계층적 베이지안 네트워크를 사용한 서비스 로봇과 인간의 상호 주도방식 의사소통

(Mixed-Initiative Interaction between Human and Service Robot using Hierarchical Bayesian Networks)

송 윤 석 † 홍 진 혁 † 조 성 배 ††
(Youn-Suk Song) (Jin-Hyuk Hong) (Sung-Bae Cho)

요 약 일상 생활에서 서비스 로봇이 효과적으로 사람들의 업무를 보조하기 위해서는 인간과의 상호작용이 매우 중요하다. 그 중 대화는 인간과 로봇이 보다 유연하고 풍부한 의사전달을 하는데 유용하다. 전통적인 로봇 연구에서는 의사소통 방법으로 명령과 같은 간단한 질의 등을 사용하였으나 실제 사람들의 대화는 보다 복잡하고 다양하며 배경 지식이나 대화의 문맥 등에 의해 중요한 정보가 생략되기도 한다. 따라서 동일한 질의라도 다양한 의미를 갖기 때문에 정확한 해석을 위해서는 이를 다루어야 한다. 본 논문에서는 계층적 베이지안 네트워크를 사용하여 '상호-주도형' 의사 소통 방식을 서비스 로봇에 구현함으로써 대화의 모호성을 처리하는 방법을 제안한다. 또한 서비스 로봇의 시뮬레이션과 사용자 평가를 통해 제안하는 방법의 유용성을 확인하였다.

키워드 : 서비스 로봇, HRI, 계층적 베이지안 네트워크, 상호 주도 방식

Abstract In daily activities, the interaction between humans and robots is very important for supporting the user's task effectively. Dialogue may be useful to increase the flexibility and facility of interaction between them. Traditional studies of robots have only dealt with simple queries like commands for interaction, but in real conversation it is more complex and various for using many ways of expression, so people can often omit some words relying on the background knowledge or the context of the discourse. Since the same queries can have various meaning by this reason, it is needed to manage this situation. In this paper we propose a method that uses hierarchical bayesian networks to implement mixed-initiative interaction for managing vagueness of conversation in the service robot. We have verified the usefulness of the proposed method through the simulation of the service robot and usability test.

Key words : Service Robot, HRI, Hierarchical Bayesian Networks, Mixed-Initiative

1. 서 론

인간-컴퓨터 상호작용(HCI)과 밀접하게 관련된 인간-로봇 상호작용(HRI)은 최근 자율로봇 분야에서 많이 연구되고 있다. 과거에는 서비스 로봇과 같은 자율로봇은 보통 사용자의 직접적인 조작에 의해 작동되었다[1]. 따라서 기존의 로봇 연구는 사람과 로봇 사이에서 이루어지는 상호작용보다는 로봇의 자율성에 더 많은 관심이 있었다. 하지만 최근 가정이나 사무실에서 서비스 로봇의 사용이 증가함에 따라 사람과 보다 자연스럽게 상호작용하는 방법이 요구되고 있다. 이들 서비스 로봇들은 인간과 함께 거주하며 주로 배달이나, 알림, 안내 등과 같은 서비스를 제공하며 효과적인 업무보조를 위해서는 정확한 수행뿐만 아니라 유연한 의사소통 기능이 필요하다[2,3].

이를 위해 자연어를 통한 대화가 사람과 로봇의 상호작용에서 효과적인 방법으로 많은 관심을 받고 있다[4,5]. 실제 대화 속에서 사람들은 문맥을 통해 그 의미를 이해하는 경우가 많기 때문에 몇 가지 정보가 생략되어도 자연스러운 대화와 정확한 의도 전달이 가능하

· 이 논문은 산업자원부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 일환으로 수행되었습니다.

† 학생회원 : 연세대학교 컴퓨터과학과
corlary@sclab.yonsei.ac.kr
hjinh@sclab.yonsei.ac.kr
†† 종신회원 : 연세대학교 컴퓨터과학과 교수
sbcho@cs.yonsei.ac.kr

논문접수 : 2005년 2월 1일
심사완료 : 2006년 2월 1일

다[6]. 따라서 로봇이 자연스러운 대화를 하기 위해서는 이 때 발생하는 모호함과 불확실성을 다루는 것이 필요하며 이를 위해 HCI 분야에서는 상호 주도적인 접근 방법이 연구되고 있다[7,8]. 이는 사람과 시스템이 지속적인 의사 소통을 통해 점진적으로 문제를 해결해 가는 방법으로써, 시스템이 상황을 추론할 때 발생할 수 있는 불확실성을 사용자와 함께 해결하게 한다.

본 논문에서는 계층적 베이지안 네트워크를 사용하여 로봇의 상호주도형 대화처리를 구현하고 시뮬레이션과 사용자 평가를 통해 제안하는 방법의 유용성을 검증하였다.

2. 서비스 로봇과 대화형 모델

서비스 로봇은 로봇의 중요한 응용 분야로서, 가정이나 사무실에서 사람과 상호작용을 하며 사용자의 업무를 보조한다[9]. 서비스 로봇의 대표적인 연구로는 홈 서비스에 대한 정보 제공, 물건 배달 서비스, 사람 안내하기 등이 있으며[2,10], 실제 환경에서는 Nursebot, HelpMate, Cero, Rhino, PaPeRo등과 같은 서비스 로봇이 개발되기도 하였다.

과거 연구에서는 로봇의 상호작용 능력보다는 자율적인 행동이나 서비스 자체에 주안점을 두었기 때문에 사람과 로봇의 상호작용은 매우 단순하였다. 하지만 로봇이 복잡하고 사회적인 기능을 하게 되면서 고성능의 로봇이라도 의사소통 문제로 성능이 저하될 수 있어 인간과 로봇의 의사소통과 관련된 연구가 중요해지고 있다[3]. 그 중 대화를 통한 의사소통은 가장 주목 받고 있는 연구 중 하나이다. Hutternrauch 등에 의해 조사된 바로는 82%의 사용자가 여러 의사소통 방법 중 '말'을 선호하는 것으로 나타났고, 최근 다양한 대화 모델이 로봇 분야에 적용되고 있다[9]. Skubic 등은 환경의 공간적 관계를 고려한 정교한 대화처리 시스템을 제안하여 대화를 통한 공간 모델링을 통해 서비스 로봇의 성능을 향상 시켰다[11]. 그밖에 대화와 관련된 다양한 기법들이 HCI분야에서 고안되어 왔으며 이와 관계된 기술로는 패턴매칭 기술, 준비된 대본을 가진 모델, 프레임 기반 모델, 유한상태 모델, 계획기반 모델 등이 있다[12, 13].

시스템은 주도권에 따라 시스템 주도형 모델과 사용자 주도형 모델로 나뉘어진다. 시스템 주도형 모델은 시스템이 필요로 하는 정보를 얻기 위해 미리 정해진 절차에 따라 사용자가 정보를 제공한다. 반면 사용자 주도형 모델은 사용자가 주도권을 가지고 선호하는 방향으로 의사 소통을 한다. 두 방법 모두 전적으로 시스템이나 사용자에게 의존하기 때문에 노이즈가 많은 환경에서 성능이 저하될 수 있고 사용자의 상황을 고려한 서비스가 어렵다는 면에서 한계를 가지고 있다. 이에 두 가지 방법을 함께 사용하는 상호 주도형 모델이 제안되었다. 이와 같은 모델은 사용자와 시스템이 서로 주도권을 가지며 필요한 경우 질의를 통해 서로의 의도를 확인할 수 있는데, 이를 대화 모델에 적용 시 기존의 방법으로는 처리 되기 어려운 '덥다' 와 같은 사용자의 의도를 알기 어려운 표현이 인식 되었을 경우에도 처리가 가능하다.

Saverio[14] 등은 개인화된 웹 검색 서비스를 사용자에게 제공하기 위해 상호주도형 모델을 사용하였다. 이 모델은 사용자가 시스템으로부터 검색 화면을 제공받은 후 관심 있는 항목에 대한 정보를 시스템에게 제공하면 이에 맞도록 시스템은 보다 축소된 정보를 사용자에게 제공한다. 이들은 상호주도형 시스템을 위해 도메인에 대한 전체적인 검색 규칙을 설계하고 이러한 규칙을 사용자의 요청에 맞는 부분적인 규칙으로 변환하여 제공하였다. 이와 같은 방법은 적은 수의 키워드를 처리할 때는 좋지만 대화처럼 많은 정보가 동시에 제공되거나 생략될 수 있는 상황에서는 키워드들의 조합에 대한 처리를 해야 하므로 규칙의 충돌 문제 등으로 인해 설계가 복잡해지며 설계되지 않는 상황 처리에 약하다[15, 16].

Meng[6] 등은 베이지안 네트워크의 후진 추론을 사용하여 외화 환전 도메인에서 발생할 수 있는 대화를 모델링하였다. 이들은 환전을 위해 필요한 정보가 빠져 있거나 불필요한 내용을 알기 위해서 미리 모델링된 베이지안 네트워크를 후진 추론하고, 추론된 결과와 제공된 질의를 비교하여 빠진 내용과 불필요하다고 생각되는 내용을 찾는다. 이를 바탕으로 필요한 질의를 사용자에게 하여 상호 주도적인 대화가 발생하도록 하였다. 이와 같은 방법은 대화 과정보다는 특정 목적에 해당되는 내용들의 관계를 모델링하는 방법으로 설계자가 대화의 추론 과정을 직관적으로 모델링하지는 않는다.

Hong[17] 등은 대화형 에이전트의 대화를 모델링하기 위해서 베이지안 네트워크와 오토메타를 함께 사용하여 대화 과정을 대화의 주제와 문형을 분석하는 두 단계로 나누어 사용자로부터 이야기를 듣고 이에 대한 답변을 찾는 방법을 제안하였다. 이 때 계층적인 구조를 갖는 베이지안 네트워크를 모델링하여 각 단계에 필요한 정보를 세부적으로 추론해 가고 필요한 경우 사용자에게 질의를 할 수 있도록 했다. 이와 같은 방법은 모델링을 용이하게 하였지만, 대화형 에이전트를 위한 모델에 한정 되었다.

본 논문에서는 실내 환경에서 서비스 로봇이 사용자와의 대화를 통해 사용자의 의도를 파악하고 서비스를 수행하기 위한 베이지안 네트워크 기반 상호 주도형 대

화 모델을 제안한다. 서비스 제공을 위한 모델을 구성하기 위해 서비스 추론 단계를 목적과 대상으로 분류하였고, 이러한 과정을 통해 최종적으로 로봇이 제공할 서비스가 결정되도록 했다. 이와 같은 방법은 베이지안 네트워크를 필요한 서비스에 맞도록 나누어 설계할 수 있게 하므로 확장성을 갖게 한다. 각 단계에서 사용되는 베이지안 네트워크는 계층적으로 설계되어 사용자의 질의를 증거로 필요한 내용을 추론한다. 이 때, 증거가 부족하다고 판단되면 베이지안 네트워크에서 하위 계층에 있는 노드의 확률 값을 통해 사용자에게 필요한 질의를 찾는 방법으로 상호주도적인 대화가 발생한다. 이와 같은 방법은 추론 과정을 직관적으로 표현하므로 베이지안 네트워크 설계가 용이하다.

3. 계층적 베이지안 네트워크를 사용한 상호 주도형 의사소통

상호 주도형 의사소통을 위해 사용된 계층적 베이지안 네트워크의 전체적인 구조는 의미 요소(primitive semantics) 노드와 목표 노드, 중간 목표 노드로 구성된다. 의미 요소 노드는 사용자 질의문과 환경 분석하여 사용자의 의도를 특정 도메인에 맞게 이해하기 위해 사용된다. 목표 노드는 베이지안 네트워크를 통해 최종적으로 추론되는 결과를 가리키는데 시스템은 제한된 응용 도메인에서 작동하기 때문에 한정된 목표를 바탕으로 동작한다. 중간 목표 노드는 상호 주도형 의사소통 방식을 위해 의미 요소층과 목표층 사이에 위치한다. 목표는 의미 요소로부터 직접 추론되지 않고 상호 주도 방식을 위해 중간 목표층을 거쳐 추론된다. 만약 서비스 목표가 결정되기에 충분한 단서를 얻지 못했다면 중간 목표들 중에서 가장 높은 확률 값을 갖는 노드를 찾고 이 노드와 연결된 목표 노드들에 대한 질문을 사용자에게 한다. 이처럼 로봇은 사용자가 제공한 증거에 따라 목표와 관련된 질문을 사용자에게 함으로써 상호 주도적 방법으로 상황을 해결한다.

사용자에 의해서 제공되는 의미 요소는 서비스 도메인에 따라 제한적이지만 다양하다. 이 때 베이지안 네트워크를 사용하여 설계하는 경우, 의미 요소와 목표들 간의 관계를 확률적으로 표현하여 사용자에 의해 제공된 정보가 적더라도 시스템이 동작할 수 있다. 반면, 이를 규칙을 기반으로 구현할 경우 규칙으로 정해지지 않은 상황에 대한 처리가 어렵다는 단점이 있다. 이를 해결하기 위해서는 각 의미 요소들이 독립적으로 발생한다는 가정하에 증거가 발생할 수 있는 모든 경우에 대한 규칙을 정해주어야 하는데, 이는 의미 요소가 많아질수록 설계가 복잡하고 수정이 어렵다[15,16]. 이에 본 논문에서는 베이지안 네트워크를 사용하여 상호 주도형 대화를 모델링하였으며 계층적인 구조를 통해 설계를 보다 용이하게 하였다. 그림 1과 표 1은 제안하는 계층적 베이지안 네트워크의 구조와 관련 내용을 보여준다.

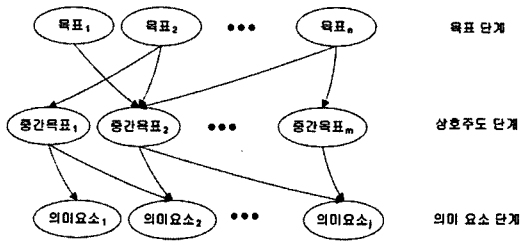


그림 1 상호 주도형 HRI를 위한 계층적 베이지안 네트워크

표 1 계층적 베이지안 네트워크

범주	단계	목적
목표(G_i)	목표 단계	최종적인 목표 결정
중간목표(SG_j)	상호 주도 단계	상호 주도 방식 문제 풀이
의미요소(Pr_k)	의미 요소 단계	추론을 위한 증거

각 단계에 있는 노드들은 0과 1사이의 확률 값을 가질 수 있으며 의미 요소들은 사용자 질의와 환경으로부터 해당되는 정보가 주어질 때 'true' 값을 갖는다. 의미 요소를 통해 각 중간 목표와 목표 노드의 확률 값이 아래에 나와 있는 식 $P(SG|Pr)$와 $P(G|SG)$를 통해 추론된다.

$$P(G_i = 1 \mid SG) = \frac{P(SG \mid G_i = 1)P(G_i = 1)}{P(SG)}$$

$$= \frac{\prod_{k=1}^{M} P(SG_k \mid G_i = 1)P(G_i = 1)}{\prod_{k=1}^{M} P(SG_k \mid G_i = 0)P(G_i = 0) + \prod_{k=1}^{M} P(SG_k \mid G_i = 1)P(G_i = 1)}$$

$$P(SG_i = 1 \mid Pr) = \frac{P(Pr \mid SG_i = 1)P(SG_i = 1)}{P(Pr)}$$

$$= \frac{\prod_{k=1}^{L} P(P_k \mid SG_i = 1)P(SG_i = 1)}{\prod_{k=1}^{L} P(Pr_k \mid SG_i = 0)P(SG_i = 0) + \prod_{k=1}^{L} P(Pr_k \mid SG_i = 1)P(SG_i = 1)}$$

추론된 목표 G는 $P(G_i|SG)$ 값 중 가장 크고 기준 값 이상인 노드이다. 만약 이와 같은 조건을 만족시키는 노드가 없다면 중간 목표 층을 추론하여 $P(SG_i|Pr)$ 중 가장 큰 값이면서 기준 값 이상을 갖는 노드를 통해 사용자의 정확한 의도를 이해하는데 필요한 질의를 한다. 상호 주도형 대화는 이와 같은 과정을 반복하며 사용자로부터 필요한 의미 요소를 얻어 목표 노드가 선택되도록 한다. 중간 노드가 선택되지 못하여 이러한 방법으로 처

리되지 않는 경우, 이를 서비스 처리 도메인 영역 바깥(Out of Domain)에 있는 것으로 간주하고 대화를 종료한다. 전체적인 상호주도형 의사소통 방식 처리 과정은 그림 2와 같으며 이와 같은 과정은 베이지안 네트워크를 사용한 모든 대화처리에 적용된다.

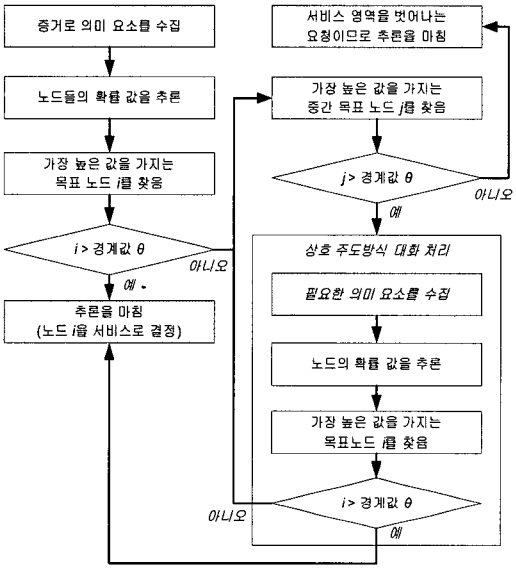


그림 2 상호주도 방식 대화 처리 과정

4. 서비스 로봇에서의 상호주도형 의사소통 처리

본 논문에서는 실내 서비스를 위해 그림 3과 같은 상호주도형 서비스 로봇을 제안한다. 서비스 로봇은 음성과 센서를 통해서 입력을 받으며 서비스 추론을 위해 두 가지 종류의 계층적 베이지안 네트워크를 사용한다. 입력으로 들어온 정보는 사용된 베이지안 네트워크의 의미 요소로 사용되며 사용된 계층적 베이지안 네트워크는 크게 서비스 종류 결정과 서비스 대상을 결정하기 위한 두 가지 종류로 나누어진다. 최종적인 서비스는 이 두 종류의 베이지안 네트워크를 차례로 추론하여 결정되며 각 과정마다 증거가 불충분할 경우 상호 주도 방식으로 사용자에게 필요한 정보를 요청한다.

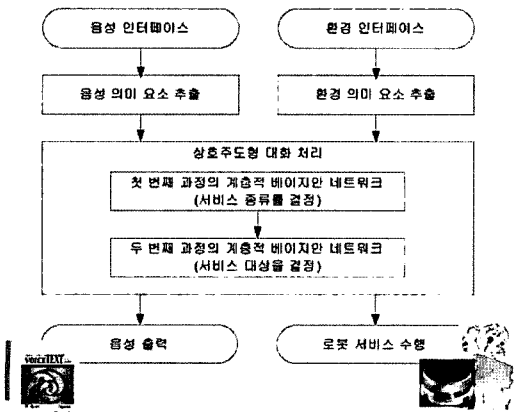


그림 3 제안하는 서비스 로봇 구조

4.1 홈 서비스 환경

실제 가정 환경에서 제공될 수 있는 서비스는 그 종류가 많다. 따라서 본 논문에서는 상호 주도형 서비스 로봇의 기능을 '켜기/끄기', '열기/닫기'로 한정하고 그와 같은 기능이 적용될 수 있는 대상들을 정하여 표 2와 같은 가정 환경을 구성하였다. 구현된 서비스 로봇이 할 수 있는 서비스 종류는 불 켜기/끄기, 오디오 켜기/끄기, TV 켜기/끄기, 에어컨 켜기/끄기, 창문 열기/닫기로 10가지이며 대상은 최대 7개까지 가질 수 있도록 하였다.

표 2 서비스 종류 및 대상

서비스 종류	서비스 대상
불 켜기	방 A, 방 B, 방 C, 방 D, 방 E, 방 F, 방 G
불 끄기	방 A, 방 B, 방 C, 방 D, 방 E, 방 F, 방 G
창문 열기	방 A, 방 B, 방 C, 방 D, 방 E, 방 F, 방 G
창문 닫기	방 A, 방 B, 방 C, 방 D, 방 E, 방 F, 방 G
오디오 켜기	방 A, 방 C, 방 F
오디오 끄기	방 A, 방 C, 방 F
TV 켜기	방 D, 방 F, 방 G
TV 끄기	방 D, 방 F, 방 G
에어컨 켜기	방 F
에어컨 끄기	방 F

4.2 서비스 결정 과정

서비스는 두 종류의 베이지안 네트워크 추론 과정을 통해 결정된다. 첫 번째 과정은 사용자의 질의를 통해 서비스 종류를 결정하기 위한 과정이며 이를 위해 사용된 계층적 베이지안 네트워크는 한 개이다. 여기에는 10가지 종류의 서비스를 위해서 10개의 목표 노드가 있으며 질의와 환경 정보에 대한 의미 요소들과 상호 주도형 의사 소통을 하기 위한 중간 목표들이 네트워크 안에 함께 구성된다. 두 번째 과정은 구체적인 대상을 결정하기 위한 과정으로써 서비스에 따라 대상을 추론하기 위해 계층적 베이지안 네트워크가 사용되었다. 이런 모델은 서비스 종류와 각 대상들의 추론을 모듈화 함으로써 서비스 종류와 대상의 증가하더라도 제안된 방법의 적용과 확장에 유용하다.

표 3은 서비스 종류 추론을 위한 계층적 베이지안 네트워크의 노드들이며 그림 4는 전체적인 구조이다. 여기서 사용된 계층적 베이지안 네트워크의 의미 요소들은 기능 의미 요소와 표현 의미 요소, 환경 의미 요소로 나

표 3 서비스 종류를 추론하기 위한 베이지안 네트워크 노드

범주		내용
의미 요소	기능요소	켜라, 꺼라, 열어라, 닫아라, 불, 창문, TV, 오디오, 에어컨,
	표현요소	춥다, 덥다, 시끄럽다, 어둡다
	환경요소	집 안/밖 온도차, 집 안/밖 밝기차, 집 안/밖 소음차, 모든 불 켜짐, 모든 창문 열림, 모든 TV 켜짐, 모든 오디오 켜짐, 모든 에어컨 켜짐
중간 목표	기능관련	켜라, 꺼라, 닫아라, 열어라, 불, 창문, TV , 오디오, 에어컨
	표현관련	밝다, 춥다1, 춥다2, 어둡다1, 어둡다2, 덥다1, 덥다2, 시끄럽다1, 시끄럽다2
목표	서비스목표	불 켜기, 불 끄기, 창문 열기, 창문 닫기, 오디오 켜기, 오디오 끄기, TV 켜기, TV 끄기, 에어컨 켜기, 에어컨 끄기

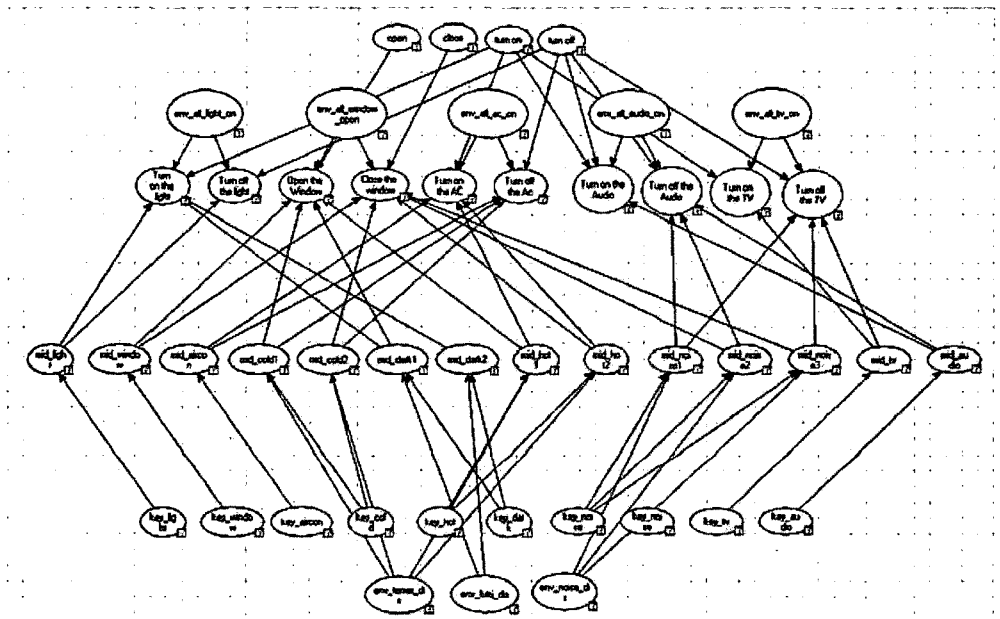


그림 4 서비스 종류 추론을 위한 계층적 베이지안 네트워크

누어 지는데, 기능 의미 요소는 직접적으로 서비스와 관련되어 사용되는 의미 요소를 뜻하고 표현 의미 요소는 사용자들이 어떤 상황이나 서비스를 위해서 사용할 수 있는 간접적인 표현들을 의미하며, 환경 의미 요소는 센서를 통해 알 수 있는 정보이다. 중간 목표는 환경 의미 요소를 다른 요소와 함께 추론하여 결정되며 기능 관련 중간 노드와 표현 관련 중간 노드로 나누어진다. 표현 관련 중간 노드는 '춥다1', '춥다2'와 같이 나누어지기도 하는데 이는 "춥다"라는 사용자 질의가 환경 정보를 통해 두 가지 의미로 해석될 수 있음을 보여준다.

실제 서비스 추론의 예로, 서비스 로봇은 "창문을 열어라"와 같은 질의를 받으면 (창문, 열어라)과 같은 의미 요소 노드에 'true' 값을 할당한다. 이 때 조건부 확률 테이블과 추론 알고리즘을 통해 중간 목표 노드와 목표 노드의 확률 값들이 정해지고 그림 2에 나온 과정을 따라 가장 높은 확률 값을 가지는 목표 노드를 찾는다. 이 경우 "창문 열기" 노드가 선택된다.

서비스 종류를 결정한 뒤 다음 추론 과정은 선택된 서비스와 관련된 구체적인 대상을 결정한다. 종류를 추론할 때와 마찬가지로 관련된 대상에 대한 질의와 환경 정보가 의미요소로 사용된다. 여기서 "창문을 열어라"라는 질의는 앞서 보았듯이 첫 번째 과정을 통해 "창문 열기" 라는 서비스 종류를 결정할 수 있게 하지만 대상에 대한 정보는 들어 있지 않다. 이 경우 상호 주도 방식으로 대화를 하며 대상을 결정하게 된다. 서비스 로봇은 가장 높은 값을 가지는 중간 목표 노드인 "열다" 노드를 추론함으로써 미리 스크립트로 정의되어 있는 "어떤 방에 있는 창문을 열까요?" 라는 질문을 할 수 있게 된다. 만약 사용자가 정확히 방의 이름을 답변하면 서비스가 이루어지고 그렇지 않다면 상호 주도 방식으로 계속 대화를 하게 되는데 모든 중간 노드의 확률 값이 기준 값 이하면 서비스 도메인을 벗어난 것으로 간주하며 대화를 종료한다. "창문 열기" 서비스의 대상을 추론하기 위한 베이지안 네트워크의 노드들이 표 4에 정리되

표 4 "창문을 열어라" 서비스 대상을 추론하기 위한 베이지안 네트워크

범주	요소	내용
의미 요소	기능요소	방A(안방), 방B(큰아들방), 방C(작은아들방), 방D(삼촌방), 방E(화장실), 방F(거실), 방G(부엌)
	표현요소	옆, 건너, 열다
	환경요소	방A 상태(열림, 닫힘), 방B 상태(열림, 닫힘), 방C 상태(열림, 닫힘), 방D 상태(열림, 닫힘), 방E 상태(열림, 닫힘), 방F 상태(열림, 닫힘), 방G 상태(열림, 닫힘)
중간 목표	기능관련	방A 옆, 방B 옆, 방C 옆, 방D 옆, 방E 옆, 방F 옆, 방G 옆, 방A 건너편, 방B 건너편, 방C 건너편, 방D 건너편, 방E 건너편, 방F 건너편, 방G 건너편
	표현관련	열다
목표	서비스 목표	방A 창문, 방B 창문, 방C의 창문, 방D의 창문, 방E의 창문, 방F의 창문, 방G의 창문

어 있다.

이와 같은 방법으로 로봇이 서비스를 수행하기 위해서 사용된 전체적인 계층적 베이지안 네트워크는 종류를 선택하기 위한 네트워크 1개와 대상을 추론하기 위한 네트워크 10개이다. 표 5에는 서비스 로봇을 위해 사용된 베이지안 네트워크들과 의미 노드 수, 중간 노드 수, 목표 노드 수, 총 노드 수가 정리되어 있다.

4.3 지능형 서비스를 위한 환경 정보 모델링

사람들은 대화할 때, 물건의 위치나 방의 상태와 같은 대화의 상황을 알고 있어 자연스럽게 대화한다. 이러한 상황을 인식하는 많은 어플리케이션들이 개발되어 사용자의 필요에 보다 적절히 반응할 수 있도록 해준다[18]. 본 논문에서는 이러한 상황들을 환경 정보로 사용하며 서비스와 관계된 정보들을 다음처럼 3가지 변수를 가지는 간단한 'first-order predicate'로 표현한다.

Context(<ContextType>,<Subject>,<Object>)

여기서 'ContextType'은 환경 정보의 속성을 의미한다. 'Subject'는 구체적인 위치를 나타내고 'Object'는 'Subject'와 'Context Type'과 관계되어 'Context Tpye'의 상태를 나타낸다. 따라서 'Subject'와 'Object'에 올 수 있는 변수 값은 'ContextType'에 의존적이다. 예를 들어 만약 'ContextType'이 '오디오'라면 'Subject'는 오디오를 가지고 있는 방만 될 수 있고 'Object'는 '켜짐' 혹은 '꺼짐'과 같은 상태만 된다. '환경 정보 서술문'의 예는 다음과 같다.

- Context(불, 방B, 켜짐)
- Context(온도, 집 안, 높음)
- Context(오디오, 방C, 꺼짐)
- Context(소음, 집 밖, 시끄러움)

비록 이런 모델은 단순하지만 기본적인 환경 정보들을 표현해주며 목표를 정확히 추론하는데 도움을 준다. 만약 에어컨이 이미 켜져있고 집 밖의 온도가 집 안보다 더 추우며, 창문이 열려 있는 상황에서 '여기가 너무 춥다' 와 같은 질의는 사용자가 창문을 닫기 원하는 것으로 해석될 수 있다. 이런 환경 정보들은 목표 노드와 직접 연결되기도 하며 이와 상충하는 서비스는 선택되지 않도록 해준다. 각 컨텍스트 서술문의 'Context-Type'과 'Subject'는 도메인에 맞도록 되어있고 'Object'는 로봇의 센서로 관측할 수 있거나 로봇에게 환경 정보로 주어지는 값이다. 표 6은 이와 같은 방법으로 표현될 수 있는 환경 속성과 값을 보여준다.

표 6 환경의 속성 값들

범주	속성	값
가정환경	밝기	"밖이 더 밝음", "안이 더 밝음"
	온도	"밖이 더 더움", "안이 더 더움"
	소음	"밖이 더 시끄러움", "안이 더 시끄러움"
가정기기 및 소품	창문	"열림", "닫힘"
	불	"켜짐", "꺼짐"
	오디오	"켜짐", "꺼짐"
	TV	"켜짐", "꺼짐"
	에어컨	"켜짐", "꺼짐"

5. 실험 결과

5.1 실험 환경

실험은 Cyberbotics사[19]에서 개발한 Webots 4 시

표 5 서비스 로봇을 위해 사용된 계층적 베이지안 네트워크 수 및 노드 수

범주	목적	의미노드 수	중간목표 수	목표노드 수	총노드 수
서비스종류	종류추론	26	13	10	49
서비스대상	불 켜기/끄기	19	15	8	42
	창문 열기/닫기	18	14	7	39
	TV켜기/끄기	13	15	5	33
	오디오 켜기/끄기	13	15	5	33
	에어컨 켜기/끄기	11	15	3	31

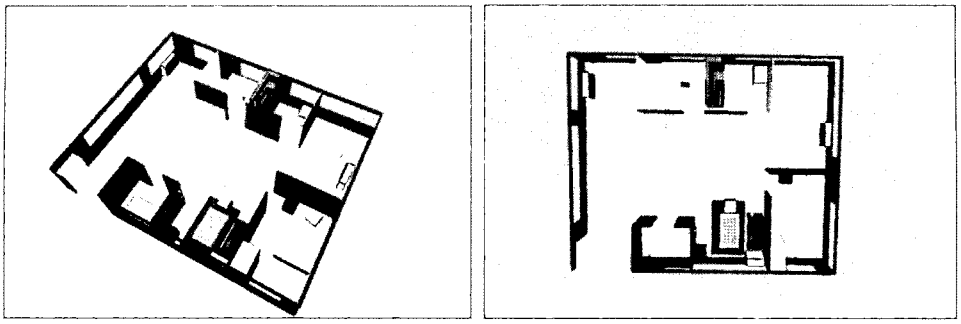
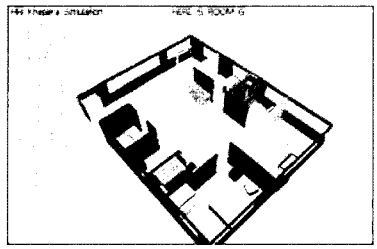


그림 5 실험 대상 환경

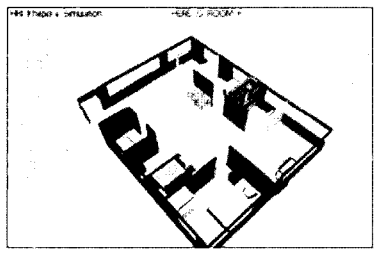
뮬레이터를 통해 그림 5와 같은 환경을 만들고 로봇을 제어하여 실험 환경을 구성하였다. Webots 4는 VRML 기반 시뮬레이터로서 3D 환경을 제공하고 로봇을 제어할 수 있는 컨트롤러를 라이브러리 형식으로 제공한다. 이를 통해 구현된 실험 환경은 4절의 표 2에 있는 것과 마찬가지로 부엌과 거실을 포함한 7개의 방과 7개의 창, 7개의 불, 3대의 TV, 3대의 오디오 및 1대의 에어컨을 포함한다. 대상 환경에서 적절한 서비스 추론을 위한 베이지안 네트워크는 전문가가 직접 설계하였다.

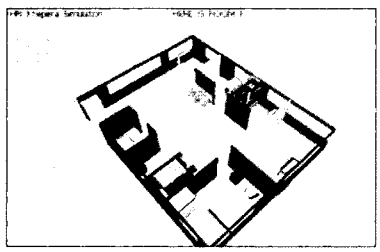
5.2 기본적인 서비스 수행 과정

대화 1은 사용자와 로봇 사이에서 이루어지는 의사소통과 관련된 간단한 예이다. 초기 로봇은 그림 6(a)와 같은 위치에서 사용자의 지시를 받고 서비스를 수행한다. 그림 6(c)는 서비스 추론 과정이다. 이 경우 사용자 질의가 명확하기 때문에 서비스 종류가 TV를 켜는 것이라는 것과 대상이 방 F에 있다는 것이 한 번에 추론

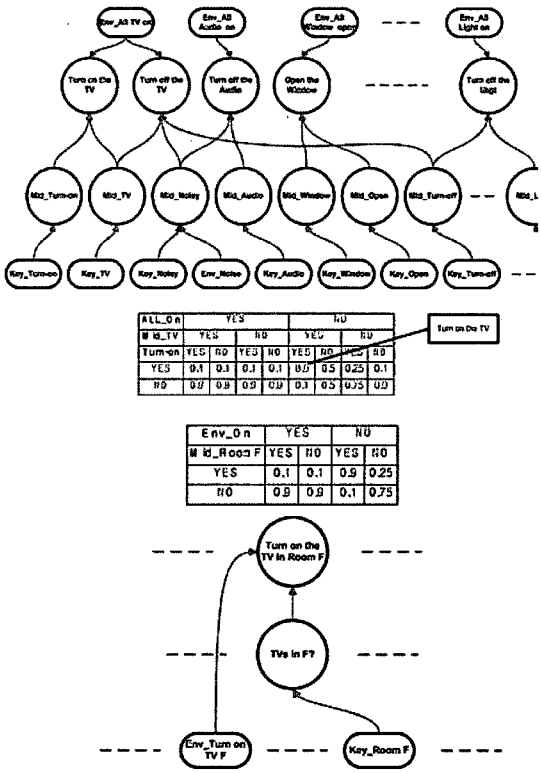


(a)





(b)



(c)

그림 6 (a) 시작 위치. (b) 방 F에 있는 TV를 켜는 서비스 과정. (c) 목표 추론 과정

사용자:	"방 F에 있는 TV를 켜줄래?"
로봇:	의미 요소(켜라, TV, 방, F)
	서비스 종류 추론(TV, 켜라)
	서비스 종류 결정(TV 켜기)
	서비스 대상 추론(방, F)
	서비스 대상 결정(방 F)
	→ 대화 내용 결정 (방 F TV 관련)
	(서비스: 방 F에 있는 TV 켜기)
	"방 F에 있는 TV를 켜겠습니다."

대화 1 명확한 목표 추론

될 수 있다.

5.3 모호한 상황 처리

대화 2와 그림 7은 모호한 상황에서 사용자의 의도를 상호 주도 방식으로 처리하는 과정을 보여준다. 사용자는 시끄러운 상황을 해결하기 위해 "시끄럽다"는 표현을 서비스 로봇에게 하였다. 이러한 경우 서비스 로봇은 사용자로부터 받은 정보만으로는 서비스 종류를 결정하지 못한다. 따라서 중간 목표 층으로부터 값을 추론하여 가장 높은 확률 값을 가진 "시끄럽다1" 노드를 통해 서비스 목표가 TV를 끄는 것인지 오디오를 끄는 것인지를 묻는다. "시끄럽다1" 노드는 "시끄럽다"라는 사용자의 질의와 바깥 소음이 더 작다는 환경 정보로부터 추론된다. 만약 바깥 소음이 더 크다는 정보가 들어온다면 "시끄럽다2" 노드가 선택되면서 창문을 닫을지, TV나 오디오를 끌지를 질문하게 된다. 이와 같은 방법은 환경 정보를 통해 사용자에게 적절한 질문을 할 수 있게 한다. 대화 2에서 서비스 로봇은 상호주도 방식으로 사용자로부터 오디오를 꺼야 한다는 정보를 얻고 두 번째 추론 과정을 통해 서비스 대상인 방C를 선택하게 된다. 이 과정에서 로봇은 오디오가 모두 꺼져있지 않다는 환경 정보를 통해 사용자에게 오디오가 존재하는 모든 방의 이름을 제공함으로써 사용자에게 대상에 대한 정보를 요청하고 응답을 통해 최종적으로 서비스를 수행한다.

대화 3은 모호한 상황에서 서비스를 추론하는 다른 예를 보여준다. 서비스 로봇은 사용자로부터 불을 꺼달라는 질의를 받고 서비스 목표를 추론한다. 이 경우 '불 끄기'가 서비스 목표로 추론되지만 서비스 대상을 결정하기 위해 로봇이 사용자에게 어느 방에 있는 불을 끌 것인지를 묻는 과정이 나와있다.

모호한 표현은 사용자의 감정 표현이나 서비스 대상이 많은 경우 발생할 수 있는데 일상적인 상황에서 사람들은 대화의 흐름이나 상황을 통해 그 의미를 구체적으로 짐작할 수 있는 경우가 많다. 하지만 이러한 모든 경우를 서비스 로봇이 이해하며 서비스를 수행하는 것

사용자:	"여기 너무 시끄럽지 않아?"
로봇:	의미 요소(시끄럽(지), 여기)
	서비스 종류 추론(시끄럽)
	→ 목표 노드 기준 값 미만
	서비스 종류 중간 노드 추론
	Context (소음, 바깥, 작다)
	→ 시끄럽다 1 중간 노드 선택
	→ 대화 내용 결정 (시끄럽다 1 노드 관련)
	"TV를 끌까요? 오디오를 끌까요?"
사용자:	"오디오 때문인 것 같은데"
로봇:	의미 요소 (오디오)
	서비스 종류 추론(오디오)
	서비스 종류 결정(오디오 끄기)
	서비스 대상 추론(NULL)
	→ 목표 노드 기준 값 미만
	서비스 대상 중간 노드 추론
	Context (오디오, 방, 모두 꺼져있지 않다)
	→ 오디오 중간 노드 선택
	→ 대화 내용 결정 (오디오 노드 관련)
	"방 A, C, F 중 어디를 말씀하시는 것입니까?"
사용자:	"방 C에 있는 것 부탁해."
로봇:	의미 요소(방 C)
	서비스 대상 추론(오디오, 방 C)
	서비스 대상 결정(방 C 오디오)
	(서비스: 방 C에 있는 오디오 끄기)
	→ 대화 내용 결정 (방 C 오디오 노드 관련)
	"방 C에 있는 오디오를 끄겠습니다."

대화 2 모호한 상황에서 서비스 추론

사용자:	불 좀 꺼줄래?
로봇:	의미 요소(불, 꺼줄(래))
	서비스 종류 추론(불, 꺼줄)
	서비스 종류 결정(불 끄기)
	서비스 대상 추론
	→ 목표 노드 기준 값 미만
	서비스 종류 중간 노드 추론
	→ 불 중간 노드 선택
	Context(불, 방 A, B, C, D, 꺼짐)
	Context(불, 방 E, F, G 켜짐)
	→ 대화 내용 결정 (불 끄기 노드 관련)
	"어떤 방에 있는 불을 끄시겠습니까?"
사용자:	"모두 부탁해."
로봇:	의미 요소(모두)
	서비스 대상 추론(모두)
	서비스 대상 결정(모두)
	(서비스: 모든 방에 있는 불 끄기)
	→ 대화 내용 결정 (모든 방 불 노드 관련)
	"켜져 있는 모든 불을 끄겠습니다."

대화 3 모호한 상황에서 서비스 대상 추론

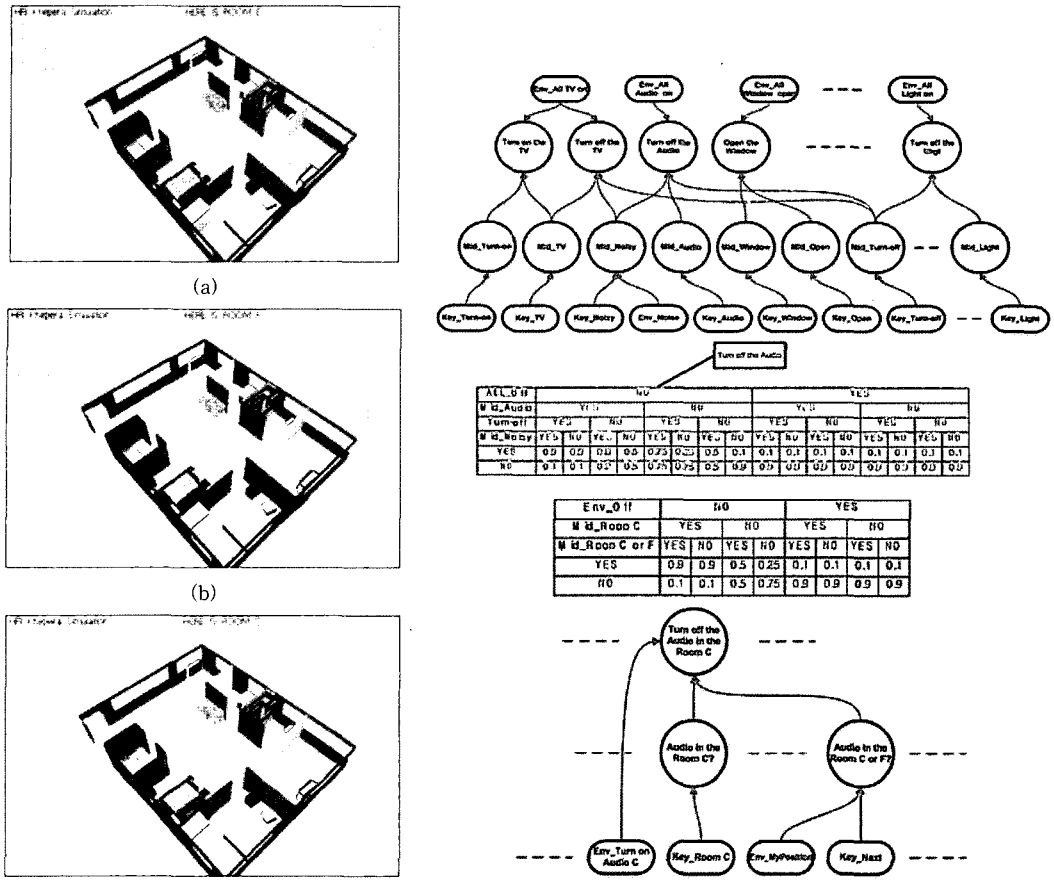


그림 7 '시끄럽다' 라는 질의 처리 및 수행과정

은 매우 어려운 일로 본 논문에서는 정해진 도메인 안에서 상호 주도적인 방법으로 사용자에게 축소된 범위의 질의를 함으로써 이러한 모호성을 줄여가는 방법을 제안하였다. 이와 같은 방법으로 처리할 수 있는 질의의 수는 각 계층적 베이지안 네트워크에 있는 중간 목표 노드 수와 같고 본 논문에서는 87가지 유형의 다른 상황을 해결할 수 있게 하였다.

6. 사용자 평가

본 논문에서는 사용자들의 시스템 사용 및 결과를 기록하고, 설문하는 방법으로 사용자 평가를 하였다. 로봇에 대한 사용자의 전문성과 문제 상황의 모호한 정도에 따라 상호 주도 방식이 어느 정도 효과를 갖는지 보기 위해 사용자 그룹을 전문가 집단과 비 전문가 집단으로 나누고 문제를 모호한 수준에 따라 세 단계로 나누었다.

6.1 사용자 평가 환경 및 방법

사용자 평가는 전문가 집단 8명, 비전문가 집단 8명으로 이루어졌고 사용자들에게는 표 7과 같이 정보가 제공되었다. 전문가 집단에 속한 사용자에게는 로봇의 사용법이 담긴 매뉴얼을 제공하고 비 전문가 집단은 로봇이 감정을 이해할 수 있으며 사람과 대화하듯이 대화해 보라는 힌트만 제공하였다.

표 7 두 집단에게 주어진 정보

	전문가	비전문가
집안 구조 및 명칭	O	O
사용자의 현재 위치	O	O
마루 가전제품 상태	O	O
로봇 사용법	O	X

각 사용자에게는 3가지 상황이 주어졌고 각 상황에 필요한 서비스를 비 상호 주도형 서비스 로봇과 상호 주도형 서비스 로봇에게 명령하도록 하였다. 사용자 평가에 사용된 문제 상황과 목표 서비스는 표 8, 9와 같다.

표 8 사용자 평가에 사용된 문제

평가자가 알 수 있는 상황 정보: *-사용자는 마루에 있는 소파에 앉아 있다* *-마루에 있는 TV는 꺼져있다* *-마루에 있는 오디오는 꺼져있다* *-마루에 있는 에어컨은 켜져있다* *-마루에 있는 창(사용자 가까이 있는)은 닫혀있다* 문제 *1. 마루에서 TV를 보고 싶습니다. 어떻게 할까요?* *2. TV를 보려고 하는데 음악 소리가 들립니다. 어떻게 할까요?* *추가로 제공된 정보:* *-사용자의 오른편에서 소리가 들린다* *3. "덥습니다. 어떻게 할까요?"* *추가로 제공된 정보:* *-집안 기온이 높다* *-바깥 온도가 집안보다 높다*

표 9 목표 서비스

번호	목표 서비스	모호성
1	마루에 있는 TV를 켜라	낮음
2	큰 아이방의 오디오를 꺼라	중간
3	삼촌방 창문을 닫아라	높음

문제들은 1번에서 3번으로 갈수록 문제 상황과 관련된 정보가 적어지면서 추측하기 어려워진다. 1번 문제는 사용자가 무엇을 목표로 하고 있고 명령을 내려야 하는지가 명확한 상황이고 2번은 정확하지 않지만 비교적 쉽게 짐작할 수 있는 상황, 3번은 많은 가능성들을 포함하고 있어 정확히 명령을 내리기가 어려운 상황이다. 제시된 세가지 문제를 통해 각 집단에 속한 사용자들이 문제의 모호함에 따라 어떻게 반응하는지를 기록하고 상호 주도형과 비 상호 주도형 로봇의 성공률과 상호작용 횟수를 기록하여 비교해 보았다.

6.2 사용자 평가 결과 및 분석

사용자 평가를 통해 얻어진 상호작용 수는 그림 8과 같다. 상호작용 수는 방법의 효율적인 측면을 보여주는 수치로 전문가 집단과 비전문가 집단 모두 상호주도형 로봇을 사용한 경우가 평균적으로는 25% 정도의 성능 향상을 보여주었다. 특히 전문가 집단의 경우 3번 문제에서는 71.4% 정도로 큰 성능 향상을 보여주었다.

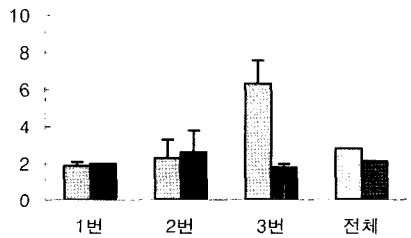
전문가 집단의 1번 문제와 2번 문제에서는 비 상호주도형이 조금 더 높은 경향을 보이는데, 이는 사용법을 알고 있는 전문가들이 상호주도형 로봇과 의사 소통 시에는 한 번에 명령을 내리기 보다 대화 하는 것을 선호하는 성향이 있기 때문인 것으로 분석된다. 비전문가 집단의 경우는 2번을 제외한 문제에서 모두 상호주도형 로봇이 더 좋은 성능을 보였다. 2번 문제의 경우는 t-test 결과 실제 두 집단의 차이가 있다고 받아들이기는 어려운 것으로 나타났는데 이는 2번 문제를 직관적으로 풀 수 있어 상호 주도와 비 상호주도에 상관없이 결과가 나온 것으로 보인다. 표 10에는 각 실험에 대한 유의도를 평가하기 위해 paired t-test를 통해 얻어진 p-값이 나와 있다.

표 10 비 상호주도와 상호주도에 대한 p-값

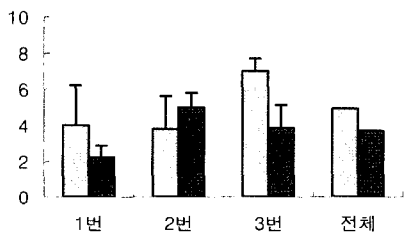
	전문가	비전문가
1번 문제	0.00048	0.047
2번 문제	0.375	0.47
3번 문제	0.0015	0.0063

테스트 결과 1번 문제와 3번 문제는 모두 유의수준 0.05에서 상호주도형과 비상호주도형일 때 차이가 있는 것으로 나타났지만 2번 문제의 결과는 그렇지 않은 것으로 나타났다.

그림 9에는 각 시스템에서의 성공률이 있다. 사용자들이 목적을 달성한 비율이다. 본 논문에서는 상호작용수가 10회 이상이거나 사용자가 문제 풀이를 그만 두겠다는 의사를 표시하면 실패로 간주하였다. 전문가 집단의 경우 방식에 상관없이 모두 문제를 풀었다. 반면 비전문가 집단에서는 비상호주도형 로봇의 경우 전체 평균 성공률이 50% 정도로 나타났고 3번 문제에서는 20%를 조금 넘는 성공률을 보였다. 이에 비해 상호주도형 로봇의 경우는 전체적으로 90%에 가까운 성공률을 보였고

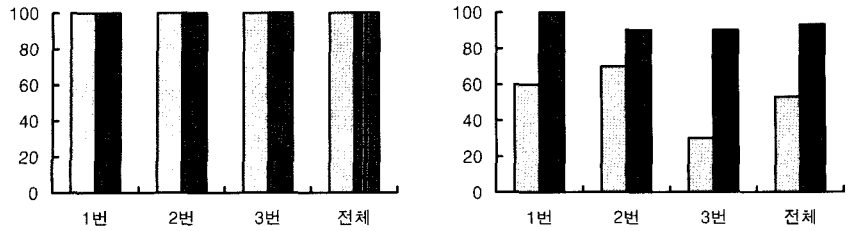


(a) 전문가 집단의 상호작용 횟수



(b) 비전문가 집단의 상호작용 횟수

그림 8 상호작용 수 ⬚ 비상호주도 ■ 상호주도



(a) 전문가 집단의 성공률 (b) 비전문가 집단의 성공률

그림 9 성공률 □ 비상호주도 ■ 상호주도

3번 문제에서도 90%정도의 성공률을 보여주었다. 이런 결과는 상호 주도형 로봇이 시스템을 잘 모르는 사용자들에게도 효과적임을 보여준다.

그림 10에서는 사용자 로그 분석을 통해 사용자가 사용한 로그가 로봇이 처리할 수 있는 도메인에 들어가 있는지 아닌지를 정리한 자료가 나와있다. 비 상호주도형 로봇의 경우 전체적으로 도메인 밖의 내용(Out Of Domain)으로 로봇과 대화를 한 비율이 46.3%인데 비해 상호주도형 로봇의 경우는 23%로 많이 줄어든 것을 알 수가 있다. 이는 로봇의 적절한 반응이 도메인에 속한 명령을 내릴 수 있도록 안내하는데도 효과적임을 보여준다. 이외에 로봇의 사용법을 모르는 사용자들은 모호한 상황에서 감정적인 표현이나 '모두'와 같은 표현을 사용하는 것으로 나타났다. 비전문가에서는 62%정도 되는 사용자가 2번 문제나 3번 문제에서 감정적인 표현을 사용했고 37.5%의 사용자가 모두라는 표현을 사용했다. 감정적인 표현이 비교적 높은 이유는 로봇이 감정적인 표현을 이해할 수 있다는 단서를 미리 제공했기 때문이라고 생각되지만 설문 조사 결과 사용자들은 대부분 로봇이 감정적인 표현을 이해할 수 있는 것이 더 좋다라고 대답했다.

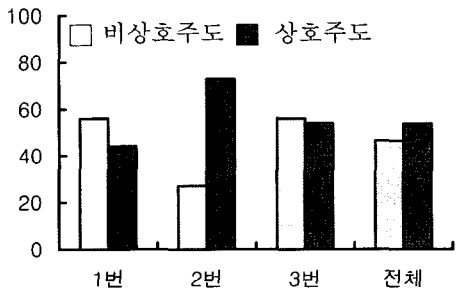


그림 10 사용자 로그 분석 결과

7. 결 론

본 논문에서는 보다 지능적인 서비스 로봇의 동작을 위한 대화형 사용자 인터페이스를 제안하였다. 사용자의 의도를 이해하고 적절한 반응을 보이는 것이 향상된 서비스 수행을 위해 필요하지만 모호한 표현이나 상황으로 인해 사용자의 의도를 정확히 알기가 어렵거나 사용자 역시 어떻게 상황에 대처해야 할지 모르는 경우가 발생하게 된다. 이 문제를 위해 계층적 베이지안 네트워크를 사용한 상호 주도 방식의 HRI 기법을 제안하였고 또한 상호 주도 방식으로 처리되는 서비스 로봇을 만들어 시뮬레이션과 사용자 평가를 통해 유용함을 확인해 보았다.

본 연구에서는 상호 주도 방식을 통해 로봇의 대화 이해 수준을 향상시키는 방법을 제안하였지만 서비스 로봇이 동작하는 실제 환경은 변화가 많고 불확실한 경우가 많아서 실제 환경을 모델링한 정도에 따라 보다 향상된 결과를 얻을 수 있다. 향후에는 좀 더 다양한 센서와 모터를 사용해서 실제 환경을 모델링하고 이를 바탕으로 서비스 로봇이 동작하도록 연구하고자 한다.

참 고 문 헌

[1] C. Breazeal, "Social interactions in HRI: The robot view," *IEEE Trans. Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 34, no. 2, pp. 181-186, 2004.

[2] A. Agah, "Human interactions with intelligent systems: Research taxonomy," *Computers and Electrical Engineering*, vol. 27, no. 1, pp. 71-107, 2001.

[3] S. Lauria, *et al.*, "Personal robots using natural language instruction," *IEEE Intelligent Systems*, vol. 16, no. 3, pp. 38-45, 2001.

[4] T. Fong, et al., "Robot, asker of questions," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 235-243, 2003.

[5] V. Zue and J. Glass, "Conversational interfaces: Advances and challenges," *Proc. of the IEEE*, vol. 88, no. 8, pp. 1166-1180, 2000.

[6] H. Meng, *et al.*, "The use of belief networks for mixed-initiative dialog modeling," *IEEE Trans. Speech and Audio 'Processing*, vol. 11, no. 6, pp. 757-773, 2003.

[7] J. Allen, "Mixed initiative interaction," *IEEE Intelligent Systems*, vol. 15, no. 4, pp. 14-23,

1999.
[8] E. Horvitz, "Uncertainty, action, and interaction: In pursuit of mixed-initiative computing," *IEEE Intelligent Systems*, vol. 14, no. 5, pp. 17-20, 1999.
[9] H. Huttenrauch, *et al.*, "Involving users in the design of a mobile office robot," *IEEE Trans. Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 34, no. 2, pp. 113-124, 2004.
[10] F. Mizoghchi, *et al.*, "Smart office robot collaboration based on multi-agent programming," *Artificial Intelligence*, vol. 114, no. 1-2, pp. 57-94, 1999.
[11] M Skubic, *et al.*, "Spatial language for human-robot dialogs," *IEEE Trans. Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 34, no. 2, pp. 154-167, 2004.
[12] E. Horvitz and T. Paek, "A computational architecture for conversation," *Proc. of the 7th Int. Conf. on User Modeling*, pp. 201-210, 1999.
[13] G. Ferguson, et al., "TRAINS-95: Towards a mixed-initiative planning assistant," *Proc. of the 3rd Int. Conf. on Artificial Intelligence Planning Systems*, pp. 70-77, 1996.
[14] S. Perugini, et al., "Personalizing Web Sites with Mixed-Initiative Interaction," *IT Professional*, vol. 5, no. 2, pp. 9-15, 2003.
[15] A. Onisko, *et al.*, "Comparison of rule-based and Bayesian network approaches in medical diagnostic Systems," *Proc. of the 8th Int. Conf. on Artificial Intelligence in Medicine*, pp. 283-292, 2001.
[16] W. W. Gibbs, "Considerate Computing," *Scientific American*, January, 2005.
[17] J.-H. Hong and S.-B. Cho, "A two-stage bayesian network for effective development of conversational agent," *Lecture Notes in Computer Science 2690*, Springer, Berlin, pp. 1-9, 2003.
[18] A. Ranganathan, *et al.*, "ConChat: A context- aware chat program," *IEEE Pervasive Computing*, vol. 1, no. 3, pp. 51-57, 2002.
[19] http://www.cyberbotics.com - Webots software

송 윤 석
2004년 8월 연세대학교 컴퓨터과학과(학사). 2004년 9월~현재 연세대학교 컴퓨터과학과 석사과정. 관심분야는 지능형 로봇, 지능형 에이전트, 인공 생명

홍 진 혁
2002년 연세대학교 기계전자공학부 정보산업전공 졸업. 2002년~2004년 연세대학교 컴퓨터과학과 석사. 2004년~현재 연세대학교 컴퓨터과학과 박사과정. 관심분야는 지능형 에이전트, 인공생명, 패턴인식, 진화게임

조 성 배
1988년 연세대학교 전산과학과(학사) 1990년 한국과학기술원 전산학과(석사) 1993년 한국과학기술원 전산학과(박사) 1993년~1995년 일본 ATR 인간정보통신연구소 객원 연구원. 1998년 호주 Univ. of New South Wales 초청연구원. 1995년~현재 연세대학교 컴퓨터과학과 정교수. 관심분야는 신경망, 패턴인식, 지능정보처리