

유비쿼터스 가정환경을 위한 계층적 베이지안 네트워크 기반 상호주도형 대화 에이전트

A mixed-initiative conversational agent using hierarchical
Bayesian networks for ubiquitous home environments

송인지, 홍진혁, 조성배
연세대학교 컴퓨터과학과

In-Jee Song, Jin-Hyuk Hong, Sung-Bae Cho
Department of Computer Science
Yonsei University

E-mail : {schunya,hjinh}@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

요 약

유비쿼터스 환경에서 다양한 서비스를 사용자에게 제공하기 위해 지능형 에이전트는 먼저 사용자의 의도를 정확히 파악해야 한다. 명령어 기반의 기존 사용자 인터페이스와는 달리, 대화는 인간과 에이전트 사이의 유연하고 풍부한 의사소통에 유용하지만, 사용자의 배경지식이나 대화의 문맥에 따라 그 표현이 매우 다양하기 때문에 본 논문에서는 '상호주도형' 의사소통을 위한 계층적 베이지안 네트워크를 이용하여 사용자와 에이전트 사이에 발생하는 대화의 모호성을 해결한다. 서비스 추론 시 정보가 부족할 경우에는 계층적 베이지안 네트워크를 이용하여 추가로 필요한 정보를 분석하고 사용자로부터 수집한다. 제안하는 방법을 유비쿼터스 가정환경에 적용하고 시뮬레이션 환경을 구축하여 그 유용성을 확인하였다.

키워드: 인공집사, 계층적 베이지안 네트워크, 상호주도 방식, 유비쿼터스 가정환경

1. 서론

최근 가정에 인터넷이 보급되고, 개인용 피시, 게임기, 텔레비전 등이 네트워크에 연결되어 이를 통합하는 새로운 서비스가 시도되고 있다. 가까운 미래에는 이러한 추세가 더욱 진행되어 유비쿼터스 컴퓨팅 환경이 가정에 구축되고, 제공 가능한 서비스 양은 더욱 증가할 것이다. 이러한 유비쿼터스 홈 환경에서 서비스를 직접 제어하는 기존의 사용자 인터페이스는 다양한 서비스를 효과적으로 제공하기 어려운 한계점을 갖는다. 이런 문제는 많은 서비스 중에서 사용자의 의도와 환경에 맞는 서비스를 적절히 선택하고, 수행하는 인터페이스 에이전트를 통해 해결할 수 있다. 에이전트가 사용자의 관심과 취향을 잘 파악한다면 더욱 효율적으로 서비스를 제공할 수 있다[1].

사용자의 관심과 취향을 파악하고, 사용자와

정보를 교환하는 데 대화가 효과적이다. 사용자가 원하는 정보를 키워드나 메뉴 등이 아닌 자연어 문장으로 입력하여 보다 친숙한 인터페이스를 제공한다[2]. 하지만 실제 대화에는 화자 사이에 암묵적으로 미리 정의되고 유지되는 문맥이 존재하여, 대화 내용 중 몇 단어가 생략되어도 의사소통이 가능하다. 인터페이스 에이전트가 사용자와 자연스럽게 정보를 주고받기 위해서는 문맥상의 모호함과 불확실성을 다루어야 하며 이를 위해 상호주도적인 방법이 연구되고 있다. 상호주도적인 방법은 사람과 시스템이 지속적인 의사소통을 통해 점진적으로 문제를 해결해 가며 에이전트가 모든 상황을 추론하기 보다는 사용자와 상호작용을 통해 불확실성을 해결한다[3].

본 논문에서는 계층적 베이지안 네트워크를 이용한 상호주도형 대화를 통해 복잡한 가정환경을 관리 및 제어하고, 그 결과를 3D 가상 환경

및 아바타를 통해 시뮬레이션 하는 인공지능 시스템을 구현하여 그 유용성을 확인한다.

2. 관련 연구

2.1 가정환경 에이전트

P. Debaty는 Cooltown 프로젝트의 일환으로 가정환경 내의 사물, 사람, 장소 등을 위한 웹 페이지를 각각 만들고, 그 목록을 관리하는 WPM(Web presence manager) 소프트웨어 아키텍처의 구현을 통해 각 개체가 웹으로 관리 가능하도록 하였고[4], 사용자가 PDA를 들고 다니면 개인 장치에서 접근 가능한 주변의 미디어 목록을 웹 인터페이스로 표시하였다. B. Carolis는 Multi-Agent System 기반의 C@sa를 지능형 가정을 모델링하고 시뮬레이션 하는 용도로 개발하였다[5]. C@sa는 물리적 위치 기준이 아닌 서비스 종류 별로 서비스 에이전트를 묶어 Influence Spheres로 정의하고, 각 관리 에이전트가 하나의 Influence Sphere를 관리하는 방법을 취하였다.

2.2 상호주도형 의사소통

대화형 에이전트는 사용자와 직접 대화로 가정환경의 다양한 서비스에서 발생하는 복잡함과 불편함을 관리하고 추상화하기 위해 사용된다. 상호주도 방식의 대화는 인간과 에이전트 사이에 주도권의 주고받음(Turn taking)을 자동적으로 처리하여 복잡한 상황을 능동적으로 해결한다. 이는 사용자와 에이전트가 서로 문제에 대한 관점을 공유하여, 문제를 가장 적합한 방식으로 풀도록 돕는다. 사용자와 에이전트의 협력을 위해서는 에이전트가 사용자에게 관한 정보를 수집하고, 이를 바탕으로 사용자의 의도나 동기 등을 추론하여 결과적으로 적합한 서비스를 추천해야 한다.

상호주도 방식을 이용한 연구로 Microsoft에서 개발한 Lookout이 있다[6]. 이 에이전트는 사용자가 이메일 메시지 확인할 때와 달력을 통해 스케줄을 확인할 때를 관찰한다. 이메일 메시지의 내용을 통해 사용자가 달력을 열고 확인을 할 날, 주, 달의 스케줄을 추론하고 정확도가 높다고 판단될 경우 사용자에게 해당 스케줄을 볼 것인지 물어본다. 추론이 정확하지 않은 경우에는 사용자에게 불분명한 부분을 묻는 과정을 통해 추론을 정교하게 한다. 이 과정에서 추론은 메시지의 제목과 내용을 패턴 매칭을 통해 미리 정의된 패턴들을 검색하여 이루어진다.

대화의 문맥을 모델링하고 모호한 자연어 처리를 위해 계층적 베이지안 네트워크가 제안되었다

[7]. 상호주도형 의사소통을 위해 사용된 계층적 베이지안 네트워크는 의미 요소 노드, 목표 노드, 중간 목표 노드로 구성된다. 목표는 의미 요소로부터 직접 추론되지 않고 상호주도 방식을 위해 중간 목표 층을 거쳐 추론된다. 만약 서비스 목표가 결정되기에 충분한 단서를 얻지 못했다면 그림 1과 같이 중간 목표들 중에서 가장 큰 값을 가지는 노드를 찾고 이를 통해 목표가 될 수 있는 노드의 범위를 사용자에게 제공하여 상호주도 방식으로 모호성을 해결한다.

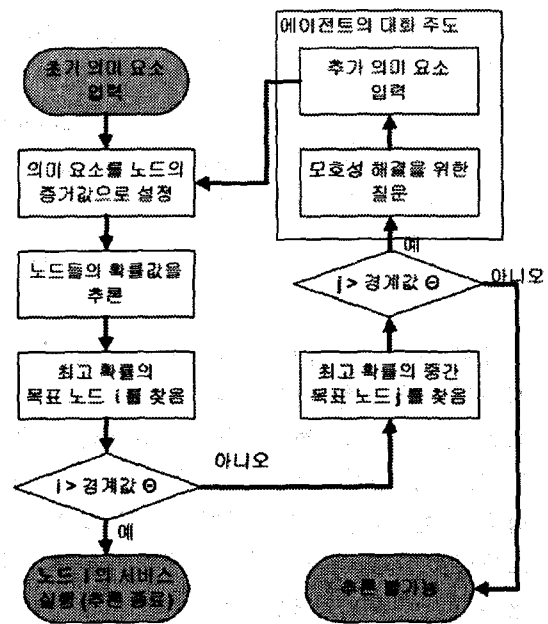


그림 1. 상호주도 대화 처리 과정

3. 상호주도형 대화 기반의 인공지능

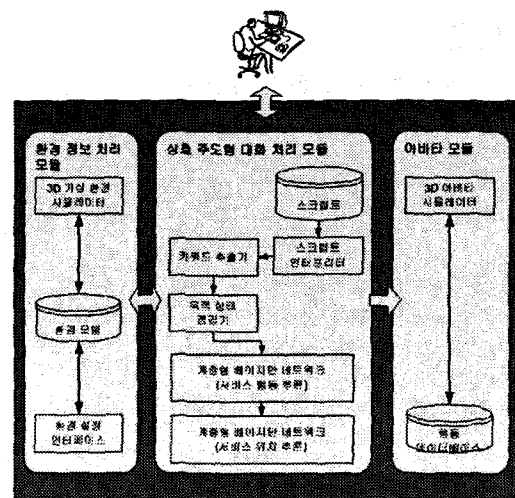


그림 2. 제안하는 인공지능 구조

인공지능은 가정환경의 복잡하고 다양한 서비

스를 사용자와 에이전트의 상호 작용을 통하여 관리하는 에이전트이다. 인공지능사는 그림 2와 같이 환경 정보를 다루는 부분, 상호주도형 대화를 처리하는 부분, 그리고 아바타의 행동을 보여주는 세 부분으로 구성된다. 환경 정보와 사용자의 질의가 상호주도형 대화 처리에 사용되고, 그 결과는 환경 정보와 아바타의 변화로 표현된다. 인공지능사에서 처리 가능한 서비스의 목록은 가정환경의 다양한 서비스를 시뮬레이션하기 위해 표 1과 같이 정의되었다.

표 1. 서비스 모델

서비스 종류	서비스 대상
불 켜기	1번 방, 2번 방, 3번 방, 4번 방, 5번 방, 6번 방, 7번 방
불 끄기	1번 방, 2번 방, 3번 방, 4번 방, 5번 방, 6번 방, 7번 방
창문 열기	1번 방, 2번 방, 3번 방, 4번 방, 5번 방, 6번 방, 7번 방
창문 닫기	1번 방, 2번 방, 3번 방, 4번 방, 5번 방, 6번 방, 7번 방
오디오 켜기	1번 방, 3번 방, 6번 방
오디오 끄기	1번 방, 3번 방, 6번 방
TV 켜기	4번 방, 6번 방, 7번 방
TV 끄기	4번 방, 6번 방, 7번 방
에어컨 켜기	7번 방
에어컨 끄기	7번 방

3.1 환경 정보처리 모듈

환경 정보처리 모듈은 사용자와 대화처리를 위해 환경정보를 확인하고 설정한다. 환경모델은 집안의 온도, 소음정도, 조도와 사용자의 현재 위치와 창문이나 조명, 가전제품의 상태(열림/닫힘, 켜짐/꺼짐)등의 환경 정보를 정의한다. 환경 정보는 환경조작 인터페이스, 환경 시뮬레이터, 대화처리 시스템을 통하여 조작이 가능하다.

3.2 상호주도적 대화처리 모듈

상호주도적 대화처리는 우선 대화를 미리 작성된 스크립트 구조를 이용하여, 분석해서 명사, 동사의 어근을 핵심 단어로 추출하고, 사전에 정의된 규칙을 기반으로 목적 상태를 검증한다. 목적 상태 검증을 통과한 핵심 단어들은 계층적 베이지안 네트워크 기반 상호주도형 대화를 통해 실행할 서비스를 선택하는데 사용된다. 베이지안 네트워크를 이용한 추론은 두 단계로 이루어진다. 첫 번째 베이지안 네트워크는 핵심 단어와 환경 정보를 이용하여 서비스 종류를 결정하고, 두 번째 베이지안 네트워크는 서비스가 행해질

위치를 결정한다. 그림 3은 서비스 종류 추론을 위한 계층적 베이지안 네트워크의 한 부분이다. key와 env로 시작되는 노드들은 핵심 단어와 환경 정보로 증거 값이 설정되는 의미요소 노드이고, mid로 시작되는 노드들은 모호한 의미 처리를 위한 중간 목표 노드이다. 나머지 노드들은 서비스 종류의 선택을 위한 목표 노드이다. 의미요소 노드 중 하단에 위치한 key_audio, key_tv, key_noise, env_noise_dis등은 중간목표 노드를 목표노드 이전에 삽입하여 모호함을 제거하였고, 상단에 위치한 env_tv_all_off, env_tv_all_on과 같이 명확한 의미처리가 가능한 부분은 직접 목표노드와 연결하였다. 이 과정에서 정보가 부족하여 추론이 정확하게 이루어지지 않으면, 그림 1에서와 같은 과정을 거쳐 추가적인 질문을 던지고 필요한 정보를 얻는다.

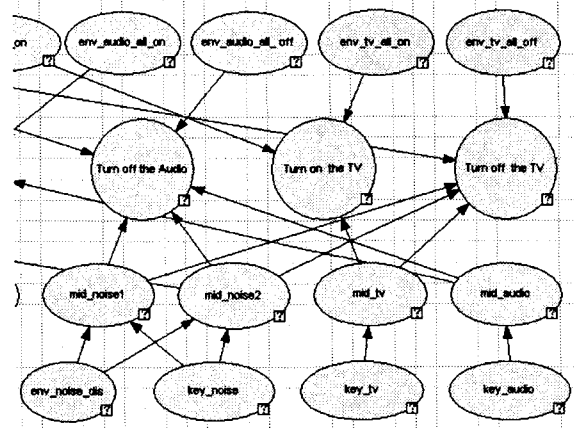


그림 3. 서비스 종류 추론을 위한 계층적 베이지안 네트워크의 부분

3.3 아바타 모듈

인공지능사에서 아바타는 사용자와의 친밀감을 향상시키고, 의사소통의 효과를 강화하기 위해 사용된다. 오류 메시지나 서비스 실행 등의 순간에 이미 정의된 행동 중 한 가지를 행하여, 사용자에게 보다 정확한 표현이 가능하다.

4. 실험 결과

4.1 실험 환경

Microsoft Foundation Class과 QAvatar, EonX 컨트롤을 이용한 윈도우 프로그램으로 그림 4와 같은 인공지능 실험 환경을 구현하였다. 환경의 제어는 좌우에 위치한 라디오 버튼과 중앙의 대화 입력창을 통하여 실행하였다. 3D 환경의 시뮬레이션은 Eon Studio를 이용하여 상호작용이 가능도록 구축되었다. 이를 통해 3D 가상 환경 상

에서 환경 정보를 쉽게 확인할 수 있을 뿐 아니라, 조명이나 가전제품의 직접적인 클릭을 통해 환경 정보의 조작이 가능하다.

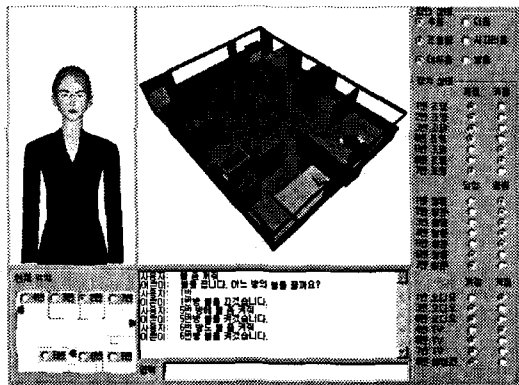


그림 4. 인공집사 실행 화면

4.2 서비스 질의에 대한 처리

사용자:	1번방에 불 좀 켜줘~
인공집사:	1번방에는 이미 불이 켜져 있습니다. 핵심단어(1번방, 불, 켜) 환경모델(1번방, 불, 켜짐) (오류 메시지 발생)

대화 1. 목적 상태에 이미 도달한 대화

대화 1은 질의의 목적 상태가 이미 완료되어 있는 경우의 대화이다. 이미 1번방의 불이 켜져 있기 때문에 사용자의 질의는 실행되지 못한다. 이 경우 인공집사는 질의가 실행될 수 없다는 오류 메시지를 출력한다.

사용자:	창문 좀 열어 줄래?
인공집사:	창을 엽니다. 어느 방의 창문을 열까요? 핵심단어(창, 열어) (서비스 추론 실패로 재질문)
사용자:	4번방 창문
인공집사:	4번방 창문을 열겠습니다. 핵심단어(4번방, 창, 열어) 환경모델(4번방, 창, 닫힘) (서비스 실행: 4번방 창문 열기)

대화 2. 서비스 위치가 모호한 대화

대화 2는 모호한 질의를 상호주도형 대화를 통해 해결하는 경우를 보여준다. 처음 사용자의 질의만으로는 서비스를 실행하기에 부족하므로, 인공집사는 질문을 통해 필요한 정보를 수집한다. 필요한 정보가 수집된 후에 다시 서비스 추론을 통해 필요한 서비스를 실행한다.

5. 결론 및 향후과제

본 논문에서는 계층적 베이지안 네트워크를 이용한 상호주도형 대화를 통해 사용자를 대신하여 서비스의 복잡함과 다양함을 다루는 인공 집사를 제안하였다. 사용자의 질의가 애매하여, 의도를 파악하기 힘들어도 질문이나 오류 메시지를 통해 대화 속의 불확실성을 제거하고 정확한 서비스의 선택을 유도하였다. 제시된 방법으로는 서비스 종류를 추론한 후에 서비스 위치를 추론하는 순으로 추론하지 않거나, 추론 시 두 개 이상의 명령이 동시에 들어오는 경우 처리가 불가능하다. 이를 해결하기 위해서는 추론 단계를 확장하기 위한 오토마타의 설계나, 계층적 베이지안 네트워크의 구조 변경을 통한 개선이 필요하다.

6. 참고문헌

- [1] P. Maes, "Agents that reduce work and information overload," *Communications of ACM*, vol. 37, no. 7, pp. 31, 1994.
- [2] M. Budzikowska et al., "A conversational interface for online shopping," *Human Language Technology Conf.*, 2001.
- [3] E. Horvitz, "Uncertainty, action, and interaction: In pursuit of mixed-initiative computing," *IEEE Intelligent Systems*, vol. 14, no. 5, pp. 17-20, 1999.
- [4] P. Debaty et al., "Integrating the physical world with the web to enhance context-enhanced services," *HP Technical Report*, 2003.
- [5] B. D. Carolis et al., "Agent-based home simulation and control," *ISMIS 2005, Lecutre Note in Artificial Intelligence*, vol. 3488, pp. 404-412, 2005.
- [6] E. Horvitz, "Principles of mixed-initiative user interfaces," *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems*, ACM press, New York, pp. 159-166, 1999.
- [7] 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.