

대화 시스템을 위한 계획 인식과 담화 스택을 이용한 효과적인 응답 생성*

강 상 우

고 영 중[†]

서 정 연

서강대학교 컴퓨터학과 동아대학교 컴퓨터공학과 서강대학교 컴퓨터학과

대화 시스템에 관한 기존 연구는 실용시스템을 위한 연구와 대화 현상 이해에 중심을 둔 연구로 나눌 수 있다. 전자는 실제 환경에서 강인한 성능을 유지하는 것에 초점을 두고 있다. 하지만 매우 제한적인 상황에서만 성능을 보장할 수 있으며 다양한 형태의 대화를 적용하기 어렵다. 후자의 연구에서는 대표적으로 계획기반 모델(plan-based model)이 제안되었는데, 이 모델은 복잡한 대화 구조를 모델링(modeling)할 수 있으며, 다양한 현상에서의 사용자 목적 추론이 가능하다. 하지만 계획기반 모델은 초기 설계가 어렵고 실용적인 대화 시스템 구축에 있어서 시스템 응답을 생성하기 위한 상호작용 모델로의 확장이 매우 어렵기 때문에 실용시스템으로의 적용이 쉽지 않다. 본 연구는 이러한 계획기반 모델의 단점을 보완하고 실용적인 대화 시스템을 구축하기 위하여, 시스템 응답을 위한 확인 대화 전략과 담화 스택(discourse stack)을 계획기반 대화 모델에 적용함으로써 효율적인 응답을 생성할 수 있는 새로운 기법을 제안 한다.

주제어 : 대화 시스템, 응답 생성, 대화 관리자, 계획기반 모델

* 이 연구(논문)는 산업자원부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능로봇 기술개발사업)의 일환으로 수행되었습니다.

† 교신저자: 고영중, 동아대학교 컴퓨터공학과, 연구세부분야: 자연어처리, HCI

E-mail: yjko@dau.ac.kr

서 론

인간은 언어를 이용하여 상호작용을 하는 과정에서 효율적인 대화(dialogue)의 방법들을 습득하였다. 따라서 인간에게는 대화를 통하여 자신의 의사를 전달하고 원하는 목적을 이루는 것이 가장 자연스러운 의사소통 방법이라 할 수 있다. 따라서 대화를 인간과 기계의 의사소통 수단으로 사용하기 위한 연구가 진행되고 있으며 최근 실제 환경에서 사용 가능한 대화 시스템(dialogue system)에 대한 연구가 활발히 이루어지고 있다. 인간과 기계가 상호작용의 도구로서 자연어(natural language)를 사용하여 특정 작업을 성취하도록 하는 프로그램을 대화 시스템이라 하며 대화 시스템이 다른 인터페이스보다 우수한 이유는 자연어가 가지고 있는 유연성(flexibility), 명료성(succinctness), 표현력(expressiveness) 등이 우수하기 때문이다.

대화 시스템은 대화의 목적에 따라 작업 수행과 정보 제공의 두 가지 역할을 수행한다. 사용자가 대화 시스템을 사용하여 특정 작업을 수행하기 위한 대화를 진행하고 있다면 그러한 대화를 작업지향 대화(task-oriented dialogue)라 하고, 정보를 얻기 위한 대화를 진행하고 있다면 그러한 대화를 정보 획득 대화(information seeking dialogue)라 한다. 인간의 대화에는 작업지향 대화와 정보 획득 대화가 혼재되어 있으며 구조적으로 매우 복잡한 형태를 갖는다. 대화 시스템을 실제 환경에서 사용하기 위해서는 이런 복잡한 형태의 대화를 처리 할 수 있어야 한다. 또한 실제 환경에서의 대화 시스템은 다양한 영역으로 적용이 가능하고 지능적인 작업 수행 능력을 갖추어야 하며, 오류를 최소화 할 수 있는 강인한 구조가 필요하다.

본 연구에서는 위와 같은 조건을 만족시키기 위해 담화 스택과 계획기반 모델을 이용한 대화 시스템을 제안한다. 계획기반 모델은 복잡한 담화 현상을 처리 할 수 있으며, 계층 구조를 갖는 영역들에 대해서도 일관성 있는 모델링이 가능하다. 하지만 계획기반 모델은 대화 현상 자체를 이해하는데 중점을 두었기 때문에 인간-기계 상호작용 모델에 바로 적용하기에는 복잡성과 확장성 등에 아직 많은 문제점을 갖고 있다. 본 연구에서 제안하는 대화 시스템에서는 계획기반 모델의 성능을 유지하면서 인간-기계 상호작용 모델로서 시스템의 목적을 수행하기 위하여 응답확인 대화 전략과 담화 스택을 계획기반 모델에 적용하여 강인한 구조의 응답모델 생성하는 기법을 제안한다.

관련 연구

대화 시스템에 관한 기존 연구는 크게 두 가지 분야로 나눌 수 있다. 그 중 첫 번째는 실제 환경에 적용할 수 있도록 강인한 성능을 보장할 수 있는 시스템을 위한 연구이다. 이 연구의 범위는 비교적 간단하고 제한적인 영역에 한정되며 대화 형태의 변이가 없거나 예측할 수 있는 형태의 대화에 적용된다. 대표적으로 스크립트기반 모델(script-based model)에 관한 연구가 있다. 다른 연구 분야로는 인간의 복잡하고 다양한 대화 현상을 설명하기 위한 대화 모델에 초점을 집중한 분야이다. 계획기반 모델은 위 분야에 대표적인 연구이며 현재까지 지속적인 연구가 진행되고 있다.

최근에는 확률기반 모델을 도입하여 각 모듈간의 에러 전파와 유지 관리 측면에서의 문제점들을 해결하려는 방법들도 시도되고 있다.

스크립트기반 대화 모델

스크립트기반 모델은 특정 목적을 수행하기 위한 작업 절차를 순차적으로 처리할 수 있도록 DTN(dialogue transition network)모델을 사용하고 일반적으로 양식기반 대화 모델(form-based dialogue model)에 사용된다[1]. DTN모델은 대화의 상태와 전이 과정을 네트워크(network)로 구현하여 예측할 수 있는 사용자 발화의 담화 구조를 표현한다. 사용자는 미리 정의된 상태(state)만을 유지할 수 있으며 다음 상태로의 이동은 네트워크에 명시된 전이를 통하여 이루어진다. 이 모델은 비교적 설계가 쉽고 강건하게 동작한다. 따라서 이 모델은 대화가 간결하고, 항상 일정한 형태를 유지하며 모호성을 배제할 수 있는 강건함이 요구되는 영역에 적합하다. 하지만 순차적 구조를 가지기 때문에 예외사항에 대처할 수 있는 유연성이 떨어지며 영역 이식성(portability) 또한 매우 낮다.

계획기반 모델

계획기반 모델은 관측된 단위 행위(action)로부터 사용자가 원하는 목적을 추론

(plan recognition)하는 것이다. 사용자 믿음(belief)의 상태에 따라 다양한 추론이 가능하고 레시피¹⁾(recipe)의 활용으로 복잡한 담화 구조를 설명할 수 있다.

Kautz와 Allen[2]은 계획 이론(plan theory)을 처음으로 대화에 적용하여 계획 인식 모델의 기반을 적립하였다. 그들은 계획 인식에 필요한 대화 모델을 담화 수준(discourse level)과 영역 수준(domain level)으로 구분하였다. 그들은 각각의 관측된 행위로부터 추론 규칙에 의하여 담화 수준 안에서 사용자의 심층 화행(deep speech act)을 파악해 내고 사용자의 현재 상태를 반영하여 사용자의 목적 영역을 추론한다. 이 과정은 계획 트리(plan-tree)로 표현되며 하나의 사용자 발화의 상황에 따라 다양한 계획 트리를 구성할 수 있다. 따라서 계획 인식이란 그 행위를 가장 잘 설명하는 계획들의 집합을 최소화하는 작업으로 설명하고 있다.

Lambert[4][5]는 보다 복잡한 담화 구조를 처리할 수 있도록 담화 수준과 영역 수준 사이에 문제-해결 수준(problem-solving level)을 추가하였다. 문제-해결 수준을 추가함으로써 Lambert는 한 개 이상의 사용자 발화가 서로 연관성을 가질 수 있으며 공통된 목적에 부합될 수 있는 모델을 제시하였다. 또한 담화 지식에 관련된 레시피만을 추가함으로써 매우 복잡한 담화 구조를 갖는 대화를 하나의 모델로 구성할 수 있으며 자연스럽게 영역 수준을 연결할 수 있게 한다.

Kautz와 Allen, 그리고 Lambert는 인간-기계의 대화를 모델링 것이 아니라 인간-인간의 대화를 설명할 수 있는 모델링의 관점에 초점을 맞추었다. 하지만 Eller[6]와 Chu-Carrol[7][8]은 사용자의 대화 목적을 인식하고 동시에 이에 부합하는 시스템 응답 생성 모델을 제시하였다. 이들은 사용자의 목적 인식을 위한 계획 트리에 대응하는 시스템 계획 트리를 생성하고 propose-evaluate-modify cycle²⁾을 통해 시스템 응답을 생성한다. 이 모델은 두 개의 계획 트리를 통해 사용자와 시스템의 상태를 독립적으로 유지할 수 있다. 따라서 대화 참여자간에 서로 다른 지식의 상태를 모델링 할 수 있으며 지식의 충돌이 일어난 경우 발생할 수 있는 협상 대화(negotiation dialogue)의 처리가 가능하다. 최근 Allen[9]은 계획기반 모델을 사용하여 학습이 가능한 시스템을 제안하였다. 하지만 이 시스템은 정형화된 특정 형태의

1) 계획 인식을 위한 담화, 영역 지식 표현 형식

2) 사용자와 시스템의 지식이 충돌할 때 시스템이 대안을 제시하고 검증을 통하여 사용자의 믿음을 변화시키는 일련의 과정

대화만을 고려하고 있다.

계획기반 모델을 한국어에 적용한 연구로는 윤철진[10], 오종건[11] 그리고 조영환[12]의 연구가 있다. 윤철진의 연구는 생략 현상을 고려하여 한국어에 적합한 계획기반 모델을 제안하였으며 오종건의 연구에서는 한국어에서 협상 대화가 가능한 모델을 제시하였다. 조영환의 연구는 작업 지향 대화에서 혼합주도(mixed-initiative) 대화에 적합한 응답 모델을 제시 하였고 시스템의 응답 성향을 조절할 수 있는 기법을 포함시켰다.

확률기반 모델

기존의 대화 시스템들이 음성 인식 시스템과의 결합이 시도 되면서 각 시스템들의 오류 전파 문제들이 제기되고, 시스템 구축과 유지관리의 효율성을 극복하기 위하여 확률기반 모델을 통한 접근 방법들이 시도되고 있다. Roy[13]는 MDP(Marcov Decision Process)을 이용하여 대화에서 일어날 수 있는 각각의 모든 상태를 모델링하여 현재 상태가 다른 상태로 전이될 확률을 예측함으로써 대화를 진행하는 방법을 제안하였고, Singh[14]은 강화 학습(Reinforcement Learning)을 사용하여 사용자의 보상(reward)을 통해 시스템의 대화 전략을 학습해 나가는 방법을 제안하였다. 이 방법들은 극히 제한된 상태들을 정의할 수 있고, 수차례의 반복 학습하는 과정이 필요하기 때문에 아직까지 실제 환경에서의 효율성은 매우 낮다고 할 수 있다.

효과적인 응답생성을 위한 계획기반 모델

실제 로봇개발 등에 사용되는 대화 시스템은 여러 제안하는 영역의 작업 지향 대화와 정보 획득 대화를 동시에 고려할 수 있어야 하며 지능적인 대화 기법을 갖추어야 한다. 또한 영역 확장이 가능하고 유지, 관리에 용이한 구조를 가져야 한다. 여기서 지능적인 대화 기법이란 시스템이 사용자가 명시적인 목적을 표현하지 않더라도 사용자의 심층 목적을 찾아내어 자연스러운 대화를 할 수 있는 전략을 갖추는 것을 말한다.

본 연구에서는 사용자의 심층적인 목적을 인식하기 위해 계획기반 모델을 사용한다. 실제 응용 환경에서 인간-기계 대화는 작업 지향 혹은 정보 획득을 위한 대화가 혼재하여 진행되지만 인간-인간의 대화보다는 낮은 수준의 대화가 발생할 것이다. 따라서 계획기반 모델을 간소화 하여 복잡성을 줄이고 영역 확장성과 유지, 관리에 용이한 강인한 구조를 갖는 실제 응용 환경에 적합한 대화 시스템 모형을 제시한다.

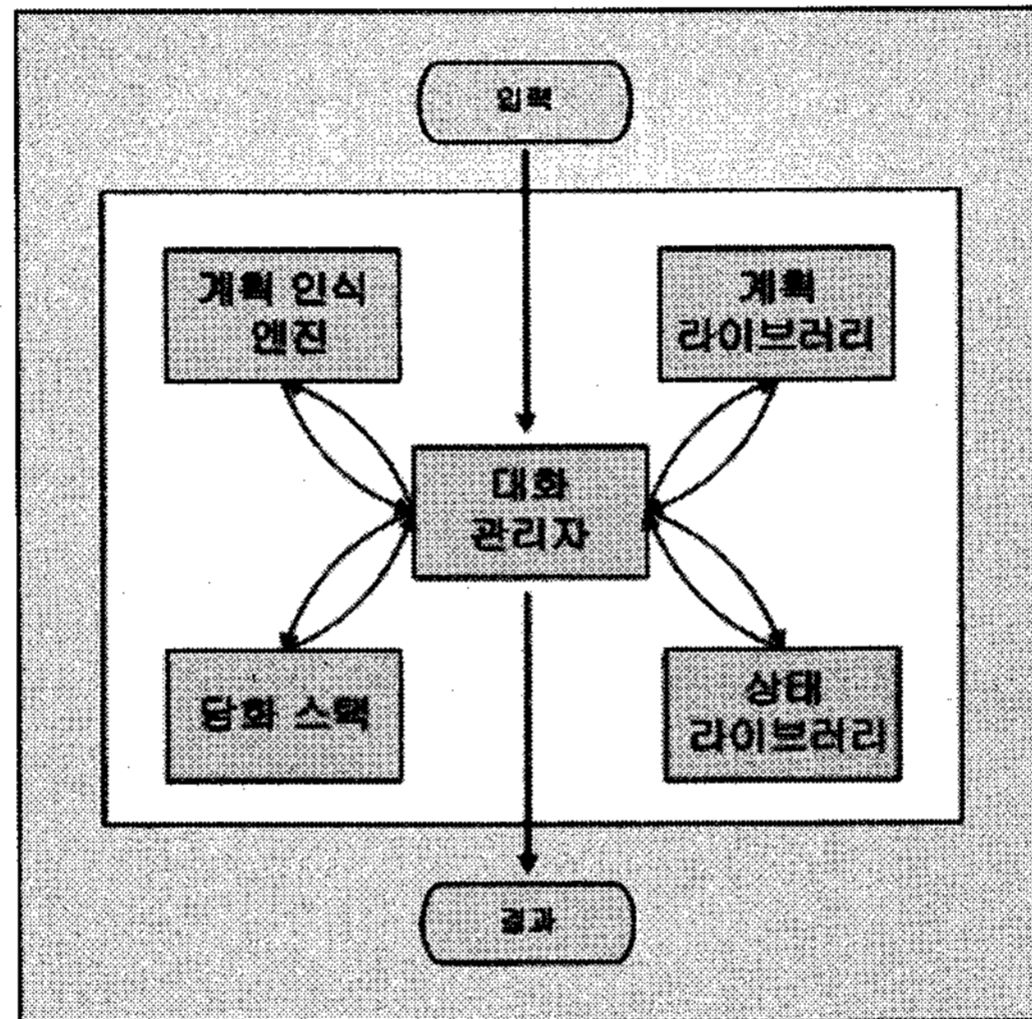


그림 1. 전체 시스템 구성도

(그림 1)에서 보는 바와 같이 제안하는 계획기반 모델에서 대화 관리자(dialogue manager)는 구축된 대화 전략을 사용하여 사용자의 목적을 인식하고 현재 상태에서 가장 적합한 응답을 선택하는 역할을 수행하며, 계획 인식 엔진은 사용자의 표층 목적을 바탕으로 가능성 있는 모든 심층 목적을 계획 트리 형태로 나타낸다. 계획 라이브러리는 계획 인식에 필요한 계획 지식들을 저장하고 있으며, 상태 라이브러리는 현재 사용자 상태와 실세계 상태를 저장한다. 대화 스택은 사용자와 시스템의 대화 구조를 저장하고 대화 관리자는 대화 전략에 따라 대화 스택을 운용한다.

계획 표현(Plan Representation)

계획 인식(plan recognition)을 위해 지식으로 사용되는 세부계획들을 표현하기 위해서 Allen[2]이 제시한 계획 라이브러리(recipe)형태의 방법을 참조하여 간략한 형태로 수정하였다(그림 2). 수정된 계획 표현의 속성들은 계획이름(plan name), 제약조건(constraint), 전제조건(precondition), 세부행위 (decomposition), 효과(effect), 그리고 목적(goal)으로 나누어진다. 계획이름은 계획의 단위가 되며 전제조건들은 해당 계획이 선택되기 위하여 반드시 필요한 조건들을 기술한다. 제약조건은 현재 선택된 계획들을 이루기 위하여 해결해야 되는 조건들이며, 세부행위는 현재 계획의 하위 계획들로서 이 하위 계획들이 모두 만족해야만 최상위에 있는 계획이 만족된다. 효과는 해당 계획이 만족되었을 때 반드시 참이 되는 명제이며 목적은 해당 계획이 만족 되었을 때 참이 되기를 희망하는 명제가 된다.

Eat- medicine (U) constraints : precondition : decomposition : eat (U, medicine), drink (U, water) effect : eaten (U, medicine) goal : eaten (U, medicine)
--

그림 2. 계획 지식의 표현

사용자 목적 인식 모델

대화 시스템에서 사용자의 대화 목적을 인식하는 것은 대화를 효율적으로 진행해 나가는데 있어서 매우 중요하다. 본 연구는 사용자 목적 인식을 위해서 2단계 계획 인식 기법을 사용한다. 본 시스템에서는 사용자의 발화 문장에 대한 언어 분석 결과를 계획 인식의 입력으로 사용한다. 언어 분석 결과는 술어 표현(predicate representation) 형식으로 변환하여 계획 인식 첫 번째 단계의 입력으로 사용되는데 화행 분석(speech act analysis)에 기반을 두고 있다.

계획 인식의 첫 번째 단계는 담화 수준(discourse level)의 추론을 처리한다. 담화

구조는 대화에서 나타나는 형식을 구조화한 것으로 일반적으로 인접 쌍(adjacency pair)의 원리를 기본으로 작성한다[15][16]. 예를 들면 ‘질문’과 ‘답변’, ‘명령’과 ‘수락’ 등과 같은 대화 쌍을 말하며 대화의 수준에 따라 담화 구조의 수준을 조절할 수 있다. 또한, 담화 수준 계획 인식 단계에서는 인접 쌍에 대한 계획 이외에도 인접 쌍을 단일한 계획으로 연결하기 위한 구조적인 계획(e.g. obtain-info-request())들도 포함된다. 두 번째 계획 인식 단계는 영역 수준(domain level)의 추론을 처리한다. 영역 수준은 단위 행위로 표현되는 모든 계획들을 포함한다. 영역 수준에는 단위 행위뿐만이 아니라 세부행위를 단위 행위로 간주하는 복잡한 형태의 행위를 추론할 수 있어야 한다. 영역 수준의 계획 인식 단계에서는 기본적으로 추론 가능한 모든 계획 트리(plan tree)를 생성한다. 그리고 대화 관리자는 상태 라이브러리를 참조하여 적절한 추론수준을 조절한다. 같은 수준에서 계획 인식 단계에서의 추론

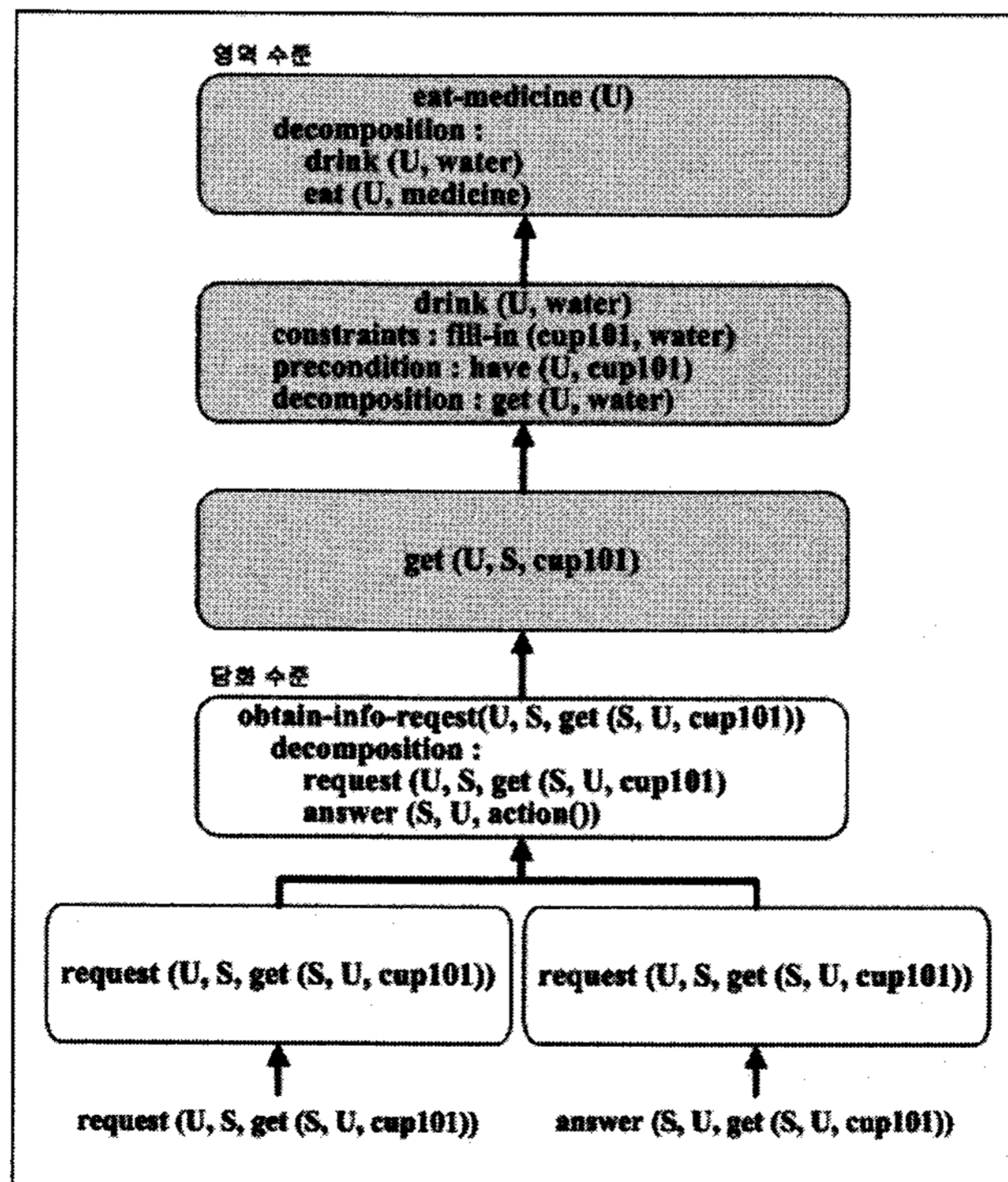


그림 3. 추론된 계획 트리의 예

규칙은 하위계획의 계획 이름이 상위계획의 세부행위에 포함되어 있으면 가능한 추론으로 정한다. 담화 수준에서 영역 수준으로의 추론 규칙은 담화 수준의 계획의 결과(goal)가 영역 수준의 계획의 전제 조건에 부합되는 경우에 가능한 추론으로 정한다. (그림 3)의 예를 통해 추론의 방법을 살펴보면 각 수준 안에서의 추론은 세부행위에 기술되어 있으며, 담화 수준으로부터 영역 수준으로의 추론은 'obtain-info-request()'의 결과(goal)가 'knowreq(U, get(S, U, cup101)' 이고 이 결과가 영역 수준의 계획 라이브러리인 'get()'의 전제조건에 포함되어 있어서 추론이 가능하다.

응답 생성 전략

앞 절에서 계획 인식엔진과 계획라이브러리를 사용하여 사용자 입력에 대한 가능한 추론 결과를 얻어내는 과정을 보였다. 본 절에서는 계획 인식엔진에서 얻어진 추론결과를 사용하여 시스템 응답을 생성하는 과정을 설명한다.

(그림 3)에서 사용자 입력에 대한 추론 결과는 3가지로 나누어진다. 첫 번째 추론 결과는 사용자가 요구한 행위에 대하여 영역 수준의 단위 행위인 'get(U, S, cup101)'까지 추론한 것이며 이것은 사용자가 명시적으로 요구한 행위에 해당하기 때문에 명백하게 참이 된다. 하지만 보다 지능적인 대화 시스템을 위해서 사용자의 심층적인 목적을 파악할 수 있어야 한다. 따라서 대화 관리자는 계획 추론 엔진을 통해 상위계획인 'drink(U, water)' 혹은 'eat-medicine(U)'까지 추론한 결과들을 얻어 낸다. 하지만 'drink(U, water)' 혹은 'eat-medicine(U)'까지 추론한 결과들은 시스템이 가능하다고 판단한 추론 결과이기 때문에 명백한 사용자의 목적으로 볼 수 없다. 따라서 시스템은 확인대화 전략을 통해 능동적으로 사용자의 목적을 파악한다.

계획기반모델에서 효과적인 시스템 응답을 생성하기 위해 본 연구에서는 일반적인 주대화와 부대화를 통합하여 처리할 수 있는 담화 스택 구조를 제안한다. 주대화는 사용자의 목적에 대하여 시스템이 명백하게 응답할 수 있는 경우에 해당하는 것이며 부대화는 현재 유지하는 문맥을 벗어난 부수적인 대화를 의미한다. 대부분 사용자의 목적을 시스템이 처리하기 위하여 부족한 정보를 요구하는 대화가

부대화에 해당하며 시스템 확인 발화도 이에 해당한다.

시스템 응답을 생성하기 위한 담화 스택 운용규칙은 다음과 같으며 공통적으로 스택의 가장 상위에 있는 내용이 시스템 응답인 경우 그 내용을 선택하여 시스템 발화를 생성한다.

담화 스택 운용규칙

1. *step1*: 사용자의 표층 목적을 저장.

2. *step2*: 심층목적 추론결과가 존재하면 영역 확인 질문을 스택에 저장.

① 대화 관리자는 사용자 상태 라이브러리의 내용 등을 활용하여 우선순위를 정하여 영역 확인 질문의 순서를 정함.

② 영역 확인 질문의 응답이 긍정이면 그 하위 영역에 대한 영역 확인 질문들은 자동 삭제.

③ 영역 확인 질문의 응답이 부정이면 다음 순서의 영역 확인 질문을 선택.

3. *step3*: 선택된 영역 계획에서 만족하지 못한 제약 조건들이 존재하면 그 조건을 확인하는 질문을 저장.

4. *step1*에서 저장된 내용이 가장 상위에 위치할 때 까지 *step2*와 *step3*를 반복.

(그림 4)는 담화 스택이 처리되는 과정을 설명한 예이다. 대화 관리자는 “저 컵 가져와라” 라는 대화를 입력으로 받았을 때 *step1*에 의하여 사용자의 표층 목적, 즉 단순히 컵을 가져오기를 원하는 경우(*get(U, S cup101)*)를 저장한다. 계획 인식 엔진은 사용자의 표층 목적 외에 심층목적인 물을 마시기 위해 컵을 가져오는 경우(*drink(U, water)*)와 약을 먹기 위해 컵이 필요한 경우(*eat-medicine(U)*)를 추론할 수 있다. 따라서 대화 관리자는 *step2*에서 사용자가 이들 대화 목적들 중 실제로 어떤 목적을 가지고 발화를 했는지를 확인하기 위하여 영역 확인 질문을 수행하게 되는데 이를 위해서 담화 스택을 사용한다. 이 때 담화 스택에 저장되는 순서가 매우 중요한데, 이는 저장되는 순서에 따라 발화가 진행되므로 좀더 효과적이고 지능적인 발화를 생성하기 위해서이다. 담화 스택에 저장되는 순서는 사용자 모델링의 결과로부터 결정 되는데 사용자 모델링의 결과는 상태 라이브러리에 저장되어 있다. 논문에서 사용하는 예의 경우에는 시스템은 사용자가 약을 먹는 시간을 사용

자 모델링을 통해 알고 있고, 현재 사용자가 약을 먹는 시간이므로 “약을 드시겠습니까?”에 해당되는 영역 확인 질문이 스택의 가장 위에 저장된다. 약 먹는 영역에 대한 확인 질문(domain-confirm-question(S, U, eat-medicine(U)))에 대해 사용자가 긍정으로 응답을 하면 현재 영역 확인 질문과 같이 스택에 저장되었던 다른 영역 확인 질문들은 (그림 4)의 스택과 같이 동시에 삭제되고 사용자가 부정의 답변을 하면 다음에 해당되는 영역확인질문을 발화하게 된다.

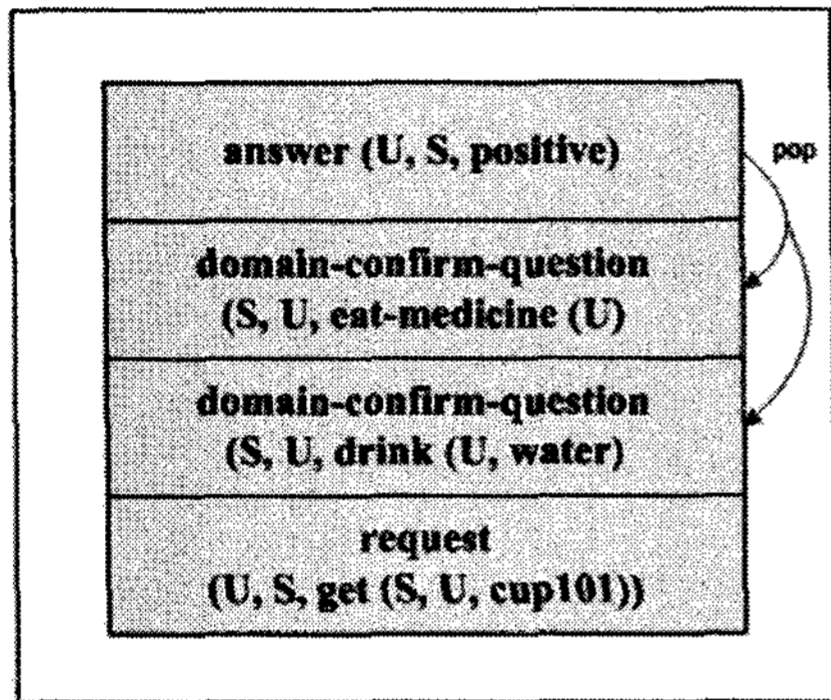


그림 4. 담화 스택 활용의 예(1)

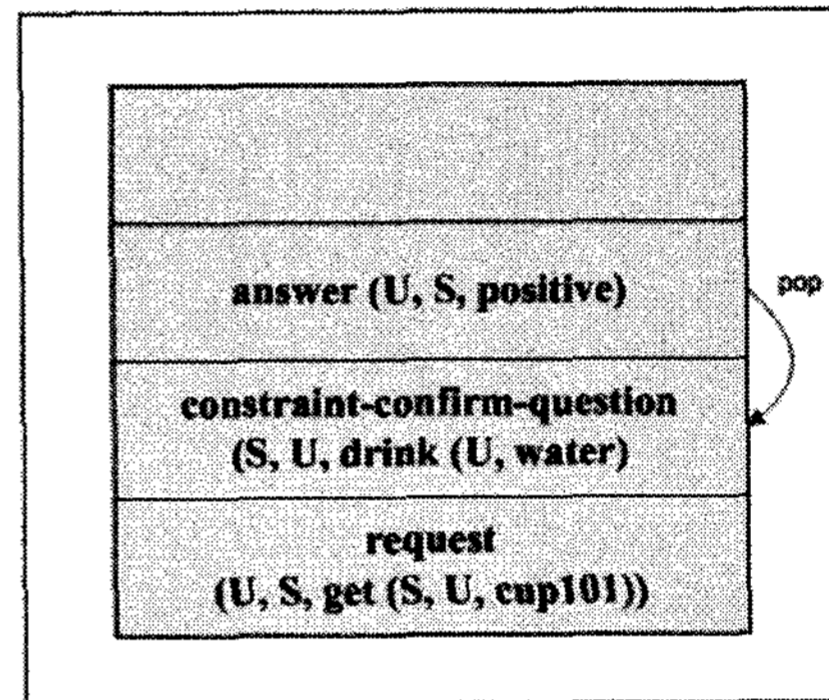


그림 5. 담화 스택 활용의 예(2)

Step3에서 대화 관리자는 현재 선택된 사용자 목적(eat-medicine(U))에 대한 계획에서 만족되지 않은 제약 조건(constraint)이 있을 경우에는 사용자에게 확인 시키는 질문을 생성할 수도 있다. (그림 3)의 예에서 계획(drink(U,water))의 제약조건에 보면 'fill-in(cup101, water)'가 있다. 즉, 컵에 물이 있어야 물을 마실 수 있다는 조건이다. 컵에 물이 없는 경우는 (그림 5)와 같이 담화 스택에 조건 확인 질문에 대한 질문 (“컵에 물을 채울까요?”)을 넣어서 제약조건에 대한 발화를 효과적으로 처리할 수 있다.

대화 관리자는 사용자의 응답에 의해 사용자의 목적을 이루기 위한 모든 계획의 조건들이 만족되었을 경우 step1에서 저장된 내용을 삭제하고 작업을 완료한다 (그림 6, 그림 7).

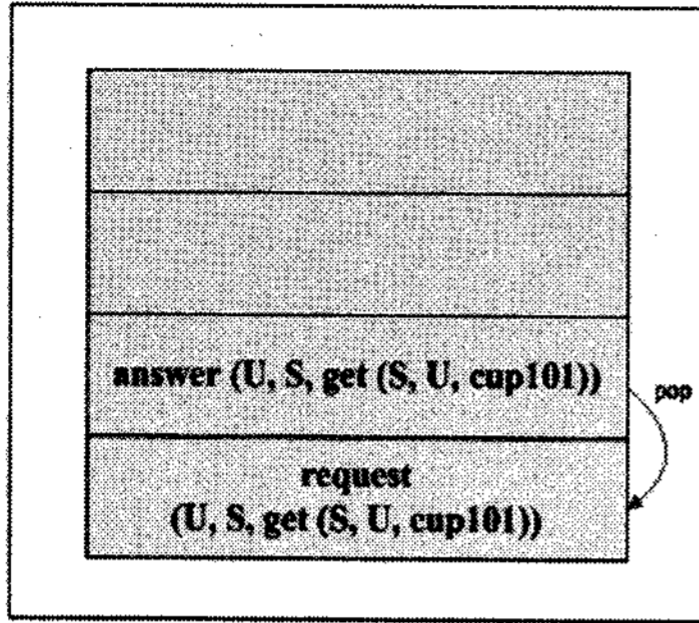


그림 6. 담화 스택 활용의 예(3)

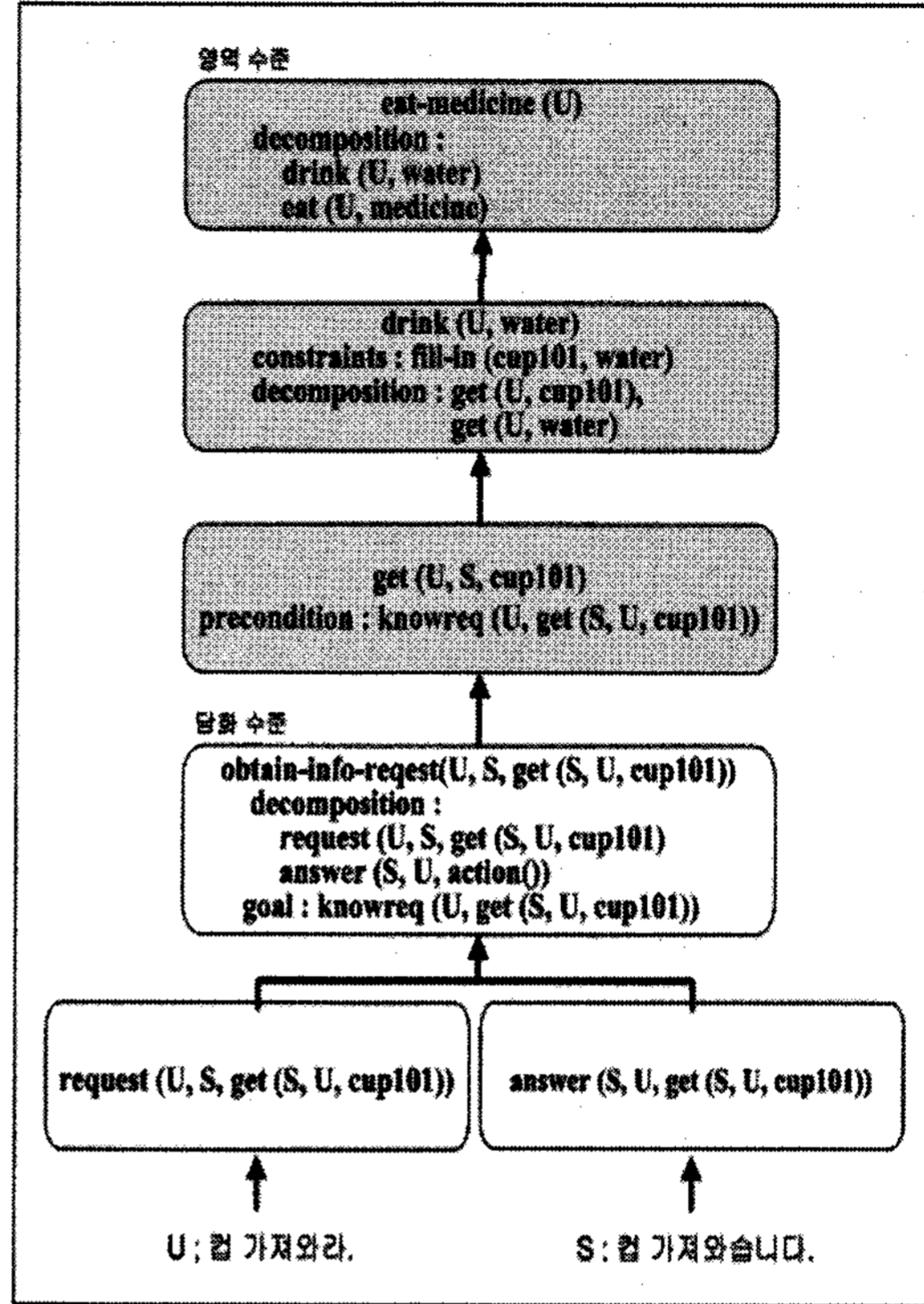


그림 7. 완성된 계획 트리

기존 시스템의 응답 모델과의 비교

본 논문에서 제안한 응답 생성 모델은 보다 효과적인 응답을 생성하기 위해서 계획기반의 새로운 모델을 제시하는 것이며 또한 대화시스템의 특성상 정량적으로 다른 시스템과의 비교가 매우 어렵다. 그러므로 본 절에서는 대화 과정에서 제안한 방법이 기존의 방법과 어떤 차이점이 있고 효율성의 측면에서 어떤 이점이 있는지를 예를 통해서 설명한다. 다음 예제들은 사용자가 약을 먹어야 하는 시간에 '컵 가져와라'라는 요청을 하는 경우를 예를 들어 설명한다. (예 1)은 기존의 사용자 주도적 대화시스템의 발화의 예로써 현재 사용자가 '약 먹기'라는 작업을 원하

고 있는지를 파악하지 못하여, 사용자가 요청한 3개의 작업을 각각 수행한 경우이다. 이 경우 사용자는 각 작업을 매번 지시하여야 하며 시스템은 각 작업을 모두 정확하게 인식하고 수행해야 하기 때문에 작업 성공률은 매우 낮아질 가능성이 높다.

(예 1)

사용자 : 컵 가져와라.

시스템 : 컵 가져왔습니다.

사용자 : 물도 가져와야지.

시스템 : 물 가져왔습니다.

사용자 : 약도 가져와.

시스템 : 약 가져왔습니다.

(예 2)에서는 시스템이 예상한 사용자의 작업을 확인함으로써 시스템 주도의 작업을 진행 된다. 이로써 사용자는 시스템의 질문에 대하여 간단한 응답만으로 대화를 진행할 수 있으며, 시스템 측면에서는 음성인식과 작업 성공률을 고려할 경우 매우 효과적인 형태의 대화를 진행할 수 있다. 대화의 내용면에서도 제안한 시스템은 조건 상태에 따라 유연하고 보다 지능적인 응답을 생성함으로써 보다 지능적인 대화가 가능하다.

(예 2)

사용자 : 컵 가져와라.

시스템 : 약을 드시려 하십니까?

사용자 : 그래.

시스템 : 컵에 물을 채울까요?

사용자 : 그래.

시스템 : 컵과 약을 가져왔습니다.

논 의

3.1절에서 논의했듯이 담화 수준(discourse level)의 대화추론과 담화 스택의 활용에 있어서 인접 쌍(adjacency pair)의 원리는 매우 중요하게 사용된다. 여기서 인접 쌍이란 ‘질문’과 ‘답변’, ‘명령’과 ‘수락’ 등과 같은 대화 쌍을 말하며 대화 모델에서는 담화 수준의 대화를 인식하고 생성하기 위한 규칙으로 사용되며 담화 스택에서는 스택에 들어간 발화를 언제 제거(pop)할 것인가를 결정하기 위한 규칙으로 사용된다. 본 연구에서는 대화모델의 입력으로 사용되는 언어 분석결과를 17개의 화행[17]을 바탕으로 생성해내는데 이들 화행의 인접 쌍을 만들기 위해 화행을 분석하여 다음 <표 1>과 같이 3가지 유형으로 나누었다.

표 1. 대화처리를 위한 화행의 3가지 유형 분류

유형	화행
요청 유형	ask-ref, ask-if, ask-confirm, offer, suggest, request
응답 유형	accept, answer, reject, acknowledge
단독사용유형	opening, introducing-oneself, correct, inform, expressive, promise, closing

결 론

본 논문은 인간-기계 상호작용의 차세대 기술로서 대두되고 있는 대화 시스템 개발을 위해서 실제 응용 환경에서 적용 가능한 복잡한 담화현상을 처리 할 수 있으면서도 견고한 구조를 가진 대화모델을 제시 하였다. 제안된 시스템은 계획기반 모델과 담화 스택을 이용하였으며 예를 통해 제안된 시스템이 어떻게 효율적인 응답 발화를 생성해 낼 수 있는지에 대해 설명하였다. 또한, 기존의 시스템들과의 차별화 된 내용은 기존 시스템들이 특정 작업이 선택되면 하향식으로 작업을 진행하는 반면에 제안한 시스템은 사용자의 발화에 내포된 실제 의도, 즉 상위 계획의 추론이 가능한 방법을 제시함과 동시에 효율성과 견고성을 증대시킬 수 있는 구조

를 제안하였다. 하지만 제안한 방법은 대화 시스템의 특성상 정량적인 평가가 매우 어렵고 국내의 비교 대상 대화 시스템이 전무하기 때문에 객관적인 평가 자료가 미약하였다. 현재 국외에서는 대화 시스템의 객관적인 평가를 위한 표준안을 확립하기 위한 노력이 진행 중이다. 예로 향후 사용자 평가를 대신할 수 있는 시뮬레이터 등의 개발을 통해 공정한 평가 자료를 제시할 수 있을 것으로 예상된다. 향후 연구로는 새로운 평가방법의 도입과 더불어, 새로운 영역의 추가가 보다 용이하도록 이식성을 증대시키고, 영역지식과 대화 모델의 명확한 분리를 통해 유지 관리에 효율적인 견고한 구조의 대화 모델을 개발하고자 한다.

참고문헌

- [1] David Goddeau, Helen Meng, Joe Polifroni, Stephanie Seneff, and Senis Busayapongchai, "A Form-Based Dialog Manager for Spoken Language Applications", *Proceedings of the ICSLP 96*, pp701-705, 1996.
- [2] H. A. Kautz and J. F. Allen. "Generalized Plan Recognition," *Proceedings of the 14th National Conference on Artificial Intelligence*, pp 32-37, 1986.
- [3] Diane J. Litman and James F. Allen, "A Plan Recognition Model for Subdialogue in Conversations," *Cognitive Science*, vol 11, pp163-200, 1987.
- [4] Lynn Lambert and Sandra Carberry, "A tripartite plan-based model of dialogue," *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pp47-54, 1991.
- [5] Lynn Lambert, *Recognizing Complex Discourse Acts : A Tripartite Plan- Based Model of Dialogue*, Ph.D. thesis, University of Delaware, 1993.
- [6] Rhonda Eller and Sandra Carberry, "A meta-rule approach to flexible plan recognition in dialogue," *User Modeling and User-Adapted interaction*, vol 2, num 1-2, pp27-53, 1992.
- [7] Jennifer Chu-Carroll and Sandra Carberry, "Generating information-sharing sub-dialogues in expert-user consultation," *Proceedings of the 14th International Conference on*

- Artificial Intelligence, pp1234-1250, 1995.
- [8] Chu-Carroll and Carberry, "Conflict Resolution in Collaborative Planning Dialogues," International Journal of Human-Computer Studies, Vol. 53(6), pp. 969-1015, 2000.
- [9] James Allen, Nathanael Chambers, George Ferguson, Lucian Galescu, Hyuckchul Jung, Mary Swift, and William Taysom, "PLOW: A Collaborative Task Learning Agent", Proceedings of the 21th AAAI Conference on Artificial Intelligence, pp1514-1519, 2007
- [10] 윤철진, 서정연, "제한된 영역의 대화에서 체언구 형태의 발화 이해를 위한 계획기반 생략 처리," 한국인지과학회 논문지, 제11권 1호, pp81-92, 2000.
- [11] 오종건, 계획에 기반한 대화 시스템의 설계, 석사학위논문, 서강대학교, 1999.
- [12] 조영환, "작업 수행 대화 시스템에서 혼합주도형 응답 생성 모형", 박사학위 논문, 한국과학기술원, 1997
- [13] Nicholas Roy, Joelle Pineau and Sebastian Thrun, "Spoken Dialog Management Probabilistic Reasoning", Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics, pp. 93-100, 2000
- [14] Satinder singh, Diane Litman, Michael Kearns and Marilyn Walker, "Optimizing Dialogue Management with reinforcement Learning", Journal of artificial Intelligence Research, 16, pp.105-133, 2002
- [15] E. Schegloff and H. Sacks, "Opening up closings," Semiotica, Vol. 7(4), pp. 289-327, 1973.
- [16] S. Levinson, Pragmatics, Cambridge University Press, Cambridge, UK, 1983.
- [17] 김경선, "개선된 자질 추출 및 가중치 부여 방법을 이용한 한국어 화행 분류 시스템," 박사학위논문, 서강대학교, 2006.

1 차원고접수 : 2007. 4. 19

2 차원고접수 : 2007. 6. 21

최종게재승인 : 2008. 4. 20

(Abstract)

Using Plan Recognition and a Discourse Stack for Effective Response Generation in a Dialogue System

Sangwoo Kang

Sogang University

Youngjoong Ko

Dong-A University

Jungyun Seo

Sogang University

The existing studies of a dialogue system can be classified into two major parts. One is a study for a practical system, and the other is a study to understand a principal of dialogue phenomena. The former focuses on robustness in real environment for dialogue systems. However, it cannot guarantee its performance in complicated dialogue environment. The latter has studied as the plan-based model typically. It has strong points that it can reflect complex dialogue phenomena and can infer user's intention in various situations. However, an initial design of this model is so complicated, and it is difficult for this model to be extended to the interaction model for response generation in a practical dialogue system. This paper proposes a new dialogue modeling using plan recognition and a discourse stack to effectively generate response in a practical dialogue system.

Keywords : dialogue system, response generation, dialogue manager, plan-based model