
VoiceXML VUI를 위한 Dialog 설계에 관한 연구

장민석 · 예상후

국립군산대학교 컴퓨터정보과학과

A Study on Design of Dialog for VoiceXML VUI

Min-seok Jang · Sang-hoo Yea

Kunsan National University, Dept. of Computer Information Science

E-mail : msjang@cs.kunsan.ac.kr, badpotat@cs.kunsan.ac.kr

요 약

현재 정보통신분야의 기업에서 많은 연구가 이루어지고 있는 VoiceXML은 웹 환경에서 훌륭한 VUI(Voice User Interface)를 제공할 수 있다. 하지만 효과적으로 사용하기 위해서는 다양한 Dialog를 연구/개발 하는 상황에 있다. 이는 2002년에 열린 "VoiceXML Conference & Expo"에서도 주요한 관심사항으로 부각되었다. 본 논문은 이러한 점을 인지하여, VoiceXML이 효과적으로 VUI로 사용되어질 수 있도록 하기 위해 VoiceXML Dialog를 설계하고 이를 실험함으로써 VoiceXML이 실제적으로 VUI로 사용될 수 있도록 하고자 한다.

ABSTRACT

Nowadays the corporations related to Information & Communication field are researching more and more on VoiceXML development. VoiceXML can provide users with more efficient interface, VUI(VoiceXML User Interface) in web environment than the existing one. But more research and development for designing the Dialog have to be done for VUI to be used in efficient way. That was a main topic in "2002 VoiceXML Conference & Expo". According to the importance this paper presents VoiceXML Dialog designed for the purpose of its efficient use and the experimental result.

키워드

VoiceXML, VUI(Voice User Interface), XML, Dialog

1. 서 론

음성인식/ 합성 관련연구의 결과로 이를 이용한 각종 음성인식 프로그램들과 가전제품 등에 응용되어 상품화 되어있으며, 일부 ARS시스템에서 음성인식을 이용한 시스템을 적용하고 있다. 또한 컴퓨팅에도 접목되어 다양한 음성인식 프로그램이 개발되어지고 있으나 아직 웹 환경에 본격적으로 활용하지 않고 있으며, 현재까지 개발되어진 음성 브라우저의 경우 사용자가 받아들이기에 어색한 작동을 보였다. 이는 웹페이지 개발자가 음성 인터페이스를 고려하지 않은 웹페이지를 단순히 음성 브라우저가 텍스트 정보를 음성으로 출력하고 또한 HTML내의 <A> 태그를 읽어 페이지 전환을 하기 때문이다. 이는 표준화된 인터페이스 방법이 없는 이유이기도 하다[1].

VoiceXML은 XML의 서브셋으로 AT&T, Lucent Technologies, Motorola, IBM 등으로 구성된 VoiceXML 포럼에서 제안하고 이를 W3C에서 공인한 대화형 마크업 언어의 표준이다. 이는 음성인식/합성 등과 같은 음성 처리 기술들을 이용하여 웹에서의 HTML을 통한 정보의 접근방법을 유무선 전화기를 이용하여 웹을 접근토록 하는 토대를 마련한다. 하지만 현재까지 VoiceXML관련 연구들은 유무선 전화기 등을 통해 이용하는 것에만 집중되어졌다. 이러한 VoiceXML을 이용하여 표준화된 VUI를 제공할 수 있다. 이를 이용하여 본 연구는 Web 환경에서 사용자에게 GUI(Graphic User Interface)환경과 더불어 VUI(Voice User Interface)를 제공하며, 보다 풍부한 인터페이스를 마련하고자 한다. 이를 위해 웹페이지 개발자들에게 음성 인터페이스를 구축할 수 있는 방법을 제시하

며, 또한 사용자에게 보다 효율적인 인터페이스를 제공할 수 있는 VoiceXML Dialog를 제시하고자 한다.

II. VoiceXML

Voice XML은 AT&T, IBM, 루슨트 테크놀로지, 모토로라 등 정보통신 분야의 4개 거대 기업체에 의해 설립된 Voice XML 포럼에서 제안한 마크업 언어로서, 음성 애플리케이션 개발을 위해 고안된 XML문서 형식의 일종이다. Voice XML포럼은 1999년 8월 Voice XML 0.9버전을 발표 후 작년 3월 이를 크게 보완한 버전 1.0을 정식 제안했다. 세계 인터넷 환경을 주도하고 있는 W3C 컨소시엄에서는 Voice XML 포럼의 제안을 받아들여 작년 5월 22일 Voice XML을 웹의 대화형 마크업 언어 표준으로 공인했다. 기존에도 WTE(The Web Telephony Engine), WAP(Wireless Application Protocol) 등이 Voice XML 이전에 개발되어져 있었고 사용되어졌으나 널리 이용되지는 못하고 있다[2].

VoiceXML은 음성 응용프로그램으로 인터넷 정보를 전달하고 음성 응용 프로그램 제작자를 하위레벨 프로그래밍이나 자원관리로부터 벗어나도록 도와준다. 이는 개발자가 음성 시나리오에 집중함으로써 개발시 다음과 같은 이점을 얻을 수 있다.

- VoiceXML은 서비스 로직에서 사용자 상호작용 코드를 분리할 수 있다.
- 음성 서비스와 DB 서비스를 통합하여 클라이언트-서버환경으로 제공할 수 있다.
- VoiceXML 웹서버는 여러 가지 서비스 로직을 수행하여 VoiceXML문서를 동적으로 제작할 수 있다.
- VoiceXML은 기존 웹기반의 네트워킹을 이용함으로써 쉽게 네트워킹 기반의 어플리케이션을 제작할 수 있다 [3][4].

III. VoiceXML Dialog 설계

본 장에서는 VoiceXML이 웹사이트에서 VUI로 사용될 수 있도록 현재의 환경을 분석하고 이를 토대로 VoiceXML Dialog를 설계한다.

현재의 웹은 GUI를 기반으로 구성되어 있다. 이러한 환경을 기계적으로 음성으로 전환하고자 할 때 대부분 어색하게 구성되진다. 이는 초기 인터페이스 구성시 음성정보를 고려한 설계를 하지 않음에 기인하고, 또한 마땅한 도구 또한 마련되어 있지 않기 때문이

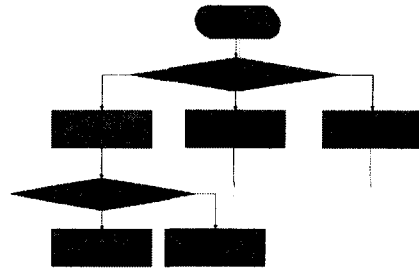
다. 이러한 문제점을 본 논문의 VoiceXML VUI를 통해 해결할 수 있도록 하기 위해 현재 바로 사용할 수 있는 XML Island기술을 이용하였다. 이는 기존의 웹 브라우저에서는 표시되지 않지만 VoiceXML 웹 브라우저에서는 사용자에게 VUI를 제공할 수 있는 토대를 마련할 수 있다. 이렇게 VoiceXML을 이용하여 VUI를 제공하기 위해 다음과 같이 현재의 웹사이트를 구분하고 특징을 알아보았다.

첫째, 일반적 메뉴방식의 웹사이트

둘째, 웹 게시판

셋째, 검색엔진

위의 세 가지가 혼합되어 대부분의 웹사이트가 구성되어 있다. 이들 각각의 구성방식 중 메뉴방식의 웹사이트를 아래의 [그림2-1]과 같은 플로차트로 나타내었다.



[그림 2-1] 메뉴방식 플로차트

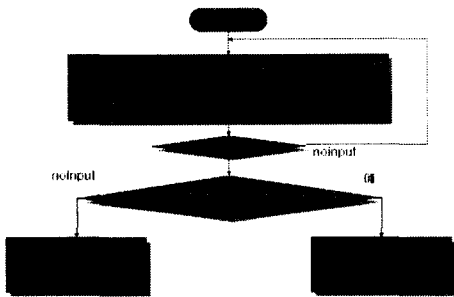
```
<XML ID="vxml">
<vxml version = "1.0">
<form>
  <field name="selection">
    <prompt>메뉴 1, 2, 3중 택1</prompt>
    <filled>
      <if cond = "selection == '1'">
        <goto nextitem = "1" />
      <elseif cond = "selection == '2'" />
        <goto nextitem = "2" />
      .....
    </if>
  </filled>
</field>
  <field name = "1">
    <prompt>메뉴1로 이동합니다.</prompt>
    <link next = "1.html"/>
  </field>
  .....
```

[그림 2-2] 메뉴방식의 VoiceXML Dialog

[그림 2-1]과 같은 메뉴방식의 웹사이트는 좌측 또는 상단 등 특정부분에 웹사이트의 카테고리를 구분하

여 사용자가 웹을 서핑 할 수 있도록 꾸며져 있다. 이와 같은 구성은 메뉴를 선택하는 즉시 페이지의 전환이 이루어지는 특징이 있다. VoiceXML Dialog는 [그림 2-2]와 같이 나타낼 수 있다.

웹 게시판의 경우는 일반적 메뉴방식과는 달리 단순히 페이지가 고정되어 있는 것이 아니라 스크립트 언어를 통해서 동적으로 생성되므로 VUI 구성시 메뉴의 중복을 피하기 위한 방법이 필요하며, 단순 열거식이 아닌 메뉴구성이기 때문에 1:1의 전환으로는 부족하다. 그러므로 좀더 기존 ARS(Automatic Response Service)시스템처럼 몇 단계를 거치거나 단축메뉴 기능 등 사용자가 원하는 작업을 수행할 수 있도록 하여야 한다. 본 논문은 이러한 VUI를 구성할 때 HTML 페이지의 즉각적인 전환 없이 하나의 VoiceXML 문서 내에서 적당한 길이의 Dialog를 작성하여 사용자에게 VUI를 제공하도록 설계하였다. 이를 [그림 2-3]과 같이 플로차트로 나타내었다.



[그림 2-3] 게시판 플로차트

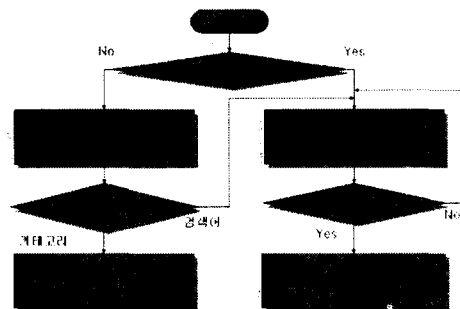
[그림 2-3]에서 게시판은 읽기, 새글쓰기, 페이지 이동, 작성자명으로 찾기, 제목으로 찾기, 내용으로 찾기 등의 기능을 갖고있으며, 이중 게시판 읽기 기능과 연관된 메일보내기 기능을 VoiceXML Dialog로 구성하였다. [그림 2-4]은 게시물을 보여주는 페이지에 추가되는 VoiceXML 문서이다. 이중 게시물 찾기 기능은 검색엔진에서도 쓰일 수 있는 방법으로 검색엔진에서는 좀더 다양한 검색방법을 추가해야 한다.

[그림 2-4]에서 각각의 게시물을 읽기 위해서 게시물 번호를 사용자가 입력하고 브라우저가 메일 보내기 여부를 물을 때 “예”를 입력하면 해당 게시물을 작성한 작성자에게 메일을 보내게 되고, 아무 입력이 없게 되면 게시물의 내용을 읽어주기 위해 게시물을 읽는 페이지로 이동하게 된다.

```
<vxml version = "1.0">
<form>
  <field name="ReadNum">
    <prompt>예상후님이 작성하신 VoiceXML요약이 1번
홍길동님이 작성하신 VoiceXML과 XML이 2번
김홍부님이 작성하신 VUI는 3번 ...
중 읽거나 메일을 보낼 게시물 번호를 말씀해주세요
  </prompt>
  <filled>
    <if cond = "ReadNum == '1'">
      <goto nextitem = "1"/>
    <elseif cond = "ReadNum == '2'">
      <goto nextitem = "2"/>
    <elseif cond = "ReadNum == '3'">
      <goto nextitem = "3"/>
    ...
  </filled>
</form>
</vxml>
```

```
</if>
</filled>
</field>
<field name = "1">
  <menu>
    <prompt>예상후님에게 메일을 보내시겠습니까?
  </prompt>
  <filled>
    <if cond = "readOk == '아니오'">
      <goto nextitem = "readOk1"/>
    </if>
  </filled>
</field>
.....
<field name = "readOk1">
  <prompt>1번게시물을 읽겠습니다</prompt>
  <link next="read.asp?idx=1"/>
</field>
<field name = "mailOk1">
  <prompt> 메일작성폼으로 이동합니다.</prompt>
  <link next =
"mailto.asp?to=badpotat@cs.kunsan.ac.kr"/>
</field>
.....
</form>
</vxml>
```

[그림 2-4] 게시판 VoiceXML



[그림 2-5] 검색엔진에서의 플로차트

[그림 2-5]는 현재 대부분의 포털사이트가 포함하고 있는 검색엔진을 VoiceXML VUI로 표현하기 위해 플로차트로 나타낸 그림이다. 검색엔진은 검색의 범위와 검색을 위한 연산자들을 포함하고 있다. 이러한 부분을 전부 음성으로 처리하기 위해서는 각각의 사이트마다 검색연산자의 음성표현을 정의해야 한다. 가령 AND, OR, NOT 등이 그것이다. 본 논문에서는 이들 각각을 “그리고”, “또는”, “빼고” 등으로 표현하였다.

```

<vxml version = "1.0">
  <form>
    <field name = "init" >
      <prompt>키워드 검색을 하시겠습니까?</prompt>
      <filled>
        <if cond = "init == '예'" >
          <goto nextitem = "Keyword" />
        </if>
      </filled>
    </noinput>
    <goto nextitem = "Category" />
  </noinput>
  </field>
  <field name = "Keyword" >
    <prompt>그리고, 또는, 빼고 등의 연결을 통해
    검색이 가능합니다.</prompt>
    <filled>
      <goto next = "keyword.asp?key=INPUT()" />
    </filled>
  </field>
  <field name = "Category" >
    <prompt>검색하고자 하는 카테고리별
    말씀해주세요.</prompt>
    <filled>
      <if cond = "Category == '컴퓨터관련'" >
        <goto next = "default.asp?category=1" />
      <elseif cond = "Category == '건강'" />
        <goto next = "default.asp?category=2" />
        .....
      </if>
      <goto next = "Category" />
    </filled>
  </noinput>
  <goto next = "Keyword" />
</noinput>
</field>
</form>
</vxml>

```

[그림 2-5] 검색엔진 VoiceXML

[그림 2-5]에서 검색엔진은 카테고리별 검색과 불린 키워드 검색이 가능한 일반적인 검색엔진을 모델로 했다. VoiceXML VUI 브라우저는 사용자가 입력한 키워드를 음성엔진에서 검색단어와 연산자를 분리해서 검색을 실시하게 되며 ECMAScript 등 추가 기능을 수행할 수 있는 스크립트를 지원해야 한다. 본 논문에서는 VoiceXML VUI 브라우저에서 지원하는 INPUT() 함수를 이용했다. 그리고 카테고리 이동 명령과 기타 메뉴의 이동은 웹사이트의 특성에 맞게 주요링크만 음성출력을 하는 것이 사용자의 혼란을 막을 수 있다.

IV. 결 론

현재 외국의 VoiceXML 관련 연구들은 이제 실용화 단계에 접어들어 이미 IBM 등과 같은 기업에서는 VoiceXML 솔루션들이 속속들이 발표되고 있는 상황이다. 또한 본 논문에서 제시하는 VoiceXML VUI와 같은 맥락에서 2001년 12월부터 W3C에서 XHTML+VOICE 라는 명칭으로 이에 관해 논의중이며 일부 스키마와, DTD를 개발하고 있다[5]. 또한 2002년 3월에 열린 VoiceXML EXPO & Conference에서 주요 주제로 논의된 VoiceXML Dialog 또한 활발한 연구가 진행되고 있는 상황이다[6]. 하지만 아직까지 국내에서는 이러한 연구는 전무한 실정이고 또한 국내의 일부 음성인식관련 기업에서 연구중인 VoiceXML 관련 연구도 단순히 ARS 등의 전화기를 이용한 서비스만을 개발하고 있는 상황이다. 하지만 본 논문은 VoiceXML 이 차세대 VUI로 충분히 활용될 수 있음을 세가지 시나리오를 바탕으로 나타내었다.

VoiceXML VUI 구성시 다음과 같은 사항을 고려해서 작성해야 사용자에게 편리한 인터페이스를 제공할 수 있다[7].

- 어플리케이션의 목적은 무엇인가?
- 사용자에게 어떤 도움을 줄 수 있는가?
- 음성제공으로 도움이 되는가?
- 사용하기 편리한가?

앞서 설계된 세 가지 시나리오는 현재 가장 많이 사용되어지는 웹사이트의 구조를 분석하여 구분한 내용이다. 하지만 대규모의 웹사이트의 경우 여러 가지 상황의 복잡하게 얽혀서 구성되어 있는 것이 현실이다. 그러므로 널리 사용되어지기 위해서는 VoiceXML을 표현할 수 있는 웹 브라우저뿐만 아니라 더욱 다양한 기능을 수행할 수 있는 VoiceXML Dialog에 대한 연구가 있어야 할 것이다.

참고문헌

- [1] 방초균, 웹환경에서 VoiceXML을 이용한 음성인터페이스 활용방안, 군산대학교 교육대학원 석사, 2002. 6
- [2] 홍용택, XML과 VoiceXML의 효율적인 통합방안 제시, 군산대학교 교육대학원 석사, 2002. 2.
- [3] <http://www.voicexml.org>
- [4] 박섭형 저, VoiceXML, 한빛미디어, 2001.12.27
- [5] 월드와이드웹 컨소시엄,
<http://www.w3.org/Submission/2001/13/>
- [6] voicexmlplanet, <http://www.voicexmlplanet.com>
- [7] Eve Astrid Andersson 외 7, early adopter VoiceXML, wrox, 2002. 8.