

## ● 목 차 ●

1. 서 론
2. 국내외 기술개발 현황
3. VXML 분석
4. Voice Browser
5. 한국어 VXML 엔진
6. 결 론

## 1. 서 론

현재 진행되고 있는 XML 응용 분야의 표준은 각 분야별로 구체적으로 진행되고 있으며 이러한 시점에서 AT&T, 루슨트 테크놀로지스, 모토롤라 등 3사는 전화와 인터넷 서버와의 연동을 음성 처리 기술을 바탕으로 하여 기존 인터넷의 다양한 정보를 검색 처리할 수 있는 VXML(Voice Extensible Markup Language)이라는 인터넷 음성처리 표준안을 마련하고 있다[8]. 이 기술을 개발하기 위해 우선 이들 3사는 VXML 포럼을 구성하였으며, 음성 처리기술에 기반한 VXML 기술 스펙을 발표하고, 웹 언어 표준화 단체인 월드와이드웹 컨소시엄에 제안한 상태이다. 3사가 추진하는 VXML은 PC, 노트북 PC 등 인터넷 접속장치 없이 전화를 통해 인터넷의 전자우편, 날씨정보, 교통정보 내용을 검색할 수 있고, 웹 페이지에 다양한 음성정보를 실을 수 있게 하는 음성기반 기술이다. 후에 IBM이 포럼에 가세하면서 제시되어진 스펙 Ver. 0.9가 소개되었다. VXML은 이전에 이미 모토롤라가 개발중인

VOXML이라는 마크업 랭귀지를 기반으로 수정하여 제시되어 졌으며, IBM에서는 자바 머신에서 돌아가는 인터프리터를 현재 Ver. 1.0 까지 구현되어 있는 상태이다. 아직까지는 XML과는 관계없이 독자적으로 돌아가는 인터프리터이지만, 곧 이를 XML과 연계하여 내용을 계획으로 알려져 있다. 현재로서는 완벽히 결정된 표준안이 제출되어진 상황이 아니므로 가시적인 형태로 보이는 브라우저나 인터프리터는 달리 제시되어진 것이 없다. 그러나 VXML이 상용화할 경우 음성을 활용한 웹 기반의 콘텐츠 서비스와 음성인식 애플리케이션 개발에 중요한 영향을 미치게 될 전망이다.

또한 이 포럼에서는 앞으로 VXML을 인터넷 음성지원 표준기술로 마련하기 위해 통신업체, 장비업체, SW업체 등과 공조하는 한편 이 기술을 기반으로 다양한 인터넷 음성지원 애플리케이션 개발에 나설 계획이므로 각 분야에 걸치는 파급효과가 크다고 할 수 있다. 이 같은 4사의 포럼 구성은 통신 서비스 사업자, 통신 장비업체, 통신 단말 기업체의 인터넷 전략이 맞아 떨어졌기 때문으로 풀이된다. AT&T는 VXML 기술 개발을 통해 자사의 통신 서비스 고객에게 다양한 인터넷 기반 통신서비스를 제공할 것으로 전망되며, 루슨트 테크놀로지

\* (주)보이스미디어텍 부설연구소 연구원

\*\* 서경대학교 컴퓨터과학과 전임강사

\*\*\* 서경대학교 컴퓨터과학과 조교수

스는 음성기반 네트워크, 통신장비 판매를 강화할 것으로 보인다. 또한 모토롤러는 이를 통해 인터넷을 활용할 수 있는 휴대폰 판매에 박차를 가할 것으로 예상된다. 현재 국내에서는 작년 말에 XML 문서를 표준으로 지정하였고, VXML에 관련된 연구는 매우 미미하게 이루어지고 있는 것으로 알려져 있다.

## 2. 국내외 기술개발 현황

VXML 연구개발이 진행되는 동안에도 전화망 서비스 업체나 음성인식 서비스 업체, 심지어는 웹 서버 시스템 업체들에 의한 VXML에 관한 관심은 날로 증폭되어 가고 있다. 그 동안 ViaVoice[7]를 내세워 VXML를 지원하겠다고 선언한 IBM은 자사의 WebSphere서버를 완성하여 현재 서비스에 들어갔으며, Microsoft는 굳이 VXML이라는 명칭을 사용하지 않고 자사의 SAPI[6]에 XML을 적용할 수 있도록 하여 간접적으로 VXML을 개발하여왔다. 이외에도 뉘앙스와 스피치 워크 등 경쟁한 음성 서비스 업체들은 VXML을 지원하기 위하여 자사의 플랫폼 및 서버를 개발 및 업그레이드하고 있는 상황이다.

VXML은 다른 XML응용언어들하고는 달리 개발에 있어서 전화기를 적용하여야 하므로 하드웨어 부분에 대해서도 고려해야 한다. VXML은 XML의 subset이므로 XML에 큰 영향을 받는다. 이는 차후 XML이 월드와이드 웹에서 사용되어질 것을 생각하면 XML 파서가 가지는 영향이 크다고 할 수 있다. 현재 소개된 파서들은 업체를 기준으로 하였을 때 크게 2가지로 분류 할 수 있다면, 마이크로소프트사의 파서 - 이는 다시 자바를 지원하는 XML 파서와 C/C++을 지원하는 XML 파서로 나뉘어 질 수 있다 - 와 일반적인 자바를 지원하는 파서를 예로 들 수 있다.

대부분의 업체들이 전화기 인터페이스를 고려하

여 위 두 가지 중 한가지의 파서를 사용하게 된다. 현재 개발되어진 자바 시스템이 가상머신을 탑재하여 여러 가지 종류의 시스템을 지원할 수 있지만 로우레벨의 하드웨어 컨트롤은 그렇게 쉽지만은 않다. 이런 이유로 IBM을 제외한 대부분의 업체들이 C/C++을 지원하는 XML 파서를 선택하는 상황이며[1], 국내의 대부분 업체들 또한 MSXML (C/C++지원)을 선택하고 있다. 이는 VXML에 필수 불가결한 음성인식 엔진 또한 C/C++기반으로 개발되어져 있는데도 그 이유가 있다. 그러나 앞으로의 웹 환경을 고려하였을 때 자바를 무시한 VXML의 개발은 있을 수 없다. 비록 마이크로 소프트가 자

### ○ 국외

업 체	세부사항	비 고
IBM (www.ibm.com)	JAVA	WebShpere Server 개발
Microsoft (www.microsoft.com)	MSXML	SAPI
Speechworks (www.speechworks.com)	비공개	Speechworks6
뉘앙스 (www.nuance.com)	비공개	Nuance Voyager

### ○ 국내

업 체	세부사항	비 고
한국통신 (www.kt.com)	MSXML 및 JAVA	VoiceXML 포럼 공식Supporter
미디어포드 (www.mediaford.co.kr)	MSXML	VoiceXML 서버, 브라우저, 에디터 개발 VoiceXML 포럼 공식Supporter
제나웨이 (www.xenaway.com)	비공개	2000년6월 VoiceXML Forum 공식 Supporters업체. 현재 개발상황 비공개
zestel (www.zestel.co.kr)	비공개	VXGate개발 ( CTI성격이 강하다 )
코난 테크놀로지	비공개	소프트웨어부분 구현 전화기 인터페이스는 제외

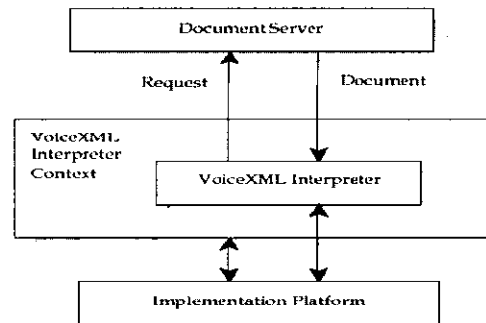
바에 대항하기 위한 닷넷 프레임워크를 개발중이라고 하지만 그 동안 개발되어온 자바 환경을 생각했을 때 그 이점을 놓칠 수 없는 것이다.

다음은 현재 VXML을 지원하거나 개발하는 국내의 대표적인 회사 및 상품 특징에 대해서 소개한다.

### 3. VXML 분석

VXML은 AT&T, IBM, 루슨트 테크놀로지, 모토라 등 정보통신 분야의 4개 거대 기업에 의해 설립된 VXML Forum에서 제안한 대화형 마크업 언어로써 음성 어플리케이션 개발을 위해 고안된 XML 문서 형식의 일종이다. VXML Forum은 지난해 8월 VXML 0.9 Version을 발표한 후 올해 3월 이를 크게 보완하여 Version 1.0을 정식 제안했다. VXML(Voice eXtensible Markup Language)은 오디오, 음성인식, DTMF, 음성입력, 전화 등 음성 기반의 사용자 인터페이스를 통해 web 정보를 검색하는 어플리케이션 개발을 위해 고안된 XML 문서 형식의 일종이다. VXML을 이용함으로써 기존의 음성서비스 제공업자들은 다양하고 쉽게 새로운 서비스를 제공할 수 있게되며 음성 서비스 저작에 대한 업무를 분할하게 된다. 즉 음성 서비스 저작자는 음성 입출력에 관한 기술적인 문제에 대해서는 알 필요가 없고 단지 저작자는 VXML 문서를 통해 서비스 자체에만 집중할 수 있게 한다.

이러한 서비스를 위해 필요한 요소들로는 VXML Parser, Interpreter, Interpreter Context를 들 수 있으며, DOM(Document Object Model)과 SAX2(Simple API for XML)를 이용하여 인터페이스하고 파싱 할 수 있다[1]. Interpreter는 dialog 단위로 사용자 요구에 대한 결과를 전달한다. VXML 인터프리터 컨텍스트는 시스템 제반 환경에 따라 VXML 인터프리터를 통제, 관리하고 platform과 인터페이스를 담당한다.



(그림 1) VXML Architectural Model

#### 3.1 VXML Parser

VXML 파서는 내부적으로 XML 파서를 내장해 VXML 태그 내포 구조를 분석하고 이 구조가 VXML DTD(Document Type Definition)에 맞는지 유효성 검사를 한다[8]. XML 파서의 경우 많은 공개 버전이 이용되고 있으며 한글 처리와 웹 접근에 문제가 없도록 유의해야 한다. 예러에 대한 핸들링은 DOM과 SAX 모두를 지원한다.

#### 3.2 VXML Interpreter

VXML 실행 환경의 가장 핵심적인 요소인 VXML 인터프리터는 문서를 적재하고 그 내용을 해석하고 실행하는 역할을 담당한다. 47종의 각 태그에 설정된 기능에 따라서 문서 실행의 순차적 흐름을 제어하고, 음성 입출력 내용을 결정해 음성 플랫폼에 필요한 명령을 내린다. 또한 문서 서버를 통해 필요한 자원을 다운 로드하거나 다른 문서로 전이하는 등 VXML 문서 실행을 총괄적으로 제어하는 엔진 역할을 수행한다. 내부적으로 XML 파서를 내장해 VXML 태그 내포 구조를 분석하고, 제안된 구조가 VXML DTD(Document Type Definition)에 맞는지 검사를 한다[8].

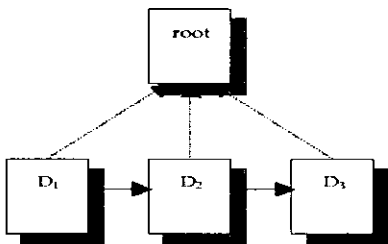
Document Server는 URI(Uniform Resource Identifier) 형태로 전송되는 VXML 인터프리터가 요청한 문서나 자원을 인터프리터에게 전송하는 역할을 담당한다. <submit> 구조에 의한 문서 전이

<표 1> XML 검증 파서 종류

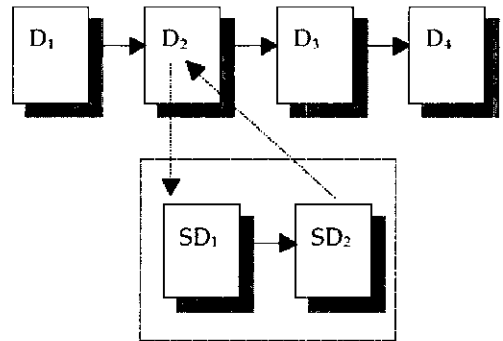
파서 이름	회사 / 기관	배포일	언어
fxd	CS Dep.Univ of Trier(DEU)	1999-06-19	SML
Java Project X TR2	Sun Microsystems	1999-09-27	Java
MSXML(preview)	Microsoft	1999-05-17	C/C++
RXP	University of Edinburgh(GBR)	1999-02-17	C
STG Validator	Brown University(USA)	1999-10-17	C++
Xj parser	Datachannel, Inc.	1999-06-21	Java
XML4C 3.1.0	IBM	2000-03-08	C++
XML4J 3.0.1	IBM	2000-03-13	Java
XML Parser for Java v2	Oracle	1999-10-17	Java
Xerces-C++	The Apache XML Project	1999-11-09	C++
Xerces-J	The Apache XML Project	1999-11-10	Java

요청의 경우 새로운 VXML 문서가 오디오나 grammar 구조의 음성 파일이나 문법 파일이 전송된다. 문서 서버는 독자적인 서버를 구축하기보다는 기존의 web 서버를 사용하는 것이 가장 간단하면서도 효율적이다. Web 서버를 문서 서버로 이용할 경우 안정적인 프로토콜에서 전송이 가능해 개발 부담을 줄일 수 있다. 또한 CGI, ASP, PHP 등을 이용할 수 있어 사용자 입력에 따른 동적 문서 생성 및 web DB와의 연동으로 실시간 정보 반영이 가능해진다는 장점이 있다.

VXML 어플리케이션은 하나의 루트(root) 문서와 이 루트 문서를 참조하는 다수의 비 루트문서로 구성된다. 루트 문서는 꼭 있어야만 하는 것은 아니고 어플리케이션에서 글로벌로 사용하고자 할 때 사용한다. 문서 전이 관계는 그림 2에 나타나 있다.



(그림 2) 문서 전이 관계

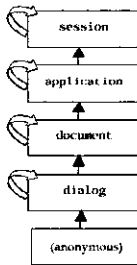


(그림 3) 다이얼로그의 전이

VXML 최상위 구조인 <vxml> 구조는 하나 이상의 대화 구조를 갖는데 다이얼로그는 주로 form과 menu로 구성되며, 사용자와의 대화에서 사용되며 입력에 대한 결과를 출력한다. 다이얼로그는 시나리오에 따라 한 문서에서 순차적으로 실행될 수도 있고 다른 문서의 다이얼로그로 전이했다가 다시 되돌아 올 수도 있다. 다이얼로그의 전이는 그림 3에 나타나 있다.

VXML 변수와 grammar들은 다섯 가지의 scope를 가지고 있다. 처음 환경변수 등 사용자가 접속되어있는 동안 모든 부분에서 이용되는 session scope(수화기를 들고 서비스를 받고 다시 수화기를 내려놓는 시점까지를 말하며, 이것은 전체 서비스

를 말하는 것은 아니다)와 여러 문서에서 함께 사용될 수 있는 application scope, 그리고 하나의 문서에서 이용될 수 있는 document scope, 하나의 form이나 menu에서 사용되는 dialog scope 하나의 element에서 사용되는 anonymous scope가 있다. VXML version 1.0에서 XML에 기초하며 정의한 태그의 수는 총 47개의 element로 구성되어 있고 각각은 10개의 범주로 분류할 수 있다.



(그림 4) 변수의 Scope

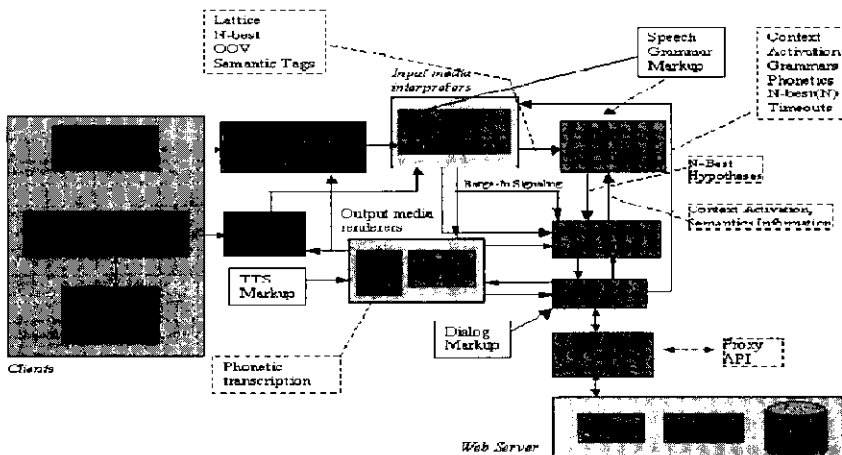
### 3.2.1 DOM (Document Object Model)

DOM(Document Object Model)은 XML 문서를 처리할 때 공통적으로 사용할 수 있는 하나의 중요한 표준이다.[2, 3, 4] 즉, 응용 프로그램이 C로 쓰여졌든 아니면 델파이나 자바, 심지어 비주얼 베이직으

로 쓰여졌다 해도 DOM을 따른다면 동일한 메소드를 사용하게 된다. XML 문서들은 보통 동적인 경우가 많기 때문에 XML 문서에 적합한 모델이 되려면, 문서의 본래 부분들을 표현할 수 있을 뿐만 아니라 새로운 내용이 추가되거나 기존 내용이 수정되어도 별다른 어려움이 없이 갱신된 문서를 표현할 수 있어야 한다. X3C의 DOM이 바로 이러한 모델이며 XML로 프로그램을 작성할 때 가장 많이 활용된다.

### 3.2.2 SAX2(Simple API for XML)

SAX2(Simple API for XML)는 이벤트 기반의 XML 분석을 위한 표준 인터페이스이며 SAX2는 DOM(Document Object Model)을 사용하는 XML 문서 처리를 빠르면서 낮은 수준의 메모리를 사용하여 처리할 수 있도록 해 준다. XML 파일 분석을 위해 DOM을 사용할 때 DOM은 완전한 문서 트리를 메모리에 구축한다. 하지만 SAX2는 문서를 이동한 후 분석 이벤트 호출을 통지해 준다. SAX2를 사용하는 가장 큰 이유 중 하나는 내용이 긴 문서의 분석을 위한 것이다. SAX2 파서를 사용하면 응용 프로그램은 이벤트 발생 시 이벤트를 모니터링 할 수 있으며 메모리에서 문서 중 필요한 부분만 읽을 수



(그림 5) 전체 VXML 시스템 구성도

있다. SAX2 파서는 ISAXXMLReader를 통해 XML 문서를 loading하고 CoCreateInstance로 인스턴스를 생성하면 생성과 파싱 작업을 수행하며 ISAX ContentHandler를 통해 수행 과정을 보여준다. 오류가 있을 경우 ISAXErrorHandler를 통해 에러 코드를 출력한다.

### 3.3 VXML Interpreter Context

VXML 인터프리터 컨텍스트는 시스템 제반 환경에 따라 VXML 인터프리터를 통제, 관리하는 역할을 담당한다[8]. 특히 ARS 서비스나 음성 포털 지원을 위한 채널 인터프리터의 경우 사용자 접속 및 해제를 감지해 인터프리터 쓰레드의 할당, 가동, 중지 등을 제어하며 시작 문서의 지정, 환경 변수와 시스템 옵션의 설정 등을 관리한다. 다양한 환경에서 인터프리터의 환경 독립성을 극대화하기 위해 설정된 외곽 요소가 VXML 인터프리터 컨텍스트이다. VXML 인터프리터를 구현하기 위해서는 Implementation Platform, VXML Interpreter Context, VXML Interpreter, Document Server 가 필요하다.

## 4. Voice Browser

### 4.1 ViaVoice와 SAPI

키보드를 사용하지 않고 음성 명령만으로 각종 데스크탑 애플리케이션을 실행시키고, 웹을 서핑할 수 있도록 지원하는 강력한 PC용 음성인식 소프트웨어인 ViaVoice[7]는 SAPI 3.0을 근간으로 하여 확장 API를 추가하여 사용하고 있다. 또한, ViaVoice는 영국식 영어, 중국어, 프랑스어, 독일어, 이탈리아어, 일본어, 스페인어, 브라질식 포르투갈어 등 다국어 언어를 수개월 안에 지원할 계획으로 알려져 있다.

IBM의 ViaVoice Millennium은 30년에 걸친 음성 인식 분야 연구개발에 기반을 둔 제품으로 음성명령만으로 문서를 작성하는 dictation 기능은 물론 인

터넷을 서핑하고, 전자우편을 송수신하는 등 편리하게 데스크탑 컴퓨터를 사용할 수 있도록 지원하고 있다. 장점으로서는 사용자의 개인적인 언어습관을 자동으로 기억하기 때문에 사용 빈도가 잦을수록 음성인식의 정확도가 높아지는 것이다. 또한 최대 12명의 음성 인식이 가능해 직장 동료나 가족들이 PC를 공유할 수도 있으며, 음성명령을 통해 AOL, 인터넷 익스플로러, 넷스케이프 브라우저 등에도 직접 접속할 수 있다.

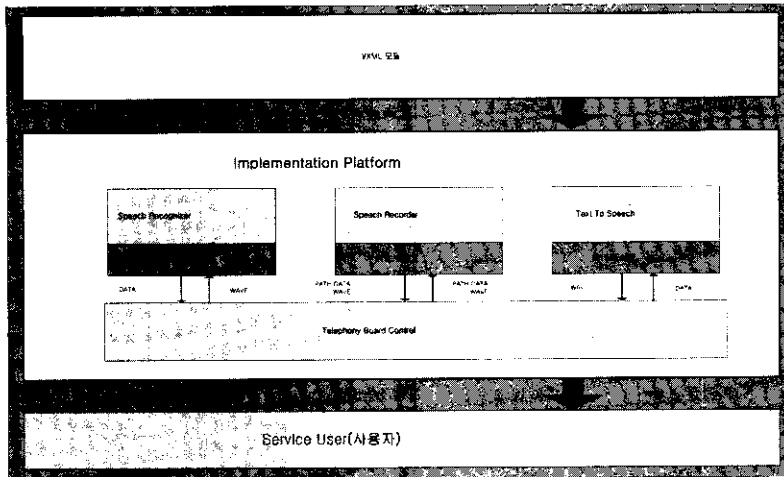
SAPI는 OLE 프로그래밍을 근간으로 하여 해당 함수들을 부르기 위한 인터페이스의 포인터를 얻을 수 있는 COM의 형태를 가지고 있다.

## 4.2 Voice Browser

### 4.2.1 VXML Browsing

VXML 인터프리터의 음성 인터페이스는 음성 처리 엔진 기능을 포함한 전화기 인터페이스 기능을 포함한다. 제안된 VXML 인터프리터로 IVR (Interactive Voice Response) 서비스 구현시 시스템의 구조를 도식화하면 그림 6와 같다. VXML 모듈에서 인터프리터 된 정보를 INAP(Intelligent Network Application Protocol)[5]을 이용하여 implementation platform과 통신함으로써 VXML에서 나타난 문서의 내용에 따라서 서비스가 되어지도록 IVR 시스템을 구현한다.

시스템의 구성을 보면 크게 VXML 모듈, implementation platform 모듈(음성인식기, TTS(Text To Speech), Speech Recorder, Telephony Board Control)로 나누어 질 수 있다. Telephony board control 모듈에서 주로 하는 일은 VXML 모듈에서 INAP에 의해서 송신된 메시지를 수신하여 전화 인터페이스 보드가 처리할 수 있는 액션을 취해주고, 처리 결과를 다시 VXML 모듈로 송신하는 것이다. 그리고 전화 인터페이스 보드의 상태에 따라서 VXML 모듈로 INAP[5] 메시지를 송신하는 것도 필요하다.



(그림 6) VXML 음성 브라우저 시스템 구조

예를 들면, 어떤 음성을 플레이하라는 VXML 문서에 대한 메시지가 오면 우선 플레이할 문서에 대한 텍스트 정보를 받아서 TTS 모듈로 보낸다. TTS 모듈에서는 이 텍스트 정보를 바탕으로 음성파일을 만들어 준다. 이 음성 파일에 대한 정보를 받아서 보드 제어 모듈에서 그대로 사용자에게 음성파일을 플레이 시켜 주면 원하는 서비스를 수행할 수 있다. 음성이 제대로 플레이되었을 경우에는 플레이가 제대로 실행되었다는 메시지를 VXML 모듈에 보내주어야 한다. 사용자의 입력을 원하는 서비스를 하는 경우에도 마찬가지로 사용자에게 안내 멘트를 전달해 주고 사용자가 DTMF톤이나 음성을 입력했을 경우에 이 메시지의 결과 정보를 VXML 모듈에 전달해 주는 일을 한다. Telephony board control 모듈에서는 VXML 모듈에서 송신하는 메시지의 내용에 따른 호 처리 및 안내 멘트 플레이 및 사용자의 음성 녹음 기능 그리고 DTMF톤 수집 등의 서비스가 언제든지 가능할 수 있도록 유연하게 구성되어야 한다.

## 5. 한국어 VXML 엔진

한국어 음성합성기(TTS)의 기술은 인간과 컴퓨

터 사이의 가장 자연스러운 의사 전달 형태인 음성 언어를 이용하여 문장을 음성으로 바꾸어 주는 문자-음성 변환 기술로서 서버가 클라이언트(사용자)에게 정보를 음성으로 변환하여 들려주고자 할 때 필요한 구성 요소이다. 여기서 서버가 하는 구체적인 일은 VXML 문서의 구문처리 중 사용자가 알아야 할 정보를 DB에서 읽어오거나 사용자에게 질문할 query를 context로 만들고 그 내용을 음성으로 변환하여 사용자에게 들려주는 것이다. 이와 같은 구성요소를 호환성있게 만들기 위해서는 분산객체 관련 COM (Component Object Model)으로 구성하여 시스템이 CTI에 연결되거나 웹에 인터넷을 통하여 연결되어도 COM의 interface를 얻어내면 어디서나 사용할 수 있게 구성하여야 한다.

일반적으로 COM요소를 생성하기 위해서는 CoCreateInstance API함수를 사용한다. 이 함수로 인해 인스턴스가 생성되면 QueryInterface를 통해 ITTSengine 인터페이스를 얻을 수 있으며, 이 인터페이스를 통해 메소드인 SpeakRendering과 GetOutputFormat을 사용할 수 있다. 이 COM요소는 NT의 MTS 또는 Windows 2000의 COM+ 서비스를 사용하여 분산객체 서비스를 받을 수 있다. MTS나 COM+에 COM을 등록하여 서비스함으로서 연계

되는 이익은 COM요소를 멀티 쓰레드를 사용하여 로컬 또는 원격으로 다수 사용자에게 서비스를 제공할 수 있는 미들웨어를 구성할 수 있다는 점이다[6].

한국어 음성인식기(Korean Speech Recognizer)는 사용자의 음성을 CTI 또는 웹 인터페이스를 통해 입력받으며, 입력된 음성의 문장의 판독된 내용과 부합되는지 여부를 결정할 수 있는 구성요소이다. 여기서 엔진은 MS의 SAPI나 IBM의 ViaVoice처럼 어떤 인식기도 사용할 수 있는 호환성을 고려하는 측면에서 COM object를 사용하여 인터페이스를 구성할 필요가 있다.

## 6. 결론

VXML의 활용은 우선적으로 기존 서비스의 질을 향상할 수 있으며, 새로운 형태의 다양한 서비스를 제공할 수 있게 되었다는 것이다. 특히, 인터넷의 활성화로 인해 매우 많은 다양한 정보가 이미 데이터베이스화하여 웹을 통하여 정보교류가 이루어지고 있고, 사용자 입장에서 자신이 사용하는 핸드폰으로 각종 정보를 검색 및 제공받을 수 있다는 것이다.

음성인식을 동반한 자동화된 음성서비스로서 여러 가지 다양한 준비된 시나리오를 통해서 친밀감을 주는 대화식 진행을 요구하는 다음과 같은 여러 가지 서비스가 제공 가능하다.

- 전자 상거래 : 상품 홍보와 판촉, 소비자 상담 및 카탈로그 소개, 주문 접수 처리, 결제 처리 등을 고급화한다. VXML 포럼에서는 음성입력 사용이 전자상거래정착을 위한 마지막 숙제라고 파악하고 있을 정도로 이 부분에 대한 관심과 기대가 크다.
- 은행거래 : 이미 은행업무는 많은 부분이 전화나 컴퓨터를 통해서 이루어지고 있다. 만약 음성인식의 기능이 추가되면 추가 가능한 고급 시나리오의 채택으로 인해 이러한 서비스를

더욱더 편리하고 유용하게 만들 수 있다.

- 실시간 정보 안내 : 날씨, 뉴스, 증권, 교통 정보 등의 실시간 정보를 제공할 수 있다. 인력에 의존하지 않으며 수집된 정보를 데이터베이스에 연결시켜 웹서버가 데이터베이스 정보로부터 VXML 문서를 자동 생성하는 방법으로 실시간 정보안내가 가능하다.
- 절차 안내 : 음식의 요리법이나 응급환자의 응급처치와 같은 절차적 내용을 갖는 여러 가지 서비스를 제공할 수 있다.
- 고객 지원 센터 운영 : 콜 센터의 운영비용은 많은 업체들에게 있어서 부담이 크다. 그렇다고 많은 인력과 자원이 투입되는 소비자와의 거래를 위한 안내 시스템을 설치하지 않을 수 없으면 이러한 사정은 조그만 중소기업들이 더욱더 심각하다.
- 설문조사 : 여론조사, 시청률조사, 소비자 정보조사 등 각종 갤럽 조사에 서비스에 유용하다.

VXML Version 1.0 이 2000년 3월에 W3C에 제출됨으로써 이제는 모든 서비스 업체들이 이를 표준으로 한 서비스 지원을 위해 기능을 구현해야 한다는 과제를 안게 되었다. VXML은 수학의 수식, 화학의 기호나 악보 등을 나타내기 위해 활용된 특수한 형태의 XML과는 전혀 차원이 다른 응용분야이다. 이는 표기의 문제뿐만 아니라 음성 처리 기술이 접목되어야 한다는 점과, 이러한 기능이 제공될 경우 사용자에게는 또 다른 차원의 인터페이스가 매우 자연스럽게 제공될 수 있다는 점에서 매우 획기적이다. 이는 입력과 출력 수단이 음성이라는 것을 의미하며 그 디바이스로서 전화를 사용한다는 것이다. 이는 다른 언어와는 달리 부가적인 장비를 제어 할 수 있는 기술이 파서 외에 더 필요하다는 사실을 말한다. 뿐만 아니라 음성의 입력/출력을 위해 음성 인식/합성 엔진과 같은 시스템이 같이 연동되어야 하므로 그 구현이 더욱 더 까다롭다.

VXML은 기존의 웹 브라우저를 사용하는 것이



아니므로 개발자를 위해서는 ECMA 등 각종 스크립트 문을 지원 할 수 있어야 한다. 기존의 웹 브라우저를 이용하게 된다면 이는 쉽게 해결할 수 있으나, VXML에서는 전화기를 이용하여 VXML 서버의 중계를 거쳐 웹 서버에 연결하는 것을 고려하므로 이러한 문제를 해결하여야 한다. 또한 전화기를 사용하므로 서비스 업체들로서는 전화기 당 채널의 한계와 전화기 인터페이스에 의한 시간 손실을 고려한 동시 접속자들을 컨트롤 할 수 있어야 한다. 마지막으로 가장 큰 문제는 VXML 포럼에서는 보다 나은 서비스를 위한 최종적 단계로 자연어 처리를 제시하고 있다. 현재 국내의 자연어 처리 기술을 보았을 때 과거와 비교하여 진일보한 것은 사실이나, 아직은 매우 미비한 상태이며 text 기반의 자연어처리 기술을 다시 음성을 고려한 자연어 처리 시스템으로 전환하고 이를 다시 한국어를 위한 VXML로 전환하기 위해서는 관련기술에 대한 습득 및 분석, 연구가 필요하다.

**참고문헌**

[1] 윤현주, 하준, 은성배, 서원균, 강상민, "VXML 인터프리터의 설계 및 구현," 제 9회 학술 발표대회 논문집, 음성과학회 2000.

[2] Document Object Model(DOM) Level 1 Specification(Second Edition) Version1.0, W3C Working Draft, 29 September 2000.

[3] Document Object Model(DOM) Level 2 Core Specification Version1.0, W3C Working Draft, 13 November 2000.

[4] Document Object Model(DOM) Level 3 Core Specification Version1.0, W3C Working Draft, 26 January 2001.

[5] DRAFT NEW ITU-T RECOMMENDATION Q.1228

[6] Guy Eddon and Henry Eddon, Inside Distributed

COM, Microsoft Press, 1998.

[7] ViaVoice SDK Documentation, <http://www-4.ibm.com/software/speech/dev/>

[8] VoiceXML FORUM([www.voicexml.org](http://www.voicexml.org)), VoiceXML Version 1.0, VoiceXML Forum, March 2000.

**저자약력**



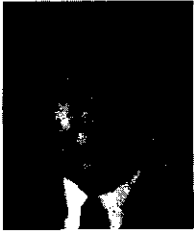
**정석영**

1999년 서경대학교 컴퓨터과학과 (학사)  
 2001년 서경대학교 컴퓨터과학과 (석사)  
 현재 고려대학교 영상정보처리 협동과정 박사과정  
 2000년-현재 (주)보이스미디어텍 부설연구소 연구원  
 관심분야: 음성인식, Human Interface,  
 e-mail : sio2@voicemediatech.com



**강선미**

1981년 고려대학교 전자공학과 (학사)  
 1983년 독일 Erlangen-Nuernberg 공과대학 전기·전자공학과 (학사)  
 1988년 독일 Erlangen-Nuernberg 공과대학 전기·전자공학과 (석사)  
 1992년 고려대학교 전자공학과 (공학박사)  
 1992년-1994년 고려대 부설 정보통신기술공동연구소 연구조교수  
 1994년-1997년 고려대 산업대학원 객원 조교수  
 1997년-현재 서경대학교 컴퓨터과학과 전임강사  
 관심분야: 음성 인식/합성, 미들웨어(VoiceXML)  
 e-mail : smkang@skuniv.ac.kr



정 태 의

1979년 고려대학교 전자공학과 (학사)  
1982년 The Ohio State Univ. EECS (공학석사)  
1989년 The Univ. of Oklahoma EECS (공학석사)  
1994년 The Univ. of Oklahoma EECS (공학박사)  
1983년-1986년 금성반도체 컴퓨터부문 선임연구원  
1986년-1987년 United Microtek, Inc. (San Jose) Senior  
Engineer  
1995년-현재 서경대학교 컴퓨터과학과 조교수  
관심분야: Computational Complexity, 알고리즘, 이동통신  
e-mail : tejeong@skuniv.ac.kr