

화행별 템플릿 기반의 지식획득 기법과 유전자 프로그래밍을 이용한 문장 생성 기법을 통한 대화형 에이전트의 대화 학습*

Learning Conversation in Conversational Agent Using Knowledge Acquisition based on Speech-act Templates and Sentence Generation with Genetic Programming

임 성 수^{**}
(Sungsoo Lim)

홍 진 혁
(Jin-Hyuk Hong)

조 성 배
(Sung-Bae Cho)

요약 지능형 시스템에서 기존의 수작업 기반의 지식구조 구축은 많은 시간과 노력이 들어가며 환경의 변화에 적절히 적용하기가 어려운 한계가 있다. 이러한 한계를 극복하기 위하여 최근 학습을 통한 동적 지식구조 구축 방법이 연구되고 있다. 자율 자아 성장(AMD; Autonomous Mental Development)은 자율적 기계 학습의 새로운 패러다임으로 지능형 시스템이 변화하는 환경에 스스로 적용하도록 시도한다. 대화형 에이전트에서의 대화 학습은 AMD와 동일한 맥락에서 해석할 수 있다. 본 논문에서는 화행별 템플릿과 유전자 프로그래밍을 이용한 대화형 에이전트의 대화 학습기법을 제안한다. 제안하는 에이전트는 화행별 템플릿을 기반으로 대화 지식을 획득하고 유전자 프로그래밍의 진화 방법을 통해 적절한 표현을 갖는 문장을 생성한다. 적용 사례와 사용자 평가를 통해서 제안하는 방법의 유용성을 확인하였다.

주제어 대화형 에이전트, 지식획득, 문장생성, 화행별 템플릿, 유전자 프로그래밍

Abstracts The manual construction of the knowledge-base takes much time and effort, and it is hard to adjust intelligence systems to dynamic and flexible environment. Thus mental development in those systems has been investigated in recent years. Autonomous mental development is a new paradigm for developing autonomous machines, which are adaptive and flexible to the environment. Learning conversation, a kind of mental development, is an important aspect of conversational agents. In this paper, we propose a learning conversation method for conversational agents which uses several promising techniques; speech-act templates and genetic programming. Knowledge acquisition of conversational agents is implemented by finite state machines and templates, and dynamic sentence generation is implemented by genetic programming. Several illustrations and usability tests show the usefulness of the proposed method.

Keywords Conversational agent, knowledge acquisition, sentence generation, speech-act template, genetic programming

* 이 연구는 산업자원부가 지원하는 뇌과학 연구 프로그램에 의해 지원되었음[10023913-2005-32].

** 연세대학교 컴퓨터과학과, 연구 세부 분야: 인공지능

서울시 서대문구 신촌동 134, 연세대학교 컴퓨터과학과 소프트컴퓨팅연구실

전화: 02-2123-4803, E-mail: lss@sclab.yonsei.ac.kr

서 론

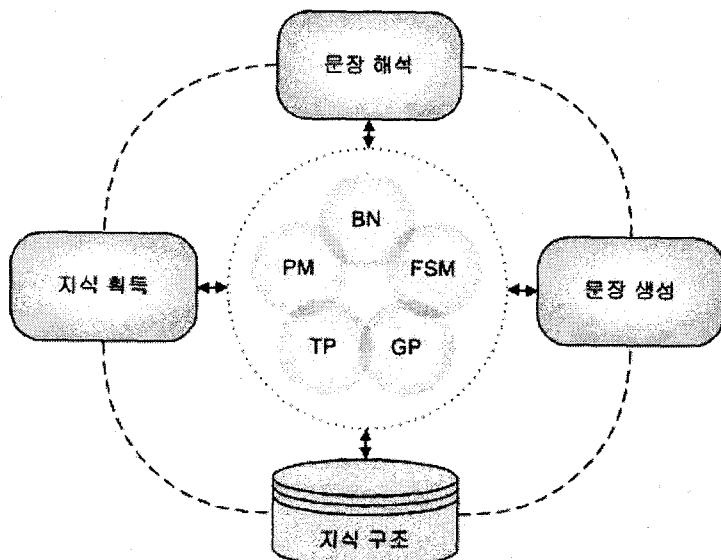
자아 성장은 인간의 중요한 특성으로 시각, 청각, 행동, 언어 등과 깊은 관련이 있다[19]. 경험이 많은 사람이 그렇지 않은 사람에 비하여 더 좋은 업무 능력을 지니는 것과 같이, 지능형 시스템에서도 많은 지식구조를 지닌 시스템이 그렇지 않은 시스템에 비하여 좋은 성능을 보인다. 하지만 수작업으로는 시스템의 모든 지식구조를 설계하는 것이 불가능하며[11, 19], 환경에 적응적인 시스템을 구축하기 어렵다. 최근 지능형 시스템의 자율 자아 성장(AMD)에 관한 연구가 진행되고 있는데, AMD는 사람과 같이 실시간으로 자신의 지식 구조를 스스로 확장하는 지능형 시스템의 새로운 설계 방법이다[9].

언어 감각은 자아 성장과 밀접하게 관련되어 있다. 어른들은 아이들에 비해 더 많은 지식과 경험을 가지고 있기 때문에, 더 세밀하-

고 자연스러운 표현의 문장을 생성할 수 있다

[3]. 지능형 에이전트의 한 분야인 대화형 에이전트는 이러한 언어 감각을 통해 사용자에게 정보를 제공하며, 에이전트의 언어 감각은 사용자 질의 패턴에 대한 답변을 표현해주는 패턴-답변 쌍을 기반으로 한다. 따라서 대화형 에이전트가 여러 상황을 고려할 수 있는 패턴-답변 쌍이 많이 가질수록 사용자에게 질 높은 정보를 제공한다[8].

대화 학습은 대화형 에이전트 설계에서 새로운 접근방법이며, 에이전트는 실시간으로 사람과의 상호작용을 통해 지식구조와 대화 표현을 학습하여 대화 능력을 향상시켜 나간다. 초기 에이전트는 “오늘 기분이 어때?”라는 질문에 적절한 대답을 하지 못하지만, 충분한 학습을 통해 “좋은데요~”와 같은 답변을 제공할 수 있다. 또한 초기의 에이전트는 “당신은 서울을 떠납니다. 그리고 당신은 부산으로 갑니다. 언제 떠나나요?”와 같이 어색한



(그림 1) 제안하는 대화형 에이전트의 대화 학습 모델

표현을 사용하지만, 자율 언어 성장을 통해 학습된 에이전트는 “당신은 언제 서울에서 부산으로 가나요?”와 같은 세련된 문장을 사용하게 된다.

본 논문에서는 (그림 1)과 같은 대화형 에이전트의 대화 학습 모델을 제안한다. 제안하는 모델은 사용자의 대화 의도 파악을 위한 문장 해석, 대화 표현을 학습하기 위한 문장 생성, 지식구조 학습을 위한 지식 획득과 대화 수행을 위한 패턴-답변 쌍을 저장하는 지식구조로 구성되며, 각 모듈을 구현하기 위해서 베이지 안 네트워크(Bayesian network; BN), 패턴 매칭(Pattern matching; PM), 유한 상태 기계(Finite state machine; FSM), 템플릿(Templates; TP), 유전자 프로그래밍(Genetic programming; GP)의 총 5개의 기술을 사용한다.

관련연구

대화형 에이전트는 대화 인터페이스를 통해 사용자는 시스템과 자연언어를 사용하여 의사

소통을 한다. 자연언어에는 생략과 축약 등의 많은 불확실성이 존재하기 때문에, 이를 해결하기 위한 많은 방법들이 연구되고 있다. ALICE는 패턴 매칭을 이용하여 불확실성을 해결하고자 하였으며, Nuance(AT&T)의 유한상태 기계, Galaxy systems의 폼 필링(Form filling), Trips의 계획기반 접근방법 등을 이용하여 대화를 처리하였다[5].

대화형 에이전트는 이러한 기술을 기반으로 웹 페이지 소개 가이드, 날씨정보 알림이, 여행 가이드 등을 위해서 적용되었다[20]. 그리고 상업적인 제품으로는 NativeMinds 사의 Nicole, Artificial Lift 사의 SmartBot, Virtual Personalities 사의 Verbot 등이 있다.

대화 학습

언어 습득은 사람의 자아 성장에 큰 역할을 한다. 단어는 여러 개의 범주를 형성하고 각 개체들을 구별하는 특징이 있다. 아이들은 자신이 알고 있는 것들을 단어를 이용하여 체계적으로 정리하며 복잡한 생각을 구별하고 유

<표 1> 언어지식 획득 과정

단계	특징
비언어(Pre-linguistic) 단계	언어의 특성을 갖추지 못함
일어문(Holophrastic) 단계	어른들이 사용하는 문장이나 구를 하나의 단어를 사용하여 표현 예) 맘마, 쉬 등
이어문(Two-word) 단계	어른들이 사용하는 문장이나 구를 두 개의 단어를 사용하여 표현 예) 더 쥐, 쉬 마려 등
간결표현(Telegraphic) 단계	중요한 의미를 가진 단어들을 조합하여 의사를 전달
중간 발달(Intermediate development) 단계	진행형, 의문형, 수동형, 과거형, 부정(否定) 변형 등을 구사할 수 있게됨
완성(Adult) 단계	완벽한 언어 구사

추한다. 뿐만 아니라 경험으로부터 배우고, 언어를 사용하여 배운 것들을 재구성한다. 따라서 인지와 언어의 상호작용을 통하여 자아 성장이 이루어진다.

<표 1>에서 보는 바와 같이 언어지식 획득 과정은 일반적으로 6개의 단계로 이루어진다고 알려져 있다. 유아는 단지 우는 것만을 하지만, 좀더 자란 아이들은 발달된 언어 감각을 통하여 보다 세련된 표현을 사용한 문장을 생성한다. 자신의 어휘와 문법이 증가함에 따라 사람은 상황에 보다 적합한 문장을 생성한다[15].

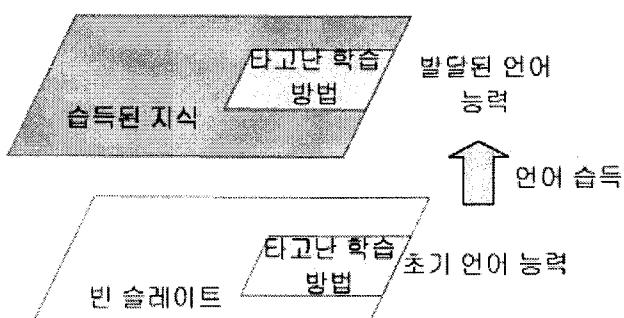
언어지식 획득 과정에서 언어 습득은 복잡한 학습 과정 중의 하나이다. 일반적으로 제약적 언어 학습이 아이들의 언어 습득 방법이라고 받아들여져 왔는데, 최근 인지 과학이 발전함에 따라 다른 의견들이 제안되고 있다. 언어 습득에 대한 초기 연구는 유아가 유전적으로 언어에 관련된 학습방법을 지니고 태어난다고 보았다. 따라서 언어의 습득은 이러한 언어에 관련된 변수들의 값을 찾아가는 일련의 과정으로 간주되었다[6]. 이와 달리 최근 연구에서는 언어 습득을 (그림 2)와 같이 타고난 학습방법을 가진 빈 슬레이트(blank slate)로 보고 있다. 유아는 타고난 학습방법을 사용하

여 환경과의 상호작용을 통해서 언어 기능과 지식을 습득한다[6, 19].

학습기법에 관한 연구는 언어 습득 연구자들에게 뇌에서 어떻게 언어를 사용하게 되는지에 대한 중요한 단서를 제공하며, 대화형 인터페이스를 이용하는 지능형 시스템의 사용성을 증대시킨다. Lauria 등은 명령어에 기반한 맞춤형 로봇 학습기법을 제안하였고[11], Dominey와 Boucher는 언어 학습을 위한 시스템을 개발하였다[4]. Green과 Eklundh는 쉬운 학습과 정보전달 방법으로 인간과 로봇간의 대화를 분석하였으며[7], Joshi와 Weng은 AMD의 개념을 소개하고 강화 학습 환경에서의 자율적인 음성 생성의 학습방법을 연구하였다[9].

자연언어 생성

자연언어 생성 분야에서는 컴퓨터 프로그램이 정보의 내부 표현 방식으로부터 어떻게 자연언어 텍스트를 생성할 수 있는가를 연구한다. 최근 사람과 컴퓨터 사이의 대화 인터페이스가 보급됨에 따라 다양한 문장 생성을 통한 자연스러운 답변 제공의 방법이 주목받고 있다.



(그림 2) 사람의 언어 학습방법

전통적인 자연언어 생성 연구에서는 주로 생성문법 기반 방식 혹은 템플릿 기반 방식을 통해 문장을 생성하였다[13]. 일반적으로 잘 작성된 생성문법 규칙은 다양한 도메인에 적용 가능성과 재사용성, 도메인 독립성을 제공 하지만, 각 단어의 속성, 분류, 정의 등 많은 부분을 설계자가 디자인해야 하여 많은 시간과 노력이 필요하다. 반면 템플릿 기반 방식이 짧은 시간에 높은 성능을 내는 시스템을 구축하기가 쉽기 때문에, 많은 연구에서는 템플릿 기반 방식을 사용하였으나, 이 방법은 특정 상황을 가정하여 구성하기 때문에 재사용성이 떨어지며, 동일하거나 비슷한 패턴의 대화가 계속되어 유연성이 부족하다.

최근 이러한 방식들의 한계점을 극복하기 위하여 학습기법을 이용한 자연언어 생성 기법이 시도되고 있다. Ratnaparkhi는 자연언어 생성의 네 가지 학습기법을 제안하였고[14], Walker 등은 항공 여행 도메인에서 문장 계획 트리를 이용한 자연언어 생성 방법을 제안하였다[18]. 그리고 Oh와 Rudnicky는 항공 여행 도메인에 통계적인 접근방법을 적용하였다 [13].

자연언어 기반의 지식 획득

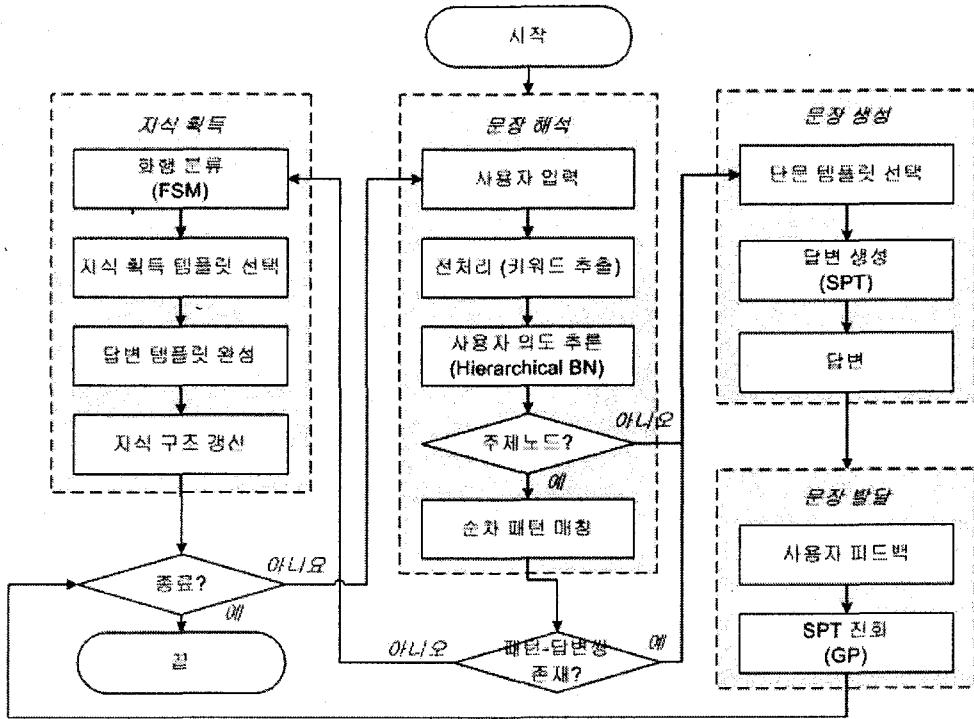
지능형 에이전트의 지식구조를 설계자가 직접 구축하지 않고 자연언어 기반 정보를 이용한다면 매우 효과적일 것이다. 보통 정형적인 자료로부터 지식을 획득하는 것은 그리 어렵지 않지만 가변적이고 특정한 형태가 제공되지 않는 자료로부터 지식을 추출하는 것은 매우 어렵다. 특히 자연언어는 상황과 사람에 따라 매우 가변적이기 때문에 이들로부터 유

용한 정보를 정확히 추출하는 것은 매우 어려운 일이다[16].

자연언어 기반의 지식 획득은 오래전부터 연구되어 왔으며, 최근 웹 문서로부터 의미정보나 구조정보를 추출하는 연구가 활발히 이루어지고 있다[1]. 모델 기반의 접근법은 전통적으로 연구되어온 지식 획득 방법으로, 자연언어 자료를 이용하여 전문가 시스템의 지식 구조를 구축할 때 많이 사용된다[16]. 대상영역의 정형적인 지식 표현을 하나의 모델로 정의하고 모델 완성을 위해 필요한 정보를 문서로부터 수집한다. 이때, 템플릿이나 프레임 등을 이용하여 정보를 효과적으로 가공하거나 저장한다. 이 방법은 정보의 처리과정이 명확하여 이해하기 쉽고, 대상영역에 적합한 템플릿이나 프레임을 설계하면 높은 성능을 가진다[2]. 최근에는 온톨로지를 이용한 지식 획득이 많은 관심을 받고 있는데, 모델 기반의 지식 획득에 비해 보다 깊은 의미수준의 지식을 추출한다. 문서 내에 존재하는 개체들 사이의 의미적 관계를 추출하고 이를 바탕으로 의미나 구조 수준의 유용한 정보를 분석한다. 특히 웹에서는 문서 내의 의미 구조뿐만 아니라 문서 간의 의미 구조도 분석하여 다양한 정보를 획득하는데 사용된다[12].

제안하는 방법

(그림 3)은 사용자 입력 문장의 처리과정을 보여준다. 문장 해석 모듈은 사용자 입력으로부터 사용자 의도를 추론하여 적절한 패턴-답변 쌍을 선택하고, 문장 생성 모듈은 가지고 있는 사용자 정보에 선택된 패턴-답변 쌍의



(그림 3) 제안하는 대화형 에이전트의 흐름도

정보를 추가하여 이 정보로부터 적절한 답변 문장을 생성하고, 문장 발달 모듈은 사용자 피드백을 받아 답변 문장의 표현을 진화시킨다. 지식 획득 모듈은 화행별 템플릿을 이용하여 대화형 에이전트의 지식을 획득한다.

문장 해석

제안하는 대화형 에이전트는 베이지안 네트워크와 패턴 매칭의 2 단계 추론을 통해서 사용자의 질의를 처리한다[8]. 베이지안 네트워크는 사용자 대화의 주제를 추론하기 위하여 사용되고, 패턴 매칭으로 선택된 주제에 적합한 패턴-답변 쌍이 선택된다. 2 단계 추론을 통해서 에이전트는 사용자의 의도를 깊이 있

게 분석할 수 있으며, 사용자의 의도에 따라서 지식구조를 나누어 구축할 수 있어, 시스템의 확장성 및 휴대성을 높여준다.

전처리 단계에서는 사용자의 입력과 사용자 입력 패턴을 인식하기 위해서 미리 정의된 템플릿들과의 패턴 매칭을 통하여 가장 유사한 템플릿을 찾아내고, 이로부터 도메인 정보와 입력 패턴을 분석한다. 예를 들어, 사용자가 ‘저는 서울에서 부산까지 여행합니다’라고 대답하였다면, 형태소 분석을 통해서 ‘저,’ ‘서울에서,’ ‘부산까지,’ ‘여행’ 등의 키워드를 추출하고 도메인 정보 데이터베이스를 통해서 (1) ‘저,’ ‘\$출발지\$에서,’ ‘\$출발지\$까지,’ ‘여행,’ (2) ‘저,’ ‘\$출발지\$에서,’ ‘\$도착지\$까지,’ ‘여행,’ (3) ‘저,’ ‘\$도착지\$에서,’ ‘\$출발지\$까지,’

```

<script> ::= "<SCRIPT>" <topic> <pattern> <response> "</SCRIPT>"
<topic> ::= "<TOPIC>" <word> "</TOPIC>"
<pattern> ::= "<PATTERN>" <word> "</PATTERN>"
<response> ::= "<RESPONSE>" <word> "</RESPONSE>"
<word> ::= word {<word>}

```

(그림 4) 패턴-답변 쌍의 정의

‘여행,’ (4) ‘저,’ ‘\$도착지\$에서,’ ‘\$도착지\$까지,’ ‘여행’의 총 4개의 후보 목록을 만든 후, 정의된 템플릿들 중에서 순차 패턴매칭을 통해 ‘저는 \$출발지\$에서 \$도착지\$까지 여행합니다’가 선택된다. 이로부터 에이전트는 사용자 정보로 \$출발지\$-[서울], \$도착지\$-[부산]의 속성·값의 쌍을 얻을 수 있으며, 입력 패턴이 “\$출발지\$, \$도착지\$, 여행”이 되는 것을 분석 할 수 있다.

베이지안 확률추론은 불충분한 지식을 가지고 정보를 표상하거나 추론하는 대표적인 방법으로, 어떤 사실이 관측되었을 경우 환경을 적절하게 표현하는 가설이 어떤 것인지 추론 한다. 네트워크의 노드는 랜덤 변수를 표현하며 아크는 변수들 간의 의존성을 표현한다 [17]. 본 논문에서는 대화도중 빈번히 발생하는 생략과 축약 등의 불확실성을 극복하기 위해 주제 계층, 개념 계층 그리고 키워드 계층을 가지는 계층적 베이지안 네트워크를 사용한다[8]. 대상영역을 계층적으로 모델링하면 네트워크 설계에 수월함을 제공하고 대화형 에이전트로 하여금 효과적으로 사용자 의도를 파악하여 적절한 대답을 하도록 한다.

베이지안 네트워크에서 입력 질의에 대한 적절한 주제가 선택되면, 선택된 주제에 관련된 지식구조와 키워드 매칭을 수행하여 적절

한 패턴-답변 쌍을 선택한다. 보통 패턴-답변 쌍이 많고 규칙이 복잡해지면 정보의 중복으로 인해 패턴 매칭의 성능이 저하되지만, 본 논문에서는 전체 패턴-답변 쌍을 주제별로 나누고 베이지안 네트워크의 추론을 통해 질의 주제를 선택함으로써 패턴 매칭에서 고려해야 하는 패턴-답변 쌍의 양과 정보의 중복을 최소화한다. 패턴-답변 쌍은 (그림 4)의 정의에 따라 XML 형식으로 저장된다. <topic>은 베이지안 네트워크 주제 추론의 결과로 선택되는 주제 노드의 이름과 같은 이름을 가지며, <pattern>은 사용자 질의문의 키워드와 패턴 매칭할 키워드의 순서이며, <response>는 패턴-답변 쌍이 선택되었을 때 제공할 답변정보를 포함한다.

키워드 매칭은 키워드의 순서를 고려하는 순차패턴매칭의 방식으로 진행되며, 점수는 문서 분류에서 자주 사용되는 F-measure을 사용한다. 본 논문에서는 정확도(precision)와 재현율(recall)이 같은 비중을 가지도록 α 값을 1로 설정한다. 키워드 매칭을 통하여 적절한 패턴-답변 쌍이 존재하면, 이 패턴-답변 쌍의 답변정보와 문장계획 트리를 통해 문장을 생성한다.

$$F-measure = \frac{(\alpha+1) \times precision \times recall}{\alpha \times precision + recall}$$

<표 2> F-measure에 사용되는 변수

		입력질의	
		포함	미포함
스크립트	포함	A	B
	미포함	C	D

$$precision = \frac{A}{A+B}, \quad recall = \frac{A}{A+C}$$

문장 생성 및 발달

문장 생성

본 논문에서는 문장계획 트리를 이용하여 단문들로부터 복합 문장을 생성한다. 문장계획 트리(Sentence Plan Tree ; SPT)는 Lavoie와 Rambow가 고안한 것으로[14], 말단 노드는 단문을 표현하고, 중간 노드는 자식 노드의 문장을 연결하는 결합연산자로 구성된다. 여기서는 SPT를 2진 트리로 구성하여 말단 노드는 단문을 만드는데 필요한 템플릿으로, 그리고 중간 노드는 두 자식 노드가 만드는 문장을 결합하는 5개의 결합연산자로 구성한다. 결합연산자로는 한국어를 분석하여 아래와 같이 $JO_{1\sim 5}$ 의 5개로 정의한다.

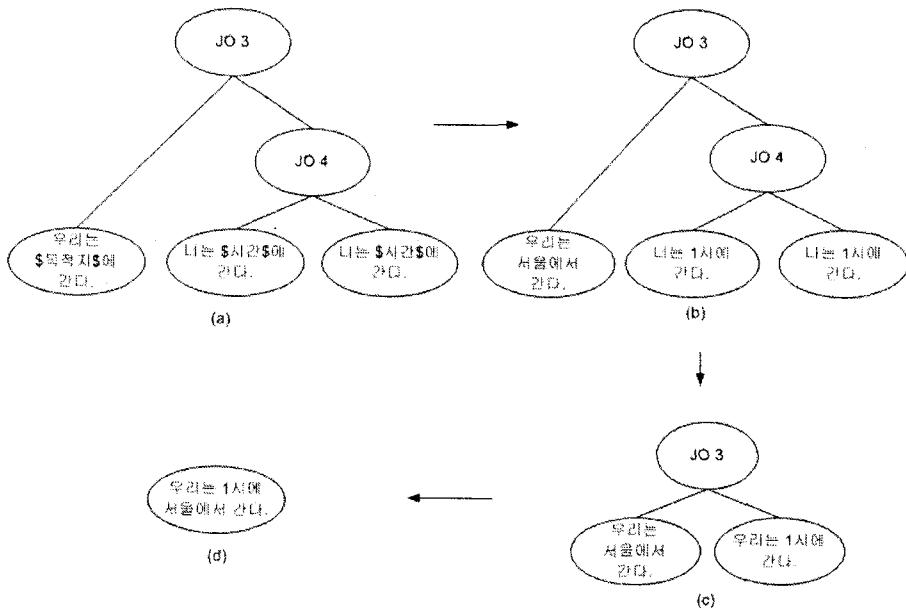
- JO_1 : 가장 간단한 방법으로 두 문장을 결합하는 연산자로 두 문장을 접속사 ‘그리고’로 연결한다.
- JO_2 : 두 문장이 동일한 주어를 가질 경우에 두 문장을 연결하는 연산자로, 하나의 주어를 생략하여 두 문장을 결합한다.
- JO_3 : 두 문장이 동일한 주어와 동일한 동사를 가질 경우에 두 문장을 연결하는 연산자로, 주어와 동사를 하나씩 생략하여 두 문장

을 결합한다.

- JO_4 : 두 문장의 주어는 다르고 나머지 부분은 같은 경우 사용되는 연산자로, 두 주어의 의미를 포함하는 주어를 사용하여 문장을 결합한다.
- JO_5 : 두 문장이 동일한 주어를 가지며 다른 동사를 가지지만 두 동사의 의미를 포함할 수 있는 하나의 동사가 존재할 경우 사용되는 연산자로, 이 동사를 사용하여 두 문장을 결합한다.

SPT가 선택되면 SPT의 말단 노드인 단문과 도메인 정보를 합쳐 하나의 문장을 생성하고 문장 골격 해석기를 통해서 하나의 복문을 완성한다. (그림 5)는 문장계획 트리의 해석과정을 보여준다.

(그림 5a)의 말단 노드는 도메인 정보 삽입기에서 \$목적지\$로 ‘서울’과 \$시간\$으로 ‘1시’의 도메인 정보를 가지고 (그림 5b)와 같이 ‘우리는 서울에 간다,’ ‘너는 1시에 간다,’ ‘나는 1시에 간다’의 세 문장으로 변환된다. (그림 5b)의 두 문장 ‘너는 1시에 간다’와 ‘나는 1시에 간다’는 결합연산자에 의해서 ‘우리는 1시에 간다’가 되며(그림 5c), 남은 두 문장은 결합연산자에 의해서 최종적으로 ‘우리는 1시에 서울에서 간다’라는 문장으로 변환된다(그림 5d).



(그림 5) STP의 해석 과정

문장 생성 모듈의 동작

(그림 3)에서 보듯이, 문장 생성 모듈은 두 가지 상황에서 활성화 된다. 한 경우는 베이지 안 주제 추론 과정에서 임계값을 넘는 주제 계층의 노드가 없는 경우이고, 다른 경우는 사용자에게 정보 제공을 위한 패턴-답변 쌍을 선택한 경우이다. 전자의 경우는 주제 추론을 위해서 좀더 많은 도메인 정보가 필요하므로 도메인의 주제 추론에 필요한 속성-값 쌍을 선택하고 이 속성의 값을 얻을 수 있는 질문을 만든다. 후자의 경우는 선택된 패턴-답변 쌍에 있는 도메인 정보 중에서 답변에 있는 도메인 정보를 사용자에게 전달하면 된다. 따라서 SPT 후보 목록에서 답변 도메인 정보를 포함하고 패턴 도메인 정보를 일부 포함하는 SPT를 선택하여 답변 문장을 생성한다.

문장 발달

유전자 프로그래밍은 컴퓨터가 문제를 풀 수 있는 프로그램을 자동으로 생성하는 기법으로 John R. Koza가 제안했다. 이는 유전자 알고리즘의 변형으로 집단의 각 개체는 컴퓨터 프로그램을 나타낸다[10]. 본 논문에서는 STP로 유전자 프로그래밍의 각 개체를 표현한다. 앞에서 언급한 바와 같이, SPT의 말단 노드는 대상영역의 정보를 가진 단문 템플릿으로 이루어지며, 부모 노드는 결합연산자로 구성된다.

유전자 프로그래밍에 사용되는 개체는 결합 연산자 집합 $\{JO_1, JO_2, JO_3, JO_4, JO_5\}$ 과 단문 템플릿의 집합 $\{r_1, r_2, \dots, r_n\}$ 으로 구성되며 트리의 생성문법은 $G = \{V = \{\text{EXP}, \text{OP}, \text{VAR}\}, T = \{JO_1, JO_2, JO_3, JO_4, JO_5, r_1, r_2, \dots, r_n\}, P, \{\text{EXP}\}\}$ 가 되며 규칙 P는 다음과 같다.

$$\begin{aligned} \text{EXP} &\rightarrow \text{EXP OP EXP} \mid \text{VAR} \\ \text{OP} &\rightarrow \text{JO}_1 \mid \text{JO}_2 \mid \text{JO}_3 \mid \text{JO}_4 \mid \text{JO}_5 \\ \text{VAR} &\rightarrow r_1 \mid r_2 \mid \dots \mid r_n \end{aligned}$$

교차 연산과 돌연변이 연산을 통하여 다양한 문장 구조를 생성한다. 교차 연산은 두 개체에서 임의로 서브 트리를 선택하여 두 서브 트리를 바꾸며, 돌연변이 연산은 임의로 선택된 노드의 값을 다른 값으로 변경한다. 이들 연산은 미리 정의된 확률에 의해 일어나며, 개체의 적합도 평가는 사용자에 의해 이루어진다. 만일 아이들이 시스템을 평가하면 아이들이 사용되는 표현을 이용한 문장이 생성되고, 어른들이 시스템을 평가하면 어른들이 사용하는 표현을 이용한 문장이 생성된다. 적합도 평가는 실시간으로 수행되기 때문에 시스템은 문장 생성의 언어적 감각을 지속적으로 발전시킨다.

지식 획득

(그림 3)과 같이 지식 획득 과정은 먼저 사용자 질의의 패턴을 분석하고 사람과의 상호작용을 통해서 이에 적합한 패턴-답변 쌍을 만든다. 스크립트의 패턴은 주제와 키워드의 집합으로 구성되며, 주제는 베이지안 네트워크 추론을 통하여 얻고 키워드는 전처리 과정을 통하여 수집한다.

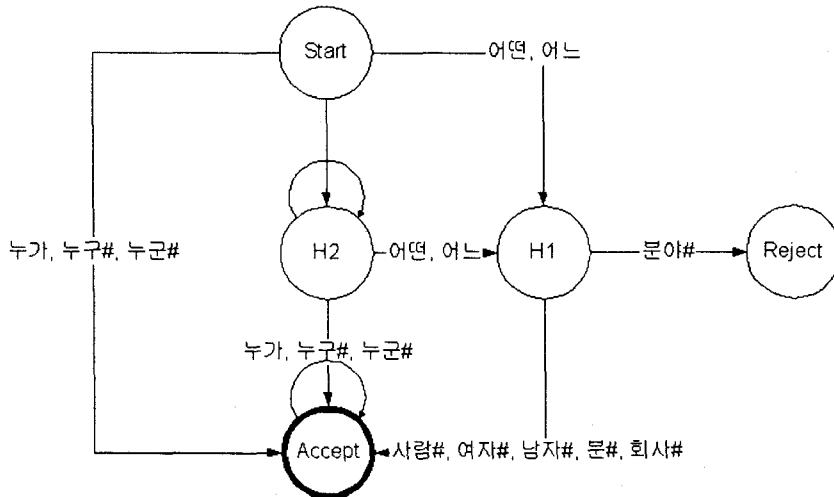
기존의 대화형 에이전트가 키워드 정보만을 이용해 패턴을 정의하는데 반해, 본 논문에서는 키워드 사이의 구조적 순서 정보인 화행을 함께 고려한다. 질의를 통해 표현되는 사용자 의도를 <표 3>과 같이 30가지로 구분하여 화행으로 정의한다. 질의는 우선 1차 질문형이나 1차 평서문형 중의 오직 하나의 화행으로 분류된다. 1차 질문/평서문형은 다시 각각 2차 질문/평서문형 중 어떤 화행으로 분류될 수 있는지 결정된다. 1차 질문/평서문형은 오직 하나의 화행으로 분류되나 2차 질문/평서문형은 여러 개의 화행으로 분류될 수 있다.

예를 들어, “학교 가는 길을 알려줄 수 있니?”와 같은 질의는 1차 분류에서 능력을 묻는 질의인 Ability 질문형으로 분류되고, 2차 분류에서 방향을 묻는 질의인 Direction 질문형으로 분류된다.

각 화행은 (그림 6)과 같이 구성된 오토마타를 통해 추출되며, 30개의 오토마타를 포섭구조로 설계하여 입력 질의의 화행이 선택되도록 한다. 예를 들어, “너는 어떤 사람을 좋아하니?”라는 문장이 들어오면 (그림 6)의 “Who” 질문형 분류 오토메타를 통해 ‘Start’ 상태에서 단어 ‘너는’을 통해 ‘H2’의 상태로 이동하고

<표 3> 화행 분류

부류	화행
1차 질문형	Ability, Description, Fact, Location, Method, Miscellaneous, Obligation, Reason, Time, WhatIf, Who
2차 질문형	Acquisition, Comparison, Confirmation, Cost, Direction, Example, More, Possession
1차 평서형	Act, Fact, Message, Miscellaneous, Possession, Status, Want
2차 평서형	Cause, Condition, Feeling, Time



(그림 6) Who 질문형 분류 오토메타

‘어떤’을 통해 ‘H1’으로 이동한다. 그리고 ‘사람을’을 통해 최종 상태인 ‘Accept’ 상태로 이동하여 이 문장의 화행이 “Who” 질문형으로 결정된다.

답변은 화행별로 정의된 답변 템플릿을 기반으로 생성되며, 답변 템플릿은 (그림 7)과 같이 정의된다. 템플릿은 정적인 정보 표현형을 가지는 동시에 변수의 값이 동적으로 변하여 다양한 형태로 정보를 표현할 수 있는 장점이 있다.

```

// Definition
(def-index template-name
  :class (dialogue-act)
  :question (question-script)
  :requirement (response primitives)
)

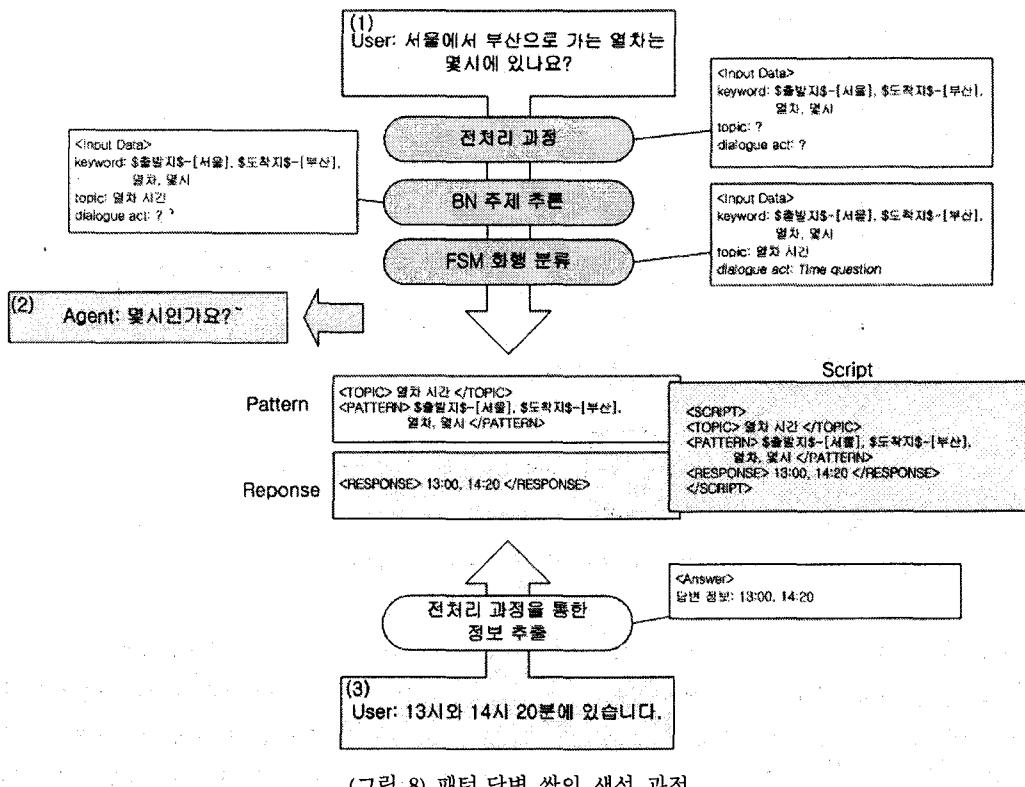
// Example
(def-index no. 10
  :class (?LocationQuestion)
  :question ("Where is it?")
  :requirement (Seoul)
)
  
```

(그림 7) 템플릿 정의 및 예

class는 화행정보를 의미하며, question은 추가로 필요한 정보를 수집하기 위해 사용자에게 던지는 질문을 가리킨다. requirement는 답변 템플릿 완성에 필요한 정보를 나타낸다. 본 논문에서는 각 질문형에 대해서 3개의 답변 템플릿을, 평서형에 대해서는 긍정적/부정적 답변 템플릿을 3개씩 정의하여 총 64개의 답변 템플릿을 설계하였다.

입력 질의의 패턴과 답변의 정의가 완료되면 이들을 결합하여 패턴-답변 쌍을 완성한다. 패턴-답변 쌍은 사용자 질의에 대해 적절한 답변을 제공하기 위한 지식구조의 요소로써 대화형 에이전트의 답변 성능을 결정한다. (그림 8)은 패턴-답변 쌍의 생성하여 지식을 획득하는 과정을 보여준다.

사용자 질의에 대한 답변을 에이전트가 찾을 수 없다면(그림 8(1)), 해당 질의를 분석하여 패턴 스크립트를 작성하고, 해당 질의에 대한 답변을 사용자에게 요구한다(그림 8(2)). 그 후, 사용자 답변이 들어오면(그림 8(3)) 전



(그림 8) 패턴-답변 쌍의 생성 과정

처리 과정을 통해 적절한 정보를 추출하여 답변 스크립트를 작성하고, 패턴 스크립트와 답변 스크립트를 결합하여 하나의 패턴-답변 쌍을 생성하여 지식 획득을 완료한다.

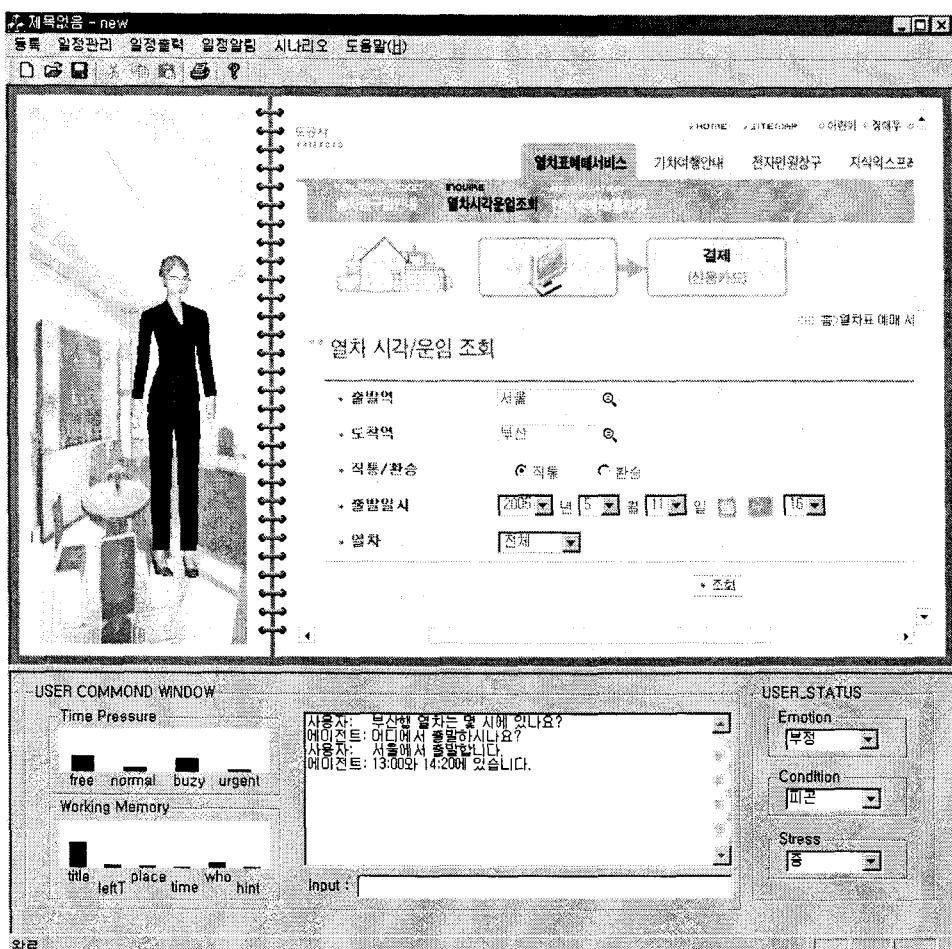
황에 적절한 움직임을 보이며 음성을 통한 답변을 해주는 아바타(avatar)로 구성된다. 아바타는 Q-avatar(<http://www.qavatar.com>)를 이용하여 아바타의 외모와 행동을 정의하였고, Voiceware(<http://www.voiceware.co.kr>)를 이용하여 음성 출력을 구현하였다.

제안하는 방법은 열차 예매를 대상영역으로 하였다. 열차예약 에이전트는 사용자와의 대화를 통하여 열차예약에 필요한 정보를 수집하고, 수집된 정보에 적합한 기차를 예약해 준다. 열차예약에 필요한 정보는 출발지, 도착지, 출발날짜, 출발시간, 도착시간 그리고 열차등급의 6가지로 구성된다.

실험 및 결과

구현

제안하는 방법의 유용성을 보이기 위해서 MFC를 이용하여 (그림 9)와 같은 인공비서 시스템을 설계하였다. 이 시스템은 주 정보를 표현해주는 창과 입력 텍스트 박스 그리고 상



(그림 9) 인공비서 시스템

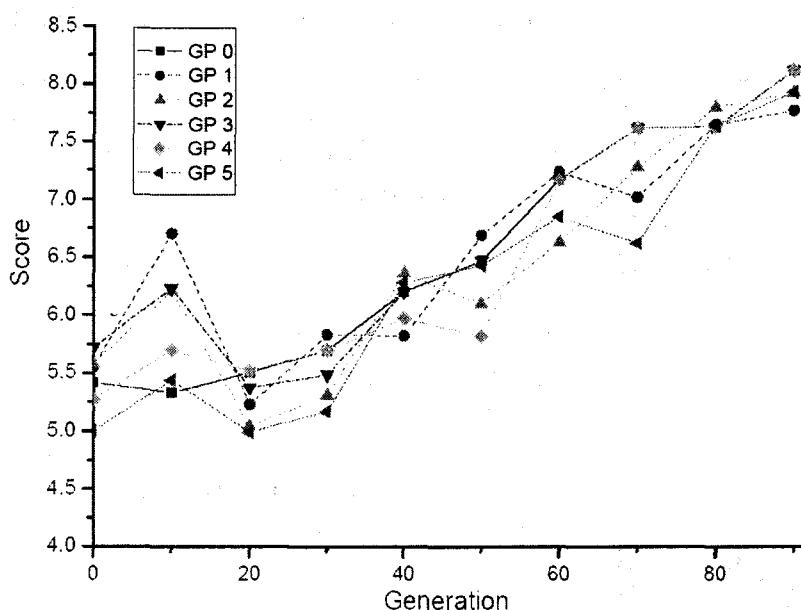
대화 학습 사용자 평가

문장 표현 진화

이 실험은 유전자 프로그래밍 기반 문장 생성 방법이 진화를 통해서 얼마나 사용자에게 만족감을 주는지 알아보기 위한 것으로 10명의 피험자들이 90세대에 걸쳐 6개의 SPT 그룹을 평가하였다. 피험자는 각 세대마다 생성되는 모든 문장을 평가해야 하므로 개체 수는 20개로 다소 적게 설정하였고, 교차연산 비율은

0.6, 돌연변이연산 비율은 0.2로 설정하였다.

(그림 10)은 진화가 진행됨에 따른 각 SPT 그룹별 평균 사용자 평가점수를 보여준다. 사용자는 각 세대에서 생성된 문장들을 0~10점으로 평가했다. 처음 10세대 동안에는 점수가 증가하다가 이후 10세대는 감소하는 경향을 보였다. 이것은 처음에 생성된 어색한 문장이 진화를 통해서 초기에는 빠르게 진화하여 좋은 점수를 얻었으나 이 후에는 진화 속도가 사용자의 기대치보다 늦어 상대적으로 낮은



(그림 10) 세대별 진화과정에 따른 사용자 평가 점수

<표 4> 진화과정 중 생성된 문장

세대	문장
0	손님은 서울에서 10월 10일에 출발하고 몇 시에 도착하길 원합니까? 그리고 몇 시에 도착하길 원하십니까? 그리고 손님은 부산에 갑니다.
30	손님은 서울에서 10월 10일에 출발합니다. 어떤 기차를 타길 원하십니까? 그리고 손님은 몇 시에 떠나십니까?
60	손님은 부산에 갑니다. 그리고 손님은 서울을 몇 시에 떠나십니까?
90	손님은 서울에서 부산으로 몇 시에 갑니까?

점수를 받았기 때문이다.

<표 4>는 도메인 정보로 \$출발지\$, \$도착지\$, \$출발날짜\$를 각각 ‘서울’, ‘부산’, ‘10월 10일’을 사용했을 때 생성된 문장을 나타낸다. <표 4>에서 보는 바와 같이 처음 세대에서는 복잡하고 어색한 문장이 생성되었지만, 진화가 진행됨에 따라 간결한 문장을 획득하였다.

지식 획득

대화기반 지식 획득의 유용성을 검증하기 위해 사용자 평가를 실시하였다. 우선 8명의 피험자들로부터 <표 5>의 3가지 과제로부터 57개의 대화를 수집하였다.

<표 6>은 수집된 대화를 이용하여 기존의 단순패턴 매칭 방법과 제안하는 방법의 성능을 비교 평가한 결과를 보여준다. 단순패턴

<표 5> 과제

과제 번호	과제 내용
과제 1	서울에서 부산으로 가는 열차 시간을 알아내시오.
과제 2	서울에서 부산과 대구로 가는 열차의 시간을 각각 알아내시오.
과제 3	에이전트의 이름, 나이, 제작자를 알아내시오.

<표 6> 비교 결과

방법	데이터	과제 1		과제 2		과제 3	
		검색률	평균 입력문장	검색률	평균 입력문장	검색률	평균 입력문장
단순패턴	학습데이터	92.5%	1.5	95%	2.7	93.2%	4.4
	테스트데이터	87.3%	1.6	86.3%	2.7	87.7%	4.6
제안하는 방법	학습데이터	100%	1.2	100%	2.1	100%	4.0
	테스트데이터	92.3%	1.4	96%	2.3	91.7%	4.5

매칭 방법은 사용자 입력 문장과 스크립트에서의 동일한 키워드의 개수를 기준으로 스크립트를 평가하는 방법으로 3.1절의 순차패턴 매칭과 같은 F-measure를 사용했다. <표 6>에서 알 수 있듯이, 단순패턴 매칭 기술은 경직된 특성으로 인해 제안하는 방법 보다 낮은 검색률을 얻었으며, 많은 입력문장이 제공되어야 원하는 정보를 얻었다. 이는 대화의 문맥 유지나 능동적 대화 등이 실제 대화에서 발생하는데, 단순패턴 매칭은 이를 처리하지 못하기 때문이다.

제안하는 대화 에이전트의 지식 획득 방법의 유용성을 평가하기 위해 수동으로 지식구조를 구축하는 경우와 비교하여 실험을 수행하였다. 임의의 질의와 답변에 대해서 제안하는 방법과 수동으로 사용자가 패턴-답변 쌍을 제작하게 하였을 때의 시간과 답변 성능을 측정하였다. 학습 데이터로 연구실에 관련된 정

보로 구성된 질의-답변 쌍 50개를 만들고 동일한 내용에 대한 질의-답변 쌍 50개를 제작하여 테스트 데이터로 사용하였다. 실험은 대화 에이전트의 스크립트 제작에 숙련된 사람과 비 숙련자를 대상으로 수행하였다. <표 7>은 실험 결과를 보여준다.

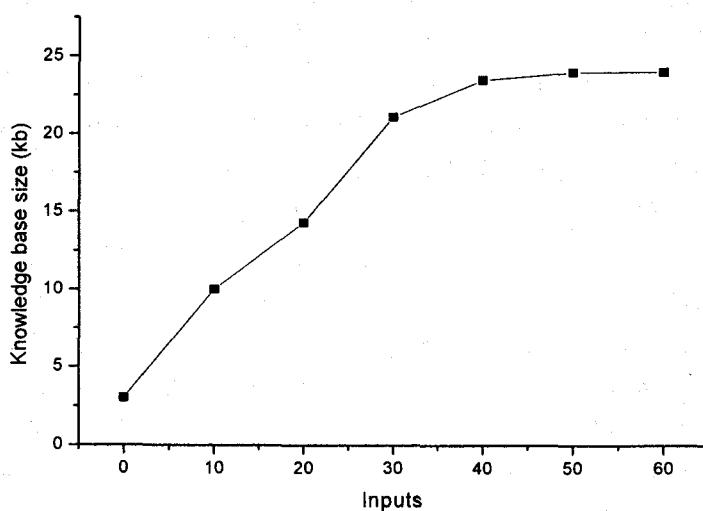
동일한 패턴을 가진 질의들이 학습데이터에 포함되어 있었으나 사용자는 이를 바로 구분하지는 못하였다. 따라서 수동 제작의 경우 50개의 패턴-답변 쌍이 생성되었다. 반면에 제안하는 방법은 동일한 패턴을 가진 질의일 경우, 지식 학습을 수행하지 않기 때문에, 보다 적은 수의 패턴-답변 쌍이 생성되었다. 제작 시간은 제안하는 방법이 월등히 우수했다. 수동 제작의 경우, 질의의 패턴을 설계자가 정의하고 모든 스크립트를 직접 작성하여야 하기 때문에 매우 시간이 많이 소요되었다. 특히 비숙련자의 경우에는 패턴의 분석과 스크립트

<표 7> 실험 결과

	수동 설계		제안하는 방법	
	숙련자	비 숙련자	숙련자	비 숙련자
패턴-답변 쌍 수	50	50	44	44
파일 크기	28 KB	30 KB	24 KB	24 KB
설계 시간	20 분	60 분	5 분	8 분
정확률	학습 데이터	92%	84%	100%
	테스트 데이터	84%	82%	88%
피드백 (0~5)	2	1	4	4

작성은 어려운 문제이기 때문에 더 많은 시간이 소요되었다. 반면에 대화를 이용하여 패턴-답변 쌍을 생성하는 경우에는 단순히 대화를 따라서 수행하기 때문에 숙련자나 비 숙련자나 별 어려움이 없이 수행하였다. 정확률에서는 패턴이 자동으로 정의되기 때문에 제안하는 방법이 매우 우수했으며, 수동 제작으로는 잘못된 패턴의 정의가 발생되었고, 비 숙련자는 더 많은 오류를 범했다. 결과적으로 사용

자 평가에서 제작 시 제안하는 방법이 매우 편리하다는 것을 확인하였으며 비 숙련자의 경우에는 이를 더 심각하게 느끼는 것을 확인하였다. 하지만 부분적으로 대화를 통해 답변 생성에 필요한 정보를 수집하는 과정에서 특정 형태의 답변문장으로부터만 필요한 정보를 추출하기 때문에 약간의 어색함이 발생하였으나, 이는 보다 다양한 오토마타를 설계하여 다양한 형태의 문장을 처리하면 해결할 수 있



(그림 11) 대화형 에이전트의 지식 학습

다. (그림 11)은 대화형 에이전트가 새로운 지식을 학습하는 양상을 보여준다. 사용자와의 대화를 통해 이전에 알지 못했던 패턴-답변 쌍을 지속적으로 생성하여 지식구조의 크기가 확장되었다.

결 론

본 논문에서는 BN, PM, FSM, TP, GP를 이용하여 자율적으로 대화 기능을 학습하는 대화형 에이전트를 제안하였다. 기존의 대화형 에이전트의 설계방법은 설계에 많은 시간과 노력이 필요하며 스스로 학습하는 기능이 부족했다. 본 논문에서는 점증적으로 지식구조를 증가시키고 문장 표현 능력을 개선하는 대화형 에이전트를 제안하고 다양한 적용 사례와 사용자 평가를 통해 그 유용성을 확인하였다.

자아 성장에 있어서 언어는 매우 중요한 위치를 차지하며, 사람은 언어발달 과정을 통해 자신의 어휘와 표현력을 증진하여 자연스러운 언어를 구사한다. 대화 학습에 관련한 연구는 인지과학 분야에서 다양한 인지 언어 모델을 설명하는데 유용하게 사용될 뿐만 아니라 지능형 에이전트와 사람간의 통신 수단의 설계에 있어서 중요한 방향성을 제시한다. 향후에는 자율 자아 성장 및 대화 학습 모델이 포함된 지능형 에이전트를 개발하고자 한다.

참고문헌

- Alani, H., Kim, S., Millard, DE, Weal, MJ, Hall, W., Lewis, PH and Shadbolt, NR (2003), Automatic ontology-based knowledge extraction from web documents, *IEEE Intelligent Systems*, 18(1), 14-21.
- Arruarte, A., Elorriaga, J. A. and Rueda, U. (2001), A template-based concept mapping tool for computer-aided learning, *Proc. IEEE Int. Conf. Advanced Learning Technologies*, 309-312.
- Clack, E. (2004), How language acquisition builds on cognitive development, *Trends in Cognitive Science*, 8(10), 472-478.
- Dominey, P. and Boucher, J. (2004), Developmental stages of perception and language acquisition in a perceptually grounded robot, *Proc. 4th Int. Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, 45-52.
- Ferguson, G., Allen, J. and Miller, B. (1996), TRAINS-95: Towards a mixed-initiative planning assistant, *Proc. Of the 3rd Int. Conf. on Artificial Intelligence Planning Systems*, 70-77.
- Gomez, R. and Gerken, L. (2000), Infant artificial language learning and language acquisition, *Trends in Cognitive Science*, 4(5), 178-186.
- Green, A. and Eklundh, K. (2003), Designing for learnability in human-robot communication, *IEEE Trans. Industrial Electronics*, 50(4), 644-650.
- Hong, J.-H. and Cho, S.-B. (2004), Domain-adaptive conversational agent with two-stage dialogue management, *Lecture Notes in Artificial Intelligence*, 3339, 1147-1153.
- Joshi, A. and Weng, J. (2003), Autonomous mental development in high dimensional

- context and action spaces, *Neural Networks*, 16(5-6), 701-710.
- Koza, J. (1994), *Genetic programming, Automatic discovery of reusable programs*, The MIT Press.
- Lauria, S., Bugmann, G., Kyriacou, T., Bos, J. and Klein, E. (2001), Training personal robots using natural language instruction, *IEEE Intelligent Systems*, 16(3), 38-45.
- Navigli, R., Velardi, P. and Gangemi, A. (2003), Ontology learning and its application to automated terminology translation, *IEEE Intelligent Systems*, 18(1), 22-31.
- Oh, H. and Rudnicky, I. (2002), Stochastic natural language generation for spoken dialog systems, *Computer Speech and Language*, 16(3-4), 387-407.
- Ratnaparkhi, A. (2002), Trainable approaches to surface natural language generation and their application to conversational dialog systems, *Computer Speech and Language*, 16(3-4), 435-455.
- Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D. and Seidenberg, M. S. (2001), How psychological science informs the teaching of reading, *Psychological Science in the Public Interest*, 2(2), 31-74.
- Schmidt, C. and Wetter, T. (1998), Using natural language sources in model-based knowledge acquisition, *Data & Knowledge Engineering*, 26(3), 327-356.
- Stephenson, T. (2000), An Introduction to Bayesian Network Theory and Usage, *IDIAP-RR00-03*.
- Walker, M. A., Rambow, O. C. and Rogati, M. (2002), Training a sentence planner for spoken dialogue using boosting, *Computer Speech and Language*, 16(3-4), 409-433.
- Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M. and Thelen, E. (2000), Autonomous mental development by robots and animals, *Science*, 291, 599-600.
- Zue, V. and Glass, J. (2000), Conversational interfaces: Advances and challenges, *Proc. of the IEEE*, 88(8), 1166-1180.

1 차원고접수: 2005. 9. 15

2 차원고접수: 2005. 11. 7

최종제재승인: 2005. 12. 3