Dominique Renaud, Cecilia Zanni-Merk, Philippe Bouché, Nathalie Gartiser, François Rousselot
LGECO – INSA de Strasbourg, Strasbourg, France
{firstname.lastname}@insa-strasbourg.fr

A STRATEGY FOR REASONING WITH CLOSE KNOWLEDGE BASES: APPLICATION TO KNOWLEDGE BASES ON SMES[1]

Abstract: One of the major difficulties encountered when building a real system for handling heterogeneous knowledge bases (KB) with close topics is to combine them without sacrificing the specificity of their contents. The resolution of this difficulty raises some issues about combining multiple ontologies. The largest part of these issues is directly related to the existing mismatches among KBs with close contents. This paper describes a system that uses a compound strategy for minimizing the number of cases of mismatch in the aggregation of close and heterogeneous knowledge. This work is conducted in the context of the MAEOS project about the modeling of the support to the organizational and strategic development of small and medium enterprises (SMEs).

Keywords: knowledge bases, ontologies, multi-agent systems, integration of ontologies.

1. Introduction

Building a real system for handling and reasoning about many knowledge bases (KB) is a complex work. To reach a fairly interesting system, the designer must avoid all the pitfalls associated with mismatch among the different KBs that he wants to use. In the context of this work, the desired system is designed to use knowledge whose scope is restricted to a particular topic – the organization of small and medium enterprises (SMEs).

It follows that the knowledge bases used by the system have specific characteristics: some subset of their concepts may be common, distinct or possibly incompatible. The main issue that appears is how to use all this knowledge together and how to permit easy visualization of this knowledge coming from potentially contradictory sources, in order to allow analysis and diagnosis of SMEs.

[1] The authors would like to thank Mr. Henri-Pierre Michaud from AEM Conseil for his support and funding.

To manipulate and reason about this knowledge in a decidable way, each knowledge base is built using an ontology defined with the description logic formalism SHOIN (D) (OWL-DL, 2004). This choice allows decidable reasoning and the use of the verification capabilities of the language.

This article is structured as follows: section 2 presents the general context of this project and the following sections present our choices about handling and reasoning on heterogeneous and close KBs (section 3). The several issues regarding the proposed strategy are discussed in section 4. Afterward, the two key points of the proposed strategy are stated (section 5). Finally, we briefly present some results and our conclusions in sections 6 and 7.

2. General context

MAEOS is a project about the modeling of the support to the organizational and strategic development of SMEs [Renaud et al. 2009]. The main objective of MAEOS is to support the improvement of the efficiency and performance of business advice to SMEs. To achieve this objective, a multidisciplinary team was created. Three main research areas are represented: artificial intelligence, software engineering and management sciences. This work aims at establishing a set of methods and software tools for SMEs analysis and diagnosis. The software tools have to be able to evolve according to the state of the art about SMEs and, in particular, their administrative or legal environments. In addition, they must also be able to reflect the richness and contradictions that are inherent to the models coming from management sciences. Finally, they must be able to increase the competences of the consultants by the use of the accurate knowledge. Although closely related, these constraints can be considered according to three aspects: the manipulation of such knowledge, the use of heterogeneous close knowledge and the transfer of that knowledge.

3. Knowledge handling

In this project, knowledge bases are designed to cover a significant portion of aspects relating to organization, production and managerial behaviors of SMEs. The targeted knowledge is separated into two kinds of expertise. On the one hand, the theoretical knowledge in the area of change in SMEs (organizational, strategic…) is used as core models, and on the other hand, expert knowledge accumulated during practice is used as complementary knowledge. Knowledge bases are designed to cover a significant portion of aspects relating to organization, production and managerial behaviors of SMEs.

Due to the context of management sciences, the pieces of knowledge handled in this project may be close, complementary and sometimes contradictory. At this point, three approaches can be used: using each KB separately, merging KBs into a homogeneous one or using a mixed approach.

Unlike current trends, which are to represent knowledge in a homogeneous knowledge base covering the domain of a problem, our choice is different. It is to keep to a maximum the plurality of each knowledge base with their fields of interest, constraints and richness; the choice was made to use a mixed approach, motivated by the drawbacks of the two other approaches. On the one hand, the use of each KB separately, implying a separated knowledge processing, provokes that the results are divided into many isolated pieces. On the other hand, the merging of KBs makes the knowledge homogeneous, erasing, in the way, the specificities of each KB.

To introduce the notion of multiple points of views, the operation process that has been chosen is similar to that of a panel of experts. Each expert has an area of knowledge and a set of skills. He examines different aspects of the business related to his area of expertise. Once the study is completed, his conclusions are shared with other experts. Finally, a report is created.

Implementation issues. In this work, the panel of experts is implemented with a Multi-Agent System (MAS) with a blackboard architecture. This system consists in several knowledge bases on areas relating to business management, which are attached to software agents with the ability to exploit their content. The agents use three kinds of knowledge bases (KB) supported by ontologies. The first one is a rule-based KB (RBR – rule-based reasoning). The second one contains cases (CBR – case-based reasoning). And the third one consists in a set mapping rules among the different used ontologies.

Each agent is associated with a particular knowledge base. Therefore, all agents are characterized by a knowledge domain, a collection of facts and/or rules and a set of meta-data. At the beginning, a set of facts about an organization is entered in the blackboard. Each agent picks information up in it. It accomplishes its deduction or mapping tasks. At the end, it adds the results to the blackboard. The triggering of an agent is made by a set of data corresponding to the characteristics of its knowledge base. The process is considered as finished when the agents have nothing new to add to the blackboard. This situation implies the production of many pieces of results related to a limited topic. To be well understood by an external user and to provide concise results that are close to the context of the subject of study, all the produced results must be aggregated by topic. The aggregation of results raises some issues about combining multiple ontologies.

4. Aggregation of results

The aggregation of results based on ontologies requires more than a mere correspondence between terms or parts of models. This is because the production of results is carried out with several facts bases and/or rules and because each of these bases is related to the specific ontology. Each ontology that is used contains its own taxonomy, roles and axioms and is built with an intention and a point of view. This results aggregation must ensure a coherent semantics. Therefore, it is, at best, the integration of several ontologies into the new one covering all the results. Finally, once aggregated, the results must also be consistent with the facts and rules implemented in the knowledge base.

Combination of ontologies. There are many tools and works on ontologies combining [Klein 2001; Choi, Song, Han 2006; De Bruijn et al. 2006; Flouris et al. 2007]. Four classes of methods are applicable to MAEOS: Merging, Mapping, Aligning and Mediating. All these methods cannot always be applied in a systematic and/or automated way. As highlighted by Noy and Musen [1999], the intervention of an expert may be required.

The combining of several ontologies implies, at least, the presence of common or related conceptual entities in them. Different criteria can be applied to identify the similarities between two conceptual entities [Maedche et al. 2002]: the similarity of terms; the similarity of properties; or the similarity of the entities subsuming or being subsumed. Some methods, such as the alignment is better suited to a merging where different ontologies are complementary or have different semantic levels. It is necessary to know the criteria for selecting a combining method as well as the limitations of these methods.

In real situations, several penalizing cases may appear at different levels. Disparities in the definitions may not only arise at the conceptual, terminological or taxonomy level but also at the syntactic level. Between two close ontologies, it is common to have the same term with different meanings or several terms referencing the same concept. Depending on the ontology author's point of view, several definitions may relate to the same concept. Mismatches among ontologies are numerous. They are summarized in [Klein 2001; Visser et al. 1997; Hameed, Preece, Sleeman 2004] with a series of examples.

These differences affect the implementation of the combining methods. The most extreme case happens when disjoint ontologies are considered and makes impossible the application of any combining method. In the case of close ontologies, a choice cannot be made if the degree of similarity among several terms is equivalent [Colomb, Ahmad 2007].

Next, even if connections are established among conceptual entities, there is no guarantee that they will be bijections. Conflicts at the semantic level may also appear. Finally, the difference of granularity between ontologies can result in the elimination or aggregation of some entities. It should be noted that the number of mismatch cases increases when ontologies are bigger. Different ways should be studied in order to minimize these mismatches.

5. The aggregation strategy

For this project, a compound strategy has been studied to reduce the incidence of mismatch in ontologies combining: selecting sub-parts of ontologies, selecting the combining method, combining ontologies and aggregating results. This part of the project is described in [Renaud, Zanni-Merk, Rousselot 2009].

The next sections present the two key points of the proposed strategy: the method selection and the selection of sub-parts of ontologies.

5.1. Method selection

In general, choosing a combining method for ontologies is a critical issue. It becomes even more problematic if these combinations have to be performed in an automated way. The use of multiple ontologies may be reduced to alignment of parts or a full ontology merging. It becomes important to determine the criteria to choose among several methods.

```
If (mapping exists)
   Then Use Mapping
   Else If (overlapping exists)
        Then Use Merging
        Else If (alignment if applicable)
             Then Use Alignment
             Else Do Nothing
```

Figure 1. Algorithm for method selection

The criteria based on the works of Flouris et al. [2007] and Colomb and Ahmad [2007] allow the following cases (Figure 1):

- merging is used when the implemented ontologies are complementary; in other words, ontologies have common parts and distinct ones;
- alignment can be achieved when ontologies are close and do not correspond to the merging situation;
- mapping is used in preparation of merging among multiple ontologies; although this technique is very reliable, it requires some previous work.

Finally, in the context of this research work, if no choice can be made, ontologies are not combined.

5.2. Using sub-parts of an ontology

It is not always possible to be in the best situation for combining ontologies. The size of the ontologies has an important influence on the possibilities of combining: big ontologies are more complex to combine. Indeed, the cases of mismatch are much more frequent if ontologies are important. Use of small complementary ontologies can facilitate the construction of a more comprehensive one. At best, a solution can be to use modular ontologies or at least, that are possible to be split. To facilitate the combining of the ontologies, the adopted strategy consists in selecting only the necessary concepts for the aggregation of the results.

The decomposition of ontologies in sub-ontologies seems to be an attractive possibility. However, it supposes several assumptions:

– there are ontologies that are modular or decomposable into partitions,
– there exist coherent sub-ontologies,
– the number of extracted concepts is sufficient for the combining of the sub-ontologies.

It is evident that these assumptions cannot apply to every combination of ontologies. The context defined by all the produced results is important. This context helps collecting close sub-parts of ontologies around a particular subject. To achieve this goal, the targeted approach is to extract, from a set of ontologies, the smallest consistent sub-ontologies with a maximum coverage of the concepts used by the results. This selection is driven by the facts that are present in the blackboard (section 3). For that purpose, an algorithm based on the properties of partitioning and modularity of

```
Step 1: Select all concepts used by the results. To maintain
        the consistency, steps two and three extend the
        selection to sibling concepts.
Step 2: Complete the previous selection with the "is-a"
        relation tree.
Step 3: Add the shortest mandatory paths between the selected
        concepts.
Main Loop (from Step 4 to Step 11): Select complementary
        concepts in relation with the current selection.
Secondary Loop: Enumerates and chooses the necessary
        relations and neighbour concepts.
The algorithm stops when no concept or no relationship
can be selected.
```

Figure 2. Algorithm for sub-ontology selection

graphs is used. The algorithm considers an ontology as a semantic network. It treats the network as a directed graph that has nodes as concepts and edges as roles with their properties and their constraints. The search for a partition in the graph of a semantic network respects certain criteria. These criteria aim at selecting the concepts intervening in the interpretation of the results and at only preserving a coherent sub-graph. They are expressed as follows:

- a sub-graph must contain all the necessary concepts to link every part of the result,
- a sub-graph can be extracted if and only if it is connected to the rest of the graph by incident edges,
- a sub-graph must preserve the hierarchy formed by the “is_a” relations,
- each node must keep its concept definition.

These facts lead to the algorithm in Figure 2.

Yet, this algorithm has some limitations. On the one hand, the extracted sub-ontology can represent all the ontology in particular cases:

- if the graph is connected or strongly connected; the high number of edges among nodes requires the extraction of a bigger sub-graph;
- if the selected concepts belong to a clique located at the bottom of the “is_a” relations tree;
- if the useful concepts are distributed in a too homogeneous way in the graph; the paths making possible to go from a selected node to another are then more important.

This strategy permits the building of a local ontology to the case being considered.

6. Experiments

We have tested our approach with different cases of organization and different requests. At present, we are focused on applications of analysis of SMEs. These requests are intended to partly reformulate and extend a profile of the organization introduced in the MAS. To facilitate the acquisition and visualization of the results, we have developed the front-end DISKO (Figure 3); it provides a user friendly interface.

The following illustration shows a use case of the MAS: a user seizes a SME profile using DISKO. The profile is about a manufacturing SME, so he uses a subset of an organization ontology and a subset of a production ontology.

The MAS contains two knowledge bases using rules, each KB being implemented by an agent. Rules are implemented with JESS [Hill 1993].

The analysis proceeds as follows: Mapping of the global profile is completed by creating cross instances between the two ontologies. The two ontology agents use the rules to improve that original set of instances. Post-processing is used to aggregate the results produced by both agents.

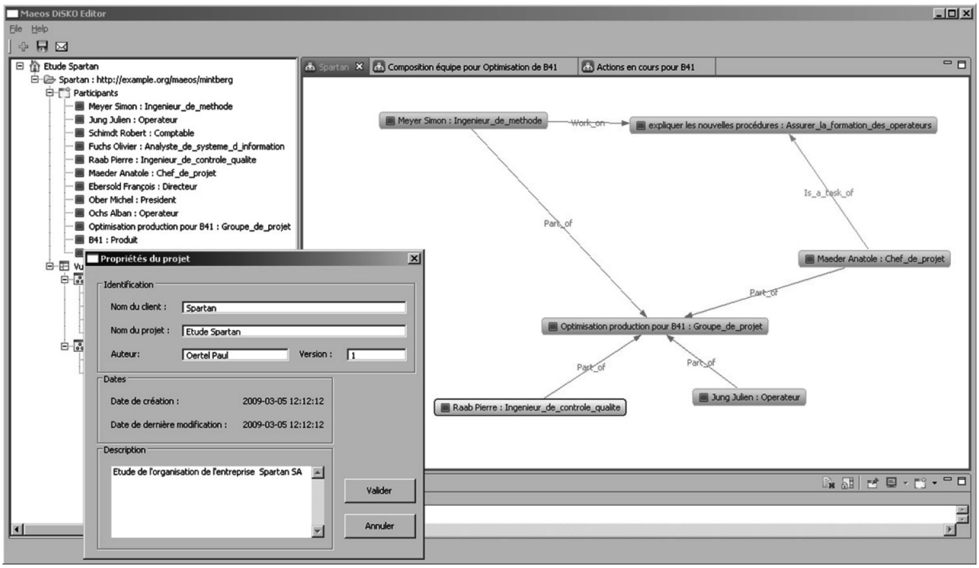

Figure 3. DISKO: the front-end to the MAS

In our experiments, we have confirmed, as expected, that the results coming from common subsets of both ontologies assemble and show additional relationships among the instances. The system also produces isolated sets of results related to the specific parts of each ontology. We have found that the weak number of performed mapping might be the main cause of these isolated sets of results. We have therefore increased the number of mappings between the ontologies. But the results still contain isolated sets of instances. We have remarked that the presence of these sets is mainly due to two factors:

– the fragmentation of the input profiles,
– the presence of specific parts that are weakly linked in an ontology.

These two points are currently under study to try and evaluate how to minimize the presence of these sets of isolated instances in our results.

7. Conclusions

We presented an approach for the aggregation of results coming from multiple ontologies. This approach aims at solving many limitations resulting from the use of ontologies whose contents are closely related. The suggested strategy is articulated around two key points: the choice of the combining method and the partitioning of ontologies.

Our experiments have shown the interest of the approach by sub-ontologies. However, the applied strategies are only efficient on close ontologies and that are slightly connected or modular.

References

Choi N., Song I.-Y., Han H., 2006, A survey on ontology mapping, *SIGMOD Record*, Vol. 35, No. 3, pp. 34-41.

Colomb R.M., Ahmad M.N. (2007), Merging ontologies requires interlocking institutional worlds, *Applied Ontology*, Vol. 2, No. 1.

De Bruijn J., Ehrig M., Feier C., Martíns-Recuerda F., Scharffe F., Weiten M. (2006), Ontology mediation, merging, and aligning, [in:] *Semantic Web Technologies*, Eds. J. Davies, R. Studer, P. Warren, John Wiley & Sons.

Flouris G., Manakanat D., Kondylakis H., Plexousakis D., Antoniou G. (2007), Ontology change: classification and survey, [in:] *The Knowledge Engineering Review*, Vol. 2, Cambridge University Press, pp. 117-152.

Hameed A., Preece A., Sleeman D. (2004), Ontology reconciliation, [in:] *Handbook on Ontologies*, Eds. S. Steffen, R. Studer, Springer, Berlin.

Hill E.F. (2003), *JESS in Action: Java Rule-Based Systems*, Manning Publications, Greenwich, USA.

Klein M. (2001), Combining and relating ontologies: An analysis of problems and solutions, [in:] *IJCAI'01, Workshop on Ontologies and Information Sharing*, Eds. G.A. Perez, M. Gruninger, H. Stuckenschmidt, M. Uschold, Seattle, WA, pp. 53-62.

Maedche A., Motik B., Silva N., Volz R. (2002), MAFRA: A mapping framework for distributed ontologies, [in:] *EKAW'2002, Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management*, Lecture Notes in Computer Science, vol. 2473, Springer Verlag.

Noy N., Musen M.A. (1999), SMART: Automated support for ontology merging and alignment, [in:] *KAW'1999, Proceedings of the Workshop on Knowledge Acquisition, Modeling and Management*, Eds. B.R Gaines., B. Kremer, M.A. Musen Banff, Alberta, Canada.

OWL-DL, 2004. http://www.w3.org/TR/owl-guide/.

Renaud D., Bouché P., Gartiser N., Zanni-Merk C., Michaud H.-P. (2009), *Knowledge Transfer for Supporting the Organizational Evolution of SMEs A Case Study*, International Conference on Innovation through Knowledge Transfer 2009 (InnovationKT), Hampton Court, UK.

Renaud D., Zanni-Merk C., Rousselot F. (2009), A compound strategy for ontologies combining, [in:] *KEOD 2009. Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, INSTICC Press, Funchal, Madeira, Portugal.

Visser P.R.S., Jones D.M., Bench-Capon T.J.M., Shave M.J.R. (1997), An analysis of ontology mismatches: Heterogeneity versus interoperability, [in:] *AAAI 1997 Spring Symposium on Ontological Engineering*, Stanford University Press, Stanford, CA.

STRATEGIA WNIOSKOWANIA W HETEROGENICZNYCH BAZACH WIEDZY: ZASTOSOWANIE W BAZACH WIEDZY MAŁYCH I ŚREDNICH PRZEDSIĘBIORSTW

Streszczenie: w artykule przedstawiono próbę rozwiązania problemu zarządzania heterogeniczną bazą wiedzy przy zachowaniu jej zawartości. Problem ten występuje w sytuacji, gdy mamy zamiar połączyć kilka ontologii o zbliżonej tematyce i różnej szczegółowości. W artykule opisano system korzystający ze strategii pozwalającej na minimalizację konfliktów w agregacji zbliżonych i heterogenicznych pojęć. Prace te wykonano w ramach projektu MAEOS wspomagającego tworzenie i modelowanie baz wiedzy dla małych i średnich przedsiębiorstw.