INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team WAM

Web, Adaptation and Multimedia

Rhône-Alpes

THEME 3A

Table of contents

1. Team

Head of project-team
Vincent Quint [Research Director]

Administrative assistant
Marion Ponsot

Scientific staff
Nabil Layaïda [Research Scientist]
Ethan Munson [Visiting scientist from University of Wisconsin]
Vincent Quint [Research Director]
Cécile Roisin [Professor P. Mendès-France university]
Jean-Yves Vion-Dury [Senior research staff]

Technical staff
Paul Cheyrou-Lagrèze [Junior technical staff, until August 2003]
Stéphane Gully [Junior technical staff, since September 2003]
Peter Hewat [Project technical staff, since October 2003]
Vincent Kober [Technical staff]
Irène Vatton [Senior technical staff]
Daniel Weck [Software development staff]

Post-doctoral fellow
Tien Tran-Thong [Industrial post-doc]

Ph. D. students
Frédéric Bes [Teaching Assistant]
Pierre Genevès [MENRT grant, since November 2003]
Tayeb Lemlouma [INRIA grant]

2. Overall Objectives

Project WAM (Web, Adaptation and Multimedia) was created in early 2003 to explore the field of adaptative multimedia on the Web, with a special focus on structured documents transformation and adaptation.

2.1.1. A Web of Diversity

Diversity on the Web increases steadily, be it the diversity of information or the diversity of access devices and communication networks. Diversity of information comes from the multimedia Web. Information shared on the Web is made of text for a significant part but also of pictures, drawings, video, animations, music, voice, etc. These media can just stand independently from each other, such as a movie or a recorded song, but they can also interact with each other in genuine multimedia documents that tightly integrate pieces of information in different media.

Devices also are multiple. The workstation or personal computer that was typically used in the early days of the Web is no longer the dominant access device. In Japan, for instance, there are now more cell phones than PCs used to access the Web. TV sets are also following this trend. The recent developments in the TV industry clearly show the convergence between television and the Web. Web users can watch at TV programs on their desktops while TV sets can now be used to access Web sites. The digital television technology is borrowing more and more techniques from the Web, such as XML, for instance. The automotive industry is also developing embedded devices that provide access to the Web. The Web is now ubiquitous and all sorts of devices with very different capabilities are involved in Web access.

Simultaneously, these devices are using new kinds of networks, ranging from personal networks such as Bluetooth to the global Internet. In the broad range of communication technologies, wireless and mobile network technologies, such as GPRS or WiFi, are taking an increasing part. Their original features make a big

change from the traditional wired Internet and have a strong impact on the way information is exchanged over the Web.

2.1.2. Need for Adaptation

The increasing diversity of the information interchanged and the hardware and channels involved makes the original scheme of the Web inefficient. The usual model of a single Web page targeted at a large color screen accessed through a high speed network does not work any more. Some information providers face this problem by developing their contents into different versions, each one suited to a specific class of devices. Another approach is to create the information in a single, universal format and to adapt it automatically to the environment where it is delivered.

Project WAM works along the second approach. It aims at developing models, methods, architectures, protocols, formats, languages that allow content to be adapted "on the fly" to the context in which it is actually used. In this approach, no restriction is put on the type of the information that has to be adapted; multimedia information is considered with the broad diversity of media that are now commonplace on the Web.

Content adaptation is not something that comes into play at the last moment, when information has to be sent to the client. To allow efficient adaptation at that time, it is important that the original information presents some features that make adaptation easier or even possible at all. This means that the production methods should also be involved in the whole process of content adaptation. Project WAM is especially interested in the authoring tools for the Web with the perspective of creating truly multimedia documents that will ease adaptation and will improve device independence. Two editors are currently under development. LimSee2 is dedicated to the production of multimedia documents in the SMIL 2.0 format, while Amaya targets multi-namespaces XML documents containing text, mathematics, and animated graphics and using CSS style sheets.

2.1.3. XML Transformations

It is clear that a major means to adapt documents is to transform them according to the current context. The project focuses on structured multimedia documents represented as XML structures. Regarding transformations, the objective is first to characterize the theoretical and practical tools needed for efficiently transforming XML structures, and then to develop models, formalisms and algorithms that are necessary for transformation languages.

A strong motivation for this research on transformations is adaptive multimedia, but transformation of XML documents and data has actually a broader range of applications. Transformations are ubiquitous in the processing of structured information on the Web, ranging from formatting to repurposing and life cycle management. Actually, XML transformations are considered as a key paradigm for document processing.

3. Scientific Foundations

3.1. Transformation Languages

Participants: Pierre Genevès, Nabil Layaïda, Vincent Quint, Jean-Yves Vion-Dury.

Key words: *document models*, *document transformation*, *transformation languages*, *path expressions*, *XML structures transformation*, *XPath*.

An important activity in the WAM project is the design of a new transformation language for XML structures called Omega.

3.1.1. Features of Transformation Languages

Structure transformation is a specific application domain that can be approached following different abstraction levels with respect to programming specifications. The lowest level is based on general purpose languages, such as Python or Java, associated with dedicated libraries and toolkits that implement a standard structure manipulation API, typically DOM. On the opposite, there are specific languages, such as XSLT, which abstract over data and control complexity through a tree based data model and a powerful execution model. Some

properties are expected from specialized languages in order to help solving most common problems in the simplest way: expressiveness, verifiability, efficiency, modularity, reusability, scalability, simplicity, clarity, robustness, completeness, etc.

For XML structure transformations, some more properties are of particular importance:

- *Type checking*: Few languages are able to use typing information related to the input document, and even fewer are able to perform type checking with respect to the output document (XDuce, CDuce, Circus, XQuery). Here, we understand types as the structural constraints over documents as expressed in a DTD or an XML schema.
  A challenge is to address "late type checking": the ability to ensure the final type soundness of a chain of transformation steps, while allowing some steps to temporarily violate the typing constraints.

- *Bidirectionality:* Transformations are usually thought in one direction only, whereas both directions could be useful in many cases. As an example, transforming a content-oriented document (e.g. a form) into a presentation-oriented document (e.g. a HTML document), may allow the reader to perceive the required information to be fulfilled, thanks to the HTML rendering. After editing the HTML document, a reverse transformation could update the original document.

- *Efficiency:* Transformation code needs powerful static analysis and deep optimizations in order to reach high efficiency. Such techniques take advantage from clear and powerful formal semantics to help compilation. Reusing well known techniques is a way to minimize the back-end development effort.

- *Streamability*: Many transformation languages assume that the whole structure to be transformed is available when the transformation process runs. In streaming applications however, documents may be very large or even endless and the transformation has to be performed on the fly, with a somewhat limited scope.

3.1.2. Basic Principles for Omega

The most basic idea of Omega is to strongly rely on XPath, more precisely, on a fragment of the XPath language that offers adequate formal properties. XPath is a node selection language for XML that brings a great simplicity by focusing on paths instead of structures, making powerful hypothesis and metaphors such as implicit context node and navigation. XPath is involved in many other languages (XSLT, XQuery, XML Schema, etc.) and widely known, adopted by many programmers and also recognized as being mathematically tractable. In most common usages, XPath is just used for node selection in XML structures. In Omega, it is also used to build and assemble tree fragments that constitute the output document.

The transformation model is based on an input document, invariant with respect to the transformation, and an output document, that is rewritten, possibly using information extracted from the input document. This model accounts both for source and target driven transformations. This clearly distinguishes Omega from languages such as XSLT, XQuery, Circus, XDuce or CDuce, as they all use a monotonic result tree construction scheme, based on recursive embedding. As a computation model, rewriting has some good properties for a structure transformation language: decomposition in small basic units (rules), structural matching/filtering, modularity, and static analysis through unification.

Omega is a rule-based language. Rules for specifying unidirectional transformations have the form $\mathrm{l}, \mathrm{l}' \rightarrow \mathrm{r}'$, where l is an XPath expression applied on the input document, l' is an XPath selection applied to the output document and r' is an XPath expression that defines updating operations that will be applied to the output document if the rule matches. Updates are a collection of node deletions and additions deduced from l and l' .

Bidirectional rules have the form $l, l' \leftrightarrow r, r'$ and can be understood as the decomposition $l, l' \rightarrow r'$ and $r', r \rightarrow l$. Such a notation puts the emphasis on structural similarities between two trees, abstracting over the transformation operation. Doing so, it should simplify the specification of symmetrical transformations, were both directions are similarly important.

Rules describe atomic transformation units. Strategies are added to rules to control the execution of a complete transformation. A strategy specifies the evolution of context nodes, depending on the results of sub-transformations. Rewriting strategies have been extensively studied in the domain of rewriting calculus. Our approach is to strongly separate rules from strategies, whereas they are tightly linked in common rewriting systems.

Deterministic transformations become quickly very complex in some applications. The programmer must handle all cases. An alternative approach is to use backtracking in order to explore a transformation space. For example, there are two ways to describe the validity of a tree fragment in a transformation. The deterministic one ensures that the structural constraints are satisfied at each step, and following a predefined rule application order. Non deterministic transformations search a way to combine sub-transformations in order to fulfill the constraints.

The final major feature of Omega is static verification. Omega aims at checking structural constraints expressed by DTDs or schemas, but also properties of the transformation process itself, such as termination, or properties of content after transformation.

3.2. Adaptation

Participants: Nabil Layaïda, Tayeb Lemlouma, Tien Tran-Thuong.

Key words: *World Wide Web*, *multimedia*, *adaptive multimedia*, *device independence*, *adaptation*, *authoring*, *document format*.

The goal of multimedia document adaptation on the Web is to take advantage of the numerous devices and networks that are now constituting the Web, and to get rid of the traditional model of a PC accessing servers through the wired Internet. In other words, the Web infrastructure has to change gradually to a device independent architecture, where information resources can be efficiently accessed through many kinds of devices and networks.

To reach this goal, project WAM is following two complementary approaches, a comprehensive, global approach that could require significant changes to various aspects of the current Web infrastructure, and a restricted, pragmatic approach that avoids such changes by using the current infrastructure and adapting existent contents.

There is no general solution today to the problem of device independence. Most efforts are rather dedicated to the development of good practices. To make progress towards a solution, three aspects are considered in the project: device independent formats, profile negotiation and transcoding.

3.2.1. Device Independent Formats

Ideally, content would be created or generated in a single universal format that could be delivered as is to any conceivable device. In practice this seems impossible, so the real techniques seek to minimize the number of variants needed, each variant being targeted at as wide a range of devices as possible. The problem becomes more complex when taking into account the various pieces of information that are needed to satisfy a user request for a Web resource. These include various types of media that constitute the content of a Web page as well as all supporting material such as style sheets that determine how the content is rendered.

The creation of such material is called authoring. Typically the authoring process must adapt material for delivery to devices. The construction of the adaptation processes themselves may also be regarded as a form of authoring, though more usefully regarded as a form of programming. Authoring of raw material, style sheets, layouts and adaptation processes may be performed by different individuals with different skills. The challenge is to facilitate this separation of concerns.

3.2.2. Profile and Negotiation

So, several variants of the same Web resource are available on a server or can be generated. The next problem is to decide which one has to be delivered on a user's request.

This problem is addressed through profiles and negotiation. A profile is a formal representation of the context in which content is used: user abilities and preferences, device capabilities and limitations, network characteristics, etc. These descriptions must be both static and dynamic as the context may change over time a Web resource is accessed and used. Although some proposals are already available to describe profiles, such as UAProf or CC/PP, more research is needed to finely and comprehensively describe a context.

Profiles are used in a negotiation in order to find the best suited form of the information to be delivered in a given context. Negotiation may be complex as it may require various pieces of profiles stored in different repositories. The manufacturer of the device or the access provider may have their own profile repositories, while the user profile will reside somewhere else. The negotiation then becomes a distributed application involving multiple information bases.

The goal of the negotiation is to establish where the adaptation must occur between three interacting entities: the user agent, the proxy and the server. A successful negotiation allows adaptive content to be delivered while optimizing the performance of the overall system in terms of latency, bandwidth, and CPU.

Negotiation is closely related to the request for the information resource. The HTTP 1.1 transport protocol, for instance, caries some negotiation information such as language preference, in the same message as the request itself. For more sophisticated negotiations, transport protocols have clearly to be extended. An alternative is to develop new negotiation protocols that interact with transport protocols.

3.2.3. Transcoding and Transformation

Negotiation allows a server to select the right form of information to be delivered. The next problem is to generate this form. It could be prepared in advance or transformed on the fly according to the outcome of the negotiation process, and based on an adaptable format.

To transform multimedia documents, one can rely on their semantics. The semantics considered here do not deal with the document content, but with the composition that is made explicit in a Web document:

- temporal semantics: in what order and when should each piece of information be presented to the user,
- spatial semantics: what are the relative positions of the document components on the display space,
- navigational semantics: how are pieces of information related in the hypertext network.

This approach allows the definition of adaptation in very general semantic terms, independently from the multimedia objects. The temporal dimension is of particular importance in multimedia documents and is probably the most challenging one.

Another way to transform Web documents is offered by the Web formats. Most of these formats are modular and are defined as profiles that assemble different sets of modules. XHTML, SMIL and SVG, for instance, have Basic profiles for mobile devices. SVG, the vector graphics format for the Web, has also a Tiny profile for low-end mobile devices. With such profiles, adaptation may be achieved by transforming a document from one profile to another, better suited to the device involved.

Finally, adaptation is also performed to cope with the quality of service (QoS) offered at any moment. This low-level adaptation is done by transcoding:

- Content can be annotated to allow the sender to pick alternative fragments of information in different encodings, or to change encoding or resolution, depending on the QoS.

- The choice of the encoding has to take into account the capabilities of the receiver as expressed in the negotiation.
- Hierarchical coding techniques can be used, as in scalable formats ala JPEG 2000.

3.3. Multimedia Documents Authoring and Formatting

Participants: Frédéric Bès, Paul Cheyrou-Lagrèze, Stéphane Gully, Peter Hewat, Vincent Kober, Nabil Layaïda, Vincent Quint, Cécile Roisin, Tien Tran-Thuong, Irène Vatton, Daniel Weck.

Key words: *multimedia*, *editing*, *authoring environment*, *structured editing*, *constraints*, *timeline*.

3.3.1. Authoring

Before a structured multimedia document can be transformed or adapted, it has to be created. Project WAM works on interactive authoring environments to help creating documents. Developing such environments is a challenging issue: structured multimedia documents are complex objects and the process of creating and updating them is complex too. Well established paradigms for office documents or static technical documents do not work. The WYSIWYG approach does not apply in a context where the final form of the document (What You Get) is multiple and unknown at creation time. Writing a description of a document in some formatting language is extremely difficult given the various levels of description that are involved: content, logical structure, layout, style, synchronization, hypertext structure, dynamic behaviours, etc. New approaches are needed.

On the Web, multimedia documents are based on XML. They are considered through several types of structures: layout, time, hypertext, animations. Following this approach, we develop techniques that allow users to manipulate all these structures in homogeneous environments. The key idea is to present simultaneously several views of the document, each view showing a particular structure and allowing the user to manipulate it directly. In addition, these views are "synchronized" to show the consequences of every change in all other views. The XML markup, although it can be accessed at any time, is handled by the tools, and the user does not have to worry about it.

Three editing tools based on this concept are currently under development, Amaya, LimSee2, and Satin. In Amaya the emphasis is put on the integration of several XML vocabularies and associated technologies, and on direct interaction with the Web: the user can edit remote documents in exactly the same way as local files. LimSee2 is dedicated to a single XML language, SMIL, and the focus is on the time dimension of multimedia document and their dynamic contents. In Satin time and synchronization are also in the focus, as well as user interaction, but the target is broadcast TV, while LimSee is Web-oriented.

3.3.2. Formatting

Formatting plays an important role in structured multimedia documents. Media objects have to be arranged both on the display space and over the time, and the computation of their position (in space and in time) depends on a large number of factors, such as screen size or network performances. Most formatters have currently many limitations and can not take all these complex constraints into account.

To cope with such limitations (failure situations, lack of flexibility), we are working in two complementary directions:

- the definition of presentation operators that allow authors to better control formatters. These operators are the basis of the XEF language [2]. They allow authors to express priorities, more abstract properties and fall-back positions. As they are independent of any presentation language, these properties can be used with any language.

- the definition of a framework for building formatting systems by composing documents specified in different presentation languages and using software components (formatters) that are mutually dependent.

5. Software

5.1. Amaya

Participants: Paul Cheyrou-Lagrèze, Stéphane Gully, Vincent Quint, Irène Vatton.

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

Work on Amaya is a joint effort with W3C that started to showcase Web technologies in a fully-featured Web client. The main motivation for developing Amaya was to provide a framework that can integrate many W3C technologies during their development. It is used to demonstrate these technologies in action while taking advantage of their combination in a single, consistent environment.

Amaya started in 1996 as an HTML editor with support for the creation and debugging of CSS style sheets. Since that time it was extended to support XML and an increasing number of XML applications such as the XHTML family, MathML (mathematical expressions), and SVG (scalable vector graphics). It allows all those vocabularies to be edited simultaneously in compound documents. Amaya includes a collaborative annotation application based on the Resource Description Framework (RDF), XLink, and XPointer.

Now that a number of technologies are implemented in the editor, the developments focus on usability. The latest extensions are oriented towards robustness and ease of use. An important redesign phase has started at the end of 2003 for improving the user interface on several platforms: Linux, MS Windows, MacOS X.

Four public releases were made in 2003, on 3 February, 23 April, 8 July, and 13 November.

5.2. LimSee2

Participants: Vincent Kober, Nabil Layaïda, Daniel Weck.

LimSee2 is an open source authoring tool for multimedia documents using the SMIL 1.0 and 2.0 formats. It features a powerful graphical user interface designed to ease the manipulation of time-based scenarios in SMIL multimedia presentations. SMIL (Synchronized Multimedia Integration Language) is an XML language, so LimSee2 is an application that constantly deals with common issues of the XML world: parsing, validating, namespaces, doctype, DTD-driven editing, encoding, etc.

The aim of LimSee2 is to keep most of the XML aspects hidden from the user, so that there is no need to manipulate raw text data: everything can be done graphically. The main "visible" specificity of the SMIL language is that it clearly separates the two main areas of a multimedia presentation:

- Spatial layout specifies where and how multimedia objects should be displayed in a 2D coordinates space.
- Media synchronization specifies how multimedia objects should be synchronized over time.

The development of LimSee2 started in October 2002. The first public release was made on 13 June 2003.

6. New Results

6.1. Transformation Languages

The full design of Omega is a long-term task. In 2003 only a few aspects of the language have been studied, namely containment of XPath expressions and tree rewriting using rules.

6.1.1. Containment in XPath Expressions

Interesting results on containment in XPath expressions have been achieved in 2003. The containment relation between two XPath expressions p_1 and p_2 (denoted $p_1 \leq p_2$) holds true when, for any XML tree t and any context node x, the set of nodes selected by $p_1^x(t)$ is included in the set selected by $p_2^x(t)$. It is worth noting that the equivalence relation between two expressions (denoted $p_1 \equiv p_2$) can be expressed using two containment relations (i.e. $p_1 \leq p_2$ and $p_2 \leq p_1$), and given a suitable algorithm, equivalence is reducible to containment.

The containment and equivalence relations for XPath expressions are essential in Omega. In addition, several fundamental problems reduce directly to containment or equivalence, such as expression optimization and keys inference. Containment is also a key component for the static analysis of XSLT transformations. It can reveal two aspects of the expressions valuable for transformation designers and query programmers: consistency and performance. In practice, complex XPath expressions turn out to be difficult to interpret, therefore errors can be easily introduced. The consistency of an expression can be verified by checking if it is contained in the empty path.

The containment and equivalence of XPath expressions have been studied using an inference system combined with a rewriting system [7]. The inference system allows us to assert and prove properties on a class of expressions. In order to keep the proof system compact, a re-writing architecture has been proposed to transform remaining expressions in a disjunctive normal form compatible with this class. In contrast with model-based approaches, the inference and rewriting systems are applied to the XPath language directly. We believe this will help understanding the underlying issues of deciding containment on the language itself.

6.1.2. Tree Rewriting using Rules

While rewriting has been applied to string and tree transformation in compiler construction, and to more complex data structures, it has been poorly applied to XML transformation. This is partly because the richness of the XML tree model requires quite complex tree pattern matching specifications. In particular, conventional tree patterns require specifying not only the sub-structures of interest, but also the ones which have to be ignored during the matching.

In Omega, rewriting techniques are used to apply XPath-based rules such as $p \rightarrow p'$, where p and p' are XPath-like expressions. In order to build such a system, we have proposed two solutions [3]:

- an elegant and simple extension to XPath in order to embed the filtering variables required for memorizing and reusing sub-structures,
- a natural interpretation of the right hand side XPath expression (p') with respect to tree construction and modification.

6.2. Adaptation

6.2.1. Device Independent Format

A new approach was proposed for generating TV-like multimedia presentations that are adapted to the target user preferences and to devices with limited resources [9]. The main contributions are:

- The encoding of video presentations from a SMIL specification

- The adaptation of the video content based on the user preferences, and
- The delivery of adapted multimedia presentations.

The architecture includes a content server, an adaptation proxy and a set of small devices in the form of personal device assistants (PDA). These devices request the content through a wireless network. In order to show how the system behaves regarding the user preferences and capabilities, two negotiation dimensions are considered: the user language and the memory capability of the device. The first dimension is used to generate a content conforming with the user profile, e.g. a video with subtitles written in the preferred language. The second dimension is used to solve the problem of the system dead locks that usually happens when limited devices access rich multimedia presentations over the network.

The main result on the semantic approach to multimedia documents adaptation was achieved jointly with project Exmo and was published at IJCAI [6]. A precise definition was given of what is expected from the adaptation of these documents, and a comparison was made between these expectations and the results given by hand made transformation. We have proposed a model-based distinction between compliant documents, refining adaptation and transgressive adaptation. This framework has been applied to the temporal dimension of the documents providing measures for sharply discriminating the possible transgressive adaptations. We have also investigated the temporal dimension of multimedia documents specified qualitatively and proposed metrics for finding the "best" adaptations. The limitations of current multimedia specifications that prohibit a better adaptation have been considered.

6.2.2. Profile and Negotiation

The negotiation and adaptation core, called NAC, is an architecture developed in order to provide a solution for the delivery of multimedia content in heterogeneous environments. For content generation, NAC uses dynamic and static adaptation techniques. The adaptation is controlled using an adaptation and negotiation module (called ANM), and an optional module, called UCM (user context module), that runs on embedded devices. UCM allows the ANM knowledge about the client description to be enriched through the use of profiles. NAC default organization is proxy-based, but the proxy entity can be omitted by using ANM at the server side.

The proxy is a third entity that stands between servers and clients. It allows better handling of the heterogeneity of clients and the variety of content representations on servers. Proxies allow content adaptation without affecting the existing components in the system. For the content generation by adaptation, the proxy is the entity responsible for retrieving client requests and profiles and performing possible adaptation on the content received from the original server. The adapted content is then sent to the target client with respect to its characteristics. The proxy can transform existing multimedia content, meaning content does not need to be authored in multiple versions. In such situations and in order to avoid costly adaptation, a negotiation strategy is used to select the best variant to be delivered. All the proxy tasks are designed to behave transparently to clients and servers.

In NAC, the environment (device capabilities and preferences, network limitation, content characteristics, etc.) is described as a set of constraints that the content provider should satisfy. In our approach, the constraints resolution strategy is achieved by adding constraints progressively to the original content.

6.2.3. Transcoding

The main result on low-level transcoding was achieved in the field of resource-driven content adaptation [10]. We proposed and implemented an extensible framework for media resources manipulation in heterogeneous environments. Media manipulation is handled in terms of semantic relationships definitions through a set of parametric adaptation techniques. In order to demonstrate how media resources can be processed in practical cases we developed an application to adapt XHTML content to mobile phones. We have shown the influence of the media selection and transformation on the overall latency and delivery time. An evaluation formula has been proposed to estimate these two factors.

The evaluation has proven useful to point out when and where content adaptation cost affects the system efficiency in terms of CPU and memory consumption. The evaluation allows also to identify precisely the threshold where some adaptations become unnecessarily costly, for example resizing an image on the proxy. In summary, the proposed framework facilitates the achievement of the two main tasks of an adapting multimedia system: resource-driven content adaptation and negotiation. The drawback is that the definition of relationships between media resources requires sometimes the user to explicitly annotate the content.

A second transcoding scheme has been designed and evaluated: a SMIL 2.0 to SMIL Basic transcoder [14]. The transcoder achivates several types of adaptations. It evaluates switch statements on the proxy or server and outputs SMIL Basic conformant documents. The transcoder also filters user agent unsupported elements and adapts media elements to the terminal. The transcoder is capable of scaling down a SMIL documents structure to a single video stream when the presentation is not too complex.

Since UPS is a general framework for content adaptation and negotiation in heterogeneous environments, a useful transformation can be the one that generates UPS profiles from some other profiles already written for mobile devices. Once UPS profiles are generated, negotiation and adaptation mechanisms can be applied in the UPS framework and then an adapted content, that takes into account the client limitations and characteristics, can be transmitted to end users. About one hundred of real devices have been described in UPS, thus providing a uniform representation of a wide variety of terminals.

6.3. Multimedia Authoring and Formatting

We have specified a framework for coupling an existing formatting system such as SMIL or Madeus with the formatting control system XEF [2]. This framework performs the coupling process at two levels: 1) the language level, which is concerned with how to link the control features of XEF and the elements of an existing formatting system, and 2) the formatter level, which deals with the creation of a new formatter by composing several formatters. The overall objective is to provide more powerful and flexible formatting services to cover new needs required by adaptive or automatically generated presentations.

Our work on audiovisual data modeling for multimedia integration and synchronization [1] has been extented in two directions: first, we have used our model and tool for aligning movies with their scripts [8] in cooperation with R. Ronfard from the MOVI project; second, we have more deeply explored the MPEG-7 standard for multimedia authoring [5]. The approach chosen consists in using description tools from MPEG-7 Multimedia Description Schemes to describe audiovisual contents and in integrating these description models into our multimedia integration and synchronization model. The resulting model provides relevant specification tools for fine-grained integration of multimedia fragments into multimedia presentations. An authoring environment illustrates the authoring features that can be obtained thanks to these integrated models.

7. Contracts and Grants with Industry

7.1. Satin

Participants: Peter Hewat, Nabil Layaïda, Vincent Quint, Cécile Roisin.

Satin (Synchronized applications in interactive digital television) is a joint project with the HTTV company, funded by the French ministry of Industry through the RIAM network.

Most interactive television services broadcast today are completely independent from the audio and video streams, as if the two worlds of television and interactive applications were only sharing the same broadcast channel. The main goal of Satin is to introduce some synchronization between the two kinds of contents. To achieve this goal, a comprehensive environment is being implemented for creating, producing, broadcasting and presenting interactive television applications that are synchronized with audio-visual contents. This project is based on both the Web standards and formats created by W3C and the digital television standards from the MPEG group. It is a contribution to the next generation of the PrimeTV product from HTTV.

The main contributions of project WAM concern the formats (based on XML and SMIL), the editing environment with its simulation component, which allows the author to immediately get feedback on the user experience.

7.2. Microsoft Research Grant

Participants: Nabil Layaïda.

Projet WAM and project Sardes received jointly a Microsoft Research Innovation Excellence Awards for Embedded Systems to work on dynamic adaptation of embedded multimedia applications using Microsoft Windows CE .Net.

Multimedia applications are increasingly deployed and run on mobile terminals, characterized by limited capacities in terms of network bandwidth, CPU, memory or display size. Therefore, managing Quality of Service for multimedia applications executed on mobile terminals is a crucial issue. Recently we experimented with component-based dynamic adaptations on a proxy node for adapting multimedia streams according to the requirements of the terminals. A perspective to this work is to enable such adaptations on the terminal, which would allow new QoS management strategies. We propose to exploit Windows CE .Net and compact .Net technologies to implement dynamic adaptation of embedded multimedia applications.

8. Other Grants and Activities

8.1. Regional Initiatives

Project WAM is involved in ISDN (*Institut des Sciences du Document Numérique*) of the Rhône-Alpes region.

As part of the partnership between INRIA and the MSH (*Maisons des Sciences de l'Homme*) network, Vincent Kober is in charge of the MSH portal. He also plays a role of technology transfer concerning multimedia document engineering towards the research groups at the MSH-Alpes. In particular, he is working there on a multimedia project aiming at presenting a comprehensive study made by historians on paintings from the 15th century in alpine chapels. Cécile Roisin is a member of the MSH-INRIA steering committee.

8.2. National Initiatives

Ministry Grant: Satin is a RIAM project carried out with the HTTV company (§7.1)

INRIA Software Development Grant: project WAM received a grant (ODL) from INRIA for the development of LimSee2. (§5.2)

8.3. International Initiatives

The CEMT Project is a CNPq-INRIA project carried out jointly by WAM and Instituto de Informatica at UFRGS, Porto Alegre, Brazil to develop a cooperative environment for the edition of multimedia documents using workflow concepts to coordinate the editing and rendering tasks. The project was completed in September 2003.

The Amaya Web editor is developed jointly with the W3C Team. The software is distributed by W3C.

9. Dissemination

9.1. Leadership within Scientific Community

Tayeb Lemlouma is a member of the W3C Device Independence working group.

Vincent Quint is a member of the W3C Advisory Committee and of the steering committee of CNRS RTP 33: "Documents and content: creation, indexing and navigation"; Cécile Roisin is a member of the AS95: "Time in digital documents".

9.2. Conferences, Meetings and Tutorial Organization

Cécile Roisin and Ethan Munson were general co-chairs of the ACM Symposium on Document Engineering 2003, Grenoble, 20-22 November 2003.

Jean-Yves Vion-Dury is a co-organizer of the workshop on High performance XML processing, to be held in New York, NY, USA, in May 2004, in conjunction with WWW2004.

9.3. Teaching

Vincent Quint has given lectures on XML technologies at the CEA-INRIA-EDF computer science summer school, St Lambert, France, 16-27 June 2003, and at CINES, Montpellier, France, 19 September 2003.

Nabil Layaïda gives lectures on Structured Multimedia Documents in the joint Master of UJF and INPG, University of Grenoble. He also teaches on XML Technologies at the third year of ENSIMAG, Grenoble.

9.4. Conference and Workshop Committees, Invited Conferences

Nabil Layaïda was chair of the program committee of the Synchronized Multimedia Integration Language European Conference, SMIL Europe, Paris, France. 12-14 February 2003. He was a member of the editorial board of the special issue of *Document numérique* on digital documents in education.

Vincent Quint is a member of the steering committee of the conference on mobile multimedia, Mcube, Montbéliard, France, 30-31 March 2004. He was on the program committee of *Journées Francophones de la Toile*, JFT'2003, Tours, France, 30 June-2 July 2003, and on the program committee of conference H2PTM'03, Paris, 24-26 September 2003.

Vincent Quint presented an invited paper at H2PTM'03, Paris, 24-26 September 2003, and gave an invited lecture at the TEI Members Meeting, Nancy, France, 7-8 November 2003.

Cécile Roisn gave an invited talk at CID (*Centre d'Informatique Documentaire*) on fine-grained integration of media in multimedia presentations, Paris, 31 March 2003.

Cécile Roisin is a member of the editorial board of *Document numérique*. She was on the program committee of *6ème Colloque International sur le Document Électronique*, Cide.6, Caen, France, 24-26 november 2003.

Cécile Roisin and Ethan Munson are members of the steering committee of the ACM Symposium on Document Engineering. They were members of the program committee of the Second International Workshop on Web Document Analysis, WDA2003, Edinburgh, UK, 3 August, 2003.

Vincent Quint, Cécile Roisin, and Jean-Yves Vion-Dury were members of the program committee of the DocEng2003, ACM Symposium on Document Engineering, Grenoble, France, 20-22 November 2003. Jean-Yves Vion-Dury was nominated co-chair of the program committee of DocEng2004, ACM Symposium on Document Engineering, Millwaukee, WI, USA, November 2004

Nabil Layaïda and Vincent Quint are members of the program committee of the workshop on High performance XML processing, New York, NY, USA, May 2004.

Members of the WAM project have participated as reviewers in the evaluation process of several journals and calls for proposals including IEEE Multimedia Magazine, ACM Multimedia Systems Journal, *Document numérique*, *ACI Masse de Données*, RNTL, and ITIC.

10. Bibliography

Doctoral dissertations and "Habilitation" theses

[1] TIEN TRAN-THUONG. *Modélisation et traitement du contenu des médias pour l'édition et la présentation de documents multimédias.* INPG, Grenoble, February, 2003, http://wam.inrialpes.fr/publications/2003/theseTien/TheseV8PDF.pdf.

Publications in Conferences and Workshops

[2] FATEH BOULMAIZ, CÉCILE ROISIN, FRÉDÉRIC BÈS. *Improving Formatting Documents by Coupling Formatting Systems.* in « Proceedings of the 2003 ACM Symposium on Document Engineering, DocEng2003 », ACM Press, pages 92-94, 2003, http://wam.inrialpes.fr/publications/2003/doceng/doceng03.pdf.

[3] JEAN-YVES VION-DURY. *XPath on Left and Right Sides of Rules: Toward Compact XML Tree Rewriting through Node Patterns.* in « Proceedings of the 2003 ACM Symposium on Document Engineering, DocEng2003 », ACM Press, pages 19-25, 2003, http://wam.inrialpes.fr/publications/2003/JYVD-DocEng2003.pdf.

[4] VINCENT KOBER, DANIEL WECK, CÉCILE ROISIN. *Le langage SMIL au service des Sciences Humaines et Sociales.* in « JRES 2003 », pages 341-353, 2003, http://wam.inrialpes.fr/publications/2003/jres/jres2003-2.pdf.

[5] TIEN TRAN-THUONG, CÉCILE ROISIN. *Multimedia Modeling Using MPEG-7 for Authoring Multimedia Integration.* in « 5th ACM SIGMM International Workshop on Multimedia Information Retrieval (ACM MIR'03) », ACM, pages 171-178, 2003, http://wam.inrialpes.fr/publications/2003/ACMSigMM/Mir03_73_TranThuong.pdf.

[6] JÉRÔME EUZENAT, NABIL LAYAÏDA, VICTOR DIAS. *A semantic framework for multimedia document adaptation.* in « Proceedings of the 18th International Joint Conference on Artificial Intelligence IJCAI'2003 », Morgan Kauffman, San-Mateo (CA USA), pages 31-36, 2003, http://wam.inrialpes.fr/publications/2003/ijcai2003/ijcai2003.pdf.

[7] JEAN-YVES VION-DURY, NABIL LAYAÏDA. *Containment of XPath expressions : an Inference and Rewriting based approach.* in « Extreme Markup Languages, Montréal, Canada », IDEAlliance, http://www.idealliance.org/papers/extreme03/, August, 2003, http://wam.inrialpes.fr/publications/2003/xtrem2003/xtrem2003.pdf.

[8] REMI RONFARD, TIEN TRAN-THUONG. *A framework for aligning and indexing movies with their script.* in « Proceedings of IEEE International Conference on Multimedia & Expo (ICME) », IEEE Computer Society, 2003, http://wam.inrialpes.fr/publications/2003/icme2003/icmeTien.pdf.

[9] TAYEB LEMLOUMA, NABIL LAYAÏDA. *Encoding Multimedia Presentation for User Preferences and Limited Environments.* in « Proceedings of IEEE International Conference on Multimedia & Expo (ICME) », IEEE Computer Society, pages 165-168, 2003, http://wam.inrialpes.fr/publications/2003/icme2003/ICME2003.pdf.

[10] TAYEB LEMLOUMA, NABIL LAYAÏDA. *Media Resources Adaptation for Limited Devices.* in « Proceedings of the 7th ICCC/IFIP International Conference on Electronic Publishing, ELPUB 2003 », Universidade deo Minho, Portugal, pages 209-218, June, 2003, http://wam.inrialpes.fr/publications/2003/elpub2003/ELPUB2003.pdf.

[11] VINCENT QUINT. *Hypermédia et technologies Web.* in « Hypertextes hypermédias, actes de H2PTM'03 », Hermès Science, pages 381-390, September, 2003, http://wam.inrialpes.fr/publications/2003/H2PTM03/H2PTM03.html.

[12] TAYEB LEMLOUMA, NABIL LAYAÏDA. *Adapted Content Delivery for Different Contexts.* in « Proc. International Symposium on Applications and the Internet (SAINT 2003) », IEEE Computer Society, pages 190-197, January, 2003, http://wam.inrialpes.fr/publications/2003/saint2003_adaptedcontent/SAINT2003.pdf.

[13] CÉCILE ROISIN, VINCENT KOBER, VINCENT QUINT, PIERRE GENEVÈS, PATRICE NAVARRO. *Editing*

SMIL with Timelines. in « Proceedings of the Synchronised Multimedia Integration Language European Conference », January, 2003, http://wam.inrialpes.fr/publications/2003/smileurope_editingtimelines/EditingSMIL.html.

[14] TAYEB LEMLOUMA, NABIL LAYAÏDA. *SMIL Content Adaptation for Embedded Devices.* in « Proceedings of the Synchronised Multimedia Integration Language European Conference », January, 2003, http://wam.inrialpes.fr/publications/2003/smileurope_smilcontent/SMILEurope2003.pdf.

Miscellaneous

[15] R. HANRAHAN, R. MERRICK, C. WONG, M. WASMUND, R. LEWIS, T. LEMLOUMA. *Authoring Techniques for Device Independence.* World Wide Web Consortium, November, 2003, http://www.w3.org/TR/2003/WD-di-atdi-20031106/.