

Doctoral Symposium

Marcus Alanen¹, Jordi Cabot², Miguel Goulão³ and José Sáez⁴

¹Åbo Akademi University, Turku, Finland
maalanen@abo.fi

²Estudis d'Informàtica i Multimèdia, Universitat Oberta de Catalunya, Spain
jcabot@uoc.edu

³Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
miguel.goulao@di.fct.unl.pt

⁴Universidad de Murcia, Spain
jsaez@um.es

Abstract: The UML 2004 Doctoral Symposium was the first Doctoral Symposium in the UML Conference series. The Doctoral Symposium sought to bring together PhD Students working in areas related to UML and modeling in general. It was a full-day workshop held in parallel with the remaining workshops of the conference. Ten students were selected and were given the opportunity to present and discuss their research goals, receiving high-quality feedback from the rest of participants of the workshop, including a number of volunteer seniors that helped making the Symposium a success. The Doctoral Symposium will also be present in the next edition of the conference.

1. Introduction

The first Doctoral Symposium in the UML conference series was held on the 10th of October at UML 2004. The aim of the Symposium was to bring together PhD students working in areas related to UML and modeling in general. It was organized by five PhD Students: Marcus Alanen (Åbo Akademi University), Jordi Cabot (Universitat Oberta de Catalunya), Miguel Goulão (Universidade Nova de Lisboa), José Sáez (Universidad de Murcia) and Devon Simmonds (Colorado State University).

The Symposium was intended for students who had already settled on a specific research proposal and had some preliminary results, but still had enough time remaining before submitting their dissertation, so that they could benefit from the Symposium discussions.

Ten students were selected to participate in the Symposium. Submissions were judged on originality, significance, technical merit, presentation quality and relevance to the conference topics. Each paper was reviewed by, at least, two reviewers of the Program Committee, composed of a group of reviewers with large experience in the UML conference: Fernando Brito e Abreu, João Araújo, Jean Bézivin, Robert France, Gonzalo Génova, Martín Gogolla, Heinrich Hussmann, Ivan Porres, Bran Selic, Friedrich Steimann, Ernest Teniente, Ambrosio Toval and Belén Vela.

The selected students had the opportunity to present and to discuss their research goals, methods and results within a constructive and international atmosphere. They received high-quality feedback from the rest of participants of the workshop, including a number of volunteer seniors that attended the Symposium or part of it.

The papers presented by the selected students covered a wide range of topics, from aspects modeling to model transformation and from model validation to new methods for Information Systems development, representing most of the major topics of the main conference. All papers are available online at the Symposium web page:

<http://ctp.di.fct.unl.pt/UML2004/phdSymp.htm>

2. Structure of the Doctoral Symposium

The workshop was structured in four different sessions where we tried to group the presentations addressing related topics. Each session was composed of two or three thirty minutes presentations where each presentation included at least ten minutes for discussion.

The first session dealt with aspects in the requirements (work of Isabel Brito), analysis and design stages (work of Y. Raghv Reddy). Then we moved on to papers related with model-driven development, focused on the transformations from Platform Independent Models (PIMs) to Platform Specific Models (PSMs) and to code, either in the specific field of Pervasive Systems Development (Javier Muñoz), in the integration of Security Patterns (Diego Ray) or in the generation of databases structures to control multiplicity constraints (H.T. Al-Jumaily).

The third session included some papers on model validation (Jörn Guy Süß), coordination diagrams (David Safranek) and model testing (Trung Dinh-Trong). Finally, two presentations proposed new methods for IS development from conceptual models. The first one proposed creating generic conceptual models for each application domain (Ruth Raventós) while the second one proposed the use of ontologies as the initial conceptual model (Jordi Conesa).

The presentations are described below, in order of presentation at the Symposium. Unfortunately two students were unable to attend the Symposium. In both cases their respective supervisor presented the work instead.

2.1 Aspect-Oriented Requirements Engineering (by Isabel Brito)

The work described focuses on a particular aspect of separation of concerns. While separating concerns into individual modules decreases software complexity and enhances understandability and traceability, there are properties which do not lend themselves to such strict modularization and are said to be cross-cutting the system. Usually they cannot be encapsulated into one component but rather are scattered throughout the system. Examples of this include non-functional requirements such as distribution, security and synchronization.

Aspect-oriented software development aims to address these issues by providing means for identification, separation, representation and composition of cross-cutting concerns. Although aspects have been used in software development, research on the use of aspects at the requirements stage is still immature. The goal of the work is to develop a framework for aspect-oriented requirements engineering that supports the identification, specification, modeling and composition of cross-cutting concerns at the requirements level.

A brief outline of how to accomplish these four tasks is given. Identifying concerns can be done using e.g. common techniques for requirements elicitation and reusing already developed catalogs. Specifying concerns is divided into four subtasks; identifying responsibilities and priorities of concerns, contributions between concerns, and dependencies among concerns. Modeling is accomplished by building a requirements analysis model as well as a behavioral model. Composition of concerns is done using match points, conflict handling and defining composition rules.

2.2 An Aspect Oriented Approach to Early Software Development (by Y. Raghu Reddy, presented by Robert France)

Locating related requirements and tracking the impact of changing requirements are two tasks that contribute significantly to the complexity in the development of large scale software systems. This work concerns the development of a rigorous aspect-oriented development approach to support these tasks during the early stages of software development.

The proposed approach is called Requirements Aspect-Oriented Modeling (RAM) and allows identifying cross-cutting concerns at the requirements level (requirements aspects). The RAM approach creates a link between requirements aspects and their corresponding design aspects by providing experience-based generic solutions that aid in the identification of conflicts between cross-cutting concerns.

The approach is expected to provide system architects with techniques to systematically identify, represent, and trace concerns throughout the software life cycle. The approach is based on the notion of viewpoints and uses the Role-Based Modeling Language (RBML) to represent cross-cutting concern solutions.

2.3 Pervasive Systems Development with the Model Driven Architecture (by Javier Muñoz)

Computing based systems growth is reaching all environments of our daily life. Pervasive systems live around us, providing services to the inhabitants of a home, the workers of an office or the drivers in a car park.

The development of this kind of systems is very hard because they have to achieve devices interoperability in a heterogeneous environment in order to satisfy system requirements. This situation requires solid engineering methods for developing robust systems. In this context, this work seeks to improve current state of the art of pervasive systems development techniques by means of an MDA based method for pervasive systems development.

It proposes the Pervasive Modeling Language (Perv-ML) a precise language for building Platform Independent Models (PIMs). Perv-ML promotes the separation of roles between analysts and architects. Analysts specify the services, structural and interaction model. Afterwards, architects specify what COTS devices and/or software systems are in charge of each service.

As a Platform Specific Model (PSM), it proposes a language for modeling an OSGi system. OSGi is a Java middleware initially created for hosting software of residential gateways. Then, it applies graph grammars to define the transformations from Perv-ML to OSGi. Finally, using a set of templates, it generates the code from the PSMs.

2.4 A Systematic Approach to Testing UML Design Models (by Trung Dinh-Trong, presented by Sudipto Ghosh)

The research proposes a systematic approach to testing design models described by UML class diagrams, sequence diagrams and activity diagrams. This approach can help in finding and removing faults in designs before these are implemented. The work suggests a dynamic testing approach in which executable forms of UML design models are exercised with test inputs. The expected behavior of a design under test is compared to the actual behavior that is observed during testing, and differences are reported.

The approach includes a set of test adequacy criteria leading to high rates of fault detection, and a technique to systematically generate test suites satisfying these criteria. It also addresses the issue of UML model execution, and the questions of how can faults be detected, and what types of faults.

None of the studied approaches for UML model execution generate test infrastructure code that supports systematic testing of models. Most of the test input generation techniques only provide test requirements without providing how to derive test inputs from the requirements. The testing approach presented can be applied to models that consist of class, sequence and activity diagrams.

The testing begins when the user provides a design under test (DUT) and a set of adequacy criteria to the system. The model is transformed into an EDUT (executable DUT) by translating it to Java, and finally, the test cases generated from the adequacy

criteria and test drivers are added to the model to form the TDUT (testable DUT). Code to detect failures checks that all variables in conditions and parameters in method calls are initialized, and that the target object of a message exists. Other checks based on OCL are delegated to an external tool.

To date, the fault detection ability of the algorithm to execute and observe UML design testing has been validated using two case studies with promising results.

2.5 Integration of Security Patterns in Software Models based on Semantic Descriptions (by Diego Ray)

This proposal aims to develop a tool-supported implementation framework for business model driven security engineering. It includes a new security engineering process that results from a fusion of software and systems engineering with security engineering and formal methods for the design and analysis of secure systems. The objective is to provide methodologies and tools for the generation of executable systems with fully configurable security infrastructures.

The proposed model is a variation of Boehm's basic spiral model for software development processes, where security engineering is integrated in the development phase of the cycle. This integration relies on the use of security patterns that represent security services with specific profiles and solutions for different environments. Each pattern will be described using XML-based meta-models. The proposal will also include the necessary pattern language.

The approach also aims to define mechanisms to automate the analysis of security-enhanced models in order to find the security patterns and to allow the inclusion of these patterns in the model, as well as defining appropriate mechanisms to validate the integrity of models with respect to security requirements.

2.6 Plugging Active Mechanisms to Control Dynamic Aspects Derived from the Multiplicity Constraint in UML (by H. T. Al-Jumaily)

This research shows how the process of transforming conceptual schemas to logical schemas in database design is sometimes subject to semantic losses. In particular, the problem is focused on the case of the cardinality constraint problem.

The issue is addressed by integrating add-ins to existing database CASE tools to automatically generate triggers which verify specific conditions on insertion, deletion and update operations in the database. The tools are used when generating the logical database schema (tables and constraints) from the UML model.

Triggers are defined according to the SQL 2003 standard, as event-condition-action rules that are activated by a database state transition. The two other important concepts are event (INSERT / DELETE / UPDATE in the database table) and activation time (BEFORE / AFTER), which defines whether the trigger is activated before or after the event.

The problem of multiplicity is split in two cases. In the first case, triggers must be generated for one-to-many associations, where two events must be considered:

- 1) deleting or updating in the one-table: the many-table must be cascade-deleted or updated, but no checks about multiplicity are carried out
- 2) deleting or updating in the many-table, minimum cardinality constraints must be checked to assure that there are enough detail items related to the master item. Insertion also needs the verification of the maximum multiplicity.

The second case deals with many-to-many associations, implemented with three tables. Here, inserting, deleting or updating in any of the three tables must be followed by the corresponding verification with the appropriate trigger.

Automatic triggers are also used to maintain the consistency in data when there are generalizations involved in the conceptual model, and these generalizations have been mapped to a three-table architecture.

The last part of the research paper describes how the add-in has been designed and integrated into the CASE tool, and explains how the user interacts with it and sets the options needed to generate the triggers.

2.7 A Standards-based UML-profile for message-based information dissemination (by Jörn Guy Süb)

Integrating information systems using message queues and XML documents has been considered difficult to design and manage, although they are reliable in operation. Combined with transformation technology they have been called the 'preferred enterprise application integration engine'. The work develops a UML profile for this domain, relying on UML and other OMG standards. Specifically, OCL is used for constraints on the UML metamodel level.

Usage of UML and modeling in general can be divided into three categories: sketches, blueprints and programs. The work looks at the viability of UML usage in blueprinting, where models are kept in a tool along with diagrams which show the different viewpoints of the system under discussion. The work first studies the available means for creating a standards-based UML profile for blueprinting business systems, and then discusses the solution in the form of the EVE framework, where OCL constraints are verified by a server separate from the actual developer's modeling environment.

Currently, the service platform, validator and local profile for analysis are implemented, while some profiles are near completion or exist only as a draft. In parallel to this development, a real industry example is modeled to ensure that the method defined by the profiles is viable.

2.8 Visual Coordination Diagrams (by David Safranek)

The objective of the work is to define a universal formal visual language for concurrent systems. Such a formalism should be capable of handling both the coordination and the behavioral aspects of concurrent systems. Properties of importance are the

suitability of the formalism for modeling heterogeneity, hierarchy and component-based structure. In general, visual formalisms for concurrent systems are split into two groups. The first group is formed by state-based formalisms, e.g. statecharts, and the second one by dataflow-based formalisms, e.g. message sequence charts. Both approaches emphasize different aspects of modeled systems and can be combined in a particular design.

The proposed research strives to incorporate both the state-based and the dataflow-based approaches into a single formalism. Based on previous research, these two approaches are going to be separated into two independent layers of the language, and at both layers heterogeneity should be achieved. The proposed language, Visual Coordination Diagrams, focuses on the coordination level whereas the behavioral level is represented by state transition diagrams, formalized as a Mealy machine. Components of the system are grouped into networks at the coordination layer, where input and output ports connect components. The network structure is hierarchical.

The goal of VCD is to build a framework for the coordination of different statecharts and other visual formalisms. The development of a graphical tool for creation and modification of VCD diagrams is in progress. It will make use of current suitable verification tools.

2.9 Model and Function Driven Development (by Ruth Raventós)

This work proposes an Information System development approach that may provide a substantial increase in Information Systems development productivity and, at the same time, facilitate changes to accommodate new functional requirements. This new approach is called Model and Function Driven Development.

Its main features are the distinction between the model and the function of a Conceptual Schema (CS) and its reuse of generic conceptual schemas.

In this approach, the conceptual schema of a particular IS, called specific CS, is not generated from scratch. It is obtained by refining a generic CS of the same domain. The generic CS of a given domain consists of the elements that should be present in all or many CSs for that domain. As an example, the generic CS for an auction domain will contain all the common elements appearing in any particular CS for an auction IS, like the concepts of auction, bidder and bid.

Both, the specific and the generic CS, include two kinds of knowledge: about its domain and about the functions it must perform. The former is called the model and the latter the function.

The work explains how the model and the function of the specific CS can be obtained by refinement of the model and function of the general CS after using a predefined set of domain-independent schema transformation operations.

To facilitate the refinement of the function part, the generic CS must be kept flexible enough. To this end, some of the ideas of the frameworks are applied when defining the CS, such as: the concepts of hot-spot, template methods and hook methods.

2.10 Ontology-Driven Information Systems: Pruning and Refactoring of Ontologies (by Jordi Conesa)

Ontology-Driven Information Systems is a method to develop conceptual schemas as refinements of the knowledge contained in general ontologies, instead of generating the conceptual schema from scratch. It differs from the previous proposal in that the starting point is not a general conceptual schema for the domain but a general ontology, which is supposed to be a *universal* conceptual schema, and thus, the initial conceptual schema is much larger. On the other hand, this proposal only addresses the static part of the IS.

The method comprises three different steps: refinement, pruning and refactoring of ontologies. During the refinement the general ontology is extended (if necessary) with the knowledge required to model the IS. After that, the ontology contains all the necessary knowledge but it is too large to be used as a final conceptual schema.

That is why we need the pruning phase. During this phase we delete from the ontology all the superfluous elements to the final conceptual schema. Afterwards, we may apply some refactoring techniques to the resulting conceptual schema to obtaining the final schema.

The main focus of the work is the development of the pruning and refactoring phases. The work describes a method that given a set of concepts of direct interest deletes all the irrelevant concepts of the extended ontology. The concepts of direct interest are those appearing in the requirements of the IS we want to specify. Taking into account this set of concepts and the relationships between them and the other elements of the ontology, the method guarantees the removal of the irrelevant elements without losing any of the needed semantics of the IS.

3. Honorary Best Student Paper Award

In conjunction with the Conference Chair of the UML 2004, the Doctoral Symposium organizers decided to promote an honorary best student paper award. The members of the program committee attending the Symposium, along with the organizers, based their judgment on the contents and actual presentation of the work. Overall participation in the workshop was also considered. Based on these criteria, Jörn Guy Süß received the award for his work titled “A Standards-based UML-profile for message-based information dissemination”. The award was sponsored by the Director of Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, and presented by Ana Moreira, the UML 2004 Conference Chair. The award consisted of a certificate and was accompanied by an attractive, fully illustrated book on Portugal.

4. Conclusions

The first Doctoral Symposium of the UML conference series can be considered a success. The participants considered the comments made by the reviewers to the submitted abstracts and the feedback received during the presentation Symposium as a positive contribution that will help them to improve the quality of their work.

The presence of a significant number of workshop attendees (over 20), including participants, committee members, organizers and several PhD students contributed to lively discussions which challenged each presenter within a very constructive environment.

Although no time was allocated explicitly from the schedule for a general discussion on common problems for PhD Students or for comparisons between different educational institutes PhD programs and working conditions, we should stress the funding problems that the students need to overcome when doing their thesis: two of the selected students and one of the organizers could not attend the Symposium due to funding problems. Most of the PhD students participating in the Symposium were not registered for the main conference. We believe that, when doing a PhD thesis, it is of the greatest importance to attend as many conferences as possible, and thus more efforts should be devoted to make this possible.

The Doctoral Symposium will continue at the Models 2005 conference, with Jeff Gray as chair handling the organization. It will have a different format, where each student will be assigned a mentor who will lead the discussion following the student's presentation. Due to the mentoring aspect of the event, the Symposium will be open only to those students and mentors participating directly in it. We expect this will improve the quality of the Symposium even further. Further details on the new edition of the Doctoral Symposium can be found at:

<http://www.cs.colostate.edu/models05/cfpDoctoralSymposium.html>

Acknowledgments

The organizers of the Symposium would like to thank all the seniors that volunteered to make the Symposium a great success. In particular, we are very grateful to all the members of our Senior Program Committee, and especially to the seniors that attended the workshop and shared some of their time helping the students to improve their work.

We would also like to thank Ambrosio Toval and Ana Moreira, for their valuable advice during the preparation of the Symposium.

Finally the organizers are grateful to all PhD students attending the Symposium for submitting and/or sharing their ideas.