Extensible Transcoder for HTML to VoiceXML Conversion

Narayan Annamalai, B. Prabhakaran, and Gopal Gupta

Anytime anywhere Internet access has become the goal for current technology vendors. Sudden increase in the number of mobile users has necessitated the need for `Internet access through mobile phones'. The existing web infrastructure was designed for traditional desktop browsers and not for hand-held devices. The data in the web is stored in HTML (Hyper Text Markup Language) format which cannot be delivered to mobile devices. The only acceptable way to present data to devices like cellular phones, is audio. Certain voice browsers are capable enough to process VoiceXML content and produce the output in the form of audio. Thus it is imperative that a transcoder which converts HTML data to VoiceXML data is developed.

The transcoder is realized in two phases: One is the parsing phase where the input HTML file is converted to HTML node tree (DOM) and other is the translation phase where the node tree is traversed in a LR recursive manner and each node mapped to its equivalent VoiceXML node. The mapping is not one to one always since both HTML and VoiceXML are Context Sensitive languages. The Transcoder is extensible in the sense it provides an easy user interface to add translation logic for new HTML tags. The transcoder also provides means for the user, to modify the translation logic while dealing with certain HTML tags.

Input HTML File

```
<html>
-----
<form next="method">
  <input type="radio" name="sex" value="male" />
  Male <br>
  <input type="radio" name="sex" value="female" />
  Female <br>
</form>
------
</html>
```

Output VoiceXML File

```
<?xml version="1.0"?>
<vxml version="2.0">
----
<field name="sex">
 <prompt> Select an Entree, what sex <enumerate/>
</prompt>
 <option dtmf="1" value="male"> Male </option>
<option dtmf="2" value="female"> Female </option>
</field>
-----
</vxml>
```

Figure: A Sample Conversion

The Transcoder finds its application in numerous roles viz. as a stand alone application in the client side (converting incoming HTML files to VoiceXML and providing the output as Voice. But for this the client computer needs a Voice browser), as a part of proxy server and as a part of the web server itself. The Transcoder plays a major role in achieving the vision of Voice Web, i.e. converting the text rich web pages to voice without having to re-author them. The extension to this project is the "Dynamic VoiceXML Generation" which imparts navigable features in the form of Voice anchors to the VoiceXML page.

Reference: N. Annamalai, "An Extensible Transcoder for HTML to VoiceXML Conversion", Masters Thesis, UTD.