APPROXIMATE REASONING IN MAS: Rough Set Approach

Andrzej Skowron

Warsaw University, Poland skowron@mimuw.edu.pl

IAT'06, December 18-22, 2006, Hong Kong

CHALLENGE

Make progress in developing methods for complex concept approximations and reasoning about approximated concepts in distributed environments.

AGENDA

ROUGH SETS
&
GRANULAR COMPUTING
&
LAYERED LEARNING
&
DOMAIN ONTOLOGY APPROXIMATION

FOR COMPLEX CONCEPT APPROXIMATION

EXAMPLES OF COMLEX CONCEPTS

- Examples of complex concepts
 - Dangerous situation
 - Identification of behavioral patterns
 - Risk patterns
 - Similarity of plans
 - Sunsopt classification
 - Outliers
 - Reinforcement learning: conditions for actions
 - Intentions, motivations, desires, beliefs,...

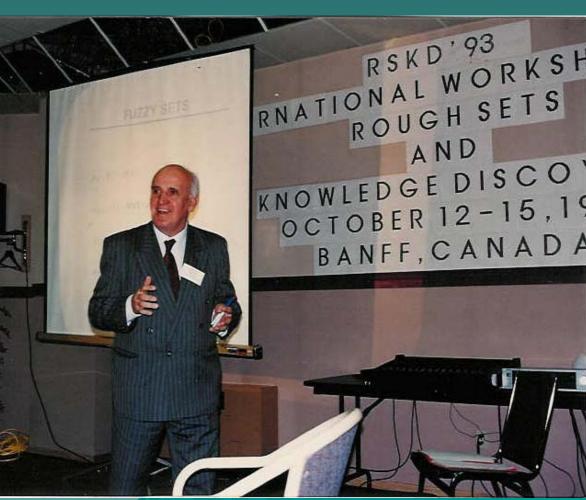
— ...

ROUGH SETS over 4000 publications

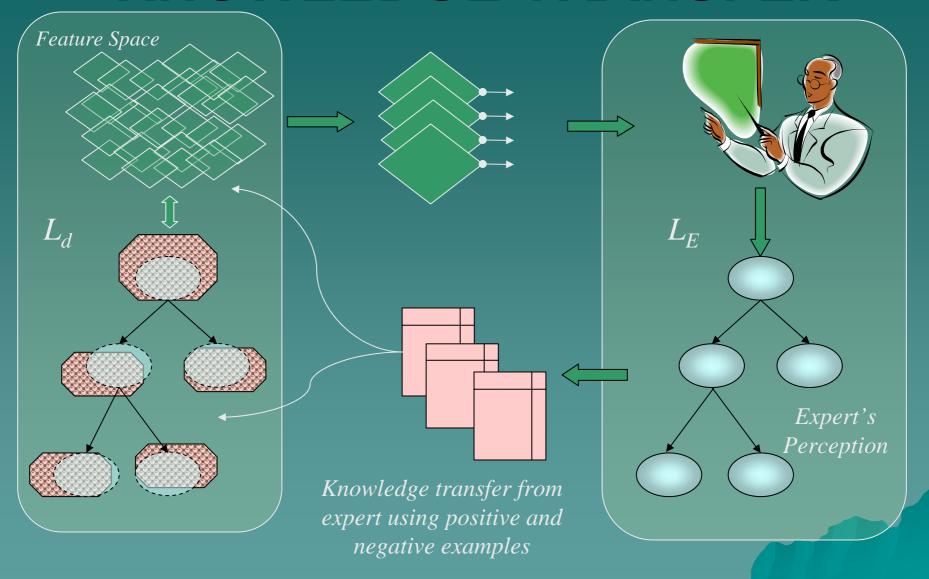


Prof. Zdzisław Pawlak

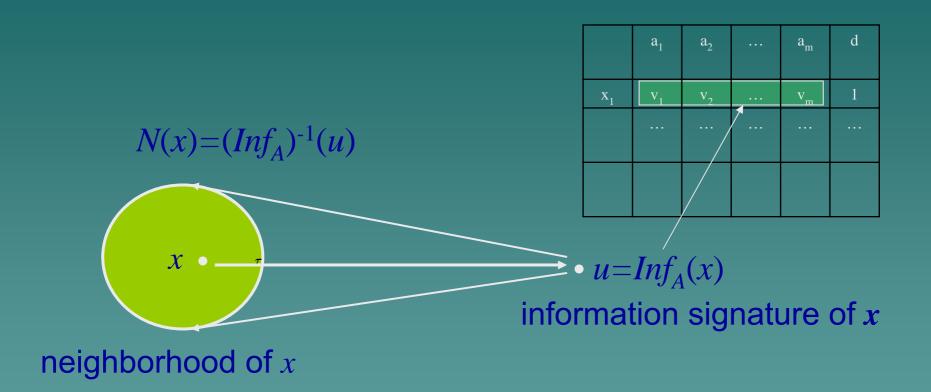
passed away on April 7, 2006



KNOWLEDGE TRANSFER

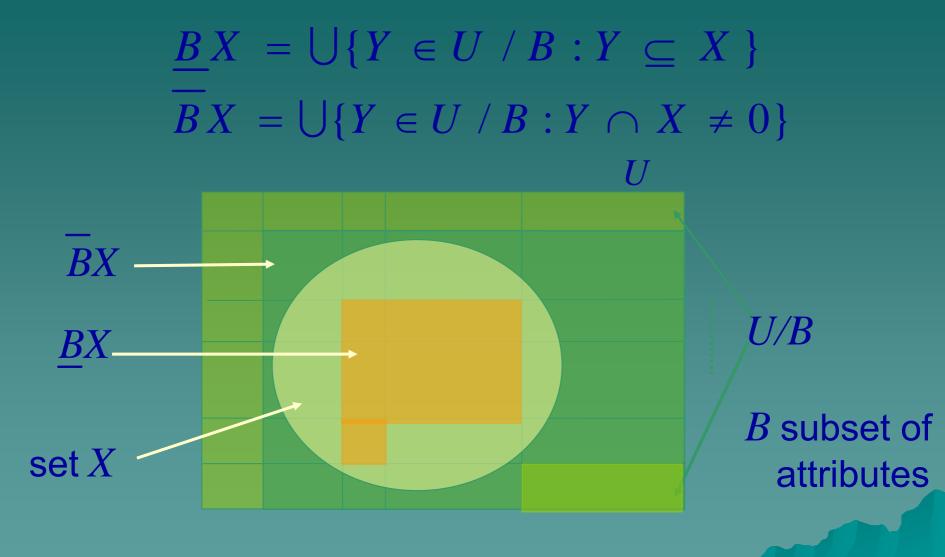


INDISCERNIBILITY FUNCTION



$$xIND(A)y iff Inf_A(x) = Inf_A(y)$$

LOWER & UPPER APPROXIMATIONS



 $BX \neq \underline{B}X$ rough sets (Zdzisław Pawlak 1982)

VAGUENESS

- Sorites (bald men) paradox, (Eubulides, ca. 400 BC)
- "Der Begriff muss scharf begrenzt sein. Einem unscharf begrenzten Begriff würde ein Bezirk entsprechen, der nicht überall ein scharfe Grenzlinie hätte, sondern stellenweise ganz verschwimmend in die Umgebung überginge"

(The concept must have a sharp boundary. To the concept without a sharp boundary there would correspond an area that had not a sharp boundary-line all around)

Gottlob Frege (1848 – 1925)

*Grundlagen der Arithmetik 2,*Verlag von Herman Phole, Jena (1893)

GENERALIZED APPROXIMATION SPACES

A. Skowron, J. Stepaniuk, Generalized Approximation Spaces 1994

$$AS = (U, N, \nu)$$

$$N:U o P(U)$$
 neighborhood function

$$\nu: P(U) \times P(U) \rightarrow [0,1]$$
 rough inclusion partial function

$$x \rightarrow Inf(x) \rightarrow N(x) = Inf^{-1}(Inf(x))$$

neighborhood of x



APPROXIMATION SPACE

$$AS = (U, N, \nu)$$

$$LOW(AS, X) = \{x \in U : \nu(N(x), X) = 1\}$$

$$UPP(AS, X) = \{x \in U : \nu(N(x), X) > 0\}$$

EXAMPLE OF ROUGH INCLUSION (Jan Łukasiewicz, 1913)

$$v_{st}(X,Y) = \begin{cases} \frac{|X \cap Y|}{|X|} & \text{if } X \neq \emptyset \\ 1 & \text{if } X = \emptyset \end{cases}$$

$$C \subseteq U, x \in U$$

 $v_{st}(N(x), C) = 1 \text{ iff } N(x) \subseteq C$
 $v_{st}(N(x), C) = 0 \text{ iff } N(x) \cap C = \emptyset$
 $v_{st}(N(x), C) \neq 0 \text{ iff } N(x) \cap C \neq \emptyset$

ROUGH MEREOLOGY

MEREOLOGY
St. LEŚNIEWSKI (1916)

x is_a_ part_of y

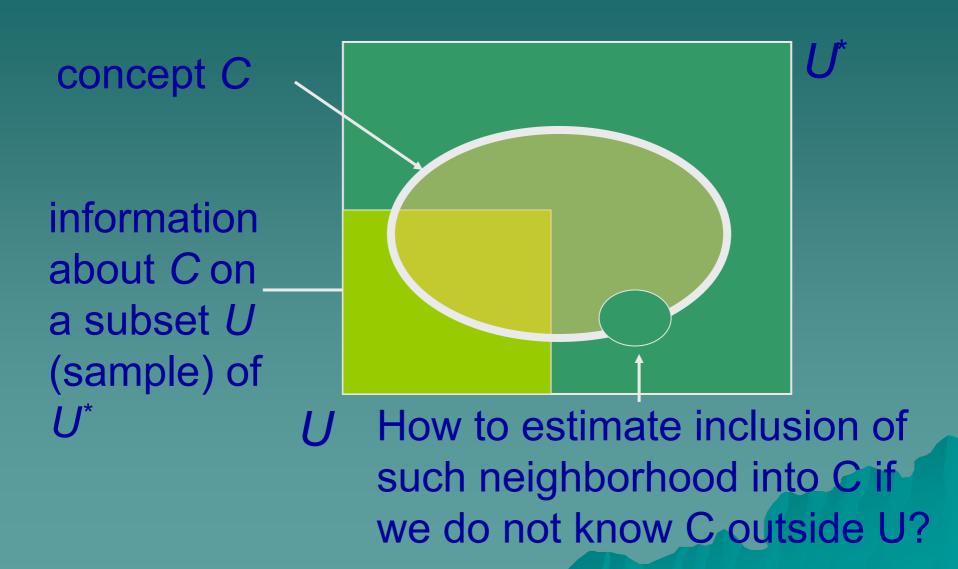
ROUGH MEREOLOGY

L. Polkowski and A. Skowron (1994-....)

x is_a_ part_of y in a degree

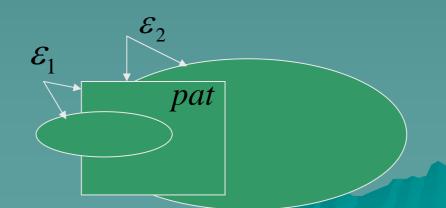
L. Polkowski, A. Skowron, Rough mereology, ISMIS'94, LNAI 869, Springer, 1994, 85-94

INDUCTIVE REASONING



GRANULATION OF ARGUMENTS FOR AND AGAINST

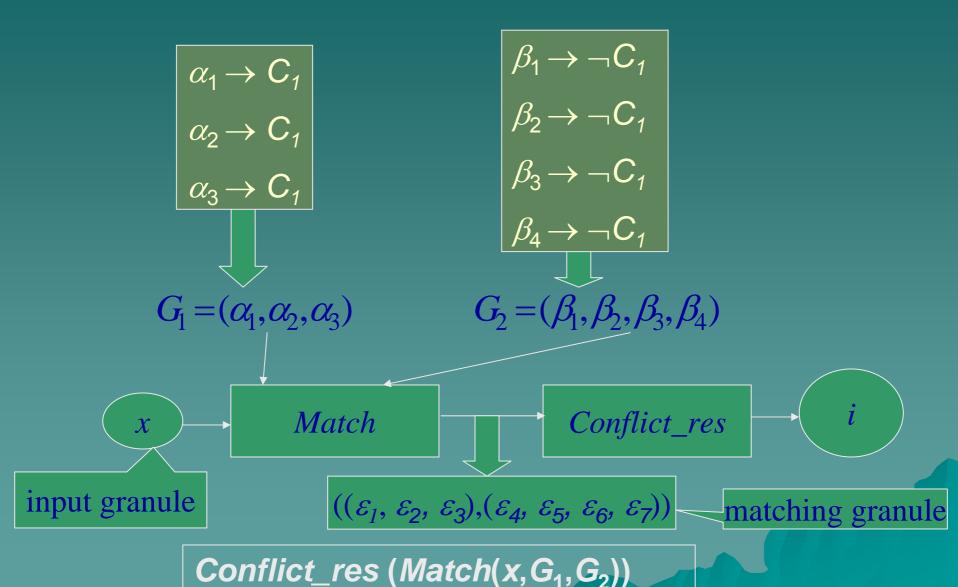
- Argument represented by
 - pat: pattern (e.g., left hand side of decision rule, e.g., $T=high \land H=Yes \rightarrow Flue=Yes$)
 - $-\varepsilon_1$: degree of object inclusion to pattern
 - $-\varepsilon_2$: degree of pattern inclusion to the decision class



ARGUMENTS ARE USED IN CALSSIFIER CONSTRUCTION

The arguments are used by a conflict resolution strategy (fusion operation) to predict the decision, i.e., to decide if the analyzed concept belongs to a concept or not.

INDUCTION OF CLASSIFIERS



Leo Breiman (Berkeley Univ.): Statistical Modeling: The Two Cultures (Statistical Science 16, 199-231, 2001)

There are two cultures in the use of statistical modeling to reach conclusions from data. One assumes that the data are generated by a given stochastic models. The other uses algorithmic models and treats the data mechanism as unknown. The statistical community has been committed to the almost exclusive use of data models. This commitment has lead to irrelevant theory, questionable conclusions, and has kept statisticians from working on a large range of interesting current problems. ... If our goal as a field is to use data to solve problems, then we need to move away from exclusive dependence on data models and adopt a more diverse set of tools.

TRANSDUCTIVE INFERENCE Statistical Learning Theory (V. N. Vapnik, Wiley 1999)

- Inductive solutions are derived in two steps: inductive step (from particular to general) and next deductive step (from general to particular)
- Transductive solutions are derived in one step, directly from particular to particular (the transductive step)

INFORMATION GRANULATION (L.A. Zadeh)

- Information granularity is a concomitant of the bounded ability of sensory organs, and ultimately the brain, to resolve detail and store information.
- Human perceptions are, for the most part, intrinsically imprecise.
- Boundaries of perceived classes are vague.
- The values of perceived attributes are granular.
- Information granulation may be viewed as a human way of achieving data compression.
- It plays a key role in implementation of the strategy of divide-and-conquer in human problem-solving.

LAYERED LEARNING & ONTOLOGY APPROXIMATION

LAYERED LEARNING

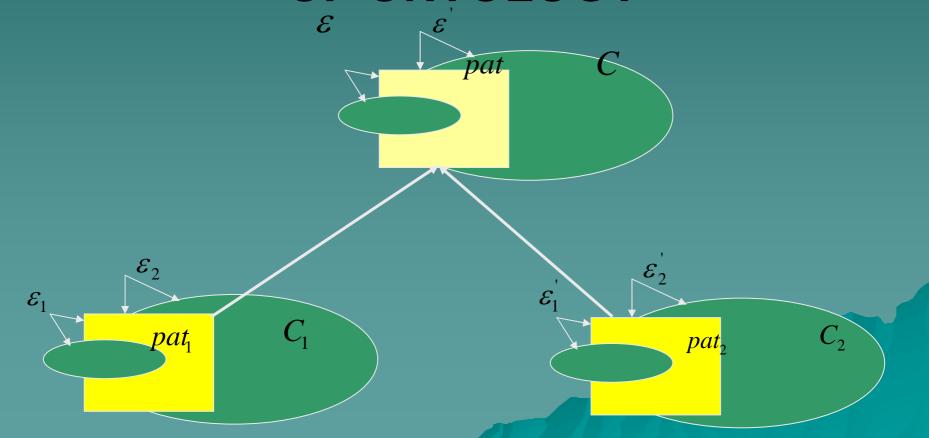
a new learning approach for teams of autonomous agents acting in real-time, noisy, collaborative and adversarial environments

Given a hierarchical task decomposition, layered learning allows for learning at every level of the hierarchy, with learning at each level directly affecting learning at the next higher level.

"Layered learning in multiagent systems: A winning approach to robotic soccer"

P. Stone 2000

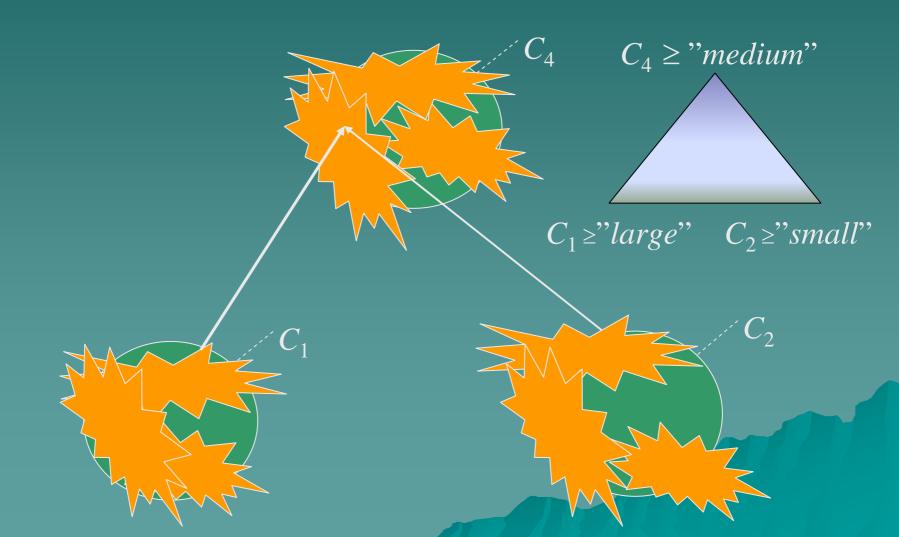
CONSTRUCTION OF ARGUMENTS FOR AND AGAINST FOR CONCEPTS ON HIGHER LEVEL FROM ARGUMENTS ON LOWER LEVEL OF ONTOLOGY



GRANULATION OF CONSTRUCTIONS TO LOCAL SCHEMES (PRODUCTION RULES)

Any local scheme corresponding to a local dependency between vague concepts can be considered as a family of transducers satisfying a monotonicity property with respect to the linear order of linguistic degrees.

PATTERN CONSTRUCTION GRANULATION



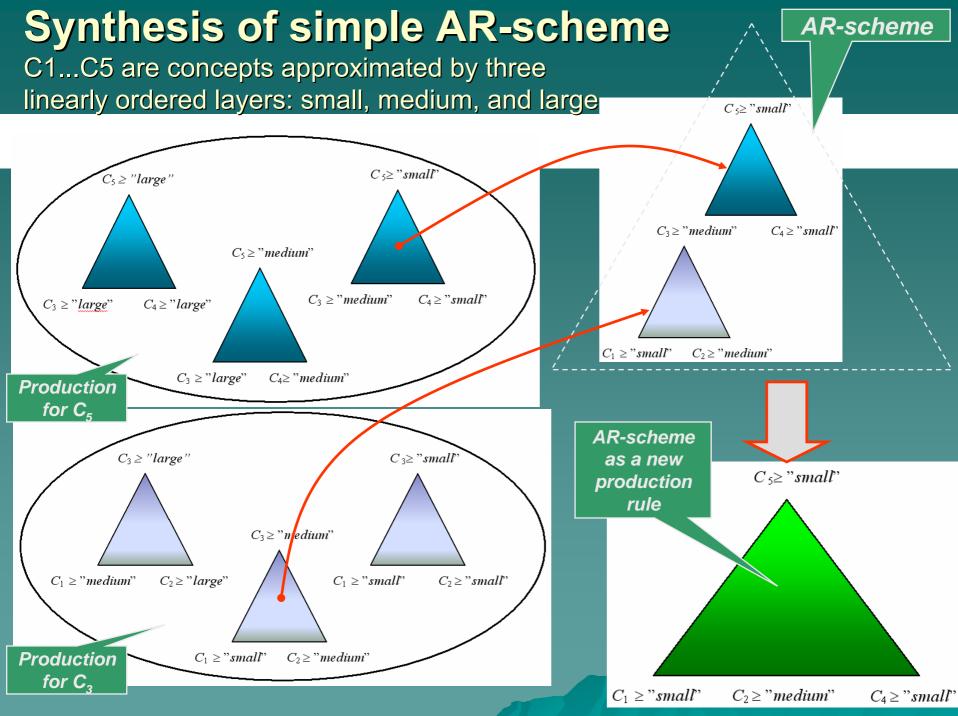
AR-SCHEMES

- Patterns for complex objects can be expressed by AR-schemes
- AR-scheme is a multi-level tree structure
- Root: inclusion degree of constructed higher level patterns into the root concept
- Leaves: inclusion degrees of primitive patterns into the concepts corresponding to leaves

 $C_4 \ge$ "medium" $C_3 \ge$ "large"

 $C_1 \geq "large" \quad C_2 \geq "small"$

 $C_5 \ge$ "small"



MAP OF GRANULES

FROM NEIGHBORHOOGS OF OBJECTS TO CLASSIFIERS, BEHAVIORAL PATTERNS AND ADAPTIVE SCHEMES

APPLICATIONS: COMPLEX CONCEPT APPROXIMATION

WITAS PROJECT

www.ida.liu.se/ext/witas/

CONTROL OF AUV



P. Doherty, W. Łukaszewicz, A. Skowron A. Szałas: Knowledge engineering: Rough set approach, Springer 2006.

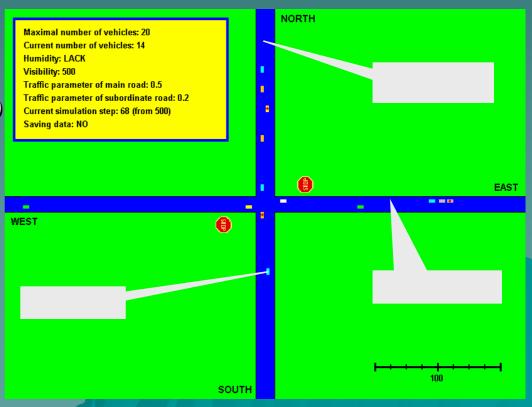
UAV

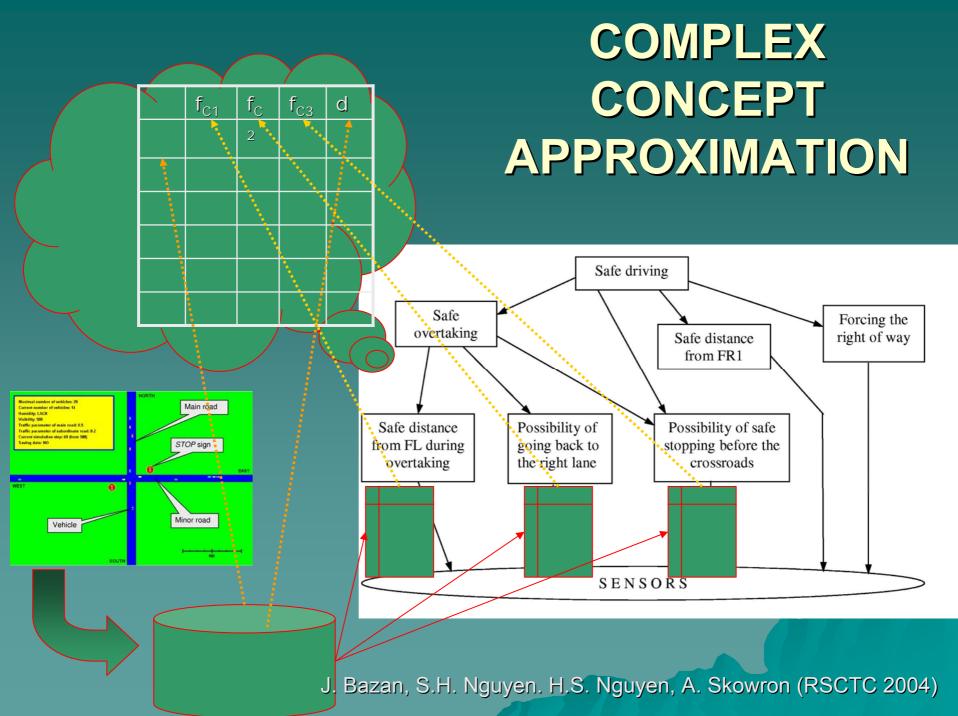


Road simulator

(http://logic.mimuw.edu.pl/~bazan/simulator/)

- The board of simulation
- Vehicles as autonomous agents
- Operations and maneuvers
- Simulation of senso
- Concepts
- Data storing





Results of experiments for concept: "Is the vehicle driving safely?"

Decision	Method	Accuracy	Coverage	Real
class				accuracy
YES	RS	0.978	0.946	0.925
	ARS	0.962	0.992	0.954
NO	RS	0.633	0.740	0.468
	ARS	0.862	0.890	0.767
All classes	RS	0.964	0.935	0.901
(YES + NO)	ARS	0.958	0.987	0.945

Real accuracy = Accuracy * Coverage

Results of experiments for concept: "Is the vehicle driving safely?"

Learning time and the rule set size

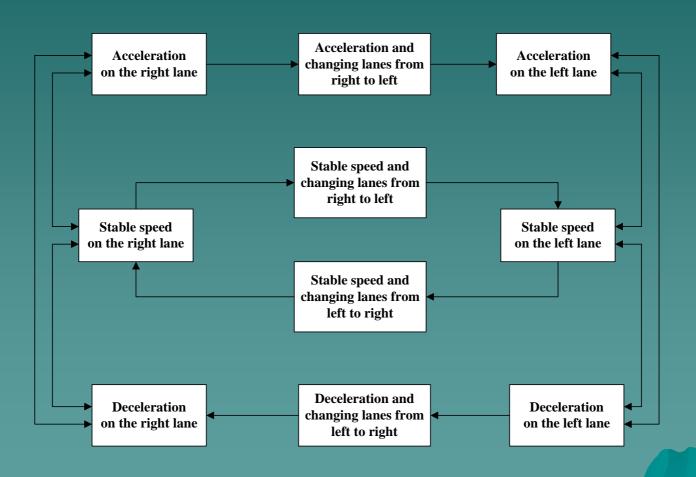
Method	Learning time	Rule set size
RS	801 seconds	835
ARS	247 seconds	189 (average number of rules from all concepts)

COMPLEX BEHAVIORAL PATTERNS

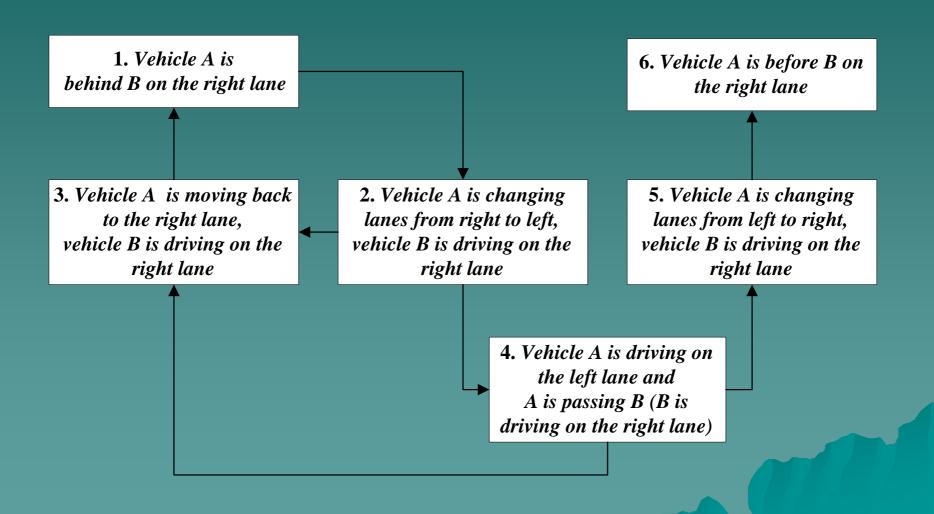
8

PERCEPTION RULES FOR **SELECTION RELEVANT PARTS:** On-line Elimination of Non-Relevant Parts of Complex Objects in Behavioral Pattern Identification (ENP method)

An example of behavioral graph for single vehicle



Behavioral graph for a group of objects (two vehicle of objects during overtaking)



Data sets

- The experiments have been performed on the data sets obtained from the road simulator.
 - A train data set consists of 17553 objects generated by the road simulator during one thousand of simulation steps.
 - A test data set consists of 17765 objects collected during another (completely different) session with the road simulator.
- Train&test method has been performed to estimate accuracy.

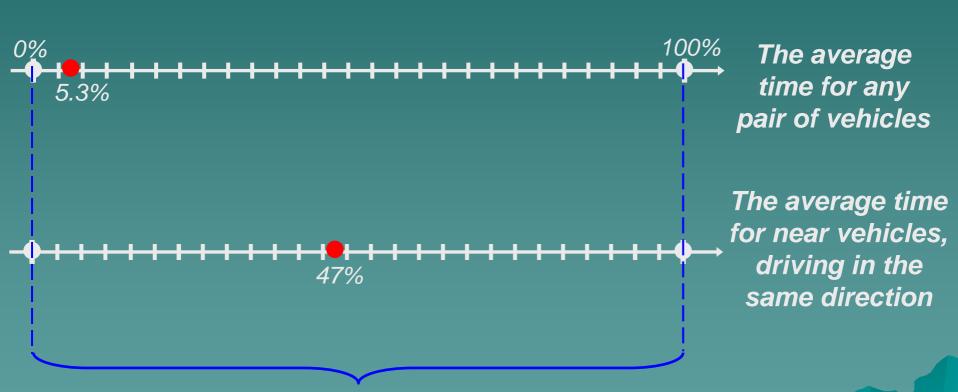
Results of experiments for the overtaking pattern

Decision	Method	Accuracy	Coverage	Real
class				accuracy
YES	RS-D	0.800	0.757	0.606
	BP	0.923	1.0	0.923
	BP-E	0.883	1.0	0.883
NO	RS-D	0.998	0.977	0.975
	BP	0.993	1.0	0.993
	BP-E	0.998	1.0	0.998
All classes	RS-D	0.990	0.996	0.956
(YES +	BP	0.989	1.0	0.989
NO)	BP-E	0.992	1.0	0.992

Real accuracy = Accuracy * Coverage

Reduction of the perception time in case of ENP method

(for decision class NO – no overtaking)



The sequence of time windows that is given for perception of behavioral patterns

COMPLEX DYNAMIC SYSTEMS (AUTONOMOUS MULTIAGENT SYSTEMS)

- Systems of complex objects with the following features:
 - objects are changing over time
 - dependencies between objects
 - cooperation between objects
 - ability to perform flexible autonomous complex actions by objects

Examples:

- Complex dynamic system: a given patient
- Complex object: a disease of the patient (e.g., respiratory failure)

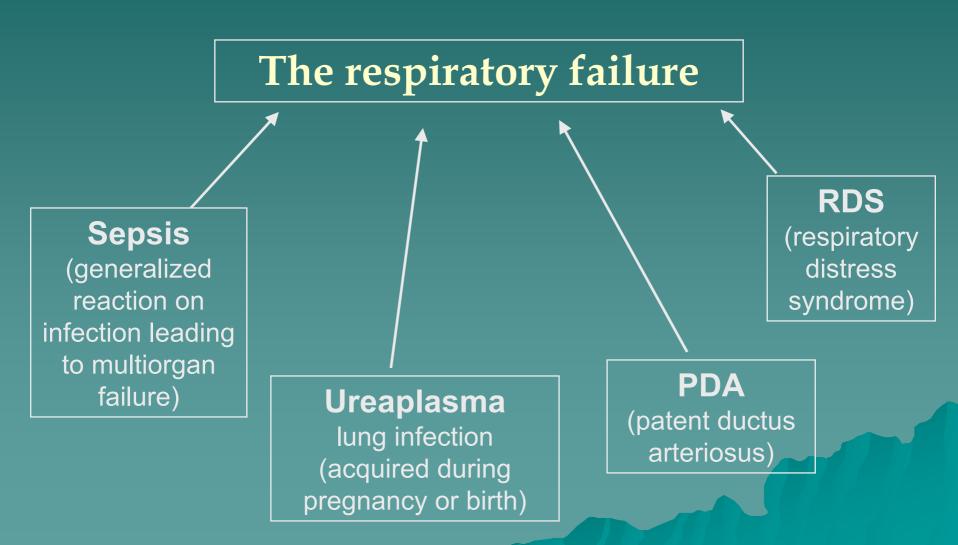
DATA

- ◆ The experiments have been performed on the data sets obtained from Neonatal Intensive Care Unit in Department of Pediatrics, Collegium Medicum, Jagiellonian University, Cracow.
 - The data were collected between 2002 and 2004.
 - The detailed information about treatment of 340 newborns:
 - perinatal history, birth weight, gestational age, lab tests results, imagine techniques results, detailed diagnoses during hospitalization, procedures and medication.

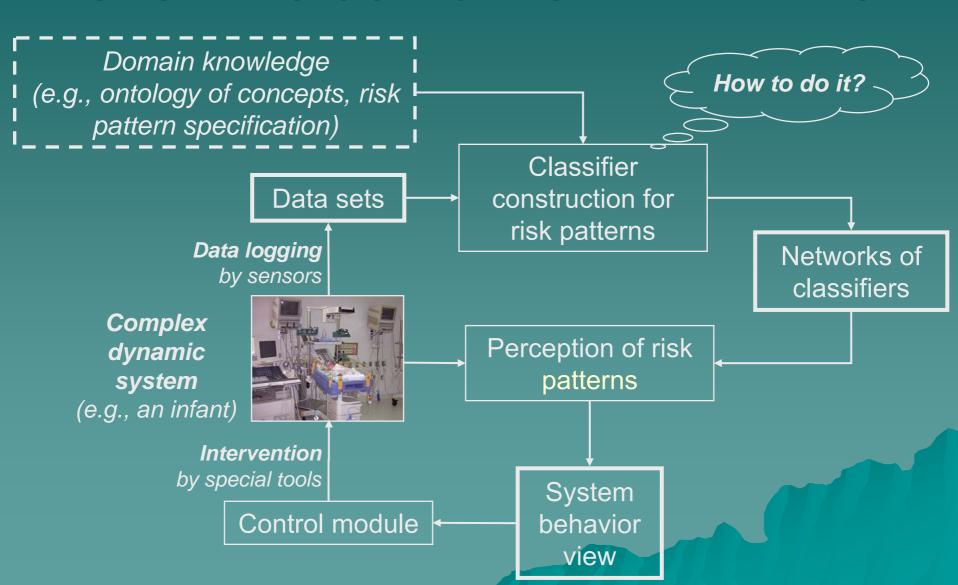
THE RESPIRATORY FAILURE

- The respiratory failure develops when the rate of gas exchange between the atmosphere and blood is unable to match the body's metabolic demands
- Arterial blood gas can be used to define respiratory failure – lower level of blood oxygen and accumulation of carbon dioxide
 - Clinical symptoms: increased rate of breathing, accessory respiratory muscles use, peripheral cyanosis
 - Other useful procedures: X-ray lung examination, lung biopsy, bronchoalveolar lavage, echocardiography

THE RESPIRATORY FAILURE AS A COMPLEX OBJECT

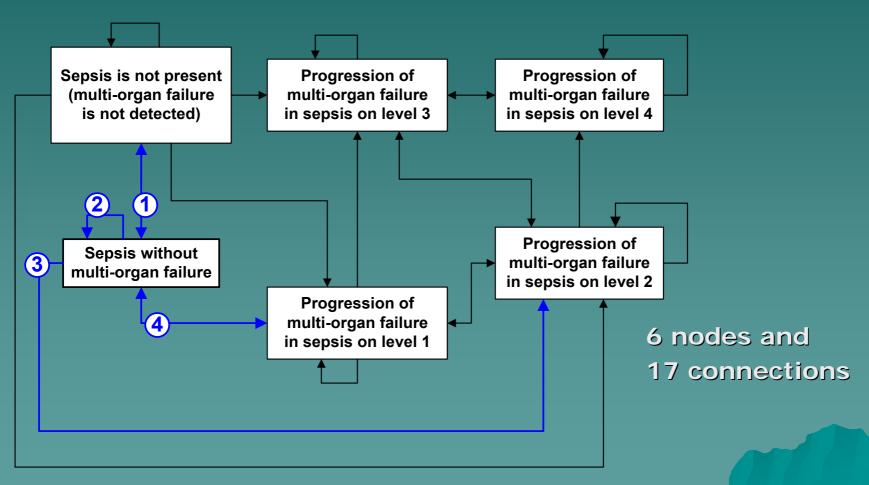


MONITORING OF COMPLEX DYNAMIC SYSTEMS USING RISK PATTERNS



AN EXAMPLE OF BEHAVIORAL GRAPH

(the simple model of behavior for a single patient in sepsis)



Four possibilities of transition from the node: Sepsis without multi-organ failure

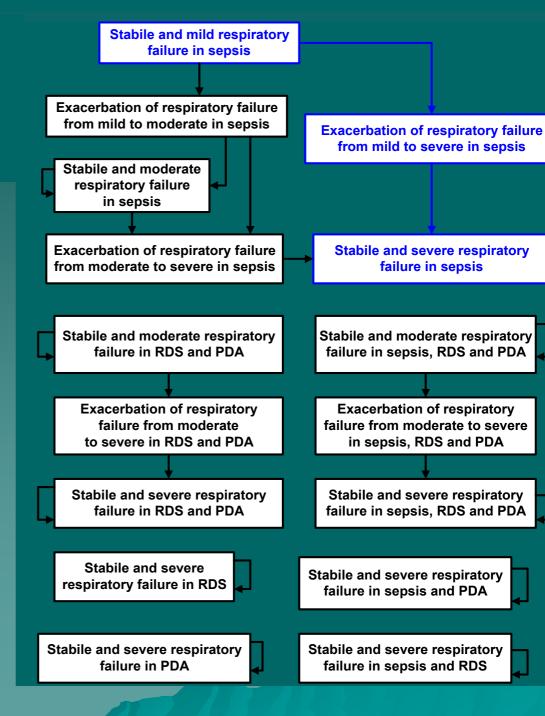
Behavioral graph as a behavioral pattern (the risk pattern of death due to respiratory failure)

- The visualization of infant behavior by a path in the behavioral graph (behavioral pattern) as a classifier:
 - Path:
 - Stable and mild respiratory failure in sepsis,
 - Stable and severe respiratory failure in sepsis
 - Exacerbation of respiratory failure from mild to severe in sepsis,

matches the graph

- Path:
 - Stable and severe respiratory failure in sepsis,
 - Exacerbation of respiratory failure from moderate to severe in sepsis,
 - Stable and moderate respiratory failure in sepsis

doesn't match the graph



Data sets

- ◆ The experiments have been performed on the data sets obtained from Neonatal Intensive Care Unit in Department of Pediatrics, Collegium Medicum, Jagiellonian University, Cracow.
 - The data were collected between 2002 and 2004.
 - The detailed information about treatment of 340 newborns:
 - perinatal history, birth weight, gestational age, lab tests results, imagine techniques results, detailed diagnoses during hospitalization, procedures and medication.
- Train&test method has been performed to estimate accuracy, sensitivity and specificity.
 - A train data set consists of 5810 objects and a test data set consists of 5289 objects

Results of experiments for the risk pattern of death due to respiratory failure

Decision class	Results		
Yes (the high risk of death)	0.992 (sensitivity)		
No (the low risk of death)	0.936 (specificity)		
All classes (Yes + No)	0.956 (accuracy)		

Measures description:

- sensitivity the proportion those cases having a positive test result of all positive cases tested,
- specificity the proportion of true negatives of all the negative cases tested,
- accuracy the ratio of the number of all properly classified cases to the total number of tested cases.

EXPERIMENTS FOR THE AUTOMATED PLANNING OF TREATMENT

- ◆ As a measure of planning success (or failure) in our experiments, we use the special classifier that can predict the similarity between two plans as a number between 0.0 and 1.0.
 - This classifier has been constructed on the basis of the ontology specified by human experts and data sets
- The average similarity between plans for all tested situations was 0.82

THE PROBLEM OF COMPARISON OF PLANS

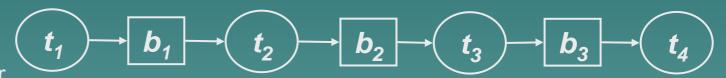
Plan 1:

(e.g., proposed by human experts)



Plan 2:

(e.g., generated automatically by our computer system)

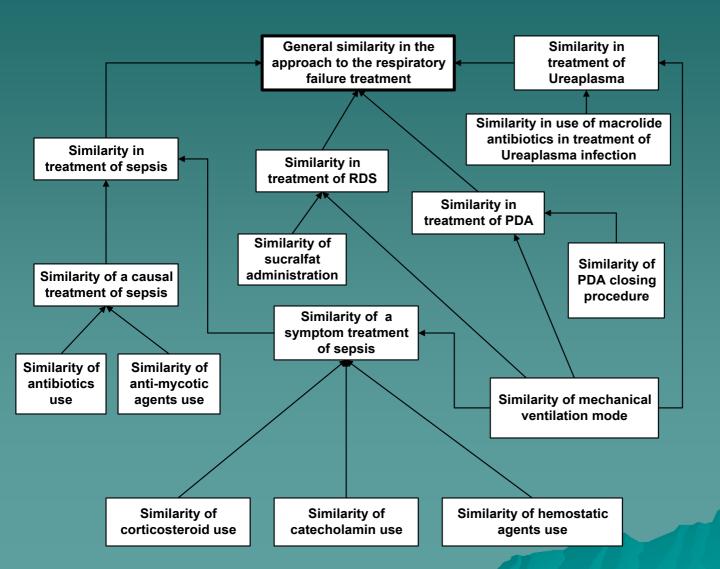


Problem: How to compare Plan 1 and Plan 2?

Solution: A tool to estimate the similarity between plans.

An example of medical ontology

to support the estimation of similarity between plans of the treatment of newborn infants with the respiratory failure



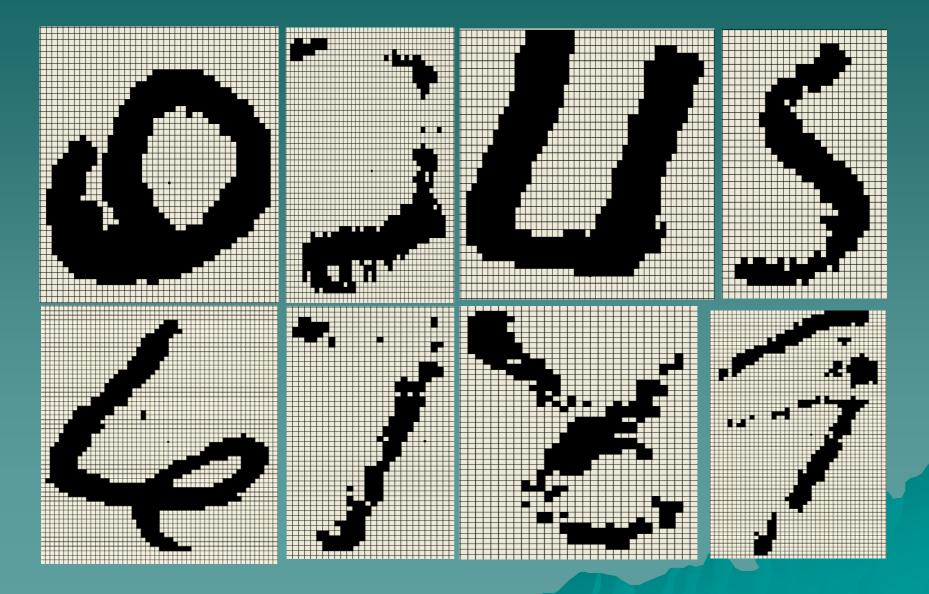
16 concepts and 18 connections

Any concept represents different aspect of similarity between medical plans

UNDERSTANDING DOMAIN KNOWLEDGE

CONCEPT APPROXIMATION USING ROUGH MEREOLOGY

HARD SAMPLES



SUNSPOT CLASSIFICATION





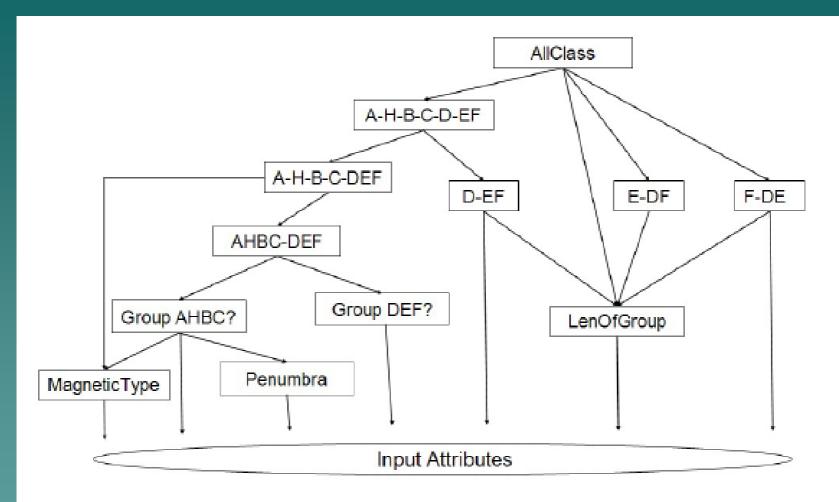


Figure 6. The concept hierarchy for sunspot classification problem

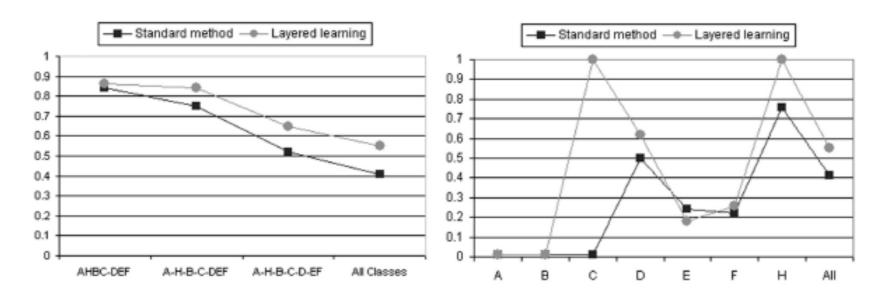
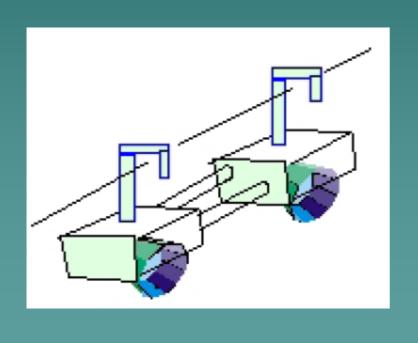
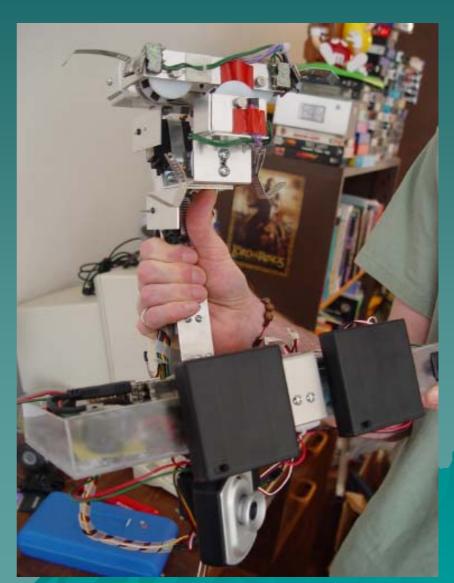


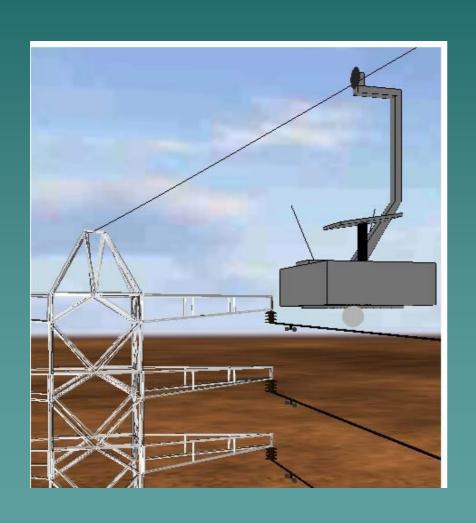
Figure 7. Left: the classification accuracy of standard and layered method for some concepts in the ontology presented in Fig. 6. Right: the classification accuracy of standard and layered method for particular decision classes.

ALICE II PROJECT



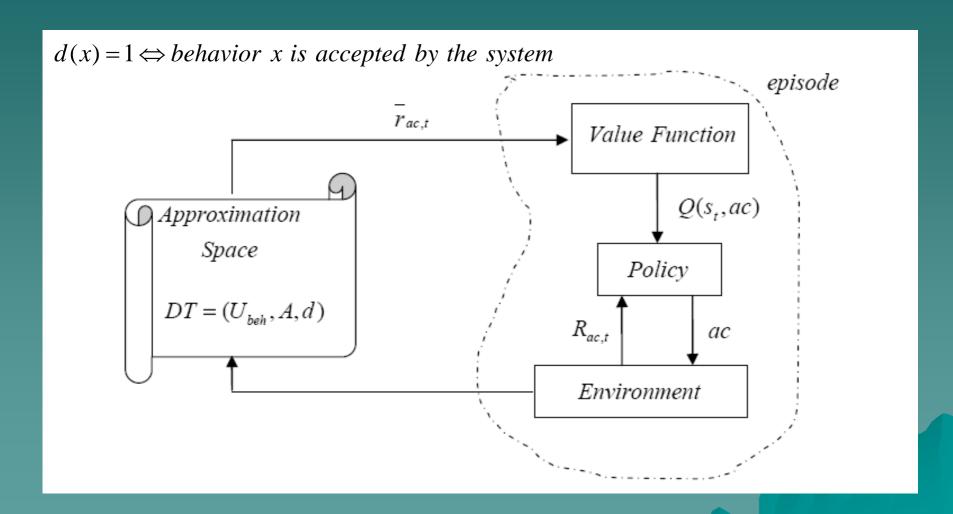


ALICE II LINE-CRAWLING ROBOT





REINFORCEMENT LEARNING



REINFORCEMENT LEARNING

 $B-lower\ approximation\ of\ D=\{x\in U:d(x)=1\}$

$$v_i = v(B_{ac}(x_{t-i}), B_*(D))$$

all B-indiscernible situations with x_{t-i} with action $ac = ac_{t-i}$

$$\bar{r}_{ac,t} = \frac{1}{n} \sum_{i=1}^{n} v_i; \ W_{ac,t} = \frac{1}{n} \sum_{i=1}^{n} \bar{r}_{ac,t-i}$$

$$Q(s_{t}, ac) = \sum_{i=1}^{n} \frac{r_{ac,t-i}}{W_{ac,t-i}} \left[R_{ac,t-i} - Q(s_{t-i}, ac) \right]$$

the sum of rewards for ac in the episode ending at t-i

REINFORCEMENT LEARNING

- ◆ REINFORCEMENT LEARNING = learning of complex concept making it possible to estimate the degree of Q(s_t,ac)
- Domain knowledge: explanations of arguments for and against performing a given action in a given situation

UNDERSTANDING THE ORGANIZATION AND PRINCIPLES OF HIGHER BRAIN FUNCTIONS: HIERARCHICAL LEARNING

- Organization of cortex for instance visual cortex

 –is strongly hierarchical.
- Hierarchical learning systems show superior performance in several engineering applications.
- This is just one of several possible connections, still to be characterized, between learning theory and the ultimate problem in natural science – the organization and the principles of higher brain functions.

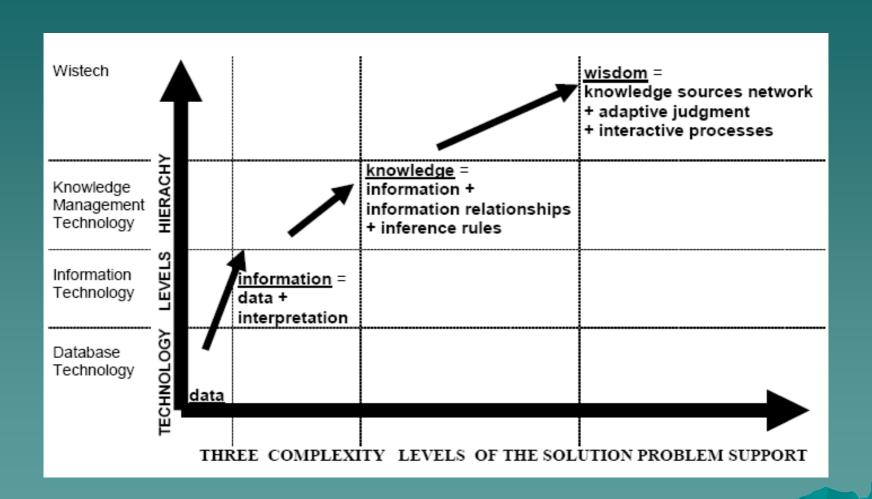
NUEROSCIENCE: T. Poggio, S. Smale Notices AMS, Vol.50,
May 2003

ONTOLOGY LEARNING

 Strategies for learning a relevant ontology for solving problems from a given (small) class of problems

NEW GRANULES

WISDOM GRANULES



WISDOM = RIGHTLY JUDGING

WISDOM - inference engine interacting with real-life environment which is able to identify important problems, find for them satisfactory solutions having in mind real-life constraints, available knowledge sources, and personal experience

WISDOM EQUATION

wisdom =

knowledge sources network +

adaptive judgement +

interactive processes

RGC

The main objective of Wistech

is to automate support for the process leading to wise actions.

These activities cover all areas of man's activities, from the economy, through medicine, education, research, development etc.

In this context, one can clearly see how large a role in the future may be played by advancing Wistech.

WISTECH: RGC for APPROXIMATE REASONING in MAS

- Learning ontology
- Knowledge representation issues
- Selection knowledge relevant for problem solving from knowledge networks
- Approximation of concepts representing intentions, desires, beliefs, social patterns,...
- Reasoning about changes
- Negotiation, conflict resolution
- Planning
- Cooperation
- Adaptation
- Interaction
- Autonomy computing

> ...

ONTOLOGY APPROXIMATION IN DISTRUBUTED ENVIRONMENTS: TOWARD PERCEPTION LOGIC

PERCEPTION LOGIC

family of approximation spaces for a local ontology of concepts

A system of LOCAL LOGICS interacting by using constrained sums performed on approximation spaces for achieving the goals.

approximation spaces relevant for

- discovery of interesting patterns and dependencies
- granulation of patterns to new concepts

PERCEPTION LOGIC

- Perception Logic: a system of local logics interacting and evolving in time based on calculi of approximation spaces
- Local Logics + Interactions + reasoning about changes
- Local Logics:
 - over ontology of concepts
 - ontology is approximated in a hierarchical way starting from sensory concepts
- Interactions modelled by constrained sums:
 - new approximation spaces, new concepts and dependencies, new local logics
 - evolutionary strategies for discovery of relevant interaction and local logics (MAS: negotiations, conflict resolution, coalition formation,...)
- Tools for reasoning about such systems

Developing Wistech under a Wistech Network (WN) co-operating with one another in accordance with open principles based on Open Innovation Principles

H. W. Chesbrough, Open Innovation: The New Imperative for Creating and Profiting from Technology, Harvard Business School Publishing, Cambridge MA, 2003.

Phase/	Summary	Spider	Conceptual	Wisdom	Wisdom Assistant
Product	Summing	Spicer	Clustering	Extraction	VVISCOIII 213313CCIIC
Document	Document	Document Spider	Document	Document	Document Wisdom
Manager	Summary	1	Conceptual	Wisdom	Assistant
J			Clustering	Extraction	
Job Market	Job Market	Job Market Spider	Job Market	Job Market	Job Market Wisdom
	Summary	•	Conceptual	Wisdom	Assistant
	_		Clustering	Extraction	
Brand	Brand Monitoring	Brand Monitoring	Brand Monitoring	Brand Monitoring	Brand Monitoring
Monitoring	Summary	Spider	Conceptual	Wisdom	Wisdom Assistant
			Clustering	Extraction	
World	World	World	World	World	World
Communicati	Communication	Communication	Communication	Communication	Communication
on	Summary	Spider	Conceptual	Wisdom	Wisdom Assistant
			Clustering	Extraction	
World Forex	World Forex	World Forex Spider	World Forex	World Forex	World Forex
	Summary		Conceptual	Wisdom	Wisdom Assistant
			Clustering	Extraction	
World Stock	World Stock	World Stock Market	World Stock Market	World Stock	World Stock Market
Market	Market Summary	Spider	Conceptual	Market Wisdom	Wisdom Assistant
			Clustering	Extraction	
World	World Tourist	World Tourist	World Tourist	World Tourist	World Tourist
Tourist	Summary	Spider	Conceptual	Wisdom	Wisdom Assistant
			Clustering	Extraction	
Physician	Physician	Physician Spider	Physician	Physician	Physician Wisdom
	Summary		Conceptual	Wisdom	Assistant
_			Clustering	Extraction	
Lawyer	Lawyer Summary	Lawyer Spider	Lawyer Conceptual	Lawyer Wisdom	Lawyer Wisdom
			Clustering	Extraction	Assistant

SUMMARY

- The approach covers the approximation of vague concepts and dependencies specified in a given ontology.
- Our ontology is presented in the framework of information granule calculi.
- The outlined methods for hierarchical construction of patterns, classifiers and vague dependency approximation create the basic step in our project aiming at developing approximate reasoning methods in distributed systems.

SUMMARY: TOWARD APPROXIMATE REASONING IN MAS AND CAS

- STRATEGIES FOR GENERATION AND ANALYSIS OF GRANULES ON DIFFERENT LEVELS OF HIERARCHY
 - **◆DATA GRANULES**
 - **♦ INFORMATION GRANULES**
 - KNOWLEDGE GRANULES
 - WISDOM GRANULES
 - ◆METAGRANULES FOR APPROXIMATE REASONING ABOUT GRANULES

The following comment from G.W. Leibniz on the idea to automate the processing of concepts representing thoughts should also not surprise us:

No one else, I believe, has noticed this, because if they had ... they would have dropped everything in order to deal with it; because there is nothing greater that man could do.

logic.mimuw.edu.pl

THANK YOU!

CALCULI OF APPROXIMATION SPACES IN DYNAMICAL SYSTEM MODELLING

- The approximation spaces are allocated in different agents and they are used to approximate concepts, discover new concepts and the dependencies between concepts.
- ◆ The agents or their teams can interact what helps them to construct new approximation spaces more relevant for understanding of their environments (e.g., to better approximate concepts or to discover new concepts) and to select more relevant actions to achieve agent goals.
- The interactions are realized by constrained sums.
- The dynamics of such a system is based on learning by individual agents and agent teams new concepts and dependencies that are next used for performing further actions by agents or their teams.

CALCULI OF APPROXIMATION SPACES IN DYNAMICAL SYSTEM MODELLING

- A system of evolving local theories of agents belonging to the system.
- These theories, are changing in time as the result of interactions between agents.
- ◆ The agents are learning to select the relevant, for their goals, behavior on the basis of properties of changes in their theories and in information about satisfiability of concepts in their theories.