INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team dahu

Verification in databases

Saclay - Île-de-France

Theme : Knowledge and Data Representation and Management

Table of contents

1. Team **1**
2. Overall Objectives **1**
3. Scientific Foundations **1**
4. Application Domains **2**
5. New Results **2**
5.1. XML specification and verification 2
5.1.1. Static analysis of query languages 2
5.1.2. Incomplete information 3
5.1.3. Data centric Workflows 3
5.1.4. Analysis of Distributed Data Management 4
5.1.5. Verification of XML updates 4
5.1.6. Probabilistic XML 4
5.2. Automata and logics for data words and data trees 4
5.3. Automata theory 5
6. Other Grants and Activities **6**
6.1. National collaborations 6
6.2. International collaborations 7
6.2.1. Cooperation within Europe 7
6.2.2. Cooperation with Tunisia 7
6.2.3. Cooperation with North America 7
7. Dissemination **7**
7.1. Thesis 7
7.2. Participation to conferences organization 8
7.3. Participation to symposia, seminars, invitations 8
7.4. Scientific Animations 8
7.5. Teaching 9
7.6. Thesis jury 9
8. Bibliography **9**

Dahu is a common project with LSV and ENS de Cachan. The team was created on January the 1st, 2008.

1. Team

Research Scientists

Serge Abiteboul [Research Director (DR), HdR]
Stéphane Demri [Research Director (DR) CNRS, HdR]
Florent Jacquemard [Research assistant (CR), Inria]
Luc Segoufin [Team leader, Research Director (DR), Inria, HdR]

Faculty Member

Cristina Sirangelo [MCF ENS Cachan]

PhD Students

Pierre Bourhis [Allocation couplée]
Diego Figueira [Cordi, until September]
Wojciech Kazana [Webdam]
Alban Galland [Webdam]
Thomas Place [Allocation couplée, until September]
Camille Vacher [Allocation couplée, until September]

Post-Doctoral Fellows

Yannis Katsis [Webdam]
Bruno Marnette [Webdam, since September]

Visiting Scientist

Victor Vianu [In Dahu since mid June. Professor, UC San Diego, USA]

Administrative Assistants

Isabelle Biercewicz [Secretary (SAR) Inria]
Valérie Lecomte [Secretary (SAR) Inria, replacing Isabelle during her maternity leave]

2. Overall Objectives

2.1. Overall Objectives

For more information see http://www.lsv.ens-cachan.fr/axes/DAHU/dahu.php.

The need to access and exchange data on the Web has led to database management systems (DBMS) that are increasingly distributed and autonomous. Data extraction and querying on the Web is harder than in classical DBMS, because such data is heterogeneous, redundant, inconsistent and subject to frequent modifications. DBMS thus need to be able to detect errors, to analyze them and to correct them. Moreover, increasingly complex Web applications and services rely on DBMS, and their reliability is crucial. This creates a need for tools for specifying DBMS in a high-level manner that is easier to understand, while also facilitating verification of critical properties.

The study of such specification and verification techniques is the main goal of Dahu.

3. Scientific Foundations

3.1. Scientific Foundations

Dahu has strong connections with the Leo project-team in Saclay, the Mostrare project-team in Lille and the Cassis project-team in Nancy.

Dahu aims at developing mechanisms for high-level specifications of systems built around DBMS, that are easy to understand while also facilitating verification of critical properties. This requires developing tools that are suitable for reasoning about systems that manipulate data. Some tools for specifying and reasoning about data have already been studied independently by the database community and by the verification community, with various motivations. However, this work is still in its infancy and needs to be further developed and unified.

Most current proposals for reasoning about DBMS over XML documents are based on tree automata, taking advantage of the tree structure of XML documents. For this reason, the Dahu team is studying a variety of tree automata. This ranges from restrictions of "classical" tree automata in order to understand their expressive power, to extensions of tree automata in order to understand how to incorporate the manipulation of data.

Moreover, Dahu is also interested in logical frameworks that explicitly refer to data. Such logical frameworks can be used as high level declarative languages for specifying integrity constraints, format change during data exchange, web service functionalities and so on. Moreover, the same logical frameworks can be used to express the critical properties we wish to verify.

In order to achieve its goals, Dahu brings together world-class expertise in both databases and verification.

4. Application Domains

4.1. Application Domains

Databases are pervasive across many application fields. Indeed, most human activities today require some form of data management. In particular, all applications involving the processing of large amounts of data require the use of a database. Increasingly complex Web applications and services also rely on DBMS, and their correctness and robustness is crucial.

We believe that the automated solutions that Dahu aims to develop for verifying such systems will be useful in this context.

5. New Results

5.1. XML specification and verification

Participants: Serge Abiteboul, Pierre Bourhis, Florent Jacquemard, Yannis Katsis, Diego Figueira, Bruno Marnette, Luc Segoufin, Cristina Sirangelo, Victor Vianu.

We have considered this year several scenarios. In all of them the flow of data is the key feature. These scenarios range from data evolving with time to static data analysis. Our tools for specification and verification are also very broad and ranges from probabilistic methods to automata theoretical methods.

5.1.1. Static analysis of query languages

XPath is arguably the most widely used XML query language as it is implemented in XSLT and XQuery and it is used as a constituent part of several specification and update languages. Hence in order to perform static analysis on a system manipulating XML data it is important to master the static analysis for XPath. Most of the important static analysis problems reduce to satisfiability checking: does a given query return a non-empty answer on some data. In general the satisfiability of XPath is undecidable, however important fragments can be shown to be decidable. In [34] we have shown that when restricting the navigational axis of XPath to the child, descendant, next sibling and following sibling relations, then satisfiability can be decided. However, its complexity is very high: non primitive recursive.

The closeness of XML navigation languages, such as XPath, to temporal logics (LTL over strings and its analog over trees) suggests that many XML static analysis tasks can also take advantage of techniques used in the field of verification, where the main focus is on the verification of temporal properties over graph-like structures. In [19] we adopt this approach and develop effective automata-based techniques for XPath static analysis. These techniques are based on a two step effective translation of XPath into automata. In particular we provide an explicit translation from XPath to a temporal logic over trees, followed by a translation from temporal logics to automata that follows the lines of classical temporal logic-to-automata translations used in verification. The automata associated to XPath formulae have exponential size, but can be constructed efficiently using the many optimization strategies currently in use in existing verification tools.

5.1.2. Incomplete information

Distribution of data comes along with the need to integrate and/or exchange XML documents conforming to heterogeneous schemas. This usually generates incomplete information in XML, due to the incompatibility of schemas and constraints among different sites. Querying an incomplete (partially specified) document generally means computing certain answers, i.e. answers that would be obtained by querying any valid completion of the incomplete document. This problem is closely related to some questions in classical pattern matching theory. In fact an incomplete document can be viewed as a pattern, and complete documents as structures where the pattern has a match. In [18] we investigated the pattern matching and pattern implication problems under a new matching semantics: each element of the pattern has to match a distinct position. Under the classical matching semantics, checking whether a set of patterns occurs in a string is solvable in polynomial time, while checking pattern implication (i.e. whether a set of patterns implies the presence of another fixed pattern) is well-known to be intractable. In [18] we showed that for the disjoint version of the problem the situation is reversed: checking whether a set of patterns occurs in a string is NP-complete, but the implication problem is solvable in polynomial time.

5.1.3. Data centric Workflows

There has recently been a proliferation of workflow specification formalisms, notably data-centric, in response to the need to support increasingly ubiquitous processes centered around databases. Prominent examples include e-commerce systems, enterprise business processes, health-care and scientific workflows. Developing suitable specification mechanisms for such workflows, understanding the expressive power of the different specification formalisms, and performing static analysis are critical issues that we address in some of our recent work [25], [30]. A prototype, named AXART, based on a model presented in [25] has been demonstrated at the VLDB 2010 conference [24].

A powerful framework for specifying and studying data-centric workflows is provided by Active XML, a high-level specification language tailored to data-intensive, distributed, dynamic Web-services. In brief, Active XML consists of XML documents with embedded function calls. The state of a document evolves depending on the result of internal function calls (local computations) or external ones (interactions with users or other services). Functions can be naturally used to model tasks in a workflow. They return documents that may be active, so may in turn activate new sub-tasks, thus having the ability to naturally specify a hierarchy of tasks. We have previously studied the verification of temporal properties of runs of Active XML systems, specified in a tree-pattern based temporal logic, Tree-LTL, that allows expressing a rich class of semantic properties of the workflow. The main results establish the boundary of decidability and the complexity of automatic verification of Tree-LTL properties. In our most recent work [25] we focus on *comparing* the specification power of various workflow control mechanisms within the Active XML framework and beyond. This is intrinsically difficult because of the lack of a standard yardstick for expressiveness. In [25], we develop a flexible framework for comparing workflow specification languages, in which the pertinent aspects to be taken into account are defined by *views*. We use it to compare the expressiveness of several workflow specification mechanisms based on automata, pre/post conditions, and temporal constraints.

The AXART system extends the model studied in [25] with the ability to have an explicit hierarchy of distributed tasks with associated control. A demonstration of the system was presented in [24], based on an example taken from the movie industry, that specifies the workflow involved in applying for a role in a film.

Another prominent formalism for specifying data-centric workflows is IBM's Business Artifacts. In collaboration with UC San Diego and IBM, we studied in previous work the verification problem for a restricted class of such workflows. However, the early results suffer from an important limitation: they fail in the presence of even very simple data dependencies or arithmetic, both crucial to real-life business processes. In [30], we extend the artifact model and verification results to alleviate this limitation. We identify a practically significant class of business artifacts with data dependencies and arithmetic, for which verification is decidable.

5.1.4. Analysis of Distributed Data Management

With the evolution of the Web and the emergence of universal standards for data exchange, data management is becoming increasingly distributed. As a first step in setting theoretical foundations for distributed data management systems, we study the equivalence of such systems [27]. To model systems, we use the Active XML framework. The resulting model is expressive enough to capture distributed systems that are recursive, utilize asynchronous communication and operate on data streams.

As our model is quite general, the equivalence problem is undecidable. We exhibit restrictions of the model, in terms of query languages used in the function calls and the presence or absence of external function calls, for which equivalence can be effectively decided. We study the computational complexity of the equivalence problem, and for a limited class of distributed systems present the axiomatization of equivalence.

5.1.5. Verification of XML updates

XQuery language has been extended to XQuery Update Facility (a W3C Candidate Recommendation of 2009) in order to provide convenient means of modifying XML documents or data. The language is a candidate recommendation from W3C and adds imperative operations that permit one e.g. to update some parts of a document while leaving the rest unchanged. This includes rename, insert, replace and delete operations at the node level. Compared to other transformation languages (such a XSLT), XQuery Update Facility is considered to offer concise, readable solutions.

In [40] we propose a model for the XML update primitives of the W3C XQuery Update Facility as parametrized rewriting rules of the form: "insert an unranked tree from a regular tree language L as the first child of a node labeled by a". We give type inference algorithms for the iteration of such rules, considering types defined by several classes of unranked tree automata. These type inference algorithms are directly applicable to XML static typechecking, which is the problem of verifying whether, a given document transformation always converts source documents of a given input type into documents of a given output type. Typechecking, for arbitrary sequences of XML update primitives, can be done in polynomial time when the unranked tree automaton defining the output type is deterministic and complete, and that it is EXPTIME-complete otherwise.

The results are applied to the verification of access control policies for XML updates. We propose in particular a polynomial time algorithm for the problem of local consistency of a policy, that is, for deciding the non-existence of a sequence of authorized update operations starting from a given document that simulates a forbidden update operation.

5.1.6. Probabilistic XML

Uncertain data are common on the Web. We study query processing for probabilistic trees (i.e., probabilistic XML), and in particular, the evaluation of aggregate queries. In [26], we provide a systematic study of the complexity of aggregate queries defined with tree-pattern queries with joins (and restrictions of that query language). We show that when only local probabilistic dependencies are considered, some particular aggregate functions (namely, *monoid* ones) are tractable. We also provide a formal model for continuous distributions in probabilistic trees and show that they essentially do not add any complexity to querying tasks, as long as a number of operations (convex sums, differentiation, integration, convolution) can be tractably performed over these distributions, either symbolically or numerically.

5.2. Automata and logics for data words and data trees

Participants: Stéphane Demri, Diego Figueira, Luc Segoufin.

Dahu aim as providing tools for specifying and verifying systems with data. This means finding a suitable logical framework for specifying such systems. A logical framework is suitable if it is expressible enough for modeling the operations of interest. Of course, for the logical framework to be useful, it must come with techniques and tools for reasoning about it, in particular it should be decidable. This can be achieved by compiling the model into some form of decidable automata manipulating data. In the presence of data, the design of appropriate classes of logic and automata with interesting complexities is an on-going research task.

Most of our new results in this direction concerns data words and data trees. Those are words and trees where each position contains a data value together with the classical label. Data words and data trees can model many systems with data with a focus on one variable flow. Data trees can also model XML data.

We have studied several extensions of the classical model of logic and automata with features that could be used for manipulating data. This is done either by using registers or memory explicitly in the model or by restricting the transitions of the automata with constraints that can involve data comparisons. Several models have been considered.

A known automata model for data words is *register automata*. In [35] we show that this automaton is closely related to a widely studied model on timed words: the *timed automaton*. Indeed we show that a run of a timed automaton can be simulated by a register automaton, and conversely that a run of a register automaton can be simulated by a timed automaton. These results allow to transfer complexity and decidability results back and forth between these two kinds of models.

As another line of investigation, we studied a top-down forward alternating 1-register automata, that can move downwards and rightwards on an unranked data tree. This model is extended with two powerful operations in [34]. This new model of automaton is shown to capture XPath with child, descendant, right-sibling and following-sibling axes. The model is decidable and hence decidability of the XPath fragment follows.

In [33], we have studied the decidability status of model-checking freeze LTL over various subclasses of counter machines for which the reachability problem is known to be decidable (reversal-bounded counter machines, vector addition systems with states, flat counter machines, one-counter machines). Runs generated from counter machines are understood as data words whose finite alphabet is the set of control states. The main shift when considering model-checking (instead of satisfiability) rests on the fact that more constraints apply to the data words (those from the counter machines). In freeze LTL, a register can store a counter value and at some future position an equality test can be done between a register and a counter value. In [33], we have completed an earlier work started on one-counter machines by considering other subclasses of counter machines, and especially the class of reversal-bounded counter machines. This gives us the opportunity to provide a systematic classification that distinguishes determinism vs. nondeterminism and we consider subclasses of formulae by restricting the set of atomic formulae or/and the polarity of the occurrences of the freeze operators, leading to the flat fragment.

5.3. Automata theory

Participants: Florent Jacquemard, Thomas Place, Luc Segoufin, Camille Vacher.

The links between models for XML and regular tree languages has been advocated in many places. Tree automata seem to be playing for semi-structured data and XML the role of the relational algebra for relational databases. As XML is central in our research we also study tree automata and regular tree languages.

In [21] we have studied the links between XPath and extensions of Tree Walking Automata. In particular we have exhibited a robust class of tree languages lying strictly between the regular tree languages and the tree languages expressible in First-Order and show that it correspond naturally to a natural extension of XPath.

Another line of research concerns the expressive power of various subclasses of regular tree languages. It is usually admitted that a fragment is completely understood, in term of expressive power, when one has a *decidable characterization* of it. That is an algorithm that given a regular tree language, presented say as a tree automata, tests whether it belongs to the class being investigated or not. This question is an active research topic that turns out to be quite challenging. We have considered the fragment of XPath without the successor

axis. In term of expressive power it is equivalent to the two-variable fragment of First-Order logic. In [41] we have shown that it is decidable whether a regular tree language is expressible in this fragment of XPath. We have also studied another fragment of First-Order logic restricting the quantification prefix. In [14] we have shown that it is decidable whether a regular tree language is expressible in First-Order logic with one quantifier alternation. In the same spirit, we also show in [36], that it is decidable whether a regular tree language of infinite trees is definable by a μ-formula.

We have also considered superclasses of regular tree languages, described by tree automata with features which extend strictly standard tree automata.

This is the case of Rigid Tree Automata (RTA), an extension of standard bottom-up tree automata with distinguished states called rigid. Rigid states define a restriction on the computation of RTA on trees: RTA tests for equality of subtrees reaching the same rigid state. In [17], we study the expressiveness of these automata properties like determinism, pumping lemma, Boolean closure, and several decision problems, and an application to the verification of security protocols. Our main contribution in this setting is a proof of decidability for the problem to know whether a given tree belongs to the rewrite closure of a RTA language under a restricted family of term rewriting systems, whereas this closure is not a RTA language. To our knowledge, this is the first result of this kind for tree automata with equality tests.

In [29], we consider another class of tree automata with global constraints much more general than RTA. These automata, called TAGC, generalize the class of tree automata with global equality and disequality constraints (TAGED) introduced by Filiot, Talbot, Tison in 2008. TAGC can test for equality and disequality between subterms whose positions are defined by the states reached during a computation. In particular, TAGC can check that all the subterms reaching a given state are distinct. This constraint is related to monadic key constraints for XML documents, meaning that every two distinct positions of a given type have different values.

The emptiness problem is proven decidable for TAGC [29]. This solves, in particular, the open question of decidability of emptiness for TAGED. This result is further extended by allowing global arithmetic constraints for counting the number of occurrences of some state or the number of different subterms reaching some state during a computation. We also allow local equality and disequality tests between sibling positions and the extension to unranked ordered trees. As a consequence of our results for TAGC, we prove the decidability of a fragment of the monadic second order logic on trees extended with predicates for equality and disequality between subtrees, and cardinality.

6. Other Grants and Activities

6.1. National collaborations

Dahu is currently participating in one ANR project:

ENUM is a research project supported by the ANR blanche (2007-2011) on algorithmic and complexity problems raised by enumerating solutions of a query. The goal is to provide formal methods to understand and compare the complexity of enumerations problems. The partners are University of Paris-7 (with Arnaud Durand), the project-team Mostrare at INRIA-Lille (with Joachim Niehren), the university of Caen (with Etienne Grandjean) and the university of Marseille (with Nadia Creignou). Dahu is involved in the ANR as part of the Paris-7 node. For more information please visit the web pages of ENUM:https://gforge.inria.fr/plugins/wiki/index.php?EnumerationProject&id=267&type=g.

Dahu is also the coordinator of one ARC INRIA

ACCESS is an ARC INRIA on Access Control for Web data, a two years project starting in 2010. The goal of this project is to study security and access control techniques for Web data exchange, and in particular the problems of the verification of access control policies (ACP) for XML and of the

enforcement of ACP for secure query evaluation and update propagation. As a case study, the results are applied to an XML-based collaborative editing system. The partners are the teams CASSIS and PAREO at the INRIA center of Nancy-Grand-Est, and the team MOSTRARE at the the INRIA center of Lille-Nord-Europe. For more information please visit the web pages:http://acxml.gforge.inria.fr.

6.2. International collaborations

6.2.1. Cooperation within Europe

Dahu is involved into two major grants funded by EU:

FOX FoX is a FET-Open project funded within the FP7 framework. The objective of FoX is to study the fundamental issues necessary in order to make the data management over the internet more efficient and more reliable. The partners of Dahu in FoX are Thomas Schwentick at the university of Dortmund, Mikołaj Bojańczyk at the university of Warsaw, Leonid Libkin at the university of Edinburgh, Georg Gottlob at the university of Oxford, Frank Neven at the university of Hasselt and Maarten Marx at the university of Amsterdam. The project started on May 1st 2009 and will last until April 30th 2012. Luc Segoufin is the coordinator of this project. For more info, please visit the web pages of FoX: http://fox7.eu.

Webdam Webdam is an ERC "Advanced investigators grants" obtained by Serge Abiteboul. It started in December 2008. The goal is to develop a formal model for Web data management. This model will open new horizons for the development of the Web in a well-principled way, enhancing its functionality, performance, and reliability. Specifically, the goal is to develop a universally accepted formal framework for describing complex and flexible interacting Web applications featuring notably data exchange, sharing, integration, querying and updating. We also propose to develop formal foundations that will enable peers to concurrently reason about global data management activities, cooperate in solving specific tasks and support services with desired quality of service. For more info, please visit the web pages of Webdam: http://webdam.inria.fr.

The Webdam project is shared between the Dahu and Leo project-teams, both from INRIA Saclay.

6.2.2. Cooperation with Tunisia

Dahu is coordinator (on the French side) of the project INRIA-DGRSRT (Tunisian universities) 10/I01 "**BRICK**" on the verification of security properties of Web services, access control policies and firewalls for XML. This project started in 2010, the other partners are the CASSIS team at INRIA Nancy-Grand-Est and the Security team at Sup'Com Tunis.

This year, DAHU has hosted in November and December the internship of Ryma Abassi (PhD in Supcom Tunis), on the problem of the type checking of W3C XQuery updates.

6.2.3. Cooperation with North America

Close links also exist with UC San Diego and the database group of Victor Vianu.

7. Dissemination

7.1. Thesis

Diego Figueira, now post-doc in Edinburgh, defended his thesis "Reasoning on words and trees with data" [11] on December 6th, 2010. The thesis was supervised by Luc Segoufin and Stéphane Demri.

Thomas Place, now post-doc in Warsaw, defended his thesis "Decidable Characterizations for Tree Logics" on December 10, 2010 [12]. The thesis was supervised by Luc Segoufin.

Camille Vacher, now ATER in Lille, defended his thesis "Tree automata global with constraints and the verification of security properties" on December 7, 2010 [13]. The thesis was supervised by Florent Jacquemard.

7.2. Participation to conferences organization

Luc Segoufin was the PC-chair of the Intl. Conf. on Database Theory (ICDT'10) that was held in Lausanne in March 2010.
Serge Abiteboul is General Program Chair of the International Conference on Data Engineering that will be in Hannover, Germany, in 2011.

Several members of the project have participated in program committees:

- Stéphane Demri: 17th International Symposium on Temporal Representation and Reasoning" (TIME), September 2010 (and member of the organization committee); 4th International Workshop on Reachability Problems (RP), August 2010; Conference on Advances of Modal Logic (AIML), August 2010; FLOC'10 Workshop on Comparing Logical Decision Methods (CLODEM), July 2010; FLOC'2010 Workshop on Logics for System Analysis (LfSA), July 2010; FLOC'2010 Workshop on Hybrid Logic and Applications (HYLO), July 2010.
- Luc Segoufin: Intl. Conf. on Database Theory (ICDT'10).
- Cristina Sirangelo: International Conference on Database Theory (ICDT '10).

In October 2010, Florent Jacquemard co-organized the first workshop on Formal Methods for Web Data Trust and Security (WTS).

Serge Abiteboul co-organized with Andreas Oberweis (KIT, Germany) and Jianwen Su (UCSB, USA) the Dagstuhl Workshop on Enabling Holistic Approaches to Business Process Lifecycle Management (04/2010) [43]

Pierre Senellart (Telecom ParisTech) and Serge Abiteboul organized the ACM SIGMOD 2010 programming contest.

7.3. Participation to symposia, seminars, invitations

Besides the presentations of our papers accepted to international conferences, the members of Dahu made the following keynote talks to international conferences or workshops.

Serge Abiteboul gave a keynote presentation on Web information management and knowledge bases at the 10th International Conference on Web Engineering, Vienna (07/2010); an invited presentation on Web data management at the Datalog 2.0 Workshop, held in March 2010, at Oxford University [23]; and an invited presentation on Object Databases at the Dagstuhl workshop on Relationships, Objects, Roles, and Queries in Modern Programming Languages (04/2010).

Stéphane Demri gave an invited presentation at JELIA'10 [31].

7.4. Scientific Animations

- Stéphane Demri is a member of the steering committee of the Tableaux conference (Intl. Conf. Automated Theorem Proving with Analytic Tableaux and Related Methods).
- Stéphane Demri is member of the publication board of the review "Technique et Science Informatiques".
- Florent Jacquemard is member of the board (general secretary) of the French Association for Information and Communication Systems (ASTI).
- Luc Segoufin is a member of the steering committee of the Intl. Conf. on Database Theory (ICDT).
- Luc Segoufin is responsible of the groupe de travail "Complexité et Modèles Finis" du GDR "Mathématique et Informatique"(http://www.gdr-im.fr/).

- Cristina Sirangelo is publicity chair for the International Conference on Database Theory (ICDT '10).

7.5. Teaching

As a Maître de conférence Cristina Sirangelo is teaching in the department of computer science of ENS de Cachan. Thomas Place and Camille Vacher are also teaching in this department as part of their allocation couplée. Serge Abiteboul is also teaching the database course in this department.

Dahu is very much involved in the Master Parisien de Recherche en Informatique (MPRI): Stéphane Demri is a participant of the M2 course 2.9, 2010–2011 (12h30), Florent Jacquemard is teaching in the first year (M1), a course on Tree Automata Techniques and Applications and Luc Segoufin is teaching in the second year (M2) a course on "Logic, descriptive complexity and theory of databases".

Stéphane Demri is "délégué aux thèses" in Computer Science for the Ecole Doctorale Sciences Pratiques (EDSP), ENS Cachan. In the master, MICR, U. de Paris-Nord, Stéphane Demri has given a course on counter systems, M2, 2010 (15h). He has been lecturer at summer school ESSLLI'2010 for the course "Decidable problems on counter systems", Copenhagen, 2010 and he has also given lectures on "Counter Systems and Temporal Logics" (12h) at the University of Buenos Aires and at the National University of Cordoba. Besides, he has been member of entrance exams committee for the Ecole Normale Supérieure(s) (ENS), 2010.

7.6. Thesis jury

Stéphane Demri was reviewer for the theses of Ida Sri Rejeki Siahaan (U. of Trento), Marco Volpe (U. of Verona) Guillaume Hoffmann (U. Henri Poincaré, Nancy 1) and was examiner for the theses of Dimitri Sustretov (U. Henri Poincaré, Nancy 1) and Yakoub Salhi (U. Henri Poincaré, Nancy 1).

8. Bibliography

Major publications by the team in recent years

[1] S. ABITEBOUL, L. SEGOUFIN, V. VIANU. *Static Analysis of Active XML Systems*, in "ACM Transactions on Database Systems", 2009, vol. 34, n^{o} 4.

[2] M. BENEDIKT, L. SEGOUFIN. *Regular tree languages definable in FO and in FO_{mod}*, in "ACM Transactions on Computational Logic", 2009, vol. 11, n^{o} 1.

[3] M. BENEDIKT, L. SEGOUFIN. *Towards a Characterization of Order-Invariant Queries over Tame Structures*, in "Journal of Symbolic Logic", 2009, vol. 74, n^{o} 1, p. 168-186.

[4] M. BOJAŃCZYK, A. MUSCHOLL, TH. SCHWENTICK, L. SEGOUFIN. *Two-variable logic on data trees and applications to XML reasoning*, in "Journal of the ACM", 2009, vol. 56, n^{o} 3.

[5] R. BROCHENIN, S. DEMRI, É. LOZES. *Reasoning about sequences of memory states*, in "Annals of Pure and Applied Logics", 2009, vol. 161, n^{o} 3, p. 305-323.

[6] B. CAUTIS, S. ABITEBOUL, T. MILO. *Reasoning about XML update constraints*, in "Journal of Computer and System Sciences", 2009, vol. 75, n^{o} 6, p. 336-358.

[7] S. DEMRI, R. GASCON. *The Effects of Bounding Syntactic Resources on Presburger LTL*, in "Journal of Logic and Computation", 2009, vol. 19, n^{o} 6, p. 1541-1575.

[8] S. DEMRI, R. LAZIĆ. *LTL with the freeze quantifier and register automata*, in "ACM Transactions on Computational Logic", 2009, vol. 10, n° 3.

[9] L. LIBKIN, C. SIRANGELO. *Reasoning about XML with temporal logics and automata*, in "Journal of Applied Logic", 2010, vol. 8, n° 2, p. 210-232, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/LS-jal10.pdf.

[10] B. TEN CATE, L. SEGOUFIN. *Transitive Closure Logic, Nested Tree Walking Automata, and XPath*, in "Journal of the ACM", 2010, vol. 57, n° 3.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] D. FIGUEIRA. *Reasoning on words and trees with data*, ENS Cachan, December 2010.

[12] T. PLACE. *Decidable Characterizations for Tree Logics*, ENS Cachan, December 2010.

[13] C. VACHER. *Tree automata global with constraints and the verification of security properties*, ENS Cachan, December 2010.

Articles in International Peer-Reviewed Journal

[14] M. BOJAŃCZYK, L. SEGOUFIN. *Tree Languages Defined in First-Order Logic with One Quantifier Alternation*, in "Logical Methods in Computer Science", 2010, vol. 6, n° 4.

[15] S. DEMRI, R. LAZIĆ, A. SANGNIER. *Model checking memoryful linear-time logics over one-counter automata*, in "Theoretical Computer Science", 2010, vol. 411, n° 22–24, p. 2298–2316.

[16] S. DEMRI, D. LUGIEZ. *Complexity of Modal Logics with Presburger Constraints*, in "Journal of Applied Logic", 2010, vol. 8, n° 3, p. 233–252.

[17] F. JACQUEMARD, F. KLAY, C. VACHER. *Rigid Tree Automata and Applications*, in "Information and Computation", 2010, In press.

[18] L. LIBKIN, C. SIRANGELO. *Disjoint pattern matching and implication in strings*, in "Information Processing Letters", 2010, vol. 110, n° 4, p. 143–147.

[19] L. LIBKIN, C. SIRANGELO. *Reasoning about XML with temporal logics and automata*, in "Journal of Applied Logic", 2010, vol. 8, n° 2, p. 210–232.

[20] A. NASH, L. SEGOUFIN, V. VIANU. *Views and queries: Determinacy and rewriting*, in "ACM Trans. Database Systems", 2010, vol. 35, n° 3.

[21] B. TEN CATE, L. SEGOUFIN. *Transitive closure logic, nested tree walking automata, and XPath*, in "Journal of ACM", 2010, vol. 57, n° 3.

International Peer-Reviewed Conference/Proceedings

[22] R. ABASSI, F. JACQUEMARD, M. RUSINOWITCH, S. GUEMARA EL FATMI. *XML Access Control: from XACML to Annotated Schemas*, in "Proceedings of the 2nd International Conference on Communications and Networking (ComNet'10)", Tozeur, Tunisia, IEEE Computer Society Press, November 2010, To appear, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/AJRG-comnet10.pdf.

[23] S. ABITEBOUL, M. BIENVENU, A. GALLAND, M.-C. ROUSSET. *Distributed Datalog Revisited*, in "Datalog 2.0 Workshop", Oxford Royaume-Uni, 2011, http://hal.inria.fr/inria-00540814/PDF/datalog20-serge.pdf.

[24] S. ABITEBOUL, P. BOURHIS, B. MARINOIU, A. GALLAND. *AXART - Enabling Collaborative Work with AXML Artifacts*, in "VLDB, demonstration", Singapour Singapour, 2010, http://hal.inria.fr/inria-00521726/PDF/D10.pdf.

[25] S. ABITEBOUL, P. BOURHIS, V. VIANU. *Comparing Workflow Specification Languages: A Matter of Views*, in "Proc. of Intl. Conf. on Database Theory (ICDT)", 2011, http://hal.inria.fr/inria-00539579/en/.

[26] S. ABITEBOUL, T.-H. H. CHAN, E. KHARLAMOV, W. NUTT, P. SENELLART. *Aggregate Queries for Discrete and Continuous Probabilistic XML*, in "Proc. of Intl. Conf. on Database Theory (ICDT)", 2010, p. 50-61, http://hal.inria.fr/inria-00537632/PDF/main.pdf.

[27] S. ABITEBOUL, B. TEN CATE, Y. KATSIS. *On the Equivalence of Distributed Systems with Queries and Communication*, in "Proc. of Intl. Conf. on Database Theory (ICDT)", 2011, http://hal.inria.fr/inria-00540199/en/.

[28] E. ANTOINE, A. GALLAND, K. LYNGBAEK, A. MARIAN, N. POLYZOTIS. *Social Networking on top of the WebdamExchange System*, in "Proc. of Intl. Conf. on Data Engineering (ICDE)", 2011, http://hal.inria.fr/inria-00536361/PDF/ICDE11_Demo_840.pdf.

[29] L. BARGUÑÓ, C. CREUS, G. GODOY, F. JACQUEMARD, C. VACHER. *The Emptiness Problem for Tree Automata with Global Constraints*, in "Proc. IEEE Conf. on Logic in Computer Science (LICS)", Edinburgh, Scotland, UK, IEEE Computer Society Press, July 2010, p. 263-272.

[30] E. DAMAGGIO, A. DEUTSCH, V. VIANU. *Artifact Systems with Data Dependencies and Arithmetic Constraints*, in "Proc. of Intl. Conf. on Database Theory (ICDT)", 2011.

[31] S. DEMRI. *Counter Systems for Data Logics*, in "Journées Européennes sur la Logique en Intelligence Artificielle (JELIA)", 2010, 10.

[32] S. DEMRI. *On Selective Unboundedness of VASS*, in "Proceedings of the 12th International Workshop on Verification of Infinite State Systems (INFINITY'10)", Y.-F. CHEN, A. REZINE (editors), Electronic Proceedings in Theoretical Computer Science, 2010, vol. 39.

[33] S. DEMRI, A. SANGNIER. *When Model checking freeze LTL over Counter Machines Becomes Decidable*, in "Foundations of Software Science and Computation Structures (FOSSACS)", 2010, p. 16–190.

[34] D. FIGUEIRA. *Forward-XPath and extended register automata on data-trees*, in "Proc. of Intl. Conf. on Database Theory (ICDT)", ACM Press, 2010, p. 230-240, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/fig-icdt10.pdf.

[35] D. FIGUEIRA, P. HOFMAN, S. LASOTA. *Relating timed and register automata*, in "Proceedings of the 17th International Workshop on Expressiveness in Concurrency (EXPRESS'10)", 2010, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/FHL-express10.pdf.

[36] G. FONTAINE, TH. PLACE. *Classes of trees definable in the μ-calculus*, in "Proceedings of the 35th International Symposium on Mathematical Foundations of Computer Science (MFCS'10)", Brno, Czech Republic, Springer, August 2010, vol. 6281, p. 381-392.

[37] A. GALLAND, S. ABITEBOUL, A. MARIAN, P. SENELLART. *Corroborating Information from Disagreeing Views*, in "International Conference on Web Search and Data Mining (WSDM)", 2010, http://hal.inria.fr/inria-00429546.

[38] A. GALLAND, S. ABITEBOUL, A. MARIAN, P. SENELLART. *Corroborating Information from Disagreeing Views*, in "International Conference on Web Search and Data Mining (WSDM)", 2010, http://hal.inria.fr/inria-00429546/PDF/document.pdf.

[39] F. JACQUEMARD, M. RUSINOWITCH. *Rewrite-based verification of XML updates*, in "12th International ACM SIGPLAN Conference on Principles and Practice of Declarative Programming", Autriche Hagenberg, ACM, Jul 2010, p. 119-130, http://hal.inria.fr/inria-00529620.

[40] F. JACQUEMARD, M. RUSINOWITCH. *Rewrite-based verification of XML updates*, in "Proceedings of the 12th International ACM SIGPLAN Conference on Principles and Practice of Declarative Programming (PPDP'10)", Hagenberg, Austria, T. KUTSIA, W. SCHREINER, M. FERNÁNDEZ (editors), ACM Press, July 2010, p. 119-130.

[41] T. PLACE, L. SEGOUFIN. *Deciding Definability in $FO_2(<_h, <_v)$ on Trees*, in "Proc. IEEE Conf. on Logic in Computer Science (LICS)", 2010, p. 253-262.

[42] N. SCHWEIKARDT, L. SEGOUFIN. *Addition-Invariant FO and Regularity*, in "Proc. IEEE Conf. on Logic in Computer Science (LICS)", 2010, p. 273-282.

Other Publications

[43] S. ABITEBOUL, A. KOSCHMIDER, A. OBERWEIS, J. SU. *Proceedings of the Dagstuhl Seminar 10151 "Enabling Holistic Approaches to Business Process Lifecycle Management"*, 2010.