

# 양식 채우기 대화에서 음성 인식 오류의 보완을 위한 대화 전략\*

## Dialogue Strategies to Overcome Speech Recognition Errors in Form-Filling Dialogue

강 상 우\*\*  
(Sangwoo Kang)

이 성 욱\*\*\*  
(Songwook Lee)

서 정 연\*\*\*\*  
(Jungyun Seo)

요 약 음성 대화 시스템에서 음성 인식 오류는 전체 시스템의 치명적인 결과를 초래한다. 음성 인식 오류가 부분적으로 발생하여 화행 분석이 실패했을 때 시스템은 원활한 대화를 진행할 수 없다. 본 논문은 양식 채우기 대화 형식에서 발생하는 음성 인식 오류 유형에 따라 시스템이 사용자 발화의 화행을 추론하기 위한 부대화 생성 전략을 제안한다. 제안하는 방법을 계획기반 대화 모델로 구현하여 실험하였고, 사용자 작업 실패 오류의 약27%를 보완하여 성능을 향상시켰으며 전체 시스템의 사용자 작업 성공률은 약 89%이다.

주제어 HCI, dialogue, NLP, UI, strategy

*Abstract* Speech recognition errors cause fatal results in a spoken dialogue system. When a system can not determine the speech-act of an utterance due to speech recognition errors, a dialogue system has a difficulty in continuing conversation. In this paper, we propose strategies for sub-dialogue generation by inferring the speech-act of an utterance with patterns of recognition errors on the field of form-filling dialogue. We used the proposed method on a plan-based dialogue model, corrected 27% of incomplete tasks, and acquired overall 89% of task completion rate.

*Keywords* HCI, dialogue, NLP, UI, strategy

---

\* 이 연구는 2005년도 서강대학교 교내연구비 지원에 의하여 이루어졌음

\*\* 서강대학교 컴퓨터학과

\*\*\* 동서대학교 컴퓨터정보공학부

부산시 사상구 주례2동 산69-1, 전화: 051-320-1744, E-mail: leesw@dongseo.ac.kr

\*\*\*\* 서강대학교 컴퓨터학과 / 바이오융합기술 협동과정

## 서론

자연언어 대화는 인간이 사용할 수 있는 가장 편리하고 효율적인 인터페이스이다. 그래서 이를 이용한 인간과 컴퓨터간의 대화가 가능한 대화 시스템에 대한 관심이 증대되고 있다. 최근의 대화 시스템은 음성 인식기와 기존의 텍스트 기반의 시스템이 결합하여 보다 편리한 사용자 인터페이스를 제공하게 되었다. 하지만, 현재 음성 인식기의 성능은 신뢰할만한 수준이 아니다. 대화 시스템의 각 모듈들은 올바른 음성 인식 결과를 가정하고 설계되었으므로 음성인식 오류로 인한 전체 시스템 성능은 더 나빠질 것이다. 따라서 음성 인식 결과의 신뢰도에 따라 대화 관리자는 동적으로 응답발화를 생성할 수 있어야 한다.

본 논문은 음성인식 단계에서 발생하는 오류 유형에 따라 시스템이 사용자의 목적을 추론하는 대화 전략을 제안한다. 기존의 음성인식 보정을 위하여 사용하는 대화 전략은 “다시 말씀해 주세요.”와 같이 사용자가 동일한 발화를 반복 하도록 한다. 하지만 이러한 대화 전략은 사용자에게 불편하며 동일한 오류가 반복될 가능성이 많다. 우리는 음성 인식에 부분적으로 실패하더라도 대화의 흐름과 상황에 따라 사용자 의도를 예측하고, 시스템이 예측한 사용자의도에 대하여 최소한의 확인요구 발화를 생성하는 효율적인 대화 시스템을 제안한다.

## 관련연구

음성인식 오류가 화행에 미치는 영향

잘못된 음성인식 결과로는 사용자의 의도를 정확하게 파악할 수 없다. 특히 화행(speech act)분석에 음성인식 오류가 영향을 미칠 경우 시스템은 사용자의 의도와는 전혀 다른 방향으로 대화를 진행시킬 수도 있다. 화행은 발화에 나타난 발화자의 의도이고 대화의 목적을 결정하는 가장 중요한 요소이다. 따라서 화행분석은 시스템 전체에서 강인한 구조를 가져야 한다. 다음은 ‘승낙’(acceptance)과 ‘거절’(reject) 두 가지 화행에 대하여 화행이 잘못 분석될 때 사용자와 시스템의 대화가 어떻게 처리되는지에 관한 연구(Marc Cavazza, 2001)의 내용이다.

(예1)은 사용자의 발화 U2에 대한 의도가 ‘거절’을 ‘승낙’으로 잘못 인식되어 시스템이 잘못된 발화 S2를 생성하고 대화가 종료된 예를 보여준다.

(예1)

U1: Do you have any movies?

S1: I found 22 programmes for your request.

I would suggest: “Twelve Monkeys”

U2: Good if I mean anything for long

(실제발화: could you find me anything funny)

S2: Here is “Twelve Monkeys”, then. Goodbye!

대화 시스템에서 음성 인식 보정을 위한 연구

음성 인식 보정에 대한 연구는 인식기 내부의 언어모델이나 음향모델을 중심으로 연구되어져 왔다. 신호처리의 한계에 벗어나 대화모델의 지식을 이용하여 인식 오류를 보정하는 연구가 최근에 시도되고 있다. Genevieve Gorrell(2003)은 인식 오류에 특성에 따라 미리

분류된 시스템 응답 발화를 12개의 범주로 분류하고 LSA(Latent Semantic Analysis)를 사용하여 해당 범주 중에 가장 알맞은 응답을 선택하는 방법을 제안하였다. 그러나 오류 패턴을 분류하는 학습 데이터로 음성 인식기의 학습 데이터를 사용하였기 때문에 학습 데이터에 나타나지 않는 오류 패턴들에 대해서는 처리가 불가능하다.

국내에서는 대화시스템에 관한 연구가 시작 단계이기 때문에 대화 모델을 이용하는 음성 인식 보정에 대한 연구가 아직까지 시도되지 않았다.

### 양식 채우기(form-filling) 대화

본 연구는 양식 채우기(form-filling) 대화에 적용 가능한 대화 전략을 제시한다. 양식 채우기 대화는 사용자가 요구하는 정보를 특정한 양식(form)으로 구축하고 양식의 각 요소를 삽입, 삭제, 수정, 검색의 범위 안에서 조작하여 사용자의 목적을 달성하는 대화형태를 의미한다. 대부분의 작업 지향 대화는 이러한 양식 채우기의 형식을 따르고 있다(David Goddeau 등, 1996).

### 객체 모델

날짜	시간	장소	상대	내용
오늘	1	매점	홍길동	식사
오늘	2	K관	박찬호	세미나
오늘	6	C관	이승엽	식사

사용자 데이터베이스

(그림 1) 일정관리영역의 객체와 데이터베이스

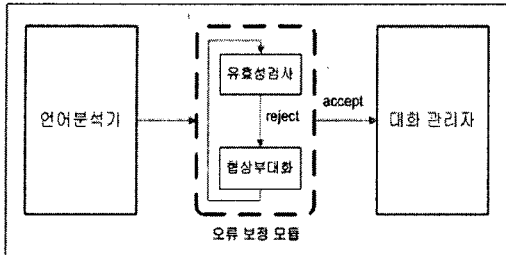
본 논문에서 제안하는 대화 관리자는 사용자의 목적 대상(target)을 객체로 처리한다.

[그림 1]은 일정관리영역에서 사용되는 객체와 사용자 데이터베이스를 나타낸다. [그림 1]의 일정관리 영역에서 생성되는 객체는 다수의 일정 객체가 생성될 수도 있으며, 시스템에서 사용자의 의도 파악을 위한 사용자 데이터베이스 검색의 제약조건으로 사용된다.

### 계획기반 대화모델

계획기반 대화모델은 레서피(recipe)라는 담화구조지식을 사용하여 사용자의 의도를 파악하고 적절한 응답을 생성한다(Diane J. Litman 등, 1987; Lynn Lambert 등, 1991; Rhonda Eller 등, 1992; Lynn Lambert, 1993; Jennifer Chu-Carroll 등, 1995; Jennifer Chu-Carroll, 1996; Jennifer Chu-Carroll, 1996). 표층적인 의미의 레서피로부터 심층적인 의미를 파악하기 위하여 상위 레서피를 추론해 나간다. 레서피에는 제약조건(constraints), 적용조건(applicable condition), 부행위(sub-action) 그리고 효과(effects)와 목적(goal)등이 기술되어 있다. 계획기반 대화모델은 이 레서피를 통하여 사용자의 지식상태를 유지하는 belief space와 시스템 행위의 순서를 정하는 obligation stack등을 갱신하면서 대화를 처리한다.

윤철진 등(2000)과 오종건(1999)은 계획기반 대화모델을 한국어에 적용하여 생략어구 처리와 협상대화가 가능한 모델을 구현하였다. 이들의 모델은 완전한 음성인식 및 언어분석 결과를 가정으로 실험하였기 때문에 인식오류가 발생하는 실제 환경에 적합하지 못하다. 따라서 인식오류에 대처할 추가적인 연구가 필요



(그림 2) 오류 보정 모듈을 추가한 대화 시스템

하다.

## 시스템 구조

기존의 음성 대화 시스템은 음성 인식기 및 언어분석기 그리고 대화 관리자로 구성된다. 언어분석기는 음성인식 결과를 통해 사용자의 의도 파악에 필요한 화행과 행위, 그리고 속성 정보들을 추출하여 그 결과를 대화 관리자에게 전달한다. 본 연구에 사용된 언어분석기는 사용자 발화문장을 제한된 영역에서 구축된 규칙을 이용하여 분석한다. 음성인식 결과가 올바르지 않다면 언어분석기는 대화 관리자에 필요한 정보들을 추출해 내지 못한다. 따라서 음성 인식기와 언어분석기의 신뢰성을 검증하고 그 결과에 따라 동적으로 대화를 관리할 수 있는 방법이 필요하다.

[그림 2]는 제안하는 시스템의 구조를 나타낸다. 오류 보정 모듈은 음성인식 및 언어분석 결과의 유효성을 검사한다. 이 모듈은 유효성 검사를 통하여 언어분석결과가 정상이라면 대화 관리자에게 그 값을 전달하고 비정상이라면 시스템 주도의 협상 부대화(negotiation sub-dialogue)를 통하여 부족한 정보를 사용자에게 얻어낸다. 음성 인식 오류가 발생했을 때

이 모듈에서 적용하는 협상 부대화 진행 전략은 다음 장에서 설명한다.

## 음성인식오류 보정을 위한 대화 전략

양식 채우기 대화에서 사용자는 기본적으로 ‘삽입’, ‘삭제’, ‘수정’, ‘검색’의 명령을 수행한다. 사용자의 구체적인 의도가 이 명령들 중 하나에 해당할 때 이를 ‘행위’라 한다. 시스템이 사용자의 ‘행위’를 알 수 없을 때 시스템은 사용자의 네 가지 ‘행위’에 근거하여 사용자의 의도를 추론할 수 있다.

시스템이 주도하는 협상 대화는 시스템이 갖고 있는 사용자 데이터베이스의 내용을 바탕으로 사용자의 의도를 추론하면서 진행된다. 시스템이 추론한 내용을 확인하는 방법은 크게 암시적 확인과 명시적 확인으로 나눌 수 있는데, 암시적 확인은 시스템이 추론한 화행으로 시스템이 처리한 결과를 알려줌으로써 사용자가 원하는 결과를 얻었는지 간접적으로 확인하는 방법이고, 명시적 확인은 시스템이 추론한 화행을 사용자에게 직접적으로 물어봄으로써 시스템이 추론한 화행을 확인하는 방법이다.

시스템은 명시적 확인을 통해 사용자에게 “삭제해” 혹은 “수정해” 등의 음성 인식률이 높은 명확한 응답을 유도함으로써 전체 시스템 성능을 향상시킬 수 있다. 그리고 암시적 확인을 통해서도 시스템이 파악한 사용자의 의도가 올바른지 확인할 수 있다.

대화 전략은 언어분석 오류패턴에 따라 크게 두 가지로 분류할 수 있다. 첫 번째는 언어분석 결과에서 사용자의 화행과 행위가 정

해지지 않는 패턴인 경우이고 두 번째는 화행이 *Wb-question*<sup>1)</sup>이고 대상이 정해지지 않는 패턴인 경우이다. 또한 두 가지 분류에 대한 세부적인 전략들은 각각 확인요구발화 생성 및 추론 전략을 갖는다.

다음 절에서는 화행과 행위가 정해지지 않는 패턴들을 처리하는 전략과 화행이 *Wb-question*이고 대상이 정해지지 않는 패턴들을 처리하는 전략을 설명한다.

**화행과 행위가 정해지지 않은 패턴**

[그림 3]의 예제에서 사용자 발화를 살펴보면 화행과 행위를 결정짓는 서술부가 인식되지 않았기 때문에 언어분석을 통해서도 화행을 결정할 수 없다. 이 경우 사용자의 대화 목적을 알 수 없기 때문에 대화를 진행시키기 어렵다. 하지만 본 연구에서는 인식에 성공한 다른 단어들과 사용자 데이터베이스를 사용하여 사용자의 목적을 예측해 나가는 대화를 시도한다.

[그림 3]에서 인식에 성공한 단어들은 “오늘”, “한시에”, “홍길동”이며 이 단어들은 영역객체의 속성 값으로 사용될 수 있다. 이 단어들을 사용하여 [그림 3]과 같이 영역객체를 생성한다. 이 영역객체는 사용자 데이터베이스의 검색조건으로 사용되며 검색결과에 따라 아래 세 가지 경우로 분류하고 각 경우에 따라 다른 전략들을 이용하여 대화를 진행시킨다.

# 화행과 행위가 정해지지 않은 경우

1. 검색결과가 없는 경우 -> 전략1

사용자 발화 : 오늘 한시에 홍길동 000 (000 : 인식 실패 단어)				
영역 객체				
날짜	시간	장소	약속상대	내용
오늘	1		홍길동	
사용자 데이터베이스				
날짜	시간	장소	약속상대	내용
오늘	1	매점	홍길동	식사
오늘	2	K관	박찬호	세미나
오늘	6	C관	이승엽	식사

(그림 3) 전략2의 사용자 데이터베이스 검색 절차

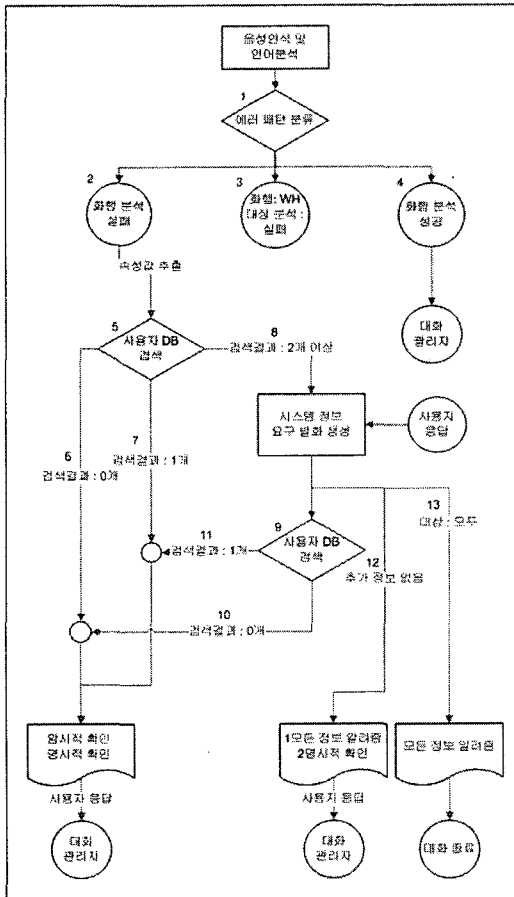
2. 검색결과가 한 개인 경우 -> 전략2
3. 검색결과가 두 개 이상인 경우 -> 전략3

전략1의 경우, 사용자가 발화한 정보들이 사용자 데이터베이스에 존재하지 않기 때문에 시스템은 사용자가 정보를 입력하는 의도를 가진다고 추측하고 대화를 진행시킨다.

전략2의 경우 시스템은 사용자가 이미 입력한 정보에 대하여 수정이나 삭제 작업을 원한다고 추측한다. 또한 추가나 삭제 작업이외에 다른 모든 가능성 있는 의도에 대해서도 대응할 수 있는 암시적 확인 발화와 명시적 확인 발화를 사용하여 대화를 진행시킨다.

전략3의 경우, 먼저 사용자에게 추가 정보를 요구하는 질문을 통하여 사용자의 응답을 얻는다. 얻어진 응답에서 영역객체의 속성 값을 추가하는 것이 가능하면 값을 추가한다. 그리고 갱신된 영역객체를 사용하여 사용자 데이터베이스를 다시 검색하고, 그 검색 결과에 따라 세 가지 전략을 재귀적으로 이용하여 대화를 진행시킨다.

1) 화행의 한 종류, 일반적으로 의문문에 해당함



(그림 4) 화행과 행위가 정해지지 않은 패턴에 적용한 전략들의 수행절차

[그림 4]는 화행과 행위가 분석되는 않는 패턴에 적용할 수 있는 세 가지 전략들의 수행 절차를 나타낸다.

# 수행절차

1. 어려 패턴 분류
2. 화행, 행위 분석에 실패한 경우
3. 화행은 *wh-question*, 대상 분석에 실패한 경우
4. 화행, 행위 분석 성공
5. 인식된 정보들을 사용하여 사용자 DB검색

6. 전략 1 적용
7. 전략 2 적용
8. 전략 3 적용: 사용자에게 추가 정보요구
9. 사용자 DB 재검색
10. 전략 1 적용
11. 전략 2 적용
- 12, 13 예외 처리

다음은 세 가지 전략 중 전략2에 관한 세부적인 내용을 예제를 통해 설명한다.

[그림 5]에서 사용자 발화의 내용은 사용자 데이터베이스(그림 3)에 이미 입력되어 있는 내용이므로 시스템은 사용자가 이미 입력한 정보에 대해서 '수정' 혹은 '삭제'의 의도를 갖는다고 추측할 수 있다. 또한 입력한 일정의 특정내용에 대하여 '질문'의 의도도 포함될 수 있다. 그 이외의 경우, 사용자가 데이터베이스에 입력한 내용에 대하여 잘못된 지식을 갖는 경우로서, 시스템은 암시적 확인 발화를 통해 사용자에게 이를 인식 시킨다.

[그림 6]은 인식 실패한 사용자 발화에 대하여 예측 가능한 발화의 리스트를 나타낸다. 시스템은 사용자의 의도가 [그림 6]의 7가지 발화 중 하나라고 추측하게 된다. 시스템은 모든 가능성 있는 발화를 만족시키기 위하여 암시적 확인과 명시적 확인을 사용한다. [그림 5]에서 시스템 발화1은 암시적 확인 발화이며 시스템은 이 사용자의 '검색' 의도(발화3~5)를

사용자 발화 : 오늘 한시에 식사 약속 000  
 시스템 발화1: 오늘 1시 홍길동님과 매점에서 식사약속을 검색 하였습니다.  
 시스템 발화2: 수결과삭제 작업 중 어느 작업을 원하십니까?

(그림 5) 전략2의 대화예제

발화1. 오늘 1시에 식사 약속 수정해줘. 발화2. 오늘 1시에 식사 약속 삭제해줘. 발화3. 오늘 1시에 식사 약속 알려줘. 발화4. 오늘 1시에 식사 약속 누구와 만나니? 발화5. 오늘 1시에 식사 약속 어디서이니? 발화6. 오늘 1시에 식사 약속 있니? 발화7. 오늘 1시에 식사 약속 등록해줘. (시스템과 사용자의 지식 충돌)
---

(그림 6) 사용자의 예상 발화 리스트

암시적으로 충족시킨다. 또한 이 시스템 발화1의 생성은 사용자와 시스템 지식과의 충돌이 발생한 발화7과 같은 경우에도 사용자에게 암시적으로 사용자의 지식이 틀렸음을 알려준다.

시스템은 시스템 발화2를 통하여 시스템이 추론한 발화1,2를 처리하기 위한 명시적 확인을 요구한다. 사용자는 시스템의 요구에 따라 “수정” 또는 “삭제”의 의미를 갖는 발화로 응답할 수 있고 시스템은 사용자의 응답에 따라 해당 작업을 수행한다. 즉, 모든 가능한 사용자 발화 의도에 대해서 암시적 확인 발화와 명시적 확인 발화를 생성함으로써 사용자 목적을 충족시키는 대화를 진행시킬 수 있다.

화행이 Wh-question이고 대상이 정해지지 않는 패턴

Diane J. Litman(1987)은 화행의 종류를 INFORM, REQUEST, INFORMREF (*Wh-question*), INFORMIF(*YN-question*) 네 가지로 정의하고 있다. 이 중에서 *Wh-question*은 다른 화행과 달리 의문의 대상(target)을 필요로 한다. 만약 *Wh-question*화행을 갖는 문장에서 그 대상을 알지 못한다면 시스템은 사용자의 의도와는 다른 응답을 생성할 것이다. 본 절에서는 이 때 시스템이 대화를 처리하는 과정을 보인다.

화행이 *WH-question*이고 대상이 정해지지 않는 패턴에서도 영역객체에 대한 사용자 데이터베이스 검색 결과에 따라 다음 세 가지 경우로 분류하여 처리한다.

# 화행은 *WH-question*이고 대상이 정해지지 않는 경우

1. 검색결과가 없는 경우 -> 전략4
2. 검색결과가 한 개인 경우 -> 전략5
3. 검색결과가 두 개 이상인 경우 -> 전략6

전략4의 경우, 사용자는 시스템에게 특정 정보<sup>2)</sup>를 요구하기 위해 그 정보의 조건<sup>3)</sup>들을 발화하였지만 조건에 맞는 정보가 존재하지 않은 경우이다. 이 경우 시스템은 특정 정보의 대상을 알지 못하여도 “정보가 없습니다.”와 같은 시스템 발화 생성이 가능하다. 또한 사용자는 존재하지 않는 정보에 대하여 질문을 하였기 때문에 잘못된 지식을 가지고 있음을 알 수 있다. 그러므로 “정보가 없습니다.”와 같은 시스템 발화는 사용자가 잘못된 지식을 가지고 있음을 확인시키는 암시적 확인 발화이다.

전략5의 경우, 사용자는 특정 정보에 대하여 특정 내용<sup>4)</sup>을 요구하는 경우이다. 이 경우 사용자가 요구하는 정보는 찾았지만 대상을 알지 못하기 때문에, 검색된 정보의 대상이 될 수 있는 모든 내용을 알려줌으로서 사용자의 목적을 달성시킨다. 또한 사용자가 이미 발화한 조건들을 제외한 다른 정보들만 알려줌으로 효율적인 대화를 이끌어 간다.

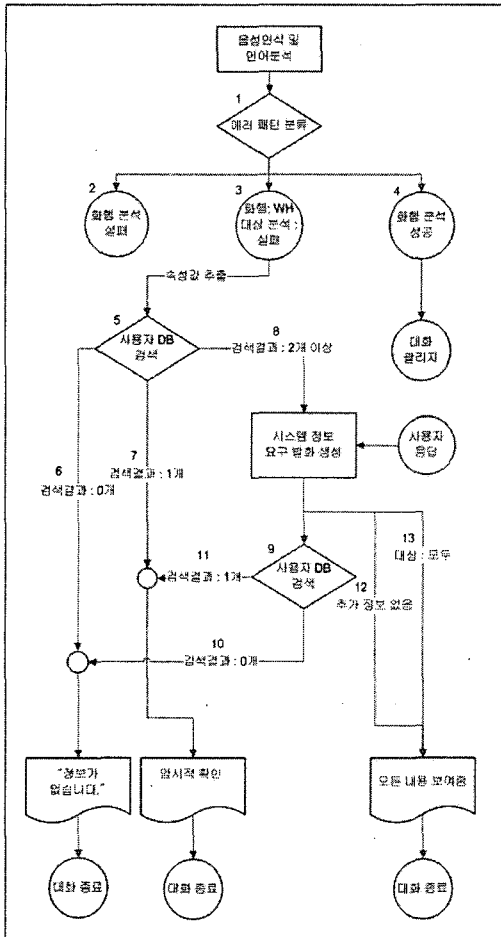
2) 여기서 특정 정보는 데이터베이스 테이블의 튜플

3) 데이터베이스 테이블의 필드값

4) 데이터베이스 테이블의 필드값

전략6의 경우, 전략3과 마찬가지로 먼저 사용자에게 추가 정보를 요구하는 질문을 통하여 사용자의 응답을 얻는다. 얻어진 응답에서 영역객체를 갱신하고 갱신된 영역객체를 사용하여 사용자 데이터베이스를 재검색한다. 그 검색결과에 따라 세 가지 전략(전략4~6)을 재귀적으로 이용하여 대화를 진행시킨다.

[그림 7]은 전략4~6에 대한 수행 절차를 나타낸 순서도이다.



(그림 7) 회행이 WH-question이고 대상이 정행지지 않은 패턴에 적용한 전략들의 수행절차

# 수행절차

1. 에러 패턴 분류
2. 회행, 행위 분석에 실패한 경우
3. 회행은 *wh-question*, 대상 분석에 실패한 경우
4. 회행, 행위 분석 성공
5. 인식된 정보들을 사용하여 사용자 DB검색
6. 전략 4 적용
7. 전략 5 적용
8. 전략 6 적용 : 사용자에게 추가 정보요구
9. 사용자 DB 재검색
10. 전략 4 적용
11. 전략 5 적용
- 12, 13 예외 처리

전략6에 관한 세부적인 내용을 예제를 통하여 살펴보자.

[그림 8]은 사용자 데이터베이스에서 두 개 이상의 정보가 검색되었을 때의 대화 예제이고, [그림 9]는 사용자 발화1에 대한 사용자 데이터베이스 검색 결과이다.

시스템은 [그림 9]에서 보인 바와 같이 두 개의 정보 중에서 작업을 진행할 한 가지의 정보를 결정하여야 한다. 그래서 [그림 8]에서처럼 시스템 발화1을 생성하여 사용자의 응답을 요구한다. 사용자 응답 발화는 [그림 8]의 a, b, c, d 중 하나일 것이다.

이 때, 패턴a는 시간 속성 값으로 '9'가 채

사용자 발화1: 오늘 식사 약속 **000** 만나지?  
 시스템 발화1: 약속시간을 알려주세요?  
 사용자 응답a: 1시  
           b: 9시  
           c: 모두 알려줘.  
           d: 모른다.

(그림 8) 전략 6의 대화예제



사용자 발화 : 오늘 식사 약속 000 만나지  
(000 : 인식 실패 단어)

영역 객체				
날짜	시간	장소	약속상대	내용
오늘				식사

사용자 데이터베이스

날짜	시간	장소	약속상대	내용
오늘	1	매점	홍길동	식사
오늘	2	K관	박찬호	세미나
오늘	6	C관	이승엽	식사

(그림 9) 전략6의 사용자 데이터베이스 검색 절차

위절 때 갱신된 영역객체로 데이터베이스를 재검색한 결과, 일치하는 정보가 존재하지 않는 경우이다. 이 경우는 전략4에 따라 “정보가 없습니다.”와 같은 시스템 응답발화를 생성한다.

패턴b는 시간 속성 값으로 ‘1’이 채워지며 갱신된 영역객체에 의해 사용자 데이터베이스 검색 결과, 한 개의 일치하는 정보가 있는 경우이다. 이 때 전략5에 따라 “오늘 매점에서 홍길동님과 약속이 있습니다.”라고 시스템은 응답한다.

패턴c의 경우, 현재 객체와 일치하는 테이블의 모든 정보를 보여줌으로써 사용자의 목적을 달성시킬 수 있다. 시스템은 “오늘 1시에 매점에서 홍길동님과, 그리고 9시에 C관에서 이승엽님과 약속이 있습니다.”라고 응답한다.

패턴d의 경우, 시스템이 유일한 일정을 결정할 수 없기 때문에 단순히 현재 영역객체와 일치하는 모든 정보를 보여준다. 결과적으로 시스템 응답은 패턴c와 동일하다.

## 실 험

본 연구에서 제안하는 전략은 양식 채우기(form-filling) 대화에서 일반적인 적용이 가능하다. 본 장에서는 제안한 대화 전략 중 전략 2(사용자 데이터베이스에서 객체가 검색되지 않은 경우)를 일정관리 영역에서 계획기반 대화모델에 적용한 예를 보인다.

### 계획기반 대화모델의 적용

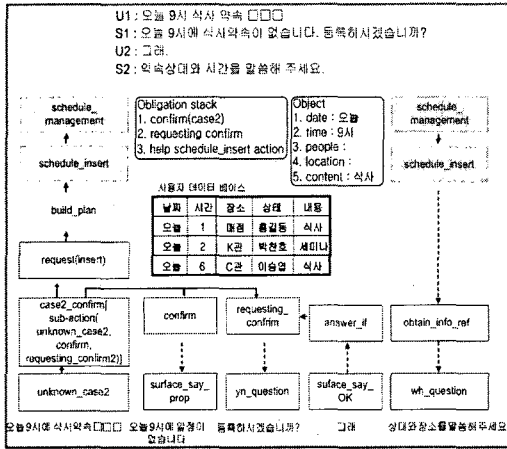
실험에는 계획기반 대화모델을 한국어에 적용한 오종건(1999)의 모델을 바탕으로 구현하였고 이 모델은 레서피(recipe)를 통하여 사용자의 의도를 파악하고 담화구조를 생성한다. 본 연구에서는 6가지 전략을 적용하기 위하여 추가적인 레서피를 작성하였다.

전략2의 경우에는, 전략2의 인식 오류 패턴을 정의하기 위해 surface\_unknown\_case2 레서피와 case2\_confirm 레서피를 추가로 작성하였다.

surface\_unknown\_case2는 화행과 행위가 정해지지 않았을 때 영역객체의 사용자 데이터베이스 검색 결과가 한 개인 경우 사용하는 입력형식이다. case2\_confirm는 surface\_unknown\_case2의 상위노드이며 시스템 주도의 부대화를 진행시킨다. case2\_confirm에 의해서 암시적 확인 발화와 명시적 확인 발화가 생성되며 사용자의 응답에 따라 사용자의 화행이 결정된다.

### 추론 형태

[그림 10]은 제안된 방법을 계획기반 대화모델에 적용한 예를 보인다. 수행절차는 아래와 같다.



(그림 10) 일정관리 영역에서 적용 예

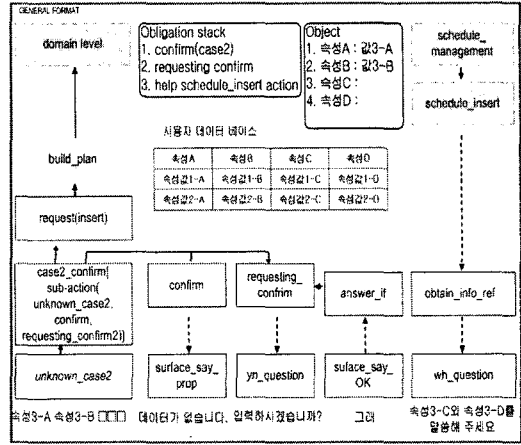
# 수행 절차

1. 인식 오류 문장 입력
2. 오류형태에 따른 레서피 선택(화행 예측)
3. 상위 레서피 추론
4. 각 레서피에서 생성되는 시스템 작업을 stack에 저장
5. stack의 맨 위에 있는 작업을 실행(confirm)
6. 시스템 confirm에 대한 사용자 응답에 의해 인식 오류 문장에 대한 화행 결정
7. 결정된 화행으로 작업 지향 대화 진행

[그림 11]은 양식 채우기 대화를 본 연구에서 제안하는 전략을 적용하여 계획 기반 대화 모델에 적용한 일반적인 모델이다. [그림 11]에서 시스템 확인발화의 형태를 영역에 적합하게 변형시킴으로써 영역 이식성을 높일 수 있다.

성능평가

음성 인식 오류에 대한 성능을 평가하기 위해 우리는 Dong-Hoon Ahn 등(2004)의 음성 인



(그림 11) 양식 채우기 대화에서 일반적 형태

<표 1> 실험 결과

	기존시스템	제안시스템	오류수정률
작업 성공률	85%	89%	27%

식을 사용하였다. 음성 인식기의 신뢰도가 80%일 때를 임계점으로 고정하고 200개의 대화에 대하여 실험한 결과를 사용하였다. 사용자 발화를 음성 인식 결과의 신뢰도가 80%이하이면 시스템은 “다시 한 번 말씀해 주세요.”라는 응답 발화를 생성하게 된다. 본 실험에서는 이러한 시스템 발화가 생성되면 작업 실패로 간주하였다. 성능 평가 방법으로는 기존 시스템에서 이와 같이 작업 실패할 경우, 즉 음성 인식기의 신뢰도가 80%이하인 경우에 제안된 6가지 전략을 적용시키고 작업 성공률<sup>5)</sup>을 측정하였다.

<표 1>에서와 같이, 일정관리 영역에서 기존 시스템[11]의 사용자 작업 성공률은 85%이며 제안된 방법을 적용하였을 때 작업 성공률

5) 사용자가 의도한 작업을 대화 시스템이 성공시킨 확률

은 89%로 약4%의 전체 시스템 성능 향상을 가져왔다. 이는 제안하는 방법이 기존 시스템에서 발생하는 오류의 약27%를 보완하는 효과를 가져왔다고 해석할 수 있다.

### 결론 및 향후 과제

본 논문은 음성 인식 결과의 신뢰도가 낮아 사용자의 의도를 분석할 수 없을 때, 시스템이 추론한 사용자의 의도를 확인하고 협상 부대화(negotiation sub-dialogue)를 통하여 정보를 사용자 데이터베이스의 양식 객체에 추가해주는 대화 전략을 제안하였다. 제안된 전략을 이용하여 실험한 결과, 음성 인식오류의 약 27%를 수정할 수 있었다.

대화 시스템에서 작업 실패는 사용자가 해당 작업을 다시 시스템에 요구해야 한다는 의미이며, 이 경우 시스템의 체감 성능은 낮게 평가 된다. 그러나 추가적인 대화가 발생하더라도 해당 작업을 성공시키게 되면 시스템의 신뢰도는 높아진다. 본 논문에서 제안한 방법을 사용하여 대화 시스템에서의 작업 성공률을 약4% 향상시킴으로써 시스템의 체감 신뢰도를 한층 높일 수 있었다.

앞으로 양식 채우기 대화 이외의 다른 대화 형태에서도 오류를 포함한 인식결과로부터 원래의 의미를 복원할 수 있는 모델에 대한 연구가 필요하다. 이를 위하여 언어학적인 접근과 함께 인지학적인 접근도 시도해야 할 것이다. 또한 음성 인식 오류에 대한 체계적인 분석과 음성인식 오류가 화행에 미치는 영향에 대한 지속적인 연구도 병행되어야 한다.

### 참고문헌

- Marc Cavazza, "An Empirical Study of Speech Recognition Errors in a Task-oriented Dialogue System", *Proceedings of the 2nd SIGdial*, 2001
- Genevieve Gorrell, "Recognition Error Handling in Spoken Dialogue System", *Proceedings of 2nd International Conference on Mobile and Ubiquitous Multimedia*, 2003
- David Goddeau and Helen Meng and Joe Polifroni, Stephanie Seneff, and Senis Busayapongchai, "A Form-Based Dialog Manager for Spoken Language Applications", *Proceedings of the ICSLP 96*, pp701-705, 1996.
- Diane J. Litman and James F. Allen, "A Plan Recognition Model for Subdialogue in Conversations", *Cognitive Science*, vol 11, pp163-200, 1987
- Lynn Lambert and Sandra Carberry, "A tripartite plan-based model of dialogue", *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pp47-54, 1991
- Rhonda Eller and Sandra Carberry, "a meta-rule approach to flexible plan recognition in dialogue", *User Modeling and User-Adapted Interaction*, vol 2, num 1-2, pp27-53, 1992
- Lynn Lambert, *Recognizing Complex Discourse Acts : A Tripartite Plan-Based Model of Dialogue*, Ph.D. thesis, University of Delaware, 1993
- Jennifer Chu-Carroll and Sandra Carberry, "Generating information-sharing sub-dialogues in expert-user consultation", *Proceedings of the 14th International Conference on Artificial Intelligence*, pp1234-1250, 1995

Jennifer Chu-Carroll, *APlan-Based Model for Response Generation in Collaborative Negotiation Dialogues*, Ph.D. thesis University of Delaware, 1996

윤철진, 서정연, “제한된 영역의 대화에서 체언구 형태의 발화 이해를 위한 계획기반 생략 처리”, 한국인지과학회 논문지, 제11권 1호, pp81-92, 2000

오종건, 계획에 기반한 대화 시스템의 설계, 석사학위논문, 서강대학교, 1999

Dong-Hoon Ahn and Minhwa Chung, “One-pass Semi-dynamic Network Decoding Using a Subnetwork Caching Model for Large Vocabulary Continuous Speech Recognition”, *IEICE Transactions on Information and Systems*, vol.E87-D, no.5, pp.1164 - 1174, May 2004

1 차원고접수: 2006. 2. 24

2 차원고접수: 2006. 5. 15

최종게재승인: 2006. 6. 14