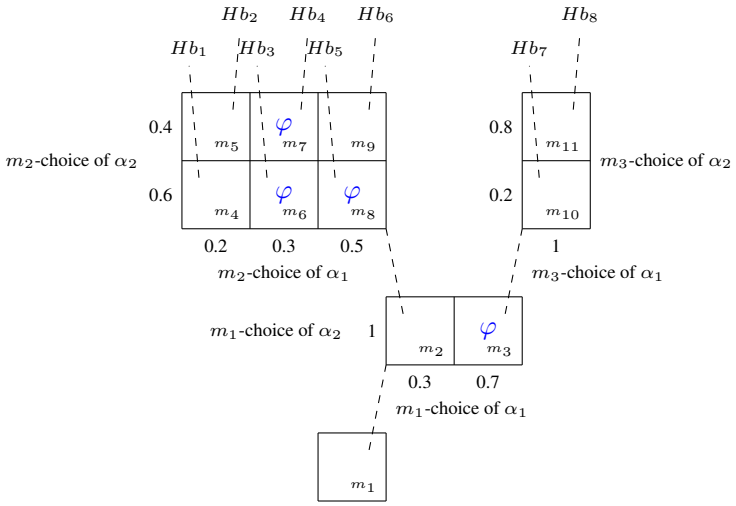


Figure 1: A partial probabilistic XSTIT model with two agents

Definition 3.8 (Possible next φ-states) The *possible next φ-states* function $PosX\colon M \times H \times \mathcal{A} \times \mathcal{L} \to \wp(\wp(H))$ which for a moment m, a history h, an agent α and a formula φ gives the possible next states obeying φ given the agent's current choice determined by h, is defined by:

$$PosX(m,h,\alpha,\varphi) = \{K \in States(m,\alpha) \mid \forall h' \in K \cap Choice(m,\alpha,h), \langle succ(m,h'),h'\rangle \models \varphi\}$$

The chance of success of an agent's action resulting to φ is the subjective belief of this agent that the choice he makes will result in a situation where φ is true. It is defined as the sum of the probabilities assigned to the states in which, given the choice made by the agent, φ will be true in the next moment.

Definition 3.9 (Chance of success) The *chance of success* function $CoS\colon M \times H \times \mathcal{A} \times \mathcal{L} \to C$ which for a moment m and a history h an agent α and a formula φ gives the chance the agent's choice relative to h is an action resulting in φ, is defined by:

$$CoS(m,h,\alpha,\varphi) = \sum_{K \in PosX(m,h,\alpha,\varphi)} B(m,\alpha,K)$$

We can now formulate the central 'responsibility function' definition which will be used to define the truth of the `resp` operator. The responsibility of an agent for an outcome can be seen as the contribution of this agent to the likelihood of this outcome. Formally, the responsibility function is defined as the difference between the chance of success given the current choice of the agent and the minimal chance of success for all the possible choices of the agent.

Definition 3.10 (Responsibility) The *responsibility* function $resp\colon M \times H \times \mathcal{A} \times \mathcal{L} \to C$ which for a moment m and a history h an agent α and a formula φ gives the responsibility of the agent α in bringing about φ by a choice relative to h is defined by:

$$resp(m,h,\alpha,\varphi) = CoS(m,h,\alpha,\varphi) - \min_{h' \in H_m} CoS(m,h',\alpha,\varphi)$$

The truth of formulas are evaluated with respect to moment/history pairs to take into account the dynamic aspect of the actions.

Definition 3.11 Relative to a model $\mathcal{M} = \langle M, H, Choices, B, V\rangle$, truth of a formula in a dynamic state $\langle m,h\rangle$, with $m \in h$, are defined as:

$$\begin{array}{lll}
\langle m,h\rangle \models p & \text{iff} & m \in V(p) \\
\langle m,h\rangle \models \neg\varphi & \text{iff} & \text{not } \langle m,h\rangle \models \varphi \\
\langle m,h\rangle \models \varphi \wedge \psi & \text{iff} & \langle m,h\rangle \models \varphi \text{ and } \langle m,h\rangle \models \psi \\
\langle m,h\rangle \models \Box\varphi & \text{iff} & \forall h' \in H_m, \langle m,h'\rangle \models \varphi \\
\langle m,h\rangle \models X\varphi & \text{iff} & \langle succ(m,h),h\rangle \models \varphi \\
\langle m,h\rangle \models [\alpha\ \texttt{xstit}^{\geq c}]\varphi & \text{iff} & CoS(m,h,\alpha,\varphi) \geq c \\
\langle m,h\rangle \models [\alpha\ \texttt{resp}^{=c}]\varphi & \text{iff} & resp(m,h,\alpha,\varphi) = c
\end{array}$$

We can go back to fig. 1 to evaluate formulas. Relative to moment m_2 and history h_5, the choice made by agent α_1 does not ensure that φ holds, since φ is not true for the second choice of agent α_2 at m_9. But φ is true at m_8 which α_1 believes has a 0.6 chance of happening. Thus we have that $\langle m_2, h_5\rangle \models [\alpha_1\ \texttt{xstit}^{\geq 0.6}]\varphi$. At moment m_2, agent α_1's choice least likely to result in φ is along history h_1 where the chance of success is 0. Thus, relative to moment m_2 and history h_5, the responsibility of α_1 for φ is the difference between the current chance of success of φ (0.6) and the minimal chance of success (0). Therefore, $\langle m_2, h_5\rangle \models [\alpha_1\ \texttt{resp}^{=0.6}]\varphi$.

In case the set C of probabilities is finite, we can express the `resp` operator from the `xstit` operator. We assume moreover that C should be closed under addition and subtraction (modulo 1). It results from this that C is necessarily of the form $C_n = \left\{\frac{k}{n} \mid 0 \leq k \leq n\right\}$ where $n \in \mathbb{N}$. We use the $[\alpha\ \texttt{xstit}^{\geq c}]$ operator as the base operator because it has better properties like additivity and monotonicity but we can define the $[\alpha\ \texttt{xstit}^{=c}]$ operator as:

Definition 3.12

$$[\alpha\ \texttt{xstit}^{=\frac{k}{n}}]\varphi \quad := \quad [\alpha\ \texttt{xstit}^{\geq \frac{k}{n}}]\varphi \wedge \neg[\alpha\ \texttt{xstit}^{\geq \frac{k+1}{n}}]\varphi$$

Proposition 3.1 *The responsibility operator can be expressed in terms of the xstit operator by:*

$$[\alpha\ \texttt{resp}^{=\frac{j-i}{n}}]\varphi \quad \textit{iff}$$

$$[\alpha\ \texttt{xstit}^{\geq \frac{j}{n}}]\varphi \wedge \neg[\alpha\ \texttt{xstit}^{\geq \frac{j+1}{n}}]\varphi \wedge \Box[\alpha\ \texttt{xstit}^{\geq \frac{i}{n}}]\varphi \wedge \neg\Box[\alpha\ \texttt{xstit}^{\geq \frac{i+1}{n}}]\varphi$$

4 Responsibility and degree of causation

Our definition of grades of responsibility can be linked to that of causal contribution defined in [2] by means of the NESS-test. The *NESS-test* says that c is a cause of e iff there is a set of events that is sufficient for e such that: (i) c is a member of the set; (ii) all elements of the set obtain; (iii) c is necessary for the sufficiency of the set. Braham and van Hees define the notion of degree of causation, which can be explained, transcribed in our notation as following.

The *degree of causation* β_i of an agent i is defined as:

$$\beta_i = \frac{|\mathcal{C}_i|}{\sum_{j\in\mathcal{A}} |\mathcal{C}_j|}$$

where $\mathcal{C}_i = \{A \mid A \subseteq \mathcal{A}$ is sufficient for φ and i is φ-critical for $A\}$.

A group of agent A is *sufficient* for φ if and only if $\forall h' \in \mathit{Choice}_G(m, A, h)$, $\langle m, \mathit{succ}(m, h')\rangle \models \varphi$. An agent i is *φ-critical* for A if and only if $\exists i \in \mathcal{A}$ such that $A \setminus i$ is not sufficient for φ.

Braham and van Hees also define a condition for a subjective notion of responsibility in [4]. This notion is not defined with a formal language in that paper, so we give here a formalization of Braham and van Hees's responsibility based on our definitions from the previous section. First, we give an operator for the NESS test.

Definition 4.1 (NESS operator) Relative to a model $\mathcal{M}$, truth of the `ness` operator in a dynamic state $\langle m, h\rangle$, with $m \in h$, is defined as:

$$\langle m, h\rangle \models [\alpha\ \texttt{ness}]\,\varphi \quad \text{iff} \quad \exists A \subseteq \mathcal{A},\ \langle m, h\rangle \models [A \cup \{\alpha\}\ \texttt{xstit}]\varphi \wedge \neg[A\ \texttt{xstit}]\varphi$$

where $\langle m, h\rangle \models [A\ \texttt{xstit}]\varphi \quad \text{iff} \quad \forall h' \in \mathit{Choice}_G(m, A, h),\ \langle \mathit{succ}(m, h'), h'\rangle \models \varphi$

Definition 4.2 (Non φ-NESS states) The *non φ-NESS states* function R: $M \times H \times \mathcal{A} \times \mathcal{L} \to \wp(\wp(H))$ which for a moment m, a history h, an agent α and a formula φ gives the next states in which α is not NESS for φ, given the agent's current choice determined by h:

$$R(m, h, \alpha, \varphi) = \{K \in \mathit{States}(m, \alpha) \mid \forall h' \in K \cap \mathit{Choice}(m, \alpha, h),\ \langle m, h'\rangle \models \neg\,[\alpha\ \texttt{ness}]\,\varphi\}$$

Definition 4.3 (Avoidance potential) The *avoidance potential* function ρ: $M \times H \times \mathcal{A} \times \mathcal{L} \to C$ which for a moment m and a history h an agent α and a formula φ gives the chance the agent's choice relative to h is an action avoiding φ, is defined by:

$$\rho(m, h, \alpha, \varphi) = \sum_{K \in R(m,h,\alpha,\varphi)} B(m, \alpha, K)$$

Definition 4.4 (Responsibility operator) Relative to a model $\mathcal{M}$, truth of the `resp_BvH` operator in a dynamic state $\langle m, h\rangle$, with $m \in h$, is defined as: $\langle m, h\rangle \models [\alpha \texttt{ resp\_BvH}]\,\varphi$ iff
(i) $\langle m', h\rangle \models [\alpha \texttt{ ness}]\,\varphi$ and
(ii) $\exists h' \in H_{m'}$, such that $\rho(m', h', \alpha, \varphi) > \rho(m', h, \alpha, \varphi)$ where $m' = prec(m, h)$.

5 Examples

We are going to compare the results of responsibility operator with those from Braham and van Hees in [2], using examples that illustrate a variety of situations and represent some modeling problems such as overdetermination or the knowledge of the agents. Even though Braham and van Hees's degree of causation is not probabilistic, we want to show that we also get correct results with our operator in non probabilistic settings when considering uniform probabilities.

Example 5.1 (Toxins [3]) Firms 1, 2 and 3 dumped different toxins in a river, denoted by T_1, T_2, and T_3 respectively. The combination of the three actions was necessary to kill all the fish in the river. Using our framework, the responsibility of each firm is therefore $\frac{1}{4}$ and is in accordance with the intuition that each agent should be equally responsible. The result obtained using the degree of causation gives $\beta_1 = \beta_2 = \beta_3 = \frac{1}{3}$ which also agree with the intuition. The value is not the same as the one we have with our framework, but the relative responsibility among the agents is the same. For simple situations like the toxin example our framework gives similar results to that of the degree of causation but it is not always the case for more complex situations.

Now, suppose that Firm 3 transfers its activities to Firm 1 so that Firm 3 ceases to dump T_3 in the river. Firm 1 now dumps toxins T_1 and T_3 into the river and Firm 2 dumps T_2. In this situation, it seems intuitive to say that Firm 1 takes over Firm 3 responsibility along with the activity and that Firm 1 is now twice as responsible as Firm 2. In our framework, the responsibility of Firm 1 is $\frac{1}{2}$ and the responsibility of Firm 2 is $\frac{1}{4}$ and agrees with the intuition. However, the degree of causation defined previously ascribes an equal responsibility to both firms. To circumvent this problem, Braham and van Hees have to extend the notion of degree of causation to complex actions and get the expected result: $\beta_1^* = \frac{2}{3}$ and $\beta_2^* = \frac{1}{3}$ with gives the same relative values as our framework.

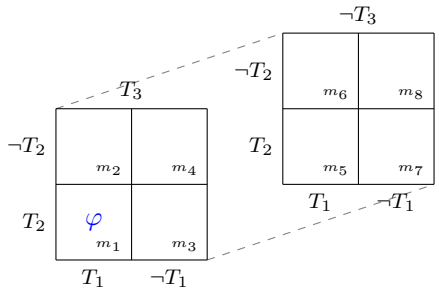


Figure 2: Three firms dumping toxins

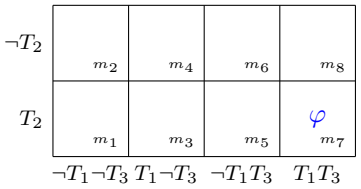


Figure 3: Two firms dumping toxins

Example 5.2 (Vote with over-determination [7]) Six voters $\{a, b, c, d, e, f\}$ have to vote in favor or against a proposal. The minimal winning coalitions in favor of the proposal are $\{a, b, c\}$, $\{a, b, d\}$ and $\{a, e, f\}$. The outcome of the vote was that $S = \{a, b, c, d, e\}$ voted in favor. This is a case of over-determination as either $\{a, b, c\}$ or $\{a, b, d\}$ was sufficient to ensure the proposal was accepted. It seems intuitive to say that each of a, b, c and d undoubtedly bears some responsibility for the outcome, but what about e? e is not a cause of the proposition being accepted; however, e's vote *could* have contributed to the proposition being accepted (had f voted in favor) and should therefore bear some part of the responsibility. Both the degree of causation and Braham and van Hees's responsibility operator consider that e is not responsible for the outcome of the vote because it is not a cause of the outcome. In our framework, a has a responsibility of $\frac{17}{32}$, b of $\frac{9}{32}$ and c, d have an equal responsibility of $\frac{3}{32}$ and e has a responsibility of $\frac{5}{32}$. f has no responsibility for the outcome because he voted against the proposition. In this situation, the results we get with our framework conform to our expectations whereas neither the degree of causation nor Braham and van Hees's definition of responsibility do.

Example 5.3 (Evening out) John and Mary want to go out together in the evening. They can either go to a restaurant, to the cinema or to a concert. They have agreed to go to a restaurant and John is certain that Mary is going to keep her promise. If John chooses to do something else, then it is reasonable to say that he is responsible for them not meeting for their evening out. Assuming that John believes that Mary has 80% chances of going to the agreed place and 10% chances to go to each of the others, according to our framework, John has a responsibility of 0.7 if he chooses not to go to the restaurant. If Mary has the same beliefs about John, but goes to the restaurant, her responsibility is 0. The degree of causation is the same of Mary and John, which does not represent accurately the responsibility in this situation where knowledge is involved. Braham and van Hees's responsibility says that John is responsible and Mary isn't, but doesn't give a grade of responsibility for each agent.

6 Conclusion

This paper extends the probabilistic XSTIT logic by adding a responsibility operator. This operator assign to each agent a level of responsibility depending on the subjective beliefs of this agent regarding the actions of the other agents. We have also shown that our responsibility operator performs as well or better than the degree of causation and the responsibility defined by Braham and van Hees. There is opportunity for further research in different directions: first, the framework supposes that each agent is aware of all his possible choices, which is not always realistic; secondly, it is assumed that the agents always know all the consequences of each outcome, which is not always the case; and finally, it would be interesting to investigate the deontic aspect of responsibility.

References

[1] N. Belnap, M. Perloff, and M. Xu. *Facing the future: agents and choices in our indeterminist world*. Oxford University Press, 2001.

[2] Matthew Braham and Martin van Hees. Degrees of causation. *Erkenntnis*, 71:323 – 344, 2009.

[3] Matthew Braham and Martin van Hees. An anatomy of moral responsibility. *Mind*, 121(483):601–634, 2012.

[4] Matthew Braham and Martin van Hees. Voids or fragmentation: Moral responsibility for collective outcomes. 2014.

[5] Jan Broersen. Probabilistic stit logic and its decomposition. *International Journal of Approximate Reasoning*, (54):467477, 2013.

[6] Jan Broersen, Andreas Herzig, and Nicolas Troquard. From coalition logic to STIT. *Electronic Notes in Theoretical Computer Science*, 157(4):23–35, 2006. Proceedings of Logic and Communication in Multi-Agent Systems (LCMAS 2005).

[7] Dan S Felsenthal and Moshé Machover. A note on measuring voters' responsibility. *Homo Oeconomicus*, 26(2):259–271, 2009.

[8] Peter Vallentyne. Brute luck and responsibility. *Politics, Philosophy & Economics*, 7(1):57–80, 2008.

Combining Combinatorial Game Theory with an α-β Solver for Clobber

Janis Griebel Jos Uiterwijk

Department of Data Science and Knowledge Engineering, Maastricht University
P.O. Box 616, 6200 MD Maastricht, The Netherlands

Abstract

Combinatorial games are a special category of games sharing the property that the winner is by definition the last player able to move. To solve such games two main methods are being applied. The first is a general NegaScout search with many possible enhancements. This technique is applicable to every game, mainly limited by the size of the game due to the exponential explosion of the solution tree. The second way is to use techniques from Combinatorial Game Theory (CGT), with very precise CGT values for (subgames of) combinatorial games. This method is only applicable to relatively small (sub)games. In this paper[1] we show that the methods can be combined in a fruitful way by incorporating endgame databases filled with CGT values into a NegaScout solver.

We apply this technique to the game of Clobber, a well-known all-small combinatorial game. Our test suite consists of 20 boards with sizes up to 18 squares. Endgame databases were created for all subgames of size 8 and less. The CGT values were calculated using the CGSUITE package. Experiments reveal reductions of at least 75% in number of nodes investigated.

1 Introduction

Clobber is a two-player board game invented by Albert, Grossman and Nowakowski in 2001 [1]. The board consists of a rectangular white and black checkerboard. The size can be varied. Each black square is occupied by a piece of the first player (conventionally called Left) and each white square by one of the second player (Right). Commonly the start position is a filled 8×8 board as shown in Figure 1.

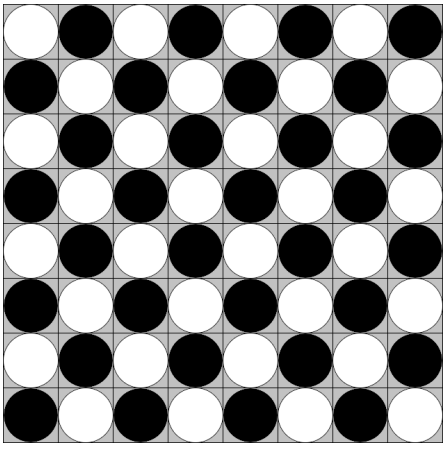


Figure 1: Common Clobber start position

The players move in alternating order. A move is done by picking one of the own stones and moving it to an orthogonally adjacent square that is occupied by an opponent stone. The opponent stone gets

[1] This work is based on the master thesis research of the first author [7].

removed from the board ("clobbered"). The player that makes the last move wins. As an example, Figure 2 shows how Left could perform a move. It can choose between two moves. In the figure it has moved its black piece to the left and removed the adjacent white piece of player Right.

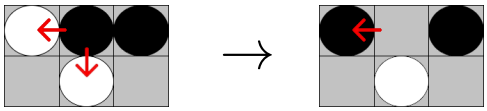


Figure 2: Move in Clobber

Important for the CGT approach of Clobber are the following facts:

- Clobber is a perfect-information, deterministic game.
- Clobber is a converging game. With each move the total number of stones gets decreased by one.
- Clobber is a partizan game. There exist stones for Left and for Right which results in the fact, that each player has its own moves.
- If a player can make a move in a specific position, the opponent player could also make a move in that position, because then two opponent stones are orthogonally adjacent. This means the game is a so-called *all-small* game [2].
- The symmetry of a position does not influence the outcome. If the board is mirrored vertically and/or horizontally, or rotated by a multiple of 90 degrees, the outcome of the position does not change.
- No draws are possible: either the first player (Left) or the second player (Right) wins.

In this article we describe the results of building endgame databases with exact CGT values for the game of Clobber. Section 2 gives an overview of related research. In Section 3 we describe the methods implemented. Next, in Section 4, we show our experiments. In Section 5 we give our conclusions and some suggestions for further research. Note that we refrain from giving an introduction to CGT, due to space restrictions. For this we refer to the literature mentioned above, in particular to [6, 4, 2].

2 Related Work

The literature regarding CGT for Clobber is scarce, except the paper introducing the game [1]. Claessen constructed for Clobber (partly filled) databases with exact CGT values, but they were not used to solve boards, but to enhance a Clobber playing engine based on the MCTS framework [5].

Other related work has been done by Müller [9] who applied CGT values to solve local endgames in Go. However, the global search is not an α-β search, and CGT values are not obtained from CGT endgame databases, but calculated on the spot. Müller and Li [11] show results for combining α-β search with CGT pruning and ordering. Their games are artificial games with very special properties, having nothing in common with "real" combinatorial games. No endgame databases were used.

Most other related work has been done in the area of the game Amazons. Müller [10] used CGT to establish bounds in a specialized divide-and-conquer approach. He was able to solve 5×5 Amazons. No CGT endgame databases were used. Snatzke [14] built CGT endgame databases for a very restricted version of Amazons, namely for subgames fitting on a 2×11 board with exactly 1 queen per player. This was extended in [15] with new results for some small databases of other shapes, with 1 to 4 queens. He did not incorporate the use of his databases in a general Amazons solver. Tegos [17] was the first to combine endgame databases for Amazons in a NegaScout-based Amazons playing program. Besides (traditional) minimax endgame databases (without CGT information) he also implemented CGT endgame databases. These contained just thermograph information, not precise CGT values, and therefore only could be used for (heuristic) move-ordering purposes. Recently, Song [16] implemented endgame databases in an Amazons solver. Again, the databases did only contain heuristic (thermograph) information useful for narrowing the bounds in the solving process. With his program 5×6 Amazons has been solved.

As far as we know the only other research reporting on combining global α-β searches with endgame databases with precise CGT values is our related work on Domineering [3, 18]. In these articles we

showed that equipping a simple α-β solver with CGT endgame databases gives reductions up to 99% in number of nodes investigated for solving a testset of 36 non-trivial rectangular boards.

3 Methods

In this section we describe the methods implemented in our Clobber solver. The core is a straightforward α-β solver, described in Section 3.1. The main enhancement consists of the construction of the CGT endgame databases (Section 3.2). Section 3.3 shows how the CGT values from the endgame databases can be used within the α-β framework.

3.1 Basic α-β solver

For the basic Clobber solver we built a NegaScout [12] implementation of a very simple α-β [8] searcher. This searcher investigates lines until the end (returning that either Left or Right made the last move and so wins the game). Since our goal was to measure the impact of using CGT on solution tree sizes, we refrained from incorporating any further enhancements to the α-β framework.

3.2 A CGT Database for Clobber

In order to use CGT values for Clobber subgames we have built all Clobber endgame databases for subgames of sizes 1–8 with arbitrary shape. Positions that exhibit mirror symmetry (left/right and/or top/bottom) are unified into a single entry.

The database is used to store pre-calculated CGT values of specific positions. These CGT values are used by the agent to get knowledge about its subgames. If the agent searches down to a position that only consists of pre-calculated subgames, the whole position can be solved by combining the CGT values.

As a preparation a database must be created, designed and filled with the appropriate values. These steps are described in the following subsections.

3.2.1 Database design

In the database the CGT values must be stored connected with the appropriate game position. The easiest way to store the CGT values is using a character string because of the different kinds of CGT value types (numbers, ups, downs, stars, etc). For the game position a hash value is generated that allows to rebuild the exact position. The format of the hash value is described in Section 3.2.2.

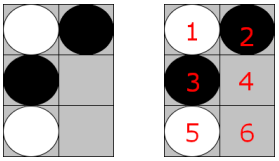
As the information that is needed is very simple, only one table is needed for the database. The hash value is used as the primary key. Since the hash value consists of two values, a string and an integer for the position width, the primary key is a combined one. A third column is added to store the CGT value associated with the position.

3.2.2 Representing a position

A game position is represented in a hash value to get stored in the database. It is important that the hash value is unique for one position. There are different approaches that can be used to generate such a unique hash value.

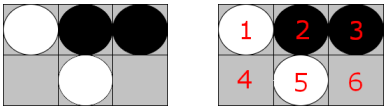
In our first approach each square is iterated sequentially and indicated in the hash value. The squares are visited from left to right through all rows from top to bottom. An empty square is indicated by the character ”0”, a left player occupied square by ”1” and a right player occupied square by ”2”. These characters are concatenated to a string and used as the hash code for the appropriate position. An example is shown in Figure 3.

The hash code is generated visiting the squares in the shown order. The issue with a hash code in this format is, that it does not include the position sizes. The hash code in Figure 3 indicates that it contains six squares. Six squares can be part of a position with the size 6×1, 1×6, 3×2 or 2×3. It is not unique. For example the hash code in Figure 4 is the same but it belongs to a different position with a different CGT value.



Hash Code = "**211020**"; width = 2; CGT value = $\{*, \uparrow | *, \downarrow\}$

Figure 3: First example for hash code generation



Hash Code = "**211020**"; width = 3; CGT value = $*$

Figure 4: Second example for hash code generation

To solve this issue the position size needs to be stored together with the hash code. For that the width of a position is added as a column to the database. The primary key consists of the combination of the position width and the generated hash code.

3.2.3 Calculating CGT values

The database is prepared to get filled with CGT values and to get used as a lookup for solved positions. The filling of the hash and the width columns are described in the previous section.

The CGT values for the Clobber subgames are calculated with the CGSUITE software tool [13]. These values are in so-called canonical form, which make them a unique representation of the values. They are stored in a third column in the database.

3.2.4 Generating positions

To fill the database a set of positions is needed for which the appropriate CGT values are calculated and stored. To generate the hash values for these positions an algorithm is implemented that iterates through all possibilities.

The algorithm gets a limit that consists of a number of connected, nonempty squares. All possible formations and fillings of the connected squares are iterated to fill the database. To get all positions with n connected and nonempty squares, an empty $n \times n$ subboard is created. In a naive approach, the algorithm would go recursively through all squares, filling each square first with nothing, then with a left player stone and at last with a right player stone.

For an $n \times n$ board it would mean that the algorithm generates $3^{(n \times n)}$ combinations of fillings. These combinations include many combinations that exceed the set limit of connected nonempty squares n. To avoid the iteration through combinations that do not fit the set limit n the algorithm counts the number of nonempty squares through the iterations. If the limit n is reached it continues just with empty squares till the last square. The resulting amount of combinations can be described with binomial coefficients. For an $n \times n$ board it means that there are

$$\binom{n \times n}{n} \times 3^n$$

combinations generated.

In a next step the algorithm is changed to ignore combinations that contain less than the set limit of n connected, nonempty squares. Since the algorithm is executed sequentially with an increasing limit n it does not need to check combinations with less than n connected, nonempty squares. This improvement is obtained by considering only positions that do not generate subgames, since it would mean that the amount of nonempty squares is less than n for each subgame. Additionally a position is not evaluated if it has less than n nonempty squares reaching the last square. The generation of subgames is prohibited

directly. This improvement leads to a smaller amount of evaluated combinations that can be described by the following formula since empty squares are not present:

$$\binom{n \times n}{n} \times 2^n$$

As a last improvement, the amount of iterated positions is drastically decreased. Consider that an $n \times n$ board is used to get all formations of n nonempty connected squares. The width n of the position is only needed if all nonempty squares are in a row. Therefore a height of 1 is enough to get all possible formations, because there is only one line with all nonempty squares in a row. If one square gets occupied that is not in a row with the others, a smaller position with the width of $n-1$ and the height of 2 can be used to generate all missing positions. For the second square that is not in row 1 or 2 it means that a position with the size of $n-2$ fits to get all possible positions with a maximum of $n-2$ squares in a row. As an example of generating all possible formations for $n = 5$ connected squares, Figure 5 shows the required board sizes.

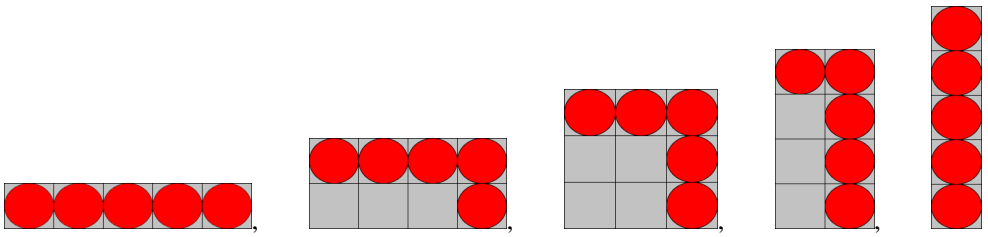


Figure 5: Board sizes needed to get all possible formations for $n = 5$ connected pieces

We replace the amount of squares, that are not in a row with any other square, but still connected with the other ones, by a variable p. Using boards with the size of $(n-p+1) \times p$, where p runs from 1 to n, we can generate all possible formations of n connected, nonempty squares. The following formula describes the amount of iterated positions for one specific p:

$$\binom{(n-p+1) \times p}{n} \times 2^n$$

So if we let p run from 1 to n and calculate the sum of the amount of all formation possibilities, the number of iterated positions is described by the following term:

$$\left(\sum_{p=1}^{n} \binom{(n-p+1) \times p}{n} \right) \times 2^n$$

To show the approach improving the generation of positions, the amount of iterated positions is listed in Table 1.

Table 1: Approach of improving position generation

n	no optimization $3^{(n \times n)}$	n nonempty squares $\binom{n \times n}{n} \times 3^n$	no subgames $\binom{n \times n}{n} \times 2^n$	final
1	3	3	2	2
2	81	54	24	8
3	19683	2268	672	48
4	4.30×10^7	1.47×10^5	29120	512
5	8.47×10^{11}	1.29×10^7	1.70×10^6	7680
6	1.50×10^{17}	1.42×10^9	1.24×10^8	1.45×10^5
7	2.39×10^{23}	1.88×10^{11}	1.10×10^{10}	3.31×10^6
8	3.43×10^{30}	2.90×10^{13}	1.13×10^{12}	8.84×10^7

The final improvement results in an intense redundancy reduction. Generating one position means that the hash code gets generated, a lookup in the database is made, the Java interface is called and sometimes a new entry is stored to the database. Now the algorithm only generates 8.84×10^7 instead of 3.43×10^{30} positions for $n = 8$ nonempty connected squares. This is an improvement of a factor 4.05×10^{22}.

3.2.5 Filling the database

The C# algorithm, that generates the positions, calculates the hash code, looks it up in the database and calls the Java interface to get the CGT values, was prepared to run with one specific set n. To fill the database it first runs with $n = 1$. After completing the generation of all positions, n is increased by 1 and the algorithm got started again. The idea is to get positions generated for an n that is as high as possible. The higher the n is the longer the algorithm runs because the amount of iterated positions increases exponentially.

Running on a 3GHz Quad-Core computer the algorithm needed around 2 hours for all positions with $n = 7$ nonempty connected squares. For $n = 8$ it took more than 3 days. Expecting a few months for $n = 9$ the generation was stopped. The number of generated entries for n connected nonempty squares is listed in Table 2. Note that for $n \geq 4$ these numbers are smaller than the ones in the right column of Table 1, since the positions generated may consist of unconnected subfragments already in the database for smaller values of n.

Table 2: Stored positions in the database

n	stored positions
1	2
2	8
3	48
4	304
5	2016
6	13824
7	97280
8	697600
	sum = 811082 entries

The database consumes 79.23 MB memory space.

3.3 Combining CGT Endgame Databases with α-β

Combining CGT endgame databases with an α-β solver can be achieved by a simple adaptation of the NegaScout algorithm. As soon as a (sub)game with at most 8 connected pieces is encountered, its value is looked up in the database and using this value it is determined whether the (sub)game is a win or loss for the player to move.

As a result, when investigating an $m \times n$ board lines are never searched deeper than $m \times n - 8$, when the position is necessarily reduced to 8 stones (since each move consists of removing exactly 1 stone). Moreover, many lines will even terminate at shallower depths when a position splits into multiple subgames. As long as there are subgames with more than 8 stones, the investigation continues, but only for these subgames. As soon as each subgame separately contains at most 8 stones, the search terminates and game logic is used to determine the winner. This ensures determining correctly the game-theoretical result.

4 Experiments and Discussion

In this section the implemented features are tested to generate knowledge about their impact on improving alpha-beta search. In the setup we used the NegaScout implementation, without and with database support, as described in the previous section.

To get comparable results, boards with different sizes are generated and searched. The size runs up to 18 squares with a maximum width of 8. The board is initially filled in the common way as a checkerboard, starting with a left (Black) stone. Since the results are not comparable on different computers considering the investigated time, the number of visited nodes is used as the value to compare.

Two remarks are in order. First, although the efficiency of α-β depends on move ordering, we have not incorporated this in the current implementation of the agents. Moreover, it is far from trivial to find a good move-ordering technique for an all-small game like Clobber. Second, who starts influences the outcome of visited nodes. Therefore each position is played twice, once with Left and once with Right as the player to move. Both amounts of visited nodes are summed and given as the results shown in Table 3.

Table 3: Comparison of visited nodes during searches

Board size	α-β_noDB	α-β_withDB	red. (%)	outcome class
1 (1×1)	0	0	100.00%	$\mathcal{P}$
2 (1×2)	2	0	100.00%	$\mathcal{N}$
3 (1×3)	2	0	100.00%	$\mathcal{N}$
4 (1×4)	3	0	100.00%	$\mathcal{N}$
4 (2×2)	6	0	100.00%	$\mathcal{N}$
5 (1×5)	15	0	100.00%	$\mathcal{P}$
6 (1×6)	42	0	100.00%	$\mathcal{P}$
6 (2×3)	68	0	100.00%	$\mathcal{P}$
7 (1×7)	49	0	100.00%	$\mathcal{R}$
8 (1×8)	60	0	100.00%	$\mathcal{N}$
8 (2×4)	162	0	100.00%	$\mathcal{N}$
9 (3×3)	222	2	99.10%	$\mathcal{N}$
10 (2×5)	654	66	89.91%	$\mathcal{N}$
12 (2×6)	16150	1003	93.79%	$\mathcal{P}$
12 (3×4)	19532	1412	92.77%	$\mathcal{P}$
14 (2×7)	45502	2849	93.74%	$\mathcal{N}$
15 (3×5)	125638	14573	88.40%	$\mathcal{N}$
16 (2×8)	304230	52384	82.78%	$\mathcal{N}$
16 (4×4)	625544	105596	83.12%	$\mathcal{N}$
18 (3×6)	24626986	6222980	74.73%	$\mathcal{P}$

In this table all investigated board sizes (with the board between brackets) are listed in the left column. The second column shows the number of visited nodes (disregarding the root in the tree) using the default alpha-beta search (denoted α-β_noDB). The third column contains the number of visited nodes when de database is used (α-β_withDB). Column 4 shows the reductions for α-β_withDB compared to α-β_noDB as percentages. Finally, the last column contains the outcome class for the board. Here, $\mathcal{P}$ means a win for the **previous** player, so a loss for the player to move, $\mathcal{N}$ means a win for the **next** player (the player to move), whereas $\mathcal{L}$ and an $\mathcal{R}$ denote wins for **Left** and **Right**, irrespective who starts. The results in this column match with the results in [1].

Keep in mind that the filling of each board starts with a Left piece and then goes further alternating between Left and Right. Therefore in the position of size 1×7, if the filling would start with Right, then Left would be determined as the winner, irrespective of who starts (i.e., outcome class $\mathcal{L}$). All other results just depend on the player to move first and do not change.

5 Conclusions and Future Research

We have shown how CGT values of subgames can be used as an enhancement of a basic NegaScout solver for Clobber. A database with exact CGT values was built for all (sub)games up to 8 connected pieces. Reductions depend on board size, going down from 100% for the boards in the database to 75% for the 3×6 board.

As future research we intend to build the databases for larger subgames. Moreover, we see opportunities to include more game-dependent knowledge, like move-ordering heuristics, to solve larger and

more complex Clobber boards. Also, well-known game-independent techniques like transposition tables can greatly enhance the efficiency. Of course it has to be investigated whether for such opimized α-β solvers for Clobber incorporating CGT endgame databases gives similar efficiency enhancements. Finally, we envision using CGT theory in solvers for other combinatorial games (besides Domineering [3] and Clobber (this work)) to get more insight into when such a combination is fruitful.

References

[1] M.H. Albert, J.P. Grossman, R.J. Nowakowski, and D. Wolfe. An introduction to Clobber. *Integers: Electronic Journal of Combinatorial Number Theory*, 5(2), 2005. 12 pp.

[2] M.H. Albert, R.J. Nowakowski, and D. Wolfe. *Lessons in Play: An Introduction to Combinatorial Game Theory*. A K Peters, Wellesley, MA, 2007.

[3] M. Barton and J.W.H.M. Uiterwijk. Combining combinatorial game theory with an α-β solver for Domineering. In F. Grootjen, M. Otworowska, and J. Kwisthout, editors, *BNAIC 2014: Proceedings of the 26th Benelux Conference on Artificial Intelligence*, pages 9–16. Radboud University, Nijmegen, 2014.

[4] E.R. Berlekamp, J.H. Conway, and R.K. Guy. *Winning Ways for your Mathematical Plays*, volume 1-2. Academic Press, London, 1982. 2nd edition, in four volumes: vol. 1 (2001), vols. 2, 3 (2003), vol. 4 (2004). A K Peters, Wellesley, MA.

[5] J. Claessen. Combinatorial game theory in Clobber. Master's thesis, Maastricht University, 2011.

[6] J.H. Conway. *On Numbers and Games*. Academic Press, London, 1976. 2nd edition (2001): A K Peters, Natick, MA.

[7] J. Griebel. Incorporating combinatorial game theory into an alpha-beta solver for Clobber. Master's thesis, Maastricht University, 2015.

[8] D.E. Knuth and R.W. Moore. An analysis of alpha-beta pruning. *Artificial Intelligence*, 6:293–326, 1975.

[9] M. Müller. Global and local game tree search. *Information Sciences*, 135:187–206, 2001.

[10] M. Müller. Solving 5×5 Amazons. In *The 6th Game Programming Workshop (GPW 2001), Hakone (Japan), 2001*, volume 14 of *IPSJ Symposium Series*, pages 64–71, 2001.

[11] M. Müller and Z. Li. Locally informed global search for sums of combinatorial games. In H.J. van den Herik, Y. Björnsson, and N. Netanyahu, editors, *Computers and Games: 4th International Conference, CG 2004*, volume 3846 of *Lecture Notes in Computer Science*, pages 273–284. Springer Int. Publ., 2006.

[12] A. Reinefeld. *Spielbaum-Suchverfahren*. Informatik-Fachberichte. Springer, 1989.

[13] A.N. Siegel. Combinatorial game suite: A computer algebra system for research in combinatorial game theory, 2003. Available from `http://cgsuite.sourceforge.net/`.

[14] R.G. Snatzke. Exhaustive search in the game Amazons. In R.J. Nowakowski, editor, *More Games of No Chance,* Proc. MSRI Workshop on Combinatorial Games, July, 2000, Berkeley, CA, MSRI Publ., volume 42, pages 261–278. Cambridge University Press, Cambridge, 2002.

[15] R.G. Snatzke. New results of exhaustive search in the game Amazons. *Theoret. Comput. Sci.*, 313:499–509, 2004.

[16] J. Song. An enhanced solver for the game of Amazons. Master's thesis, University of Alberta, 2013.

[17] T. Tegos. Shooting the last arrow. Master's thesis, University of Alberta, 2002.

[18] J.W.H.M. Uiterwijk and M. Barton. New results for Domineering from combinatorial game theory endgame databases. *Theoret. Comput. Sci.*, 592:72–86, 2015.

What can the PGM community contribute to the 'Bayesian Brain' hypothesis?

Johan Kwisthout[†ab]

[a] Department of Artificial Intelligence
[b] Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
PO Box 9104, 6500HE Nijmegen, The Netherlands

Abstract

Despite the now common view amongst neuroscientists that the brain effectively approximates Bayesian inferences, there are only few researchers in the Probabilistic Graphical Models (PGM) community currently working in this research area. We believe that this is partially due to a misunderstanding of the theoretical challenges that theoretical neuroscience currently faces and the potential contribution that the PGM community can offer in interdisciplinary research. With this paper we hope to remedy such misunderstandings and invite the community to contribute to the mutual benefit of neuroscience and AI alike.

1 Introduction

When discussing recent advances in neuroscience–that postulate that the human brain is at its essence just a Bayesian inferential machine–with scholars in the Probabilistic Graphical Models (PGM) community, our research group occasionally receives lukewarm responses that can best be paraphrased as "I'm just not interested in the brain as an application area of my research". Although there are few things as personal as a research agenda, we still feel that this lack of interest may be at least partially due to a) a misconception of the questions that are currently being addressed in neuroscience and b) lacking some 'insider's insight' in the contribution that the PGM community can offer in interdisciplinary research. With this paper we hope to remedy both. We will give a short overview of the increasingly popular 'Bayesian Brain' hypothesis in neuroscience, in particular its 'predictive processing' manifestation. We will then identify three research areas within this topic where contributions from the PGM community can actually have a huge scientific impact. After identifying potential pitfalls in such interdisciplinary research, including a discussion of the specific (and sometime peculiar) connotations of the neuroscience community with respect to concepts like 'Bayesian', 'uncertainty', and 'prior', we will conclude with an invitation to the community to contribute.

2 The Brain as 'Application Area'

Herman von Helmholtz [39] is traditionally seen as the originator of the view of human perception as (statistical) inference to the best explanation of the causes of the perceptual input. The suggestion that the human brain can be seen as performing some approximate Bayesian inference (integrating prior expectations with newly arriving information) was coined as early as 1957 by Edwin T. Jaynes (first published in [17]). Peter Dayan and colleagues further explored these ideas and proposed the notion of the *Bayesian Brain*, emphasizing on the basis of psychophysical evidence that human perception

† j.kwisthout@donders.ru.nl.

actually is 'Bayes optimal' in combining priors and new signals. The *Bayesian coding* hypothesis postulates that the brain indeed encodes probability distributions in populations of neurons.

In recent years, the *Bayesian Brain* hypothesis has become increasingly popular due to the emergence of Karl Friston's *free energy principle*, providing for a biological and physical foundation; the *predictive processing* view of the brain as a 'prediction machine' that minimizes computational effort by trying to predict its inputs, and the *spiking neural network* research area that shows that probability distributions can be encoded and sampled from using power-efficient networks of spiking neurons. We will elaborate more on these three important recent developments.

2.1 The free energy principle

Friston's *free energy principle* [9, 10] postulates that any biological system that 'resists a tendency to disorder' – be it a single cell or a social network – effectively aims to minimize free energy. In thermodynamics, free energy is the amount of energy that is potentially available, but not put to effective use. In information theory, it is a measure on the discrepancy between our observation of the world and our model of the world, which becomes manifest as the *prediction error* between predicted and observed world state. A biological system that aims to defy disorder seeks to lower entropy (the average of surprise of outcomes). It can do so by minimizing prediction error, that is, aiming to make the predicted world state match the observed world state (adapting one's models of the world), or vice versa (changing one's sensory input by acting upon the world). Because biological systems must remain within certain boundaries to exist, their models of what the world should look like (e.g., have access to a sufficient, but not excess, amount of oxygen to maintain homeostasis) and how they currently perceive the world (e.g., shortage of oxygen) should match, and if not, actions are taken to minimize this prediction error (e.g., breathe faster and deeper). Friston [9, p.295] summarizes this by postulating that *(i) agents resist a natural tendency to disorder by minimizing a free-energy bound on surprise; (ii) this entails acting on the environment to avoid surprises, which (iii) rests on making Bayesian inferences about the world.*

2.2 Predictive processing

The Predictive Processing account proposes that the brain continuously predicts its inputs in a hierarchical cascade of (increasingly more concrete) probabilistic predictions [4, 5, 16]. For example, when observing a bowler on a bowling lane, contextual information ("this bowler already hit three strikes in this game") will generate predictions for the result of the throw ("many pins will fall down"). Based on that expectation, more specific predictions will be made for the throwing kinematics, the ball trajectory, where the ball will hit the pins, etc. Violations of predictions will yield prediction errors that need to be 'explained away' by updating ones hypotheses ("even good bowlers will sometimes fail to throw a strike"), taking new contextual information into consideration ("the bowler seems to have injured his wrist whilst throwing") etc. The computations 'under the hood' of this conceptual description can be described and analyzed as various computations on causal Bayesian networks, such as the computation of posterior probability distributions and the tuning of parameters of the network [24]. The rationale behind this account is that processing only the prediction error is less computationally demanding as processing the entire input; however for exact computations, it was shown that this assumption does not hold in general, since processing even a single bit of prediction error is an NP-hard problem [21]; whether *approximate* Bayesian inference is tractable when the prediction error is low is currently an open problem. Despite its popularity as a unifying theory, it is far from clear what the brain's approximation algorithms actually look like; in Clark's [4, p.201] words: *What do the local approximations to Bayesian reasoning look like as we depart further and further from the safe shores of basic perception and motor control? What new forms of representation are then required, and how do they behave in the context of the hierarchical predictive coding regime?*

2.3 Networks of spiking neurons

One of the most promising computational models of neuronal computation in general is the recurrent *network of spiking neurons* model [28]. These biologically inspired networks mimic Boltzmann machines (neural networks that represent a probability distribution that can be sampled from), with a key difference that the neurons are not outputting a zero or one state, but a *spike*; a brief burst of

energy. These networks are energy-efficient and stochastic in nature and they can represent, and reason with, arbitrary probability distributions by means of stochastic sampling in winner-take-all microcircuits [2, 13, 33]. It has been proposed that such sampling methods (like MCMC sampling) are the most promising techniques to describe actual stochastic inferences in the brain [36]. Because of their efficiency – the brain uses a mere 25W of energy – these networks are potentially crucial for future generations of computer hardware by utilizing (rather than trying to filter) the noise that is inherent at the nano-scale [14]. No free lunch is offered, though: As approximate Bayesian inference is an intractable problem [7, 23], there will be problem instances where the convergence time of the network will grow exponentially with the input size, in particular in networks with extreme probability distributions [28].

In terms of Marr's levels of explanation [29], one can see the free energy principle as aiming to answer the 'why' of the Bayesian Brain hypothesis, the predictive processing account describes 'what' is actually being computed, whereas the 'spiking neurons' community studies the 'how' aspect of approximate Bayesian computations in the brain. Where the free energy/predictive processing and the networks of spiking neurons communities were traditionally relatively isolated – as a proxy, one could see them as exponents of the *UK*, respectively *Continental* approach towards theoretical neuroscience – there have been recent mutual research events (for example at the European Institute for Theoretical Neuroscience in Paris) that try to bridge the gap between both communities.

2.4 Organization of this paper

All these developments support the 'Bayesian' view of the brain as it is currently dominant in contemporary neuroscience. We believe that this opens up a significant area of research for the PGM community. Where graphical models are currently used in neuroscience, their role is typically limited to association, clustering, or classification of brain data [1], i.e., as a data-analysis tool rather than as a process-level description of the brain's mechanisms for information processing. Yet, the emergence of the Bayesian brain hypotheses opens up a whole new area of research. In the remainder of this paper we will further elaborate on this. We will show how a formal and computational background can help to bring conceptual clarity and formal rigidity to the field; how neuroscience is in urgent need for new algorithms, implementations, and complexity analyses that computer scientists and AI practitioners can provide, and where new questions in the 'meta'-theory of learning and modifying Bayesian networks emerge.

3 Conceptual Clarity and Rigidity

An important area where researchers with a strong background in computational and formal modeling can make vital contributions is in offering conceptual clarity and formal rigidity, translating verbal theories into complete and consistent computational models, thus exposing ambiguities and gaps in the theory and explicating 'design choices' and their computational consequences [31]. Examples are in the formal explication of the role and nature of the underlying principles of predictive processing [20,34,37], critically assessing the validity of simplifying assumptions [30], and in exposing the consequences of alternative readings of vague or conflicting verbal models [25]. On top of this, the specific background of researchers in the PGM community can contribute significantly to the theory itself, generating new theoretical and empirical questions. The following case study will further exemplify this.

In the predictive processing theory, stochastic predictions are compared with actual observations and only the residual (non-predicted) signal is processed by prediction error minimization. This prediction error, however, is dependent on the state space of the prediction and its granularity; for example, when we predict and observe a (non-specified) tree, or when we predict to see an oak and see a chestnut tree. This observation – made from an information-theoretic point of view – led to a further refinement of the predictive processing account with the notion of *levels of detail* of models and predictions (Figure 1), and spawned various research projects. One particular empirical result that is based on these insights is the development of a predictive processing account of how psychedelics effect the brain's information processing [35]. Here it was proposed that psychedelics such as psilocybin hyper-activate $5HT_{2A}$ receptors in layer-5 pyramid cells, effectively leading to over-detailed, diffuse predictions that cause many of the sympoms associated with psilocybin administration.

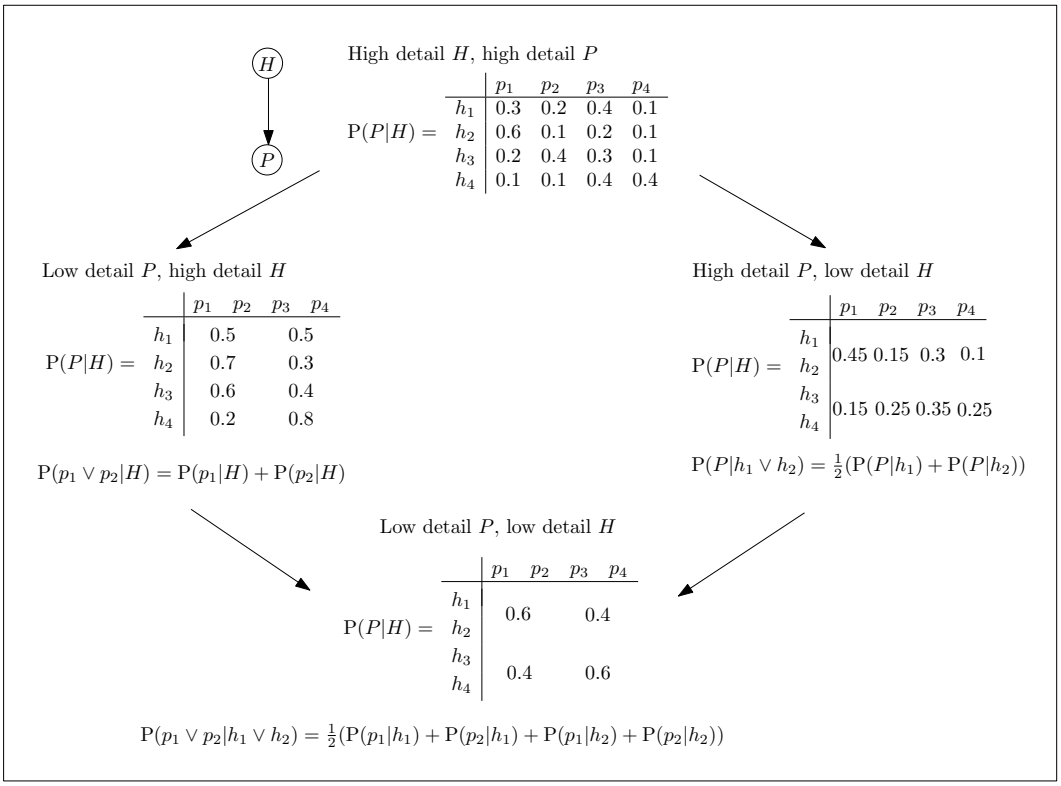


Figure 1: A formalization of the relationship between different levels of detail of hypotheses and predictions. In this example we have a singleton hypothesis and a singleton prediction, each with vour different values. Observe that the actual hypotheses, as well as the predictions, can be *clustered*, redefining the conditional probability distributions in a straightforward way. We can thus lower the detail of the predictions (leftmost CPT), lower the detail of the hypotheses (rightmost CPT), or both (bottom CPT).

4 Theory, Algorithms, and Analysis

The Bayesian Brain hypothesis stipulates that the brain approximates probabilistic inference by means of variational Bayes methods [12] or sampling approaches [36]. As exact inference is PP-complete [26] there cannot exist efficient approximation algorithms in general, unless BPP equals PP. That means that there must be constraints on the inputs in order to render this approximate computation tractable; the field of *parameterized complexity* studies such constraints. Recent developments in this area allow for the analysis of stochastic computations where the probability of answering incorrectly is parameterized, rather than the computation time [22, 23]. This allows for the study of so-called *fixed error randomized tractable* approximations, relative to 'ecologically valid' parameters, viz. parameters that can plausibly be assumed to be small in the computations as performed by the brain. In particular the parameterized complexity of approximate inference, as parameterized by the prediction error of the generative model, is an important open problem that would significantly contribute to the Bayesian Brain hypothesis in general and predictive processing in particular: It is claimed [4] that the brain can be efficient *because* it tries to minimized prediction error and thus that inference can be tractable when prediction error is low.

Apart from process-level considerations (under what constraints can the approximations postulated by predictive processing be tractable), one can study the properties and plausibility of neuronal implementations of such approximations using networks of spiking neurons. Crucial properties here are the power efficiency of such networks [28], the nature of the *noise* in the brain and its consequences for efficient sampling [13], and the general question how many resources are needed for effective computations [27]. Computational complexity theory offers an indication of the resources needed for a particular computational problem to be solved, as a function of the input size of a problem. These resources most

notably, time and memory are typically fairly coarse and built on a theoretical abstract model of computation: Turing machines. Here, the 'time' resource refers to the number of state transitions in the machine, and the 'memory' resource refers to the number of memory cells on the tape that are used. It has been proposed by a working group at the Dagstuhl seminar on Resource-Bounded Problem Solving (seminar 14341) to have a more refined, brain-focused model of computation in the brain, based on networks of spiking neurons, and have complexity measures based on brain resources, such as spiking rates, network size, and connectivity [15]. The development of such a model of computation would allow for seminal contributions to the Bayesian Brain hypothesis by analyzing the fundamental limits of brain computations.

5 Meta-theory of Bayesian Networks

When learning a Bayesian network from data one might reconstruct the structure of the network, the probability distributions, and even the distributions over hidden variables. Crucially, though, one needs to settle beforehand on the variables and their state space. This is to be contrasted with how generative models in the Bayesian brain hypothesis are actually constructed: Here, one somehow needs to 'learn' new variables and the values they can take, both for potential causes and their observable manifestations. The question then arises *when* a Bayesian learner realizes that the current model is insufficient and new hypotheses should be formed, as well as *what* these hypotheses should look like [3]. This problem comes on top of 'normal' model revision by Jeffrey updating [19], where just the probability distributions are updated in the light of new evidence; this aspect of model revision can be elegantly related to predictive processing concepts such as precision-weighted prediction error [32]. The problem of adding new variables and values of variables to a network in the light of unresolvable prediction error is a major open problem.

When a prediction error is to be accounted for, one can either update ones current beliefs about the actual hypotheses or try to reduce uncertainty by observing hidden variables. These predictive processing sub-processes (belief revision and adding observations) correspond to aspects of parameter tuning and sensitivity analysis [6] and selecting evidence [38]. Algorithmic and analytical aspects of these problems are of direct relevance to the Bayesian Brain hypothesis.

A vital open problem in the predictive processing account relates to the trade-off between making predictions that are very *detailed* and predictions that are likely to be *correct*. For example, when predicting the outcome of a throw at a bowling lane, a prediction over a distribution containing values like 'pin four will be hit by the ball from the left side and will topple over pins seven and eight' is very detailed, but probably always gives a huge prediction error. On the other hand, a prediction like 'the ball will hit the pins and some will fall' is likely to be correct, but as a prediction not very informative. There are reasons to believe that particular neurotransmitters (in particular serotonin) control this *level of detail* [35], but from a more meta-perspective it is completely open how causal Bayesian models can be 'flexible' in their granularity and how algorithms on such models may trade-off information gain and prediction error.

6 Potential Pitfalls

In the previous sections we highlighted several research areas and tentative research questions where the PGM community can substantially contribute to the 'Bayesian Brain' with a potential for considerable impact. Notwithstanding this potential, there are also pitfalls to avoid that are inherent risks of interdisciplinary work, in particular when the research fields have different cultures and tradition and use specific terminology that may be misunderstood. Here we enumerate a few potential pitfalls.

- '**Terminology**' — An informal quiz at the interdisciplinary Lorentz Center workshop 'Perspectives on Human Probabilistic Inference'[2] on the association that participants had with the word 'Bayesian' was illuminative to us. For some participants *Bayesian* was a synonym of *probabilistic*, for others it concerned the semantics of probability distributions (*subjective*, as contrasted with *frequentist*), yet others associated *Bayesian* with *Bayes' rule* for updating distributions. Despite the traditional interpretation of 'Bayesian' as 'subjective degrees of belief' [18], it is not

[2] http://www.lorentzcenter.nl/lc/web/2014/627/info.php3?wsid=627&venue=Oort

uncommon for proponents of the Bayesian Brain hypothesis to have a strong frequentist view on probabilities as describing the objective state of the world [8]. Similarly diverse (and sometimes counterintuitive) associations could be elicited for terms like 'prior', 'uncertainty', 'information', and 'structure'. The bottom line is to be aware of potential misunderstandings and to be explicit of one's intended meaning of such terms in communication with neuroscientists.

- '**Culture and tradition**' — In computer science and artificial intelligence, acceptance of a paper to a prestigious conference such as AAAI, UAI, FOCS or STOC is distinctive. Many scholars focus their publication strategy on such conferences, rather than journal papers. In neuroscience, a conference publication is close to irrelevant when it comes to evaluating research output; much more emphasis is put on the impact factor of the journals one is publishing in. Culture and tradition put emphasis on different 'golden standards' of excellence in research, validity of research methodology, and importance of research topics. Awareness of such issues and an open mind may help avoid or solve misunderstandings.

- '**Interdisciplinary**' — Members of interdisciplinary teams have different backgrounds and distinct areas of expertise; that is exactly the main benefit of having interdisciplinary collaborations at all. There is a fine line between 'nitpicking on details' versus 'allowing crucial misconceptions to exist' in interdisciplinary collaborations, and it requires some expertise to see what is important and what not. For example, it is rarely important to insist on the distinction between NP-hardness and NP-completeness of a problem, but the difference between an observation and an intervention in (causal) Bayesian networks may well be important to clarify. Don't assume your neuroscience collaborators share your background, and don't be afraid to ask for clarification about what seems obvious to them.

- '**Selling your work**' — An elegant intractability proof or a new formalization of a verbal theory is typically not sufficient for publication in neuroscience outlets. In order to get published one should aim to understand the problems that neuroscientists care about, make clear why your contribution is instrumental in solving these problems, and write in a way that connects to their background and expectations. It might be difficult to convince one's departmental chair or (grant) reviewers of the relevance of this work. Our approach is to seek for niches that both allow for a significant PGM contribution *and* solve crucial problems with respect to the Bayesian Brain.

7 Conclusion

Despite the potential pitfalls we identified in the previous section, we strongly believe computer scientists and AI practitioners working in the PGM area can make a vital interdisciplinary contribution to contemporary theoretical neuroscience. With this paper we hope to have given an overview of crucial open problems in the Bayesian Brain hypothesis and a sketch of the contributions that the PGM community can offer. We conclude this paper with this quote from Karl Friston [11] that (probably inadvertently) illustrates the importance of research on probabilistic graphical models for theoretical neuroscience: *Life (. . .) is an inevitable and emergent property of any (ergodic) random dynamical system that possesses a Markov blanket.* We would like to invite the community to bring their toolbox of computational and formal modeling and help to advance this fascinating research area — who knows what else may emerge!

References

[1] C. Bielza and P. Larrañaga. Bayesian networks in neuroscience: A survey. *Frontiers in Computational Neuroscience*, 8:Article 131, 2014.

[2] L. Buesing, J. Bill, B. Nessler, and W. Maass. Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLoS Computational Biology*, 7(11):e1002211, 2011.

[3] C. D. Carroll and C. Kemp. Hypothesis space checking in intuitive reasoning. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, 2013.

[4] A. Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204, 2013.

[5] A. Clark. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, 2015.

[6] V. M. H. Coupé, F. V. Jensen, U. B. Kjærulff, and L. C. van der Gaag. A computational architecture for n-way sensitivity analysis of Bayesian networks. Technical report, Aalborg University, 2000.

[7] P. Dagum and M. Luby. Approximating probabilistic inference in Bayesian belief networks is NP-hard. *Artificial Intelligence*, 60(1):141–153, 1993.

[8] C.D. Fiorillo. Beyond Bayes: On the need for a unified and Jaynesian definition of probability and information within neuroscience. *Information 2012, 3(2)*, 3(2):175–203, 2012.

[9] K.J. Friston. The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7):293–301, 2009.

[10] K.J. Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[11] K.J. Friston. Life as we know it. *Journal of the Royal Society Interface*, 10(86):20130475., 2013.

[12] K.J. Friston and K. E. Stephan. Free-energy and the brain. *Synthese*, 159:417–458, 2007.

[13] S. Habenschuss, Z. Jonke, and W. Maass. Stochastic computations in cortical microcircuit models. *PLoS Computational Biology*, 9(11):e1003037, 2013.

[14] T.J. Hamilton, S. Afshar, A. van Schaik, and J. Tapson. Stochastic electronics: A neuro-inspired design paradigm for integrated circuits. *Proceedings of the IEEE*, 5:843–859, 2014.

[15] Y. Haxhimusa, I. van Rooij, S. Varma, and H. T. Wareham. Resource-bounded problem solving (dagstuhl seminar 14341). *Dagstuhl Reports*, 4(8), 2014.

[16] J. Hohwy. *The Predictive Mind*. Oxford University Press, 2013.

[17] E.T. Jaynes. How does the brain do plausible reasoning? In G. J. Erickson and C. R. Smith, editors, *Maximum-Entropy and Bayesian Methods in Science and Engineering*, 1988.

[18] E.T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[19] R.C. Jeffrey. *The Logic of Decision*. University of Chicago Press, 1965.

[20] J. W. Kay and W. A. Phillips. Coherent infomax as a computational goal for neural systems. *Bulletin of Mathematical Biology*, 73(2):344–372, 2011.

[21] J. Kwisthout. Minimizing relative entropy in hierarchical predictive coding. In L.C. van der Gaag and A.J. Feelders, editors, *Proceedings of PGM'14*, LNCS 8754, pages 254–270, 2014.

[22] J. Kwisthout. Tree-width and the computational complexity of map approximations in bayesian networks. *Journal of Artificial Intelligence Research*, 53:699–720, 2015.

[23] J. Kwisthout. The parameterized complexity of approximate inference in Bayesian networks. In *Proceedings of Machine Learning Research*, volume 52, 2016.

[24] J. Kwisthout, H. Bekkering, and I. van Rooij. To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions. *Brain and Cognition*, in press.

[25] J. Kwisthout and I. van Rooij. Predictive processing and the Bayesian brain: Intractability hurdles that are yet to overcome. *Journal of Mathematical Psychology*, under review.

[26] M. L. Littman, J. Goldsmith, and M. Mundhenk. The computational complexity of probabilistic planning. *Journal of Artificial Intelligence Research*, 9:1–36, 1998.

[27] W. Maass. Neural computation: a research topic for theoretical computer science? Some thoughts and pointers. In *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, volume 72. 2000.

[28] W. Maass. Noise as a resource for computation and learning in networks of spiking neurons. *Proceedings of the IEEE*, 102(5):860–880, 2014.

[29] D. Marr. *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman, 1982.

[30] M. Otworowska, J. Kwisthout, and I. van Rooij. Counter-factual mathematics of counterfactual predictive models. *Frontiers in Consciousness Research*, 5:801, 2014.

[31] M. Otworowska, J. Riemens, C. Kamphuis, P. Wolfert, L. Vuurpijl, and J. Kwisthout. The robohavioral methodology: Developing neuroscience theories with FOES. In *Proceedings of the 27th Benelux Conference on AI (BNAIC'15)*, 2015.

[32] M. Otworowska, L. Zaadnoordijk, E. de Wolff, J. Kwisthout, and I. van Rooij. Causal learning in the crib: A predictive processing formalisation and babybot simulation. In *Proceedings of the Sixth Joint IEEE International Conference on Developmental Learning and Epigenetic Robotics*, 2016.

[33] D. Pecevski, L. Bueling, and W. Maass. Probabilistic inference in general graphical models through sampling in stochastic networks of spiking neurons. *PLoS Computational Biology*, 7(12):1–25, 2011.

[34] W.A. Phillips. Cognitive functions of intracellular mechanisms for contextual amplification. *Brain and Cognition*, in press.

[35] S. Pink-Hashkes and J. Kwisthout. A predictive processing account of psychedelia. In *Interdisciplinary Conference on Psychedelics Research*, 2016.

[36] J. B. Tenenbaum, C. Kemp, T.L. Griffiths, and N.D. Goodman. How to grow a mind: Statistics, structure, and abstraction. *Science*, 331:1279–1285, 2011.

[37] C. Thornton. Predictive processing simplified: The infotropic machine. *Brain and Cognition*, in press.

[38] L. C. van der Gaag and H. L Bodlaender. On stopping evidence gathering for diagnostic Bayesian networks. In W. Liu, editor, *Proceedings of the Eleventh European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 6717 of *LNCS*, pages 170–181, 2011.

[39] H. von Helmholtz. *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss, 1867.

The Dual Codebook: Combining Bags of Visual Words in Image Classification

Jonathan L. Maas [a] Emmanuel Okafor [a] Marco A. Wiering [a]

[a] *Institute of Artificial Intelligence and Cognitive Engineering. University of Groningen, Nijenborgh 9, Groningen, The Netherlands*

Abstract

In this paper, we evaluate the performance of two conventional bag of words approaches, using two basic local feature descriptors, to perform image classification. These approaches are compared to a novel design which combines two bags of visual words, using two different feature descriptors. The system extends earlier work wherein a bag of visual words approach with an L2 support vector machine classifier outperforms several alternatives. The descriptors we test are raw pixel intensities and the Histogram of Oriented Gradients. Using a novel Primal Support Vector Machine as a classifier, we perform image classification on the CIFAR-10 and MNIST datasets. Results show that the dual codebook implementation successfully utilizes the potential contributive information encapsulated by an alternative feature descriptor and increases performance, improving classification by 5-18% on CIFAR-10, and 0.22-1.03% for MNIST compared to the simple bag of words approaches.

1 Introduction

In this paper, we propose and evaluate the use of a Dual Bag Of visual Words model (Dual-BOW) in a relatively conventional framework to perform image classification. Within computer vision, there are many approaches that have been used to create image recognition systems [12]. The challenge which renders classic conventional machine learning techniques inaccurate revolve around representing and encapsulating the essential and unique features of an object or entity, which may occur rotated, scaled, illuminated, or oriented differently.

A popular approach which can encapsulate this is known as the bag of visual words (BOW) [4], which has been shown to reach good performances on multiple tasks [2, 3] and is also simple in design. Recently, improvement with the use of a bag of visual words with local feature descriptors has been applied in domains such as facial recognition, character, animal and object recognition. This is evident in the works of [16] whereby the bag of visual words with the histogram of oriented gradient (HOG-BOW) showed a superior performance relative to other local feature descriptors on different character datasets. Also, the authors in [8] showed that the use of HOG-BOW outperforms several classical based methods such as HOG, SIFT, and a multi-subregion based correlation filter bank (MS-CFB) on a facial dataset (FERET). Though studies by authors in [14] have shown that the combination of several feature descriptors which they called the Joint learning framework outperforms the BOW [14] and HOG-SIFT-BOW [11] approaches.

In this paper, we show the superiority of a bag of visual words with the combination of two local feature descriptors by creating a dual codebook which contains both local features (Dual BOW) compared to the conventional bag of visual words methods (BOW and HOG-BOW) with a single codebook. Our goal of this study is to research the additional effect of combining two bags of words, using different local feature descriptors (LFD). Under the notion that different feature descriptors may encapsulate different essential information, we will assess the performance increase (if any) of combining this information with respect to the conventional bag of visual words, which utilizes only single feature descriptors.

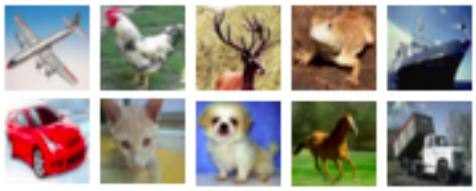


Figure 1: Samples from the CIFAR-10 dataset.

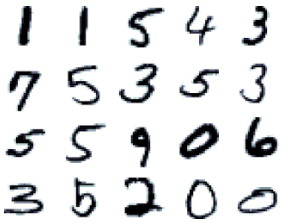


Figure 2: Samples from the MNIST dataset.

Outline This paper is organized as follows: in Section 2, we describe the datasets used in our experiment. Hereafter, we discuss the system design, local feature descriptors, bag of word models, and the used classifier in Section 3. Having covered the basis of our implementation, we will discuss our experiments and results in Section 4. Lastly, we discuss our findings in the conclusion in Section 5.

2 Datasets

For our experiment, we decided to use two datasets to achieve a more reliable assessment of the potential benefit of our proposed approach. Two popular, and diverse, benchmarks datasets often used in this field are the MNIST and CIFAR-10 datasets. MNIST [10] (see Figure 2) consists of 70,000 (60,000 training, 10,000 testing) 28 x 28 pixel images of 10 classes of digits. Though often considered a simplistic dataset, it remains a popular benchmark and provides plenty research to compare with. CIFAR-10 [9] (see Figure 1) consists of 60,000 (50,000 training, 10,000 testing) 32 x 32 colour images, constructed from 10 more diverse classes (ranging from animals to vehicles).

The bag of words approach we work with relies on the extraction of so called patches, sub-parts of the image, that can be extracted using a sliding window of a fixed size. For MNIST, the images were rescaled (using cubic interpolation) to an image resolution of 48 x 48 pixels, after which patches of 14 x 14 pixels were extracted. For CIFAR-10, smaller patchsizes of 8 x 8 were more appropriate, as patch size appeared to have a large impact depending on the dataset used. The image size remained unchanged at 32 x 32 pixels.

3 System Design

The system design builds upon the framework used in [16], wherein a bag of visual words is used, and the performance of several different local feature descriptors was evaluated. Herein, they also compare the performance of several types of support vector machines.

We designed our system with flexibility in mind, as such that it enables swapping different local feature descriptors[1] to be used in combination with bag of word approaches, allowing different patch sizes, and implementation methodologies.

3.1 Conventional Bag of Words

The Bag of Visual Words has been a popular tool in computer vision and classification [4], wherein an image is represented by regarding the patches that it is composed of. Patches are described by an appropriate local feature descriptor, which is used to construct their patch-features. Using this methodology, one can create a bag of words by applying an unsupervised algorithm (such as K-means clustering [13]) on a random collection of patches, extracted from images from the training set. The resulting centroids are intended to represent generalized patches, or visual words, and as a whole act as a dictionary (which we refer to as a codebook within the context of this paper), representing which visual elements are acknowledged to exist and occur in the data [16].

Once the codebook is constructed, it can be used to represent a new image. This is done by partitioning a given image N into S (non-overlapping) segments, of equal size. Within every segment, n

[1] For the experiment, however, only two local feature descriptors were used. We also intended to include a local binary patterns feature descriptor, but at the time did not possess the computational resources to include it in our research.

patches are extracted using a sliding window of a custom size and shift. The derived set of patches are then described by feature vectors using the appropriate local feature descriptor.

Hereafter, the activations are computed in the following fashion. For every patch-feature $p_i \in \mathbb{R}^n$ from the collection of patches within a segment, distances are computed to each word $w_j \in \mathbb{R}^n$ from a codebook $C^l = \{w_1, w_2 ... w_K\}$ (where $l \in \{IMG, HOG, DUAL\}$ denotes the appropriate feature descriptor), using a distance function $d\left(p_i, w_j\right)$. In our experiment, we used the Euclidean distance as distance function:

$$d\left(p_i, w_j\right) = \sqrt{\sum_{x=1}^{n} \left(p_i^x - w_j^x\right)^2} \tag{1}$$

to represent the distance from a patch p from an image to centroid w from the codebook, over all elements of its feature vector length.

Computing the distance to all words allows us to compute the mean distance of patch p_i to all words:

$$\bar{d}(p_i, w) = \frac{\sum_{j=1}^{K} d(p_i, w_j)}{K} \tag{2}$$

We will compute the cluster activations according to the Soft-Assignment function [3], by updating the activation vector $a_j \in \mathbb{R}^K$, which denotes the activations of the codebook centroids with respect to the patches within the segment. For every patch $p_i \in \mathbb{R}^n$, the activation value (a_j) of word w_j is updated by:

$$a_j = \begin{cases} a_j, & \text{if } \delta \leq 0 \\ a_j + \delta, & \text{if } \delta > 0 \end{cases} \tag{3}$$

Where $\delta = \bar{d}(p_i, w) - d(p_i, w_j)$ (and corresponds to a similarity measure between a patch and a word). Repeating this procedure for every patch within segment $s \in \mathbb{R}^S$ gradually generates an activation vector for segment s:

$$A_s(K) = \{a_1, a_2, ... a_K\} \tag{4}$$

To create the final feature vector, $x_{_N}^l$, representing a given image N, using codebook l (and its corresponding local feature descriptor), the activations of all S segments of the image are concatenated:

$$x_{_N}^l(s) = \{A_1; A_2; ... A_S\} \tag{5}$$

and standardised once.

The resulting final feature vector can be used as training and testing data for any classifier of choice. Obviously, computational complexity in this approach grows with feature descriptor size, and the number of centroids used. The dimensionality of the final feature vector of the image, corresponds to $S * K$, where S corresponds to the number of segments the image is partitioned in, and where K is the number of centroids in the codebook used. The codebooks we used in our approach are generated using 200,000 patches randomly extracted from the dataset used, and clustered (using K-means Clustering) using 150 iterations. Having described the bag of words approach, we now describe the methods used in our experiment.

3.1.1 Bag of Visual Words with Pixel Intensities (BOW)

In its most conventional implementation, the bag of visual words approach uses patches described by their raw pixel intensities. The raw pixel intensities method directly uses the RGB intensities of the pixels within a patch. Simple as it may be, its successes in several tasks have shown its potential [15], and show that raw pixel intensities within patches can be used to represent interesting features. Nevertheless, the feature vector length can grow very large when larger patches are used, especially in colour images (which is the case for the 3-channel CIFAR-10 dataset, as opposed to the single-channel MNIST dataset).

In our experiments, for MNIST, the patch size of 14 x 14 pixels results in a patch-feature size of 196 elements. For CIFAR-10, however, we need to track three colour channels of a 8 x 8 pixel patch, which results in a patch-feature length of 192. After computing the patch-feature vector, it is standardised.

Though we included modules for performing different levels of pre- and postprocessing, we settled on using only standardisation where appropriate. Standardisation of a vector is performed by computing the mean of its elements:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{6}$$

Then, the deviation is computed by:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (\overline{x} - x_i)^2}{n} + e} \tag{7}$$

Where e is used as a small constant to avoid a zero standard deviation. Then the standardised vector is obtained by updating the vector values:

$$x'_i = \frac{x_i - \overline{x}}{\sigma} \tag{8}$$

We used this standardisation scheme on several occasions within the design. For our experiment, we ran two configurations of the IMG implementation, one using 400 centroids for its codebook, the other using 800. Images are partitioned into 9 segment (3 x 3), and this results in a final feature dimensionality of 3,600 and 7,200 for the 400 and 800 centroids approach respectively.

3.1.2 Bag of Visual Words with Histogram of Oriented Gradients (HOG-BOW)

An alternative to the raw pixel intensities is to use the Histogram of Oriented Gradients (known as HOG) to describe patches. The Histogram of Oriented Gradients [5] has been a popular feature descriptor for a long while, and knows several different uses [17, 16]. To compute the descriptor, gradient components are computed for the horizontal and vertical gradient (G_x and G_y respectively) for every pixel in the patch. Though multiple masks can be used, the simple kernel $[-1, 0, +1]$ bears preference [1]. The gradients are computed with:

$$G_x = f(x+1, y) - f(x-1, y) \tag{9}$$

$$G_y = f(x, y+1) - f(x, y-1) \tag{10}$$

where $f(x, y)$ is the pixel intensity at coordinate x,y. The final Magnitude $M(x, y)$ (intensity of change) and orientation $\theta(x, y)$ (direction of change) are computed as:

$$M(x, y) = \sqrt{G_x^2 + G_y^2} \tag{11}$$

$$\theta(x, y) = \tan^{-1} \frac{G_y}{G_x} \tag{12}$$

After computing the magnitudes and orientations for every pixel, the patch is segmented into four quadrants.Within each quadrant, the magnitudes of all pixels are binned using linear interpolation (thus the binned magnitude is distributed over the neighbouring bins) into a histogram by the corresponding orientations, which produces the Histogram of Oriented Gradients. After computing the histograms of all four quadrants, these are concatenated to produce the feature vector representing the patch.

For our experiment, we used 9 bins to represent orientations in a range of $0 - 180^o$ (thus a bin width of 20 degrees). Since the patch sizes do not determine the HOG's feature vector size, the feature vector length for MNIST is 36. For the tri-colour channel CIFAR-10, it is 108.

For MNIST, a patch size of 14 x 14 pixels is reduced to 12 x 12 to cope with padding, after which HOG is computed for four 6 x 6 pixel cells. For CIFAR-10, a patch size of 8 x 8 pixels is reduced to 6 x 6 for the same reason, and the HOG is computed for four 3 x 3 pixel cells. As with the raw pixel intensities local feature descriptor, the HOG feature vector is also standardised.

For our experiment, we also ran two configurations of the HOG-BOW implementation, mirrorring the IMG runs with one implementation using 400 centroids for the codebook, the other 800. The final feature dimensionality remains unchanged at 3,600 (for 400 centroids) and 7,200 (for 800 centroids) since the final feature length is unaffected by the length of the local feature descriptor used.

3.2 Dual Bag of Visual Words: combining Pixel Intensity and Histogram of Oriented Gradients (Dual BOW)

We propose the combination of both the raw pixel intensities and HOG features to develop a dual codebook. This enigma of combining features within the scope of the visual bag of words approach knows little prior research [6]. In essence, the dual codebook is the combination of two codebooks, which may have been generated either using the same local feature descriptor (possibly under a different configuration), or an entirely different one. The configuration of the second codebook is not bound by those used in the first, and thus may also operate with a different number of centroids.

In this fashion, given two codebooks C^{IMG} and C^{HOG} (generated using raw pixel intensities, and the histogram of oriented gradients respectively), an image N is represented by computing the activations, x_N^l, for both codebooks towards this image. The activation vectors obtained, x_N^{IMG} and x_N^{HOG} are then concatenated:

$$x_N^{DUAL} = \left\{ x_N^{IMG}; x_N^{HOG} \right\} \tag{13}$$

to create the final feature vector of the image under the dual codebook approach.

This approach effectively allows the combination of two different local feature descriptors, which can aid classification accuracy by the inclusion of potentially essential information which may be encapsulated by the one, but not the other feature descriptor.

In our experiment, the dual codebook was evaluated under the same configurations as its singular alternatives, and combines two codebooks of 400 centroids each. This configuration therefore results in a final feature vector with a dimensionality of 7,200. Based on the dual codebook used in this section, the new bag of visual word formed is referred to as Dual-BOW.

3.3 Classifier

For classification, we designed an L2 'primal' support vector machine (one for each class) using a revised objective function:

$$\min_{\omega,b} L = \|\omega\|^2 + C \cdot \sum_N \xi_N^2 \tag{14}$$

and output function:

$$g\left(x_N\right) = \omega \cdot x_N + b \tag{15}$$

where $x_N = x_N^l$ denotes the centroid activations from the bag of words, using descriptor l, and the error is represented as:

$$\xi_N = \max(0, 1 - y_N \cdot g(x_N)) \tag{16}$$

$y_N \in \langle -1, 1 \rangle$ represents whether the target label of example x_N belongs to the class which this SVM represents.

Training is done in iterations, and all training data are presented in each iteration. For every iteration, if the output doesn't perfectly predict the class ($y_N \cdot g(x_N) < 1$), then the weights are adjusted using the formula:

$$\Delta w_j = -\lambda \cdot (\frac{w_j}{C} - (y_N - g(x_N)) \cdot x_N^j) \tag{17}$$

Where λ denotes the learning rate. At the end of every iteration, the bias b is updated to represent the mean error $y_N - g(x_N)$ of all examples where $y_N \cdot g(x_N) < 1$.

We used an L2 primal Support Vector Machine, with a learning rate λ of 0.0000001, and performed 2000 training iterations before testing. The initial weight values are 0.000002, and C is set to 2048.

4 Results

In total, for both MNIST and CIFAR-10, we designed 5 experiment configurations. For the single bag of word approaches (BOW and HOG-BOW) we performed runs with codebooks of 400 and 800 centroids, whereas the dual codebook implementation was run with two codebooks of 400 centroids each.

We performed 10-Monte Carlo cross validation runs for every of the 5 configurations (BOW-400, BOW-800, HOG-BOW-400, HOG-BOW-800, DUAL-2x400). The results are shown in Table 1.

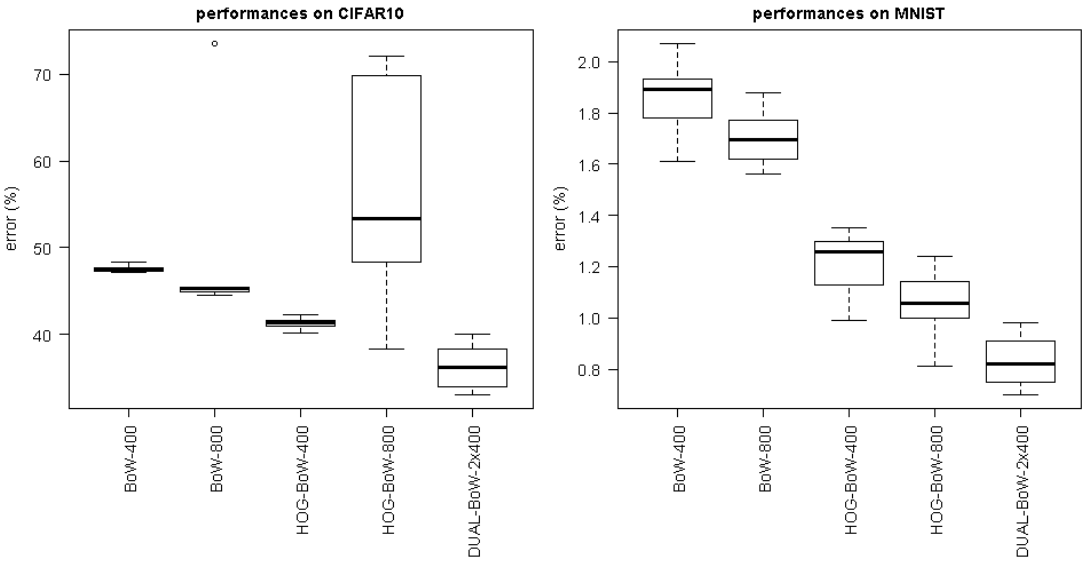


Figure 3: Error rates for CIFAR-10(left) and MNIST(right)

Methods	**MNIST**		**CIFAR-10**	
	Mean	SD	*Mean*	SD
BOW-400	1.85	0.14	47.59	0.42
BOW-800	1.71	0.10	47.96	9.00
HOG-BOW-400	1.22	0.12	41.28	0.61
HOG-BOW-800	1.05	0.13	54.98	12.64
Dual-BOW-2x400	0.83	0.09	36.20	2.60

Table 1: Classification Error (in %) on test-sets of MNIST and CIFAR-10, 10-fold Monte Carlo Cross Validations.

4.1 Evaluation of the CIFAR-10 Dataset

The results of classification on the CIFAR-10 dataset can be seen in Table 1 (and is visualized in Figure 3). As shown, the dual codebook reaches commendable classification performance. Though not stellar nor exceeding present state-of-the-art performance [7], the results still reflect the added value of the dual codebook, resulting in a significant performance increase compared to all single codebook variants.

Student's T-tests shows the dual codebook (Dual-BOW) performs better than the Histogram of Oriented Gradients with 400 centroids ($t = 6.01$, $p < 0.05$), outperforms the 800-centroid variant ($t = 4.60$, $p < 0.05$), and surpasses both 400 and 800-centroid raw pixel intensities (conventional BOW) implementations ($t = 13.26$, $p < 0.05$ and $t = 3.97$, $p < 0.05$, respectively).

Therefore, on CIFAR-10, the Dual-BOW approach, which employs the dual codebook, appears superior to both BOW and HOG-BOW which use only a single codebook, because it obtains the lowest error rate.

4.2 Evaluation of the MNIST Dataset

Though performance improvements may not be as pronounced as those in CIFAR-10, the dual codebook again significantly outperforms all single codebook configurations (see Table 1, and Figure 3).

Student's T-test indicate that the dual codebook approach (Dual-BOW) displays significant improvements over HOG-BOW-400 and HOG-BOW-800 ($t = 8.26$, $p < 0.05$ and $t = 4.50$, $p < 0.05$ respectively). With regard to raw pixel intensities, the Dual-BOW approach significantly outperforms both the BOW-400 ($t = 19.01$, $p < 0.05$) and BOW-800 ($t = 19.97$, $p < 0.05$) implementations.

Thus, the results on MNIST confirm those of CIFAR-10, showing that the Dual-BOW again outperforms conventional BOW approaches utilizing only single codebooks.

4.3 Discussion of Results

In this paper, we have demonstrated the dual codebook's superiority over comparable single codebook approaches, showing a consistent performance improvement over two substantially different datasets. This implies the capability of successfully combining the essential information encapsulated by different local feature descriptors, improving classification performance.

Though both the datasets and the approach used may be considered simplistic by current standards, it does not appear that the dual codebook approach would perform worse with alternative datasets, than single codebook alternatives would.

5 Conclusion

Though performance on either dataset is not present state-of-the-art, it should be kept in mind that many of the data-preprocessing enhancements and excessive parameter tuning conventionally performed for these datasets were not applied, as we intended to study the exclusive benefit of the dual codebook approach, with regard to conventional bag of words approaches that utilize only a single codebook. Therefore, these results say little about the limits of the dual codebook approach, which was used in a quite simple configuration in this experiment. Under slightly more computationally demanding configurations of the primal SVM, performance for CIFAR-10 for the dual codebook reached scores up to 73.18%, and for MNIST up to 99.3%. However, these results were discarded under the need to perform cross validations with limited computational resources, and time constraints.

With regard to future research, there are many possibilities. We intend to expand the design to an N-codebooks implementation, which will be able to combine N bags of words in order to investigate if this can increase performance further.

Additionally, it might be worth investigating the potential value of combining codebooks of the same feature descriptor, but under different configurations (for example, Histogram of Oriented Gradients with a different segmentation grid, or different bin distributions). Other grounds for further research could focus on the necessary sizes of the codebooks in regard to feature vector dimensionality, as it would be ideal if one were able to improve performance by incorporating a mere 100-centroid small extra codebook, which might be based on a local feature descriptor with a computational complexity or intensity too high to consider for larger codebooks.

In regard to the use of the L2 primal support vector machine as classifier, it proved to be more efficient to train than the conventional support vector machine implementation. Though a drawback still remains in an undeniable necessity for parameter optimization. Concerning computational intensity, one might consider the learning rate used (0.0000001) in combination with the number of iterations (2000).

We hope to develop an open framework[2] which combines not only easy modularity and flexibility of combining a number of codebooks, but also remains open to recycling of codebooks, exporting and importing centroids derived from previously trained codebooks, to allow the user to avoid the need to re-train the entire codebook.

References

[1] Jon Arrspide, Luis Salgado, and Massimo Camplani. Image-based on-road vehicle detection using cost-effective histograms of oriented gradients. *Journal of Visual Communication and Image Representation*, 24(7):1182 – 1190, 2013.

[2] A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. J. Wu, and A. Y. Ng. Text detection and character recognition in scene images with unsupervised feature learning. In *Proceedings of the 2011 International Conference on Document Analysis and Recognition*, pages 440–445, 2011.

[2]The framework is currently available online at https://github.com/JonathanMaas/nCodebooks

[3] A. Coates, H. Lee, and A.Y. Ng. An analysis of single-layer networks in unsupervised feature learning. In Geoffrey Gordon, David Dunson, and Miroslav Dudk, editors, *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *JMLR Workshop and Conference Proceedings*, pages 215–223. JMLR W&CP, 2011.

[4] G. Csurka, C. Bray, C. Dance, and L. Fan. Visual categorization with bags of keypoints. *Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22, 2004.

[5] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893 vol. 1, 2005.

[6] Huilin Gao, Wenjie Chen, and Lihua Dou. Image classification based on support vector machine and the fusion of complementary features. *CoRR*, abs/1511.01706, 2015.

[7] Benjamin Graham. Fractional max-pooling. *CoRR*, abs/1412.6071, 2014.

[8] Mahir Faik Karaaba, Olarik Surinta, L. R. B. Schomaker, and Marco A. Wiering. Robust face identification with small sample sizes using bag of words and histogram of oriented gradients. In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pages 582–589, 2016.

[9] A. Krizhevsky and G. Hinton. Learning multiple layers of features from tiny images. *Master's thesis, Department of Computer Science, University of Toronto*, 2009.

[10] C. LeCun, Y. Cortes. The mnist database of handwritten digits. 1998.

[11] Kart-Leong Lim and Hamed Kiani Galoogahi. Shape classification using local and global features. In *Image and Video Technology (PSIVT), Fourth Pacific-Rim Symposium on*, pages 115–120, 2010.

[12] D. Lu and Q. Weng. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5):823–870, 2007.

[13] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, pages 281–297, Berkeley, Calif., 1967. University of California Press.

[14] Bharath Ramesh, Cheng Xiang, and Tong Heng Lee. Shape classification using invariant features and contextual information in the bag-of-words model. *Pattern Recognition*, 48(3):894–906, 2015.

[15] O. Surinta, L. Schomaker, and M. Wiering. A comparison of feature and pixel-based methods for recognizing handwritten Bangla digits. In *12th International Conference on Document Analysis and Recognition*, pages 165–169, 2013.

[16] Olarik Surinta, Mahir F. Karaaba, Tusar K. Mishra, Lambert R. B. Schomaker, and Marco A. Wiering. *Recognizing Handwritten Characters with Local Descriptors and Bags of Visual Words*, pages 255–264. Springer International Publishing, Cham, 2015.

[17] Kazuhiko Takahashi, Sae Takahashi, Yunduan Cui, and Masafumi Hashimoto. *Remarks on Computational Facial Expression Recognition from HOG Features Using Quaternion Multi-layer Neural Network*, pages 15–24. Springer International Publishing, Cham, 2014.

Predicting civil unrest by categorizing Dutch Twitter events

Rik van Noord [a] Florian A. Kunneman [b] Antal van den Bosch [b]

[a] *Institute of Artificial Intelligence, University of Groningen*
[b] *Centre for Language Studies, Radboud University*

Abstract

We propose a system that assigns topical labels to automatically detected events in the Twitter stream. The automatic detection and labeling of events in social media streams is a 'big data' problem. The early detection of future social events, specifically those associated with civil unrest, has a wide applicability in areas such as security, e-governance, and journalism. We used machine learning algorithms and encoded the social media data using a wide range of features. Experiments show a high-precision (but low-recall) performance in the first step. We designed a second step that exploits classification probabilities, boosting the recall of our category of interest, social action events.

1 Introduction

Many instabilities across the world stem from civil unrest, often involving crowd actions such as mass demonstrations and protests. A prime example of a mass crowd action in the Netherlands was the *Project X* party in Haren, Groningen, on September 21st 2012. A public Facebook invitation to a birthday party of a 16-year old girl ultimately led to thousands of people rioting.[1] The riots could only be stopped by severe police intervention, resulting in more than 30 injuries and up to 80 arrests. Afterwards it was concluded that the police were insufficiently prepared and that they were not well enough informed about the developments on social media. An evaluation committee recommended the development of a nation-wide system able to analyze and detect these threats in advance.[2] In this paper, we describe a system that leverages posts on Twitter to automatically predict such civil unrest events before they happen.

To facilitate this objective we start from a large set of open-domain events that were automatically detected from Twitter from a period spanning multiple years by the approach described in [7]. From this set we aim to identify the events that might comprise civil unrest, henceforth *social action events*. Instead of focusing on this event type only, we categorize all events into a broad categorization of events, and distinguish social actions as one of the event types.

2 Related Work

2.1 Predicting social action events

Only a few works aim to detect social action events. [3] try to predict civil unrest in South America based on Twitter messages. In contrast to our approach, they predict such events directly from tweets, by matching them with specific civil unrest related keywords, a date mention, and one of the predefined locations of interest. Their system obtains a precision of 0.55 on a set of 283 predefined events. The main drawback of their approach is that it has no predictive abilities. For example, the system is not able to detect social action events that use newly emerging keywords for a specific event, or take place in a new location. As a consequence, their system likely has a low recall; many future social actions are likely to go undetected.

A more generic approach to detecting social action events is the EMBERS system by [13]. They try to forecast civil unrest by using a number of open source data sources such as Facebook, Twitter, blogs,

[1] http://www.nu.nl/binnenland/2915769/facebook-feest-haren-ontaardt-in-chaos-en-rellen.html
[2] http://nos.nl/artikel/482043-cohen-fouten-politie-burgemeester.html

news media, economic indicators, and even counts of requests to the TOR browser.[3] Using multiple models, the system issues a warning alert when it believes a social action event is imminent. Tested over a month, the EMBERS system obtained a precision and recall of respectively 0.69 and 0.82. In follow-up work, [10] report on the results of EMBERS when only taking Twitter information into account, mentioning a precision of 0.97 but a recall of 0.15.

2.2 Categorizing events

Some approaches based on Twitter perform some form of broad categorization (e.g. [15]; [12]). In these approaches there is no event detection procedure before dividing the data into different categories. The described approaches either identify which topics are often talked about on Twitter, or focus on the categorization of users instead of events. To our knowledge, the only approach that focuses on the categorization of automatically detected events is by [14]. They apply Latent Dirichlet Allocation [2] to a set of 65 million events to generate 100 topical labels automatically. Manual post-annotation winnowed these down to a set of 37 meaningful categories. 46.5% of the events belong to one of these categories, while 53.5% of the events are in a rest category. [14] compared their unsupervised approach to categorizing Twitter events to a supervised approach. They selected the best 500 events (detected with the highest confidence) and manually annotated them by event type. Their unsupervised approach obtained an F1-score of 0.67, outperforming the supervised approach which obtained an F1-score of 0.59. However, they do show that the F1-score of the supervised approach steadily increases when using more training instances.

3 Experimental Set-up

Our study starts with a set of automatically detected events from Twitter, described in Section 3.1.1. We manually annotate a subset of these events by type, and subsequently train a machine learning classifier on several feature types extracted from these events. Performance is both evaluated on the annotated event set and on the larger set of remaining events.

3.1 Data

3.1.1 Event set

To perform automatic event categorization, we use the event set described in [6] which was extracted based on the approach described in [7]. As this approach was applied to Dutch tweets, the set mainly comprises Dutch events. The event detection approach is based on the method of [14], who used explicit future time expressions to identify events. Each event has a set of attributes, such as the date, keywords and event score. The event score is linked to the size and popularity of an event. For the exact calculation of this score, we refer to [7, pp.13]. Over a 6-year period (2010-2015), [6] ultimately obtained 93,901 events. This event set is used for our categorization system.

3.1.2 Event annotations

We select two sets of events for categorization. Our first event set contains the 600 events with the highest event score in the output of [6]. This enables us to make an approximate comparison to [14], who evaluated their system on the basis of their best 500 events. We refer to the set of events with the highest event scores as the *best event set*.

Our second event set is created by randomly selecting an event from the ranked total event set for intervals of 155 events (with all events ranked by event score), excluding the best 600 events of the best event set. We refer to this event set as the *random event set*. Non-Dutch events were manually removed from both event sets, leaving 586 of the best events and 585 of the random events suitable for annotation.

Seven different annotators were involved in the annotation process, who all at least annotated 40 and at most 175 of the events in the best event set. 195 of the 586 best events received a double annotation so that we are able to calculate inter-annotator agreement. The other 390 events, as well as the 585 random

[3] An indication of the number of people who chose to hide their identity and location from the online community.

events, were annotated by one annotator. Similar to [14], the annotator is asked two questions for each event:

- Is this an actual event according to the definition?
- What is the category of this event?

An event in our full event set is not necessarily a proper event according to the definition, as the detection procedure makes errors. Since we are not interested in the category of a non-event, the events that are annotated as a non-event are filtered from the event set.

We defined ten possible categories after an initial manual inspection of about 200 events. They are listed in Table 1. *Social action* is the category of interest. As arguably less straightforward categories we included *special day* and *advertisement* because manual inspection of the data suggested that those types of events were frequent enough to deserve their own category.

Table 1: The ten different categories with examples.

Category	**Example event**
Sport	Soccer match, local gymnastics event
Politics	Election, public debate
Broadcast	Television show, premiere of a movie
Public event	Performance of a band, festival
Software	Release of game, release of new iPhone
Special day	Mother's Day, Christmas
Social action	Strikes, demonstrations, flashmobs
Celebrity news	Wedding or divorce of a celebrity
Advertisement	Special offers, retweet and win actions
Other	Everything that does not fit in one of the other categories

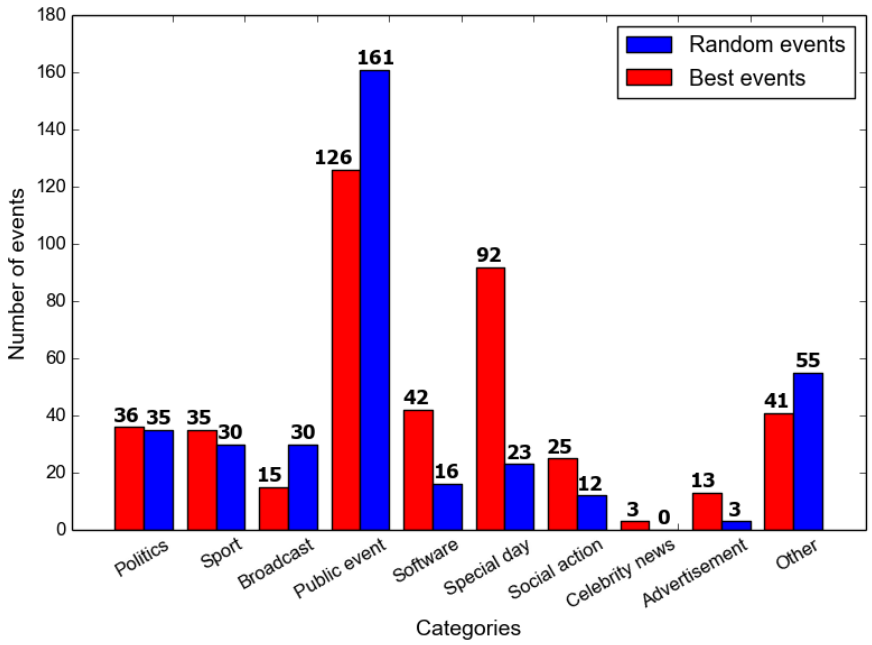


Figure 1: The ten different categories with the number of annotated examples in the best and random event set.

The 195 events annotated by two coders yielded a Krippendorff's alpha [5] of 0.81 on judging whether or not it was an actual event and 0.90 on categorizing events. These scores can be consid-

ered excellent and show that we can reliably view the events that were annotated once as if they were annotated correctly. Therefore, the 586 random events could be annotated once by two annotators.

Events that were (at least once) annotated as a non-event are removed from the event set, as well as events where annotators disagreed on the category. 27.4% of the best events were a non-event, leaving 425 of the best events. In the random event set 38.1% were discarded as non-events (leaving 362 events).

The annotations by event category are shown in Figure 1. *Public event* is the dominant category, comprising 29.6% of the best events and 44.5% of the random events. Most other event categories occur regularly, with the exception of *advertisement* and *celebrity news*. The latter category was so infrequent that it was removed from both event sets. *Advertisement* was removed from the random event set, but was retained for the best event set.

3.2 Training and testing

Based on the annotated events we trained a machine learning classifier to distinguish the ten event types. We describe the event features, classification approaches and evaluation below.

3.2.1 Feature extraction

Table 2: Types of extracted features

Tweet features	Tweet count	Bag-of-words	Sentiment		
Event features	Event-score	Keyword-score	Event date	Periodicity	Wikipedia

To enable the classifier to learn the specific properties of each event category we extract several types of features from each event. They are listed in Table 2, and can roughly be divided into 'Tweet features' and 'Event features'. Tweet features comprise characteristics of the tweets that refer to an event. We extract three types of features from them. First, the number of tweets, which might reflect the popularity of an event. We also distinguish between the numbers of tweets before, during or after[4] an event. Second, we extract each word used in the event tweets as a feature, jointly referred to as Bag-of-words features. As a third feature type we scored the average subjectivity and polarity of each event tweet, using the approach by [4]. The subjectivity and polarity score of the event are averaged over all event tweets. Some event types might be referred to fairly objectively in tweets, while others might stir more sentiment.

The first three types of Event features are derived from [7]. The first feature type, the event score, describes the link between the event keywords and the date of the event. This score gives an indication of the confidence that the set actually represents an event. Second, the keyword scores give an indication of the commonness of each event keyword, based on the commonness score as described in [8]. Third, the event date might help to recognize event types that are linked to big events, such as elections. The fourth feature type indicates whether an event is of a periodic nature. This feature is based on the output from a periodicity detection system described by [6]. Fifth, we employ DBpedia [1] in order to generalize over the different named entities present in the events. Since we want to generalize over the different terms, we are especially interested in the `type` attribute of the entity in DBpedia. This gives us a broader description of the terms than just a single value would have provided us with. For example, Feyenoord is a *SoccerClub*, *SportsTeam* and *Organisation*, while Justin Bieber is an *Artist*, *MusicalArtist*, *MusicGroup*, and a *NaturalPerson*. We extract the different DBpedia `types` for each event keyword. The keywords are linked to DBpedia using Wikification [9].

3.2.2 Classification

Based on the extracted feature sets along with the annotated categories, we train a Naive Bayes classifier[5] using the Python module Scikit-learn[6] [11]. We applied two methods to increase the performance of the

[4]Since we wanted to provide our system with as much training data as possible, we also extracted relevant tweets that were posted after the event took place. Obviously, when predicting events in the future, this type of data will be unavailable.

[5]In addition to Naive Bayes, we experimented with Support Vector Machines and K-nearest neighbors. We will only report on the outcomes of Naive Bayes, which yielded the best performance.

[6]`http://scikit-learn.org`

classifier: down-sampling the dominant *public event* class, and performing bag-of-words classification as a first-step classification. The first method simply reduces the number of *public events* in the training set to ensure it does not hinder the performance of the minority classes. The number of *public events* is reduced to the same frequency as the second-most frequent class in the event set, resulting in the deletion of 34 events in the best event set and 106 events in the random event set.

The second method only feeds the bag-of-words features to the classifier in an initial stage, and subsequently adds the resulting classification to the set of other features. The advantage of this method is that it reduces the feature space, while acknowledging the information in the word features. Also, it enables us to measure the impact of the non bag-of-words features in comparison to a bag-of-words baseline.

3.2.3 Evaluation

The performance on categorizing events is evaluated in two ways. The first is to apply 5-fold cross validation on the annotated sets of events. We do this for both the best event set (for a comparison with [14]) and the random event set, calculating the average precision, recall and F1-score.[7] The second way is to evaluate the results on a set that was never used in the training phase. The classifiers are trained on the two sets of annotated events and subsequently applied to the remaining 92,701 events. Performance on these unseen events is evaluated by manually inspecting a subset of them. In order to evaluate the precision of each classification category we randomly selected 50 events per category for evaluation, except for our category of interest, social action events, for which we include all 93 events classified with this category. *Advertisement* could only be evaluated for 25 events, as it was only predicted 25 times. This ultimately resulted in a total set of 468 events, which we refer to as the *Evaluation set*.

4 Results

4.1 Annotated set

Table 3 shows the most important results of the 5-fold cross validation. Averaged over all categories, the best event set obtained an F1-score of 0.65, while the random event set received an F1-score of 0.58. It appears to be easier to classify events with a higher event score.[8] However, we found no significant effect of event score when doing a least-squares logistic regression test for the random event set ($r(360) = -0.05$, $p = 0.39$). This suggests that there is a small subset of events with a very high event score that is easier to classify, but that there is no significant effect of event score in general.

Comparing the setting where only bag-of-words is used as a feature with the setting where the classification based on bag-of-words is added as a feature to the other features, the latter setting yields the best outcomes.

Social action is predicted at a high precision in the best events set, but the scores for the random events are poor. This might be due to the low number of instances in this set (12), in comparison with the 25 social actions in the best event set.

Table 3: The results of the 5-fold cross validation for the Naive Bayes algorithm while down-sampling the dominant *public event* class.[9]

		All categories			**Social actions**		
		Prec	**Rec**	**F1**	**Prec**	**Rec**	**F1**
Best events	Only bag-of-words	0.67	0.59	0.55	0.68	0.41	0.52
	Bag-of-words added as a feature	**0.67**	**0.67**	**0.65**	0.79	0.44	0.56
Random events	Only bag-of-words	0.61	0.60	0.57	0.40	0.17	0.24
	Bag-of-words added as a feature	**0.64**	**0.60**	**0.58**	0.40	0.17	0.24

[7] This was calculated by using the *weighted* setting in scikit-learn.

[8] Recall that the event score indicates a high confidence that the automatically detected unit is an event.

[9] Down-sampling increased the F1-score by 0.05 for the best event set and 0.06 for the random event set.

4.2 Evaluation set

Table 4 shows the results on the Evaluation set, listing the precision per category. In general, these scores are high for a 9-class classification task. The precision per class is even 1.00 for *sport* and *politics*, meaning that if the classifier predicted those categories, it did so perfectly. The categories *public event* and *advertisement* score below 0.70, however. The low precision for *public event* impacts the overall performance of the classification system substantially. As 81,538 out of 92,701 events were classified as a *public event*, a precision of 0.57 leads to about 35 thousand incorrectly classified events.

We should keep in mind that the non-events were not excluded from the full event set. It was estimated that 38.1% of all detected events are not events. In the training phase these non-events were excluded, so it is likely that the classifier will assign the non-events in the full event set to the most frequent category. A large part of the bias to *public event* may be due to the occurrence of non-events in the full event set. This leads us to conclude that if there were a more reliable way to automatically exclude non-events, the results of the general categorization would considerably improve.

The results for the *Social Action* category are promising, since the 93 *social actions* in this set were predicted with a precision of 0.80. However, we estimate that the recall of this category will be low. Only 93 out of 92,701 events (0.1%) were predicted as a *social action*, while 3.3% of events were annotated as a *social action* in the random event set.

Table 4: The precision and number of predicted instances per category.

Category	**Instances**	**Precision**	**Category**	**Instances**	**Precision**
Sport	2,771	1.00	Special day	1,722	0.78
Politics	2,170	0.86	Social action	93	0.80
Broadcast	206	1.00	Advertisement	25	0.51
Public event	81,538	0.57	Other	1,535	0.70
Software	1,630	0.96			

5 Analysis

5.1 Increasing the recall for Social Action events

Our main goal is to detect *social action events* and possibly alerting the authorities when such an event will take place. Therefore, we rather show a large list of events that might be a social action event that actually includes most of the actual events, than a system that often misses them. Since we are not talking about thousands of events daily, an analyst can annotate the set of possible social action events manually. We thus prefer a high recall to a high precision. Therefore, we propose a method to increase the recall of social action events, at minimal precision costs.

In order to increase recall we make use of the Naive Bayes classifier probability by category that is assigned to each event. Events for which *social action* obtained the second highest probability are now completely ignored, One way to remedy this is to classify all events where *social action* was the second most probable class. We refer to these events as **secondary social action events**.

By doing this we were able to expand this set with 226 additional events, which we annotated manually. 26 of the 226 *secondary social actions* were annotated as a non-event and were thus excluded from the set. 130 of the remaining 200 events were indeed annotated as a social action, resulting in a precision of 0.65. Adding the 200 events to the *social action* events in the evaluation set results in a drop of total precision from 0.80 to 0.69. However, recall was increased by **232%** while the precision only dropped by **14%**. Hence, including the *secondary social action events* seems a useful method for increasing the recall, while only mildly hurting the precision.

5.2 Most informative features

In order to achieve some insight from the most informative features for the two event sets, we calculated the chi-squared value for each feature in relation to the category label. These are listed in Table 5. The

most informative features are generally intuitive. They include words such as *stemmen (to vote)* and *stem (vote)* as indicators of a political event, but also specific hashtags such as *#VVD* and *#CDA*; CDA and VVD are political parties in The Netherlands. The best predictors for *sport* are the DBpedia type features *SoccerClub* and *ClubOrganization*. The most indicative features of the category *social action* are the words *protest* and *demonstratie (demonstration)*. Although these words almost exclusively occurred in *social action events*, due to their low frequency they do not rank in the feature top 100.

The polarity, subjectivity and periodicity features turned out to be less valuable, ranking in the bottom 25% of all features. This is surprising, since *special days* are often periodic, while it is, for example, uncommon for *social action events* to be periodic.

Table 5: The eight best features for the best and random event set, based on their chi-squared value. Non-word features are in italics. Features are only included if they occurred at least ten times in their event set.

Best events		**Random events**	
Feature	**Category**	**Feature**	**Category**
stemmen (vote)	Politics	*ClubOrganization*	Sport
stem (vote)	Politics	*SoccerClub*	Sport
19-03-2014	Politics	wint (wins)	Sport
SoccerClub	Sport	wedstrijd (match)	Sport
#vvd	Politics	2015	Politics
wedstrijd (match)	Sport	seizoen (season)	Broadcast
ClubOrganization	Sport	tv	Broadcast
#cda	Politics	tegen (against)	Sport

6 Conclusion and discussion

In this study we presented a generic event categorization system which we evaluated particularly on its ability to predict civil unrest. The general categorization system has a bias towards the dominant category *public event*, but has a high precision for the other categories, including *social action*. The recall for *social action* was low; a follow-up step that exploited the specific per-class probabilities generated by the Naive Bayes classifier led to a considerable improvement in recall of 232%, at the minor cost of a 14% decrease in precision.

The study by [14] is the only related work in the literature that also produced an extensive evaluation of event categorization, evaluating their system on a set of 500 events with the highest association (similar to the event score by which we selected a set of best events). Their 37-class approach ultimately obtained a precision, recall and F1-score of 0.85, 0.55 and 0.67. Our system offered a comparable performance: a precision, recall and F1-score of 0.67, 0.67 and 0.65.

A comparable approach to predicting civil unrest is the EMBERS system by [13]. They evaluated their system over a period of a month, resulting in a precision and recall of respectively 0.69 and 0.82. In comparison, we obtained a higher precision while our estimated recall is lower. It is interesting to note how they received this recall score. They obtained a gold standard set of *social action events* by an independent organization that had human analysts survey newspapers and other media for mentions of civil unrest; arguably a reliable way of calculating recall in the real world. Our approach is only able to recall events that were present in the set of [6]. We have not explored ways to evaluate to what extent [6] detected all *social action events* that actually happened. We should consider the possibility that we might still miss *social action events* that were never detected as events in the first place, lowering our estimated recall.

Using the ranking of the Bayesian probabilities helped to increase the recall of *social action events* by 232%. We did not use the actual probabilities to influence the classification process, but used only the ranking of these probabilities. A potential direction for future research is to use the per-class probabilities generated by the Naive Bayes classifier in a more sophisticated manner. For example, it is possible to learn a certain probability threshold for *social action* and classify events that exceed this threshold as *social action*, regardless of the probability of other categories. The actual implementation of such a

method requires a search for the best threshold setting. The main advantage of this approach is that this allows us to specify a specific precision-recall trade-off that is the most suitable for predicting social action events.

References

[1] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. Dbpedia-a crystallization point for the web of data. *Web Semantics: science, services and agents on the world wide web*, 7(3):154–165, 2009.

[2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[3] R. Compton, C.-K. Lee, T.-C. Lu, L. de Silva, and M. Macy. Detecting future social unrest in unprocessed twitter data:emerging phenomena and big data. In *Intelligence and Security Informatics (ISI), 2013 IEEE International Conference On*, pages 56–60. IEEE, 2013.

[4] T. De Smedt and W. Daelemans. Pattern for python. *The Journal of Machine Learning Research*, 13(1):2063–2067, 2012.

[5] A. F. Hayes and K. Krippendorff. Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1):77–89, 2007.

[6] F. Kunneman and A. van den Bosch. Automatically identifying periodic social events from twitter. *Proceedings of the RANLP 2015*, pages 320–328, 2015.

[7] F. Kunneman and A. van den Bosch. Open-domain extraction of future events from twitter. *Natural Language Engineering*, 2016.

[8] E. Meij, W. Weerkamp, and M. de Rijke. Adding semantics to microblog posts. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 563–572. ACM, 2012.

[9] R. Mihalcea and A. Csomai. Wikify!: linking documents to encyclopedic knowledge. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, pages 233–242. ACM, 2007.

[10] S. Muthiah, B. Huang, J. Arredondo, D. Mares, L. Getoor, G. Katz, and N. Ramakrishnan. Planned protest modeling in news and social media. In *AAAI*, pages 3920–3927, 2015.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[12] D. Ramage, S. T. Dumais, and D. J. Liebling. Characterizing microblogs with topic models. *ICWSM*, 10:1–1, 2010.

[13] N. Ramakrishnan, P. Butler, S. Muthiah, N. Self, R. Khandpur, P. Saraf, W. Wang, J. Cadena, A. Vullikanti, G. Korkmaz, et al. 'beating the news' with embers: Forecasting civil unrest using open source indicators. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1799–1808. ACM, 2014.

[14] A. Ritter, O. Etzioni, S. Clark, et al. Open domain event extraction from twitter. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1104–1112. ACM, 2012.

[15] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li. Comparing twitter and traditional media using topic models. In *Advances in Information Retrieval*, pages 338–349. Springer, 2011.

An Analysis on Better Testing than Training Performances on the Iris Dataset

Marten Schutten and Marco A. Wiering

Institute of Artificial Intelligence and Cognitive Engineering
University of Groningen, The Netherlands

Abstract

The Iris dataset is a well known dataset containing information on three different types of Iris flowers. A typical and popular method for solving classification problems on datasets such as the Iris set is the support vector machine (SVM). In order to do so the dataset is separated in a set used for training and a set used for testing. The error rate, after training, for the training set should be lower than the error rate on the test set. However, in this paper we show that when solving the classification problem for the Iris dataset with SVMs this is not the case. Therefore, we provide an analysis of the Iris dataset and the classification models in order to find the origin of this interesting observation.

1 Introduction

The Iris dataset [7] is a well known dataset used for classifying different types of Iris flowers (the Iris Setosa, Iris Versicolor and Iris Virginica). A version of this dataset can be found in the UCI repository [1], with two slight deviations from the original set [2]. Support Vector Machines (SVMs) [5, 11] are a well known and popular method to solve classification problems such as to be solved for the Iris dataset. Methods such as the Support Vector Machine aim to find a hyperplane that separates observations from different classes from each other. This hyperplane is learned from observations from which it is known to which class they belong, and should then generalize to instances that have not been observed before. One of the risks in learning this hyperplane on the basis of a limited number of examples is the problem of over fitting: The hyperplane is constructed specifically to separate the specific observations that are in the training set, even the ones that would be more likely to belong to a different class, when the class were to be unknown. As a result the hyperplane generalizes very poorly and is unable to correctly classify instances in the test set that have not been observed before. The result is then a very high classification accuracy on the train set and a much lower accuracy on the test set.

In this paper, we show that when using SVMs to solve the classification problem for the Iris dataset, a very different problem occurs: rather than over fitting on the known data, the hyperplane is 'under fitted' and generalizes better to unseen instances than it is able to classify the known instances, which is unknown to happen for any classification problem until now. The aim of this paper is to give an analysis of the Iris dataset and the obtained classification models in order to provide insights into this problem.

The paper is organized as follows. First a short explanation of support vector machines is given, followed by the methods through which the Iris dataset is analyzed. Finally, the dataset and classification models are analyzed and discussed.

2 Support Vector Machines

First, a short introduction to support vector machines [5, 11] will be given. Support vector machines can be used both for regression and classification problems, however, due to the nature of the Iris dataset, the explanation will be limited to their use for classification problems. For the explanation, initially a linear support vector machine will be considered for the separation of two classes. For a more thorough explanation of SVMs, see Burges [3], on which this explanation is based.

Consider a dataset $\mathcal{D}$ containing N instances of one of two classes, where the i-th (with $1 \leq i \leq N$) instance of $\mathcal{D}$ is given by $\{\mathbf{x}_i, y_i\}$, where $\mathbf{x}_i \in \mathbb{R}^d$ and the target class $y_i \in \{1, -1\}$, and where d is the number of features that describe each instance in the dataset. The goal is to construct a hyperplane of the form given by the weight vector $\mathbf{w} \in \mathbb{R}^d$ and offset (or bias) $b \in \mathbb{R}$, such that for each instance i in the dataset it holds that $\mathbf{w} \cdot \mathbf{x}_i + b \geq 1$ for $y_i = 1$, and $\mathbf{w} \cdot \mathbf{x}_i + b \leq -1$ for $y_i = -1$. Given that this holds for all points in $\mathcal{D}$, these formulae can be combined to the constraint given by equation 1.

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \quad \forall i \tag{1}$$

Now the points that lie closest to the separating hyperplane, are the points given either by $\mathbf{w} \cdot \mathbf{x}_i + b = 1$ (if $y_i = 1$) or $\mathbf{w} \cdot \mathbf{x}_i + b = -1$ (if $y_i = -1$). The points that satisfy these constraints are called the *support vectors*. Note that the support vectors for both classes all lie at the same distance from the separating hyperplane, which is given by $\frac{1}{\|\mathbf{w}\|}$. Let the margin of the separating hyperplane be defined as the sum of the distances to support vectors of both classes, being $\frac{2}{\|\mathbf{w}\|}$. Now the goal is to maximize this margin, by minimizing $\|\mathbf{w}\|^2$, subject to the constraint given by equation 1. Thus the goal is to find a separating hyperplane that maximizes the distance between itself and the support vectors of both classes.

Unfortunately, it is not always the case that two classes are perfectly separable, and no solution can be found according to these specific constraints. In order to deal with this non-separability it is required to relax the constraints of the classification problem. This is done with the introduction of a positive slack variable ξ_i for all examples in $\mathcal{D}$. With the introduction of this slack variable the constraints are now given by:

$$\mathbf{w} \cdot \mathbf{x}_i + b \geq 1 - \xi_i \quad \text{if } y_i = 1 \tag{2}$$

$$\mathbf{w} \cdot \mathbf{x}_i + b \leq -1 + \xi_i \quad \text{if } y_i = -1 \tag{3}$$

$$\xi_i \geq 0 \quad \forall i \tag{4}$$

Note that with these constraints an instance is falsely classified whenever $\xi_i > 1$. However, since miss classifications are undesirable, the objective function that is to be minimized is rewritten from $\|\mathbf{w}\|^2$ to the formula given in equation 5. In this formula C defines the cost parameter that determines the severity of errors. Note that minimizing $\frac{1}{2}\|\mathbf{w}\|^2$ yields the same weight vector as minimizing $\|\mathbf{w}\|^2$, but $\frac{1}{2}$ is added to simplify future equations.

$$\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i \tag{5}$$

In order to solve this problem it can be rewritten in the form of a Lagrangian formulation, introducing a Lagrange multiplier α_i for each of the i constraints provided by equations 2 and 3. The problem can then be defined according to a Wolfe dual formulation, defining a primal and a dual Lagrangian formulation, given respectively by formulas 6 and 7. In the former formula μ_i represent the Lagrange multipliers that are required to enforce positivity of ξ_i, in accordance with constraint 4.

$$L_P \equiv \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\{y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 + \xi_i\} + \sum_{i=1}^{N}\mu_i\xi_i \tag{6}$$

$$L_D \equiv \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \tag{7}$$

The solution can then be found by either maximizing L_D or by minimizing L_P, which yield the same solutions. The constraints of both problems are given in Table 1 and the solution is given by formula 8. Using this weight vector a classification output can be given for any unseen example $\mathbf{x}$ by using equation 9.

$$\mathbf{w} = \sum_{i=1}^{N}\alpha_i y_i \mathbf{x_i} \tag{8}$$

Table 1: Constraints for the primal and dual Lagrangian formulations L_P and L_D.

L_P	L_D
$\frac{\delta L_P}{\delta \mathbf{w}} = \frac{\delta L_P}{\delta b} = \frac{\delta L_P}{\delta \xi_i} = 0$	$0 \leq \alpha_i \leq C$
$\xi_i, \alpha_i, \mu_i \geq 0$	$\sum_{i=1}^{N} \alpha_i y_i = 0$
$\alpha_i\{y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 + \xi_i\} = 0$	
$\mu_i \xi_i = 0$	

$$f(\mathbf{x}) = sign(\sum_{i=1}^{N} \alpha_i y_i \mathbf{x_i} \cdot \mathbf{x} + b) \tag{9}$$

Now this solution holds for linear separating hyperplanes. However, often non-linear hyperplanes are better suited to solve the classification problem. These non-linear hyperplanes can be described using so-called kernel functions. Rather than using the linear form of $\mathbf{x_i} \cdot \mathbf{x}$ that can be found in the solution given by equation 9, the solution can then be found using an alternate form defined by the kernel function $K(\mathbf{x_i}, \mathbf{x_j})$, as given by equation 10.

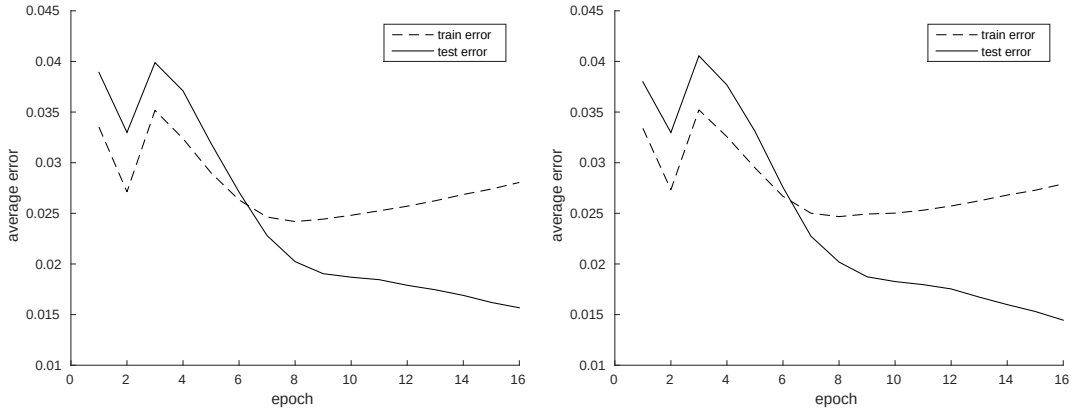
$$f(\mathbf{x}) = sign(\sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x_i}, \mathbf{x}) + b) \tag{10}$$

A well-known kernel function is the radial basis function (RBF), which computes similarities between two examples according to equation 11. In this paper, a truncated version of this kernel was used, so that whenever the output of the kernel function was below a certain threshold θ, it was set to 0. The value of θ was optimized in the same manner as the other parameters for the SVM, as discussed in the following section.

$$K(\mathbf{x_i}, \mathbf{x}) = e^{-\|\mathbf{x_i} - \mathbf{x}\|^2 / 2\sigma^2} \tag{11}$$

3 Methods

In Figure 1, we show the development of the error rate over the 16 epochs for which the SVM was trained using both a regular RBF kernel (Figure 1a) and a truncated RBF kernel (Figure 1b). The error rate is the average error rate over 10000 runs that were performed.



(a) regular RBF kernel

(b) truncated RBF kernel

Figure 1: Learning curves for both the training and the testing set as developed over the 16 epochs for which the SVM was trained. The error rates are averaged over 10000 runs.

The error rate is slightly lower for the truncated RBF kernel than for the regular RBF kernel (0.0144 versus 0.0157 resp. after 16 epochs for the testing set and 0.0279 and 0.0281 resp. after 16 epochs for the training set). Interestingly, and in contrast to expectations, the error rate for the training set starts increasing after 7 epochs for both kernels, while the error on the test set keeps decreasing.

In order to examine how the Iris dataset is classified, the first two principal components were determined using Principal Component Analysis (PCA) [10, 8]. These principal components were used to get an insightful look at the separability of the data. Furthermore, each of the individual features, laid out against each other is examined as well. Since the overall training error is higher than the testing error for the Iris dataset, there must be a number of points in the dataset for which the same holds. In order to identify these points the data were being trained and tested upon using the SVM with the truncated RBF kernel for 10000 times, in which the data were randomly separated in a training set (90%) and a testing set (10%). Particle Swarm Optimization (PSO) [9] was used to find the optimal parameters for the SVM. As SVM, we used the gradient descent SVM explained in [4].

During the runs with the support vector machines a number of things were kept track of for each data point: (1) the number of times the point was added to the testing set; (2) the number of errors that was made when the point was added to the training set and (3) the number of errors that was made when the point was added to the testing set. From this, the error rates can be computed both for when the data point was added to training set and when it was added to the testing set. In order to determine whether the error rates in the testing set were significantly higher than in the training set, Fisher's exact test [6] was used with $\alpha = 0.01$. Finally records were kept for the α-coefficients for each point.

4 Description and Analysis of the Data

The dataset contains 3 classes of types of Iris flowers: the Iris Setosa, Iris Versicolor and the Iris Virginica. Each class has 50 instances in the set of which one is linearly separable from the other classes, while the other classes are not linearly separable from each other. Each instance in the dataset has the following attributes: 1) sepal length in cm; 2) sepal width in cm; 3) petal length in cm; 4) petal width in cm. There are no instances in the dataset that miss any of these attributes. Figure 5 shows how the data is scattered for all permutations of the different features.

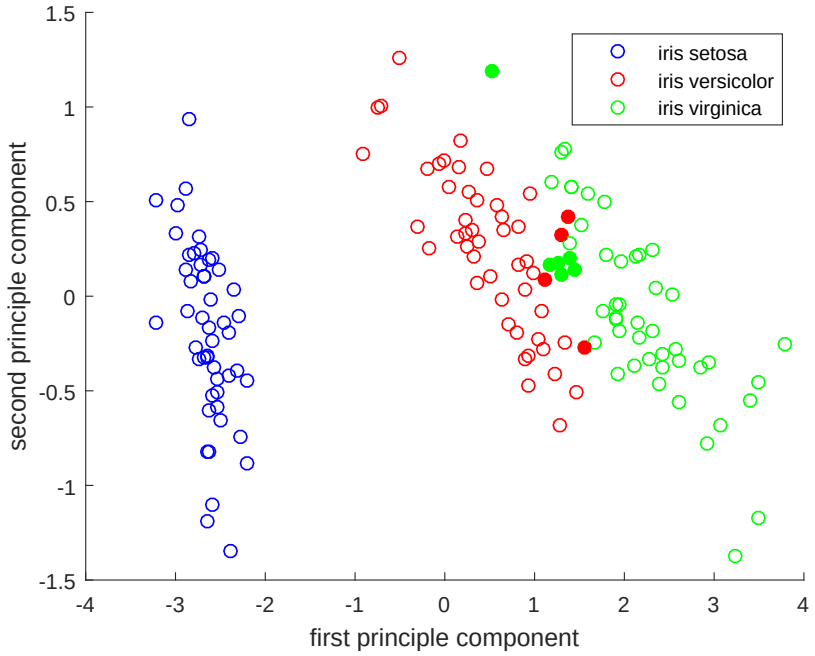


Figure 2: The examples plotted using the first two principal components of the Iris dataset. The filled points represent the examples for which the error rate is significantly higher when the example is used in the training set than in the testing set.

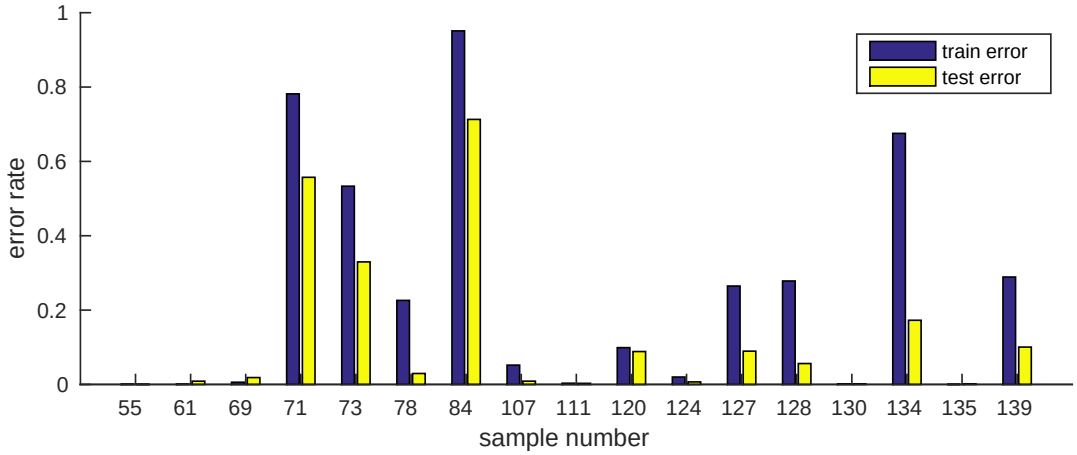


Figure 3: The error rates for all points for which at least one error was made in either the training set or the testing set.

#	true class	error in training set	error in testing set	p	#	true class	error in training set	error in testing set	p
55	ver.	0.000	0.000	1.000	120	vir.	0.099	0.089	0.332
71	ver.	0.782	0.557	$< 2.2e-16$	*124*	vir.	0.020	0.007	0.002
73	ver.	0.533	0.330	$< 2.2e-16$	*127*	vir.	0.265	0.090	$< 2.2e-16$
78	ver.	0.226	0.029	$< 2.2e-16$	*128*	vir.	0.278	0.056	$< 2.2e-16$
84	ver.	0.951	0.713	$< 2.2e-16$	130	vir.	0.001	0.001	1.000
107	vir.	0.052	0.011	$7.939e-13$	*134*	vir.	0.675	0.173	$< 2.2e-16$
111	vir.	0.003	0.003	1.000	*139*	vir.	0.289	0.101	$< 2.2e-16$

Table 2: Error rates for all examples in which the error rate in the training set is higher than the error rate in the testing set, along with the probability that this error rate is due to non-random chances. The points for which this difference is significant (for $p < 0.01$) are emphasized. The true classes were abbreviated to ver. (for the Iris Versicolor) and vir. (for the Iris Virginica). No instances of the Iris Setosa class need to be examined.

The Iris dataset was first published by Fisher [7]. Over time different versions of the dataset have been published [2] with slight alterations. For this paper the version that can be found in the UCI repository [1] was used. This version contains two deviations from the original dataset [2] (in the 35th and 38th example), which were corrected.

Figure 2 shows the data plotted using the first two principal components that can be obtained for this dataset. It can be seen that the Iris Setosa class lies far apart from the other two classes and is easily separable. The other two classes lie more closely to each other and even though most examples should be easy to separate, around the boundary area it might be hard to differentiate between the two of them. Note that a similar observation can be made when looking at the different features plotted together in Figure 5.

The results of the 10000 runs performed with the SVM can be seen in Figure 3. This plot shows the error rate for all points in both the training and testing sets. Note that only points are included in which at least one error is made in either the training or testing set. The first 50 examples in the dataset belong to the Iris Setosa class and as expected no errors are made when classifying instances of it. For each of the points that have a higher error rate in the training set than in the testing set, Table 2 shows the error rates for both sets and the results of the Fisher test (with $\alpha = 0.01$). It can be seen that for four of the examples (55, 111, 120 and 130) the difference is not significant while in all other points the error rate in the training set is significantly higher than in the testing set. Figure 3 together with Table 2 show that only for three examples (61, 69, and 135), the error is higher when the example is used in the test set compared to being part of the training set.

Both in Figure 5 and 2 the examples that score better in the testing set than in the training set are

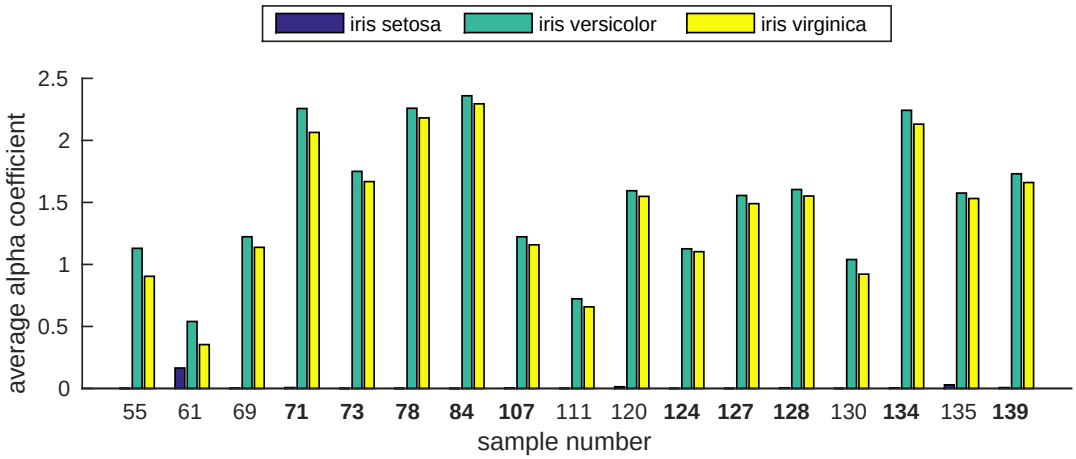


Figure 4: α-coefficient for all points in which an error was made in either the training or testing set. Points printed bold are points in which the error rate was higher in the training set than in the test set.

highlighted. Especially in the plot containing the principal components (Figure 2) it can be seen that all of these 10 points lie around the boundary between the Iris Versicolor and Iris Virginica class. A similar observation can be made in Figure 5, albeit not as clear for all feature combinations.

Finally, Figure 4 shows the α-coefficients for all points in which an error was made in either the training or the testing set. The points that are printed bold are the points in which the error rate was higher in the training set than in the testing set. This figure shows that the alpha coefficients for one class (Iris Versicolor) are generally higher than those for Iris Virginica, which indicates a preference of the model to classify examples as being part of the former class. Due to other examples and the kernel function that uses distances to all examples, this does not mean that these examples are all classified as Iris Versicolor.

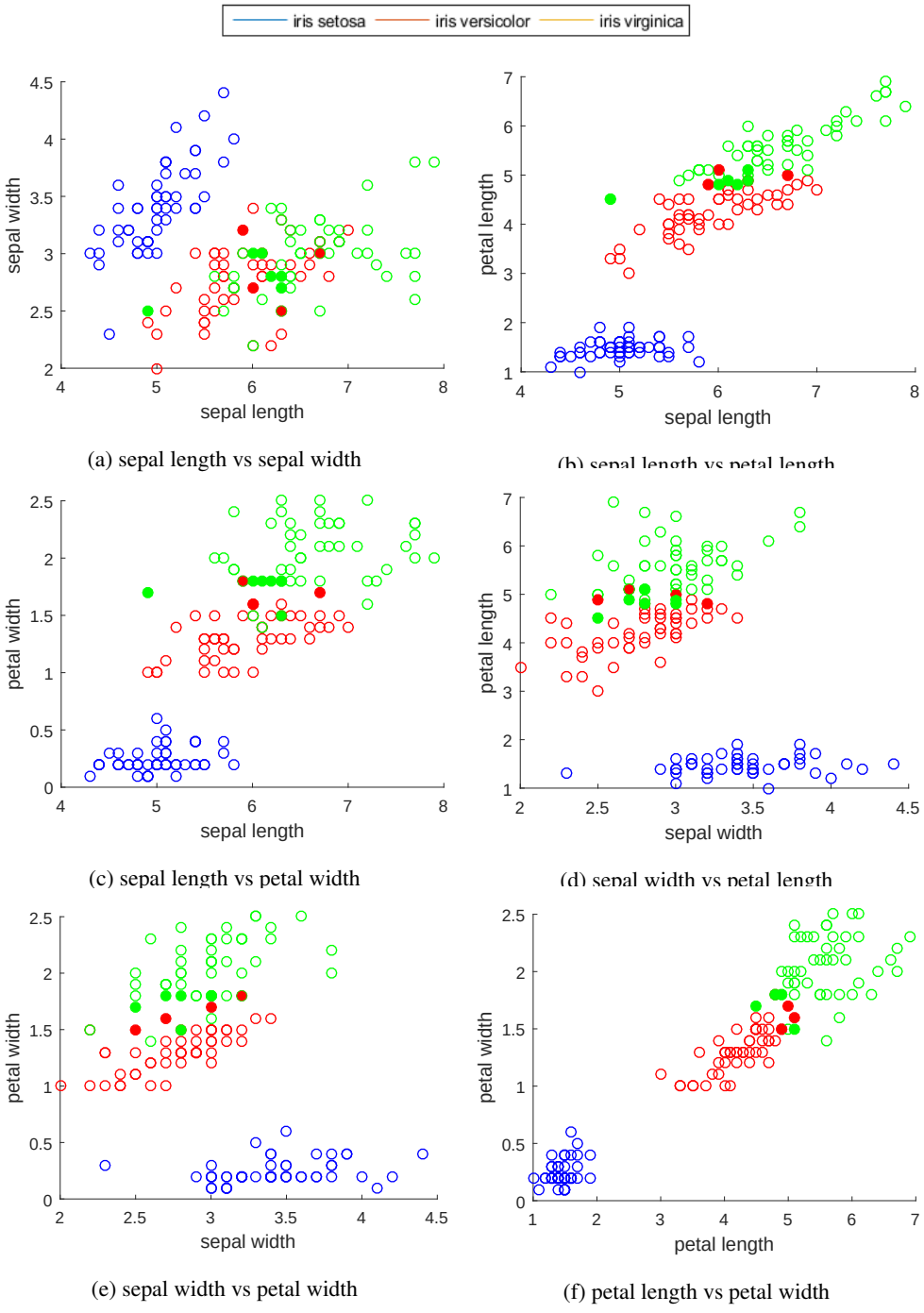
5 Discussion

The aim of this paper is to identify and clarify the unusual behavior in the Iris dataset with regards to the error rates in the training and error phases. Insight has been given how the SVM classifies different data points, and crucial data points that are hard to classify correctly in both the training and testing phase have been identified. We showed that particular data points close to the decision boundary are more often classified correctly when they are in the test dataset than in the train dataset.

One of the ways this behaviour can be explained is through the use of PSO as an optimization method. With PSO a wide variety of parameter settings is tried to find the setting that optimizes the results for their respective test sets using many different random cross validation runs. This means that parameter settings might be chosen that neglect performance in the training phase, in order to be able to generalize better to unseen cases in the testing phase. Because the Iris dataset is very small, meta-parameters were found by PSO that are useful to miss classify training points in order to obtain better results on the test data. Somehow, the extensive parameter search with PSO has led to overfitting on the cross validation results.

This is an important finding, and shows the importance of using a separate unpolluted dataset for the final test. However, the Iris dataset is much too small to be split into a train-validation-test set. Therefore, this option is not available for all datasets (also some datasets with for example fMRI scans are very small and do not contain more than 150 examples).

We found that models can be trained that sacrifice train performance in order to obtain better results on test examples. Still, it is complex to understand how this is exactly done in the training process, as the conventional way of overfitting is exactly the other way around. Therefore, this effect can not be fully accounted for by the use of extensive parameter tuning with PSO and cross validation. Furthermore, no other cases of this phenomenon have been reported to the best of our knowledge. One of the main reasons for this effect to occur in the Iris dataset and not in other datasets, is probably the small amount



(a) sepal length vs sepal width

(b) sepal length vs petal length

(c) sepal length vs petal width

(d) sepal width vs petal length

(e) sepal width vs petal width

(f) petal length vs petal width

Figure 5: all different permutations of the four features in the Iris dataset plotted against each other. The filled points represent the examples for which the error rate is significantly higher when the example is used in the training set than in the testing set.

of examples as well as the limited number of features for each example. However, we would gladly be pointed to other datasets and/or cases where the testing results outperform the training results.

References

[1] K. Bache and M. Lichman. UCI machine learning repository, 2013.

[2] James C Bezdek, James M Keller, Raghu Krishnapuram, Ludmila I Kuncheva, and Nikhil R Pal. Will the real Iris data please stand up? *IEEE Transactions on Fuzzy Systems*, 7(3):368–369, 1999.

[3] Christopher J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.

[4] V. Codreanu, B. Dröge, D. Williams, B. Yasar, P. Yang, B. Liu, F. Dong, O. Surinta, L.R.B. Schomaker, J.B.T.M. Roerdink, and M.A. Wiering. Evaluating automatically parallelized versions of the support vector machine. *Concurrency and Computation, Practice and Experience*, 28(7):2274–2294, 2016.

[5] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995.

[6] Ronald A Fisher. On the interpretation of χ 2 from contingency tables, and the calculation of p. *Journal of the Royal Statistical Society*, 85(1):87–94, 1922.

[7] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.

[8] Ian Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[9] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Neural Networks, Proceedings of the IEEE International Conference on*, volume 4, pages 1942–1948, 1995.

[10] Karl Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6*, 2(11):559–572, 1901.

[11] Vladimir Vapnik. *The nature of statistical learning theory*. Springer Science & Business Media, 2013.

Offline Handwriting Recognition Using LSTM Recurrent Neural Networks

Yaroslav Shkarupa [a] Roberts Mencis[a] Matthia Sabatelli [a]

[a] *University of Groningen*
Department of Artificial Intelligence and Cognitive Engineering
Nijenborgh 4, 9747 AG Groningen
The Netherlands

Abstract

Handwriting recognition is a notoriously difficult problem in Machine Learning. Despite decades of research and development, modern handwriting recognition systems still exhibit suboptimal performance in the real world applications. Recent studies show great potential of Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) for unsegmented handwritten word recognition. In this paper we evaluate two approaches based on RNN with LSTM for the recognition of historic handwritten Latin texts. The first approach makes use of a Connectionist Temporal Classification output layer, and the second approach is based on Sequence-to-Sequence Learning. To test these approaches we have built a handwriting recognition system which takes an unsegmented word image as an input and provides a decoded string as an output. Both approaches show promising results with 78.10% and 72.79% word-level accuracy on the test dataset for respective methods. The Connectionist Temporal Classification approach consistently outperforms the Sequence-to-Sequence Learning approach in terms of generalization and prediction of long words.

1 Introduction

While optical character recognition (OCR) can be considered as a mature field with very high accuracy rates for printed documents, recognition of handwritten text remains a much harder challenge. The best performing teams of the ICDAR2015 competition were able to achieve just 69.8% word level accuracy, and 84.5% character level accuracy rates [13]. High variability of handwritten characters and limited accessibility of language models, especially for historical texts, are two big hurdles for existing OCR methods. While for a long time the main methods were based on carefully extracted features and probabilistic models, such as the Hidden Markov Models, recently, better results were achieved by using deep learning techniques, like the Convolution Neural Networks (CNN) and Recurrent Neural Networks (RNN) with Long Short Term Memory (LSTM).

In this paper we evaluate two RNN architectures - Connectionist Temporal Classification (CTC) and Sequence-to-Sequence Learning (Seq2Seq) for recognizing word-level labels in handwritten medieval Latin texts. We focus on balancing the performance and architectural simplicity of handwriting recognizing system so that the training of the models could be completed on average PC within reasonable time (less than 24 hours).

As a direct comparison of internal configuration (e.g. number of hidden units) of CTC and Seq2Seq models could be difficult due to considerable differences of both architectures, we evaluated different configurations for each model individually with an aim to maximize accuracy within reasonable training time. At the end the best performing configuration of each model was evaluated on the separate test dataset.

The research was done as part of the course ”Handwriting Recognition” given by prof. dr. L. Schomaker at the University of Groningen in Spring 2016. The full pipeline of the recognition system is described alongside the achieved results. We used OpenCV for image preprocessing and Tensorflow framework for all machine learning tasks.

2 Dataset

The training dataset consisted of 118 scanned pages of handwritten medieval Latin texts from two sources - KNMP Chronicon Boemorum [1] (72 pages) and Stanford CCCC [3] (46 pages). The dataset was made available through the Monk System [2], a search engine for handwritten document collections developed at the University of Groningen by a research group at the Artificial Intelligence institute ALICE, under supervision by prof. dr. L. Schomaker [14]. Figures 1 and 2 show examples of handwritten text from both sources.

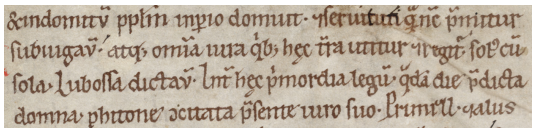


Figure 1: Example of KNMP handwritten text

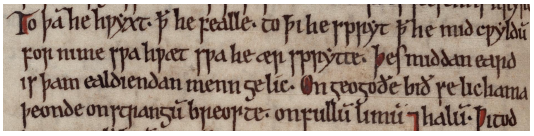


Figure 2: Example of Stanford handwritten text

Each scanned page was accompanied by XML file specifying positions of line, word and character rectangles as well as text labels for words and individual characters. In total 4988 labeled words and 14443 labeled characters were available for training, with 42.3 labeled words on average per page. Average length of a word was 2.9 characters with one or two character words making up around half of all labels. Figures 3 and 4 show some examples of segmented words with accompanying labels.

Figure 3: Examples of labeled KNMP words

Figure 4: Examples of labeled Stanford words

Additional 528 words from both sources were manually labeled to increase the number of multi-character words available for training. These labels were used only for training the Seq2Seq model as character-level labeling is not needed using this approach.

The final test dataset contained approximately the same number of pages as the dataset made available for training. The evaluation dataset was small subset of final test dataset and was made available few weeks before the final test. The test dataset itself was not available during training, but a lexicon of words with corresponding frequencies was provided.

3 Methods

3.1 Preprocessing

The preprocessing phase can be considered as the first stage of the recognition system. The main goal of this step is to modify the images in a way that will make it easier and faster for the recognizer to learn from them.

The word images were cut out of high-resolution RGB images. The average width and height of a word image was 90x60 pixels, with maximum width reaching 300-400 pixels for very long words. Thus 90x60x3=16200 features would be needed to describe an average word image. While images of similar size are nowadays commonly used in image recognition with convolutional networks our main goal was rapid prototyping of various RNN architectures and therefore the size of input data was important factor affecting the total time of training.

Taking into account above mentioned considerations, we applied following preprocessing steps to each word image:

1. Dataset augmentation
2. Conversion to grayscale
3. Resizing

Firstly, we generated 10 additional augmented images by applying random shear (with shear factor from -0.20 to 0.20) to each original image. Secondly, the augmented images were converted to greyscale, on the basis of the assumption that color information provides little discriminative power in the case of handwritten text. Finally, every image was resized to the fixed height (28 pixels) and to the 50% of the original width. This was done to convert the image into format suitable for feeding into RNN where image height would correspond to the size of feature vector and image width to the number of timesteps in the input sequence.

We considered using convolutional layers for feature extraction, but initial experiments showed that it provided only minor improvement of accuracy and considerably increased the training time. We also tried simple image enhancement techniques (Gaussian blur for noise reduction and binarization) as additional preprocessing steps, which provided small increase of accuracy for Seq2Seq approach, but did not substantially affect accuracy of CTC approach.

3.2 Recurrent Neural Networks (RNN)

The core part of the approach presented in this work is based on Recurrent Neural Networks (RNN), a particular type of Neural Network that is characterized by the presence of self-connected nodes. These nodes enable the network to memorize and keep track of previous inputs, allowing it to store and access information over long periods of time. This particular ability allows the RNN to learn the context of the labels which turns out to be a big advantage in handwriting recognition since the characters of the words can not be considered as really independent components. In order to use RNNs for handwriting recognition images with handwritten text are interpreted as time sequences along single or multiple axis.

Even though classical RNN provide the network with some memory they are still quite limited in their ability to preserve old content due to the vanishing gradient problem [9]. In fact, as explained by the authors, it is hard for a RNN to bridge gaps of more than about 10 time steps between the moment in which the relevant input is presented and the relative target events. To address this problem we have used Long Short-Term Memory (LSTM) [10]. LSTMs are specially constructed RNN nodes to preserve long lasting dependencies. They consist of self-connected memory cell that can be compared to the classical RNN node and three gates that control output and input of the node. Each gate is in fact a sigmoid function of the input to the LSTM node. The first gate is an input gate which controls whether new input is available for the node. The second gate is a forget gate which makes possible for the node to reset activation values of the memory cell. he last gate is an output gate controlling which parts of the cell output are available to the next nodes.

Further improvement to the RNN models based on LSTM is achieved by the use of opposite two-directional layers or so-called Bidirectional Long Short-Term Memory (BLSTM). The goal of the forward layer is to learn context of the input by processing the sequence from the beginning to the end, while the backwards layer performs the opposite operation by processing the sequence from the end to the beginning. It was demonstrated [8] that this architecture performs better than a simple uni-directional LSTM.

3.3 Connectionist Temporal Classification approach

The feed-forward approach is similar to the original recurrent neural networks by the fact both architectures require a direct alignment between the input features and target variables. However in the real-world handwriting recognition problems it is much easier to segment text into words rather than individual characters. Achieving direct alignment between the image of input character and character target label would require a prior segmentation step. Being a very hard problem by itself its complexity keeps increasing in time since there is a high tendency of encountering possible errors already in the segmentation step. As a result this would also limit the context of the data learned by the RNN.

To target this problem the Connectionist Temporal Classification (CTC) approach was introduced, originally for speech recognition [6] and afterwards also for handwriting recognition [7]. CTC makes it

possible to avoid the previously mentioned direct alignment between the input variables and the target labels by interpreting the output of the network as a probability distribution over all possible label sequences on the given input sequence.

The last layer of the network is a softmax output with $N+1$, where N is the total number of labels (in case of handwriting recognition - characters). These outputs define the probability of observing each label or not observing a label at a given moment in time. By using single character probabilities it is possible to obtain a set of probabilities that correspond to a possible output sequences given an input one. This logic is formalized by equation

$$p(\pi|x) = \prod_{t=a}^{T} y^t_{\pi_t}, \forall \pi$$

where x is the input sequence, π a possible output sequence and $y^t_{\pi_t}$ the probability of a given label from π at time t.

A Special operator β is defined to match sequences from previous step π to the final output sequences l by removing blank labels and merging repeated labels. For example a possible sequence as β (dd_a_b) = (dab). Since more than one sequence π can correspond to a single final sequence then probability of a particular final sequence l is computed as follows:

$$p(l|x) = \sum_{\beta(\pi)=l} p(\pi|x).$$

For each given input x and sequence l, the maximum probability $p(l|x)$ is considered as the final output.

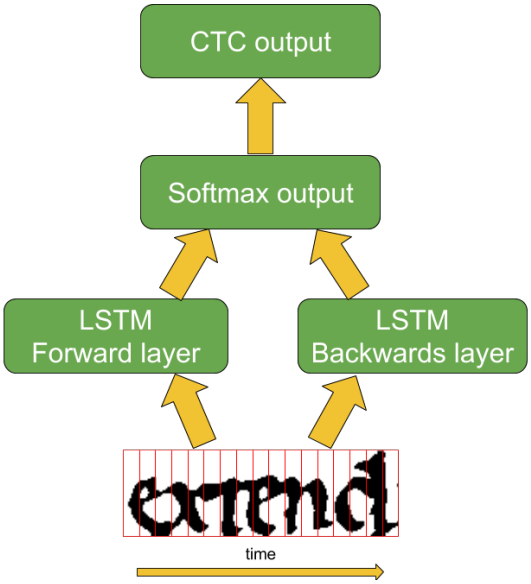


Figure 5: RNN architecture with Connectionist Temporal Classification output layer

For handwriting recognition problem a word image is interpreted as a time sequence where time is modeled along the width of the image with 1px slice as a single unit of time. Pixel values after preprocessing were used as input features for the network. The Figure 5 shows the general recognizer structure.

3.4 Sequence-to-Sequence Learning approach

Sequence-to-Sequence Learning (Seq2Seq) is based on the approach developed by Google researchers for the automatic sentence translation from English to French [15]. The main idea is to use two connected RNNs, the first RNN for reading an input sentence and mapping it to a fixed-dimensional vector

representation and the second RNN for decoding an output sequence from that representation. The approach is very suitable for machine translation, as input and output sequences can be of various lengths and ordering of words in each language can be different.

While in handwriting recognition it can generally be assumed that handwritten characters can be mapped to labeled characters in the same order, Seq2Seq approach still provides the advantage that no specific mapping between positions of characters in input image and target character labels is required.

The model consists of two connected RNNs, namely: the Image-RNN for the image encoding and the Label-RNN for the generation of the text label as shown in Figure 6. Each RNN is constituted of one or more LSTM or BLSTM layers. Weighted cross-entropy loss between a target sequence and a predicted sequence of characters is used as a cost function.

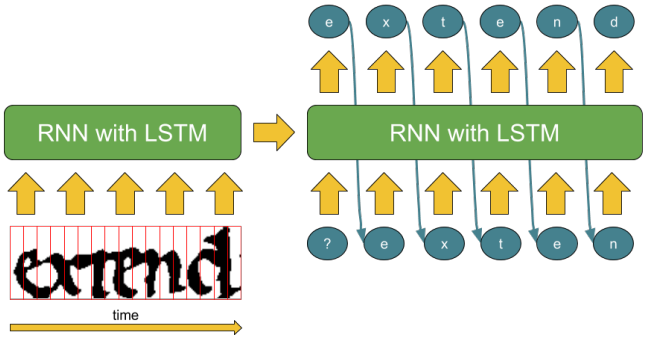


Figure 6: RNN architecture with Sequence-to-Sequence Learning

An input image is presented to the Image-RNN as a separate 1px slices with features corresponding to pixel values after preprocessing. The final output state of the Image-RNN is used as an initial state for the Label-RNN. As a number of hidden units in both RNNs can differ, a fully connected layer provides the mapping from the output state of the Image-RNN to the initial state of the Label-RNN.

At the first time step of the Label-RNN the ”start word” character is presented. The Label-RNN generates an output character, which is used as an input for the next time step and thus recursively the whole word is generated until a white space character is encountered. Both input and output characters are encoded as one-hot vectors.

4 Experiments

Initially we performed number of experiments to determine optimal configurations for each approach. The models were trained on both document sources (KNMP and Stanford) separately. The dataset was randomly divided into training data (90%) and hold-out data (10%) for validation and hyperparameter tuning. Mini-batch approach was used for training with each batch consisting of 128 word images. In each epoch all images were presented to the neural network in random order.

The CTC model was trained using a decaying learning rate. The learning rate was initially set to 0.01 and gradually decreased to 0.0001. For the Seq2Seq model we discovered that setting the learning rate to 0.1 ensured the best convergence in spite of occasional jumps of the loss function value. Other learning rates (0.01, 0.001) led to much slower convergence and considerably increased the total time of training. The moderate 0.75 dropout keeping probability was found to achieve an optimal balance between training speed and prevention of over-fitting as suggested by Pham et al [12].

BLSTM was used in all CTC experiments, but in case of Seq2Seq it turned out that BLSTM offered no increase in the performance, therefore most of Seq2Seq experiments were performed with uni-directional LSTM to achieve faster training. Limited improvement provided by BLSTM can be explained by prevalence of short words in the dataset as noted by Lee et al [11].

It was found that simply increasing the number of LSTM nodes in forward and backwards layers did not considerably influence the accuracy of the model, but rather lead to overfitting on the training set. We found that 128 nodes on each layer was an optimal configuration for the CTC model and 96 nodes in both Image-RNN and Label-RNN were optimal choice for the Seq2Seq model with good generalization performance and reasonable training time.

As it could be assumed that first characters of the word are more important for predicting the whole word, the cost function weights for the Seq2Seq model were chosen so that higher significance (weight values - 10.0, 7.0, 4.9, ..., 1.0) were assigned to the first letters of the word. The use of weighted cross-entropy loss function reduced convergence time of the model.

Initially we used preprocessed grayscale images to achieve optimal configuration of the models. Subsequent experiments showed that using binarized images with Gaussian blur offered consistent 5-7% increase in word-level accuracy for the Seq2Seq model, but did not affect accuracy of the CTC model.

We also examined the use of convolutional layers for feature extraction. 1-2 convolutional layers with up to 16 filters and sliding window with patch size from 3x3 to 7x7 were used to extract 1px slices of features from word images. Unfortunately, contrary to expectations, this approach did not provide increase of accuracy and thus was not used for the final recognizer.

After finding optimal configurations for both models we performed a number of experiments to compare them. Two test datasets which were not included in training process were used to compare accuracies of both models. We used results from evaluation dataset to improve our models before the final test. Using Levenshtein distance to find the closest word in provided lexicon (instead of using decoded word straight from the model) also improved accuracy on the final test dataset.

5 Results

5.1 Metrics

Two metrics were used to evaluate the performance of the handwriting recognition system:

- Word accuracy rate (WAR) - percentage of words with all characters correctly recognized
- Character accuracy rate (CAR) - average Levenshtein distance normalized by the length of the longest word and subtracted from 1 (thus 0 means no similarity and 1 means complete similarity).

5.2 Accuracy

The following table shows the accuracy of both approaches on the training dataset as well as on the evaluation and the final test dataset:

	Connectionist Temporal Classification		Sequence-to-Sequence Learning	
	Character Accuracy Rate (CAR)	Word Accuracy Rate (WAR)	Character Accuracy Rate (CAR)	Word Accuracy Rate (WAR)
KNMP training dataset	96.0%	88.2%	86.9%	73.4%
Stanford training dataset	96.2%	94.5%	88.7%	82.0%
Evaluation dataset	-	78.7%	-	68.3%
Final test dataset	-	**78.10%**	-	**72.79%**

Table 1: Accuracy of the CTC and Seq2Seq model

5.3 Impact of word length on accuracy

Figure 7 shows how accuracy of both models depends on the length of the word using the training dataset without any lexicon. The performance of the CTC model decreases faster at the beginning, but with longer words it performs slightly better than the Seq2Seq model.

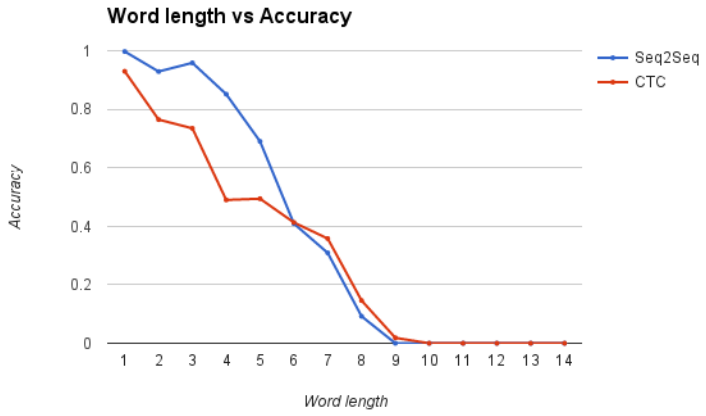


Figure 7: Impact of word length on accuracy (on the training dataset)

6 Discussion & Conclusion

In this paper we have evaluated two RNN architectures for handwritten text recognition based on Connectionist Temporal Classification and Sequence-to-Sequence Learning approach. The obtained results (see Table 1) are comparable with 81.5% average recognition rate over all manuscripts in Monk System [14] and in similar range as 69.8% word accuracy rate achieved in ICDAR2015 competition [13]. However it should be noted that our models were fine-tuned to specific data sources and might show varying performance on different datasets.

Although both methods showed promising results, the CTC model consistently outperformed the Seq2Seq model on both training and test datasets. During the experiments we observed that the Seq2Seq model had tendency to overfit on the training data and was not generalizing well on the validation data, while the CTC model was more robust against overfitting and as a result had higher accuracy on previously unseen data.

Figure 7 shows that Seq2Seq was better at predicting short words, but CTC model had higher accuracy predicting longer words. That could be explained by the fact that the CTC model predicts characters ”on the fly” while the Seq2Seq model first represents the whole image of handwritten word as a fixed-size vector and then decodes it to generate predicted label. As a result the Seq2Seq model has a tendency to learn whole words instead of separate character sequences. Thus CTC approach seems more natural choice for problems where an input sequence has the same ordering as an output sequence.

Various experiments during training phase showed that image preprocessing and dataset augmentation by additional image transformations can improve accuracy rates and prevent overfitting. Some additional preprocessing steps could be suggested, for example, vertical alignment of letters and removal of unnecessary white-space at the top and the bottom of word images. The use of more advanced features instead of raw pixel intensities (for example, HOG [5]) could be considered, although experiments with automatic feature learning using convolutional layers showed that other factors are probably more important for RNN performance than feature engineering.

Even though our models showed good results on provided datasets it is hard to infer the performance of examined methods on different handwritten text sources. In the future work the use of more advanced methods like Multi-Dimensional Long-Short Term Memory (MDLSTM) and attention mechanisms could be considered [4][11]. However usage of these methods also means increased complexity and training time of the models and therefore was not used in this research.

7 Acknowledgments

We are grateful to prof. dr. Lambert Schomaker and Sheng He for valuable advise during this research and for providing an access to the KNMP and Stanford dataset. For any questions regarding the dataset please contact prof. dr. L. Schomaker (email: *schomaker@ai.rug.nl*).

References

[1] KNMP Chronicon Boemorum dataset in Monk System. http://application02.target.rug.nl/cgi-bin/monkweb?cmd=showbook&pagid=*&db=1201.

[2] Monk System. http://www.rug.nl/society-business/target/projects-and-rd/target_projects/monk?lang=en.

[3] Stanford CCCC dataset in Monk System. http://application02.target.rug.nl/cgi-bin/monkweb?cmd=showbook&pagid=*&db=61200.

[4] Théodore Bluche and Ronaldo Messina. Scan, attend and read: End-to-end handwritten paragraph recognition with mdlstm attention. *arXiv preprint arXiv:1604.03286*, 2016.

[5] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893. IEEE, 2005.

[6] Alex Graves, Santiago Fernández, Faustino Gomez, and Jürgen Schmidhuber. Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. In *Proceedings of the 23rd international conference on Machine learning*, pages 369–376. ACM, 2006.

[7] Alex Graves, Marcus Liwicki, Santiago Fernández, Roman Bertolami, Horst Bunke, and Jürgen Schmidhuber. A novel connectionist system for unconstrained handwriting recognition. *IEEE transactions on pattern analysis and machine intelligence*, 31(5):855–868, 2009.

[8] Alex Graves and Jürgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural Networks*, 18(5):602–610, 2005.

[9] Sepp Hochreiter, Yoshua Bengio, Paolo Frasconi, and Jürgen Schmidhuber. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies, 2001.

[10] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[11] Chen-Yu Lee and Simon Osindero. Recursive recurrent nets with attention modeling for OCR in the wild. *arXiv preprint arXiv:1603.03101*, 2016.

[12] Vu Pham, Théodore Bluche, Christopher Kermorvant, and Jérôme Louradour. Dropout improves recurrent neural networks for handwriting recognition. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 285–290. IEEE, 2014.

[13] Joan Andreu Sanchez, Alejandro H Toselli, Veronica Romero, and Enrique Vidal. ICDAR 2015 competition HTRtS: Handwritten text recognition on the transcriptorium dataset. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*, pages 1166–1170. IEEE, 2015.

[14] Lambert Schomaker. Design considerations for a large-scale image-based text search engine in historical manuscript collections. *it-Information Technology*, 58(2):80–88, 2016.

[15] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.

Solving the Travelling Umpire Problem with Answer Set Programming

Joost Vennekens

KU Leuven, Dept. of Computer Science @ Technology Campus De Nayer, J.-P. De Nayerlaan 5, 2860 St-Katelijne-Waver, Belgium

Abstract

In this paper, we develop an Answer Set Programming (ASP) solution to the Travelling Umpire Problem (TUP). We investigate a number of different ways to improve the computational performance of this solution and compare it to the current state-of-the-art. Our results demonstrate that the ASP solution is superior to other declarative solutions, such as Constraint Programming, but that it cannot match the most recent special-purpose algorithms from the literature. However, when compared to the earlier generation of special-purpose algorithms, it does quite well.

1 Introduction

The Travelling Tournament Problem is a well-known optimisation problem, where the goal is to schedule a series of games at different venues such that the travel distance for the participating teams is minimized. The *Travelling Umpire Problem (TUP)* is a related problem, in which the games and their location are given and the goal is to assign umpires to each of these games. This assignment must satisfy a number of hard constraints (e.g., the same umpire should not referee the same team two weeks in a row) and the overall distance travelled by the umpires should be minimised.

This problem, which abstract a real-life task for the Major League Baseball (MLB), was first introduced in 2007 [8]. In the last five years, it has received a good deal of attention within the scientific community [9, 10, 11, 2, 1, 13, 7, 6, 14]. One of its interesting characteristics is that the TUP has a compact description, but is computationally very challenging (the associated decision problem is NP-complete [2]).

In this paper, we investigate whether it is possible to solve the problem using the declarative paradigm of Answer Set Programming (ASP). For this, we will make use of the state-of-the-art ASP solver Clasp [4], which is a regular winner of the ASP competition. As we will show below, the TUP can be represented in ASP in an elegant and modular representation. Such a representation has the advantage that it can easily be extended with additional constraints, in order to better capture the real-life MLB problem. The computational performance of our method is superior to that of other declarative paradigms, and matches that of the first special-purpose algorithm to be developed for TUP [9]. However, it cannot match the performance of more recent special-purpose algorithms, such as the state-of-the-art method from [7]. Nevertheless, it is of course a benefit of our declarative approach that improvements to the general-purpose Clasp server automatically speed-up our approach as well.

2 Preliminaries: Answer Set Programming

In this section, we briefly recall the Answer Set Programming (ASP) language and it semantics [5].

An *atom* is an expression $p(T_1, \ldots, T_n)$ where p is an n-ary predicate symbol ($n \geq 0$) and the T_i are either constants (starting with a lower case letter) or variables (starting with an upper case letter). A *(normal) rule* is an expression of the form:

$$A \ :- \ B_1, \ldots, B_n, \texttt{not}\ C_1, \ldots, \texttt{not}\ C_m, \tag{1}$$

Here, A and all of the B_i and C_j are atoms. The atom A is called the *head* of the rule and $B_1, \ldots, B_n$, `not` $C_1, \ldots,$ `not` C_m its *body*. A *(normal) logic program* is a finite set of such rules. An atom, rule or program is called *ground* if it does not contain any variables.

A *belief set* X is a set of ground atoms. Such a set X *satisfies* a ground rule r of form (1) if A belongs to X or there exists an $i \in [1, n]$ such that $B_i \notin X$ or a $j \in [1, m]$ such that $C_j \in X$. A belief set is a model of a ground program P if it satisfies all rules $r \in P$.

For a ground rule r of the form (3) and a belief set X, the reduct r^X is defined whenever there is no atom C_j for $j \in [1, m]$ such that $C_j \in X$. If the reduct r^X is defined, then it is the rule: A :– $B_1, \ldots, B_n$. The reduct P^X of the ground program P consists of all r^X for $r \in P$ for which the reduct is defined. A belief set X is a *stable model* or *answer set* of P, denoted $X \models_{st} P$, if it is the least model of P^X.

The answer sets of a non-ground program are defined as the answer sets of the *grounding* of the program. This grounding is constructed by replacing each non-ground rule r by the set of all ground rules that can be constructed by replacing all the variables in r by constants in all possible ways.

An answer set solver is a program that computes the stable models of a given program. Typically, these solvers extend the basic ASP language with some additional constructs to make it easier to represent interesting problems. In this paper, we will use the input language of the solver Clasp [3]. This language allows the following additional constructs.

A rule of the form

$$:- B_1, \ldots, B_n, \texttt{not}\ C_1, \ldots, \texttt{not}\ C_m, \tag{2}$$

is called a *constraint*. It is seen as an abbreviation for a rule

$$f\ :- B_1, \ldots, B_n, \texttt{not}\ C_1, \ldots, \texttt{not}\ C_m, \texttt{not}\ f. \tag{3}$$

where f is a fresh atom. The effect of such a rule is that no answer set may contain both all of the positive literals B_i and none of the negative literals C_i. A constraint therefore expresses that its body may *not* be satisfied.

Another common pattern in ASP programming is the use of a loop over negation to represent a choice between two alternatives. For instance, to express that either p or q must hold, the following two rules may be used:

```
p :- not q.      q :- not p.
```

The answer sets of this small programs are precisely {`p`} and {`q`}. However, because this style of programming quickly grows cumbersome, Claps also allows to represent choices as:

```
1 { p, q} 1.
```

This expresses that precisely one of the two atoms p and q must holds (the leftmost 1 states that at least one of these atoms must hold and the rightmost 1 states that at most one may hold). It has the same two answer sets as the above two rules (but the two programs are not strongly equivalent—for that, an additional constraint `:- p, q` should be added to the two rules).

This construct is made even more powerful by the fact that more complex set expressions may be used instead of a simple enumeration of atoms. For instance, `1{p(X) : q(X)}1` states that precisely one atom $p(X)$ must hold out of all atoms $p(X)$ for which X is such that $q(X)$ holds. If we add to this rule the following three facts:

```
q(1).  q(2).  q(3).
```

the answer sets will be precisely {`p(1)`}, {`p(2)`} and {`p(3)`}. The above three facts may also be conveniently abbreviated as: `q(1..3).`

3 The Travelling Umpire Problem

The TUP is the problem of assigning n umpires to a double round-robin tournament of $2n$ teams. Since each team plays every other team twice, there are $r = 2(2n - 1)$ different rounds, in which each team plays against exactly one other team. Each game is played at the home venue of one of the two participating teams. The assignment of umpires is subject to three hard constraints:

Round	Home	Away	Umpire
1	Team 1	Team 2	?
	Team 3	Team 4	?
2	Team 1	Team 3	?
	Team 4	Team 2	?
3	Team 2	Team 3	?
	Team 4	Team 1	?
4	Team 2	Team 1	?
	Team 4	Team 3	?
5	Team 3	Team 1	?
	Team 2	Team 4	?
6	Team 3	Team 2	?
	Team 1	Team 4	?

→

Round	Home	Away	Umpire
1	Team 1	Team 2	Umpire 1
	Team 3	Team 4	Umpire 2
2	Team 1	Team 3	Umpire 2
	Team 4	Team 2	Umpire 1
3	Team 2	Team 3	Umpire 1
	Team 4	Team 1	Umpire 2
4	Team 2	Team 1	Umpire 2
	Team 4	Team 3	Umpire 1
5	Team 3	Team 1	Umpire 2
	Team 2	Team 4	Umpire 1
6	Team 3	Team 2	Umpire 1
	Team 1	Team 4	Umpire 2

Figure 1: An instance of the TUP for $n = 2, q_1 = 2, q_2 = 1$ and its optimal solution.

- Each umpire should visit each team's home venue at least once.
- An umpire should not visit the same home venue more than once in any sequence of q_1 rounds.
- An umpire should not referee the same team more than once in any sequence of q_2 rounds.

The parameters q_1 and q_2 control the tightness of the constraints: higher values are more difficult. Clearly, it only makes sense to have $q_2 \leq q_1 \leq r/2$.

In addition to these three hard constraints, there is also an objective function that must be minimised, namely, the total distance travelled by all of the umpires (assuming they start in the location where their first game is played). To calculate this objective function, a (symmetric) distance matrix is given which enumerates the distances between the home venues of each pair of teams.

Figure 1 shows an example instance for $n = 2$ and a solution for the problem with $q_1 = 2$ and $q_2 = 1$ (note that this problem has no solutions for $q_2 > 1$; taking $q_2 = 1$ means that the third constraint can simply be ignored). The solution shown is also the optimal solution (regardless of which distance matrix is used), since the only other solution that exists (modulo the symmetry between the two umpires) is that in which they needlessly move around between rounds 4 and 5.

4 Representing the TUP in ASP

Problem instance. The parameters of an instance are represented by a set of ASP facts

`limit_big`(q_1). `limit_small`(q_2).
`team`$(1..2n)$. `umpire`$(1..n)$. `round`$(0..r)$.

The tournament schedule is given by a set of facts of the form `plays`(i, j, r), which denote that home-team i plays away-team j in round r. The distance matrix is represented by a set of facts `distance`(i, j, δ), indicating that the distance between the home venue of team i and that of team j is δ. Only facts for $i \leq j$ are included. The symmetric information is represented by a rule

```
dist(X,Y,D) :- dist(Y,X,D).
```

Finally, it is also convenient to include a unary predicate `last`$(r-1)$ which indicates the number $r-1$ of the last round and a unary predicate `max_dist`(δ_{max}) which indicates the largest number that occurs in the distance matrix.

The example shown in Figure 1 corresponds to the following set of facts:

```
limit_big(2).     plays(2,3,2).     dist(1,2,745).    dist(3,3,0).
limit_small(1).   plays(4,1,2).     dist(1,3,665).    dist(3,4,380).
team(1..4).       plays(2,1,3).     dist(1,4,929).    dist(4,1,929).
umpire(1..2).     plays(4,3,3).     dist(2,1,745).    dist(4,2,337).
round(0..5).      plays(3,1,4).     dist(2,2,0).      dist(4,3,380).
plays(1,2,0).     plays(2,4,4).     dist(2,3,80).     dist(4,4,0).
plays(3,4,0).     plays(3,2,5).     dist(2,4,337).    max_dist(929).
plays(1,3,1).     plays(1,4,5).     dist(3,1,665).    last(5).
plays(4,2,1).     dist(1,1,0).      dist(3,2,80).
```

Generating the search space We represent the possible solutions to the problem by a set of atoms `move`(u, t, r), meaning that umpire u moves to the home venue of team t in round r. For instance, the solution shown in Figure 1 would be represented as:

```
move(1,1,0).    move(1,4,3).    move(2,3,0).    move(2,2,3).
move(1,4,1).    move(1,2,4).    move(2,1,1).    move(2,3,4).
move(1,2,2).    move(1,3,4).    move(2,4,2).    move(2,1,4).
```

Obviously, in a valid solution these atoms must be such that team t actually plays a home game in round r. We first define which teams this are.

```
home_team(Home,R) :- plays(Home,Away,R).
```

We can then generate the search space by the following choice rule:

```
1 { move(X,Y,T) : home_team(Y,T) } 1 :- umpire(X), round(T).
```

Testing the hard constraints Each game must be assigned precisely one umpire. We represent this requirement as:

```
% Each game is officiated by at most one umpire
:- move(U1,T,R), move(U2,T,R), U1 != U2.

% Each game is officiated by at least one umpire
:- home_team(T), round(R), { move(U,T,R) : umpire(U) } 0.
```

The expression `{move(U,T,R) : umpire(U)} 0` holds if the set of all `move(U,T,R)` atoms for which `round(R)` holds is at most zero; in other words, if no umpires are assigned to the round R game in which team T is the home team. Given the way in which the search space is generated, it would actually be sufficient to add only of these two constraints. However, adding both of them helps the solver to solve the problem more quickly.

Next, we handle the constraint that each umpire should visit each venue . First, we define when an umpire U visits the home venue of team T:

```
been_to(U,T) :- round(R), move(U,T,R).
```

Now, the constraint is that this predicate must hold for all U and T:

```
:- umpire(U), team(T), not been_to(U,T).
```

Of the two constraints regarding repeated games with the same umpire, the constraint concerning home venues is easiest to represent (recall that `limit_big` contains the paramater q_1 of the instance):

```
:- move(U,T,R1), move(U,T,R2), R1 < R2, limit_big(B), R2 - R1 + 1 <= B.
```

To express the constraint regarding q_2, we first need to define when an umpire U officiates a game in which team T is involved (as either home or away team) in round R:

```
officiates(U,Home,R) :- move(U,Home,R).
officiates(U,Away,R) :- move(U,Home,R), plays(Home,Away,R).
```

Using this predicate, the constraint for the parameter q_1 (given by the predicate `limit_small`) is represented as follows:

```
:- officiates(U,T,R1), officiates(U,T,R2),
   R1 < R2, limit_small(S), R2 - R1 + 1 <= S.
```

Optimisation The rules and constraints that we have so far suffice to produce feasible solutions. To find the optimal solution, we first define the distance D that a given umpire U moved in round R (i.e., to go from his venue in round $R-1$ to his venue for round R):

```
moved(U,R,D) :- umpire(U), team(T), round(R), R > 0,
                move(U,T,R), move(U,Tp,R-1), dist(T,Tp,D).
```

The following statement instructs Clasp to minimize the sum of all these distances D:

```
#minimize { D,U,R : moved(U,R,D) }.
```

5 Improving the performance

The program that we have discussed so far correctly generates optimal solutions to the TUP. In this section, we discuss two additions to the program that may improve its computational performance.

A typical method for improving the efficiency of such a program is to introduce additional constraints to break symmetries. For the TUP, only one symmetry is known is literature, namely the fact that the umpires are all identical. We therefore introduce the following symmetry breaking rule, which forces one particular assignment in round 0:

```
:- move(U1,T1,0), move(U2,T2,0), U1 < U2, T1 >= T2.
```

A second possibility for improving the efficiency is to choose a suitable heuristic to guide the search process. Similar to, e.g., SAT solvers, an ASP solver works by iteratively selecting some atom that is currently neither true nor false, assigning it one of these two truth values, and then propagating the effects of this assignment. A heuristic is used to decide, first of all, which atom (among those that do not yet have a truth value) should be selected and, second, whether this atom should be made true or whether it should be made false. Clasp already contains a number of built-in heuristics, but it also allows users to provide a domain-specific heuristic by defining a predicate `_heuristic`(a, x, w). Here, a is an atom, w is a weight and x is a special constant. If $x =$`sign`, then w must be either 1 or -1. The meaning of such an atom `_heuristic(`a`,sign,-1)` is that whenever atom a is chosen, it will be given the value false; if instead $w = 1$, then a is assigned the value true.

For the TUP, a good heuristic might be to make choices that assign an umpire to a particular game (i.e., that make a `move(U,T,R)` atom true). It seems likely that these are the kind of choices that will lead to the most propagation: indeed, making a single such atom true will have at least the effect of forcing all $n - 1$ such atoms for the same game to be false. We therefore include the following rule:

```
_heuristic(move(U,T,R),sign, 1) :- round(R), team(T), umpire(U).
```

It still remains to decide *which* `move(U,T,R)` atom to choose first. To get the most propagation, it seems reasonable to handle the different rounds in order, and since the assignment in the first round is fixed by the symmetry breaking rule, it makes sense to work in ascending order. Within a round, we greedily attempt to send each umpire to the venue with the smallest travel distance. In this way, we hope to bias the search towards more optimal solutions. This is the same strategy as followed in [7]. We implement this strategy by assigning each `move(U,T,R)` atom a weight of $n\delta_{max} + (\delta_{max} - \delta)$, where δ_{max} is the maximum distance between venues, δ is the distance that the umpire would have to travel to get to the home venue of team T in round R, and n is the number of rounds left after the Rth round.

```
_heuristic(move(U,T,R),level,(L-R+1) * M + (M-D) ) :-
      last(L),round(R),team(T),umpire(U),
      move(U,T1,R-1),dist(T,T1,D),max_dist(M).
```

In addition to changing the heuristic, Clasp also has numerous configuration options that determine how it precisely conducts the search for a solution. Because the space of all possible configurations is huge, a tool called Piclasp has been developed, which tries to automatically deduce an optimal configuration for a given problem using the SMAC method. It does this by first experimenting with different settings on a number of training instances and then applying Machine Learning techniques to surmise an optimal configuration. This tool does not yet offer support for optimisation problems, however. Nevertheless, we have used it to determine a configuration for the problem of finding a feasible solution (i.e., one that satisfies all the hard constraints, without taking the optimisation criterium into account).

Finally, Clasp is also able to take advantage of multiple processors by running certain computations in parallel. It offers two strategies for this. The *compete* strategy launches a number of independent executions of the solver, each with different configuration parameters. The output is then simply the result produced by the best configuration. The *split* strategy essentially runs a single instance of the solver, splitting its search tree over different parallel threads. This has the advantage that the different threads actually help each other, rather than competing with each other and possible duplicating a lot of the work done by the other threads. The disadvantage, however, is that the different threads now need to communicate with—and thus wait for—each other, which may cause some of the capacity to go unused.

Benchmark	def-comp	dom-split-π	def-1-π	def-1	dom-split	dom-1	dom-comp	dom-1-π	def-split	dom-comp-π
14-5-3	0.14	0.18	0.20	0.15	0.13	0.15	**0.13**	0.15	0.14	0.17
14-6-3	0.10	0.13	0.19	0.12	0.11	0.11	0.11	0.16	**0.09**	0.13
14-7-3	0.07	0.10	0.14	0.06	0.07	0.07	0.07	0.13	**0.06**	0.07
14A-5-3	0.17	0.21	0.23	0.18	**0.16**	**0.16**	0.16	0.21	0.16	0.19
14A-6-3	0.10	0.14	0.23	0.12	0.12	**0.10**	0.13	0.18	0.11	0.15
14A-7-3	**0.06**	0.12	0.20	0.11	0.09	0.09	0.07	0.11	0.08	0.12
14B-5-3	0.14	0.17	0.18	0.16	**0.13**	0.15	0.14	0.18	0.14	0.17
14B-6-3	0.12	0.14	0.15	0.13	**0.11**	0.12	0.12	0.14	0.13	0.16
14B-7-3	**0.07**	0.10	0.16	0.10	0.09	0.09	0.08	0.11	0.09	0.11
14C-5-3	0.16	0.18	0.22	0.18	**0.15**	0.17	0.16	0.21	0.19	0.18
14C-6-3	0.15	0.16	0.21	0.15	**0.10**	0.13	0.13	0.13	0.13	0.18
14C-7-3	**0.11**	0.13	NoSol	0.12	0.11	0.12	0.11	0.16	0.11	0.16

Table 1: Comparison of different variants for size $n = 14$.

6 Experimental results

In this section, we present some experiments that were conducted to determine which ways of running Clasp perform best for the TUP and how our solution compares to the state of the art. For the first point, we want to answer three questions:

- Whether the domain-specific heuristic discussed in the previous section performs better than Clasp's default heuristic (which it selects based on some properties of the problem instance);
- Whether the configuration settings learned by Piclasp perform better than the default configuration settings (which Clasp again selects based on properties of the problem instance);
- Whether using multiple processors in parallel leads to better results and, if so, which of the two possible strategies (*compete* or *split*) is better.

The results are shown in Table 1. In the name of the different variants, `def` and `dom` refer to the use of, respectively, Clasp's default heurstic versus our domain-specific heuristic; `1`, `comp` and `split` refer to whether only a single processor was used, or whether all available processors were used with either the *competitive* or *split* strategy; finally, π is added to the name whenever the settings learned by Piclasp where used instead of the default settings. All results in this table were computed on a Linux machine containing eight Intel(R) Core(TM) 3.2 GHz processors. A timeout of 10 minutes was used. The table reports the value for the optimisation criterion that was reached at the timeout. The benchmarks instances come from [9].

A first observation about these results is that the default settings consistently perform better than those learned by Piclasp, regardless of which heuristic is being used and regardless of whether parallel processing is used. This may be due to the fact that Piclasp does not support optimisation problems, or to the fact that we supplied it with relatively small training instances (because it needs to run a large number of experiments). Second, on a single processor, our domain-specific heuristic is almost always (in 10 of the 12 benchmarks) better than the default heuristic. In competitive multi-processor mode, the default heuristic has a slight edge (7 of the 12 benchmarks), whereas in split mode, the domain dependent heuristic has the edge (also 7 out of 12). When the default heuristic is used, competitive mode beats split mode (8 out of 12), whereas split mode always beats competitive mode for the domain-dependent heuristic. Regardless of the heuristic, using all 8 processors instead of just a single processor leads to better results *if* the appropriate multi-processor strategy is chosen (e.g., dom-1 beats dom-comp, but loses to dom-split in 9 benchmarks).

We conclude from these results that the best variants are dom-split and def-comp. In the rest of our experiments, we use these two variants, together with dom-1, the best single-processor variant. We now compare these three variants with results from [9]. This article reports on three different approaches: a Constraint Programming (CP) approach, an Integer Programming (IP) approach (both using ILOG OPL Studio 3.7 as a solver), and the special-purpose search algorithm GBNS developed by its authors. The results reported for these three approaches are taken from [9] and were performed on an Intel(R) Xeon(TM) processor which ran at the same clock speeds of 3.2 GHz as the Intel(R) Core(TM) processor we used to test our own approach.

Table 2 shows results for $n = 14$. All results in this paper are given as the percentage difference with the best known solution from [12]. Our own approaches were benchmarked with a timeout of *10 minutes*, while the results reported for the approaches from [9] use a timeout of *3 hours*. Even taking into account the fact that our multiprocessor variants use 8 processors instead of one, this still puts our approach at a disadvantage. Nevertheless, as can be seen in Table 2, it performs quite well. All of our variants beat Trick et al.'s CP solution on all benchmarks. Against Trick et al.'s IP solution, dom-split performs best, winning 9/12 benchmarks; our single-processor variant wins 8, while dom-split wins

Benchmark	def-comp	IP [9]	dom-split	GBNS [9]	CP [9]	dom-1
14-5-3	0.14	0.13	0.13	**0.08**	0.18	0.15
14-6-3	**0.10**	0.15	0.11	0.11	0.16	0.11
14-7-3	0.07	0.15	**0.07**	0.12	0.09	0.07
14A-5-3	0.17	0.12	0.16	**0.10**	0.20	0.16
14A-6-3	0.10	0.13	0.12	0.12	0.20	**0.10**
14A-7-3	**0.06**	0.17	0.09	0.11	0.09	0.09
14B-5-3	0.14	0.12	0.13	**0.08**	0.19	0.15
14B-6-3	0.12	0.16	**0.11**	0.17	0.20	0.12
14B-7-3	**0.07**	0.21	0.09	0.09	0.10	0.09
14C-5-3	0.16	**0.12**	0.15	0.14	0.20	0.17
14C-6-3	0.15	0.10	**0.10**	0.14	0.24	0.13
14C-7-3	**0.11**	0.16	0.11	0.22	0.19	0.12

Table 2: Comparison for $n = 14$ between our methods (timeout: 10min) and those of [9] (timeout: 3h).

Benchmark	def-comp	CP [9]	GBNS [9]	dom-split	IP [9]	dom-1
16-7-2	0.27	0.34	**0.11**	0.26	0.14	0.26
16-7-3	0.18	0.32	0.17	**0.14**	NoSol	0.16
16-8-2	0.16	0.24	0.11	0.18	**0.10**	0.18
16-8-4	**UNSAT**	NoSol	NoSol	NoSol	NoSol	**UNSAT**
16A-7-2	0.25	0.36	**0.09**	0.22	0.17	0.23
16A-7-3	0.17	0.18	0.17	**0.15**	NoSol	0.15
16A-8-2	0.18	0.29	**0.10**	0.16	0.11	0.19
16A-8-4	**UNSAT**	NoSol	NoSol	NoSol	NoSol	**UNSAT**
16B-7-2	0.26	0.36	**0.11**	0.24	0.14	0.25
16B-7-3	0.18	0.25	NoSol	**0.18**	NoSol	0.18
16B-8-2	0.17	0.22	**0.11**	0.14	0.12	0.17
16B-8-4	**UNSAT**	NoSol	NoSol	NoSol	NoSol	NoSol
16C-7-2	0.21	0.23	**0.09**	0.21	0.18	0.21
16C-7-3	0.14	0.20	NoSol	**0.13**	0.18	0.14
16C-8-2	**0.11**	0.21	0.12	0.12	0.15	0.12
16C-8-4	**UNSAT**	NoSol	NoSol	NoSol	NoSol	**UNSAT**

Table 3: Comparison for $n = 16$ between our methods and those of [9] (both timeouts: 3 hours).

only 7. Against their special-purpose GBNS algorithm, our single-processor variant and dom-split both win 8 benchmarks, while def-comp wins 7.

Table 3 shows experimental results for $n = 16$. Here, our own approaches were benchmarked with the same timeout of 3 hours as in [9]. An entry of `UNSAT` in this table means that the system was able to report that no feasible solution exists; an entry of `NoSol` means that the system was unable to decide whether a feasible solution exists within the time limit. Als for $n = 16$, the CP solution does not win a single benchmark against any of our approaches. Against IP, our def-comp approach wins 9 out of 16 benchmarks, while dom-split wins only 5 and loses 7. Against GBNS, def-comp wins half of the benchmarks against GBNS and loses the other half. Our dom-split approach does worse, winning only 5 benchmarks against GBNS and losing 7.

In conclusion, our approach clearly outperform the CP approach and does slightly better than IP. It is also on par with the GBNS special-purpose algorithm. However, the same cannot be said for more recent special-purpose algorithms. The currently state-of-the-art the approach of [7] is able to find the optimal solution for each of the $n = 14$ benchmarks and most of the $n = 16$ ones. The best results reported in Tables 2 and 3 above are typically around 10% worse than those of [7].

7 Conclusions

A declarative approach allows computational problems to be solved without requiring the development of special-purpose algorithms. An advantage of such an approach is that the problem specification can easily be changed, by adding or replacing certain constraints. Moreover, developments in solver technology immediately improve the performance of declarative solutions for numerous different problems. A disadvantage is that the computational performance of declarative solvers often lags behind that of special-purpose algorithms, since it may take some time to lift improvements that were made to specific algorithm for one specific problem to the level of a generic reasoning tool. However, features such as domain-specific heuristics allow domain knowledge to be used to improve the performance of the solver.

In this paper, we examined a declarative solution for the Travelling Umpire Problem, which is a challenging optimisation problem that has recently received a lot of attention. As we have shown, it can be formulated in an elegant and modular way in the declarative ASP paradigm. Because the TUP as

considered in the literature is only an abstraction of a real-life problem in Major League Baseball, the flexibility to easily add or change constraints is a useful feature for this application.

Using the state-of-the-art ASP solver Clasp, we found that the performance of our approach improves on that of previous declarative solutions and is on par with the first special-purpose algorithm published for this problem. However, it cannot match the performance of current special-purpose algorithms. This suggests that these algorithms use more advanced techniques which have not yet found their way into general-purpose ASP solvers. We believe the TUP might therefore be an interesting benchmark to guide future developments in ASP solver technology.

References

[1] L. de Oliveira, C. de Souza, and T. Yunes. Improved bounds for the traveling umpire problem: A stronger formulation and a relax-and-fix heuristic. *EJOR*, 2:592–600, 2014.

[2] L. de Oliveira, C. de Souza, and T. Yunes. On the complexity of the traveling umpire problem. *Theoretical Computer Science*, pages 101–111, 2015.

[3] M. Gebser, A .Harrison, R. Kaminski, V. Lifschitz, and T. Schaub. Abstract gringo. *TPLP*, 2015.

[4] M. Gebser, R. Kaminski, B. Kaufmann, M. Ostrowski, T. Schaub, and M. Schneider. Potassco: The Potsdam answer set solving collection. *AI Communications*, 24(2):107–124, 2011.

[5] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In R. Kowalski and K. Bowen, editors, *ICLP*, pages 1070–1080, Cambridge, MA, 1988. MIT Press.

[6] T. Toffolo, S. Van Malderen, T. Wauters, and G. Vanden Berghe. Branch-and-price and improved bounds to the traveling umpire problem. In *PATAT*, 2014.

[7] T. Toffolo, T. Wauters, S. Van Malderen, and G. Vanden Berghe. Branch-and-bound with decomposition-based lower bounds for the traveling umpire problem. *EJOR*, 2015.

[8] M. Trick and H. Yildiz. Benders' cuts guided large neighborhood search for the traveling umpire problem. In *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, pages 332–345, 2007.

[9] M. Trick and H. Yildiz. Benders' cuts guided large neighborhood search for the traveling umpire problem. *Naval Research Logistics (NRL)*, 8:771–781, 2011.

[10] M. Trick and H. Yildiz. Locally optimized crossover for the traveling umpire problem. *EJOR*, 2:286–292, 2012.

[11] M. Trick, H. Yildiz, and T. Yunes. Scheduling major league baseball umpires and the traveling umpire problem. *Interfaces*, 3:232–244, 2012.

[12] T. Wauters. `http://benchmark.gent.cs.kuleuven.be/tup/` (Accessed: 27 oct 2016).

[13] T. Wauters, S. Van Malderen, and G. Vanden Berghe. Decomposition and local search based methods for the traveling umpire problem. *EJOR*, 3:886–898, 2014.

[14] Li Xue, Zhixing Luo, and Andrew Lim. Two exact algorithms for the traveling umpire problem. *EJOR*, 3:932–943, 2015.

Reasoner-Aid Research: Potentials and Popularity

Shuai Wang

ILLC, University of Amsterdam, The Netherlands

Abstract

This paper presents the first study on the popularity of reasoner-aid research among students. To understand how students are accepting and conducting research with the help of reasoners, a survey was conducted among students in two Dutch universities. The analysis shows that the majority has little knowledge about this approach and a small amount of students have experience with reasoner-aid research. This paper summarise the benefit of reasoner-aid research, reasons of its limited popularity and discuss the potentials of this approach.

1 Introduction

Since the first use of computers for research to find large prime numbers, computer software and hardware have been playing an important role in research activities. Among them, reasoners have a long history in assisting research activities. Reasoners are programs made to assist theorem proving, verify proofs, find counter examples and test and discover properties of mathematical and logical problems. After decades of development, these programs are more powerful than ever before and capable to free researchers from many repetitive and tedious calculations and reasoning tasks. However, not every mathematician, logician, and computer scientist knows how to use such computer programs to assist their research. Many have never heard about this class of programs and this approach. The project investigates the popularity of reasoners and the gap between the use and the development of reasoners in logic, mathematics and computer science research. This paper presents a qualitative study by summarising the history and recent advances of this approach, and a quantitative analysis by conducting a survey among students about the following question.

How much do mathematics, logic and computer science students know using reasoners in their research?

In this paper, we take a general definition of reasoners. We include all interactive and automatic, domain-specific and general-purpose reasoners, developed or maintained in academia or industry. Despite the development and increasing use of reasoners during the past decades, not many scientists and students know how to use reasoners to assist their research. Moreover, the quality, reliability and usability of these programs are questioned by many.

Hypothesis: For the students in the survey, the majority have little knowledge about this approach and a minority have some limited knowledge and experience with reasoners for research purposes.

This paper is organised as follows. Section 1 summarise the history and recent advances of reasoner-aid research and presents the motivation to conduct the survey. Section 2 is about the method of the survey. Section 3 study the results of the survey and Section 4 further discuss the results and the reasons of its limited popularity.

1.1 Background

Human reasoning has been playing a central role in innovation, especially mathematical and logical research. Deductive reasoning has been considered the key of mathematical and logical research and has

been one of the major interest in psychology and cognitive science for decades [5, 15]. In mathematics and logic, as defined in [15], a formal proof is a finite sequence of sentences in which each sentence is either a premise, an axiom of the logical system, or a sentence that follows from preceding sentences by one of the system's rules. An argument is deducible in the system if there is a proof whose final sentence is the conclusion of the argument. Theorem proving and proof checking are among the main tasks of mathematicians and logicians. Without the aid of computers, proofs are generated and checked manually. Thanks to the development of reasoners over the past decades, some part of these reasoning activities can be automated [1]. Depending on the degree of automation, we distinguish between automated and interactive theorem proving. Once the axioms, inference rules, premises and the argument are specified, if a computer program can carry out the proof without interacting with the user, it is generally considered an automated reasoner or Automated Theorem Prover (ATP). If interaction with a user is necessary, such reasoners are named Interactive Theorem Provers (ITPs), or Proof Assistants. Depending on the system, the inference rules and some axioms are either embedded in a reasoner or specified while producing the proof of a statement or check a proof in a certain setting [9, 16]. The former is considered universal or general-purpose, the latter is more domain specific. While most reasoners are maintained in academia, there are some reasoners involved in solving problems in industry and are supported or maintained by organisations outside of academia or companies.

The past few decades witnessed the development and wide use of reasoners. Reasoners are more user-friendly and able to find harder proofs and faster than before. The proof of four color theorem was formalised in the Coq proof assistant to check the correctness by Georges Gonthier and Benjamin Werner in 2005 [7]. A milestone is the automatically found proof of the Robbins Conjecture and its postprocessing into a human-comprehensible proof [1]. More recently, the Flyspeck project, a collaborative project to produce a complete formal proof of the Kepler Conjecture, was announced completed on August 10, 2014 [8]. In recent years, a large number of proofs were generated and many of them are grouped in libraries. The Mizar Mathematical Library (MML) consists of 50,000 theorems and more than 100,000 premises [1]. A recent evaluation of ATP systems on MML shown that Vampire can reprove 39% of the proofs in MML when necessary premises are selected [1]. Most recently, Machine Learning methods are being used to train theorem provers on existing proofs for better automation of proof discovery [11]. In addition, SAT solver are being use to find counter examples for mathematical hypothesis. For example, Boris Konev and Alexei Lisitsa proved that by encoding the problem of Erdős discrepancy conjecture into Boolean satisfiability and applying SAT solvers, one can obtain a discrepancy 2 sequence of length 1600 and a proof sequence of length 1161 [14]. One of the most popular applications outside of proving mathematical and logical theorems is model checking and verification. Temporal model checkers, Binary Decision Diagram (BDD) solvers, SAT solvers and reasoners of Quantified boolean Formulas (QBF) are playing important roles in these tasks. Recent years, SMT solvers have contributed to the research of many domains including synthetic biology and structural biology [20], social choice theory [6], railway system design [10], etc. Once generated from a reasoner, the correctness of such proofs can be trivial to validate within the system itself. This leads to programs designed specifically to verify proofs. For example, Dedukti [16] is a universal proof checker for this purpose. Recently, there is an increasing interest in modelling human reasoning within automated reasoning systems using answer set programming, abductive logic programming, declarative programming, etc [3, 13]. These logic-based attempts cuts across cognitive modelling based on cognitive architectures, artificial intelligence, logic itself, psychology of reasoning and research design, and connect various disciplines [3]. This logic-based approach has many potentials in empirical research. Scientific knowledge represented in formal languages are with sufficient semantics and less ambiguity [13, 17]. For example, there is an increasing use of ontology for medicine and physiological research such as patient information modelling, cancer research and drug design [2, 12, 19]. A natural extension of this question is the idea of a robot scientist [13]. A prototype work is the robot scientists Adam and Eve [13, 18].

The Netherlands is home to many reasoners (e.g. Automath [4]) as well as users of reasoners. As reasoners get more and more important in research, they appear more and more often in students' research projects. All this lead to a necessary rethink about the role of reasoners and a study about how students are understanding and accepting this approach and using them as tools. This paper is to study the popularity among students in two Dutch universities.

2 Method

2.1 Design

A survey was designed to study mathematics, logic and computer science students' knowledge of reasoners for research. Considering Bachelor's students' research experience is very limited, we study only that of Master's and PhD students in *Universiteit van Amsterdam* (known as University of Amsterdam in English, UvA in short) and the *Vrije Universiteit Amsterdam* (VU University Amsterdam in English, VU in short). More specifically, the survey collected data from Master's students and PhD students registered at the Mathematics department of both universities, the Informatics Institute of UvA, the Computer Science department of VU, as well as the Institute for Logic, Language and Information (ILLC) of UvA. Here, by PhD students, we mean the *promovendi* since in the Dutch education system, PhD students are officially not students. The Master's students within the study domain consists of around 300 students from 10 master programmes of UvA and around 250 students from 9 master programmes of VU. As for doctoral level, there are around 87 students from the Informatics Institute of UvA, 32 from the Mathematics department, and 66 PhD students from ILLC. For VU, there are around 63 PhD students from the Computer Science department and 31 PhD students from the Mathematics department. This study excluded post-doctoral researchers, university researchers with more experience (associate professors, professors, etc) and researchers and engineers in industry. The questions were grouped into three sections. **Section 1** took some general information from the participants: every participant was required to specify if he or she is a Master's student or a PhD student. **Section 2** studied participants' knowledge about this approach. Moreover, the research activities the participants have involved and the reasoners used were asked. **Section 3** collected participants' view about the learning and usability of reasoners as well as the trust of results and willingness to learn more about this approach. More specifically, the three sections of questions were as follows. Most questions were multiple choice questions, except 3a), 3d) and 3e), which were with a linear scale of 1 to 5. Due to the limit of space, a summary of the survey organised by section is as follows. A copy of the complete survey is online at `http://goo.gl/forms/kOTOV38HIPTfJ98e2` and a summary of the survey questions is attached in Appendix A.

2.2 Survey Procedure

The questionnaire was constructed and managed using Google Forms. The survey was conducted from the morning of Tuesday 21st June until Midnight 22nd June. It was spread in two ways. The link to the questionnaire was first posted on Facebook groups of master and PhD students of UvA and VU. Following that, emails were sent to many (randomly selected) PhD students within the study domain. The data was exported to a spreadsheet which is available online at `https://github.com/airobert/C3GI`. This survey method was chosen since many Master's students no longer have courses during this survey. While most PhD students work in their own offices. For this reason, conducting an online survey is less time-consuming compared with a face-to-face approach. More specifically, the survey received 63 responses on the first day and 18 on the second day.

3 Results

The survey obtained 81 responses. Among which, 76 are valid[1]. More specifically, 47 participants are Master's students and 29 are PhD students. In other words, around 8.5% of the Master's students participated in the survey and around 10.4% for that of PhD students. Around 61.8% of the participants were Master's students and th7e rest (around 38.2%) were PhD students. The sample was therefore representative based on the amount of valid entries obtained.

Section 2 consisted of three questions about the participants' knowledge about using reasoners for research. The first question asked about the user's knowledge about reasoners. 43.4% of the participants have never heard about reasoners of any kind and 19.7% know reasoners, but not for research. That was, 48 out of 76 (63.1%) participants knew little or nothing about this approach. An additional 22.4% had some awareness about how a reasoner could help by reading an article, a report or attending a presentation. Only two out of 9 mathematics Master's students had some awareness about this approach. At least 9.2% of those have learnt using a reasoner from courses by the universities and 5.2%

[1]Entry 13, 55 and 70 are not within the study domain and entry 56 and 59 are incomplete or self-contradictory. They are filtered out during data analysis.

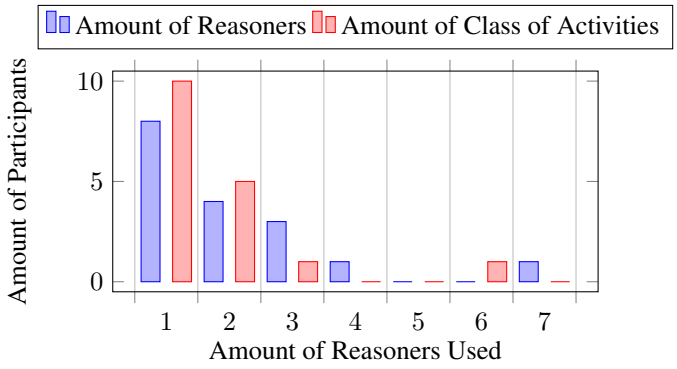


Figure 1: Amount of Reasoners/Class of Activities

obtained some more knowledge and experience using at least one reasoner for research. Question 2 further studied what the popular reasoners are. The result reflected the findings of Question 1. 56 out of 76 participants have never used any reasoners so far in their studies or research. The most popular reasoners were SAT solvers (13.2%), inductive reasoners (9.2%) and type inference systems (5.3%). Aside from these, MAXSAT solvers, first-order theorem provers, temporal model checkers had a popularity of 3.9% each. Abductive reasoners and BDD solvers were less popular in comparison (2.6% each). Only one student had experience with answer set programming and another student have used higher order theorem prover. No participant had experience with SMT solver, QBF solver or other modal logic solvers (apart from those of temporal logic). Question 3 studied the actual research research tasks involved. Out of 17 valid responses, 7 had experience with model checking using reasoner (around 41.1%), making it the most popular research activity with the reasoners. 5 participants (29.4%) showed that they have used reasoners to prove mathematical theorems and 6 have had experience verifying existing proofs (35.2%). Reasoning about networks (including semantic web) was the fourth most popular research activity in study (4 participants, 23.5%). Robot and agent systems and speech and language technology were equally popular (2 participants each, 11.7%). Another two participants reported their experience with cryptography, type inference in compilers, and finding hard instances of Sudoku, respectively. As for the amount of reasoners these 17 participants have used, most of participants have used less than three reasoners. More specifically, half of them have used only one reasoners for research. 4 of them have experience with 2 reasoners. Another 3 participants have used 3 reasoners. Finally, one has used 4 reasoners so far and another has experience with 7 reasoners in total. By comparing the results of Question 2b) and 2c), as in Figure 1, the amount of reasoners one can use does not determine how many activities the reasoner(s) can be used. For example, one participant can use ASP solvers for both model checking and network reasoning.

Section 3 consisted of 5 questioned about the views of participants on using reasoners for research. Figure 2 indicated that most people find the difficulty modest. Out of 21 valid responses of question 3b), 18 of them were valid. 57.1% agreed that learning a reasoner takes some time. 71.4% admitted that some mathematical, logic, computer science knowledge was required to use a reasoner for research. 42.9% of the participants have been introduced to using reasoners for research in their university education (so far). In comparison, 9.5% of them stated that they have not been introduced (yet). 38.1% complained that some reasoners lack of documentation or tutorials, which made it hard to learn. Question 3c) studied the usability of reasoners in comparison with traditional pen and paper approach. It received 18 valid responses. 8 of them agreed that this approach speed up problem solving process in general and makes research easier. In comparison, 3 of them thought it could make some research problems even more complicated. Half of them admitted that it was the only approach for some research problems. Another two indicated that this approach was not applicable in many domains. One suggested that this approach was not applied in a large scale yet. Question 3d) got 24 responses (Figure 3). The average for this question was 3.58, indicating that the proofs generated by reasoners were well trusted. The last question (Figure 4) was compulsory for all participants. A bit less than 30% of them did not see themselves using reasoners more for research. Only just over ten percent of the participants were more certain that they will use reasoners more in research in the future. The overall average is 2.3. For those 16 participants who have used reasoners for research, the average was slightly higher (3.1).

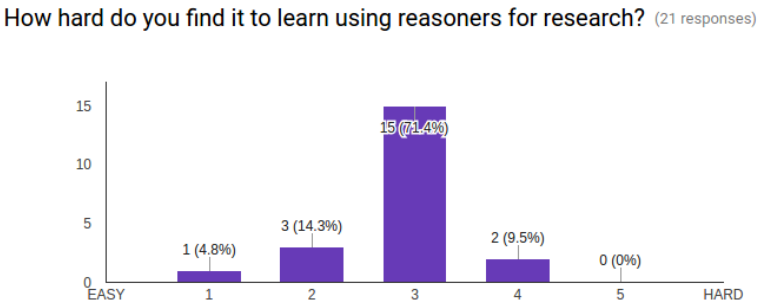


Figure 2: Reply Distribution of Question 3a)

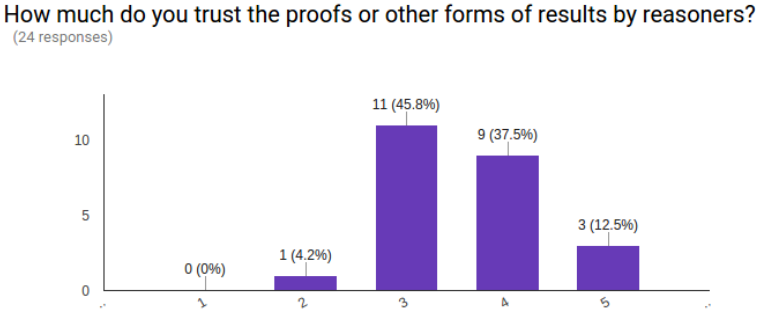


Figure 3: Reply Distribution of Question 3d)

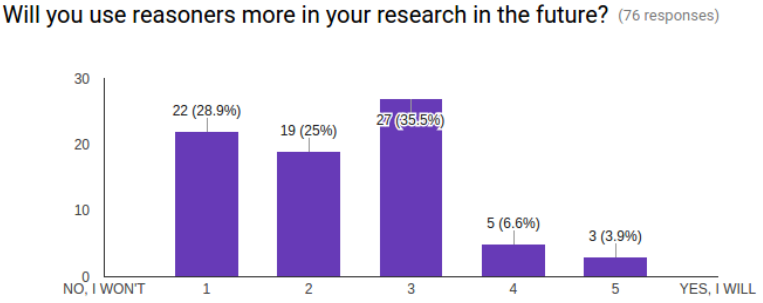


Figure 4: Reply Distribution of Question 3e)

4 Discussion

The hypothesis of the survey was that the majority had little or no knowledge about this approach while a minority had a limited amount of knowledge and experience. The survey results confirmed the hypothesis. Among those who had some experience using reasoners for research, the overall average about their willingness to use reasoner for their future research was 3.1, which was not supportive enough to conclude that using reasoners was getting more popular (among students). Both UvA and VU offered a good amount of courses using, or at least introducing reasoners. In UvA, courses such as *Knowledge Representation*, *Knowledge Representation on the Web* and *Functional Specification of Algorithms* included assignments or projects using reasoners as part of the course. Other courses such as *Computational Complexity* and *Combinatorics with Computer Science Applications* were highly likely to introduce the use of reasoners for some research tasks. For Master's students of VU, *Protocol Validation*[2], *Logical Verification* and *Term Rewriting Systems* were available courses where reasoners are introduced. Overall, at least 9.0% of those have learnt using a reasoner from courses by the universities. Considering some of the courses were compulsory for some Master's students, this rate showed that this approach is still to be promoted. However, in both university's mathematics master programme, there was no course using reasoners of any kind as far as the author knew from the description of courses on the websites. This was reflected in the result, only two out of nine mathematics Master's students

[2]This course was available to many students in UvA also.

had some basic awareness about this approach. No mathematics PhD student had awareness about this approach. As for doctoral level, many research groups' work took advantage of reasoners. In VU, the Knowledge Representation Reasoning group and the Theoretical Computer Science group were actively involved in developing and using reasoners. As for UvA, ILLC was a major player of logic research. However, there were too few responses from PhD students in ILLC to conclude anything. The Qualitative Reasoning Group of the Theory of Computer Science section was the major user of reasoners in the Informatics Institute. As a consequence of active research in using and developing reasoners, VU and UvA computer science PhD students' knowledge about reasoners was better than that of Master's students and some were involved developing reasoners. The survey also showed that ITPs are less popular than ATPs in general. There were more classes of research activities involving automated provers and many were specialised for certain tasks only (e.g. MAXSAT). In comparison, ITPs were mostly used for proving theorems. The study also showed that, proving mathematical and logical theorem may no longer be limited to first order or higher order theorem provers, etc. Type inference systems, ASP solvers and SAT solvers were also being used for such activities. SAT solver was the most popular reasoner due to its fundamental role in the development of reasoners and assignments in courses. However, no participant had experience with SMT solver. Instead, a more specialised class of reasoners, MAXSAT solvers, were relatively familiar to students. Taking into consideration that two of the three who have used MAXSAT solvers have also used SAT solvers, a possible reason could be that MAXSAT implements similar functions as SAT solvers and of similar use. Another output worth addressing was that inductive reasoners were very popular among both logic and computer science Master's and PhD students. As far as the course category showed, there was no course about inductive reasoning at Master's level. The best guess is that Prolog was taught as a part of their Bachelor's education and it was natural for some to use some Prolog programs as inductive reasoners. Another factor could be that Amsterdam had a big Prolog community. However, the actual reason remained unclear to the author.

It was shown that the majority agreed that there was an overhead of learning if this approach was chosen. Many agreed that learning reasoners took time and required specific knowledge. ATP Users usually need to specify one problem to the reasoner. This process is called encoding. The output of the reasoner is then decoded to meaningful solutions of the original problem. This learning process can be time consuming and the encoding and decoding process can be confusing to many. However, this encode-solve-decode process is a general pattern for many automated reasoners. Once one reasoner is learnt, learning and using the others are of similar style. A barrier is that many reasoners lack of documentation or user manual, which makes some participants suffer while learning. No participant admited that mailing lists help understanding reasoners. In fact, the mailing list of a reasoner is usually a platform where developers and users help each other. Considering only one participant have experience developing a reasoner, it was not representative enough to conclude anything valuable. In addition, over half of those who have used reasoners agreed that it is the only approach for some problems, which emphasised the status of the use of reasoners in research.

There were several drawbacks of this survey. Question 1b) and ac) may be merged. It would be better to study the source of knowledge about this approach (i.e. if it was taught in a course as a separate question (see Question 2a) and 3b)). As for Question 3e), the question itself suggested that participants suppose to use reasoners more, which can be misleading. Also it was compulsory, which made little sense to those who did not even know what a reasoner is. Following this work, future research may include more studies about senior researchers and details about user experience and trends. The results showed that most students, especially the mathematics students lack of knowledge of this approach. As reasoners get more usable and powerful, introducing reasoner-aid research may be done systematically in some Master's courses or even Bachelor's courses.

5 Conclusion

This paper provides a summary of recent advances of interactive and automated reasoning and reasoner-aid research. Although there were several aspects to be improved in the questionnaire, the survey results successfully provided a general understanding of the popularity of reasoner-aid research among Master and PhD students in UvA and VU. The research results confirmed the hypothesis that most students lack of the knowledge of this approach and a minority have some limited experience.

References

[1] Jesse Alama, Daniel Kühlwein, and Josef Urban. Automated and human proofs in general mathematics: An initial comparison. In *International Conference on Logic for Programming Artificial Intelligence and Reasoning*, pages 37–45. Springer, 2012.

[2] Matt-Mouley Bouamrane, Alan Rector, and Martin Hurrell. Using OWL ontologies for adaptive patient information modelling and preoperative clinical decision support. *Knowledge and information systems*, 29(2):405–418, 2011.

[3] Selmer Bringsjord. Declarative/logic-based computational cognitive modeling. *The Handbook of Computational Cognitive Modeling, Cambridge University Press, Cambridge (forthcoming)*, 2008.

[4] Nicolaas Govert De Bruijn. The mathematical language AUTOMATH, its usage, and some of its extensions. In *Symposium on automatic demonstration*, pages 29–61. Springer, 1970.

[5] Jonathan St BT Evans. Logic and human reasoning: an assessment of the deduction paradigm. *Psychological bulletin*, 128(6):978, 2002.

[6] Christian Geist and Ulrich Endriss. Automated search for impossibility theorems in social choice theory: Ranking sets of objects. *Journal of Artificial Intelligence Research*, 40(1):143–174, 2011.

[7] Georges Gonthier. *The Four Colour Theorem: Engineering of a Formal Proof*, pages 333–333. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[8] Thomas Hales, Mark Adams, Gertrud Bauer, Dat Tat Dang, John Harrison, Truong Le Hoang, Cezary Kaliszyk, Victor Magron, Sean McLaughlin, Thang Tat Nguyen, et al. A formal proof of the kepler conjecture. *arXiv preprint arXiv:1501.02155*, 2015.

[9] John Harrison. HOL Light: an overview. In *International Conference on Theorem Proving in Higher Order Logics*, pages 60–66. Springer, 2009.

[10] Alexei Iliasov, Ilya Lopatkin, and Alexander Romanovsky. The SafeCap platform for modelling railway safety and capacity. In *International Conference on Computer Safety, Reliability, and Security*, pages 130–137. Springer, 2013.

[11] Cezary Kaliszyk and Josef Urban. FEMaLeCoP: Fairly efficient machine learning connection prover. In *Logic for Programming, Artificial Intelligence, and Reasoning*, pages 88–96. Springer, 2015.

[12] Ross D King, Stephen Muggleton, Richard A Lewis, and MJ Sternberg. Drug design by machine learning: The use of inductive logic programming to model the structure-activity relationships of trimethoprim analogues binding to dihydrofolate reductase. *Proceedings of the National Academy of Sciences*, 89(23):11322–11326, 1992.

[13] Ross D King, Jem Rowland, Stephen G Oliver, Michael Young, Wayne Aubrey, Emma Byrne, Maria Liakata, Magdalena Markham, Pinar Pir, Larisa N Soldatova, et al. The automation of science. *Science*, 324(5923):85–89, 2009.

[14] Boris Konev and Alexei Lisitsa. A SAT attack on the erdős discrepancy conjecture. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 219–226. Springer, 2014.

[15] Lance J Rips. *The psychology of proof: Deductive reasoning in human thinking*. Mit Press, 1994.

[16] Ronan Saillard. Dedukti: a universal proof checker. In *Foundation of Mathematics for Computer-Aided Formalization Workshop*, 2013.

[17] Larisa N Soldatova and Ross D King. An ontology of scientific experiments. *Journal of the Royal Society Interface*, 3(11):795–803, 2006.

[18] Andrew Sparkes, Wayne Aubrey, Emma Byrne, Amanda Clare, Muhammed N Khan, Maria Liakata, Magdalena Markham, Jem Rowland, Larisa N Soldatova, Kenneth E Whelan, et al. Towards robot scientists for autonomous scientific discovery. *Automated Experimentation*, 2(1):1, 2010.

[19] Dennis G Thomas, Rohit V Pappu, and Nathan A Baker. NanoParticle Ontology for cancer nanotechnology research. *Journal of biomedical informatics*, 44(1):59–74, 2011.

[20] Boyan Yordanov, Christoph M Wintersteiger, Youssef Hamadi, and Hillel Kugler. SMT-based analysis of biological computation. In *NASA formal methods symposium*, pages 78–92. Springer, 2013.

A Survey Questions

Due to the limit of space, a summary of the survey organised by section is as follows. A copy of the complete survey is online at `http://goo.gl/forms/kOTOV38HIPTfJ98e2`.

1. General Information
   (a) What is your education status?
   (b) If you are a PhD student (promovendus or promovenda), what is your main research area?
   (c) If you are a Master's student, what is your programme?
2. Participants' knowledge about this approach
   (a) How much do you know about using reasoners of any kind to assist research of any field at all?
   - I have never heard about reasoners of any kind
   - I know reasoners but I have never heard that they can be used to assist research
   - I have some awareness about how a reasoner could help

   ⋮

   (b) What kind of reasoners have you used?
   - SAT solver (for satisfiability problems)
   - SMT solver (for satisfiability modulo theories problems)

   ⋮

   (c) What kind of research activities have you conducted that takes advantage of these reasoners?
   - prove mathematical or logical theorems
   - verify existing proofs

   ⋮

3. Participants' view about this approach
   (a) How hard do you find it to learn using reasoners for research?
   (b) Check all statements that apply to you concerning learning the use of reasoners for research
   - It takes some time to learn.
   - It requires some mathematical, logical or computer science knowledge.

   ⋮

   (c) What do you think about this approach in terms of usability, compared with traditional pen and paper approach?
   - It can make some research problems even more complicated.
   - It makes research easier and accelerates the problem solving process in general.

   ⋮

   (d) How much do you trust the proofs or other forms of results by reasoners?
   (e) Will you use reasoners more in your research in the future?

Rationalisation of Profiles of Abstract Argumentation Frameworks[1]

Stéphane Airiau[a] Elise Bonzon[b] Ulle Endriss[c] Nicolas Maudet[d] Julien Rossit[b]

[a] *LAMSADE, Univiversité Paris-Dauphine, France*
[b] *LIPADE, Univiversité Paris Descartes, France*
[c] *ILLC, University of Amsterdam, The Netherlands*
[d] *LIP6, Université Pierre et Marie Curie, France*

Abstract

Different agents may have different points of view. This can be modelled using different abstract argumentation frameworks, each consisting of a set of arguments and a binary attack-relation between them. A question arising in this context is whether the diversity of views observed in such a profile of argumentation frameworks is consistent with the assumption that every individual framework is induced by a combination of, first, some basic factual attack-relation between the arguments and, second, the personal preferences of the agent concerned. We treat this question of *rationalisability* of a profile as an algorithmic problem and identify tractable and intractable cases. This is useful for understanding what types of profiles can reasonably be expected to come up in a multiagent system.

1 Introduction

The model of abstract argumentation introduced by Dung [3] is at the root of a vast amount of work in artificial intelligence and multiagent systems. In a nutshell, this model abstracts away from the content of an argument, and thus sees argumentation frameworks as directed graphs, where the nodes are arguments and the edges are attacks between arguments—in the sense that one argument undercuts or contradicts another argument. Different semantics provide principled approaches to selecting sets of arguments that can be viewed as coherent when taken together.

Starting with the work of Coste-Marquis et al. [2], in recent years, a number of authors have addressed the problem of aggregating several argumentation frameworks, each associated with the stance taken by a different individual agent, into a single collective argumentation framework that would appropriately represent the views of the group as a whole. This is an interesting and fruitful line of research, bringing together concerns in abstract argumentation with the methodology of social choice theory, but it raises one important question: For a given profile of argumentation frameworks, is it in fact conceivable that that profile would manifest itself? Intuitively speaking, it may often seem more natural to encounter a profile with similar individual attack-relations. So, how do we explain the differences in perspective of the individual agents for a given profile?

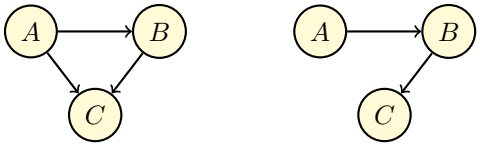
The point that the attack-relation should not be viewed as absolute and objective, but may very well depend on the individual circumstances of the agent considering the arguments in question, is central to the study of argumentation. Frameworks for modelling this phenomenon have been proposed by several authors. Here we adopt a preference-based approach, in the *value-based* variant originally due to Bench-Capon [1]. In his model, whether argument A ultimately defeats argument B does not only depend on whether A attacks B in an objective sense, but also on how we rank the importance of the social or moral values attached to A and B: If we rank the value associated with B strictly above that associated with A, we may choose to ignore any attacks of A on B.

[1] This is an extended abstract of a paper that appears in the *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2016)*.

At the technical level, we thus ask the following question: Given a profile of argumentation frameworks $(AF_1, \ldots, AF_n)$, one for each agent, can this profile be explained in terms of a single master argumentation framework, an association of arguments with values, and a profile of preference orders over values $(\succcurlyeq_1, \ldots, \succcurlyeq_n)$, one for each agent? In other words: Can the profile of argumentation frameworks observed be *rationalised*? To be able to answer this question in the affirmative, for every agent i, we require AF_i to be exactly the argumentation framework we obtain when the master argumentation framework with its associated values is reduced using the preference order $\succcurlyeq_i$. We may wish to impose any number of *constraints* on the solutions for this *rationalisability problem* we are interested in: e.g., constraints on the attack relation of the master argumentation framework, constraints on the number of values used for rationalisation, or constraints on the preference relations admitted.

2 Example

Suppose we observe two agents with the following argumentation frameworks:



Thus, they disagree on whether argument A attacks argument C (agent 1 says it does and agent 2 says it does not). As preferences can only cancel attacks rather than creating new ones, our best chance at rationalisation is to assume that the argumentation framework of agent 1 is also the master argumentation framework. Whatever values we end up associating with the arguments, if we assume that agent 1's preference relation is such that she is indifferent between all values, then we won't have to cancel any of the attacks for her, and rationalisation works correctly as far as she is concerned.

Hence, our original rationalisability problem for this example now reduces to the question of whether we can rationalise the righthand argumentation framework under the constraint that the lefthand argumentation framework is the master framework. If we associate each argument with a distinct value and if we permit incomplete preferences, then rationalisation is possible: If agent 2 prefers the value of C to the value of A and does not express a preference between any other pair, then the attack between A and C gets cancelled and all other attacks remain in place. Can we also rationalise using only *two* distinct values? No! We need two distinct values for A and C to cancel the attack between them. If B gets the same value as A, then we must also cancel the attack between B and C. And if B gets the same value as C, then we must also cancel the attack between A and B. Finally, can we also rationalise using a *complete* preference order? No! As we have seen before, agent 2 must strictly prefer the value of C to the value of A. Now, if we place the value of B anywhere above the value of A in the preference order, then we must also cancel the attack between A and B. On the other hand, if we place the value of B anywhere below the value of C, then we must also cancel the attack between B and C.

3 Results

Besides the introduction of the rationalisability problem itself, our contribution in the full paper consists in the development of algorithms to efficiently solve the rationalisability problem for a range of different constraints, and, for one choice of constraints, a complexity result showing that for those constraints the problem it is intractable. We also discuss possible applications in some detail.

References

[1] T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[2] S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex, and P. Marquis. On the merging of Dung's argumentation systems. *Artificial Intelligence*, 171(10–15):730–753, 2007.

[3] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.

A Contextualised Semantics for `owl:sameAs` [1]

Wouter Beek[a] Stefan Schlobach[a] Frank van Harmelen[a]

[a] *Dept. of Computer Science, VU University Amsterdam, NL*

1 Problem statement

Identity relations are at the foundation of the Semantic Web and the Linked Data Cloud. Standardized as part of the OWL 2 specification, the Semantic Web definition of identity follows Leibniz' principle (Definition 1). $\mathcal{I}$ is the interpretation function mapping terms to resources and EXT is the extension function mapping properties to pairs of resources. Under this semantics identical entities are indistinguishable and indistinguishable entities are identical.

Definition 1 (Semantics of `owl:sameAs`)**.**

$$\langle \mathcal{I}(a), \mathcal{I}(b) \rangle \in \mathit{EXT}(\mathcal{I}(\texttt{owl:sameAs})) \iff \mathcal{I}(a) = \mathcal{I}(b)$$

In many instances the classical interpretation of identity is too strong for practical purposes. This is particularly the case when two entities are considered the same in some but not all contexts. Unfortunately, modeling the specific contexts in which an identity relation holds is cumbersome and on the web it is impossible to anticipate all contexts in which an entity will be used. We propose an alternative semantics for `owl:sameAs` that partitions the original relation into a hierarchy of subrelations. The subrelation to which an identity statement belongs depends on the dataset in which the statement occurs. Adding future assertions may change the subrelation to which an identity statement belongs, resulting in a context-dependent and non-monotonic semantics.

2 Outline of our approach

We start with a given identity relation $\sim$ that partitions the universe of discourse U into equivalence classes. We can reinterpret $\sim$ as if it were an indiscernibility relation $\approx_\Phi$ whose set of properties Φ is implicit in the data. For each identity set we can form the set of shared properties $P+$, which we call the indiscernibility properties. Since OWL also allows entities to be asserted dissimilar, we can also form the set of dissimilarity properties $P-$. Using the indiscernibility and discernibility properties we defined the indiscernibility relation (Definition 2).

Definition 2 (Indiscernibility relation)**.**

$$x \approx_\Phi y \iff P^+(\{x, y\}) = \Phi \wedge P^-(\{x, y\}) \cap \Phi = \emptyset$$

Two pairs of resources are *semi-discernible* iff their indiscernibility properties are the same. When we look at the pairs that constitute (the extension of) a given identity relation $\sim$, all identity assertions look the same. But when we redefine identity in terms of indiscernibility and semi-discernibility, we see that within a given identity relation there are pairs that are indiscernible with respect to different properties. Stating this formally, semi-discernibility is an equivalence relation on pairs of resources that induces a partition of the Cartesian product of the universe of discourse (Definition 3).

[1] The full paper has been published in *The Semantic Web. Latest Advances and New Domains*, LNCS 9678, pp. 405-419, 2016

Definition 3 (Semi-discernibility)**.**

$$\langle x_1, y_1\rangle \equiv_\Phi \langle x_2, y_2\rangle \iff P^+(\{x_1, y_1\}) = P^+(\{x_2, y_2\}) = \Phi \\ \wedge P^-(\{x_1, y_1\}) \cap P^-(\{x_2, y_2\}) = \emptyset$$

The partitions obtained by $\equiv_\Phi$ contain but are not limited to the original identity pairs. Therefore, for sets of pairs closed under semi-discernibility we can distinguish between the following three categories:

1. All pairs in the set are identity pairs. This characterizes a consistent subrelation of the identity relation, since no semi-discernible pair is left out. (Green node in Figure 1.)
2. Only some pairs in the set are identity pairs. This characterizes a subrelation of the identity relation that is not applied consistently with respect to the semi-discernibility relation that can be observed in the data. (Red node in Figure 1.)
3. No pairs in the set are identity pairs. This characterizes a subrelation of the collection of pairs that is consistently kept out of the identity relation. (Not displayed in Figure 1.)

Now that we have determined the subrelations of identity we go on to define how these subrelations are related. Borrowing insights from Formal Concept Analysis we take U^2 as our set of FCA objects and the set of properties Φ as our set of FCA attributes. The mapping from the former to the latter is $M(\langle x, y\rangle) = \Phi(\{x, y\})$. Because the number of FCA objects is quadratic in the size of the universe of discourse it is not practical to calculate the full concept lattice. However, we are only interested in the identity subrelations and how they are related to one another. Indeed, for every pair $\langle x, y\rangle \in \sim$ we can calculate the formal concept $\langle \{\langle x, y\rangle\}'', \{\langle x, y\rangle\}'\rangle$ by using the polars $(\cdot)'$. What FCA adds to the picture is a partial order $\leq$ between the identity subrelations (Definition 4).

Definition 4 (Indiscernibility lattice)**.** *For a given identity relation $\sim$, the poset of indiscernibility subrelations is $\langle B, \leq\rangle$ with:*

- $B = \{\langle \{\langle x, y\rangle\}'', \{\langle x, y\rangle\}'\rangle \mid \langle x, y\rangle \in \sim\}$
- $\langle \{\langle x_1, y_1\rangle\}'', \{\langle x_1, y_1\rangle\}'\rangle \leq \langle \{\langle x_2, y_2\rangle\}'', \{\langle x_2, y_2\rangle\}'\rangle \iff \Phi(\{x_1, y_1\}) \subseteq \Phi(\{x_2, y_2\})$

3 Example

The approach outlined in Section 2 is implemented and tested on datasets published in the instance matching track of the Ontology Alignment Evaluation Initiative. Figure 1 shows an indicative example of an indiscernibility lattice that is calculated for such a dataset. Each rectangular box represents an indiscernibility relation. The set notation shows the indiscernibility properties Φ for each subrelation.

For each box the precision quantifies how many of the pairs that are indiscernible with respect to Φ are in the original identity relation, i.e., $|\sim \cap \approx_\Phi| \ / \ |\approx_\Phi|$. For each box the recall quantifies how much of the original identity relation is characterized by Φ, i.e., $|\sim \cap \approx_\Phi| \ / \ |\sim|$.

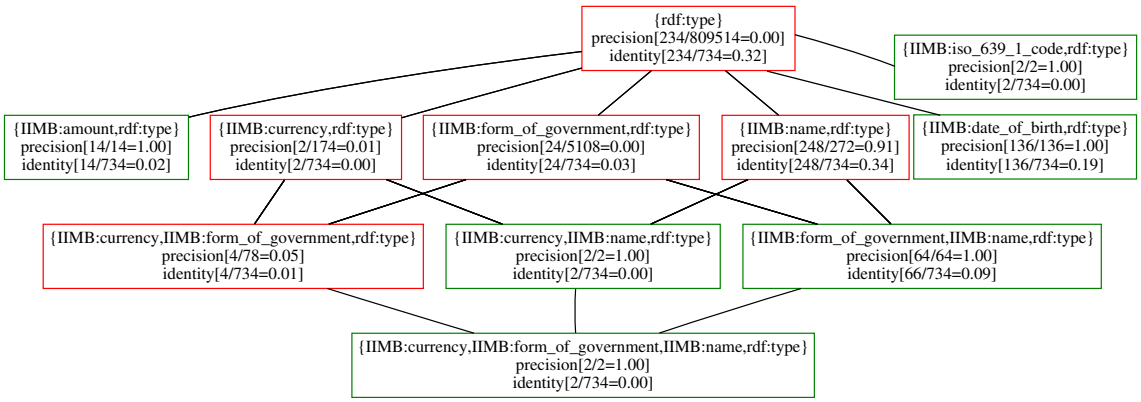


Figure 1: Example of the identity subrelations for a dataset in the instance matching track of the Ontology Alignment Evaluation Initiative. This is the 16^{th} variant of the IIMB datasets in the 2012 challenge.

A Temporal-Causal Modelling Approach to Integrated Contagion and Network Change in Social Networks (extended abstract)[1]

Romy Blankendaal, Sarah Parinussa and Jan Treur

Vrije Universiteit Amsterdam, Behavioural Informatics Group
De Boelelaan 1081, 1081HV Amsterdam, The Netherlands
Email: romy.blankendaal@gmail.com, sparinussa@gmail.com, j.treur@vu.nl

In today's world being successful and popular is mostly influenced by your social capabilities and how you interact with the people you know and work with. These social interactions are heavily investigated over the last few decades, in which analysis and prediction of the behaviour of humans in social situations play a major role. The area of Social Networks has already a longer tradition, starting in the Social Sciences over 40 years ago. More recently, it has gradually developed in other disciplines as well; see, for example [2, 4]. This development also involves computational methods to analyse and simulate networks both from the perspective of network structure and of dynamics.

Two main types of dynamics in relation to networks can be distinguished: dynamics *within* a given network structure (e.g., social contagion), and dynamics *of* a network (evolving networks). In the former case the network stays the same, but *states* (nodes) in the network change their *level* over time. In the latter case the network *connections* change, for example, their *weights* may increase or decrease. An example of this, is a network of adolescents followed over the years with the individuals' opinions about alcohol drinking as states and their friendships as connections. These states and connections both change over time. In many cases the two types of dynamics are addressed computationally as separate phenomena. This paper will address both types of dynamics and their interactions in an integrated manner.

The computational modelling approach used in this paper is the Network-Oriented Modelling approach based on temporal-causal networks described in [10, 11]. This approach is a generic, dynamic AI modelling approach based on networks of causal relations but it differs from most other causal approaches (e.g., [8]) in that it incorporates a continuous time dimension to model dynamics in an adaptive manner, both of the states and of the network itself. This temporal dimension enables causal reasoning and simulation for cyclic and adaptive causal networks, such as networks for connected mental or brain states, or for social interactions, or both. The modelling approach can incorporate ingredients that are sometimes used in specific types of (continuous time, recurrent) neural network models, and ingredients that are often used in probabilistic or possibilistic modelling. It is more generic than such methods in the sense that a much wider variety of modelling elements are provided, enabling the modelling of many types of dynamical systems, as described in [10, 11].

For the dynamics within a network an existing approach to social contagion was adopted. For the dynamics of the network, two different principles were considered and integrated, namely the *homophily* principle and the *more becomes more* principle (in a wide context also known under names such as Matthew effect and preferential attachment). The main objective of this paper was to explore how combining of these three models for social contagion and network evolution can be used to analyse and predict human behaviour in social situations. The *contagion* principle indicates that the more a person interacts with someone else, the more their opinions or beliefs or emotions or other states will converge; e.g., [3]. The *homophily* principle in a sense indicates the converse of this: the similarity of states such as opinions or beliefs or emotions of persons affects the strength of the connection between them; e.g., [6, 9]. The *more becomes more principle* describes the phenomenon that whenever someone is popular (in the sense of the strength of connections), he or she will become more popular over time because it is believed that this person is worth relating to; e.g., [1, 7].

[1] The full paper was published in *Proceedings of the 22nd European Conference on Artificial Intelligence*, *ECAI'16*, 2016.

The obtained computational model for dynamics in social networks integrating three models for the dynamics of states and relations of persons in a social network relates to and differs from existing work as follows. The first two models (for the contagion and homophily principle) were adopted as variations on existing models from [3, 9]; the paper focuses further on the third model (for the more become more principle) and the integration of these three models.

The integrated model was evaluated in three different manners: by various simulation experiments, by verification based on mathematical analysis, and by validation against an empirical data set. By all three methods a positive evaluation was found. The simulation results were as expected or at least were well explainable, the verification showed that the model provides the outcomes as predicted by the mathematical analysis, and the validation provided outcomes of the model with deviations from the empirical data within a 6% range of the [0, 1] interval. As part of the validation it was found that for the considered network of adolescents [5], according to the tuned model the dynamics of connections was determined by the homophily principle based on similarity in alcohol drinking for about 25%, and by the more becomes more principle for about 75%.

For future research variations of the model can be analysed. For example, for the homophily sub-model, as an alternative a quadratic variant as introduced in [9] can be used. Moreover, homophily with respect to multiple states can be explored, for example, not only similarities for drinking alcohol but also for smoking and sports behaviour. Also alternatives for the more becomes more and social contagion sub-models can be analysed, for example by choosing other combination functions from the collection of possible combination functions shown in [10, 11].

Finally, validation can be performed for more empirical data on social networks and their dynamics over time. However, often data sets for social networks only provide data for one point in time and only about the connections. It is not easy to find data sets that include data about connections over time and in addition also data about states of the members of the network over time. In this sense the Glasgow data set [5] used here is very valuable and has a rather unique position.

References

[1] A.L. Barabási, and R. Albert. Emergence of Scaling in Random Networks. Science, 286: 509-512, 1999.

[2] S. Boccalettia, V. Latorab, Y. Morenod, M. Chavez, and D.-U. Hwanga. Complex networks: Structure and dynamics. Physics Reports, 424: 175 – 308, 2006.

[3] T. Bosse, R. Duell, Z.A. Memon, J. Treur, and C.N. van der Wal. Agent-Based Modelling of Emotion Contagion in Groups. Cognitive Computation Journal, 7: 111-136, 2015.

[4] J. Giles. Computational Social Science: Making the Links. *Nature*, 488: 448-450, 2012.

[5] Glasgow Empirical Data https://www.stats.ox.ac.uk/~snijders/siena/Glasgow_data.htm, 2016

[6] M. McPherson, , L. Smith-Lovin, and J.M. Cook. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol*., 27 : 415–444, 2001.

[7] M.E.J. Newman. The structure and function of complex networks. *SIAM Review*, 45: 167–256, 2003.

[8] J. Pearl. *Causality*. Cambridge University Press, 2000.

[9] A. Sharpanskykh, and J. Treur. Modelling and Analysis of Social Contagion in Dynamic Networks. *Neurocomputing*, 146: 140–150, 2014.

[10] J. Treur. Dynamic Modelling Based on a Temporal/Causal Network Modelling Approach. *Biologically Inspired Cognitive Architectures*, 16: 131-168, 2016.

[11] J. Treur. *Network-Oriented Modelling: Addressing Complexity of Cognitive, Affective and Social Interactions*. Understanding Complex Systems Series, Springer Publishing, in press, 2017.

Two problems for sophistication[1]

Peter Bloem [a] Steven de Rooij [b]

[a] *Dept. of Computer Science, VU University Amsterdam, NL*
[b] *Mathematical Institute, University of Leiden*

Abstract

Kolmogorov complexity measures the amount of information in data, but does not distinguish structure from noise. Kolmogorov's definition of the *structure function* was the first attempt to measure only the structural information in data, by measuring the complexity of the smallest model that allows for optimal compression of the data. Since then, many variations of this idea have been proposed, for which we use *sophistication* as an umbrella term. We describe two fundamental problems with existing proposals, showing many of them to be unsound. Consequently, we put forward the view that the problem is fundamental: it may be impossible to objectively quantify the sophistication of a dataset

Kolmogorov complexity formalizes the idea of *information content* of data. It states that the size of the smallest description from which the data, represented as a bitstring, can be recovered is the amount of information contained in the data. The theory is invariant, up to a constant term, to the choice of description language: so long as the description language is effective (a computer program can translate from descriptions to the data they describe) and Turing-complete (equal in expressive power to a Turing machine), the Kolmogorov complexity will not differ more than a constant term between description languages. For sufficiently large data, two versions of Kolmogorov complexity will eventually agree.

While Kolmogorov complexity is a very succesful theory, it clashes with intuitive ideas of complexity: data seems most complex to us when it contains structure, like language or interesting visual information. Yet, the data that is most complex by Kolmogorov complexity is pure random noise. For instance, a sequence of bits generated by flipping a fair coin. Moreover, by flipping a *biased* coin, we can generate entirely uninteresting data that has exactly the same Kolmogorov complexity as a dataset that contains rich, deep structure (such as an encoding of a novel, or a film). It is an open question whether we can find a second measure of complexity, complementary to Kolmogorov complexity, and built on the same principles, which captures this notion of the "interestingness" of the data.

One approach is to assume that an object consists of *structural* and *stochastic* information, and that only the former determines the complexity. A dataset x is separated into structural and stochastic parts by representing it as a two-part code: the first part a computable function ϕ, the *model*, and the second part a string y, the *input*, such that $\phi(y) = x$. The aim is that for efficient two-part representations, ϕ contains all structure, and y contains all stochastic information. Thus, the simple dataset generated from a biased coin has a simple model, whereas the deeper linguistic or visual data requires more complexity in the model, even though both two-part codes have the same total length.

Many proposals, including the *structure function* [2], the *algorithmic minimal sufficient statistic* [3], the *effective complexity* [4] and *sophistication* [5, 1] fit this mould. We use the latter as a catch-all term.

We highlight two crucial issues with this general approach. First, it is essential to correctly measure the information content of the model. All treatments rely on a *numbering*; a canonical description of the computable functions. We show how to construct two effective numberings, such that the sophistication built on the first will differ from that built on the second by an exponential amount in the size of the data. This in stark contrast to the Kolmogorov complexity, where we see that switching numberings will not affect the complexity by more than a constant.

The problem of efficient numberings can be resolved, but many published treatments do not address it at all, leading to measures of sophistication that are highly dependent on the choice of numbering.

[1] The full paper has been published in *Algorithmic Learning Theory*, pages 379–394, 2015.

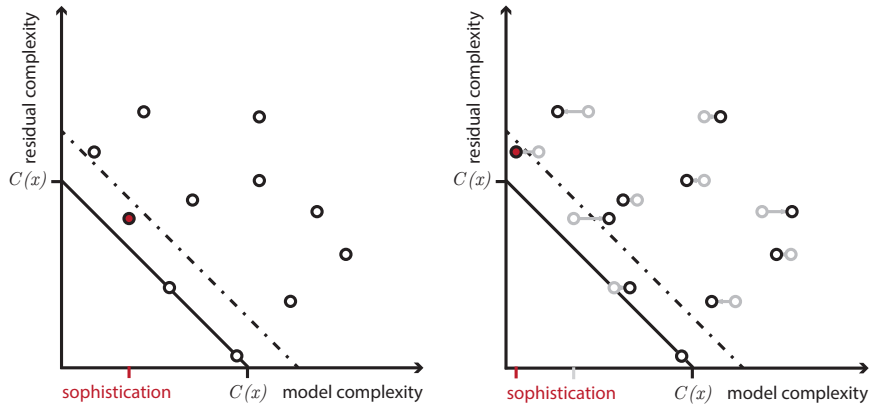


Figure 1: (left) Two-part representations of x by the two components of their code. The Kolmogorov complexity $C(x)$, appearing as a black diagonal, provides a lower bound on the total codelength. We consider only representations that are close to this optimum—called *candidates*—with the threshold represented by a dashed line. The size of the smallest model below the threshold is the sophistication of the data. (right) The same image, after a constant perturbation in the model complexity caused by a change in numbering.

Second, we discuss the problems of over- and underfitting. To give an extreme example of underfitting: we can use as a model ϕ the function computed by a universal Turing machine. A representation with this model has all information in the input, yet is maximally efficient, up to a constant. We show that for some numberings, such models are *always* preferred, leading to a sophistication that is bounded by a contant for all datasets.

Underfitting is a known issue which many treatments seek to circumvent by restricting the class of models to total functions. However, the reverse can also be achieved: we can find two-part representations with all information in the model part of the code. Again, we can constuct numberings so that the sophistication is always determined by these models, so that the sophistication is always equal to the Kolmogorov complexity, up to a constant.

Figure 1 illustrates the heart of the problem: if we change the numbering, the complexity of the model in each two-part description may change only by a constant—and consequently the Kolmogorov complexity changes by no more than a constant—but after such a perturbation, a different candidate may determine the sophistication, changing that value by much more than a constant.

Only one formalization [6] escapes these issues by allowing non-self-delimiting descriptions, limited to the total functions. The question remains open whether this form is invariant, but we show that numberings exist that cause it to be bounded for all strings but those whose shortest descriptions take vast amounts of time to unpack. Thus, it certainly does not fit the original idea behind sophistication.

We draw three conclusions. First, sophistication is an interesting problem that leads to diverse questions about statistics, computability and compression. Second, we show that all currently published approaches are either incorrect, or do not behave as desired. Finally, we propose that a sound definition of sophistication, capturing the right intuition may be impossible, and we put forward several informal reasons for this point of view.

References

[1] L.F.C. Antunes and L. Fortnow. Sophistication revisited. *Theory Comput. Syst.*, 45(1):150–161, 2009.

[2] T.M. Cover. Kolmogorov complexity, data compression, and inference. In *The Impact of Processing Techniques on Communications*, pages 23–33. Springer, 1985.

[3] P. Gács, J. Tromp, and P.M.B. Vitányi. Algorithmic statistics. *IEEE Tr. Inf. Th.*, 47(6):2443–2463, 2001.

[4] M. Gell-Mann and S. Lloyd. Information measures, effective complexity, and total information. *Complexity*, 2(1):44–52, 1996.

[5] M. Koppel and H. Atlan. An almost machine-independent theory of program-length complexity, sophistication, and induction. *Inf. Sci.*, 56(1-3):23–33, 1991.

[6] P.M.B. Vitányi. Meaningful information. *IEEE Tr. Inf. Th.*, 52(10), 2004.

An Intelligent System for Aggression De-escalation Training*

Tibor Bosse[a] Charlotte Gerritsen[b] Jeroen de Man[a]

[a] *Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam*
[b] *Netherlands Institute for the Study of Crime and Law Enforcement, P.O. Box 71304, Amsterdam*

A train conductor was assaulted at Nuneaton railway station after asking to see a man's ticket. The 48-year-old victim was working on a service from Crewe to London when he asked to see a passenger's ticket before he boarded the train in Nuneaton. The man became abusive and started to push the conductor before leaving the station." [6]

Although it is just one example, this incident illustrates the vulnerability of employees in the public sector to aggressive behaviour of people like customers, patients, travellers or other citizens. Aggressive behaviour against public service workers (e.g. police officers, ambulance personnel, public transport employees) is an ongoing concern in many countries [2, 5]. According to a national safety investigation in the Netherlands in 2011, almost 60% of the employees is confronted with unwanted behaviour on a daily basis [1]. Most incidents of aggression are of a verbal nature (e.g., insulting, swearing, intimidating), but in about 10% of the cases the conflicts escalate into physical aggression (e.g., threatening, abusing, robbing).

To better prepare them for these incidents, professionals in the public domain often receive dedicated resilience training. Such training is typically performed in a group setting based on role-play, where employees learn to communicate with aggressive clients in a de-escalating manner. Although this form of training has shown to be successful, it is quite expensive with respect to both money and time. Furthermore, the training is not always easy to control or repeat systematically.

As a complementary approach, the aim of the STRESS project [7] was to develop a simulation-based training system for aggression de-escalation. This is in line with a number of recent initiatives that show promising results regarding the possibility to train social and communicative skills based on simulated environments involving virtual humans [3, 4].

The main idea of the current system is that employees in the public domain can practice their aggression de-escalation skills by engaging in conversations with aggressive virtual characters. By designing the scenarios in such a way that the characters calm down if they are being approached correctly, but become more aggressive if they are being treated inappropriately, trainees will receive immediate feedback on their performance. Meanwhile, they are monitored by intelligent software that observes and analyses their behaviour and physiological state (e.g., heart rate, skin conductance, brain activity) and provides tailored feedback. Feedback consists of two categories, namely hints and prompts on the one hand, and run-time modifications in the scenarios on the other hand. By using such a system, employees have the ability to practice their aggression de-escalation skills in a cost-effective, personalised and systematic manner. A high-level overview of the system's architecture is depicted in Figure 1.

* The full paper has been published in: Proc. of the Prestigious Applications of Intelligent Systems conference (PAIS), IOS Press.

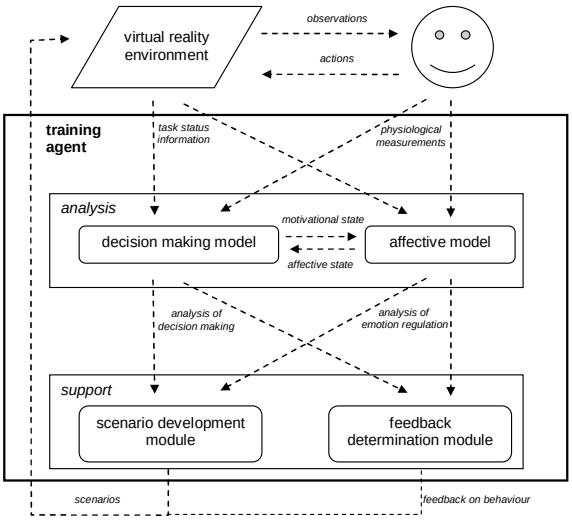


Figure 1. Global architecture of the training system

The preliminary results of evaluation studies among employees of the public transport company Amsterdam (n=24) and students of the Dutch Police Academy (n=41) indicate that with respect to user satisfaction, participants were generally positive about the content of the virtual scenarios, the mechanisms to interact with the characters, and the potential of the system as a learning tool. Nevertheless, also a number of points for improvement were identified, which mainly have to do with the emotional aspect of the system: for several participants, the perceived sense of presence was limited because they did not 'feel' the emotion in the virtual conversation partner. One interesting way to improve this situation, which we are currently considering, is to combine the scenarios with haptic feedback (e.g., by using a vibrating vest designed for video games). Based on such technology, a situation can be created in which an (aggressive) virtual character can actually 'touch' the user. Another obvious possible extension would be to use a head-mounted display instead a flat video screen. Our expectation is that such extensions will in the future lead to a more engaging and therefore more effective training tool.

In the full version of this paper, the overall architecture of the system, some details of its various components, and the preliminary results of evaluation studies with potential end users are described. Many of these elements have been published in previous papers [7], but this is the first time in which all of them are combined into our coherent description. Hence, the main contribution of the paper is in explaining how the various parts of the STRESS project come together in a practical application.

References

[1] M. Abraham, S. Flight, and W. Roorda. Agressie en geweld tegen werknemers met een publieke taak. Research for the program 'Veilige Publieke Taak 2007 - 2009 – 2011'. Amsterdam: DSP.

[2] J. DelBel. De-escalating workplace aggression. Nursing Management (USA) 34, 31-34, 2003.

[3] M. Hays, J. Campbell, M. Trimmer, J. Poore, A. Webb, C. Stark, and T. King,. "Can Role-Play with Virtual Humans Teach Interpersonal Skills?. In Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), 2012.

[4] J. Jeuring, F. Grosfeld, B. Heeren, M. Hulsbergen, R. IJntema, V. Jonker, N. Mastenbroek, M. van der Smagt, F. Wijmans, M. Wolters, and H. van Zeijts. Communicate! – A Serious Game for Communication Skills. In: G. Conole et al. (Eds.): EC-TEL 2015, Springer LNCS, vol. 9307, pp. 513-517, 2015.

[5] M.J. Smith and R.V. Clarke. Crime and public transport. In R. Clarke (Ed.), Crime and Justice: A Review of Research (pp. 169-234). Chicago, IL: The University of Chicago Press, 2000.

[6] http://www.coventrytelegraph.net/news/local-news/conductor-assaulted-platform-railway-station-10757452.

[7] http://stress.few.vu.nl/

Exploring and Comparing Machine Learning Approaches for Predicting Mood over Time[1]

Ward van Breda [a] Johnno Pastor [a] Mark Hoogendoorn [a]

Jeroen Ruwaard [b] Joost Asselbergs [b] Heleen Riper [b]

[a] *VU University Amsterdam, Department of Computer Science, De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands*

[b] *VU University Amsterdam, Department of Clinical Psychology, De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands*

1 Introduction

Mental health problems have a high impact on the lives of patients, their social surrounding, and the society in general. It obstructs patients to learn, work or participate in society. Many affected therefore turn to professional help. The costs for society have been estimated at a startling amount of 2.5 trillion dollars per year worldwide in 2015, and will rise to 6 trillion dollars per year by 2030 [2]. These problems are the driving force behind health-related research that aims to provide more effective therapies for patients and their social environment. Major developments in mobile technologies offer new possibilities for mental health interventions, such as performing Ecological Momentary Assessments (EMA) more effectively. Using new technology you can more frequently assess the mental state of the user as well as the context in which it is measured. The context can be collected unobtrusively using sensors. Such measurements provide highly detailed insights into the behavior and mental state of the patient and can be a driver for more personalized and effective therapies.

Although studies that involve EMA are increasing in number, only very few studies try to fully take advantage of the wealth of data that results. Predictive modeling on a more detailed level, e.g. predicting mood level changes in terms of hours, are rarely seen while they can be a great driver for more real-time (semi) automated forms of therapy. Most predictive modeling endeavors focus on more long term predictions such as therapeutic effectiveness or long term recovery. There are some exceptions, such as [**?**] or [1]. However, these studies do not take advantage of more recent developments in the area of machine learning that can lead to more accurate results and as a consequence show relatively poor performance.

2 Experimental Setup

Dataset The data used in this research paper originates from the VU Unobtrusive Ecological Momentary Assessment pilot study data (for more info see [1]). A total of 33 participants were selected for the pilot of which 27 contributed enough data for meaningful analysis. The data consists of 76 variables and 1249 observations, running for 52 days. The data was obtained through two applications installed on the participant's smartphone: the eMate EMA application and the iYouVU application.

Experiment using mood only For predicting mood exclusively using the mood series for each individual, we start with time series modeling using Auto Regressive Integrated Moving Average (ARIMA). The ARIMA model explores if the mood signal is a 'stationary' signal, i.e. the model tries to find statistical properties that make the mood series constant over time. For each participant the ARIMA models

[1] The full paper has been published in *Innovation in Medicine and Healthcare 2016*, pages 37–47, 2016.

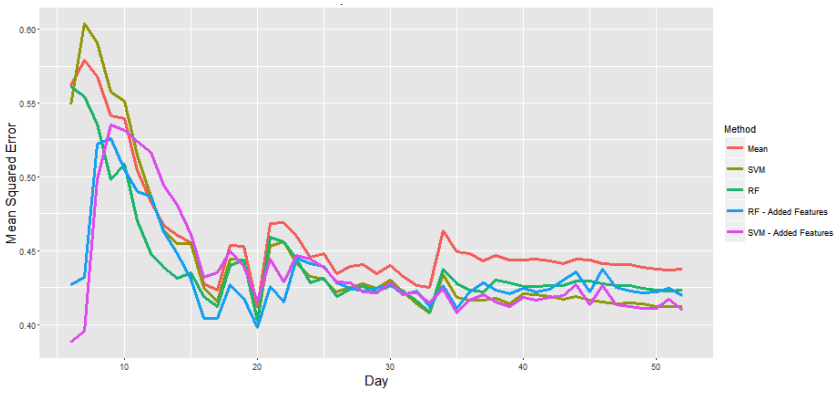


Figure 1: The cumulative performance (MSE) over time for all participants. The x-axis shows the number of days of history considered.

use the whole series as training data. This first step, i.e. to reproduce the data, is important because we want to know the potential power of the method. Thus, the time series model has maximum potential to find statistical attributes to explain the dynamics for each individual. After this, for each individual the tuned model is fed the data related to each time point in an attempt to predict the target variable (the mood value) for each next time point. The final MSE for this method is calculated by taking into account the MSE for all individuals and all time points.

Experiment using similarity measures To explore similarities in the mood series the Dynamic Time Warping (DTW) algorithm was applied. This algorithm compares two time-series for similarities, even if they differ in terms of time or speed. The similarity measure uses the whole series as training data as well, because we want to see the extent of similarities that are present between participants in the most ideal circumstance. In such a way we can better assess the potential of similarity measures for predicting mood series. However, as discussed in Section 3, we do not believe this has much potential yet. Therefore no evaluation setup is necessary for this method.

Experiment using full predictive modeling For full predictive modeling we use two techniques: Support Vector Machines (SVM) and Random Forests (RF). For each participant the SVM and RF methods are trained based on the past days that are available and predict the target variable, the mood value, for the next time point. This way, as time increases, more training cycles become available for the machine learning method to train on. The final MSE for each method is calculated by taking the MSE for all individuals for all time points.

3 Results and Discussion

Predicting mood based on past mood alone turned out to have low predictive performance. Mood generally does not seem to have intrinsic statistical properties that can explain the dynamics of mood to a large degree. Also, looking at similarities between mood series of different individuals, it was found that certain pairs of individuals do share more similarities than others, but we found the similarities too limited to be of use for prediction. Finally, the as can be seen in Figure 1, predictive models using SVM and RF that leveraged all data did result in better performance compared to the naive benchmark and the time-series models. This finding indicates that contextual data about the individual is useful and needed to increase predictive performance, but, given the attributes we had to our disposal, is not yet at the level to meaningfully employ within an E-health application.

References

[1] J Asselbergs, J Ruwaard, M Ejdys, N Schrader, M Sijbrandij, and Heleen Riper. Smartphone-based unobtrusive ecological momentary assessment of day-to-day mood: An explorative study. *J Med Internet Res*, forthcoming.

[2] Thomas R Insel, Pamela Y Collins, and Steven E Hyman. Darkness invisible: The hidden global costs of mental illness. *Foreign Aff.*, 94:127, 2015.

Using Defeasible Information to Obtain Coherence[1]

Giovanni Casini [a] Thomas Meyer [b]

[a] *Université du Luxembourg, Luxembourg*
[b] *CAIR (CSIR and University of Cape Town)South Africa*

In this paper we consider the problem of obtaining coherence in a propositional knowledge base (KB) using techniques from Belief Change. Our motivation comes from the field of formal ontologies where coherence is interpreted to mean that a concept name has to be satisfiable. In the propositional case we consider here, this translates to a propositional formula being satisfiable.

Consider a scenario in which we are presented with the following information about red blood cells: *Avian red blood cells and mammalian red blood cells are vertebrate red blood cells. Vertebrate red blood cells have a nucleus, but mammalian red blood cells don't.* In a propositional KB, this information can be represented as follows: $\{a \rightarrow v, m \rightarrow v, v \rightarrow n, m \rightarrow \neg n\}$. An obvious problem with this formulation is the consequence that mammalian red blood cells don't exist ($m \rightarrow \bot$). Even without knowing much more about the intentions of the knowledge engineers who designed the KB, we can be reasonably certain that this type of modelling is not what they had in mind. In the Description Logic (DL) community the analogous problem (i.e. when DL concept names are not satisfiable) is referred to as *incoherence* [13], and we shall adopt the same terminology for the propositional context. Our goal here is to consider a generalised version of coherence for which a designated complex concept has to be satisfiable. In the propositional case this translates to ensuring that a designated formula has to be satisfiable.

In literature, the standard approach for achieving coherence is closely related to classical belief base contraction [7]. Coherence w.r.t. the atom p is obtained via a base contraction of the formula $\neg p$, thereby ensuring the satisfiability of p. Also in this paper we investigate mechanisms for obtaining propositional coherence by applying techniques developed in the area of Belief Change, but, unlike the standard approaches, our version of belief change is applied to a logic for nomonotonic reasoning. For example, when stating that vertebrate red blood cells have a nucleus, what the knowledge engineers probably had in mind is that vertebrate red cells *usually* have a nucleus, with mammalian red blood cells being an exception. While (classical) propositional logic is not able to deal with exceptions, there are a range of nonmonotonic formalisms capable of doing so, thereby enabling us to obtain coherence.

To summarise, the problem can be stated as follows: we have a KB forcing us to conclude that a specific proposition cannot be satisfied (like m in the above scenario), and we want to make such a proposition satisfiable. In our proposal, instead of eliminating a piece of information to make m satisfiable (e.g. eliminating $v \rightarrow n$), we 'weaken' the information in the KB, allowing for the existence of exceptions. In the example above such a result could be obtained substituting the conditional $v \rightarrow n$ with a *defeasible conditional* $v \mathrel{|\!\sim} n$, interpreted as '*Usually*, vertebrate red blood cells have a nucleus'. In such a way, using the reasoning tools developed in the nonmonotonic reasoning framework, we have the possibility to obtain coherence with a minor loss of information.

We develop our proposal referring for the defeasible part of our reasoning to Lehmann and Magidor's Rational Closure (RC) [10].We choose RC as our closure operation because it is well defined from a semantical point of view and there are nice decision procedures, and, due to some properties of the main inferential closures in the preferential approach, if we solve the kind of problems we want to model for

[1]The full paper has been published in *Proceedings of the Proceedings of the 15th International Conference on Principle of Knowledge Representation and Reasoning (KR-16)*, pp. 537-540, AAAI Press, 2016.

RC, we automatically solve them also for the other main closure operations of the preferential family, e.g. [8, 10, 9, 3, 2].

In the paper we present two belief change operations, *Simple Weakening* and *Nuanced Weakening*.

Simple Weakening is an operation that can be easily implemented, also with some languages with different expressivity w.r.t. Propositional Logic (e.g., Description Logics). The limit of such an operation is that the outcome depends heavily on the syntactic form of the KB and success is not guaranteed.

Nuanced Weakening is a more refined operation. The entire operation can be reduced to a series of of classical Propositional Logic decision problems, plus a classical form of base-contraction, *kernel-contraction* [7]. Compared to Simple Weakening, the outcome of Nuanced Weakening depends less on the syntactic form of the KB and every instance is successful in restoring coherence. Also, once we reformulate for the present context the properties characterising kernel contraction in the classical context (*Success, Inclusion, Core-retainment, Uniformity*, [7]), we obtain a representation theorem proving the correspondence between the family of Nuanced Weakening operations and the class of operations satisfying such properties.

To summarise, in this paper we combine formal tools taken both form the Belief Revision and the Nonmonotonic Reasoning communities in order to define belief change operators that are appropriate for enforcing coherence in conditional KBs, minimising the loss of information. The proposed operators are easily implementable. Our proposal can be developed in various other ways. First of all, since the focus is on problems strongly connected to the development of ontologies, we plan to extend these results to Description Logics (DLs). Much of the content of this paper is already applicable to DLs, and the decision procedures for the main inference operations in the framework of preferential reasoning have already been defined for DLs [12, 1, 3, 2, 6, 4].

References

[1] K. Britz, G. Casini, T. Meyer, K. Moodley, and I. Varzinczak. Ordered interpretations and entailment for defeasible Description Logics. Tech. report, CAIR, CSIR and UKZN, ZA, 2013.

[2] G. Casini, T. Meyer, K. Moodley, and R. Nortje. Relevant closure: A new form of defeasible reasoning for description logics. In *Proceedings of JELIA 2014*, LNCS vol. 8761, pp. 92–106. Springer, 2014.

[3] G. Casini and U. Straccia. Defeasible inheritance-based Description Logics. *J. Artif. Intell. Res. (JAIR)*, 48:415–473, 2013.

[4] G. Casini and U. Straccia. Lexicographic closure for defeasible description logics. forthcoming.

[5] L. Giordano, N. Olivetti, V. Gliozzi, and G.L. Pozzato. A non-monotonic Description Logic for reasoning about typicality. *Artificial Intelligence*, 195:165–202, 2013.

[6] L. Giordano, N. Olivetti, V. Gliozzi, and G.L. Pozzato. Semantic characterization of rational closure: From propositional logic to description logics. *Artificial Intelligence*, 226:1–33, 2015.

[7] S.O. Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers, 1999.

[8] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[9] D. Lehmann. Another perspective on default reasoning. *Annals of Mathematics and Artificial Intelligence*, 15(1):61–82, 1995.

[10] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

[11] K. Lehrer. *Theory of Knowledge*. Routledge, 1990.

[12] T. Łukasiewicz. Expressive probabilistic description logics. *Artif. Intell.*, 172:852 – 883, 2008.

[13] Guilin Qi and Anthony Hunter. Measuring incoherence in description logic-based ontologies. In *The Semantic Web*, LNCS vol. 4825, pp. 381–394. Springer, 2007.

Multi-Objective Optimization for Information Retrieval[1]

Joost van Doorn[a] Daan Odijk[ac] Diederik M. Roijers[ab]
Maarten de Rijke[a]

[a] *University of Amsterdam, Informatics Institute*
[b] *Vrije Universiteit Brussel, Artificial Intelligence Laboratory*
[c] *Now at Blendle*

The primary goal of information retrieval (IR) systems is to satisfy the information need of a user. To achieve this most search engines are optimized to rank documents based on their topical relevance to the query. However other relevance criteria, such as credibility, reputability and readability can play a role depending on the information need of the user [7]. The importance of these criteria may differ, and it might not be known a priori how these criteria need to be balanced. E.g. a child might need more easily readable documents in comparison to other users.

We propose to mitigate this by viewing multiple relevance criteria as objectives and learning a set of rankers that provide different tradeoffs w.r.t. these objectives. We combine the multi-objective technique *Optimistic Linear Support (OLS)* [5] with multiple utility-based metrics in a learning-to-rank setting.

Scalarization function If we can measure a value for every relevance criterion, then V_i is the value for relevance criteria i. Following [4] we use a *scalarization function* f, that collapses the value vector to a scalar utility: $f(\mathbf{V}, \mathbf{w})$, where $\mathbf{w}$ is a vector that parameterizes f. As we cannot observe the set of all possible rankers directly, as such we aim to find a *coverage set* [4] of rankers, that contains an optimal trade-off for each possible preference (i.e., f and $\mathbf{w}$) that a user might have, see Figure 1.

We assume that f is linear: $f(\mathbf{V}, \mathbf{w}) = \mathbf{w}^T\mathbf{V}$, i.e., the utility for the user is a *weighted* sum ($\sum_i^C w_i = 1$) over relevance criteria.

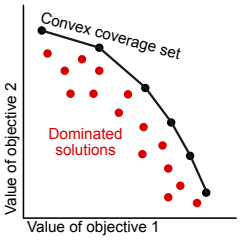


Figure 1: The points on the line represent solutions in the coverage set, the others are dominated.

Metrics as objectives To formulate our own scalarization function we can follow the standard notion of gain based utility functions [1], and adapt it to a scalarized value function:

$$V_i = \frac{1}{N_i} \sum_{k=1}^{K} gain_i(doc_k) \times discount(k) \quad (1)$$

Where k is the rank, $discount(k)$ lowers the weight of lower ranked documents, and each of the C relevance criteria has a separate gain function, which are based on expert annotations.

Convex coverage set Because each criterion contributes positively to the scalarization function, and we are interested in the *relative* importance of each criterion, we can assume that $\mathbf{w}$ is a positive vector that sums to 1 in order to determine a coverage set. A coverage set that covers all possible linear scalarizations is called a convex coverage set (CCS) [4]. To compute the CCS, we use OLS (see [5]); OLS computes a CCS by solving a multi-objective problem as a series of single-objective problems, i.e., problems that are scalarized using different $\mathbf{w}$. At each iteration, OLS tries to find a new ranker, thereby incrementally building a CCS. OLS finishes when no new rankers are found. Figure 2 illustrates

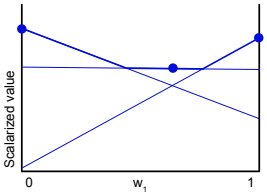


Figure 2: CCS with three rankers (dots), and their respective scalarized value functions (lines).

[1] Compressed contribution of the paper "Balancing Relevance Criteria through Multi-Objective Optimization" in *SIGIR '16*, pages 769-772. 2016. [8]

the CCS with three rankers. We can use existing single-objective optimization techniques to find rankers for a given $\mathbf{w}$. We use Dueling Bandit Gradient Descent (DBGD) [9]. As each iteration of OLS requires re-running the learning to rank algorithm, an easy way to improve convergence speed is by reusing previous solutions [6]. For each new corner weight, we multi-start DBGD, initializing the parameters based on the rankers that were found at the three closest corner weights so far. Additionally we employ Iterated Local Search [2] to avoid getting stuck in local minima, as OLS requires the single-objective optimization method to be approximately optimal [6].

In practice the ranker presented to the user is selected based on: $\max_{\mathbf{V} \in S} f(\mathbf{V}, \mathbf{w})$, where S are the values of the rankers, and $\mathbf{w}$ is determined based on the user profile, or set in the user interface.

To demonstrate how multi-objective optimization for balancing multiple relevance criteria works in practice, we perform experiments on two datasets: (i) balancing readability and topical relevance in a health setting (CLEF eHealth 2015 task 2 [3]), and (ii) balancing diversity and topical relevance in a web search dataset annotated for sub-topic relevance (TREC 2012 Web Track diversity task).

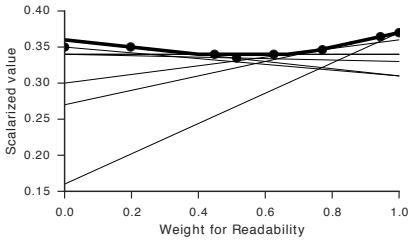


Figure 3: The CCS found on CLEF eHealth 2015.

On CLEF eHealth 2015 task 2, both topicality, and readability is evaluated using nDCG. While on TREC 2012 diversity task, we simultaneously optimize nDCG and α-nDCG. The CCSs, constructed using OLS, are shown in Figure 3 and Figure 4. Fewer solutions were found for the CCS compared to the readability task, suggesting a large correlation between the metrics nDCG and α-nDCG, from which we conclude that this setting is less suitable for our method. Whereas for readability and relevance more different alternative rankings can be presented to the user.

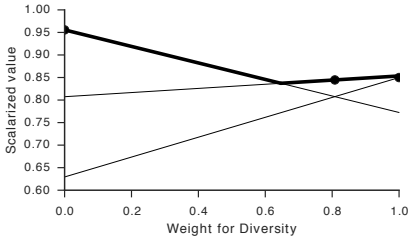


Figure 4: The CCS found on the TREC 2010 and 2011 datasets.

We demonstrated how to optimize rankings for multiple objectives. Using this approach, we have found multiple optimal rankers on the CLEF eHealth 2015 task 2 and on the TREC diversity task that offer different trade-offs w.r.t. different relevance criteria. These multiple optimal rankers are more flexible than a one-size-fits-all ranker produced by a standard learning to rank approach, and our work therefore forms an important step for flexibly optimizing search when multiple criteria are in play.

References

[1] Ben Carterette. System effectiveness, user models, and user utility: a conceptual framework for investigation. In *SIGIR'11*, pages 903–912. ACM, 2011.

[2] Helena R. Lourenço, Olivier C. Martin, and Thomas Stützle. *Iterated local search*. Springer, 2003.

[3] Joao Palotti, Guido Zuccon, Lorraine Goeuriot, Liadh Kelly, Allan Hanbury, Gareth JF Jones, Mihai Lupu, and Pavel Pecina. CLEF eHealth evaluation lab 2015, task 2: Retrieving information about medical symptoms. In *CLEF '15*. Springer, 2015.

[4] Diederik M. Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research*, 48:67–113, 2013.

[5] Diederik M. Roijers, Shimon Whiteson, and Frans A. Oliehoek. Computing convex coverage sets for faster multi-objective coordination. *Journal of Artificial Intelligence Research*, 52:399–443, 2015.

[6] Diederik M Roijers, Shimon Whiteson, and Frans A Oliehoek. Point-based planning for multi-objective POMDPs. In *IJCAI'15*, 2015.

[7] Linda Schamber and Judy Bateman. User criteria in relevance evaluation: Toward development of a measurement scale. In *ASIS'96*, volume 33, pages 218–25. ERIC, 1996.

[8] Joost van Doorn, Daan Odijk, Diederik M. Roijers, and Maarten de Rijke. Balancing relevance criteria through multi-objective optimization. SIGIR '16, pages 769–772. ACM, 2016.

[9] Yisong Yue and Thorsten Joachims. Interactively optimizing information retrieval systems as a dueling bandits problem. In *ICML'09*, pages 1201–1208. ACM, 2009.

A Conversational Agent that Reacts to Vocal Signals[1]

Daniel Formolo Tibor Bosse

Vrije Universiteit Amsterdam - Department of Computer Sciences, De Boelelaan 1081a, 1081HV Amsterdam, The Netherlands

Abstract

Conversational agents are increasingly being used for training of social skills. One of their most important benefits is their ability to provide natural interaction with humans. This work proposes to extend conversational agents' benefits for social skills training by analysing the emotion conveyed by the user's speech. For that, we developed a new system that captures emotions from the human voice and, combined with the context of a particular situation, uses this to influence the internal state of the agent and change its behaviour. An example of the system's use is shown and its limitations and advantages are discussed, together with the internal workflow of the system.

1 Introduction

Embodied Conversational Agents (ECAs) can be defined as computer-generated characters 'that demonstrate many of the same properties as humans in face-to-face conversation, including the ability to produce and respond to verbal and nonverbal communication' [1]. As research into ECAs is becoming more mature, conversations with ECAs are increasingly being perceived as natural, or at least 'believable'. As a result, there is a growing interest in the use of ECAs for the training of communicative skills, such as negotiation, conflict management or leadership skills, e.g. [2]. The main motivation is that a training system based on conversational agents provides a cost-effective method to replace (or at least complement) human actors, as it can be used anytime, anywhere.

Despite this promising prospect, developing effective conversational agents for communication training is far from easy. An important requirement for effective ECAs is their ability to react to the behaviour of the trainee in a similar manner as a human interlocutor would do. Otherwise, there is a risk that the system reinforces the wrong behaviour. Although most ECAs respond to *what* the user says, they often do not respond to *how* the user says it. This is a serious limitation, as the style of a person's speech is very important during social interactions: as discussed in [3], humans heavily rely on vocal cues (such as volume, or speed of talking) to infer other people's emotions. For example, the phrase 'sorry sir, we cannot accept 100 Euro bills' can be perceived as very friendly when it is uttered calmly and gently, but it can be perceived as offensive when it is uttered with a quick and monotone voice.

Hence, this paper proposes the use of ECAs for social skills training that adjust their behaviour based on vocal signals that are extracted from the user's speech[2]. The paper first presents a global architecture to develop such feature in the agents, followed by a discussion on how the system can be used for specific types of communication training.

[1] The full paper has been published in *The 8th International Conference on Intelligent Technologies for Interactive Entertainment*, 2016.

[2] Obviously, vocal signals are not the only aspect of behaviour that is relevant for communication training. Other aspects include facial expression, gestures, and posture, among others. However, these aspects are beyond the scope of this paper.

2 The System

The proposed system is expected to be easily integrated within different serious games or other specialised systems. Fig. 1A shows an overview of the system (i.e., the ECA) while Fig. 1B shows some outputs of the system. The system includes an interface to capture the user's speech, the off-the-shelf openSmile tool to process this speech [4], and a module to generate a response to the user. The openSmile provides some models to infer emotion, each one has its own set of emotions. The used model in this work is emodbemotion and the set of emotions is: Anger, Boredom, Disgust, Fear, Happiness, Neutral, Sadness. For each emotion is assigned a value. That information is processed by the Context Awareness Module, which deals with ambiguous outputs received by the previous module through a decision tree algorithm combined with context information provided by the ECA's beliefs.

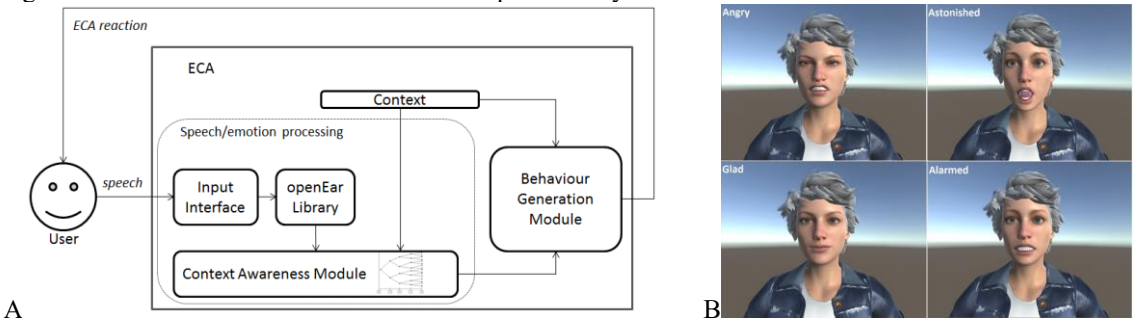


Fig. 1. (A) Flow diagram; (B) Example of recognised emotions visualised through an ECA

3 Discussion

This paper proposes the use of vocal signals that are extracted from the user's speech as one additional component to adjust ECAs' behaviour. The system can be easily plugged in into ECAs or other specialised systems that can enrich user experience. Especially for ECAs, the emotional information of a person's voice provides a new element to model their internal behaviour, which may make the interaction between ECAs and humans more natural and effective for training applications. A second innovation is the use of context information to extract emotions from human speech more accurately. Often, context conveys crucial information that is neglected by systems and serious games. Nevertheless, there are circumstances that might limit the use of the proposed system; for example, when the user's environment is noisy or has more than one person speaking at the same time, the system cannot provide precise information. Despite these limitations, the system is an important addition to the state-of-the-art of the development of ECAs. For future work, it is necessary to refine the system, analyse its accuracy in different contexts, and test it in real world applications.

Acknowledgments. This research was supported by the Brazilian scholarship program Science without Borders - CNPq {scholarship reference: 233883/2014-2}.

References

[1] J. Cassell, J. Sullivan, S. Prevost and E. Churchill. Embodied Conversational Agents, MIT Press, Cambridge, MA. 2000.

[2] T. Bosse and S. Provoost. Towards Aggression De-escalation Training with Virtual Agents: A Computational Model. In: Proceedings of the 6th International Conference on Human-Computer Interaction, HCI'14. Springer Verlag : 375-387 2014.

[3] P.N. Juslin and K.R. Scherer. Vocal expression of affect. In: J.A. Harrigan et al. (Eds.): The New Handbook of Methods in Nonverbal Behavior Research, Oxford Press, Oxford, 2005.

[4] F. Eyben, F. Weninger, F. Gross and B. Schuller. Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor, Proc. ACM Multimedia (MM), Barcelona, Spain, ACM, ISBN 978-1-4503-2404-5, pp. 835-838, October 2013. doi:10.1145/2502081.2502224 2013.

Two Aspects of Relevance in Structured Argumentation: Minimality and Paraconsistency[1]

Diana Grooters [a] Henry Prakken [b]

[a] *ORTEC Finance Rotterdam, The Netherlands*
[b] *Department of Information and Computing Sciences, Utrecht University, Faculty of Law, University of Groningen, The Netherlands*

Abstract

This paper studies two issues concerning relevance in structured argumentation in the context of the *ASPIC*$^+$ framework, arising from the combined use of strict (i.e., deductive) and defeasible inference rules. One issue arises if the strict inference rules correspond to classical logic. A longstanding problem is how the trivialising effect of the classical Ex Falso principle can be avoided while satisfying consistency and closure postulates. In this paper, this problem is solved by replacing classical logic by a paraconsistent logic as the source of strict rules and then disallowing chaining of strict rules. Another issue is minimality of arguments. If arguments can apply defeasible inference rules, then the minimality requirement of premise sets should be restricted to strict-rule applications, since defeasible rules based on more information may well make an argument stronger. Moreover, circular arguments should be excluded. The resulting *ASPIC** framework is shown to satisfy closure and consistency postulates, to have some computationally attractive properties and to be a proper extension of classical-logic argumentation with preferences and defeasible rules.

One tradition in the logical study of argumentation is to allow for arguments that combine strict (i.e., deductive) and defeasible inference rules [6, 9, 7, 5]. Strict inference rules are intended to capture deductively valid inferences, where the truth of the premises guarantee the truth of the conclusion. Defeasible inference rules are meant to capture presumptive inferences, where the premises create a presumption in favour of the conclusion, which can be refuted by evidence to the contrary. In this tradition, two issues arise concerning relevance, namely, minimality of arguments and paraconsistency in strict-rule application. We study both issues in the context of the *ASPIC*$^+$ framework [7, 5].

In deductive approaches to argumentation [1] arguments are required to have a subset-minimal set of premises. However, if arguments can apply defeasible inference rules, then this requirement is undesirable, since defeasible rules that are based on more information may well make an argument stronger. For example, *Observations done in ideal circumstances are usually correct* is stronger than *Observations are usually correct*. In this paper we therefore restrict the minimality condition to applictions of strict rules. We also exclude circular arguments since, although logically valid, they are in argumentation theory regarded as fallacious. We prove that these changes do not affect the results on [3]'s rationality postulates of consistency and strict closure.

Another relevance issue arises if the strict inference rules are chosen to correspond to classical logic. Then a longstanding unsolved problem originally identified by [6] is how the trivialising effect of the classical Ex Falso principle can be avoided when two arguments that use defeasible rules have contradictory conclusions. The problem is especially hard since any solution should arguably preserve satisfaction of the rationality postulates of consistency and strict closure [3]. In a nutshell, the problem is as follows. Suppose two arguments have contradictory conclusions q and $\neg q$. If the strict inference rules

[1] The full paper has been published in *Journal of Artificial Intelligence Research*, 56 (2016): 197–245. Diana Grooters currently works at AEGON.

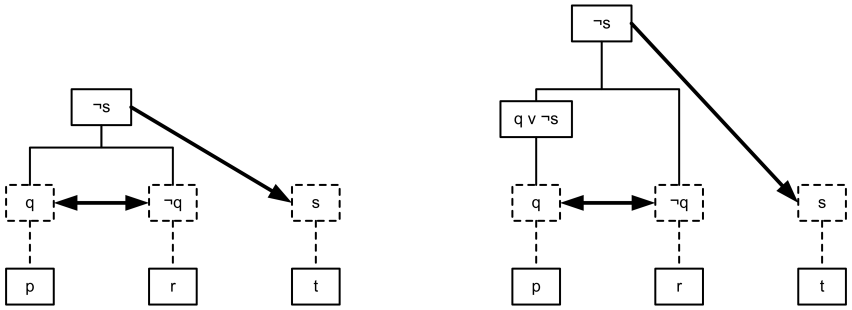


Figure 1: Illustrating trivialisation

include the Ex Falso principle that an inconsistent set implies any formula, then these two arguments can be combined into an argument for $\neg s$ for any formula s. This combined argument can potentially defeat any argument for s by applying the Ex Falso inference rule to their joint conclusions. So when there are arguments for contradictory conclusions, any other argument is potentially under threat, which is clearly undesirable, since the conflict about q is in general unrelated to s. Figure 1 (on the left) visualises this situation. Solid, respectively, dotted lines indicate strict, respectively, defeasible inferences and dotted boxes indicate rebuttable conclusions, while arrows indicate defeat relations. The figure on the right shows that restricting strict-rule application to consistent set of premises does not help, since then $\neg s$ can be derived in two consistent strict steps.

Pollock [6] thought that he had avoided such 'trivialising' arguments with a specific use of preferred semantics [4] but Caminada [2] showed that Pollock's solution does not fully avoid them. This paper solves the problem by replacing classical logic as the source for strict rules with [8]'s paraconsistent consequence notion and by disallowing chaining of strict rules (which rules out both arguments in Figure 1. This change in turn motivates a new interpretation of [3]'s strict-closure postulate and a new rationality postulate of 'logical closure'.

The combined results of this paper are then shown to be a proper extension of classical-logic argumentation with preferences and defeasible rules. Moreover, we show that if the knowledge base and set of defeasible rules are finite, then the *ASPIC** framework generates finitary argumentation frameworks in the sense of [4], which is computationally beneficial.

References

[1] Ph. Besnard and A. Hunter. *Elements of Argumentation*. MIT Press, Cambridge, MA, 2008.

[2] M. Caminada. Contamination in formal argumentation systems. In *Proceedings of the Seventeenth Belgian-Dutch Conference on Artificial Intelligence (BNAIC-05)*, Brussels, Belgium, 2005.

[3] M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171:286–310, 2007.

[4] P.M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and *n*–person games. *Artificial Intelligence*, 77:321–357, 1995.

[5] S. Modgil and H. Prakken. A general account of argumentation with preferences. *Artificial Intelligence*, 195:361–397, 2013.

[6] J.L. Pollock. *Cognitive Carpentry. A Blueprint for How to Build a Person*. MIT Press, Cambridge, MA, 1995.

[7] H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1:93–124, 2010.

[8] N. Rescher and R. Manor. On inference from inconsistent premises. *Journal of Theory and Decision*, 1:179–219, 1970.

[9] G.A.W. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence*, 90:225–279, 1997.

Combining Conflicting Environmental and Task Requirements in Evolutionary Robotics[1]

Evert Haasdijk

Vrije Universiteit Amsterdam

Abstract

The MONEE framework endows collective adaptive robotic systems with the ability to combine environment- and task-driven selection pressures: it enables distributed online algorithms for learning behaviours that ensure both survival and accomplishment of user-defined tasks. This paper explores the trade-off between these two requirements that evolution must establish when the task is detrimental to survival. To this end, we investigate experiments with populations of 100 simulated robots in a foraging task scenario where successfully collecting resources negatively impacts an individual's remaining lifetime. We find that the population remains effective at the task of collecting pucks even when the negative impact of collecting a puck is as bad as halving the remaining lifetime. A quantitative analysis of the selection pressures reveals that the task-based selection exerts a higher pressure than the environment.

1 Introduction

Imagine a collective of robots that is released in an uncharted, possibly changing, environment. The robots have to learn to operate in that environment, of which the particulars are unknown at design time. Thus, the robots have to adapt to circumstances as they find them. Of course, the robot collectives must also serve the purpose of its designers, and so must satisfy their preferences and tasks as well. Typical examples for such scenarios include monitoring, patrolling, surveying, mining or harvesting in remote, inaccessible and possibly hostile environments where human oversight is unfeasible like space, deep mining or undersea [1]. We envision such robot collectives that autonomously adapt through evolution: they evolve controllers that enable them to survive and to perform their tasks.

The environment in which robots operate indirectly circumscribes goals for the population of robots to survive and evolve, but does so without specifying objective functions. Robots must, for instance, move about to spread their genomes, or they must maintain their energy levels by regularly visiting charging stations, but these goals need not be defined explicitly: it is just that robots that display this behaviour get more opportunities to procreate. By virtue of its similarly unbounded nature, biological evolution has resulted in the high levels of adaptability and robustness that we see in natural living organisms. To exploit this creative potential in a system of evolving robots (or robot controllers), we would want to give evolution as much freedom as possible, pushing for open-ended, unbounded adaptivity, unconstrained by user-defined objective functions.

On the other hand, if the system is to be of any practical relevance, the robots must of course also perform user-defined tasks, pushing for specific, crisply defined task-related objectives.

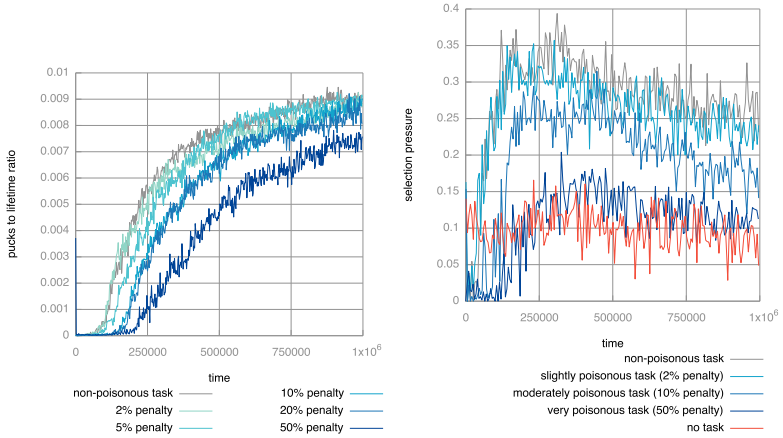
Balancing these two aspects of evolution –environment-driven adaptation and task-driven optimisation– represents a vital step towards implementing our vision of autonomous, functional, responsive and self-sufficient robot collectives. The autonomy that our vision implies prohibits centrally orchestrated evolution, so genomes and performance must be assessed, exchanged and used for selection locally, by the robots themselves.

[1]The full paper has been published in the proceedings of *IEEE 9th International Conference on Self-Adaptive and Self-Organizing Systems (SASO 2015)*, pp. 131–137, 2015.

The principal idea behind MONEE [2] is to provide this balance by concurrently employing two selection mechanisms in different roles: environmental selection for open-ended evolution and parent (or mate) selection for task-driven adaptation, both operating solely at local level. The results reported there show that the strategy of adding explicit task-based selection to an environment-driven evolutionary system yields a system where robots evolve behaviour that allows them to procreate in the environment as well as perform their tasks.

This paper investigates the effects of a scenario where the task requirements oppose those of the environment. In particular, cases where executing the task is detrimental to a robots life expectancy, e.g., because it implies a physical risk or simply drains batteries rapidly. The environment and task are essentially the same as in [2], but in this case, picking up a puck incurs a penalty and the robot's remaining lifetime is reduced by some percentage: the pucks are poisonous. This causes a complex interaction between the requirements posed by task and environment. Disregarding the task, the robots would avoid the pucks to maximise their lifetime and so have more opportunities to spread their genome. However, robots that perform the task well are more likely to be selected as parents once they have spread their genomes.

A series of simulation experiments with varying poison levels showed that the explicit task-based parent selection outweighs the implicit environmental survivor selection until the penalty of performing the task becomes very high. Figure 1(a) shows that the robots, in fact, barely reduce their puck-collecting efforts - the rate at which robots pick up pucks while they are active develops almost identically, or at least reaches identical levels, for poison levels up to 20% as it does when the pucks are not poisonous at all. Only when the poison level reaches 50%, i.e., when picking up a puck *halves* the remaining lifespan, do we see an appreciable effect. A quantitative analysis of selection pressure considers the correlation between task performance (when there is no task defined: distance covered during the individual's lifetime) and number of offspring per individual. Figure 1(b) shows that pressure that derives from the environment (red) is clearly lower than that from task-based parent selection for poison levels up to 20%. For a 50% poison level, the selection pressure from puck collection is still, albeit slightly, higher than the environmental selection pressure.



(a) The median (N=32) number of pucks collected, normalised for lifetime, per 1,000 time-steps for poison levels ranging from a 0 (grey) to 50 (dark blue) percent lifetime penalty.

(b) Quantitative analysis of selection pressure over time.

Figure 1: Experimental results

References

[1] James G Bellingham and Kanna Rajan. Robotics in remote and hostile environments. *Science*, 318(5853):1098–1102, 2007.

[2] Evert Haasdijk, Nicolas Bredeche, and A.E. Eiben. Combining environment-driven adaptation and task-driven optimisation in evolutionary robotics. *PLoS ONE*, 9(6):e98466, 2014.

[3] Christopher G. Langton. *Artificial Life: Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.

Utilizing uncoded consultation notes from electronic medical records for predictive modeling of colorectal cancer[1]

Mark Hoogendoorn[a,b] Peter Szolovits [b] Leon M.G. Moons [c]

Mattijs E. Numans [d]

[a] *VU University Amsterdam, Department of Computer Science*
De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands,
m.hoogendoorn@vu.nl
[b] *Massachusetts Institute of Technology,*
Computer Science and Articial Intelligence Lab
32 Vassar Street, Cambridge, MA, USA
psz@mit.edu
[c] *Utrecht University Medical Center,*
Department of Gastroenterology and Hepatology,
Heidelberglaan 100, 3584CX Utrecht, the Netherlands
l.m.g.moons@umcutrecht.nl
[d] *Leiden University Medical Center,*
Department of Public Health and Primary Care,
Hippocratespad 21, 2333 ZD Leiden, the Netherlands
m.e.numans@lumc.nl

1 Introduction

Electronic medical records (EMRs) are a valuable resource in the development of predictive models for diseases. The increasing level of integration of information from different caretakers into single EMR systems increases the possibilities even more. The step from an EMR to a predictive model is, however, far from trivial and requires dedicated processing techniques for a number of reasons: First of all, EMRs are typically ambiguous because different caretakers use different coding conventions. Furthermore, some information stored in the system might require background knowledge or context to be sufficiently usable in the development of predictive models (e.g., a raw lab value). Third, information stored in EMRs is of a highly temporal nature, whereas traditional predictive modeling techniques are unable to take advantage of this temporal dimension. Finally, not all EMR data is always coded; uncoded notes written by a physician are frequently seen as part of EMRs.

In previous research (cf. [5]) we have developed a pre-processing pipeline that includes components to handle the first three characteristics of EMRs, allowing for the application of off-the-shelf machine learning algorithms while benefiting from the rich content of the EMRs. However, the pipeline does not yet include a natural language processing (NLP) component which is able to distill useful information from uncoded notes. Research has shown (see e.g. [6]) that such notes can be beneficial when it comes to the development of predictive models, even when coded data are present. In this paper, we study

[1] The full paper has been published in *Artificial Intelligence in Medicine*, volume 69, pages 78–83, 2016.

three different NLP approaches and investigate their added value: (1) a simple bag-of-words approach (seen as a benchmark), (2) a topic modeling approach using both latent dirichlet allocation (LDA, cf.[3]) and hierarchical dirichlet processes (HDP, cf. [7]) where topics of text descriptions are identified in an unsupervised way using Bayesian learning, and (3) a dedicated approach which matches the text with a medical ontology (UMLS [4] and an alternative coding scheme called ICPC [1]). Although there are numerous studies aimed at extracting knowledge from medical text, hardly any has tried to compare a range of techniques. In addition, the notes we study are brief, and more keyword oriented and not full blown reports, making the application of known NLP tools for processing medical uncoded text less appropriate. We use NLTK [2] for the pre-processing of the Dutch text.

We study the performance of the different NLP techniques in the context of a large anonymized primary care dataset we have access to, covering around 90,000 patients in the region of Utrecht, the Netherlands. Specifically, we focus on predictive modeling of colorectal cancer (CRC), a disease known for its nonspecific symptoms. The dataset consists of coded data (on lab measurements, diagnoses during consultations, medications, and referrals) and includes uncoded doctor's notes associated with each consultation/visit of a patient.

2 Results

The results show that all techniques are able to distill interesting attributes that have predictive value, but when it comes to a true benefit compared to coded data, only the more knowledge-driven approaches make a real difference. The UMLS-based approach was even shown to make a substantial improvement in predictive performance on top of the coded data when simple counts of occurrences of attributes are used. An important remark in this context is that the combination of age and gender is already highly predictive for CRC, making the benefits of using additional predictors relatively modest. We move from an Area Under the ROC Curve (AUC) of 0.831 using age and gender alone to 0.890 with coded data and 0.900 with the coded data and consultation notes.

While we are satisfied with the results obtained so far, there are ample avenues for future work. First, we want to study other diseases and see how well the results obtained for this setting generalize across diseases. Secondly, we want to study datasets originating from other GP information systems to explore whether our results generalize across the different systems. We foresee that the resulting predictive model can assist the general practitioner (GP) in identifying patients with an elevated risk. We plan to provide the GP with the risk score determined by the model for patients that present relevant complaints. We will evaluate the benefits of showing the risk score in a clinical setting in the near future.

References

[1] B.G. Bentsen. International classification of primary care. *Scandinavian journal of primary health care*, 4(1):43–50, 1986.

[2] S. Bird. Nltk: the natural language toolkit. In *Proceedings of the COLING/ACL on Interactive presentation sessions*, pages 69–72. Association for Computational Linguistics, 2006.

[3] D.M. Blei, A.Y. Ng, and M.I. Jordan. Latent Dirichlet allocation. *The Journal of Machine Learning research*, 3:993–1022, 2003.

[4] O. Bodenreider. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic acids research*, 32(suppl 1):D267–D270, 2004.

[5] R. Kop, M. Hoogendoorn, A. ten Teije, F. L. Buechner, P. Slottje, L.M.G. Moons, and M.E. Numans. *Computers in Biology and Medicine*, 76:30–38, 2016.

[6] K.P. Liao, T. Cai, V. Gainer, S. Goryachev, Q. Zeng-treitler, S. Raychaudhuri, P. Szolovits, S. Churchill, S. Murphy, I. Kohane, et al. Electronic medical records for discovery research in rheumatoid arthritis. *Arthritis care & research*, 62(8):1120–1127, 2010.

[7] Y.W. Teh, M.I. Jordan, M.J. Beal, and D.M. Blei. Hierarchical Dirichlet processes. *Journal of the american statistical association*, 101(476), 2006.

A Topic-centric Approach to Detecting New Evidences for Evidence-based Medical Guidelines (Extended Abstract)

Qing Hu, Zhisheng Huang, Annette ten Teije, and Frank van Harmelen

Department of Computer Science,VU University Amsterdam, The Netherlands
{qhu400,huang,annette,Frank.van.Harmelen}@cs.vu.nl

1 Introduction

Evidence-based Medical guidelines are developed based on the best available evidence in biomedical science and clinical practice. Such evidence-based medical guidelines should be regularly updated, so that they can optimally serve medical practice by using the latest evidence from medical research. The usual approach to detect such new evidence is to use a set of terms from a guideline recommendation and to create queries for a biomedical search engine such as PubMed, with a ranking over a selected subset of terms to search for relevant new evidence. However, the terms that appear in a guideline recommendation do not always cover all of the information we need for the search, because the contextual information (e.g. time and location, user profile, topics) is usually missing in a guideline recommendation. Enhancing the search terms with contextual information would improve the quality of the search results. *In this paper, we propose a topic-centric approach to detect new evidence for updating evidence-based medical guidelines as a context-aware method to improve the search.* Our experiments show that this topic centric approach can find the goal evidence for 12 guideline statements out of 16 in our test set, compared with only 5 guideline statements that were found by using a non-topic centric approach. The full version of this paper appears in the proceedings of 9th International Joint Conference on Biomedical Engineering Systems and Technologies (HealthInf2016), 2016.

2 Topic-centric Approach for finding New Evidences

In [1], we propose a semantic distance measure to rank terms for finding relevant evidence from a Biomedical search engine such as PubMed. Our semantic distance measure is based on the (widely shared) assumption that more frequently co-occurring terms are more semantically related.

Contextualization has been considered to be a useful approach to improve the quality of search, because the context can provide more precise information for users to make queries and to reduce the size of search results. Personalization can be considered as a special case of contextualization. The same scenario can be also applied to the topic that the search is concerned with, since this is usually also not stated explicitly.

For medical guidelines, it is quite convenient to obtain this topic information, because each guideline recommendation or conclusion is always covered in a section or a subsection with a specific title. Again we can use the semantic distance measure to rank the terms appearing in the topic. Therefore, a topic centric approach to rank the terms can be done as follows: i) Obtain the terms which appear in the title of section and subsection of a guideline conclusion. ii) Rank the topic terms by using the semantic distance measure. The first term in the ranking is considered to be the center term. iii) Add non-topic terms which appear in the guideline statement one by one, based on their semantic distance to the center term. iv) Create a search query based on the merged set of the topic terms and non-topic terms. v) Search over PubMed to find relevant evidence by using the generated queries. vi) Select the best query answer based on the heuristic function.

3 Experiments and Discussion

We have conducted several experiments for finding relevant evidence for guideline updates. We selected the Dutch breast cancer guideline (version 1.0, 2004) [2] and the Dutch breast cancer guideline (version 2.0, 2012) [3] as test data. For our experiments we have selected 16 conclusions which appear in both versions of the guidelines. Thus, the evidence items appearing in the second version of the guideline can serve as a gold standard to test the proposed approach in this paper. Namely, we want to know whether or not finding relevant evidence for the first version of the guideline can really find the target evidence (alternatively called *goal evidence items*) which was used on the second version of the guideline. For the non-topic-centric approach, one of our experiments in [1] is using the Central Distance Ranking. We compare the results of the topic centric approach with the results of the non-topic centric approach to see whether or not it can get a better result, namely, finding goal evidence items for more guideline statements.

One of the experiments is to use the topic centric approach to find relevant evidence for the sixteen selected guideline conclusions and use the same heuristic function to guide the search with the same weights on the three criteria. The experiments show that the topic centric approach can find goal evidence items for 12 guideline statements out of 16, while the non-topic centric approach can find goal evidence items for only 5 guideline statements. Across the entire corpus of guideline items, the percentage of found goal evidences doubles from 18% to 41%.

References

1. Q. Hu, Z. Huang, A. den Teije, and F. van Harmelen. Detecting new evidence for evidence-based guidelines using a semantic distance method. In *Proceedings of the 15th Conference on Artificial Intelligence in Medicine(AIME 2015)*, 2015.
2. NABON. Guideline for the treatment of breast carcinoma 2004. Technical report, Nationaal Borstkanker Overleg Nederland (NABON), 2004.
3. NABON. Breast cancer, dutch guideline, version 2.0. Technical report, Integraal kankercentrum Netherland, Nationaal Borstkanker Overleg Nederland, 2012.

LOTUS: Adaptive Text Search for Big Linked Data[1]

Filip Ilievski Wouter Beek Marieke van Erp Laurens Rietveld
Stefan Schlobach

The Network Institute, VU University Amsterdam, Amsterdam, The Netherlands

1 Problem statement

A wealth of information is potentially available from Linked Open Data (LOD) sources such as those found in the LOD Cloud[2] or LOD Laundromat [1]. However, it is surprisingly difficult to find things on the Semantic Web today. With no centralized query service and limited support for natural language access, a resource is typically 'found' by memorizing its resource-denoting IRI.

The Semantic Web currently relies on four main strategies to find relevant resources: datadumps, IRI dereferencing, Linked Data Fragments, and SPARQL, but none of these is particularly suited to global text-based search. On the other hand, previously developed text search tools over the LOD Cloud, such as Sindice [2], have operated on much smaller scale and are not available today. Additionally, none of these text-search tools have allowed its users to customize the retrieval process.

The lack of a global entry point to resources through a flexible text index is a serious obstacle for Linked Data consumption.

2 LOTUS

We present LOTUS: Linked Open Text UnleaShed, a text-based entry point to a large subset of today's LOD Cloud found in LOD Laundromat. LOD Laundromat is a Semantic Web cleaning architecture and centralized data collection, that spans hundreds of thousands of data documents and tens of billions of ground statements. LOTUS allows RDF statements from the LOD Laundromat collection to be findable through approximate string matching on natural language literals.

In order to support the approximate matching and linguistic access to the LOD Laundromat data through literals, LOTUS uses an inverted ElasticSearch index. As indexing of big data in the range of billions of RDF statements is expensive, the inverted index of LOTUS is created offline. This also allows the approximation model to be efficiently enriched with various precomputed retrieval metrics.

The LOTUS system architecture is depicted in Figure 1. It consists of two main components: the Index Builder (IB) and the Public Interface (PI). The role of the IB is to index strings from LOD Laundromat in ElasticSearch; the role of the PI is to expose the indexed data to users for querying.

3 Customizability

Because 'finding the right resource' is often a domain-dependent task, LOTUS provides an adaptive framework in which researchers and developers are able to combine multiple matching and ranking algorithms. The text search in LOTUS works in two steps: a subset of the results are first *matched* and then *ranked* according to some relevance metric. LOTUS currently includes four different matching

[1] The full paper has been published in *Proceedings of the International Semantic Web Conference*, pages 470–485, 2016.
[2] `http://lod-cloud.net/`

algorithms and eight ranking algorithms, resulting in 32 retrieval combinations in total, which leverage both textual features and relational information from the RDF graph. Some matchers focus on fuzzy matching to cover typo's while others perform conjunctive or phrase-based matching. Some rankers allow concepts to be prioritized over instances, while others prioritize IRIs that occur in multiple documents. It is also possible to match literals in a specific language, or filter results by subject/predicate.

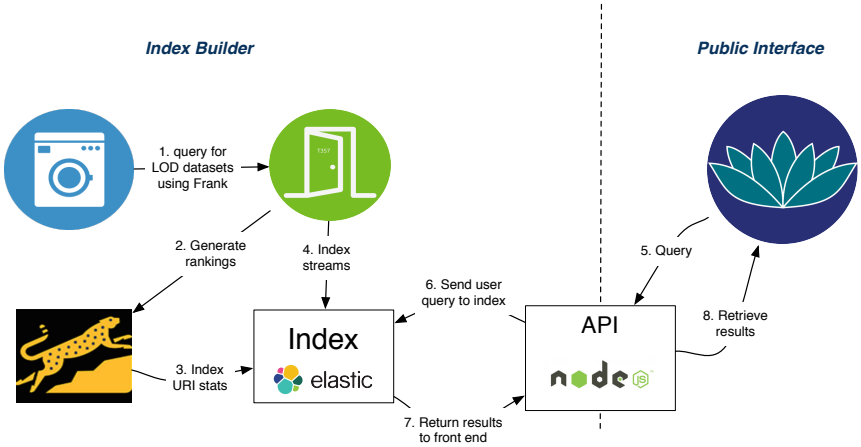


Figure 1: System architecture of LOTUS

4 Scalability and Accessibility

The index of LOTUS contains 4.33 billion entries, allowing over two billion distinct LOD Laundromat URIs to be found. LOTUS is implemented on a 5-server ElasticSearch/Lucene cluster. Each entry in LOTUS is replicated to enable high availability. Because the indexes of LOTUS are very large, they have to be created with Big Data tools, specifically Hadoop[3] and RocksDB.[4] Cached versions of the data are stored in Header Dictionary Triples (HDT)[5] and are exposed through a Linked Data Fragments (LDF)[6] API. Metadata about the source documents is stored in a Virtuoso triple store and exposed through a SPARQL endpoint. Because all components of the LOTUS framework are exposed using standards-compliant web APIs, it is easy for developers to extend LOTUS' functionality.

We show high scalability and availability of LOTUS on 18k queries. We observe that most of the queries are answered within a second. Analysis of the obtained results shows that the number and the order of the results can vary notably based on the choice of matching and ranking algorithms.

LOTUS is accessible both through human-friendly interface[7] and a machine-readable API. A comprehensive list of query parameters[8] and the code of the LOTUS API and indexer[9] can be found online.

References

[1] Wouter Beek, Laurens Rietveld, Hamid R Bazoobandi, Jan Wielemaker, and Stefan Schlobach. Lod laundromat: a uniform way of publishing other peoples dirty data. In *ISWC 2014*, pages 213–228. 2014.

[2] Giovanni Tummarello, Renaud Delbru, and Eyal Oren. Sindice.com: Weaving the open linked data. In *Proceedings ISWC'07/ASWC'07*, pages 552–565, 2007.

[3] `http://hadoop.apache.org`
[4] `http://rocksdb.org`
[5] `http://www.rdfhdt.org/`
[6] `http://linkeddatafragments.org/`
[7] Try it at `http://lodsearch.org`
[8] `http://lodsearch.org/docs`
[9] See `https://github.com/filipdbrsk/LOTUS_Search` and `https://github.com/filipdbrsk/LOTUS_Indexer`

Adaptive importance sampling for control and inference

H.J. Kappen and H.C. Ruiz

June 27, 2016

Stochastic optimal control theory (SOC) considers the problem to compute an optimal sequence of actions to attain a future goal. The optimal control is usually computed from the Bellman equation, which is a partial differential equation. Solving the equation for high dimensional systems is difficult in general, except for special cases, most notably the case of linear dynamics and quadratic control cost or the noiseless deterministic case. Therefore, despite its elegance and generality, SOC has not been used much in practice.

In [Fleming and Mitter, 1982] it was observed that posterior inference in a certain class of diffusion processes can be mapped onto a stochastic optimal control problem. These so-called Path integral (PI) control problems [Kappen, 2005] represent a restricted class of non-linear control problems with arbitrary dynamics and state cost, but with a linear dependence of the control on the dynamics and quadratic control cost. For this class of control problems, the Bellman equation can be transformed into a linear partial differential equation. The solution for both the optimal control and the optimal cost-to-go can be expressed in closed form as a Feyman-Kac path integral. The path integral involves an expectation value with respect to a dynamical system. As a result, the optimal control can be estimated using Monte Carlo sampling. See [Todorov, 2009, Kappen, 2011, Kappen et al., 2012, Theodorou and Todorov, 2012] for earlier reviews and references.

In this contribution we review path integral control theory in the finite horizon case. Important questions are: how to compute and represent the optimal control solution. In order to efficiently compute, or approximate, the optimal control solution we discuss the notion of importance sampling and the relation to the Girsanov change of measure theory. As a result, the path integrals can be estimated using (suboptimal) controls. Different importance samplers all yield the same asymptotic result, but differ in their efficiency. We show an intimate relation between optimal importance sampling and optimal control: we prove a Lemma that shows that the optimal control solution is the optimal sampler, and better samplers (in terms of effective sample size) are better controllers (in terms of control cost) [Thijssen and Kappen, 2015]. This allows us to iteratively improve the importance sampling, thus increasing the efficiency of the sampling.

In addition to the computational problem, another key problem is the fact that the optimal control solution is in general a state- and time-dependent function $u(x,t)$ with u the control, x the state and t the time. The state dependence is referred to as a feed-back controller, which means that the execution of the control at time t requires knowledge of the current state x of the system. It is often impossible to compute the optimal control for all states because this function is an infinite dimensional object, which we call the representation problem . Within the robotics and control community, there are several approaches to deal with this problem.

This paper is organised as follows. We present a review of the main ingredients of the path integral control method. We define the path integral control problem and state the basic Theorem of its solution in terms of a path integral. We then prove the Theorem by showing that the Bellman equation can be linearised by a log transform and that the solution of this equation is given in terms of a Feyman-Kac path integral. We discuss how to efficiently estimate the path integral using the idea of importance sampling. We show that the optimal importance sampler coincides with the optimal control.

Thus, a good control solution can be used to accelerate the computation of a better control solution. Such a solution is a state-feedback controller, i.e.. a function of t, x for a larger range of t, x values. This leads to the issue how to compute and represent

such a solution. The path integral Theorem shows how to compute the solution u(t, x) for a given t, x, but repeating this computation for all t, x is clearly infeasible.

A solution to this problem was first proposed in [Zhang et al., 2014] to use the cross entropy method to improve importance sampling for diffusion processes. Their approach follows quite closely the original cross entropy method by [De Boer et al., 2005]. In particular, they restrict themselves to a control function that is linearly parametrised so that the optimisation is a convex problem. In our work, we generalise this idea to arbitrary parametrisation, resulting in a gradient based method. We review the cross entropy method, as an adaptive procedure to compute an optimised importance sampler in a parametrised family of distributions over trajectories. In order to apply the cross entropy method in our context, we reformulate the path integral control problem in terms of a KL divergence minimisation and we apply this procedure to obtain optimal samplers/controllers to estimate the path integrals. We refer to this method as the path integral cross entropy (PICE) method. We illustrate the PICE method to learn a time-independent state-dependent controller for some simple control tasks involving a linear and a non-linear parametrisation.

We consider the reverse connection between control and sampling: We consider the problem to compute the posterior distribution of a latent state model that we wish to approximate using Monte Carlo sampling, and to use optimal controls to accelerate this sampling problem. In neuroscience, such problems arise, e.g. to estimate network connectivity from data or decoding of neural recordings. The common approach is to formulate a maximum likelihood problem that is optimised using the EM method. The E-step is a Bayesian inference problem over hidden states and is shown to be equivalent to a path integral control problem. We illustrate this for a small toy neural network where we estimate the neural activity from noisy observations.

References

[De Boer et al., 2005] De Boer, P.-T., Kroese, D. P., Mannor, S., and Rubinstein, R. Y. (2005). A tutorial on the cross-entropy method. Annals of operations research , 134(1):19–67.

[Fleming and Mitter, 1982] Fleming, W. H. and Mitter, S. K. (1982). Optimal control and nonlinear filtering for nondegenerate diffusion processes. Stochastics: An International Journal of Probability and Stochastic Processes , 8(1):63–77.

[Kappen, 2005] Kappen, H. (2005). Linear theory for control of non-linear stochastic systems. Physical Review letters, 95:200201.

[Kappen, 2011] Kappen, H. (2011). Optimal control theory and the linear Bellman equation. In Barber, D., Cemgil, T., and Chiappa, S., editors, Inference and Learning in Dynamic Models, pages 363–387. Cambridge University press.

[Kappen et al., 2012] Kappen, H. J., Gómez, V., and Opper, M. (2012). Optimal control as a graphical model inference problem. Machine learning , 87(2):159–182.

[Theodorou and Todorov, 2012] Theodorou, E. and Todorov, E. (2012). Relative entropy and free energy dualities: connections to path integral and kl control. In Decision and Control (CDC), 2012 IEEE 51st Annual Conference on, pages 1466–1473.

[Thijssen and Kappen, 2015] Thijssen, S. and Kappen, H. J. (2015). Path integral control and state-dependent feedback. Phys. Rev. E, 91:032104. http://arxiv.org/abs/1406.4026 .

[Todorov, 2009] Todorov, E. (2009). Efficient computation of optimal actions. Proceedings of the National Academy of Sciences, 106:11478–11483.

[Zhang et al., 2014] Zhang, W., Wang, H., Hartmann, C., Weber, M., and Schutte, C. (2014). Applications of the cross-entropy method to importance sampling and optimal control of diffusions. SIAM Journal on Scientific Computing 36(6):A2654–A2672.

A Dynamic Logic of Norm Change

Max Knobbout [a] Mehdi Dastani [a] John-Jules Meyer [a]

[a] *Utrecht University*

Abstract

Norms are effective and flexible means to control and regulate the behaviour of autonomous systems. Adding norms to a system changes its specification which may in turn ensure desirable system properties. As of yet, there is no generally agreed formal methodology to represent and reason about the dynamics of norms and their impacts on system specifications. In this abstract, we report on a paper that introduces various types of norms, such as state-based or action-based norms, and develops a dynamic modal logic to characterize the dynamics of such norms in a formal way. The logic can be used to prove various properties of norm dynamics and their impacts on system specification. It is shown that the logic is sound and complete.

1 Introduction

Norms are widely proposed as an effective and flexible means to control and regulate the behaviour of autonomous systems. Generally, norms specify the standards of behaviours such as which actions or states should be achieved or avoided. Adding norms to a system changes its specification which may in turn incentivize/inhibit specific behaviours and thereby ensure some desirable system level properties. For example, consider the norm "individuals entering a train station should have valid tickets" being introduced in a train station. The addition of this norm incentivize having a valid ticket before entering the train station and ensures that the train station is not getting unnecessarily crowded. We assume that the addition of norms incentivizes/inhibits behaviours by various enforcement means such as regimentation (e.g. by placing ports at the entrance gates of the train station) or by means of sanctions (e.g. by random inspection of individuals at the train station and issuing fines for those who has no valid ticket). In this research, we ignore the issue of norm enforcement and focus on how norms update system specifications, i.e., which system states or actions are considered as good/bad after the system is updated with norms. A system to which a set of norms is added, i.e., a system that is governed by a set of norms, is referred to as a normative system [7, 8].

A lot of research has focused on deciding (or proving) correctness of a normative system. A normative system, i.e., a system with a set of norms, is correct if the objectives of the system designer are satisfied after the norms have been added to the system [8, 1, 5]. As of yet, there is no generally agreed formal methodology to represent and reason about the norm change in normative systems. Such a methodology allows us to formally investigate the dynamics of various types of norms, such as state-based or action-based norms, in normative systems and their impact on the system specification. For example, the methodology enables us to reason about the introduction of the above-mentioned norm in a train station before and after an individual has entered the train station, either with or without a valid ticket. It also enables us to reason about the impact of different order of norm change on system specifications, and the interaction between norm change and agents' behaviours.

One may identify two possible methodological approaches when dealing with norm change. On the one hand we can identify the syntactic approach [3], where norm change is considered as an operation on the underlying "code" that constitutes the system. On the other hand we can identify the semantic approach [2, 6] which aims to look at norm change as an update of the model. This work falls in the second category, but is novel because (1) instead of providing just a semantic analysis we provide a new dynamic logic to represent and reason about norm change, and (2) we provide an accompanying (sound and complete) proof system for the logic.

The view we adopt is that normative systems can be modelled by pointed labelled transition systems, which show which facts become true under execution of which actions. Moreover, we assume that updating a system with a norm modifies the system specification and thereby its behaviour. The specific problem we address in this paper is how to represent and reason about these norm updates. We introduce new types of norms that are expressive enough to model existing norm types. We propose a new dynamic norm logic with norm update operations and an accompanying proof system to reason about norm updates in normative systems. Inspired by dynamic logic [4, 9], the effect of a norm update operation is an update of the normative system. The kind of updates (norms applied to normative systems) and its effects (normative systems that are aligned with the norm) are completely novel. The contribution of this work is significant because it paves the way for the development of formal tools that can be used to prove correctness of a normative system. From a practical point of view such tools are essential to investigate the interaction between norm change and the system behaviour, in particular, whether or when the addition of some norms satisfy some desired system properties.

The reported paper introduces a formal framework to model normative systems and various types of norms. It explores norms of the 'to-be' variant, which may forbid (or permit) certain states to occur, and norms of the 'to-do' variant, which may forbid (or permit) certain actions to occur. A dynamic logic with update operation for 'to-be' norms and a dynamic logic with update operation for 'to-do' variant are devised. For the proposed dynamic logics accompanying proof systems are provided.

References

[1] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge. Normative system games. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, pages 881–888, 2007.

[2] Natasha Alechina, Mehdi Dastani, and Brian Logan. Reasoning about normative update. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI 2013)*, pages 20–26, 2013.

[3] Guido Boella, Gabriella Pigozzi, and Leendert van der Torre. Normative framework for normative system change. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, pages 169–176, 2009.

[4] Michael J. Fischer and Richard E. Ladner. Propositional modal logic of programs. In *Proceedings of the Ninth Annual ACM Symposium on Theory of Computing*, pages 286–294, 1977.

[5] Max Knobbout and Mehdi Dastani. Reasoning under compliance assumptions in normative multi-agent systems. In *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, pages 331–340, 2012.

[6] Max Knobbout, Mehdi Dastani, and John-Jules Ch. Meyer. Reasoning about dynamic normative systems. In *Logics in Artificial Intelligence - 14th European Conference (JELIA 2014)*, pages 628–636, 2014.

[7] Marek J. Sergot. Action and agency in norm-governed multi-agent systems. In *Proceedings of the Engineering Societies in the Agents World VIII (ESAW 2007)*, pages 1–54, 2007.

[8] Yoav Shoham and Moshe Tennenholtz. On the synthesis of useful social laws for artificial agent societies. In *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI 1992)*, pages 276–281, 1992.

[9] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic, volume 337 of Synthese Library Series*. Springer, 2007.

Automating Failure Detection in Cognitive Agent Programs[1]

Vincent J. Koeman Koen V. Hindriks Catholijn M. Jonker

Delft University of Technology, Mekelweg 4 2628CD Delft, The Netherlands

Debugging is notoriously difficult and extremely time consuming, but also essential for ensuring the reliability and quality of a software system. Manual testing, using, for example, a debugger for single-step execution to identify differences between observed and intended behaviour [4], however, is not an efficient failure detection method and heavily relies on the programmer to identify the failure. In order to reduce debugging effort and enable automated failure detection, we propose an *automated testing framework for cognitive agent programs*. Automated testing yields a reduction in the effort needed to detect a failure and is more effective than code inspection methods [5].

The aim of our work is to introduce and develop a test framework that supports *automated failure detection* for programs written in rule-based agent programming languages. A *failure* is an event in which a system does not perform a required function within specified limits [3]. Failures thus are manifestations of undesired behaviour. They are caused by a *fault*, an incorrect step, process, or data definition in a program [3] or mistake in a program [6]. Upon detecting a failure, a programmer needs to locate and correct the fault that causes the failure. Our focus is on automating the detection of failures and on dynamic analysis, i.e., the process of evaluating a system or component based on its behaviour during execution [3]. The test framework introduced also provides support for fault localization.

The main contribution is a novel *automated test framework* for cognitive agent programs that provides support for detecting frequently occurring failure types. We argue that program modules instead of the top-level agent program are the most natural unit for testing, and that test conditions should be associated with modules. Second, we introduce two basic temporal operators that in practice are sufficient for specifying test conditions to detect failures. Third, using this generic framework, we propose test templates for failure types that have been identified in a previously developed taxonomy in [6]. Finally, we introduce a test approach for deriving test templates given some initial functional requirements. In order to empirically evaluate and demonstrate that our framework is expressive enough to detect all failure types, we verify that we can reproduce and identify all failures found in the sample used in [6]. The test templates that we introduce can be considered as a refinement of the failure taxonomy proposed in [6]. We show that by automating testing, we are able to identify more failures.

Test Framework The automated testing framework for GOAL [1] has been embedded in the GOAL agent programming plug-in[2] for Eclipse. When provided with a test program, the framework will initialize and set up the infrastructure for running an agent system and (external) environment in which the test will be performed. A module can be associated with a *pre-condition*, a *post-condition*, and *in-conditions*. The pre-condition of a module is a state condition that should hold when a module is entered. Similarly, a post-condition is a state condition that should hold when a module is exited. An in-condition is a temporal condition that specifies which behaviour is expected of a module. A temporal property or condition is a statement of the form **`never`** ψ or ψ **`leadsto`** ψ, where ψ is some condition on the cognitive state of an agent. Conditions **`never`** ψ can be used to specify safety conditions, i.e., things that never should occur. Conditions ψ **`leadsto`** ψ can be used to specify liveness conditions, i.e., things that are supposed to occur sooner or later after something else has happened. An example of

[1] The full paper has been published in the proceedings of the *2016 International Conference on Autonomous Agents & Multi-agent Systems (AAMAS '16)*, pages 1237-1246, International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[2] See `http://goalhub.github.io/eclipse` for a demonstration(video) of the testing framework implementation and instructions on how to install GOAL in Eclipse.

what a test for a GOAL agent looks like is given below; it verifies that the `stackBuilder` module of the `stackBuilderAgent` eventually moves block `b8` to some arbitrary location (within 1 second).

```
use BlocksWorld as mas.
use stackBuilder as module.
timeout = 1.

test stackBuilder with
   in { eventually done(move(b8,X)). }

stackBuilderAgent {
   do stackBuilder.
}
```

Test Templates and Approach A test template consists of one or more templates for individual test conditions. Test templates facilitate writing tests. Test templates also help increase the coverage of aspects that need testing. We introduce test templates for all aspects of an agent program, as based on the taxonomy of [6], the test templates are split into three main categories: templates for percepts with labels that start with **P**, templates for goals with labels that start with **G**, and templates for actions with labels that start with **A**. The test templates provide a useful starting point for writing tests. They facilitate a *structured approach* to testing an agent. The main steps of this approach are (i) to define success in terms of functional requirements, (ii) to test cognitive state updating, and (iii) to classify failures that concern actions and goals. We also provide guidelines for instantiating the templates for a specific application. These guidelines suggest ways, for example, for finding specific reasons for instantiating the state conditions that are required for the *G*- and *A*-templates.

Evaluation and Conclusion By analysing different agent program samples, we show that the framework is sufficient for detecting all failures in these programs. In particular, we are able to reproduce and detect all failures that were manually identified in [6] using our automated test framework. We demonstrate that our approach is not biased towards a specific sample of agent programs by applying the framework to other sample programs, and in a different environment. In all of these cases, we are able to detect all failures by means of the automated test framework, i.e., all agents eventually met all functional requirements after fixing the detected failures. Even though our results are encouraging in that fault localization was facilitated by the test framework, we found that faults related to actions that are performed but should not have been performed are difficult to locate. Tools that can explain why these actions were performed might be useful here [2]. In addition, more work is needed to identify how often a test should be repeated in order to find all failures in an agent program. Most uncertainty in the test results of an agent is caused by the dynamics of the specific environment it operates in. Therefore, techniques such as fuzzing that model agent environments are an interesting research direction

References

[1] Koen V. Hindriks. Programming Rational Agents in GOAL. In *Multi-Agent Programming: Languages, Tools and Applications*, pages 119–157. Springer US, 2009.

[2] Koen V. Hindriks. Debugging Is Explaining. In *PRIMA 2012: Principles and Practice of Multi-Agent Systems*, volume 7455 of *LNCS*, pages 31–45. Springer Berlin Heidelberg, 2012.

[3] ISO. Iso/iec/ieee 24765:2010 systems and software engineering - vocabulary. Technical report.

[4] Vincent J. Koeman and Koen V. Hindriks. Designing a Source-Level Debugger for Cognitive Agent Programs. In *PRIMA 2015: Principles and Practice of Multi-Agent Systems*, volume 9387 of *LNCS*, pages 335–350. Springer International Publishing, 2015.

[5] P. Runeson, C. Andersson, T. Thelin, A. Andrews, and T. Berling. What do we know about defect detection methods? *Software, IEEE*, 23(3):82–90, May 2006.

[6] Michael Winikoff. Novice Programmers' Faults & Failures in GOAL Programs. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, pages 301–308, Richland, SC, 2014. IFAAMS.

A Data Driven Similarity Measure and Example Mapping Function for General, Unlabelled Data Sets [1]

Damien Lejeune Kurt Driessens [a]

[a] *Maastricht University, email: kurt.driessens@maastrichtuniversity.nl*

While raw data is abundant, the difficulty with using this data is the lack of any structured way of representing the data. This leads to many different, human chosen, but unmatching representations of similar or related, but most of the time not identical[2] data. Examples of this are numerous. They range from data stemming from medical questionnaires, where almost never the same questions are asked, but the topics are often similar, over gene transcription data where old style microarray data and more recent RNAseq measurements exist over a pool of intersecting but not identical gene sets, to control oriented data where samples of system behaviour of a number of control problems exist, but almost never match in the chosen representation. Existing work on transfer learning and in domain adaptation for supervised tasks [5, 3] mainly focuses on the shift of the probability distributions observed between different tasks and how to correct for those, but not on the issue of different representations. When they did, defining the relation between the old task and the new, e.g. feature mapping, goal mapping, translating the model or policy to match the representations was handled by a human expert [8].

The contribution of this work is twofold, as we introduce: (i) a **data driven difference measure** for comparing data sets and (ii) an **automatically derived inter-task mapping** that can be used to compare any two data sets and thus learning tasks, whether they are supervised, unsupervised or reinforcement learning. The approach takes the shape of a pipeline employing well studied and tested techniques: (i) deep belief networks [4], (ii) manifold alignment [10] without correspondence and (iii) t-SNE [9]. The pipeline (see Figure 1) relies on deep belief nets to generate expressive features for both data repositories and on unsupervised manifold alignment to find the best mapping between these features. Applying one deep belief net to generate hidden feature activations, a forward and backwards projection into the alignment space and the other deep belief net to reconstruct the visible node activations from the projected hidden node activations, it becomes possible to translate learning examples from a source task into examples for the target task. By comparing the embedding of the original and the translated examples in a low dimensional projection built using the t-Distributed Stochastic Neighbour Embedding, we define a difference measure based on the Kullback-Leibler divergence between the two example distributions. Experiments show that the pipeline is able to autonomously find meaningful analogies between data-sets that match human intuition.

Since there is no ground truth when comparing different datasets with respect to the inter domain mapping and the computed similarity measure, we selected domains that can be matched using human intuition, i.e. the MNIST dataset of written characters [6] and the pen-digits dataset [1], as well as domains from reinforcement learning where related work generated a base for comparison. While the domain of the MNIST and pen-digits datasets is hand-written digits, the representation of the data is very different: first of all, the pen-digit samples have only 16 dimensions and represents coordinates over time, while MNIST uses 784 pixels intensities. While in this case, there seems to be a ground truth for the similarity measure to discover, without any background information about the two data-sets, they are very difficult, if not impossible for humans to match. Most of the time, the pen-digits and their corresponding MNIST digit turn out to be the most similar by a huge margin when compared to others.

[1] The full paper has been published in Proceedings of the European Conference on Artificial Intelligence (ECAI), 2016.

[2] With identical, we refer to the representation and domain of the data, not the examples measured.

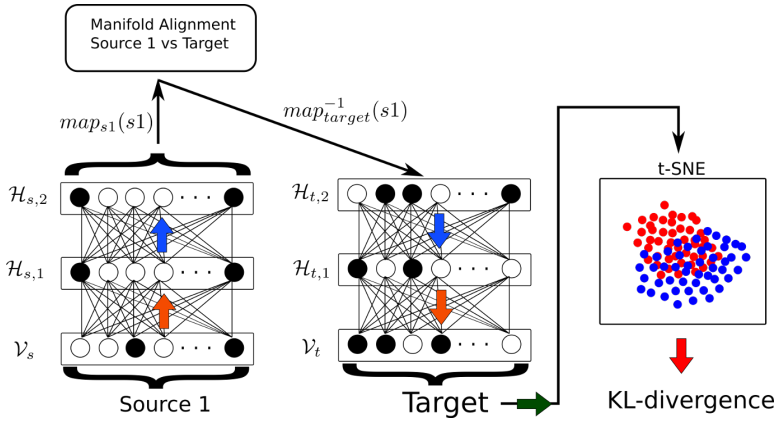


Figure 1: Pipeline using DBNs, manifold alignment and t-SNE.

In future work we would like to approach even more challenging tasks such as measuring the similarity of phonemes in speech and test the robustness by using, for example, affNIST (i.e. MNIST with affine transformations). We would also like expand the technique to e.g. bigger images that require convolutional networks. This leads to additional complexity caused by the challenge in reversing convolution and pooling layers. However, related work exists for reversing these networks and using them to reconstruct data [7, 2] that could be useful to extend the pipeline presented.

References

[1] F. Alimoglu and E. Alpaydin. Combining multiple representations and classifiers for pen-based handwritten digit recognition. In Document Analysis and Recognition, 1997., Proceedings of the Fourth International Conference on, volume 2, pages 637–640 vol.2, Aug 1997.

[2] A. Dosovitskiy, T.J. Springenberg, and T. Brox. Learning to generate chairs with convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1538–1546, 2015.

[3] Boqing Gong, Kristen Grauman, and Fei Sha. Learning kernels for unsupervised domain adaptation with applications to visual object recognition. International Journal of Computer Vision, 109(1-2):3–27, 2014.

[4] Geoffrey Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. Neural computation, 18(7):1527–1554, 2006.

[5] Judy Hoffman, Erik Rodner, Jeff Donahue, Brian Kulis, and Kate Saenko. Asymmetric and category invariant feature transformations for domain adaptation. International Journal of Computer Vision, 109(1-2):28–41, 2014.

[6] Yann Lecun and Corinna Cortes. The MNIST database of handwritten digits.

[7] A. Mahendran and A. Vedaldi. Understanding deep image representations by inverting them. In Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, pages 5188–5196, June 2015.

[8] Matthew E. Taylor, Shimon Whiteson, and Peter Stone. Transfer via inter-task mappings in policy search reinforcement learning. In Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), pages 156–163, May 2007.

[9] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. Journal of Machine Learning Research, 9(2579-2605):85, 2008.

[10] Chang Wang and Sridhar Mahadevan. Manifold alignment without correspondence. In IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009, pages 1273–1278, 2009.

Energy consumption profiling using Gaussian Processes[1]

Christiaan Leysen [a] Mathias Verbeke [b] Pierre Dagnely [b]
Wannes Meert [a]

[a] *Dept. Computer Science, KU Leuven, Belgium*
[b] *Data Innovation Team, Sirris, Belgium*

1 Introduction

Accurate prediction of household energy consumption is of great value to energy companies as it is crucial for estimating the aggregated energy that needs to be bought by these companies. Often, however, the variables impacting the electricity consumption behaviour of a household (e.g., family size, presence of airconditioning) are not known or it is unclear what variables should be investigated and measured. It is in this context that we propose an algorithm for clustering household energy consumption data. These clusters can be used to derive spending profiles of the households or to recommend households certain energy packages. Not only the relations between households but also the relationship between consumption periods are interesting to investigate, e.g. to detect anomalies.

Energy consumption is naturally represented as a time series. The households that end up in the same cluster do not necessarily have an identical consumption but their spending behaviour follows the same temporal patterns. A model of this behaviour can be learned by auto-regression methods to predict future values. In this work we will compare the learned patterns to group households together rather than to predict future values. These clusters could be used to recommend energy packages or to select a predictive model adapted to the cluster behaviour. Since we do not know upfront what time periods are relevant or how many patterns are superimposed, a method that learns this model based on the available data is preferred. For this reason and the fact that there is no need for parameter tuning by the user, we selected Gaussian Processes (GPs) [1].

Out-of-the-box GP implementations requires a single function as input while we consider a set of functions, as we want to learn a model for multiple households together. To overcome this, we present a new approach to jointly learn a GP model and optimize the hyperparameters over a set of functions, which is the first contribution of this paper. Second, we embed this in a hierarchical clustering method for time series. Finally, we present an interactive application and show how this can be used by energy providers to analyse and predict customer behaviour.

2 Method

As mentioned above, GP can't learn over multiple time series (households). To overcomes this difficulty, we learn a generalized model with the maximum likelihood over all time series, i.e. we (automatically) tune the GP parameters to have a model for which the sum of the likelihood with each time series is maximal. The parameters of the GP model are therefore extracted from the data.

This generalized model is then embedded in a clustering method. The intuition behind this method is that time series with similar behaviour will have similar RMSE (compared to the generalized model). Therefore, our method starts by learning a generalized model over all time series and then recursively

[1] The full paper has been published in *Proceedings of the 8th IEEE International Conference on Intelligent Systems*, to appear.

split the cluster of all time series based on the RMSE and the clustering parameters provided by the user (similarity threshold, minimum cluster size and split ratio).

The complexity of our approach is $O(SN^3)$, with S the number of time series and N the length of the time serie, while usual pair-wise algorithm (such as K-medoids) have a $O(S^2N^2)$ complexity. Therefore, our approach is more scalable for large number of time series (households).

3 Results and discussion

In order to test our approach, we propose two experiments. In our main experiment we cluster the energy consumption of 71 distinct households to investigate which households have similar spending habits. In a second experiment, we test our method on data from four different weeks per households (for three households) in order to find households with steady spending habits.

3.1 Households with similar consumption behaviour

We select the same week for every household and apply the selected clustering algorithms. Selecting the same week allows us to assume that these households experience similar weather conditions due to their geographical distribution. When applying GP clustering we obtain the clusters represented as a tree, where e.g. the left branch contains households with different spending habits during the weekend and in the right branch households with the same day/night habits. The result of our method was validated by a domain expert. The visualization allows to interpret distance in the tree as a proxy for the distance between two spending profiles.

As our complexity is only $O(SN^3)$, for $O(S^2N^2)$ for pair-wise methods. A complexity comparison with K-medoids DTW and agglomerative DTW clustering showed that our method was faster for more than 50 time series, because we do not need to do a pairwise comparison of all the time series. This shows the scalability of our approach.

3.2 Households with (in)consistent consumption behavior

For our second experiment we investigate the spending habits of households during the year. This is achieved by selecting multiple weeks of the same household and by checking if they end in the same cluster. Our experiment was able to label households as consistent (all weeks are clustered together), inconsistent (the weeks are clustered separately) and subject to seasonality (autumn and winter weeks are clustered together while spring and summer weeks are in the same cluster).

4 Practical use in decision support

The proposed method can support companies operating in the energy sector in tackling a number of common tasks more efficiently. The first experiment allows one to benchmark the electricity consumption per cluster, and only compare households to peers which have a similar consumption behaviour. Another use of the first experiment is to identify inefficient or problematic heating systems. Such inefficiencies can be detected by looking for households that are alone in a cluster.

The second experiment shows how to detect if the household exhibits a smooth and stable behaviour over time. For such a household, forecasting techniques to predict the consumption day-ahead are more accurate. Furthermore, it can be useful information when recommending certain energy packages. In addition, our method is scalable and it does not requires any parameter tuning by the user to create a model (but does to create the clusters). This supports its practical usefulness.

References

[1] Carl Edward Rasmussen and Christopher K. I. Williams. Gaussian processes for machine learning. 2006.

Modeling Internalizing and Externalizing Behaviour in Autism Spectrum Disorders (extended abstract)[1]

Laura M. van der Lubbe, Jan Treur, and Willeke van Vught[2]

Vrije Universiteit Amsterdam, Behavioural Informatics Group
De Boelelaan 1081, 1081HV Amsterdam, The Netherlands

Over the years, much research has been performed in the area of Autism Spectrum Disorders (ASD); e.g., [4, 10]. Most people think of people with autism as shy or socially disabled, but ASD can occur in different forms. It is sometimes difficult to find a suitable way of counseling somebody with ASD. Persons with ASD often need extra counseling, for example during high school. For such counseling to be effective it is important to have insight in the specific variant the person has, and which specific mental processes and behaviours occur. A computational model can be a basis for someone to get more understanding of such mental processes and behaviours.

One distinction that can be made within ASD is between persons who are internalizing versus those who are externalizing; e.g., [7]. The former type of persons may show anxiety whereas the latter type may show aggression. The computational model presented here enables to model both internalizing persons and externalizing persons with ASD, depending on the settings of certain parameters representing characteristics of the person. Besides this, by adopting elements of an earlier model from [13], the model also covers other characteristics that can occur in persons with ASD, such as enhanced sensory processing sensitivity (e.g., [1]), reduced mirror neuron activation (e.g., [6]), imperfect self-other distinction (e.g., [3]), and reduced emotion integration (e.g., [5], pp. 73-74).

This new model extends the earlier model from [12, 13] with different behavior types and contributes to the understanding externalizing and internalizing behavior of persons with ASD. The model was inspired by findings and theories from Cognitive and Social Neuroscience and designed as a network of mental states according to the temporal-causal network modeling approach presented in [14, 15]. The presented model specifically addresses findings and theories concerning internalizing and externalizing behaviour types within ASD. However, as elements from the model described in [12, 13] were adopted as well, the model also integrates some other aspects of ASD, addressed by theories on reduced mirror function or poor self-other distinction; e.g., [6, 3].

The model represents specific personal characteristics by specific values of parameters included in the model, such as connection weights. By using proper choices for these connection weights, the model can either simulate an internalizing person or an externalizing person. Moreover, it was shown how based on a given data set concerning a person's behaviour, by parameter estimation methods the characteristics determining the behaviour type of this person can be identified automatically. By simulation experiments and mathematical analysis the model was verified.

The computational model proposed here can be used as an ingredient to develop human-aware or socially aware computing applications (e.g. [8,9,11]) that can provide help in getting more understanding of the different behaviour types and their influence on the behaviour of a person with ASD. Such applications can be designed with knowledge of human and/or social processes as a main ingredient represented by a computational model of these processes which is embedded within the application.

---

[1] The full paper was published in *Proceedings of the 8th International Conference on Computational Collective Intelligence, ICCCI'16.*

[2] The authors made a comparable contribution to the paper

As an example, in [2] this design approach is illustrated to obtain a human-aware software agent supporting professionals in attention-demanding tasks, based on an embedded dynamical numerical model for attention. Because the computational model enables an agent to reason about the behaviour of a person with ASD, it can be used in a similar manner to design a human-aware or socially aware software agent to support persons with ASD. This might be helpful in particular for those who are counseling or supervising persons with ASD. The method shown to identify the behaviour type of a person based on empirical behavioural data can be useful, for example, in choosing a counseling approach.

References

[1] Baker, A.E.Z., Lane, A.E., Angley, M.T., and Young, R.L., The Relationship Between Sensory Processing Patterns and Behavioural Responsiveness in Autistic Disorder: A Pilot Study. *J Autism Dev Disord*, 38: 867–875 (2008)

[2] Bosse, T., Memon, Z.A., Oorburg, R., Treur, J., Umair, M., and Vos, M. de, A Software Environment for an Adaptive Human-Aware Software Agent Supporting Attention-Demanding Tasks. *Journal of AI Tools*, 20: 819-846 (2011)

[3] Brass, M., and Spengler, S., The Inhibition of Imitative Behaviour and Attribution of Mental States. In: Striano, T., Reid, V., *Social Cognition: Development, Neuroscience, and Autism*, pp. 52-66. Wiley-Blackwell (2009)

[4] Frith, U., *Autism, Explaining the Enigma*. Blackwell (2003)

[5] Grèzes, J., and Gelder, B. de, Social Perception: Understanding Other People's Intentions and Emotions through their Actions. In: Striano, T., Reid, V.: *Social Cognition: Development, Neuroscience, and Autism*, pp. 67-78. Wiley-Blackwell (2009)

[6] Iacoboni, M., *Mirroring People: the New Science of How We Connect with Others*. Farrar, Straus & Giroux (2008)

[7] Noordhof, A., Krueger, R.F., Ormel, J., Oldehinkel, A.J., and Hartman, C.A., Integrating Autism-Related Symptoms into the Dimensional Internalizing and Externalizing Model of Psychopathology. The TRAILS Study. *Abnormal Child Psychology,* 43: 577-587 (2015)

[8] Pantic, M., Pentland, A., Nijholt, A., and Huang, T.S., Human Computing and Machine Understanding of Human Behavior: A Survey, *Proc. of the Int. Conf. on Multimodal Interfaces*, 239-248 (2006)

[9] Pentland, A., Socially aware computation and communication, *IEEE Computer*, 38: 33-40 (2005)

[10] Richer, J., and Coates, S., (eds.), *Autism, The Search for Coherence*. London: Jessica Kingsley Publishers (2001)

[11] Treur, J., On Human Aspects in Ambient Intelligence. In: *Proc. of the First Int. Workshop on Human Aspects in Ambient Intelligence*. In: M. Muehlhauser et al. (eds.), *Constructing Ambient Intelligence: AmI-07 Workshops Proceedings*. Comm. in Computer and Information Science (CCIS), vol. 11, Springer Verlag, pp. 262-267 (2008)

[12] Treur, J., A Cognitive Agent Model Displaying and Regulating Different Social Response Patterns. In: Walsh, T. (ed.), *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, IJCAI'11*, pp. 1735-1742 (2011)

[13] Treur, J., Displaying and regulating different social response patterns: a computational agent model. *Cognitive Computation*, 6: 182-199 (2014)

[14] Treur, J., Dynamic Modeling Based on a Temporal-Causal Network Modeling Approach. *Biologically Inspired Cognitive Architectures*, 16: 131-168 (2016)

[15] Treur, J., *Network-Oriented Modeling: Addressing Complexity of Cognitive, Affective and Social Interactions*. Understanding Complex Systems Series. Springer Publishers, in press, 2017.

Monitoring Opportunism in Multi-Agent Systems (Extended Abstract)[1]

Jieting Luo [a] John-Jules Meyer [a] Max Knobbout [b]

[a] *Utrecht University, Utrecht, the Netherlands*
[b] *Delft University of Technology, the Netherlands*

1 Introduction

Consider a common social scenario. A seller sells a cup to a buyer and it is known by the seller beforehand that the cup is actually broken. The buyer buys the cup without knowing it is broken. Since the buyer's value gets demoted, the behavior performed by the seller is usually forbidden by the legal norm. Such a social behavior intentionally performed by the seller is first named opportunistic behavior (or opportunism) by economist Williamson [3]. It is a typical social behavior that is motivated by self-interest and takes advantage of knowledge asymmetry about the behavior to achieve own gains, regardless of the principles [1]. This definition implies that, given a social context, opportunistic behavior results in promoting own value while demoting social value. Therefore, it is prohibited by norms in most societies. Because opportunistic behavior cannot be observed directly, there has to be a monitoring mechanism that can detect the performance of opportunistic behavior in the system. This paper provides a logical framework based on the specification of actions to specify monitoring approaches for opportunism. We investigate how to evaluate agents' actions to be opportunistic with respect to different forms of norms when those actions cannot be observed directly, and study how to reduce the monitoring cost for opportunism.

2 Framework

Since monitors cannot observe the performance of opportunism directly, the action can only be identified through the information about the context where the action can be performed and the property change in the system, which is called *action specification* or *action description*. Usually an action can be specified through its precondition and its effect (postcondition): the precondition specifies the scenario where the action can be performed whereas the postcondition specifies the scenario resulting from performing the action. Therefore, we assume that every action has a set of pairs of the form $\langle \psi_p^a, \psi_e^a \rangle$, where ψ_p^a is the precondition of action a and ψ_e^a is the effect of performing action a in the context of ψ_p^a. The models that we use are transition systems, which consist of agents, states, actions, transitions between states by actions and a valuation function mapping a state to a set of propositions that hold in that state. When an action is performed in a certain state s, the system might progress to a different state in which different propositions might hold. Since we have already introduced the notion of action specification $\langle \psi_p^a, \psi_e^a \rangle$, all the possible state transitions are defined such that they go from a ψ_p^a-state to a ψ_e^a-state. We also extend the standard framework with a monitor relation $\mathcal{M}$, which represents the indistinguishability of a monitor over different states. The logical language we use in this paper is propositional logic extended with action modality for reasoning about dynamic worlds. We simply consider a norm as a subset of all the state transitions that is decided by designers of the system. In other words, if a norm is denoted as η, a state transition is an η-violation if and only if it is in the set η.

[1] The full paper has been published in *Proceedings of the Coordination, Organizations, Institutions and Norms workshop@ECAI2016*, to appear.

3 Monitoring Opportunism

In our previous paper [1], we emphasized opportunistic behavior is performed by intent rather than by accident. However, monitors cannot read agents' mental states, so for monitoring we assume that agents violate the norms always by intention from a pragmatic perspective. For example, we always assume that speeding is performed with intention. In this paper we remove all the references to the mental states from the formal definition of opportunism in our previous paper [1], assuming that the system can tell agents' value promotion/demotion causing by an action. In a sentence, from the perspective of the system, opportunistic behavior performed by an agent in a social context can be simply defined as a behavior that causes norm violations and promotes his own value.

In this paper, a monitor is considered as an external observer to evaluate a state transition with respect to a given norm. We first define a state monitor $m_{state}(\varphi)$, which can evaluate the validity of a given property in a given state: $m_{state}(\varphi)$ holds in state s if and only if for all s' $s\mathcal{M}s'$ implies φ in state s'. We can combine state monitors to check how things change in a given state transition and evaluate it with respect to a given set of norms. Based on our definition of opportunism above, we design monitoring opportunism approaches m_{opp} with respect to norm $\eta(\varphi, \psi)$, $\eta(\varphi, a)$ and $\eta(\varphi, a, \psi)$. In order to check whether action a', which has already occurred, was opportunistic behavior (violates norm η and promotes own value), we verify if action a' is performed in a φ-state. Besides, when a norm η is specified with an action a (such as $\eta(\varphi, a)$), we also need to check if action a' was indeed action a. Since the monitors cannot observe the performance of action a' directly, we only can identify action a' to be possibly action a with pair $\langle \psi_p^a, \psi_e^a \rangle$ by checking if formulas ψ_p^a and ψ_e^a are successively satisfied in the state transition by action a'. We then investigate in which case or with what requirement action a' that is detected to be opportunistic was indeed opportunistic behavior a. All these issues are elaborated and discussed in the full paper [2].

4 Monitoring Cost for Opportunism

For designing a monitoring mechanism, we not only think about whether it can perfectly detect agents' activities, but also consider if it is possible to decrease the cost involved in the monitoring process. Intuitively, given a state property denoted by a propositional formula φ, function $c(\varphi)$ returns a positive real number representing the cost that it takes to verify φ. Given a set of propositional formulas X, the cost of monitoring X is the sum of the cost of verifying each element in X. However, if it holds for $\varphi, \varphi' \in X$ that $\varphi \neq \varphi'$, and $\varphi \rightarrow \varphi'$, then monitoring $X \backslash \{\varphi'\}$ is actually the same as monitoring X. For this issue, we define the notion of *Largest Non-inferential Subset* $X_{\mathcal{I}}$ such that for all $\varphi \in X_{\mathcal{I}}$ there is no $\varphi' \in X_{\mathcal{I}}$ with $\varphi \neq \varphi'$ such that $\varphi \rightarrow \varphi'$. Given a set of propositional formulas we want to verify, we always get its largest non-inferential subset before checking anything in order to reduce the monitoring cost. Besides, given the fact that a conjunction of propositions returns false if and only if there exists at least one false proposition, we can stop monitoring X once a proposition is detected to be false, because it has already made the conjunction false, regardless of the truth value of the rest of the propositions. Therefore, it is sensible to sort the propositions in X in ascending order by cost before checking anything, when the sorting cost is much lower than the monitoring cost. We then apply the general idea above to monitoring opportunism based on our monitoring approaches, and propose steps to reduce the monitoring cost for opportunism.

References

[1] Jieting Luo and John-Jules Meyer. A formal account of opportunism based on the situation calculus. *AI & SOCIETY*, pages 1–16, 2016.

[2] Jieting Luo, John-Jules Meyer, and Max Knobbout. Monitoring opportunism in multi-agent systems. In *Proceedings of the Coordination, Organizations, Institutions and Norms workshop @ECAI2016*, page to appear. Springer.

[3] Oliver E Williamson. Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization. 1975.

User-Independent Recognition of Sports Activities From a Single Wrist-Worn Accelerometer: A Template-Matching-Based Approach

Jenny Margarito, Rim Helaoui, Anna M. Bianchi, Francesco Sartor, and Alberto G. Bonomi

Abstract

1. Introduction

The discrimination of sports exercises using an unobtrusive monitoring system represents a challenge to real life applications [1].
Several researchers have approached the problem of recognizing human activities such as walking, running and cycling from wearable sensors by applying statistical learning classifiers [2, 3]. Previous works have shown that an accurate classification of different activities needs information extracted from multiple sensors located at different body parts, while solutions adopting one sensor are able to discriminate only few large clusters of activities [4].
Our work investigated the use of template matching for recognizing eight sports activities using one single accelerometer worn at the wrist. In particular, we have evaluated the ability of a template-matching classification algorithm to generalize on a population of overweight subjects and compare it to popular statistical-learning based classifiers.

2. Methods

A population of 48 subjects, 29 normal weight and 19 overweight, was recruited to perform eight common sports activities, such as cycling, cross trainer, rowing, squatting, stepping, running, walking, and weight lifting, while body movement was measured using a triaxial accelerometer placed at the wrist.
Data from 50% of the normal weight subjects (randomly selected) were used for training the classification algorithms, whereas the remaining 50% of the normal weight group and the entire overweight group were used for testing the performance.
The overweight subjects were excluded from the training set and used for testing in order to investigate whether different biomechanical characteristics due to excess weight had an impact on the performance of the activity classification system built on data extracted from normal subjects.
Template-based activity recognition was compared to statistical learning classifiers such as naive bayes (NB), decision tree (DT), logistic regression (LR) and artificial neural network (ANN), which were trained with traditional time- and frequency-domain features [5].
The template matching approach consisted of two major steps: 1) generation of templates for each target activity type based on training data, and 2) classifying unseen data by comparing it to the set of generated templates and searching for the best match.

One user- and axis-independent acceleration signal template was extracted with an automatic procedure for each activity type. Such procedure consisted of calculation of the template length, equivalent to twice the period of a specific activity, and averaging of largely correlated waveforms extracted from different acceleration signals of the same class.
The performance of five different template matching similarity measures were compared: Euclidean distance, dynamic time warping (DTW), derivative DTW [6, 7], correlation and an innovative index calculated as the ratio of correlation coefficient to Euclidean distance (*Rce*). Such index represents an additional measure to find shape similarity between signal segments. It combines correlation and distance measure thus overcoming issues related to noise and different scaling, which usually affect correlation and distance measure, respectively.

3. Results

Template matching well discriminated most of the sport activities by using user independent and automatically extracted template prototypes.
Correlation-based matching techniques and *Rce* generally outperformed other similarity measures such as Euclidean distance, DTW, and DDTW. Sensitivity for the correlation metric and *Rce* was above 80% for the majority of activities on both test datasets, normal weight and overweight subjects.
Not discriminative template was extracted for cycling due to the aperiodic nature of wrist motion during the execution of this activity. However, the low amplitude of the acceleration signal typical of cycling showed to be a powerful feature for distinguishing such activity from other sports. A poorer performance was obtained for cross trainer and rowing classification (*Rce* ~50%) likely due to the subject-specific interaction with the gym equipment employed in the execution of the activity. Indeed, unforeseen wrist acceleration patterns may have been present in the test datasets due to peculiar preference in the usage of supporting handlebars.
Machine learning approach and specifically LR and ANN classifiers achieved the highest overall recognition accuracy (~85%), thus showing to be more accurate than the other proposed techniques.
However, statistical learning algorithms accuracy decreased when evaluated on the overweight subjects test dataset, while increased with template matching. These results confirmed the robustness of the template-matching approach against deterioration in classification accuracy due to previously unseen data and subjects. This aspect may increase the likelihood of obtaining large recognition accuracy in free-living conditions, where the performance of statistical-learning algorithm usually decreases.

4. Conclusion

In this paper, we investigated the use of template matching for classifying sport activities using one single accelerometer worn at the wrist and compared with statistical classifiers. User-independent signal templates were created and five distance- and correlation-based matching techniques were proposed for activity classification.
Template matching showed to be well suited to classify sports activities due to their inherent periodic nature.

The overall classification accuracy of the best (*Rce*) template-matching metric was lower than that offered by the statistical learning algorithms, which can be imputed to the lower accuracy in classifying cross trainer and rowing activities.
However, a template-matching framework preserved classification accuracy during testing, whereas several statistical-learning algorithms did not.
Future work should focus on exploring the benefit of measurement-axis-specific templates, on investigating the advantages of combining template-matching metrics with statistical classifiers and testing generalizability of the proposed solutions on larger datasets including data collected in free living conditions.

References

[1] M. Chen et al., *Body area networks: A survey*, Mobile Netw. Appl., vol. 16, no. 2, pp. 171–193, Apr. 2011.

[2] A. Avci et al, *Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey*, in Proc. 23rd Int Conf. Archit. Comput. Syst., 2010, pp. 1–10.

[3] L. Gao et al., *Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems*, Med. Eng. Phys., vol. 36, no. 6, pp. 779–785, 2014.

[4] L. Atallah et al., *Sensor positioning for activity recognition using wearable accelerometers*, IEEE Trans. Biomed. Circuits Syst., vol. 5, no. 4, pp. 320–329, Aug. 2011.

[5] S. J. Preece et al., *Activity identification using body-mounted sensors: A review of classification techniques*, Physiol. Meas., vol. 30, no. 4, pp. R1–R33, 2009.

[6] E. J. Keogh and M. J. Pazzani, *Derivative dynamic time warping*, in Proc. SIAM Int. Conf. Data Mining, 2001, vol. 1, pp. 5–7.

[7] R. Muscillo et al., *Classification of motor activities through derivative dynamic time warping applied on accelerometer data*, in Proc. IEEE 29th Annu. Int. Conf. Eng. Med. Biol. Soc., 2007, pp. 4930–4933.

Apology and forgiveness evolve to resolve failures in cooperative agreements[1]

Luis Martinez-Vaquero [ab] The Anh Han [c] Luis Moniz Pereira [d]

Tom Lenaerts [ab]

[a] *AI lab, Vrije Universiteit Brussel, Brussels, Belgium*
[b] *MLG, Université Libre de Bruxelles, Brussels, Belgium*
[c] *School of Computing, Teesside University, Middlesbrough, UK*
[d] *NOVA-LINCS,Universidade Nova de Lisboa, Caparica, Portugal*

Abstract

When interactions are repeated mistakes, whether intentionally or not, tend to occur. Researchers have argued that revenge, apology and forgiveness are mechanisms that humans have acquired to ensure that intentional mistakes are avoided and that mutually beneficial relationships can continue. We have shown in the context of the iterated prisoners dilemma wherein agents can decide to make cooperative agreements that these three behaviours emerge spontaneously. Concretely our work reveals that apology and forgiveness are very efficient even in a very noisy environment and ensure long lasting relationships. Yet in order for apology to work, it needs to be sufficiently costly otherwise taking revenge by defecting is the most profitable behaviour. This research has direct implications for online socio-technological systems who's success depends on the trust users (and agents) have in the other users (or agents) participating in the system.

Commitment deals – defined as prior agreements with potentially posterior compensations in case the agreements fail – are most often established to ensure favourable interactions over longer time periods. Experiments have shown that commitment facilitates cooperation in long-term interactions [2], especially when it is voluntary. Moreover, long-term commitments as opposed to one-shot ones [3] are most likely more cost-efficient as the cost of setting up the agreement is paid only once for the entire duration of the agreement. Interestingly, as suggested in [4], commitment in long-term relationships may induce behavioural differences as they may remove the need for reciprocal behaviour like tit-for-tat (TFT).

In [1] we provide for the first time analytical and numerical results for the viability of the commitment strategy within the context of an iterated prisoners dilemma (IPD) [5], expanding our prior work on commitment in the one-shot prisoners dilemma and the public goods game [3, 6]. To study commitment within the IPD context the strategy space needed to be expanded: First, as it is possible for agreements to end before the interaction is finished, strategies need to take into account how to behave when the agreement is present and when it is absent, on top of proposing, accepting or rejecting such agreements in the first place. Second, individuals need to decide whether to continue the agreement when a mistake is made, or end it collecting the compensation. In a cooperative agreement this occurs when a player defects even though she agreed to cooperate with her partner. All these choices define the strategy space of each individual player.

Our work reveals first of all how the detrimental effect of having a large arrangement cost, which was observed for one-shot games [3], is limited as the setup cost is only paid when starting the interaction. Individuals that propose commitments (and are willing to pay their cost) and, following the agreement, cooperate unless a mistake occurs are now the most successful players. But if the agreement is violated through a defection before the IPD is finished then these individuals take revenge by defecting in the

[1] The full paper has been published in the journal *Scientific Reports*, 5, Article number: 10639, June 2015 [1]."

remaining rounds. This observation confirms analytically what has been argued in [7]. Moreover, although defection leads here to the withholding of a the benefit from both players, this revenge-taking behaviour leads to a more favourable outcome for cooperation as opposed to the well-know TFT.

Yet, as mistakes may not be intentional and stopping a mutually beneficial interaction (i.e. cooperation in the IPD) may be strategically a bad idea, individuals may decide not to end the agreement. It might be better to apologise and forgive, continuing the agreement without taking revenge. To study this question the commitment model was extended with an apology-forgiveness mechanism, where apology was defined either as a systemic or individual parameter in the model. In both cases, we have shown that forgiveness is effective if it takes place after receiving an apology from the co-players. However, to play a promoting role for cooperation, apology needs to be sincere, in other words, the amount offered in the apology has to be high enough (yet not too high), which is also corroborated by a recent experimental psychology work [7]. This extension to the commitment model produces even higher cooperation levels than in the revenge-based outcome. In the opposite case, fake committers that propose or accept to commit with the intention to take advantage of the system (defecting and apologising continuously) will dominate the population. In this situation, the introduction of the apology-forgiveness mechanism is detrimental for the level of cooperation level that is produced by commitment and compensation model. Hence there is a lower-limit on how sincere apology needs be as below this limit apology and forgiveness even reduce the level of cooperation further as to what one would expect from simply taking revenge.

As we argued in [8] these results may have direct implications in Artificial Intelligence research: In the context of hybrid socio-technical systems which use reputation scoring to ensure trust [9], apology and forgiveness have been shown to provide additional gains in the transactions. As violations of trust may also occur within systems of interacting non-human autonomous agents [10], our behavioural results may also provide essential mechanisms to ensure that such systems will survive the critical evaluation of their users.

References

[1] Luis A Martinez-Vaquero, The Anh Han, Luís Moniz Pereira, and Tom Lenaerts. Apology and forgiveness evolve to resolve failures in cooperative agreements. *Scientific reports*, 5(10639), 2015.

[2] Robert Kurzban, Kevin McCabe, Vernon L Smith, and Bart J Wilson. Incremental commitment and reciprocity in a real-time public goods game. *Personality and Social Psychology Bulletin*, 27(12):1662–1673, 2001.

[3] The Anh Han, Luís Moniz Pereira, Francisco C. Santos, and Tom Lenaerts. Good agreements make good friends. *Scientific Reports*, 2013.

[4] R. M. Nesse. *Evolution and the capacity for commitment*. Russell Sage Foundation series on trust. Russell Sage, 2001.

[5] Robert Axelrod and William Donald Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[6] The Anh Han, Luís Moniz Pereira, and Tom Lenaerts. Avoiding or Restricting Defectors in Public Goods Games? *Journal of the Royal Society Interface*, page 20141203, 2014.

[7] Michael E McCullough. *Beyond Revenge, the evolution of the forgiveness instinct*. Jossey-Bass, 2008.

[8] Tom Lenaerts, Luis Martinez-Vaquero, The Anh Han, and Luís Moniz Pereira. Conditions for the evolution of apology and forgiveness in populations of autonomous agents. In *2016 AAAI Spring Symposium, Technical Reports*, pages 242–248. AAAI Press, 2016.

[9] Cynthia L Corritore, Beverly Kracher, and Susan Wiedenbeck. On-line trust: concepts, evolving themes, a model. *International journal of human-computer studies*, 58(6):737–758, 2003.

[10] Stephen Marsh and Pamela Briggs. Examining trust, forgiveness and regret as computational concepts. In *Computing with social trust*, pages 9–43. Springer-Verlag, 2009.

Empirical Analysis of Social Support Provided via Social Media[1]

Lenin Medeiros Tibor Bosse

Behavioural Informatics Group, Vrije Universiteit Amsterdam
De Boelelaan 1081, 1081 HV Amsterdam – The Netherlands

Peer support, which can also be referred to as social support or peer-to-peer support, can be defined as supportive information from people with strong social ties, i.e., friends, within a given network. Typically it occurs when a given person talks about their needs for support with others, which leads to a response expressing that one is loved and cared for. An example of that would be to share a stressful personal situation with friends on Facebook. Such friends may give some responses as a way to help the one who is stressed.

Nevertheless, helpful peer support is not always available for all social media users, for the simple reason that some people have fewer friends than other. Moreover, even if they have many friends, users do not always want to share their problems online, especially in cases where such problems are very personal. To improve upon this situation, our research explores the potential of an 'artificial friend': an intelligent software agent with the ability to provide social support via online social networks, thereby helping people to deal with their personal 'everyday problems'.

This paper is part of a project that studies the role of peer support via online social networks. The project is motivated by the observation that everyday problems like broken relationships, difficult work situations, and loss of family members are important sources of stress. Moreover, peer support was found to be a promising means for improving various aspects of (mental) health and well-being.

As a first step in that direction, this paper makes an analysis of the requirements of such a system. Its main goals are to: 1) categorize the types of personal problems people share via social networks, 2) categorize the types of support people offer in such cases (and relate them to the categories found in the first step), and 3) provide a conceptual description of the proposed artificial friend.

In our work, we attempt to categorize classes of peer support by relating them to emotion regulation strategies. We aimed to check to what extent such strategies, which are originally defined in the scope of emotional self-regulation, are also used in our context (friends trying to regulate peers' emotion).

Based on the literature (and, posteriorly, in our data) we came up with the following types of support: 1) *situation selection* – it means to take actions to make it more likely that potentially undesired situations will be avoided (e.g., not going to the party where you might encounter your ex), 2) *situation modification* – it means to take actions in order to try to change a given situation aiming to reduce or even to eliminate its undesired consequences (e.g., keeping some physical distance from your ex when you meet him or her), 3) *attentional deployment* – it is an "internal version of situation selection" in which the individual tries to stop thinking about a given undesired situation (e.g., trying to forget your ex), 4) *cognitive change* – it means to change one or more appraisals regarding a given situation in order to try to modify the situation's emotional significance (e.g., convincing yourself that you can find a better partner), and 5) *general emotional support* – in this type of support the friends try to provide social support only by providing empathy, love, trust and/or caring (e.g., by saying how much their friend means to them).

[1]The full paper was accepted to be published in the proceedings of the *8th International Conference on Social Informatics - SocInfo 2016* and it will be available soon.

Using a crowdsourcing platform, we performed a survey in order to ask people what kind of stressful situations they shared and the respective responses they received from their friends as support. We collected answers from participants located in 26 different countries. After filtering and analyzing our data, the results were processed (see Table 1).

Table 1: Frequency table for all types of shared stressful situations and the respective supports identified. The last column and row (in white) are, respectively, the amount of responses given by the participants' peers for each type of stressful situation and for each support strategy.

	general emotional support	cognitive change	attentional deployment	situation selection	situation modification	
relationship	8	6	5	4	5	28
work	3	8	3	0	4	18
death	8	4	0	0	3	15
financial	2	2	0	0	1	5
disease	1	5	2	0	1	9
exams	0	1	0	0	3	4
other	5	2	0	1	1	9
	27	28	10	5	18	

Even though we don't have enough data yet to state conclusions with statistical significance, this is sufficient to give us some initial insights and state some hypotheses that need to be checked in future works. Such hypotheses are described below.

Hypothesis 1 *The most typical stressful situation shared by people via their social networks concerns their own relationships.*

Hypothesis 2 *In case people share stressful situations about relationships or death, the most frequently used support strategy is general emotional support.*

Hypothesis 3 *In case people share stressful situations about work or diseases, the most frequently used support strategy is cognitive change.*

The results also have provided us insights regarding the software we aim to develop. A conceptual framework was put forward that describes the functioning of the proposed support agent. The main idea is to process incoming messages in online social networks in three steps: analysis of the incoming message, selection of a support strategy, and generation of support messages. This approach makes it possible to generate personalized messages that are based on generic knowledge about support strategies, yet being sufficiently specific to address each individual case separately.

It should be noted that the presented approach focuses on simple interactions consisting of one incoming message and one response message. Nevertheless, it would be interesting to extend it to more complex types of interaction, eventually resulting in entire human-agent conversations.

Finally, once the support agent is actually developed, the next step will be a systematic evaluation. Even though human peer support has already shown to be beneficial in various situations, such an evaluation is crucial to test the assumption that an agent providing such support also has an added value. This will be done both through usability studies (to investigate how users perceive the added value of the agent) and on the longer term, by actually measuring the effectiveness of the system in reducing people's experienced stress.

Acknowledgements

The authors would like to state that Lenin Medeiros' stay at Vrije Universtiteit Amsterdam was funded by the Brazilian Science without Borders program. This work was realized with the support from CNPq, National Council for Scientific and Technological Development - Brazil, through a scholarship which reference number is 235134/2014-7. We also would like to thank all the people who gave us important data by answering our questionnaire.

Incremental multi-class semi-supervised clustering regularized by Kalman filtering[1]

Siamak Mehrkanoon Oscar Mauricio Agudelo Johan A.K. Suykens

KU Leuven, ESAT-STADIUS, Kasteelpark Arenberg 10, B-3001 Leuven (Heverlee), Belgium

Abstract

An on-line semi-supervised learning algorithm formulated as a regularized kernel spectral clustering (KSC) approach is proposed. We consider the case where new data arrive sequentially but only a small fraction of it is labeled. The available labeled data act as prototypes and help to improve the performance of the algorithm to estimate the labels of the unlabeled data points. Given a few user-labeled data points the initial model is learned and then the class membership of the remaining data points in the current and subsequent time instants are estimated and propagated in an on-line fashion. Furthermore we show how the tracking capabilities of the Kalman filter can be used to provide the labels of objects in motion and thus regularizing the solution obtained by the MSS-KSC algorithm.

1 Introduction

In many real-life applications obtaining the labels of input data is often cumbersome and costly. Therefore in many cases one encounters a large amount of unlabeled data while the labeled data are rare. Semi-supervised learning (SSL) is a framework in machine learning that aims at learning from both labeled and unlabeled data points [1]. Most of the SSL algorithms, operate in batch mode, hence requiring a large amount of computation time and memory to handle data streams which exhibit distributional changes like the ones found in real-life applications such as voice and face recognition as well as object tracking in computer vision. Therefore designing SSL algorithms that can operate in an on-line fashion is necessary for dealing with such data streams.

2 Formulation of the method

Consider training data points $\mathscr{D} = \{x_1, ..., x_{n_u}, x_{n_u+1}, .., x_n\}$, where $\{x_i\}_{i=1}^n \in \mathbb{R}^d$. The first n_u data points do not have labels whereas the last $n_L = n - n_u$ points have been labeled. Assume that there are Q classes, then the label indicator matrix $Y \in \mathbb{R}^{n_L \times Q}$ is defined as follows: $Y_{ij} = +1$, if the ith point belongs to the jth class and -1 otherwise. The formulation of Multi-class semi-supervised KSC (MSS-KSC) in primal is given as follows [2]:

$$\begin{aligned} \min_{w^{(\ell)}, b^{(\ell)}, e^{(\ell)}} \quad & \frac{1}{2}\sum_{\ell=1}^{Q} w^{(\ell)^T} w^{(\ell)} - \frac{\gamma_1}{2}\sum_{\ell=1}^{Q} e^{(\ell)^T} V e^{(\ell)} + \frac{\gamma_2}{2}\sum_{\ell=1}^{Q} (e^{(\ell)} - c^{(\ell)})^T \tilde{A} (e^{(\ell)} - c^{(\ell)}) \\ \text{subject to} \quad & e^{(\ell)} = \Phi w^{(\ell)} + b^{(\ell)} 1_n,\ \ell = 1, \ldots, Q, \end{aligned} \tag{1}$$

where $c^{(\ell)}$ is the ℓ-th column of the matrix C defined as $C = [c^{(1)}, \ldots, c^{(Q)}]_{n \times Q} = \left[\begin{array}{c} 0_{n_u \times Q} \\ \hline Y \end{array}\right]_{n \times Q}$. Here $0_{n_u \times Q}$ is a zero matrix of size $n_u \times Q$ and Y is defined as previously. The matrix $\tilde{A}$ is defined as follows:

$$\tilde{A} = \left[\begin{array}{c|c} 0_{n_u \times n_u} & 0_{n_u \times n_L} \\ \hline 0_{n_L \times n_u} & I_{n_L \times n_L} \end{array}\right],$$

[1] The full paper has been published in *Neural Networks*, pages 88–104, 2015.

where $I_{n_L \times n_L}$ is the identity matrix of size $n_L \times n_L$. V is the inverse of the degree matrix. In the dual the solution is obtained by solving a system of linear equations [2]:

$$\gamma_2 \left(I_n - \frac{R 1_n 1_n^T}{1_n^T R 1_n} \right) c^{(\ell)} = \alpha^{(\ell)} - R \left(I_n - \frac{1_n 1_n^T R}{1_n^T R 1_n} \right) \Omega \alpha^{(\ell)}, \tag{2}$$

where $R = \gamma_1 V - \gamma_2 \tilde{A}$. The proposed on-line MSS-KSC consists of two stages [2]. In the first stage, one trains the MSS-KSC algorithm to obtain the initial solution vectors α_i and the cluster memberships. Assuming that N_c clusters are detected, the initial cluster representatives are then obtained. The aim of the second stage is to predict the membership of the new arriving data points using the updated solution vectors α_i. The score variables of the arrived batch of new data points is estimated by means of the out-of-sample extension property of the MSS-KSC method. Then the estimation of the projection of the points in the α-space and calculating the membership of the points are carried out. Finally the cluster representatives in both α and original spaces are updated. The interaction between Kalman filter and I-MSS-KSC algorithm is shown in Fig. 1.

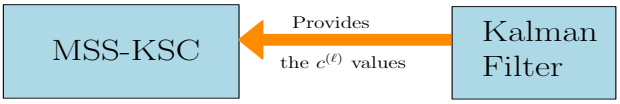


Figure 1: Kalman filter acts as a regularizer for the MSS-KSC algorithm

The applicability of the approach for online video segmentation are shown in the Fig. 2. In this case the segments represent the clusters. The membership of the pixels act as label. The data point x is the local color histogram of the pixels. The number of data points is therefore equals to the number of pixels in the given image.

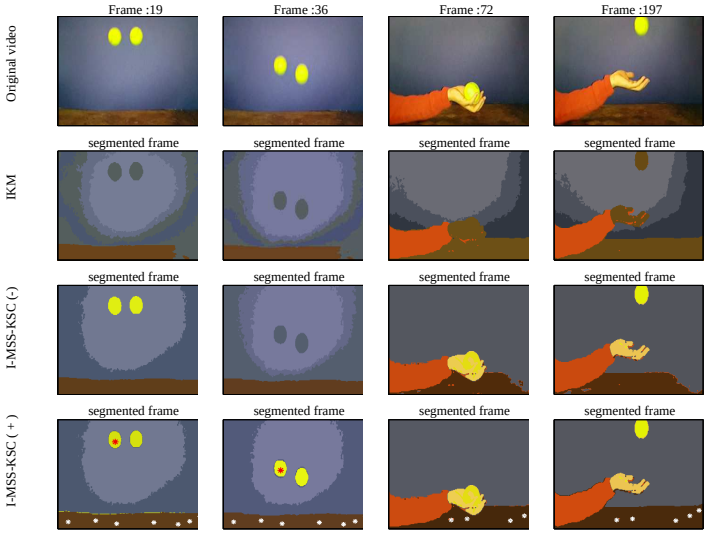


Figure 2: Bouncing balls and Siamak's hand video. **First row:** Original video, **Second row:** Incremental Kmeans, **Third row:** Incremental MSSKSC(-): the user labels are only provided for the first frame, **Fourth row:** Incremental MSSKSC(+): the user labels are provided for the entire frames.

Acknowledgments. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC AdG A-DATADRIVE-B (290923). This letter reflects only our views: The EU is not responsible for any use that may be made of the information in it. The research leading to these results received funds from the following sources: Research Council KUL: GOA/10/09 MaNet, CoE PFV/10/002 (OPTEC), BIL12/11T; PhD/Postdoc grants; Flemish Government: FWO: PhD/Postdoc grants, projects: G.0377.12 (Structured systems), G.088114N (Tensor based data similarity); IWT: PhD/Postdoc grants, projects: SBO POM (100031); iMinds Medical Information Technologies SBO 2014; Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017). Siamak Mehrkanoon is a postdoctoral researcher at KU Leuven, Belgium. Johan Suykens is a full professor at KU Leuven, Belgium.

References

[1] Mikhail Belkin and Partha Niyogi. Semi-supervised learning on Riemannian manifolds. *Machine learning*, 56(1-3):209–239, 2004.

[2] Siamak Mehrkanoon, Oscar Mauricio Agudelo, and Johan AK Suykens. Incremental multi-class semi-supervised clustering regularized by Kalman filtering. *Neural Networks*, 71:88–104, 2015.

Methods for causal inference from gene perturbation experiments and validation[1]

Nicolai Meinshausen [a] Alain Hauser [b] Joris M. Mooij [c] Jonas Peters [d]
Philip Versteeg [c] Peter Buhlmann [a]

[a] *Seminar for Statistics, Eidgenösische Technische Hochschule (ETH) Zurich, CH-8092 Zurich, Switzerland*
[b] *Department of Engineering and Information Technology, Bern University of Applied Sciences, CH-3400 Burgdorf, Switzerland*
[c] *Informatics Institute, University of Amsterdam, 1090 GH Amsterdam*
[d] *Max Planck Institute for Intelligent Systems, D-72076 Tuebingen, Germany*

1 Introduction

In many fields of science, research focuses not only on exposing associative relations between variables but on identifying causal effects. Colloquially, a causal effect is the effect that an external perturbation has on the system as a whole. In this work we focus on the task of predicting the outcome of unseen and new interventions by using heterogeneous data from both observational measurements and perturbation experiments. For example, in genetics a typical task is to predict what the effect is of a new gene knock-out on a phenotype of interest. Causal predictions can be plagued with various problems, especially for large-scale systems with many measured variables in different experimental environments, i.e. in the genetic case above one can have combined data from wild-type observations and many different knock-out experiments. We present software and validation of a newly developed method for causal prediction using heterogeneous data.

2 Causal Modeling

One powerful way to represent the causal structure that generates the observational data at hand is using Structural equation models (SEM's) [1]:

$$X_j \leftarrow f_j\left(X_{\mathrm{pa}(j)}, \varepsilon_j\right)(j = 1, \ldots, p+1), \tag{1}$$

where each node j is caused by its the parent set pa_j through a function f_j in the underlying directed graph and the error terms ε_j are jointly independent. Interventional effects are commonly modeled with the do-operations[4] where $do\left(X_j = x\right)$ represents setting variable X_j to value x, corresponding to replacing the relevant structural equation with $X_j \leftarrow x$ in the SEM.

2.1 Invariant Causal Prediction

Assuming a linear model, eq. 1 for one prediction target Y can be written as

$$Y \leftarrow \sum_{k \in S^*} \gamma_k^* X_k + \varepsilon_Y, \tag{2}$$

[1] The full paper has been published in *Proceedings of the National Academy of Sciences* **113**(27):7361-7368, 2016.

where $S^* = \mathrm{pa(Y)}$ denotes the parent set or the 'direct causes' of Y and γ^* corresponds to the coefficients (edge weights) in the SEM (directed graph). We further assume that we have data (X^e, Y^e) for target and predictor variables in different environments $e \in \mathcal{E}$. Invariant Causal Prediction (ICP) [5] is a recently developed algorithm that exploits these different environments $\mathcal{E}$ to find common components of the regression vector γ^e that remain invariant across all e. The method returns an estimate of the parent set $\hat{S}$ for a given Y with a predetermined confidence level of $1 - \alpha$, when an **Invariance Assumption**[2] on the true coefficients γ^* holds. As an example, the invariance assumption holds under an environment of do-interventions as long as the interventions do not directly target the target variable Y. ICP crucially depends on data from different experimental settings and has a higher degree of identifiability of causal effects with increasingly diverse sources $|\mathcal{E}|$.

3 Validation with Gene Perturbation Experiments

We consider empirical validation of the ICP method using large-scale genome-wide gene perturbation data on the species *Saccharomyces cerevisiae* [2]. Genome-wide mRNA expression levels are measured for 6170 genes: 160 observational data points from wild-type individuals and 1479 interventional data arising from single-gene deletions. Notating X_i as the expression level for gene i, the goal is to predict whether the response single-gene expression level $Y \in \{X_1, \ldots, X_{p+1}\} \setminus \{X_j\}$ changes under intervention j for all $p = 6169$ possible targets. Applying the ICP method, we use a simple labeling of experimental settings: $\mathcal{E} = \{1, 2\}$ where $e = 1$ corresponds to observational data and $e = 2$ to all perturbation data, and have a confidence level of $\alpha = 0.01$ for $\gamma_j \neq 0$.

A fivefold validation split of the data is made: repeatedly leaving one fold of interventional data out and predicting the parent set of those missing interventions with the remainder. A score is made on repeated bootstrapped samples of the data by ranking the edges in order of decreasing frequency. Results from ICP are comparing favorably to other methods by using an internal score, the Strong Intervention Effects (SIE) as a ground-truth[3]. Additionally we defined a score based on external data from transcription factor measurements (TF) [3] and six scores (A-F) that were extracted from a bioinformatics source on yeast interactions at **`yeastgenome.org`**. The top 20 of most selected ICP pairs shows significant overlap with both the SIE and external measures but not with TF score.

4 Conclusions

The newly developed ICP method for causal inference uses invariance in heterogeneous data and is equipped with confidence bounds for inferential statements. We validate the algorithm empirically on single-gene deletion data from *Saccharomyces cerevisiae* internally and across various additional datasets. This demonstrates the successful prediction of the outcome of an unseen new intervention.

References

[1] K.A. Bollen. *Structural Equations with Latent Variables*. John Wiley & Sons, New York, NY, 1989.

[2] P. Kemmeren et. al. Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors. *Cell*, 157:740–752, 2014.

[3] K.D. MacIsaac, T. Wang, D.B. Gordon, D.K. Gifford, G.D. Stormo, and E. Fraenkel. An improved map of conserved regulatory sites for Saccharomyces cerevisiae. *BMC Bioinformatics*, 7:113, 2006.

[4] J. Pearl. *Causality: models, reasoning and inference*. Cambridge University Press, New York, NY, 2000.

[5] J. Peters, P. Bühlmann, and N. Meinshausen. Causal inference using invariant prediction: identification and confidence intervals, 2015. Preprint arXiv:1501.01332, to appear in JRSSB.

[2] As defined in the paper.

[3] See **fig. 2** in the paper.

Functional Representation of Prototypes in LVQ and Relevance Learning[1]

Friedrich Melchert [ab] Udo Seiffert [b] Michael Biehl [a]

[a] *University of Groningen, Johann Bernoulli Institute for Mathematics and Computer Science, P.O. Box 407, 9700AK, Groningen, The Netherlands*
[b] *Fraunhofer Institute for Factory Operation and Automation IFF, Sandtorstr. 22, 39106 Magdeburg, Germany*

1 Introcution

A large number of unsupervised and supervised machine learning techniques are based on the use of distances or dissimilarity measures. Such measures can be employed for pairwise comparison of feature vectors, as for instance in the well-known K-Nearest-Neighbor (KNN) classifier. Prototype-based methods replace the reference data by a number of typical representatives. In the popular Learning Vector Quantization (LVQ) [1], prototypes serve as characteristic exemplars of the classes and, together with a distance measure, parameterize the classification scheme. Prototype- and distance-based systems are generally intuitive and straightforward to implement.

Here, we examine the use of prototype-based systems for functional data [2], where feature vectors do not simply comprise a set of more or less independent quantities, but represent a functional relation. This is the case, for instance, in the presence of temporal or other dependencies, which impose a natural ordering of the features. Functional data is found in a large variety of practical application areas [2]. Perhaps, time series and sequences come to mind first in the context of, e.g., bioinformatics, meteorology or economy. Similarly, densities or histograms can be used to represent statistical properties of observations. Another important example is that of spectral data, for instance optical or mass spectra obtained in various fields ranging from remote sensing to chemistry and bioinformatics.

2 Methodology

We consider a standard classification problem, where a set of d-dimensional labeled feature vectors serve as examples for the target classification. This training set of feature vectors is used to parameterize a Generalized Matrix Relevance LVQ (GMLVQ) system [3]. GMLVQ is an extension of LVQ, employing an adaptive quadaratic distance meassure, that is optimized during the training process.

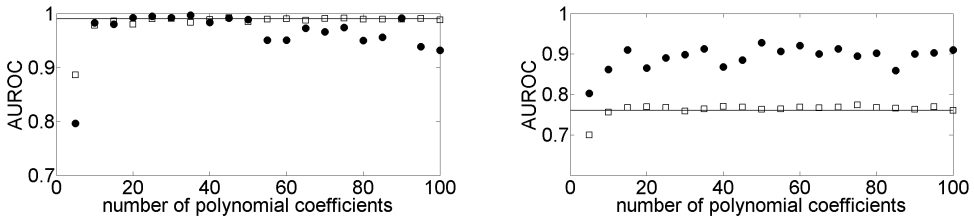
To obtain a data representation that incorporates its functional nature, we perform a data approximation as a weighted sum of basis functions:

$$f_{\mathbf{c}}(x) \approx \sum_{i=0}^{n} c_i g_i(x). \tag{1}$$

For the illustration of the approach we selected Chebyshev polynomials of first kind as a particular functional basis, due to their straightforward implementation and numerical advantages.

Three different classification scenarios were considered: In scenario (A) a GMLVQ system is adapted using the original d-dimensional feature vectors. In scenario (B) the classification system is trained on the lower dimensional vectors of approximation coefficients obtained from the functional expansion of the data as given in Eq. (1). To investigate whether a potential increase of classification

[1] The full paper has been published in *Advances in Self-Organizing Maps and Learning Vector Quantization Volume 428* of the series *Advances in Intelligent Systems and Computing* pages 317 – 327, Springer, 2016.



(a) Wine dataset (b) Orange Juice dataset

Figure 1: Comparison of the achieved validation performance, i.e. the area under ROC for different datasets in dependence of the degree of the polynomial curve fitting. The solid line represents the value for the classification using the original feature vectors (A). Filled circles represent results achieved using polynomial curve fitting, open squares correspond to polynomial smoothing of the input data only. All results are displayed as a function of the number of coefficients n in the polynomial expansion.

performance is the result of the smoothening implicit to the approximation, in the third scenario (C) the data is first approximated as in (B) and the resulting approximation is resampled equidistantly to the original input space, thus creating smoothened copies of the original data. A GMLVQ system is then trained on this set of smoothened data. To evaluate the influence of the approximation quality, the number of approximation coefficients is varied from $n = 5, 10, ..., 100$ in scenarios (B) and (C).

3 Results and Discussion

In Figure 1 the results for two of the examined datasets are shown. The classficiation performance is determined as average Area under the ROC (AUROC) over ten random splits of the data, selecting 90% training data and 10% validation data.

The obtained results demonstrate that the proposed approach has potential to improve classification performance with functional data significantly, as verified by the *Orange Juice*[2] dataset (Fig. 1b). The prediction accuracy as measured by the AUROC is significantly larger when applying GMLVQ in the polynomial representation. For a wide range of degrees n, the AUROC robustly exceeds that of the system trained from the raw datasets. This is not the case for training according to scenario (C), which shows that the improvement in (B) cannot be explained as an effect of the smoothing only.

In the other depicted dataset[2] (*Wine*, Fig. 1a) the accuracy of the system with polynomial representation is comparable or nearly identical to that in original and smoothed feature space. Clearly, the positive effect of the functional representation on the performance will depend on the detailed properties of the data and the suitability of the basis functions. However, even if accuracy does not increase, one advantage of the polynomial representation remains: The dimensionality of the problem can be reduced significantly in all considered cases without adverse effects on performance. For approximations with, say, 20 polynomial coefficients, which performs well in all datasets, the number of input dimensions is reduced by approx. 92% for *Wine* and by 97% for the *OJ* dataset. This leads to a drastically reduced number of free parameters, which is quadratic in the number of feature dimensions in GMLVQ, and results in a massive speed-up of the training process. Additionally the smaller number of free parameters reduces the risk of over-fitting, avoids convergence problems and requires less computational effort.

References

[1] M. Biehl, B. Hammer, and T. Villmann. Prototype-based models in machine learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 7(2):92–111, 2016.

[2] J. Ramsay and B. Silverman. *Functional Data Analysis*. Springer, 2006.

[3] P. Schneider, M. Biehl, and B. Hammer. Adaptive relevance matrices in Learning Vector Quantization. *Neural Computation*, 21:3532–3561, 2009.

[2]Dataset available at: University of Louvain, `http://mlg.info.ucl.ac.be/index.php?page=DataBases`

Interventions on Contextualized Decision Making: an Agent-Based Simulation Study [1]

Rijk Mercuur [a] Frank Dignum [a] Yoshihisha Kashima [b]

[a] *Utrecht University*
[b] *University of Melbourne*

Despite both governmental and scientific interventions, greenhouse gas (GHG)-emissions continue to rise, almost surely leading to more frequent extreme weather events, a rise in sea levels, an increase in widespread infectious diseases, and decreases in crop yields and water quality [3]. Climate change is primarily driven by GHG-emitting human behavior and could therefore be largely mitigated by interventions in human behavior. For example, a global transition to a healthy low meat-diet would have a substantial impact on lowering GHG-emissions [5, 9]. However, human behavior is the least understood aspect of the climate change system and many interventions have not had the effect hoped for [4]. One possible major defect in such interventions might be the focus on changing intentions rather than context [1]. Whether one's prior aims, i.e. intentions, actually lead to action is influenced by one's setting, i.e. context. For example, information on the environmental consequences of meat-eating might not induce behavior change [e.g. 6]. Firstly, because one's context might not give the opportunity to do a new action [8]. For example, one might have an intention to eat a vegetarian dish, but there are no nice vegetarian dishes on the menu. Secondly, one's context can trigger a *habit* opposing one's intention. Habits are learned dispositions to repeat past behavior that emerge when context co-varies with behavior [10]. For example, one might have the intention to eat a vegetarian dish, but upon entering the pub with mates one automatically orders that good old hamburger. This raises the main research question: how can we improve interventions given the described influence of context on decision making? In this study we focus on the choice to eat meat or vegetarian dishes. Some results will only apply to this domain, but the study aims to be general enough for some results to be extrapolated to other domains.

To study this question we first ask ourselves a subquestion: can we explain the limited success of interventions that target intentions rather than context, with a model that uses the concepts of agents and social practices? Agents represent acting individuals, e.g. robots, bacteria, or in this case, humans. A social practice is a way of (inter)acting that is shared between these agents. The concept 'social practice' originates from sociology [7] and aims to integrate actors with their context. It could thus be a good starting point to model contextualized decision making. The challenge in this study lies in translating this originally macroscopic notion to the microscopic deliberation of agents. In this study we present such an an agent-based model that uses social practices on the micro level to give a parsimonious representation of contextualized decision making. In this model each agent has a personal representation of the social practice of dining (see figure 1).

The social practice firstly differs in two embodiments of the practice of dining (green and red in figure 1): meat dining and vegetarin dining. For each of these embodiments the agent keeps track of some important information:

Opportunities describe features of the context that enable an embodiment (e.g. V:Vegatarian Venue);

Triggers describe features of the context that habitually trigger an embodiment (e.g. A1:agent1);

Purpose refers to the meaning of the practice, i.e. the value it furthers (e.g. self-transcendence);

[1] An extended abstract has been published in *Advances in Social Simulation 2015 (in press)* and the full master thesis can be found in the Utrecht University Repository.

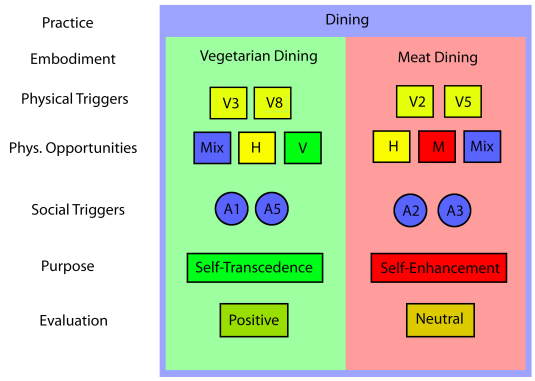


Figure 1: A possible instantiation of the social practice of dining.

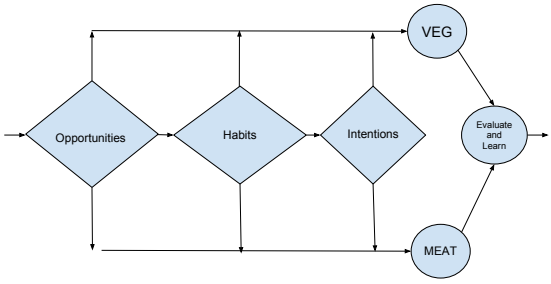


Figure 2: Schematic overview of the deliberation cycle of an agent.

Evaluation captures the agent's own judgements about past enactments of the practice (e.g. positive);

The agent uses this information in its deliberation cycle to decide if it will eat meat or vegetarian (see figure 2). There are several stages in the deliberation that act as a filter on the embodiments: if only one embodiment remains it is enacted, as long as multiple remain the agent will continue with the next stage in the deliberation. In the first stage of the deliberation an agent filters the impossible embodiment(s) out by comparing its current context with the practices' opportunities. For example, one cannot eat vegetarian in a meat venue. In the second stage of the deliberation the agent filters the least salient embodiment(s) out by comparing its current context with the practices' triggers. What context features trigger an embodiment depends on past enactment and evaluation. In the third stage of the deliberation an agent will use his intentions to make a choice out of the remaining possible embodiments. The agents differ in these initial intentions based on the values they support. (Values as in 'ideals of worth' . Please refer to the full paper for the role of values in the model.) After the enactment an agent will learn. Firstly, by updating the habitual triggers of the practice based on hebbian learning. Secondly, it will evaluate the enactment based on its alignment with intentions and what the others did. To validate the model we firstly made sure that it replicates the limited succes of interventions that target intentions rather than context. In addition to this 'replicative validity', we payed careful attention to relating the mechanisms of the model to intuition and literature on habits and intentions, i.e. 'structural validity'.

This model is firstly relevant because it helps in gaining insight in our main research question. For example, it showed us that in our scenario the limited success of interventions targeting intentions is a result of meat-eating habits and not so much a lack of opportunities to eat vegetarian. Furthermore, it showed us that although agents do evaluate meat-eating lower because it is not in line with their intentions, they do not escape their habits because they see that other agents keep doing the same behavior. This model is not only relevant because it helped us gain such insights, but also because it answer the calls in several fields for an adequate model of contextualized decision making (e.g. social psychology [3] and agent based modeling [2]). Although as we discuss in the full thesis there is more work to be done on the topic of learning and the interconnectedness of social practices. We continued by using our model to gain more insight in our second subquestion: what happens if we aim interventions at the context? We simulated opening extra vegetarian venues, changing the menu of the meat venue, and organizing a 'vegetarian week' at a new location. We showed that this model has great potential in explaining why some interventions on context might work where others do not. As we discuss in the full thesis we already gained a few insights from these experiments. Firstly, providing new opportunities does not always have a positive effect, especially if it allows people to fall back in their old habits. Temporarily breaking someone's main routine and evaluation can have a long lasting effect. And lastly, a crucial opportunity to break habits is to appeal to people's intentions to change their context. We combined these insights in a 'vegetarian week' intervention that promises great effects. We ended our discussion by explaining that such results largely depends on how habit transcends context and directed further work towards this crucial topic. We conclude this study by stating that we hope this model will form the basis for many context-aware social agent models and that the resulting insights will help policy makers to help prevent one of the most important problems humanity faces today: global warming.

References

[1] Wokje Abrahamse, Linda Steg, Charles Vlek, and Talib Rothengatter. A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25:273–291, 2005. ISSN 02724944. doi: 10.1016/j.jenvp.2005.08.002.

[2] Virginia Dignum and Frank Dignum. Contextualized Planning Using Social Practices. In Aditya Ghose, Nir Oren, Pankaj Telang, and John Thangarajah, editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems X: COIN 2014 International Workshops, COINAAMAS, Paris, France, May 6, 2014, COINPRICAI, Gold Coast, QLD, Australia, December 4, 2014, Revised Selected Papers*, pages 36–52. Springer International Publishing, Cham, 2015. ISBN 978-3-319-25420-3.

[3] Robert Gifford. Environmental psychology matters. *Annual review of psychology*, 65:541–79, 2014. ISSN 1545-2085. doi: 10.1146/annurev-psych-010213-115048.

[4] Intergov. Panel Climate Change. Climate Change 2014 Synthesis Report Summary Chapter for Policymakers, 2014.

[5] Christopher M Jones and Daniel M Kammen. Quantifying carbon footprint reduction opportunities for U.S. households and communities. *Environmental science & technology*, 45(9):4088–95, may 2011. ISSN 1520-5851. doi: 10.1021/es102221h.

[6] Mirra Seigerman. Manipulating Meat-Eating Justifications: A Novel Approach to Influencing Meat Consumption, 2014.

[7] Elizabeth Shove, Mike Pantzar, and Matt Watson. *The Dynamics of Social Practice: Everyday Life and How it Changes*. SAGE Publications, London, 2012. ISBN 9780857020437. doi: 10.4135/9781446250655.n1.

[8] Linda Steg and Charles Vlek. Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology*, 29(3):309–317, 2009. ISSN 02724944. doi: 10.1016/j.jenvp.2008.10.004.

[9] Elke Stehfest, Lex Bouwman, Detlef P. Van Vuuren, Michel G J Den Elzen, Bas Eickhout, and Pavel Kabat. Climate benefits of changing diet. *Climatic Change*, 95(1-2):83–102, 2009. ISSN 01650009. doi: 10.1007/s10584-008-9534-6.

[10] Wendy Wood and David T Neal. A new look at habits and the habit-goal interface. *Psychological review*, 114(4):843–863, 2007. ISSN 0033-295X. doi: 10.1037/0033-295X.114.4.843.

Incentivizing Intelligent Customer Behavior in Smart-Grids: A Risk-Sharing Tariff & Optimal Strategies — Extended Abstract[1]

Georgios Methenitis [ab] Michael Kaisers [a] Han La Poutré [ab]

[a] *Centrum Wiskunde & Informatica, Amsterdam, The Netherlands*
[b] *Delft University of Technology, Delft, The Netherlands*

Abstract

Current electricity tariffs for retail rarely provide incentives for intelligent demand response of flexible customers. Such customers could otherwise contribute to balancing supply and demand in future smart grids. This paper proposes an innovative risk-sharing tariff to incentivize intelligent customer behavior. Within a game-theoretical analysis, we capture the strategic conflict of interest between a retailer and a customer in a two-player game, and we present optimal, i.e., best response, strategies for both players in this game. We show analytically that the proposed tariff provides customers of varying flexibility with variable incentives to assume and alleviate a fraction of the balancing risk, contributing in this way to the uncertainty reduction in the envisioned smart-grid.

1 Introduction

Energy systems are in transition towards more sustainable but fluctuating generation portfolios, which must be matched with more flexible behavior in the demand side. In the current electricity systems, retailers are the balancing responsible parties, pooling customers into larger portfolios. Retailers are responsible to cover the balancing needs resulting from deviations from the predicted demand. The main strategy for retailers to control balancing costs is to avoid the need to purchase expensive balancing power, incurred by fast ramping generators, by reducing uncertainty in the demand side.

Most existing tariffs by electricity retailers, especially in Europe, do not provide incentives for intelligent customer behavior, which could otherwise contribute in balancing supply and demand, thereby improving the efficiency of the power system. Dynamic pricing is a means to encourage favorable changes in demand patterns by the customers, e.g., time of use (ToU), critical peak price (CPP), and real-time pricing. However, dynamic pricing approaches may introduce disruptive and unfavorable market behavior [2], and thus planning and ahead prices as presented here are required.

2 The Risk-Sharing Tariff

We propose the risk-sharing tariff, where the customer can choose the amount of risk to assume from the retailer. We consider as risk the expected costs of the retailer for participating in the balancing market with regards to the default demand behavior of the customer, considering settings where one customer has a direct or a representative influence on the balancing requirements of the retailer.

We capture the strategic interactions between the retailer and the customer in a two-player game, where each player responds optimally to the actions of the other player. We assume a two-step market. The retailer, to satisfy the expected demand of the customer, first procures electricity in the ahead market with a low unit price p, and later pays for any absolute deviation, between the observed demand of the

[1] The full paper appears in the *Proceedings of the 25th International Joint Conference on Artificial Intelligence* [1].

customer and the procured quantity, in the balancing market with a high unit price p', where $p' > p$. We assume that the default behavior of the customer can be described by a probability distribution function that is known to the retailer.

Similarly to the closest state-of-the-art work [3], the risk-sharing tariff comprises two price functions. The pre-commitment price function $p_c(\tau)$ for the anticipated demand, and the imbalance price function $p'_c(\tau)$ for any absolute deviation from the anticipated demand. The risk-sharing tariff requires the customer to pay for his anticipated load with the price $p_c(\tau)$, and later pay any deviation from the expected load with the price $p'_c(\tau)$. Let $\tau \in [0, 1]$ denote the share of risk that remains with the retailer, $(1 - \tau)$ is the share of risk that is assumed by the customer. Figure 1 illustrates the two price functions, which are computed based on the optimal procurement quantity of the retailer in the ahead market and the prices p, p'. Note that for $\tau = 1$ the balancing risk stays with the retailer, on the contrary for $\tau = 0$ the customer holds all the risk. The pre-commitment price decreases as the risk for the customer increases. In this context, we interpret demand response as the ability of the customer to reduce the uncertainty of his demand, in response to the financial incentives (decreasing price with risk assumption) offered by the risk-sharing tariff. Furthermore, the price functions embrace two desired properties with regards to the demand response ability of the customer: (1) a customer that cannot reduce the uncertainty of his demand has no incentive to assume any risk, (2) given the quadratic form of the imbalance price function $p'_c(\tau)$, the optimal choice of the risk is proportional to the ability of the customer to reduce the uncertainty of his demand.

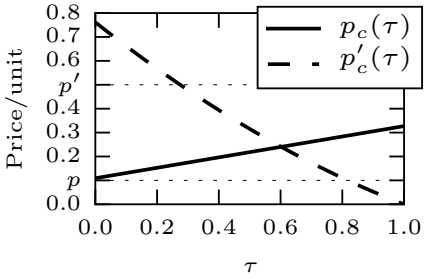


Figure 1: The price functions $p_c(\tau)$ and $p'_c(\tau)$ for all values of τ.

3 Contributions & Conclusion

We propose a novel tariff where the balancing risk can be shared between the retailer and the customer, contributing to the uncertainty reduction in the demand. We show that the proposed tariff is acceptable, yielding only benefits for both players. We further study how social welfare is improved due to the uncertainty reduction, and we provide arguments why the proposed tariff elicits all freely available demand response. We show the existence of Nash-equilibria within the two-player game when the retailer has full access to the information related to the activation costs of the demand response. Last, we show how a bounded-rational customer can affect the tariff selection by the retailer: the retailer is keen to offer more incentives to a customer that may not choose the cheapest risk option available.

The proposed work provides the basis for tariffs that enable customers to choose their own participation in the balancing risk the retailers face in the balancing markets, enabling in this way uncertainty reduction on the demand side.

Acknowledgments

This work is part of the research programme Uncertainty Reduction in Smart Energy Systems (URSES) with project number 408-13-012, which is partly financed by the Netherlands Organisation for Scientific Research (NWO).

References

[1] Georgios Methenitis, Michael Kaisers, and Han La Poutré. Incentivizing intelligent customer behavior in smart-grids: A risk-sharing tariff & optimal strategies. In *Proceedings of the 25th International Conference on Artificial Intelligence*, pages 380–386. AAAI Press, 2016.

[2] M. Roozbehani, M.A. Dahleh, and S.K. Mitter. Volatility of power grids under real-time pricing. *Power Systems, IEEE Transactions on*, 27(4):1926–1940, Nov 2012.

[3] Meritxell Vinyals, Valentin Robu, Alex Rogers, and Nicholas R. Jennings. Prediction-of-use games: A cooperative game theory approach to sustainable energy tariffs. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '14, pages 829–836, Richland, SC, 2014. International Foundation for Autonomous Agents and Multiagent Systems.

A topological insight into restricted Boltzmann machines (extented abstract)[1]

Decebal Constantin Mocanu [a] Elena Mocanu [a] Phuong H. Nguyen [a]
Madeleine Gibescu [a] Antonio Liotta [a]

[a] *Eindhoven University of Technology, Dep. of Electrical Engineering, Netherlands*

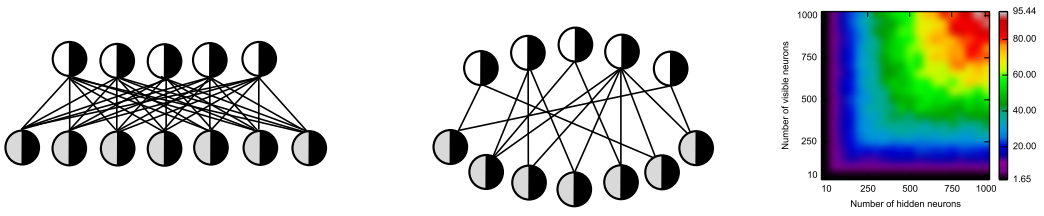
1 Introduction

Restricted Boltzmann Machines (RBMs) and models derived from them have been successfully used as basic building blocks in deep neural networks for automatic features extraction, unsupervised weights initialization, but also as standalone models for density estimation, activity recognition and so on. Thus, their generative and discriminative capabilities, but also their computational time are instrumental to a wide range of applications. The main contribution of this paper [4] is to study the above problems by looking at RBMs and Gaussian RBMs (GRBMs) [2] from a topological perspective, bringing insights from network science, an extension of graph theory which analyzes real world complex networks [6].

2 The proposed method

Firstly, we study the topological characteristics of RBMs and GRBMs, showing that these exhibit a small-world topology. We then hypothesize that by constraining the topology to be also scale-free it is possible to reduce the size of ordinary RBMs and GRBMs models, as it has been shown in [1] that scale-free networks are sparse. Thus, we introduce a three stages method to create RBMs and GRBMs with small-world, scale-free topologies while still considering local neighborhoods and data distribution. In the first stage, a scale-free bipartite graph is generated; in the second one, the graph is adjusted to be also small-world; and in the third stage, the graph topology is fitted to the data distribution. We dub the resulting models as compleX Boltzmann Machine (XBM) (see Fig. 1a) and Gaussian compleX Boltzmann Machine (GXBM), respectively.

An interesting finding is that constraining such XBM and GXBM topologies at their inception leads to intrinsically sparse networks, a considerable advantage to typical state-of-the-art methods in which sparsity is enforced as an aftermath, that is during testing (exploitation) phase (e.g. [3]). In turn, XBM



(a) (b)

Figure 1: (a) Schematic architecture of: RBM (left) and XBM (right). (b) Studying the relation between the number of weights in RBM and XBM (the heatmap values are given by n_w^{RBM}/n_w^{XBM}), where n_w is the number of weights of the specific model.

[1]The full paper has been published in *Machine Learning*, Volume 104, 2016, Pages 243–270, ISSN 1573-0565, http://dx.doi.org/10.1007/s10994-016-5570-z.

Table 1: Estimation of the average log-probabilities on the training and testing data obtained from the MNIST digits dataset using AIS [5] on fully connected RBM, XBM and two state-of-the-art sparse models (i.e. $\text{RBM}_{\text{FixProb}}$, $\text{RBM}_{\text{TrPrTr}}$). The RBM result is taken from [5].

No. of CD steps during learning	No. of weights	Model	No. of hidden units	Average shortest path	Average cluster coefficient	No. of pruning iterations	Average train log-probabilities	Average test log-probabilities
	392000	RBM	500	1.52	1	0	-83.10	-86.34
	387955	XBM	27000	2.05	0.156	0	-86.12	**-85.21**
	391170	$\text{RBM}_{\text{FixProb}}$	27000	2.87	0.053	0	-107.23	-106.78
from 1 to 25	3262957	$\text{RBM}_{\text{TrPrTr}}$	27000	2.18	0.076	50	-349.87	-376.92
(variable)	10790	XBM	500	2.44	0.082	0	-121.26	-120.43
	10846	$\text{RBM}_{\text{FixProb}}$	500	3.12	0.039	0	-136.27	-135.89
	36674	$\text{RBM}_{\text{TrPrTr}}$	500	2.35	0.071	50	-134.25	-135.76

and GXBM have a considerably smaller number of weights, which further on contributes to considerably faster computational times (proportional to the number of weights in the model), both in the training and testing phases. What is more, we found that the proposed topology imposes an inductive bias on XBMs and GXBMs, which leads to better statistical performance than RBMs and GRBMs. Our comparative study is based on both simulated and real-world data, including the Geographical origin of music dataset, the MNIST digits dataset, CalTech 101 Silhouettes dataset, and the 8 datasets from UCI evaluation suite. We show that, given the same number of hidden neurons, XBM and GXBM have similar or relatively close capabilities to RBM and GRBM, but are considerably faster thanks to their reduced amount of weights. For instance, in a network of 100 visible and 100 hidden neurons, the reduction in weights was by one order of magnitude. A network with 1000 visible and 1000 hidden neurons led to a reduction in weights by two orders of magnitude, as depicted in Fig. 1b. Additionally, we show that given the same amount of weights, XBMs and GXBMs achieve better generative capabilities than fully connected RBMs or GRBMs, due to their higher number of hidden neurons. For the sake of illustration, Table 1 presents a snapshot of XBM performance on the MNIST dataset.

3 Conclusion

In this paper [4], we look at RBMs and GRBMs from a topological perspective, bringing insights from network science. Firstly, we point out that RBMs and GRBMs are small-world bipartite networks. Secondly, by introducing scale-free constraints, we devised two novel sparse models, namely XBMs and GXBMs. These sparse models exhibit much faster computational time than their fully connected counterparts thanks to a smaller number of parameters which have to be computed, at almost no cost in performance. We believe that this will lead to the ability to tackle problems having much higher dimensional data - something that is today unfeasible without performing dimensionality reduction - and we intend to tackle this research direction in the near future.

References

[1] Charo I. Del Genio, Thilo Gross, and Kevin E. Bassler. All scale-free networks are sparse. *Phys. Rev. Lett.*, 107:178701, Oct 2011.

[2] G. E. Hinton and R. R. Salakhutdinov. Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786):504–507, July 2006.

[3] Honglak Lee, Chaitanya Ekanadham, and Andrew Y. Ng. Sparse deep belief net model for visual area v2. In J.C. Platt, D. Koller, Y. Singer, and S.T. Roweis, editors, *Advances in Neural Information Processing Systems 20*, pages 873–880. Curran Associates, Inc., 2008.

[4] Decebal Constantin Mocanu, Elena Mocanu, Phuong H. Nguyen, Madeleine Gibescu, and Antonio Liotta. A topological insight into restricted boltzmann machines. *Machine Learning*, 104(2):243–270, 2016.

[5] Ruslan Salakhutdinov and Iain Murray. On the quantitative analysis of deep belief networks. In *In Proceedings of the International Conference on Machine Learning*, pages 872–879, 2008.

[6] Steven H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, March 2001.

Hope and Fear in Reinforcement Learning Agents: Extended Abstract[1]

Thomas M. Moerland Joost Broekens Catholijn M. Jonker

Interactive Intelligence, Delft University of Technology

Abstract

We study models of emotion generation in reinforcement learning agents, focussing on anticipatory emotions. Taking inspiration from the psychological Belief-Desire Theory of Emotions (BDTE), our work specifies models of hope and fear based on best and worst forward traces. Results illustrate the plausibility of these signals, for example in the game *Pacman*. Our models enable learning agents to elicit hope and fear, and moreover, explain what anticipated event caused the emotion.

1 Introduction

This paper studies models of anticipatory emotions in reinforcement learning (RL) agents. Computational emotion models are usually derived from the agent's decision making architecture, of which RL is an important subclass. Studying emotions in RL-based agents is useful for different research fields. For machine learning (ML) researchers, emotion models may improve learning efficiency, e.g. by influencing action selection. From a human-robot interaction (HRI) perspective, emotions may communicate agent state (i.e. transparency) or create empathy and enhance user investment. This can help robots and agents transition to domestic environments in the forthcoming years.

To this end we first need plausible models of emotion generation in RL agents (i.e. Markov Decision Process (MDP)-based agents), which is the topic of this paper. Previous work on emotion elicitation in RL agents focussed on model-free learning [1]. However, important emotions like hope and fear are anticipatory, i.e. they require explicit forward simulation. Moreover, forward simulation also allows the agent to explain which anticipated event caused the emotion, i.e. to aid transparency. This work introduces the first anticipatory models of hope and fear in a RL agent. In particular, we show how hope and fear can be efficiently estimated from the best and worst forward traces. Our results show the plausibility of these signals, for example in the game *Pacman*.

2 Method

Models are grounded in the psychological Belief-Desire Theory of Emotion (BDTE) [2]. According to BDTE, hope and fear originate when the *belief* $b(s) \in [0, 1]$ about a state is smaller than 1, while the *desirability* $d(s) \in \mathbb{R}$ is larger or smaller than 0, respectively. We estimate the belief from the path probability towards it, and a positive or negative desirability from the positive or negative temporal difference (TD) error, respectively (see full paper for details). We write s, a, r, γ for state, action, reward and discount parameter, $P(\cdot|s, a)$ for the state transition function, $\pi(s, a)$ for the policy, and $V(s)$ for the state value. Moreover, let a trajectory of depth d from the current node s_0 be given by $g_d = \{s_0 a_0 s_1 a_1 ... s_{d-1} a_{d-1} s_d\}$. The fear in the current node s_0 is then estimated as:

[1]The full paper has been published in: Moerland TM, Broekens J, Jonker CM. Fear and Hope Emerge from Anticipation in Model-Based Reinforcement Learning. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* 2016, pages 848-854.

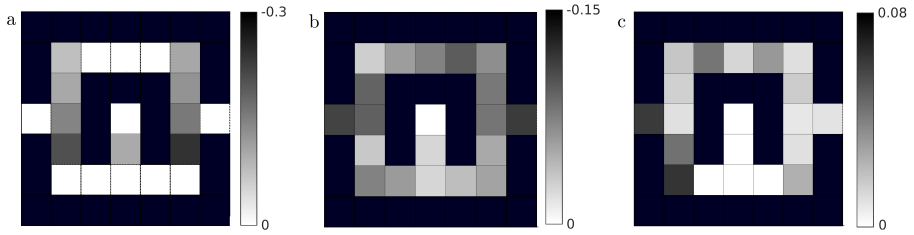


Figure 1: Pacman. a) Fear per location for a ghost *below* and ϵ-greedy(0.10) policy. b) Fear per location for *no* ghost and ϵ-greedy(0.10) policy. c) Hope per location for *no* ghost and softmax(τ=10) policy.

$$F(s_0) = \min_{s'} \left[b(s'|s_0) \times d(s'|s_0) \right]$$
$$= \min_{s'} \sum_{d, g_d | s_d = s'} \left[\prod_{t=0}^{(d-1)} \pi(s_t, a_t) P(s_{t+1}|s_t, a_t) \cdot \left(\left[\sum_{t=0}^{(d-1)} \gamma^t r(s_t, a_t, s_{t+1}) \right] + \gamma^d V(s_d) - V(s_0) \right) \right]^- \quad (1)$$

where $[\cdot]^-$ denotes the negative part. We effectively identify the successor node with highest product of path probability and TD, summing over all possible paths towards it. The model for hope is similar, replacing the $\min$ for a $\max$ and the negative part for the positive. Traces are identified by integrating a Monte Carlo Tree Search procedure (UCT) into the model-based RL architecture.

3 Results

We evaluate our models in *Pacman* (Figure 1). Pacman starts from the center-top, needs to reach the center ($r = +1$), but is chased by a ghost that starts from the center ($r = -1$ for a capture). Pacman can see whether a ghost exists in each cardinal direction, but not the distance to it. We train until Pacman has estimated a stable policy and world model, and then evaluate the plausibility of Pacman's emotions in different scenarios (Figure 1). Figure 1a shows fear when a ghost is seen below. We see how Pacman smoothly learned to be most afraid near the corridor bottoms, as the ghost must then be close. Figure 1b shows fear when Pacman does *not* observe a ghost in any direction. At nearly all locations Pacman fears a capture, but this is most prominent just before the corners (e.g. at the most left and right). Pacman start to feel safer when he approaches the center.

Finally, Figure 1c shows hope when Pacman does not see a ghost. Although we expected Pacman to be hopeful near the center, it turns out to be most hopeful just before the corners, hoping to step around it and still not see the ghost. This implies a jump in the value function, and happens on both sides of the top corridor, at the bottom of the left corridor, and also on the far left (note how Pacman has developed a left-wing tactic). Altogether, these unexpected results illustrate plausibility of the signals, as Pacman identified locations where things might change for better or worse.

4 Conclusion

This paper introduced the first anticipatory models of hope and fear in a RL (i.e. MDP-based) agent. Our emotions emerge from the agent's functionality and a (very sparse) reward signal, without the need for any pre-wired (and ad-hoc) solutions. Future work includes *using* the elicited emotions, for example to benefit learning, e.g. by biasing action selection, or to benefit human-agent/robot interaction.

References

[1] Joost Broekens, Elmer Jacobs, and Catholijn M. Jonker. A reinforcement learning model of joy, distress, hope and fear. *Connection Science*, pages 1–19, 2015.

[2] Rainer Reisenzein. Emotional experience in the computational belief–desire theory of emotion. *Emotion Review*, 1(3):214–222, 2009.

Design of a support agent for habit formation and breaking: extended abstract*

Pietro Pasotti M. Birna van Riemsdijk Catholijn M. Jonker

Interactive Intelligence, Delft University of Technology

1 Motivation

Habits are goal-independent *learned associations between responses and features of performance contexts*. Man, then, is a creature of habit. While people display a fascinating variety of behaviours even across relatively simple domains, from day to day most people are fixed in their ways. Carrying out habitual activities is mostly unproblematic and even desirable. However at times unforeseen circumstances make our habitual choices unavailable or their outcomes undesirable. Other times we wish to adopt or break a habit, and both are difficult enterprises. While many of us normally have little or no difficulty in dealing with these challenges, the actual amount of nuisance is subjective. To some, even small disruptions of daily routines may cause anxiety and distress, whereas for others, such as people suffering from depression, breaking habits can be beneficial. We take the first steps in developing a concrete implementation of *HabInt*, a Socially Adaptive Electronic Partner ([2]) that supports its user in trying to adopt, break or maintain habitual behaviours. In particular, we describe its user model's knowledge structures.

2 Habit support agent: what and how

The core functions of *HabInt* require it to capture mainly two aspects of its user: what is her current behaviour, and what she would like her behaviour to be. That is, the *actual* and *desired behaviours* of the user. These are then matched to identify the situations where an intervention is called for. Finally, *HabInt* determines what to do to support the user, thereby closing the interaction loop (see Figure 1).

To support the user's habits, *HabInt* must first represent the *activities* that are habitual to her and the *goals* that those activities usually involve [3]. These can be specified down to the level of the individual *actions* that the user must execute to achieve her goals. Additionally, *HabInt* can understand the motives behind adopting a goal and those behind the choice of the activities and actions that are supposed to achieve them. This is achieved by modelling what *values* promoted and demoted by the actions the user

*The full paper was presented as: Pasotti P, van Riemsdijk MB, Jonker CM. *Habit formation and breaking: a support agent design as an extended mind.* at COIN++@ECAI2016. *To appear* in the postproceedings.

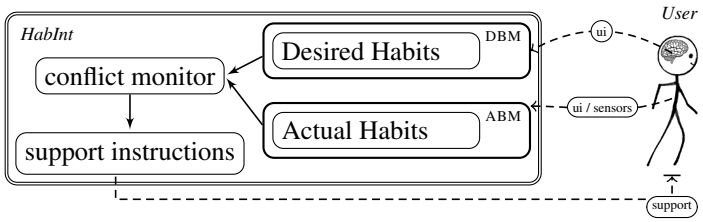


Figure 1: *HabInt* high-level architecture.

takes. All information pertaining to the user's behaviours, as they currently are, is stored in the Actual Behaviour Model (ABM).

Secondly, *HabInt* keeps track of the intentions to change the user has expressed. Hence, *HabInt* knows what needs to change in the user's actual behaviour for the behaviour to become *ideal*, that is, better suited to the user's desires. This information is stored in the Desired Behaviour Model (DBM).

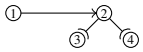
The information contained in the DBM is treated by *HabInt* as having a normative role: the agent's goal is then to help the user to conform to the rules she has set to herself. The user also needs to tell *HabInt* how he wishes to be supported, i.e. what should the agent do in case an undesired behaviour is occurring or predicted to occur. This way *HabInt*'s proactiveness is limited to only those actions that are explicitly requested.

3 Example

Suppose the user needs support with forming a habit concerning two activities: going from home to work and waking up. The first step is to explain those to *HabInt*, so it knows what is going on. The goals she wishes to achieve, in her own words, are *'be at work'*, ②, and *'be awake'*, ①, respectively. Furthermore, *'go by bike'* and *'go by car'*, are names of activities whose corresponding declarative goals are

'be at work' ∧ *'biked to work'* = ③ and *'be at work'* ∧ *'drove to work'* = ④

It has also recorded how going by bike/by car are ways of going to work, but going to work seems not to be a way to do something else. We represent the 'is a way of' relation by drawing ③ and ④ as child nodes of ② (cf. the following figure).



If the user mentions that waking up requires having set the alarm, an action α achieving ② (i.e. with ② as postcondition) is specified, which requires *'alarm set'* to be true. Then, α is linked to ①, to represent how achieving ① involves doing α.

HabInt's world and user models are explicitly represented, and readily understandable by the user, as they store information using the user's own vocabulary of goals, values, activities and actions.

4 Conclusion and future work

The knowledge structures we present model habits on the basis of a representation of the goals, activities and actions relevant to the user given her concerns. These concepts can be linked to personal values, essential for helping the user to choose desired behaviours that are in line with her motivations [1]. We advance the notion of a desired behaviour model as the basis for supporting a user's habit change.

A key feature of *HabInt* is that it adheres strictly to the user's vocabulary for expressing goals, activities, actions and values, and that the fine-grainedness of the behaviour model and of the actions involved in the user's activities are tailored to the user's needs.

In future work we intend to investigate relations to deontic logics and extend our model with temporal information and qualitative reasoning. Also in future work we plan on describing the conflict monitor and the generation of support action, implement the whole agent and perform user studies to prove its validity.

References

[1] T.L. van der Weide, F. Dignum, J.-J. Ch. Meyer, H. Prakken, and G.A.W. Vreeswijk. Practical reasoning using values: Giving meaning to values. In *Argumentation in Multi-Agent Systems*, volume 6057 of *Lecture Notes in Computer Science*, pages 79–93. Springer, 2009.

[2] M.Birna van Riemsdijk, Catholijn M. Jonker, and Victor Lesser. Creating socially adaptive electronic partners. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, 2015.

[3] Wendy Wood and Dennis Rünger. Psychology of habit. *Annual Review of Psychology*, 2015.

The role of values

Klara Pigmans, Huib Aldewereld, Neelke Doorn, Virginia Dignum

August 29, 2016

Accepted for the workshop Coordination, Organisations, Institutions and Norms (COIN) 2016, during ECAI 2016.

1 Abstract

Decision-making processes in policy making with multiple-stakeholders can be complex because of the technically advanced systems that are decided upon, e.g. water management systems. But the complexity is often also caused by social dependencies and interests at stake. This can lead to processes in which stakeholders firmly defend their interests, resulting in behavior that is so defensive that it severely blocks the process. For this reason, policy makers are trying to find a different approach.

Research suggest that making values explicit during the process could influence the process by increasing the social acceptance of the outcomes [3, 4, 6]. Still, the role of values in multi-stakeholder decision-making processes has not been described in clear terms. In this research we present a conceptual model as the first step to understand the role of values in these processes. In this model we identify the relevant concepts and their relations.

In the field of normative multi-agent systems, values have been discussed by [1] and [2]. We elaborate on this discussion by describing the explicit use of values in multi-stakeholder decision-making processes for policy making.

Based on [7], we identified the following concepts as part of a multi-stakeholder decision-making process to come to a policy:

- Values: enduring beliefs that certain end-states are personally or socially preferable to an opposite end-state [8], e.g. justice vs injustice.
- Value conceptions: interpretation of the value, the arguments for how the value should be interpreted in practice [5]
- Context: physical and social setting in which one lives or in which something happens or develops, including culture and people and institutions one interacts with.
- Vision: institutional objective
- Collective decision-making process: the process of getting to a decision, which can take decades and can include meetings, discussions, deliberations, one-to-one meetings, newsletters, informative events and compensation negotiations.
- Collective action: action that has been agreed upon by the stakeholders at the end of the decision making process.
- Norms: regulate the behavior of gatens by describing the actions they must (or must not) execture in specific situations.
- Agent: represents stakeholders or groups of stakeholders
- Individual action: action taken by the agent based on its personal norms

Relating these concepts to one another to depict the multi-stakeholder process, results in the conceptual model depicted in Figure 1. The conceptual model has both an individual structure, describing the concepts that are relevant for the individual agents, as well as a collective structure representing the collective

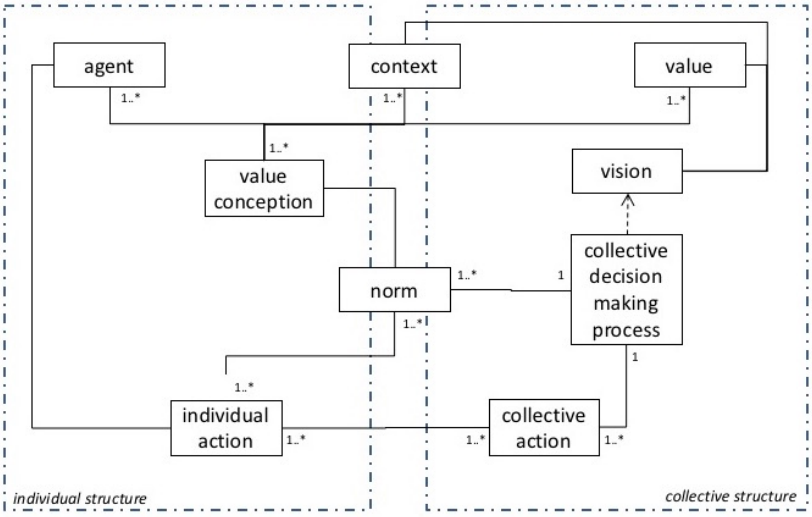


Figure 1: Conceptual model

concepts of the decision-making process. The concepts 'context' and 'norms' are part of both structures and connect the two structures.

This conceptual model is the first step to identify the relvant concepts and to understand the relation between these concepts. The next step will be to add institutions to the model, including roles that agents have and rules that are relevant. This will support us in developing a formal model of values in multi-stakeholder decision-making proceses, which we need in order to be able to reason about such processes.

References

[1] K. Da Silva Figueiredo and V. Torres da Silva. Identifying conflicts between norms and values. In *Coordination, Organizations, Institutions, and Norms in Agent Systems IX*. Springer International Publishing., 2013.

[2] F. Dechesne, G. Di Tosto, V. Dignum, and F. Dignum. No smoking here: values, norms and culture in multi-agent systems. *Articficial intelligence and law*, 21:79–107, 2013.

[3] N. Doorn. Governance experiments in water management: From interests to building blocks. *Science and Engineering Ethics*, 22(3):755–774, 2016.

[4] L. Glenna. Value-laden technocratic management and environmental conflicts: The case of the new york city watershed controversy. *Science, Technology & Human Values*, 35(1):81–112, 2010.

[5] M. Jacobs. Sustainable development as a contested concept. In A. Dobson, editor, *Fairness and futurity*. Oxford university press, 1999.

[6] I. Oosterlaken. Applying value sensitive design (VSD) to wind turbines and wind parks: An exploration. In *Science and Engineering Ethics*, volume 21. Springer, 2015.

[7] K. Pigmans, N. Doorn, H. Aldewereld, and V. Dignum. Decision-making in water governance: from conflicting interests to shared values. In M.J. Van den Hoven and L. Asveld, editors, *Responsible innovation 4*. Springer, 2017. forthcoming.

[8] M. Rokeach. *The nature of human values*. New York: Free Press, 1973.

The Game of Reciprocation Habits[1]

Gleb Polevoy Mathijs de Weerdt Catholijn Jonker

Delft University of Technology, email: g.polevoy@tudelft.nl

Abstract

People have reciprocal habits, almost automatically responding to others' actions. A robot who interacts with humans may also reciprocate, to come across natural and be predictable because of the reacting. We aim to advise which habits are utility-efficient in these interactions. To this end, given a model for reciprocation behavior with parameters that represent habits, we define a game that describes which habit one should adopt to increase the utility of the process. This paper concentrates on two agents. The used model defines that an agent's action is a weighted combination of the other's previous actions (reacting) and either i) her innate kindness, or ii) her own previous action (inertia). We define a game where an agent may choose her habit, which is either her reciprocation attitude (i or ii), or both her reciprocation attitude and the weight of the combination. We characterize the Nash equilibria of these games and their efficiency. We find that the less kind agents should adjust to the kinder agents to improve both their own utility as well as the social welfare. This constitutes advice on improving cooperation and explains real life phenomena in human interaction, such as the societal benefits from adopting the behavior of the kindest person, or becoming more polite as one grows up.

1 Introduction

Interaction is central in human behavior, e.g., at school, in file sharing over networks, and in business. In particular, people, countries and companies tend to reciprocate, i.e., react on the past actions of others [1, 3]. Virtual assistants also need to be reciprocal in order to be credible. We can help people and artificial agents obtain more from the interaction by supporting decision how to reciprocate. This requires a model that is amenable to analytical analysis and allows predicting interaction.

Some studies explain the reasons for reciprocation, such as the evolutionary model of Axelrod [1]. Other works assume reciprocation and analyze what happens in an interaction where being reciprocal pays off [8]. We use a model from [6], which considers a lengthy reciprocal interaction of extent varying with time where agent choose their own attitude to reciprocation in order to receive more and do less. That work represents an action by a single number, representing the action's value; the bigger the number, the better the action influences its recipient. Assume all the agents act on times $T = \{0, 1, \dots\}$. Following the example of an arms race [2] or arguing spouses [4], they define an action of agent i on agent j at time t by a linear combination of the action of agent j on i at time $t-1$ and of either one's own kindness or own action at time $t-1$ (mental inertia). The kindness k_i is a parameter defining the internal inclination to act. This linear combination is determined by the agent's reciprocation coefficient r_i. The first reciprocation attitude is called *fixed*, while the second attitude is called *floating*.

A reciprocation process converges, and the actions in the limit are found in [6]. Aiming to provide decision support we analyze how the agents can strategically set their own reciprocation attitude and coefficient so as to maximize their own utility in the reciprocation process. We find the social welfare, which is the sum of the individual utilities. Assuming people strategically choose each action is unrealistic, since people usually act on habits [5], so a real choice consists of choosing a habit of reciprocation.

Setting habits resembles bounded rationality, like the procedures of choice [9, Chapter 2], but with a rational step. This also reminds metagames, such as Rubinstein's [9, Chapter 8] *machine game*, where an agent chooses a well-paying and simple strategy. We, however, consider equally simple habits, and anyway, reciprocation can not be modeled by a finite automaton.

[1] The full paper has been published in [7].

2 Reciprocation Attitude Game

We define the utility of an agent as the received action minus a constant times the performed action and analyze the one-shot game where the strategies are own reciprocation attitudes (*fixed* or *floating*). The strategy profile defines a reciprocation process, and the limits of the above defined utilities are the utilities in the game.

As a preparation to analyzing the game, we show how an agent can maximize her own utility by setting her reciprocation coefficient or her reciprocation attitude. To prepare for analyzing the efficiency when the agents act egoistically, we find the maximum possible social welfare. We see that if acting is easy, relatively to receiving an action, than the less kind agent should align to the kinder one by being *floating*, in order to maximize both own utility and the social welfare.

We characterize the Nash equilibria (NE) of the reciprocation attitude game and study the most and the less efficient NE, giving the price of anarchy and stability. We conclude that when acting is easy, all the NE are socially optimal, so regulation is not needed. The NE are also achievable by best responding.

3 Reciprocation Attitude and Coefficient Game

We now analyze the game when an agent simultaneously sets her own reciprocation attitude and coefficient r_i. We characterize the NE and their efficiency. Here, unlike the case of reciprocation attitude game, when acting is easy, then in addition to a socially optimal NE, where exists an NE with the social welfare of $\frac{1}{2}\left(\frac{k_1}{k_2}+1\right)$ from the maximum possible social welfare. This means allowing more choice may reduce the efficiency, requiring regulation.

The original reciprocation model is defined for a general number of agents and the convergence still holds. This allows for similar games with $n \geq 2$ agents.

This work demonstrates that if acting is not too hard, then following the kindest makes the individual and the society thrive, as we often observe. This includes people becoming more polite with experience.

One of the interesting directions for future work is the reciprocation coefficient and/or attitudes changing throughout the interaction. To summarize, we predict reciprocation phenomena and advise how to behave.

References

[1] R.M. Axelrod. *The evolution of cooperation*. Basic books. Basic Books, 1984.

[2] William J. Dixon. Reciprocity in united states-soviet relations: Multiple symmetry or issue linkage? *American Journal of Political Science*, 30(2):pp. 421–445, 1986.

[3] Armin Falk and Urs Fischbacher. A theory of reciprocity. *Games and Economic Behavior*, 54(2):293 – 315, 2006.

[4] John Gottman, Catherine Swanson, and James Murray. The mathematics of marital conflict: Dynamic mathematical nonlinear modeling of newlywed marital interaction. *Journal of Family Psychology*, 13:3–19, 1999.

[5] D. Kahneman. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.

[6] G. Polevoy, M.M. de Weerdt, and C.M. Jonker. The convergence of reciprocation. In *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '16. Intern. Foundation for Autonomous Agents and Multiagent Systems, 2016.

[7] G. Polevoy, M.M. de Weerdt, and C.M. Jonker. The game of reciprocation habits. In *Proceedings of the 2016 European Conference on Artificial Intelligence*, volume Volume 285: ECAI 2016 of *Frontiers in Artificial Intelligence and Applications*, pages 417–425, 2016.

[8] Matthew Rabin. Incorporating fairness into game theory and economics. *The American Economic Review*, 83(5):pp. 1281–1302, December 1993.

[9] Ariel Rubinstein. *Modeling Bounded Rationality*, volume 1. The MIT Press, 1 edition, 1997.

On the Need for a Coordination Mechanism to Guarantee Task Completion in a Cooperative Team[1]

Chris Rozemuller [a] Koen V. Hindriks [a] Mark A. Neerincx [a]

[a] *Delft University of Technology, Mekelweg 4 Delft*

Abstract

To design good cooperative team members in robotics it is important to know what coordination mechanisms are required. Our approach to explore the need for a coordination mechanism is based on a systematic methodology to identify team coordination requirements. We show that a team combined of robots that each individually can solve a task not always is able to guarantee task completion as a team. In these cases some mechanism for coordination is required and we formally identify various problem classes that impose different requirements. We introduce a formal task model and distinguish between no, implicit and explicit coordination mechanisms. This model is used to study which mechanisms guarantee task completion. It allows us to prove some empirical findings reported in the literature such as that a simple foraging task does not require coordination.

1 Introduction

Our interest is in designing effective robot interaction. To do so, it is important to determine the need for coordination mechanisms in robot teams. Teamwork is crucial to many applications of robotics and foraging tasks have become the golden standard to empirically benchmark teamwork in such applications, with an extensive amount of results [1]. An interesting common observation is that for completing a foraging task only minimal coordination is required, however it remains unclear what the minimum is. In this study, we formally determine the *coordination requirements that guarantee task completion* for a range of foraging tasks. Moreover, we will do so by introducing a formal model that can represent foraging tasks and that is fully compatible with the joint activity testbed proposed in [2] used for experimental and simulated studies of human and/or robot teamwork.

We introduce a formal model that allows us to investigate which conditions require coordination of agents to ensure task completion for increasingly complex foraging problems. To systematically explore whether, and, if so, which coordination mechanisms are required for task completion we use a four step methodology. The idea is to investigate in step (1) which tasks can be completed by a single agent. With this step we restrict ourselves to problems that are single agent solvable so coordination is not imposed by the task. In steps 2-4 it is incrementally checked whether respectively (2) *no coordination*, (3) *implicit coordination*, or (4) *explicit coordination* is sufficient to guarantee task completion. Also we investigate what level of team-awareness needs to be implicitly embeded in agent's the decision function to guarantee task completion.

2 Formal coordination model

The model that we introduce consists of an environment that is modeled as a *network of locations* with *resources* distributed over the network. Locations on a network can either be regular or sink locations.

[1] The full paper has been published in *In Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on (pp. 2817-2822). IEEE.* Please refer to the full paper for all model definitions and proofs

Sinks are special locations where resources are irreversibly consumed. A *task* represents a desired resource distribution, and an *agent* model represents agents with a resource inventory that can perform actions *take*, *drop*, *move* or *skip*. We specify three execution semantics that represent our coordination mechanisms. *no coordination*: all agents in parallel only observe their local state, then decide and then act. *turn-based coordination*: same as no coordination, but only one agent can perform an action at the time. *communication*: same as no coordination, but all observations are shared before agents make a decision. We also introduce three decision functions that assume different levels of team awareness: greedy, careful and team-aware. A greedy agent has the assumption it is alone, a careful agent takes in account that other agents could also make progress, and a team-aware agent also takes in account that symmetry needs to be broken sometimes. Finaly we define a notion of task progression - based on spatial distance, task distance, inventory and exploration - that is used to establish various results.

3 Results

Initially we assume that actions have no preconditions. First we show that all feasible tasks in our model can be achieved by a *single greedy* agent, i.e. the tasks do not inherently require coordination. We show that only for a subset of these tasks a team of *multiple greedy* agents can guarantee task completion. In this set of tasks there can be no ambiguity about which resources need to be manipulated, For example: all resources need to be removed. Also we give an example of a task that cannot be completed by multiple careful agents, because agents keep performing identical actions. For example: two agents dropping a resource where only is required, so both pick up again. We show that the remaining tasks can be successfully completed by *multiple team-aware* agents that implicitly coordinate their actions. Since they will perform random skips these agents will eventually break symmetry. By extension this proofs that careful agents with a turn based coordination mechanism are also able to solve all tasks. As a result, it follows that the only tasks in this model that require explicit coordination are tasks where agents need to perform actions that depend on the actions of other agents. Finally, we show that multiple team-aware agents that use explicit coordination can solve at least some of these tasks.

4 Conclusion and Future Work

This paper introduces a formal model for studying coordination mechanisms required to guarantee task completion. Although inspired by search and retrieval tasks it can also model other planning problems. A simple methodology is introduced to explore requirements on coordination. We show that garbage cleaning tasks do not require coordination and that implicit coordination mechanisms are sufficient to guarantee task completion if action preconditions are minimal.

We have only been able to partially address the many questions related to cooperative teams in this paper. There are many open problems that we would like to study formally in future work. We mention only one open question that we think is particularly interesting: do strongly-cooperative tasks always require explicit coordination, i.e., do tasks that require more than one agent also require explicit coordination to be able to guarantee task completion.

Our work has focused on *robustness* of teamwork but we are also interested in the *resilience* of a team to cope with unexpected events and failures. In future work we want to investigate which mechanisms can ensure success when communication may fail for various task environments. Also we want to investigate the coordination requirements of a wider range of tasks that can be formulated with linear temporal logic (LTL)

References

[1] A. Farinelli, L. Iocchi, and D. Nardi. Multirobot systems: a classification focused on coordination. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 34(5):2015–2028, Oct 2004.

[2] Matthew Johnson, Catholijn Jonker, Birna Van Riemsdijk, Paul J Feltovich, and Jeffrey M Bradshaw. Joint activity testbed: Blocks world for teams (bw4t). In *Engineering Societies in the Agents World X*, pages 254–256. Springer, 2009.

Enhancements for Real-Time Monte-Carlo Tree Search in General Video Game Playing[1]

Dennis J. N. J. Soemers Chiara F. Sironi Torsten Schuster
Mark H. M. Winands

Games and AI Group, Department of Data Science and Knowledge Engineering, Maastricht University, The Netherlands

General Video Game Playing (GVGP) [6] is a field of Game AI where the goal is to develop agents that are able to play a variety of real-time video games that are unknown in advance. It is closely related to General Game Playing (GGP) [4], which focuses on abstract games instead of video games. The wide variety of games in GGP and GVGP makes it difficult to use domain-specific knowledge, and promotes the use of generally applicable techniques. There are two main frameworks for GVGP. The first framework is the Arcade Learning Environment [1] for developing agents that can play games of the Atari 2600 console. The second framework is GVG-AI [8], which can run any real-time video game described in a Video Game Description Language [10]. This framework is used in the GVG-AI Competition [9]. Past GVG-AI competitions only ran a single-player Planning Track, where agents were ranked based on their performance in single-player games. Since 2016, a second track exists, where agents play adversarial and cooperative two-player games. This research focuses on the single-player Planning Track of the GVG-AI competition.

Monte-Carlo Tree Search (MCTS) [3, 5] is a popular technique in GGP [2] because it does not rely on domain-specific knowledge. It is a best-first search algorithm that gradually builds up a search tree and uses Monte-Carlo simulations to approximate the value of game states. MCTS has also performed well in GVGP in 2014 [8] – the first year of the GVG-AI competition – but was less dominant in 2015 [9]. In the paper we have discussed and evaluated eight enhancements for MCTS to improve its performance in GVGP: *Progressive History*, *N-Gram Selection Technique*, *Tree Reuse*, *Knowledge-Based Evaluations*, *Deterministic Game Detection*, *Breadth-First Tree Initialization*, *Novelty-Based Pruning*, and *Loss Avoidance*. The first five are known from the existing MCTS literature, which are either extended or introduced in the context of GVGP, and the last three are novel enhancements to MCTS.

Experiments showed that most enhancements significantly increase the average win percentage over sixty different games when added individually to MCTS. All enhancements combined increase the win percentage of our basic MCTS implementation from 31.0 ± 1.2 to 48.4 ± 1.5. This final performance is relatively close to the win percentage of the winner at the IEEE CEEC 2015 conference; YBCRIBER, with a win percentage of 52.4 ± 1.3. An optimized version of our MCTS implementation, called MAASTCTS2, has won subsequently the GVG-AI Single-Player Planning Championship 2016 (out a total of 20 bots). Although not designed for two-player games, MAASTCTS2 also participated in the GVG-AI Two-Player Planning Championship 2016, where it finished second (out of 13).

For future research, the overall performance of the agent can also likely be improved by incorporating more features that are commonly seen among the top entries in past competitions, such as the use of influence maps [7]. Finally, some of the new enhancements for MCTS, such as Loss Avoidance and Novelty-Based Pruning, can be evaluated in domains other than GVG-AI.

[1]The full version of this paper is published in: *IEEE Computational Intelligence and Games Conference (CIG 2016)*, pp. 436-443, IEEE, 2016.

Acknowledgement

This work is partially funded by the Netherlands Organisation for Scientific Research (NWO) in the framework of the project GoGeneral, grant number 612.001.121.

References

[1] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling. The Arcade Learning Environment: An Evaluation Platform for General Agents. *Journal of Artificial Intelligence Research*, 47:253–279, 2013.

[2] Y. Björnsson and H. Finnsson. CadiaPlayer: A Simulation-Based General Game Player. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(1):4–15, 2009.

[3] R. Coulom. Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search. In H. J. van den Herik, P. Ciancarini, and H. H. L. M. Donkers, editors, *Computers and Games*, volume 4630 of *Lecture Notes in Computer Science*, pages 72–83. Springer Berlin Heidelberg, 2007.

[4] M. Genesereth, N. Love, and B. Pell. General Game Playing: Overview of the AAAI Competition. *AI Magazine*, 26(2):62–72, 2005.

[5] L. Kocsis and C. Szepesvári. Bandit Based Monte-Carlo Planning. In J. Fürnkranz, T. Scheffer, and M. Spiliopoulou, editors, *Machine Learning: ECML 2006*, volume 4212 of *Lecture Notes in Computer Science*, pages 282–293. Springer Berlin Heidelberg, 2006.

[6] J. Levine, C. B. Congdon, M. Ebner, G. Kendall, S. M. Lucas, R. Miikulainen anad T. Schaul, and T. Thompson. General Video Game Playing. In S. M. Lucas, M. Mateas, M. Preuss, P. Spronck, and J. Togelius, editors, *Artificial and Computational Intelligence in Games*, volume 6 of *Dagstuhl Follow-Ups*, pages 77–83. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013.

[7] I. Millington and J. Funge. *Artificial Intelligence for Games*. Morgan Kaufmann, 2nd edition, 2009.

[8] D. Perez, S. Samothrakis, J. Togelius, T. Schaul, S. M. Lucas, A. Couëtoux, J. Lee, C.-U Lim, and T. Thompson. The 2014 General Video Game Playing Competition. *IEEE Transactions on Computational Intelligence and AI in Games*, 8(3):229–243, 2016.

[9] D. Perez-Liebana, S. Samothrakis, J. Togelius, S. M. Lucas, and T. Schaul. General Video Game AI: Competition, Challenges and Opportunities. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pages 4335–4337. AAAI Press, 2016.

[10] T. Schaul. A Video Game Description Language for Model-based or Interactive Learning. In *Proceedings of the IEEE Conference on Computational Intelligence in Games*, pages 193–200, Niagara Falls, 2013. IEEE Press.

Rapid Adaptation of Air Combat Behavior[1]

Armon Toubman [a] Jan Joris Roessingh [a] Pieter Spronck [b]
Aske Plaat [c] Jaap van den Herik [c]

[a] *Netherlands Aerospace Centre NLR, P.O. Box 90502, 1006 BM Amsterdam*
[b] *Tilburg University, P.O. Box 90153, 5000 LE Tilburg*
[c] *Leiden University, P.O. Box 9500, 2300 RA Leiden*

Abstract

Fighter pilots train in simulations to practice real-life missions. However, writing the behavior for the virtual opponents is a costly task. In this paper, we use machine learning to provide the virtual opponents with the ability to adapt their behavior to that of the human trainees. This has two advantages: (1) we support the experts that create the behavior models, and (2) the trainees can be continuously challenged. To this end, we adapted the dynamic scripting technique to generate finite-state machines. The effectiveness of the method is shown through automated air-to-air combat simulations.

1 Introduction

Fighter pilots require continuous training to maintain their level of readiness. Training is mainly done in live exercises where the pilots engage in mock fights. However, due to the limited availability of jets and supporting personnel, a substantial part of training has moved to simulations.

Air combat training simulations require virtual agents that aid or oppose the trainees. While other pilots could be employed to control these agents, they might not be experts on emulating the necessary enemy tactics. Therefore, an AI solution is preferred. Currently, however, training developers still need to write new behavior for each new opponent they need to present, which is a costly undertaking.

We aim to alleviate the problem of writing behavior for the virtual agents in training simulations using machine learning. By automating the generation of behavior, we accomplish two main goals: (1) we support the subject matter experts (SMEs) that develop behavior for the virtual agents, and (2) we are able to challenge trainee fighter pilots with more interesting opponents. However, these goals require a machine learning method that is able to produce behavior models that are readable by the experts (so they can easily inspect and edit the behavior). Furthermore, the method has to be able to provide adaptation rapidly and online (i.e., during training sessions).

In this paper, we present a method for generating air combat behavior for virtual entities in training simulations. This method uses a version of dynamic scripting (DS) (a rule-based reinforcement learning technique) [2] that has been modified to work with finite-state machines (FSMs) (Section 2). We show using automated air-to-air combat simulations that our method is capable of rapidly adapting the behavior of virtual agents to that of their opponents (Section 3).

2 Method

Traditional DS selects pre-written rules from a rule base to form a script. This script is used to control an agent in some environment. Iterative feedback from the environment leads to changes in the probabilities that rules will be selected the next time a script is formed.

Scripts are relatively simple and readable behavior models. However, scripts only provide reactive behavior, and the behavior resulting from combinations of rules is hard to verify. Another behavior

[1] The full paper has been published in *ECAI 2016 - 22nd European Conference on Artificial Intelligence* (IOS Press).

model, the FSM, allows more structured behavior while remaining readable by experts. This is also the model used in Smart Bandits [1], one of the softw
opponents in Royal Netherlands Air Force (RNLAI

We modified the DS technique to generate FSMs. States and transitions in FSMs can be easily translated to rules: e.g., for a *Patrol* state, *[if my state is Patrol, fly between points A and B]*. For each state and transition, variations can be made that are interchangeable with the original states and transitions: e.g., a variation on the previous example, *[if my state is Patrol, fly between points C and D]*. This way, a rule base can be filled with rules that are the constituent parts of an FSM, including variations. We altered the DS algorithm such that for each state and transition, only one of its variations can be selected. This way, DS will always generate a completely valid FSM, while still being able to vary its contents.

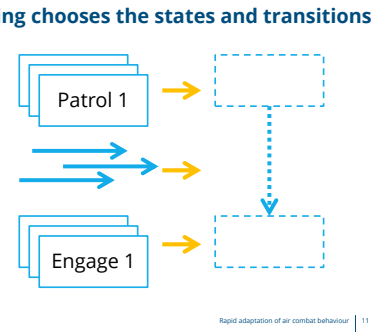


Figure 1: We altered DS to generate FSMs by selecting states and transitions.

3 Experiments and results

We used AI-versus-AI experiments to validate our method. We conducted four experiments in which two adaptive agents (AAs) had to defeat two non-adaptive agents (NAAs). The results of each experiment were averaged over ten runs. In our experiments, we used two tactics (T1 and T2). T1 was based on a tactic originally designed for Smart Bandits. T2 was newly designed to defeat T1.

In experiments 1 and 2, the NAAs used T1 and T2 respectively. The AAs were each given a rule base that contained T1 in rule form, along with several expert-made variations on the states and transitions in that tactic. As a baseline, another pair of NAAs using T1 fought the NAAs.

The results of the first two experiments are shown in Figure 2. In both experiments, the AAs reached (near-)peak performance after around 15 episodes.

Experiments 3 and 4 further demonstrate the AAs' adaptive capabilities. In these experiments, the AAs fought the NAAs using T2 after having adapted to the NAAs using T1, and vice versa.

At the point of the transfer, the win rates of the AAs quickly fall below 0.2, yet recover to >0.5 in 8 episodes. Peak performance is reached again after 35 and 28 episodes in experiments 3 and 4 respectively.

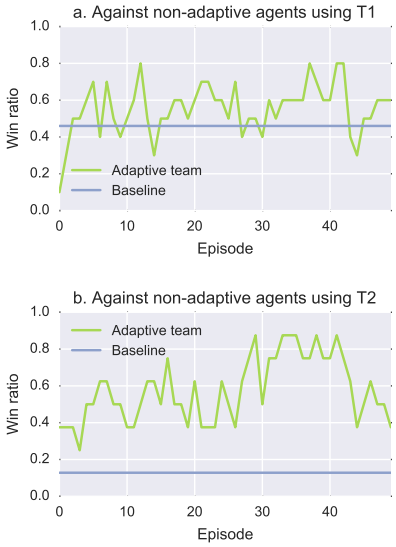


Figure 2: Win rates of the adaptive agents.

In conclusion, we have generated air combat behavior using machine learning in a way that is both transparent for the SMEs developing the virtual opponents for training simulations (i.e., with behavior in FSM form), and fast enough for use during training sessions. Human-in-the-loop experiments with RNLAF fighter pilots are currently underway.

References

[1] Jan Joris Roessingh, Robbert-Jan Merk, Pieter Huibers, Remco Meiland, and Roel Rijken. Smart Bandits in air-to-air combat training: Combining different behavioural models in a common architecture. In *21st Annual Conference on Behavior Representation in Modeling and Simulation*, Amelia Island, Florida, USA, 2012.

[2] Pieter Spronck, Marc Ponsen, Ida Sprinkhuizen-Kuyper, and Eric Postma. Adaptive game AI with dynamic scripting. *Machine Learning*, 63(3):217–248, 2006.

Network-Oriented Modeling and its Conceptual Foundations (extended abstract)[1]

Jan Treur

Vrije Universiteit Amsterdam, Behavioural Informatics Group
De Boelelaan 1081, 1081HV Amsterdam, The Netherlands
https://www.researchgate.net/profile/Jan_Treur http://www.few.vu.nl/~treur

To address the complexity of modeling the world's processes, over the years different strategies have been used. From these strategies separation assumptions are quite common in all scientific disciplines and have often turned out very useful. They traditionally serve as means to address the complexity of processes by some strong form of decomposition. This also holds for classical disciplines such as Physics, where, for example, for mechanical modeling for building construction only forces from objects on earth are taken into account and not forces from all other objects in the universe, that still do have some effects as well. For such cases within Physics such an isolation assumption may be a reasonable choice, but in how far is it equally reasonable to address complexity of human mental and social processes? Over the years within the Behavioural and Social Sciences also assumptions have been made in the sense that some processes can be studied by considering them as separate or isolated phenomena. However, within these human-directed sciences serious debates have occurred time and time again on such assumptions. Examples of such separation assumptions are mind versus body, cognition versus emotion, individual processes versus collective processes, non-adaptive processes versus adaptive processes, and earlier versus later (temporal separation). It can be questioned whether, for example, mind can be studied while ignoring body, or cognition while ignoring emotion, or sensory processing without action preparation.

Put more general, in how far are these traditional means to address complexity by separation still applicable if the complexity of human mental and social processes has to be addressed? Do we need to break with such traditions to be able to make more substantial scientific progress in this area addressing human processes? And, not unimportant, are there adequate alternative strategies to address human complexity? This is discussed in this paper, and it is pointed out that Network-Oriented Modeling can be considered an alternative way to address complexity. In this paper, the five separation assumptions mentioned above are discussed in more detail. Next, it is discussed how as an alternative, interaction in networks can be used to address complexity and a Network-Oriented Modeling perspective is discussed. Moreover, the need for a temporal dimension to address the dynamics is discussed, in particular to handle cyclic causal connections and realistic timing in mental and social processes, and the Network-Oriented Modeling approach based on temporal-causal networks presented in [7, 8] is briefly pointed out.

A conceptual representation of a temporal-causal network model in the first place involves representing in a declarative manner states (indicated by X and Y) and connections between them that represent (causal) impacts of states on each other, as assumed to hold for the application domain addressed. The states are assumed to have (activation) levels that vary over time. In reality not all causal relations are equally strong, so some notion of *strength of a connection* is used: *connection weights* $\boldsymbol{\omega}_{X,Y}$. Furthermore, when more than one causal relation affects a state, some way to *aggregate multiple causal impacts* on a state is used: *combination functions* $\mathbf{c}_Y(..,..)$. Moreover, as not every state has the same extent of flexibility, a notion of *speed of change* of a state is used for timing of processes: *speed factors* $\boldsymbol{\eta}_Y$. These three notions are covered by elements in the Network-Oriented Modeling approach based on temporal-causal networks, and are considered part of both a conceptual and numerical representation of a temporal-causal network model.

[1] The full paper is published in *Proceedings of the 8th International Conference on Social Informatics*, *SocInfo'16*, 2016.

Combination functions in general are similar to the functions used in a static manner in the (deterministic) Structural Causal Model perspective described, for example, in (Pearl, 2000), but in the Network-Oriented Modeling approach described here they are used in a dynamic manner, as is pointed out below briefly, and in more detail in [7, 8]. Combination functions can have different forms. The more general issue of how to combine multiple impacts or multiple sources of knowledge occurs in various forms in different areas, such as the areas addressing imperfect reasoning or reasoning with uncertainty or vagueness. For example, in a probabilistic setting, for modeling multiple causal impacts on a state often independence of these impacts is assumed, and a product rule is used for the combined effect; e.g., [2]. In the areas addressing modeling of uncertainty also other combination rules are used, for example, in possibilistic approaches minimum- or maximum-based combination rules; e.g., [2]. In another area, addressing modeling based on neural networks yet another way of combining effects is used often. In that area, for combination of the impacts of multiple neurons on a given neuron usually a logistic sum function is used: adding the multiple impacts and then applying a logistic function; e.g., [3, 4]. All of these types of combination functions can be used in temporal-causal networks as part of their conceptual representation. A list of such standard combination functions is available and a number of relevant properties of such combination functions have been identified (see [7, 8]). To design a temporal-causal network model a choice can be made from this list, but for specific purposes it is also possible to define other combination functions.

A conceptual representation of temporal-causal network model can be transformed in a systematic or even automated manner into a numerical representation of the network model, thus obtaining the following *difference* and *differential equation* for each state Y with impacts from states X_1, …, X_k:

$$Y(t+\Delta t) = Y(t) + \boldsymbol{\eta}_Y [\mathbf{c}_Y(\boldsymbol{\omega}_{X_1,Y} X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y} X_k(t)) - Y(t)] \, \Delta t$$

$$\mathrm{d}Y(t)/\mathrm{d}t = \boldsymbol{\eta}_Y [\mathbf{c}_Y(\boldsymbol{\omega}_{X_1,Y} X_1(t), \ldots, \boldsymbol{\omega}_{X_k,Y} X_k(t)) - Y(t)]$$

For modeling processes as adaptive networks, some of the parameters (such as connection weights) are handled in a similar manner, as if they are states. For more detailed explanation, see [7, 8].

Concerning the scope of applicability, it has been shown [7, 8] that any smooth dynamical system described as a state-determined system (or by a set of first-order differential equations) can also be modeled by temporal-causal networks, by choosing suitable parameters such as connection weights, speed factors and combination functions. In this sense it is as general as modeling approaches put forward, for example, in [1, 6]. To facilitate applications, dedicated software is available supporting the design of models in a conceptual manner, automatically transforming them into an executable format and performing simulation experiments. A variety of example models that have been designed illustrates the applicability of the approach in more detail, varying from network models for more complex mental processes to network models for social interaction processes (e.g., social networks) and integration of both types. In [8] a large number of such applications can be found, and pointers to even more of them.

References

[1] W.R. Ashby. *Design for a Brain*, Chapman and Hall, London (second extended edition), 1960. First edition, 1952.

[2] D. Dubois and H. Prade. Possibility Theory, Probability Theory and Multiple-valued Logics: A Clarification, *Annals of Mathematics and Artificial Intelligence* 32: 35–66, 2002.

[3] S. Grossberg. On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics*, 1: 319-350, 1969.

[4] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci. (USA)*, 81: 3088-3092, 1984.

[5] J. Pearl. *Causality*. Cambridge University Press, 2000.

[6] R.F. Port and T. van Gelder. *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press, 1995.

[7] J. Treur. Dynamic Modeling Based on a Temporal-Causal Network Modeling Approach. *Biologically Inspired Cognitive Architectures*, 16: 131–168, 2016.

[8] J. Treur. *Network-Oriented Modeling: Addressing Complexity of Cognitive, Affective and Social Interactions*. Understanding Complex Systems Series. Springer Publishing, 2016.

Using Distributed Representations to Disambiguate Biomedical and Clinical Concepts*

Stéphan Tulkens Simon Šuster Walter Daelemans

CLiPS, University of Antwerp, Prinsstraat 13, 2000, Antwerp

1 Introduction

Disambiguation of concepts is an important component of applications operating on clinical and biomedical text, in which the detection of different concepts might lead to physicians taking a completely different course of action. The word "cold", for example, might refer to a temperature, or the disease known as the *common cold*.

Compounding this problem is the fact that clinical text is, in general, noisier than text from other domains. Doctors from different wards will, for example, use different words to denote the same concept, depending on their specialization. Additionally, there is an absence of large volumes of annotated clinical text, even for English, which presents a problem for supervised approaches. For other languages, such as Dutch, there exist no such freely available resources.

A first step towards solving this problem could be the use of distributed representations. Where a more traditional word representation, such as a bag-of-words (BoW) representation, just indicates which words are present in a given unit of text, distributed representations encode semantic information by representing words as a function of the contexts in which they occur. Using this type of semantic information has been shown to improve performance on a variety of tasks when compared to more traditional BoW representations.

We hypothesize that these kinds of distributional representations are well-suited for Word Sense Disambiguation in the clinical and biomedical domain. We present a knowledge-based approach to Word Sense Disambiguation which combines definitions from the Unified Medical Language System (UMLS) with distributed representations. We test our hypothesis on the MSH-WSD [2], which is a well-known dataset for WSD (Word Sense Disambiguation) in the biomedical domain, and compare our approach to other state of the art approaches.

2 Approach

Our approach creates *concept vectors* from concepts defined in the UMLS ontology [1] by retrieving definitions of concepts from the ontology, and then replacing each word in every definition by the vector representation of that word. This creates an $M \times n$ matrix for each definition, where M is the dimensionality of the word vectors, and n the number of words contained in that definition. Following this, for each definition, we then obtain a single vector of dimensionality M by applying a columnwise sum to the matrix. Each concept can then be represented by a $M \times d$ matrix, where d is the number of definitions that a concept has in the UMLS. Finally, we apply a columnwise average to this matrix, thereby obtaining a single vector of dimensionality M which represents the combined meaning of all definitions for that concept, i.e. a *concept vector*.

For each abstract in the MSH-WSD, we first locate each ambiguous term through a simple lookup. For each located term in the abstract we create a vector representation by retrieving all words in a window

*Originally published in the proceedings of the 15th Workshop on Biomedical Language Processing (BIONLP), Berlin, 2016.

	med	mim	bio	MRD	2-MRD	0-step	2-step	r-step	UMLS::SenseRelate
Accuracy C	0.80	0.69	0.84	0.81	0.78	0.82	0.86	0.89	0.75
Accuracy U	0.72	0.63	0.75	-	-	-	-	-	-

Table 1: Results using constrained (C) and unconstrained (U) terms.

of size w surrounding the ambiguous term, and replacing the words by their vectors. The idea behind this is that the direct context in which a term resides is most informative to its meaning. Note that this window does not include the ambiguous term itself. These collections of vectors are then combined into M-dimensional vectors using the same composition function as above. This is done separately for each term occurrence within a single document, creating a $M \times x$ matrix, where x is the number of times the ambiguous term occurs in a single document. These are then combined in an M-dimensional *term vector* using the same composition we used for the concepts, above.

3 Results

We evaluate our approach using three sets of word vectors: The first set was trained on a small set of Medline abstracts (220 million words), the second set on a part of the MIMIC-III corpus (130 million words) [3]. For both sets, we used the `word2vec` implementation from `gensim` [4], using skipgram with negative sampling, a frequency cutoff of 5 and a negative sampling of 15.

Additionally, we used a third set of vectors, available from the BioASQ organisers[1], which was trained on a much larger set of Medline abstracts.

The accuracy scores obtained by our models using the different word vectors are displayed in Table 1. med, mim and bio denote the small Medline vectors, the Mimic-III vectors and the BioASQ vectors, respectively. We consider both a constrained (C) and an unconstrained (U) version of the task. For each word, the constrained version of the task only considers the senses present in the MSH-WSD as possible targets. The unconstrained version considers all concepts which are denoted by the ambiguous term in the 2015AB version of the UMLS as possible targets. The term `cortex`, for example, only has 2 concepts associated with it in the MSH-WSD, while in the 2015AB UMLS release it can denote 5 separate concepts. Because the unconstrained version of the task considers all words, it therefore gives a better indication of real-life performance.

All reported scores use a window size of 6, which was optimized on a randomly selected set of 20 terms from the MSH-WSD set. Where possible, we display the self-reported scores from the relevant papers on the same dataset. We outperform similar methods (MRD, and UMLS::SenseRelate), while the more complicated step models still outperform our models.

References

[1] Olivier Bodenreider. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic acids research*, 32(suppl 1):D267–D270, 2004.

[2] Antonio J Jimeno-Yepes, Bridget T McInnes, and Alan R Aronson. Exploiting MeSH indexing in MEDLINE to generate a data set for word sense disambiguation. *BMC bioinformatics*, 12(1):1, 2011.

[3] AEW Johnson, TJ Pollard, L Shen, L Lehman, M Feng, M Ghassemi, B Moody, P Szolovits, LA Celi, and RG Mark. MIMIC-III, a freely accessible critical care database. *Scientific Data*, 2016.

[4] Radim Řehůřek and Petr Sojka. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta, May 2010. ELRA. `http://is.muni.cz/publication/884893/en`.

[1] Available on the BioASQ website.

Arguments, Scenarios and Probabilities:
Connections between Three Normative Frameworks for Evidential Reasoning[1]

Bart Verheij[a] Floris Bex[b] Sjoerd T. Timmer[b] Charlotte S. Vlek[a]
John-Jules Ch. Meyer[b] Silja Renooij[b]
Henry Prakken[b,c]

[a]*Institute of Artificial Intelligence, University of Groningen*
[b]*Department of Information and Computing Sciences, Utrecht University*
[c]*Faculty of Law, University of Groningen*

Abstract

Due to the uses of DNA profiling in criminal investigation and decision-making, it is ever more common that probabilistic information is discussed in courts. The people involved have varied backgrounds, as fact-finders and lawyers are more trained in the use of non-probabilistic information, while forensic experts handle probabilistic information on a routine basis. Hence, it is important to have a good understanding of the sort of reasoning that happens in criminal cases, both probabilistic and non-probabilistic. In the present paper, we report results on combining three normative reasoning frameworks from the literature: arguments, scenarios and probabilities. We discuss a hybrid model that connects arguments and scenarios, a method to probabilistically model possible scenarios in a Bayesian network, a method to extract arguments from a Bayesian network, and a proposal to model arguments for and against different scenarios in standard probability theory. These results have been produced as parts of research projects on the formal and computational modelling of evidence. The present paper reviews these results, shows how they are connected and where they differ, and discusses strengths and limitations.

1 Summary

In the full paper [8], of which the present text is a compressed contribution at BNAIC 2016, we study connections between arguments, scenarios and probabilities as normative frameworks in reasoning with evidence. Such a study is relevant given the different backgrounds of the people involved in criminal investigation and decision-making: Arguments and scenarios are familiar among fact-finders and lawyers, whereas probabilities are prominent in reports by forensic experts. By studying connections between arguments, scenarios and probabilities, we aim to enhance the understanding of these three tools for evidential reasoning in the law, thereby contributing to the reduction of reasoning errors and miscommunication caused by these different backgrounds.

Our work builds on recent developments to study reasoning with forensic evidence probabilistically, and in particular using Bayesian networks [5, 2]. Since it is known that it is easy to misinterpret Bayesian networks, for instance causally, we have started the exploration of the combined modelling of arguments and scenarios. Our approach continues earlier work on the design of structured probabilistic models and their explanation [4, 3].

[1]The full paper of this compressed contribution at BNAIC 2016 has been published in 2016 in *Law, Probability & Risk*, 15, 35-70. http://dx.doi.org/10.1093/lpr/mgv013.

We review research on the formal and computational connections between three normative frameworks for evidential reasoning based on arguments, scenarios and probabilities, respectively. We study pairwise connections, and connections between all three.

Specifically, we discuss a hybrid model connecting arguments and scenarios, showing how reasoning with arguments can be combined with reasoning to the best explanatory scenario (see also [1]). We show how scenarios can be embedded in Bayesian networks, thereby connecting the role of the global coherence of scenarios with degrees of uncertainty (see also [9] where a case study evaluation is performed). We show how arguments can be extracted from a Bayesian network (see also [6]). We propose a view on arguments to and from scenarios in the context of probability theory (see also [7]). The paper explains the motivation for the different parts of the research and discusses strengths and weaknesses of each part.

There are many remaining hard questions about the safe handling of probabilistic and non-probabilistic evidence in criminal investigation and decision-making. Still we hope that the lessons that we have learnt by studying the different connections between arguments, scenarios and probabilities, will gradually contribute to the prevention of reasoning errors, and a reduction of miscommunication between fact-finders and forensic experts.

Acknowledgments

The research reported has been performed in the context of the project 'Designing and Understanding Forensic Bayesian Networks with Arguments and Scenarios', funded in the NWO Forensic Science program (`http://www.ai.rug.nl/~verheij/nwofs/`).

References

[1] F. J. Bex. *Arguments, Stories and Criminal Evidence: A Formal Hybrid Theory*. Springer, Berlin, 2011.

[2] N. E. Fenton. Science and law: Improve statistics in court. *Nature*, 479:36–37, 2011.

[3] N. E. Fenton, M. D. Neil, and D. A. Lagnado. A general structure for legal arguments about evidence using Bayesian Networks. *Cognitive Science*, 37:61–102, 2013.

[4] A. B. Hepler, A. P. Dawid, and V. Leucari. Object-oriented graphical representations of complex patterns of evidence. *Law, Probability and Risk*, 6(1–4):275–293, 2007.

[5] F. Taroni, C. Aitken, P. Garbolino, and A. Biedermann. *Bayesian Networks and Probabilistic Inference in Forensic Science*. Wiley, Chichester, 2006.

[6] S. T. Timmer, J. J. Meyer, H. Prakken, S. Renooij, and B. Verheij. Explaining Bayesian Networks using argumentation. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty - 13th European Conference, ECSQARU 2015, Compigne, France, July 15-17, 2015. Proceedings*, pages 83–92. Springer, Berlin, 2015.

[7] B. Verheij. To catch a thief with and without numbers: Arguments, scenarios and probabilities in evidential reasoning. *Law, Probability and Risk*, 13:307–325, 2014.

[8] B. Verheij, F. J. Bex, S. T. Timmer, C. S. Vlek, J. J. Meyer, S. Renooij, and H. Prakken. Arguments, scenarios and probabilities: Connections between three normative frameworks for evidential reasoning. *Law, Probability and Risk*, 15:35–70, 2016.

[9] C. S. Vlek, H. Prakken, S. Renooij, and B. Verheij. Building Bayesian Networks for legal evidence with narratives: a case study evaluation. *Artifical Intelligence and Law*, 22(4):375–421, 2014.

ProofCloud: A Proof Retrieval Engine for Verified Proofs in Higher Order Logic

Shuai Wang

ILLC, University of Amsterdam, The Netherlands

Abstract

ProofCloud is a proof retrieval engine for verified proofs in higher order logic. It provides a fast proof searching service for mathematicians and computer scientists for the reuse of proofs and proof packages. In addition, it includes the first complete proof checking results and benchmarks of the OpenTheory repository.

1 Introduction

Higher order logic is also known as simple type theory. It is an extension of simply typed λ-calculus with additional axioms and inference rules [3]. Interactive Theorem Provers (ITPs) of higher order logic have been playing an important role in software verification, hardware verification and formal mathematics. Among them, the HOL family consists of HOL Light [6], HOL4 [10] and ProofPower [8], etc. These ITPs implemented the same higher order logic, namely Church's simple type theory [7]. However, they each contain significant theory formalizations that are not accessible to each other. Proof libraries from one ITP may contribute to the proof automation of another ITP [4, 5]. For the sake of proof sharing between ITPs with different theories, OpenTheory [7] was developed as a cross-platform proof package manager for proofs in the HOL family. It consists of a standard library and many proof packages including lists, natural numbers, functions, etc. ITPs may not be bug-free and may lead to errors in generated proofs while not being apparent within the proof systems themselves. Together with possible mistakes in the process of importing and exporting, OpenTheory is not guaranteed to be reliable. Even worse, proofs can be huge, making them difficult or even impossible to be checked by hand. The demand of reliability of such systems leads to the necessity of proof checking, especially independently from the ITPs involved. ProofCloud is a better interface of OpenTheory and presents the first complete proof checking results of OpenTheory. This abstract is to be submitted as a Type-B paper to BNAIC'16. The original paper is as attached and has been accepted by the 12th International Workshop on User Interfaces for Theorem Provers (UITP'16) workshop and has been presented during IJCAR in Coimbra, Portugal.

2 Proof Retrieval with ProofCloud

ProofCloud[1] is a search engine of higher order logic proofs customised from Swiftype. While OpenTheory groups proofs up, ProofCloud unpacks them and displays a description of each package (on the index pages) as well as all the theorems contained individually. It has been populated by over 1,800 proofs from 6 packages of OpenTheory. As far as the author knows, it is the only online proof search engine of its kind. ProofCloud presents also the proof checking results by Holide and Dedukti. Further more, it tracks the origin of classicism. With the ability to classify classical proofs, it illustrates the axioms and constructive/classical lemmas involved for each theorem as well as the amount of constructive/classical proofs for the package page.

[1] `airobert.github.io/proofcloud/`

3 Proof Transformation and Checking with Holide and Dedukti

Dedukti [9] is a logical framework based on $\lambda\Pi$-calculus Modulo [2] for defining logics and checking proofs. It has been widely used as a universal proof checker for ITPs including Coq, Matita, HOL, FoCaLiZe, etc[2]. To transform higher order logic proofs from the OpenTheory format to the Dedukti format, we need a translator, namely Holide [1]. Holide uses a modular translation of higher order logic which makes the translation possible to extend [1]. Recent updates of the OpenTheory resulted in an upgrade of the Holide program. More specifically, OpenTheory expanded its logic kernel and added some new inference rules and some other features. Holide is therefore upgraded to version 2 to capture corresponding features and the new format. The standard library in OpenTheory grouped theorems into packages, including the standard library and theorems of booleans, sets, lists, etc. Previous work of Holide has checked only the standard library. Following the upgrade of Holide, for the first time, Holide and Dedukti performed proof checking on all packages of OpenTheory. This provides evidence that OpenTheory is a reliable platform for higher order logic proofs and validated the upgrade of Holide. In addition, the structural proof analyses by ProofCloud shows that the proportion of constructive theorems varies from package to package. Apart from maintaining Holide as OpenTheory evolves, future work also includes adding more packages to ProofCloud and further improve the user interface and the searching accuracy.

References

[1] Ali Assaf and Guillaume Burel. Translating HOL to dedukti. In *Proceedings Fourth Workshop on Proof eXchange for Theorem Proving, PxTP 2015, Berlin, Germany, August 2-3, 2015.*, pages 74–88, 2015.

[2] Denis Cousineau and Gilles Dowek. Embedding pure type systems in the lambda-pi-calculus modulo. In *Typed lambda calculi and applications*, pages 102–117. Springer, 2007.

[3] William M Farmer. The seven virtues of simple type theory. *Journal of Applied Logic*, 6(3):267–286, 2008.

[4] Thibault Gauthier and Cezary Kaliszyk. Matching concepts across HOL libraries. In Stephen Watt, James Davenport, Alan Sexton, Petr Sojka, and Josef Urban, editors, *Proc. of the 7th Conference on Intelligent Computer Mathematics (CICM'14)*, volume 8543 of *LNCS*, pages 267–281. Springer Verlag, 2014.

[5] Thibault Gauthier and Cezary Kaliszyk. Sharing HOL4 and HOL Light proof knowledge. In *Logic for Programming, Artificial Intelligence, and Reasoning (LPAR'15)*, LNCS, 2015. to appear.

[6] John Harrison. HOL Light: An overview. In *Theorem Proving in Higher Order Logics, 22nd International Conference, TPHOLs 2009, Munich, Germany, August 17-20, 2009. Proceedings*, pages 60–66, 2009.

[7] Joe Hurd. The opentheory standard theory library. In *NASA Formal Methods - Third International Symposium, NFM 2011, Pasadena, CA, USA, April 18-20, 2011. Proceedings*, pages 177–191, 2011.

[8] Marcel Oliveira, Ana Cavalcanti, and Jim Woodcock. Unifying theories in ProofPower-Z. In *Unifying Theories of Programming*, pages 123–140. Springer, 2006.

[9] Ronan Saillard. Dedukti: a universal proof checker. In *Foundation of Mathematics for Computer-Aided Formalization Workshop*, 2013.

[10] Konrad Slind and Michael Norrish. A brief overview of HOL4. In *Theorem Proving in Higher Order Logics*, pages 28–32. Springer, 2008.

[2] `https://www.rocq.inria.fr/deducteam/software.html`

Estimating the use of higher-order theory of mind using computational agents[1]

Harmen de Weerd Denny Diepgrond Rineke Verbrugge

Institute of Artificial Intelligence, University of Groningen

1 Introduction

In social interactions, people often reason about the beliefs, goals, and intentions of others. People use this so-called *theory of mind* [5] to understand why others behave the way they do, as well as to predict the future behavior of others. People can even use their own theory of mind to reason about the way others make use of their theory of mind. For example, people make use of *second-order theory of mind* to understand a sentence such as "Alice *knows* that Bob *knows* that Carol is throwing him a party".

Both empirical and simulation studies suggest that the use of higher-order theory of mind may be particularly useful in simple competitive games [1, 3, 4]. In such games, however, it is difficult to distinguish between theory of mind reasoning and simpler, behavior-based strategies. In this paper, we combine a Bayesian model selection technique introduced by [7] with our theory of mind agents [1, 2] to estimate the extent to which participants make use of theory of mind in simple strategic games.

We consider behavioral data from two empirical studies in which participants play a simple repeated game known as matching pennies. In this game, two players simultaneously choose one of two possible actions. The first player wins if both players selected the same action, while the second player wins if the two players chose different actions. Sher et al. [6] let 69 children play a total of 12 rounds each against a confederate who always selected the action that would have won in the last round, while Devaine et al. [3] let 29 adult participants play this game against software agents that followed a theory of mind strategy for a total of 480 rounds per participant.

2 Agents as generative models of behavior

We made use of a technique known as group-level random-effects Bayesian model comparison (RFX-BMS), introduced by Stephan et al. [7]. This technique assumes that different participants may be best described by different models. Given a set of models of participant behavior, RFX-BMS determines the relative proportions of these models that best describes the aggregate behavior of participants.

Our model set consists of nine models of participant behavior. The *biased* model of behavior ignores all past behavior and selects each action with a fixed probability. The *other-regarding* model takes past behavior into account by playing the best response to the opponent's last observed action. The *self-regarding* model, on the other hand, repeats the action it performed in the previous round with a fixed probability. Finally, the *Nash* model of behavior selects an action to play at random.

In addition to the heuristics described above, we also included the theory of mind agents we developed to evaluate the benefits of higher-order theory of mind [1, 2]. A zero-order theory of mind (ToM_0) agent predicts the future behavior of the opponent based on her past behavior. A first-order (ToM_1) agent also considers the game from the perspective of his opponent, and determines what he would do himself in her position. Each additional order of theory of mind provides a theory of mind agent with an additional hypothesis of his opponent's future behavior. For example, a second-order theory of mind (ToM_2) agent believes that his opponent may be using first-order theory of mind to predict his behavior.

[1]The full paper has been published in *Proceedings of the Twelfth Conference on Logic and the Foundations of Game and Decision Theory*, 2016.

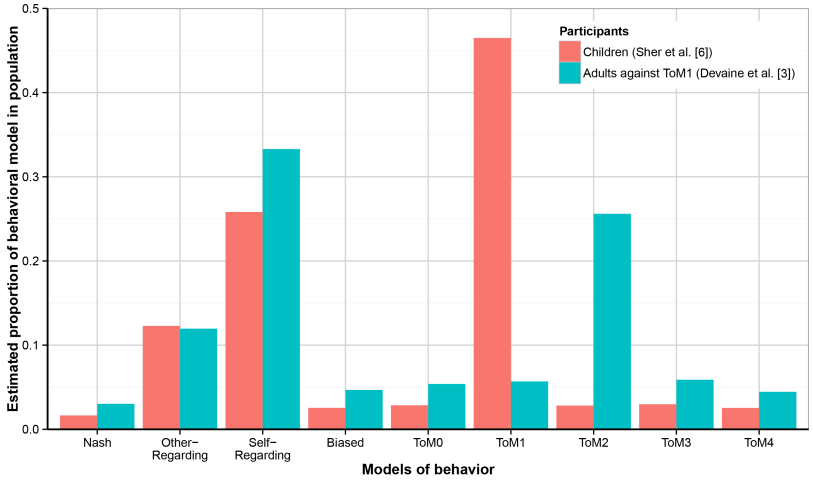


Figure 1: Estimated proportions of behavioral models in the matching pennies game.

3 Results

Figure 1 shows the estimated proportions of the nine models of participant behavior. These results show that Bayesian RFX-BMS estimation classifies over 45% of the children in the study by Sher et al. [6] as using first-order theory of mind. However, the behavior of many children is better described by the either the other-regarding or the self-regarding model of behavior.

Figure 1 also shows that many adult participants playing against a first-order theory of mind agent in the study by Devaine et al. [3] are classified as using second-order theory of mind. However, over 30% of the participants are better described by the self-regarding model of behavior. This suggests that a sizable proportion of participants may rely on simple heuristics when playing matching pennies.

Our results suggest that both children and adults engage in theory of mind in simple games such as matching pennies, but also that many participants are better described by simpler models of behavior. In future research, we aim to determine whether this lack of theory of mind reasoning is due to difficulties participants may experience in classifying the behavior of their opponent. Alternatively, human players may not use their theory of mind abilities primarily to compete with others. Bayesian RFX-BMS estimation could provide more insight in whether participants are more likely to use theory of mind in more cooperative settings, or in settings in which both cooperation and competition play a role.

References

[1] Harmen de Weerd, Rineke Verbrugge, and Bart Verheij. How much does it help to know what she knows you know? An agent-based simulation study. *Artificial Intelligence*, 199-200:67–92, 2013.

[2] Harmen de Weerd, Rineke Verbrugge, and Bart Verheij. Higher-order theory of mind in the Tacit Communication Game. *Biologically Inspired Cognitive Architectures*, 11:10–21, 2015.

[3] Marie Devaine, Guillaume Hollard, and Jean Daunizeau. The social Bayesian brain: Does mentalizing make a difference when we learn? *PLoS Computational Biology*, 10(12):e1003992, 2014.

[4] Adam S. Goodie, Prashant Doshi, and Diana L. Young. Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making*, 25(1):95–108, 2012.

[5] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(04):515–526, 1978.

[6] Itai Sher, Melissa Koenig, and Aldo Rustichini. Children's strategic theory of mind. *Proceedings of the National Academy of Sciences*, 111(37):13307–13312, 2014.

[7] Klaas E. Stephan, Will D. Penny, Jean Daunizeau, Rosalyn J. Moran, and Karl J. Friston. Bayesian model selection for group studies. *Neuroimage*, 46(4):1004–1017, 2009.

Fair task allocation in transportation[1]

Qing Chuan Ye [a] Yingqian Zhang [b] Rommert Dekker [a]

[a] *Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam*
[b] *Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven*

Overview In [1], we study a task allocation problem in which we take fairness into account in addition to the standard minimum cost criterion. This work was inspired by an actual transportation situation in the port of Rotterdam. Due to an increase in inter-terminal transport, an idea was proposed to use already present trucks to execute open jobs, in which daily repeated auctions will be used. In this task allocation problem, the port tries to give jobs to each company according to their availability and their cost for each job. Jobs have designated time slots and can only be allocated to one bidder, while bidders have a limited number of trucks available per time slot. Due to the repetitive nature, we do not only look at optimizing the costs in the task allocation, but we also incorporate fairness in the task allocation. For the fairness criterion we use the notion of max-min fairness. The max-min fairness principle entails that given a total number of jobs, the number of jobs for any company cannot be increased by at the same time decreasing the number of jobs of the other companies that have the same number of jobs or less. Intuitively speaking, we want to allocate jobs among the companies as evenly as possible.

We formulate this problem as a "max-min fair minimum cost allocation problem" (MFMCA), in which we aim for a polynomial-time solution. The difficulty of our problem lies in the additional fairness criterion, which requires the developed algorithm to satisfy three criteria: allocation maximization, fairness, and cost minimization. To the best of our knowledge, no existing polynomial-time algorithm can be directly applied to solve our problem. In [1], we propose polynomial-time algorithms to solve MFMCA as a two-level optimization problem. First, we aim at a fairest allocation among companies while ensuring that a maximal set of tasks can be allocated for execution. We call this the "max-min fair allocation problem" (MMFA). Second, because there might be an exponential number of allocations that are considered max-min fair, we would like to determine which of these fair allocations has the lowest cost. The resulting allocation is max-min fair with minimum cost. Using computational results, we provide insights into situations in which fairness can be incorporated without giving up too much efficiency.

Algorithms First, to solve MMFA, we propose an algorithm called IMaxFlow, which iteratively applies a maximum flow algorithm in a greedy fashion. In this way, the flow assigned to each company is increased step by step until no more flows can be assigned. Given an instance of the MMFA problem, we can construct a flow network and apply IMaxFlow to obtain the optimal max-min fairness vector, which specifies the fairest task distribution among agents given a maximal set of tasks that can be allocated.

Once we know the fairness vector from IMaxFlow, we want to minimize the associated cost. However, we cannot apply a standard minimum-cost maximum-flow algorithm to our flow network as it may violate the max-min fairness condition while looking for the minimum cost. The obtained fairness vector tells us in what quantities the jobs will be distributed in the fairest allocation. However, we do not know *which* company would be assigned *which* number of jobs such that the total cost is smallest. Therefore, we propose the FairMinCost algorithm, which adjusts the flow network in such a way that a standard minimum-cost maximum-flow algorithm simultaneously decides which company is assigned which quantity from the fairness vector, and thus will not violate the max-min fairness condition. This is done by introducing a set of dummy jobs and corresponding additional dummy nodes, dependent on the

[1] The full paper has been accepted in *Omega*, 2016 [1].

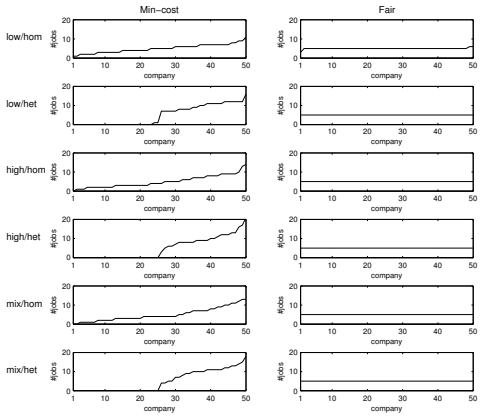


Figure 1: Job distributions in the minimum-cost and fair solutions with *10%* capacity, sorted in ascending order by number of assigned jobs.

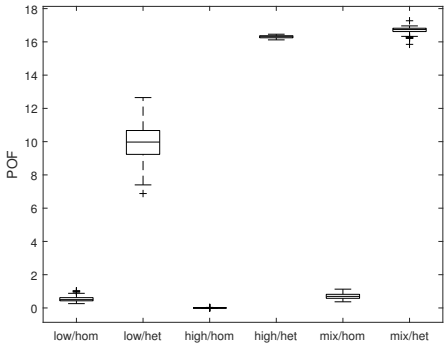


Figure 2: Boxplot of the price of fairness (POF) for different scenarios with *10%* capacity.

optimal max-min fairness vector, to the flow network. We prove in [1] that the solution obtained by applying any standard minimum-cost maximum-flow algorithm on this altered flow network is equivalent to the optimal solution to MFMCA.

Experiments We investigate the performance of the proposed algorithms compared with an optimal algorithm for the task allocation problem with a sole objective of minimum cost through numerical experiments. We are interested in two performance measurements: (1) the price of fairness (POF), i.e., the relative increase in the total cost under the fair solution with regard to the minimum-cost solution; (2) the effect of fairness on the job distribution, i.e., the number of jobs assigned to each company. We look at several scenarios for both the jobs that are being auctioned (*homogeneous* and *heterogeneous* costs), and the companies who are bidding on the jobs (*low*, *high* or *mixed* competition). We also look at different capacities for the companies (*5%* and *10%* capacity). The experimental results show that the job allocation in the minimum-cost solution is very skewed, where some companies obtain many jobs, while others end up with none. The fair allocation allocates the jobs much more evenly over the companies (see Figure 1). However, this of course comes at a cost (see Figure 2). In general, when the costs among companies are similar (*homogeneous* cost), implementing fair allocations comes with almost no extra cost for the task owner. When the prices are highly volatile however (*heterogeneous* cost), the auctioneer may need to pay more for the fairness. In additional experiments, we observe that when there are relatively few jobs, the POF will usually be relatively high due to lack of flexibility in reallocating jobs, but it stabilizes as the number of jobs increases. Similarly, the more companies are participating, the more bids there will be, resulting in more flexibility for reallocation and a lower POF. When there are only few companies, there is lack of flexibility for reallocation, resulting in a fair allocation similar to the minimum-cost allocation. This means that the number of participants should be sufficient in order to have the desired flexibility needed for reallocation. All in all, among the majority of test instances, fairness in the job distribution comes with a very small price in terms of cost.

Although the additional fairness criterion introduces a problem that cannot be solved by known standard algorithms, our proposed algorithms solve MFMCA in polynomial time. Furthermore, our algorithms can be applied to many other task allocation problem in which the centralized planner wants to enforce some kind of fairness among the agents. This is especially relevant in the upcoming sharing economy and collaborative consumption in transport, e.g., car-sharing.

References

[1] Q. C. Ye, Y. Zhang and R. Dekker, *Fair task allocation in transportation*, Omega, 2016, In Press, http://dx.doi.org/10.1016/j.omega.2016.05.005

Generalizing the Detection of Internal and External Interactions in Clinical Guidelines[1]

Veruska Zamborlini[ab] Rinke Hoekstra[ac] Marcos da Silveira [b]

Cedric Pruski [b] Annette ten Teije [a] Frank van Harmelen [a]

[a] *Dept. of Computer Science, Vrije Universiteit Amsterdam, The Netherlands*
{v.carrettazamborlini, rinke.hoekstra, annette.ten.teije, frank.van.harmelen}@vu.nl
[b] *Luxembourg Institute of Science and Technology - LIST, Luxembourg*
{marcos.dasilveira, cedric.pruski}@list.lu
[c] *Faculty of Law, University of Amsterdam, The Netherlands*

Abstract

This paper presents a method for formally representing Computer-Interpretable Guidelines to deal with multimorbidity (patients suffering from multiple simultaneous diseases). Our main contribution is twofold: (i) we provide general models and rules for representing guidelines that express evidence as causation beliefs; (ii) we introduce a mechanism to exploit external medical knowledge acquired from Linked Open Data to detect interactions between recommendations. We evaluate it combining (part of) three guidelines (Osteoarthritis, Diabetes, and Hypertension).

1 Introduction

Clinical Guidelines (CG) are developed for supporting physicians' decision, e.g. specifying what treatment work best in what situation. Ideally, the **recommendations** are based on **evidence** from clinical researches. Since an evidence is not a fact, a multitude of evidence rating systems are adopted by CGs authors. Epistemologically, an evidence reflects a ***belief*** in the existence of a ***causal*** relation between e.g. *administering aspirin and gastrointestinal bleeding*. This paper reports on improvements to the TMR (Transition-based medical recommendation) models and its implementation (Zamborlini et al., 2015) to better address the issue of multimorbidity by providing: (i) a more generic version of the models with respect to recommendations, beliefs and event types; (ii) a formalization of the improved models and rules in FOL; (iii) a Semantic Web framework for representing and reasoning about recommendations and beliefs using standard vocabularies, including (iv) a flexible mechanism for reusing external knowledge bases to enhance the detection of interactions (showcased using Drugbank and Sider[2]).

2 Detecting Recommendations Interactions using External Sources

We illustrate the application of our approach through a (reduced) case study on combining (parts of) three guidelines on Osteoarthritis (OA), Diabetes (DB) and Hypertension (HT) and applying the following generic rules for detecting side-effect (re)using knowledge from the external source Sider:

```
IF R1 is a positive recommendation to action A1 in order to change situation S1
& R2 is a positive recommendation to action A2 to change another situation (not S1)
& Action A2 (or other one in subsuming relationship) is believed to bring about S1
THEN R1 and R2 have 'external' incompatible effect
```

[1] The full paper is published as (Zamborlini et al., 2016).
[2] http://www.drugbank.ca/ and http://sideeffects.embl.de/

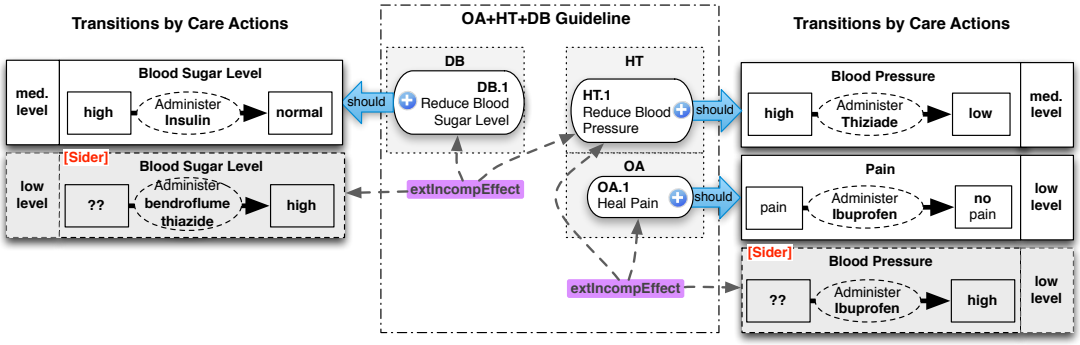


Figure 1: Case study on combining guidelines for OA+HT+DB

The diagram depicted in Figure 1 illustrates our case study. The inner dotted box shows the three guidelines each of which with one recommendation (e.g. DB.1). Those recommendations are based on causation beliefs (depicted outside that box) about actions causing a certain transition. The gray-shade ones are imported from the external source. The top-left causation belief can be read as: *Administer Insulin causes a transition that brings the blood sugar level from high to normal (medium level of confidence)*. By applying the side-effect rule in our case study, two 'external' side-effects can be detected:

1. DB.1 and HT.1 interact if the *thiazide* prescribed is *bendroflumethiazide*, because it has *high blood sugar level* as side-effect according to *Sider*, which is meant to be changed to normal in DB.1.
2. HT.1 and OA.1 interact since the latter has *high blood pressure* as side effect according to *Sider*, which is meant to be changed in HT.1.

This method[3] requires the knowledge from Sider to be reinterpreted according to the causation structure in the TMR model. In this line, the knowledge from any other dataset about side-effects can be used, given that it is also reinterpreted into the causation structure. It means that the side-effect rule (as well as the other ones – external alternative and incompatible actions) does not change if new datasets are added. Moreover, as in the previous work, the rules do not mention any particular guideline, action or drug. It means that they also do not change if new guidelines, actions or drugs are added. This favors scalability, maintanability and reusability.

3 Conclusion

This paper offers a generic and scalable way to represent and reason over multiple clinical guidelines for detecting interactions[4]. Both the conceptual model and the Semantic Web-based implementation were adapted and extended with respect to our previous work (Zamborlini et al., 2015). Incorporating the epistemological nuance of beliefs in the Semantic Web representation, improves the ability to (i) handle knowledge from different sources and (ii) select the reliable ones; (iii) to allow different, incompatible beliefs about the same event to co-exist; and (iv) to provide reusable formal rules that are applicable regardless of specific regulations, guidelines or external sources. This has a favorable effect on reusability, maintainability, and scalability beyond the guidelines we currently covered. An evaluation in collaboration with experts was performed and is under review process. The conclusion is that the automatically detected interactions are relevant, and more interaction types will be addressed in future work.

References

Zamborlini, V., Hoekstra, R., da Silveira, M., Pruski, C., ten Teije, A., and van Harmelen, F. (2015). Inferring Recommendation Interactions in Clinical Guidelines: Case-studies on Multimorbidity. *Semantic Web Journal*.

Zamborlini, V., Hoekstra, R., da Silveira, M., Pruski, C., ten Teije, A., and van Harmelen, F. (2016). Generalizing the Detection of Internal and External Interactions in Clinical Guidelines. In *Proceedings of the 9th International Conference on Health Informatics (HealthINF)*, Rome, Italy.

[3] SWI-Prolog implementation provided in `http://preview.tinyurl.com/BNAIC-Guidelines`

[4] Extended version entitled *Generalizing the Detection of Clinical Guideline Interactions enhanced with LOD*, in Biomedical Engineering Systems and Technologies, Springer (to appear)

BNAIC
Anticipating Habit Formation: A Psychological Computing Approach to Behavior Change Support [1]

Chao Zhang[a] Arlette van Wissen[b] Daniël Lakens[a] Joaquin Vanschoren[a]
Boris de Ruyter[b] Wijnand A. IJsselsteijn[a]

[a] *Eindhoven University of Technology, PO Box 513 5600 MB Eindhoven*
[b] *Philips Research High Tech Campus 34 5656 AE Eindhoven*

Extended Abstract

Introduction: Mobile computing systems have been proposed as promising tools to support lifestyle behavior change. One big challenge in the field is the problem of behavioral consolidation - even with the awareness and motivation to change, it requires someone's effortful self-control to initiate and sustain the right actions in order to form a new habit. Following the recent development of area of computational cognitive modeling for intelligent behavior change support, we present here a novel mobile computing approach that utilizes a computational model of habit formation to meet this challenge. Based on the formal model, the intelligent system can facilitate behavioral consolidation in two ways. Firstly, the system presents the computed information to the users as new self-knowledge, e.g., the current habit strength, to motivate the change. Secondly, it uses the model to predict user behaviors, to anticipate future changes, and to reason about intervention options based on these predictions.

System description and requirements: The system consists of three components: input, processing, and actuation. The *input component* collects three types of data: (1) the target behavior monitored by system sensors; (2) contextual variables measured by the sensors that are potentially impactful on the target behavior, including both internal (e.g., mood, emotions) and external (e.g., weather, social environment) ones; (3) user's cognitive states measured by self-report. Equipped with the data and a formal model of habit formation, *the processing component* computes two kinds of information at each step of the change process: (1) the target behavior predicted by the contextual variables and the cognitive states; (2) cognitive states updated based on their temporal dynamics and behaviors observed. When contextual variables can be anticipated for some future period

[1] Zhang, C., van Wissen, A., Lakens, D., Vanschoren, J., De Ruyter, B., & IJsselsteijn, W. A. (2016). Anticipating habit formation: a psychological computing approach to behavior change support. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (pp. 1247-1254). ACM. doi: 10.1145/2968219.2968439

(e.g., weather forecast), the processing component also simulates future cognitive states and behaviors. Finally, the *actuation component* utilizes the information provided to select appropriate interventions to influence the users's cognitive states and behaviors.

Theoretical basis and a conceptual model: A conceptual model of habit formation is proposed as the basis for computational implementation. Based on learning and decision-making theories in psychology, we model habit formation from two closely related perspectives: the making of individual daily decisions, and the temporal dynamics of the decisions. Individual decisions are modeled in two sequential processes: *option generation* and *option selection*. When making a decision, e.g., to choose what to eat, behavioral options (e.g., eating tasty fast-food or health salad) need to be generated from the user's memory through effortful retention, environmental cues (e.g., McDonald sign), or higher-level goals (e.g., to lose weight). When the salience of one option dominates the others, it leads to automatic execution of the corresponding behavior. When two or more options are equally salient, the user deliberates more on the options to choose the one with the highest utility. Two sources of information are used for the utility calculation: value functions for each attribute of the options (e.g., how do health benefits translate into the overall value), and the probabilistic beliefs about the option-attribute associations (e.g., how likely it is that salad will benefit my health). Contextual variables and states of the user may moderate the utility calculation by changing the values functions or changing the deliberation time.

The dynamics of three cognitive states that influence decisions are considered in this model. Firstly, the accessibility of an option naturally decays over time, but can be enhanced by external reminders and by performing the behavior. Secondly, habit strength, or the strength of association between cues and the behavior, increases whenever the behavior is performed with presence of the cues, and decreases otherwise. Thirdly, the value functions and the probabilistic beliefs (measured as attitude) are shaped by experience and by persuasive interventions.

Computational implementation: The conceptual model is currently implemented computationally as a time-discrete agent-based model. First of all, the salience of each behavioral option is calculated as the sum of the values of *accessibility* and *habit strength*. Secondly, if the salience difference between two options does not pass a threshold, an expected utility calculation with linear value functions and probabilistic beliefs as inputs is used to select the option with the highest utility. Finally, the temporal dynamics of accessibility and habit strength are modeled as the linear combinations of proportional decays and proportional gains conditioned on the states of relevant events (e.g., the presence or absence of the target behavior and reminders). The dynamics of attitudes are not currently implemented yet.

Future work and Conclusion: We are currently in the phase of verifying and improving the computational implementation of the model. Agent-based simulations are used to check if the dynamics of the cognitive states and behaviors generated by the model make sense based on existing empirical data and intuitive expectations. Future work includes (1) collecting additional data for model validation, (2) designing actuation algorithms to map interventions to the model outputs, and (3) evaluating the effectiveness of our approach. Although our approach is still in early development, we believe that it will contribute greatly to the promising enterprise of model-based behavior change support.

Agents Sharing Secrets[1]

Franziska Burger Joost Broekens Mark A. Neerincx

Interactive Intelligence Group, TU Delft,
Mekelweg 4, 2628 CD Delft, NL

Abstract

Companion agents in health care can only be effective if the user engages with them. Human factors theories posit that motivation may be maintained by creating a bond between user and agent through reciprocal self-disclosure. To study this, we developed a disclosure dialog module and embedded it in a mobile application with avatar presence for diabetic children. The app was used by 11 children in an exploratory field study over the course of approximately two weeks. It was found that the amount of disclosures that children made to the avatar was an indicator for the relatedness children felt towards it at the end of the study. While more related children showed a decline in their usage over time, they used the application more and more consistently than less related children.

1 Introduction

Companion agents in health care are developed for long-term use. Particularly children are susceptible to novelty but lose interest quickly. Self Determination Theory (SDT) [3] posits that one possibility for sustaining motivation is by leveraging relatedness. Social Penetration Theory (SPT) [1] provides a method for establishing relatedness: the reciprocal sharing of personal information with increasingly intimate content. Within the framework of the Horizon 2020 PAL project[2], we explored the possibilities of creating a bond between diabetic children and a virtual companion agent through self-disclosure with the goal of increasing the motivational capacity of the agent. The related literature on the links between disclosure and liking and between liking and motivation in child-robot interaction is sparse. We thus developed a small module to explore the chain *self-disclosure* $\rightarrow$ *relatedness* $\rightarrow$ *motivation* further. Upon a trigger event, the avatar discloses information about itself to the child and encourages the child to disclose in return. The child may then type a response (*active disclosure*) or choose not to reply (*passive disclosure*). In either case, the avatar ends the dialog kindly by thanking the child for the disclosure or for simply listening. We integrated the module into the PAL project application, MyPal.

2 Methods

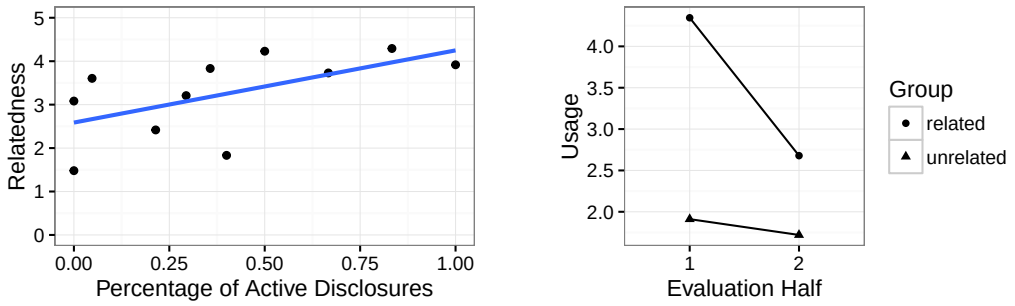
Eleven Dutch diabetic children between the ages of 8 and 12 ($Mean_{age} = 9.91$, $SD_{age} = 1.08$, 6 girls) participated in the exploratory field study. Children were given tablet computers with the application, including the module, and were instructed to use it for two weeks as they pleased. All children had used the application in a prior (one month earlier) three-week field study without the module. Relatedness at the end of the study[3] was assessed with the mean (range: 1-5) of questions from the applicable subscales[4] of the Friendship Qualities Scale [2]. The active and passive disclosures were counted and motivation was operationalized by regarding the average amount of added content within MyPal per day and the consistency (percentage of active app usage days) with which children used the application.

[1] The full thesis was submitted in partial fulfillment of the requirements for the degree of Master of Science in Artificial Intelligence at Radboud University Nijmegen in August of 2016.

[2] `http://www.pal4u.eu/`

[3] For comparability, we had originally employed the same questionnaire as in the prior study to assess relatedness. Due to ceiling effects on the questions, however, the pre-evaluation relatedness could not be determined.

[4] *Companionship*, *Reliable Alliance*, and *Closeness*



(a) (b)

Figure 1: 1a—The relationship between the relative amount of active disclosures of children within the application and their relatedness. 1b—The usage (average daily contributions to MyPal) of more and less related children over time (evaluation half corresponds to approximately one week, but varies slightly per child.)

3 Results

We modeled the relationship between disclosures and relatedness using a linear regression with *the percentage of active disclosures* and *the sum of active and passive disclosures* as predictors. This was chosen to account for both the relative amount of active (or passive) disclosures and children's overall engagement with the module. We found the percentage of active disclosures to be a significant predictor ($b = 1.79, t(8) = 2.690, p = .028$) of relatedness. Thus, controlling for the total amount of disclosures, a unit increase in the percentage of active disclosures corresponds to a 1.79 increase on the relatedness scale (Figure 1a). Furthermore, relatedness correlated significantly with the average amount of added content (one-tailed, $\rho(9) = .64$, $p = .019$) and marginally[5] significantly with the consistency measure (one-tailed, $\rho(9) = .59$, $p = .030$). A visual inspection of the behavior of more and less related children over time (compare Figure 1b), though, shows that children feeling closer to the avatar could not maintain their high initial usage[6].

4 Conclusion

The results of our study show that for children with the same overall engagement with the disclosure module, those disclosing more actively indicated feeling more related to the avatar at the end of the study. More related children used the application more intensely and more consistently than less related children but showed a sharp decrease in usage over time nonetheless. This may also be a sign that the other pillars of intrinsic motivation (autonomy and competence) according to SDT are not optimally met by the application. Since only the post-evaluation relatedness could be assessed, the causal relationships in the chain *self-disclosure* $\rightarrow$ *relatedness* $\rightarrow$ *motivation* remain unexplored. We leave this to future research.

References

[1] I. Altman and D.A. Taylor. *Social penetration: The development of interpersonal relationships.* Holt, Rinehart & Winston, 1973.

[2] W.M. Bukowski, B. Hoza, and M. Boivin. Measuring friendship quality during pre-and early adolescence: The development and psychometric properties of the friendship qualities scale. *Journal of social and Personal Relationships*, 11(3):471–484, 1994.

[3] E.L. Deci and R.M. Ryan. Overview of self-determination theory: An organismic dialectical perspective. *Handbook of self-determination research*, pages 3–33, 2002.

[5] Bonferroni-corrected significance threshold $\alpha = \frac{.05}{2} = .025$

[6] statistical analyses were not conducted, since values are based on only 5-6 children

Performance Indicators for Online Geography in Secondary Education[1]

Pepijn van Diepen
pepijnvandiepen@gmail.com

Bert Bredeweg
B.Bredeweg@uva.nl

Informatics Institute, Faculty of Science
University of Amsterdam, The Netherlands

Abstract

There is little consensus about what variables extracted from learner data are the most reliable indicators of learning performance. The aim of this study was to determine those indicators by taking a wide range of variables into consideration concerning overall learning activity and content processing. A Genetic Algorithm was used for the selection process and the variables were evaluated based on their predictive power. Variables extracted from exercise activities turned out to be the most informative.

1 Introduction

The analysis of learner data can provide insight into the progress of students and their learning performance. Students and teachers can benefit from feedback on the learning process whereas publishers could benefit from feedback on their content [1]. The prediction of learning performance can support this, however, scientists are not confident about what data is most suited for this purpose [1, 6]. Quantitative data concerning *resource use*, *time spent on resources* and *grades* have been used for the prediction of learning performance in the reviewed literature [4, 6]. The aim of this study was to determine what aspects of learning behavior can be extracted from a Learning Management System (LMS) in secondary education and are reliable indicators of learning performance.

2 Method

The data for this research was provided by educational publisher ThiemeMeulenhoff and was extracted from the online geography course De Geo[2]. It consisted of chronological activity logs and exercise results of 226 first grade secondary education students. The course material included reading material (also referred to as *theory*), online exercises and self-assessment tests. Each exercise was categorized according to Bloom's taxonomy for learning objectives [3]. Exercise activity was therefore analyzed separately for each category. Since all data was anonymous and no final grades were made available due to privacy constraints, learning performance had to be determined based on alternative sources. The self-assessment tests were designed to provide the students an indication of their learning performance, therefore results on self-assessment tests were found to be the most appropriate measure of learning performance. First, variables concerning overall online activity were considered (e.g. number of clicks, time online, theory/exercise time spent ratio). Subsequently, content specific variables extracted from reading and exercise activities were analyzed. All data was categorized in terms of exercise processing and theory processing. Subsequently an extensive set of variables was composed for each category

[1]The resources for this project were provided by ThiemeMeulenhoff. Special thanks go to Joost Borsboom, Gilian Halewijn, Wouter van Rennes, Emiel Ubink and Johan Verhaar (in alphabetic order). The full BSc thesis can be found at: https://staff.science.uva.nl/b.bredeweg/pdf/BSc/20152016/Van%20diepen.pdf

[2]De Geo is part of Edition by ThiemeMeulenhoff. More information: https://www.thiememeulenhoff.nl/voortgezet-onderwijs/mens-en-maatschappij/aardrijkskunde/de-geo-onderbouw-9e-editie

based on the type of variables that were found to be reliable in the reviewed literature. Reading activity variables concerned the number of clicks and time spent. Exercise activity variables covered the number of incompletes (wrong answer), completes and the time spent. A selection was made using a univariate variable selection method based on linear correlation with learning performance. Subsequently, multivariate variable selection was applied on the remaining variables using a Genetic Algorithm (GA). By using the prediction performance as fitness and variable subsets as individuals, the GA selected the strongest combination of variables. Since usage of GAs for variable selection is prone to overfit [2], the entire selection process was 10-fold cross validated. Each fold had an optimal variable subset as output and a voting mechanism was used to make a final selection. The predictive power of the selected variables was evaluated in two learning performance classification tasks: fail/pass and fail/sufficient/excellent. Classification algorithms such as Support Vector Machine (SVM), Gaussian Naive Bayes (GNB) and K-Nearest Neighbour (KNN) provided by the Scikit-Learn library [5] were used. All classifications were evaluated using repeated 10-fold cross validation. The accuracy of the classifications was examined alongside with the recall of the poor performing class. Finally, a baseline classifier that was set to always predict the most common class was included in the evaluation.

3 Results and Conclusion

The selection process resulted in a final set of fifteen variables coming from almost all categories. Two belonged to the overall activity category: the total number of clicks and the exercise-theory time spent ratio. Two variables belonged to the reading activity category: the number of theory accesses and amount of time spent on theory while doing exercises. Eleven variables were of the exercise activity category, especially exercises from the *apply* category (five variables) and *understand* category (three variables) from Bloom's taxonomy were found to be reliable indicators. No variables came from the *remember* and *analyze* category and one variable from both the *evaluate* and *create* category. From the exercise processing variables the number of incompletes (wrong answer provided) was most informative, followed by the mean and total time spent on exercises. The last variable of the list concerned the mean time spent on an exercise over all of Bloom's categories together.

The baseline classifiers achieved an accuracy of 0.51 and 0.50 in the two and three class classification task respectively. The SVM classifier predicted most accurately for both classification tasks followed by GNB and KNN. An accuracy of 0.80 and recall of 0.84 of the *fail* class was achieved at the classification of two classes. For the classification of three classes an accuracy of 0.67 and recall of 0.67 was achieved. Other classification algorithms were also evaluated but resulted in less accurate predictions. However, all classifiers did perform significantly better (0.05) than the baseline classifiers.

Concluding, the indicators concerning exercise processing were found to be the most reliable. Variables extracted from exercise activities that were designed to train students in *understanding* and *applying* material were found to be especially informative. However, a combination of features concerning overall activity, theory- and exercise-processing was needed to achieve the best prediction results. Therefore it is important to capture as many aspects of the learning process as possible in order to make accurate predictions.

References

[1] J. Esmeijer and A. van der Plas. Learning analytics en zelfsturend leren. *TNO R10373*, 2013.

[2] I. Guyon et al. An introduction to variable and feature selection. *JMLR*, 3:1157–1182, 2003.

[3] D.R. Krathwohl. A revision of bloom's taxonomy: An overview. *Theory Pract*, 41(4):212–218, 2002.

[4] L.P. Macfadyen and S. Dawson. Mining lms data to develop an early warning system for educators: A proof of concept. *Comput Educ*, 54(2):588–599, 2010.

[5] F. Pedregosa et al. Scikit-learn: Machine learning in Python. *JMLR*, 12:2825–2830, 2011.

[6] D.T. Tempelaar, B. Rienties, and B. Giesbers. In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Comput Human Behav*, 47:157–167, 2015.

Detecting Interesting Outliers

Active Learning for Anomaly Detection

Floris den Hengst (florisdenhengst@gmail.com)
Mark Hoogendoorn (m.hoogendoorn@vu.nl)

Vrije Universiteit, De Boelelaan 1105 1081 HV Amsterdam

1 Introduction

This work treats the problem of identifying a subset of outliers which are interesting to a domain expert. Advancements in technology and a continuously growing adoption of technology has resulted in an increasing growth of data volume [2]. A characteristic of this data is the lack of control over the generating process and/or a statistical rigour therein, rendering much of the statistical outlier detection methods inapplicable.

We assume a difference between outliers that are ‘interesting‘ and ‘noisy‘ by judgement of a domain expert. Interesting can be thought of as erroneous, requiring further inspection. Noisy can be thought of as highly unlikely in an uninformative manner. We propose a framework for finding the interesting outliers with little expert effort. We perform measurements on three datasets.

The notion of separating noise and anomalies through a combination of unsupervised and supervised methods as presented is novel to this domain. Similar works either use an unsupervised model once or contain a single semi-supervised model to achieve active learning [5] [3].

2 Approach

We propose the active learning framework depicted in Figure 1. The *unsupervised component* serves to identify outliers while the *supervised component* serves to form a definition of ‘interestingness’ of outliers. The *selection mechanism* selects an entry to be labelled by the expert based on output from the unsupervised and supervised components. We aim to select informative entries to minimize expert effort (selecting the most informative entry to label next, as in Active Learning). The label for this entry is retrieved before starting a new iteration. This leads to a growing number of labelled inputs to aid the supervised component in separating ‘interesting outliers’ from ‘noisy’ ones. Component outputs are combined for final classification once the stopping criterion has been met.

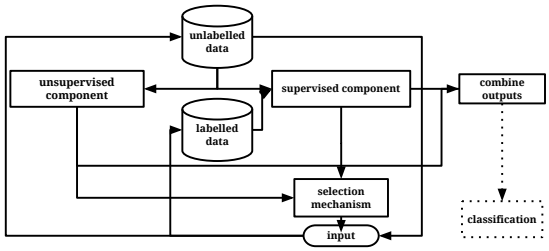


Figure 1: Framework for detecting interesting outliers.

We measure Precision, Recall and F1 score for a varying number of hints provided by the expert to asses learning capabilities. In order to assess *exploratory* qualities, we measure performance on training data. In this scenario, all data is available for querying but not all labels are known. In order to assess *generalisability*, we measure performance on test data using five fold cross-validation. The latter represents a scenario where entries are added to an already assessed data set.

We perform experiments on adapted versions of the ‘Abalone’ and ‘Thyroid’ data sets from the UCI repository and on a novel set from the Human Resource (HR) domain. For benchmark data, ground truth was established by appointing one out of multiple minority classes as ‘interesting’. For the domain

data set, a set of previously labelled records was enriched by a second pass in which the top-50 outliers according to various methods was revisited by the domain expert.

3 Results

We find Local Outlier Factor (LOF) [1] and a general SVM classifier to yield the highest F1 scores out of the tested unsupervised and supervised methods respectively as preliminary results. Out of time considerations, these are exclusively used in further experiments.

For the selection mechanism, we propose a combination of the *margin strategy* for SVM classifiers [4] with rank-based disagreement $D(i)$ which is defined as: $D(i) = |R_U(i) - R_S(i)|$ for ranks R_U and R_S produced by unsupervised component U and supervised component S for each item i. The intuition for this selection mechanism is to gradually refine the supervised model without ignoring promising but previously unseen regions of data.

From Figure 2 we see the benefits of using an informed query mechanism over a random query mechanism. The learning curves for the proposed query mechanism (solid lines) dominate those for the random query mechanism (dashed lines) for the domain and Allhypo Thyroid data sets. The proposed query mechanism yields higher maximum F1 scores and requires less labels to do so. The performance on the Abalone data sets is comparable amongst query mechanisms.

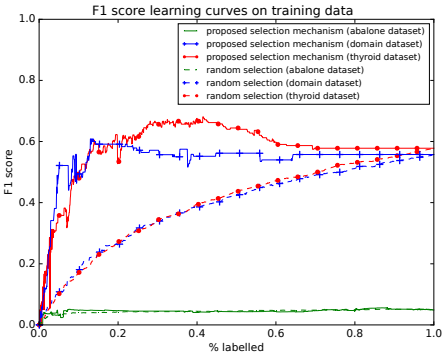


Figure 2: Learning curve based for F1 score on training data.

The upper half of Table 1 contains a comparison of the novel method to an unsupervised method on training data: the gains are mostly in less false positives when all positive items are to be found, with up to roughly nine times as little false positives.

The lower half of Table 1 shows performance on test data using five-fold cross validation. It shows that the novel method is unable to find all positive instances in a test set. We conclude that the novel method succeeds in helping a domain expert find all interesting outliers in an exploratory setting, yet fails to finds all interesting outliers in an autonomous setting.

	Novel Method			Unsupervised	
	$max(Recall)$	at $min(Precision)$	at % labelled	Precision	Data
Abalone	0.775	0.026	69.3%	**0.029**	
Allhypo Thyroid	1.0	**0.403**	78.2%	0.184	Training
Domain	1.0	**0.367**	63.9%	0.038	
Abalone	0.633	0.091	66.3%	0.029	
Allhypo Thyroid	0.182	0.013	42.2%	0.184	Test
Domain	0.785	0.591	100.0%	0.038	

Table 1: Comparison of Precision scores at maximum Recall on both training and test test data in novel method and in unsupervised method. Max Recall for the unsupervised method on all data sets is 1.0.

References

[1] Markus M Breunig, Hans-Peter Kriegel, Raymond T Ng, and Jörg Sander. Lof: identifying density-based local outliers. In *ACM Sigmod Record*, volume 29, pages 93–104. ACM, 2000.

[2] Adam Jacobs. The pathologies of big data. *Communications of the ACM*, 52(8):36–44, 2009.

[3] Dan Pelleg and Andrew W Moore. Active learning for anomaly and rare-category detection. In *Advances in Neural Information Processing Systems*, pages 1073–1080, 2004.

[4] Simon Tong and Daphne Koller. Support vector machine active learning with applications to text classification. *The Journal of Machine Learning Research*, 2:45–66, 2002.

[5] Cui Zhu, Hiroyuki Kitagawa, Spiros Papadimitriou, and Christos Faloutsos. Obe: outlier by example. In *Advances in Knowledge Discovery and Data Mining*, pages 222–234. Springer, 2004.

A Fuzzy Logic Approach for Anomaly Detection in Energy Consumption Data *

BNAIC 2016

B.Sc. Abstract. †

Muriel Hol [a] Aysenur Bilgin [b]

[a] *University of Amsterdam, 1098 XH Amsterdam*

[b] *Institute for Logic, Language and Computation, UvA, 1090 GE Amsterdam*

1 Introduction

Deficient building energy management [4] induces the need for improved automated fault detection and diagnostics (FDD). In real world energy consumption data, there exist many sources of uncertainties, such as changing weather conditions and varying building features. Fuzzy Logic (FL) enables modelling knowledge-based non-linear systems that can handle these uncertainties. As opposed to binary membership of classical logic, FL uses continuous membership to a set, which are represented by fuzzy membership functions [10]. In other words, FL allows for more degrees of freedom in uncertainty modelling that can be suited for real-world applications. Furthermore, FL uses linguistic terms, which facilitate modelling human interpretable systems. The aim of this study is to investigate how FL can improve energy management systems. In particular, the focus is placed on how FL performs in (1) forecasting, (2) anomaly detection and (3) modelling large scale systems involving multiple buildings.

2 Method

Based on data used by Nadai and van Someren [3], Lodewegen [5] investigated gas consumption forecasting with an Artificial Neural Network (ANN). This approach allows a good reference point for investigating the performance of FL versus ANN. The same data is provided for this project. A data frame was constructed for both the NTH and KMH building. Time-series data based on weather and gas consumption is used. For example, the variables include temperature at time t, gas consumption at time t-1 and day of the week. Additionally, the Seasonal Trend Decomposition was applied with the LOESS method [3]. The residuals of the daily and yearly gas consumption trends are used as features. In order to test if the Fuzzy Inference System (FIS) is scalable, a generic database containing four different buildings was constructed by adding building characteristics such as the year of construction [5].

The 5 components of a FIS are (1) the Fuzzifier, (2) Rulebase, (3) Database, (4) Inference System and (5) Defuzzifier [7]. The rulebase and database were generated with the frbs package in R [6] using the method of Wang and Mendel (WM) [8]. The outputs are a fuzzy rulebase and a database that holds the membership function values and types. In order to modify and fine-tune the system, the fuzzifier and inference system are implemented outside of the frbs package. This is done because the WM method is based on partitioning of the data and the frbs package only allows the same number of equal ranged partitions for each feature. However, some features clearly show different patterns than others. Based on examination of the data plots, some feature values are represented by multiple membership functions. Based on Wang and Mendel [8], the centroid method is used for the defuzzification. The output of the defuzzification is the predicted gas consumption.

The FIS is tuned for (1) membership function type: Gaussian, triangle or trapezoid, (2) T-norm: MIN- or product operator [8] (3) and the number of fuzzy sets: 5,7,9, or 11. The forecasting performance measures used for tuning the system include MAE and RMSE. For tuning, the FIS is trained on 'normal' data from 2008 to 2012 en tested on 'normal' data in 2013. The fuzzy system, which uses 9 Gaussian membership functions with MIN t-norm (Gaussian-MIN-9), has given the lowest forecasting errors.

Wijayasekera et al. [9] address that the strength with which a data point fires the rules indicates to what extent the data point behaves normally. Therefore, the rules need to be inferred from normal data. The data is labelled as either 'normal' or 'anomalous' with the 3 sigma rule [2, p.30].

In all data sets, the mean firing strength of the 'normal' data was higher than the of the 'anomalous' data. Therefore, the threshold is calculated by taking the mean of the firing strengths of the 'anomalous' data. If the firing strength of a data point is lower than the threshold, it is classified as an anomaly.

3 Results

The forecasting experiments were conducted with the Gaussian-MIN-9 system, using 5 fold cross validation, where 80% of the data was used for training and the remaining 20% was used for testing. The results are shown in Figure 1. A t-test with a significance level of 0.25% has proved the results of the fuzzy approach to be significantly better than the ANN[1].

*This project is partially supported by the Marie Curie Initial Training Network (ITN) ESSENCE, grant agreement no. 607062.

†Hol, M. & Bilgin, A. (2016). A Fuzzy Logic Approach for Anomaly Detection in Energy Consumption Data (Unpublished B.Sc thesis). University of Amsterdam, Netherlands.

[1]Results of the ANN are obtained with a single hidden layer. For the NTH building, 150 hidden neurons 150 epochs were used. For the KMH building these were 100 and 100 respectively [5].

The generic forecasting results are retrieved by training the system on the data of 4 buildings using data points from 2008 to 2012 and testing it on data of 2013 from one specific building. The MAE and RMSE, when trained on the generic database and tested on the NTH building, are 30.22 and 24.46, respectively. This is an increase of approximately 100%.

For the anomaly classification, the FIS is trained on normal data from 2008 to 2012 and tested on the year 2013. In this manner, results reflect performance in practice and the ability to classify anomalies that lay near the boundary of being normal or abnormal is tested, which is one of the biggest challenges in this field [2]. Since there is no reference system for comparing the anomaly detection results of the data at hand, a baseline system has been created. The baseline system implements that all data points are an anomaly. The F1 [2] score of the baseline system is 0.201, whereas the Gaussian-MIN-9 gives 0.226.

NTH Building

System	MAE	RMSE
ANN	12.01	15.58
Fuzzy	**9.61**	**13.60**

KMH Building

System	MAE	RMSE
ANN	1.82	3.04
Fuzzy	**1.73**	**2.70**

Figure 1: The Forecasting results of the ANN and FIS.

4 Conclusion and Future Work

The proposed fuzzy approach using the WM method has led to significantly better forecasting performance in comparison with ANN approach [5] regarding the energy consumption data of the NTH and KMH buildings. For anomaly classification, the challenges in this field have been met and results above the baseline were achieved. The preliminary results of the proposed approach for large scale applications are not yet sufficient. More research is required to draw conclusions about the performance.

As part of future work, advanced methods for establishing an anomaly classification threshold and reliable data labelling can be investigated. Furthermore, it would be valuable to investigate the options of amplifying the system to a hybrid system, for example by using the ARIMA method as proposed in Nadai and van Someren [3]. In order to reduce the size of the Rule Base, the system would be improved by using the Cooperative Rule version of Wang and Mendel [1].

References

[1] J. Casillas, O. Cordón, and F. Herrera. Improving the Wang and Mendel's Fuzzy Rule Learning Method by Inducing Cooperation Among Rules. *Proceedings of the 8th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference*, pages 1682–1688, 2000.

[2] V. Chandola and V. Kumar. Anomaly detection: A survey. *ACM Computing Surveys*, 51:0–58, 2009.

[3] M. De Nadai and M. van Someren. Short-term anomaly detection in gas consumption through ARIMA and Artificial Neural Network forecast. *Procs. 2015 ieee workshop on environmental, energy, and structural monitoring systems*, pages 250–255, 2015.

[4] S. Katipamula and R.B Brambley. Review article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I. *HVACR Research*, 11:3–25, 2005.

[5] J. Lodewegen. Saving energy in buildings using an Artificial Neural Network for outlier detection. *B.Sc Thesis. University of Amsterdam.*, June 2015.

[6] L.S. Riza, C. Bergmeir, F. Herrera, and J.M. Benítez. frbs: Fuzzy Rule-based systems for Classification and Regression in R. *Journal of Statistical Software*, 65(6):1–30, 2015.

[7] F. Sanz, J. Ramírez, and R. Correa. Fuzzy Inference Systems Applied to the Analysis of Vibrations in Electrical Machines, Fuzzy Inference System - Theory and Applications. 2012.

[8] L.X. Wang and J.M. Mendel. Generating fuzzy rules by learning from examples. *Systems, Man and Cybernetics*, 22 6:1414–1427, 1992.

[9] D. Wijayasekera, O. Linda, M. Manic, and C. Rieger. Mining Building Energy Management System Data Using Fuzzy Anomaly Detection and Linguistic Descriptions. *IEEE Transactions on Industrial Informatics*, 10:1829–1839, 2014.

[10] L.A. Zadeh. Fuzzy Sets. *Information and Control*, 9:338–353, 1965.

[2] $F1score = 2 \cdot \frac{precision \cdot recall}{precision + recall}$

Multi-Objective Deep Reinforcement Learning with Optimistic Linear Support[1]

Hossam Mossalam

University of Oxford, Department of Computer Science

Sequential decision-making problems with multiple objectives arise naturally in practice and pose unique challenges for research in reinforcement learning. While, reinforcement learning has largely focused on single-objective settings, most real-world problems have multiple objectives, and it is not always clear how to evaluate different available trade-offs between these objectives a priori. Therefore, it is often highly desirable to produce a so-called *coverage set*, i.e., a set containing at least one optimal policy (and policy value vector) for each possible utility function that a user might have. Recently, a lot of progress has been made in this field. However, until now, deep learning methods have not yet been developed for multi-objective reinforcement learning problems, while they have proven very effective for single-objective reinforcement learning problems. One main reason for this is that neural networks — which is the underlying data structure for deep learning — cannot account for unknown preferences and the resulting sets of value vectors. In this thesis, we aim to circumvent this issue by taking a so-called outer loop approach [3] to multi-objective reinforcement learning, i.e., we aim to learn an approximate coverage set of policies, represented by a neural network, by evaluating a sequence of scalarized (that is, single-objective) problems. In order to do so effectively, we expect that it is necessary to share parts of the neural networks from earlier scalarized problems, in later iterations of the algorithm, which poses interesting algorithmic challenges.

In this thesis, we restrict ourselves to the case in which a *convex coverage set (CCS) is the optimal solution*, which is the optimal solution when the utility of the user can be expressed with a linear scalarisation function, $f(\mathbf{V}^{\pi}, \mathbf{w}) = \mathbf{w} \cdot \mathbf{V}^{\pi}$, i.e., the scalarised value is a convex combination of the value, $\mathbf{V}^{\pi}$, of a policy, π, in each objective. To this end, we use the *Optimistic Linear Support (OLS)* [3, 4] framework for multi-objective planning, in combination with deep reinforcement learning to create a novel multi-objective reinforcement learning algorithm. We call the resulting algorithm *Deep OLS Learning (DOL)*. DOL learns a CCS by learning policies for different $\mathbf{w}$ sequentially. These $\mathbf{w}$ are selected in a smart way via the OLS framework. By selecting the appropriate *Deep Q-Network (DQN)* architecture for different MOMDP problems (e.g., we use convolutional neural networks for image-based problems), we can learn near-optimal approximate CCSs, for multi-objective RL problems of unprecedented size.

While DOL can already tackle very large MOMDPs, re-learning the parameters for the entire network when we move to the next $\mathbf{w}$ in the sequence is rather inefficient. Gladly, we can make use of the following observation: the optimal value vectors (and thus optimal policies) for a scalarised MOMDP with a $\mathbf{w}$ and a $\mathbf{w}'$ that are close together, are typically close as well [5]. Because deep Q-networks (auto-)encode those features of a problem that are relevant for the optimal value (and thereby policy) of an MOMDP, we could probably speed up computation by reusing the neural networks that we found earlier in the sequence. We therefore extend DOL with the ability to reuse already learnt neural networks to speed up the learning of new ones leading to our new algorithm *DOL with reuse (DOL-R)*. DOL-R has two variations: in *DOL Full Reuse DOL-FR*, we start learning for a new scalarisation weight $\mathbf{w}'$, using the complete DQN we optimised for the previous $\mathbf{w}$ that is closest to $\mathbf{w}'$ in the sequence of scalarisation weights that OLS generated so far; in *DOL Partial Reuse DOL-PR*, we take the same network as for full reuse, but we reinitialise the last layer of the network randomly, in order to escape local optima.

[1] Extended abstract of the MSc thesis "Multi-objective Deep Reinforcement Learning" from the University of Oxford [1]. A more extended version of this work is available in [2].

We use Deep Sea Treasure world problem which is a well known multi-objective benchmark [6] to test the potential of DOL and DOL-R. In Deep Sea Treasure world problem, the agent controls a submarine searching for 10 treasures (terminal states) which are at different distances from the starting state. The two objectives in this problem are fuel usage (-1 per step) and the value of the treasures. We adjusted the standard rewards of the problem so that the optimal policies for finding the different treasures constitute a CCS. The main reason for using this problem is that it has a well-known structure, and we have a ground-truth CCS that we can compare our results to. We test on both the standard raw version, in which the state is a tuple of grid coordinates, and a new version, in which we use an image of the deep sea treasure problem as input for the learner. While the input for the learner (i.e., an image) of the latter is incomparably higher than the raw version (i.e., two integers), the ground truth CCS stays the same. In order to solve the image version of the problem, deep learning methods are essential, as the input would be too large for tabular reinforcement learning methods. For the raw version of the problem we use a feed-forward neural network for DOL(-R), while for the image version we employ a convolutional neural network.

To measure the effectiveness of our algorithms, we employ the Max CCS Difference, ε, i.e., we compare the output set of our algorithms with the ground truth CCS, and measure the maximum difference in (linearly) scalarised value [5]. The max CCS error is defined as $\varepsilon_i = \max_{w \in W} |(\max_{c \in CCS} w.c) - (\max_{s \in S_i} w.s)|$, where i represents the i^{th} iteration.

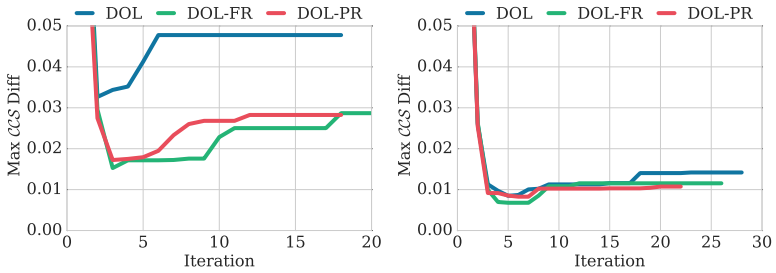


Figure 1: The Max CCS difference error (ε) after the i-th iteration of DOL and DOL-R: (left) Raw version; (right) Image version.

First, we evaluate our proposed methods, in a simple scenario, where the agent has direct access to s_t (Figure 1 (left)). Hence, we employ a simple feed forward neural network architecture, to measure the maximum error in scalarised value with respect to the true CCS. We trained the neural network for 4000 episodes and we observe that DOL suffers from the worst performance amongst the three proposed algorithms. Also, the difference in performance between DOL and DOL-R is considerably high indicating the benefit of reuse on the overall accuracy. For the image version, we employed a convolutional neural network which was trained for 6000 episodes. We observe that DOL and DOL-R achieve a very similar performance with with DOL-PR achieving the best results amongst all three algorithms. The results from both experiments indicate that DOL-R consistently achieves better results than DOL.

We therefore conclude that DOL can successfully learn a CCS, in problems (such as image deep sea treasure) that are intractable with traditional multi-objective RL algorithms. Furthermore, DOL-PR leads to a significant improvement over DOL without reuse. To our knowledge, DOL and DOL-R are the first multi-objective RL methods that can harness the power of deep reinforcement learning.

References

[1] Hossam Mossalam. Multi-objective deep reinforcement learning. Master's thesis, MSc in Computer Science, University of Oxford, 2016.

[2] Hossam Mossalam, Yannis M Assael, Diederik M Roijers, and Shimon Whiteson. Multi-objective deep reinforcement learning. *arXiv preprint arXiv:1610.02707*, 2016.

[3] Diederik M. Roijers. *Multi-Objective Decision-Theoretic Planning*. PhD thesis, University of Amsterdam, 2016.

[4] Diederik M. Roijers, Shimon Whiteson, and Frans A. Oliehoek. Computing convex coverage sets for faster multi-objective coordination. *Journal of Artificial Intelligence Research*, 52:399–443, 2015.

[5] Diederik M. Roijers, Shimon Whiteson, and Frans A. Oliehoek. Point-based planning for multi-objective POMDPs. In *IJCAI 2015: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, pages 1666–1672, July 2015.

[6] P. Vamplew, R. Dazeley, A. Berry, E. Dekker, and R. Issabekov. Empirical evaluation methods for multiobjective reinforcement learning algorithms. *Machine Learning*, 84(1-2):51–80, 2011.

Deep Reinforcement Learning for Coordination in Traffic Light Control

Elise van der Pol [a]

[a] *University of Amsterdam*

Abstract

This thesis combines the Deep Q-learning algorithm with max-plus coordination in transfer planning in order to find optimal coordinated policies in traffic light control with multiple agents, an approach that outperforms earlier work.

1 Introduction

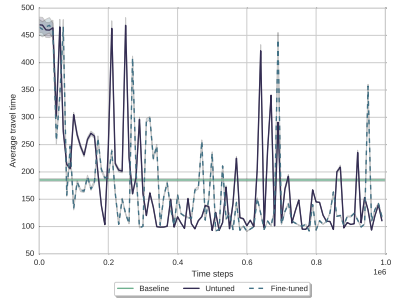
The cost of traffic congestion in the EU is estimated to be 1% of the EU's GDP [2], and good solutions for traffic light control may reduce traffic congestion, saving time and money and reducing pollution. To find optimal traffic light policies, reinforcement learning (RL) uses reward signals from the environment to learn to make optimal decisions. The Deep Q-learning algorithm [4] (DQN) has shown good results on Atari games, and we research whether it can also be succesfully applied to traffic light control. In particular, this thesis [6] applies DQN to the problem of traffic light control and extends earlier work [7] to the multi-agent case.

2 Deep Reinforcement Learning for Traffic Light Control

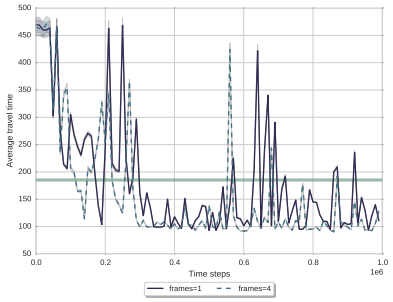
A traffic light agent is trained using the DQN algorithm. The state representation is a binary matrix of vehicle positions, which is fed into a convolutional network. Different settings are compared, by varying amongst others replay memory size, state representation and the use of prioritized experience replay [8]. While the DQN algorithm may suffer from stability issues, caused in part by its convolutional network struggling with non-i.i.d. datasets and moving targets, it can learn directly from image data and find very good policies. However, in traffic light control, taking suboptimal actions may lead to traffic jams, making exploration more costly than in Atari games. To deal with these issues, the best found settings are combined into a fine-tuned DQN agent, which is compared with the untuned agent and a baseline agent that uses a linear approximator in Figure 1a, on the average travel time found during policy evaluations at different points in the training process. While both DQN agents outperform the linear baseline, there is clear oscillation in the graph - the training process is not stable. Fine-tuning, especially adding more information to the state, helps - see for example Figure 1b, where adding past position matrices to the state improves stability.

3 Coordination of Deep Reinforcement Learners

First, the DQN algorithm is used to learn local joint Q-value functions of two neighbouring agents. Then, using *transfer planning* [5], the learned functions are used in the max-plus coordination algorithm [1] to make decisions in a three-agent chain and four-agent cycle. In Figure 2, the fine-tuned agent is compared to the untuned agent and an earlier RL approach [3], evaluated at the best policy found. In both the three-agent and four-agent case, the fine-tuned agent is most stable, while the best policies found by both agents are close in performance. However, in the four-agent scenario, the untuned agent has some huge travel time spikes, suggesting that the error in estimating the Q-value function stacks with the lack of guarantees for max-plus in cyclical graphs, resulting in suboptimal coordinated policies.

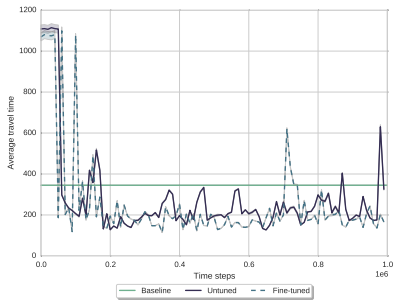


(a) Average travel time. Fine-tuned versus untuned DQN agent, linear baseline.

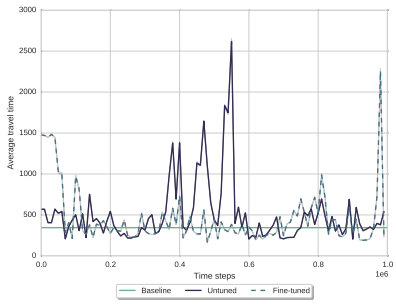


(b) Average travel time. One versus four state frames.

Figure 1



(a) Average travel time, three-agent scenario



(b) Average travel time, four-agent scenario

Figure 2

4 Conclusion

Traffic light control presents unique challenges not necessarily present in the benchmarks used in earlier work. Moreover, while it can outperform earlier work, DQN for traffic light control suffers from stability issues. However, in principle the approach of using deep reinforcement learning for learning source problems in transfer planning is promising and there are many directions for future research that can make this approach more reliable, both in general and for traffic light control.

Acknowledgements

The Tesla K40 used for this research was donated by the NVIDIA Corporation. The thesis was supervised by Frans A. Oliehoek.

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag New York, Inc., 2006. ISBN: 0387310738.

[2] Commission of the European communities. *White paper European transport policy for 2010: time to decide*. 2001.

[3] L. Kuyer et al. “Multiagent reinforcement learning for urban traffic control using coordination graphs”. In: *ECML-PKDD*. 2008.

[4] V. Mnih et al. “Human-level control through deep reinforcement learning”. In: *Nature* 518.7540 (2015), pp. 529–533.

[5] F. A. Oliehoek, S. Whiteson, and M. T. Spaan. “Approximate solutions for factored Dec-POMDPs with many agents”. In: *AAMAS*. 2013.

[6] E. van der Pol. “Deep reinforcement learning for coordination in traffic light control”. Master’s Thesis. University of Amsterdam, 2016.

[7] T. Rijken. “DeepLight: Deep reinforcement learning for signalised traffic control”. Master’s Thesis. University College London, 2015.

[8] T. Schaul et al. “Prioritized experience replay”. In: *ICLR*. 2016.

Conditional Return Policy Search for TI-MMDPs with Sparse Interactions[1]

Joris Scharpff[a] Diederik M. Roijers[b] Frans A. Oliehoek[c]
Matthijs T. J. Spaan[a] Mathijs M. de Weerdt[a]

[a] *Delft University of Technology*
[b] *University of Oxford*
[c] *University of Amsterdam, University of Liverpool*

When cooperative teams of agents are planning in uncertain domains, they must coordinate to maximise their (joint) team value. In several problem domains, such as maintenance planning [6], the full state of the environment is assumed to be known to each agent. Such *centralised* planning problems can be formalised as multi-agent Markov decision processes (MMDPs) [1], in which the availability of complete and perfect information leads to highly-coordinated policies. However, these models suffer from exponential joint action spaces as well as a state that is typically exponential in the number of agents. This is especially an issue when *optimal* policies are required. In this paper, we identify a significant MMDP sub-class whose structure we compactly represent and exploit via locally-computed upper and lower bounds on the optimal policy value. We exploit both the compact representation, and the upper and lower bounds to formulate a new branch-and-bound policy search algorithm we call *conditional return policy search (CoRe)*. CoRe typically requires less runtime than the available alternatives and finds solutions to previously unsolvable problems [5].

We consider *transition independent* MMDPs (TI-MMDPs). In TI-MMDPs, agent rewards depend on joint states and actions, but transition probabilities are *individual*. Our key insight is that we can exploit the reward structure of TI-MMDPs by decomposing the *returns* of all execution histories – i.e., all possible state/action sequences from the initial time step to the planning horizon – into components that depend on local states and actions. To do so, we build on three key observations. 1) Contrary to the optimal value function, returns *can* be decomposed without loss of optimality, as they depend only on local states and actions of execution sequences. This allows a compact representation of rewards and efficiently computable bounds on the optimal policy value via a data structure we call the *conditional return graph* (CRG). 2) In TI-MMDPs agent interactions are often sparse and/or local, typically resulting in very compact CRGs. 3) In many problems the state space is transient, i.e., states can only be visited once, leading to a directed, acyclic transition graph. With our first two key observations this often gives rise to *conditional reward independence* – the absence of further reward interactions – and enables agent decoupling during policy search.

In order to represent the returns compactly with local components, we first partition the reward function into additive components $\mathcal{R}_i$ and assign them to agents. The *local* reward for an agent $i \in N$ is given by $\mathcal{R}_i = \{R^i\} \cup \mathcal{R}_i^e$, where R^i is the reward function that only depends on agent i and $\mathcal{R}_i^e$ is the set of interaction reward functions assigned to i (restricted to a subset of those R^e where $i \in e$, i.e., those functions that depend on i have at least one other agent in its scope, e). The sets $\mathcal{R}_i$ are disjoint sub-sets of the reward functions, $\mathcal{R}$. Given a disjoint partitioning $\bigcup_{i \in N} \mathcal{R}_i$ of rewards, the *Conditional Return Graph (CRG)* ϕ_i is a *directed acyclic graph* with for every stage t of the decision process a node for every reachable local state s_i, and for every local transition $(s^i, a^i, \hat{s}^i)$, a tree compactly representing all transitions of the agents in scope in $\mathcal{R}_i$. The tree consists of two parts: an action tree that specifies all dependent local joint actions, and an influence tree, that contains the relevant local state transitions included in the respective joint action.

[1] Full version published in the proceedings of AAAI 2016 [5].

An example CRG for one time step is given in Fig. 1. The graph represents all possible transitions that can effect the rewards in $\mathcal{R}_i$, given the local transitions of the state of agent i (in this case only from s_0^1 to s_1^1). The labels on the path to a leaf node of an influence tree, via a leaf node of the action tree, sufficiently specify the joint transitions of the agents e in scope of the functions $R^e \in \mathcal{R}^i$, such that we can compute the reward $\sum_{R^e \in \mathcal{R}_i} R^e(s^e, \vec{a}^e, \hat{s}^e)$. The wildcard, $*^2$, represents any action of agent 2 for which there is no interaction reward, i.e., all reward functions depending on both agent 1 and agent 2 yield 0. In the paper, we prove that CRGs are indeed a compact representation of histories, and even more so when interactions are sparse.

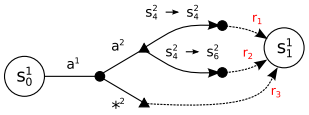


Figure 1: *Example of a CRG for the transition for one agent of a two-agent problem, where R_1 only depends on a^2.*

In addition to storing rewards compactly, we use CRGs to bound the optimal expected value. Specifically, the maximal (resp. minimal) attainable return from a joint state s_t onwards, is an upper (resp. lower) bound on the value. Moreover, the sum of bounds on local returns bounds the global return and thus the optimal value. We define them as $U(s^i) = \max_{(s^e, \vec{a}_t^e, \hat{s}^e) \in \phi_i(s^i)} \left(\mathcal{R}_i(s^e, \vec{a}_t^e, \hat{s}^e) + U(\hat{s}^i)\right)$, such that $\phi_i(s^i)$ denotes the set of transitions available from state $s^i \in s^e$ (ending in $\hat{s}^i \in \hat{s}^e$), in the corresponding CRG. The bound on the optimal value for a joint transition $(s, \vec{a}, \hat{s})$ of all agents is $U(s, \vec{a}_t, \hat{s}) = \sum_{i \in N} \left(\mathcal{R}_i(s^e, \vec{a}_t^e, \hat{s}^e) + U(\hat{s}^i)\right)$, and lower bound L is defined similarly over minimal returns. Note that a bound on the joint returns automatically implies a bound on the value.

We combine the above, together with conditional reward independence, in our *Conditional Return Policy Search (CoRe)* algorithm. CoRe performs a branch-and-bound search over the joint policy space, represented as a DAG with nodes s_t and edges $\langle \vec{a}_t, \hat{s}_{t+1} \rangle$, such that finding a joint policy corresponds to selecting a subset of action arcs from the CRGs (corresponding to $\vec{a}_t$ and $\hat{s}_{t+1}$). First, however, the CRGs ϕ_i are constructed for the local rewards $\mathcal{R}_i$ of each agent $i \in N$, assigned heuristically to obtain balanced CRGs. The generation of the CRGs follows a recursive procedure, during which we store upper and lower bounds on the local returns. During the subsequent policy search, CoRe detects when subsets of agents become conditionally reward independent, and recurses on these subsets separately.

When we compare CoRe to previously available methods, we observe that CoRe can both solve instances that could not previously be solved [5], and that CoRe can solve instances that could be solved by existing methods a lot faster. For example, in the sample of our results presented in Figure 2 we compare the runtime of CoRe to that of SPUDD [2] using a problem-tailored encoding [6] for instances of the maintenance planning problem.

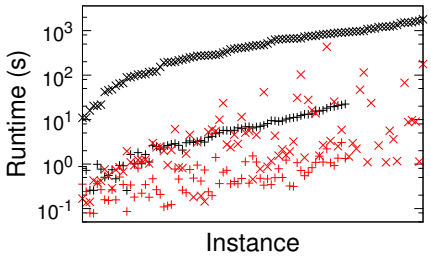


Figure 2: *Experimental results: the runtime of CoRe (red) versus that of SPUDD (black), on the same 2-(+) and 3-agent (×) instances.*

Finally, inspired by the success of CoRe for single-objective TI-MMDPs, we have shown [3] that we can extend our earlier work on multi-objective (TI-)MMDPs [4], using CoRe as a subroutine, to solve significantly larger multi-objective problem instances as well. We thus conclude that CoRe is vital to keeping both single- and multi-objective TI-MMDPs tractable.

References

[1] Craig Boutilier. Planning, learning and coordination in multiagent decision processes. In *TARK*, pages 195–210, 1996.

[2] Jesse Hoey, Robert St-Aubin, Alan Hu, and Craig Boutilier. SPUDD: Stochastic planning using decision diagrams. In *UAI*, pages 279–288, 1999.

[3] Diederik M. Roijers. *Multi-Objective Decision-Theoretic Planning*. PhD thesis, Univ. of Amsterdam, 2016.

[4] Diederik M. Roijers, Joris Scharpff, Matthijs T. J. Spaan, Frans A. Oliehoek, Mathijs de Weerdt, and Shimon Whiteson. Bounded approximations for linear multi-objective planning under uncertainty. In *ICAPS*, pages 262–270, 2014.

[5] Joris Scharpff, Diederik M. Roijers, Frans A. Oliehoek, Matthijs T. J. Spaan, and Mathijs M. de Weerdt. Solving transition-independent multi-agent MDPs with sparse interactions. In *AAAI*, pages 3174–3180, 2016.

[6] Joris Scharpff, Matthijs T. J. Spaan, Mathijs M. de Weerdt, and Leentje Volker. Planning under uncertainty for coordinating infrastructural maintenance. In *ICAPS*, pages 425–433, 2013.

Radical Reddits - Into the mind of online Radicalised Communities

Paul Verhaar[a] Maya Sappelli[b]

[a] *Utrecht University, humanities department. Domplein 29, 2512 JE Utrecht*
[b] *TNO The Hague. Anna van Buerenplein 1, 2595 DA Den Haag*

Abstract

We have investigated language use in radical communities using automated approaches. A Naive Bayes classifier was capable of predicting whether a Reddit post was radical or not with 75% accuracy. Moreover, the language use showed characteristics of a virtual community, based on its salient language features. This work is a step towards automatic detection of radicalism online.

1 Introduction

Nowadays there is an increasing amount of radical communication that is taking place on online platforms such as Reddit, Facebook or Twitter. Recently, Twitter closed more than 200.000 accounts for violating policies related to violent threats and promotion as terrorism. I investigate the use of online media for radical communications from the perspective of new media and linguistics.

2 Method

The conveyance of radical beliefs has been widely researched offline [2]. However, the amount of research for online sources remains scarce [7]. Recent studies looking into the effect of the internet on radicalism have shown that non-US based extremists were more likely to radicalise through virtual tools [4]. The web enhances possibilities for radicalised groups to communicate, organise and plan activities [3]; they use the power and the freedom of social media platforms to exercise pressure, power and influence across the globe [5]. Using social media in such a way mostly leads to a virtual community, which shows that all information that is spread (online) contributes to the shaping of communities and a "we" sense [6]. It shows that several studies have sought to understand how radicalism online comes to exist and grow. Typically, these studies are focused on keyword analysis. In this paper we also explore machine learning algorithms for analysing and predicting salient language features within online radicalised media.

In this paper, we demonstrate a proof of concept for a classifier that predicts whether a post is radical or not. The characteristics of the classifier give insight in salient language features within online radical discourse. Moreover, we analysed the language use in radical discourse from a linguistic perspective. Data was gathered from the social media platform Reddit as part of the VOX-Pol project. The data were crawled via the Reddit API from October 2007 to May 2015 and was divided in two classes: radical and non-radical in nature. It contains 105,930,239 collected from 239,773 threads. The radical class contained clear distinguishable radicalised sub-reddits such as FeministHate, nationalism and WhiteRights, which clearly promotes radicalism in a variety of beliefs such as anti-feminism or white supremacy. For this paper, a million posts per class were used for analysis, making the two classes equal in size. The posts were divided into a 40/60 ratio for the train and test set for accuracy measurements. Data and the adherent labels were loaded into Python. Pre-processing was executed using NLTK and SciKit. The pre-processing pipeline consists of stemming, stopwords removal, and filtering of non-English posts. Unigrams were weighted using TF-IDF to train the Naive Bayes classifier.

3 Results and Conclusion

Results showed that the classifier achieved an accuracy of 75% on correctly labelling documents as belonging to the radical or non-radical discourse. The radical class contains salient language features, such as “white male”, “rape victim” or “black person”. These words are frequently used in combination throughout all radical posts. The data display a stark difference in the distribution of collocations such as “sexual assault” and “hate women”. The collocations within the radical dataset contain important features typical of their discourse within their virtual community [6]. The linguistic analysis shows that the language within the radicalised communities are significantly different from the language in non-radicalised communities. The notion of community is less apparent in the non-radical discourse, which is shown by the difference in language patterns compared to the radical discourse [1]. This is shown in the combined word use of “nigger” and “black” over time within the radical community for example.

This paper shows that multidisciplinary methods yield interesting results and should be further explored. This work successfully builds on previous theories from the field of new media and linguistics. This work sheds light on research methods applicable for detecting online radicalism within the Reddit domain. The techniques used in this paper allow for use on other (radicalised) communities as well as building towards automatically detecting radicalism online.

References

[1] B. Anderson. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London, UK: Verso, 1991.

[2] D. Barton and C. Lee. 2013. Language Online: Investigating Digital Texts and Practices. *Studies in Second Language Acquisition, 224*, 2013.

[3] C. Easttom and J. Taylor. Computer crime, investigation, and the law. *Course Technology,* Boston, MA, 15-28, 2011.

[4] P. Gill, C. Emily and A. Thornton. 2015. What Are The Roles Of The Internet In Terrorism?

[5] C. Graham. Terrorism.com: Classifying Online Islamic Radicalism as a Cybercrime, 2013.

[6] H. Rheingold. *The Virtual Community*. MIT Press, 2000.

[7] H.A. Schwartz, J.C. Eichstaedt, M.L. Kern, L. Dziurzynski, S.M. Ramones, M. Agrawal and L.H. Ungar. Personality, gender, and age in the language of social media: the open-vocabulary approach. *PloS One, 8(9), e73791*, 2013.

Content-Aware Image Resizing to improve the Object Detection rate in Aerial Imagery *

Anouk Visser

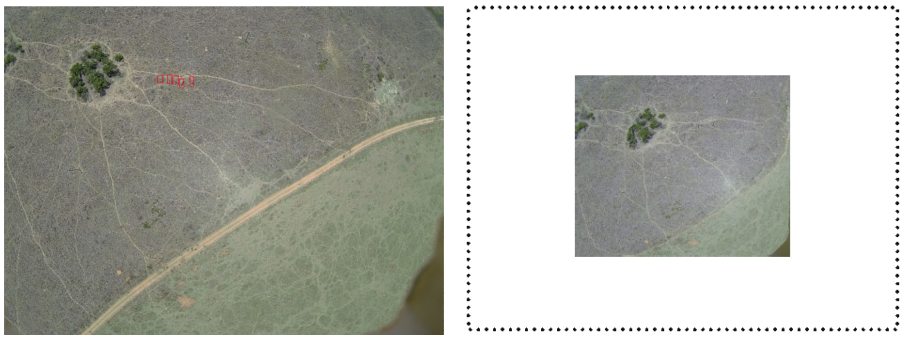
Universiteit van Amsterdam

Abstract

Unmanned aerial vehicles allow rangers to survey much larger areas of game reserves than possible on foot. However, from their high viewpoint, less than 1% of the pixels contain the wildlife the rangers like to preserve. This means that more than 99% of the pixels are not relevant. Yet, reducing the size of the image without being aware of the content can remove precisely the important pixels which contain wildlife. Content-aware image resizing avoids the removal of important pixels. The reduction in image size makes it possible to apply advanced object detection algorithms on the images and transfer the images much faster to the ranger's base station.

1 Introduction

To combat poaching or perform game counts nature conservationists need to inspect areas that are very large and hard to reach by car or foot. Recently, nature conservationists have been able to inspect these areas more easily by using unmanned aerial vehicles equipped with downward looking cameras (see for an example Fig. 1). Despite of the ease with which these systems can be deployed, the recorded imagery still needs to be analyzed manually. Automatic object detection algorithms greatly reduce the time spent looking for the object of interest, benefiting the conservation work.



(a) (b)

Figure 1: (a) Example output of an automatic object detection algorithm on an image taken from our dataset recorded in Welgevonden Game Reserve, South-Africa. The red boxes indicate the rhinos to be found (covering 0.07% of the image). (b) The same image 50% resized. Note the yellow road at the lower-right corner, with non-informative pixels, is nearly gone from the image.

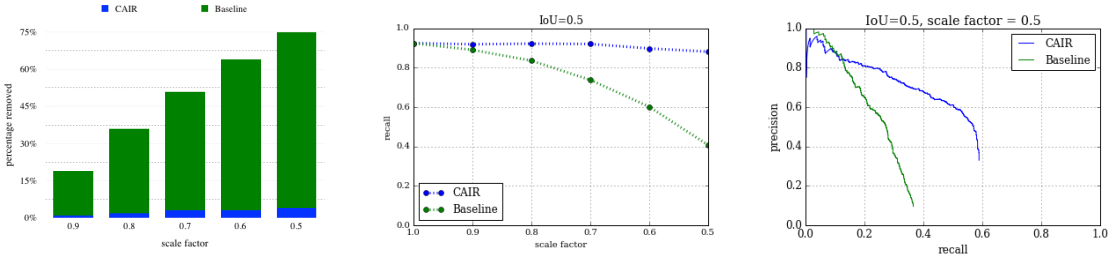
*This is an abstract of a thesis that will be defended as *Master thesis, Universiteit van Amsterdam, October 2016.*

2 Method

State-of-the-art object detection algorithms such as Fast R-CNN [2] rely heavily on object proposals that are used to provide a speedup in the object detection pipeline. Yet, proposal methods such as Edge Boxes [4] are significantly slower when applied to high resolution images, which affects the detection rate [3]. To maintain a good object detection rate, we apply a content-aware image resizing method (CAIR) [1] that resizes the image without compromising on the content.

3 Results

As indicated in Fig. 2a, the percentage of removed important pixels is much higher when we resize the image with bilinear interpolation (the not content-aware baseline) and CAIR (content-aware). The baseline removes 75% of the important pixels when resizing with scale factor 0.5, CAIR only removes 4%. This has effect on the recall of Edge Boxes; if we compare resizing the image with the baseline and CAIR the recall is far less affected for CAIR, as indicated in Fig. 2b. As a result, CAIR outperforms resizing with bilinear interpolation in terms of precision and recall of the Fast R-CNN object detector as shown in Fig. 2c.



(a) (b) (c)

Figure 2: (a) Comparison of the percentage of important pixels that were removed from the image by CAIR to rescaling with bilinear interpolation (not content-aware). (b) Recall of object proposals after resizing the image with CAIR against rescaling with bilinear interpolation. (c) Precision-recall curves after resizing the image with CAIR against rescaling with bilinear interpolation.

4 Conclusion

By applying content-aware resizing one can reduce the image size in a smarter way, with less effect on the quality of the object detection methods. This makes it possible to apply advanced object detection algorithms with the limited computing power of unmanned aerial vehicles and transfer the images much faster back to the ranger's basestation.

References

[1] Shai Avidan and Ariel Shamir. Seam carving for content-aware image resizing. In *ACM Transactions on graphics (TOG)*, volume 26, page 10. ACM, 2007.

[2] R. Girshick. Fast r-cnn. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1440–1448, Dec 2015.

[3] Jan C. van Gemert, Camiel R. Verschoor, Pascal Mettes, Kitso Epema, Lian Pin Koh, and Serge Wich. *Nature Conservation Drones for Automatic Localization and Counting of Animals*, volume 8925 of *Lecture Notes in Computer Science*, pages 255–270. Springer, 2015.

[4] C Lawrence Zitnick and Piotr Dollár. Edge boxes: Locating object proposals from edges. In *Computer Vision–ECCV 2014*, pages 391–405. Springer, 2014.

Lack of effort or lack of ability? Robot failures and human perception of agency and responsibility[1]

Sophie van der Woerdt[a] Willem F.G. Haselager[b]

[a]Dpt. of Psychology, [b]Dpt. of Artificial Intelligence, Radboud University, Comeniuslaan 4, 6525 HP, Nijmegen

Abstract

Research on human interaction has shown that attributing agency to another agent has substantial consequences for the way we perceive and evaluate its actions. Specifically, considering an agent's actions related to either effort or ability can have important consequences for the attribution of responsibility. This study indicates that participants' interpretation of a robot failure in terms of effort –as opposed to ability– significantly increases their attribution of agency and –to some extent– moral responsibility to the robot. However a robot displaying lack of effort does not lead to the level of affective and behavioural reactions of participants normally found in reactions to other human agents.

1 Introduction

Currently, much debate is devoted to the question of how we should deal with harm caused by robots. Research on anthropomorphism, blame and examples of media and pop culture speaking of 'robot laws' underline the possibility of humans –perhaps inappropriately– attributing moral responsibility to automated systems. Although legal solutions have been proposed for dealing with such conflicts [1], in daily life this may still have undesired implications. Owners and developers of robots may (unknowingly) distance themselves from potential harms caused by their robots, causing responsibility to become diffused. Therefore, it is relevant to find out what factors contribute to the attribution of agency and responsibility in robots.

Extensive work on attributional processes in human interaction reveals that the perception of an agent's effort and abilities are central determinants in the attribution of agency and moral responsibility [2]. This, in turn, is strongly related to fundamental affective and behavioural reactions such as sympathy, rejection, altruism and aggression [2, 3]. Yet, with regard to human robot interaction (HRI), little is known about the attribution of agency and moral responsibility.

In this study, we applied Weiner's *Theory of Social Conduct* [2] to HRI by showing participants videos of robots (Aldebaran's NAO) failing tasks in ways that could be interpreted as due to either *lack of ability* (LA-condition; e.g. dropping an object) or *lack of effort* (LE-condition; e.g. throwing away an object). We expected that a display of *lack of effort* would incite the illusion of a robot having *agency* over its actions. In addition, we expected that a robot's *lack of effort* would have little effect on the attribution of moral *responsibility* to the robot, compared to a display of *lack of ability*.

2 Method

In an online survey, sixty-three participants (M_{Age} = 25,5, *SD* = 9,7; drawn from a university population) were shown a video of about 30-60 seconds portraying a situation in which a NAO robot was shown failing a task either due to *lack of ability* or *lack of effort*. Seven of such scenarios were presented[2]. After each video, participants were asked to fill in a questionnaire containing scales of *agency* (five questions about the robot's control over the situation and its ability to make its own decisions), and *responsibility*

(twelve questions on attributed blame and kindness, affective and behavioural reactions). Additionally, scales were included measuring the participant's estimate of the robot's *experience* (e.g. having beliefs, desires, intentions, emotions), *predictability, propensity to do damage, trustworthiness* and *nonanthropomorphic features* (e.g. strength, efficiency, usefulness)[3].

For analysis, mean scores of each scale (range 1-5) were calculated and transposed to Z-scores. Since reliability and goodness-of-fit for the scale of *responsibility* was questionable, items of this scale were analyzed separately. In order to answer our main questions, a GLM multivariate analysis was performed with the composite means of *agency*, *experience*, *predictability*, *propensity to do damage*, and each item related to *responsibility* as dependent variables. *Condition* (LA/LE) was indicated as between-subject factor.

3 Results

According to what was expected, participants attributed more agency to a NAO robot after seeing videos in which it displayed *lack of effort* (M = 2.80, SD = 0.82) compared to videos in which it displayed *lack of ability* (M = 2.12, SD = 0.61). Univariate tests expressed significant and large effects for the composite scores of *agency* ($F(1,61) = 13.601$, $p = .000$, $eta^2 = .182$), *experience* ($F(1,61) = 12.235$, $p = .001$, $eta^2 = .168$), and *predictability* ($F(1,61) = 14.040$, $p = .000$, $eta^2 = .187$). The results for the items of *responsibility* were mixed. While univariate tests for *blame* and *disappointment* revealed significant, medium effects (respectively: $F(1, 61) = 5.757$, $p = .019$, $eta^2 = .086$; $F(1, 61) = 9.704$, $p = .003$, $eta^2 = .137$), effects for the items *anger, put away, sell, kindness, pity, sympathy, help* and *try again* were not significant.

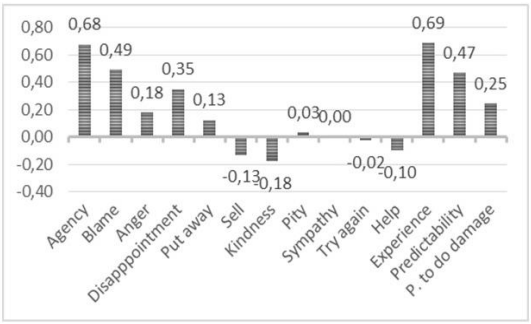


Fig. 1: Difference scores of means within the *lack of ability* and *lack of effort* (LA - LE) conditions.

4 Conclusion

Similar to findings related to human interaction, the results of our study reveal that, in case of robots displaying behaviour that can be interpreted as lack of effort, humans tend to explain robotic behaviour by attributing agency. In case of failure, a robot displaying lack of effort -essentially refraining from 'trying'- may lead to blame and disappointment. However, it does not necessarily lead to negative affective and behavioural reactions such as anger, or wanting to shut the robot off and put it away. Results like these emphasize that we should be aware of potential diffusion of human responsibility when (advanced) robots create the impression that they are agents in the sense of actually controlling and intending their own actions. Our results also suggest that –in case of NAO robots- failure, or even reluctance for doing tasks is received well, illustrating a promisingly positive view on robots.

References

[1] P. Asaro, P. A body to kick, but still no soul to damn: Legal perspectives on robotics. In N.P. Lin, K. Abney, and G. Bekey, editors, *Robot Ethics: The Ethical and Social Implications of Robotics*. MIT Press, Cambridge, MA, 2013.

[2] B. Weiner. *Judgments of responsibility: A foundation for a theory of social conduct*. Guilford Press, New York/London, 1995.

[3] U. Rudolph, S.C. Roesch, T. Greitemeyer, and B. Weiner. A meta-analytic review of help giving and aggression from an attributional perspective. *Cognition and Emotion*, 18(6):815–848, 2004.

[1] The full thesis was submitted in fulfilment of the requirements for the degree of Bachelor of Science (Honours) in Psychology at the Radboud University in August 2016.

[2] Videos and complete survey can be found online: https://www.youtube.com/playlist?list=PLSQsUzV48QtG__YPY6kVcgCM8-YOcNqja; https://eu.qualtrics.com/jfe/preview/SV_6y4TuTii0CFnpch

[3] In this abstract we chose to focus on our main questions only. Therefore, additional analyses and results regarding these variables will not be discussed.

A Library of Meta-Level Diagnostics and Repairs for Fluid Construction Grammar

Katrien Beuls [a] Paul Van Eecke [ba] Luc Steels [ca]

[a] *AI-lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel*
[b] *Sony Computer Science Laboratory Paris, 6, Rue Amyot, 75005 Paris*
[c] *Institute for Evolutionary Biology (IBE, UPF-CSIC), Dr. Aiguader 88, 08005 Barcelona*

Abstract

When processing natural language, one of the main challenges is to handle input that cannot be routinely processed by the language model. In order to deal with this challenge, it is necessary that problems of any kind (e.g. unknown words, deviant word ordering) are accurately diagnosed. These diagnostics can then trigger repairs, which try to find a solution for the problem. If the proposed solution proves to be appropriate, it can be stored and reused in the future. We will give a live demonstration of the use of a new library of diagnostics and repairs for language processing, that can handle many common issues. The library is implemented in Common Lisp and designed to be used in combination with the Fluid Construction Grammar platform.

1 Routine and Meta-level Processing

Any computational system that deals with human language needs to be able to handle unexpected input, including lexical or grammatical innovations, incomplete utterances, or even downright errors. One commonly used architecture for handling problems consists of a routine layer for normal processing, and a meta-layer for problem solving and learning [2]. In this architecture, processing always starts in the routine layer. At regular intervals, a set of diagnostics is run on the current state, and problems are reported in detail. If there are any problems, the system jumps to its meta-layer, where repairs try to handle the problem. Once it is solved, the system jumps back to the routine layer and normal processing can continue. If a repair turns out to have solved a problem successfully, the solution can be stored and later be reused in routine processing.

Fluid Construction Grammar (FCG)[1] [3, 4], a fully operational computational formalism for semantic parsing and production, has such an integrated meta-layer with support for diagnostics and repairs. There is a large body of previous work on this topic [1, 7, 6], providing various examples of how FCG's meta-layer can be used for making processing more robust and for learning new grammar rules.

We have further integrated this meta-layer into the core of FCG, making it easier and clearer for the user. Furthermore, we have designed and implemented a library of diagnostics and repairs that addresses many common issues in language processing. The general diagnostics and repairs in the library can be directly used in FCG grammars, function as templates for new, more grammar-specific diagnostics, or serve as inspiration for problem-solving in other grammar formalisms.

At BNAIC, we will give a live demonstration of (i) how FCG's meta-layer can be used in concrete grammars, (ii) how the library of diagnostics and repairs makes language processing more robust, and (iii) how new constructions can be learned using this library[2].

[1] Fluid Construction Grammar is freely available at `http://www.fcg-net.org/download`

[2] An interactive web demonstration with part of what will be shown at BNAIC can be found at `http://www.fcg-net.org/demos/bnaic-demo`

2 Library of Diagnostics and Repairs

For many common issues in language processing, the library contains `diagnose` methods to diagnose the problem and `repair` methods to create a suitable solution. All repairs create solutions in the form of constructions, which can at the end of processing easily be consolidated by adding them to the grammar. The following diagnostics and repairs are part of the library:

	Problem	Diagnostic	Repair
Out-of-Vocabulary word	In comprehension, one of the words in the input cannot be processed by the grammar.	No more constructions can apply, but there are some words left in the input.	A new, general lexical construction is created for this word. Word class and semantic category are initially underspecified. For its meaning, a new predicate is introduced.
Out-of-Vocabulary meaning	In production, one of the meaning components in the input cannot be expressed using the grammar.	No more constructions can apply, but there are some unexpressed meanings left in the input.	A new, general lexical construction is created for this meaning. Word class and semantic category are initially underspecified and a new word is introduced.
Non-integrated meaning components	At the end of processing, the analysis yields a semantic network with a non-integrated component.	No more constructions can apply and the resulting semantic network is not fully connected.	A new phrasal construction is introduced. It captures the observed word order and integrates the meaning component by making it co-referencing.
Matching conflicts (e.g. gender mismatch, deviant word order)	At the end of processing, the resulting semantic network is not fully connected because some grammatical constructions could not apply.	No more constructions can apply, the resulting semantic network is not fully connected and the other diagnostics did not trigger.	The grammatical constructions are anti-unified with the current state and the anti-unified construction with the lowest cost is applied. Then, the resulting construction constrained towards the current state using pro-unification [5].

Figure 1 shows the application process of an utterance in which an out-of-vocabulary word is diagnosed. The orange node signals that a problem is diagnosed and the yellow node shows that a repair has been active. In this case a new lexical construction for the word ”cat” has been created. The following green nodes show that routine processing could continue and the final dark green node signals that a solution has been found.

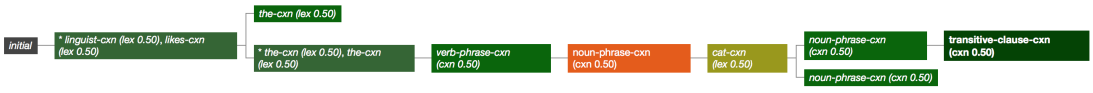


Figure 1: Construction application process for the utterance ”the linguist likes the cat”, in which ”cat” is an unknown word for the grammar. In the orange node, the ”out-of-vocabulary word” diagnostic triggers, in the yellow node a lexical construction for ”cat” is introduced by the repair.

3 Availability and System Requirements

Fluid Construction Grammar is released under an Apache 2.0 license and is freely downloadable at `http://www.fcg-net.org/download`. The library of meta-level diagnostics and repairs will be added to the release soon. Both FCG and the library are implemented in Common Lisp and are compatible with Clozure Common Lisp (CCL), Steel Bank Common Lisp (SBCL) and LispWorks (LW). It runs on Linux, OSX and Windows.

References

[1] Katrien Beuls, Remi Van Trijp, and Pieter Wellens. Diagnostics and repairs in fluid construction grammar. In *Language grounding in robots*, pages 215–234. Springer, 2012.

[2] John E Laird. *The Soar cognitive architecture*. MIT Press, 2012.

[3] Luc Steels. *Design patterns in fluid construction grammar*, volume 11. John Benjamins Publishing, 2011.

[4] Luc Steels. Basics of fluid construction grammar. *Submitted*, 2016.

[5] Luc Steels and Paul Van Eecke. Insight grammar learning using pro- and anti-unification. *Under Review*, 2016.

[6] Luc Steels and Remi van Trijp. How to make construction grammars fluid and robust. In *Design patterns in fluid construction grammar*, volume 11. John Benjamins, 2011.

[7] Remi van Trijp. A reflective architecture for robust language processing and learning. In *Computational issues in fluid construction grammar*, pages 51–74. Springer, 2012.

SWISH DataLab: a web interface for data exploration and analysis[1]

Tessel Bogaard Jan Wielemaker Laura Hollink
Jacco van Ossenbruggen

Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands

1 Introduction

Data is ubiquitous, and so are tools supporting data analysis. More often than not, different tools are used for different stages of the analysis. For example, the exploration and preprocessing of data is handled in one tool using one programming language and the analysis in a completely different environment, with scripts spread out in different files stored locally. SWISH DataLab[2] is a shared web interface combining both exploration, preprocessing and analysis of data, supporting transparency of the choices made. It combines the clarity of Prolog with the statistical computing power of R[3]. Data cleaning and creating concepts and abstractions over the data benefits from the elegance of Prolog's rule-based logic programming paradigm; the statistical analysis and visualization are the strength of R. Using SWISH DataLab, it is easy to quickly try out different data abstractions on a sample and evaluate the expected impact on the results. SWISH DataLab is currently being developed as an environment for responsible data science using data from the National Library of the Netherlands.

2 Case study: analysis of online user search behavior

We have collected six months of server logs[4] from the full text search platform[5] providing acces to combined collections from the National Library of the Netherlands and other national heritage and research institutes, spanning a period from October 2015 to March 2016. This dataset is through DataLab connected to a second dataset, the historical newspaper collection with over 1M documents across four centuries. This collection is—like other digital libraries—characterized by bibliographic data describing the content (e.g. publication date, type of document, distribution zone) and represented in the user interface by facets. These metadata values are linked to the clicked and downloaded documents enabling a comparison between clicked and searched metadata categories.

3 Demo scenario: rule-based data exploration

In this demo we will show how SWISH DataLab was designed for exploration, (pre-)processing, visualization and analysis of data. We will explain how the interactive analysis and visualization helps to come to better concepts and abstractions over the data; and makes the process of data cleaning more transparent.

[1]The development of SWISH DataLab was partially supported by the VRE4EIC project, a project that project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 676247

[2]A version of SWISH for teaching Prolog is available online: `http://swish.swi-prolog.org/`.

[3]`https://www.r-project.org/`

[4]Data received under strict confidentiality agreement

[5]`http://www.delpher.nl/`

The goal for this use case is to describe user search behavior, and this has motivated an exploration of the server logs based on user interactions within sessions. Sessions are defined based on IP address and a 30 minute timeout. In order to recognize usage patterns, we have visualized sessions in graphs (see Fig. 1). As an added benefit, the graphs help in conceptualizing the data. For example, the graphs made visible that users visit the same search engine results page often in a session, leading to a Prolog rule where we abstract this to a single node with more than one incoming arrow. We removed repeated visits of the same web page, where this is likely a reload of the web browser and not a new interaction by the user. This has resulted in a better count of the number of clicks and a clearer definition of the dwell time on a document (as we time this from the first load of the page until a new interaction and not as separate shorter dwell times).

These incremental rules to clean the data and visualize the graphs can be traced in the web environment. All intermediate versions of the code used for data cleaning are stored by user id. A researcher can not only review the changes made before saving them, but can also return to a previous stage, making it easier to evaluate the impact on the results and supporting transparency and reproducibility.

SWISH DataLab enables a researcher to perform statistical analysis and visualization of the data using R, integrated into the web interface. We can use all the available libraries for R, such as ggplot2 for vizualization and can also evaluate the impact of concepts and abstractions with R. For example, we inspect the statistical effect of different symbolic definitions of repeated visits to the same search engine results page.

4 Conclusions and future work

With SWISH DataLab we are moving closer to a single environment for responsible data science shared between researchers. The use of graph visualization in combination with concepts and abstractions defined as rules makes data cleaning more transparent and more thorough, and patterns in user interactions more insightful. Future work on the use case will include a better definition of a session based on the graphs and machine learning for clustering the graphs. For SWISH DataLab future work also includes an integration with Monet DB.

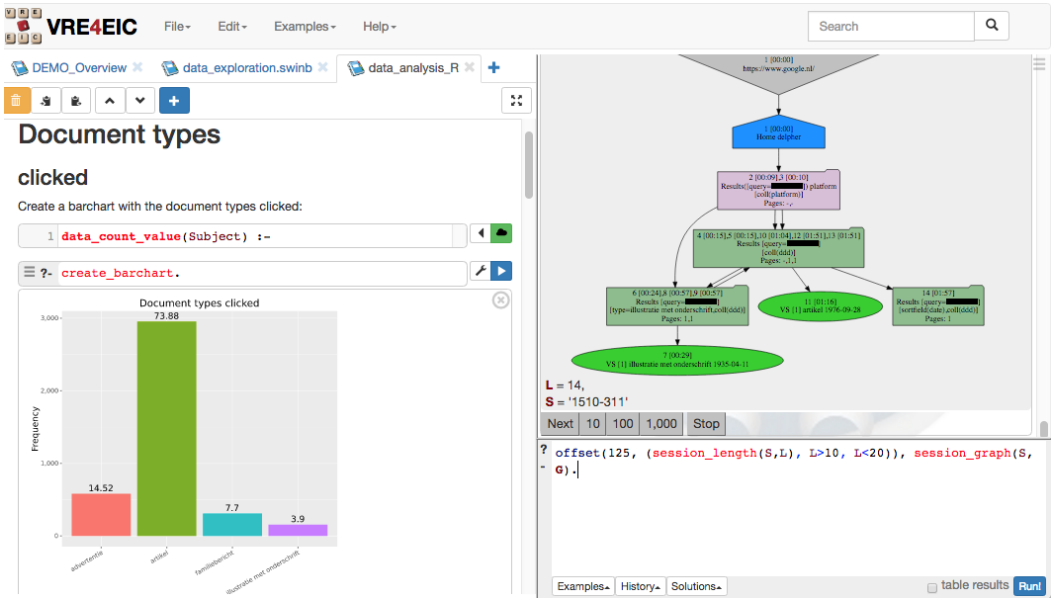


Figure 1: Screenshot from Swish DataLab, showing visualizations of the data using Prolog and R. On the right hand side we can see a graph visualizing a single session; on the lefthand side we can see a bar chart created using ggplot2 showing the document types clicked in the data sample.

TINKR
A skill-based online learning and career development platform

Quincy Dalh[a] John C. Lokman[a,b] Kishan Nirghin[a,b]

[a] *FRISSR BV. De Boelelaan 1081 1081HV Amsterdam*
[b] *Vrije Universiteit Amsterdam Department of Computer Science De Boelelaan 1081 1081HV Amsterdam*

Abstract

FRISSR is a startup company established by young entrepreneurs and scientists. Its collaborators include Vrije Universiteit Amsterdam and the municipality of Amsterdam. FRISSR focuses on the development of new technologies for academic education, career development (lifelong learning), and also provides matchmaking services between students and companies based on possessed and sought skills. FRISSR's approach to matters of education, career, and job market is to see individuals, universities, and job vacancies as a network of skills. Just like an individual can be seen as a network of skills, so can an organization or university (i.e., a network of skills possessed by an individual, a network of skills being sought by a company, and a network skills being taught by a university). FRISSR's unique value proposition is to map these three networks onto each other with a high level of granularity. This would mean constructing a detailed '*skill map*' individuals, universities, and vacancies. Such a detailed skill map would allow unparallelled precision in recommending individuals the educational content they need, and direct them to companies that seek their skills. It is aimed to accomplish this goal through development of an AI-powered learning environment, 'Tinkr', and its counterpart 'FRISSR Careers'.

1 Background

Since the beginning of 1990s, the World has begun to experience a transition to a knowledge economy[1] [1]. This transition brings two important changes to society and nature of jobs: (a) Primary workforce in Western societies are shifting from service sector to knowledge and tech sector [3] and (b) jobs increasingly require specific technical and practical skills (i.e., employers are rather looking for a person with C++ programming skills than a person who has obtained a Masters degree in computer science) as well as a higher level of education in general [4].

As a knowledge economy is based on specialized skills, its success largely depends on a successful match between skills of the labor force and the skills required by organizations. Indeed, a major problem for the knowledge economy is the mismatch between the skills individuals possess and the skills required by employers: the so-called skill gap. It is estimated that in the European ICT sector alone, the skill gap will result in a shortage of 900,000 ICT professionals in 2020 [5]. This mismatch is expected to significantly hamper economic recovery in Europe.

In recent years, several studies [6, 7, 8, 9, 10] have tried to identify solutions to tackle the skill gap. Most studies agree that the following actions should be taken:

- Develop mechanisms that can identify and predict required skills in the market;
- Develop frameworks which offer standardized means of certification and validation of skills across EU member countries[2];
- Develop innovative educational systems which include e-learning. Rapid technology change requires rapid ways to obtain new skills. E-learning needs to play a larger role in supporting the diffusion of advanced skills by allowing for modularised educational offerings, that do not require workers to complete an extensive academic programme to update their skills;
- Develop matchmaking mechanisms which would allow employers and employees to 'match' the skills needed versus the skills required.

In order to address the issue of skill gap and other demands of the the rapidly emerging knowledge economy, FRISSR is developing a new learning environment, '*Tinkr*' and its accompanying career service. '*FRISSR Careers*'. While Tinkr is aimed at delivering education through a modern, interactive, and gamified e-learning platform; FRISSR Careers is aimed at using the learning history of Tinkr users to match them with vacancies that fit their skills with unparalleled accuracy.

2 How Does Tinkr Work?

2.1 Courses and Learning Experience

Evolving and adaptive courses. Both course creators and users are allowed to contribute learning material to Tinkr courses. This way, courses will be be improved / or edited by users and creators alike. For

[1] A *knowledge economy* can be defined as a social system where the primary medium of value generation is knowledge. It differs from economies of other eras such as Agricultural economy (during Pre-industrial Age, whose primary workforce had been agrarian sector), industrial economy (during Industrial age, whose primary workforce had been manufacturing sector), and mass-production economy (during mid 1900s, whose primary workforce had been service sector) [1]. Although it can be argued that all previous economic eras has been knowledge and specialized skill-based to a certain extent, the importance of specialized knowledge never been as pronounced as it is in today's economy [2].

[2] Through the New Skills for New Jobs initiative#, the EU already started to implement this solution for some sectors. For the ICT sector, a European e-Competence Framework has been developed.

purposes of of quality control, each educational material contributed will be treated in the same way apps are treated in Google's Play Store or Apple's App Store, and their ranking will be determined by community ratings. Such an open system will also allow adding (and filtering) course material in different formats such as documents, lectures, animations, documentaries, sandboxes, and so on. This will enable users to learn in a way that is compatible with their learning style and preferences. In later stages, usage of ontologies (in order to establish a map for skills in each domain), learning analytics and neural networks (in order to provide an adaptive learning mechanism to match educational materials to user's skill level, and for creating a personalized learning experience through predicting the format of educational materials individual users would prefer) is planned.

Learning experience. Despite the large increase in demand for online education, lack of student motivation and dropout rates are an important challenge online education services must overcome [11, 12]. Tinkr aims to reduce dropout rates by (a) delivering educational material that is highly relevant to users, (b) creating a gamified learning environment, (c) implementing storytelling elements (e.g., an AI assistant TinkrBell), and (d) cooperation with universities and organizations (for a discussion, see [11] and [12]).

2.2 Skill ID

Once a Tinkr lesson is successfully completed, the newly acquired skill is (e.g., skill in NodeJS runtime) is added to a personal skill map of the learners—the so-called '*Skill ID*[3]'. Because users are provided with a map of their own unique network of skills, much like a mirror of their minds, they can explore their skill space, make informed learning decisions, and maximize their talents in less time.

2.3 Career Advice and FRISSR Careers

Vacancy ID*, *Company ID, and career recommenders. FRISSR Careers will create maps associated with vacancies and organizations (through web crawlers). The Vacancy ID will allow users to see what skills a vacancy boils down to in a visual graph. This graph will be directly comparable to users' own Skill IDs and highlight overlaps and gaps. Similarly, employers will also be able to browse users' Skill IDs. Beyond allowing individuals and companies to explore each other in a (micro-)skill-based manner, Tinkr will also serve as a matchmaking platform. Through above mentioned IDs (which will be partly created by user histories and partly by web crawlers) and recommenders, it will be able to match individuals with vacancies and organizations that are most likely to fit their network of skills, and vice versa. Finally, in case of skill gaps, FRISSR Careers and Tinkr will be able to provide strategic learning advice (i.e., what skills to learn next) depending on which companies or careers individuals are interested in. Furthermore, as the landscape of a profession changes, Tinkr users will be able to see what skills they are missing out on, and update their skills in a more strategic manner. As the system is developed further. such functions are expected to be enhanced with usage of semantic web technologies[4].

At BNAIC 2016, a MVP(+) of Tinkr will be presented.

Links

Alpha version of Tinkr:
http://alpha.tinkr.nl

Interactive Presentation of FRISSR and Tinkr:
https://prezi.com/vl-tjxn7ewuf/frissr/?utm_campaign=share&utm_medium=copy#

[3] The Skill ID is based on the European e-Competence Framework (http://www.ecompetences.eu/; accessed 26.10.2016.), a platform that helps in mapping the micro-level skills of individuals.

[4] For a similar system that uses semantic web technology to connect individuals and employers in based on skills, see TextKernel.com (http://www.textkernel.com/).

References

[1] Amidon, Debra M., Piero Formica, and Eunika Mercier-Laurent, eds.*Knowledge economics: emerging principles, practices and policies*. Tartu, Estonia: Tartu University Press, 2005.

[2] Rotherham, Andrew J., and Daniel T. Willingham. “21st-Century” Skills. *American Educator,* 17: 2010.

[3] Frey, Carl Benedikt, and Michael A. Osborne. The future of employment: how susceptible are jobs to computerisation, 2013.

[4] Carnevale, Anthony P., Nicole Smith, and Jeff Strohl. Recovery: Job growth and education requirements through 2020, 2013.

[5] Neef, Dale, Gerald Anthony Siesfeld, and Jacquelyn Cefola. *The economic impact of knowledge*. Routledge, 1998.

[6] Delfino, Manuela, Giuliana Dettori, and Valentina Lupi. *Task-based learning and ICT: creative activities in the context of a European project.*

[7] Dettori, Giuliana, and Valentina Lupi. Using a narrative blog to support reflection in a blended course. *International Conference on Hybrid Learning and Education*. Springer Berlin Heidelberg, 2009.

[8] de Leeuw, K. Serious Gaming in Nederland En in Het Buitenland : Opleidingen En Perspectieven. Amsterdam: Universiteit van Amsterdam, 2013.

[9] Glor, Eleanor D. Key factors influencing innovation in government. *The Innovation Journal* 6(2): 1-20, 2001.

[10] Majdán, Gábor. "The success and sustainability of Online Communities: a Social Network Analysis approach." *Unpublished manuscript, Utrecht University,* 2012.

[11] Park, Ji-Hye, and Hee Jun Choi. Factors Influencing Adult Learners' Decision to Drop Out or Persist in Online Learning. *Educational Technology & Society,* 12(4)*:* 207-217, 2009.

[12] Lee, Youngju, and Jaeho Choi. "A review of online course dropout research: Implications for practice and future research." *Educational Technology Research and Development,* 59(5): 593-618, 2011.

RoboTutor: A Robotic Teaching Assistant [1]

Koen V. Hindriks [ab] Joost Broekens [ab] Joachim de Greeff [ab]

[a] *Delft University of Technology, Mekelweg 4, 2628CD, Delft*
[b] *Interactive Robotics B.V., Delft*

Abstract

Robots are being introduced into more and more classroom environments. They are envisaged as tools that will stimulate pupils to learn about programming and computational thinking. In our demonstration we aim to show that robots can also be effectively used as teaching assistants in the classroom. Towards this end we demonstrate the expressive capabilities of the Nao robot, in particular mood expression and the robot's interactive potential as a teaching assistant. This not only supports the teacher but also allows pupils to familiarize themselves with rapidly advancing technology.

1 Introduction

The use of robots in education has been promoted by many, e.g. [1, 4]. Robots have mostly been used as tools to teach pupils about programming. We propose the innovative use of robots as teaching assistants in the classroom in addition to their use as teaching tools for programming. Such a use as teaching assistants serves two important goals in addition to the robot's use as tools for teaching about programming.

First, robotic teaching assistants can provide support to a teacher by presenting a subject topic. The teacher is not replaced but complemented by the robot. We have developed a RoboTutor for frontal instruction, i.e. a robot that can provide presentations in front of a classroom. In this setting, the teacher no longer has to focus on presenting a topic but instead can pay attention to social cues and interact with individual pupils when needed. This enables 'flipping the classroom' in an innovative way and allows the teacher to focus more on the social interaction in the classroom and the way the content is absorbed by the students. Furthermore, the RoboTutor provides easy to use scripting integrated with presentation tools such as Microsoft Powerpoint, enabling teachers to develop content for the robot.

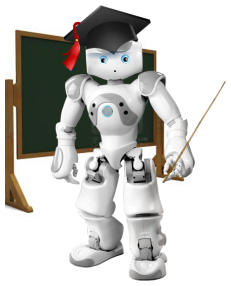


Figure 1: The Nao RoboTutor

Second, robotics teaching assistants can familiarize pupils with interacting with technologies that get smarter over time. In a society that will be dependent more and more on machines that are able to demonstrate some level of adaptation and reasoning, it is important that children learn how to engage and interact with such technology. Currently, children's perceptions of robots are biased by movies and cartoons, and in general an overly optimistic view of robot capabilities exists. A robotic teaching assistant provides a very suitable and engaging alternative for making pupils familiar with both the limitations as well as the possibilities of machines that are getting more advanced every day.

The use of robots in the daily environments of humans such as classrooms requires these robots to be able to smoothly interact with people. The challenge here is to address interaction of a robot with an *audience* of more than one person. Where most robot interaction studies have focused on one-on-one

[1] See for videos illustrating a lecture about robotics with the RoboTutor's mood modulated as either sad or happy `http://ii.tudelft.nl/SocioCognitiveRobotics/index.php/RoboTutorMood` and for a general impression of the RoboTutor `https://youtu.be/L6M8KSP7urg`.

interaction, a robot in the classroom needs to interact with a group. An important aspect of interaction consists in a robot being able to express emotions. In previous work [3], we have investigated the effects of mood expression in the classroom. To create a channel for content interaction, we have integrated both presentation as well as polling software. The presentation software allows the robot to display content on a screen, while the polling software allows students to provide input to the robot and the robot to respond to these inputs.

2 Demonstration

In our demonstration we focus on two important aspects of robotic teaching assistants: interaction and mood expression. The expression of mood is important in order to make the presentation of the robot more lively and engaging. It has been shown that students appreciate the robot more if the robot expresses a positive mood [3]. A robot that interacts with its audience during a lecture is also more engaging and more able to hold the attention of an audience.

We will demonstrate the Nao RoboTutor (see Figure 1) that we have developed. The robot will present a lecture about the question what a robot is. It will not simply present a definition but will establish one together with the robot's audience by means of question-answer interaction with the audience. This allows different audience groups to develop their own understanding of robots, and leaves room for variation such as whether a robot should be mobile or not.

Our demonstration will need a large (TV) screen for displaying the presentation and a table to position the robot on, with power sockets for screen, robot, and laptop. In addition, we need access to WiFi or a router to connect to the robot.

3 Conclusion

The use of robots as aids in the classroom rather than as tools is a quickly becoming a reality, as our demonstration will show. Teachers will benefit from these developments as they will be able to actively engage their pupils with the help of a robotic assistant. This will enable a teacher to focus more on the social cues and interaction in the classroom. By using the robot also as a tool, pupils will also be able to learn about the technical limitations and possibilities of current robot technology.

The robot platform developed as RoboTutor also has great potential for presenting professional presentations and for interacting in contexts outside the classroom, such as conferences and other types of events. The challenge is to adapt the robot's interaction style and behaviour to these new audiences.

The design and development of robot interaction with an audience creates many challenges. We aim in particular to make the interaction more engaging. Future developments we are thinking of are particularly interesting include asking and answering open questions from pupils and more responsive behaviour of the robot to classroom situations where pupils attention is dropping, indicated by e.g. the level of noise in a classroom [2].

References

[1] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, 19(1):61–84, June 2004.

[2] T. Salter, K. Dautenhahn, and R. Bockhorst. Robots moving out of the laboratory - detecting interaction levels and human contact in noisy school environments. In *Robot and Human Interactive Communication (ROMAN 2004)*, pages 563–568, Sept 2004.

[3] J. Xu, J. Broekens, K.V. Hindiriks, and M.A. Neerincx. Effects of bodily mood expression of a robotic teacher on students. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014.

[4] Zhen-Jia You, Chi-Yuh Shen, Chih-Wei Chang, Baw-Jhiune Liu, and Gwo-Dong Chen. A robot as a teaching assistant in an english class. In *Advanced Learning Technologies, 2006. Sixth International Conference on*, pages 87–91, July 2006.

ION Demonstrator: an interactive exploration of a multimodal news corpus[1]

Laura Hollink[a] Adriatik Bedjeti[a] Desmond Elliott[b]

[a] *Centrum Wiskunde & Informatica, Amsterdam, The Netherlands*
[b] *ILLC, University of Amsterdam, The Netherlands*

1 Introduction

Images are a powerful tool in the framing of a story in a news article [8]. Several studies have analyzed the content of news images to identify how a particular topic is framed [3, 6]. For example, in [4], 192 news photos were coded to determine the visual framing of conflict in magazines; in [2], a team of coders watched 45.3 hours of televised news to compare news culture in four countries. Since manual analysis of visual materials is time consuming, many of these studies are done on relatively small datasets. We hypothesise these types of studies will eventually be accompanied by automatic analysis of larger quantities of visual material. With this demo we aim to contribute to this research direction.

In the *Images in Online News* (ION) project, we explore to what extent the we can support studies into (visual) framing of a particular event or topic. For that purpose, we have collected a large corpus of online news articles, including images, and we designed a web demonstrator that allows a user to quickly and easily explore this corpus. The demo can be accessed at http://ion.project.cwi.nl/.

2 ION Dataset

The demo runs on the ION corpus [5], a collection of news articles of five news publishers: the US-based The New York Times, The Washington Post and The Huffington Post, and the UK-based The Daily Mail and The Independent. The articles were collected during a period of one year - from 13 August, 2014 to August 13, 2015 - and contains approximately 323,000 articles, making it larger than other datasets of news images and text (e.g. [7, 1, 11]). The complete 1-year coverage over multiple publishers ensures a broad scope in terms of topics, image quality, and editorial / political viewpoints.

Enrichment with text and image analysis tools The article texts are processed using the TextRazor API, a commercial NLP tool. We use it to obtain two types of annotations: (1) high-level categories (called "Coarse Topics" in TextRazor), which are selected from a limited number of Wikipedia Category pages, and (2) more specific entities and topics, which can be any type of Wikipedia entry. These topic annotations facilitate the selection of subsets of the corpus for separate analysis or evaluation. In addition, in collaboration with the University of Amsterdam, we have explored automatic content-based labelling of the news photos using ImageNet, with a vocabulary of about 15k WordNet concepts.

Format and availability of the corpus The corpus is available as JSON-LD, a specification for representing Linked Data in the popular JSON format [10]. For each article, it contains the original URL of the article, the date of publication, the headline of the article, the URL of the image displayed with the article (if any), and the caption of that image (if present). The dataset can be downloaded from the project website. In the public version of the corpus, neither the article text nor the images

[1] This tool was also demonstrated at ICT.Open2016, and was awarded the 3rd prize in the meet-the-demo contest there.

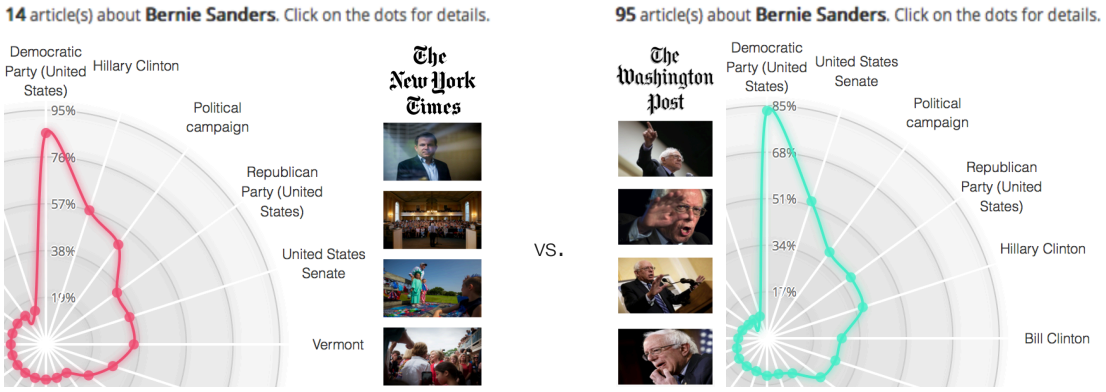


Figure 1: Screenshot with news coverage of Sanders in the New York Times vs. the Washington Post.

are included. Instead, the images are distributed as high-dimensional feature vectors extracted from a Convolutional Neural Network object recognition model [9], anticipating their use in computer vision tasks. The article text is represented as a list of entity and topic annotations extracted with TextRazor.

3 ION Demonstrator functionality

We aim to support an interactive exploration of the choices that different media make with regard to topic selection, topic coverage, and images. We focus on comparative questions. For example, how do two newspapers compare with regard to the frequency with which a topic is reported? In which context do they cover these topics? And how do they compare with respect to the images that they select with a given topic? A split screen design allows a user to visualize two newspapers and compare them side by side. Figure 1 shows a detail of a screenshot where a user selected the New York Times and the Washington Post to compare how they cover Bernie Sanders. It shows (a) differences in how often the topic was reported, (b) differences with regard to which other topics are discussed in those articles and (c) different choices with regard to the photos that are displayed with these articles.

References

[1] D. Elliott and F. Keller. Image Description using Visual Dependency Representations. In *EMNLP*, 2013.

[2] Frank Esser. Dimensions of political news cultures: Sound bite and image bite news in france, germany, great britain, and the united states. *The International Journal of Press/Politics*, 13(4):401–428, 2008.

[3] Shahira Fahmy and Daekyung Kim. Picturing the iraq war: Constructing the image of war in the british and us press. *International Communication Gazette*, 70(6):443–462, 2008.

[4] Keith Greenwood and Joy Jenkins. Visual framing of the syrian conflict in news and public affairs magazines. *Journalism Studies*, 16(2):207–227, 2015.

[5] Martin van Harmelen Laura Hollink, Adriatik Bedjeti and Desmond Elliott. A corpus of images and text in online news. In *Proc. of LREC2016*, pages 207–213, 2016.

[6] Katy Parry. A visual framing analysis of british press photography during the 2006 israel-lebanon conflict. *Media, War & Conflict*, 3(1):67–85, 2010.

[7] Cyrus Rashtchian, Peter Young, Micah Hodosh, and Julia Hockenmaier. Collecting image annotations using amazon's mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 139–147, 2010.

[8] L. Rodriguez and D.V. Dimitrova. The levels of visual framing. *J. of Visual Literacy*, 30(1):48–65, 2011.

[9] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *ICLR*, 2015.

[10] Manu Sporny, Dave Longley, Gregg Kellogg, Markus Lanthaler, and Niklas Lindström. Json-ld 1.0. *W3C Recommendation (January 16, 2014)*, 2014.

[11] C. L. Zitnick and D. Parikh. Bringing semantics into focus using visual abstraction. In *CVPR*, 2013.

Autonomous robot soccer matches

Caitlin Lagrand [a] Patrick M. de Kok [a] Sébastien Negrijn [a]
Michiel van der Meer [a] Arnoud Visser [a]

[a] *University of Amsterdam*

Abstract

The Dutch Nao Team develops fully autonomous soccer behaviors for bipedal robots. In the process of doing so, several fields of artificial intelligence have been explored such as: behavior representations, sound recognition, visual search algorithms and motion models. To showcase the latest developments a penalty shoot-out will be demonstrated.

1 Introduction

Emulating human behaviour and motion is not an easy task. In an effort to incentivize research in these areas, RoboCup offers various platforms at which teams compete, one of which is robot soccer. Using robots, a team is required to program them in such a way that they will autonomously play a match, usually in a five versus five format. One of these teams is the Dutch Nao Team (DNT), consisting of a collaboration between the Universiteit van Amsterdam, Universiteit Maastricht and previously the Technische Universiteit Delft. To reach the goal of playing autonomous soccer with a robot, several main topics of artificial intelligence and robotics have to be combined such as: behavior representations, sound recognition, visual search algorithms and motion models.

2 RoboCup & Standard Platform League

The RoboCup was originally founded in 1997, and has the goal to win with a fully autonomous humanoid robot soccer team against the 2050 FIFA world champion. In order to accomplish this goal, RoboCup has divided itself in multiple leagues, each of which has a different research focus. In the Standard Platform League (SPL), in which DNT participates since 2010, contestants play with identical robots (hence the Standard Platform), which means that the team with the best algorithms in theory wins. The robots used in this league are the bipedal NAO robots as can be seen in Figure 1.

Figure 1: NAO robots in a game at the RoboCup Leizig 2016.

3 The demonstration

The demonstration will feature a display of one attacker versus a goalkeeper on a smaller field, who will try to autonomously score in the goalkeepers goal. Except for the field size, the demonstration will follow the RoboCup SPL rules of next year's competition in Nagoya, Japan. This includes a ground surface of 6 mm high artificial grass, which is a big difference in respect to previous years' flat carpet field. This means that our team has to demonstrate their latest developments in the walking engine. The demonstration requires an area of at least three by six metres. Another challenge is natural lighting conditions; currently a team of honours students is improving the ball-detection algorithm to cope with dynamic lighting (in 2016 the first games outdoor were played).

4 Showcased work

The demonstration showcases work from several publications and theses[1]. It includes, but is not limited to:

- automatic whistle detection [1] to start matches;
- localization [3] supported by a visual compass that updates over time [4];
- robot detection based on Haar-like features [5].

For a more detailed overview of the team's work, we refer to the latest team qualification document [2]. In the 2016 competition the team had the best result so far, it became ex-aequo second in the first-round and was later eliminated in a shoot-out of the play-in round[2].

5 Conclusion

The demonstration allows to show the challenges, research and application of several fields of artificial intelligence behind the RoboCup initiative, which will be explained in more detail by one of the keynote speakers.

References

[1] Niels W. Backer and Arnoud Visser. Learning to recognize horn and whistle sounds for humanoid robots. In *Proceedings of the 26th Belgian-Netherlands Conference on Artificial Intelligence (BNAIC 2014)*, November 2014.

[2] Patrick de Kok, Sébastien Negrijn, Mustafa Karaalioğlu, Caitlin Lagrand, Michiel van der Meer, Jonathan Gerbscheid, Thomas Groot, and Arnoud Visser. Dutch Nao Team team qualification document for RoboCup 2016 - Leipzig, Germany, 2016.

[3] Amogh Gudi, Patrick de Kok, Georgios K. Methenitis, and Nikolaas Steenbergen. Feature Detection and Localization for the RoboCup Soccer SPL. Project Report, Universiteit van Amsterdam, February 2013.

[4] Georgios Methenitis, Patrick M. de Kok, Sander Nugteren, and Arnoud Visser. Orientation finding using a grid based visual compass. In *Proceedings of the 25th Belgian-Netherlands Conference on Artificial Intelligence (BNAIC 2013)*, November 2013.

[5] Duncan S. Ten Velthuis. Nao detection with a cascade of boosted weak classifier based on Haar-like features. Bachelor's thesis, Universiteit van Amsterdam, July 2014.

[1] `http://dutchnaoteam.nl/publications/`
[2] `https://www.youtube.com/watch?v=kQlIMMRLfMI`

Video Demo: Deep Reinforcement Learning for Coordination in Traffic Light Control

Elise van der Pol [a] Frans A. Oliehoek [a,b]

[a] *University of Amsterdam*
[b] *University of Liverpool*

Abstract

This video demonstration contrasts two approaches to coordination in traffic light control using reinforcement learning: earlier work, based on a deconstruction of the state space into a linear combination of vehicle states, and our own approach based on the Deep Q-learning algorithm.

1 Overview

The cost of traffic congestion in the EU is estimated to be 1% of the EU's GDP [2], and good solutions for traffic light control may reduce traffic congestion, saving time and money and reducing pollution. To find optimal traffic light policies, reinforcement learning uses reward signals from the environment to learn to make optimal decisions. This demo[1] shows the difference in policies for a multi-agent traffic network, between Deep Q-learning (DQN) with transfer planning [6] and earlier work by Kuyer et al. [3].

2 Approach & Experimental Setup

Earlier reinforcement learning approaches to traffic light control relied on simplifying assumptions over the state and manual feature extraction, so that potentially vital information about the state is lost. Techniques from the field of deep learning can be used in deep reinforcement learning to enable the use of more information over the state and to potentially find better traffic light policies.

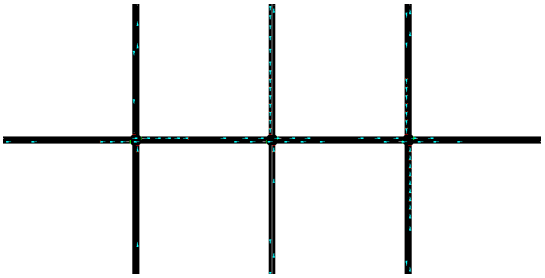
The algorithms compared are the approach used by Kuyer et al. [3], with the adjustment that we learn the joint local Q-value functions, and then use transfer planning with max-plus to coordinate. Similarly, in the DQN approach, we use the DQN algorithm to learn joint local Q-functions, and then transfer planning and max-plus to coordinate. The difference is in the state representation and reward function. In the Kuyer approach, the state is decomposed into a linear combination of lane positions. The reward function penalizes each halted vehicle on the agent's roads. In the DQN approach, a binary matrix of vehicle positions is used, following earlier work [7], and the reward function balances different objectives.

3 Results

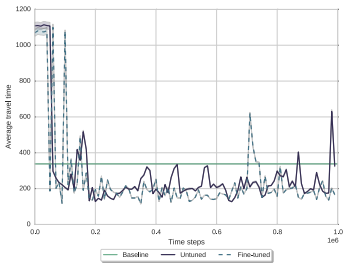
The state is represented as a matrix of vehicle positions, and fed into a convolutional network. We learn the joint Q-value functions between neighbouring agents with the DQN algorithm [4] and use transfer planning [5] with the max-plus coordination algorithm [1] to find an optimal joint action for each traffic network state. The algorithms are compared on, amongst others, the three-agent scenario in Figure 1a.

[1] http://www.fransoliehoek.net/trafficvideo

To illustrate the behavior of the DQN approach, Figure 1b shows the average travel time for an untuned and a fine-tuned agent, by evaluating the greedy policy at different times during the training process. The best version of the Kuyer algorithm is plotted as a horizontal line.



(a) Three agent traffic scenario: each junction is regulated by a traffic light agent.



(b) Average travel time in the three-agent scenario, for an untuned and fine-tuned DQN agent. The baseline is the Kuyer algorithm.

Figure 1

The video demo shows the behavior of both the DQN approach and the Kuyer approach. The Kuyer agents switch traffic light configurations very rapidly, whereas DQN allows a queue to build up. Because of Kuyer's rapid switching, vehicles are almost always accelerating or decelerating. This is a result of the reward function used, which assigns a penalty for each halted vehicle. In contrast, DQN uses a reward function that balances multiple objectives, one of which penalizes rapid decelerations (emergency stops).

After a few minutes of the simulation the Kuyer agents start getting congested, whereas the DQN agents see less congestion and no rapid switching. This suggests that our approach will be able to handle a higher traffic density. However, near the end of the simulation, the Kuyer agents eventually empty all their roads, but the DQN agents seem 'stuck'. This may be caused by a loss of information when converting from the continuous traffic situation to a discrete position matrix. If the matrix misses some vehicles, or the DQN agent's convolutional filters do not trigger for these vehicles, the result is that the DQN agent behaves as though there are no vehicles on the road.

4 Conclusion

While the DQN approach is promising and finds better policies than earlier work, there are state instances where it fails to take the right action. Thus, more work is needed to increase its robustness.

Acknowledgements

Research supported (in part) by NWO Innovational Research Incentives Scheme Veni #639.021.336. The Tesla K40 used for this research was donated by the NVIDIA Corporation.

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag New York, Inc., 2006. ISBN: 0387310738.

[2] Commission of the European communities. *White paper European transport policy for 2010: time to decide*. 2001.

[3] L. Kuyer et al. "Multiagent reinforcement learning for urban traffic control using coordination graphs". In: *ECML-PKDD*. 2008.

[4] V. Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (2015), pp. 529–533.

[5] F. A. Oliehoek, S. Whiteson, and M. T. Spaan. "Approximate solutions for factored Dec-POMDPs with many agents". In: *AAMAS*. 2013.

[6] E. van der Pol. "Deep reinforcement learning for coordination in traffic light control". Master's Thesis. University of Amsterdam, 2016.

[7] T. Rijken. "DeepLight: Deep reinforcement learning for signalised traffic control". Master's Thesis. University College London, 2015.

PGMs4SDA: a public repository for Probabilistic Graphical Models

María Dolores Sánchez-García [a] José del Sagrado [b] Antonio Salmerón [a]
Rafael Rumí [a]

[a] *Dpt. of Mathematics, University of Almería, 04120 Almería (Spain)*
[b] *Dpt. of Informatics, University of Almería, 04120 Almería (Spain)*

Abstract

PGMs4SDA repository has as major goals the dissemination of results and the information exchange between researchers interested in the field of Probabilistic Graphical Models. It has been developed as a dynamic web application with two views: a public one, where any user can browse and download algorithms, and a private one, in which any interested researcher can submit a contribution.

1 Motivation

Probabilistic graphical models (PGMs) [1, 2] provide a well-founded and principled approach for handling complex domains endowed with uncertainty. The Probabilistic Graphical Models for Scalable Data Analytics (PGMs4SDA) project has as main purpose to generate new methodological developments in the field of PGMs and to make them available to the research community as software implementations sufficiently sound and innovative as to place them as reference tools. Methods and algorithms in this project are being developed by research groups belonging to the universities of Almería, Castilla-La Mancha and Granada, using different systems and software development environments. This scenario made us decide to develop a web repository as a way of disseminating the results of the project and, at same time, encouraging cooperation between researchers interested in the field of PGMs or in their different uses.

2 Design and major features

The PGMs4SDA repository has two major goals: the first one is to store the key algorithms that have been designed in the project and the other one is to allow to download and use these algorithms. As an open community repository it also has additional characteristics, such as a contributors registration and login system and a peer-reviewed process for assessing a minimum quality in the contributions. Taking into account the aforementioned goals and characteristics, we decided to develop the repository as a dynamic web site based on Apache, PHP and MySQL. Any interested user can access the PGMs4SDA repository through the URL *http://www2.ual.es/PGMs4SDA/Repository/repository.php*.

Users can interact in two ways (see Fig. 1): by browsing through contents or by submitting contributions. Let us start with the public part in which PGMs4SDA project's contributions and software implementations are displayed grouped according to the tasks associated to PGMs (i.e. learning, inference, classification and use cases).

Any interested researcher can submit a contribution by registering and logging in to the private part of the repository. After registering as contributor, providing a valid email and a user defined password, the user can upload a contribution related to one of the tasks associated with PGMs. For each contribution it is mandatory to indicate: *i)* the *algorithm name* which is a unique, short and representative name for your software package; *ii)* a plain text *description* of the contribution; *iii)* a publication or *reference*

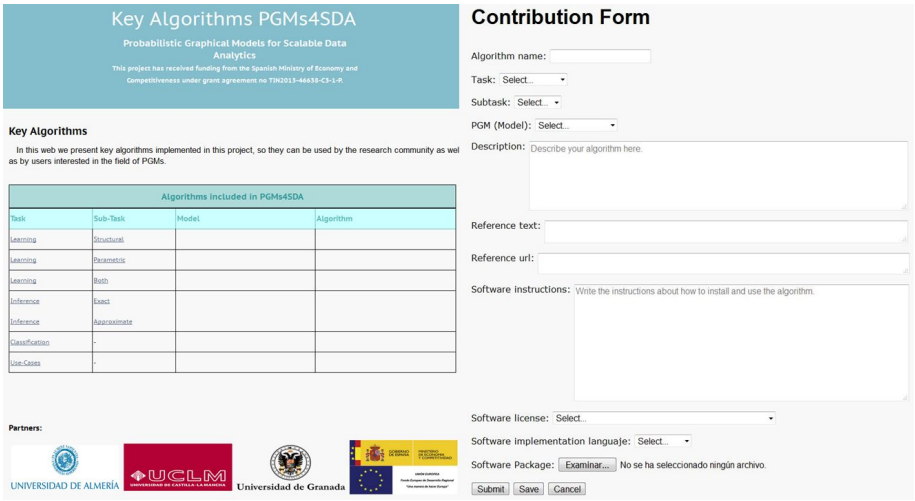


Figure 1: PGMs4SDA repository: navigation and contribution views.

text which should be cited when someone uses your package; *iv)* an URL link to the reference text (i.e. *reference URL*); *v) software instructions* briefly describing hardware and software requirements together with any other details on how to install your contributed software after downloading it and how it is intended to operate, and, *vi)* the *license* under which your software is distributed. You must upload all required files (including software instructions and code) into a single compressed file (only rar and zip files are accepted, with a maximum allowed size of 7 MB) so your code can be downloaded, installed and afterwards be fully operational.

Once a contribution is submitted, a one-round reviewing process starts. It is sent to a reviewer, who may require you to make some modifications on your contribution. Then according to the reviewer comments you must revise your contribution and make the appropriate modifications. Finally, you are asked to answer the reviewer's message, so she/he can check the contribution for its acceptance. Note that in this process the state of the contribution changes as well as the actions that can be performed on it (i.e. read/answer messages, consult/edit the contribution). This is indicated by a list of actions displayed as icons in the application. Only when the contribution is finally accepted, it will appear on the main web page of the repository.

3 Demonstration

Demonstration of the developed tool needs a browser and an Internet connection. After introducing the main page of the PGMs4SDA repository, audience can interact and navigate through the available contributions and software implementations. It also will be shown how to register and contribute to the repository, while getting familiar with the different steps of the reviewing process. The demo [1] will take approximately 9 minutes.

Acknowledgments

This research has been funded by the Spanish Ministry of Economy and Competitiveness under project TIN2013-46638-C3-1-P.

References

[1] N. Friedman D. Koller. *Probabilistic Graphical Models. Principles and Techniques*. MIT Press, 2009.

[2] A. Nicholson K. Korb. *Bayesian Artificial Intelligence*. Chapman Hall/CRC, 2010.

[1] https://drive.google.com/file/d/0B8a-PuFD3ZfPU3hzRzNGQ2ZXWHM/view?usp=sharing

Crowdsourcing ground truth data for analysing brainstem tumors in children

Benjamin Timmermans Zoltán Szlávik Robert-Jan Sips

Vrije Universiteit, Amsterdam

Abstract

Brainstem tumors are a rare form of childhood cancer for which there is currently no cure. The Semmy Foundation and IBM Netherlands are developing a cognitive system for quicker analysis of MRI-scans and better detection of anomalies in the brainstem. In order to gather training data for this system, the Semmy Annotation Tool was developed that allows untrained participants to draw the exact shape of the brainstem and tumor onto an MRI-scan. The tool was tested in an experiment at the Lowlands 2016 festival, in which 823 participants annotated 5.152 images.

1 Purpose

Diffuse intrinsic pontine glioma is a rare type of tumor located in the center of the brainstem. Yearly around 18 children die in The Netherlands through this disease, and there is currently no known cure. The Semmy Foundation[1] was founded with the goal of increasing the survival chance of children with this type of cancer. Building on the work of Kaspers et al. [2], the Semmy Foundation has started a new initiative together with the Center for Advanced Studies at IBM Netherlands to develop a cognitive system that should allow doctors and researchers to quicker analyse MRI-scans and better detect anomalies in the brainstem. In order to train this system large amount of annotated data is required. The Semmy Annotation Tool was developed in order to gather this data through crowdsourcing.

The Semmy Annotation Tool[2] was developed and tested at the science fair of the festival Lowlands 2016. The participation in this fair allowed: 1) to increase the awareness for the Semmy Foundation and this type of cancer, 2) to gather large amounts of annotations on the shape of the brainstem and cancer cells, and 3) to measure the quality of the annotations that a participant made in relation to the well being of that participant. The hypothesis is that people under influence can still make valuable contributions, but that these are of lower quality than sober people. This study may give insight into the reliability of online crowd workers and the effects of various aspects on the annotation quality. This is a problem in crowdsourcing that to our knowledge has not yet been investigated.

2 Methods

The dataset used for the experiment contained 1.711 anonymized images of MRI-scans on children with brainstem cancer, which were made available by the Semmy Foundation. The Semmy Annotation Tool was developed so multiple untrained participants could simultaneously draw the edges of the brainstem precisely on these medical images in a controlled environment. The tool uses the Laravel PHP framework, and Semantic UI for the interface. When an image is shown for annotation it is scaled up to the maximum size within the canvas, after which drawing of the shape on an image is made possible through custom javascript code (See Figure 1). This differs from approaches like [1], where the placing of independent polygons is used and the assumption is made that the annotator can always identify the shape that is to be found.

[1] http://www.stichtingsemmy.nl/

[2] http://github.com/CrowdTruth/Semmy/

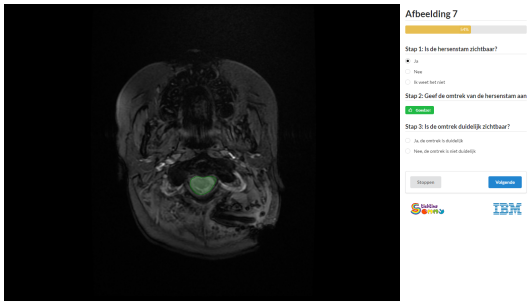


Figure 1: Semmy Annotation Tool used to draw the shape of the brainstem and tumor.

Figure 2: Dashboard used to motivate the crowd to participate.

In the crowdsourcing task[3] the participant was first asked for age, gender, education level, and whether alcohol or drugs had been used. This should allow the precision of the annotations to be matched to the registered well being of the participant. In order to guarantee correctness of this input, moderators were assigned to assist and give the initial instructions for the annotation task. After this the participant was given two minutes to annotate images. The images were ordered randomly, but with priority for images with the fewest annotations yet. For each image the participant was asked whether the brainstem or tumor could be seen. If the answer was yes, the participant was asked to draw the shape onto the image, and to indicate whether the edges were clearly visible in the image. Based on the CrowdTruth methodology [3] for capturing human interpretation, the goal was to gather 10 to 15 annotations for images where the tumor was indicated to be seen.

In order to motivate the audience to participate, a dashboard[3] was created (Figure 2) that shows live aggregated statistics such as the amount of participants and the amount of annotations gathered. Also, a 3D visualization of the brainstem was added to the dashboard in order to increase the understanding of what the brainstem looks like. In the experimental setup at Lowlands the participants could annotate simultaneously on two laptops and the dashboard was shown on a large tv screen.

3 Results

In total 823 people participated with an average age of 28.3 years. Of the participants 58% were female, and 69% had finished at least a bachelors degree. In total they annotated 5.152 images, of which 2.440 were said to contain a brainstem or tumor. This indicates almost half of the images were either not clear or did not contain parts of the brainstem. Within the images that were said to contain the brainstem 558.887 polygons were drawn, resulting in an average resolution of 230 polygons per image. The developed tool proved useful for gathering the data, but more annotations are needed for further analysis of the annotation quality and classification of brainstem cancer.

References

[1] Alberto Chávez-Aragón, Won-Sook Lee, and Aseem Vyas. A crowdsourcing web platform-hip joint segmentation by non-expert contributors. In *Medical Measurements and Applications Proceedings (MeMeA), 2013 IEEE International Symposium on*, pages 350–354. IEEE, 2013.

[2] GJ Kaspers. A twenty-year review of diagnosing and treating children with diffuse intrinsic pontine glioma in the netherlands. *Diffuse Intrinsic Pontine Glioma Clinical Aspects and Imaging*, 1:41, 2014.

[3] O. Inel et al. CrowdTruth: machine-human computation framework for harnessing disagreement in gathering annotated data. In *The Semantic Web–ISWC*. 2014.

[3] Screencast available at: `https://youtu.be/8VhZ7SGn0IE`

Multi-agent Multi-contact Path Planning with HPP

Shuai Wang [ab] Steve Tonneau [a] Nicolas Mansard [a]

[a] *LAAS-CNRS, Toulouse, France*
[b] *ILLC, University of Amsterdam, The Netherlands*

Abstract

This demonstration paper presents the first attempt using HPP, a new path planner, for multi-agent multi-contact path planning. More specifically, a few HyQ robots, loaded on a Boeing 747-400 airplane, need to move to the emergency exits. The trajectories were first generated, based on which the contact sequences were specified. Finally the motion was illustrated in animation. This work was also a first attempt introducing Agent-Oriented concepts to multi-contact path planning in HPP.

1 Introduction

HPP (Humanoid Path Planner) is a new open-source software designed for complex classes of motion planning problems, such as navigation among movable objects, manipulation and contact-rich locomotion [1]. It adopts a clear object oriented architecture and can be extended to path planning for other types of robots. Rbprm [2] is a module that extends HPP with a reachability-based algorithm that generates acyclic contacts from a specified trajectory. HPP was developed mostly to handle planning tasks for a single robot. This project aimed at exploring possible extension of HPP as a Multi-agent path planning system. More specifically, multiple HyQ robots [3] were to move to the emergency exits of a Boeing 747-400 airplane without collision with the airplane or each other. Our attempt followed a three-phase hierarchical approach: the trajectories were first generated, based on which, the contacts were specified. Then the motion was illustrated in animation.

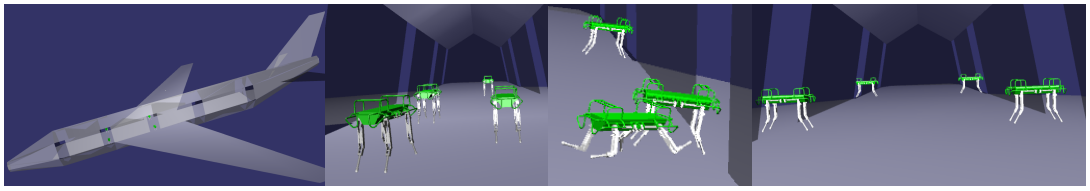
2 Design and Implementation

2.1 Trajectory Planning

A simplified Boeing 747-400 was first initialised in HPP (Fig. 1-a). Following that agents were loaded with the initial and goal configuration specified (Fig. 1b and 1d). Before planning, each agent was asked to generate a default trajectory. They were then registered on the platform to whom they propose plans based on the location of other agents. Once the platform detects a collision (e.g. Fig. 1-c), the agents were asked to re-plan from then on. If the agent failed to generate a valid plan, its default plan was loaded. If there is a collision when all agents use their default plans, we backtrack to a validate previous moment.

2.2 Contact Planning

Based on the trajectory, sequences of acyclic contacts were generated using a reachability-based planner [2]. The contact configurations were then exported to a file in preparation for the animation.



(a) The airplane (external view) (b) The initial configuration (internal view) (c) Collision avoidance (d) The goal configuration

Figure 1: Planning and collision avoidance of 4 HyQ robot avatars in a simplified Boeing airplane

2.3 Animation

After loading the contact configuration of each robot, the platform can demonstrate the locomotion by repeatedly refreshing the agents' configuration in display. Users may choose to have an external view or internal view.

2.4 Implementation and Evaluation

HPP follows a client-server based design patten. A *client* consisted of a *robot*, a *problem* as well as some *obstacles*. At the trajectory planning stage, the agent class we as inherited from the client class for better management of interaction with the platform and reasoning tasks. For every planning task, a new problem was defined and the obstacles were reloaded.

This project used the Python API of HPP. A complete 3-step planning procedure may take a few minutes. The project is available online [1].

3 Conclusion and Acknowledgement

This project was a first attempt using HPP for the task of multi-agent path planning with contacts. The project unveiled possible changes of HPP for better management of multiple agents. The current HPP implementation cannot perform reasoning about multiple planning problems in parallel. Reuse of intermediate planning results are possible but not implemented. The project implemented a naive collision avoidance algorithm, which could be further optimised. The project was a part of Shuai Wang's internship supervised by Dr. Nicolas Mansard. Detailed help and guidance from Dr. Steve Tonneau and Mr. Joseph Mirabel were also much appreciated.

References

[1] Joseph Mirabel, Steve Tonneau, Pierre Fernbach, Anna-Kaarina Seppälä, Mylène Campana, Nicolas Mansard, and Florent Lamiraux. Hpp: a new software for constrained motion planning. In *IEEE/RSJ Intelligent Robots and Systems*, October 2016.

[2] Steve Tonneau, Nicolas Mansard, Chonhyon Park, Dinesh Manocha, Franck Multon, and Julien Pettré. A reachability-based planner for sequences of acyclic contacts in cluttered environments. In *International Symposium on Robotics Research*, 2015.

[3] Barkan Ugurlu, Ioannis Havoutis, Claudio Semini, and Darwin G Caldwell. Dynamic trot-walking with the hydraulic quadruped robothyq: Analytical trajectory generation and active compliance control. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 6044–6051. IEEE, 2013.

[1] The demonstration videos were embedded on the project page: `https://airobert.github.io/MMPP-BNAIC/`. The source code and instructions were also made available there.

Dynamic, context-aware behavior change support using distributed reasoning and central processing

Arlette van Wissen [a] Annerieke Heuvelink [a] Cliff J. R. H. Laschet [a]
Charlotte Vinkers [a]

[a] *Philips Research, High Tech Campus 34, 5656 AE Eindhoven, the Netherlands*

1 Developing an adaptive, context-aware BCSS

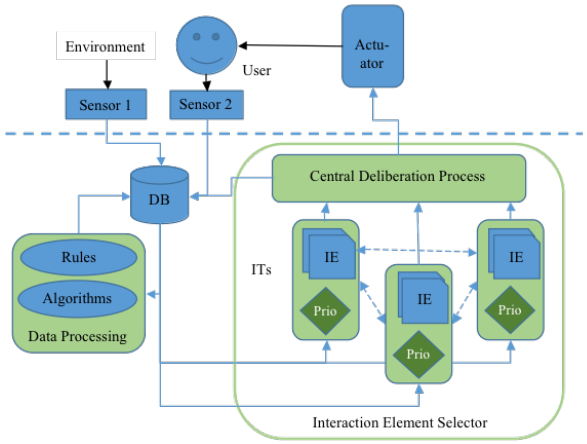
Behavior change support systems (BCSSs) are technologies that support their users in changing behaviors and cognitions related to (mental) health and wellness [3]. For a BCSS to choose the most suitable interaction elements (IEs) for a user, it should consider user preferences and context (i.e., knowledge of the user), as well as which intervention technique (IT) to use. These are techniques to deliver interventions to help people to change their behavior, such as giving feedback to the user based on measurements, or verbally persuading them about their capability. Current BCSSs are constrained in their capability to successfully support behavior change, as they have limited ability to tailor or adapt the components of their intervention (i.e., which, how and when they offer them) to the state and context of users in real-time and over time.There are three main ways in which our system differs from current BCSSs and why it is able to dynamically adapt interaction elements and intervention techniques to the user. First, ITs are declaratively specified in the system, that is, the technique is annotated and its dependencies and consquences specify under what circumstances the IT should be activitated. Importantly, and this is the second way in which our system differs from other BCSSs, this means that which ITs are provided and when is not predefined nor prescheduled by developers or behavior change experts. Instead, the ITs receive different activation values that depend on the user's current state and context, which makes our system flexible to changing circumstances during the intervention program. Finally, knowledge about the user (and context) is frequently updated, which enables dynamic and real-time coaching based on the changing needs of the user over time and in different contexts.

2 Proof of concept

The approach is inspired by Brooks' subsumption architecture [1], which couples sensory information to action selection by decomposing the complete behavior into sub-behaviors, implemented as a hierarchy of layers. The layers, which all receive sensor-information, work in parallel and generate outputs. This architecture emphasizes the strength of distributive and parallel control to come to action selection. Our approach proposes similar distributed processes combined with a final centralized process where it is identified which of the activated ITs will be used and send to the user (see Fig. 1). As a first reasoning step, distributed reasoning elements process updates in user states. Each process corresponds to one IT and in parallel they determine their priority, based on their applicability and relevance, which is derived from several features (i.e. characteristics) of the user input. In the second reasoning step, a central deliberation process identifies which from the applicable and relevant ITs should be executed, based on their priorities. It receives the priority values from the distributed reasoning processes and returns an IT (or a list of ITs) and corresponding IEs that will be executed by the actuator.

As a proof of concept, the two-step approach was used to generate coaching content in an app to support patients with physical activity. The implemented ITs consist of the following: review behavior

Figure 1: Combining distributed reasoning and central processing for behavior change



goal, feedback on behavior, discrepancy between behavior and goal, problem solving, and commitment (see [2]). For each of the techniques, there is a reasoning process that determines their priority to be executed, given the incoming data about the current user state (such as activity and location) and current data in the system. In our implementation, the priority depends on four different features of the ITs: efficacy, capability, opportunity and urgency. How the features are combined to calculate the final priority, can differ per IT. In our solution the priority is determined by a weighted sum of the features, whose weights vary between ITs. The proposed approach is more scalable than one in which conditions and triggers are defined for each IT in isolation by the designer, because in our approach the system automatically handles interactions effects and competition (i.e. the hierarchy) between the ITs.

Consider a scenario where Annie's goal is to walk 45 minutes every day, but her motivation is low. The system sensors (e.g., accelerometer, iBeacons) detect that Annie has walked 15 minutes and that she is at home. The system will be triggered by this new data and starts to reason about what is the most appropriate intervention to support Annie (if any). The parallel reasoning processes calculate the feature values and combine them into one priority value. In our scenario, the priority values per IT could be: feedback – 0.73, goal discrepancy – 0.58, and commitment – 0.32 (other ITs do not become active at all, because, for instance, there is not enough data). Annie has indicated that her preferred frequency of interactions is 'many'; she does not mind getting multiple interventions. Therefore, the threshold for priorities to get selected is low, e.g. 0.5. The central processing unit compares the priorities between the ITs and selects those with a priority above the threshold, in this case both feedback and goal discrepancy. Annie will receive both ITs with corresponding IEs.

During the demo we will demonstrate an implementation of this design approach which is (i) adaptive to the user state and context, and (ii) scalable over different users and contexts, as well as over the number and types of ITs and IEs they can offer. The aim of demo is to show how behavior change intervention techniques and elements can be dynamically offered on a smartphone, so that users receive content that is not just personalized but also tailored to their current state. Future work will include evaluation of this approach by including it in a working coaching program with actual users.

References

[1] R. Brooks. A robust layered control system for a mobile robot. *IEEE journal on robotics and automation*, 2(1):14–23, 1986.

[2] S. Michie, M. M. van Stralen, and R. West. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(1):1, 2011.

[3] H. Oinas-Kukkonen. Behavior change support systems: A research model and agenda. In *International Conference on Persuasive Technology*, pages 4–14. Springer, 2010.

BNAIC2016

THE 28TH BENELUX CONFERENCE ON ARTIFICIAL INTELLIGENCE

NOVEMBER 10-11 2016, AMSTERDAM

ORGANISED BY

VU VRIJE UNIVERSITEIT AMSTERDAM

UNIVERSITY OF AMSTERDAM

UNDER THE AUSPICES OF

SIKS

BNVKI AIABN