

ERCIM “Alain Bensoussan” Fellowship Scientific Report

Fellow: Hanna Kozankiewicz-Sielska
Visited Location: ATLAS Group, INRIA and LINA University of Nantes
Duration of Visit: 9 months

I - Scientific activity

My research during ERCIM fellowship was focused on applying Model-Driven Engineering to problems from the area of databases. The work was conducted basing on the platform AMMA (Atlas Model Management Architecture) developed by ATLAS group at INRIA/LINA University of Nantes. The platform supports model-to-model transformations. The part of AMMA that has been used during work is AMW (Atlas Model Weaver). AMW supports identifying and managing correspondences between models (i.e., information on sets of elements from woven models that are related to one to another).

The research during the fellowship based on AMW prototype mainly focused on the two main aspects:

- (1) *Analysing the problem of the inverting schema mappings in AMW.* The problem of inverting schema mappings is very important in the area of model management. It has multiple applications in schema evolution, data exchange, etc. However, to the best of our knowledge only one approach to the problem exists (and it is a theoretical one). Therefore, our goal was to propose and implement a practical solution. It consists in extending a metamodel of the weaving model in such a way to include information about all types of relationship between woven models' elements and information about inverses of such operations. Based on this metamodel we are later able to automatically generate a global inverse (inverse weaving model). Later, inverse weaving model can be transformed into executable transformation so we are able to transform data form one form to another. The results of this research have been described in the paper [1].
- (2) *Analyzing the problem of data integration using AMW.* We studied how Model-Driven Engineering can support problems of integration of schemas/metamodels/ontologies. The proposed scenario of integration in AMW prototype looks as follows. First, weaving metamodel should be created (if a basic AMW weaving metamodel appears not sufficient). Next, weaving links should be evaluated either manually or basing on provided matching transformations. Finally, the integrated metamodel should be calculated on the basis of woven metamodels, and the weaving model. The study on data integration has been carried on context of various technologies like OWL, XML, etc. The results of this research have been described in paper [2].

During fellowship I also continued my earlier research in core database area. This resulted in publications [3].

During the fellowship I was also a reviewer as a member of the Programme Committees of:

1. 18th International Conference on Database and Expert Systems Applications – DEXA 2007
2. First International Workshop on Semantic Technology Adoption in Business – STAB 2007

II- Publication(s) during your fellowship

1. *Implementing Inverse Mappings in a Model Management Framework*. Co-authors: M.Didonet Del Fabro, J.Bézivin, P.Valduriez. The paper has been submitted to ICDE 2007.

Abstract: Schema mapping is critical to many data management applications. Recently, much work has focused on automating the process using a set of basic operators on schemas in order to support more complex operations. In this paper, we concentrate on one important basic operator – inverting of mappings. There has been some theoretical work on inverse mappings, but to the best of our knowledge, there is still a lack of a practical approach for cases where a mapping can be more complex than 1:1. Our motivation is to fill this gap and propose an implementation of inverse mappings. In this paper, we show how we can build a metamodel of correspondences between source and target data that includes information about inverses of operations. Using this information, we are able to semi-automatically find inverse mappings. To validate our approach, we use the Atlas Model Weaver tool on a Sales Force Automation system and we also verify performance and scalability of our solution on bigger models.

2. *Applying Model Engineering to Data Integration*. Hanna Kozankiewicz , Marcos Didonet Del Fabro, Patrick Valduriez. To be submitted to SAC 2007.

Abstract: Since many years the amount of resources on the Web constantly grows, what makes the problem of data integration remains in a focus of research community. In this paper, we show that Model-Driven Engineering (MDE) paradigm can be successfully applied to deal with problems of data integration (a cases of horizontal, vertical as well as hybrid data fragmentation). To support this thesis we introduce an Atlas Model Weaver prototype that allows performing tasks related to data integration. Main advantage of applying of MDE is that we are able to easily integrate data having different representations.

3. Work partially prepared during fellowship. *Executing Distributive Queries to Horizontally Fragmented Data*. Co-authors: J.Matusz, K.Stencel, K.Subieta. The paper has been submitted to EuroMicro 2007.

Abstract: In this paper we analyze query execution strategies for object-oriented and XML-oriented distributed databases that are horizontally fragmented. We introduce several scenarios of query decomposition based on the property of the query language – fully compositional and precise formal semantics. For this purpose we use the Stack-Based Approach (SBA) and its query language SBQL, which is well prepared to distributed query optimization. The presented methods have been successfully validated within distributed database prototype LoXiM and they will be one of methods of optimization in a virtual repository ODRA being implemented within our 6th EC Framework project.

III -Attended Seminars, Workshops, and Conferences

Given talks:

1. *Updatable Object Views in Data Integration*, internal seminar of Atlas team at the University of Nantes
2. *Querying Workflows over Distributed Systems*. 3rd International Conference on Web Information System and Technologies, Barcelona, Spain, 2007

Attended conference:

1. 3-6 March 2007. 3rd International Conference on Web Information System and Technologies, Barcelona, Spain

Attended workshop:

1. 20 June 2007. Le développement de logiciels avec Eclipse. Workshop at the university of Nantes

Attended seminars:

1. 16 November 2006, University of Nantes. *Distributed Data Management – Web, Clusters & Grid*. Marta Mattoso, UFRJ, Rio de Janeiro
2. 16 November 2006, University of Nantes. *ParGRES - an open-source database cluster middleware for OLAP applications*. Alexandre Lima, UFRJ, Rio de Janeiro
3. 14 December 2007, University of Nantes. *Data Privacy in Distributed Systems*. Patricia Serrano Alvarado, University of Nantes
4. 18 January 2007, University of Nantes. *StrangerDB: Safe Data Management with Untrusted Servers*. Dennis Shasha, University of New York
5. 18 January 2007, University of Nantes. *The Nature of Invention in Computer Science -- a collaborative reflection*. Dennis Shasha, University of New York
6. 25 January 2007, University of Nantes. *Database Approaches for Biological Data*. Sérgio Lifschitz, PUC-Rio, Rio de Janeiro
7. 21 March 2007, University of Nantes. *Ubiquitous computing, models and the informatic future*. Robin Milner, Cambridge University
8. 11 May 2007, University of Nantes. Database summary management in P2P systems. Rabab Hayek, , University of Nantes.
9. 24 May 2007, University of Nantes. *Build One, Get One Free: Leveraging the Coexistence of Multiple P2P Overlay Networks*. Anne-Marie Kermarrec, INRIA/IRISA, Rennes
10. 24 May 2007, University of Nantes. *PlugDB: Querying Visible and Hidden Data without Leaks*. Philippe Pucheral, INRIA, Roquencourt and PRISM, Versailles
11. 1 June 2007, University of Nantes. *Bibliothèques numériques ouvertes*. Marcos Sunyé, LIP6
12. 7 June 2007, University of Nantes. *Data currency in replicated DHTs*. Reza Akbarinia, University of Nantes
13. 5 July 2007, University of Nantes. *How to have a successful und research career – a summary of the Young Faculty Symposium which was held during SIGMOD 2007*. Patrick Valduriez