INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team WAM

Web, Adaptation and Multimedia

Grenoble - Rhône-Alpes

Theme : Knowledge and Data Representation and Management

Table of contents

1. Team

Research Scientists

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

Faculty Members

Jacques Lemordant [Associate Professor, on leave from Université de Grenoble, HdR]
Cécile Roisin [Professor, Université de Grenoble, HdR]

Technical Staff

Fabien Cazenave [Engineer]
Audrey Colbrant [Engineer]
Yohan Lasorsa [Engineer]
David Liodenot [Engineer]
Mathieu Razafimahazo [Engineer]

PhD Students

Everardo Bárcenas Patiño
Melisachew Chekol [jointly with project-team EXMO]

Post-Doctoral Fellow

Nils Gesbert [from September 2010]

Administrative Assistant

Ahlem Zammit-Boubaker [shared with project-teams MOAIS and EXMO]

2. Overall Objectives

2.1. Objectives

Research in the WAM team aims at making it easier to use and develop rich multimedia contents and applications on the web.

There are already a number of specialized web sites for sharing pictures; there are other sites for video clips, still others for music, and so on, but on these sites each modality stands for itself, independently from the others. As opposed to this approach, our vision of the multimedia web emphasizes the tight integration of multiple modalities in a consistent set of resources, that takes advantage of the many possible interactions between different types of content. Available on the web, these multimedia resources are distributed, linked together, and use platform-neutral formats, that make them usable by anyone through any kind of terminal or network. In this vision, the web is not restricted to a one-way publishing medium for single-media content, but it becomes an open platform for producing, sharing, transforming, using and reusing multimedia contents and applications. In other words, the web is not only a giant repository, but it is first and foremost a platform for processing multimedia documents and data through many different kinds of applications. To realize this vision, a number of challenges have to be faced:

- A key issue is the formats used for processing and sharing multimedia contents. Formats must represent the many facets of multimedia contents. The ability to process these contents and to use it in various contexts strongly depends on the richness and versatility of formats. A research theme in WAM is dedicated to models and languages for representing rich multimedia web documents.
- XML is the ground on which these formats are built. Designed for the web, XML offers many features for taking advantage of the specificities of the web. The language is flexible and open enough for representing a wide variety of contents and data, and it comes with many accompanying languages and technologies that address issues such as defining schemas, mixing XML languages

(document formats), linking resources, querying and transforming data and documents, formatting and presenting documents, etc. As data and document formats for the web are based on XML, it is crucial to better understand how XML structures can be processed, and what are the theoretical tools that may help to develop an effective framework for safely processing XML structures. This is another research theme in WAM.

- Once rich formats are available, documents and data encoded in these formats have to be created. Given the richness of the formats, this is a challenge. Authors need help from specialized tools to benefit from the many, complex possibilities offered by web multimedia formats. Methods used for static, textual documents, such as WYSIWYG or direct manipulation, do not work for dynamic multimedia web documents. New approaches have to be developed and experimented. This is the third research theme in WAM.
- A fourth theme has grown recently in WAM, that consists in exploring a new family of multimedia applications, namely Augmented Environments. The idea is to build on the results of the previous three themes and to develop technologies and tools that make it possible to combine web resources and multimedia contents with data from the real world that surrounds the user. This will enable a broad variety of new applications where users interact with their environment, taking advantage of mobile devices with their various sensors and the huge amount of information available on the web.

3. Scientific Foundations

3.1. XML Processing

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

Since its introduction more than a decade ago, Extensible Markup Language (XML) has 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 standard type languages such as DTDs, XML Schemas, and Relax NG. XML processing can be seen as transforming these structures using tree-oriented query languages such as XPath and XQuery within full-blown transformation languages such as XSLT.

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 biggest 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. 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 web formats. 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.

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.

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 services based on a scientific approach. 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. Services 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.

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 of 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.

For our work on interactive audio and augmented environments, the application domains we address currently are mainly audio guides. In the area of accessibility, our objective with these guides is to help people to move around in a city, with the goal of bringing autonomy to people with disabilities, for instance to people who are blind. This also has a significant impact on sustainable development, as it eases and encourages complex multimodal commuting. In the area of cultural heritage, audio guides are also used to help people visit areas and buildings of particular interest.

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 researches 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. We are working with the French National Institute of Audiovisual (INA) for exploring new multimedia models and formats in this domain.

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 [5] is a tool for the static analysis of XPath queries and XML schemas based on the latest theoretical advances. 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, and Relax-NG).

The system has been 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 a declarative SMIL markup. Using a declarative language makes sense for all 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.

This project is a follow-up of the LimSee3 authoring tool, where a first version of the SMIL Timesheets scheduler was provided. The plan is also to address authoring needs related to main classes of HTML5 documents such as HTML subtitles, media annotations, slideshows and course or conference recordings.

5.4. ARIA

Participants: Yohan Lasorsa, Jacques Lemordant.

The ARIA library (A2ML Reference Implementation API) is an open source API for using A2ML documents on mobile devices. It makes it easy to add advanced interactive audio to any mobile application through the use of the A2ML format. It can be used in various application domains such as games or auditory guidance, for example. Synchronization of the sound manager is done using a SMIL-compatible scheduler with additional event-driven functions to export sound events back to the application. SMIL-like animations are also supported through the same synchronization paradigm. The current version of the API supports iOS and Mac OS X. Further support for other mobile platforms (Android, Symbian) is planned.

6. New Results

6.1. Static Analysis Techniques for XML Processing

6.1.1. An IDE for XQuery

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. A peculiarity of XQuery code is that XQuery programs are very commonly written against a given XML schema that defines constraints with which the queried documents must comply. This provides a dead-code prone setting, as XQuery path expressions may contain navigational information that contradict requirements expressed by the schema. In that case, the result of the path expression is always empty (no matter what the input data are), and all XQuery instructions that depend on this sub-expression are dead code.

We have developed an IDE [4] as an extension to the XQuery Development Toolkit (XQDT), a popular XQuery Eclipse plugin with full support for XQuery 1.1. The extension consists in static analysis features for assisting the programmer. These features are capable of identifying and eliminating dead code automatically. The tool is based on formal programming language verification techniques we are developing, which are now mature enough to be introduced in the process of software development. The IDE has the capabilities of statically detecting inconsistent paths. It provides code completion and code templates, as-you-type validation, and integration with existing XQuery evaluation engines. For a given path expression e and a schema S, the IDE checks for the inconsistency of e in the presence of S. The analysis functions mark inconsistent paths and dead-code with syntax coloring capabilities offered by the IDE plugin.

Theses functions operate on the abstract syntax tree of the program to identify XPath expressions. Then the evaluation context of each XPath expression is built. Since XPath host languages (XSLT, XQuery) allow variables to be defined and then used in paths, the tool maintains the correspondence between variables occurring in paths by their values. When the static verification is triggered, each XPath expression and its evaluation context, as well as the schema chosen by the programmer, are transmitted to the plugin. The analysis is then performed and the UI updated with the corresponding analysis result: line numbers of dead code, syntax coloring of dead code and highlighting of inconsistent paths.

6.1.2. Schema Evolution

In the document world, a driving concern is long term access to content, i.e. the ability to process content, for example a web page or a scalable vector graphic, written today, in say several decades. This major concern gave birth to SGML, and more recently to XML, where the idea is to separate data structures from processor-specific instructions. For this purpose, the essence of XML consists in organizing information in tree-tagged structures conforming to some constraints, which are expressed using standard type definition languages such as DTDs, XML Schema and Relax NG. A set of constraints define a class of documents.

When writing a normative schema such as a W3C Recommendation (XHTML, SVG, MathML, SMIL, etc.), an important issue is then to make sure that schemas written in different languages are equivalent, that is they describe the same structure, possibly with some minor differences due to the expressivity of the language for defining tree grammar based constraints. Another issue is to precisely identify the differences between two versions of the same schema expressed in different schema languages. Moreover, the issues of forward and backward compatibility of document instances obviously are posed when schema languages change from a version to another.

We have developed a tool [6] for helping XML schema designers to obtain a high quality level for their specifications. It has been implemented as a Java/JSP web application and user interaction is offered through a web user interface in a web browser. The tool allows one to analyze relations between classes of XML documents and to formally prove them. When such a relation does not hold, the tool provides the reasons and reports detailed counter-examples that exemplify the problem.

The tool is based on a set of predicates offered to XML application and schema designers for assessing the correctness of schema evolutions, with respect to the properties described above. These predicates can be combined with logical connectives (e.g. disjunction, conjunction, negation...) in order to put the focus on peculiarities of a schema and its variants.

The tool was used to investigate forward and backward compatibilities of very popular W3C recommendations [1]. We have been able to successfully identify inconsistencies in normative documents for HTML, MathML and SVG. We believe this type of tools can be of great value for standardization bodies that define specifications using various XML type definition languages (such as W3C specifications), and are concerned with quality assurance for their normative recommendations.

6.1.3. Counting in Trees

Regular tree grammars and regular path expressions constitute core constructs widely used in programming languages and type systems. Nevertheless, there has been little research so far on frameworks for reasoning about path expressions where node cardinality constraints occur along a path in a tree.

Our work in this area is motivated by the need to tackle one of the major drawbacks of schema languages such as XML Schema and Relax NG with respect to counting constraints. The reason behind this lies in the fact that schema designers usually describe types in the forward direction starting from the root and by declaring the content models of every sub-schema fragment. If the type of these fragments is sensitive to the context, or if the schema has occurrence constraints, the context and counting information is propagated across the nesting elements. This process yields schemas and sub-schemas that can be hardly reused and extremely verbose. This is typically the reason why the standard DTD for HTML does not syntactically prevent the nesting of anchors, whereas this nesting is actually prohibited in the standard. Instead, it may appear tempting to rely on another kind of formal notation that just describes a simple pattern and additional constraints on it, which are both intuitive and compact. For instance, one could imagine a notation where the additional constraint is described using XPath notations:

$$\left(x \rightarrow (a[x] \mid b[x] \mid c[x])^*\right) \ \wedge \ \text{count(/descendant-or-self::}b) \leq 2$$

We have designed a logic [12] capable of efficiently expressing deep counting along paths which may include arbitrary recursive forward and backward navigation. The counting extensions can be seen as a generalization of graded modalities that count immediate successor nodes. While the combination of graded modalities, nominals, and inverse modalities yields undecidable logics over graphs, we have shown that these features can be combined in a decidable tree logic whose main features can be decided in exponential time. The logic being closed under negation, it may be used to decide typical problems on XPath queries such as satisfiability, type checking with relation to regular types, containment, or equivalence.

One of the advantages of this logic is its ability to capture both paths and types. It can then be used to develop schema languages that use types and path expressions. An interesting alternative is to design new schema languages where the context is captured by means of backward path expressions declared from within the sub-schema. Such languages are more modular in the sense that schema fragments can be extracted and re-used in other contexts while retaining their contextual constraints. This will also help building some kind of modular schema libraries that can be used as building blocks for more involved ones.

6.2. Multimedia Authoring

In cooperation with the Media group of EPFL, we have developed a framework [10] for editing, publishing and sharing XML content on the web directly from the browser. The two main components of this framework are the XTiger XML language and the AXEL library.

XTiger XML is a slight extension to the original XTiger language. The extension is a very simple mechanism that specifies how an XTiger structure is translated into an XML structure. With this extension, XTiger becomes a language for creating XML document models. These models contain the rules for creating both a XML structure and its XHTML representation. The latter allows a web browser to display the document, while the former can guide an editor for building an XML document with a predefined structure specified by the XTiger XML template.

AXEL is a Javascript library that implements an editor for any XML structure defined by a XTiger XML template. It turns the document template into a document editing application running in a web browser.

This framework is targeted at non XML speaking end users, since it preserves authors from XML syntax during editing and does not require them to master any other XML technology, such as schemas or transformations. Its current implementation creates a pseudo-WYSIWYG user interface where the document template provides a document-oriented editing metaphor, or a form-oriented metaphor, depending on the template and the data to be edited.

The AXEL library is complemented with a mechanism for XML schema generation [9]. This is based on XSLT transformations that turn the XML structure definition represented by the XTiger template into the equivalent XML schema. This schema can then be used to validate the XML content created by the AXEL library, or for any other purpose.

This tool provides an easy way to edit XML content directly on the web, with the usual benefit of valid XML content. It makes it possible to create and publish content targeted for lightweight web applications, and the template-driven editing approach allows any web user to easily enter content while schemas automatically generated from templates make sure various applications can safely process this content.

6.3. Augmented Environments

6.3.1. Augmented Reality

We have been working on audio as a primary modality for augmented reality applications. The term Augmented Reality Audio (ARA) characterizes techniques where a physically real sound and voice environment is extended with virtual, geolocalized sound objects [8]. The application we use for testing is a global (indoor and outdoor) navigation system for visually impaired people. One of our ideas was to extend the use of OpenStreetMap to navigation inside buildings. This offers a continuum to the navigation, and existing tools for outdoor GIS authoring can be used for indoor authoring.

We have shown [7] that the authoring of an ARA scene, sometimes called mobile mixing, can be done through an iterative process composed of two stages:

- in the first stage the author has to go to the rendering zone, to apprehend the audio spatialization and the chronology of audio events;
- in the second stage, a textual editing of the sequencing of sound sources and DSP acoustics parameters is done.

This authoring process is based on the joint use of two XML languages, OpenStreetMap for maps, and A2ML for interactive 3D audio. A2ML being a format for a cue-oriented interactive audio systems, we have found that audio POIs (Points of Interest) can be computed from OpenStreetMap through the use of a tag-based cue dispatching XML language.

An interactive audio system allows changes in input behavior to modify the audio behavior, whereas a reactive system simply plays back static audio events without any adaptation to the user stimulus. An ARA guidance application cannot be only reactive and the behavior of sound objects has to be modified according to the context. Our A2ML language for interactive audio supports this nicely.

We have implemented an experimental ARA system running on the iPhone which provides the user with a continuous sense of relevant information for indoor navigation. We are in the process of building a catalog of sound objects corresponding to OpenStreetMap features in collaboration with partners of the Autonomy project (see section 7.1.1). At the rendering level, an external head-tracker, used to stabilize the soundscape, has been designed for the iPhone in the Apple Made-For-iPhone (MFi) program. The first prototypes will be out in January 2011.

6.3.2. Augmented Virtuality

Augmented Virtually (AV) is the natural complement to Augmented Reality (AR) when considering the authoring of AR content or the field of collaborative AR. A2ML being very close to SVG at the conceptual level (embedded SMIL animation, def/use mechanism, etc.), it was natural to consider using A2ML in synchronization with SVG, A2ML being used for the Augmented Reality aspect and SVG for the Augmented Virtually aspect.

We have shown [3] that the best way to use SVG and A2ML together is, like with OpenStreetMap, to synchronize them through a tag-value dispatching language. Synchronization of animation has also been considered. To test the potentiality of the joint use of both languages, we have implemented an SVG remote console for a navigation system for visually impaired people based on OpenStreetMap. A demo has been given at W3C TPAC in November 2010: the remote console in Lyon was used to control a semi-autonomous visually impaired pedestrian inside the INRIA building in Grenoble.

Panoramic images are another type of media with many uses in both fields of Augmented Reality and Augmented Virtuality. Panoramic images can be used in the background remotely (augmented virtuality) or instead of live video (augmented reality). We have developed tools for the navigation of panoramic scenes augmented with 3D audio data. A2ML sound objects have been located into panoramic images, allowing us to support the authoring of mixed reality experiences that leverage the live channel data in various ways, at the physical site and for remote viewing. The use of panoramic images has been extended to the concept of geospots and remote panoramic servers for a mixed reality browser.

6.3.3. Mixed Reality Browser

The term Mixed Reality Browser (MRB) refers to systems that combine physical and virtual worlds as Mixed Reality (MR). MR techniques have now been explored in a wide variety of domains, including navigation, maintenance and repair, air traffic control, urban design, office work, games, and medicine. One of the keys to success is whether MR models (including geographical models, sound objects, structured content and semantic models for POIs) can be authored, aggregated and distributed through adequate channels.

Our work on Mixed Reality Browsers is based on a vision of a distributed world of POIs having contextual metadata associated with them and linked to web pages containing textual, visual and auditory information objects. Metadata can be used to decide when and where to present the information objects to users. Such a vision is sometimes referred as the Real-World Wide Web (RWWW) and interfaces to it as RWWW browsers. We prefer to use the term Mixed Reality Browser. In a MRB, web pages are accessed through real-world locations by moving in the real world and using POIs coming from specific or semantic channels as proxies to the web pages. An MRB accesses new servers that deliver specialized content, such as panoramic images, POIs servers, and SPARQL aggregating servers.

We have built a prototype MRB to re-engineer the guidance application for visually impaired people developed in the Autonomy project (see section 7.1.1). This complex guidance application is very useful to understand what are the basic components required by an MRB. The component for Sensors Stream Fusion (SSF) is the key component of a framework on which an MRB can be built. The output of SSF is triples composed of position, orientation and behavior. Its input is made of streams of data arriving from embedded sensors and external sensors specified by URIs, like any other web resource. We are now in the process of designing an XML language to control this network of streams. Our prototype SSF module is based on physiological models of walking.

7. Other Grants and Activities

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.

Large scale experimentation of augmented reality technologies, Rhône-Alpes Region Call for Projects 2010.

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 POAP 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 inside 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 A2ML audio POIs,
- switching between Augmented Reality and Augmented Virtuality through the use of panoramic images and the concept of groups of POIs.

The POAP 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

We are working with the MEDIA group at EPFL (Lausanne, Switzerland) on XML editing, more specifically on the template-driven approach presented in section 6.2. We have jointly published two papers on this topic this year: [9], [10].

7.4. International Initiatives

7.4.1. World Wide Web Consortium

Members of the WAM project-team participate in several activities of 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.

7.4.2. Audio Community

Jacques Lemordant is a member of the AES (Audio Engineering Society).

8. Dissemination

8.1. Conferences, Meetings and Tutorial Organization

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

8.2. Teaching

Pierre Genevès gives Master lectures on Core XML technologies and their theoretical foundations at Grenoble Universities (Master of Science in Informatics at Grenoble, final year).

Nabil Layaïda gives Master lectures on SMIL and Multimedia Principles at Grenoble Universities (Master of Science in Informatics at Grenoble, M2R SIGAL: UE IST, final year).

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] P. GENEVÈS, N. LAYAÏDA, V. QUINT. *Impact of XML Schema Evolution*, in "ACM Transactions on Internet Technology (TOIT)", 2011, to appear.

[2] S. LABORIE, J. EUZENAT, N. LAYAÏDA. *Semantic adaptation of multimedia documents*, in "Multimedia Tools and Applications", June 2010, p. 1-20 [*DOI :* 10.1007/s11042-010-0552-9], http://hal.inria.fr/hal-00496695/en.

International Peer-Reviewed Conference/Proceedings

[3] A. COLBRANT, Y. LASORSA, J. LEMORDANT, D. LIODENOT, M. RAZAFIMAHAZO. *Sound Objects for SVG*, in "SVG Open 2010", Paris, France, Telecom Paris, August 2010, https://www.svgopen.org/2010/papers/34-Sound_Objects_for_SVG/, http://hal.inria.fr/inria-00540356/en.

[4] P. GENEVÈS, N. LAYAÏDA. *Eliminating Dead-Code from XQuery Programs*, in "32nd ACM/IEEE International Conference on Software Engineering", Cape Town, South Africa, ACM, May 2010, p. 305-306 [*DOI :* 10.1145/1810295.1810363], http://hal.inria.fr/hal-00494233/en.

[5] P. GENEVÈS, N. LAYAÏDA. *XML Reasoning Made Practical*, in "26th IEEE International Conference on Data Engineering, ICDE 2010", Long Beach, CA, USA, IEEE, March 2010, p. 1169-1172 [*DOI :* 10.1109/ICDE.2010.5447786], http://hal.inria.fr/hal-00494261/en.

[6] N. LAYAÏDA, P. GENEVÈS. *Debugging Standard Document Formats*, in "19th International Conference on World Wide Web, WWW 2010", Raleigh, NC, USA, M. RAPPA, P. JONES, J. FREIRE, S. CHAKRABARTI (editors), ACM, April 2010, p. 1269-1272 [*DOI :* 10.1145/1772690.1772899], http://hal.inria.fr/hal-00494245/en.

[7] J. LEMORDANT, Y. LASORSA. *Augmented Reality Audio Editing*, in "128th AES Convention", London, UK, May 2010, paper 8143, http://hal.inria.fr/hal-00494239/en.

[8] J. LEMORDANT. *Basic Concepts in Augmented Reality Audio*, in "W3C Workshop: Augmented Reality on the Web", Barcelona, Spain, June 2010, http://hal.inria.fr/hal-00494276/en.

[9] V. QUINT, C. ROISIN, S. SIRE, C. VANOIRBEEK. *From templates to schemas: bridging the gap between free editing and safe data processing*, in "10th ACM symposium on Document engineering", Manchester, UK, ACM, September 2010, p. 61-64 [*DOI :* 10.1145/1860559.1860572], http://hal.inria.fr/inria-00522175/en.

[10] S. SIRE, C. VANOIRBEEK, V. QUINT, C. ROISIN. *Authoring XML all the Time, Everywhere and by Everyone*, in "XML Prague 2010", Prague, Czech Republic, Institute for Theoretical Computer Science, Prague, March 2010, p. 125-149, http://hal.inria.fr/hal-00494255/en.

National Peer-Reviewed Conference/Proceedings

[11] J. MIKÁČ, C. ROISIN, J.-M. WOLFF. *Instrumentation des interrogations orales Support multimédia à l'évaluation et à la capitalisation des interrogations orales*, in "TICE2010 : 7ème Colloque Technologies de l'Information et de la Communication pour l'Enseignement", December 2010, http://hal.inria.fr/inria-00540393/en.

Research Reports

[12] E. BARCENAS, P. GENEVÈS, N. LAYAÏDA, A. SCHMITT. *On the Count of Trees*, INRIA, August 2010, n^{o} RR-7251, http://hal.inria.fr/inria-00473160/en.