

# 단어추출을 기반으로 한 음성 대화처리 시스템

송창환<sup>○</sup> 유하진 오영환  
한국과학기술원 전산학과

## Spoken Dialogue Management System based on Word Spotting

Chang-Hwan Song<sup>○</sup> Ha-Jin Yu Yung-Hwan Oh  
Department of Computer Science  
Korea Advanced Institute of Science and Technology

### 요 약

본 연구에서는 인간과 컴퓨터 사이의 음성을 이용한 대화 시스템을 구현하였다. 특별히 음성을 인식하는데 있어서 단어추출(word spotting) 방법을 사용하는 경우에 알맞은 의미 분석 방법과 도표 형태의 규칙을 기반으로 하여 시스템의 응답을 생성하는 방법에 대하여 연구하였다. 단어추출 방법을 사용하여 음성을 인식하는 경우에는 형태소분석 및 구문분석의 과정을 이용하여 사용자의 발화 의도를 분석하기 어려우므로 새로운 의미 분석 방법을 필요로 한다. 본 연구에서는 퍼지 관계를 사용하여 사용자의 발화 의도를 파악하는 새로운 의미 분석 방법을 제안하였다. 그리고, 사용자의 발화 의도에 적절한 시스템의 응답을 만들고 응답의 내용을 효율적으로 관리하기 위한 방법으로 현재의 상태와 사용자의 의도에 따른 응답 규칙을 만들었다. 이 규칙은 도표의 형태로 구현되어 규칙의 갱신 및 확장을 편리하게 만들었다. 대화의 영역은 열차 예매에 관련된 예매, 취소, 문의 및 관광지 안내로 제한하였다. 음성의 오인식에 의한 오류에 적절히 대처하기 위해 시스템의 응답은 확인 및 수정 과정을 포함하고 있다. 본 시스템을 문자 입력과 음성 입력으로 각각 실험한 결과, 사용자는 시스템의 도움을 받아 자신이 의도하는 목적을 달성할 수 있었다.

### 1. 서론

인간과 기계 사이에 음성을 통하여 의사를 교환하고자 하는 연구가 계속되어 왔으며, 이를 위한 보다 진보된 기술의 필요성이 계속해서 증가되고 있다. 특별히 서로간의 주고 받는 대화는 정보들 한편에서 다른 한편으로 일방적으로 전달하는 경우보다 더욱 확실한 정보를 전달할 수 있으므로 인간과 기계 사이의 대화 시스템을 구현하고자 하는 노력들이 계속되어 왔다.

음성 대화 시스템은 단순한 받아쓰기 시스템(dictation system)과는 많은 차이가 있다. 받아쓰기 시스템은 화자의 발화 내용을 그대로 인식하는 것이 목적이지만, 대화 시스템은 여기에 그치지 않고 자연스러운 음성을 통한 인간과 기계 사이의 의사 교환을 목적으로 하고 있다. 지금까지 화자독립으로 대응할 어휘를 처리할 수 있는 고성능의 음성인식 시스템이 많이 개발되어 왔다. 그러나 이러한 시스템들은 주위의 잡음이 첨가되고 자연 발생음(spontaneous speech)에 대해 처리하게 되면 그 성능이 급격히 감소한다. 그리고, 대부분의 시스템들은 문법적으로 정확하게 발생된 문장만을 해석할 수 있다[1].

뿐만 아니라, 지금까지 음성 인식 기술이 많은 발전을 거듭해 온 것이 사실이지만, 현재의 기술로서는 인식 오류를 피할 수는 없으며, 인식된 결과를 그대로 수용할 수 있는지 아니면 거절을 해야 하는지도 알 수 없다. 따라서 기계가 인식한 결과를 확인하는 과정이 꼭 필요하다[2][3]. 또, 시스템에 따라서는 음성의 인식 오류를 감소시키기 위해 사용자가 응답으로 할 수 있는 단어의 수를 매우 작은 수로 제한하기도 한다. 이 때는 사용자가 제한된 단어 외의 단어를 사용하지 않도록 시스템의 응답 형식을 설계해야 한다[4][5].

그리고, 자연발화(spontaneous utterance)에는 다음과 같이 그 내용의 인식을 어렵게 만드는 특성들이 있다. 그 특성에는 머뭇거림(hesitation), 잘못된 시작점(false starts), 재시작(restarts), 자기수정(self-correction), 비문법적 구조(ungrammatical structures) 등이 있다. 그리고, 발화의 내용을 인식하는데에는 쓸모가 없지만 대화의 구조를 결정하는 데는 유용한 단어들도 있다. 또한, 자연 발화에는 단어를 이루지 못하는 음향학적 절(acoustic phrases)도 있고, 음소 기호로 옮겨지지 않는 음향학적 흐름(acoustic stream)도 있다[2]. 따라서, 현재의 기술로는 자연 발화로부터 정확한 내용을 인식하는 것이 쉽지 않다.

그러나, 지금까지 인간과 기계 사이의 대화에 관련된 연구들은 많은 경우 문자 입력 상태의 문장들을 이용하여 진행되어 왔다. 특히, 국내에서의 대화에 관한 연구는 아직 음성을 입력력으로 사용하는 시스템 상에서 구현된 예를 찾아보기 힘들다.

외국의 경우 [1]에서는 키워드를 추출한 후 키워드 격자 구문분석기(a semantic keyword lattice parser)를 사용하여 의미를 분석하는 방식의 대화관리를 구현하였다. [6]에서는 사용자의 발화에 대해 SR-계획의 정보와 두 스택의 내용을 참조하여 사용자의 다음 발화 내용을 미리 예측함으로써 효과적으로 대처하고 있다. [3],[4],[5]에서는 인식의 오류를 감소시키기 위하여 시스템의 응답 형식을 이용하여 사용자가 사용하는 단어를 제한하도록 유도하고 있다. 또, [7]에서는 언어 의존적인 부분과 다른 부분을 분리하여 여러 언어로 사용 가능한 안내 시스템을 개발하고 있다.

본 연구에서 구현된 음성 대화처리 시스템은 음성인식부에서 단어추출(word spotting) 방법을 사용하였으며, 단어와 의도 사이의 퍼지 관계를 이용하여 사용자의 의도를 분석하고 도표 형식의 응답 규칙을 사용하여 시스템의 응답을 만들었다.

## 2. 시스템 개요

구성하고자 하는 전체 시스템의 개요는 그림 1과 같다.

사용자의 발화에 의해 음성 데이터가 입력되면 이것으로부터 음성 인식기에 의한 키워드 추출이 일어난다. 단어 추출에 의해 인식된 키워드에 대한 정보들은 대화관리부로 옮겨진다. 이 때 대화관리부가 넘겨 받는 키워드에 대한 정보는 키워드의 내용, 시작점과 끝점, 그리고 인식 정도에 대한 점수이다.

대화관리부에서는 각종 데이터베이스를 이용하여 적절한 응답을 문장으로 만들어 낸다. 이 때 사용되어지는 데이터베이스로는 단어 사전, 열차시간표와 남은 표, 가격, 소요 시간 등에 관한 데이터베이스, 대화처리의 규칙(열차예매 영역에서의 규칙, 관광지 안내 영역에서의 규칙), 각 단어와 사용자의 의도간의 퍼지 관계를 나타내는 상관 관계표가 있다.

이 시스템에서 제공하고 있는 서비스는 크게 열차예매에 관한 것과 관광지 안내에 관한 것으로 나눌 수 있다. 열차예매의 경우는 당일 열차 예매에 관한 것으로 예매, 취소, 문의, 변경의 기능을 제공해 준다. 또, 관광지 안내의 경우는 주로 문의에 관한 것으로 관광지가 있는 지역의 역명과 역 근방의 관광지들 이름, 관광에 걸리는 시간 및 관광지에 대한 개략적인 소개를 제공해 준다.

키워드에 대한 인식 결과를 이용하여 사용자의 발화 의도를 분석하는데에는 각 키워드와 사용자 의도 사이의 퍼지 관계가 사용된다. 사용자의 의도에 대한 분석은 퍼지 관계를 사용하여 먼저 대화의 영역을 결정된 후에 결정된 대화 영역에서의 사용자 의도를 점차 세부적으로 분석하게 된다.

대화관리부에서는 사용자의 의도에 합당한 서비스를 제공한 후 그 결과를 보고하거나 대화 도중에 일어나는 정보의 제공 또는 추가 정보를 요구하는 등의 기능을 제공해 준다. 대화관리부의 응답은 대화 처리의 규칙을 따른다. 이 규칙들은 도표(table)와 형식으로 만들어져 있으며 대화관리부는 이 규칙들을 참조하여 자신이 취해야 하는 행동을 결정하게 된다. 이러한 기능들은 결과적으로 사용자의 말에 대하여 적절하게 대처하는 응답을 문장으로 만들어 내게 된다. 대화관리부에서 문장으로 만들어진 응답은 음성합성기를 통하여 음성으로 합성되어 사용자에게 제공된다.

## 3. 퍼지 관계를 이용한 의도 분석

사용자의 요구가 문자열로 된 문장으로 입력된 경우는 보통 구문 분석(syntactic analysis)의 과정을 거쳐 의미분석(semantic analysis)을 함으로써 사용자의 의도를 파악한다. 그러나, 음성인식기는 대화의 내용을 문자 입력과 같이 정확한 문장으로 바꾸어 주지 못한다. 또, 대화에 사용되는 언어는 문법적으로 완전하지 못할 때가 많다. 따라서, 구문분석을 거쳐 의미분석을 하기가 용이하지 않다.

또, 음성인식기가 한정된 단어만을 키워드로 가지고 있으면서 단어 추출 방식을 이용하여 사용자의 발화 내용을 처리하는 경우에는 문장 전체를 처리하는 경우와는 다른 형태로 의미를 파악해야 한다. 여기서는 이러한 방식에 적용할 수 있는 방법으로 퍼지 관계(fuzzy relation)를 사용하여 사용자의 의도를 파악하는 방법을 사용하였다.

인식기에 의하여 인식할 수 있는 단어들 중에는 사용자의 의도를 파악하는데 사용할 수 있는 단어들이 있다. 예를 들어, 화자의 말 속에 '예매'라는 단어가 나오면, 화자는 예매를 하고자 하는 가능성이 크다. 그러나, "예매를 취소하고 싶습니다.", "예매를 변경하고 싶습니다."와 같이 그 의도가 취소나 변경일 가능성도 있다. 또, 그보다는 빈도가 작지만 "예매 됩니까?"와 같이 사용자의 의도가 예매의 가능성을 확인하는 문의일 수도 있다. 따라서, '예매'와 화자의 의도

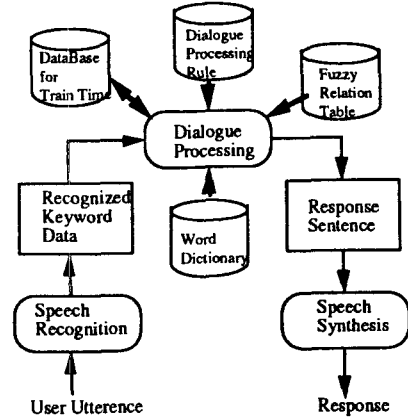


그림 1. 전체 시스템의 개요

와의 상관 관계 정도는 예매에 대한 값이 가장 크고 취소와 변경에 대한 값은 그 다음이고 문의에 대한 값이 그 다음이 되어야 한다. 본 연구에서는 예매에 0.8, 문의에 0.4, 취소에 0.5, 변경에 0.5의 값을 부여하였다. 이와 같은 방법으로 각 단어가 가지고 있는 의도의 값이 결정된다.

사용자의 의도에 대한 결정은 아래의 수식과 같이 인식된 각 단어의 정확도에 그 단어가 가지고 있는 의도의 값을 곱한 값들의 합으로 결정된다.

$$TotalInten = \sum_{i=1}^N RcgScore(word[i]) \cdot IntenScore(word[i])$$

이와 같이 네 가지 의도에 대한 총합을 구하여 그 값이 가장 큰 의도를 사용자의 의도로 간주한다. 이와 같은 방법으로 사용자의 의도를 결정하는 것은 사용자의 제한없는 자연 발화에 대해 적절하게 대처할 수 있다. 아래는 퍼지의 개념을 이용한 결과 사용자의 의도에 적절히 대처한 예이다.

사용자:	“서울행 음... 언제였더라... 오후 2시로 예매해 주세요.”
인식결과:	(서울 .856) (언제 .725) (오후 .912) (2시 .920) (예매 .895) (주세요 .923)
의도점수:	예매: 0.5033 / 문의: 0.3742 / 취소: 0.3513 / 변경: 0.3513
응답:	“서울행 통일호 오후 2시 1장 예매하십니까?”

위의 예에서 사용자는 말을 하는 도중 시스템에게 ‘음... 언제였더라...’라는 불필요한 말을 함으로써 시스템으로 하여금 화자의 의도를 파악하는데 어려움을 가져다 주었다. 반면, 사용자의 말 속에 ‘예매’와 ‘주세요’라는 단어가 있음으로 해서 화자의 의도를 알 수 있게 해 준다. 인식기에 의한 인식의 결과와 단어와 의도 사이의 퍼지 관계를 사용하여 계산한 결과 ‘예매’를 나타내는 의도의 점수가 가장 높은 것을 볼 수 있다. 이와 같이 퍼지 개념을 사용함으로써 시스템은 비문법적이고 불필요한 정보들이 섞여 있는 경우에 대해서도 화자의 의도를 파악할 수 있다.

앞에서 이미 언급한 것과 같이 본 시스템에서 대화 영역은 크게 열차 예매와 관광지 안내로 나누어진다. 열차 예매 영역에서 사용자의 의도는 크게 예매, 문의, 취소, 변경으로 구분한다. 그리고, 문의로 파악이 된 경우에는 또 다시 가능성, 차종, 시간, 소요시간, 요금으로 세분하여 구분한다. 또, 관광지 안내 영역에서 사용자의 의도는 역명, 관광지명, 관광지 내용, 소요 시간, 가능성으로 구분한다.

따라서 시스템은 대화의 영역(열차 예매, 관광지 안내)과 열차 예매 영역에서의 대분류(예매, 문의, 취소, 변경)와 문의상 분류(가능

성, 차종, 시간, 소요시간, 요금), 그리고 관광지 안내 영역에서의 의도(역명, 관광지명, 관광지 내용, 소요시간, 가능성)에 대한 퍼지관계를 나타내는 표를 가지고 있게 된다.

다음은 이러한 퍼지 관계를 나타내는 표 중 일부를 보인 것이다. 영역의 분류에서

	얼차예약	관광지안내
예대	1	0
어디	0.3	0.7

얼차 예약의 대부분류에서

	예대	문의	취소	변경
예대	0.7	0.4	0.5	0.5
얼마	0.2	0.8	0.2	0.2

얼차 예약에 관한 문의상의 분류에서

	가능성	차종	시간	소요시간	요금
얼마	0.1	0.1	0.1	0.3	0.9
있습니까	0.65	0.6	0.6	0.1	0.1

관광지 안내시의 의도 분류에서

	역명	관광지명	관광지내용	소요시간	가능성
다른	0	0.7	0.3	0	0.2
알고	0	0.3	0.7	0.3	0.2

현재 대화관리부에서는 156개의 단어를 키워드로 사용하고 있다. 이 중 98개는 항목에 관련된 단어들이고 그 외는 사용자의 의도를 분석하기 위한 퍼지 관계에 사용된 단어들이다. 이들 키워드들은 서로 중복되기도 한다. 표 1은 사용된 키워드의 수를 나타낸 것이다.

표 1. 키워드의 수

항목 분류	의도의 분류
역명	6
차종	3
정오	6
시간	13
시각	13
매수	16
확인	13
시간류	8
관광지	20

#### 4. 상태코드를 이용한 응답 규칙

사용자의 발화 의도에 적절한 시스템의 응답을 만들 뿐 아니라, 응답의 기능을 확장하고 변경하기에 용이한 응답 규칙을 만드는 것은 시스템의 성능 향상을 위해서 매우 중요하다. 이를 위하여 본 연구에서는 도표 형태의 규칙을 만들었다. 대화관리부에서 사용자 의도에 대한 분석이 끝나면, 현재의 입력에 포함되어 있는 각 항목들에 대한 데이터와 이전의 대화 속에 포함되어 있었던 의도 및 항목에 대한 기록을 이용하여 현재의 상태 코드(condition code)를 만든다. 그리고 시스템의 응답 규칙을 참조하여 현재의 상태 코드 상에서 수행해야 할 루틴들을 찾아내어 수행하게 된다. 이 응답 규칙들은 도표의 형태로 구현되어 있으므로 도표의 내용만을 바꾸어줌으로써 시스템의 응답을 다르게 할 수 있다. 따라서 규칙의 갱신 및 확장이 용이하다.

그림 2에서와 같이 얼차예매 영역은 요구(request), 확인(confirm), 실행(execution)의 3가지 상태(state)를 가지고 있다. 음성인식기를 통하여 각 키워드들에 대해 인식된 결과는 요구 상태와 확인 상태에서 참조된다. 그리고 시스템의 응답은 요구, 확인, 실행 모두에서 문장의 형태로 만들어지며 이 결과가 음성합성

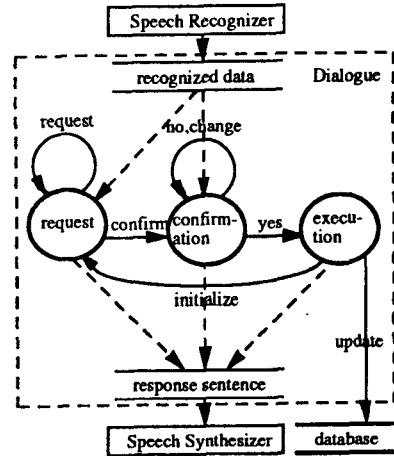


그림 2. 얼차예매를 위한 대화처리의 과정

기로 넘겨져서 음성으로 바뀌게 된다. 또, 각 얼차의 남은 차표의 수에 대한 데이터베이스는 요구와 확인 상태에서는 참조가 일어나고 실행 상태에서는 차표의 수를 갱신하는 작업이 일어난다.

사용자와의 대화가 시작되면 시스템은 요구 상태에서 시작하게 된다. 그리고 사용자가 요구하는 행위(예대, 취소, 변경)를 하는데 꼭 필요로 하는 정보가 다 채워질 때까지 시스템은 사용자에게 도움을 줄 수 있는 정보를 제공해 주거나 채워지지 않은 필수 정보에 대한 요구를 하게 된다. 필수적인 정보가 다 채워지면 시스템은 사용자에게 현재까지 인식된 결과가 사용자의 요구와 일치하는지 사용자에게 확인해 줄 것을 요구하는 확인 상태로 바뀐다.

확인 상태에서 사용자에 의해 요구 사항이 확인되면 실행 상태로 넘어간다. 그렇지 않고 사용자에 의해 새로운 정보가 추가되거나 변경되면 변경된 사항을 가지고 사용자에게 다시 확인을 요구한다. 그 외에도 대화 중간에 있을 수 있는 사용자의 문의 행위에 대해서는 상태의 천이(state transition) 없이 그것에 관련된 정보를 제공해 준다.

얼차예매 영역에서의 상태코드를 이용한 응답 규칙은 이러한 상태의 천이와 관련된 시스템의 응답을 규칙으로 만든 것이다. 얼차예매와 관련된 상태 코드는 10개의 변수로 되어 있다. 이 중 5개는 항목의 유무에 관한 것이고, 시간의 종류 구분에 1개, 사용자의 의도에 2개, 실행 가능한 유무에 1개, 사용자에 의한 확인의 유무에 1개가 매정되어 있다. 표 2는 상태코드의 매정에 관한 표이다. 항목의 유무에서 0은 지금까지 관련 정보가 입력되지 않았음을 의미하고, 1은

표 2. 얼차예매 영역에서의 상태코드

목적지	항목의 유무				가능	확인
	차종	정오	시간	매수		
0	0	-	0	-	0	0
1	1	1	1	1	1	1

의도					무시 (don't care)
예대	문의	취소	변경		
1	2	3	4		-1

문의류				
가능	차종	시간	소요시간	요금
1	2	3	4	5

시간류		
정시	부근시	구간시
1	2	3

입력된 자료가 있음을 의미한다. 정오와 매수는 처음부터 초기값이 주어지므로 0을 가질 수 없다. 상태코드를 보고 적용할 대화처리 규칙을 찾을 때, 코드의 값이 -1 이면 무시하게 된다.

사용자의 발화에 대한 분석이 끝나면 표 2를 이용한 상태 코드가 만들어 진다. 상태 코드에 따른 시스템의 응답 규칙은 도표 형식으로 만들어져 있으므로, 현재의 상태코드가 완성되면 해당 규칙을 찾아 관련 함수를 실행하게 된다. 다음은 이러한 대화 규칙의 몇가지 예를 보인 것이다.

목차정시 메시의 문가 확 적 지 중 오 간 수 류 도 류 능 인	실 함 수 값 들
1 0-1 1-1 1 1-1-1-1	안내(차중)→확인(예매).
1 1-1 0-1-1 1-1-1-1	안내(시간).
1 1-1 1-1 1 1-1 0-1	안내(차중 시간)→확인(예매).
1 0-1 1-1 2 1-1-1-1	안내(차중,시간).
1 0-1 1-1 3 1-1-1-1	안내(차중,시간).

여기에서 숫자로 된 부분은 상태 코드에 해당하고 우측 하단은 이러한 상태 코드를 만족할 때 실행하게 되는 함수들이다. 실행할 함수는 두 단어 이상으로 되어 있는데 앞의 단어는 함수에 해당하고 괄호 안은 그 함수의 파라미터에 해당한다. 위에서 '안내(차중)'의 경우 안내는 함수의 이름이 되고 차중은 파라미터가 된다. 또, 실행하게 되는 함수가 둘 이상인 경우는 '-'로 연결되어 있다. 예를 들어 '안내(차중)→확인(예매)' 는 차중을 파라미터로 하는 안내 함수가 실행된 후 예매를 파라미터로 하는 확인 함수가 실행되는 것을 의미한다. 이와 같이 시스템이 제공해 주는 서비스 및 응답은 이러한 도표를 이용한 규칙들을 기반으로 하여 동작한다.

관광지 안내 영역에서의 시스템 응답 규칙은 열차 예매 영역에서의 응답 규칙과 같은 형식이며, 단지 제공해주는 응답의 내용만 다를 뿐이다. 이 영역에서 제공해 주는 서비스는 관광지가 있는 역의 이름, 특정 역 부근의 관광지들 이름, 특정 관광지에 대한 간단한 소개, 관광하는데 걸리는 예상 소요시간, 시간 내에 관광이 가능한지의 여부이다. 표 3은 관광지 안내 영역에서의 상태코드를 나타낸 것이다. 여기서 항목은 역명, 관광지명, 소요시간의 3개 뿐이며 사용자의 의도는 5가지로 분류된다.

표 3. 관광지 안내 영역에서의 상태코드

항목의 유무		역명	관광지	소요시간	의도	무시
0	1					
0	0	0				
1	1	1		1		
의도		역명	관광지명	관광지내용	소요시간	가능성
1	2					
1	2	3	4	5		-1

## 5. 그 외의 고려사항

키워드의 인식 결과에 모호한 데이터가 있는 경우에는 사용자에게 좀 더 확실한 입력을 요구하게 된다. 가령 목격지도 서울과 부산이 비슷한 정도로 인식이 되어 사용자가 원하는 목적지를 알 수 없는 경우, 시스템은 사용자에게 서울행인지 부산행인지를 묻는 질문을 만들게 된다. 또한, 시스템은 사용자가 의도하는 행위를 도출 수 있도록 각종 안내를 해주어야 한다. 사용자의 요구 내용이 데이터베이스에 있는 내용과 일치하지 않아서 시스템이 처리할 해 줄 수 없는 경우 시스템은 사용자에게 그와 관련된 정보를 제공해 줌으로써 사용자가 원하는 행위를 돕게 된다.

시스템에서 시간에 대한 정보는 간단한 구문분석에 의하여 정확한 시간, 부근 시간, 구간 시간으로 나뉜다. '2시', '3시' 등은 정확한 시간에 해당되고, '2시 경', '3시 부근', '4시 근처' 등은 부근 시간에 해당된다. 또, '2시에서 4시 사이', '1시 부터 3시 까지' 등은 구간 시간에 해당된다.

현재 시스템의 인식 결과가 사용자의 요구와 맞는가를 확인하는 과정은 시스템의 필수적인 요소이다. 그러나 사용자의 편의를 위해서는 확인하는 횟수를 최소로 줄이지 않으면 안된다. 따라서 여기서는 예매나 취소와 같이 시스템이 사용자의 요구를 잘못 인식한 상태에서 그 결과를 처리했을 때 치명적인 오류를 범하게 되는 경우에 한해서만 사용자에게 확인을 요구하게 된다. 사용자의 요구에 대하여 시스템이 해 주는 응답 속에는 인식된 결과를 포함하고 있어서 사용자는 시스템이 응답할 때마다 시스템이 인식한 결과들을 간접적으로 집할 수 있게 되므로 잘못된 인식 결과는 즉시 확인할 수 있다.

사용자는 시스템과의 대화 도중 대화의 영역을 변경할 수 있다. 이렇게 영역이 바뀔 때, 항목들 중 일부는 바뀌기 전의 대화 영역에서 이야기되던 항목을 이어 받는다. 가령 대화의 영역이 열차예매에서 관광지 안내 영역으로 바뀐다면, 열차예매 영역의 목적지를 관광지 안내 영역의 역명으로 이어받는다.

## 6. 실행 예

실험은 문자 입력과 음성 입력으로 각각 행해졌다. 현재 음성 상에서 구현된 시스템은 음성 인식기에 의해 인식 가능한 키워드의 수가 31단어로 한정되어 있다. 다음은 그 실행 예이다.

### 6.1 문자 입력에서의 실행 예

<사용자> 부산행 열차는 2시 경에 어떤 것이 있습니까?  
 상태코드: 1 0 1 1 1 2 2 3 0 0  
 <시스템> 부산행 열차는 2시 경에는 오후 1시 새마을호 오후 3시 통일호가 있습니다.  
 <사용자> 어떤 관광지가 있는 데요?  
 상태코드: 1 0 0 2  
 <시스템> 부산의 관광명소는 용두산공원 해운대 태종대 남천동 앞바다 송정 동래천 뱀어사 자갈치시장이 있습니다.  
 <사용자> 해운대를 구경 하는데 얼마나 걸립니까?  
 상태코드: 1 1 0 4  
 <시스템> 해운대를 관광하는데에는 약 2시간 정도 소요됩니다.  
 <사용자> 예매해 주세요.  
 상태코드: 1 0 1 1 1 2 1 3 0 0  
 <시스템> 부산행 열차는 2시 경에는 오후 1시 새마을호 오후 3시 통일호가 있습니다.  
 <사용자> 1시 새마을호요.  
 상태코드: 1 1 1 1 1 1 1 3 1 0  
 <시스템> 부산행 새마을호 오후 1시 1장 예매하시겠습니까?  
 <사용자> 아뇨 2장입니다.  
 상태코드: 1 1 1 1 1 1 1 3 1 0  
 <시스템> 부산행 새마을호 오후 1시 2장 예매하시겠습니까?  
 <사용자> 예.  
 상태코드: 1 1 1 1 1 1 1 3 1 1  
 <시스템> 부산행 새마을호 오후 1시 2장 예매되었습니다.  
 감사합니다.

### 6.2 음성 입력에서의 실행 예

<사용자> 대전행 통일호 오후 3시 1장 주세요.  
 <입력> [대전| station: 0- 3:1.000000]  
 [오후| pm: 2- 2:1.000000]

[ 통일 | class: 5-10:0.980000 ]  
 [ 9시 | hour:10-16:0.973333 ]  
 [ 10시 | hour:14-18:0.940000 ]  
 [ 3시 | hour:14-18:0.950000 ]  
 [ 1시 | hour:18-20:0.960000 ]

<시스템> 시간을 다시 말씀해 주세요.  
 <사용자> 3시요.  
 <입력> [ 3시 | hour: 3- 4:1.000000 ]  
 [ 10시 | hour: 3- 7:0.940000 ]  
 <시스템> 대전행 통일호 오후 3시 1장 예매원하십니까?  
 <사용자> 아뇨, 2시요.  
 <입력> [ 2시 | hour: 1- 5:0.987500 ]  
 <시스템> 대전행 오후 2시에는 열차가 없습니다. 대전행 열차는 2시 경에는 오후 1시 새마을호 오후 3시 통일호가 있습니다.  
 <사용자> 1시 새마을호요.  
 <입력> [ 1시 | hour: 0- 3:1.000000 ]  
 [ 3시 | hour: 1- 4:0.950000 ]  
 [ 새마을 | class: 5-11:0.966667 ]  
 [ 1장 | num: 7- 8:1.000000 ]  
 [ 오후 | pm:11-12:1.000000 ]  
 <시스템> 대전행 새마을호 오후 1시 1장 예매원하십니까?  
 <사용자> 예.  
 <입력> [ 예 | yes: 0- 6:0.950000 ]  
 <시스템> 대전행 새마을호 오후 1시 1장 예매되었습니다. 감사합니다.

## 7. 결론

본 연구에서는 인간과 컴퓨터 사이에 음성을 이용하여 정보를 교환 및 처리할 수 있는 시스템을 구현하였다. 특별히 사용자의 자연 발성음에 대처하기 위한 방법으로 음성인식부에서 단어 추출 방식을 사용하였다. 이러한 단어 추출 방법을 사용하는 음성 인식 시스템에 대해서는 문자열로 입력된 문장에 대하여 흔히 사용되는 구문 분석기와 의미 분석기를 그대로 사용할 수 없다. 키워드들만을 사용하는 구문 분석기와 의미 분석기를 만드는 경우 구현하는데 많은 어려움이 따르며, 구현된 경우에도 추출 가능한 단어의 집합이 바뀌는 경우 변경된 단어 집합에 적합하도록 구문 분석기와 의미 분석기를 갱신해야 하는 문제가 발생한다. 또, 단순히 키워드의 유무만을 점검하여 사용자의 의도를 분석할 경우에는 서로 다른 의도를 담고 있는 단어들 이 함께 있는 경우 시스템의 대처 능력이 떨어지게 된다. 반면, 본 논문에서 제안한 퍼지 관계를 이용한 의도 분석 방법은 각 키워드에 대한 퍼지 점수만을 부여하면 되므로 구현 방법이 간편할 뿐 아니라 갱신하는데도 큰 어려움이 없다. 따라서 소규모 어휘를 인식하는 단어 추출 시스템의 사용자 의도 분석 방법으로 유용하게 사용되어 질 수 있다.

시스템의 기능 향상 및 확장을 고려해 볼 때, 시스템의 응답은 사용자의 의도에 적합할 뿐만 아니라 그 내용을 갱신하기에도 용이해야 한다. 또, 시스템에 대한 실험 결과 적절치 못한 응답이 있는 경우 손쉽게 응답의 내용을 바꾸어 줄 수 있어야 한다. 따라서, 시스템의 응답 내용을 쉽게 갱신할 수 있는 응답 규칙의 관리 방법이 중요하다. 본 논문에서는 응답 규칙을 유용하게 관리하기 위한 방법으로 상태 코드를 기반으로 한 도표 형식의 응답 규칙을 제안하였다. 이 응답 규칙은 도표의 내용만을 변경함으로써 시스템의 응답을 변경할 수 있다. 따라서 시스템의 응답 내용을 쉽게 바꿀 수 있으며, 응답 기능을 갱신하고 확장하는 작업을 쉽게 할 수 있다.

그러나, 앞의 실험 예에서 볼 수 있는 것처럼 음성 상에서 구현된 시스템은 인식가능 어휘 수의 제한과 오인식에 의하여 문자 입력 상

에서 구현된 시스템에 비해 그 기능이 한정되어 있다. 따라서, 앞으로는 음성인식부에서의 어휘 확장과 이에 따라 예상되는 오인식의 증가를 방지하기 위한 방법에 대한 연구를 진행시킬 계획이다. 또한, 대화관리부에서 담화기록(discourse history)에 대한 분석과 이를 이용한 시스템의 대처 방법에 대한 연구가 필요하다고 판단된다.

## 참고 문헌

- [1] Yoichi Takebayashi, Hiroyuki Tsuboi, Hiroshi Kanazawa, "A Real-Time Speech Dialogue System Using Spontaneous Speech Understanding", *IEICE Trans. Inf. & Syst.*, Vol.E76-D, No.1, 1993, pp.112-120
- [2] Yasuhisa Niimi "How Might One Comfortably Converse with a Machine?", *IEICE Trans. Inf. & Syst.*, Vol.E76-D, No.1, 1993, pp.9-16
- [3] Hiroyuki Nishi, Mikio Kitai, "A New Confirmation Method using the Statistical Relationship between Likelihood and Accuracy", *Technical Report of IEICE*, SP93-101 (1993-11)
- [4] K. Arai, M. Kitai, S. Nakajima, H. Nishi, "Effects of Question Style on Speech Dialog", *Technical Report of IEICE*, SP93-100 (1993-11)
- [5] Ronald A. Cole, David G. Novick, "Rapid Prototyping of Spoken Language Systems: The Year 2000 Census Project", *Proc. ISSD-93 International Symposium on Spoken Dialogue*, 1993, pp.19-23
- [6] T. Yamamoto, Y. Ohta, Y. Yamashita, O. Kakusho, R. Mizoguchi, "MASCOTS: Dialog Management System for Speech Understanding System", *IEICE Trans. Inf. & Syst.*, Vol.E74, No.7, 1991, pp.36-43
- [7] V. Zue, S. Seneff, S. Sakai, J. Polifroni, M. Phillips, C. Pao, D. Goodine, D. Goddeau, G. Flammia, J. R. Glass, "Multilingual Human-Computer Interactions: A Case Study in the VOYAGER Domain", *ATR IWST '93*, November, 1993