

## 유무선 전화 음성 기반 VoiceXML 학습 평가 시스템

이인숙\* · 홍기형\*\*

### 1. 서 론

인터넷 기술의 발달로 시간과 공간의 제약을 받지 않고 언제 어디서나 학습을 할 수 있는 온라인 멀티미디어 원격교육이 본격화되었다. 원격교육 시스템은 인터넷상에서 컴퓨터 관련 기술을 활용하여 다양한 형태의 교육자료 제공, 개인별 학습 성취도의 실시간 평가, 교육 시간 및 공간에 무관한 교육 및 평가의 가능성 등의 장점을 갖고 있어 많은 연구가 진행되고 있다.

원격 교육에서 학습 평가 시스템은 기존 교육 시스템에서 제공된 평가 방법을 실시간으로 처리하여, Web 상에서 학생들이 제출한 답안을 실시간으로 평가 가능하도록 개발된 시스템이다. 원격 교육에서 학습 평가 시스템은 실제로 개인의 성적을 평가하기 위한 도구라기보다는 개인의 학습 성취도 평가를 통하여 알맞은 학습 진행과정을 선택하게 도와줄 수 있는 보조 수단으로 매우 중요하다[1].

기존의 평가 시스템은 대부분 텍스트나 그래픽 기반으로 개발되어 시각장애자들을 위한 평가나 듣기 평가에 적합하지 못하며, 대부분 PC 사용자를 가정하고 있어 피교육자의 공간이 PC를 사용할 수 있는 곳으로 제한된다. 듣기 평가나, 시각

장애인, 그리고 현재 가장 널리 보급되어 있는 휴대 전화를 단말로 이용할 수 있는 평가 시스템은 음성 기반으로 이루어 져야 하며, PC 사용자에게도 인간의 가장 자연스러운 인터페이스인 음성이 기존 텍스트와 그래픽 기반 교육 및 평가와 결합되어 제공된다면, 더욱 큰 원격 교육 효과를 가져올 수 있다.

현재 음성 인식과 음성 합성 기술은 최근에 급속히 발전하여 실용화 단계에 이르러 음성 포털을 비롯한 다양한 분야로 확대되어 가고 있고, 실생활의 편의성과 생산성의 향상을 제공해 주고 있다. 특히, VoiceXML[2]은 음성 인식 및 합성 기술을 기반으로 음성 정보 서비스의 웹 기반 개발을 가능하게 하는 XML 응용으로 음성 합성과 음성 인식을 결합한 차세대 음성 서비스를 실현하기 위한 가장 유력한 방안으로서 기대를 모으고 있다. 기존의 음성서비스는 인터넷의 다양한 정보를 제공하지 못하며, 시나리오 표준의 부재로 인해 정보 제공의 제약이 있었다. 그러나, VoiceXML은 대화형 음성 어플리케이션 개발을 위해 고안된 XML 문서 형식의 일종으로 이 문서 형식에 맞추어 작성된 VoiceXML 문서는 음성 어플리케이션에서 대화의 진행 방식을 지정하는 일종의 시나리오 역할을 하게 된다. 정보 제공자는 VoiceXML 표준에 따라 음성 입출력이 가능한 시나리오를 만들고 유무선 전화 및 웹 기반의 사용자는

\* (주)미디어포드

\*\* 성신여자대학교 컴퓨터정보학부(중신회원)

VoiceXML 인터프리터에 의해 구동되는 VoiceXML 문서 시나리오에 따라 음성 서비스를 제공받을 수 있게 된다.

VoiceXML은 AT&T, IBM, 루슨트 테크놀로지, 모토로라 등 정보통신 분야의 4개 기업에 의해 설립된 VoiceXML 포럼[3]에서 제안한 대화형 마크업 언어이다. 대화형 음성정보 서비스를 위해 표준화된 음성 애플리케이션 개발을 목적으로 하여 표준화가 진행 중인 XML 문서 형식의 일종이다. VoiceXML 포럼은 1999년 8월 VoiceXML 0.9 버전을 발표한 후 2000년 3월 이를 크게 보완한 버전 1.0을 정식 제안했고, 5월 22일 W3C 컨소시엄의 Voice Browser Working Group[4]에서 VoiceXML 포럼의 제안을 받아들여 VoiceXML을 중심으로 음성 Web의 표준화를 진행 중이다. 이에 따라 VoiceXML은 유무선 인터넷 및 전화망을 이용한 음성 정보 서비스 환경의 새로운 주역으로 그 입지가 더욱 확고해졌고, 최근에 VoiceXML 포럼에서 2.0 버전[5]을 발표하였다.

본 논문에서는 텍스트나 간단한 그래픽, 또는 녹음된 음성 파일로 구현된 평가 시스템에서 벗어나, 평가 시스템에 음성 기술을 도입하여, 음성 합성과 음성 인식을 통해 유무선 전화를 이용한 실시간 학습 평가가 가능한 VoiceXML 기반의 평가 시스템을 구현하는데 있다. 본 논문에서는 Web 서버에 연결된 CGI나 DB 등을 통하여 멀티미디어 PC를 이용한 평가 모델뿐만 아니라, 유무선 전화의 대중성과 컴퓨터의 정보 처리 능력을 통합한 CTI(Computer Telephony Integration)기술 [6]을 이용하여 전화망을 이용한 평가 IVR(Interactive Voice Response) 시스템을 설계 및 구현하였다. 평가에 있어서는 사용자의 개인차를 고려한 수준별 평가가 가능하도록 하였으며, 평가 결과는 사용자의 수준을 재평가하는 피드백 자료로 이용하였다. 본 논문에서 설계 구현한 시스템은 현재

의 음성 인식 기술로 현실적으로 실현 가능한 객관식형, OX형, 팔호넣기형, 단답형의 4 가지문제 유형을 통한 평가가 가능하다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구로서 기존 교육평가 시스템의 분석과, 전화를 이용한 음성 기반 평가를 위한 IVR 시스템의 기반 기술인 CTI 기술, 음성 인식과 음성 합성 기술, VoiceXML에 대해 살펴본다. 3장에서는 VoiceXML 기반의 평가 시스템의 구조와 평가 시스템에 필요한 데이터베이스를 설계하였다. 그리고, PC 및 전화용 음성 기반 평가 IVR 시스템의 설계와 구현 환경 및 실행 화면 등을 살펴보았다. 마지막으로 4장에서는 본 논문의 연구 결과를 요약하고, 향후 연구 과제를 제시하였다.

## 2. 관련 연구

### 2.1 교육 평가 시스템 사례 분석

Web을 이용한 학습 평가는 인터넷 기술의 발달과 함께 새로운 방법을 모색해가며 기존의 Text 중심에서 멀티미디어 기술을 결합한 다양한 방법으로 진행되고 있다. 현재 교육 솔루션에서 제공하고 있는 평가 시스템은 Text 중심의 시각적 평가 시스템, 녹음된 음성 파일을 이용한 평가 시스템, 음성 기술을 이용한 평가 시스템으로 나누어 살펴 볼 수 있다.

▶ Text 중심의 시각적 평가 시스템: Text에 그래픽 또는 오디오 파일을 추가하여 문제를 제시하는 대부분의 평가 시스템에서 사용되는 방식으로 듣기 평가나 시각 장애인에게는 사용이 불가능하다는 단점을 가지고 있다. Text 중심의 평가 시스템의 예로는 짝기. 킴[7]와 서강대학교의 멀티넷[8] 등이 있다.

▶ 녹음된 음성 파일을 이용한 평가 시스템: 평

가를 위해 사람이 문제를 직접 녹음하여, 녹음된 파일을 이용하여 문제를 제시하고 평가를 수행하는 시스템으로 주로 영어 회화나 토익 등의 어학 평가 시스템에서 많이 사용하고 있다. 녹음된 음성 파일을 이용한 평가 시스템은 평가를 위한 문제를 녹음하기 위해 녹음을 위한 별도의 장비와 장소가 필요하며, 녹음과정이 복잡하여 문제 제작이 쉽지 않고, 제출자 마다 음성이 다르기 때문에 평가 시 사용자에게 혼돈을 줄 수 있다는 단점을 가지고 있다. 대표적인 사이트로는 에듀넷의 어학 평가 시스템[9]과 영산 정보 통신의 Yes English [10] 등이 있다.

▶ 음성 기술을 이용한 평가 시스템: 좀더 발전된 단계의 평가 시스템으로 음성 녹음이 아닌 Text로 제출된 문제를 음성 합성 기술을 이용하여 합성된 음성으로 변환해 문제를 제시하는 평가 시스템이다. 이 평가 시스템은 음성기반의 문제 제공이 편리하다는 장점을 가지고 있으나, 음성 합성만을 이용한 평가로 평가를 위해 음성 인식 기술을 사용하지 않고 있으며, PC 기반의 평가 시스템으로 전화를 사용한 평가가 불가능하다는 단점이 있다. 음성 합성 기술을 이용한 평가 시스템의 예로는 하이드림의 인터넷 교육평가 시스템 [11]이 있다.

2.2 CTI(Computer Telephony Integration)

우리 생활에서 빼놓을 수 없는 가장 대표적인 통신수단은 전화와 컴퓨터이다. 사용 방법이 매우 쉽고 간단하다는 장점을 가진 전화는 그 기능이 크게 변하지 않는 가운데 최근에는 휴대전화 급속히 보급되어 핵심적이고 대중적인 통신수단으로 자리잡고 있다. 특히, 우리나라의 경우, PC를 이용한 인터넷에 비하여 IMT2000 등을 기반으로 하는 휴대전화를 이용하는 무선 인터넷이

더 큰 사용자 수를 확보할 것으로 보인다. 서비스가 복잡해지고 다양해지면서 음성 사서함, 무인 자동 교환 안내, 통합메시지 시스템 등과 같이 컴퓨터와 전화가 연결된 서비스 제공 요구가 커짐에 따라 전화의 대중성과 컴퓨터의 정보 처리 능력을 통합할 수 있는 CTI 기술[6]이 필요하다.

CTI란 컴퓨터와 전화망이 상호 협조하여 서비스를 제공하는 것이다. CTI를 위해서는 컴퓨터와 교환기 또는 컴퓨터와 전화기 사이에 물리적인 연결이 필요하며 컴퓨터 기반의 애플리케이션이 전화 서비스를 요청하고 전화 호를 추적하는 등 전화 서비스 관련 동작을 수행할 수 있어야 한다.

그림 1은 CTI 개념을 설명하고 있다. 컴퓨터 영역과 통신망 영역의 기능을 CTI로 통합하고 이를 기반으로 UMS나 텔레마케팅 서비스, Help 데스크 등 다양한 서비스를 제공할 수 있다.

CTI 응용분야는 매우 다양하며, 몇 가지를 살펴보면 다음과 같다.

▶ Audiotex: 제공하고자 하는 정보를 시스템 내의 대용량 기억장치에 음성으로 녹음해 놓고 전화를 건 사람에게 미리 녹음된 정보를 제공해주는 것이다.

▶ IVR(Interactive Voice Response) 시스템: 전화를 건 고객이 누른 전화기의 번호버튼에 즉각

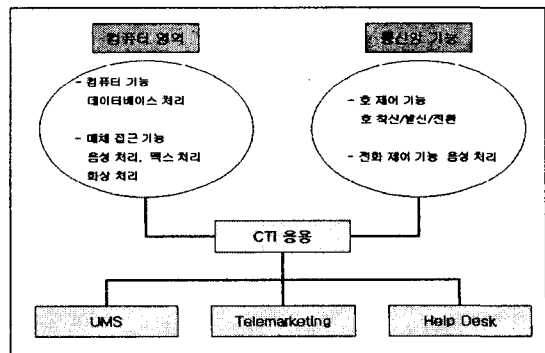


그림. 1 CTI 개념

적으로 반응하여 해당 입력에 대한 응답 처리가 데이터베이스를 통해 작동 되도록 한다. 이는 일반적으로 특정 레코드들을 조회하여 처리 수행되거나 스크립트로 문서화된 대화내용이 미리 지정된 처리 단계를 통해 수행된다. 예를 들어 은행업무에서 잔액조회, 계좌이체와 같은 서비스를 말한다.

▶ Voice Mail: 메시지를 전하고자 하는 상대방이 통화중이거나 부재중일 경우 효율적인 간접 통신 수단으로서, 전화 사용자에게 음성 메시지를 녹음하고 이를 저장, 수정하는 기능을 제공한다.

▶ Call Center: 상담원이나 교환원에게 걸려온 전화를 부분적으로 혹은 전체적으로 자동화 해준다. 오늘날의 Call Center는 CTI기술, 디지털 음성 처리 기술, 서로 다른 노드간의 음성/팩스자원 공유기술, 팩스이미지 처리기술, 디지털 스위칭 기술 등을 접목시켜 고객과 상담원 모두에게 다양한 서비스를 신속정확하고도 편리하게 제공하는 'One-Stop', One-Call' 서비스의 첨단 솔루션으로 정의될 수 있다[12].

### 2.3 음성 인식과 음성 합성

음성 인식이란 인간의 음성을 컴퓨터가 분석해 이를 인식 또는 이해하는 기술로, 발음에 따라 입 모양과 혀의 위치 변화로 특정한 주파수를 갖는 인간의 음성을 전기신호로 변환한 후 음성 신호의 주파수 특성을 추출해 발음을 인식하는 기술이다[13].

음성 인식은 적용 화자, 발음 방식, 인식 어휘로 분류할 수 있다.

#### (1) 적용 화자에 의한 분류

▶ 화자 종속(Speaker Dependent) 음성 인식 기술은 사용자가 먼저 자신의 목소리로 인식기를 훈련시키고, 인식기는 훈련시킨 특정 화자의 목소리만을 인식하는 기술로 비교적 쉬운 기술이며

인식률이 높은 반면 사용자가 훈련 과정을 거쳐야 하므로 사용상 불편함이 있다.

▶ 화자 독립(Speaker Independent) 음성 인식 기술은 사용자에게 제한이 없으며 훈련 과정도 필요 없이 모든 사람의 음성을 인식할 수 있는 기술이다. 화자 종속 음성 인식에 비해 난이도가 높은 기술이며, 요즘 많이 활용되고 있는 '음성 인식 PC' 등에 적용되고 있다.

#### (2) 발음 방식에 의한 분류

▶ 고립 단어 음성 인식(Isolated Word Recognition) 기술은 고립 음성 형태로 발음된 단어만 인식할 수 있다. 예를 들어 '우리집', '불꺼', 'TV 켜' 등 고립된 발성 형태를 인식하는 것이다. 음성의 시작점과 끝점을 알 수 있으므로 비교적 쉬운 음성 인식 기술이라고 할 수 있다.

▶ 연결 단어 음성 인식(Connected Word Recognition) 기술은 고립 단어의 연결 형태를 인식할 수 있는 기술로, '우리집 TV' 등을 예로 들 수 있다.

▶ 핵심어 인식 기술(Keyword Spotting)은 연속으로 발음된 음성에서 특정 단어만을 검색하여 인식하는 기술이다. 예를 들어 "서울에서 부산까지 비행기표 부탁드립니다"라고 발성했을 때 '서울', '부산'이라는 핵심어만 검색하여 인식하는 기술이다.

▶ 연속 음성 인식 기술(Continuous Speech Recognition)은 말 그대로 연속된 낭독체 형태의 음성을 인식하는 것으로, 음성 구술기(Dictation) 등에 적용할 수 있는 기술이다.

▶ 대화 음성 인식 기술은 일반적으로 사람과 대화하듯 대화체로 발성을 하면 이를 인식하는 기술이다. 현재 연구중인 분야로서 대화 음성 인식 기술을 가진 업체는 전 세계적으로도 아직 전무한 실정이다.

#### (3) 인식 어휘에 의한 분류

▶ 소용량 어휘 인식 기술은 수십 단어 수준

소용량의 어휘 세트에 대해서만 인식할 수 있는 기술로 음성 인식 단말기, PDA나 대화형 장난감 등에 적용되는 인식기가 이에 해당된다.

▶ 대용량 어휘 인식 기술은 수천, 수만 단어 수준의 대용량 어휘를 인식하는 기술로 음성 구술기(Dictation) 등에 적용된다.

▶ 고정 어휘 인식 기술은 인식 대상 단어가 특정 시점에는 고정된 것으로 대상 단어가 변경될 때에는 음성 인식기의 데이터베이스를 재 구축해야 하기 때문에 많은 시간과 비용이 든다.

▶ 가변 어휘 인식 기술은 대상 단어의 변경이 손쉬워 원하는 대상 단어를 자유롭게 사용자가 입력할 수 있으며, 이런 장점으로 많은 응용 분야에 활용되고 있다.

음성 합성 기술은 문자 정보 또는 기호를 인간의 음성으로 변환하여 들려주는 기술이다. 초기의 음성 합성에 대한 연구는 대부분 기계적 또는 전자회로를 이용하여 인간의 발성 기관을 모의하는 것이었지만 컴퓨터의 연산 속도 및 기억용량이 급속히 발전하면서 음성 합성에 대한 연구는 문서 처리 기술을 포함한 문자-음성 변화 기술로 확장되었다[13].

음성 합성 기술은 응용 방식에 따라 제한어휘 합성과 문자-음성 합성으로 구분될 수 있다.

(1) 제한 어휘 합성

제한 어휘 음성 합성 기술은 필요한 음성 조각을 미리 녹음했다가 이를 연결시켜 연속된 음성을 만들어 내는 기술로 현재 114의 전화번호 안내 등 자동응답 장치(ARS) 서비스에 주로 응용된다. 기술적으로 매우 간단하지만 문장 형태가 제한되며 음성 연결 부위가 부자연스러운 단점이 있다.

(2) 문자-음성 합성(TTS ; Text-to-Speech)  
TTS 기술은 합성 대상 어휘에 제한이 없으며

일반적인 문자 형태의 정보를 음성으로 변환하는 것이다. 문자-음성 합성 방법은 한국어의 모든 음소에 대한 발음 데이터베이스를 구축하고 이를 연결시켜 연속된 음성을 생성하게 되는데, 이 때 음성의 크기, 길이, 높낮이 등을 조절해 자연스러운 음성을 합성해 낸다.

2.4 VoiceXML

VoiceXML은 음성 대화 시나리오 구성, 음성 합성, 오디오 파일의 발송, 음성 및 DTMF(Dual Tone Multi Frequency) 입력 인식, 음성 녹음, 콜 전환, mixed interactive conversation, 하이퍼 링크에 의한 타 서비스 연결, 웹 서버 및 백 엔드 응용 서버 연결 등의 기능을 제공하고, 문서 내부에 스크립트 코드 및 연산 기능을 갖도록 허용함으로써 입력 정보에 대한 복잡한 정보 처리가 가능해 복잡한 시나리오도 표현할 수 있다[14]. 초기에 AT&T, IBM, 루슨트 테크놀로지, 모토로라의 4사가 전화 통신 애플리케이션을 위해 준비해 왔던 것인데, 현재는 음성 전문 업체뿐만 아니라 3Com, 휴렛패커드, 노텔 네트워크, 유니시스, 브리티시 텔레콤 등 222개의 VoiceXML 포럼 정보 기술 업체가 기술 지원을 하고 있다.

2.4.1 VoiceXML 엘리먼트

VoiceXML은 음성 입출력 기반의 음성서비스 시나리오의 표준으로 XML에 기초하며 다이얼로그, 문법, 이벤트, 오디오 출력, 오디오 입력, 콜 제어, 흐름 제어 등에 관련된 엘리먼트들로 구성된다[5]. 표 1은 VoiceXML가 제공하는 엘리먼트를 보이고 있다.

2.4.2 VoiceXML 시스템

VoiceXML 서버 시스템(그림 2)은 크게 네 가지 요소로 구성되는데, VoiceXML 인터프리터의 요청에 따라 서비스 문서 및 오디오 화일, 문법

표. 1 VoiceXML의 Element 분류

No	분류항목	상세 내용
1	Event Handling	<noinput>, <nomatch>, <help> 등의 이벤트들을 처리하고 throw 해주는 Element 집합
2	Executable Contents	변수 선언, 제어문 등의 처리하는 Element 집합
3	Forms	Form item들을 논리적 단위로 묶는 Dialog 엘리먼트
4	Form Items	사용자의 입력 등을 받아들이는 Dialog로그 요소 집합
5	Grammars	음성 인식기에서 사용되는 Grammar 지정 및 DTMF를 표현하는 Element 집합
6	Links	<link>의 범위에 있는 모든 Dialog에 적용되는 전이를 표현
7	Menus	Dialog의 한 형식인 메뉴를 표현하는 Element의 집합
8	Prompt	음성 합성기 혹은 출력 단에 보내질 Prompt를 가공, 출력하는 Element 집합

하는 구현 플랫폼, 마지막으로 가장 핵심적인 요소인 VoiceXML 문서를 파싱하기 위한 파서와 VoiceXML 문서를 해석하여 수행하는 VoiceXML 인터프리터로 구성된다. 구현 플랫폼은 VoiceXML 인터프리터의 지시에 따라 음성 합성(Text To Speech), 음성 인식(Speech To Text), 오디오 파일 출력, 음성 입력 녹음, DTMF 입력 인식 등의 음성 입출력을 수행하고 VoiceXML 인터프리터에게 결과를 보고한다[15].

2.4.3 VoiceXML로 작성한 문제지

음성 인식을 이용한 대화요소를 갖는 간단한 VoiceXML 문서의 예는 다음과 같다.

이 문서를 VoiceXML인터프리터로 실행시키면 “문제와 보기를 잘 듣고 맞는 답의 번호를 눌러

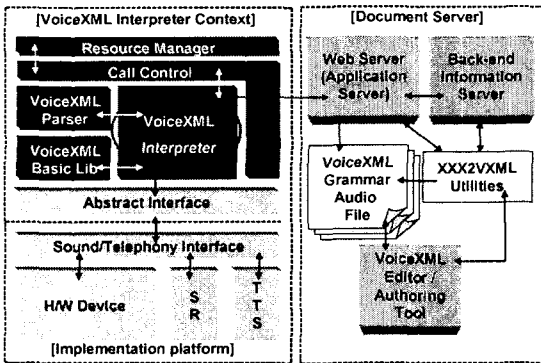


그림. 2 VoiceXML 시스템

화일 등을 제공하며 전체 서비스 및 업무 로직을 유지하는 문서 서버와 시스템의 제반 환경에 따라 VoiceXML 인터프리터를 통제, 관리하는 역할을 하는 VoiceXML 인터프리터 컨텍스트, 그리고 음성 인식 엔진과 음성 합성 엔진, 전화 액세스 지원을 위한 전화 음성 보드 및 인터페이스 등을 포함

```

<?xml version="1.0" encoding="ISO8859-1"?>
<vxml version="1.0">
  <form id="problem">
    <block>
      문제와 보기를 잘 듣고 맞는 답의 번호를 눌러주세요.
    </block>
    <field name="answer">
      <prompt>
        문제1번, 포드사에서 제공하는 펠터가 아닌 것은?
        <enumerate>
          <value expr="_dtmf"/><value expr="_prompt"/>
        </enumerate>
        <prompt>
          <option dtmf="1" value="1">정규분포에 의한 흐리기</option>
          <option dtmf="2" value="2">수채화 느낌 주기</option>
          <option dtmf="3" value="3">후추가루 같은 잠을 제거하기</option>
          <option dtmf="4" value="4">물결처럼 흔들려 보이기</option>
          <option dtmf="5" value="5">답없음</option>
        </option count="2">
        <submit method="post" next="http://Quiz/ans.asp" namelist="answer"/>
      </noinput>
      <nomatch count="2">
        <submit method="post" next="http://Quiz/ans.asp" namelist="answer"/>
      </nomatch>
    </field>
    <filled>
      <submit method="post" next="http://Quiz/ans.asp" namelist="answer"/>
    </filled>
  </form>
</vxml>
  
```

주세요. 문제1번, 포토샵에서 제공하는 필터가 아닌 것은? 1. 정규분포에 의한 흐리기 2. 수채화 느낌 주기 3. 후추가루 같은 잡음 제거하기 4. 물결처럼 흔들려 보이기 5. 답없음”이라는 음성 메시지가 나온다. <enumerate>구조 안의 내용은 일종의 템플릿으로 각 항목에 대해서 반복적으로 적용된다. <value>는 expr 속성값을 계산해 출력하는 기능을 갖는데, 특히 “\_dtmf”와 “\_prompt”는 대응되는 <option> 선택항목에서 사용할 버튼 번호나 인식 문자열을 나타내는 시스템 정의 변수이다.

문제에 대한 답을 입력하면 입력된 사용자 답은 answer에 저장된 후 http://Quiz/ans.asp에 answer를 post방식으로 전달하여 서버 쪽에서 처리되어 온 VoiceXML 문서를 수행하게 된다. 정답을 입력하지 않거나 1에서 5사이의 값을 입력하지 않으면 2회 문제를 다시 제시하고 2회 이상 입력하지 않으면 answer에는 아무 값도 저장되지 않은 채 http://Quiz/ans.asp에 answer를 전달한다.

문서에 사용된 <vxml> tag는 VoiceXML 문서임을 알려 주고, <form> tag는 단위 대화를 나타내는 것이고, <block> tag는 음성 출력을 비롯한 각종 실행 명령을 저장하는 역할을 한다.

<field> tag는 <prompt> tag에 의한 입력 안내문 출력하고, 입력 값이 들어오길 대기한 후 <filled> tag를 이용하여 입력 정보의 처리 등을 행한다.

<submit> tag는 입력 결과를 지정된 URL의 웹서버로 전송하여 다음 문서를 전송 받아 실행할 것을 지정한다. <nomatch>은 grammar에서 정의된 입력이 들어오지 않았을 경우 ‘nomatch’ event 발생시의 처리를 지정한 것이며, <noinput>은 아무런 입력이 들어오지 않았을 때 발생하는 ‘noinput’ event를 처리하기 위한 것이다.

### 3. VoiceXML 기반 평가시스템

VoiceXML 기반 평가 시스템은 음성 인식과 음성 합성을 이용한 대화형 평가 시스템으로, 기존의 PC에서의 평가뿐만이 아니라 유무선 전화를 통한 평가가 가능하도록 설계하였다. PC 또는 전화를 이용한 피교육자에게 VoiceXML로 작성된 문제지를 VoiceXML인터프리터가 해석하여 음성으로 변환하여 전달하며, 피교육자는 DTMF나 키보드 또는 음성으로 답을 입력하여 평가를 수행한다.

이러한 VoiceXML 기반의 평가 시스템의 장점은 평가를 위해 가장 대중적인 통신 수단이 유무선 전화를 사용할 수 있어 이동중이거나 컴퓨터를 사용할 수 없는 환경에서도 장소와 시간에 구애 없이 평가가 가능하며, 컴퓨터를 잘 사용하지 못하는 사람, 인터넷 사용이 불가능한 사람들도 평가를 받을 수 있다. 또, VoiceXML이 음성 입출력에 대한 표준을 제시하고 있으므로 음성 입출력의 기술적인 지식이 없이도 신속하게 개발이 가능하며, 기존의 Text기반의 평가시스템을 쉽게 VoiceXML 기반 평가 시스템으로 적용할 수 있다는 것이다.

본 논문에서 개발한 VoiceXML 기반의 평가 시스템에서는 화자독립, 가변어휘의 음성 인식 기술과 TTS 음성 합성 기술을 사용하였다. VoiceXML 기반의 평가 시스템에서는 객식형, OX형, 괄호넣기형, 단답형의 문제유형의 평가가 가능하다. 이 시스템은 방대한 양의 문제들을 문제 유형별로 분류하여 데이터베이스에 등록해 놓고 필요시 일부의 문제를 추출하여 학습자의 평가에 이용할 수 있도록 하는 문제 은행 시스템을 기반으로 하고 있다.

#### 3.1 평가 시스템의 설계

그림 3은 본 논문에서 설계한 평가 시스템의

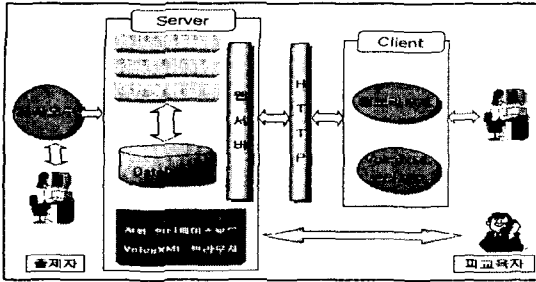


그림. 3 VoiceXML 기반 교육평가 시스템 구조

의 구조를 보이고 있다. 서버는 데이터베이스와 웹서버를 포함하고 있으며, 전화를 사용한 평가가 가능하도록 전화인터페이스 기능을 가진 음성처리 보드와 이 보드를 제어하는 VoiceXML 기반의 IVR 서버를 내장하고 있다. 클라이언트는 웹브라우저, 또는 유무선 전화기이다.

(1) 사용자 등록 모듈

사용자 등록 모듈에서는 사용자의 시스템의 사용 가능 여부를 판단하고, PC와 전화를 통하여 신규 사용자의 등록을 할 수 있다. 그림 4는 사용자 등록 모듈의 흐름도이다.

(2) 평가 모듈

평가 모듈은 문제에 대한 평가를 실시하는 부분으로 VoiceXML 문서로 작성되어 있는 문제를

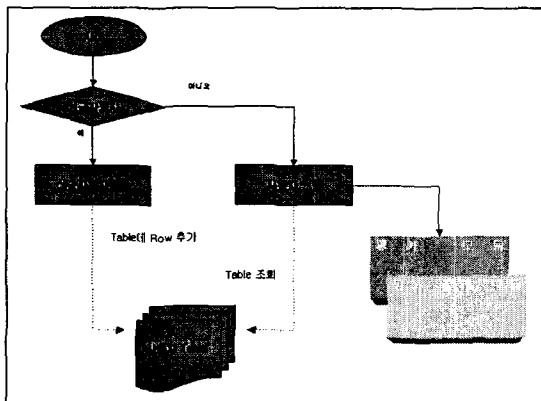


그림. 4 사용자 등록 모듈 흐름도

평가 도구를 통해 음성으로 들려주고, 키 또는 음성으로 답안을 입력한다. 문제 제시는 1번부터 순차적으로 1회씩 들려주며, 자동으로 다음 문제로 넘어 가게 된다. 마지막 문제까지 제시되면 답안지를 채점하여 출제된 총 문제 수와 정답 수, 획득 점수를 음성으로 통보한다.

평가를 위해서는 사용자가 문제지 선택 시, 이미 평가를 치른 문제지에 대해서는 다시 평가를 치를 수 없다. 따라서 사용자가 평가 문제지 선택 시 사용자 과목별 Level을 이용하여 사용자 Level에 맞는 평가가 가능한 문제지 리스트를 동적으로 생성하여 선택이 가능하도록 한다. 그림 5는 평가 모듈의 흐름도이다.

(3) 평가 결과 모듈

평가 결과 모듈에서는 평가 모듈에서 채점한 결과를 문항별로 자세히 표시해 주고, 채점 결과 값을 이용하여 사용자 개인 평균, 전체사용자 평균 등의 통계 처리와, 사용자 수준의 재결정에 사용한다. 또, 사용자가 평가받은 문제에 대한 풀이를 해준다.

3.2 평가 데이터베이스 설계

VoiceXML기반 평가시스템에서 평가를 위한

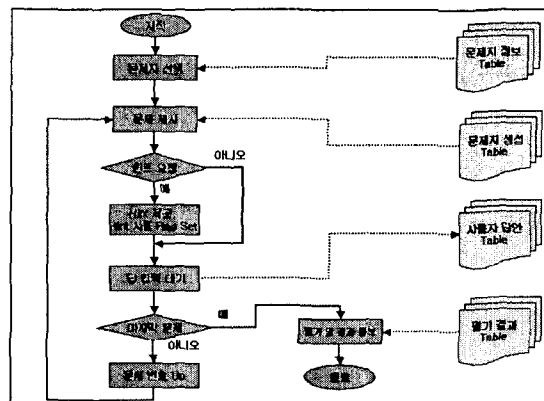


그림. 5 평가 모듈 흐름도



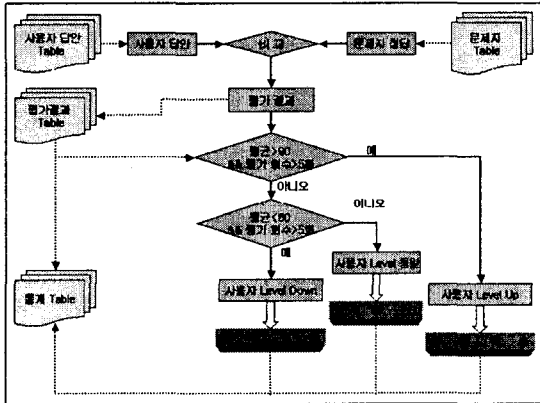


그림. 6 평가 결과 모듈 흐름도

전체 데이터 베이스는 14개의 테이블로 구성된다. 4지선다형, OX형, 괄호넣기형, 단답형의 문제 유형 별로 문제를 저장하기 위한 SELECTION QNA, OXTTYPE, PARENTH의 문제 테이블, 문제 제출자의 정보를 저장하고 있는 MAKER 테이블, 시험에 대한 정보를 저장하는 문제지 테이블인 PAPER와 Paper에 해당하는 선택된 문제 List를 저장하기 위한 MAKEPROBLEM, 사용자 정보를 관리하는 사용자 테이블인 USER, 사용자 답안을 저장하기 위한 USER\_ANSWER 테이블, 사용자의 과목별 통계 점수를 저장하기 위한 통계 테이블인 RESULTDATA, 평가 결과를 저장하기 위한 RESULT 테이블이다. 이외에 과목과 Chapter를 나타내기 위한 LESSON, CHAPTER, SECTION 테이블이 있다.

그림 7는 평가시스템의 각 모듈과 데이터베이스 테이블과의 연계를 나타낸다. 사용자가 등록을 하게되면 USER 테이블에 저장하고, 또 Login시 USER 테이블을 사용하여 사용자 인증을 하게된다. 과목 선택과 사용자의 과목별 수준에 따라 문제지를 선택하게되면, PAPER 테이블의 정보를 참조하여 각 문제 유형 테이블에서 문제를 가져와 VoiceXML 평가 도구를 사용하여 음성으로 문제를 출제하고, 사용자가 Key 또는 DTMF와 음성

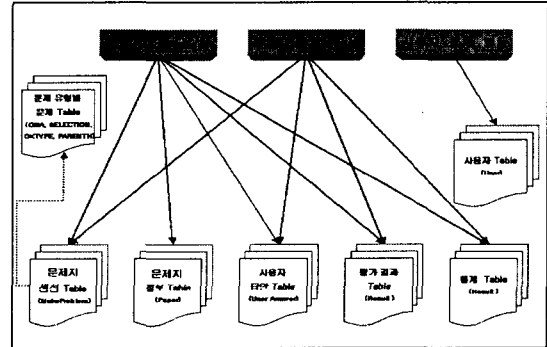


그림. 7 모듈과 데이터베이스 연계

으로 입력한 답을 USERANSWER 테이블에 저장하여 채점을 하게 된다. 채점한 결과는 평가 결과 테이블에 저장되고, 통계처리를 행한 후 통계 테이블에 저장된다.

### 3.3 평가 시스템의 구현

시스템의 개발 환경은 윈도우 NT Server 4.0을 사용하였고, 웹서버는 윈도우 NT 전용 Web 서버인 IIS(Internet Information Server) 4.0를 사용하였다. Web 운용프로그램의 구성요소로서는 HTML 파일, 데이터의 조작을 위하여 ASP(Active Server Page)를 이용하여 시스템을 구축하고, 사용자 관리 및 시험 결과 등을 저장하기 위한 데이터베이스 서버로는 SQL 서버 7.0을 사용하였다.

VoiceXML 서버는[15]에서 구현한 것을 이용하였으며, 전화용 평가 IVR System의 Call 제어를 위해서는 Dialogic사의 음성처리보드인 D/160SC-LS의 16채널용 아날로그 보드를 사용하였다.

#### (1) 신규 사용자 등록 화면(그림 8)

VoiceXML 기반 평가 시스템을 사용하기 위해서는 사용자 등록을 해야한다. 사용자가 PC를 이용한 평가뿐만이 아니라 전화를 이용한 평가가 가능하도록 4자리 이상 8자리 이하의 숫자만으로

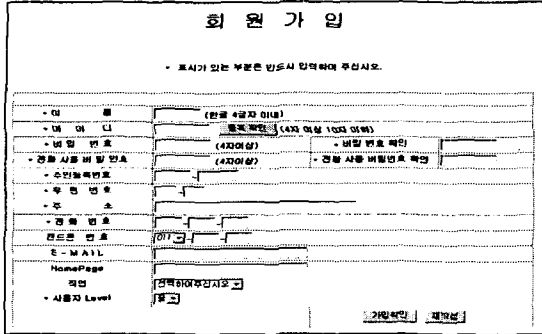


그림. 8 신규 회원 등록 화면

구성된 비밀 번호를 입력하도록 하였다. 사용자 Level은 사용자에게 따른 수준별 평가를 위하여 Default를 중으로 하여 상·중·하로 입력하도록 하였다.

(2) 평가 IVR System 관리

IVR System을 실행하려면 우선 Dialogic사에서 제공한 Voice Driver를 실행한 후, 실행 프로그램을 실행해야 한다. 표 2는 전화용 평가 시스템에서의 각 회선별 상태를 나타내기 위해 사용되는 Bitmap이다.

그림 9는 전화용 평가 시스템의 실행 관리를 위한 모니터화면을 나타낸다. 이 프로그램은 모두

표. 2 전화용 평가 IVR System 상태 표시

NO	Bitmap	설 명
1		회선 초기화가 안되어 사용 불가능 상태(녹색)
2		회선 초기화 후 사용 가능상태(청색)
3		사용자로부터 전화가 걸려와 전화를 받은 상태
4		평가 시스템에서 음성을 제시하는 상태
5		사용자로부터 DTMF 입력을 대기하는 상태
6		사용자로부터 음성입력을 대기하는 상태

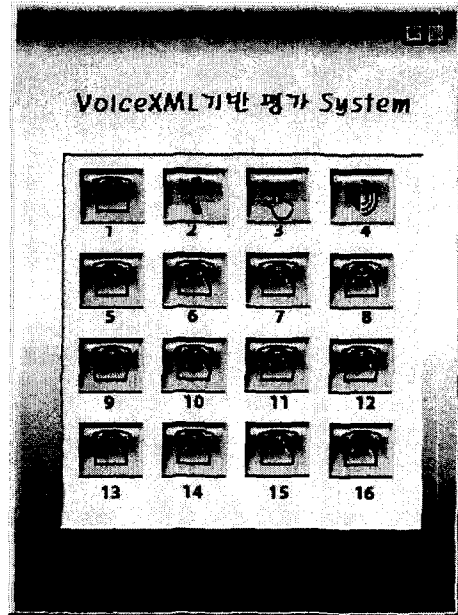


그림. 9 평가 IVR System 관리 화면

16개의 회선까지 지원이 가능하다. 만약 회선 초기화 시 오류가 발생한 경우는 표 2의 Bitmap 1에 해당하는 그림이 오류가 발생한 채널에 표시된다. 본 시스템에서는 전체 회선을 관리하기 위하여 각 회선마다의 Log데이터를 저장하여 이를 장애 처리용과 통계 처리용으로 사용할 수 있다.

(3) 평가 결과 인터페이스

그림 10은 사용자가 평가한 문제지에 대해 문항별 평가 결과를 보여 준다. 각 문항의 정답과 사용자가 작성한 답안을 표시하여 정답 여부를 확인할 수 있도록 해 주고, 각 문항의 배점을 계산하여 100분률로 표시한다. Paper 테이블에 설정된 CutLine 점수와 비교하여 Pass 여부를 표시한다.

그림 11은 평가한 문제지에 대한 풀이과정 화면이다. 각 문제를 Web Page를 이용하여 보여주고, 문제에 대한 해설을 요구하는 경우에는 문제 출제자가 문제 출제시 작성한 해설부분을 사용자에게 보여준다. 물론 이러한 평가 결과 및 문제

평가 일시		2000.11.25		
과 목		전산		
문제지 ID		1		
문제 번호	정답	사용자 답	정답여부	배점
1	3	3	○	5
2	2	2	○	5
3	1	1	○	5
4	4	4	○	5
5	2	2	○	5
6	3	2	×	5
7	1	1	○	5
8	1	2	×	5
9	3	3	○	5
10	4	4	○	5
맞은 갯수	획득 점수	백분율	CutLine	Pass 여부
8 / 10	40 / 50	80	35	P

그림. 10 문제지 평가 결과 화면

평가 일시		2000.11.25	
과 목		전산	
문제지 ID		1	
문제 1 대한민국과 'Korea'를 표현하기 위해 앞으로 하는 바이트 수는? <b>해설</b>			
1. 4.5			
2. 4.10			
3. 8.10			
V 4. 8.5			
문제 2 데이터 베이스를 운영하도록 구성된 소프트웨어를 무엇이라고 하는가? <b>해설</b>			
1. DBL			
2. DDL			
V 3. DBMS			
4. DBA			

그림. 11 문제 풀이 (해설)

풀이는 VoiceXML 문서로 재작성 되어 전화를 통하여 서비스된다.

그림 12는 사용자의 과목별 통계 처리를 나타낸다. 선택한 과목별로 사용자의 평가 내역을 그래프로 표시해 주고, 해당 과목에서 사용자의 최고 점수와 최저 점수, 평균을 표시해 준다. 또, 평가시스템의 다른 사용자와 상대적 비교를 할 수 있도록 과목의 최고 점수와 그 과목의 평가를 응했던 모든 사용자들의 평균을 함께 표시해 준다.

#### 4. 결론 및 향후 과제

본 논문에서는 VoiceXML을 이용하여 음성 기반의 평가 시스템을 설계 및 구현하였다. 평가

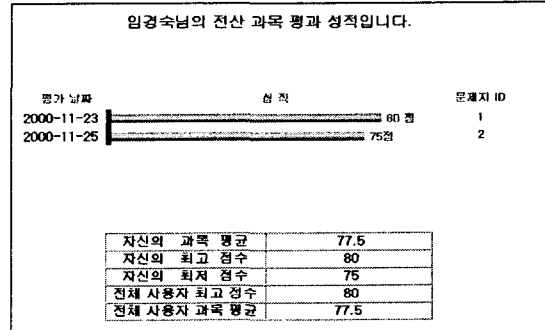


그림. 12 평가 결과 통계 처리

시스템은 PC상에서의 평가뿐만이 아니라 유무선 전화의 대중성과 사용법이 간편하다는 점을 고려하여, Dialogic사의 음성처리보드를 이용하여 전화를 통한 평가가 가능한 전화용 평가 IVR System을 구현하였다. 평가 시스템에서는 학습자 자신의 수준에 적합한 평가를 받을 수 있도록 하였다. 이러한 수준별 학습 평가를 통하여 평가에 대한 자신감을 얻어, 학습에 대한 흥미와 자신감을 갖게 되고 해당 수준에서 최고의 학습 효과를 얻을 수 있다.

VoiceXML 기반의 평가 시스템의 장점은 평가를 위해 가장 대중적인 통신 수단인 유무선 전화를 사용할 수 있어 이동중이거나 컴퓨터를 사용할 수 없는 환경에서도 장소와 시간에 구애 없이 평가가 가능하며, 컴퓨터를 잘 사용하지 못하는 사람, 인터넷 사용이 불가능한 사람들도 평가를 받을 수 있다. 또, VoiceXML이 음성 입출력에 대한 표준을 제시하고 있으므로 음성 입출력의 기술적인 지식이 없이도 신속하게 개발이 가능하며, 기존의 Text기반의 평가시스템을 쉽게 VoiceXML 기반 평가 시스템으로 적용할 수 있다는 것이다.

향후 연구 과제로는 XSL(eXtensible Stylesheet Language)를 이용하여 VoiceXML 문서로 작성된 평가서가 음성 기반만의 평가 시스템이 아닌, 음성과 함께 Web상에서 시각적으로도 문제를 제

시할 수 있는 모듈의 개발이 필요하다. 또, 데이터 마이닝 기술을 이용하여 사용자의 평가 결과와 문제를 분석하여 사용자에게 가장 적합한 차기 평가 방법을 제시하는 기술이 개발되어야 한다.

## 참 고 문 헌

- [1] 이석호, "인터넷 환경에서 가상 학습평가 시스템 설계 및 구현", 멀티미디어학회 논문지 제 1권 2호, pp.204-214, 1998.12.
- [2] Bob Edgar, The VoiceXML Handbook: Understanding and Building the Phone-Enabled Web, CMP Books, 2001.
- [3] VoiceXML Forum, <http://www.voicexml.org>.
- [4] Voice Browser Activity, <http://www.w3c.org/voice>.
- [5] VoiceXML Version 2.0, <http://www.w3.org/TR/2001/WD-voicexml20-20011023>, W3C Working Draft, 2001.10.
- [6] William A. Yarberry, Jr., Computer Telephony Integration, CRC press, 2000.
- [7] 짝기.컴, <http://jjikki.com>.
- [8] 서강대학교 멀티넷, <http://multinet.sogang.ac.kr>.
- [9] 에듀넷 어학평가시스템, <http://www.edunet4u.net>.
- [10] Yes English, <http://www.youngsang.co.kr>, 영상정보통신.
- [11] 인터넷 교육평가시스템, <http://www.highdream.co.kr>.
- [12] 김영복, "CTI 콜센터를 이용한 마케팅 전략에 관한 연구," 홍익대학교 국제경영대학원 석사논문, 1998.
- [13] 최성순, "음성 기술의 현주소", 마이크로소프트 9월호. 2000.9.
- [14] 김경란, 홍기형, "VXML 편집기와 음성 브라우저의 설계 및 구현," 한국정보과학회 2000년 춘계학술발표 논문집, 제 27권 1호, pp414-416,

- 2000.4.
- [15] 김경란, 홍기형, "다중 사용자 지원 VoiceXML 서버 설계 및 구현," 한국음향학회 2001년 하계 학술발표 논문집, 제 20권 1(S)호, pp175-178, 2001.7.



이 인 속

- 1995년 성신여자대학교 전자계산학과 학사
- 2001년 성신여자대학교 전자계산교육학 석사
- 1995년 3월 ~ 2000년 8월 FKM(주) API 개발실 주임연구원
- 2000년 9월 ~ 현재 미디어포드(주) 선임연구원
- 관심분야 : 음성 응용 시스템, CTI, VoIP



홍 기 형

- 1985년 서울대학교 공과대학 전자계산기공학과 학사
- 1987년 한국과학기술원 전산학과 석사
- 1994년 한국과학기술원 전산학과 박사
- 1991년 ~ 1994년 (주)컴퓨터응용기술 개발실장
- 1994년 ~ 1998년 한국전자통신연구원 선임연구원
- 1998년 ~ 현재 성신여자대학교 컴퓨터정보학부 조교수
- 관심분야 : VoiceXML, XML 스키마, 음성 미들웨어, 데이터베이스 시스템