

VoiceXML을 이용한 운세정보 서비스 시스템 구현

최재원, 김동규, 노용완, 홍광석
성균관대학교 정보통신공학부

An Implementation of Fortune Information Service System using VoiceXML

Jae-Won Choi, Dong-Gyu Kim, Yong-Wan Roh, Kwang-seok Hong
School of Information and Communication Engineering, Sungkyunkwan University

요약

운세정보 서비스는 전화로 일반 사용자들이 많이 이용하고 있지만 사용자들이 이용하기에 너무 비싸고 서비스 제공자가 항상 대기해야한다는 단점이 있다. 이점을 보완하는 웹서비스가 있지만 언제 어디서나 이 서비스를 활용할 수는 없다. 그래서 VoiceXML을 이용한 운세정보서비스를 구현하였다. VoiceXML은 음성을 웹 페이지처럼 작성할 수 있는 언어이고 웹과도 호환성이 좋아 기존 웹 서비스의 기능을 모두 지원하면서 독립적으로 음성 인터페이스를 제공해 전화와 웹 서비스의 장점을 모두 취할 수 있다. 본 연구에서는 웹 서비스중 오늘의 운세, 궁합, 바이오리듬, 타로카드점을 VoiceXML을 이용하여 구현했다. 12인의 피 실험자가 각 메뉴를 모두 2번씩 실험한 결과 95%이상의 동작 성공률을 보였다. 그리고 실험자에 대한 설문 조사 결과 기존 웹 서비스에 비해 범용성과 경제성 면에서 장점을 갖는다는 결과가 나왔다. 앞으로 입력부분에 대화형 서비스 메뉴를 보완하여 더욱 사용이 편리한 서비스로 개선할 계획이다.

I. 서론

사람의 음성신호를 인식하는 기술은 크게 음성인식과 화자인식으로 나뉜다. 음성인식은 인간의 음성을 기계나 컴퓨터에서 자동으로 인식하는 것으로 좁은 의미로는 음성 신호로부터 발생된 문자열을 인식하고 출력할 때까지를 의미하고, 넓은 의미로는 말하는 사람(화자)의 인식과 자연어 처리를 포함한 음성이해와 의미추출까지를 의미한다. 이러한 음성인식 시스템은 특정 화자를 인식하는 '화자종속'시스템과 화자와는 상관없이 이루어지는 '화자 독립' 시스템으로 다시 분류된다.

음성인식 관련 표준은 VoiceXML, CCXML(Call Control XML), SALT(Speech Application Language Tags)등이 있다. 이중 VoiceXML은 XML에 기반을 둔 Markup language로서, HTML이 graphical web page를 만드는 데 사용되는 것처럼 VoiceXML은 spoken dialog를 정의하는데 사용되는 언어이다. XML(eXtensible Markup Language)은 W3C(World Wide Web Consortium)가

HTML의 한계를 극복하기 위해 만든 마크업 언어이다. VoiceXML은 이러한 XML을 이용한 특화된 마크업 언어의 일부이다[1].

AT&T, IBM, Lucent Technologies, Motorola 등이 설립한 VoiceXML 포럼에서는 1999년 8월 음성 응용 프로그램 개발을 위해 VoiceXML(Voice eXtensible Markup Language) 0.9 버전을 발표하였고, 2000년 3월 이를 보완한 1.0 버전을 정식으로 제안하였다. W3C에서는 이 제안을 받아들여 2000년 5월 웹의 대화형 마크업 언어로서 VoiceXML을 표준으로 공인하였다. 현재 1.0 버전의 문법 부분을 보완한 VoiceXML 2.0버전이 나와 있다. 적용 가능 분야는 UMS(Unified Messaging System), ASR(Automatic Speech Reponse)같은 음성 관련 서비스 분야에 사용될 수 있다.

VoiceXML은 대화형 음성 서비스의 인터넷 기반 개발을 위한 문서형 인터페이스로 XMLspec 1.0에 기초하여 다이얼로그, 문법, 오디오 출력, 오디오 입력, 콜 제어, 흐름제어 등에 관련된 엘리먼트들로 구성된다[2].

II. VoiceXML 의 구조

그림1은 VoiceXML의 구조를 보여주고 있다. 여기서 Document Server는 웹서버를 말하며 HTTP 클라이언트 응용 프로그램에 해당하는 VoiceXML Interpreter가 VoiceXML Interpreter Context를 통하여 요청하는 것을 처리한다. Document Server는 VoiceXML Interpreter의 요청에 따라 적절한 처리 과정을 거쳐서 (VoiceXML) Document를 만들고 그것을 HTTP 요청에 의한 응답으로 VoiceXML Interpreter에 전달한다[3].

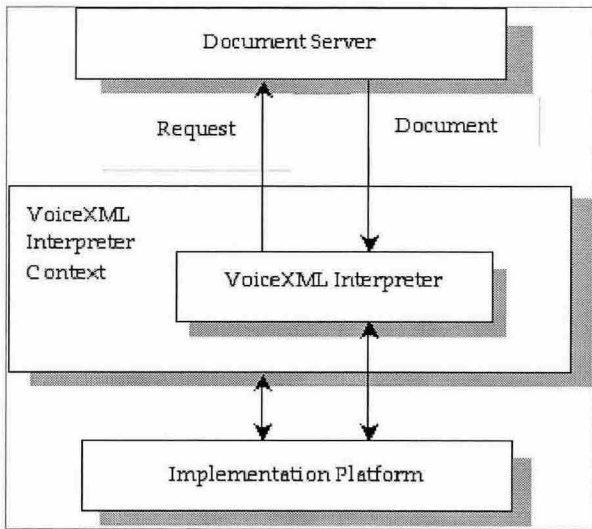


그림 1. VoiceXML 구조

- ◆VoiceXML Interpreter : 서버가 보낸 Document를 해석하고사용자와의 상호작용(interaction)을 수행하기 위해 implementation platform을 제어하는 컴퓨터 프로그램을 말한다.
- ◆VoiceXML Interpreter Context : 웹서버(document server)와 HTTP를 통해 데이터를 주고받는 HTTP통신을 담당하며, VoiceXML interpreter가 VoiceXML문서를 해석하게 하고, VoiceXML Interpreter와는 독립적으로 Implementation Platform과 상호작용을 한다.
- ◆Implementation Platform : H/W와 S/W를 모두 포함하며, 전화 수신기능, 전화상호전환(call transfer)기능, 음성인식기능, 음성합성기능, 음성과 오디오 재생기능, 음성과 오디오 녹음 기능 등을 수행한다[4][5][6].

III. 운세정보 시스템 구조

본 논문에서는 오늘의 운세, 궁합, 바이오피드, 타로 카드 점의 4가지 메뉴로 운세정보 서비스를 구현하였다. 처음에 사용자는 인사말과 함께 오늘의 운세, 궁합보기, 바이오피드 보기, 타로 카드 점 보기의 4가지 메뉴의 선택을

택을 강요받아 각 메뉴로 이동한다.

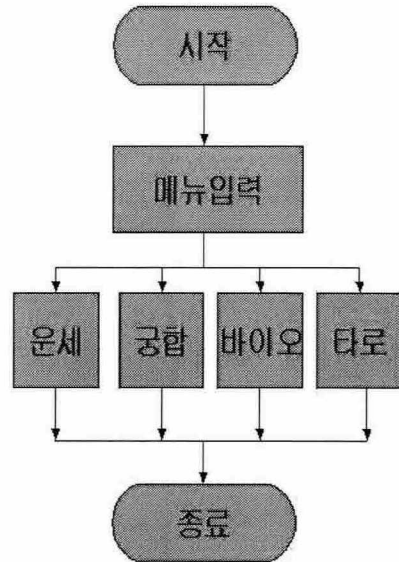


그림 2. 메뉴 구성

3.1 오늘의운세

오늘의 운세는 날마다 결과가 바뀌어야 한다. 본 논문에서는 오늘 날짜와 사용자의 생년월일로 DB에서 결과를 불러오도록 설계하였다. 구성은 아래 그림 3과 같이 처음에 주민등록번호를 입력 받은 후, 생년월일의 각 2자리 숫자와 오늘날짜(년, 월, 일) 각 2자리 숫자를 모두 더해 50으로 나눈 결과를 검색하는 방식으로 구성되어 있다.

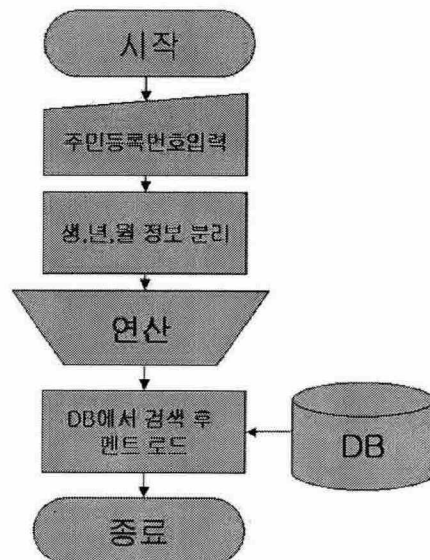


그림 3. 오늘의 운세

3.2 궁합

궁합은 2개의 주민등록번호를 입력받아 각각의 생년월

일 정보로 DB에서 일치하는 결과를 불러오도록 설계했다. 구성은 아래 그림 4와 같이 사용자의 번호를 입력한 후, 궁합을 볼 상대의 번호를 입력받도록 설계되었다.

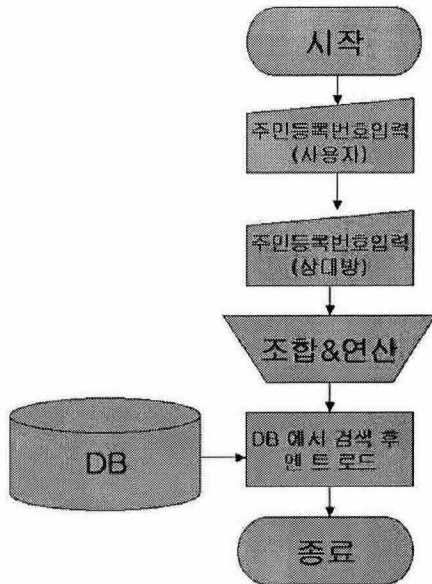


그림 4. 궁합

3.3 바이오리듬

바이오리듬은 오늘의 운세와 비슷하게 구성되며 오늘의 운세의 DB와 같은 위치에 있는 숫자와 오늘 날짜, 생일을 조합해서 지성리듬, 감성리듬, 신체리듬을 계산한다. 구성은 아래와 같다.

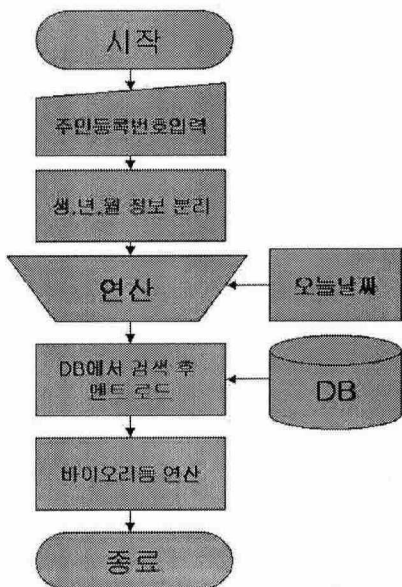


그림 5. 바이오 리듬

3.4 타로카드점

타로 카드는 서로 다른 2장의 카드로 구성되어 있으며, 사용자는 이중 각각 하나씩의 카드를 뽑아 두 카드의 조

합을 이용해 보는 점이다.

카드 종류는 메이저카드와 마이너 카드가 있는데, 메이저 카드는 전체의 운세를, 마이너 카드는 메이저 카드에 추가되는 사항이다.

매번 뽑을때마다 다른 카드가 나와야 하므로 서비스 이용시마다 다른 결과가 나와야 한다.

본 서비스에서는 1~9까지 숫자와 시각을 이용해 뽑는 행위를 대신했다.

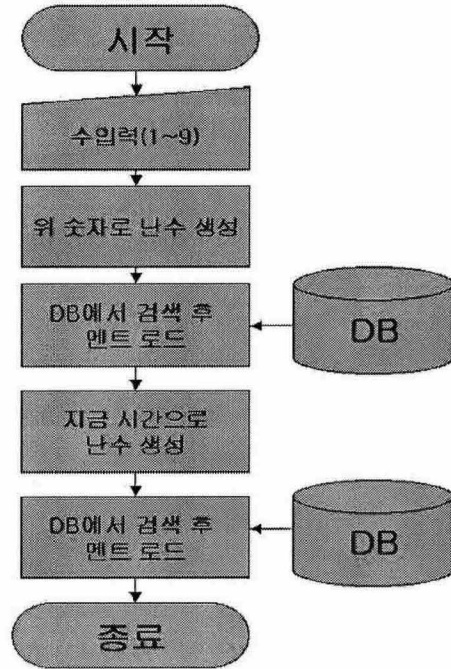


그림 6. 타로 카드 점

3.5 개발 환경

시스템의 개발 환경은 윈도우 2000 professional을 사용하였고, 웹 서버는 윈도우 NT전용 web 서버인 IIS(Internet Information Server) 5.0을 사용하였다. 저장된 운세정보를 불러오기 위해 데이터베이스로서 ASP(Active Server Page)를 이용하여 시스템을 구축하고, 사용자에게 발송할 목록을 DB에 저장하기 위해 Microsoft Access를 사용하였다. Call 제어를 위해서 Intel Dialogic 41JCT/LS를 사용하였고, Interpreter, 음성 인식기와 합성기는 KT의 HUVOICE 1.0을 사용하였다. Intel(R) Celeron(R) CPU 1.70GHZ 2G RAM 컴퓨터를 사용했다.

IV. 실험결과

4.1 실험 환경

연구실에 있는 VoiceXML서버에 접속하기 위해 일반 유선 전화기 및 휴대폰을 이용했으며, CTI보드는 인텔

다이얼로그 4채널모드를 사용하였다. 실험 중 VoiceXML 서버는 Wondow2000 professional을 운영체제로 동작하였다.

4.2 실험방법

본 시스템은 평가자 12인에게 오늘의 운세, 궁합, 바이오리듬, 타로 카드 점의 4가지 메뉴를 순서대로 한번씩 실행해 보는 방식으로 실험했으며, 각각 2번씩 평가한 결과를 기술했다.

평가자는 본 학부의 재학 인원 중 일부를 선발하였으며 평가의 공정성을 위해 평가가 끝날 때 까지 서로간의 의사소통을 금지하였다.

4.3 실험결과

4개의 서브시스템의 각 분기 성공률은 아래 표와 같다.

표 1. 각 파트별 성공률

	메뉴(%)	입력(%)	출력(%)
오늘의 운세	91	100	100
궁합	91	100	100
바이오리듬	100	100	100
타로	100	100	100

실험자 12인 모두 본 시스템으로부터 응답을 받았으며 오류는 한 차례도 나오지 않았다.

단, 처음부분 메뉴 입력 시(운세|궁합|바이오|타로) '운세'와 '궁합'을 못 인식해 사용자가 다시 한번 발언했던 경우가 4차례 있었다.

이것은 사용자의 발음이 부정확했거나, 목소리가 작았거나, 아니면 VoiceXML 자체에서 특정 사용자의 발음을 인식하지 못한 경우로 추정된다.

4.4 평가

VoiceXML을 이용한 운세정보서비스의 실험자에게 설문조사를 한 결과 서비스 이용 편의성이 상-중 이상으로 평가되었다. VoiceXML을 이용한 운세정보서비스가 기존 운세정보시장에서도 충분히 경쟁력 있음을 도출할 수 있다. 또한 피 실험자들은 기존서비스와 비교 시, 본 시스템이 더 나은 이유로 비용을 들었는데, 이는 최적의 경쟁력을 갖추기 위해서는 서비스를 실제 시장에 적용할 때 철저한 경제적 검증이 필요함을 보여준다.

표 2. 사용자 만족도

	운세정보이용횟수	재사용의향	기존서비스와 비교
상			○
중	○	○	
하			

V. 결론

본 논문은 운세정보를 주제로 VoiceXML 문서를 사용해 효과적인 서비스를 구현했다. 그 결과 기존 서비스의 단점이었던 범용성을 개선할 수 있었으며, 추가로 사용자 측면에서는 경제성도 갖출 수 있어 피 실험자들에게 좋은 평가를 받았다.

본 서비스의 메뉴부분과 입력부분에 대화형 알고리즘을 보완한다면 더욱 편리한 시스템 구축이 가능하다. 앞으로는 VoiceXML로 대화형 알고리즘을 개발해 운세정보 서비스에 적용시킬 계획이다.

또한 웹 상의 여러 가지 다른 관련 서비스들도 VoiceXML 문서로 컨버전하여 본 서비스에 추가해 더욱 경쟁력 있는 서비스를 구축할 계획이다.

감사의 글

본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구 결과로 수행되었음 (IITA-2005-C1090-0501-0019).

참고 문헌

- [1] VoiceXML 포럼 <http://www.voicexml.org>
- [2] Eve Astrid 외, early adopter [VoiceXML]
- [3] 홍용택, "XML과 VoiceXML의 효율적인 통합 방안 제시", 군산대학교 교육대학원 석사논문, 2001
- [4] 홍광석, "음성인터페이스 위한 VXML 기술", 한국인터넷정보학회지, 제2권, 1호, 2001.
- [5] 이인숙, 홍기형, "유무선 전화 음성 기반 VoiceXML 학습 평가 시스템," 한국멀티미디어학회지, vol.005, no.004, 2001.
- [6] Quiane Ruiz, J.A.; Manjarrez Sanchez, J.R, " Design of a VoiceXML gateway," Computer Science, 2003. ENC 2003. Proceedings of the Fourth Mexican International Conference on, 49 53, pp. 8-12 Sept. 2003.