

XML-Based Languages for Multimodality in Mobile Environments

Danilo Avola

Istituto di Ricerche sulla Popolazione e le Politiche Sociali, Italy

Maria Chiara Caschera

Istituto di Ricerche sulla Popolazione e le Politiche Sociali, Italy

Fernando Ferri

Istituto di Ricerche sulla Popolazione e le Politiche Sociali, Italy

Patrizia Grifoni

Istituto di Ricerche sulla Popolazione e le Politiche Sociali, Italy

INTRODUCTION

The development of multimodal tools and mobile devices in particular is producing great interest, especially for accessing Web information, performing transactions, and use of services in general. This article considers the different markup languages proposed by the working groups of W3C (World Wide Web Consortium) to manage multimodal interaction and perspectives of multimodal applications and services.

The trend toward the convergence of various methodologies and technologies has developed new devices providing complex services, contributing to the sharing of experiences, and promoting the inclusion of people as community members (Paternò, 2004). This trend is based on the development of mobile devices and their usability, accessibility, portability, and versatility (Kvale, Warakagoda, & Knudsen, 2003).

The usefulness and usability of services, and the ability to access them and information, are the basic elements in the diffusion of Web systems and development of Web multimodal languages. The diffusion and implementation of multimodal services is supported by the activities of the World Wide Web Consortium, aimed at extending interaction modes for different devices and particularly devoted to solving various problems connected with: (1) multimodal Web interaction through the different devices, and (2) practice Web navigation from different devices.

Some W3C working groups focus their activities on issues such as independence from devices, multimodal Web access, and types of contents for multimodal messaging. These specifications allow rich multimodal contents to be transmitted, and are based on the power and extensibility of XML (eXtensible Markup Language) (Bray, Paoli, Sperberg-McQueen, Maler, & Yergeau, 2004).

XML is highly important in a mobile application environment, as many applications have to manage multimedia

contents and need dedicated tools for this. SMIL (Synchronized Multimedia Integration Language) (Solon, McKeivitt, & Curran, 2004) was proposed to achieve this goal.

In the early years W3C-MMI (W3C—MultiModal Interaction) focused on multimodal interaction modes such as speech and pen interaction, and providing users with W3C technologies.

W3C develops these technologies by orienting individual interaction modes in order to create mixed-namespace XML documents, such as SVG (scalable vector graphics) (Chatty, Lemort, Sire, & Vinot, 2005) and XHTML (eXtensible HyperText Markup Language) (Musciano & Kennedy, 2003) for visual interaction, and VoiceXML (Voice Extensible Markup Language) (Lucas, 2000) for voice interaction.

However, many other XML-derived languages have helped in the development of mobile services.

The next target is the consideration of the mobile network as an extension of the global Internet network. This article explains the importance of XML and its dialects in a mobile application environment to enable their use by the “various applications/services” (today available on the Web). In fact, different dialects may be needed for different mobile devices depending on their characteristics (such as memory, CPU speed, integrated software engine, etc.). For example, two SVG profiles are defined for cellular phones and PDAs (personal digital assistants): SVG Tiny (SVGT) is suitable for the next generation of cellular phones especially, while SVG Basic (SVGB) is aimed at high-tech devices such as PDAs or smart phones (Andersson et al., 2003).

The pervasive use of mobile devices will be the target for the near future (Branco, 2001), given the trend towards considering the mobile network as an extension of the Internet global network. This scenario promotes the development of new dialects for multimodal interaction through mobile devices. The dialects developed for speech, sketch, and visual

interaction are discussed next. An area for future development might focus on interaction through gestures.

XML (eXtensible Markup Language) is a simple, flexible, and powerful markup language, based on text format that allows the development of a potentially unlimited number of innovative multimodal services and applications. It was derived from the more complex, complete SGML (Standard Generalized Markup Language, ISO 8879) (Chamberlin & Goldfarb, 1987), designed for more general purposes. However, XML language is easier to manage, and is genuinely Web oriented and mobile oriented. In other words, XML language is an optimal subset of SGML, constructed in consideration of the possible Web services and applications.

XML can be used to develop several languages taking the specific working context into account. It also plays an important role in the exchange of a wide variety of data, making them available and accessible by Web using computers and mobile devices.

BACKGROUND

The World Wide Web is undergoing continuous development. This has enabled a great expansion in a wide variety of applications and services, covering every human activity. In addition, advanced technology mobile systems are becoming ever more complete, complex devices, which can offer a broad range of Web applications and services originally conceived for personal computers. These two factors explain the need to introduce various multimodal systems (services and applications oriented) to interact with the Web using mobile devices.

In this context, the use of XML-based technology to create powerful, multi-purpose system interfaces is a winning choice. Meta-language XML allows ad hoc language solutions to be developed according to the specific argument.

A multitude of XML “dialects” for multimodal solutions have already been developed, while any others are under current or future development. An exhaustive point of view on XML-based languages is thus not a simple matter. This section provides a panorama of the XML-based languages, considering the different interaction modes, multimodality, and multimodality features.

SMIL (Solon et al., 2004) is a basic, developing technology that allows several multimodal environments to be implemented. It is not an “out and out” multimodal language. As in many multimodal environments, it is necessary to interact with several kinds of multimedia content; SMIL works in the background with many different multimodal applications and services.

SMIL enables interactive audiovisual presentations to be easily produced. It is typically used to choreograph complex multimedia presentations, where audio and video streaming, images, text, graphics, and other media types

are combined in real time. In other words, SMIL makes it possible to manage the temporal and spatial constraints of multimedia presentations. The current SMIL conception offers modules for animation, content control, layout, linking, media objects, meta information, timing and synchronization, and transition effects. This modular approach allows reuse of SMIL syntax and semantics in other XML-based languages, especially those used for timing and synchronization. These fundamental features play a leading role in multimodal user interaction. SMIL is an easy-to-learn, HTML-like language and the World Wide Web Consortium (W3C) recommendation to achieve synchronized multimedia. In this context, it simplifies the creation of time-based multimodal interfaces with a high portability factor. SMIL is also exploited to aid the construction of powerful mobile-oriented multimodal applications and services.

A classic example of SMIL use is to enable authors to specify and control the precise time a sentence is spoken and make it coincide with the display of a given image. This simple “technical pattern” is at the base of several multimodal general systems.

To consider the different languages and problems connected with multimodality, we must take into account the different interaction approaches (visual, voice, etc.). For a visual approach, we must consider SVG. The SVG language (Chatty et al., 2005) is another important basic technology that works in background mode to resolve different types of problems in the multimodal interfaces. It describes two-dimensional graphics and graphical applications in XML, providing facilities for document structuring, shape definition, painting, clipping and masking, compositing, text manipulation, styling, linking, scripting, animation, interactivity, integration of multimedia content, alpha masks, filter effects, and template objects. It supports object zooming, interaction and manipulation, and scene annotation, among others. SVG element features can be static or dynamic, and each complex element can be used in interactive mode. A strong point of this language is the ability to interact with script languages (such as JavaScript, ECMAScript, JScript, etc.), which provide complete access to all elements, attributes, and properties necessary to develop powerful, complete 2D graphical representations. However, this latter point is not a binding factor. In fact, for example, an animation can be defined and triggered either declaratively (i.e., by embedding SVG animation elements in SVG content) or via scripting.

Because of SVG’s ability to produce high-quality rich graphical displays, enable the development of highly interactive user interfaces, and manipulate the contents and structure of an SVG document, it is very well suited for the development of interactive multimodal applications and services.

In fact, it is Web accessible by both personal computer and innovative mobile-oriented systems, and needs sophisticated multimodal environments to satisfy the user’s

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/xml-based-languages-multimodality-mobile/17218

Related Content

Visualization-Driven Approach to Fraud Detection in the Mobile Money Transfer Services

Evgenia Novikova and Igor Kotenko (2019). *Algorithms, Methods, and Applications in Mobile Computing and Communications* (pp. 205-236).

www.irma-international.org/chapter/visualization-driven-approach-to-fraud-detection-in-the-mobile-money-transfer-services/208462

An Electronic Auction Service Framework Based on Mobile Software Agents

Sheng-Uei Guan (2009). *Mobile Computing: Concepts, Methodologies, Tools, and Applications* (pp. 1640-1652).

www.irma-international.org/chapter/electronic-auction-service-framework-based/26613

Mobile Data Offloading Using Opportunistic Communication and AP Deployment: A Case Study

Sanjit Kumar Dash, Sasmita Mishra and Jibitesh Mishra (2017). *International Journal of Mobile Computing and Multimedia Communications* (pp. 66-84).

www.irma-international.org/article/mobile-data-offloading-using-opportunistic-communication-and-ap-deployment/193260

Business and Technology Issues in Wireless Networking

D. Wright (2007). *Encyclopedia of Mobile Computing and Commerce* (pp. 90-95).

www.irma-international.org/chapter/business-technology-issues-wireless-networking/17058

FlexRFID Middleware in the Supply Chain: Strategic Values and Challenges

M. E. Ajana, H. Harroud, M. Boulmalf and M. Elkoutbi (2011). *International Journal of Mobile Computing and Multimedia Communications* (pp. 19-32).

www.irma-international.org/article/flexrfid-middleware-supply-chain/55082