

한국어 모바일 대화형 에이전트 시스템

홍금원*, 이연수*, 김민정*, 이승욱*, 이주영*, 임해창**

A Korean Mobile Conversational Agent System

Gumwon Hong*, Yeon-Soo Lee*, Min-Jeoung Kim*, Seung-Wook Lee*, Joo-Young Lee*,
Hae-Chang Rim**

요약

본 논문에서는 한국어 정보처리 기술을 사용한 모바일 환경의 대화형 에이전트 시스템에 대해 논한다. 대화형 에이전트 시스템 구축의 목적은 인간 사용자와 시스템 에이전트간의 자연어 인터페이스를 제공하여 보다 편리한 상호작용을 가능하게 하는 데 있다. 모바일 환경의 대화형 에이전트를 구축하기 위해서는 구어체 발화에 특화된 다양한 언어 처리 및 언어 이해 요소들이 필요하다. 본 시스템은 입력 문장의 오류처리, 형태소 분석 및 품사 태깅, 양태 분석, 논항 인식 및 의미프레임 생성, 그리고 유사 발화 검색 및 응답 생성으로 구성된다. 주어진 사용자 발화에 적절한 응답을 생성하기 위해서 본 시스템은 사용자 발화와 예제 발화 간의 어휘적, 통사/구문적, 의미적 유사도 정보를 활용하여 예제기반 응답 검색을 수행한다.

Abstract

This paper presents a Korean conversational agent system in a mobile environment using natural language processing techniques. The aim of a conversational agent in mobile environment is to provide natural language interface and enable more natural interaction between a human and an agent. Constructing such an agent, it is required to develop various natural language understanding components and effective utterance generation methods. To understand spoken style utterance, we perform morphosyntactic analysis, shallow semantic analysis including modality classification and predicate argument structure analysis, and to generate a system utterance, we perform example based search which considers lexical similarity, syntactic similarity and semantic similarity.

▶ Keyword : 대화형 에이전트(Conversational Agent), 챗봇(Chatbot), 자연어처리(Natural Language Processing)

• 제1저자 : 홍금원

• 접수일 : 2008. 9. 25, 심사일 : 2008. 10. 16, 심사완료일 : 2008. 11. 26.

* 고려대학교 컴퓨터·전파통신공학과 박사과정 **고려대학교 컴퓨터·전파통신공학과 교수

※이 연구에 참여한 연구자(의 일부)는 '2단계 BK21사업의'의 지원비를 받았음.

I. 서론

대화형 에이전트(conversational agent)는 인간 사용자와 시스템 에이전트 사이에 정보를 주고받는 시스템으로, 자연어(natural language) 인터페이스를 통해 자연스러운 대화를 수행하는 목적을 가지고 있다. 이러한 대화형 에이전트가 사용되는 애플리케이션은 주위에서 쉽게 찾아볼 수 있다. 예를 들어 마이크로소프트사의 제품군(윈도우즈, 오피스 등)에서는 MS Agent라는 기술을 사용하여 대화형 에이전트를 구현하고 있는데, 일종의 캐릭터 에이전트가 사용자가 궁금해하는 사항에 대한 답변을 검색해서 제공해 주는 기능을 수행하고 있다. 최근 웹의 발달과 아울러 이러한 대화형 에이전트는 항공권 구매, 고객 서비스, 기술 지원, 게임 및 엔터테인먼트 등 여러 분야에서 활용이 되고 있다.

이와 같은 대화형 에이전트를 구축하는데 필요한 핵심 기술은 언어처리 기술이다. 에이전트는 사용자의 질문이나 발화를 이해할 수 있어야 하며 상황에 맞는 적절한 답변을 제공해 주어야 한다.

본 논문에서는 모바일 환경에서 사용자와 에이전트간의 일방적 채팅수행을 목적으로 구축된 대화형 에이전트 시스템에 대하여 설명한다. 사용자-에이전트 간에 수행하는 대화의 예는 다음 <그림 1>과 같다.



그림 1. 사용자-에이전트 간 대화의 예
Fig. 1. Chat between human and an agent

위와 같은 예에서 사용자 발화가 입력되면 에이전트는 우선 주어진 발화에 존재하는 오류를 정규화(normalization)

하고, 발화를 형태적, 구문적으로 분석해야 한다. 더 나아가, 에이전트는 발화를 문형, 양태, 논항 등 의미적으로 분석하여 적절한 행동(action)을 취해야 한다. 이때 에이전트가 취하는 행동은 애플리케이션에 따라 달라지겠지만 일반적으로 답변(response)과 웹 페이지 등의 정보 제공 및 데이터베이스에 기록(writing)등이 수행된다.

기존의 대화시스템은 주로 문어체 환경을 바탕으로 구축되었기 때문에 약어, 신조어, 도치, 생략 등 모바일 환경의 구어체 언어현상에 관한 처리가 고려되지 않았다. 또한 휴대폰 등 모바일 장치의 입력 환경은 데스크탑에 사용되는 입력장치에 비해 많이 제한되어 있기 때문에, 문장의 길이가 상대적으로 짧고, 조사와 마침표의 생략, 띄어쓰기 오류 및 철자 오류 등 언어처리 시 고려해야 할 어려운 점들을 포함하고 있다. 본 논문은 이러한 모바일 환경의 구어체 언어처리에 중점을 두고 대화형 에이전트 시스템을 구축을 위한 요소들을 살펴본다.

본 논문의 2장에서는 대화형 에이전트와 관련한 기존 시스템들에 대해 살펴보고, 3장에서는 제안하는 시스템의 구성 및 각 하위 모듈들에 대해 설명한다. 4장에서는 제안된 시스템의 구현 및 성능 평가 결과를 제시하고 5장에서 결론을 내린다.

II. 관련연구

최초의 대화형 에이전트는 1966년 Joseph Weizenbaum 이 개발한 ELIZA로서, 환자의 심리치료를 목적으로 사용되었다(1). 이후 편집증 환자의 치료를 목적으로 1973년 Colby에 의해 Parry (2)가 개발되었고, 최근에는 웹의 발달과 더불어 MSN messenger에 등록 가능한 챗봇의 개수만해도 60여개를 넘는다(3).

이러한 대화형 에이전트 기술들은 음성합성, 음성인식 기술과 접목되어 최근 각광받는 HUBO, ASIMO등 인간형 로봇들에 설치되어 인간-로봇 간의 다양한 양상의 인터페이스로 활용이 시도되고 있다.

대화형 에이전트는 기본적으로 패턴매칭기술을 사용하여 사용자 발화에 대한 적절한 응답을 찾아서 제공하는 기술이 사용된다. 미리 설계된 질의에 대해서 정확히 일치하는 경우에만 답변을 제공하기 때문에 대화 유형이 단순하며, 얼마나 많은 양의 답변을 보유하느냐로 그 답변 성능이 결정되기 때문에 많은 양의 질의 패턴을 미리 구축해 두어야 한다. ALICE(Artificial Linguistic Internet Computer Entity)의 경우, ELIZA에서 단순 패턴 매칭 기법을 개선하여 AIML(Artificial Intelligence Mark-up Language)라는 XML을 확장한 마크업 언어를 사용한 패턴 매칭 기법을

사용한다[4,5]. AIML 패턴은 단어와 공백 문자 그리고 와일드카드 구성되며, 기본 원리는 최적, 최장일치에 기반한 패턴 매칭을 구현하고 있으며, ALICE외에도 Dawnstar, Ailysse, Alan 등 다수가 존재한다.

하지만 이와 같은 방법도 사용자 발화에 대한 패턴 매칭 수준을 향상시키긴 하지만, 발화에 대한 고도화된 언어분석을 수행하지 못하며, 형태적으로는 유사하지만 의미적으로 관련이 없는 발화를 검색해 줄 수 있다는 단점이 있다.

본 논문에서 제시된 시스템도 패턴매칭 기술에 기반을 두고 있다. 그러나 단순한 어휘수준의 패턴매칭이 아닌, 자연어를 각 단계별로 처리하고 분석하여 통사적, 구문적, 그리고 의미적인 정보들을 결합하여 주어진 발화와 유사한 발화를 검색한다. 또한, 모바일 디바이스의 입력환경을 고려한 GUI 기반의 언어 처리 및 언어 분석을 수행을 한다는 특징을 가지고 있다.

III. 한국어 모바일 대화형 에이전트 시스템

1. 시스템 구성

(그림 2)는 모바일 대화형 에이전트 시스템의 전체 구성도를 보여준다. 전체시스템은 크게 통사/구문분석 모듈, 의미분석 모듈, 그리고 예제기반 응답 생성 모듈로 구성된다. 사용자 발화가 입력되면 통사/구문 분석 모듈이 오류가 포함된 입력문장을 정규화하고, 정규화된 문장을 대상으로 형태소 분석 및 품사 부착을 수행한다. 통사/구문 분석모듈의 자세한 설명은 2.2절에서 다룬다. 다음 단계인 의미 분석 모듈은 품사 부착된 사용자 발화로부터 주요 키워드를 추출하고 발화의 문형, 시제 및 긍/부정을 인식한다. 의미 분석모듈의 하위 모듈인 양태 분석 모듈에서는 입력 발화에 담긴 화자의 심리적 태도, 즉 양태를 정해진 분류 체계로 분류한다. 양태 분석과는 별도로, 논향 인식 모듈에서는 입력 문장의 슬어와 각 논향들을 인식한다. 구체적인 의미 분석 방법 및 각 분류 체계 등은 2.3절에서 논한다. 마지막으로, 예제기반 응답 생성 모듈에서는 사용자 입력 발화의 어휘, 품사, 구문 및 의미정보를 활용하여 유사도 기반 유사 발화 검색을 수행하며 검색된 발화에 대응하는 적절한 응답 발화를 출력한다. 응답생성 모듈에 관해서는 2.4절에서 설명한다.

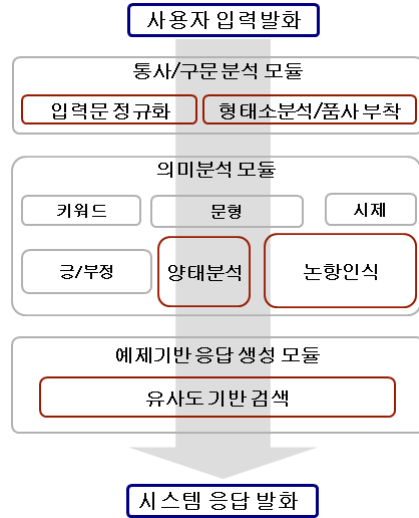


그림 2. 모바일 대화형 에이전트의 전체 시스템 구조도
Fig. 2. System architecture of mobile conversational agent

(그림 3)은 본 시스템에서 사용자 입력발화 “점심 먹을까”를 처리하여 응답을 출력하는 과정의 예를 보여준다.

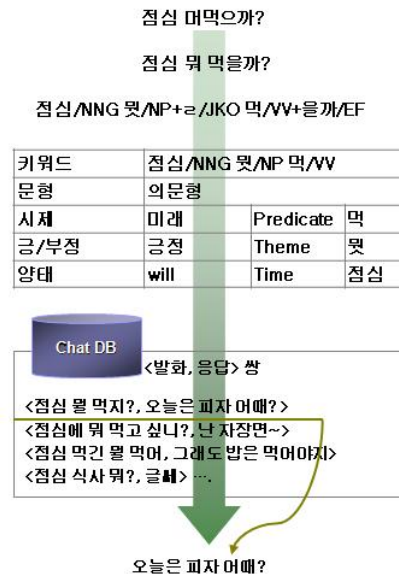


그림 3. 사용자 입력발화가 처리 되는 예
Fig 3. Example of processing user utterance

(그림 3)에서, Chat DB는 사용자발화와 응답발화의 쌍을 저장해 놓은 데이터베이스를 의미하며, 시스템은 주어진 발화

와 가장 유사한 “점심 뭘 먹지”라는 발화를 검색하여 미리 정해진 “오늘은 피자 어때?”라는 응답을 사용자에게 제공하여 준다.

2. 통사/구문 분석 모듈

통사/구문 분석 모듈은 크게 입력문장 정규화 모듈과 형태소 분석 및 품사 부착 모듈로 구성된다.

2.1. 입력문장 정규화

입력문장 정규화는 발화에 존재하는 띄어쓰기 오류 및 철자오류의 제거, 과도한 반복어구의 제거, 분해된 음소의 조합, 이모티콘의 인식과 같은 정규화 과정을 수행함으로써 다음 단계인 형태소 분석 모듈로 이어지는 오류전파를 사전에 막는 기능을 수행한다. <그림 4>는 입력문장 정규화 과정의 예를 보여준다.



그림 4. 입력문장 정규화 과정의 흐름도
Fig. 4. Example of text normalization

본 시스템에 적용된 띄어쓰기 오류 교정 기술은 [7]에 소개된 HMM(Hidden Markov Models)기반의 띄어쓰기 모형을 이용하였다. 본 기술은 입력문장의 각 음절이 띄어쓰기 태그(1로 표시)와 붙여쓰기 태그(0으로 표시)을 가진다고 가정하며, 입력문장이 주어지면 각 음절에 띄어쓰기 태그가 부착된 <그림 5>와 같은 음절/태그 확률을 최대로 하는 태그열을 찾는다.

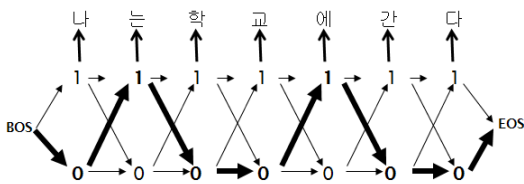


그림 5. 음절 발생 및 띄어쓰기 상태전이도
Fig. 5. State transition of word spacing

철자 오류 교정기술은 [9]을 적용하였으며, 자소단위와 음절단위 규칙을 이용하여 규칙기반 오류 교정을 수행한다. 본 기술을 사용하면 사용자 입력 발화와 교정 발화로 구성된 학습 말뭉치에서 문자열 수정 규칙을 자동으로 추출하고, 이를 바탕으로 새로운 사용자 입력 발화에 대해 문자열을 수정하게 된다. 문자열 수정 규칙은 “[좌문맥]#[수정전]#[우문맥]->[수정후]”의 형식으로 구성된다. 예를 들어, “오#바#쟁->버”의 규칙을 이용하여 “오바쟁이”를 “오버쟁이”로 교정하게 되고, “시#ㅂ#ㅇ->교”의 규칙을 이용하여 “십어서”, “십었니”, “십으니” 등을 “싫어서”, “싫었니”, “싫으니”로 교정하게 된다.

2.2. 형태소 분석 및 품사 부착

본 시스템에서는 모바일 환경의 구어체에 적합한 형태소 분석 및 품사 부착의 성능 향상을 위하여 품사 부착 말뭉치를 직접 구축하고 이를 통해 파라미터를 학습하였다. 모바일 환경의 구어체 품사부착 말뭉치 구축은 본 시스템의 실험 부분(4장)에 소개된 약 11만 사용자 발화에 대하여 정규화를 수행한 교정본을 대상으로 언어학 전공자 4명이 3개월에 걸쳐 수행하였다.

형태소 분석 및 품사 부착 기술은 [10]을 적용하였다. 본 방법의 특징은 어절 단위 분석 모형과 형태소 단위 분석 모형의 결합을 통하여 형태소 분석을 수행한다는 점이다. 만일 학습 말뭉치에 분석하고자 하는 어절이 존재한다면 어절 단위 모형을 적용하여 형태소 분석을 수행하며, 그렇지 않다면 형태소 단위 모형에서 음운 복원 후보를 모두 생성하고 생성된 후보에 대해서 형태소 분석을 수행한다. 품사 부착 과정은 위의 형태소 분석에서 생성된 모든 분석 후보들 중 최적의 한 후보만을 선택하는 과정으로 구성된다.

3. 의미 분석 모듈

의미 분석 단계에서는 품사 부착된 발화를 입력으로 우선 키워드를 추출한다. 키워드는 차후 유사발화 검색을 위한 기본 단위로서 사용이 된다. 본 논문에서는 키워드를 발화 내 중요 내용어(content word)로 정의한다. 기본적으로 품사 정보를 바탕으로 체언, 용언, 독립언은 키워드로 인식하며, 수식언, 조사, 어근, 어미, 접사, 기호 등은 제외시킨다. 하지만 경우에 따라서 부사어, 부사격 조사, 보조용언 등이 발화의 형태, 의미, 화자의 의도 등을 결정하는 주요 요소로 작용할 수 있기 때문에, 본 시스템은 제거리스트(black list)와 제거 예외리스트(white list)를 별도 사전으로 작성하여 운용한다. 예를 들어, “근데 너 이름이 뭐야”라는 문장의 경우, ‘너/NP’, ‘이름/NNG’, ‘뭐/NP’의 세 개의 키워드가 추출이 된다.

키워드 인식 외에도 의미 분석 모듈의 전체 하위모듈인 문형, 시제, 긍/부정, 양태, 논항 인식 모듈을 사용하기 위해서 반드시 필요한 것이 술어부 인식이다. 술어부 인식은 문장에서 주절을 서술하는 부분을 찾는 것인데, 술어부는 문장의 의미, 발화의 의도를 결정하는 중요한 정보가 된다. 아래 문장에서 밑줄 친 부분이 술어부에 해당한다.

“오늘 어디 놀러갈 데 없나”

술어부 인식을 위해 본 시스템은 형태소 분석 정보를 이용하며, 술어부 패턴을 표현하는 오토마타를 정의하여 술어부에 해당하는 형태소/품사열을 인식한다. 술어부가 인식되면 이 결과를 활용하여 문형, 시제, 긍/부정인식을 차례로 수행한다.

3.1. 문형 인식

본 논문에서 정의하고 있는 문형은 평서, 명령, 청유, 의문, 비종결형의 5가지이다. 문형은 문맥사전, 술어부, 종결/비종결 어미 등을 이용하여 사전과 규칙에 기반하여 결정한다. 이 중, 의문형은 다시 판정형과 설명형으로 나뉘는데, 판정형은 예/아니오의 답변을 요구하는 경우이며, 설명형은 육하원칙(六何原則)에 근거한 의문사를 포함한 의문문을 뜻한다. 따라서, “너 밥 먹었니?”는 판정형, “너 어제 어디 갔었니?”는 설명형 의문문에 해당한다.

3.2. 시제 인식

시제인식은 국문법의 문장 종결법을 반영한 규칙 기반 알고리즘을 통해 이루어지며, 문장부호, 어미열의 어휘/품사열, 문맥의 어휘/품사열 사전을 이용하여 현재, 과거, 미래 중 하나로 분류한다. 따라서 “나 밥 먹었어”는 ‘었’이라는 과거형 어미에 근거하여 과거시제로 분류한다.

3.3. 긍/부정 인식

긍/부정 인식은 원문이 긍정문인지 부정문인지를 판별하는 것이다. 긍/부정은 문장의 종결법과 특정 부사어의 사용으로 표현되며, 이를 파악하기 위하여 긍/부정 패턴 사전을 구성하고, 시제, 문형 인식과 같이 규칙기반의 인식 방법을 사용한다.

3.4. 양태 분류

양태(modality)는 일반적으로 발화에 명제에 대해 발화자와의 관련성을 의미한다[8]. 예를 들어, “나 학교에 가고 싶어”라는 발화는 화자의 ‘희망’을 나타낸다. 일반적으로 발화자의 의도를 파악하기 위해 화행(speech act)을 사용하기도 하지만, 화행은 현재 발화와 이전 발화 사이의 문맥정보가 충분한 상황에서 분류가 가능하며, 1문 1답 위주의 채팅에서는 그러한 정보가 불충분하기 때문에 본 시스템에서는 양태 정보를 사용한다. 양태 분류를 위해서 우선 24개의 체계를 정의하

였다. <표 1>은 본 시스템에서 사용한 양태 분류 체계를 수록하고 있다.

표 1. 양태 분류 체계
Table 1. Modality classes

지각, 추측, 희망, 금지, 명령, 가정, 의무, 의지, 능력, 전언, 질문, 선호, 비선호, 칭찬, 축하, 감사, 사과, 인사, 기쁨, 슬픔, 회상, 감탄, 이유, 긍정대답, 부정대답

양태 분류를 위해서 우선 입력 발화의 형태소 분석 및 품사 부착 결과를 이용하여 보조용언, 선어말어미, 어말어미, 의존명사를 추출하고, 이들을 미리 구축된 양태 규칙사전과 비교하여 누적된 스코어 값 중 가장 높은 값을 가지는 양태로 분류하게 된다. 스코어 계산을 위해 본 논문에서는 <표 2>와 같은 가중치 계산 방법을 사용한다.

표 2. 양태 분석 규칙의 예
Table 2. Example of modality classification rules

형태소/품사열	양태	가중치
누구/NP	질문	2
이/VCP+겠/EP+지/E	추측	2
*~/V+고/EC 싶/VX+어/E	희망	2

3.5. 논항 인식

논항 인식은 문장의 술어가 그 의미를 완성하기 위해 필요로 하는 문장의 각 문법적 요소들을 찾는 작업이다. 예를 들어 “나 어제 피자 먹었어”라는 발화에서 추출된 논항은 다음과 같다.

표 3. 논항 인식의 예
Table 3. Example of identified semantic roles

술어	역/VV
행위자	나/NP (1인칭 단수)
행위대상	피자/NNG (음식)
시간	어제/NNG (과거)

논항 인식은 3단계의 인식 규칙을 사용한다. 첫째, 술어인식 단계에서는 술어부 인식 결과를 이용하여 맨 뒤에서 처음 나오는 본용언을 술어로 사용한다. 둘째, 논항 후보 선택 단계에서는 논항 사전과 명사구 청킹(chunking) 사전과 논항 패턴을 바탕으로 논항 후보가 될 명사구를 후보로 추출한다. 마지막으로, 의미역 결정단계에서는 미리 정의된 격들 사전을 바탕으로

세종 말뭉치에 정의된 14가지의 의미역 카테고리중 하나로 분류를 한다. <표 4>는 본 시스템에서 사용한 논항의 종류이다.

표 4. 논항 종류
Table 4. Categories of semantic roles

시간, 행위주, 경험주, 동반주, 대상, 장소, 도착점, 결과상태, 출발점, 도구, 영향주, 기준치, 목적, 기타

4. 예제기반 응답 생성 모듈

예제기반 응답 생성 모듈은 통사/구문 분석 결과와 의미 분석 결과 도출된 정보를 바탕으로 사용자 발화와 가장 유사한 예제 발화를 검색하여 이에 대응하는 응답 발화를 출력하는 기능을 수행한다. <그림 6>는 예제기반 응답 생성 모듈의 기능적 흐름을 보여주고 있다.

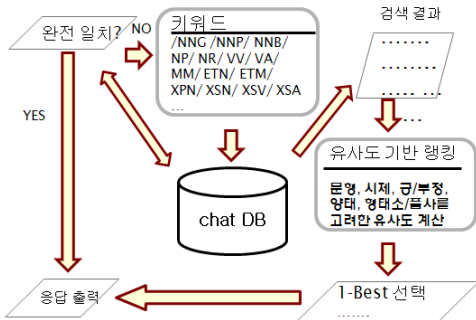


그림 6. 응답 생성의 흐름도
Fig. 6. Flow diagram of the response generation

예제 기반 응답 생성 모듈에서는 첫째, 사용자 발화와 정확히 일치(exact match)하는 발화를 chat DB에서 검색하여 이에 대응하는 응답을 출력한다. 둘째, 정확히 일치하는 발화가 없을 경우, 발화에서 추출된 키워드를 이용하여 chat DB를 검색하고 이 결과들을 유사도 측정 모델을 이용하여 순위(ranking)한 후 가장 상위의 발화를 찾아 이에 대응하는 응답을 선택한다. 유사도 측정은 [11]에서 정의한 모델을 적용하였으며 이는 수식 (3.1)과 같다.

$$\sim(x, y) = \frac{1}{z} \left[\sum_i \lambda_i f_i(x, y) \right]$$

f_1 : 문형 유사도
 f_2 : 시제 유사도
 f_3 : 긍/부정 유사도
 f_4 : 문형, 시제, 긍/부정 결합 유사도
 f_5 : 양태 유사도
 f_6 : 품사간 편집거리

..... (3.1)

여기에서 $\text{sim}(x, y)$ 는 입력 발화 x 와 후보 발화 y 간의 유사도를 의미하며 이는 6가지 세부 유사도 함수의 가중합으로 결정된다. $f_1 \sim f_5$ 는 두 발화간 각 세부 유사도 함수를 의미한다. 이 세부 유사도 함수를 구하기 위하여 다음과 같은 말뭉치 기반의 통계정보를 추출하였다. 우선 약 7,000 개의 발화 쌍을 대상으로 두 발화가 유사한 정도에 따라서 0점(완전히 다름)부터 3점(거의 동일)까지 4단계로 나누어 자연어처리 전공자 2인이 유사 발화 말뭉치를 구축하였다. 두 발화가 완전히 동일한 경우는 유사도 측정 단계 이전에 미리 출력되기 때문에 고려하지 않는다. 다음으로 2, 3점 유사도를 가진 발화 쌍과 0, 1점 유사도를 가진 발화 쌍에서 통계를 추출하여 수식 (3.2)와 같이 각 세부 함수의 유사도를 측정하였다.

$$f_1(x, y) = \frac{A2(x_{st}, y_{st}) + A3(x_{st}, y_{st})}{A0(x_{st}, y_{st}) + A1(x_{st}, y_{st})}$$

x_{st} : x 의 문형
 y_{st} : y 의 문형
 A_n : 말뭉치 내 유사도 n 점의 빈도

..... (3.2)

수식 (3.2)는 두 발화의 문형 사이에 존재하는 세부 유사도 함수 f_1 을 측정하는 수식이며, 특정 문형 조합이 0점과 1점에서 적게 나타나고 2점, 3점에서 많이 나타날수록 세부 유사도가 높다고 결정하는 방법이다. f_2 에서 f_5 까지의 세부 유사도 함수는 위와 동일한 방식으로 구하며, <표 5>는 이와 같이 구축된 유사도의 예를 보여준다.

표 5. 세부유사도 함수값의 예
Table 5. Example of individual similarity values

f_1	값	f_2	값	f_3	값
청유, 청유	3.25	미래, 미래	0.71	부정, 부정	0.63
의문, 의문	0.82	과거, 과거	0.67	긍정, 긍정	0.56
명령, 명령	0.67	현재, 현재	0.64	긍정, 부정	0.05
명령, 평서	0.21	현재, 미래	0.18		

f_6 은 사용자 입력 발화의 형태소 품사열과 후보 발화의 형태소 품사열 간의 편집 거리(edit distance)를 측정한 값이며 자세한 측정 방법은 [10]을 참고하였다.

IV. 실험 결과

1. 실험 환경

실험에 사용된 말뭉치는 실제로 휴대폰을 사용하여 채팅을 수행한 결과를 기록한 약 11만 메시지로 구성되어 있다. 본 말뭉치를 이용하여, 띄어쓰기 오류 교정, 철자 오류 교정, 형태소 분석 및 품사 부착, 양태 분석, 논항 인식 등을 위한 학습 말뭉치를 수작업으로 구축하였다. <표 6>은 통사/구문 및 의미 분석 모듈에 사용된 말뭉치 통계를 보여주고 있다.

표 6. 언어분석 모듈에 사용된 말뭉치
Table 6. Corpus used for the linguistic analysis modules

언어분석 단계	말뭉치 크기 (발화수)
원본 메시지	109,080
철자 교정	109,080
띄어쓰기 교정	109,080
형태소 분석 및 품사부착	109,080
술어부, 서술어, 문형, 시제, 긍부정 인식	2,280
양태 분석	4,650
논항 인식	1,515

응답 생성 모듈에 사용된 말뭉치는 위 말뭉치에서 단순한 응답 및 완전히 동일한 발화를 제외한 나머지 발화 쌍을 대상으로 유사도 0에서 유사도 3까지를 수작업으로 분류하여 구축하였다. <표 7>은 응답생성 모듈에 사용된 말뭉치의 통계를 보여주고 있다.

표 7. 응답 생성 모듈 실험에 사용된 말뭉치
Table 7. Corpus used for the response generation module

유사도	발화 쌍 개수 (비율 %)
유사도 0	2,365 (44.84)
유사도 1	1,022 (19.00)
유사도 2	1,027 (19.47)
유사도 3	880 (16.69)
계	5,274 (100.0)

2. 모듈 별 성능 평가

본 논문에서는 각 시스템 구성 모듈의 성능을 평가하기 위하여 다음과 같은 정확도와 재현율을 기본으로 한 측정 기준을 적용하였다.

$$\text{정확도} = \frac{\text{시스템이 출력한 결과 중 정답의 개수}}{\text{시스템이 출력한 결과 수}}$$

$$\text{재현율} = \frac{\text{시스템이 출력한 결과 중 정답의 개수}}{\text{전체 정답의 개수}}$$

$$\text{형태소분석 성공율} = \frac{\text{분석 성공}}{\text{분석 성공} + \text{분석 실패}}$$

<표 8>은 시스템을 구성하는 각 모듈단위의 성능을 표시하고 있다. 이 중, 발화단위는 발화(메시지) 단위로 개수를 세어 평가한 것이고, 어절단위는 어절을 단위로 옳고 그름을 세어 평가한 것이다. 따라서 원문의 경우 100개의 발화 중 약 7개 정도만이 각종 오류가 없는 문법적으로 올바르게 입력된 발화임을 알 수 있다.

표 8. 각 모듈의 성능
Table 8. Performance of each module

언어분석 단계	성능 (평가 방법)	
원문(오류 포함)	40.0	어절단위 정확도
	7.2	발화단위 정확도
입력문 정규화	81.7	어절단위 정확도
	73.4	발화단위 정확도
형태소 분석 및 품사 태깅	98.4	분석 성공율
	95.6	어절단위 정확도
술어부 인식	88.5	발화단위 정확도
	76.7	정확도
서술어 인식	75.1	정확도
문형 인식	79.8	정확도
시제 인식	81.7	정확도
긍부정 인식	98.0	정확도
양태 분석	82.3	정확도
	85.6	재현율
논항 인식	65.3	정확도
	61.8	재현율
응답 생성	74.0	유사 발화 검색 정확도

본 연구에서 측정된 응답 생성의 정확도는 유사도 검색 모듈의 정확도로 측정하였으며, 이는 다음과 같다.

$$\frac{\text{응답 생성 정확도} = \text{시스템 결과가 1위 발화를 맞힌 경우의 수}}{\text{시스템 검색 결과의 수}}$$

응답 생성 모듈의 정확도는 시스템 출력 결과가 사용자가 정한 1위 발화만을 맞힌 경우이며 이는 전체의 약 74%를 보이고 있다. 이 측정 방법은 다소 엄격한 방법으로 2위나 3위로 검색된 유사 발화를 검색 성공으로 간주할 경우, 그 성능은 높아질 것으로 예상된다. 또한, 본 시스템에서는 유사 발화 검색의 자질로서 논항을 제외하였는데, 이는 <표 8>에서 볼 수 있듯이, 논항인식의 정확도와 재현율이 기대만큼 충분히 높지 않아서이다. 논항인식의 성능이 높지 않은 이유는 우선, 논항 인식 자체가 언어처리에서 매우 어려운 분야라고 할 수 있으며, 또한, 고도화된 논항인식을 위해서는 정교한 구문분석과 청킹 기술이 필요한데, 본 연구에서는 효율성을 목적으로 규칙 기반의 청킹 기법을 사용하였기 때문에 복잡하고 긴 논항 인식이 실패하는 경우가 많이 발생하였다고 판단된다.

하지만, 표에서 볼 수 있듯이 오류교정 이전 원문의 정확도가 현저히 낮은 모바일 환경의 구어체 발화의 특징을 감안할 경우, 각 언어분석 모듈은 비교적 높은 수준의 분석 성능을 보인다고 할 수 있다.

V. 결론

모바일 환경에서 대화형 에이전트는 다양한 유형의 사용자 발화를 잘 분석하고 이에 적절한 응답을 출력해야 한다. 이러한 모바일 환경에서는 구어체 언어의 다양한 현상들을 잘 고려한 언어처리 기술이 요구된다. 본 논문에서는 정교한 언어처리 기술에 기반을 둔 모바일 환경의 대화형 에이전트 시스템을 제안하고 구현하였다.

통사/구문 모듈에서는 사용자 발화의 다양한 오류들을 교정하고 정규화 하여 구어체 문장의 형태소 분석 및 품사 부착 성능을 고도화 하였다. 또한, 의미 분석 모듈에서는 주어진 사용자 발화와 가장 유사한 발화의 검색에 특화된 양태 분류와 논항 인식을 제안하였다. 마지막으로 응답 발화 생성 모듈에서는 각 언어분석 모듈의 결과를 이용하여 어휘, 통사/구문, 의미적 자질을 결합한 유사도 측정 모델을 적용하였다. 이를 위하여 전체 시스템을 구성하는 각 언어 분석단계의 모듈을 각각 구현하고 실험 및 평가 하였다.

향후 연구로는 다음과 같은 네 가지 관점에서의 진행이 필요하다고 판단된다. 첫째, 유사 발화 검색 모듈과 응답 생성 모듈을 분리하여 사용자 프로파일을 고려한 개인화된 응답생성을

연구를 수행할 계획이다. 둘째, 담화 문맥(discourse history)를 고려하여 발화 매칭 및 응답 생성을 개선할 계획이다. 셋째, 현재와 같은 사용자-시스템간의 어휘적(lexical) 발화 쌍 외에 의미적(semantic) 발화 쌍을 고려한 chat DB를 구축하여, chat DB의 부족 및 매칭 실패율을 개선할 계획이다. 마지막으로, 전체 시스템 평가를 위해서 주어진 사용자 발화에 에이전트가 출력한 응답을 가지고 사용자 만족도를 평가할 수 있는 적절한 평가 척도에 대한 연구를 수행할 계획이다.

참고문헌

- [1] Joseph Weizenbaum, "ELIZA-A computer program for the study of natural language communication between man and machine", Communications of the ACM 10.8: 36-45, 1966.
- [2] Kenneth Colby, "Simulation of belief systems", Computer Models of Thought and Language, 251-286, 1973.
- [3] Chatbots.org, <http://www.chatbots.org>, 2008.
- [4] ALICE, A.L.I.C.E AI Foundation, <http://www.alicebot.org/>, 2002
- [5] Richard Wallace, "The Elements of AIML style", ALICE AI Foundation, 2003.
- [6] Min Jeong Kim, Sang-Bum Kim, Kyoung-Soo Han and Hae-Chang Rim, "Modality Analysis for Spoken Language Processing", The First Europe-Korean Workshop on Spoken Dialog System Technology, 2008.
- [7] Do-Gil Lee, Hae-Chang Rim, and Dong-Suk Yook, "Automatic word spacing using probabilistic models based on character n-gram", IEEE Intelligent Systems, vol. 22, pp. 28-35, 2007.
- [8] F.G. Garcia, "Modulating Grammar through Modality: A Discourse Approach", ELIA 1:193-136, 2000.
- [9] 변증현, "교정 말뭉치를 이용한 철자오류 교정규칙 자동 추출", 석사학위 논문, 고려대학교 2008.
- [10] 이도길, "한국어 형태소 분석과 품사 부착을 위한 확률 모형", 박사학위 논문, 고려대학교, 2005.
- [11] 이연수, "다양한 의미 자질을 고려한 발화간 유사도 측정 방법", 석사학위논문, 고려대학교 대학원, 2008.

저 자 소 개



홍금원 (Gumwon Hong)
 2000년: 고려대학교 컴퓨터학과 이학사
 2002년: 고려대학교 컴퓨터학과 이학 석사
 2007년 - 현재: 고려대학교 컴퓨터·전파통신공학과 박사과정



이연수 (Yeon-Soo Lee)
 2000년: 고려대학교 컴퓨터학과 이학사
 2008년: 고려대학교 컴퓨터학과 이학 석사
 2008년~현재: 고려대학교 컴퓨터·전파통신공학과 박사과정



김민정 (Min-Jeoung Kim)
 2005년: 고려대학교 컴퓨터학과 이학사
 2007년: 고려대학교 컴퓨터학과 이학 석사
 2007년 - 현재: 고려대학교 컴퓨터·전파통신공학과 박사과정



이승욱 (Seung-Wook Lee)
 2006년: 수원대학교 컴퓨터학과 이학사
 2008년: 고려대학교 컴퓨터학과 이학 석사
 2008년 - 현재: 고려대학교 컴퓨터·전파통신공학과 박사과정



이주영 (Joo-Young Lee)
 2003년: 고려대학교 컴퓨터학과 이학사
 2005년: 고려대학교 컴퓨터학과 이학 석사
 2005년 - 현재: 고려대학교 컴퓨터·전파통신공학과 박사과정



임해창 (Hae-Chang Rim)
 1981년: Missouri 주립대학 학사
 1983년: Missouri 주립대학 석사
 1990년: Texas 주립대학 박사
 1991년~1994년: 고려대학교 전산과학과 조교수
 1994년~1999년: 고려대학교 컴퓨터학과 부교수
 1999년~현재: 고려대학교 컴퓨터·전파통신공학과 교수