An Agent-Based Knowledge Acquisition Platform

David Sánchez, David Isern, Antonio Moreno

University Rovira i Virgili (URV)
Department of Computer Science and Mathematics (DEIM)
Av. Països Catalans, 26. 43007 Tarragona, Catalonia, Spain.
{david.sanchez, david.isern, antonio.moreno}@urv.net

Abstract. Accessing up-to-date information in a fast and easy way implies the necessity of information management tools to explore and analyse the huge number of available electronic resources. The Web offers a large amount of valuable information, but its human-oriented representation and its size makes extremely difficult any kind of computer-based processing. In this paper, a combination of distributed AI and information extraction techniques is proposed to tackle this problem. In particular, we have designed a *multiagent system* that composes ontologies from *taxonomies* of terms. Moreover, the obtained ontology is used to represent, in a structured way, the currently available web resources. The paper analyses the application of this approach in some examples in the medical domain.

1 Introduction

Researchers typically assess the evolution of their discipline by reading scientific journals, attending conferences or, quite often, by hearsay. On the other side, the Web offers a way for fast data access and information exchange that could represent a great help but, unfortunately, it is impossible to analyse manually due to the huge amount of available resources and their weak structure.

These electronic repositories are usually accessed by means of keyword-based search engines (e.g. Google, Altavista) allowing a user to retrieve information by stating a combination of words that have to appear in the retrieved documents. This type of search usually suffers from two problems derived from the nature of the query and the lack of structure in the documents: a) the difficulty to set the most appropriate and restrictive search query, and b) the tedious evaluation of the huge amount of potential resources obtained. Moreover, while search engines provide support for the automatic retrieval of information, the tasks of extracting and structuring relevant information and its further processing remain to be done by the human user. In this sense, a methodology for representing the web resources in a structured way depending on the main topics of a desired domain can be a great help for any researcher. Thus, the important point is to find a way of creating and representing the domain's knowledge structure efficiently. Here is where ontologies [4] become indispensable.

In general, ontologies allow organizing and centralizing knowledge in a formal, machine, and human understandable way, making themselves an essential component to many knowledge-intensive services like the Semantic Web [3], knowledge management, and electronic commerce. However, they are traditionally built entirely by hand, capturing the knowledge in a static way. This knowledge is usually evolvable (especially in technological domains) and an ontology maintenance process is required to keep the ontological knowledge up-to-date. A computer-based ontology construction process becomes a very important deal for information engineers, especially with highly dynamic domains like the Web. The building process has been described by several authors [6, 10] by five main steps: *i*) identification of concepts and instances, *ii*) word sense disambiguation, *iii*) taxonomic construction, *iv*) identification of non taxonomic relations, and *v*) ontology population.

However, the processing of a huge repository like the Web is a very time consuming task. In order to handle this problem, the agent paradigm is a promising technology for information retrieval. *Multiagent systems* provide some advantages with respect to traditional systems such as scalability, flexibility and autonomy [18] and they are very suitable for implementing dynamic and distributed systems. Several projects applying MAS to information retrieval and knowledge acquisition such as [7, 13] indicates that agents can provide domain independence and flexibility to this type of systems. Although agents could operate in a completely autonomous way, in our case the supervision of a human expert is recommended to limit the search only to the knowledge areas that are really interesting for the desired domain, maximizing the throughput of the learning process.

So, in this paper we present *a combination of new AI methodologies to extract knowledge from the Web to build semi-automatically an ontology of concepts and web resources for a given domain, through a distributed agent-based platform*.

The rest of the paper is organized as follow. Section 2 introduces the methodology used to obtain taxonomies of terms, discover instances and propose non-taxonomical relations. Section 3 describes the agent-based platform used to compose a final ontology from individual taxonomies of terms, under the supervision of a human expert. Section 4 explains the way of representing the results and discusses the evaluation against other information retrieval systems in relation to *precision* and *recall*. The final section contains the conclusions and proposes lines of future work.

2 Taxonomy creation from unstructured documents

The basis of the proposed ontology construction process is the intensive use of a methodology [16] for constructing taxonomies of terms and web resources that are relevant for a domain. The most important characteristic of the method is that the whole process is performed automatically and autonomously directly from the Web.

The algorithm is based on analysing a large number of web sites in order to find important concepts for a domain by studying the neighbourhood of an initial keyword that characterizes the desired searched domain. Concretely, in the English language, the immediate anterior word for a keyword is frequently classifying it (expressing a semantic specialization of the meaning), whereas the immediate posterior one represents the domain where it is applied [8]. So, on the one hand, the previous word for a specific key*word* is used for obtaining the taxonomical hierarchy of terms (e.g. *breast* *cancer* will be a subclass of *cancer*). The process is repeated recursively in order to create deeper-level subclasses (e.g. *metastatic breast cancer* is a subclass of *breast cancer*). On the other hand, the *posterior word* for the specific *keyword* is used to categorise the web resources, considered as a tag that expresses the context in where the searched domain is applied (e.g. *colorectal cancer research* will be a domain of application of *colorectal cancer* covered on a specific web document). Moreover, particular examples (i.e. proper names) for a discovered concept are found based on the way that they are represented in the text, considering them as *instances* in the defined hierarchy. In both cases, the most representative web sites for each class or instance are also retrieved and categorised according to the specific topic covered. Finally, a *polysemy detection algorithm* is performed in order to disambiguate polysemic domains. This algorithm performs a clusterisation of the discovered sub-classes, in order to group the most similar ones, detecting automatically different sets of terms that correspond with different word senses.

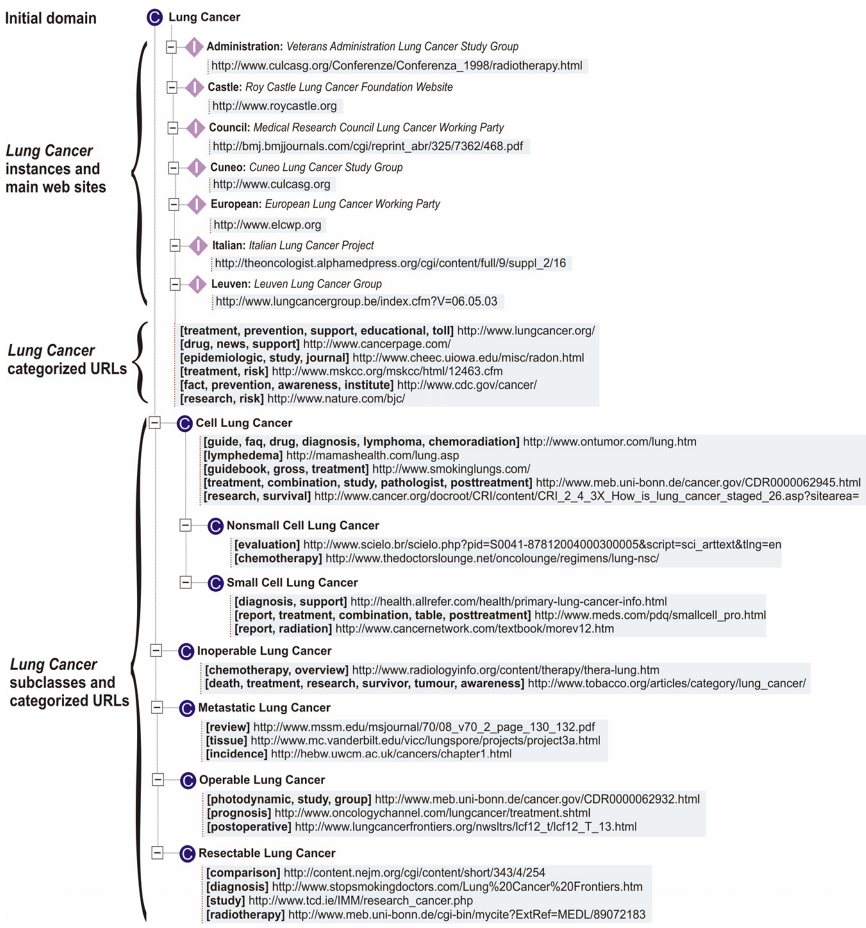

Fig. 1. Examples of classes, instances and URLs discovered for the *lung cancer* domain.

The result of the process is a hierarchical and categorized organization of the available resources according to the *classes* –concepts- and *instances* -particular examples- discovered for the given domain (see Fig. 1 for an example of the obtained hierarchy for the *lung cancer* domain).

In order to detect and extract relevant information from the Web, the method relies on a search engine for searching and accessing the available web resources. It constructs dynamically the appropriate search queries for the search engine, obtaining the most adequate corpus of web resources at each time. Moreover, the search engine is also used for checking the relevance of the extracted terms and evaluating the strength of the taxonomical relationships between them through a statistical analysis based on the number of estimated results available in the Web.

As an additional step, the methodology analyses and extracts the sentences from where concepts of the taxonomy are extracted. This knowledge could be very useful in a latter stage of the ontology construction process to obtain more complex relations like non-taxonomical ones. Concretely, analysing the subject and the object of a sentence we can infer a new relationship between them according to the verb (which can express a taxonomical or non taxonomical relationship). Some authors such as [9] have been used that approach successfully for ontology learning.

In order to ease a future evaluation of the discovered set of sentences, the analyser applies several syntactic processing tools to obtain a simplified but meaningful view of the original sentence. Then it uses several *Natural Language Processing* tools to analyse the sentence syntactically and select only those ones that express knowledge in a simplified way, excluding ambiguous syntactical constructions like conditionals or futures (some examples of selected sentences can be found in Table 1). Sentences that accomplish a set of simplicity rules like the described ones are typically called *text nuggets,* and they are commonly used for different knowledge acquisition and information extraction tasks [14].

As will be shown in section 3, the additional knowledge acquired by the analysis of those sentences will be used by the proposed agent-based platform (with the supervision of a human expert) in order to discover non-taxonomical relationships, expand the analysis and build a final ontology for the domain.

3 Distributed ontology building process

Once an initial taxonomy of terms that are relevant for the domain is obtained through the described methodology we can use the discovered knowledge to find more complex relationships and performs further analysis. Concretely, using the extracted sentences for particular discovered terms, we can find other concepts that are related to the domain with a certain relationship (in many cases, a non-taxonomical one). With that new concept, a new taxonomical analysis in the same way as has been described can be performed, obtaining a complex network of knowledge with the rich semantics that ontologies require. In this case, some of the previously obtained knowledge (e.g. the concept from which this one has been obtained) can be added to the queries performed in order to restrict and contextualize the search, obtain the most suitable resources and, in consequence, improve the throughput of the learning process.

However, the described process for creating taxonomies is a very time consuming task, especially when dealing with such a general and enormous repository as the Web. Concretely, accessing and downloading web documents online overheads the execution and affects seriously on the system's performance. So, if we plan to perform several taxonomical analysis iteratively, they may suppose a computational cost that is hard to be assumed by a centralized approach.

However, as several tasks of the learning process can be performed concurrently (e.g. construct different taxonomies, evaluate text nuggets, cluster terms, etc.) a distributed approach can promise a great improvement over a centralised one. However, as the execution requirements are very dynamic as they are defined by the knowledge acquired at execution time, coordination, flexibility and dynamicity are fundamental. In order to handle this problem, the agent paradigm is a promising technology for information retrieval [18].

Therefore, we present a supervised agent-based platform for building ontologies from the combination of taxonomies of several interrelated terms that are selected into a semiautomatic way.

3.1 Multiagent system architecture

The system is composed of several autonomous entities (agents) that could be deployed around a network. Each agent can be considered as an execution unit that follows a particularly modeled behavior and interacts (communicates) with other ones, coordinating their execution for achieving a common goal. The inter-agent communication allows them to share partial results and coordinate their efforts in order to construct the final ontology. Concretely, there are three kinds of agents in the MAS:

a) *User Agent* (UA): allows the human expert to interact with the system. Through this agent, she can configure, initialize and control the construction process in order to obtain an ontology that fits with her interests.

b) *Internet Agent* (IA): It implements the taxonomy construction methodology described in section 2. For a specific initial query, it performs the web search process and returns the result. The coordinated action of several IAs with specific queries allows obtaining a set of partial results (taxonomies) that can be joined and interrelated adequately in order to build the final ontology. As this construction process is very time consuming, it is important that these agents could execute concurrently (or in parallel using a computer network) and co-ordinately in order to achieve the highest efficiency. They are highly reusable components as they are initialized and finished dynamically depending on the ontology construction's requirements.

c) *Coordinator Agent* (CA): it coordinates the ontology construction process by receiving orders from the user and creating, configuring and finalising the appropriated IAs to explore web domains. Partial results obtained from the execution of each IA are received and composed to create the final ontology. Note that although the ontology construction is centralised by this agent, its work load in relation to the IAs (even with several machines available) is quite reduced.

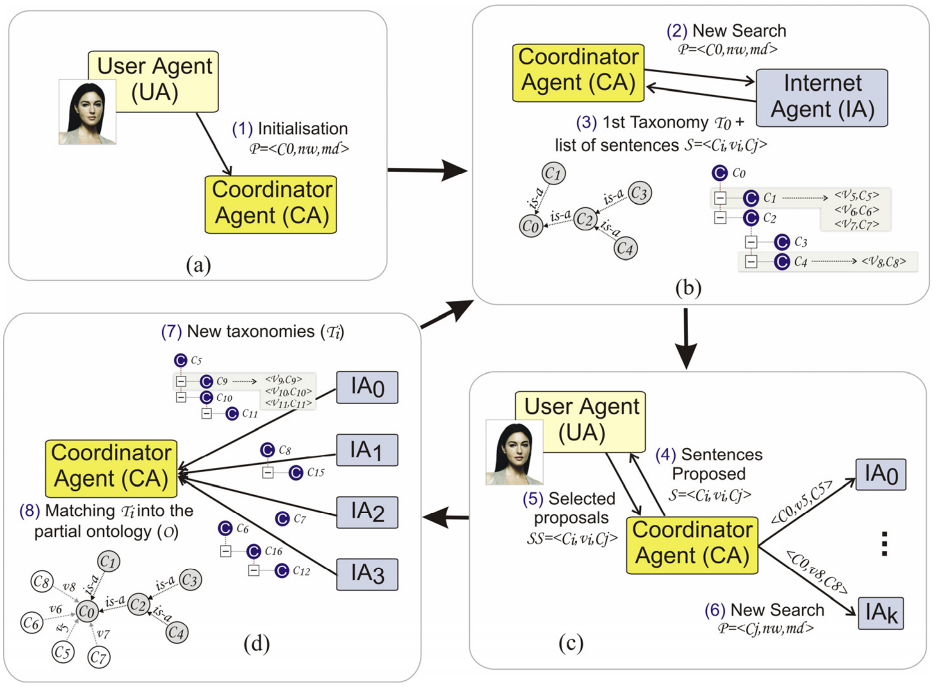

Fig. 2. Steps of the agent-based ontology building process.

3.2 Ontology construction steps

The system is composed of several types of agents that coordinate their work to model and solve the ontology construction process in an efficient and scalable way. Moreover, the supervision of a human expert is required to drive the search towards the knowledge areas in which she is interested, retrieving the highest amount of useful knowledge with the maximum efficiency. As shown in Fig. 2, the steps of the process are:

- The process begins when the user introduces through the UA (see Fig. 2a) the initial parameters of the search: the concept that represents the domain in which she is interested, number of web sites to evaluate and maximum depth of each hierarchy (C_0, *nw* and *md* respectively). The UA sends this information to the CA that will start to construct the associated ontology.
- Then, as it is shown in Fig. 2b, the CA creates and initializes the first IA that starts building the taxonomy that corresponds with the initial concept by creating the initial search query. This IA executes the methodology described in section 2 and returns the following results: the taxonomy T_0 associated with the initial concept, a set of related web pages for each concept and instance found, and a set of relevant sentences (*S*) involving the discovered concepts in the form *Subject (Ci)* + *Verb (Vi)* + *[Preposition]* + *Object (Cj)*. Either the *Subject* or the *Object* must contain a concept in-

cluded in the taxonomy; the other will be a new term. Both are related with the relationship specified by the verb (which can be taxonomical or non-taxonomical) and will be used to retrieve new knowledge for the domain. Examples of sentences for the *cancer* domain are listed in Table 1.

- At this point the CA includes T_0 in the ontology (only composed of the initial concept) and sends the set of sentences to the UA. Then the human expert must consider which of these proposals are correct and/or interesting for her interests in order to continue the search. So, the user selects a subset of sentences (SS) to be evaluated (see Fig. 2c), sending them to the CA. Concretely, each one expresses a relationship between a concept (Ci) that is included in the current partial ontology and a new one (Cj) that defines another domain to be explored; the relation between them will be labelled with the verb Vi (optionally a prepositional verb). For instance, for the discovered concept *breast cancer* for the *cancer* domain, the term *radiotherapy* could be found in a proposed sentence with the verb *receives*, representing a non-taxonomical relation in the ontology. It is also possible to specify which of those concepts and relations will be expanded through new web searches (e.g. *polyp*) and which ones will be included directly into the ontology without evaluation as they represent simple facts (e.g. *hair loss*, *smokers*).
- For each new concept (Cj) extracted from each selected sentence in SS that should be evaluated and has not been considered yet, the CA creates and initializes a set of IAs to build several taxonomies associated to those concepts. In order to maintain the initial context, the CA attaches the root concept (C_0) to all queries specified for each IA, as shown in Fig. 2c. Future improvements can consider also Ci or other previously acquired knowledge for the domain as a bootstrap for contextualizing new searches.
- Again, the IAs builds taxonomies for each new concept concurrently with the specific execution conditions, and finally they send a new set of results (terms, relationships, web resources, instances, etc) to the CA.
- The CA joins the new taxonomies obtained (T_j) by the IAs into the global ontology, by relating the new concepts with the existing ones through the relationships specified by the verbs of the sentences from which they were extracted. Note that the direction of the relation will depend on the role of each concept into the original sentence (subject or object). This is a very critical process in order to obtain a coherent and useful representation of the knowledge. As will be described in the final section, several questions that are not fully developed at this moment regarding the joining of partial results and the processing of verb labels in a semantic way should be considered carefully.
- In addition to the taxonomies, new sentences sets are received by the CA that will send them to the UA to perform a new user-centred evaluation. The process will be repeated while the user selects new interesting terms to be evaluated.

At the end of the process, the CA will store the final ontology composed of several *is-a* taxonomies interrelated with verb-labelled relationships. This knowledge does not give a complete view of the full domain and all its relationships, but it offers a representation of the knowledge in which the user is interested over a specific domain (e.g. *treatments* and *symptoms* of certain types of *cancers*).

4 Case study: *Cancer* domain

For illustrative purposes, in this section we will show the results obtained for a medical domain such as *cancer*, and how it could be useful for a medical researcher to extract or structure this specific knowledge from the Web.

The ontology building process begins with a concept given by the user, in this case, the term *cancer.* As a result of this first search, a taxonomy for that term containing several concepts related to *cancer* is obtained in a tree-based way. For instance, at the first level of the taxonomy, different kinds of *cancer* are identified (e.g *lung cancer*, *colorectal cancer*, *breast cancer*). Recursively, these concepts can contain different subclasses such as *metastatic*, *metachronous* or *nonpolyposis*, in the case of *colorectal cancer*. In addition to this hierarchy, a list of sentences that can contain non taxonomic relations are presented to the user (see examples in Table 1).

Table 1. Examples of sentences obtained from web sites. **Bold** represents the user's choice.

Concept	Sentences
breast cancer	**[breast_cancer][receives][radiotherapy]** **[the pill][protects][against][breast_cancer]** [most breast_cancers][are][ductal carcinomas]
colon cancer	**[colon_cancers][start][as][polyps]**
colorectal cancer	**[most colorectal_cancers][begin][as][a polyp]** [all colorectal_cancer patients][require][a colostomy] [most colorectal_cancers][start][in][the glandular cells]
lung cancer	**[lung_cancer][causes][paraneoplastic syndromes]** **[spiral_scans][find][lung_cancer]** **[lung_cancer][tend to develop][in][smokers]** [asbestos exposure][increases][lung_cancer risk] [lung_cancer treatment][depends][on][tumor size]
cervical cancer	**[cervicography or colposcopy][screening][for][cervical_cancer]**
skin cancer	[ozone depletion][increases][skin_cancer risk] [fair-skinned people][develop][skin_cancers]
cranial radiotherapy	**[cranial_radiotherapy][causes][hair loss]**
beam radiotherapy	[external beam_radiotherapy][include][x-ray therapy]
hyperplastic polyp	**[hyperplastic_polyps][occur][in][normal gastric mucosa]** [colorectal hyperplastic_polyps][are][benign lesions]

The user can select the ones that she considers correct and covers a desired related topic (marked in **bold**). For instance if she is interested in treatment and prevention for *breast cancer* she could select the first two sentences of the list that are related to *radiotherapy* and *the pill* respectively. In this case, *radiotherapy* is evaluated for obtaining a new taxonomy and *the pill* is considered as a simple fact. All those partial results are joined creating a final complete ontology as shown in Fig. 3.

It is important to note that each concept of the final ontology stores valuable web information for the domain, like the categorized list of available web sites and a set of instances that represent proper names like healthcare organizations or institutions (see some examples for the *lung cancer* class in Fig.1).

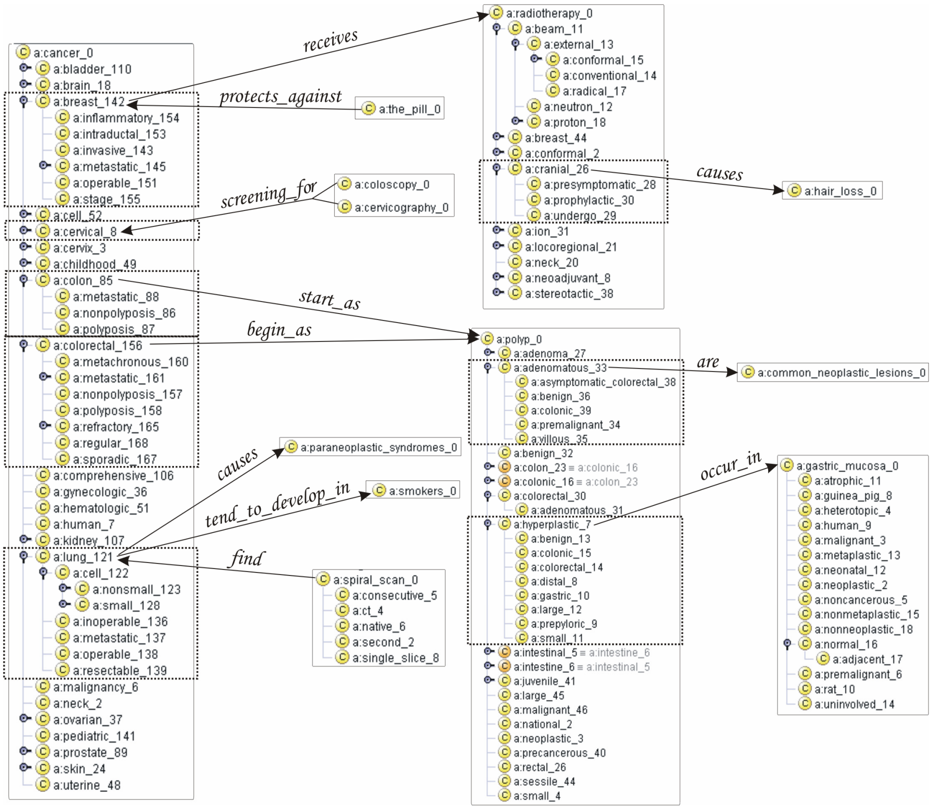

Fig. 3. Part of the ontology obtained for the *cancer* domain.

5 Ontology representation and evaluation

The final ontology composed of the CA is stored in the standard representation language OWL: a semantic markup language for publishing and sharing ontologies on the Web [3]. It is supported by many ontology visualizers and editors, like Protégé, allowing the user to explore, understand, analyse or even modify the ontology easily.

Concerning the evaluation of ontologies, it is recognized to be an open problem [5]. Regarding our proposal, the evaluation is a task that, at this moment, is a matter of current research. However, some evaluations have been performed manually and others are planned to be performed automatically comparing the results against other methodologies or available semantic repositories such as WordNet.

In our case, as the non-taxonomical relations have been selected by a human expert, we have centred the evaluation in the automatically obtained taxonomies. Their evaluation is performed manually at this moment. Whenever it is possible, a representative human made classification is taken as the ideal model of taxonomy to achieve

(*Gold Standard*). Then, several tests for that domain are performed with different sizes of search. For each one, we apply the standard measures *recall* and *precision* (see an example of evaluation for the *Cancer* domain in Fig. 4). As a measure of comparison of the quality of the obtained results against similar available systems, we have evaluated *precision* and *recall* against hand-made web directory services and taxonomical search engines (a detailed evaluation of those system can be found in [17]). For the first case, we have used Yahoo, as it can be considered the most popular human-made directory-based search engine. For the second case, we have selected the taxonomical search engine Clusty as it seems to be the one that provides best and more complete results from the available ones and AlltheWeb as it provides high quality query refinements involving domain related terms. This comparison can also give us an idea of the potential improvement of our structuring and representation of web resources (as shown in section 2) in relation to the results presented by currently available web search engines.

Some conclusions from the evaluation of the results for several domains are:

- The performance of the candidate selection procedure is high as the number of mistakes (incorrectly selected and rejected items) is maintained around 15-20%.
- The growth of the number of discovered concepts (and in consequence the *recall*) follows a logarithmic distribution in relation to the size of the search due to the redundancy of information and the relevancy-based sorting of web sites presented by the search engine. Moreover, when a certain point in which a considerable amount of concepts has been discovered is reached, precision tends to decrease due the growth of false candidates. As a consequence, analysing a large amount of web sites does not imply obtaining better results than with a more reduced corpus due to, among other reasons, the ranking algorithms of the web search engines that potentially present the most relevant and useful web sites in first place.
- Comparing the performance to Yahoo, we see that although its *precision* is the highest, as it has been made by humans, the number of results (*recall*) is quite limited. In relation to the automatic search tools, Clusty and AlltheWeb, the first one tends to return more results (*recall*) but with low *precision* whereas the second one offers an inverse situation. In both cases, the performance (*recall-precision* compromise) is easily surpassed by our proposal.

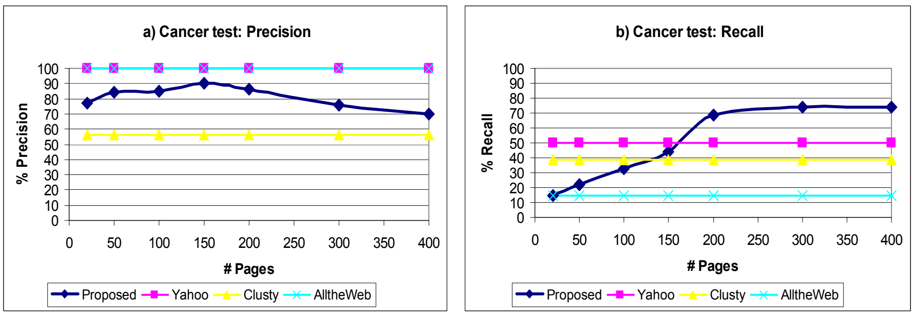

Fig 4. Evaluation of results for the *Cancer* taxonomy.

6 Conclusions

Some authors have been working on ontology learning from different kinds of structured information sources (like data bases or knowledge bases [12]). However, taking into consideration the amount of resources available easily on the Internet, we believe that ontology creation from unstructured documents like web sites is an important line of research. In this sense, many authors [1, 2, 11] are putting their efforts on processing natural language texts. In most cases, an ontology of basic relations (WordNet) is used like a semantic repository. Moreover, in most cases, a relevant corpus of documents carefully selected is used as a starting point. In consequence, these approaches have problems with very concrete domains, or dynamic environments like the Web.

In contrast, our proposal does not start from any kind of predefined knowledge, like WordNet and, in consequence, it can be applied over domains that are not typically considered in semantic repositories as, for example, *Biosensors* [15]. Moreover, the supervision of a human expert that allows driving the search only to the desired areas of knowledge, and the distributed execution based on agents, improve the efficiency and the throughput. These facts result in a scalable and suitable method for acquiring knowledge from a huge and dynamic repository as the Web. The final ontology and the structured list of web sites can be a great help for many knowledge intensive tasks as the Semantic Web [3] and for improving the accessing of web resources in relation to the currently available web search engines.

As future lines of research, some topics can be proposed:

- In the current version, the number of potential sentences returned for human evaluation could be quite high. To assist the user in this task, we plan to perform a filtering process to obtain only the most common or relevant sentences or those whose verbs could be the most important or frequent for the searched domain.
- A detailed study on the casuistry of the taxonomy joining process should be considered carefully in order to detect implicit relationships from redundant or equivalent concepts among the obtained taxonomies.
- Regarding to the discovered verb labelled relationships, in order to obtain an easily usable and interoperable knowledge base, verb expressions should be simplified (e.g. the verbal form *"is usually included into"* expresses a *"part of"* type relationship). In this sense, some basic relationship types can be considered for this task (e.g. *is-a, part-of, related-to, similar-to, cause/effect, result*, etc).
- Several executions from the same domain and parameters in different moments can give us different ontologies, maintaining the results up-to-date. A study about the changes between them can tell us how a domain evolves.
- As mentioned before, ways for automate or at least easing the evaluation of every step of the ontology learning process will be studied.

Acknowledgements

The work has been supported by *Departament d'Universitats, Recerca i Societat de la Informació* of Catalonia.

References

1. Alfonseca E. and Manandhar S.: An unsupervised method for general named entity recognition and automated concept discovery. In Proceedings of the 1st International Conference on General WordNet, Mysore, India, 2002.
2. Ansa O., Hovy E., Aguirre E., Martínez D.: Enriching very large ontologies using the WWW. In Proceedings of of the Ontology Learning Workshop of the European Conference of AI (ECAI-00), Berlin, Germany, 2000.
3. Berners-lee T., Hendler, J., Lassila O.: The semantic web. Scientific American (284):5, pp.34-43. 2001.
4. Fensel D.: Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce. Springer-Verlag, 2001.
5. Fensel D. and Gómez-Pérez, A.: A Survey on Ontological Tools. Deliverable 1.3. OntoWeb: Ontology-based Information Exchange Management, 2000.
6. Fernández-López M., Gómez-Pérez A., Juristo N.: METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In Proceedings of Spring Symposium on Ontological Engineering of AAAI. Stanford University. USA, 1997.
7. Gibbins N., Harris S. and Shadbolt N.: Agent-based semantic web services. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, (1):2, pp 141-154, 2003.
8. Grefenstette G.: SQLET: Short Query Linguistic Expansion Techniques: Palliating One-Word Queries by Providing Intermediate Structure to Text. In Proceedings of: Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology. RIAO'97. LNAI 1299, pp. 97-114. Montreal, Quebec, Canada, 1997.
9. Kavalec, M., Maedche, A. and Skátek, V.: Discovery of Lexical Entries for Non-taxonomic Relations in Ontology Learning. In: Proceedings of SOFSEM 2004: Theory and Practice of Computer Science, LNCS 2932, pp. 249-256. Merin, Czech Republic, 2004.
10. Lamparter S., Ehrig M., Tempich C.: Knowledge Extraction from Classification Schemas. In Proceedings of CoopIS/DOA/ODBASE 2004, LNCS 3290, pp. 618-636, Larnaca, Cyprus, 2004.
11. Maedche A., Staab S.: Discovering conceptual relations from text. In Proceedings of ECAI 2000. IOS Press, Berlin, Germany, 2000.
12. Manzano-Macho D., Gómez-Pérez A.: A Survey of Ontology Learning Methods and Techniques. Deliverable 1.5. OntoWeb: Ontology-based Information Exchange Management, 2000. Available at www.ontoweb.org.
13. Moreno A., Riaño D., Isern D., Bocio J., Sánchez D., Jiménez L.: Knowledge Explotation from the Web. In Proceedings of 5th International Conference on Practical Aspects of Knowledge Management (PAKM 2004). LNAI 3336. Viena, Austria, 2004.
14. Paşca, M.: Finding Instance Names and Alternative Glosses on the Web: In Proceedings of Computational Linguistics and Intelligent Text Processing: 6th International Conference, (CICLing 2005), LNCS 3406, pp.280-292. Mexico City, Mexico, 2005.
15. Sánchez, D. and Moreno, A.: Creating ontologies from Web documents. Recent Advances in Artificial Intelligence Research and Development. IOS Press, vol. 113, pp.11-18. 2004.
16. Sánchez, D. and Moreno, A.: Automatic Generation of Taxonomies from the WWW. In: Proceedings of the 5th International Conference on Practical Aspects of Knowledge Management. LNAI 3336, Vienna, Austria, 2004.
17. Yeol Yoo, S. and Hoffmann, A.: A New Approach for Concept-Based Web Search. In Proceedings of AI 2003: Advances in Artificial Intelligence: 16th Australian Conference on AI, LNAI 2903, pp. 65–76. Perth, Australia, 2003.
18. Wooldridge, M.: An Introduction to multiagent systems. John Wiley and Sons, Ltd., West Sussex, England, 2002.