

도시 관광용 영어 대화 시스템

최승권*, 권오욱*, 노유희*, 이기영*, 김영길*

*한국전자통신연구원 언어처리연구팀

e-mail : {choisk, ohwoog, yhroh, leeky, kimyk}@etri.re.kr

English Dialogue System for City Tour

Sung-Kwon Choi*, Oh-Woog Kwon*, Yoon-Hyung Roh*, Ki-Young Lee*, and Young-Gil Kim*

*Natural Language Processing Team, Electronics and Telecommunications Research Institute

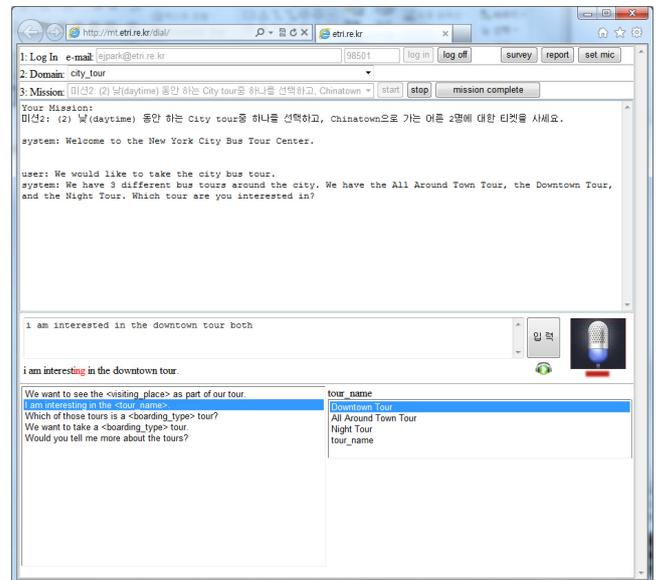
요 약

본 논문은 한국전자통신연구원(ETRI)에서 2010 년부터 2015 년까지 5 년간에 걸쳐 개발 예정인 영어 교육용 대화 시스템 중 2011 년에 개발된 관광용 영어 교육용 대화 시스템 중 도시 관광 도메인을 대상으로 한 도시 관광용 영어 대화 시스템을 소개하는 것을 목표로 한다. 도시 관광용 영어 대화 시스템은 크게 대화 이해 모듈, 대화 관리 모듈, 대화 생성 모듈, 대화 모델링 구축/관리 모듈, 대화 지식 구축 도구로 구성된다. 도시 관광용 영어 대화 시스템 평가를 위해 평가자를 초급, 중급, 고급의 3 개 그룹으로 나누어 평가를 실시하였다. 평가자는 원격 웹 평가 도구에 접속하여 4 개의 대화 미션에 대해 영어로 대화 시스템과 대화를 실시하였으며 평가는 태스크 성공률로 측정되었다. 태스크 성공률은 82.5%로 측정되었다.

1. 서론

영어 대화 시스템이란 사용자와 시스템 간에 영어로 대화를 하는 시스템을 말한다. 대화 시스템의 응용 영역은 모바일 대화형 정보 서비스, 차량용 네비게이션, 외국어 교육, TV 프로그램 가이드, 노인 의료 계획, 자동 예약 등에 이를 정도이다[1].

한국전자통신연구원(이후 ETRI)에서는 영어 학습용 응용 시스템 개발을 목표로 2010 년부터 2015 년까지 5 년간에 걸쳐 영어 교육용 대화 시스템을 개발하고 있다. 1 차년도인 2010 년에는 학습자의 의도를 이해하여 자연스런 영어 대화를 유도하는 프로토타입 수준의 영어 대화 시스템을 개발하였다. 2 차년도인 2011 년에는 1 차년도의 프로토타입 수준의 영어 대화 시스템을 여행과 관련된 4 개의 도메인인 도시 관광, 입국 심사, 호텔 예약, 호텔 체크인을 위한 웹 기반 영어 대화 시스템을 개발하였다.¹ 이 중 본 논문에서는 도시 관광과 관련하여 기술하도록 하겠다. 아래의 그림은 개발된 다중 사용자용 웹 기반 영어 음성 대화 시스템의 스크린샷의 예이다.



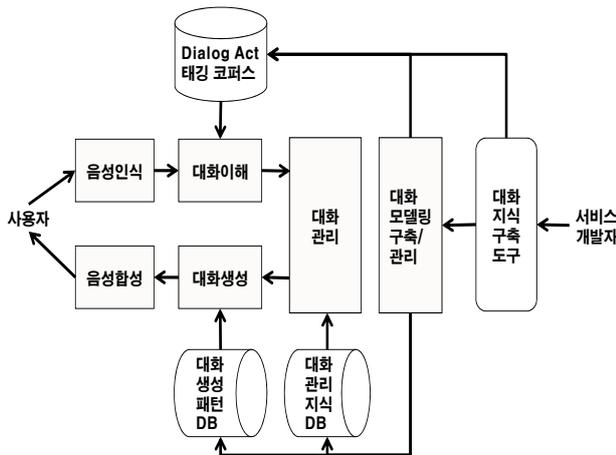
(그림 1) 웹 기반 영어 대화 시스템

2. 도시 관광용 영어 대화 시스템 구성도

도시 관광을 대상으로 한 영어 대화 시스템의 개략적인 구성도를 그림으로 나타내면 다음과 같다.

¹ 본 논문은 지식경제부의 산업융합원천기술 개발사업의 일환인 “ 모바일 플랫폼 기반 대화모델 적용 자연어 음성 인터페이스 기술 개발 ” 결과임을 밝힙니다.

공하는 세부 기능은 다음과 같다.



(그림 2) 영어 대화 시스템 구성도

본 시스템의 구조는 논리적 구조를 나타낸 것으로, 크게 사용자의 음성 발화 문장을 분석하여 이해하는 대화 이해 모듈과 사용자 발화에 알맞게 시스템이 대화를 생성하는 대화 관리 및 대화 생성 모듈로 구성된다. 대화 모델링 구축/관리 모듈은 적용 도메인에 맞는 대화 모델링을 구축 및 관리하는 모듈이다. 대화 이해와 대화 생성 및 대화 관리 모듈의 지식은 대화 지식 구축 도구와 연계되는 대화 모델링 구축/관리 모듈에 의해 구축된 Dialog Act 태깅 코퍼스, 대화 생성 패턴 DB, 대화 관리 지식 DB를 이용한다.

2.1. 대화 이해 모듈

대화 이해 모듈은 사용자의 발화를 음성 인식한 결과를 입력으로 하여 사용자 발화 의도를 생성하는 역할을 한다. 사용자 발화 의도는 언어 분석 결과를 이용해 대화 행위(Dialog Act) 형태로 표현된다. 대화 이해 모듈이 제공하는 기능은 다음과 같다.

- 음성 인식 결과를 문장 단위로 받아, 언어 분석을 수행하여 형태소 단위 분석 결과 생성
- 문장 내 등장하는 용언/기능어에 대한 정규화 표현 변환
- 대화 문장에 포함된 명사를 대상으로 의미 표현 정보 부착
- 대화 처리를 위한 NE (Named Entity) 단위 인식
- 입력된 발화에 대한 대화 의도 생성

2.2. 대화 관리 및 대화 생성 모듈

대화 관리 및 생성 모듈은 사용자 의도 분석 모듈의 결과인 사용자 의도에 대한 의미 표현(semantic representation)을 입력으로 하여, 사용자의 대화 문맥에 적합한 시스템 응답을 생성한다. 대화 관리의 도메인에 적합한 대화 계획 지식을 바탕으로 사용자 의도에 부합하여 대화를 진행하여 사용자가 원하는 목적을 달성하도록 한다. 대화 관리 및 생성 모듈이 제

- 생략/조용 처리: 대화 흐름에서 이미 말한 것에 대하여 생략하거나 조용으로 표현한 경우에 이에 대한 처리가 필요하다. 일반적으로 대화 관리 모듈에서는 도메인 슬롯(domain slot)과 슬롯값, 상태에 대하여 대화 이력(dialog history)으로 관리하여, 대화 중에 생략 또는 조용한 것에 대하여 유추하도록 설계한다.
- 사용자 요구 분석 및 평가: 대화에 따른 사용자 요구의 적합성과 완성도를 계산하여, 사용자 요구가 완전히 이해하도록 또는 애매성이 없도록 대화를 유도할 수 있게 한다.
- 오류 탐지와 복구: 음성 인식과 사용자 의도 분석의 오류에 대해 탐지하고 복구할 수 있어야 한다.
- 응용 DB 검색: 사용자 의도에 적합한 정보를 제공하기 위해서, 외부 데이터베이스로부터 도메인 정보를 가져와야 한다.
- 시스템 응답: 시스템은 사용자 의도에 맞는 대답이나 의도를 파악 또는 이해하는 대화를 진행해야 한다.

2.3. 대화 모델링 구축/관리 모듈

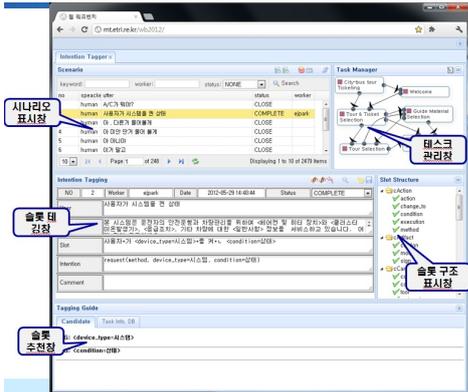
도메인 대화 모델링 구축 및 관리 모듈은 도메인에 대한 대화 지식인 대화 계획(Dialog Plan) 지식을 관리하는 기능을 제공한다.

- 대화 계획 지식: 특정 태스크를 수행하기 위해서 사용자가 시스템과 대화를 진행하는 대화 흐름에 대한 지식과 사용자의 대화에 대한 시스템의 응답 방법에 대해 기술한 것이다.
- 대화 계획 지식 관리: 대화 계획 지식을 외부로부터 로딩하고, 사용자와 대화를 진행 또는 응답하기 위해서, 미리 정의된 대화 계획 지식으로부터 태스크 진행 흐름을 검색, 태스크 진행 흐름 간의 관계 검색, 대화 응답 방법 검색 등을 수행한다

3. 대화 지식 구축 도구 및 Dialog Act 태깅 코퍼스

3.1. 대화 지식 구축 도구

대화 지식 구축 도구에서는 대화 시스템의 베이스 지식(슬롯 체계, 태스크 관리, 태스크 정보)과 Dialog Act 태깅 작업을 할 수 있다. 대화 지식 구축 도구는 다음과 같은 모습을 가진다.



(그림 3) 영어 대화 시스템의 대화 지식 구축 도구

3.2. Dialog Act 태깅 코퍼스

ETRI의 Dialog Act는 “Dialog Act Type”과 “Slot”의 조합으로 구성된다. Dialog Act Type은 알림/요청/질문과 같은 언표 내적 행위(illocutionary act)[2]에 해당하며 슬롯은 언표 내적 행위의 대상이 되는 내용을 의미한다. ETRI의 Dialog Act는 기본적으로 [3]를 토대로 구축되었다. ETRI의 Dialog Act의 BNF 형태를 기술하면 다음과 같다.

```
actset = act { "," act }
act = acttype "(" [item { "," item } ] ")"
item = bareattr | attrvalue | barevalue ;
bareattr = attr
attrvalue = attr eq value
barevalue = eq value
eq = "=" | "!="
attr = { qual "." } name
qual = name
value = string | subtype_name
```

예)

문장: How much are the tickets for the Night Tour?
Dialog Act: request(ticket_price, tour_name="Night Tour")

기존의 도메인 비종속적 Dialog Act Tag Set([4], [5])이나 도메인 종속적 Dialog Act Tag Set([6], [7], [8])에 비해 ETRI에서 도시 관광용 영어 대화 시스템에 적용하고 있는 Dialog Act Tag Set은 그 수를 제한하여 다른 도메인에도 쉽게 적용할 수 있도록 하였다. 따라서 도메인 특수성은 슬롯에 표현된다.

Dialog Act를 태깅하는 방법은 대화 코퍼스 원문에는 슬롯을 부착하고 원문 끝에 “# Dialog Act” 식으로 부착하였다. 다음의 도표는 도시 관광용 영어 대화 시스템에 사용하고 있는 Dialog Act Type의 일부이다.

<표 1> 도시 관광용 영어 대화 시스템의 Dialog Act Type

Dialog Act Type	설명
시작	hello() 또는 hello(a=x,b=y,...)
종료	bye() 또는 bye(a=x,b=y,...)

도움	repeat()	마지막 화행의 반복 요청
	help()	도움 요청
	restart()	다시 시작 요청
알림	inform(a=x,b=y,...)	a=x, b=y, 등등의 사실을 알리기
질문/요청	request(a)	a의 값(value) 요청
	request(a,b=x,...)	b=x이며 a의 값(value) 요청
	confirm(a=x,b=y,...)	a=x,b=y,..라는 확인 yes/no 대답을 요청
대답	affirm()	yes라는 단순 대답
	negate()	no라는 단순 대답
	thankyou()	감사 표현

다음의 그림은 도시 관광 대화 시나리오에 Dialog Act가 부착된 예를 보여준다.

```
S: Welcome to the <tour_center_place=New York City Bus Tour Center>. # greet() // 뉴욕 시티 버스 관광 안내소에 오신 것을 환영합니다.
U: I want to buy tickets for me and my child. # inform(task=buy_ticket) // 저와 아이 티켓을 사려고 해요.
S: What kind of tour would you like to take? # request(tour_name) // 어떤 종류의 관광을 원하세요?
U: We would like to go on a tour during the <day_type=day>. # inform(day_type="day") // 주간에 관광을 하려고 합니다.
S: We have <tour_name_num=two> <day_type=daytime> tours: the <ANA1=tour_name=Downtown Tour> and the <ANA2=tour_name=All Around Town Tour>. # inform(tour_name_num="two", day_type="daytime", tour_name=("Downtown Tour", "All Around Town Tour")) // 네, 두 가지 유형이 있습니다. 자세히 설명을 드리면, 도시 중심가 관광과 도시 전역을 관광하는 두 가지 종류가 있습니다.
U: Which tour goes to the <visiting_place=Statue of Liberty>? # request(tour_name,
```

(그림 4) 도시 관광 대화 시나리오의 Dialog Act 태깅 코퍼스 예

4. 평가

도시 관광용 영어 대화 시스템을 평가하기 위해 다음과 같은 방법으로 평가를 실시하였다.

- 평가자
 - 평가자를 초급, 중급, 고급의 3개 그룹으로 나누어 평가를 실시함 (초급은 토익 500점 이하, 중급은 토익 501점~900점 이하, 고급은 900점 이상)
- 평가 방법
 - 평가자가 원격 웹 평가 도구에 접속하여 아래의 도표와 같은 4개의 대화 미션에 대해 영어로 대화 시스템과 대화를 실시함

<표 2> 평가용 대화 미션

대화 목표: Buying New York City Bus Tour Tickets	
미션 1	(1) 어른 1명에 대한 Night tour 티켓을 사세요. (2) 5세 미만 아이 1명에 대한 Downtown tour 티켓을 사세요.
미션 2	(1) City tour 중 자유의 여신상(Statue of Liberty)으로 가는 하나를 선택하고 어른 1명에 대한 티켓을 사세요. (2) 낮(daytime) 동안 하는 City tour 중 하나를 선택

	택하고 Chinatown 으로 가는 어른 2 명에 대한 티켓을 사세요.
미션 3	(1) 밤(nighttime)에 있는 City tour 를 알아보고 어른 2 명 티켓을 구입하고 또한 한국어 가이드북(Korean tour booklet)을 얻으세요. (2) 낮(daytime)동안 하는 City tour 중에 이틀 사용권(two-day ticket)이 가능한 City tour 어른 1 장을 구매하고 1 개의 오디오 가이드(audio guide)를 구매하세요.
미션 4	자유 미션 (단, 위에서 시도하지 않은 미션을 생각해 할 것)

- 대화 미션에 대해 태스크 성공률을 측정함.
태스크 성공률은 주어진 미션에 따라 사용자가 시스템과 대화를 한 후, 주어진 미션을 성공하면 1, 실패하면 0 를 표시하여 그 합계에 의하여 나온 평가 방법을 말함. 태스크 성공률은 다음과 같은 식에 의하여 구할 수 있음

$$\text{태스크 성공률} = \frac{\text{성공한 태스크 수}}{\text{전체 태스크 수}}$$

● 평가 결과

- 평가 결과는 다음과 같았다.

미션	초급		중급		고급		평균	
	태스크 성공 (%)	대화 턴	태스크 성공 (%)	대화 턴	태스크 성공 (%)	대화 턴	태스크 성공 (%)	대화 턴
1	66.67	15	100	15.6	100	22.5	90	16.8
2	100	24.67	80	20.6	50	15	80	20.7
3	66.67	26.33	100	18.6	100	24	90	22
4	33.33	19.33	80	15.6	100	20	70	17.6
계	66.67	21.33	90	17.6	87.50	20.38	82.5	19.28

- 4 개의 미션에 대한 태스크 성공률 평균은 82.5% 였음
- 평균 대화턴 수는 19.28 이었음

5. 결론

본 논문은 한국전자통신연구원에서 2010 년부터 2015 년까지 5 년간에 걸쳐 개발 예정인 영어 교육용 대화 시스템 중 2011 년에 개발된 관광용 영어 교육용 대화 시스템 중 도시 관광 도메인을 대상으로 한 도시 관광용 영어 대화 시스템을 소개하는 것을 목표로 하였다. 본 논문에서는 도시 관광용 영어 대화 시스템을 구성하는 대화 이해 모듈, 대화 관리 모듈, 대화 생성 모듈, 대화 모델링 구축/관리 모듈, 대화지식 구축 도구를 소개하였다.

도시 관광용 영어 대화 시스템 평가를 위해 평가자를 초급, 중급, 고급의 3 개 그룹으로 나누어 평가를 실시하였다. 평가자는 원격 웹 평가 도구에 접속하여 4 개의 대화 미션에 대해 영어로 대화 시스템과 대화를 실시하였으며 평가는 태스크 성공률로 측정되었다. 태스크 성공률은 82.5%로 측정되었다.

향후 계획은 도시 관광 뿐만 아니라 입국 심사, 호텔 예약, 호텔 체크인에 대해서도 더욱 높은 태스크 성공률을 가지는 영어 대화 시스템을 만드는 것이다.

참고문헌

- [1] Jason D. Williams. "Spoken dialogue systems: challenges, and opportunities for research", In IEEE Automatic Speech Recognition and Understanding Workshop, 2009.
- [2] Searle, J. R. Speech Acts. Cambridge University Press, Cambridge, 1969.
- [3] Steve Young, Milica Gasic, Simon Keizer, Francois Mairesse, Jost Schatzmann, Blaise Thomson, Kai Yu. "The Hidden Information State Model: A practical framework for POMDP-based spoken dialogue management", Computer Speech & Language, Volume 24, pp.150-174, 2010.
- [4] Jurafsky, Daniel, Noah Coccaro, Rachel Martin, Marie Meteer, Klaus Ries, Elizabeth Shriberg, Andreas Stolcke, Paul Taylor and Carol Van Ess-Dykema. "Switchboard Discourse Language Modeling Project Final Report", In John Hopkins LVCSR Workshop-97, 1997.
- [5] 김민정, 한경수, 박재현, 송영인, 임해창. "도메인에 비종속적인 대화에서의 화행 분류", 제 18 회 한글 및 한국어 정보처리 학술대회. pp.246-253, 2006.
- [6] Wahlster, Wolfgang. editor. Verbmobil: Foundations of Speech-to-Speech Translation. Springer. 2000.
- [7] Choi, Won Seun, Jeong-Mi, and Jungyun Seo. "Analysis system of speech acts and discourse structures using maximum entropy model", In Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics, pp. 230-237, 1999.
- [8] Harry Bunt, Jan Alexandersson, Jean Carletta, Jae-Woong Choe, Alex Chengyu Fang, Koiti Hasida, Kiyong Lee, Volha Petukhova, Andrei Popescu-Belis, Laurent Romary, Claudia Soria, David Traum. "Towards an ISO standard for dialogue act annotation", In LREC' 10, pp.2548-2555, 2010.