11.5 Voice Interfaces for Web

Dr. K. Samudravijaya, TIFR Mumbai

Voice Interfaces for Web

Samudravijaya K
chief@tifr.res.in
 Tata Institute of Fundamental Research

Speech enabled Web Browsing

Visually oriented medium → multi-modal medium
- Browsers specifically designed for people with [visual] disabilities.
- Screen-readers allow visually impaired people to navigate with standard applications, used in conjunction with mainstream browsers.
- Voice browsers allow voice-driven navigation, some with both voice-in and voice-out.

Alternative Web Browsing
W3C Speech Interface Framework
VoiceXML
CSS: Audio Features in Style Sheets

Speech enabled Web Browsing

Visually oriented medium → multi-modal medium
- Browsers specifically designed for people with [visual] disabilities.
- Screen-readers allow visually impaired people to navigate with standard applications, used in conjunction with mainstream browsers.
- Voice browsers allow voice-driven navigation, some with both voice-in and voice-out.
- Lists of above:
  http://www.w3.org/WAI/References/Browsing
Speech Interface Framework

- **The User**
  - Speech Recognition
  - Speech Synthesis
  - SRGS/SI
  - SSML
- **Dialogues**
- **VoiceXML**
- **The Network**
  - XSL/HTTP/DOM...
  - CCXML

Voice XML

- Convergence of telecommunications and the Web
- Language for writing Web pages with speech interface
- Enables spoken dialog

Voice XML documents describe:
- spoken prompts (synthetic speech)
- output of audio files and streams
- recognition of spoken words and phrases
- recognition of touch tone (DTMF) key presses
- recording of spoken input
- control of dialog flow
- telephony control (call transfer and hangup)

An example ...

- Offers a menu of three choices: sports, weather or news.

```
<xml version="1.0">
<xml version="2.0">
<xml>
  <prompt>
    Say one of: <enumerates/>
  </prompt>
  <choice next="http://www.sports.example/start.xml"> Sports
  </choice>
  <choice next="http://www.weather.example/intro.xml"> Weather
  </choice>
</xml>
```

An example of a VoiceXML application

```
<choice next="http://www.news.example/news.xml"> News
  </choice>
<choice>
  <input>Please say one of <enumerates/>.</input>
</xml>
```

A typical dialog

- **Computer**: Say one of: Sports, Weather, News.
- **Human**: Astrology
- **Computer**: I did not understand what you said.
  (a platform-specific default message.)
- **Computer**: Say one of: Sports, Weather, News.
- **Human**: Sports
- **Computer**: (proceeds to http://www.sports.example/start.xml)
Dialog requirements for the voice markup language

- must have
- should have
- nice to have
- future revision

Aural style sheets

- Controlling Volume
- Basic Speech
- Talking headers

Aural style sheets

- Controlling Volume
- Basic Speech
- Talking headers