Activity Report 2014

Project-Team TYREX

Types and Reasoning for the Web

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

Table of contents

Project-Team TYREX

Keywords: Formal Methods, Programming Languages, Multimedia, Web, Augmented Reality

Creation of the Team: 2012 November 01, updated into Project-Team: 2014 July 01.

1. Members

Research Scientists
Nabil Layaïda [Team leader, Inria, Senior Researcher, HdR]
Pierre Genevès [CNRS, Researcher, HdR]

Faculty Members
Nils Gesbert [INP Grenoble, Associate Professor]
Jacques Lemordant [Univ. Grenoble I, Associate Professor, until Jun 2014]
Cécile Roisin [Univ. Grenoble II, Professor, HdR]

Engineers
Guillaume Dupraz-Canard [Univ. Grenoble I]
Nicolas Hairon [Inria, Caisse des Dépôts et Consignations, until Apr 2014]
Mathieu Razafimahazo [Inria, FP7 VENTURI project]

PhD Students
Abdullah Abbas [Univ. Grenoble I, from Feb 2014]
Damien Graux [Univ. Grenoble I]
Nicola Guido [Inria]
Louis Jachiet [ENS Paris]
Muhammad Junedi [Univ. Grenoble I, until Sep 2014]
Thibaud Michel [Univ. Grenoble I]

Post-Doctoral Fellow
Hyeonseung Im [Inria, ANR Typex, from May 2014]

Visiting Scientist
Joel Ferreira Dos Santos [Universidade Federal Fluminense Brazil, from Oct 2014]

Administrative Assistant
Marion Ponsot [Inria]

Others
Marti Bosch Padros [Inria, Intern, from Feb 2014 until Jul 2014]
Massinissa Boudraham [Inria, Intern, from Jun 2014 until Aug 2014]

2. Overall Objectives

2.1. Objectives

The Tyrex team aims at developing a vision of a web where content is enhanced and protected, applications made easier to build, maintain and secure. It seeks to open new horizons for the development of the web, enhancing its potential, effectiveness, and dependability. In particular, we aim at making contributions by obtaining fundamental results, by building advanced experimental applications showcasing these results and by contributing to web standards. One fundamental problem of our time is a lack of formalisms, concepts and tools for reasoning simultaneously about content or data, programs and communication aspects. One scientific goal is to establish a unifying framework for designing advanced (robust, flexible, rich, efficient and novel) web applications. We address this objective along two complementary research directions: (i) modeling, and (ii) analysis for verification and optimisation. We further describe these directions in the next section. The overall goal is to enable richer, more reliable, secure, and efficient systems.

3. Research Program

3.1. Modeling

Modeling consists in capturing various aspects of document and data processing and communication in a unifying model. Our modeling research direction mainly focuses on three aspects.

The first aspect aims at reducing the impedance mismatch. The impedance mismatch refers to the complexity, difficulty and lack of performance induced by various web application layers which require the same piece of information to be represented and processed differently. The mismatch occurs because programming languages use different native data models from those used for documents in browsers and for storage in databases. This results in complex and multi-tier software architectures whose different layers are incompatible in nature. This, in turn, results in expensive, inefficient, and error-prone web development. For reducing the impedance mismatch, we will focus on the design of a unifying software stack and programming framework, backed by generic and solid logical foundations similar in spirit to the NoSQL approach.

The second aspect aims at harnessing heterogeneity. Web applications increasingly use diverse data models: ordered and unordered tree-like structures (such as XML), nested records and arrays (such as JSON), graphs like (e.g. RDF), and tables. Furthermore, these data models also involve a variety of languages for expressing constraints over data (e.g. XML schema, the well-founded RelaxNG, and RDFS to name just a few). We believe that this heterogeneity is here to stay and is likely to increase. These differences in representations imply loads of error-prone and costly conversions and transformations. Furthermore, some native formats (e.g. JSON) are diverted from a programming construct to a data exchange one. This often results in a loss of information and in errors that need to be tracked and corrected. In this context, it is important to seek methods for reducing risks of information loss during data transformation and exchange. For harnessing heterogeneity, we will focus on the integration of data models through unified formal semantics and in particular logical interpretation. This allows using the same programming language constructs on different data models. At the programming language level, this is similar to languages such as JSonIq for JSON and XML.

Finally, the third aspect aims at making applications and data more compositional. Most web programming technologies are currently limited from a compositional point of view. For example, tree grammars (like schema languages for XML) are monolithic in the sense that they require the full description of the considered structures, instead of allowing the assembly of smaller and reusable building blocks. More generally, this need is illustrated in the industry by the increasing development of W3C specifications organised in ad-hoc modules. So far, these various attempts have failed to provide an acceptable mechanism for composition. For example, HTML5 has been specified in a monolithic way despite the fact that it relies on several other existing specifications (such as HTML, SVG, SMIL, CSS, etc.). As a consequence, this translates into monolithic web applications, which makes their automated verification harder by making modular analyses more difficult. For making applications and data more compositional, we will focus on the design of modular schema and programming languages. For this purpose, we will notably rely on succinct yet expressive formalisms (like two-way logics, polymorphic types) that ease the process of expressing modular specifications.

One major scientific difficulty in this overall direction consists in taking into account the specificities of the web, which require new programming models and supporting theoretical tools that do not exist today.

3.2. Analysis, verification and optimization

This research direction aims at guaranteeing two different kinds of properties: safety and efficiency.

The first kind of properties concern safety of web applications. Software development was traditionally split between critical and non-critical software. Advanced (and costly) formal verification techniques were reserved to the former whereas non-critical software relied almost exclusively on testing, which only offers a "best-effort" guarantee (removes most bugs but some of them may not be detected). The central idea was that in a non-critical system, the damage a failure may create is not worth the cost of formal verification. However as web applications grow more pervasive in everyday life and gain momentum in corporates, various

social organizations, and touch larger numbers of users, the potential cost of failure is increasing rapidly and significantly. Despite this fact, it is more obvious, in healthcare for instance, to qualify as a critical component a pacemaker than the hospital's information system. Of course, a failure of such a device would directly cause death, however a general failure of the hospital's information system may cause deaths as well and possibly even incur greater damages. In that sense, we can consider that web applications are becoming more and more critical. The growing dependency on the web as a tool, combined with the fact that some applications involve very large user bases, is becoming problematic as it seems to increase rapidly but silently. Some errors like crashes and confidential information leaks, if not discovered, can have massive effects and incur significant financial or reputation damage.

The second kind of properties concern efficiency of web applications. One particular characteristic of web programming languages is that they are essentially data-manipulation oriented. These manipulations rely on query and transformation languages whose performance is critical. This performance is very sensitive to data size and organization (constraints) and to the execution model (e.g. streaming evaluators). Static analysis can be used to optimize runtime performance by compile-time automated modification of the code (e.g. substitution of queries by more efficient ones). One major scientific difficulty here consists in dealing with problems close to the frontier of decidability, and therefore in finding useful trade-offs between programming ease, expressivity, complexity, succinctness, algorithmic techniques and effective implementations.

4. Application Domains

4.1. Web Programming Technologies

Despite the major social and economic impacts of the web revolution, current web programming methods and content representation are lagging behind and remain severely limited and in many respects archaic. Dangerously, designing web applications even becomes increasingly complex as it relies more and more on a jungle of programming languages, tools and data formats, each targeted toward a different application layer (presentation, application and storage). This often yields complex and opaque applications organized in silos, which are costly, inefficient, hard to maintain and evolve, and vulnerable to errors and security holes. In addition, the communication aspects are often handled independently via remote service invocations and represent another source of complexity and vulnerability. We believe that we reached a level where there is an urgent need and a growing demand for alternative programming frameworks that capture the essence of web applications: advanced content, data and communication. Therefore, successful candidate frameworks must capture rich document formats, data models and communication patterns. A crucial aspect is to offer correction guarantees and flexibility in the application architecture. For instance, applications need to be checked, optimized and managed as a whole while leveraging on the consistency of their individual components and data fragments. For all these reasons, we believe that a new generation of tools must be created and developed in order to overcome the aforementioned limitations of current web technologies.

4.2. Multimedia and Augmented Environments

The term Augmented Environments refers collectively to ubiquitous computing, context-aware computing, and intelligent environments. The goal of our research on these environments is to introduce personal Augmented Reality (AR) devices, taking advantage of their embedded sensors. We believe that personal AR devices such as mobile phones or tablets will play a central role in augmented environments. These environments offer the possibility of using ubiquitous computation, communication, and sensing to enable the presentation of context-sensitive information and services to the user. AR applications often rely on 3D content and employ specialized hardware and computer vision techniques for both tracking and scene reconstruction and exploration. Our approach tries to seek a balance between these traditional AR contexts and what has come to be known as mobile AR browsing. It first acknowledges that mobile augmented environment browsing does not require that 3D content be the primary means of authoring. It provides instead a method for HTML5 and audio content to be authored, positioned in the surrounding environments and manipulated as freely as in modern web browsers.

The applications we develop to guide and validate our concepts are pedestrian navigation techniques and applications for cultural heritage visits. Features found in augmented environments are demanding for the other activities in the team. They require all kinds of multimedia information, that they have to combine. This information has to be processed efficiently and safely, often in real time, and it also, for a significant part, has to be created by human users.

5. New Software and Platforms

5.1. XML Reasoning Solver

Participants: Pierre Genevès, Nabil Layaïda, Nils Gesbert, Louis Jachiet, Nicola Guido.

The XML Reasoning Solver is a tool for the static analysis of queries and schemas based on our theoretical advances [9]. It allows automated verification of properties that are expressed as logical formulas over trees. A logical formula may for instance express structural constraints or navigation properties (like e.g. path existence and node selection) in finite trees.

The reasoner is built on top of a finite tree logic solver for a new modal logic equipped with recursion and backward axes. The solver is very fast in practice and uses symbolic techniques (Binary Decision Diagrams). The solver has been recently extended to support functions, parametric functions and polymorphic subtyping. One notable difficulty was to elaborate many advanced optimizations with symbolic implementation techniques. The logical solver significantly advances the state of the art. In particular, it is the first implementation that effectively solves the query containment problem for a large fragment of the XPath query language. It supports all navigation axes and regular tree constraints. Although researchers had studied XPath satisfiability before, such prior works were either unimplementable or deemed to explode even for tiny examples. As of 2014, it is still the only implementation actually capable of solving this problem in practice for real world instances.

The reasoner includes compilers and various static analyzers for web query and schema languages. This includes compilers for XPath, for XML schemas (DTDs, XML Schemas, Relax NGs) into logical formulas, parsers, benchmarks, and libraries for automated testing. Various difficulties reside in the compilation of real-world queries, including compiling XPath queries into fixed-point logics, developing specific implementation techniques in order to avoid worst case blow-ups as much as possible when e.g. supporting unordered XML attributes among (ordered) XML elements, etc. The reasoner also generates counter-examples that allow program defects to be reproduced independently from the analyzer.

The off-line version of the solver (with a native library) is fast and up-to-date with the latest advances. We developed and deployed an interactive web interface to make the solver available to the international scientific community. For this purpose, we redesigned the libraries used for the manipulation of binary decision diagrams (BDDs) so that they could used in a fully concurrent and multithreaded manner. This is in order to allow several instances of the logical solver to run concurrently for several users on a web server (GWT-based), while decreasing performance as less as possible.

The reasoner helps us to guide and validate our approach. We continue to develop, maintain and use it on an almost-daily basis.

5.2. XQuery type-checker

Participants: Pierre Genevès, Nabil Layaïda, Nils Gesbert.

This prototype implements a sound static type-system for XQuery, which, as of december 2014, is the most precise type system known for XQuery. It supports the static typing of backward axes that no other does nor is supported in the XQuery recommendation. It also includes precise typing for conditional statements which is challenging as such statement are usually sensitive to the program context. Our type checker successfully verifies complex programs for which existing type-checkers (either known from the literature or those developed in commercial software) fail by reporting false alarms. One major benefit is to allow the cost of validation to be deferred from runtime to compile-time (once only). This prototype is implemented in Scala and interacts with the solver by issuing externals calls for deciding complex subtyping relations. This prototype is described in preprint [20]

5.3. CSS Analyzer

Participants: Pierre Genevès, Nabil Layaïda, Marti Bosch Padros.

This software now consists in two distinct prototypes: two static analyzers (with a different purpose) that share a common compiler for CSS. The first prototype is used for bug detection and verification of a cascading style sheet (CSS) file. It involves a compiler for CSS rules (and in particular selectors) into logical formulas, adapted for the semantics of CSS (see the initial WWW'12 paper). The second prototype performs automated refactoring for size reduction of CSS style sheets. It reuses the first compiler and the logical solver for detecting which rules can be refactored and how. It implements various optimisation techniques (like early pruning), for the purpose of dealing with large-size real CSS files. This prototype reduces the size of CSS files found in the most popular websites (such as CNN, facebook, Google Sites, Apple, etc.) by up to 30% while preserving their semantics [13].

5.4. ClaireCourseMaker Library

Participants: Nicolas Hairon, Cécile Roisin, Nabil Layaïda.

The goal of the ClaireCourseMaker is to provide direct and visual editing tools for structuring, annotating and timeline-based authoring of continuous content such as audio or video. I it is mainly devoted to the synchronisation and layout of pedagogical material (video, slides, chaptering, etc.) and enables the incorporation of rich media content in MOOCs. The underlying technology is based on Web standards and relies on the open source JavaScript Popcorn library and Popcorn Maker web application developed by the Mozilla Foundation. The tool is a wysiwyg web-based authoring tool which benefits from the generic features of Popcorn and offers structuring methods such chaptering and container-based synchronisation.

ClaireCourseMaker is the direct follow-up tool of the Timesheet library developed in the project. Timesheet library is a cross-browser JavaScript implementation for scheduling the dynamic behaviour of HTML5 content. It uses and provides a reference implementation for declarative synchronisation markup such as SMIL Timing and Synchronization and SMIL Timesheets.

ClaireCourseMaker is developed in collaboration with the OpenClassrooms company in the context of the Claire project (see section 7.1.1).

5.5. Interactive eXtensible Engine (IXE)

Participants: Nabil Layaïda, Pierre Genevès, Thibaud Michel, Mathieu Razafimahazo.

PDRTrack is a localization utility running on iOS or Android smartphones used for recording and playing data sets (accelerometer, gyroscope, barometer and magnetometer values) to study the effect of different pedometer and map matching parameters on indoor and outdoor localization accuracy. This application uses the PDR library, written in C++, which provides the user's location in real time based on the interpretation of mobile phone sensors. Three main modules have been designed to build this localization system:

- a pedometer that estimates the distance the user has walked and his speed
- a motion manager that enables data set recording and simulation but also the creation of virtual sensors or filters (e.g gyroscope drift compensation, linear acceleration, altimeter)
- a map-matching algorithm that provides location estimates on a given OpenStreetMap description and the current user's trajectory

The PDR library is a central component of the VENTURI project. It has been used for applications such as guiding a visually impaired people. Others partners have used this localisation system for retrieving a scale factor needed for the computer vision part (i.e SLAM).

GPS navigation systems, when used in an urban environment, are limited in precision and can only give instructions at the level of the street and not of the pavement or corridor. GPS is also limited to outdoor navigation and requires some transitioning system when switching to indoor navigation.

PDRTrack is embedded in IXE. IXE is an urban pedestrian navigation system based on Inertial Measurement Units (IMU) and running on mobile phones with onboard geographic data and a routing engine. IXE allows augmented reality queries on customised embedded geographical data. Queries on route nodes or POIs, on ways and relations are predefined for efficiency and quality of information. Following a web paradigm, IXE can be seen as web browser for XML documents describing navigation networks. by using the micro-format concept, one can define inside OpenStreetMap a complex format for pedestrian navigation networks allowing navigation at the level of pavements or corridors.

The big advantage of IXE is that it relies on a standard OpenStreetMap editor called JOSM to create navigation networks and augmented reality content. IXE browser reads OSM documents and produces from them visible or audible navigation information. IXE is composed of three engines, one for dead-reckoning navigation, one for interactive audio and the last one for Augmented Reality visual information.

6. New Results

6.1. Automated Refactoring for Size Reduction of CSS Style Sheets

Cascading Style Sheets (CSS) is a standard language for stylizing and formatting web documents [17]. Its role in web user experience becomes increasingly important. However, CSS files tend to be designed from a result-driven point of view, without much attention devoted to the CSS file structure as long as it produces the desired results. Furthermore, the rendering intended in the browser is often checked and debugged with a document instance. Style sheets normally apply to a set of documents, therefore modifications added while focusing on a particular instance might affect other documents of the set.

We present a first prototype and a new CSS semantic analyzer and optimizer that is capable of automatically detecting and removing redundant property declarations and rules. We build on earlier work on tree logics to locate redundancies due to the semantics of selectors and properties. Existing purely syntactic CSS optimizers can be used in conjunction with our tool, for performing complementary (and orthogonal) size reduction, toward the common goal of providing smaller and cleaner CSS files. We have been able to detect large numbers of unnecessary property declarations in complex web pages; and we have also found mistakes in the style sheets of some of the most popular web sites. The number of safe modifications can easily grow as more components of CSS are supported and more features are implemented, such as property inheritance, translation of pseudo-classes into query languages, analysis of media queries, merging of equivalent selectors or containment involving grouped selectors.

6.2. Equipping IDEs with XML-Path Reasoning Capabilities

One of the challenges in Web development is to achieve a good level of quality in terms of code size and runtime performance for popular domain-specific languages such as XQuery, XSLT, and XML Schema. We developed an IDE augmented with static detection of inconsistent XPath expressions that assists the programmer with simplifying development and debugging of any application involving XPath expressions [12]. The tool is based on newly developed formal verification techniques based on expressive modal logics, which are now efficient enough to be used in the process of software development. We applied this to a full XQuery compiler for which we introduced an analysis for identifying and eliminating dead code automatically.

6.3. XQuery and Static Typing: Tackling the Problem of Backward Axes

XQuery is a functional language dedicated to XML data querying and manipulation. As opposed to other W3C-standardized languages for XML (e.g. XSLT), it has been intended to feature strong static typing. Currently, however, some expressions of the language cannot be statically typed with any precision. This is due to a discrepancy between the semantics of the language and its type algebra: namely, the values of the language are (possibly inner) tree nodes, which may have siblings and ancestors in the data. The types on the other hand are regular tree types, as usual in the XML world: they describe sets of trees. The type associated to a node then corresponds to the subtree whose root is that node and contains no information about the rest of the data. This makes navigation expressions using "backward axes," which return e.g. the siblings of a node, impossible to type.

We show how to handle this discrepancy by improving the type system. We describe a logic-based language of extended types able to represent inner tree nodes and show how it can dramatically increase the precision of typing for navigation expressions. We describe how inclusion between these extended types and the classical regular tree types can be decided, allowing a hybrid system combining both type languages. The result is a net increase in precision of typing [20].

6.4. A Core Calculus for XQuery 3.0: Combining Navigational and Pattern Matching Approaches

XML processing languages can be classified according to whether they extract XML data by paths or patterns. The strengths of one category correspond to the weaknesses of the other. In this work, we propose to bridge the gap between these two classes by considering two languages, one in each class: XQuery (for path-based extraction) and CDuce (for pattern-based extraction). To this end, we extend CDuce so as it can be seen as a succinct core λ-calculus that captures XQuery 3.0. The extensions we consider essentially allow CDuce to implement XPath-like navigational expressions by pattern matching and precisely type them. The elaboration of XQuery 3.0 into the extended CDuce provides a formal semantics and a sound static type system for XQuery 3.0 programs [18].

6.5. Session Types as Generic Process Types

Behavioural type systems ensure more than the usual safety guarantees of static analysis [15]. They are based on the idea of "types-as-processes", providing dedicated type algebras for particular properties, ranging from protocol compatibility to race-freedom, lock-freedom, or even responsiveness. Two successful, although rather different, approaches, are session types and process types. The former allows to specify and verify (distributed) communication protocols using specific type (proof) systems; the latter allows to infer from a system specification a process abstraction on which it is simpler to verify properties, using a generic type (proof) system. What is the relationship between these approaches? Can the generic one subsume the specific one? At what price? And can the former be used as a compiler for the latter? This work is a step towards answers to such questions. Concretely, we have defined a stepwise encoding of a pi-calculus with sessions and session types (the system of Gay and Hole) into a pi-calculus with process types (the Generic Type System of Igarashi and Kobayashi). We encode session type environments, polarities (which distinguish session channels end-points), and labelled sums. We show forward and reverse operational correspondences for the encodings, as well as typing correspondences. To faithfully encode session subtyping in process types subtyping, one needs to add to the target language record constructors and new subtyping rules. This work shows how the programming convenience of session types as protocol abstractions can be combined with the simplicity and power of the pi-calculus, taking advantage in particular of the framework provided by the Generic Type System.

6.6. Personal Shopping and Navigator System for Visually Impaired People

We have developed a a personal assistant and navigator system for visually impaired people [14]. This system has been built using a set of domain specific languages based on XML such as OpenStreetMap extended

for Augmented Reality. It demonstrate how partially sighted people could be aided by the technology in performing an ordinary activity, like going to a mall and moving inside it to find a specific product. We propose an Android application that integrates Pedestrian Dead Reckoning and Computer Vision algorithms, using an off-the-shelf Smartphone connected to a Smart-watch. The detection, recognition and pose estimation of specific objects or features in the scene derive an estimate of user location with sub-meter accuracy when combined with a hardware-sensor pedometer. The proposed prototype interfaces with a user by means of Augmented Reality, exploring a variety of sensorial modalities other than just visual overlay, namely audio and haptic modalities, to create a seamless immersive user experience. The interface and interaction of the preliminary platform have been studied through specific evaluation methods. The feedback gathered will be taken into consideration to further improve the proposed system.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. Investissements d'avenir

CLAIRE

> Title: Community Learning through Adaptive and Interactive multichannel Resources for Education
>
> Call: Technologies for e-education
>
> Duration: March 2012 - February 2014
>
> Coordinator: OpenClassrooms, ex-SimpleIT
>
> Others partners: LIRIS
>
> See also: http://www.projet-claire.fr/
>
> Abstract: Project CLAIRE aims at developing an open-source tool for collaborative authoring in an e-learning environment (Learning Content Management System), targeting teachers and students in high-school and universities. Its innovative features include:
>
> - a platform for collaborative structured editing of rich media and "semantic" content, e.g.: tools for chaptering video, and for generating interactive evaluation tests
> - processes for continuous enhancement of content, e.g.: social annotation, behaviour analysis, accessible multi-support publishing, e.g.: web, PDF, ODT, LaTeX, smartphones, tablets.

Datalyse

> Title: Entrepôt Intelligent pour Big Data hétérogènes. Investissements d'Avenir Développement de l'Economie Numérique.
>
> Call: Cloud Computing, num 3 – Big Data.
>
> Duration: May 2013 - November 2016
>
> Coordinator: Business & Decision Eolas
>
> Others partners: Groupement des Mousquetaires, Inria Saclay (OAK EPC), LIG (Hadas and Erods teams), LIRMM (Montpellier), LIFL (Lille).
>
> See also: http://www.datalyse.fr/
>
> Abstract: Project Datalyse aims at designing and deploying an infrastructure for big data storage, collection, certification, integration, categorisation, enrichment and sharing over very large heterogeneous data sets. It relies on an industrial platform, to be made available on the cloud, and focuses on three flagship applications, showcasing three uses of big data over different data sets:

- Data-Center Monitoring: The goal of this application is to provide features such as traceability, reporting, optimisation and analysis of abnormal behaviour regarding energy efficiency and security issues. The application will be built with an existing application called ScopeBR (Eolas) and will be deployed in two different green data centers, those of Eolas and GDF SUEZ.
- "Territoire de données ouvertes et liées": This application aims at extracting and provisioning public open data collected from the city of Grenoble and its suburbs. The goal is to make public data available to third-party application developers and to federate local actors around a single platform.
- Real-time Business Intelligence for the management and processing of points of sale: this application will focus on real-time data analytics and will be deployed within "Groupement des Mousquetaires" in support of their business intelligence platforms.

7.1.2. ANR

Typex

Title: Typeful certified XML: integrating language, logic, and data-oriented best practices
Call: Programme Blanc
Duration: January 2012 - December 2014
Coordinator: PPS (CNRS - Paris 7 Diderot)
Others partners: LRI (Orsay)
See also: http://typex.lri.fr
Abstract: The highly ambitious and final goal of this project is to produce a new generation of XML programming languages stemming from the synergy of integrating three approaches into a unique framework:
- a logical approach based on solvers
- a programming language (PL) approach
- a data-oriented approach

These languages will feature precise and polymorphic type systems that merge PL typing techniques with logical-solver-based type inference. They will be implemented efficiently using the latest research on tree automata and formally certified using modern theorem prover technology. They will offer the capacity to specify and formally verify invariants, business rules, and data integrity, and will have a direct and immediate impact on standardization processes.

7.2. European Initiatives

7.2.1. FP7 & H2020 Projects

VENTURI

Title: immersiVe ENhancemenT of User-woRld Interactions
Type: Cooperation (ICT)
Call: FP7-ICT-20111.5 Networked Media and Search Systems
Instrument: Specific Targeted Research Project (STREP)
Duration: October 2011 - September 2014
Coordinator: Fondazione Bruno Kessler (Italy)
Others partners: Fraunhofer Heinrich Hertz Institute (Germany), ST Microelectronics (Italy), ST-Ericsson (France), Metaio (Germany), e-Diam Interactive (Spain), Sony-Ericsson (Sweden)
See also: https://venturi.fbk.eu/
Abstract: Venturi aims to create a pervasive Augmented Reality paradigm, where available information will be presented in a user- rather than device-specific way. The goal is to create an experience that is always present whilst never obstructing. Venturi will exploit, optimize and extend current and next generation mobile platforms; verifying platform and QoE performance through life-enriching use cases and applications to ensure device-to-user continuity.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

Prof. Boualem Benatallah, Professor at the School of Computer Science and Engineering (CSE), the University of New South Wales (UNSW) in Sydney, Australia, visited our group for one week in July 2014. The goal of this visit was to initiate common work on the formal verification of web services orchestration and verified API-driven web programming.

7.3.1.1. Internships

Martì Bosch Padros from Universitat Politècnica de Catalunya (UPC) Spain spent six months in the team to work on Automated Refactoring for Size Reduction of CSS Style Sheets.

Joel Ferreira Dos Santos from Universidade Federal Fluminense, UFF, Brasil is spending a one year sandwich PhD in the team to work on the formal verification of multimedia presentations.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

8.1.1.1. Member of the organizing committee

C. Roisin is a member of the steering committee of the ACM Symposium on Document Engineering.

8.1.2. Scientific events selection

8.1.2.1. Chair of conference program committee

P. Genevès is the PC Chair of the ACM Symposium on Document Engineering 2015.

8.1.2.2. Member of the conference program committee

P. Genevès is member of the external review committee for the 42nd ACM Symposium on Principles of Programming Languages (POPL'15).

P. Genevès and N. Layaïda are program committee members for the 14th ACM Symposium on Document Engineering (DocEng'14).

8.1.2.3. Reviewer

P. Genevès has been referee for International Colloquium on Automata, Languages and Programming (ICALP)

P. Genevès has been referee for International Conference on Relational and Algebraic Methods in Computer Science (RAMiCS).

8.1.3. Journal

8.1.3.1. Reviewer

N. Layaïda has been a referee for the Communications of the ACM.

P. Genevès has been a referee for Engineering Science and Technology (Elsevier).

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Licence: C. Roisin, Computer Networks, 100h/year, L2, University of Grenoble, France

Licence: C. Roisin, Pedagogical manager of the Professional Licence "MESSI" and Academic-side tutoring of several apprentices of this licence, University of Grenoble, France

Licence: N. Gesbert, Logic, 45h/year, L3, University of Grenoble, France

Licence: N. Gesbert, Algorithmics, 60h/year, L3, University of Grenoble, France

Licence: N. Gesbert, Academic-side tutoring of an apprentice, 10h/year, University of Grenoble, France

Master: N. Gesbert, Logic, 45h/year, M1, University of Grenoble, France

Master: N. Gesbert, Web development, 22h30/year, M1, University of Grenoble, France

Master: N. Gesbert, Software analysis, conception and validation, 45h/year, M1, University of Grenoble, France

Master: P. Genevès, Foundations for XML: logics and automata, 18h/year, M2 (Mosig), University of Grenoble, France

8.2.2. Supervision

HdR: P. Genevès, Static Analysis for Data-Centric Web Programming, Université de Grenoble Alpes, 21 November 2014.

PhD in progress: D. Graux, Large scale evaluation of semantic web queries, since November 2013, N. Layaïda and P. Genevès

PhD in progress: N. Guido, Semantic Query-Update Independence Analysis under Constraints, since October 2012, C. Roisin and P. Genevès

PhD in progress: A. Abbas, Web query rewriting for heterogeneous data sources, since October 2014, N. Layaïda and P. Genevès

PhD in progress: T. Michel, Mobile Augmented Reality Applications for Smart Cities, since October 2014, P. Genevès, N. Layaïda and H. Fourati

PhD in progress: L. Jachiet, Reasoning with NoSQL Data Flows in Massively Parallel Systems, since October 2014, N. Layaïda and P. Genevès

8.2.3. Juries

N. Layaïda has been reviewer for the PhD thesis of Houari Mahfoud, Efficient Access Control to XML Data: Querying and Updating Problems, Université de Lorraine, supervisors: Abdessamad Imine and Michael Rusinowitch. 18 February 2014.

N. Layaïda has been committee member of the HdR of P. Genevès, Static Analysis for Data-Centric Web Programming, Université de Grenoble Alpes, 21 November 2014.

C. Roisin has been reviewer for midterm PhD thesis of Mira Sarkis, Supervised by Jean-Claude Dufourd and Cyril Concolato, Télécom Paristech, Paris, 2014.

8.3. Popularization

M. Razafimahazo developed a mobile application for the pupils of the école de la paix, Grenoble. The application uses Augmented Reality for cultural heritage tour of Grenoble City with content produced by pupils.

M. Razafimahazo Animated a brainstorming during CitizenTIC around OpenStreetMap crowd-sourcing: "Les cartes changent la donne autour des parcours accessibles dans l'agglomération Grenobloise".

N. Layaïda animated several workshops on Open Linked Data for Grenoble City and Grenoble Agglomeration (La métro).

We received two pupils from secondary school, Jeremy Andreoletti and Hugo Kersaudy, for a one week initiation to research. We hosted several pupils from the primary school for a one day for a presentation and demos.

8.3.1. Institutional commitment:

C. Roisin Board member of the University Pierre Mendes-France, University of Grenoble, in charge of IT systems and TICE.

C. Roisin is President of the disciplinary commission of the University Pierre Mendes-France, University of Grenoble.

N. Layaïda is member of the Scientific Board of Advanced Data-mining of the Parsyval Labex.

N. Layaïda is member of the experts pool (selection committee) of the minalogic competitive cluster.

9. Bibliography

Major publications by the team in recent years

[1] G. CASTAGNA, N. GESBERT, L. PADOVANI. *A theory of contracts for Web services*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", 2009, vol. 31, n° 5, pp. 1-61 [*DOI :* 10.1145/1538917.1538920], http://hal.archives-ouvertes.fr/hal-00434459

[2] F. CAZENAVE, V. QUINT, C. ROISIN. *Timesheets.js: When SMIL Meets HTML5 and CSS3*, in "DocEng 2011: Proceedings of the Eleventh ACM Symposium on Document Engineering", Mountain View, États-Unis, September 2011, 10 p. , http://hal.archives-ouvertes.fr/hal-00619382

[3] M. W. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA. *SPARQL Query Containment under RDFS Entailment Regime*, in "6th International Joint Conference on Automated Reasoning (IJCAR)", Manchester, United Kingdom, B. GRAMLICH, D. MILLER, U. SATTLER (editors), Springer, June 2012, pp. 134-148, http://hal.inria.fr/hal-00749087

[4] M. W. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA. *SPARQL Query Containment Under SHI Axioms*, in "26th AAAI Conference on Artificial Intelligence", Toronto, Canada, AAAI Press, July 2012, pp. 10-16, http://hal.inria.fr/hal-00749080

[5] S. J. GAY, N. GESBERT, A. RAVARA, V. T. VASCONCELOS. *Modular Session Types for Objects*, May 2011, http://hal.archives-ouvertes.fr/hal-00700635

[6] S. J. GAY, V. T. VASCONCELOS, A. RAVARA, N. GESBERT, A. Z. CALDEIRA. *Modular Session Types for Distributed Object-oriented Programming*, in "Proceedings of the 37th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", New York, NY, USA, POPL '10, ACM, 2010, pp. 299–312, http://doi.acm.org/10.1145/1706299.1706335

[7] P. GENEVÈS, N. LAYAÏDA, V. QUINT. *Impact of XML Schema Evolution*, in "ACM Transactions on Internet Technology", July 2011, vol. 11, n° 1 [*DOI :* 10.1145/1993083.1993087], http://hal.inria.fr/inria-00619225

[8] P. GENEVÈS, N. LAYAÏDA, V. QUINT. *On the Analysis of Cascading Style Sheets*, in "WWW '12 - 21st international conference on World Wide Web - 2012", Lyon, France, April 2012, pp. 809-818 [*DOI :* 10.1145/2187836.2187946], http://hal.inria.fr/hal-00690899

[9] P. Genevès, N. Layaïda, A. Schmitt. *Efficient Static Analysis of XML Paths and Types*, in "Proceedings of the 2007 ACM SIGPLAN conference on Programming language design and implementation", San Diego, États-Unis, 2007, pp. 342–351 [*DOI :* 10.1145/1250734.1250773], http://hal.archives-ouvertes.fr/hal-00189123

[10] N. Gesbert, P. Genevès, N. Layaïda. *Parametric Polymorphism and Semantic Subtyping: the Logical Connection*, in "International conference on functional programming", Tokyo, Japon, ACM SIGPLAN, September 2011, pp. 107-116 [*DOI :* 10.1145/2034773.2034789], http://hal.inria.fr/inria-00585686

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] P. Genevès. *Static Analysis for Data-Centric Web Programming*, Université Grenoble Alpes, November 2014, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01102401

Articles in International Peer-Reviewed Journals

[12] P. Genevès, N. Layaïda. *Equipping IDEs with XML-Path Reasoning Capabilities*, in "ACM Transactions on Internet Technology", July 2014, vol. 13, n° 4, 20 p. [*DOI :* 10.1145/2602573], https://hal.inria.fr/hal-00868723

International Conferences with Proceedings

[13] M. Bosch, P. Genevès, N. Layaïda. *Automated Refactoring for Size Reduction of CSS Style Sheets*, in "Proceedings of the 2014 ACM symposium on Document engineering", Fort Collins, Denver, United States, September 2014 [*DOI :* 10.1145/2644866.2644885], https://hal.inria.fr/hal-01081876

[14] P. Chippendale, V. Tomaselli, V. D'Alto, G. Urlini, C. Maria Modena, S. Messelodi, S. Mauro Strano, G. Alce, K. Hermodsson, M. Razafimahazo, T. Michel, M. F. Giovanni. *Personal Shopping Assistance and Navigator System for Visually Impaired People*, in "ACVR2014: Second Workshop on Assistive Computer Vision and Robotics", Zurich, Switzerland, September 2014, https://hal.inria.fr/hal-01102707

[15] S. J. Gay, N. Gesbert, A. Ravara. *Session Types as Generic Process Types*, in "Proceedings of the Combined 21st International Workshop on Expressiveness in Concurrency and 11th Workshop on Structural Operational Semantics (EXPRESS/SOS)", Rome, Italy, Electronic Proceedings in Theoretical Computer Science, September 2014, vol. 160, pp. 94 - 110 [*DOI :* 10.4204/EPTCS.160.9], https://hal.inria.fr/hal-01102349

Scientific Popularization

[16] M. Razafimahazo, N. Layaïda, P. Genevès, T. Michel. *Mobile Augmented Reality Applications for Smart Cities*, in "ERCIM News", 2014, 2 p. , https://hal.inria.fr/hal-01102797

Other Publications

[17] M. Bosch, P. Genevès, N. Layaïda. *Automated Refactoring for Size Reduction of CSS Style Sheets*, July 2014, https://hal.inria.fr/hal-01021332

[18] G. CASTAGNA, H. IM, K. NGUYỄN, V. BENZAKEN. *A Core Calculus for XQuery 3.0: Combining Navigational and Pattern Matching Approaches*, September 2014, https://hal.inria.fr/hal-01104872

[19] S. J. GAY, N. GESBERT, A. RAVARA, V. T. VASCONCELOS. *Modular session types for objects*, January 2015, https://hal.archives-ouvertes.fr/hal-00700635

[20] P. GENEVÈS, N. GESBERT, N. LAYAÏDA. *XQuery and Static Typing: Tackling the Problem of Backward Axes*, November 2014, https://hal.inria.fr/hal-01082635

[21] P. GENEVÈS, N. LAYAÏDA, A. SCHMITT, N. GESBERT. *Efficiently Deciding μ-calculus with Converse over Finite Trees*, January 2015, https://hal.inria.fr/hal-00868722