Scientific Report: ERCIM Fellowship November 2005 – July 2006 (Second Period)

Centre de Recherche Public Henri Tudor, Centre for IT Innovation

Monika Lanzenberger

Vienna University of Technology,
A-1040 Vienna, Austria
lanzenberger@ifs.tuwien.ac.at

1 Results and Ongoing Work

In continuation of my research work during the first ERCIM fellowship period, visual ontology alignment was the main topic of the second period as well. Furthermore, I published some results from other areas of interest. In the following I give a summary of visual ontology alignment, some conclusions, and an outlook to future work.

Within the next decade many people hope to see the data removed from the World Wide Web. Instead they want to populate it with knowledge. Such paradigm shift emphasizes the need for knowledge representation methods like the introduction of ontologies to the semantic web portfolio. However, in order to make data accessible for tool-supported manipulation, integration, and usage, which are significant steps in humans' effort of gaining knowledge, interoperability among different knowledge representations and knowledge domains is essential. Beyond technical aspects we need to focus on semantic interoperability issues. To that end, ontology alignment allows for relating metadata and bringing ontologies into mutual agreement.

Along our framework for ontology alignment quality we have identified a number of important aspects for assessing alignment results. Quality of alignment results is described in terms of physical, syntactic, semantic, pragmatic, perceived semantic, social, and taxonomic properties. Quality alignment can be partly supported by tools but human evaluation is still required. Current alignment tools usually create long lists of pairs of entities which are difficult to interpret in detail. Moreover, it seems impossible to assess the alignment results in a comprehensive way considering both ontologies as one knowledge resource. With AlViz, our alignment visualization tool, we help the user to explore ontology alignment results in both ways, in detail and with the overall consequences. Small world graphs within a multiple-view system represent both source ontologies on different levels of detail applying a cluster algorithm. The color of the nodes indicate whether equal or similar entities were found in both ontologies. When relating two ontologies we might have following questions in mind: Where

do most of the mappings between ontologies occur? (Location) Do the mapping choices directly or indirectly affect parts of the ontology we are concerned about? (Impact) What kinds of alignments occur between the ontologies? (Type) How different is the generated ontology (which combines parts of both source ontologies and is created by the alignment) to the individual source ontologies? (Extent) Finally, we are interested in the degree of additional knowledge derived when aligning two ontologies. (Benefit) Assessing and improving alignment results means to integrate new knowledge in a careful and goal-oriented way. Appropriate ontology alignment visualization offers practical means to support humans in achieving such challenging tasks. With AlViz we hope to offer a major step forward. However, some issues need further investigation. For example, if combining several object properties within one graph visualization, we want to categorize object properties by using specific weights. Weights of edges determine the distance between entities and influence the clustering results. Such categories are for instance transitive, non-transitive, or functional. In order to thoroughly understand the consequences of introducing such weight categories, I need to evaluation different types of ontologies. Another ongoing work is the search for automatic support of the alignment assessment exploiting approaches of graph theory. Finally, I need to do detailed user testing of AlViz.

2 Participation in Conferences

Interoperability for Enterprise Software and Applications Conference, March 22 - 24, 2006, Bordeaux, France.

CAISE '06, Conference on Advanced Information Systems Engineering, June 5 - 9, 2006, Luxembourg.

Information Visualisation Conference, July 5 - 7, 2006, London, UK.

3 Other Contributions to the Scientific Community

Scientific reviewer :

IEEE Information Visualisation, London, UK, July 5 - 7, 2006

IEEE Visualization 2006, Baltimore, Maryland, USA, October 31 - November 3, 2006

IEEE Symposium on Visual Analytics Science and Technology, Baltimore, Maryland, USA, October 31 - November 2, 2006

Publication chair:

ARES, International Conference on Availability, Reliability, and Security Bridging Theory and Practice, Vienna, Austria, April 20 - 22, 2006

Program committee member and scientific reviewer:

ARES, International Conference on Availability, Reliability, and Security Bridging Theory and Practice, Vienna, Austria, April 20 - 22, 2006

References

Accepted Publications, November 2005 - July 2006

1. Lanzenberger, M.; Sampson, J.: **AlViz - A Tool for Visual Ontology Alignment** In: *Proceedings of the 10th Information Visualisation Conference*, July 5 - 7, 2006, London, UK, forthcoming.

   Abstract:
   We introduce a multiple-view tool called AlViz, which supports the alignment of ontologies visually. Ontologies play an important role for interoperability between organizations and for the semantic web because they aim at capturing domain knowledge in a generic way and provide a consensual understanding of a domain. Alignment is the process where for each entity in one ontology we try to find a corresponding entity in the second ontology with the same or the closest meaning. Existing ontology alignment tools do not adequately provide a way for users to analyse the results. While many alignment tools generate lists of mappings it is difficult to analyse these alignments without examining every pairwise correspondence in the output files and even then it is an overwhelming task. We propose the use of visualization techniques to facilitate user understanding of the ontology alignment results. AlViz is implemented as a tab plug-in for Protégé.

2. Sampson, J.; Veres, C.; Lanzenberger, M.: **An Integrated Approach for Organizational Data Interoperability** In: *Proceedings of the Interoperability for Enterprise Software and Applications Conference*, March 22 - 24, 2006, Bordeaux, France, forthcoming.

   Abstract:
   Organizational interoperability is facilitated by mapping legacy database schemas to ontology models and using techniques for the alignment of ontologies. Refactoring of database schemas is important in this context because they repre- sent complex agreements established between the designers, users and stakeholders within an organization. We describe an integrated set of tools and techniques for sup- porting the complete life cycle including validation of legacy schemas, transforming them into ontologies, and ensuring quality ontology alignment.

3. Sampson, J.; Lanzenberger, M.: **Visual Ontology Alignment for Semantic Web Applications** In: *Proceedings of the ER 2006 Workshops, 25th International Conferences on Conceptual Modeling*, November 6 - 9, 2006, Tucson, Arizona, Springer, forthcoming.

   Abstract:
   Ontologies play an important role for the semantic web because they aim at capturing domain knowledge in a generic way and provide a consensual understanding

of a domain. Due to the number of ontologies available the need for mapping or bringing them into alignment has prompted a surge of tools and algorithms to be developed for this purpose. We describe some of these ontology alignment tools and discuss issues and shortcomings in current state of the art. From this analysis we propose the use of visualization techniques to facilitate user understanding of the ontology alignment results. Finally we briefly describe AlViz, our visual ontology alignment tool.

4. Pohl, M.; Lanzenberger, M.: **Participation of Female Computer Science Students in Austria**, In: Eileen M. Trauth (ed): *Encyclopedia of Gender and Information Technology*, Idea Group, ISBN: 1- 59140-815-6, 2006, pp. 970-975.

   Abstract:
   The under-representation of women in Computer Science seems to be quite a complex phenomenon. Many different reasons have been offered to explain the lack of participation of women in computing but we still do not know enough about the relative importance of these reasons. In Austria, some of the explanations given for the under-representation of women in Computer Science do not hold anymore, as, for example, lack of access to computer technology. Our own questionnaire and the micro census data both indicate that there is a tendency to more equal access among younger age groups. The future will show whether these changes will lead to a higher percentage of female students of Computer Science in the long run. Another problem are the Bachelor Studies with a high percentage of female students. Women apparently feel attracted by Media Informatics and Medical Computer Science and deterred by Computer Engineering and Information & Software Engineering. This conforms to general stereotypes. From a feminist point of view the introduction of subjects like Medical Computer Science is contradictory because on the one hand it attracts women to Computer Science but on the other hand it reinforces gender differences. The higher demand for (female) IT specialists is also difficult to interpret. It is highly probable that this demand will fluctuate in the future in accordance with the development of the economy. As a consequence the proportion of female IT specialists will probably fluctuate as well.

Submitted Publications, November 2005 - July 2006

(Currently with the reviewers / editors)

5. Sampson, J., Krogstie, J. and Lanzenberger, M.: **Facilitating Understanding and Agreement in Ontology Alignment**, Submitted to: *ISWC 2006, 5th International Semantic Web Conference.*

   Abstract:
   We describe how to improve the understanding and agreement of ontology alignment results. First we briefly present a comprehensive framework for understanding quality in ontology alignment. The framework was developed as a mechanism for analyzing current state of the art alignment results. We found that while performance results are relatively high for current tools, understanding and interpretation of the results by ontology engineers is low. Through our findings we recommend a number of techniques, tools and methods for facilitating understanding and agreement of the alignments between ontology engineers. We briefly describe our

development towards such a tool AlViz, using a real world ontology alignment example.

6. Lanzenberger, M., Sampson, J., Rester, M., Naudet, Y., Latour, T. : **Visual ontology alignment for semantic interoperability**, Submitted to: *Journal of Knowledge Management, Special Issue on Competencies Management: Integrating Semantic Web and Technology Enhanced Learning approaches for effective Knowledge Management* .

Abstract:
Ontology alignment plays an important role in the context of semantic interoperability . Usually ontology alignment tools generate results that are difficult to understand or assess. In order to enable users to check and improve alignment results and to understand their consequences we use information visualization techniques. Moreover, we discuss the relevant quality aspects in ontology alignment as well as current activities and available tools. Based on literature study we identify quality measures for ontology alignment and define requirements for visual ontology alignment. As a proof of concepts we developed a prototype called AlViz. Information visualization offers appropriate methods for the assessment of ontology alignment results. Different levels of detail and overview help the user to navigate and understand the alignments. Along a comprehensive framework we identify alignment assessment tasks and we introduce and apply a visualization tool which aims at making ontology alignment results manageable and comprehensible.