

음성정보 서비스를 위한 VoiceXML 저작도구 설계

김성범^{*)} 홍현술^{**)} 한성국^{*}

^{*}원광대학교 컴퓨터공학과, ^{**}원광보건대학 컴퓨터응용개발과
{piscs^{*)}, skhan^{*}}@wonkwang.ac.kr, hshong@wkhc.ac.kr^{**}

Design of VoiceXML authoring tool for Voice Information Service

Sung-pum Kim^{*)} Hyeun-sool Hong^{**)} Sung-kook Han^{*}

^{*}Dept. of Computer Engineering, Wonkwang University

^{**}Dept. of Computer Application & Development, Wonkwang Health Science College

요 약

음성정보 기술의 발달은 음성 마크업 언어인 VoiceXML1.0의 등장으로 인하여 기존의 음성 정보 기술은 보다 많은 발전 기회를 가지게 되었다. 그러나 현재 음성정보 서비스를 위한 기술은 많이 발전된 반면 이를 위한 마크업 언어인 음성 마크업 언어의 활용에 대한 연구는 부족한 실정이다.

이에 본 논문에서는 음성연구의 한 분야로 VoiceXML의 필요성과 기반 기술에 대하여 알아보고, 이를 활용하여 음성 정보 서비스를 위해 설계 요구사항을 정립하여, 저작도구의 구성요소를 기능별로 설계하였고 도 프로토타입으로 검증하였다.

1. 서 론

VoiceXML(Voice Extensible Markup Language)은 음성정보 서비스를 위한 문서 개발의 표준으로 제공되고 있다. 이에 따라 국내·외 관련 업체에서 다양한 서비스를 제공하기 위해 VoiceXML 서비스 시스템을 구축하고 있다.[1,2] 또한 VoiceXML 문서를 저작하는데 많은 비용과 시간의 낭비를 해결하기 위하여 VoiceXML 저작도구를 개발하고 있으나 몇가지 문제점들이 있다.

첫째, VoiceXML 저작도구의 형태가 텍스트 기반으로 이루어져 있어 비용과 시간적인 낭비를 해결하지 못하고 있다.

둘째, 계층적인 형태로 제공되고 있는 저작도구를 제공하고 있으나, VoiceXML 문서를 작성하는데 VoiceXML 문서 형태를 숙지하고 있어야 한다.

이에 본 논문에서는 VoiceXML 문서를 작성하는데 있어서 VoiceXML 문서 형태를 숙지하지 않더라도 쉽게 작성할 수 있는 GUI(Graphic User Interface) 환경의 저작도구를 분석 설계하여 위의 문제점을 해결하도록 하였다.

본 논문의 구성은 다음과 같이 작성 되었다. 먼저, 제2장에서는 관련연구로 VoiceXML의 활용과 XML 파서에 대하여 알아보고, 3장에서는 VoiceXML 저작도구 기능과 구성요소의 설계에 대해서 보이며, 제4장에서는 향후 연구 내용 및 결론으로 끝맺는다.

2. 음성정보 관련 기술

2.1 음성정보 서비스

웹은 인프라가 풍부하고 확장성이 좋기 때문에 다양한 서비스가 구축되어 왔으며, 현재도 웹 환경으로 이전하려는 노력이 계속되고 있다. 그러나 웹 서비스의 한계점은 인터넷 접속이 가능한 도구를 필요로 하기 때문에 많은 사용자에게 다양한 웹 서비스를 제공하는데 제한점을 갖고 있으며, 이를 극복하기 위해 새로운 웹 서비스 접근 방법을 찾게 되었다.[2]

이와 같은 문제점들을 해결하기 위해 XML(Extensible Markup Language)을 기반으로 한 VoiceXML을 통하여 음성정보 서비스를 실현해 나갈수 있는 기틀을 제공하였다. 음성정보 서비스를 실현은 인터넷을 통한 실시간, 또는 비실시간 정보서비스를 이용하여 웹 정보의 활용을 극대화 한다. 그림 1은 음성정보 서비스의 과정을 나타내고 있다.

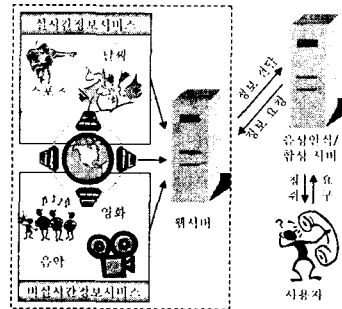


그림 1 음성정보 서비스

2.2 VoiceXML

VoiceXML은 AT&T, IBM, 루슨트와 모토로라가 VoiceXML 포럼을 설립하여 제안하고, W3C(World Wide Web Consortium)에 VoiceXML 포럼의 제안을 받아들여 대화형 마크업 언어 표준으로 공인하였다.[3,4,5] 이로 인해 국내·외적으로 VoiceXML에 대한 많은 연구가 활발히 진행중이다. 비보컬, 모토로라, 루슨트, IBM, 텔미, VoiceGenie 등 국외 기업에서는 VoiceXML 해석기를 개발하였으며, VoiceXML 서비스를 제공하고 있다.[6,7] 대부분 서비스를 목적으로 연구되고 있어, VoiceXML 문서의 작성은 미비하다. VoiceXML이 대중화 되지 않는 가장 큰 이유는 VoiceXML 문서 작성을 할 수 있는 전문인력의 부족이다. 따라서 VoiceXML 문서를 쉽게 작성할 수 있는 저작도구가 요구된다.

2.3 XML 파서

XML은 구조화된 문서를 웹에서 전송 가능하도록 설계한 표준화된 형식이다. 따라서 XML 문서를 처리하기 위해서 파싱 과정이 필수적이다. 파서는 XML 문서를 객체화 하여 트리를 구성하고, 트리에 대한 다양한 API(Application Programming Interface)를 제공하여 문서의 접근을 가능하게 한다. 이러한 파서의 종류에는 DOM과 SAX가 있다.

2.3.1 DOM(Document Object Model)

DOM은 W3C에서 제안하였고, 버전은 레벨3이다. DOM은 트리 구조기반의 API를 제공하여 XML 문서를 파싱 한 후 메

모리에 트리 형태의 자료구조를 형성한다. 때문에 트리구조를 생성하지 않는 SAX에 비해 시간이 오래걸릴뿐 아니라 큰 용량의 문서 처리에는 적당하지 않다. 그러나 XML의 계층구조를 파싱 후 DOM 트리가 그 형태를 가지고 있기 때문에 저작도구 등의 응용프로그램 작성시 많이 활용된다.[8.9.10]

2.3.2 SAX(Simple API for XML)

SAX는 이벤트 기반의 XML 분석을 위해 제안된 표준 API로 버전은 2.0이다. SAX는 XML 문서를 파싱 해 나가는 과정에 특정 이벤트를 발생시킨다. 이 이벤트를 통해 XML 데이터를 처리한다. 따라서 DOM을 사용한 문서 처리보다 빠르고 낮은 메모리를 사용하여 처리하게 된다. SAX는 이벤트를 기반으로 하는 API를 제공하여 문서의 구조, 엘리먼트 타입 이름, 속성 값 등의 정를 얻을 수 있는 콜백 기능의 핸들러로 등록후 파싱을 하면서 데이터를 처리하게 된다. 따라서 동일한 XML 문서를 반복하여 사용할 경우 여러번 파싱을 하는 단점이 있다.[11.12]

3. VoiceXML 저작도구 설계

음성정보 서비스 구축을 위한 VoiceXML 저작도구는 개발자와 일반인들이 편안하게 사용할 수 있도록 GUI 환경으로 구성하였다. 그림 2는 VoiceXML 문서 작성의 일반적인 흐름도이다.

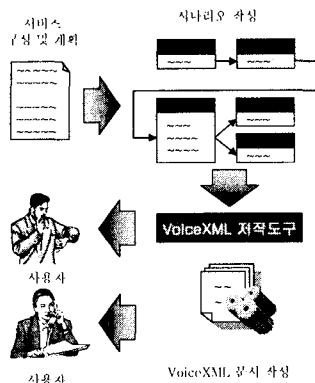


그림 2 VoiceXML 문서 작성의 흐름도

3.1 VoiceXML 저작도구의 기본 기능

VoiceXML 저작도구는 VoiceXML 문서를 작성하기 위해서 VoiceXML 문서의 구성요소인 엘리먼트와 에트리뷰트등을 손쉽게 입력하거나 수정할 수 있어야 한다. 또한 작성한 문서에 대한 유효성 검사를 할 수 있어야 한다.

표 1 VoiceXML 저작도구 기능

기능	내용
문서 편집 기능	- 엘리먼트 뷰는 VoiceXML DTD를 분석하여 트리뷰 형태로 제공 - 속성 뷰는 VoiceXML 엘리먼트를 참조하여 속성리스트뷰로 제공 - 소스뷰는 VoiceXML 엘리먼트와 속성 리스트를 참조하여 텍스트뷰로 디스플레이
유효성 검사 기능	- 유효성 검사는 VoiceXML 문서를 파서인 DOM(Document Object Model)을 이용하여 분석

3.1.1 저작 창

VoiceXML 저작도구는 기본적으로 계층(엘리먼트)도 편집과 속성(에트리뷰트)값 입력 창을 제공한다. 계층도 편집은 XML 문서의 구조적인 특성을 이용한다. VoiceXML 문서를 편집하

기 위하여 트리뷰 형태로 보여지게 된다. 또한 사용자를 위하여 VoiceXML DTD를 분석 저장하여 사용자가 원하는 정보를 제공하게 된다. 에트리뷰트 입력은 VoiceXML 엘리먼트를 참조하여 속성리스트뷰에 값을 입력 하도록 한다. 소스뷰는 저장된 VoiceXML 문서를 참조하여 텍스트 형태로 보여준다.

3.1.2 유효성 검사

사용자가 VoiceXML 문서를 작성하면 작성된 VoiceXML 문서에 대해 문법적인 오류가 존재하는지를 검사할 수 있다. 이는 VoiceXML 문서를 열때와 VoiceXML 문서를 새로 작성하거나 편집하거나 저장할 때 검사를 한다. 또한 소스뷰를 통해 VoiceXML 문서를 보여줄때 파서를 통하여 문서의 유효성을 검사한다.

3.2 VoiceXML 저작도구의 구성 요소

VoiceXML 저작도구의 구성 요소는 그림 3과 같이 크게 3부분으로 나누어지며, 본 장에서는 각각의 요소에 대해서 기술한다.

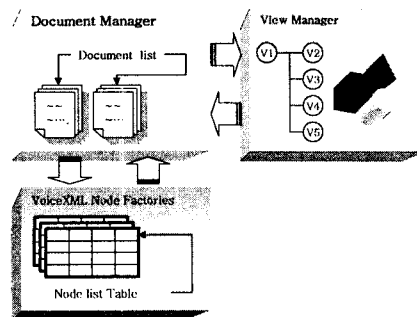


그림 3 VoiceXML 저작도구 구성요소

3.2.1 View Manager

사용자와의 상호작용을 담당하는 View Manager는 VoiceXML 문서 작성시 UI(User Interface)를 제공한다.

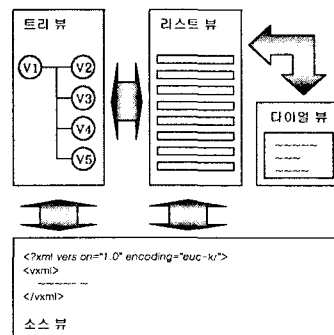


그림 4 View Manager 상세도

View Manager는 그림 4와 같이 트리뷰, 리스트뷰, 소스뷰, 다이얼 뷰로 이루어져 있다. 먼저 트리뷰는 VoiceXML 엘리먼트를 구조화하여 보여진다. 각 엘리먼트는 각각의 에트리뷰트를 갖게 되는데 이는 리스트 뷰에 나타난다. 각각의 에트리뷰트는 값을 갖게 되는데, 일반적으로 다이얼뷰를 이용하여 값을 입력하게 된다. 소스뷰는 엘리먼트와 에트리뷰트의 값이 추가 변경 되었을때 내용을 텍스트로 보여준다.

3.2.2 Document Manager

Document Manger는 작성된 VoiceXML 문서의 실제적인 데이터 저장을 담당한다. 이는 문서들의 효율적인 관리를 목적

으로 구성되었다. 따라서 View Manager와 VoiceXML Node Factories의 중간단계로 View Manager를 통한 사용자 입력 정보를 저장 관리하며, VoiceXML 문서의 엘리먼트 정보와 에트리뷰트 정보의 전달과 저장을 담당한다.

3.2.3 VoiceXML Node Factories

VoiceXML Node Factories는 VoiceXML 저작도구의 효율성을 높이기 위해 디자인되었다. 이는 VoiceXML 문서 구조를 분석하여 VoiceXML Node를 관리하게 된다.

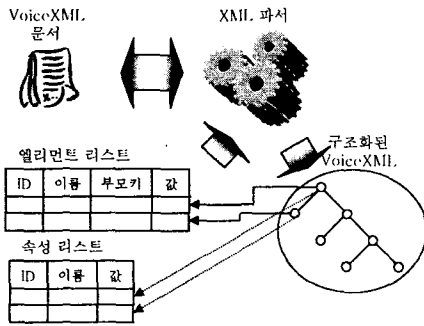


그림 5 Document Manager 상세도

그림 5는 VoiceXML Node Factories의 흐름도이다. 파서는 VoiceXML 문서를 구조화하여 엘리먼트와 속성리스트를 추출한다. 추출된 정보는 Document Manager의 요청에 따라 엘리먼트 정보와 에트리뷰트 정보를 전달하게 된다. VoiceXML 문서와 VoiceXML DTD를 분석하여 테이블과 이들의 관계 집합 정보를 갖게 된다. 이는 VoiceXML 정보들을 효율적으로 관리하며, 해당 정보를 능동적으로 제공하기 위한 구조를 갖고 있다. 즉, Document Manager의 요청을 받게 되면 Node Factories에서는 해당사항을 분석하여 데이터 테이블에서 원하는 정보를 전달하게 된다.

3.3 VoiceXML 저작도구의 사용자 인터페이스

그림 6은 저작도구의 프로토타입을 보여주고 있다. 화면 구성은 엘리먼트의 계층도와 속성값 입력창으로 구성되어 있으며, 여러 문서를 작성하기 위해 멀티창으로 디자인하였다.

화면 왼쪽의 VoiceXML 문서의 엘리먼트 구조를 보여주고, 각각의 엘리먼트를 선택하면 오른쪽 화면의 속성값 입력창에 해당 엘리먼트의 속성리스트가 보여지게 된다.

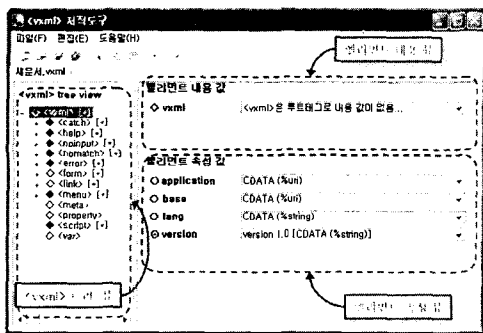


그림 6 VoiceXML 저작도구의 프로토타입

문서를 작성중이나 작성후 VoiceXML 문서를 볼때는 소스뷰를 통하여 VoiceXML 문서를 보여주게 된다. 그림 7은 프로토타입의 결과로 생성된 VoiceXML 문서 뷰를 나타낸다.

```

<?xml version="1.0" ?>
<!-- Root document -->
<xml version="1.0" ?>
  <meta name="author" content="Sung-pum, Kim" />
  <meta name="copyright" content="Kim Publishing 2001" />
  <meta name="description" content="Root document for our book store
    to access call (xxx) - xxx - xxx" />
  <!-- properties -->
  <property name="tetchaudio" value="tetchaudio.wav" />
  <!-- variables -->
  <var name="total_order" expr="0" />
  <var name="total_book" expr="0" />
  <!-- links -->
  <link next="calloperator.xml" tetchaudio="tetchaudio.wav" />
  <link
    <grammar type="application/x-jsgf" operator/>grammar)
</xml>
    
```

그림 7 생성된 VoiceXML 문서 뷰

4. 결론

기존의 저작도구는 개발자나 일반인들이 편안하게 사용할 수 있는 기능을 제공하지 못하고 있다. 이는 아직까지 VoiceXML의 구조가 일반인들이 편히 사용할 수 있는 환경을 만들기엔 부족한점이 많기 때문이다. 이에 음성정보 서비스를 위한 기술의 발전 과정과 핵심기술에 대해서 알아보았고, 음성정보 서비스를 효과적으로 제작하기 위한 VoiceXML 저작도구를 설계하였다. 분명한 점은 음성정보 기술이 앞으로도 계속 발전해 나갈 것이라는 점이며, 우리는 이를 활용할 수 있는 방안에 대해 연구가 지속되어야 할 것이다.

향후 음성정보 서비스 제공을 위한 VoiceXML 저작도구를 구현 하여야 할 것이다. 또한, 다양한 음성정보를 이용하기 위한 멀티미디어 데이터의 표현 방법 및 효율적인 정보 관리에 대한 연구가 필요하다.

참고문헌

- [1] 홍광석, 음성 인터페이스를 위한 VXML 기술, 한국인터넷정보학회, 제2권 1호, 69~74쪽, 2001.3.
- [2] 홍준희, 최중인, VoiceXML을 이용한 텔레포니 인터넷 시스템, 제 11회 한국정보과학회 인간과 컴퓨터 상호작용 연구회 학술대회, 2002.
- [3] VoiceXML Forum, <http://www.voicexml.org>
- [4] 정태의, VoiceXML, 한국정보처리학회지, 제8권 3호, 17~26쪽, 2001.5.
- [5] W3C, "Voice Extensible Markup Language(Voice XML) version 1.0", <http://www.w3.org/TR/voicexml>, W3C Note 05 May 2000.
- [6] 정채영, 이승호, 배종민, XML 에디터, 한국정보처리학회, 제8권 3호, 10~16쪽, 2001.5.
- [7] 김정란, 홍기영, VXML 편집기와 음성 브라우저의 설계 및 구현, 한국정보과학회 국제학술대회, 27권 1호(B), 414~416페이지, 2001.2.
- [8] Document Object Model(DOM) Level1 Specification(Second Edition) Version1.0, W3C Working Draft, October 02, 2000.
- [9] Document Object Model(DOM) Level2 Core Specification Version 1.0, November 14, 2000.
- [10] Document Object Model(DOM) Level3 XPath Specification Version 1.0, August 30, 2001.
- [11] The Simple API for XML(SAX) Version 1.0, May 15, 1998(<http://www.saxproject.org/>)
- [12] The Simple API for XML(SAX) Version 2.0, May 05, 2000(<http://www.saxproject.org/>)