

레스토랑 영역에서의 자질기반 대화시스템 구현

양현석*, 김동주**, 설용수*, 정성훈*, 김한우*

*한양대학교 컴퓨터공학과

**안양대학교 컴퓨터공학과

e-mail : iexisto@hanyang.ac.kr, djkim@anyang.ac.kr,

iamtbm@hanyang.ac.kr, ishkkman@gmail.com, kimhw@hanyang.ac.kr

Implementation of Feature-based Dialog System in Restaurant domain

Hyeon-Seok Yang*, Dong-Joo Kim**, Yong-Soo Seol*,

Sung-Hun Jung*, Han-Woo Kim*

*Dept of Computer Engineering, Hanyang University

**Dept of Computer Engineering, Anyang University

요 약

서비스 로봇과 펫 로봇 등 사람과 직접 상호작용하는 로봇기술의 필요성이 증가하고 있다. 대화시스템은 자연언어처리 기술을 활용하여 음성인식 기술과의 결합을 통해 현재 로봇에서 주로 사용되고 있는 버튼과 터치스크린 위주의 HRI(Human-Robot Interface)보다 자연스러운 HRI를 제공한다. 이러한 자연스러운 HRI를 수행할 수 있는 로봇을 구성하기 위해서는 로봇이 서비스를 제공할 실제 영역에 맞는 대화시스템의 연구가 필요하다. 본 논문에서는 자질사전, 단일화 문법(unification grammar), 대화 흐름도(dialogue flow diagram)를 사용한 레스토랑 영역의 자질기반(feature-based) 대화시스템을 제시한다. 자질 정보는 형태소, 시제, 어휘의 의미구조 등을 나타내며 화행(speech act) 결정에 사용하고 문장 자질과 구문 자질을 파서에서 활용한다. 자질기반 대화시스템을 통하여 레스토랑 영역에서 사용자 화행 이해 및 주문, 안내 등의 서비스를 성공적으로 수행할 수 있음을 보인다.

1. 서론

서비스 로봇과 펫 로봇 등 사람과 직접 상호작용하는 로봇기술의 필요성이 증가하고 있다. 대화시스템은 자연언어처리 기술을 활용하여 음성인식 기술과의 결합을 통해 현재 로봇에서 주로 사용되고 있는 버튼과 터치스크린 위주의 HRI(Human-Robot Interface)보다 유연성, 명료성, 표현력 면에서 더 뛰어난 HRI를 제공한다. 대화는 상황과 문맥에 의존적이므로 모든 대화 영역(dialogue domain)을 대상으로 하는 대화 시스템의 구현은 현실적으로 어렵다 [1]. 자연스러운 HRI를 수행할 수 있는 대화시스템을 활용한 로봇을 구성하기 위해서는 로봇이 서비스를 제공할 실제 영역에 맞는 대화시스템의 연구가 필요하다. 최근까지 대화시스템은 호텔예약[2, 3], 항공예약[4], 의학정보제공 [5] 등이 연구되었으며 주로 전화예약이나 웹사이트에의 적용에 적합했다. 본 논문에서는 실제 로봇이 서비스를 제공할 수 있는 서비스 로봇에의 활용을 가정하고 안내, 주문, 계산 등의 다양한 서비스가 필요한 레스토랑 영역에서의 대화시스템을 제시한다.

대화 시스템은 자연어 기반 대화를 통해서 사용자의 의도를 이해하고 적절한 행동을 취하여 사용자를 돕는다 [6]. 사용자와 상호작용을 하기 위해서 대화시스템은 사용자 발화가 입력되고 응답 문장을 생성한다. 대화시스템은

크게 언어해석부, 문맥처리부, 대화생성부가 필요하다. 언어해석부는 형태소 분석, 패턴매칭 혹은 구문구조 분석을 수행한다. 문맥처리부는 사용자 입력과 이전 대화 흐름을 고려하여 사용자의 의도를 파악하고 그에 따른 행위 수행, 정보 제공, 정보 요청 등의 시스템 화행(speech act)을 결정한다. 대화생성부는 결정된 화행과 영역 지식(domain knowledge)를 활용하여 적절한 응답 발화를 생성한다.

본 논문에서 제시하는 대화시스템은 코퍼스를 분석하여 얻어진 자질사전과 단일화 문법(unification grammar)을 사용하여 자연언어이해를 수행하고 대화시스템에서 많이 사용되는 대화 흐름도(dialogue flow diagram)와 양식(form)을 사용한다. 또한 레스토랑 영역의 서비스를 처리하기 위한 화행을 정의하며 화행의 결정에 형태소, 시제, 의미 자질등을 사용하고 파서에 문장 자질과 구문 자질을 활용한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 시스템과 관련된 연구들에 대해서 소개한다. 3장에서는 대화시스템을 구성하기 위해 필요한 정의와 모듈 설명을 한다. 4장에서는 현재 시스템에 대해 평가와 논의를 한다.

2. 관련 연구

대화시스템은 대화 흐름도(dialogue flow diagram) 모

델, 예제기반(example-based) 모델, 계획기반(plan-based) 모델 등이 있다.

예제기반 모델은 대화 설계에 있어서 대화 예제를 수집하고 의미정보를 태깅한 후, 사용자 질의문과의 태깅된 정보 간의 유사도를 계산하여 대화를 수행하는 방법이다. 상대적으로 적은 양의 예제만으로도 높은 성능의 대화 시스템을 구축할 수 있다. 그러나 이 방법은 수집된 대화를 분석하여 의미정보를 태깅을 하여야 하는 번거로움이 있고 코퍼스에서 유사한 문장을 찾으므로 목적지향의 논리적인 대화에는 적합하지 않다.

계획기반 모델은 복잡한 담화 현상을 처리할 수 있고 심층적인 의도 추론이 가능하다. 반면 초기 설계가 매우 복잡하며 유지 관리가 어렵다는 단점이 있다. 계획기반 대화시스템의 대표적인 모델은 Litman 모델과 Caberry 모델이 있다.[7, 8, 9]

대화 흐름도 모델은 순차적인 실행 구조를 가지며 비교적 설계가 간단하다. 반면 영역에 의존적이며 다양한 상황 대응이 어렵다.

본 논문에서는 비교적 단순하고 목적지향대화에 적합한 대화 흐름도 모델을 사용한다.

양식 채우기(form-filling) 대화는 사용자가 요구하는 정보를 특정한 양식으로 구축하고 양식의 각 요소를 삽입, 삭제, 수정, 검색의 범위 안에서 조작하여 사용자의 목적을 달성하는 대화 형태를 의미한다. 대부분의 작업 지향 대화는 이러한 양식 채우기의 형식을 따르고 있다. [10]

3. 구현 시스템

제시하는 대화시스템은 레스토랑 영역을 대상으로 한다. 레스토랑 영역에서는 안내, 주문, 예약, 계산, 자리 안내 등의 서비스가 제공되어야 한다. 안내, 주문, 예약, 계산을 수행하기 위해서는 레스토랑에서 제공하는 음식의 명칭, 종류, 가격 등의 영역 지식이 필요하다.

대화시스템에는 사용자 발화 문장이 입력된다. 입력된 각 문장에 대해 원형복원을 수행한다. 원형복원이 수행된 결과는 형태소 분석기를 사용하여 분석된다. 또한 형태소 분석된 결과를 파싱하여 파스 트리를 생성한다. 파스 트리 와 정의된 대화 흐름도를 사용하여 화행을 결정하고 그 화행에 맞는 응답 화행에 따라 응답 문장을 생성한다.

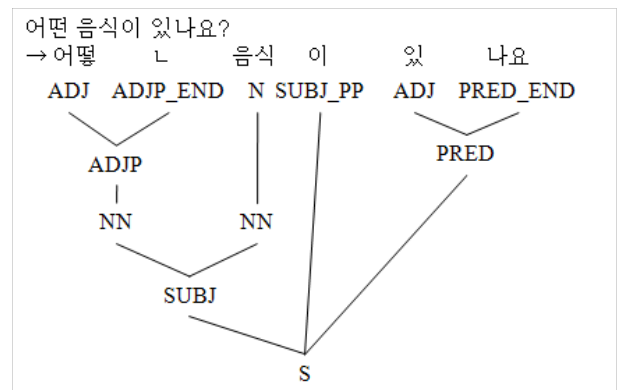
대화시스템을 구축하기 위해서는 먼저 대화시스템이 인식할 발화와 생성할 발화를 구성하기 위한 코퍼스가 필요하다. 본 시스템에서는 총 329문장으로 이루어진 시나리오 40개를 분석하여 사전과 화행을 정의한다.

사전은 코퍼스를 분석하여 자질사전의 형태로 구축한다. 형태소 자질 태그는 체언(N), 동사(V), 형용사(ADJ), 관형사(ADNOM), 부사(ADV), 조사(PP), 어미(END), 감탄사(II), 보조용언(PX), 서술격조사(JCP)로 총 10가지이다. 또한 형태소의 의미정보를 나타내기 위하여 각 단어의 유형별로 필요한 자질을 추가한다. 추가된 자질은 체언의 경우 음식종류, 장소, 수량, 음식속성 등이며 동사는 주문

하(다), 주(다), 굶(다) 등이며 어미는 시제를 나타낸다. 의미자질은 동의어의 표현과 의미 계층을 표현한다. 표 1은 의미 자질 값의 계층 태그의 예시이다. 위 자질들은 화행 결정의 중요한 기준으로 활용된다. 형태소 분석 결과는 파서에 입력된다. 파서는 문장 자질과 구문 자질을 활용한다. 파서는 형태소 분석 결과를 주어부, 목적어부, 형용사, 부사 등을 나타내는 파스 트리로 만든다. 파스 트리에 대한 검색을 통하여 문장의 주어부, 목적어부 등을 획득한다. 그림 1은 파싱한 결과의 예시이다.

<표 1> 의미 자질 값의 계층 태그 일부

상위 자질	하위 자질				
FOOD	MEAT	RICE	STEW
SERVICE	ORDER	GIVE
PLACE	RESTAURANT	TOILET



(그림 1) 파스 트리의 예시

문맥처리부는 사용자 화행의 결정, 정보 추출, 응답 화행의 결정의 역할을 수행한다. 사용자 화행의 결정과 정보 추출은 언어해석부의 결과로 얻어진 파스 트리를 이용하여 이루어진다. 응답 화행은 결정된 사용자 화행과 추출된 정보를 사용하여 결정한다.

코퍼스를 분석한 결과로 사용자 화행과 시스템 화행을 총 13가지로 정의한다. 정의된 사용자의 화행은 요구(Request), 문의(Inquiry), 동의(Agree), 거부(Disagree), 항의(Complain)이며 시스템의 화행은 인사(Greeting), 요구 수락(Accept), 요구 거절(Refuse), 안내(Descript), 제안(Suggest), 주문 확인(Confirm), 사과(Apology), 약속(Commiss)이다. 사용자는 요구와 문의로 서비스를 요청하며 동의와 거부, 항의 등으로 시스템의 응답에 반응한다. 시스템은 요구 수락, 요구 거절, 안내로 사용자의 서비스 요청에 반응하며 주문 확인을 통해 주문 내용을 확인한다. 또한 처리할 수 없는 요청에 대하여 제안, 사과, 약속으로 응답한다. 사용자의 요구(Request)나 문의(Inquiry), 시스템의 안내(Descript)의 경우 내용에 따라 다르게 처리한다. 정의된 화행에 해당하지 않는 사용자 발화에 대해서는 안내(Descript)를 수행한다. 표2는 정의된 화행들이다.

<표 2> 화행의 정의

화행	Actor	설명
Request	User	요구
Inquiry	User	문의
Agree	User	동의
Disagree	User	거부
Complain	User	항의
Greeting	System	인사
Accept	System	요구 수락
Refuse	System	요구 거절
Descript	System	안내
Suggest	System	제안
Confirm	System	주문 확인
Apology	System	사과
Commiss	System	약속

입력 문장의 정보는 형태소 분석과 구문분석을 거쳐 얻어진 파스 트리로 표현된다. 입력 문장의 화행은 코퍼스를 분석한 결과로 얻어진 규칙들을 사용해 결정한다.

레스토랑 대화시스템에서 사용자의 주문을 처리하기 위해서는 음식의 명칭, 종류, 가격 등에 대한 영역 지식이 필요하다. 영역 지식을 표현하기 위하여 음식의 종류, 음식의 속성을 나타내는 사전을 구축하고 양식 채우기 방법을 사용한다. 또한 자질태그를 사용하여 요청된 음식의 종류를 인식하고 음식에 대한 설명에 응답을 할 수 있도록 구현하였다. 표 3은 양식의 예시이다.

<표 3> 양식(form) 예시

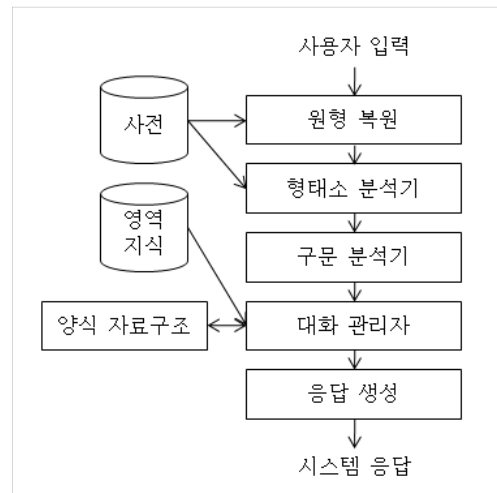
항목	값
음식명	스테이크
음식 수량	1
음식 종류	MEAT
음식 크기	NULL
음식 맛	NULL
구운 정도	미디엄

시스템은 사용자의 요청에 맞는 정보를 제공하고 필요한 추가정보를 요구하며 요구를 처리하였음을 나타내기 위한 발화를 생성해야 한다. 응답은 대화 흐름도를 이용하여 현재 사용자의 화행과 이전 화행 등을 고려하여 문맥 처리부에서 시스템의 화행을 결정된 뒤 시스템 발화 생성 모듈을 통하여 이루어진다. 이 과정에서 사용자의 발화를 분석하여 양식 자료구조로 추출한 정보와 음식의 명칭, 종류, 가격 등의 영역 지식을 사용한다. 생성 문장의 형태는 각 시스템 화행마다 몇 가지 패턴을 사용한다. 표 4는 생성 문장 패턴의 일부이다.

<표 4> 생성 문장 패턴의 일부

화행	패턴 예시
Greeting	안녕하세요. 한양식당입니다. 무엇을 드시겠습니까?
Confirm	주문 확인해 드리겠습니다. [음식명] [수량] 맞으신가요.
Apology	죄송합니다.

그림 2는 전체 시스템 구조도이다. 사용자 입력이 들어오면 원형 복원이 먼저 수행되고, 형태소 분석과 구문 분석을 수행한다. 그 결과 생성된 파스 트리를 사용하여 대화 관리자가 문맥을 처리한다. 대화 관리자는 정보를 추출하여 양식을 채우고 사용자 화행을 결정하며 그에 대응되는 시스템 화행을 결정한다. 대화 생성부는 결정된 시스템 화행의 문장 패턴과 양식을 사용하여 문장을 생성한다.



(그림 2) 시스템 구조도

4. 평가

제시한 대화시스템의 사전과 대화 흐름도 구성을 위하여 총 329문장으로 이루어진 시나리오 40개를 분석하였다. 구현된 대화시스템에서는 처리되지 않은 문장에 대하여 재 발화를 유도하였으며 사용자 화행에 대응하는 응답을 생성하여 주문 안내 등의 서비스를 수행한다.

그림 3은 구현된 시스템의 대화처리 예시이다. 영역 지식을 사용하여 없는 메뉴에 대해 해당 메뉴가 없다는 안내가 제시되고 사용자의 문의에 따라 음식 종류가 안내되었다. 최종적으로 된장찌개를 주문하여 주문에 성공한다. 자질 정보를 사용하여 사용자 화행을 결정하고 인식되지 않거나 양식에 값이 입력되지 않아 처리할 수 없는 문장에 대해서는 안내를 통해 처리 가능한 발화를 유도하여 성공적인 처리를 수행했다.

```

SYSTEM : 안녕하세요. 한양식당입니다. 무엇을 드시겠습니까?
USER : 볶음밥을 주세요.
LEX : 볶음밥, POS :N, SEM : UNK
LEX : 을, POS :PP, SEM : NULL
LEX : 주, POS :U, SEM : GIVE
LEX : 세요, POS :E, SEM : NULL
SYSTEM : 죄송합니다. 볶음밥은 없습니다.
다시 주문하시겠습니까?
USER : 어떤 음식이 있나요?
LEX : 어떨, POS :ADJ, SEM : HOWDO
LEX : 나, POS :E, SEM : NULL
LEX : 음식, POS :N, SEM : FOOD
LEX : 이, POS :PP, SEM : NULL
LEX : 있, POS :ADJ, SEM : EXIST
LEX : 나요, POS :E, SEM : NULL
SYSTEM : 음식은 된장찌개, 김치찌개, 비빔밥, 매운탕이 있습니다.
다시 주문하시겠습니까?
USER : 된장찌개를 주세요.
LEX : 된장찌개, POS :N, SEM : STEW
LEX : 를, POS :PP, SEM : NULL
LEX : 주, POS :U, SEM : GIVE
LEX : 세요, POS :E, SEM : NULL
SYSTEM : 주문 확인해 드리겠습니다. 된장찌개 한 개 맞으신가요.
USER : 네
LEX : 네, POS :II, SEM : AGREE
SYSTEM : 잠시만 기다려 주십시오
    
```

(그림 3) 레스토랑 영역의 대화처리 일례

제시된 대화시스템은 레스토랑 영역의 서비스로봇에의 적용을 가정하였고 대화로 모든 서비스를 제공함을 전제한다. 음성인식의 오류에 대해서는 고려되지 않았으며 로봇의 터치스크린 등의 별도 장치와의 상호작용은 고려되지 않았다.

5. 결론 및 향후 과제

본 논문에서는 자질 사전과 통합 문법, 대화 흐름도를 사용한 레스토랑 영역의 자질기반(feature-based) 대화시스템을 제시하였다.

차후 연구로 문법과 자질을 활용하는 대화 생성부, 감정에 대한 고려가 필요하다. 또한 음성인식에서 발생하는 오류와 대응어, 생략어 등의 언어처리를 더 고려하여야 한다.

참고문헌

[1] 김용재 “데이터베이스 검색을 위한 한국어 대화 인터페이스 시스템의 설계” 서강대학교 석사학위논문, 1997.
 [2] 김영길 “자연언어에 의한 질의응답 시스템의 설계” 한양대학교 석사학위 논문, 1992.
 [3] 이현정 “한국어 대화체 문장의 화행 분석” 서강대학교 석사학위 논문, 1996.
 [4] 김세중, 나승훈, 이종혁 “계층적 정보유형과 문맥정보를 사용한 항공여행대화시스템에서의 예약정보 추출” 한국컴퓨터종합학술대회 논문집 Vol. 34, 2007.
 [5] 이수경, 채영문, 강건욱, 박경모, 박찬, 김근희 “효율적인 암정보 제공을 위한 온톨로지 기반 대화시스템 개발” 한국보건정보통계학회지 제30권 제2호 31~39page, 2005.
 [6] P. Nuguesm et al. “A conversational agent to navigate in virtual worlds” Proc. of the 11th Workshop on Language Technology pp.23-33, 1996.
 [7] Lynn Lambert, Sandra Caberry “A Tripartite Plan-Based Model of Dialogue” ACL’90.
 [8] Sandra Caberry “Plan Recognition in Natural

Language Dialogue” MIT Press: Cambridge, Ma, 1990.
 [9] Diane J. Litman & James F. Allen “A Plan Recognition Model for Subdialogues in Conversations” Cognitive Science 11, pp. 163-200, 1987.
 [10] 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.
 [11] 윤종원, 임성수, 조성배 “행동 네트워크를 이용한 감정형 제스처 기반 대화 관리 시스템” 정보과학회논문지 : 소프트웨어 및 응용 제37권 제10호, pp. 723-801, 2010.
 [12] 강상우, 이성욱, 서정연 “양식 채우기 대화에서 음성인식 오류의 보안을 위한 대화전략” 인지과학 제17권 제2호 pp 139~150, 2006.