

대화시스템을 위한 계획인식과 담화스택을 이용한 효과적인 응답 생성

강상우¹ 고영중² 서정연³

서강대학교 컴퓨터학과¹

동아대학교 컴퓨터공학과²

서강대학교 컴퓨터학과/바이오융합기술 협동과정³

swkang@sogang.ac.kr¹, yjko@dau.ac.kr², seojy@sogang.ac.kr³

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

Sangwoo Kang¹ Youngjoong Ko² Jungyun Seo³

Dept. of Computer Science and Engineering, Sogang University¹

Dept. of Computer Engineering, Dong-A University²

Dept. of Computer Science and Interdisciplinary Program of Integrated Biotechnology,
Sogang University³

요 약

대화 시스템에 관한 기존 연구는 대화 현상에 대한 원리를 이해 하는데 초점을 맞춘 연구와 매우 제한적인 상황에서 동작하는 실용적인 시스템 구축에 관한 연구로 이루어져 왔다. 전자의 연구를 위해서 계획기반모델(plan-based model)이 제안되었는데, 이는 복잡한 대화 구조를 모델링(modeling)할 수 있으며, 다양한 현상에서의 사용자 목적 추론이 가능하다. 하지만 계획기반모델은 초기 설계가 어려우며 실용적인 대화 시스템 구축에 있어서 시스템 응답을 생성하기 위한 상호작용 모델로의 확장이 매우 어렵다는 단점이 있다.

본 연구는 이러한 계획 기반 모델의 단점을 보완하고 실용적인 대화시스템을 구축하기 위하여 시스템 응답을 위한 확인 대화 전략과 담화스택(discourse stack)을 계획기반 대화 모델링에 적용하여 효율적인 응답을 생성할 수 있는 기법을 제안한다.

1. 서론

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 대화시스템(dialogue system)에 대한 연구가 활발히 이루어지고 있다. 인간과 기계가 상호작용의 도구로서 자연언어(natural language)를 사용하여 특정 작업을 성취하도록 하는 프로그램을 대화 시스템이라 하며 대화시스템이 다른 인터페이스

보다 우수한 이유는 자연어가 가지고 있는 유연성(flexibility), 명료성(succinctness), 표현력(expressiveness) 등이 우수하기 때문이다.

대화시스템은 대화의 목적에 따라 작업수행과 정보제공의 2가지 역할을 수행한다. 사용자가 대화시스템을 사용하여 특정 작업을 수행하기 위한 대화를 진행하고 있다면 그러한 대화를 작업지향대화(task-oriented dialogue)라 하고 정보를 얻기 위한 대화를 진행하고 있다면 그러한 대화를 정보획득대화(information seeking dialogue)라 한다. 인간의 대화에는 작업지향 대화와 정보획득대화가 혼재 되어 있으며 구조적으로 매우 복잡한 형태를 갖는다. 대화 시스템을 실제 환경에서 사용하기 위해서는 이런 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.

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

본 연구에서는 위와 같은 조건을 만족시키기 위해 담화스택(discourse stack)과 계획기반 모델(plan-based model)을 이용한 대화시스템을 제안한다. 계획 기반 모델은 복잡한 담화 현상을 처리 할 수 있으며 계층 구조를 갖는 영역들에 대해서도 일관성 있는 모델링이 가능하다. 하지만 계획 기반 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 대화시스템에서는 계획기반모델의 성능을 유지하면서 인간-기계 상호작용 모델로서 시스템의 목적을 수행하기 위하여 응답확인 대화전략과 담화스택을 계획기반모델에 적용하여 강인한 구조의 응답모델 생성하는 기법을 제안한다.

2. 관련 연구

현재까지의 대화 시스템에 관한 기존의 연구는 제한적이지만 실용적인 시스템 구축에 관한 연구와 대화 현상을 이해 하는데 초점을 맞춘 분야로 분류할 수 있다. 실용적인 시스템의 대표적인 모델인 스크립트 기반 모델은 일반적으로 순차적인 실행 구조를 가지며 비교적 쉬운 설계가 가능하지만 영역에 매우 의존적이며 다양한 상황에 대응이 어렵다. 계획 기반 모델은 대화 현상을 이해 하는데 초점을 맞춘 연구에 속하며 현재 상태에 따라 다양한 사용자 목적 추론이 가능하며 복잡한 담화 구조의 처리가 가능하다.

Kautz와 Allen[1][2]은 처음으로 계획 기반 모델에 관한 구체적인 이론은 제시하였다. 그들은 대화를 인식하기 위한 모델을 담화수준과 영역수준으로 분리하였다. 그들의 이론에서 각각의 관측된 행위는 담화수준을 통하여 하나 이상의 영역 수준의 일부로 간주한다. 그리고 계획인식이란 그 행위를 가장 잘 설명하는 계획들의 집합을 최소화하는 작업으로 설명하고 있다.

Lambert[3][4]는 보다 복잡한 담화 구조를 처리할 수 있도록 담화수준과 영역수준 사이에 메타수준을 추가하였다. Lambert의 모델은 사용자의 복잡한 여러 개의 발화가 1개의 목적에 부합될 수 있도록 설계되었다.

Kautz와 Allen, 그리고 Lambert는 인간-기계의 대화를 모델링 것이 아니라 인간-인간의 대화를 이해 할 수 있는 모델링 하는 것에 초점을 맞추었지만 Eller[5]와 Chu-Carrol[6][7]은 사용의 목적에 부합하는 시스템 응답을 생성하기 위해 시스템 계획을 추가하여 협상대화를 처리할 수 있는 모델을 제시 하였다.

계획 기반 모델을 한국어에 적용한 연구로는 윤철진[8]과 오종건[9]의 연구가 있다. 윤철진의 연구는 생략현상을 고려하여 한국어에 적합한 계획기반 모델을 제안하였으며 오종건의 연구에서는 한국어에서

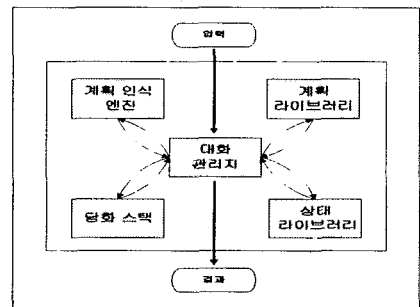
협상 대화가 가능한 모델을 제시하였다.

3. 효과적인 응답생성을 위한 계획 기반 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 계획 기반 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 계획 기반 모델을 간소화 하여 복잡성을 줄이고 영역 확장성과 유지, 관리에 용이한 강인한 구조를 갖는 실제 응용 환경에 적합한 대화 시스템 모형을 제시한다.

[그림 1]에서 보는 바와 같이 제안하는 계획 기반 모델에서 대화관리자(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 실 세계 상태를 저장한다. 담화스택은 사용자와 시스템의 담화구조를 저장하고 대화 관리자는 대화 전략에 따라 담화스택을 운용한다.

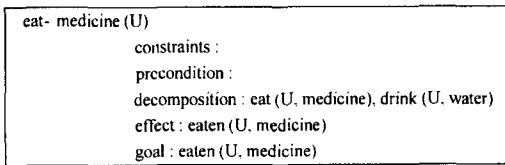


[그림 1] 전체 시스템 구성도

3.1 계획 표현(Plan Representation)

계획 인식(plan recognition)을 위해 지식으로 사용되는 세부계획들을 표현하기 위해서 Allen[1]이

제시한 계획라이브러리(recipe)형태의 방법을 참조하여 간략한 형태로 수정하였다 ([그림 2] 참조). 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.



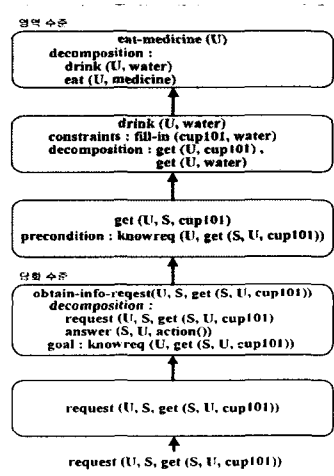
[그림 2] 계획 지식의 표현

3.2 사용자 목적 인식 모델

대화시스템에서 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 화행분석(speech act analysis)에 기반을 두고 있다.

계획인식의 첫 번째 단계는 담화수준 (discourse level)의 추론을 처리한다. 담화 구조는 대화에서 나타나는 형식을 구조화한 것으로 일반적으로 인접 쌍(adjacency pair)[10][11]의 원리를 기본으로 작성한다. 예를 들면 ‘질문’과 ‘답변’, ‘명령’과 ‘수락’ 등과 같은 대화 쌍을 말하며 대화의 수준에 따라 담화 구조의 수준을 조절할 수 있다. 또한, 담화수준 계획인식 단계에서는 인접 쌍에 대한 계획 이외에도 인접 쌍을 단일한 계획으로 연결하기 위한 구조적인 계획(e.g., obtain-info-request)들도 포함된다. 두 번째 계획인식 단계는 영역수준(domain level)의 추론을 처리한다. 영역수준은 단위 행위로 표현되는 모든 계획들을 포함한다. 영역수준에는 단위 행위뿐만이 아니라 세부행위를 단위 행위로 간주하는 복잡한 형태의 행위를 추론할 수 있어야 한다. 영역 수준의 계획 인식 단계에서는 기본적으로 추론 가능한 모든 계획트리(plan tree)를 생성한다. 그리고, 대화관리자는

상태 라이브러리를 참조하여 적절한 추론수준을 조절한다. 같은 수준에서 계획인식 단계에서의 추론규칙은 하위계획의 계획 이름이 상위계획의 세부행위에 포함되어 있으면 가능한 추론으로 정한다. 담화수준에서 영역수준으로의 추론규칙은 담화수준의 계획의 결과(goal)가 영역수준의 계획의 전제 조건에 부합되는 경우에 가능한 추론으로 정한다. 아래의 [그림 3]의 예를 통해 추론의 방법을 살펴보면 각 수준 안에서의 추론은 세부행위에 기술되어 있으며, 담화수준으로부터 영역수준으로의 추론은 ‘obtain-info-request’의 결과(goal)가 ‘knowreq(U, get(S, U, cup101))’이고 이 결과가 영역수준의 계획 라이브러리인 get의 전제조건에 포함되어 있어서 추론이 가능하다.



[그림 3] 추론된 계획트리의 예

[그림 3]은 사용자가 시스템에게 컵(cup101)을 가져오라고 요청하는 문장을 입력으로 사용하여 담화수준의 계획부터 최상위 영역 수준의 계획까지 추론한 형태를 보여준다. 결론적으로 “저 컵 가져와라”는 문장은 ‘단순히 컵을 가지고 오는 계획’과 ‘물을 마시는 계획’ 그리고 ‘약을 먹는 계획’을 활성화시킨다.

3.3 응답 생성 전략

3.2절에서 계획인식엔진과 계획라이브러리를 사용하여 사용자 입력에 대한 가능한 추론 결과를 얻어내는 과정을 보였다. 본 절에서는 계획인식-엔진에서 얻어진 추론결과를 사용하여 시스템응답을 생성하는 과정을 설명한다.

[그림 3]에서 사용자 입력에 대한 추론 결과는 3가지로 나누어진다. 첫 번째 추론 결과는 사용자가 요구한 행위에 대하여 영역 수준의 단위 행위(get(U, S, cup101))까지 추론한 것이며 이것은 사용자가

명시적으로 요구한 행위에 해당하기 때문에 명백하게 참이 된다. 하지만 보다 지능적인 대화시스템을 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 확인대화전략을 통해 능동적으로 사용자의 목적을 파악한다.

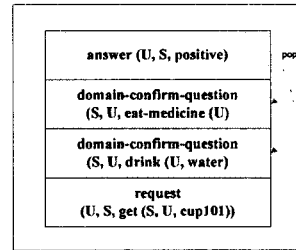
계획기반모델에서 효과적인 시스템응답을 생성하기 위해 본 연구에서는 일반적인 주대화와 부대화를 통합하여 처리할 수 있는 담화스택 구조를 제안한다. 주대화는 사용자의 목적에 대하여 시스템이 명백하게 응답할 수 있는 경우에 해당하는 것이며 부대화는 현재 유지하는 문맥을 벗어난 부수적인 대화를 의미한다. 대부분 사용자의 목적을 시스템이 처리하기 위하여 부족한 정보를 요구하는 대화가 부대화에 해당하며 시스템 확인 발화도 이에 해당한다.

시스템응답을 생성하기 위한 담화스택 운용규칙은 다음과 같으며 공통적으로 스택의 가장 상위에 있는 내용이 시스템응답인 경우 그 내용을 선택하여 시스템발화를 생성한다.

- 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) step1에서 저장된 내용이 가장 상위에 위치할 때까지 step2와 step3를 반복

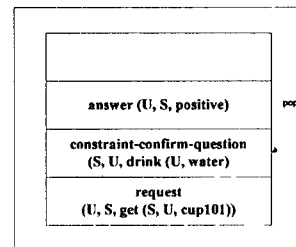
[그림 4]는 담화스택이 처리되는 과정을 설명한 예이다. 대화 관리자는 “저 컵 가져와라” 라는 대화를 입력으로 받았을 때 step1에 의하여 사용자의 표층 목적, 즉 단순히 컵을 가져오기를 원하는 경우(get(U, S cup101))를 저장한다. 계획 인식 엔진은 사용자의 표층 목적 외에 심층목적인 물을 마시기 위해 컵을 가져오는 경우(drink(U, water))와 약을 먹기 위해 컵이 필요한 경우(eat-medicine(U))를 추론할 수 있다. 따라서 대화 관리자는 step2에서 사용자가 이들 대화 목적들

중 실제로 어떤 목적을 가지고 발화를 했는지를 확인하기 위하여 영역 확인 질문을 수행하게 되는데 이를 위해서 담화스택을 사용한다. 이 때 담화스택에 저장되는 순서가 매우 중요한데, 이는 저장되는 순서에 따라 발화가 진행되므로 좀더 효과적이고 지능적인 발화를 생성하기 위해서이다. 담화스택에 저장되는 순서는 사용자 모델링의 결과로부터 결정되는데 사용자 모델링의 결과는 상태 라이브러리에 저장되어 있다. 논문에서 사용하는 예의 경우에는 시스템은 사용자가 약을 먹는 시간을 사용자 모델링을 통해 알고 있고, 현재 사용자가 약을 먹는 시간이므로 “약을 드시겠습니까? (eat-medicine(U))”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 왼쪽 스택과 같이 동시에 삭제되고 사용자가 부정의 답변을 하면 다음에 해당되는 영역확인질문을 발화하게 된다.



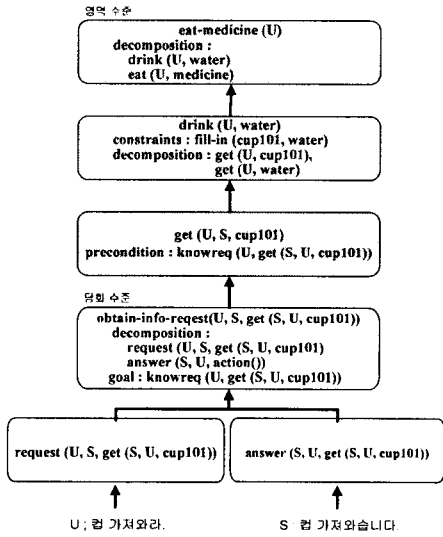
[그림 4] 담화스택 활용의 예(1)

Step3에서 대화관리자는 현재 선택된 사용자 목적(e.g. 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 담화스택에 조건확인 질문에 대한 질문(“물을 컵에 채울까요?”)을 넣어서 제약조건에 대한 발화를 효과적으로 처리할 수 있다.

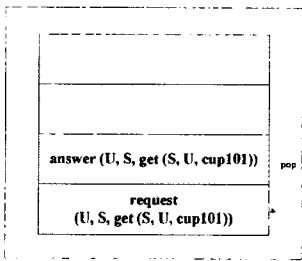


[그림 5] 담화스택 활용의 예(2)

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] 완성된 계획 트리



[그림 7] 담화스택 활용의 예(3)

4. 논의(Discussion)

3.2절에서 논의했듯이 담화 수준(discourse level)의 대화추론과 담화스택의 활용에 있어서 인접쌍(adjacency pair)의 원리는 매우 중요하게 사용된다. 여기서 인접쌍이란 ‘질문’과 ‘답변’, ‘명령’과 ‘수락’ 등과 같은 대화 쌍을 말하며 대화모델에서는 담화수준의 대화를 인식하고 생성하기 위한 규칙으로 사용되며 담화스택에서는 스택에 들어간 발화를 언제 제거(pop)할 것인가를 결정하기 위한 규칙으로 사용된다. 본 연구에서는 대화모델의 입력으로 사용되는 언어분석결과를 17개의 화행[12]을 바탕으로 생성해내는데 이들 화행의 인접쌍을 만들기 위해

화행을 분석하여 다음 [표1]과 같이 3가지 유형으로 나누었다.

[표 1] 대화처리를 위한 화행의 3가지 유형 분류

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

[표1]에서 보는 바와 같이 ‘요청유형’은 ‘응답유형’의 화행을 만나야만 담화스택에서 제거될 수 있다. 그리고 단독사용유형은 인접쌍이 존재하지 않는 유형의 화행이므로 담화스택에 관리될 필요가 없는 화행들이다.

5. 결론

본 논문은 인간-기계 상호작용의 차세대 기술로서 대두되고 있는 대화시스템 개발을 위해서 실제 응용 환경에서 적용 가능한 복잡한 담화현상을 처리 할 수 있으면서도 견고한 구조를 가진 대화모델을 제시 하였다. 제안된 시스템은 계획기반모델과 담화스택을 이용하였으며 예를 통해 제안된 시스템이 어떻게 효율적인 응답 발화를 생성해 낼 수 있는지에 대해 설명하였다.

향후 연구로는 제안된 시스템을 많은 응용영역에 적용하여 더욱 견고하고 인간의 대화에 가까운 발화를 생성해 낼 수 있는 대화모델을 개발하고자 한다.

참고문헌

- [1] H. A. Kautz and J. F. Allen. "Generalized Plan Recognition," *Proceedings of the 14th National Conference on Artificial Intelligence*, pp 32-37, 1986.
- [2] Diane J. Litman and James F. Allen, "A Plan Recognition Model for Subdialogue in Conversations," *Cognitive Science*, vol 11, pp163-200, 1987.
- [3] 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.
- [4] Lynn Lambert, *Recognizing Complex Discourse Acts : A Tripartite Plan-Based Model of Dialogue*, Ph.D. thesis, University of Delaware, 1993.
- [5] 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.
- [6] 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.

- [7] Chu-Carroll and Carberry, “Conflict Resolution in Collaborative Planning Dialogues,” *International Journal of Human-Computer Studies*, Vol. 53(6), pp. 969-1015, 2000.
- [8] 윤철진, 서정연, “제한된 영역의 대화에서 체언구 형태의 발화 이해를 위한 계획기반 생략 처리,” *한국인지과학회 논문지*, 제11권 1호, pp81-92, 2000.

[9] 오종건, *계획에 기반한 대화 시스템의 설계*, 석사학위논문, 서강대학교, 1999.

[10] E. Schegloff and H. Sacks, “Opening up closings,” *Semiotica*, Vol. 7(4), pp. 289-327, 1973.

[11] S. Levinson, *Pragmatics*, Cambridge University Press, Cambridge, UK, 1983.

[12] 김경선, “개선된 자질 추출 및 가중치 부여 방법을 이용한 한국어 화행 분류 시스템,” 박사학위논문, 서강대학교, 2006.