

대화형 홈네트워크 제어 시스템

김현정⁰, 은지현, 손단영, 장두성
KT 미래기술연구소 HCI 연구담당
{hyunj⁰, jh06, dyson, dschang}@kt.co.kr

Dialogue Management System for Home Network Control

Hyunjeong Kim, Jihyun Eun, Dan-young Son, Du-seong Chang
Human Centric Interface Research Department
Advanced Technology Laboratory, KT

요 약

현재 서비스되고 있는 홈네트워크는 가전기기들에 대한 홈오토메이션 서비스로서 웹이나 TV 등을 통한 메뉴선택 방식으로 제어되고 있다. 본 논문에서는 홈네트워크에 SMS를 통한 대화형 인터페이스를 도입함으로써 시간과 장소의 제약이 없는, 가전기기 제어를 가능하게 하였다. 대화형 인터페이스의 구현을 위해 홈네트워크 영역의 요구분석을 통하여 23개의 화행과 22개의 의미성분으로 이루어진 의미구조를 수립하였으며, WOZ(Wizard-of-OZ)방식으로 9만 6천 대화쌍을 수집하였다. 대화 처리를 위하여 예제기반 대화추론 및 응답생성 방법을 사용하였으며, SMS 영역에서의 성능향상을 위하여 전처리를 도입하였다. 구현된 서비스에 대한 성능평가 실험을 통해 사용자 만족도를 조사하였다.

1. 서 론

유무선 네트워크를 기반으로 가정내의 가전기기들을 제어하는 홈네트워크는 사용하는 단말, 네트워크, 사용자 인터페이스를 다양하게 제공하여 사용자의 편의를 극대화하고 시간, 장소에 관계없이 스마트한 서비스를 제공하도록 진보하고 있다. 이 논문에서는 홈네트워크 사용자에게 다양한 입출력 방식을 제공하고, 편리한 가전기기 제어와 재미를 더한 서비스를 제공하기 위해 자연어를 이용한 대화관리시스템을 제안한다. 대화관리시스템은 대화 제어 방법에 따라 유한 상태 기반 (Finite State-based) 대화관리시스템, 프레임 기반 (Frame-based) 대화관리시스템[1], 정보상태 기반(Information state-based) 대화관리시스템[2], 예제기반 대화관리시스템[3][4] 등이 있다. 본 논문에서는 자연어로 홈네트워크를 제어하기 위해 해석된 의미구조와 대화예제를 기반으로 대화추론 및 응답생성을 하는 대화시스템을 구축하였다. 이를 위해 사용자 요구분석을 하였으며, 분석결과로 홈네트워크 제어에 필요한 의미구조를 설계하고 이를 기반으로 홈네트워크 제어 영역과 일반 채팅영역의 대화를 수집하였다. 이렇게 수집된 대화는 대화관리시스템의 언어이해 모듈과 대화추론 및 응답생성모듈에서 사용되었다. 본 논문은 다음과 같이 구성되어 있다. 2장에서는 홈네트워크 제어를 위한 요구분석 결과로

설계된 의미 구조, 시스템 응답을 소개하고, 3장에서는 대화를 수집하기 위한 환경구축과 대화수집 방법론을 설명한다. 4장에서는 제안된 대화관리시스템의 대화처리 절차를 설명하고 5장에서는 대화관리시스템을 사용자 인터페이스로 사용하여, 입출력 방식으로 SMS를 이용한 홈네트워크 제어시스템에 대해 소개하고 6장과 7장에서는 대화관리시스템의 성능평가 결과와 향후 연구방향에 대해 설명한다.

2. 요구분석 및 의미프레임 설계

KT에서 현재 서비스중인 홈네트워크[5]는 가전기기들에 대한 홈오토메이션 서비스로써 음성인식 ARS 방식과 웹, TV, 휴대폰을 이용한 메뉴선택 방식으로 제어되고 있다. 이러한 제어방식은 새로운 가전기기의 추가나 제어방법의 추가 요청에 유연하게 대처하기 어렵고, 다양한 입출력 방식을 제공하지 못하고 있다. 본 연구에서는 기존의 메뉴선택 방식의 단점을 개선하고 사용자에게 편리한 서비스를 제공하기 위해 자연어 대화관리시스템을 홈네트워크 제어를 위한 사용자 인터페이스로 제시하고자 한다. 홈네트워크 서비스에 대한 요구분석결과를 바탕으로 제어대상 가전기기를 선정하고 가전기기의 상태확인과 기기제어를 목표로 설정하였고 문형은 단문 형태의 대화를 하는 것으로 설정하였다[6].

2.1 의미프레임 설계

요구분석 단계를 위해 WOZ 환경에서 시험적으로 수집된 30개의 대화를 분석하여 발화에 표현된 일반적인 의도를 8개로 분류하였고, 홈네트워크 영역에 한정적인 사용자의 목적의도를 15개로 분류하였다. 기기제어를 위한 의미성분은 22개로 정의하였고, 사용자의 발화에 대한 시스템의 응답은 16개의 행위로 정의하였다. 의미 구조 설계는 향후 가전기기의 추가나 제어 기능의 추가 시에 쉽게 대화를 추가하여 훈련할 수 있도록 기기의 특성에 따라 그룹별로 설계하였고 시스템의 응답은 홈네트워크 제어영역에만 한정적으로 사용되지 않고 범용으로 사용할 수 있도록 설계되었다. 의미성분 설계에서는 홈네트워크 제어를 위해 예약 시각 설정을 할 수 있도록 시간 및 날짜에 관한 의미성분을 절대시각, 상대시각으로 구분하여 설계하였으며, 온도 제어도 '1도 높게', '10도 낮게' 등으로 자연스러운 발화를 허용하도록 설계하였다.

<표1> 화행(Speech Act) 분류

일반의도	목적의도
REQUEST	QUERY(온도/모드/시간)
WH_QUESTION	TURN_ON/OFF
YN_QUESTION	OPEN/CLOSE
ASSERT	RESET(온도/모드/시간)
AGREE	CANCEL
REJECT	ALARM
EXPRESS	HELP
OTHER	

<표2> 의미성분 분류

의미성분	설명
OBJ	가전기기의 명칭
LOC	가전기기의 지정위치
PERFORMER	확인대상
PATIENT	알람대상
TIME_START_ABS	절대시작시각
TIME_START_REL	상대시작시각
TIME_END_ABS	절대완료시각
TIME_END_REL	상대완료시각
TIME_REPEAT	반복간격
TIME_DURATION	시간
DATE_START_ABS	절대시작날짜
DATE_START_REL	상대시작날짜
DATE_END_ABS	절대완료날짜
DATE_END_REL	상대완료날짜
DATE_REPEAT	반복간격
DATE_DURATION	반복기간
...	...

<표3> 시스템의 행위(System Act) 분류

동작명	의미
GREET/BYE	대화시작/종료
SPECIFY(R/D)	제어속성을 요구
RELAX(R/D)	완화속성을 요구
CONFIRM(R/D/Q)	사용자의 의도를 확인
REJECT(1/T)	이전발화를 오류로 인정
RESULT	결과안내
EXECUTE	제어시스템 실행요청
ERROR	시스템 제어 오류
FAIL	시스템 제어 실패
CANCEL	발화 취소

3. 대화수집 방법

언어이해 모듈과 대화추론 및 응답생성용 예제 대화 수집을 위해 실제 사용자와 시스템간의 대화를 가정하고 사용자 역할 수집자와 시스템 역할 수집자를 선정하여 홈네트워크 제어영역의 대화 4만6천개와 채팅을 위한 대화 5만개를 수집하였다.

3.1 홈네트워크 제어 영역의 대화수집

홈네트워크 제어에 대한 대화를 수집하면서 최대한 자유 발화도를 가진 대화를 수집하기 위해 사용자 역할 수집자에게 영역 내의 기기와 제어방법에 대한 사전 지식만을 가지고 자유롭게 홈네트워크 시스템과 대화하는 자유발화 세션과 대화초기에 가전기기 제어에 필요한 의미성분 몇 개를 제시하고 사용자가 자유롭게 이 의미성분을 채워나가도록 하는 의미성분 제시형 세션으로 구분하여 WOZ 방식으로 대화를 수집하였다. 시스템 역할 수집자는 영역 지식을 참고하여 사용자의 발화에 대해 적절한 응답을 하도록 하였다. 사용자와 시스템간의 상호주도형 대화를 수집하기 위해 WOZ 환경을 지원하는 워크벤치를 개발하여 사용하였다[6].

3.2 대화 분석

수집된 대화를 언어이해를 위한 훈련용 말뭉치로 사용하고 대화 추론 및 응답 생성용 예제 DB로 사용하기 위해 2장에서 설계된 의미구조 성분을 <표 4>와 같이 대화에 부착하였다.

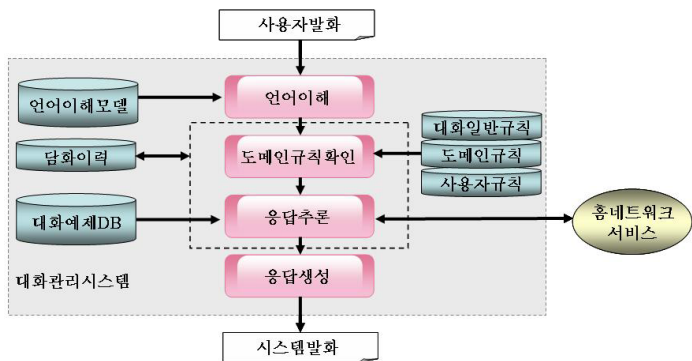
<표 4> 의미구조 성분 부착 예

사용자발화 : 안방의 전등이 켜져 있어? Speech Act = REQUEST/QUERY LOC = 안방 OBJ = 전등
--

System Act = RESULT
 시스템응답 : 안방의 전등은 꺼진상태입니다.
 발화 템플릿 : \$LOC의 \$OBJ은 \$SYSTEM_STATE입니다.

4. 대화관리시스템

본 연구의 대화관리시스템은 언어이해 모듈, 대화추론 모듈, 응답생성 모듈로 구성되어 있다.



<그림 1> 대화관리시스템 구조

4.1 언어이해 모듈

3장에서 수집된 대화DB를 언어이해를 위한 훈련용 말뭉치로 구축하여, 사용자의 입력을 품사분석 하고 분류(classification) 기법을 통하여 사용자의 의도, 문형, 의미성분 등으로 구성된 의미구조를 추출하였다. 각 의미구조를 추출하기 위해 적용한 분류 방식은 조건부 랜덤필드(Conditional Random Field)[7]이다. 추출된 의미구조는 어휘 정규화를 통해 가전기명이나 장소명 등을 표준 명칭으로 대체하고, 상대시간, 상대온도 등을 절대 시간, 절대 온도로 변환하여 대화처리에 사용될 수 있도록 하였다. 어휘 정규화는 음절 단위로 입력 어휘와 표준 어휘 간의 레벤스타인 편집 거리 (Levenshtein edit distance)[8]를 측정함으로써 수행하였다.

4.2 대화추론 모듈

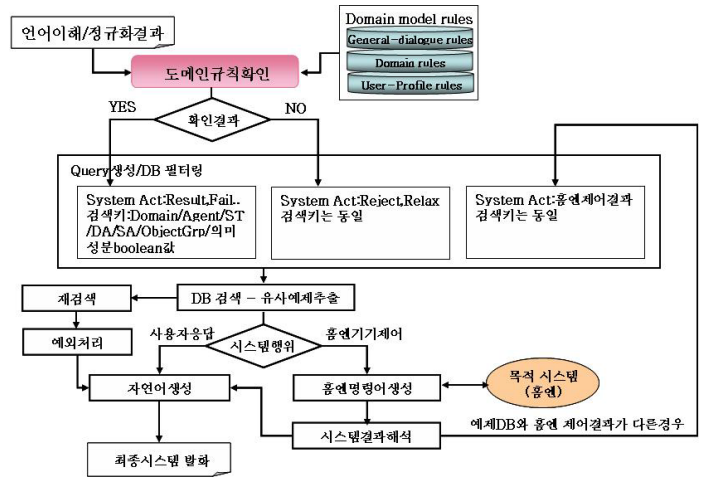
4.2.1 도메인 규칙 확인

사용자의 대화를 추론하여 시스템의 응답을 찾기 위해서 예제기반 대화추론 방법을 사용하며, 예제DB 검색 전 단계에서 여러 턴에 걸친 사용자와의 대화 기록을 누적한 의미구조를 가지고 일반적인 대화규칙이나, 영역 공통 규칙, 사용자 프로파일 규칙에 의한 제약 사항을 먼저 확인하고 그 결과를 대화추론에 반영함으로써, 검색의 크기를 줄이고 정확도를 높였다. 영역 공통 규칙에는 특정 기기를 기준으로 적용 가능한

목적의도, 의미 성분, 특정 목적의도에 따라서 적용 가능한 의미성분 등에 대해 기술된다. 사용자 프로파일 규칙은 해당 세대에서 특정 기기가 위치한 장소, 사용자가 특정 기기에 대해서 특별히 허용하지 않고자 하는 목적 의도, 사용자가 설정한 의미 성분의 기본값 등에 대해 기술된다.

4.2.2 응답추론

사용자 발화의 분석된 결과와 구축된 예제DB와의 유사도를 계산하여 시스템의 응답을 추론하였다. 추론된 시스템의 행위가 홈네트워크 시스템에 대하여, 조회, 제어 등의 실행인 경우에는 어플리케이션 서버로 질의 또는 제어 명령을 보내고, 추론된 시스템의 행위가 사용자에 대한 동작인 경우는 응답 생성 모듈을 통하여 사용자에게 추가 정보를 요청하거나, 속성의 완화를 요청하거나 사용자의 의도를 확인하도록 한다. 예제DB 검색 결과가 없을 때는 검색 조건을 완화하여 누적된 의미슬롯의 벡터 값과 사용자의 의도로 예제DB를 검색하고 완화된 검색에서도 결과가 없을 때는 예외처리 모듈에서 예외처리 응답 템플릿을 최종 결과로 선출하게 된다. 이와 같은 대화추론 방법은 아래 <그림 2>에서 보여주고 있다.



<그림 2> 대화추론 방법

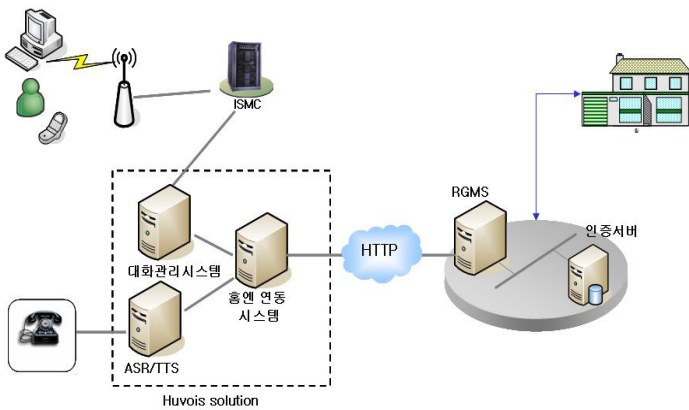
4.3 응답생성모듈

대화추론 모듈로부터 추론된 시스템의 행위가 사용자에 대한 동작인 경우는 시스템의 발화 템플릿을 이용하여 사용자가 발화한 실제 의미성분 값으로 변수를 치환하여 응답을 생성하며, 홈네트워크 시스템에 대한 동작인 경우, 질의나 제어 후에 시스템으로부터 제어결과 또는 질의결과를 포함하여 <표 4>와 같이 시스템의 발화 템플릿을 이용하여 응답을 생성한다. 시스템의 응답을 시스템의 행위와 발화 템플릿 형태로 분석하여 예제DB를 구축하고 시스템의 응답을 발화

템플릿을 이용하여 생성함으로써, 사용자 발화의 실제 값에 따라 상황에 맞는 적절한 응답발화를 생성할 수 있다.

5. 대화형 홈네트워크 제어 시스템 개발

본 연구에서 개발된 대화형 홈네트워크 제어 시스템은 <그림 3>과 같이 사용자로부터 SMS 단문 메시지 형태로 자연어 문장을 입력받아 대화처리를 하는 대화관리시스템과 홈네트워크 제어를 위한 홈엔 연동시스템, 가정내의 홈네트워크를 제어하는 게이트웨이 시스템과 인증서버로 구성되어 있으며 기존의 메뉴선택 방식으로 제어하는 음성인식 ARS 시스템과 함께 전체 홈네트워크 제어 시스템을 구성하고 있다.



<그림 3> 홈네트워크 제어 시스템 구성

5.1 SMS문장 전처리

SMS를 이용하여 사용자와 시스템간의 대화를 진행하기 위해, 대화관리시스템의 전처리 단계를 도입하였다. 여기에서는 SMS 영역에서 빈번하게 발생하는 변형된 문자열의 인식과 이모티콘 처리, 정규화 및 띄어쓰기 등의 전처리를 하여 문법에 맞는 문자열로 변화하여, 형태소 분석 오류를 최소화 한다. 통신언어 어휘집 및 구축된 대화DB에서 빈도가 높은 변형 문자열 등의 통계 데이터를 추출하여 문자열 변형 사전, 문법 형태소 오류, 음소 단위 어미 오류, 어절단위 오류 교정 사전을 추출하였으며, 띄어쓰기 교정 기능을 같이 사용하여 SMS영역에서 가능한 다양한 형태의 사용자 입력에 대응하고 있다.

6. 성능평가

6.1 언어이해모듈

언어이해모듈의 성능은 의미구조를 기준으로 정확률(precision), 재현률(recall), 이들의 조화평균인

F1 으로 평가한다. 3장에서 수집된 대화 말뭉치 중 4000개로 훈련을 하였고, 200개로 테스트를 하였다. 홈엔 영역의 언어이해 성능은 <표 5>와 같이 평가되었다<표 5>.

<표 5> 언어이해모듈 성능

	분류개수	정확률	재현률	F1
일반의도	8	99.02	99.02	99.02
목적의도	15	89.27	89.27	89.27
의미성분	22	90.44	90.69	90.57

6.2 응답추론 모듈

앞서 설명한대로 사용자의 발화가 들어왔을 때 언어이해 모듈에서 추출된 일반의도, 목적의도, 의미성분 등의 의미구조와 누적된 의미프레임의 벡터값으로 예제DB와 유사도 계산 결과를 검색했을 때의 성능을 측정했다. 평가 척도로는 일치도와 정확도를 사용했는데, 일치도는 정답 말뭉치의 응답과 시스템의 응답 추론이 일치할 경우를, 의미 정확도는 응답추론이 일치하지는 않지만, 의미상 납득 가능한 경우를 포함하여 평가하였다<표 6>.

<표 6> 응답추론 모듈 성능

	일치도(Acc.,%)	의미정확도(Acc.,%)
응답추론	85	91

6.3 사용자 성능평가

사용자 성능평가를 위해서 10명의 실험자에게 자연어를 이용하여 가전기기를 자유발화를 통해 제어하도록 하였고, 사용자당 5개의 대화를 하도록 하고 한 개의 대화 내에는 하나 이상의 작업을 수행하도록 하여 요구 분석 시에 수집하였던 대화와 비교하였다<표 7>.

<표 7> 대화당 발화수와 어절수

	요구분석시	사용자성능평가
대화개수	30	50
평균 발화(turn)수	1.83	2.76
평균 어절수	3.26	2.38

사용자 분석 결과 실제 시스템을 이용하여 하나의 대화 내에 원하는 작업을 전부 완료한 작업 완료도는 82%이었으며, 실제 이러한 시스템을 구매할 의사가 있느냐는 질문에 대한 구매 의향도는 5점 만점에 평균 3.8점의 점수를 받았다.<표 8>

<표 8> 사용자 성능평가

실험 결과		
작업 완료	41 / 50 (82%)	
작업 미완료	응용시스템 오류	2 / 50
	대화시스템 오류	5 / 50
	사용자 중도 포기	2 / 50
만족도	3.8 / 5	

7. 결론

본 연구에서는 홈네트워크 제어를 위한 대화관리시스템 개발을 위해 WOZ 방식으로 대화를 수집하여 요구분석을 하였으며, 요구분석 결과 설계된 의미 프레임과 사용자의 의도 등을 기반으로 대화를 수집하여, 언어이해모듈, 대화추론 및 응답생성모듈에서 예제 DB로 사용하였다. 홈네트워크 서비스를 고려하여 도메인 규칙 확인 모듈이 포함된 대화추론 모듈을 제안하였으며, 시스템의 발화 템플릿을 이용하여 시스템의 응답을 사용자의 발화 상황에 맞게 생성하는 응답생성 모듈을 구현하였다. 이 대화관리시스템은 실제 운용중인 홈네트워크 제어시스템 연동모듈과 SMS 송수신 모듈에 연결되어 사용자로부터 SMS를 통해 제어명령을 자연어로 입력받아 홈네트워크 제어 명령으로 생성하고 그 결과를 응답생성 모듈에 포함하여 사용자에게 시스템의 응답을 전달하는 전체 시스템에 포함되어 구현되었다. 개발된 시스템은 언어이해부, 응답추론부를 개별적으로 성능평가를 실시하였으며, 사용자 성능평가 실험을 통해 만족도 평가를 시행하였다.

현재 구현된 시스템을 다중영역으로 확장하고 있으며, 통계적 대화추론 방법을 기반한 대화관리시스템[9] 등을 보완적으로 적용하여 대화관리시스템의 성능을 향상시키는 연구가 진행중이다.

참고문헌

[1] M. McTear, "Spoken Dialogue Technology", Springer, 2004

[2] S. Larsson and D. Traum, "Information State and Dialogue Management in the TRINDI Dialogue Move Engine Toolkit," Natural Language Engineering, pp. 323-340, 2000.

[3] H. MURAO, N. KAWAGUCHI, S. MATSUBARA, Y. YAMAGUCHI, Y. INAGAKI, "Example-based Spoken Dialogue System using WOZ System Log," 4th SIGdial Workshop on Discourse and Dialogue, 2003

[4] C. Lee, S. Jung, J. Eun, M. Jeong, G. Lee. "A Situation-based Dialogue Management using Dialogue Examples". In Proceedings of the 2006 IEEE international conference on acoustics. speech and signal proceeding May 2006 Toulouse

[5] KT,2006, 홈엔서비스, <http://www.homen.co.kr/>

[6] 김현정, 은지현, 장두성, 최준기, 구명완, "홈네트워크 제어를 위한 대화관리시스템 설계" 대한음성과학회 추계학술대회 논문집, 2006.

[7] J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," ICML, 2001.

[8] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions and reversals". Doklady Akademii Nauk SSSR 163(4) p845-848, 1965.

[9] 은지현, 최준기, 장두성, 김현정, 구명완, "마르코프 의사결정 과정에 기반한 대화 관리 시스템," HCI 학술대회 논문집, 2007.