IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université Pierre Mendes-France (Grenoble 2)

Université Joseph Fourier (Grenoble 1)

Activity Report 2011

Project-Team WAM

Web, adaptation and multimedia

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

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Knowledge and Data Representation and Management

Table of contents

Project-Team WAM

Keywords: Web, Multimedia Document, XML, Logics, Verification, Augmented Reality

1. Members

Research Scientists

Pierre Genevès [Junior Researcher Cnrs]
Nabil Layaïda [Junior Researcher Inria]
Vincent Quint [Senior Researcher Inria, Team Leader, HdR]

Faculty Members

Nils Gesbert [Associate Professor, Université de Grenoble INP – since 1 September 2011]
Jacques Lemordant [Associate Professor, on leave from Université de Grenoble UJF, HdR]
Cécile Roisin [Professor, Université de Grenoble UPMF, HdR]

Technical Staff

Fabien Cazenave [Engineer – until 13 July 2011]
Audrey Colbrant [Engineer]
Yohan Lasorsa [Engineer]
David Liodenot [Engineer]
Mathieu Razafimahazo [Engineer]

PhD Students

Everardo Bárcenas Patiño [until 28 February 2011]
Melisachew Chekol [jointly with project-team EXMO]

Administrative Assistants

Ahlem Zammit-Boubaker [shared with project-teams Moais and Exmo – until 30 September 2011]
Françoise de Coninck [shared with project-teams Ibis, Licit, Nano-D and Opale – since 1 October 2011]

2. Overall Objectives

2.1. Objectives

Project WAM aims at making it easier to develop and use *rich multimedia* contents and applications *on the web*.

Many web sites are specializing in a single type of content, such as Picasa and Flickr for photographs, YouTube and Dailymotion for videos, iTunes and Deezer for music, etc. Some other sites offer web pages that contain text, pictures, videos and audio simultaneously (newspapers sites, for instance). So, different types of contents coexist on the web, even on the same web page, but this does not really make a multimedia web or multimedia pages. The web has demonstrated how links, relations, connections, interactions between pieces of information can enhance the raw content of each piece. We are not there yet with multimedia content. Integrating and connecting heterogeneous contents on the web still have to be explored.

That is the reason why we pay a particular attention to documents and applications that *tightly integrate* different types of media objects, be they discrete (text, images, equations) or continuous (video, audio, animations). Continuous contents add a time dimension to documents that mix various sorts of contents. This extra dimension raises new issues. It has to be combined with other, more traditional points of view on documents, such as their layout and style (spatial dimension), their organization (logical dimension) often represented as a hierarchical structure, etc.

In the context of the web, multimedia resources are distributed and can be assembled in various ways to make different documents and to be processed by multiple applications, running on all sorts of computers, devices and networks. For this reason, they have to be represented in platform-neutral formats.

This approach to web multimedia content and applications raises a number of issues. We have chosen to address four categories of problems:

2.1.1. Multimedia Models and Formats

For a long time, most multimedia web pages have isolated continuous content behind the fences of add-ons or plug-ins, thus preventing real interaction between these contents and the rest of their host page and the whole web. In addition, the many interactive features that are available with discrete content have no equivalent within plug-ins, where users are limited to the same level of control they have with a VCR.

New models are required to represent the many dimensions of multimedia documents. Ideally, such models should keep the aspects of traditional documents that have proven useful, and extend them with the specificities of the web environment and continuous contents. The key issue here is to allow all these aspects to be present simultaneously for representing a single document. If this goal is achieved, models can be rich and versatile enough to offer many possibilities to a broad range of applications handling multimedia contents.

To be used in real applications, such multimedia document models have to be instantiated in actual formats and languages. As documents have to be part of the web, these formats must be compatible with existing web formats. They could be extensions of existing formats, or new languages that share as many features as possible with the existing ones. The goal is not to create a separate web for multimedia content, but to seamlessly extend the web as we know it.

2.1.2. XML Processing

XML was created for representing documents and data on the web in a secure and rigorous way. XML is now the ground on which web formats are built. If we want to propose new formats for the web, they have to be based on XML, and we need to make sure new applications will be able to take advantage of these formats. It is therefore crucial to better understand how XML structures can be handled, and what are the theoretical tools that may help to develop an effective framework for processing XML structures.

This is of course an ambitious and long-term goal that requires intermediate steps. The first specialized languages for handling XML structures were transformation-oriented (XSLT, XDuce, CDuce, etc.). Typically, programs written in these languages read an XML structure and produce another XML structure as their output, after performing some transformations. Query languages can also be considered as behaving that way. So, the transformation paradigm is an interesting intermediate step towards general XML processing. Actually, a number of applications can be built as transformations: document formatting (XSLT was initially developed as a part of the XSL formatting language), filtering, merging, conversion, re-purposing, data query, etc.

A major component in an XML transformation language is the part that allows a programmer to select in the input structure the data of interest for a given transformation. We have then focused on this part of XML processing languages, and we have in particular studied the XPath language, which is used in a variety of other languages for XML (XSLT, XQuery, XML Schemas). The main goal of this work is to find the theoretical tools and formalisms that are needed for static analysis of XPath expressions, in order to help programmers develop better and more reliable code for XML data and documents.

2.1.3. Multimedia Authoring

Before they can be processed, multimedia XML documents have to be created. A significant part of web documents are generated by programs from other documents and data (see XML processing above), but another part is created by human authors using authoring tools. For multimedia formats to be really used, it is important that authoring tools be available.

Our work in the area of multimedia authoring tools aims at developing editing techniques for creating rich multimedia documents that take advantage of the many new dimensions of multimedia formats. The challenge is to keep these tools simple enough for average web users. Methods used for static, textual documents do not work for dynamic, multimedia web resources. New approaches have to be developed and experimented.

Research in this area is strongly connected with software development projects, with the goal of creating real tools that can be deployed on the web and that real users can use.

2.1.4. Augmented Environments

For the previous three objectives, we have chosen Augmented Reality as an application domain that helps us focus our work in accordance with application requirements.

To recreate or augment our perception of the real world, all modalities may be involved. For visual perception, the media that come to mind are text, graphics, photographs, video (live or recorded). But augmented reality is not restricted to the visual space. The auditory space also contributes to re-creating the user environment. Moreover, the visual and auditory spaces are connected: events happening in one space often have consequences in the other, and all this is synchronized.

The geographical space is important in augmented environments. The location of the users in the real or virtual world plays a key role, as well as the moves they make. This involves mobility and specific kinds of information, such as maps or points of interest (PoIs). A number of information resources needed to build augmented environments are available on the web. Applications have then not only to capture a lot of information about the local environment of their user (mainly through various sensors), but they also need to access additional information on the web.

All these features of augmented environments are very 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 has also, for a large part, to be created by human users.

3. Scientific Foundations

3.1. XML Processing

Participants: Everardo Bárcenas Patiño, Melisachew Chekol, Pierre Genevès, Nils Gesbert, Nabil Layaïda, Vincent Quint.

Extensible Markup Language (XML) has now gained considerable interest from industry, and now plays a central role in modern information system infrastructures. In particular, XML is the key technology for describing, storing, and exchanging a wide variety of data on the web. The essence of XML consists in organizing information in tree-tagged structures conforming to some constraints which are expressed using type languages such as DTDs, XML Schemas, and Relax NG.

There still exist important obstacles in XML programming, especially performance and reliability. Programmers are given two options: domain-specific languages such as XSLT, or general-purpose languages augmented with XML application programming interfaces such as the Document Object Model (DOM). Neither of these alternatives is a satisfactory answer to performance and reliability issues, nor is there even a trade-off between the two. As a consequence, new paradigms are being proposed and all have the aim of incorporating XML data as first-class constructs in programming languages. The hope is to build a new generation of tools that are capable of taking reliability and performance into account at compile time.

One of the major challenges in this line of research is to develop automated and tractable techniques for ensuring static type safety and optimization of programs. To this end, there is a need to solve some basic reasoning tasks that involve very complex constructions such as XML types (regular tree types) and powerful navigational primitives (XPath expressions). In particular, every future compiler of XML programs will have to routinely solve problems such as:

- XPath query emptiness in the presence of a schema: if one can decide at compile time that a query is not satisfiable then subsequent bound computations can be avoided
- query equivalence, which is important for query reformulation and optimization
- path type-checking, for ensuring at compile time that invalid documents can never arise as the output of XML processing code.

All these problems are known to be computationally heavy (when decidable), and the related algorithms are often tricky.

We have developed an XML/XPath static analyzer based on a new logic of finite trees [13]. This analyzer consists in compilers that allow XML types and XPath queries to be translated into this logic, and an optimized logical solver for testing satisfiability of a formula of this logic.

The benefit of these compilers is that they allow one to reduce all the problems listed above, and many others, to logical satisfiability. This approach has a couple of important practical advantages. First of all, one can use the satisfiability algorithm to solve all of these problems. More importantly, one could easily explore new variants of these problems, generated for example by the presence of different kinds of type or schema information, with no need to devise a new algorithm for each variant.

3.2. Multimedia Models and Languages

Participants: Yohan Lasorsa, Nabil Layaïda, Jacques Lemordant, Vincent Quint, Cécile Roisin.

We have participated in the international endeavor for defining a standard multimedia document format for the web that accommodates the constraints of different types of terminals. SMIL is the main outcome of this work. It focuses on a modular and scalable XML format that combines efficiently the different dimensions of a multimedia web document: synchronization, layout and linking. Our current work on multimedia formats follows the same trend.

With the advent of HTML5 and its support in most popular browsers, HTML is becoming an important multimedia language. Video and audio can now be embedded in HTML pages without worrying about the availability of plugins. However, animation and synchronization of a HTML5 page still require programming skills. To address this issue, we are developing a scheduler that allows HTML documents to be animated and synchronized in a purely declarative way. This work is based on the SMIL Timing and Synchronization module and the SMIL Timesheets specification. The scheduler is implemented in JavaScript, which makes it usable in any browser. Timesheets can also be used with other XML document languages, such as SVG for instance.

Audio is the poor relation in the web formats family. Most contents on the web may be represented in a structured way, such as text in HTML or XML, graphics in SVG, or mathematics in MathML, but sound was left aside with low-level representations that basically only encode the audio signal. Our work on audio formats aims at allowing sound to be on a par with other contents, in such a way it could be easily combined with them in rich multimedia documents that can then be processed safely in advanced applications. More specifically, we have participated in IAsig (Interactive Audio special interest group), an international initiative for creating a new format for interactive audio called iXMF (Interactive eXtensible Music Format). We are now developing A2ML, an XML format for embedded interactive audio, deriving from well-established formats such as iXMF and SMIL. We use it in augmented environments (see section 3.4), where virtual, interactive, 3D sounds are combined with the real sonic environment.

Regarding discrete media objects in multimedia documents, popular document languages such as HTML can represent a very broad range of documents, because they contain very general elements that can be used in many different situations. This advantage comes at the price of a low level of semantics attached to the structure. The concepts of microformats and semantic HTML were proposed to tackle this weakness. More recently, RDFa was introduced with the same goal. These formats add semantics to web pages while taking advantage of the existing HTML infrastructure. With this approach new applications can be deployed smoothly on the web, but authors of web pages have very little help for creating and encoding this kind of semantic markup. A language that addresses these issues is developed and implemented in WAM. Called XTiger, its role is to specify semantically rich XML languages in terms of other, less expressive XML languages, such as HTML. Recent extensions to the language make it now usable also to edit pure XML documents and to define their structure model (see section 3.3).

3.3. Multimedia Authoring

Participants: Audrey Colbrant, Yohan Lasorsa, Jacques Lemordant, David Liodenot, Vincent Quint, Mathieu Razafimahazo, Cécile Roisin.

3.3.1. Structured editing

Multimedia documents are considered through several kinds of structures: logical organization, layout, time, linking, animations. We are working on techniques that allow authors of such documents to manipulate all these structures in homogeneous environments. The main objective is to support new advances in document formats without making the authoring task more complex. The key idea is to present simultaneously several views of the document, each view putting the emphasis on a particular structure, and to allow authors to manipulate each view directly and efficiently. As the various structures of a document are not independent from each other, views are "synchronized" to reflect in each of them the consequences of every change made in a particular view. The XML markup, although it can be accessed at any time, is handled by the tools, and authors do not have to worry about syntactical issues.

3.3.2. Template-driven editing

We have recently experimented another way to edit highly structured XML documents without the usual complexity of the most common XML editors. The novelty of the approach is to use templates instead of XML schemas or DTDs, and to run the editor as a web application, within the browser. This way, it is much easier to create new document types and to provide an editing environment for these document types, that any web user can instantly use. This lightweight approach to XML editing complements the previous approach by covering new categories of XML applications.

3.4. Augmented Environments

Participants: Audrey Colbrant, Yohan Lasorsa, Jacques Lemordant, David Liodenot, Mathieu Razafimahazo.

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 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.

Augmented reality applications often rely on 3D content and employ specialized hardware and computer vision techniques for both tracking and scene reconstruction. 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.

Many service providers of augmented environments desire to create innovative services. Accessibility of buildings is one example we are involved in. However, service providers often have to strongly rely on experience, intuition, and tacit knowledge due to lack of tools on which to base a scientific approach. Augmented environments offer the required rigorous approach that enables the creation of Evidence-Based Services (EBS) if adequate tools for AR technologies are designed. Service cooperation through exchange of normalized real-time data or data logs is one of these tools, together with sensor data streams fusion inside an AR mobile browser. EBS can improve the performance of real-world sensing, and conversely EBS models authoring and service operation can be facilitated by real-world sensing.

The applications we use to elaborate and validate our concepts are pedestrian navigation for visually impaired people and applications for cultural heritage visits. On the authoring side, we are interested in interactive indoor modeling, audio mobile mixing, and formats for Points of Interest. Augmented environment services we consider are, among others, behavior analysis for accessibility, location services, and indoor geographical information services.

4. Application Domains

4.1. Application Domains

Broadly speaking, the main application domain of our research is the web and its numerous applications. This includes the recent evolutions of the web, with a special attention paid to the mobile web, the multimedia web, and the web as a platform for applications. The goal of our research is to enable new multimedia applications that can be deployed easily on the web, taking advantage of the existing infrastructure and the web technology.

Work on XML processing is related to one of the foundations of web architecture, i.e. resource representation. As such, it applies to a large part of web technology, be it used on the web or in other settings. At the moment, it has strong connections with research in other areas of computer science such as data bases and programming languages, where XML structures play an increasingly important role.

A highly challenging area for experimenting multimedia models and tools is the access to large audiovisual collections. The use of discrete information (text, images, graphics) tightly synchronized with continuous contents (audio, video) is the main way to develop new applications for exploiting the cultural heritage stored in radio and TV archives.

For our work on augmented environments, the application domain we address currently are pedestrian navigation and AR (Augmented Reality) information systems. A pedestrian navigation system has to cope with many cooperating tasks referring to different levels of precision, from micro-navigation to global navigation, including macro-navigation.

Micro-navigation builds upon embedded software ability to create a greater awareness of the immediate environment, using texture-based tracking or vision algorithms and relating this information to map and IMU (Inertial Measurement Units) data. Micro-navigation includes avoiding obstacles, locating a clear path in the proximate surroundings or at a complex crossing, finding objects and providing absolute positioning using known landmarks or beacons. Micro-navigation works at a precision level of a few centimeters by using predefined landmarks.

Macro-navigation refers to the actions required to find a way in a larger, not immediately perceptible environment, and builds upon carefully designed pedestrian-ways incorporating speech instructions, audio guidance, environmental queries and IMU instructions among other things. Macro-navigation works at a precision level of one step using carefully designed routes with map-matching instructions.

There is a duality relation between micro-navigation and macro-navigation. Micro-navigation is based on a localization system giving an absolute position which allows to compute a relative position with respect to the planned route. Macro-navigation is based on a localization system giving a relative position which allows to compute an absolute position on the route through a map-matching process. As a consequence, these two kinds of navigation complement and enhance each other.

Global navigation is based on an absolute global localization system like the GPS. Its precision is that of a few meters if used in a adequate geographical environment where data from external sensors are accessible. It can be used to bootstrap macro-navigation through remote sight guidance for example.

MMG navigation, i.e. the join use of micro, macro and global navigation, allows to build richer and more precise AR mobile applications in such fields as cultural heritage visits, outdoor games and visually impaired people guidance.

5. Software

5.1. Amaya

Participant: Vincent Quint.

Amaya is an open source web editor, i.e. a tool for creating and updating documents directly on the web. Browsing features are seamlessly integrated with editing features in a uniform environment that allows users to save files locally and on remote servers as well. This follows the original vision of the web as a space for collaboration and not just a one-way publishing medium.

Amaya started as a joint effort with W3C to showcase web technologies in a fully-featured web client. The main motivation for developing Amaya was originally to provide a framework that can integrate many web technologies during their development, with the goal of demonstrating these technologies in action while taking advantage of their combination in a single, consistent environment.

Amaya now implements a number of web technologies, such as HTML and the XHTML family, CSS style sheets, generic XML, MathML (for mathematical expressions), and SVG (for vector graphics). It allows all those document formats to be edited simultaneously in compound documents. It also includes a collaborative annotation application based on RDF, XLink, and XPointer.

It is a unique tool for manipulating simultaneously different kinds of content through a formatted representation of documents, while closely following standard formats. Developed jointly with W3C, the software is distributed world wide through the W3C servers and many mirrors. It is also part of several Linux distributions.

Amaya is also used as a platform for experimenting and distributing new editing techniques and document formats developed in WAM. It provides a full implementation of the XTiger language and its constraint-driven editing feature. It also helps users to create their own document types defined as XTiger templates.

5.2. XML Reasoning Solver

Participants: Pierre Genevès, Nabil Layaïda.

The XML Reasoning Solver is a tool for the static analysis of XPath queries and XML schemas based on the latest theoretical advances [13]. 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 tool can solve many fundamental XML problems such as satisfiability of XPath expressions in the presence of XML schemas, containment and equivalence of XPath expressions, and many other problems that can be formulated with XPath expressions and schemas (DTDs, XML Schemas, Relax-NG).

The system is implemented in Java and uses symbolic techniques (binary decision diagrams) in order to enhance its performance. It is capable of comparing path expressions in the presence of real-world DTDs (such as the W3C SMIL and XHTML language recommendations, for instance). The cost ranges from several milliseconds for comparison of XPath queries without tree types, to several seconds for queries under very large, heavily recursive, type constraints, such as the XHTML DTD. These measurements shed light for the first time on the cost of solving static analysis problems in practice. Furthermore, the analyzer generates XML counter-examples that allow program defects to be reproduced independently from the analyzer.

5.3. Timesheets Library

Participants: Fabien Cazenave, Cécile Roisin.

The goal of the Timesheets library is to synchronize HTML5 content using declarative synchronization languages defined by W3C standards (namely, SMIL Timing and Synchronization and SMIL Timesheets).

With the raise of HTML5 which natively supports continuous content (audio, video) there is a dramatic need for handling synchronization, animation and user interaction in an efficient and homogeneous way. As web browsers do not support SMIL, except for SVG Animation (which rely on the SMIL BasicAnimation module), multimedia web authoring remains difficult and relies on code-based, non-standard solutions.

Therefore we are developing a generic, cross-browser JavaScript implementation for scheduling the dynamic behavior of HTML5 content that can be described with declarative SMIL markup. Using a declarative language makes sense for the most common tasks, that currently require JavaScript programming:

- it is much easier for web authors and for web authoring tool developers;
- it is a much better way to achieve good accessibility and indexability;
- it is easier to maintain, since no specific JavaScript code is used.

5.4. Mobile Audio Language

Participants: Yohan Lasorsa, Jacques Lemordant.

5.4.1. MAUDL library

The MAUDL library (Mobile AUDio Language) [15] is an evolution of the ARIA library whose primary target was games on mobile.

Augmented Reality Audio applications use sound objects to create a soundscape. A sound object is a time structure of audio chunks whose duration is on the time scale of 100 ms to several seconds. These sound objects have heterogeneous and time-varying properties. In order to describe Interactive Audio (IA) contents, we created MAUDL, an XML language inspired by iXMF that is well adapted to the design of dynamic soundtracks for navigation systems.

MAUDL prevents audio information overwhelming through categorization at the declarative level and the use of priority queues at the execution level. This allows to take account of speed when walking, and rapid hand gestures when interrogating the environment for example. MAUDL can be used as an authoring time interchange file format for interactive mobile applications or as a runtime file format that is actually loaded through the web and played directly in the mobile. MAUDL is a cue-oriented interactive audio system, audio services being requested using named events and the systems response to each event being determined by the audio artist. The current version of the API supports iOS and further support for other mobile platforms (Android) is planned.

5.4.2. 3D Audio Pointer

A virtual 3D audio pointer provides an intuitive guide to the user of a mobile application, reducing the need for cognitive work when compared to vocal instructions. We have built such a pointer using the MAUDL language. It gives the user the azimuth using HRTF spatialized audio cues, with additional hints taking the form of variations in the sound used. It allows to superpose other kinds of audio contents, such as voice while the pointer is active, to indicate distance for example. This audio object is suitable for different sorts of navigation systems, such as POIs browsers, self-guided audio tours, or predefined route following applications.

5.5. Mixed Reality Browser

Participants: Audrey Colbrant, Yohan Lasorsa, Jacques Lemordant, David Liodenot, Mathieu Razafimahazo.

The Mixed Reality Browser (MRB) is a geolocalized web browser running on mobile devices.

The concept of Mixed Reality comes from the fact that the real/virtual dichotomy is not sharp, but interpolatively smooth over a virtuality continuum. Idealized notions of reality and virtuality can be thought of as endpoints on a continuum, an instance of the former approach corresponding for example to a see-through display with natural sounds, an instance of the latter to texture-mapped image-based rendering (panoramas) with synthetic sound objects.

Augmented Reality (AR) mode refers to all cases in which the auditory or visual display of an otherwise real environment is augmented by means of virtual sound or graphic objects. The converse case on the virtuality continuum is Augmented Virtuality (AV), where a virtual world, one that is generated primarily by computer, like with synthetic 3D graphic or synthetic panoramic, is being augmented with the audio-visual content of points of interest (POIs).

The introduction of mobile augmented reality browsers has forced a rethink on what kind of reality should be offered. Mobility induces a need for telepresence and simulation to free the user or the developer of the necessity to go every time in the real world. Mobility is the main reason behind the concept of Mixed Reality Browsers. By its intrinsic characteristics, MRB supports advance MR applications like mobile remote maintenance and assisted navigation.

6. New Results

6.1. Multimedia Models and Formats

6.1.1. SMIL timesheets

With the advent of HTML5 and its support in most popular browsers, HTML is becoming an important multimedia language. Video and audio can now be embedded in HTML pages without worrying about the availability of plugins. However, a major issue is to specify the dynamic behavior of documents (user interactions, timing and synchronization with continuous contents). This is done usually by writing (often complex) scripts, which require programming skills from the authors.

To address this issue, we have created the timesheets.js library, [1] a scheduler that allows HTML documents to be animated and synchronized in a purely declarative way. This work is based on the SMIL Timing and Synchronization module and the SMIL Timesheets specification, with a few extensions.

The library is implemented in JavaScript, which makes it usable in any browser. Authors can specify the dynamic behavior of HTML5 (+CSS3) documents [6]. They can thus develop multimedia applications without writing a single line of JavaScript. Timesheets can also be used with other XML document languages, such as SVG for instance. This approach was validated in a class with students learning web multimedia [5].

6.1.2. Multimedia content adaptation

Multimedia documents may have to be played on multiple devices such as mobile phones, tablets, desktop computers, set-top boxes, etc. Usage and platform diversity requires documents to be adapted according to execution contexts, sometimes unpredictable at design time. In a joint work with project-team Exmo, we have designed a semantic adaptation framework for multimedia documents. This framework captures the semantics of document composition and transforms the relations between media objects according to adaptation constraints [3].

6.2. XML Processing

Mature results about XML processing were obtained along three main directions: the formalization and implementation for checking the impact of schema evolution on validation and queries; logical extensions supporting a notion of counting and the shuffle operator in trees; and the decision of a subtyping relation for a very expressive type algebra supporting a notion of polymorphism.

[1] http://wam.inrialpes.fr/timesheets/

In addition, preliminary results were obtained on the definition of a rigorous logical framework for the static analysis of semantic web languages, on the static analysis of cascading style sheets, and on the equipment of an IDE with new static analysis features for XQuery.

6.2.1. Impact of XML schema evolution

In the ever-changing context of the web, XML schemas continuously change in order to cope with the natural evolution of entities they describe. Schema changes have important consequences. First, existing documents valid with respect to the original schema are no longer guaranteed to fulfill the constraints described by the evolved schema. Second, the evolution also impacts programs manipulating documents whose structure is described by the original schema.

We have proposed a unifying framework for determining the effects of XML Schema evolution both on the validity of documents and on queries [2]. The system is very powerful in analyzing various scenarios in which forward/backward compatibility of schemas is broken, and in which the result of a query may not be anymore what was expected. Specifically, the system offers a predicate language which allows one to formulate properties related to schema evolution. The system then relies on exact reasoning techniques to perform a fine-grained analysis. This yields either a formal proof of the property or a counter-example that can be used for debugging purposes. The system has been fully implemented and tested with real-world use cases, in particular with the main standard document formats used on the web, as defined by W3C. The system identifies precisely compatibility relations between document formats. In case these relations do not hold, the system can identify queries that must be reformulated in order to produce the expected results across successive schema versions.

6.2.2. Counting in trees

A major challenge of query language design is the combination of expressivity with effective static analyses such as query containment. In the setting of XML, documents are seen as finite trees, whose structure may additionally be constrained by type constraints such as those described by an XML schema. We have considered the problem of query containment in the presence of type constraints for a class of regular path queries extended with counting and interleaving operators [1]. The counting operator restricts the number of occurrences of children nodes satisfying a given logical property. The interleaving operator provides a succinct notation for describing the absence of order between nodes satisfying a logical property. We have proposed a logic supporting these operators, which can be used to solve common query reasoning problems such as satisfiability and containment of queries in exponential time [4].

6.2.3. Typing higher-order programs

We have considered a type algebra equipped with recursive, product, function, intersection, union, and complement types together with type variables and universal quantification over them. We have defined the subtyping relation between such type expressions, and have proved its decidability[9].

This has solved an open problem that was attracting a considerable research effort. The novelty, originality and strength of our solution reside in introducing a logical modeling for the semantic subtyping framework. We have modeled semantic subtyping in a tree logic and use a satisfiability-testing algorithm in order to decide subtyping. We have shown how the subtyping relation can be decided in EXPTIME. We have reported on practical experiments made with a full implementation of the system. This has provided a powerful polymorphic type system aiming at maintaining full static type-safety of functional programs that manipulate trees, even with higher-order functions, which is particularly useful in the context of XML.

6.2.4. Detection of inconsistent paths and dead code in XML IDEs

One of the challenges in web software development is to help achieving a good level of quality in terms of code size and runtime performance, for increasingly popular domain specific languages such as XQuery. We have presented an IDE equipped with static analysis features for assisting the programmer [8]. These features are capable of identifying and eliminating dead code automatically. The tool is based on newly developed formal programming language verification techniques, which are now mature enough to be introduced in the process of software development.

6.2.5. Static analysis of semantic web languages

We work with the Exmo project-team on the static analysis of semantic web languages such as RDF, OWL and SPARQL by investigating modal logics over graphs. We seek to build a rigorous logical reasoning framework based on μ-calculus adapted for the web semantic languages [7] [11]. In particular, we studied the containment problem for SPARQL queries: determining whether, for any graph, the answers to a query are contained in those of another query. Our approach consists in encoding RDF graphs as transition systems and queries as μ-calculus formulas and then reducing the containment problem to testing satisfiability in the logic.

6.2.6. Static analysis of style sheets

Developing and maintaining cascading style sheets (CSS) is an important issue to web developers as they suffer from the lack of rigorous methods. Most existing means rely on validators that check syntactic rules, and on runtime debuggers that check the behavior of a CSS style sheet on a particular document instance. However, the aim of most style sheets is to be applied to an entire set of documents, usually defined by some schema. To this end, a CSS style sheet is usually written w.r.t. a given schema. While usual debugging tools help reducing the number of bugs, they do not ultimately allow to prove properties over the whole set of documents to which the style sheet is intended to be applied.

We have proposed a novel approach to fill this lack [14] by analyzing CSS style sheets using the same logic and compile-time verification technique we use for other XML problems. We have developed an original tool based on our XML Reasoning Solver (see section 5.2). The tool is capable of statically detecting a wide range of errors (such as empty CSS selectors and semantically equivalent selectors), as well as proving properties related to sets of documents (such as coverage of styling information), in the presence or absence of schema information. This new tool can be used in addition to existing runtime debuggers to ensure a higher level of quality of CSS style sheets.

6.3. Multimedia Authoring

6.3.1. C2M project

The C2M project (see section 7.2.2) aims at developing industrial solutions that allow multimedia developers to achieve mass production with high quality results. It uses the SCENARI platform for document production and we have proposed a solution for dealing with multimedia content in such a framework [16]. Indeed, automatic tools are not always sufficient for generating high quality documents; manual editing of documents in their publishing format is often necessary to tune a number of details.

Our approach consists in providing a post-editing service to allow authors to adjust their multimedia presentations directly on the final form of documents. The first step is to provide a web rendering engine based on the latest advances in web standards, as described in section 6.1.1. The second step consists in designing web-aware authoring tools based on this library, thus providing authors with direct editing services for producing high quality multimedia documents while preserving the advantages of using an XML production workflow. We have developed a prototype of this authoring tool in which all editing templates are described with XUL (XML User interface Language) and XBL (XML Binding Language) elements that we have defined for handling time-based content and widgets (timeContainer, timeNode, timeLine, etc.).

With such a solution, we can combine two worlds: a semantic-oriented authoring approach, as provided by an XML document workflow, and a direct web-based editing system. The first guarantees homogeneous rendering while the latter enables direct adjustments on the final form of the document. Bridging these two worlds is made possible by using declarative web languages (namely HTML5, SMIL and CSS) and implementing their timing part in the browser (with the Timesheets.js library, see section 6.1.1). The authoring components are directly mapped to the document structures.

6.3.2. On-line editing of multimedia web content

In cooperation with EPFL (Lausanne) we have continued to explore the concept of template-driven editing for XML multimedia contents (see section 3.3.2). This year, we have carried out more experiments with very different types of contents, including structured documents, factual data, and multimedia objects [17].

These experiments have been done with the AXEL library developed by EPFL, based on our joint work on template languages. AXEL is an innovative client-side authoring tool that runs in the browser for editing XML documents, driven by an XTiger template. It allows average web users to easily edit XML content on web servers with no specific knowledge of XML. Our experiments have shown that the template-based approach significantly enhances the ability for web users to directly feed various applications with structured content.

6.4. Augmented Environments

A large part of the research on augmented environments specializes in the use of visual media. In WAM, we focus on the use of audio media and we put a strong emphasis on mobility.

We have developed the first indoor augmented reality audio navigation application running on personal AR devices such as mobile phones. The main idea behind the development of this application was a joint use of three concepts:

- Continuous localization by using embedded sensors together with physiological models of walking and assumptions about walking in structured indoor environments.
- OpenStreetMap Indoor Mapping used for map-aided positioning, assistive routing for visually impaired people, and environmental queries through audio panoramics.
- Guidance and navigation through AR audio, both 3D and environmental, with mixing of synthetic and natural sounds and support for timely audio information presentation.

We have demonstrated that these concepts are inter-dependent [12], and that bringing them together is a way to find new solutions to problems which are difficult to tackle when looking at them separately. These three concepts are implemented with web technologies we use XML languages and XML tools for interactive audio, building modeling, and personal navigation module configuration. This enables easy authoring of sound objects or audio icons used for building sonification [15], indoor navigation maps and panoramics, and walking models. Adaptability of navigation to preferences of people is based on the concept of audio stylesheets for OpenStreetMap data rendering, XML configuration of the Pedestrian Dead Reckoning module, and assistive routing specification.

We have developed two mobile browsers and a framework for generic navigation:

1. The Mixed Reality Browser [10] that we have developed can display PoI content either remotely through panoramics with spatialized audio, or on-site by walking to the corresponding place. MRB is the only browser of geolocalized data to use a declarative XML format for PoIs, panoramics, 3D audio and to be based on HTML5 both for the iconic and full information content of PoIs. MRB can be used for any kind of augmented reality visits. A cultural heritage visit of Grenoble (see section 7.1.2) has been realized with the tourist office of Grenoble and the CCSTI (Centre de Culture Scientifique Technique et Industrielle de Grenoble).
2. The Pedestrian Way Browser that we have developed can be used for indoor-outdoor navigation with assistive audio technology for visually impaired people on pedestrian ways with precise geospatial description. Its main characteristic is to be based entirely on the OpenStreetMap XML format for the representation of the route. We anticipate that in the context of the european project Venturi (see section 7.3.1), we will have a convergence of the MRB and PWB, allowing visually impaired people to undertake cultural heritage visit. An demonstration showing the use of the PWB in a structured outdoor environment is available online: http://www.youtube.com/watch?v=h2b8yfCauZ8
3. We have created an extensible client-server framework named TARA which allows navigation on an OpenStreetMap XML graph (indoor and outdoor) by computing routes in real-time. User preferences, like stairs versus lift, are supported through a ponderation of paths in the routing algorithm. The client is an HTML5 running in the browser on mobiles and desktop computers. It can therefore be used for simulation, to test or learn a route before the navigation in the real world. The user interface is based on three modalities, touch, audio and visual and can be operated by visually impaired people through VoiceOver using only touch and audio. Localization through

embedded or external sensors is not mandatory as step by step instructions can be accessed through touch modality. The server is a full REST server (Sling-Apache) giving priority to the representation of geospatial resources and allowing environmental queries through the use of XQuery.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. Autonomy

Participants: Audrey Colbrant, Yohan Lasorsa, Jacques Lemordant, David Liodenot, Mathieu Razafimahazo.

Autonomy is a 22-month project funded by the global competitiveness cluster Minalogic (6th call for R&D projects) that started in March 2010, and is coordinated by ST Ericsson. Other partners are ST Microelectronics, Raisonance, Université de Grenoble, and Ivès.

The goal of the project is to develop high-tech tools to improve autonomy for people with disabilities. These tools are integrated in mobile devices such as cell phones or special-purpose devices, to improve the quality of life of people with disabilities. These devices access remote dedicated services to help geolocation and guiding. They take advantage of the latest advances in embedded systems: cameras, audio, video, GPS, RFID, compass, accelerometer, gyroscope. Two major application areas are addressed: software tools on cell phones for sight disabled people, and guiding and information tools for moving around in a city.

7.1.2. Grenoble Augmented City

Participants: Audrey Colbrant, Yohan Lasorsa, Jacques Lemordant, David Liodenot, Mathieu Razafimahazo.

Grenoble Augmented City is a large scale experimentation of augmented reality technologies, funded by the Rhône-Alpes Region for 12 months. Coordinated by CCSTI Grenoble, the project includes GRESEC (Stendhal University, Grenoble), the public libraries of Grenoble, the tourist office of Grenoble, and musée Dauphinois.

This project uses XML formats for multimedia content (HTML5), interactive audio (A2ML), and points of interest (W3C POI) in complex mixed reality applications. As a consequence, the authoring of a specific application is greatly simplified.

The MRB browser developed by WAM is a Mixed Reality Browser whose main features are:

- use of an XML format for Points of Interest (POI) issued from an on-going discussion within the W3C Points of Interest working group,
- use of HTML5 for the multimedia content of POIs, allowing easy authoring inside a standard HTML5 browser,
- navigation between POIs at the level of the format using MAUDL audio POIs,
- switching between Augmented Reality and Augmented Virtuality through the use of panoramic images and the concept of groups of POIs.

The MRB browser is running on the iPhone 4 and the iPad.

A cultural heritage visit of Grenoble can be downloaded from the web and played on site in Augmented Reality or remotely in Augmented Virtuality. This visit will be tested by visually impaired people.

7.2. National Initiatives

7.2.1. Codex

Participants: Everardo Bárcenas Patiño, Nils Gesbert, Pierre Genevès, Nabil Layaïda.

Codex is a project funded by ANR as part of its Emerging Domains program (DEFIS). It started in March 2009 for a duration of 36 months. WAM is working with five partners: INRIA Saclay Île-de-France (project-team Leo), INRIA Lille Nord-Europe (project-team Mostrare), University Paris Sud, Centre universitaire de Blois, and Innovimax SARL.

Codex seeks to push the frontier of XML technology innovation in three interconnected directions:

- Languages and algorithms: prototypes are developed for efficient and expressive XML processing, in particular advancing towards massively distributed XML repositories.
- Codex considers models for describing, controlling, and reacting to the dynamic behavior of XML corpora and XML schemas with time.
- The project proposes theories, models and prototypes for composing XML programs for richer interactions, and XML schemas into rich, expressive, yet formally grounded type descriptions.

7.2.2. C2M

Participants: Fabien Cazenave, Cécile Roisin.

Multimedia Cooperative Publishing Chain (C2M) is a project funded by ANR as part of its Digital Contents and Interaction program (CONTINT). It started in September 2009 for a duration of 24 months. WAM is working with five partners: Université de Technologie de Compiègne, Kelis Conseil et Développement, Amexio, Heudiasyc laboratory (CNRS), Institut National de l'Audiovisuel (INA).

The project aims at integrating XML publishing chains, Enterprise Content Management (ECM), and multimedia creation tools, in order to design a complete digital system for multimedia creation, management and publishing.

The main challenge lies in the convergence of several approaches:

- storage and management of document fragments,
- structured editing,
- maintaining and repurposing content,
- planning, cooperation and production.

Convergence is made possible by the maturity of XML technologies and by the collaborative practices popularized by the web.

7.3. European Initiatives

7.3.1. Venturi

Participants: Yohan Lasorsa, Jacques Lemordant, David Liodenot, Mathieu Razafimahazo.

immersiVe ENhancemenT of User-woRld Interactions, FP7-ICT-2011-7, STREP, duration: 36 months starting in October 2011, Partners: Fondazione Bruno Kessler (Italy), Fraunhofer Heinrich Hertz Institute (Germany), ST Microelectronics (Italy), ST-Ericsson (France), Metaio (Germany), e-Diam Interactive (Spain), Sony-Ericsson (Sweden).

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.

7.4. International Initiatives

7.4.1. INRIA International Partners

We are working with the MEDIA group at EPFL (Lausanne, Switzerland) on XML editing, more specifically on the template-driven approach.

Members of the WAM project-team participate in several activities of the World Wide Web Consortium (W3C):

- Vincent Quint is a member of the W3C Advisory Committee.
- Nabil Layaïda is a member of the W3C Synchronized Multimedia working group.
- Jacques Lemordant is a member of the W3C Points of Interest Working Group.
- The Amaya web editor is developed jointly with W3C. The software is distributed by W3C.

8. Dissemination

8.1. Animation of the Scientific Community

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

N. Layaïda is a member of INRIA Grenoble-Rhône-Alpes Committee for Employment of Scientists: INRIA hiring of Post-docs, Delegates and Secondments.

N. Layaïda is an invited expert of the Qatar Foundation and Qatar Computing Research Institute.

P. Genevès was a project evaluator for the Qatar Foundation.

Jacques Lemordant is on the jury of the international ST Innovation Cup.

8.2. Teaching

License & Master

Licence: Computer Networks, 130h/year, L2, University of Grenoble – C. Roisin

Licence: Programming in Java, 20h/year, L3, University of Grenoble – C. Roisin

Licence: Web Technologies (XML, XSLT), 20h/year, L3, University of Grenoble – C. Roisin

Licence: Language Theory, 18h/year, L3, University of Grenoble – N. Gesbert

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

Master: Multimedia systems, 6h/year, M2, University of Grenoble – N. Layaïda

PhD & HdR

PhD: Everardo Bárcenas Patiño, Automated reasoning on trees with cardinality constraints, Grenoble University, 14 February 2011, advisors: Vincent Quint and Nabil Layaïda

PhD in progress: Melisachew Chekol, Reasoning on regular queries for graphs, since October 2009, advisors: Jérôme Euzenat and Nabil Layaïda

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] E. BARCENAS-PATINO. *Automated reasoning on trees with cardinality constraints*, Université de Grenoble, February 2011, http://hal.inria.fr/tel-00578972/en.

Articles in International Peer-Reviewed Journal

[2] P. GENEVÈS, N. LAYAÏDA, V. QUINT. *Impact of XML Schema Evolution*, in "ACM Transactions on Internet Technology", July 2011, vol. 11, n° 1, p. 4:1-4:27 [*DOI :* 10.1145/1993083.1993087], http://hal.inria.fr/inria-00619225/en.

[3] S. LABORIE, J. EUZENAT, N. LAYAÏDA. *Semantic adaptation of multimedia documents*, in "Multimedia Tools Appl.", 2011, vol. 55, n° 3, p. 379-398, http://dx.doi.org/10.1007/s11042-010-0552-9.

International Conferences with Proceedings

[4] E. BARCENAS-PATINO, P. GENEVÈS, N. LAYAÏDA, A. SCHMITT. *Query Reasoning on Trees with Types, Interleaving, and Counting*, in "22nd International Joint Conference on Artificial Intelligence IJCAI'2011", Barcelone, Spain, July 2011, http://hal.inria.fr/inria-00609556/en.

[5] F. CAZENAVE, V. QUINT, C. ROISIN. *Timesheets.js: Tools for Web Multimedia*, in "ACM Multimedia 2011", Scottsdale, Arizona, United States, November 2011, 4, http://hal.inria.fr/hal-00619396/en.

[6] F. CAZENAVE, V. QUINT, C. ROISIN. *Timesheets.js: When SMIL Meets HTML5 and CSS3*, in "DocEng 2011, 11th ACM Symposium on Document Engineering", Mountain View, United States, September 2011, 10, http://hal.inria.fr/hal-00619382/en.

[7] M. W. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA. *PSPARQL Query Containment*, in "DBPL 2011 The 13th International Symposium on Database Programming Languages", Seattle, United States, August 2011, 8, http://hal.inria.fr/hal-00619342/en.

[8] P. GENEVÈS, N. LAYAÏDA. *Inconsistent Path Detection for XML IDEs*, in "33rd International Conference on Software Engineering, ICSE 2011", Honolulu, United States, ACM, May 2011, p. 983-985 [*DOI :* 10.1145/1985793.1985967], http://hal.inria.fr/inria-00609513/en.

[9] N. GESBERT, P. GENEVÈS, N. LAYAÏDA. *Parametric Polymorphism and Semantic Subtyping: the Logical Connection*, in "International conference on functional programming", Tokyo, Japan, ACM SIGPLAN, September 2011, p. 107-116 [*DOI :* 10.1145/2034773.2034789], http://hal.inria.fr/inria-00585686/en.

[10] J. LEMORDANT. *Mixed Reality Browsers*, in "International workshop on AR Standards", February 2011, http://wam.inrialpes.fr/publications/2011/ARWorshopBarcelona2011.pdf.

Research Reports

[11] M. W. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA. *PSPARQL Query Containment*, INRIA, June 2011, http://hal.inria.fr/inria-00598819/en.

[12] A. COLBRANT, Y. LASORSA, J. LEMORDANT, D. LIODENOT, M. RAZAFIMAHAZO. *One Idea and Three Concepts for Indoor-Outdoor Navigation*, INRIA, December 2011, n° RR-7849, http://hal.inria.fr/hal-00654106/.

[13] P. GENEVÈS, N. LAYAÏDA, V. KNYTTL. *XML Reasoning Solver User Manual*, INRIA, June 2011, n° RR-6726, http://hal.inria.fr/inria-00339184/en.

[14] P. GENEVÈS, N. LAYAÏDA, V. QUINT. *On the Analysis of Cascading Style Sheets*, INRIA, November 2011, n^o RR-7808, http://hal.inria.fr/hal-00643075/.

[15] Y. LASORSA, J. LEMORDANT. *Le language audio pour mobiles MAUDL et son moteur de rendu audio interactif IXE*, INRIA, December 2011, n^o RR-7847, http://hal.inria.fr/hal-00654067.

[16] C. ROISIN, F. CAZENAVE, V. QUINT, L. GAILLARD, D. SAINT-MARTIN. *Enhancing an XML Publishing Workflow for Web Multimedia*, INRIA, July 2011, n^o RR-7836, http://hal.inria.fr/hal-00651543/.

[17] C. VANOIRBEEK, V. QUINT, S. SIRE, C. ROISIN. *A Lightweight Framework for Authoring XML Multimedia Content on the Web*, INRIA, December 2011, n^o RR-7848, http://hal.inria.fr/hal-00653965/.