Building the Tower of Babel:
Converting XML Documents to VoiceXML for Accessibility[1]

(Extended Abstract)

Gopal Gupta

Laboratory for Logic, Databases, and Advanced Programming
Department of Computer Science
New Mexico State University
Las Cruces, NM 88003
http://www.cs.nmsu.edu/lldap

1 XML and Marking up Voice Data

Hyper Text Markup Language (HTML) is the language popularly used for marking up information on the World Wide Web so that it can be displayed using commercially available browsers (e.g., Netscape and Internet Explorer). People involved in the development of HTML (the W3 Consortium) have realized that HTML is a weak mark-up language if a user wants to process WEB-pages further. The problem with HTML is that to infer any kind of semantic information *automatically* requires that one examine the content of the web-page and analyze it. The HTML tags do not provide any help, beyond how the document is structured. Thus, writing programs that will further process information contained in the WEB page at the user sites is extremely difficult. This limits the use of information that is available on the WEB. To alleviate this problem, the W3C has come up with an eXtensible Markup Language (or XML) for marking-up information in the WEB pages. XML is indeed flexible in that it allows developers/writers of WEB pages to define their own tags. These tags can have some well-defined meanings that others can understand. W3C expects that groups of users will evolve their own XML for their particular needs. For example, toy store owners and manufacturers will evolve their own XML that will facilitate the selling, distribution and description of toys. This has already started to happen and a plethora of XML based mark-up languages have appeared. These include: BioML for describing biopolymer sequence information, MathML for marking up math for display on the web, SMIL: an XML for describing multimedia presentations for the Web, CML: an XML for describing chemical information. A partial list of the numerous XMLs being proposed by various industry groups, research organizations, and the Government, can be seen at http://www.xml.org/xmlorg_catalog.htm.

2 Universally Accessible WEB: VoiceXML and VoxML

The most significant development, as far as the visually impaired community is concerned is the development of VoxML and VoiceXML, two XML based languages for marking up voice data. The goal behind VoxML and VoiceXML is to make the internet accessible through the phone. VoxML is being developed by Motorola (http://voxml.mot.com), while VoiceXML is being developed by a consortium of IBM, Lucent, Motorola and AT&T. VoiceXML and VoxXML require a voice browser (already developed by designers of VoiceXML and VoxML) that understands VoiceXML or VoxXML and that reads out the information contained in the VoiceXML or VoxML marked-up web page. Both VoiceXML and VoxXML also allow the user to interact with the web-page. Of course, these companies are developing VoiceXML and VoxML with the intent of making the web-pages and the Internet accessible via the phone (or access the WEB in a car while driving), and in the process doing a great service to the visually-impaired community. Thus, if all documents in the WEB were written in VoiceXML or VoXML, they will be accessible to the visually impaired community as well. However, this is not so, since given the tens (or hundreds) of XML-based languages that are being designed, only some documents will be coded in VoiceXML. Most will be coded using the XML the document most closely relates to (for example, all legal documents will be perhaps marked up with Legal XML, rather than with VoiceXML). Thus, while XML is going to allow web-page writers to pack more semantic information about the WEB pages, it is also going to lead to a plethora of XMLs each of which is different from the other, leading to the following problem: WEB-pages prepared by one community (that uses its own XML to prepare it) cannot be understood by another community,

[1]This work has been supported by NSF HRD 9800209 and INT 9904063.

because this other community has its own language. This leads to the compromise of the dream of building a seamless Internet, just as in the biblical story of the tower of Babel where the dream of the people to build a tower that reaches the heaven was compromised when God created a plethora of spoken languages and people could no longer communicate with each other.

3 Making XML documents Universally Accessible

In this paper we propose an approach to make XMLs of all types universally accessible. Our idea is to develop a methodology that will allow us to *rapidly* build translation filter software to automatically translate one type of XML into another. Our main interest is in using this framework to translate a document written in a particular type of XML automatically to VoiceXML or VoxML, thus making that document, that was otherwise inaccessible, accessible to visually impaired individuals. It is important that the methodology that we devise allows for rapid development of the translator software for automatically translating XML to VoiceXML or VoxML. This is because the number of XMLs is probably going to be in the 100s, and if it takes a long time to build these automatic translators, they will never be built.

To build a translator to translate one formal notation to another, traditional compiler technology [4] is normally used. However, a solution based on traditional compiler technology will take a long time, as building translator in such a way will be quite cumbersome. Instead we propose a novel approach that is based on denotational semantics [2] and logic programming [1].

Denotational Semantics [2] is a well-established methodology for design and analysis of computer languages. In the denotational approach, the semantics or meaning of a computer language/notation is specified in terms of mathematical objects (such as sets and functions). A denotational specification has three components: *syntax specification* which maps sentences of the language to parse trees; *semantic algebras* that specify these mathematical objects, and *valuation functions* that map parse trees to semantic algebras. It turns out [3] that the syntax specification, the semantic algebra, and the valuation functions can be specified using logic programming [1]. This logic programming specification is also executable [3]. To obtain a translator from a language $\mathcal{L}_1$ to $\mathcal{L}_2$ we choose the semantic algebra consisting of parse trees of $\mathcal{L}_2$. The logic programming specification then automatically turns into a translator which can be executed to translate sentences of language $\mathcal{L}_2$ to sentences of language $\mathcal{L}_2$.

In our case, the language $\mathcal{L}_1$ will be the XML we want to translate for universal accessibility, while the language $\mathcal{L}_2$ will be VoiceXML or VoxML. Using the technology of denotational semantics and logic programming, the translation filters can be developed very rapidly, e.g., within one or two weeks. In the full paper, we will show how MathML can be automatically translated to VoiceXML using our framework so that MathML documents become accessible to visually impaired people. Essentially, using the denotational approach, the syntax of MathML will be expressed using the Definite Clause Facility of logic programming. In fact, the DCG can be automatically obtained from the Document Type Definition (DTD) of an XML document. Next, the semantics of MathML will be given in terms of VoiceXML or VoxML constructs. Both the syntax specification and the semantics specification will be expressed using logic programming. The logic programming specification of syntax and semantics is executable, yielding a parser automatically. In the full paper, we will also discuss how existing HTML coded documents can be automatically converted into VoiceXML or VoxML document using our methodology. Such a converter for a subset of HTML is already operational.

Our methodology, based on logic programming and denotational semantics, will permit individuals to rapidly construct a filter between any type of XML to VoxML or VoiceXML, making XML documents universally accessible. Our methodology can also be used for rapid construction of a translator between one type of XML to another, for example, to translate BioML documents to documents in CML (Chemical Markup Language).

References

[1] L. Sterling & S. Shapiro. The Art of Prolog. MIT Press, '94.

[2] D. Schmidt. *Denotational Semantics: a Methodology for Language Development.* W.C. Brown Publishers, 1986.

[3] G. Gupta. Horn Logic Denotations and Their Applications. In *The Logic Programming Paradigm: A 25 year perspective.* Springer Verlag. pp. 127-160. May 1999.

[4] A. Aho, J.D. Ullman, R. Sethi. Compilers: Principles, Techniques, and Tools. Addison Wesley. 1986.