

Linköping Studies in Science and Technology

Thesis No. 1225

Aligning and Merging Biomedical Ontologies

by

He Tan



Submitted to Linköping Institute of Technology at Linköping University in partial fulfilment of the requirements for the degree of Licentiate of Engineering

Department of Computer and Information Science
Linköpings universitet
SE-581 83 Linköping, Sweden

Linköping 2006

Aligning and Merging Biomedical Ontologies

by

He Tan

January 2006

ISBN 91-85497-01-0

Linköping Studies in Science and Technology

Thesis No. 1225

ISSN 0280-7971

LiU-Tek-Lic-2006:6

ABSTRACT

Due to the explosion of the amount of biomedical data, knowledge and tools that are often publicly available over the Web, a number of difficulties are experienced by biomedical researchers. For instance, it is difficult to find, retrieve and integrate information that is relevant to their research tasks. Ontologies and the vision of a Semantic Web for life sciences alleviate these difficulties. In recent years many biomedical ontologies have been developed and many of these ontologies contain overlapping information. To be able to use multiple ontologies they have to be aligned or merged. A number of systems have been developed for aligning and merging ontologies and various alignment strategies are used in these systems. However, there are no general methods to support building such tools, and there exist very few evaluations of these strategies. In this thesis we give an overview of the existing systems. We propose a general framework for aligning and merging ontologies. Most existing systems can be seen as instantiations of this framework. Further, we develop SAMBO (System for Aligning and Merging Biomedical Ontologies) according to this framework. We implement different alignment strategies and their combinations, and evaluate them in terms of quality and processing time within SAMBO. We also compare SAMBO with two other systems. The work in this thesis is a first step towards a general framework that can be used for comparative evaluations of alignment strategies and their combinations.

This work has been supported by CUGS (the Swedish national graduate school in computer science), CENIIT (the Center for Industrial Information Technology) and the EU Network of Excellence REWERSE (Sixth Framework Programme project 506779).

Abstract

Due to the explosion of the amount of biomedical data, knowledge and tools that are often publicly available over the Web, a number of difficulties are experienced by biomedical researchers. For instance, it is difficult to find, retrieve and integrate information that is relevant to their research tasks. Ontologies and the vision of a Semantic Web for life sciences alleviate these difficulties. In recent years many biomedical ontologies have been developed and many of these ontologies contain overlapping information. To be able to use multiple ontologies they have to be aligned or merged. A number of systems have been developed for aligning and merging ontologies and various alignment strategies are used in these systems. However, there are no general methods to support building such tools, and there exist very few evaluations of these strategies. In this thesis we give an overview of the existing systems. We propose a general framework for aligning and merging ontologies. Most existing systems can be seen as instantiations of this framework. Further, we develop SAMBO (System for Aligning and Merging Biomedical Ontologies) according to this framework. We implement different alignment strategies and their combinations, and evaluate them in terms of quality and processing time within SAMBO. We also compare SAMBO with two other systems. The work in this thesis is a first step towards a general framework that can be used for comparative evaluations of alignment strategies and their combinations.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Associate Professor Patrick Lambrix, who has introduced me to the research. Thank you for your invaluable guidance, discussions and comments during the work. I appreciate your support and encouragement. Your steady, reliable and attentive presence provides such comfortable work environment. I would like to express my appreciations to Professor Nahid Shahmehri for providing useful comments and pointing out important aspects of the research world.

I am thankful to my colleagues at the Laboratory for Intelligent Information Systems (IISLAB) and at the Division for Database and Information Techniques (ADIT).

Thanks to my friends. I appreciate the time we share challenges and celebrate joys of lifes. Thanks to Gunilla Sunnerud for the warmth of family you give me in the foreign land. Last but not least, I feel grateful to my parents for your understanding, support and endless love.

We acknowledge the financial support from the Center for Industrial Information Technology, the Swedish national graduate school in computer science, and of the EU Network of Excellence REWERSE (Sixth Framework Programme project 506779).

*Tan, He
Linköping, Sweden
December, 2005*

List of Enclosed Papers

This thesis contains the following papers:

1. P. Lambrix, H. Tan, SAMBO - A System for Aligning and Merging Biomedical Ontologies, submitted to *Journal of Web Semantics, special issue on semantic web for the life sciences*.
2. H. Tan, V. Jakonienė, P. Lambrix, J. Aberg, N. Shahmehri, Alignment of biomedical ontologies using life science literature, submitted to *the International Workshop on Knowledge Discovery in Life Science Literature*.

List of Additional Papers

In addition to these papers, the research conducted in this thesis has resulted in the following related publications that are not included in the thesis. The contributions from the following publications are covered by paper 1.

3. P. Lambrix, H. Tan, A Framework for Aligning Ontologies, *Proceedings of the Third International Workshop on Principles and Practice of Semantic Web Reasoning*, LNCS 3703, Springer, 2005, pp. 17-31.
4. P. Lambrix, H. Tan, Merging DAML+OIL Ontologies, *Barzdins and Caplinskas, eds., Databases and Information Systems - Selected Papers from the Sixth International Baltic Conference on Databases and Information Systems*, IOS Press, 2005, pp. 249-258.
5. P. Lambrix, A. Edberg, C. Manis, H. Tan, Merging DAML+OIL bio-ontologies, *Proceedings of the International Workshop on Description Logics*, Rome, Italy, 2003.

Table of Contents

Introduction	1
Motivation	3
Problem Statement	4
Contributions	6
Paper Summaries	7
Related Work	8
Conclusion and Future Work	9
References	11
Paper 1	15
Paper 2	47

Introduction

1 Motivation

Nowadays the sequences of genomes for many organisms from microbes to plants to humans have been completed [CMP03, FJTO03]. In the genomics research efforts in the last decade, massive quantities of data and knowledge have been generated, and an enormous number of tools for data analysis, informatics tools and other related techniques have been developed. In the post-genomic era these data, knowledge and tools are used by researchers from different areas, like biological research, clinical medicine, health, environment and agriculture, to find solutions for the challenges across the life sciences, such as, detecting systemic functional behaviors of the cell and the organism, and developing genomics-based approaches in the practice of health care [CGGG03]. To support the work of researchers techniques from areas such as data source integration, information retrieval, knowledge discovery, and decision support systems are needed [KB03, MK03]. Due to the explosion of the amount of biological data and tools, often publicly available over the Web, a number of difficulties are experienced in these areas. For instance, as discussed in [Lam05], it is difficult to locate the relevant data sources and tools which are being continuously updated, retrieve the relevant information based on the current syntax based approaches, and integrate the heterogeneous data sources. Ontologies and the vision of a Semantic Web for life sciences alleviate these difficulties [REW]. The Semantic Web can be seen as an extension of the current Web in which information is given a well defined meaning by annotating Web content with ontology terms.

Ontologies (e.g. [Gom99, Lam04]) can be seen as defining the basic concepts and relations of a domain of interest, as well as the rules for combining these concepts and relations. The benefits of the use of ontologies include reuse and sharing of knowledge across platforms, and improved maintainability, documentation, maintenance, and reliability, portability [JU99]. Ontologies are used for communication between people and organizations by providing a common terminology over a domain. They provide the basis for interoperability between systems. They can be used for making the content in information sources explicit and serve as an index to a repository of information. Further, they can be used as a basis for integration of information sources and as a query model for information sources. They also support clearly separating domain knowledge from application-based knowledge as well as validation of data sources. Overall, ontologies lead to a better understanding of a field and to more effective and efficient handling of information

in that field.

In biomedical research the importance of ontologies has been recognized. There is much international research cooperation for the development of biomedical ontologies, e.g. Gene Ontology (GO) [GO], Open Biomedical Ontologies (OBO) [OBO], and Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) [SNOMED]. Also there are many efforts to advance the practical use of biomedical ontologies and ontology-based tools. For instance, the National Center for Biomedical Ontology [BO] was founded in October 2005, and will focus on the use of biomedical ontologies to facilitate access, review and integration of disparate knowledge resources in all aspects of biomedical investigation and clinical practice.

Although the technology is not yet mature [BMM03], a vast amount of biomedical ontologies have been created. Many of these ontologies contain overlapping information. Therefore applications often have to deal with multiple ontologies. For example, applications using their own but similar ontologies need cooperation. Application developers often link custom-specific and smaller ontologies to the standard ontologies. Applications need to use ontologies from different areas or from different views on one area. Ontology builders may use already existing ontologies as the basis for the creation of new ontologies by extending the existing ontologies or by combining knowledge from different smaller ontologies. In each of these cases, the relationships between the different ontologies need to be known. It has been accepted that defining the relationships between the different ontologies is one of the main issues in making biomedical ontologies work [NRM04], and some organizations have started to deal with it. For instance, the organization for Standards and Ontologies for Functional Genomics (SOFG) [SOFG] developed the SOFG Anatomy Entry List (SAEL) which defines cross species anatomical terms relevant to functional genomics and which can be used as an entry point to anatomical ontologies. The objective of our work involves developing tools to find these inter-ontology relationships.

2 Problem Statement

In this thesis we present our work on aligning and merging biomedical ontologies. We use the following definitions in our work.

Aligning Ontologies is to define the relationships between concepts or relations from two different source ontologies. The relationships can

be equivalence as well as *is-a*, *part-of* or any other kind of relation.

Merging Ontologies is to derive a new ontology which is a (minimal) union of two source ontologies based on the identified alignments between them.

A number of systems have been developed for aligning and merging ontologies. However, there are no general methods to support building such tools, and there exist no comparisons between these systems. Various approaches are used to discover the alignments between different ontologies in these systems and very few evaluations of these strategies exist. Also, how different strategies can be combined to produce better results, needs to be investigated. Therefore, there is a need to develop a framework that can be used to integrate different alignment strategies and evaluate them. Thus, in this work we have chosen to focus on the following problems.

1. Development of a general framework for aligning and merging ontologies. The framework can be used to integrate different alignment strategies and to evaluate the strategies and their combinations.
2. Development of an ontology aligning and merging system based on the framework.
3. Implementation and evaluation of different alignment strategies as well as their combinations using the developed system.

In the rest of this section we define the scope of our work. Ontologies can be classified into different types depending on which of the components (e.g. [SGB01]) are represented and the kind of information they can represent [Lam04]. A simple type of ontology are the *controlled vocabularies*, which are essentially lists of concepts that are enumerated explicitly. When the concepts are organized into a hierarchical structure, we obtain a *taxonomy*. The relations in a taxonomy can be *is-a* and *part-of*. *Thesauri* are more complex ontologies, in which there are also relations that connect concepts across the hierarchy. Most currently available biomedical ontologies are of these three types of ontologies. Other types of ontologies are *data models* and *knowledge bases*. In *data models* axioms which describe facts that are always true in topic area of the ontology, are used in addition to relations. *Knowledge bases* can represent the other types of ontologies and are often based

on a logic. In this thesis we focus on *controlled vocabularies, taxonomies and thesauri*.

The heterogeneity of ontologies may occur at different levels [Kle01]. On language level ontologies may have different representation formats, such as RDF(s) [RDF], F-Logic [KLW95] and OWL [OWL]. On terminology level, ontologies mismatch due to, for example, synonyms, homonyms and different encodings. On concept level we encounter the heterogeneity of ontologies due to different scope, granularity and perspective. On pragmatics level the discrepancy of ontologies is determined by the context related to the knowledge represented by the ontologies. The focus of the work in this thesis is to overcome the heterogeneity of ontologies on the terminology and concept levels. We also assume that the ontologies to be aligned and merged are represented in DAML+OIL [DAML] and OWL, and that these biomedical ontologies are in one context.

In the definition of *aligning ontologies* the relationships between concepts or relations from different ontologies can be equivalence, *is-a*, *part-of* or any other kind of relation, but in this work we only focus on identifying equivalence and *is-a* relations.

3 Contributions

The work presented in this thesis contributes to aligning and merging biomedical ontologies as follows.

1. As a first step in dealing with problem 1 stated in the previous section, we studied the existing systems for aligning ontologies and schema matching from literature and their public distributions. Some of the systems also support ontology merging. We identified the types of the various approaches used in these systems for the discovery of the alignments between the different ontologies, and investigated the procedure to perform aligning and merging ontologies in these systems. An overview of the existing ontology aligning and merging systems and the strategies they use is given in paper 1.
2. As the second step in coping with problem 1, we proposed, based on contribution 1, a general framework for aligning and merging ontologies. Most of the existing systems can be seen as an instantiations of our framework. The framework also provides support for experiment-

ing with different alignment strategies and their combinations. This contribution is found in paper 1.

3. Considering problem 2, we developed SAMBO (System for Aligning and Merging Biomedical Ontologies) according to the general framework. In SAMBO the alignment process is separated into two steps, aligning relations and aligning concepts. The user can perform the alignment manually or based on the suggestions. Suggestions are the alignment candidates proposed by the alignment strategies as well as their combinations. We implemented several different matchers that are representatives for the different types of alignment strategies in SAMBO. The strategies can be combined in different ways by giving different weights for each matcher. The user can set a threshold to filter the suggestions. After the user accomplishes the alignment process, the system can be asked to compute a merged ontology from the source ontologies and their identified alignments. Furthermore, SAMBO uses a DIG description logic reasoner (e.g. Racer [Racer] or FaCT [FaCT]) to provide a number of reasoning services. This contribution is discussed in paper 1.
4. Regarding problem 3, we evaluated different alignment strategies as well as their combinations within SAMBO. In the evaluation we focused on the quality of the suggestions. The lessons learned from the evaluation can provide recommendations for the improvement of the existing alignment strategies, the development of new alignment algorithms, and choosing the appropriate combinations. This is also the first step towards our ambition to develop a framework for comparative evaluations of different alignment strategies. This contribution is found in paper 1 and 2.

4 Paper Summaries

There are two papers enclosed in this thesis. In the following we give short summaries of the two papers.

Paper 1: SAMBO - A System for Aligning and Merging Biomedical Ontologies

In this paper we present a general framework for aligning and merging ontologies. In the framework the alignments between two source ontologies can

be determined under the supervision of the user based on the suggestions generated by various alignment strategies as well as their combinations. Further, we identify the types of alignment strategies used in the current alignment systems. We describe SAMBO which exemplifies the use of the framework. In SAMBO we implement several matchers that are representatives for the different types of alignment strategies, and also experiment with the different combinations of these matchers. We present the results of evaluations in which we compare the quality of suggestions generated by the different matchers and their combinations with different thresholds, and the amount of time they take to generate the suggestions. We also compare SAMBO with two other alignment systems with respect to the quality of suggestions. We use several well-known biomedical ontologies in the evaluations.

Paper 2: Alignment of biomedical ontologies using life science literature

In this paper we present one of our experiments on alignment strategies in SAMBO. The matchers investigated in this experiment utilize life science literature related to concepts from the source ontologies. The basic matcher uses the Bayesian learning method, and its extensions take the structure of the ontologies into account. We compare the basic matcher with its extensions, and with other alignment strategies implemented in SAMBO in terms of the quality of suggestions and the amount of time they take to generate the suggestions. We also present the results of evaluations on the different combinations of these matchers with other implemented alignment strategies.

5 Related Work

There exist many systems devoted to ontology alignment, such as PROMPT [NM03], Chimaera [MFRW00], ArtGen [MW02], FCA-Merge [SM01], OntoMapper [SYT02], GLUE [DMDH03], QOM [ES04] (FOAM), ASCO [LKG04], HCONE [KV04], IF-Map [YM03], iMapper [SHB04], and S-Match [GSY04]. Some of the systems also support merging ontologies. A number of different alignment strategies are used in these systems. An overview of the alignment strategies is given in paper 1.

Another related work is evaluations of systems for ontology alignment and merging. Up to date only three comparative evaluations of systems for ontology alignment and merging have been performed. The EU OntoWeb

project [OntoWeb] evaluated the systems iPROMPT [NM00] based on Protégé (with extension Anchor-PROMPT [NM01]), Chimaera (described, not evaluated), FCA-Merge and ODEMerge. This evaluation focused on such things as functionality, interoperability and visualization, but did not include tests on the quality of the alignment. In [LE03, LT05] PROMPT, Chimaera and a former version of SAMBO were evaluated in terms of the quality of the alignment as well as the time it takes to align ontologies with these tools. Further, an ontology alignment contest was held at EON-2004 [Euz04]. FOAM, OLA [ELTV04], SCM [HYNT04] and PROMPT participated. The main goal of the contest was to show how ontology alignment tools can be evaluated and a follow-up was planned for 2005. Other tools have not appeared in comparative evaluation studies.

6 Conclusion and Future Work

This thesis makes a contribution to an essential problem, aligning and merging ontologies, of the use of ontologies in the biomedical community. It proposes a general framework for aligning and merging ontologies, in which various different alignment strategies can be integrated. A system (SAMBO) is developed as an example of the use of the framework. A set of evaluations is performed on comparison between the alignment strategies and their combinations implemented in SAMBO, and those in other publicly available alignment systems. Although in this thesis we focus on the types of ontologies that currently are available in the biomedical domain, the work can be extended to ontologies in other domains.

In the future the work will continue towards different directions. Although a number of systems have been developed for aligning and merging ontologies, and different alignment strategies are implemented in these systems, there are few comparative evaluations of these systems and no tool is available to perform such evaluations. The main task of the future work is to develop a framework that can be used to integrate different alignment strategies and evaluate these techniques using different types of ontologies which can be from various domains. The results of the evaluations would be recommendations on which (combinations of) strategies are well suited for aligning which kinds of ontologies, and how to improve the (combinations of) alignment algorithms, or develop new alignment approaches.

Another track is to further develop SAMBO. We will improve the existing matchers, develop new alignment algorithms in SAMBO, and evaluate

them. We are working on integrating a highly interactive ontology visualization tool into SAMBO, which not only improves the user interface to visualize information, but also may lead to the development of new alignment algorithms.

The ontologies used in this work are currently available in the biomedical domain. Most of these ontologies are *controlled vocabularies*, *taxonomies* or *thesauri*. We will evaluate our current strategies on more complex ontologies. Also, we will investigate alignment strategies, e.g. constraint-based approaches, that take into account all kinds of information, such as axioms, that the complex ontologies provide.

References

- [BO] The National Center for Biomedical Ontology.
<http://www.bioontology.org/>.
- [BMM03] O. Bodenreider, J.A. Mitchell, A.T. McCray, Biomedical Ontologies, *Proceedings of 2003 the Pacific Symposium on Biocomputing*, 2003, 8:562-564.
- [CGGG03] F. Collins, E. Green, A. Guttmacher, M. Guyer, A vision for the future of genomics research, *Nature*, 2003, 422:835-847.
- [CMP03] F.S. Collins, M. Morgan, A. Patrinos, The Human Genome Project: Lessons from Large-Scale Biology, *Science*, 2003, 300(5617):286-290.
- [DAML] DAML+OIL. <http://www.w3.org/Submission/2001/12/>.
- [DMDH03] A. Doan, J. Madhavan, P. Domingos, A. Halevy, Ontology matching: A machine learning approach, *Staab and Studer, eds., Handbook on Ontologies in Information Systems*, Springer, 2003, 397-416.
- [ELTV04] J. Euzenat, D. Loup, M. Touzani, P. Valtchev, Ontology alignment with OLA, *Proceedings of the 3rd international Workshop on the Evaluation of Ontology-based Tools*, 2004.
- [ES04] M. Ehrig, S. Staab, QOM - quick ontology mapping, *Proceedings of 3rd International Semantic Web Conference*, LNCS 3298, Springer, 2004, 683-697.
- [Euz04] J. Euzenat, Introduction to the EON ontology alignment context, *Proceedings of the 3rd international Workshop on the Evaluation of Ontology-based Tools*, 2004.
- [FaCT] FaCT. <http://www.cs.man.ac.uk/~horrocks/FaCT/>.
- [FJTO03] M.E. Frazier, G.M. Johnson, D.G. Thomassen, C.E. Oliver, A. Patrinos, Realizing the Potential of the Genome Revolution: The Genomes to Life Program, *Science*, 2003, 300(5617):290-293.
- [GO] The Gene Ontology Consortium, Gene Ontology: tool for the unification of biology, *Nature Genetics*, 2000, 25(1):25-29.
<http://www.geneontology.org/>.

- [Gom99] A. Gómez-Pérez, Ontological Engineering: A state of the Art, *Expert Update*, 1999, 2(3):33-43.
- [GSY04] F. Giunchiglia, P. Shvaiko, M. Yatskevich, S-Match: an algorithm and an implementation of semantic matching, *Proceedings of the 1st European Semantic Web Symposium*, LNCS 3053, Springer, 2004, 61-75.
- [HYNT04] T. Hoshiai, Y. Yamane, D. Nakamura, H. Tsuda, A semantic category matching approach to ontology alignment, *Proceedings of the 3rd international Workshop on the Evaluation of Ontology-based Tools*, 2004.
- [JU99] R. Jasper, M. Uschold, A Framework for Understanding and Classifying Ontology Applications, *Proceedings of the 12th Workshop on Knowledge Acquisition, Modeling and Management*, Banff, Canada, 1999.
- [KB03] M. Kanehisa, P. Bork, Bioinformatics in the post-sequence era, *Nature Genetics*, 2003, 33:305-310.
- [Kle01] M. Klein, Combining and Relating Ontologies: An Analysis of Problems and Solutions, *Workshop on Ontologies and Information Sharing, IJCAI'01*, 2001.
- [KLW95] M. Kifer, G. Lausen, J. Wu, Logical Foundations of Object-Oriented and Frame-Based Languages, *Journal of the ACM*, 1995, (42):741-843.
- [KV04] K. Kotis, G.A. Vouros, The HCONE Approach to Ontology Merging, *Proceedings of the 1st European Semantic Web Symposium*, LNCS 3053, Springer, 2004, 137-151.
- [Lam04] P. Lambrix, Ontologies in Bioinformatics and Systems Biology, *Dubitzky and Azuaje, eds., Artificial Intelligence Methods and Tools for Systems Biology*, Springer, 2004, chapter 8, 129-146.
- [Lam05] P. Lambrix, Towards a Semantic Web for Bioinformatics using Ontology-based Annotation, *Proceedings of the 14th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises*, 2005, 3-7. Invited talk.
- [LE03] P. Lambrix, A. Edberg, Evaluation of ontology merging tools in bioinformatics, *Proceedings of the Pacific Symposium on Biocomputing*, 2003, 8:589-600.

- [LT05] P. Lambrix, H. Tan, Merging DAML+OIL Ontologies, *Barzdins and Caplinskas, eds., Databases and Information Systems - Selected Papers from the Sixth International Baltic Conference on Databases and Information Systems*, IOS Press, 2005, 249-258.
- [LKG04] B.T. Le, R.D. Kuntz, F. Gandon, On ontology matching problem (for building a corporate semantic web in a multi-communities organization), *Proceedings of the 6th International Conference on Enterprise Information Systems*, 2004.
- [MFRW00] D. McGuinness, R. Fikes, J. Rice, S. Wilder, An Environment for Merging and Testing Large Ontologies, *Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning*, 2000, 483-493.
- [MK03] V. Maojo, C.A. Kulikowski, Bioinformatics and Medical Informatics: Collaboration on the Road to Genomic Medicine? *Journal of the American Medical Informatics Association* 2003, 10(6):515-522.
- [MW02] P. Mitra, G. Wiederhold, Resolving terminological heterogeneity in ontologies, *Proceedings of the ECAI Workshop on Ontologies and Semantic Interoperability*, 2002.
- [NM00] N.F. Noy, M.A. Musen, PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment, *Proceedings of the Seventeenth National Conference on Artificial Intelligence*, 2000, 450-455.
- [NM01] N.F. Noy, M.A. Musen, Anchor-PROMPT: Using Non-Local Context for Semantic Matching, *Proceedings of the IJCAI Workshop on Ontologies and Information Sharing*, 2001, 63-70.
- [NM03] N.F. Noy, M.A. Musen, The PROMPT Suite: Interactive Tools For Ontology Merging And Mapping, *International Journal of Human-Computer Studies*, 2003, 59(6):983-1024.
- [NRM04] N.F. Noy, D.L. Rubin, M.A. Musen. Making Biomedical Ontologies and Ontology Repositories Work, *IEEE Intelligent Systems*, 2004, 19(6):78-81.
- [OBO] Open Biomedical Ontologies. <http://obo.sourceforge.net/>.

- [OntoWeb] OntoWeb Consortium, Deliverables 1.3 (A survey on ontology tools) and 1.4 (A survey on methodologies for developing, maintaining, evaluating and reengineering ontologies), 2002. <http://www.ontoweb.org>
- [OWL] OWL. <http://www.w3.org/2004/OWL/>.
- [Racer] Racer. <http://www.cs.concordia.ca/~haarslev/racer/>.
- [REW] REVERSE EU Network of Excellence on Reasoning on the Web with Rules and Semantics, Deliverables of the A2 Working Group. <http://www.reverse.net>.
- [RDF] RDF. <http://www.w3.org/RDF/>.
- [SGB01] R. Stevens, C. Goble, S. Bechhofer, Ontology-based knowledge representation for bioinformatics, *Briefings in Bioinformatics*, 2001, 1(4):398-414.
- [SHB04] X.M. Su, S. Hakkarainen, T. Brasethvik, Semantic enrichment for improving systems interoperability, *Proceedings of the 2004 ACM Symposium on Applied Computing*, 2004, 1634-1641.
- [SM01] G. Stumme, A. Mädche, FCA-Merge: Bottom-up merging of ontologies, *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, 2001, 225-230.
- [SNOMED] SNOMED-CT. <http://www.nhsia.nhs.uk/snomed/pages>.
- [SYT02] P. Sushama, P. Yun, F. Timothy, Using Explicit Information To Map Between Two Ontologies, *Proceedings of the AAMAS Workshop on Ontologies in Agent Systems*, 2002.
- [SOFG] SOFG, Standards and Ontologies for Functional Genomics. <http://www.sofg.org/>.
- [YM03] Y. Kalfoglou, M. Schorlemmer, IF-Map: an ontology mapping method based on information flow theory, *Journal of data semantics*, 2003, 1:98-127.



LINKÖPINGS UNIVERSITET

Avdelning, institution
Division, department

Institutionen för datavetenskap
Department of Computer
and Information Science

Datum
Date

2006-01-23

Språk

Language

- Svenska/Swedish
 Engelska/English

Rapporttyp

Report category

- Licentiatavhandling
 Examensarbete
 C-uppsats
 D-uppsats
 Övrig rapport

ISBN

91-85497-01-0

ISRN

LiU-Tek-Lic-2006:6

Serietitel och serienummer
Title of series, numbering

ISSN

0280-7971

Linköping Studies in Science and Technology

Thesis No. 1225

URL för elektronisk version

Titel
Title

Aligning and Merging Biomedical Ontologies

Författare
Author

He Tan

Sammanfattning

Abstract

Due to the explosion of the amount of biomedical data, knowledge and tools that are often publicly available over the Web, a number of difficulties are experienced by biomedical researchers. For instance, it is difficult to find, retrieve and integrate information that is relevant to their research tasks. Ontologies and the vision of a Semantic Web for life sciences alleviate these difficulties. In recent years many biomedical ontologies have been developed and many of these ontologies contain overlapping information. To be able to use multiple ontologies they have to be aligned or merged. A number of systems have been developed for aligning and merging ontologies and various alignment strategies are used in these systems. However, there are no general methods to support building such tools, and there exist very few evaluations of these strategies. In this thesis we give an overview of the existing systems. We propose a general framework for aligning and merging ontologies. Most existing systems can be seen as instantiations of this framework. Further, we develop SAMBO (System for Aligning and Merging Biomedical Ontologies) according to this framework. We implement different alignment strategies and their combinations, and evaluate them in terms of quality and processing time within SAMBO. We also compare SAMBO with two other systems. The work in this thesis is a first step towards a general framework that can be used for comparative evaluations of alignment strategies and their combinations.

Nyckelord

Keywords

ontologies, ontology engineering, biomedical ontologies, ontology alignment, ontology merging

Linköping Studies in Science and Technology
Faculty of Arts and Sciences - Licentiate Theses

- No 17 **Vojin Plavsic:** Interleaved Processing of Non-Numerical Data Stored on a Cyclic Memory. (Available at: FOA, Box 1165, S-581 11 Linköping, Sweden. FOA Report B30062E)
- No 28 **Arne Jönsson, Mikael Patel:** An Interactive Flowcharting Technique for Communicating and Realizing Algorithms, 1984.
- No 29 **Johanny Eckerland:** Retargeting of an Incremental Code Generator, 1984.
- No 48 **Henrik Nordin:** On the Use of Typical Cases for Knowledge-Based Consultation and Teaching, 1985.
- No 52 **Zebo Peng:** Steps Towards the Formalization of Designing VLSI Systems, 1985.
- No 60 **Johan Fagerström:** Simulation and Evaluation of Architecture based on Asynchronous Processes, 1985.
- No 71 **Jalal Maleki:** ICONStraint, A Dependency Directed Constraint Maintenance System, 1987.
- No 72 **Tony Larsson:** On the Specification and Verification of VLSI Systems, 1986.
- No 73 **Ola Strömfors:** A Structure Editor for Documents and Programs, 1986.
- No 74 **Christos Levcopoulos:** New Results about the Approximation Behavior of the Greedy Triangulation, 1986.
- No 104 **Shamsul I. Chowdhury:** Statistical Expert Systems - a Special Application Area for Knowledge-Based Computer Methodology, 1987.
- No 108 **Rober Bilos:** Incremental Scanning and Token-Based Editing, 1987.
- No 111 **Hans Block:** SPORT-SORT Sorting Algorithms and Sport Tournaments, 1987.
- No 113 **Ralph Rönquist:** Network and Lattice Based Approaches to the Representation of Knowledge, 1987.
- No 118 **Mariam Kamkar, Nahid Shahmehri:** Affect-Chaining in Program Flow Analysis Applied to Queries of Programs, 1987.
- No 126 **Dan Strömberg:** Transfer and Distribution of Application Programs, 1987.
- No 127 **Kristian Sandahl:** Case Studies in Knowledge Acquisition, Migration and User Acceptance of Expert Systems, 1987.
- No 139 **Christer Bäckström:** Reasoning about Interdependent Actions, 1988.
- No 140 **Mats Wirén:** On Control Strategies and Incrementality in Unification-Based Chart Parsing, 1988.
- No 146 **Johan Hultman:** A Software System for Defining and Controlling Actions in a Mechanical System, 1988.
- No 150 **Tim Hansen:** Diagnosing Faults using Knowledge about Malfunctioning Behavior, 1988.
- No 165 **Jonas Löwgren:** Supporting Design and Management of Expert System User Interfaces, 1989.
- No 166 **Ola Petersson:** On Adaptive Sorting in Sequential and Parallel Models, 1989.
- No 174 **Yngve Larsson:** Dynamic Configuration in a Distributed Environment, 1989.
- No 177 **Peter Åberg:** Design of a Multiple View Presentation and Interaction Manager, 1989.
- No 181 **Henrik Eriksson:** A Study in Domain-Oriented Tool Support for Knowledge Acquisition, 1989.
- No 184 **Ivan Rankin:** The Deep Generation of Text in Expert Critiquing Systems, 1989.
- No 187 **Simin Nadjim-Tehrani:** Contributions to the Declarative Approach to Debugging Prolog Programs, 1989.
- No 189 **Magnus Merkel:** Temporal Information in Natural Language, 1989.
- No 196 **Ulf Nilsson:** A Systematic Approach to Abstract Interpretation of Logic Programs, 1989.
- No 197 **Staffan Bonnier:** Horn Clause Logic with External Procedures: Towards a Theoretical Framework, 1989.
- No 203 **Christer Hansson:** A Prototype System for Logical Reasoning about Time and Action, 1990.
- No 212 **Björn Fjellborg:** An Approach to Extraction of Pipeline Structures for VLSI High-Level Synthesis, 1990.
- No 230 **Patrick Doherty:** A Three-Valued Approach to Non-Monotonic Reasoning, 1990.
- No 237 **Tomas Sokolnicki:** Coaching Partial Plans: An Approach to Knowledge-Based Tutoring, 1990.
- No 250 **Lars Strömberg:** Postmortem Debugging of Distributed Systems, 1990.
- No 253 **Torbjörn Näslund:** SLDF-A-Resolution - Computing Answers for Negative Queries, 1990.
- No 260 **Peter D. Holmes:** Using Connectivity Graphs to Support Map-Related Reasoning, 1991.
- No 283 **Olof Johansson:** Improving Implementation of Graphical User Interfaces for Object-Oriented Knowledge-Bases, 1991.
- No 298 **Rolf G Larsson:** Aktivitetsbaserad kalkylering i ett nytt ekonomisystem, 1991.
- No 318 **Lena Srömbäck:** Studies in Extended Unification-Based Formalism for Linguistic Description: An Algorithm for Feature Structures with Disjunction and a Proposal for Flexible Systems, 1992.
- No 319 **Mikael Pettersson:** DML-A Language and System for the Generation of Efficient Compilers from Denotational Specification, 1992.
- No 326 **Andreas Kägedal:** Logic Programming with External Procedures: an Implementation, 1992.
- No 328 **Patrick Lambrix:** Aspects of Version Management of Composite Objects, 1992.
- No 333 **Xinli Gu:** Testability Analysis and Improvement in High-Level Synthesis Systems, 1992.
- No 335 **Torbjörn Näslund:** On the Role of Evaluations in Iterative Development of Managerial Support Systems, 1992.
- No 348 **Ulf Cederling:** Industrial Software Development - a Case Study, 1992.
- No 352 **Magnus Morin:** Predictable Cyclic Computations in Autonomous Systems: A Computational Model and Implementation, 1992.
- No 371 **Mehran Noghabai:** Evaluation of Strategic Investments in Information Technology, 1993.
- No 378 **Mats Larsson:** A Transformational Approach to Formal Digital System Design, 1993.
- No 380 **Johan Ringström:** Compiler Generation for Parallel Languages from Denotational Specifications, 1993.
- No 381 **Michael Jansson:** Propagation of Change in an Intelligent Information System, 1993.
- No 383 **Jonni Harrius:** An Architecture and a Knowledge Representation Model for Expert Critiquing Systems, 1993.
- No 386 **Per Österling:** Symbolic Modelling of the Dynamic Environments of Autonomous Agents, 1993.
- No 398 **Johan Boye:** Dependency-based Groudnness Analysis of Functional Logic Programs, 1993.

- No 402 **Lars Degerstedt:** Tabulated Resolution for Well Founded Semantics, 1993.
- No 406 **Anna Moberg:** Satellitkontor - en studie av kommunikationsmönster vid arbete på distans, 1993.
- No 414 **Peter Carlsson:** Separation av företagsledning och finansiering - fallstudier av företagsledarutköp ur ett agent-teoretiskt perspektiv, 1994.
- No 417 **Camilla Sjöström:** Revision och lagreglering - ett historiskt perspektiv, 1994.
- No 436 **Cecilia Sjöberg:** Voices in Design: Argumentation in Participatory Development, 1994.
- No 437 **Lars Viklund:** Contributions to a High-level Programming Environment for a Scientific Computing, 1994.
- No 440 **Peter Loborg:** Error Recovery Support in Manufacturing Control Systems, 1994.
- FHS 3/94 **Owen Eriksson:** Informationssystem med verksamhetskvalitet - utvärdering baserat på ett verksamhetsinriktat och samskapande perspektiv, 1994.
- FHS 4/94 **Karin Pettersson:** Informationssystemstrukturer, ansvarsfördelning och användarinflytande - En komparativ studie med utgångspunkt i två informationssystemstrategier, 1994.
- No 441 **Lars Poignant:** Informationsteknologi och företagsetablering - Effekter på produktivitet och region, 1994.
- No 446 **Gustav Fahl:** Object Views of Relational Data in Multidatabase Systems, 1994.
- No 450 **Henrik Nilsson:** A Declarative Approach to Debugging for Lazy Functional Languages, 1994.
- No 451 **Jonas Lind:** Creditor - Firm Relations: an Interdisciplinary Analysis, 1994.
- No 452 **Martin Sköld:** Active Rules based on Object Relational Queries - Efficient Change Monitoring Techniques, 1994.
- No 455 **Pär Carlshamre:** A Collaborative Approach to Usability Engineering: Technical Communicators and System Developers in Usability-Oriented Systems Development, 1994.
- FHS 5/94 **Stefan Cronholm:** Varför CASE-verktyg i systemutveckling? - En motiv- och konsekvensstudie avseende arbetssätt och arbetsformer, 1994.
- No 462 **Mikael Lindvall:** A Study of Traceability in Object-Oriented Systems Development, 1994.
- No 463 **Fredrik Nilsson:** Strategi och ekonomisk styrning - En studie av Sandviks förvärv av Bahco Verktyg, 1994.
- No 464 **Hans Olsén:** Collage Induction: Proving Properties of Logic Programs by Program Synthesis, 1994.
- No 469 **Lars Karlsson:** Specification and Synthesis of Plans Using the Features and Fluents Framework, 1995.
- No 473 **Ulf Söderman:** On Conceptual Modelling of Mode Switching Systems, 1995.
- No 475 **Choong-ho Yi:** Reasoning about Concurrent Actions in the Trajectory Semantics, 1995.
- No 476 **Bo Lagerström:** Successiv resultatavräkning av pågående arbeten. - Fallstudier i tre byggföretag, 1995.
- No 478 **Peter Jonsson:** Complexity of State-Variable Planning under Structural Restrictions, 1995.
- FHS 7/95 **Anders Avdic:** Arbetsintegrerad systemutveckling med kalkylprogram, 1995.
- No 482 **Eva L Ragnemalm:** Towards Student Modelling through Collaborative Dialogue with a Learning Companion, 1995.
- No 488 **Eva Toller:** Contributions to Parallel Multiparadigm Languages: Combining Object-Oriented and Rule-Based Programming, 1995.
- No 489 **Erik Stoy:** A Petri Net Based Unified Representation for Hardware/Software Co-Design, 1995.
- No 497 **Johan Herber:** Environment Support for Building Structured Mathematical Models, 1995.
- No 498 **Stefan Svenberg:** Structure-Driven Derivation of Inter-Lingual Functor-Argument Trees for Multi-Lingual Generation, 1995.
- No 503 **Hee-Cheol Kim:** Prediction and Postdiction under Uncertainty, 1995.
- FHS 8/95 **Dan Fristedt:** Metoder i användning - mot förbättring av systemutveckling genom situationell metodkunskap och metodanalys, 1995.
- FHS 9/95 **Malin Bergvall:** Systemförvaltning i praktiken - en kvalitativ studie avseende centrala begrepp, aktiviteter och ansvarsroller, 1995.
- No 513 **Joachim Karlsson:** Towards a Strategy for Software Requirements Selection, 1995.
- No 517 **Jakob Axelsson:** Schedulability-Driven Partitioning of Heterogeneous Real-Time Systems, 1995.
- No 518 **Göran Forslund:** Toward Cooperative Advice-Giving Systems: The Expert Systems Experience, 1995.
- No 522 **Jörgen Andersson:** Bilder av småföretagares ekonomistyrning, 1995.
- No 538 **Staffan Flodin:** Efficient Management of Object-Oriented Queries with Late Binding, 1996.
- No 545 **Vadim Engelson:** An Approach to Automatic Construction of Graphical User Interfaces for Applications in Scientific Computing, 1996.
- No 546 **Magnus Werner:** Multidatabase Integration using Polymorphic Queries and Views, 1996.
- FiF-a 1/96 **Mikael Lind:** Affärsprocessinriktad förändringsanalys - utveckling och tillämpning av synsätt och metod, 1996.
- No 549 **Jonas Hallberg:** High-Level Synthesis under Local Timing Constraints, 1996.
- No 550 **Kristina Larsen:** Förutsättningar och begränsningar för arbete på distans - erfarenheter från fyra svenska företag, 1996.
- No 557 **Mikael Johansson:** Quality Functions for Requirements Engineering Methods, 1996.
- No 558 **Patrik Nordling:** The Simulation of Rolling Bearing Dynamics on Parallel Computers, 1996.
- No 561 **Anders Ekman:** Exploration of Polygonal Environments, 1996.
- No 563 **Niclas Andersson:** Compilation of Mathematical Models to Parallel Code, 1996.
- No 567 **Johan Jenvald:** Simulation and Data Collection in Battle Training, 1996.
- No 575 **Niclas Ohlsson:** Software Quality Engineering by Early Identification of Fault-Prone Modules, 1996.
- No 576 **Mikael Ericsson:** Commenting Systems as Design Support—A Wizard-of-Oz Study, 1996.
- No 587 **Jörgen Lindström:** Chefers användning av kommunikationsteknik, 1996.
- No 589 **Esa Falkenroth:** Data Management in Control Applications - A Proposal Based on Active Database Systems, 1996.
- No 591 **Niclas Wahllöf:** A Default Extension to Description Logics and its Applications, 1996.
- No 595 **Annika Larsson:** Ekonomisk Styrning och Organisatorisk Passion - ett interaktivt perspektiv, 1997.
- No 597 **Ling Lin:** A Value-based Indexing Technique for Time Sequences, 1997.

- No 598 **Rego Granlund:** C³Fire - A Microworld Supporting Emergency Management Training, 1997.
- No 599 **Peter Ingels:** A Robust Text Processing Technique Applied to Lexical Error Recovery, 1997.
- No 607 **Per-Arne Persson:** Toward a Grounded Theory for Support of Command and Control in Military Coalitions, 1997.
- No 609 **Jonas S Karlsson:** A Scalable Data Structure for a Parallel Data Server, 1997.
- FiF-a 4 **Carita Åbom:** Videomötesteknik i olika affärssituationer - möjligheter och hinder, 1997.
- FiF-a 6 **Tommy Wedlund:** Att skapa en företagsanpassad systemutvecklingsmodell - genom rekonstruktion, värdering och vidareutveckling i T50-bolag inom ABB, 1997.
- No 615 **Silvia Coradeschi:** A Decision-Mechanism for Reactive and Coordinated Agents, 1997.
- No 623 **Jan Ollinen:** Det flexibla kontorets utveckling på Digital - Ett stöd för multiflex? 1997.
- No 626 **David Byers:** Towards Estimating Software Testability Using Static Analysis, 1997.
- No 627 **Fredrik Eklund:** Declarative Error Diagnosis of GAPLog Programs, 1997.
- No 629 **Gunilla Ivefors:** Krigsspel och Informationsteknik inför en oförutsägbar framtid, 1997.
- No 631 **Jens-Olof Lindh:** Analysing Traffic Safety from a Case-Based Reasoning Perspective, 1997
- No 639 **Jukka Mäki-Turja:** Smalltalk - a suitable Real-Time Language, 1997.
- No 640 **Juha Takkinen:** CAFE: Towards a Conceptual Model for Information Management in Electronic Mail, 1997.
- No 643 **Man Lin:** Formal Analysis of Reactive Rule-based Programs, 1997.
- No 653 **Mats Gustafsson:** Bringing Role-Based Access Control to Distributed Systems, 1997.
- FiF-a 13 **Boris Karlsson:** Metodanalys för förståelse och utveckling av systemutvecklingsverksamhet. Analys och värdering av systemutvecklingsmodeller och dess användning, 1997.
- No 674 **Marcus Bjärelund:** Two Aspects of Automating Logics of Action and Change - Regression and Tractability, 1998.
- No 676 **Jan Håkegård:** Hierarchical Test Architecture and Board-Level Test Controller Synthesis, 1998.
- No 668 **Per-Ove Zetterlund:** Normering av svensk redovisning - En studie av tillkomsten av Redovisningsrådets rekommendation om koncernredovisning (RR01:91), 1998.
- No 675 **Jimmy Tjäder:** Projektledaren & planen - en studie av projektledning i tre installations- och systemutvecklingsprojekt, 1998.
- FiF-a 14 **Ulf Melin:** Informationssystem vid ökad affärs- och processorientering - egenskaper, strategier och utveckling, 1998.
- No 695 **Tim Heyer:** COMPASS: Introduction of Formal Methods in Code Development and Inspection, 1998.
- No 700 **Patrik Hägglund:** Programming Languages for Computer Algebra, 1998.
- FiF-a 16 **Marie-Therese Christiansson:** Inter-organisatorisk verksamhetsutveckling - metoder som stöd vid utveckling av partnerskap och informationssystem, 1998.
- No 712 **Christina Wennestam:** Information om immateriella resurser. Investeringar i forskning och utveckling samt i personal inom skogsindustrin, 1998.
- No 719 **Joakim Gustafsson:** Extending Temporal Action Logic for Ramification and Concurrency, 1998.
- No 723 **Henrik André-Jönsson:** Indexing time-series data using text indexing methods, 1999.
- No 725 **Erik Larsson:** High-Level Testability Analysis and Enhancement Techniques, 1998.
- No 730 **Carl-Johan Westin:** Informationsförsörjning: en fråga om ansvar - aktiviteter och uppdrag i fem stora svenska organisationers operativa informationsförsörjning, 1998.
- No 731 **Åse Jansson:** Miljöhänsyn - en del i företags styrning, 1998.
- No 733 **Thomas Padron-McCarthy:** Performance-Polymorphic Declarative Queries, 1998.
- No 734 **Anders Bäckström:** Värdeskapande kreditgivning - Kreditriskhantering ur ett agentteoretiskt perspektiv, 1998.
- FiF-a 21 **Ulf Seigerroth:** Integration av förändringsmetoder - en modell för välgrundad metodintegration, 1999.
- FiF-a 22 **Fredrik Öberg:** Object-Oriented Frameworks - A New Strategy for Case Tool Development, 1998.
- No 737 **Jonas Mellin:** Predictable Event Monitoring, 1998.
- No 738 **Joakim Eriksson:** Specifying and Managing Rules in an Active Real-Time Database System, 1998.
- FiF-a 25 **Bengt E W Andersson:** Samverkande informationssystem mellan aktörer i offentliga åtaganden - En teori om aktörsarenor i samverkan om utbyte av information, 1998.
- No 742 **Pawel Pietrzak:** Static Incorrectness Diagnosis of CLP (FD), 1999.
- No 748 **Tobias Ritzau:** Real-Time Reference Counting in RT-Java, 1999.
- No 751 **Anders Ferntoft:** Elektronisk affärskommunikation - kontaktkostnader och kontaktprocesser mellan kunder och leverantörer på producentmarknader, 1999.
- No 752 **Jo Skåmedal:** Arbete på distans och arbetsformens påverkan på resor och resmönster, 1999.
- No 753 **Johan Alvehus:** Mötets metaforer. En studie av berättelser om möten, 1999.
- No 754 **Magnus Lindahl:** Bankens villkor i låneavtal vid kreditgivning till högt belånade företagsförrävar: En studie ur ett agentteoretiskt perspektiv, 2000.
- No 766 **Martin V. Howard:** Designing dynamic visualizations of temporal data, 1999.
- No 769 **Jesper Andersson:** Towards Reactive Software Architectures, 1999.
- No 775 **Anders Henriksson:** Unique kernel diagnosis, 1999.
- FiF-a 30 **Pär J. Ågerfalk:** Pragmatization of Information Systems - A Theoretical and Methodological Outline, 1999.
- No 787 **Charlotte Björkegren:** Learning for the next project - Bearers and barriers in knowledge transfer within an organisation, 1999.
- No 788 **Håkan Nilsson:** Informationsteknik som drivkraft i granskningsprocessen - En studie av fyra revisionsbyråer, 2000.
- No 790 **Erik Berglund:** Use-Oriented Documentation in Software Development, 1999.
- No 791 **Klas Gäre:** Verksamhetsförändringar i samband med IS-införande, 1999.
- No 800 **Anders Subotic:** Software Quality Inspection, 1999.
- No 807 **Svein Bergum:** Managerial communication in telework, 2000.

- No 809 **Flavius Gruian:** Energy-Aware Design of Digital Systems, 2000.
 FiF-a 32 **Karin Hedström:** Kunskapsanvändning och kunskapsutveckling hos verksamhetskonsulter - Erfarenheter från ett FOU-samarbete, 2000.
- No 808 **Linda Askenäs:** Affärssystemet - En studie om teknikens aktiva och passiva roll i en organisation, 2000.
 No 820 **Jan Paul Meynard:** Control of industrial robots through high-level task programming, 2000.
 No 823 **Lars Hult:** Publika Gränssytor - ett designexempel, 2000.
 No 832 **Paul Pop:** Scheduling and Communication Synthesis for Distributed Real-Time Systems, 2000.
 FiF-a 34 **Göran Hultgren:** Nätverksinriktad Förändringsanalys - perspektiv och metoder som stöd för förståelse och utveckling av affärsrelationer och informationssystem, 2000.
- No 842 **Magnus Kald:** The role of management control systems in strategic business units, 2000.
 No 844 **Mikael Cäker:** Vad kostar kunden? Modeller för intern redovisning, 2000.
 FiF-a 37 **Ewa Braf:** Organisations kunskapsverksamheter - en kritisk studie av "knowledge management", 2000.
 FiF-a 40 **Henrik Lindberg:** Webbaserade affärsprocesser - Möjligheter och begränsningar, 2000.
 FiF-a 41 **Benneth Christiansson:** Att komponentbasera informationssystem - Vad säger teori och praktik?, 2000.
 No. 854 **Ola Pettersson:** Deliberation in a Mobile Robot, 2000.
 No 863 **Dan Lawesson:** Towards Behavioral Model Fault Isolation for Object Oriented Control Systems, 2000.
 No 881 **Johan Moe:** Execution Tracing of Large Distributed Systems, 2001.
 No 882 **Yuxiao Zhao:** XML-based Frameworks for Internet Commerce and an Implementation of B2B e-procurement, 2001.
- No 890 **Annika Flycht-Eriksson:** Domain Knowledge Management in Information-providing Dialogue systems, 2001.
 FiF-a 47 **Per-Arne Segerkvist:** Webbaserade imaginära organisationers samverkansformer, 2001.
 No 894 **Stefan Svarén:** Styrning av investeringar i divisionaliserade företag - Ett koncernperspektiv, 2001.
 No 906 **Lin Han:** Secure and Scalable E-Service Software Delivery, 2001.
 No 917 **Emma Hansson:** Optionsprogram för anställda - en studie av svenska börsföretag, 2001.
 No 916 **Susanne Odar:** IT som stöd för strategiska beslut, en studie av datorimplementerade modeller av verksamhet som stöd för beslut om anskaffning av JAS 1982, 2002.
- FiF-a-49 **Stefan Holgersson:** IT-system och filtrering av verksamhetskunskap - kvalitetsproblem vid analyser och beslutsfattande som bygger på uppgifter hämtade från polisens IT-system, 2001.
 FiF-a-51 **Per Oscarsson:** Informationssäkerhet i verksamheter - begrepp och modeller som stöd för förståelse av informationssäkerhet och dess hantering, 2001.
- No 919 **Luis Alejandro Cortes:** A Petri Net Based Modeling and Verification Technique for Real-Time Embedded Systems, 2001.
 No 915 **Niklas Sandell:** Redovisning i skuggan av en bankkris - Värdering av fastigheter. 2001.
 No 931 **Fredrik Elg:** Ett dynamiskt perspektiv på individuella skillnader av heuristisk kompetens, intelligens, mentala modeller, mål och konfidens i kontroll av mikrovärlden Moro, 2002.
- No 933 **Peter Aronsson:** Automatic Parallelization of Simulation Code from Equation Based Simulation Languages, 2002.
 No 938 **Bourhane Kadmiry:** Fuzzy Control of Unmanned Helicopter, 2002.
 No 942 **Patrik Haslum:** Prediction as a Knowledge Representation Problem: A Case Study in Model Design, 2002.
 No 956 **Robert Sevenius:** On the instruments of governance - A law & economics study of capital instruments in limited liability companies, 2002.
- FiF-a 58 **Johan Pettersson:** Lokala elektroniska marknadsplatser - informationssystem för platsbundna affärer, 2002.
 No 964 **Peter Bunus:** Debugging and Structural Analysis of Declarative Equation-Based Languages, 2002.
 No 973 **Gert Jervan:** High-Level Test Generation and Built-In Self-Test Techniques for Digital Systems, 2002.
 No 958 **Fredrika Berglund:** Management Control and Strategy - a Case Study of Pharmaceutical Drug Development, 2002.
- FiF-a 61 **Fredrik Karlsson:** Meta-Method for Method Configuration - A Rational Unified Process Case, 2002.
 No 985 **Sorin Manolache:** Schedulability Analysis of Real-Time Systems with Stochastic Task Execution Times, 2002.
- No 982 **Diana Szentiványi:** Performance and Availability Trade-offs in Fault-Tolerant Middleware, 2002.
 No 989 **Iakov Nakhimovski:** Modeling and Simulation of Contacting Flexible Bodies in Multibody Systems, 2002.
 No 990 **Levon Saldamli:** PDEModelica - Towards a High-Level Language for Modeling with Partial Differential Equations, 2002.
- No 991 **Almut Herzog:** Secure Execution Environment for Java Electronic Services, 2002.
 No 999 **Jon Edvardsson:** Contributions to Program- and Specification-based Test Data Generation, 2002
 No 1000 **Anders Arpteg:** Adaptive Semi-structured Information Extraction, 2002.
 No 1001 **Andrzej Bednarski:** A Dynamic Programming Approach to Optimal Retargetable Code Generation for Irregular Architectures, 2002.
- No 988 **Mattias Arvola:** Good to use! : Use quality of multi-user applications in the home, 2003.
 FiF-a 62 **Lennart Ljung:** Utveckling av en projektivitetsmodell - om organisationers förmåga att tillämpa projektarbetsformen, 2003.
- No 1003 **Pernilla Qvarfordt:** User experience of spoken feedback in multimodal interaction, 2003.
 No 1005 **Alexander Siemers:** Visualization of Dynamic Multibody Simulation With Special Reference to Contacts, 2003.
- No 1008 **Jens Gustavsson:** Towards Unanticipated Runtime Software Evolution, 2003.
 No 1010 **Calin Curescu:** Adaptive QoS-aware Resource Allocation for Wireless Networks, 2003.
 No 1015 **Anna Andersson:** Management Information Systems in Process-oriented Healthcare Organisations, 2003.
 No 1018 **Björn Johansson:** Feedforward Control in Dynamic Situations, 2003.
 No 1022 **Traian Pop:** Scheduling and Optimisation of Heterogeneous Time/Event-Triggered Distributed Embedded Systems, 2003.
- FiF-a 65 **Britt-Marie Johansson:** Kundkommunikation på distans - en studie om kommunikationsmediets betydelse i affärstransaktioner, 2003.
 No 1024 **Aleksandra Tesanovic:** Towards Aspectual Component-Based Real-Time System Development, 2003.

- No 1034 **Arja Vainio-Larsson:** Designing for Use in a Future Context - Five Case Studies in Retrospect, 2003.
No 1033 **Peter Nilsson:** Svenska bankers redovisningsval vid reservering för befarade kreditförluster - En studie vid införandet av nya redovisningsregler, 2003.
- Fif-a 69 **Fredrik Ericsson:** Information Technology for Learning and Acquiring of Work Knowledge, 2003.
No 1049 **Marcus Comstedt:** Towards Fine-Grained Binary Composition through Link Time Weaving, 2003.
No 1052 **Åsa Hedenskog:** Increasing the Automation of Radio Network Control, 2003.
No 1054 **Claudiu Duma:** Security and Efficiency Tradeoffs in Multicast Group Key Management, 2003.
Fif-a 71 **Emma Eliasson:** Effekttanalys av IT-systems handlingsutrymme, 2003.
No 1055 **Carl Cederberg:** Experiments in Indirect Fault Injection with Open Source and Industrial Software, 2003.
No 1058 **Daniel Karlsson:** Towards Formal Verification in a Component-based Reuse Methodology, 2003.
FiF-a 73 **Anders Hjalmarsson:** Att etablera och vidmakthålla förbättringsverksamhet - behovet av koordination och interaktion vid förändring av systemutvecklingsverksamheter, 2004.
No 1079 **Pontus Johansson:** Design and Development of Recommender Dialogue Systems, 2004.
No 1084 **Charlotte Stoltz:** Calling for Call Centres - A Study of Call Centre Locations in a Swedish Rural Region, 2004.
FiF-a 74 **Björn Johansson:** Deciding on Using Application Service Provision in SMEs, 2004.
No 1094 **Genevieve Gorrell:** Language Modelling and Error Handling in Spoken Dialogue Systems, 2004.
No 1095 **Ulf Johansson:** Rule Extraction - the Key to Accurate and Comprehensible Data Mining Models, 2004.
No 1099 **Sonia Sangari:** Computational Models of Some Communicative Head Movements, 2004.
No 1110 **Hans Nässla:** Intra-Family Information Flow and Prospects for Communication Systems, 2004.
No 1116 **Henrik Sällberg:** On the value of customer loyalty programs - A study of point programs and switching costs, 2004.
FiF-a 77 **Ulf Larsson:** Designarbete i dialog - karaktärisering av interaktionen mellan användare och utvecklare i en systemutvecklingsprocess, 2004.
No 1126 **Andreas Borg:** Contribution to Management and Validation of Non-Functional Requirements, 2004.
No 1127 **Per-Ola Kristensson:** Large Vocabulary Shorthand Writing on Stylus Keyboard, 2004.
No 1132 **Pär-Anders Albinsson:** Interacting with Command and Control Systems: Tools for Operators and Designers, 2004.
No 1130 **Ioan Chisalita:** Safety-Oriented Communication in Mobile Networks for Vehicles, 2004.
No 1138 **Thomas Gustafsson:** Maintaining Data Consistency in Embedded Databases for Vehicular Systems, 2004.
No 1149 **Vaida Jakonienė:** A Study in Integrating Multiple Biological Data Sources, 2005.
No 1156 **Abdil Rashid Mohamed:** High-Level Techniques for Built-In Self-Test Resources Optimization, 2005.
No 1162 **Adrian Pop:** Contributions to Meta-Modeling Tools and Methods, 2005.
No 1165 **Fidel Vascós Palacios:** On the information exchange between physicians and social insurance officers in the sick leave process: an Activity Theoretical perspective, 2005.
FiF-a 84 **Jenny Lagsten:** Verksamhetsutvecklande utvärdering i informationssystemprojekt, 2005.
No 1166 **Emma Larsdotter Nilsson:** Modeling, Simulation, and Visualization of Metabolic Pathways Using Modelica, 2005.
No 1167 **Christina Keller:** Virtual Learning Environments in higher education. A study of students' acceptance of educational technology, 2005.
No 1168 **Cécile Åberg:** Integration of organizational workflows and the Semantic Web, 2005.
FiF-a 85 **Anders Forsman:** Standardisering som grund för informationssamverkan och IT-tjänster - En fallstudie baserad på trafikinformationstjänsten RDS-TMC, 2005.
No 1171 **Yu-Hsing Huang:** A systemic traffic accident model, 2005.
FiF-a 86 **Jan Olausson:** Att modellera uppdrag - grunder för förståelse av processinriktade informationssystem i transaktionsintensiva verksamheter, 2005.
No 1172 **Petter Ahlström:** Affärsstrategier för seniorbostadsmarknaden, 2005.
No 1183 **Mathias Cöster:** Beyond IT and Productivity - How Digitization Transformed the Graphic Industry, 2005.
No 1184 **Åsa Horzella:** Beyond IT and Productivity - Effects of Digitized Information Flows in Grocery Distribution, 2005.
No 1185 **Maria Kollberg:** Beyond IT and Productivity - Effects of Digitized Information Flows in the Logging Industry, 2005.
No 1190 **David Dinka:** Role and Identity - Experience of technology in professional settings, 2005.
No 1191 **Andreas Hansson:** Increasing the Storage Capacity of Recursive Auto-associative Memory by Segmenting Data, 2005.
No 1192 **Nicklas Bergfeldt:** Towards Detached Communication for Robot Cooperation, 2005.
No 1194 **Dennis Maciuszek:** Towards Dependable Virtual Companions for Later Life, 2005.
No 1204 **Beatrice Alenljung:** Decision-making in the Requirements Engineering Process: A Human-centered Approach, 2005
No 1206 **Anders Larsson:** System-on-Chip Test Scheduling and Test Infrastructure Design, 2005.
No 1207 **John Wilander:** Policy and Implementation Assurance for Software Security, 2005.
No 1209 **Andreas Käll:** Översättningar av en managementmodell - En studie av införandet av Balanced Scorecard i ett landsting, 2005.
No 1225 **He Tan:** Aligning and Merging Biomedical Ontologies, 2006.