

■ 2010년도 학생논문 경진대회 수상작

# 행동 네트워크를 이용한 감정형 제스처 기반 대화 관리 시스템

(An Emotional Gesture-based Dialogue Management System  
using Behavior Network)

윤 종 원 <sup>†</sup>                      임 성 수 <sup>†</sup>                      조 성 배 <sup>\*\*</sup>  
(Jong-Won Yoon)              (Sung-Soo Lim)              (Sung-Bae Cho)

**요 약** 최근 로봇의 다양한 활용과 더불어 로봇-사람간의 상호작용에 대한 연구 또한 활발하게 진행되고 있으며, 대표적으로 자연어 처리나 대화중 제스처 생성 등의 방법이 로봇과 사람간의 의사소통에 적용되었다. 그러나 기존의 로봇-사람간의 의사소통을 위한 방법을 통해서서는 정적인 의사소통만을 수행할 수 있다는 한계가 존재하며, 보다 자연스럽고 사실적인 의사소통 방법에 대한 연구가 필요하다. 본 논문에서는 보다 수준 높은 의사소통을 위해 감정형 제스처 기반 대화 관리 시스템을 제안한다. 제안하는 시스템은 베이저안 네트워크와 패턴 매칭을 이용하여 로봇과 사람간의 대화를 수행함과 동시에 로봇-사람간의 대화 도중 실시간으로 상황에 어울리는 로봇의 감정형 제스처를 생성한다. 감정형 제스처 생성을 통해 로봇은 사람에게 대화를 보다 효과적으로 전달함과 더불어 사실적인 상호작용을 수행할 수 있다. 제스처 생성은 동적으로 변화하는 대화 상황에 유연하게 대처할 수 있도록 행동 네트워크를 사용하였다. 이후 제안하는 감정형 제스처 기반 시스템의 유용성을 검증하기 위해 사용성 평가를 통해 감정이나 제스처를 사용하지 않는 기존의 대화 관리 시스템과의 비교를 수행하였다.

키워드 : 휴머노이드, 대화 관리, 감정형 로봇, 제스처 생성

**Abstract** Since robots have been used widely recently, research about human-robot communication is in process actively. Typically, natural language processing or gesture generation have been applied to human-robot interaction. However, existing methods for communication among robot and human have their limits in performing only static communication, thus the method for more natural and realistic interaction is required. In this paper, an emotional gesture based dialogue management system is proposed for sophisticated human-robot communication. The proposed system performs communication by using the Bayesian networks and pattern matching, and generates emotional gestures of robots in real-time while the user communicates with the robot. Through emotional gestures robot can communicate the user more efficiently also realistically. We used behavior networks as the gesture generation method to deal with dialogue situations which change dynamically. Finally, we designed a usability test to confirm the usefulness of the proposed system by comparing with the existing dialogue system.

**Key words** : Humanoid, Dialogue management, Emotional robot, Gesture generation

\* 이 논문은 2010년도 교육과학기술부의 재원으로 한국연구재단 뇌과학원천기술 개발사업의 지원을 받아 수행된 연구임(No. 2010-0018948)

<sup>†</sup> 학생회원 : 연세대학교 컴퓨터과학과  
jwyoons@slab.yonsei.ac.kr  
lss@slab.yonsei.ac.kr

<sup>\*\*</sup> 종신회원 : 연세대학교 컴퓨터과학과 교수  
sbcho@cs.yonsei.ac.kr

논문접수 : 2010년 5월 31일

심사완료 : 2010년 9월 2일

Copyright©2010 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 소프트웨어 및 응용 제37권 제10호(2010.10)

1. 서론

최근 로봇과 사람 간의 상호작용을 위한 방법으로 의 사소통이 주목받기 시작하면서, 사람과 로봇 사이의 의 사소통을 위한 방법으로 자연어를 이용한 방법이 주목받고 있다[1,2]. 이로 인해 HCI 분야에 있어서 자연어를 이용한 대화관리에 대한 연구가 진행되고 있으며, HRI 분야에서도 활용될 수 있는 가능성을 보여주고 있다[3,4].

로봇과 사람과의 유연하고 사실적인 상호작용을 제공하기 위해 최근 대화 시스템에서는 자연어 생성[5-8]과 같은 문장 생성 메커니즘을 적용하려는 시도가 이루어지고 있다. 이와는 달리 본 논문에서는 보다 유연하고 사실적인 상호작용을 위한 다른 방법으로 로봇의 제스처 생성에 주목하였다. 사람은 대화 도중 무의식적으로 상대방에게 제스처를 취함으로써 보다 순조로운 상호작용을 수행할 수 있다[9]. 이러한 이유로 대화형 에이전트의 사실적인 상호작용을 위해 대화 내용에 따라 적합한 제스처를 생성하는 연구가 진행되고 있다[11,12].

사람은 대화 중 대화 내용뿐만 아니라 다양한 상황을 고려하여 가장 자연스럽다고 판단되는 제스처를 취한다. 따라서 로봇의 경우에도 자연스러운 제스처 생성을 위해서는 대화 상황과 화자의 의도 등의 다양한 상황이 고려되어야 한다. 그러나 기존 대화 중 제스처 생성에 관한 연구의 경우 제한된 수의 상황만을 이용하여 간단한 규칙 기반 방법을 통해 제스처를 생성하는 데에 그치기 때문에 복잡하고 동적인 환경 속에서 진행되는 대화에는 적용하기가 힘들다는 한계가 존재한다.

또한 보다 사실적인 대화 중 제스처 생성을 위해 로봇이 내부 상태를 가짐으로써 보다 자율적인 행동을 수행할 수 있다[12]. 일반적으로 널리 사용되는 내부 상태 중 하나는 감정 상태로써, 감정 상태를 고려한 제스처는 상대방에게 화자 및 청자의 의도를 효율적으로 전달할 수 있으며 사실적인 상호작용을 이끌어낼 수 있다.

본 논문에서는 유연하고 사실적인 사람-로봇 간의 상호작용을 위해 제스처 기반 대화 관리 시스템을 제안한다. 제안하는 시스템은 대화 시스템과 감정 시스템, 그리고 행동 시스템으로 구성된다. 대화 시스템은 대화의 흐름을 파악하고 베이저안 네트워크를 통한 주제 추론과 키워드 매칭을 이용한 답변 선택을 수행한다. 감정 시스템은 OCC모델을 이용하여 현재 상황에 적합한 로봇의 감정을 생성한다. 행동 시스템에서는 대화 시스템과 감정 시스템에서 생성된 정보를 바탕으로 적합한 제스처를 생성한다.

2. 관련 연구

2.1 대화 시스템

표 1 대화형 에이전트에 사용되는 기술

기술	수행작업 예	기술 난이도	쉬움 ↕ 어려움
단순패턴매칭	웹검색엔진	단순질의응답	
제한대본 기반모델	게임	단순한 선택을 통한 대화진행	
유한상태모델	장거리전화	상태전환을 통한 제한된 유연성	
프레임기반모델	열차정보검색	단순패턴매칭을 이용한 정보기입	
예제기반모델	TV, 날씨, navi	대화 예제를 통한 대화 수행	
계획기반모델	부엌디자인 도우미	동적주제변화	

대화 시스템은 자연어 기반 대화를 통해서 사용자의 의도를 이해하고 적절한 행동을 취하여 사용자를 돕는다 [13]. 자연어를 통한 사용자 인터페이스는 유연성, 명료성, 표현력 면에서 장점을 가지고 있기 때문에 인간과 컴퓨터 간의 의사소통 수단으로 사용하기에 적합하다[14].

대화 에이전트의 연구를 위해서 기존의 연구에서 표 1에서 보여주는 기술들이 제안되었다[15]. 계획기반모델을 제외한 대부분의 방법은 대화 에이전트를 설계하는 일반적인 기법이다. 이들은 단순한 대화의 유형들에 기초하여 설계되기 때문에 단순한 작업에 대해서는 설계가 쉽고 성능도 상당히 뛰어나지만, 복잡한 상황의 대화를 처리하기에는 한계가 있다. 예제기반모델은 대화 설계에 있어서 대화 예제를 수집하고 의미정보를 태깅한 후, 사용자 질의문과의 태깅된 정보 간의 유사도를 계산하여 대화를 수행하는 방법이다[16]. 이 방법은 수집된 대화를 분석하여 의미정보를 태깅해야하는 번거로움이 있지만, 상대적으로 적은 양의 예제만으로도 높은 성능의 대화 시스템을 구축할 수 있다[16]. 계획기반모델은 대상영역의 질의모델링에만 한정되어 있는 이전의 기법들과는 다르게, 사용자의 계획을 파악하고 그 계획을 성취하기 위한 동작 등을 결정한다. 좀 더 실세계의 대화에 가깝고, 사용자의도를 파악하여 대화를 진행하기 때문에 다른 방법들에 비해 사용자가 원하는 정보나 동작에서 성능이 우수하다. 하지만 시스템의 구현과 유지가 어려운 것이 문제점이다.

본 논문에서는 대화 에이전트를 보다 유연하고 자연스럽게 만들기 위해서 계획 기반 모델을 이용해 분석 모듈을 설계한다. 또한 베이저안 네트워크 주제 추론, 키워드 패턴 매칭의 2단계 추론을 통해서 보다 정확하게 사용자의 의도를 추론한다.

2.2 행동 기반 시스템

행동 기반 시스템은 명확한 지식을 표현하는 전통적

인 인공지능 방법과는 달리 에이전트와 환경과의 직접적인 상호작용을 통해 행동을 생성한다. Maes가 제안한 행동 네트워크[17]는 행동 기반 시스템에 목표를 부여하여 전체적인 목적에 따라 자율적으로 행동을 생성하는 방법이다. 행동 네트워크는 다양한 분야의 에이전트의 행동 생성에 적용되어 사용된다. Nicolescu 등은 행동 기반 구조를 바탕으로 독자적으로 확장시킨 행동 네트워크 시스템을 제안하였다[18]. 이후 계층적 구조를 가지는 행동 네트워크 시스템을 제안하여 모바일 로봇의 행동 생성에 적용시켰다[19]. 또한 Khoo와 Zubek은 행동 네트워크를 이용하여 컴퓨터 게임의 캐릭터 행동을 생성하였다[20].

행동 네트워크는 일반적으로 간단한 규칙 기반 시스템에 비해 행동들을 보다 자연스럽게 상황에 유연하게 대처하여 생성할 수 있다는 장점을 지니고 있다. 이에 착안하여 본 논문에서는 행동 네트워크를 이용하여 로봇의 제스처를 생성하는 방법을 제안한다.

2.3 감정형 로봇

로봇에게 있어서 감정을 표현하는 기능은 로봇이 사회의 구성원으로, 혹은 좀 더 지적인 기계로 인정받기 위한 가장 중요한 요소이다[21]. 이러한 이유로 로봇으로 하여금 스스로 감정을 인지, 또는 생성하고 표현하게끔 하는 기술에 관한 연구는 꾸준히 진행되어왔다.

감정형 로봇 연구에 있어서 로봇의 역할이자 주목받고 있는 연구이슈 중 한 가지는 로봇 스스로의 감정 생성이다. 로봇의 감정 생성은 주로 OCC 모델[22], Weiner 모델[23], FLAME 모델[24] 등의 감정 모델을 이용하여 이루어졌다. 감정 모델들을 이용한 로봇의 감정 생성은 인간의 인지 모델에 이론적 기반을 두고 감정을 생성할 수 있다는 장점이 존재한다. 또한 로봇은 생성된 감정을 사용자에게 효과적으로 전달할 수 있어야 한다. 이를 위해 로봇의 감정을 표정, 혹은 행동으로 사용자에게 표현하는 것에 대한 연구가 진행되었다[25].

로봇의 감정 요소는 제스처 기반 대화 관리에 있어서도 유용하게 사용될 수 있다. 이에 착안하여 본 논문에서는 감정형 제스처를 이용하여 로봇이 스스로의 상태를 표현하게끔 하며, 사용자로 하여금 로봇을 더욱 친근하게 느끼게 해 보다 사실적인 상호작용을 유도한다.

3. 제안하는 방법

로봇의 보다 유연하고 사실적인 제스처 생성을 위해 본 논문에서는 대화 시스템, 감정 시스템, 행동 생성 시스템으로 구성된 시스템을 제안한다. 그림 1은 제안하는 시스템의 모식도를 나타낸다. 로봇은 대화로부터 감정을 생성하고 대화 도중 대화 상황과 생성된 감정 상태 표현에 적합한 제스처를 자연스럽게 생성한다.

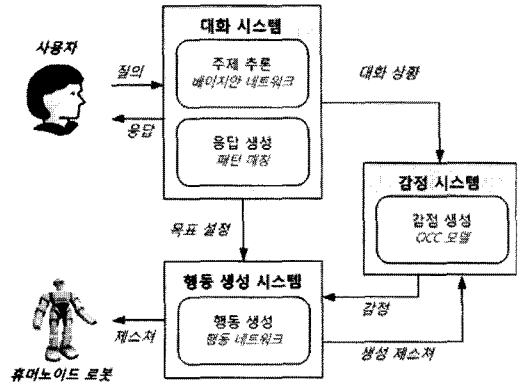


그림 1 제안하는 시스템 구조

3.1 대화 시스템

대화 시스템은 사용자와의 대화를 통해서 필요한 정보를 추출하고 사용자가 어떠한 답을 원하는지 그 의도를 추론하여 에이전트가 적절한 서비스를 제공할 수 있도록 돕는다. 본 논문의 대화 시스템에서는 베이지안 네트워크와 패턴 매칭의 2 단계 추론을 통해서 사용자의 질의를 처리한다. 베이지안 네트워크는 사용자 질의를 수행하기 위한 패턴-답변쌍의 범주인 사용자 대화의 주제를 추론하기 위하여 사용되고, 키워드를 이용한 패턴 매칭[26]은 선택된 주제에 적합한 패턴-답변쌍을 선택된다. 에이전트는 2 단계 추론을 통해서 사용자의 의도를 깊이 있게 분석하며, 주제별로 지식구조를 나누어 시스템의 확장성과 휴대성을 높여준다.

3.1.1 베이지안 네트워크를 이용한 주제 추론

베이지안 네트워크는 불충분한 정보를 표상하거나 추론하는 대표적인 방법이다. 본 논문에서는 대화 도중 발생하는 생략과 축약 등의 불확실성을 극복하기 위해 계층적 베이지안 네트워크를 사용한다. 그림 2는 대상영역을 계층적으로 설계한 베이지안 네트워크를 보여준다.

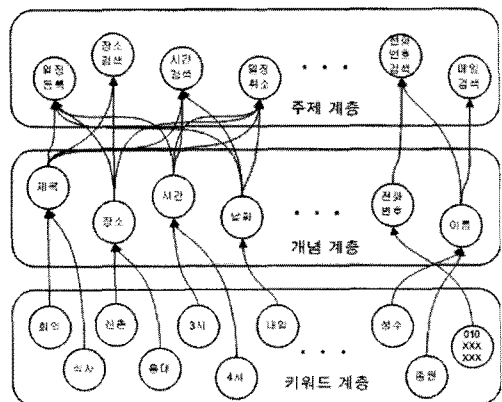


그림 2 베이지안 네트워크의 계층적 주제 설계

키워드 계층의 노드들은 대상영역에서 사용되는 키워드들로 구성되어, 그 값은 각 키워드와 현재 대화와의 연관성을 나타내며, 개념(Concept) 계층의 노드들은 대상영역에 존재하는 속성-값의 쌍으로 구성되고, 그 값은 속성에 대한 값의 존재 가능성을 의미한다. 그리고 주제 계층의 노드들은 최종적으로 사용자가 알고자 하는 정보를 의미한다. 개념 계층과 키워드 계층 간의 연결은 개념 계층의 속성-값의 쌍과 이와 관련된 키워드들과 연결이 되고, 주제 계층과 개념 계층의 연결은 주제 노드의 정보를 제공하기 위해 필요한 속성-값의 쌍과 연결된다.

대상영역의 주제를 계층적으로 분류하여 모델링하면 사용자 의도를 몇 단계로 분석하여 사용자 의도를 세부적으로 파악하도록 대화를 유도할 수 있으며, 공유지식을 설계하여 계층 간의 이동시에 발생하는 각 계층에서의 사용자의 의도에 대한 모호함을 최소화한다[27]. 뿐만 아니라 사람이 직접 대상영역을 분석하고 네트워크를 설계할 때 개념적 이해의 용이함을 제공한다. 그림 3은 언어 해석 모듈에서의 베이지안 네트워크의 추론과정이다.

입력 질의에서 포함된 키워드는 베이지안 네트워크의 추론을 위한 증거변수로 사용되며, 단계 1에서 보듯이, 이전 대화를 기억하기 위해 입력 질의에는 포함되지 않은 키워드는 그 확률값을 일정 비율로 줄인다. 추론이 끝나면 확률값이 가장 높은 주제 노드가 선택되고 그 확률값이 임계치를 넘는다면 입력 질의의 주제가 된다.

- 대상영역:  
키워드집합  $K = \{k_1, k_2, \dots, k_n\}$   
개념집합  $C = \{c_1, c_2, \dots, c_m\}$   
주제집합  $T = \{t_1, t_2, \dots, t_l\}$
- 추론과정:  
단계 1: 사용자 질의에 포함된 키워드 노드의 확률값을 1로, 포함되지 않은 것을 이전 값에  $\omega$ 를 곱함(문맥 유지 가능,  $0 \leq \omega < 1$ )  
단계 2: 베이지안 추론 알고리즘 실행, 노드의 확률값 계산  
a) 개념 노드의 확률값 계산:  
$$P(c_i|pa(c_i)) = \prod_{j=1}^n P(c_i|k_j),$$
  
(단,  $1 \leq i \leq m$ ,  $k_j$ 는  $c_i$ 의 부모)  
b) 주제 노드의 확률값 계산:  
$$P(t_i|pa(t_i)) = \prod_{j=1}^m P(t_i|c_j),$$
  
(단,  $1 \leq i \leq l$ ,  $c_j$ 는  $t_i$ 의 부모)  
단계 3: 임계값을 넘는 가장 높은 주제 계층의 노드를 선택, 선택되는 노드가 없다면 주제 추론 실패

그림 3 베이지안 네트워크의 동작과정

### 3.1.2 키워드 패턴 매칭을 통한 답변 선택

베이지안 네트워크에서 입력 질의에 대한 주제가 선택되면, 관련된 지식구조와 키워드 매칭을 수행하여 적절한 패턴-답변쌍을 찾는다.

보통 패턴-답변쌍이 많고 규칙이 복잡해지면 정보의 중복으로 인해 패턴 매칭의 성능이 저하되지만, 본 논문에서는 전체 패턴-답변쌍을 주제별로 나누고 베이지안 네트워크의 추론을 통해 질의 주제를 선택함으로써 패턴 매칭에서 고려해야 하는 패턴-답변쌍의 양과 정보의 중복을 축소한다. 패턴-답변쌍은 그림 4의 정의에 따라 XML 형식으로 저장된다. <topic>은 베이지안 네트워크의 주제 노드의 이름이며, <pattern>은 사용자 질의문의 키워드와 패턴 매칭할 키워드의 순서이고, <response>는 패턴-답변쌍이 선택되었을 때 제공할 답변 정보이다.

키워드 매칭의 점수는 문서 분류에서 자주 사용되는 F-measure로 계산한다. 본 논문에서는 정확도(precision)와 재현율(recall)이 같은 비중을 가지도록  $\alpha$  값을 1로 설정한다. 키워드 매칭을 통하여 적절한 패턴-답변쌍이 존재하면, 이 패턴-답변쌍의 답변정보와 문장계획 트리를 사용하여 답변 문장을 생성한다.

패턴-답변 쌍의 정의	
<pre>&lt;script&gt; := "&lt;script&gt;" &lt;topic&gt; &lt;pattern&gt; {&lt;pattern&gt;} &lt;responses&gt; {&lt;responses&gt;} "&lt;/script&gt;" &lt;topic&gt; := "&lt;topic&gt;" &lt;word&gt; "&lt;/topic&gt;" &lt;pattern&gt; := "&lt;pattern&gt;" &lt;word&gt; "&lt;/pattern&gt;" &lt;response&gt; := "&lt;response&gt;" &lt;action&gt; "&lt;/response&gt;" &lt;action&gt; := &lt;action&gt; {&lt;action&gt;}   "&lt;action&gt;"            {&lt;goal&gt;   &lt;answer&gt;   &lt;function&gt;} "&lt;/action&gt;" &lt;word&gt; := word {&lt;word&gt;} &lt;goal&gt; := "&lt;GOAL&gt;" &lt;word&gt; "&lt;GOAL&gt;" &lt;answer&gt; := "&lt;ANSWER&gt;" &lt;word&gt; "&lt;ANSWER&gt;" &lt;function&gt; := function call</pre>	
패턴-답변 쌍의 예	
<pre>&lt;script&gt; &lt;topic&gt; 전화번호검색 &lt;/topic&gt; &lt;pattern&gt; \$name\$의 전화번호는 어떻게 되지? &lt;/pattern&gt; &lt;response&gt; &lt;goal&gt; 대답하기 &lt;/goal&gt; &lt;function&gt; GetTel\$name\$tel\$ &lt;/function&gt; &lt;answer&gt; \$name\$의 전화번호는 tel\$입니다. &lt;/answer&gt; &lt;/response&gt; &lt;/script&gt;</pre>	

그림 4 패턴-답변쌍의 정의 및 예

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

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

		입력질의	
		포함	미포함
스크립트	포함	A	B
	미포함	C	D

3.2 감정 시스템

감정은 로봇의 자기 표출에 있어서 핵심적인 부분으로써, 로봇의 구현에 있어서 특히 중요시 된다[28]. 본 논문에서는 감정 생성을 위해 OCC모델[22]을 사용한다. OCC모델은 감정 생성을 위해 제안된 표준 모델로써 감정의 인지적 평가 이론에 기반을 두고 있으며, 강한 이론적 근거를 가지고 널리 사용되고 있다[29]. 본 논문에서는 그림 5와 같이 총 14가지의 감정 유형을 지닌 모델을 사용한다.

$$e_t = \begin{cases} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_m \end{bmatrix}, \forall i \in [1, m]: \beta_i \geq 0 & \text{if } t > 0 \\ 0 & \text{if } t = 0 \end{cases} \quad (1)$$

$$c_t = \begin{bmatrix} w_1 f_1 \\ \vdots \\ w_i f_i \end{bmatrix}, \forall i \in [1, n]: w_i \in [0, 1]: f_i = 0, 1 \quad (2)$$

OCC모델의 목표는 각각의 감정 유형의 감정 강도를 시간에 따라 외부 요인을 이용하여 변화시키고 변화된 감정 강도를 구하는 데에 있다[22].  $m$ 개의 감정 유형이 존재할 때, 시간  $t$ 에 대해  $m$ 차원의 감정  $e_t$ 는 수식 (1)과 같이 정의할 수 있다. 각 감정의 강도  $\beta_1 \dots \beta_m$ 은 3가지 유형의 외부 요인인 사건, 행동, 그리고 물건에 의해 정해지며, 각각 고유의 가중치를 가지게 된다.  $n$ 가지의 외부요인에 대해 시간  $t$ 에 따른 외부요인집단  $c_t$ 는 식 (2)와 같이 정의할 수 있다.

외부요인의 값은 시간  $t$ 에서 외부요인의 존재 여부  $f_i$ 값과 각각의 외부요인에 할당된 고유 가중치  $w_i$ 값의 곱으로 표현된다.  $h_t$ 는 시간  $t$ 까지의 외부요인 내력을 의미한다. 이 때, 시간  $t+1$ 에서 새로운 감정  $e_{t+1}$ 은 다음과 같이 계산된다.

$$e_{t+1} = e_t + \Phi(c_t, h_t) + \Pi(e_t) \quad (3)$$

감정 강도의 변화량은 외부요인에 의한 변화량을 계산하는 감정 변화 함수  $\Phi$ 와 내부 요인에 의한 감정의 쇠퇴량을 결정하는 감정 쇠퇴 함수  $\Pi$ 의 합으로 계산된다.

감정 변화 함수는 우선적으로 로봇이 미리 가지고 있는 목표, 표준, 태도를 이용하여 외부요인에 대한 평가를 수행한다. 사건은 로봇의 목표에 부합되는 바람직한 사건인지 여부가 만족/불만족으로 평가되고 행동은 로봇의 표준에 의거하여 동의/비난으로 평가한다. 또한 물건은 로봇의 태도를 반영해 선호/혐오로 평가된다. 특정 외부요인에 대응되는 대한 로봇의 평가 결과는 사전에 정의되며, 외부요인이 입력되면 로봇은 정의된 대응 관계를 통해 평가 결과를 맵핑시킨다. 평가된 외부요인은 특정 감정 유형의 감정 강도에 영향을 미치게 된다. 또한 감정 변화 함수는 외부요인 내력과 내력 함수를 이용하여 바로 전에 얻어진 외부요인에 의한 감정 강도 변화량을 조절한다. 감정 쇠퇴 함수는 시간이 지남에 따라 감정의 강도를 감소시키는 역할을 한다. 각각의 감정 유형은 고유의 감정 쇠퇴 함수를 이용해 다음 단위 시간에 감소될 감정 강도 변화량을 구한다.

몇 개의 감정들이 동시에 활성화될 경우에는 해당 감정들이 결합되어 새로운 감정이 활성화된다. 로봇의 현재 감정은 위의 과정을 거쳐 계산된 감정 강도들 중 강도가 가장 높은 감정을 택하게 된다.

3.3 행동 생성 시스템

행동 생성 시스템은 행동 네트워크[17]를 이용하여 상황에 가장 적합한 행동들을 자연스럽게 생성한다. 행동 네트워크는 행동들 사이의 연결, 외부 환경, 목표를 구성 요소로 하여 현재 상황에 가장 적합한 행동을 선택하는 모델로, 기본 행동, 조건, 그리고 목표로 구성된다[17].

행동 네트워크의 각 행동 노드는 선행조건, 추가조건, 삭제조건, 활성도를 가지고 있다. 행동 간의 내부 연결

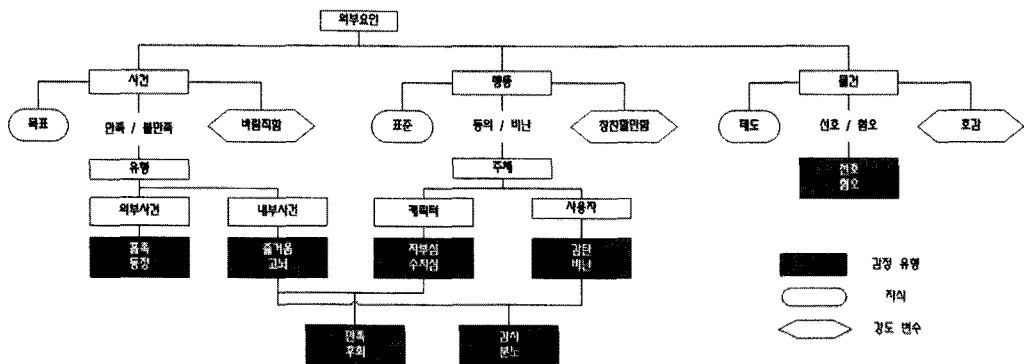


그림 5 제안하는 OCC 감정 모델

은 다음과 같이 정의된다.

선행자 연결 :  $B_i$ 의 선행조건  $p$ 가 거짓이고  $p$ 가  $B_j$ 의 추가조건이다.

후임자 연결 :  $B_i$ 의 추가조건  $p$ 가 거짓이고  $B_j$ 가 실행 가능하며  $p$ 가  $B_j$ 의 선행조건이다.

충돌자 연결 :  $B_i$ 의 선행조건  $p$ 가 참이고  $p$ 가  $B_j$ 의 삭제조건이다.

행동의 활성화도는 우선 외부 환경과 목표로부터 유도되며,  $i$  번째 행동의 활성화도  $A_i$ 의 변화를 다음과 같이 표현할 수 있다.  $w_e$ 와  $w_g$ 는 각각 환경과 목표로부터 활성화도를 유도하기 위한 가중치이고,  $E_{i,n}$ 과  $G_{i,m}$ 은 각각  $i$ 번째 행동에 대해  $n$ 번째 환경과  $m$ 번째 목표의 연결 여부를 나타내며 0과 1의 값을 가진다.

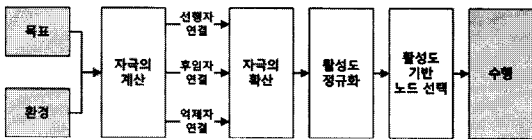


그림 6 행동 네트워크의 행동 생성 과정

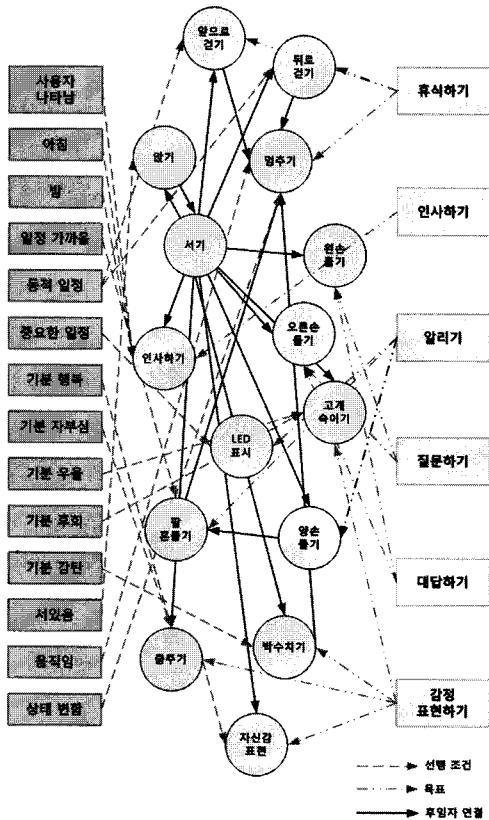


그림 7 설계한 행동 네트워크의 일부분

$$A_i = A_i + \sum_n w_e E_{i,n} + \sum_m w_g G_{i,m} \quad (E_{i,n}, G_{i,m} = 0,1) \quad (4)$$

이후 행동은 다른 행동과의 연결 관계를 이용하여 다음과 같이 행동기리 활성화도를 교환한다.  $w_p$ ,  $w_s$ ,  $w_e$ 는 각각 선행자, 후임자, 충돌자 연결을 통해 활성화도를 교환하기 위한 가중치이며,  $P_{i,j}$ ,  $S_{i,j}$ ,  $C_{i,j}$ 는 각각  $B_j$ 에서  $B_i$ 로의 각각의 연결 여부를 나타낸다.

$$A_i = A_i + \sum_j (w_p P_{i,j} + w_s S_{i,j} - w_c C_{i,j}) \quad (5)$$

$(i \neq j, P_{i,j}, S_{i,j}, C_{i,j} = 0,1)$

행동 네트워크에는 실행 가능한 행동을 결정하기 위한 임계치가 있으며 선행조건이 모두 참인 행동 중 활성화도가 임계치를 넘는 행동을 선택하게 된다. 임계치를 넘는 행동이 없을 경우에는 행동이 선택될 때까지 임계치를 일정량 감소시켜가며 행동을 선택한다. 그림 6은 행동 네트워크의 행동 생성 과정을 보여준다.

본 논문에서 설계된 행동 네트워크는 사용자 인식 상태, 시간, 일정, 로봇 상태 등을 환경 요소로 하며 휴식하기, 인사하기, 알리기, 질문하기, 대답하기, 감정 표현하기의 6가지 목표를 가지고 총 29가지의 체스처를 생성한다. 그림 7은 설계한 행동 네트워크의 일부분을 보여준다.

### 3.4 목표 관리

로봇의 목표는 대화 내용이나 감정 상태에 어울리는 대화 생성을 위해 대화 시스템과 감정 시스템에 의해 설정된다. 표 2는 상황에 따른 목표들을 보여준다. 목표는 한 번에 한 가지, 또는 여러 가지를 가질 수도 있다. 로봇은 현재 상황에서 필요한 것을 판단하고 목표로 설정한다.

표 2 상황에 따른 목표

목표	상황
휴식하기	대화 하고있지 않음
인사하기	사용자가 처음 나타남
알리기	중요 알림 사항 발생
질문하기	사용자와 대화중 (질문)
대답하기	사용자와 대화중 (대답)
감정 표현하기	특정 감정이 활성화

### 4. 실험 및 결과

제안하는 시스템은 Microsoft Visual C++ 2008과 Microsoft Visual C# 2008을 이용하여 제작되었으며, Speecys 휴머노이드 로봇 상에서 동작한다. 사용자는 키보드 입력을 통해 로봇과 대화를 수행한다.

대화 시스템의 패턴-답변 쌍은 일정과 연락처의 검색, 추가, 수정, 삭제와 관련한 총 13개의 주제로 구성되며, 각 주제별로 2개에서 10개까지의 총 74개의 패턴-답변

쌍이 정의되었다. 또한, 베이지안 네트워크의 주제 추론을 위해서 패턴-답변 쌍에 정의된 주제에 맞춰서 총 13개의 주제 노드를 구성하였으며, 일정과 연락처에 관련된 총 8개의 속성-값 쌍에 해당하는 개념 노드로 구성된다.

제안하는 시스템의 동작 과정을 분석하기 위해 대화 시나리오를 설계하고 시나리오 상에서 생성되는 제스처를 관찰하였다. 표 3은 설계한 대화 시나리오를 보여준다.

로봇이 처음 사용자를 발견하게 되면 대화 시스템은 행동 생성의 목표로 '인사하기'를 설정하고, '인사' 행동의 활성도가 높아지게 되어 인사로 사용자를 맞이하게 된다.

인사를 마친 후 사용자는 로봇에게 오늘 일정에 대해서 문의하게 된다. 로봇은 문의 내용을 인지하고 오늘 일정을 검색한다. 로봇은 오늘 일정의 종류가 정적인 '미팅'인 것을 파악하고 행동 네트워크의 환경 요소 중 '정적 일정'을 설정한다. 또한 사용자에게 대답하기 위해 '대답하기' 목표가 설정된다. 행동 생성 시스템을 통해 로봇은 현재 상황에 일반적인 대답 제스처가 적합하다는 것을 인지하게 되고, '손 들기' 행동이 활성화되어 한 손을 들면서 사용자에게 대답을 수행한다.

사용자는 일정 확인 후 새로운 일정을 등록하려 하지만, 해당 시간에 이미 다른 일정이 등록되어 있다. 이 상황에서 로봇은 등록 오류가 발생했음을 알려주어야 한다. 따라서, '알리기' 목표가 설정되고, '중요 일정' 환경이 설정된다. 이후 로봇은 양 손을 들어 올림으로써 일정 등록 오류를 알려준다. 또한 현재 존재 일정의 중요도를 고려하여 양 팔을 흔들어준다.

이후 최종적으로 로봇은 시간을 조정하여 일정을 성공적으로 등록한다. 이 때, 로봇은 일정 시간을 조정하고 성공적으로 등록한 행동에 대해 스스로 평가를 내려 '자부심'의 감정이 활성화 되고, '감정 표현' 목표가 설정되며 기존에 설정되어 있던 '알리기' 목표는 해제되어

표 3 대화 시나리오

화자	대화
Speecys	안녕하세요, 무엇을 도와드릴까요?
사용자	오늘 일정 좀 알려줄래?
Speecys	오후 3시에 교수님과 미팅이 있습니다.
사용자	여자친구와 저녁 약속을 잡으려고 하는데, 오후 5시에 가능할까?
Speecys	죄송하지만 해당 시간에는 수업 일정이 있습니다. 해당 일정은 중요 일정으로 잡혀 있기 때문에 다른 일정으로 대체할 수 없습니다. 대신 내일 오후 7시는 가능한데, 오후 7시로 잡아드릴까요?
사용자	그래.
Speecys	내일 오후 7시 여자친구와의 저녁식사 일정이 등록되었습니다.

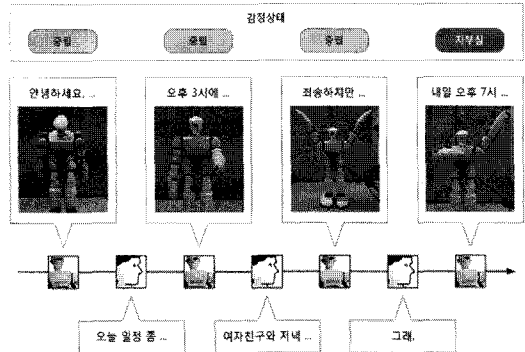


그림 8 대화 흐름에 따른 생성 제스처

'자부심' 제스처를 수행할 확률이 증가하며 '양 팔 흔들기' 제스처의 활성도는 낮아진다. 그림 8은 대화 시나리오의 흐름에 따른 로봇의 생성 제스처를 보여준다.

이후 제안하는 시스템을 사용하였을 때 감정형 제스처가 실제 시스템의 유용성에 미치는 영향을 분석하기 위해 사용성 평가를 수행하였다. 사용성 평가는 10명의 대학원생이 참가하여 진행되었다. 평가를 위해 참가자는 표 4와 같이 다섯 가지의 태스크를 제스처를 사용하지 않는 일반적인 대화 관리 시스템, 일반적 제스처만 사용하는 시스템, 그리고 제안하는 시스템에서 각각 수행한 뒤 시스템에 대한 사용성 평가 질의를 수행하였다.

표 4 사용성 평가를 위해 수행한 태스크

No	수행 태스크
1	로봇이 사용자에게 다가오는 일정을 알림
2	등록된 오늘 일정을 조회
3	등록된 일정을 변경
4	새로운 일정을 등록 (겹치는 일정 없음)
5	새로운 일정을 등록 (겹치는 일정이 존재하여 등록하려는 일정 시간을 변경)

사용성 평가를 위해 피험자들을 대상으로 SUS질의 [30]가 제공되었다. 점수는 각 문항 당 "강한 부정"을 나타내는 1점부터 "강한 긍정"을 나타내는 5점까지 5단계로 나누어서 평가되었다. 표 5는 사용성 평가에 사용된 SUS질의를 보여준다.

사용성 평가 결과, 제안한 감정형 제스처 기반 대화 관리 시스템의 경우 평균 72.1점을 얻었지만, 제스처를 사용하지 않는 일반적인 대화 관리 시스템의 경우 51.3점, 감정이 없이 일반적 제스처만 사용하는 시스템의 경우 63.7점을 얻었다. 이를 통해 기존의 시스템에 비해 제안하는 시스템이 사용성 측면에 있어서 보다 우수함을 증명할 수 있다. 그림 9는 사용성 평가 결과 두 시스템의 비교 결과를 보여준다.

표 5 SUS 질의

No	질의
1	이 시스템을 자주 사용할 것이다.
2	이 시스템이 불필요하게 복잡하다고 생각한다.
3	이 시스템이 사용하기 쉽다고 생각한다.
4	이 시스템을 사용하기 위해서는 전문가의 도움이 필요하다고 생각한다.
5	이 시스템의 다양한 기능이 조직적으로 잘 결합되어 있다고 생각한다.
6	이 시스템이 너무 불안정하다고 생각한다.
7	많은 사람들이 이 시스템의 사용법을 빠르게 익힐 것이라고 생각한다.
8	이 시스템이 방해가 된다고 생각한다.
9	이 시스템을 사용하는 데에 큰 자부심을 느낄 것이다.
10	이 시스템을 계속 사용하기 위해서 많은 것들을 배워야 한다고 생각한다.

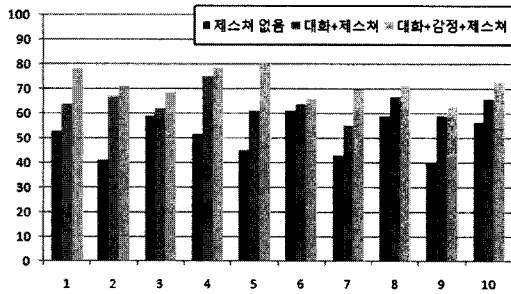


그림 9 사용성 평가 결과

## 5. 결론

본 논문에서는 로봇과 사람 간의 유연하고 사실적인 상호작용을 위해 감정형 제스처 기반 대화 관리 시스템을 제안하였다. 제안한 시스템은 대화 상황에 따라 로봇의 감정 상태를 생성하고, 상황과 감정에 적합한 제스처를 행동 네트워크를 통해 생성한다. 제안하는 시스템의 동작 과정을 분석하기 위해 시나리오 기반 생성 제스처 분석이 이루어졌으며, 유용성을 평가하기 위해 기존의 감정형 제스처를 사용하지 않는 대화 관리 시스템과의 SUS 테스트 결과 비교가 이루어졌다.

그러나 아직은 로봇이 스스로의 감정을 생성하고 반영할 뿐, 사용자의 감정을 인식하여 대화에 반영하고 대화를 이끌어가는 것은 불가능하다. 향후에는 감정 생성뿐만 아니라 대화를 통해 사용자의 감정을 인식하여 대화 및 제스처 생성에 반영하는 연구가 이루어져야 할 것이다. 또한 제시된 감정형 제스처가 사용성에 미치는 영향에 대한 분석뿐만 아니라 각 모듈의 성능 및 정확도에 대한 분석 역시 이루어져야 할 것이다.

## 참고 문헌

- [1] A. Green and K. Eklundh, "Designing for learnability in human-robot communication," *IEEE Trans. Ind. Electron.*, vol.50, no.4, pp.644-650, 2003.
- [2] T. Fong, C. Thorpe, and C. Baur, "Robot, asker of questions," *Robot. Auton. Syst.*, vol.42, no.3-4, pp.235-243, 2003.
- [3] H. Prendinger and M. Ishizuka, "Let's talk! Socially intelligent agents for language conversation training," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol.31, no.5, pp.465-471, 2001.
- [4] D. Sanford and J. Roach, "A theory of dialogue structures to help manage human-computer interaction," *IEEE Trans. Syst., Man, Cybern.*, vol.18, no.4, pp.567-574, 1988.
- [5] A. Ratnaparkhi, "Trainable approaches to surface natural language generation and their application to conversational dialog systems," *Computer Speech and Language*, vol.16, no.3-4, pp.435-455, 2002.
- [6] M. A. Walker, O. C. Rambow and M. Rogati, "Training a sentence planner for spoken dialogue using boosting," *Computer Speech and Language*, vol.16, no.3-4, pp.409-433, 2002.
- [7] E. Levin, R. Pieraccini and W. Eckert, "A stochastic model of human-machine interaction for learning dialog strategies," *IEEE Trans. on Speech and Audio Processing*, vol.8, no.1, pp.11-23, 2000.
- [8] I. Bulyko and M. Ostendorf, "Efficient integrated response generation from multiple targets using weighted finite state transducers," *Computer Speech and Language*, vol.16, no.3-4, pp.533-550, 2002.
- [9] N. Mitsunaga, C. Smith, T. Kanda, H. Ishiguro, and N. Hagita, "Adapting robot behavior for human-robot interaction," *IEEE Trans. on robotics*, vol.24, no.4, pp.911-916, 2008.
- [10] W. Breitfuss, H. Prendinger, and M. Ishizuka, "Automatic generation of gaze and gestures for dialogues between embodied conversational agents," *International Journal of Semantic Computing*, vol. 2, no.1, pp.71-90, 2008.
- [11] S. Franklin, and A. Graesser, "Is it an agent, or just a program? : A taxonomy for autonomous agents," *Lecture Notes in Computer Science*, vol.1193, pp.21-35, 1997.
- [12] S. Macskassy and S. Stevenson, "A conversational agent," Master essay, Rutgers University, 1996.
- [13] P. Nugues et al., "A conversational agent to navigate in virtual worlds," *Proc. of the 11th Workshop on Language Technology*, pp.23-33, 1996.
- [14] J.-H. Hong and S.-B. Cho, "A two-stage bayesian network for effective development of conversational agent," *Lecture Notes in Computer Science*, vol.2690, pp.1-9, 2003.
- [15] C.-J. Lee, S.-K. Jung, S.-H. Kim, and G. C.-B. Lee, "Example-based dialog modeling for practical



multi-domain dialog system," *Speech Communication*, vol.51, no.5, pp.466-484, 2009.

[16] P. Maes, "How to do the right thing," *Connection Science Journal*, vol.1, no.3, pp.291-323, 1989.

[17] M. N. Niculescu and M. J. Mataric, "Extending behavior-based systems capabilities using an abstract behavior representation," *AAAI Fall Symposium on Parallel Cognition*, pp.27-34, 2000.

[18] M. N. Niculescu and M. J. Mataric, "A hierarchical architecture for behavior-based robots," In *Proc. of First Int. Joint Conf. on Autonomous Agents and Multi-Agent Systems*, pp.227-233, 2002.

[19] A. Khoo and R. Zubek, "Applying inexpensive AI techniques to computer games," *IEEE Intelligent Systems*, vol.17, no.4, pp.48-53, 2002.

[20] D. Evans, "Emotion," Oxford University Press, 2001.

[21] A. Ortony, G. Clore, and A. Collins, "The cognitive structure of emotions," Chicago University Press, 1998.

[22] C. D. Elliott, "The affective reasoner: A process model of emotions in a multi-agent system," Northwestern University, 1992.

[23] M. El-Nasr, T. R. Loerger, and J. Yen, "Emotionally Expressive Agents," 1998.

[24] C.-T. Cheng, Y.-T. Yang, S.-H. Miao, and C.-C. Wong, "Motion and emotional behavior design for pet robot dog," *Lecture Notes in Computer Science*, vol.5744, pp.13-22, 2009.

[25] S.-S. Lim, S.-B. Cho, "Automatic construction of hierarchical Bayesian networks for topic inference of conversational agent," *Journal of Korea Information Science Society: Software and Applications*, vol.33, no.10, pp.877-885, Oct. 2006.

[26] E. Horvitz and T. Paek, "A computational architecture for conversation," *Proc. 7th Int. Conf. User Modeling*, pp.201-210, 1999.

[27] C. Bartneck, "Integrating the OCC model of emotions in embodied characters," In *Proc. of the Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges*, 2002.

[28] B. R. Steunebrink, M. Dastani, and J.-J. Ch. Meyer, "A formal model of emotion-based action tendency for intelligent agents," *Lecture Notes in Computer Science*, vol.5816, pp.174-186, 2009.

[29] J. Brooke, "SUS: A quick and dirty usability scale," *Usability evaluation in industry*, 1996.



윤 종 원

2009년 연세대학교 컴퓨터과학과(학사)  
2009년~현재 연세대학교 컴퓨터과학과  
석사과정. 관심분야는 패턴 인식 및 지능  
형 에이전트

임 성 수

정보과학회논문지 : 소프트웨어 및 응용  
제 37 권 제 1 호 참조

조 성 배

정보과학회논문지 : 소프트웨어 및 응용  
제 37 권 제 1 호 참조