

제한된 영역의 대화에서 체언구 형태의 발화 이해를 위한 계획기반 생략 처리

Plan-based Ellipsis Resolution for Utterances in Noun-Phrase-Form in Restricted Domain Dialogues

윤철진* 서정연*
(Cheol-jin Yoon)(Jung-yun Seo)

요약 인간의 자연스러운 대화에서는 생략 현상이 빈번하게 일어난다. 생략 어구로부터 화자의 의도를 파악하는 것은 쉽지가 않다. 생략 어구 자체만 가지고는 그것의 의미와 화자의 의도를 파악하기 힘들며 이전 발화들로부터 구성된, 혹은 영역에 내재된 맥락과의 연관성을 살펴보아야 하기 때문이다. 본 연구에서는 Lambert가 제안한 3단계 계획기반 대화이해모델을 확장함으로써 한국어 대화에서 나타나는 생략 어구로부터 화자의 의도를 인식하는 모델을 제안한다. 먼저 Lambert의 모델에서 고려하지 않은, 생략 어구 형태의 발화를 통해 화자가 의도하는 담화 행위를 새로운 담화 recipe로 정의하여 추가하였다. 한국어에서는 조사가 화자의 의도를 나타내므로 이러한 특성을 이용하여 생략 어구를 표층 화행으로 표현할 때, 조사 정보를 포함하여 화자의 의도 인식을 용이하게 함을 볼 수 있었다. 또한 객체와 초점화 이론을 제안하여, 생략 어구를 통해서 화자가 두 개의 계획을 비교하여 고려하는 의도를 인식할 수 있도록 했다.

주제어 계획기반 대화이해모델, 체언구, 생략 처리

Abstract Elliptical fragments are common in natural language dialogues between humans. Since most elliptical fragments should be interpreted within the context, it is not easy for computers to recognize the speaker's intention from the elliptical fragments. In this paper, we propose a model to recognize speaker's intention from elliptical fragments in Korean by expanding the tripartite plan-based model proposed by Lambert. We add new discourse recipes to define user's discourse actions through elliptical fragments. In order to use plan inference process, we must represent utterances as actions, e.g., elliptical fragments are represented as surface speech acts. In surface speech act representation, we include the information of 'Josa' (case markers in Korean), because the information of 'Josa' plays a very important role in analysing speakers' intention in Korean. Finally, by using an object and discourse focus theory, the system can recognize the intention that a user is trying to compare between two plans by uttering elliptical fragments.

1. 서론

인간의 자연스러운 대화에서는 생략 현상이 빈번하게 일어난다. 그러므로, 인간의 대화를 잘 이해하려면 생략 어구의 처리가 중요하다. 그런데, 여기서 생략 어구의 처리란 단순히 표면적인 의미의 복원이어서는 안 된다. 화자의 생략 어구 발화에 적절하게 대응하

* 서강대학교 전자계산학과
Dept. of Computer Science, Sogang University)
연구 세부분야 : 자연언어처리, 한국어 처리
주소 : (121-742)서울시 마포구 신수동 1
서강대학교 전자학과 자연어 처리 연구실
TEL : 706-8954
E-Mail : yoonej@nlprep.sogang.ac.kr,
seojoy@ccs.sogang.ac.kr

기 때문이다. 그러나, 생략 어구로부터 화자의 의도를 인식하는 것은 쉽지 않다. 생략 어구 자체만으로는 그것의 의미와 화자의 의도를 파악하기 힘들며 이전 발화들로부터 구성된, 혹은 영역에 내재된 맥락과의 연관성을 살펴보아야 하기 때문이다.

- (1) 고객 : 5월 9일에 방 하나 예약하고 싶은데요.
 (2) 호텔 : 어떤 방으로 하시겠어요?
 (3) 고객 : 어떤 방이 있나요?
 (4) 호텔 : 침대방과 온돌방이 있습니다.
 (5) 고객 : 침대방은 값이 얼마죠?
 (6) 호텔 : 몇 분이 묵으실 건데요?
 (7) 고객 : 두 명이오.
 (8) 호텔 : 하루에 4만 5천원입니다.
 (9) 고객 : 온돌방은요?

(그림 1) 호텔 예약 영역에서 대화의 예

그림 1은 호텔 예약 영역에서 고객이 방을 예약할 때 호텔 직원과 나누는 대화의 예다. 밑줄 친 발화는 본 논문에서 처리하고자 하는 생략 어구이다. (7)번 생략 어구는 (6)번의 질문에 대한 대답으로 (6)번 발화를 고려하지 않고 (7)번 발화만 가지고는 고객의 의도가 무엇인지 알 수 없다. (9)번 발화는 (5)번 발화에 의해서 침대방의 가격과 온돌방의 가격을 비교해서 묻고 있는 것을 알 수 있다. 그러나, 이 경우 고객은 단순히 일반적인 온돌방의 가격을 묻고 있는 것이 아니다. 이전에 주고 받은 발화들에 의하면 고객이 원하는 온돌방은 가능한 날짜가 5월 9일이고, 가격은 2인실인 온돌방이다. 이렇게 이전 발화들에 의한 문맥을 고려하지 않고 단순히 온돌방의 가격을 묻는 것이라고 판단한다면 고객이 투숙하고자하는 날짜와 투숙할 사람의 수를 다시 물어봐야 올바로 대답할 수 있을 것이다. 이처럼 생략 어구로부터 화자의 올바른 의도를 파악하기 위해서는 생략 어구 자체가 이전의 몇개의 발화들만 보아서는 안되며 대화 전체의 문맥과 영역에 관련된 맥락을 살펴봐야 한다.

화자의 생략 어구에 대해서 적절하고 협조적으로 응답하려면 생략 어구에 내재된 화자의 의도를 인식하는 것이 중요하다. 그러므로, 본 논문에서 제안하는 생략 어구의 처리란 단순히 생략 어구의 표면적인 의미 복원이 아니며 생략 어구에 내재된 화자의 의도를 인식하는 것을 의미한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 화자의 발화로부터 의도를 인식하기 위해 Lambert가 제안한 3단계 계획기반(Tripartite Plan-Based) 대화 이해모델을 살펴보고, 이 모델에 한국어 대화문장에서 나타나는 생략된 체언구를 처리할 수 있도록 수정, 보완시킨 내용에 대해 설명한다. 제 3장에서는 생략 어구가 의도하는 담화 행위의 결정 조건으로 생략 어구와 맥락과의 연관성에 대해 알아본다. 제 4장에서는 실제 대화 예제를 가지고 생략 어구로부터 화자의 의도를 인식하는 과정을 알아보고, 마지막으로 제 5장에서는 결론과 향후 과제에 대해 논한다.

2. 대화이해모델

본 장에서는 발화로부터 화자의 의도를 인식하는 모델인 계획에 기반한 대화이해모델에 대해 살펴본다. 2.1절에서는 Lambert가 제안한 3단계 계획기반 대화이해모델에 대해 간략히 설명하며, 2.2절에서는 생략 어구로부터 화자의 의도를 인식하기 위해 Lambert의 모델에 수정 및 추가된 내용에 대해 살펴본다.

2.1 3단계 계획기반 대화이해모델

3단계 계획기반 대화이해모델은 화자의 의도를 영역 수준(domain level), 문제해결 수준(problem-solving level), 담화 수준(discourse level)의 세가지 수준으로 세분하여 모델링하였다[14]. 영역 수준은 화자의 영역 목적(domain goal)을 이루기 위한 영역 행위들을 표현하며 문제해결 수준은 계획을 세우거나 영역 행위의 인수값을 채우는 행위 등 문제해결 행위들을 표현하고 담화 수준은 발화를 함으로써 의사소통을 하려는 담화 행위들을 표현한다.

계획기반 대화이해모델은 행위에 대한 시스템의 지식을 나타내는 일반적인 계획 라이브러리(plan library)를 가지고 있다. 일반적인 계획은 어떤 행위가 어느 조건하에서 일어나는가, 어떤 효과(effect)와 목적(goal)을 가지는가, 어떤 세부행위(subaction)로 이루어지는가 등의 지식을 나타낸다. 시스템은 사용자의 행위를 관찰하고 이러한 일반적인 계획들을 이용하여 그 행위의 목적을 추론해낸다. 이런 일반적인 계획은 recipe[14, 17], plan schemas[16], plan operators[9]라고 부르며 본 논문에서는 Pollack[17]의 표기를 따라 recipe라고 부른다.

3단계 계획기반 대화이해모델에서는 화자의 발화로부터 recipe 라이브러리, 계획추론규칙, 조건 만족,

초점화 휴리스틱(focusing heuristics) 등을 이용하여 대화모델을 구축한다. 우선, 화자의 발화로부터 얻은 표층화행을 recipe 라이브러리와 계획추론규칙에 적용하여 일련의 담화 행위들을 추론한다. 그리고, 추론된 일련의 담화 행위들로부터 다시 문제 해결 행위를 추론하고 마지막으로 추론된 문제 해결 행위로부터 영역 행위를 추론하여, 화자가 발화를 함으로써 의도하고자 하는 바를 세가지 수준으로 세분하여 모델링하게 된다. 그런데, 일련의 계획추론 과정에 의해 제시된 모든 행위를 발화자의 의도로 볼 수 있는 것은 아니다. 추론된 행위들의 조건 만족을 수행하여 계획추론규칙에 의해 추론된 행위들이 올바른가 확인해야 한다. 만약 조건 만족이 되지 않는다면 규칙에 의해 제시된 행위는 거절되며 올바른 행위들이 추론될 때까지 계획추론규칙에 의해 새롭게 제시된 행위들을 살펴보게 된다. 초점화 휴리스틱은 각 수준에 있는 초점(focus of attention)을 이용해서 이전의 발화들로 구축된 대화모델과 새로운 발화의 가장 결속성이 있는 관계를 제시한다. 여기서 초점이란 현재에 각 수준에서 대화참여자가 주의집중을 하고 있는 행위를 나타낸다.

2.2 추가 및 개선된 recipe

Lambert의 3단계 계획기반 대화이해모델은 생략 어구와 같은 특수 표현의 처리는 지원하지 않는다. 본 절에서는 생략 어구로부터 화자의 의도를 인식할 수 있도록 Lambert 모델에 추가 및 개선된 내용에 대해 살펴본다.

2.2.1 한국어의 특성을 고려한 생략 어구의 표층화행 표현

화행(Speech Act)이란 해당 발화 속에 포함되어 있는 화자에 의해 의도된 언어적 행위(알림, 요구, 명령, 축하 등)이다(8). 표층화행이란 문맥을 고려하지 않고 해당 발화의 언어적인 표층 정보만 가지고 정의한 화행을 말한다. 표층화행을 결정하기 위한 정보로는 문장 유형, 본용언, 보조 용언, 어미, 양상, 단서 단어 등이 있다(6). 발화를 행위로 표현하는 것, 즉 표층화행으로 표기하는 것은 계획에 기반한 화자의 의도 인식과정에서 가장 기본적인 일이다. 화자의 의도를 인식하는 기본적인 과정은 어떤 행위로부터 계획추론규칙에 의하여 일련의 행위들을 추론함으로써 화자가 의도한 행위를 추론하는 것이기 때문이다. 본 절에서는 생략 어구를 표층화행으로 표현하는 문제에

대해 살펴본다. 이를 위해, 한국어에서 조사와 화자의 의도와의 관계에 대해 알아보고, 조사를 표층화행에 표현함으로써 화자의 의도 인식에 도움을 주는 것을 보인다.

한국어는 교착어로서 어미의 활용과 조사의 곡용에 의하여 화자의 의도를 자유자재로 나타내는 것이 특색이다(4). 화자의 의도에 따라 조사의 활용이 다르게 나타나므로 본 논문에서는 화자의 의도를 인식하기 위해서 조사라는 언어적 단서(linguistic clues)를 이용한다.

(10) 시스템 : 손님의 이름을 대기리스트에 놓을까요?

(11) 사용자 : 대기리스트라니요?

종결형 서술격 조사 '(이)라니'는 상대방의 말이 의심쩍거나 놀랐다는 뜻을 나타낼 경우에 쓰이는 조사이다(1). 위 대화에서 시스템은 사용자의 (11)번 발화로부터 '라니'라는 언어적 단서를 통해서 사용자가 '대기리스트'라는 단어를 이해를 하지 못했고 그것의 의미를 알려 달라는 의도를 파악할 수 있다. 즉, '라니'라는 조사를 통해서 사용자의 담화 행위가 정체 확인 요구(Ask-Identify)라는 것을 인식할 수 있다. 이처럼 한국어에서 조사는 화자의 의도를 나타내며 생략 어구처럼 언어적 정보가 적은 발화로부터 화자의 의도를 잘 파악하려면 조사와 같은 언어적 단서를 잘 이용해야 화자의 의도를 인식할 수 있다.

보조사 '은/는'은 단순히 격조사가 아니고 기능적, 문맥적 또는 정서적인 의미를 갖고 있는, 이른바 보조조사로서 그 중에서도 가장 많이 쓰이고 있는 낱말이다(4). 보조사 '은/는'은 일반적으로 '대조성', '총칭성(generic)', '비대조-비총칭성'의 3가지 의미가 분석된다고 알려져 있는데, 그중에 '대조성'이 고유한 의미이고, 나머지는 이 의미와 관련된 것이거나, 아니면 상황적 의미에서 드러나는 것이라고 본다(3).

그림 1의 (9)번 발화에서는 보조사 '은/는'으로 말미암아 생략 어구의 의미가 '대조성'임을 보여준다. (9)번의 생략 어구에 포함된 보조사 '은/는'은 화자가 이전에 관심이 있던 객체와 '비교', '대조' 등의 의미를 나타낸다. 즉, 이전에 고려했던 객체인 '침대방'과 비교하여 온돌방의 가격을 묻는 것임을 알 수 있다. 이는 고객이 단순히 일반적인 온돌방의 가격을 묻는 것이 아니다. 침대방에서 고려했던 여러가지 속성들과 비교하여, 5월 9일에 두 명이 묵을 온돌방의

가격을 묻고 있는 것이다. 화자의 이러한 의도를 파악하는 것은 협조적이고 효율적인 응답 발화를 생성하는 데 있어서 중요하다. 고객이 묵을 날짜와 사람의 수를 모른다면 묵을 날짜에 여분의 '온돌방'이 있는지, 있다면 묵을 사람의 수에 따라서 가격은 정확히 얼마인지를 가르쳐 줄 수 없기 때문이다. 만약, 고객의 마지막 발화에서 보조사 '은/는'이 쓰이지 않는다면 발화의 의미는 크게 달라지게 된다. 예를 들면, "온돌방이오?"라고 발화한다면 "온돌방의 가격이 하루에 4만 5천원입니까?"라고 명확성 요구(Ask-Clarify) 행위가 될 수 있지만 문맥상 자연스럽게 않은 발화가 된다. 만약, 앞의 대화처럼 고객이 이전에 고려했던 침대방과 비교하여 온돌방의 어떤 속성을 물어보고자 한다면 보조사 '은/는'의 쓰임이 필수적이며 이러한 의도를 파악하기 위해서 생략 어구에 나타난 조사를 파악해야 한다.

앞에서 언급했듯이 보조사 '은/는'이 '비교', '대조'의 의미로만 쓰이는 것은 아니다. 보조사 '은/는'이 그 고유 의미인 '대조성'으로 말미암아 한정성의 의미를 가지고 주제를 지칭하는 데 쓰인다(3). 다음 대화에서 사용자의 마지막 발화에서는 '은/는'이 앞의 체언을 주제화하는 경우를 보여준다.

- (12) 시스템 : 그날 침대방은 있습니다.
- (13) 사용자 : 가격은요?

주제화란 문장 속에서 '설명의 대상'으로 두드러지게 부각시키는 것이다(3). 앞의 대화에서 '은/는'은 침대방의 여러 속성 중에서(예를 들면, 경관, 크기, 가격 등) '가격'을 부각시키는 역할을 한다.

종결형 서술격 조사, '이오', '요'는 평서문에는 무언을 단정하여 알려주는 뜻으로, 의문문에서는 단순히 묻는 뜻으로 나타낼 때 쓰인다(1). '이오', '요'는 '(이)라니', '은/는'처럼 특별한 의미를 가지고 있지 않으므로 생략을 해도 화자의 의도는 변하지 않는다.

- (14) 시스템 : 온돌방의 가격은 4만 5천원입니다.
- (15) 사용자 : 4만 5천원이오?

위의 대화에서 사용자의 마지막 발화는 확인 요구(Seek-Confirm)의 담화 행위를 나타낸다. 조사 '이오'를 생략해서 "4만 5천원?"이라고 발화해도 화자의 확인 요구 담화 행위는 변하지 않으며 단지, 존칭의 뜻이 없어졌을 뿐이다.

본 논문에서는 한국어의 특성을 활용하여 생략 어구에 나타난 조사를 표층화해에 표현함으로써 생략 어구로부터 의도파악을 도와주고자 한다. 생략 어구를 크게 평서형, 의문형에 따라 Surface-Say-Noun(_agent1, _agent2, _noun), Surface-Noun-Question(_agent1, _agent2, _noun)으로 구분한다. 전자는 _agent1이 _agent2에게 체언 _noun을 평서형으로 말하는 것이고, 후자는 _agent1이 _agent2에게 체언 _noun을 의문형으로 질문하는 것이다. 의문형의 생략 어구는 특정 조사가 포함된 경우 3가지를 포함해서 생략 어구를 총 4가지의 표층화행으로 분류하여 표현한다.

• Surface-Say-Noun(_agent1, _agent2, _noun)

종결형 서술격 조사 '이오', '요'가 생략되거나 포함된 평서형의 생략 어구를 표현한다. "홍길동이오.", "온돌방요.", "더블룸."과 같은 생략 어구가 이에 해당된다.

• Surface-Noun-Question(_agent1, _agent2, _noun)

종결형 서술격 조사 '이오', '요'가 생략되거나 포함된 의문형의 생략 어구를 표현한다. "4만 5천원이오?", "홍길동?", "오일이오?"와 같은 생략 어구가 이에 해당된다.

• Surface-Noun-QuestionRN(_agent1, _agent2, _noun)

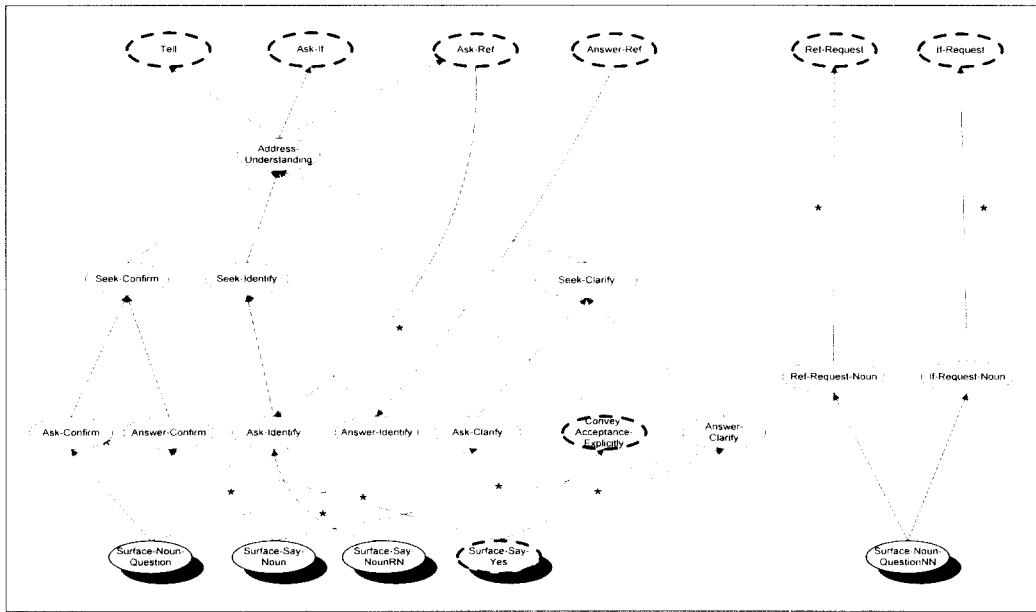
종결형 서술격 조사 '(이)라니'가 포함된 의문형의 생략 어구를 표현한다. "대기리스트라니요?", "더블룸이라니요?"와 같은 생략 어구가 이에 해당된다.

• Surface-Noun-QuestionNN(_agent1, _agent2, _noun)

보조사 '은/는'이 포함된 의문형의 생략 어구를 표현한다. "온돌방은요?", "가격은요?", "침대방은?"과 같은 생략 어구가 이에 해당된다.

2.2.2 추가된 담화 recipe

견고한 대화처리모델을 위해서는 담화 행위를 파악하는 것이 중요하다. 담화 행위를 인식하는 것은 다음 발화에 대한 예측을 가능하게 하며 결속성 있고 협조적인 응답 발화를 할 수 있게 한다. Lambert가 제안한 3단계 계획기반 대화이해모델에서 가장 역점을 둔 것은 화자의 담화 행위를 알아내는 것이며,



(그림 2) 생략 어구 처리를 위해 추가된 담화 행위들의 상관관계도

을 둔 것은 화자의 담화 행위를 알아내는 것이며, Lambert의 모델에는 26가지 담화 recipe가 정의되어 있다.

본 논문에서는 Lambert가 정의한 담화 recipe 외에 생략 어구의 처리를 위해서 16 가지의 담화 recipe를 추가하였다. 이는 Carberry(10)가 생략 처리를 위해 정의한 담화목적규칙들(discourse goal rules)중에 일부를 3단계 계획기반 대화이해모델에서 계획인식기법을 이용하여 추론할 수 있도록 recipe 형태로 변형시킨 것들과 그 외에 추가한 recipe, 그리고 생략 어구를 한국어의 특성을 고려하여 표현한 표층화행들이다.

화자가 생략 어구를 발화함으로써 어떠한 담화 목적을 가지는가를 알아내기 위해서 Carberry의 모델은, 담화 예측이 제시한 순위화된 담화 목적의 규칙을 생략 어구가 만족하는가를 살펴본다. 담화목적규칙은 생략 어구가 의문문인가 평서문인가 등의 생략 어구의 문장 유형과 생략 어구와 이전의 발화들로 구축된 혹은 영역에 내재된 맥락과의 연관 방식, 대화 참여자의 믿음(belief) 등의 조건을 기술한다[10]. 그러나, Carberry의 모델은 생략 어구의 담화 목적만 인식할 수 있으며 완전한 문장의 담화 목적은 인식하지 못한다. 또한 담화 목적을 인식하는 방법이 계획기반이 아니기 때문에 기존의 계획인식기법을 적용하

지 못하는 문제가 있다.

본 논문에서는 생략 어구를 표층화행으로 표현하고 생략 어구로부터 화자가 의도하는 담화 행위를 recipe로 정의하여 기존의 계획인식기법을 이용함으로써 별도의 생략 어구 처리 모듈이 필요없이 생략 어구로부터 화자의 의도를 인식하고자 한다. 본 논문에서는 2.2.1절에서 설명한 생략 어구의 표층화행과 생략 어구가 의도하는 여러 담화 행위를 담화 recipe로 정의하여 Lambert의 모델에 추가하였다. 그림 2는 추가된 담화 행위들의 상관관계를 도식적으로 표현한 것이다. 그림에서 각 노드는 담화 행위의 이름을 나타내며 행위에 필요한 인수는 공간문제로 생략했다. 아크는 담화 행위들의 관계를 나타내는 것으로 '*'는 specialization 관계를 나타내며 이를 제외한 모든 아크는 decomposition 관계를 나타낸다. 그림에서 점선으로 표시된 노드는 Lambert의 모델에 이미 정의되어 있는 담화 행위이고 나머지 노드들은 본 논문에서 추가된 담화 행위이며 그중에서 그림자 처리가 된 것은 표층화행이다.

3. 담화 행위의 결정 조건 - 생략 어구와 맥락과의 연관성

생략 어구로부터 화자의 의도를 인식하기 위해서 표층화행으로 표현된 생략 어구로부터 계획추론규칙

을 이용하여 화자가 의도하는 일련의 행위들을 추론한다. 그러나, 계획추론규칙에 의해 제시된 모든 행위를 발화자의 의도로 볼 수 있는 것은 아니다. 추론된 행위들의 조건 만족을 수행하여 계획추론규칙에 의해 추론된 행위들이 올바른가 확인해야 한다.

예를 들어 살펴보면, 그림 2에서 표층화행 Surface-Noun-Question은 "2인실이오?"와 같은 생략 어구를 나타낸다. 그림에서 Surface-Noun-Question은 확인 요구를 나타내는 Ask-Confirm 행위의 하위 행위일 뿐만 아니라, 명확성 요구를 나타내는 Ask-Clarify의 하위 행위이기도 하다. 즉, Surface-Noun-Question은 확인 요구, 또는 명확성 요구의 행위가 될 수 있다.

(16) 시스템 : 2인실의 방값은 4만 5천원입니다.

(17) 사용자 : 2인실이오? (확인 요구)

(18) 시스템 : 온돌방 값은 하루에 4만원입니다.

(19) 사용자 : 2인실이오? (명확성 요구)

위의 두 대화는 "2인실이오?"라는 똑같은 생략 어구가 맥락에 따라서 다른 행위를 의도하는 것을 보여준다. 첫번째 대화의 (17)번 발화에서는 '2인실'이라는 단어가 제대로 전달되었나를 확인 요구하는 행위를 나타내고, 두번째 대화의 (19)번 발화에서는 2인실 온돌방의 값이 4만원인가를 물어보는, 명확성을 요구하는 행위를 나타낸다. 이처럼 생략 어구가 의도하는 행위를 올바르게 인식하려면 생략 어구와 맥락과의 관계를 살펴보아야 한다.

따라서, 본 장에서는 생략 어구가 의도하는 담화 행위의 결정 조건으로 생략 어구와 맥락과의 연관성에 대해 살펴본다. 생략 어구와 이전 발화에 의해서 구성된, 혹은 영역에 내재된 맥락과 연관하는 방식에 따라서 생략 어구가 의도하는 행위가 달라지므로 이러한 연관성을 담화 행위의 결정 조건으로 제시한다. 먼저, 3.1절에서는 이전 발화들로 구축된 문맥과 영역에 내재된 맥락을 표현하는 방법으로 '객체'에 대해서 설명한다. 3.2절에서는 화자가 의도하는 담화 행위를 결정하기 위해서 생략된 문장 성분의 복원이 필요한 경우를 처리하기 위해서 초점화 이론을 이용하여 생략 어구의 의미를 복원하는 방법을 살펴본다. 마지막으로, 3.3절에서는 생략 어구가 의도할 수 있는 각각의 담화 행위의 결정 조건에 대해 살펴본다.

3.1 객 체

Allen은 바로 이전 문장의 문맥으로 그 문장의 구문 구조(syntactic structure), 의미 구조(semantic structure), 담화 요소(discourse entities) 등을 일컬어 지역 담화 문맥(local discourse context)이라 하였다(8). 여기서 담화 요소는 이전 발화에서 언급된 객체들로서 다음 발화의 대명사나 명사구에 의해 지시될 수 있는 것들을 말한다. 예를 들면, "5월 8일에 방을 예약하고 싶은데요."라는 발화에서 '5월 8일', '방'은 대명사나 명사구에 의해 언급될 수 있는 객체들이다. 객체는 일반적으로 발화에 포함된 각각의 명사구에 대해서 생성되나, 본 논문에서는 제한된 영역 대화만을 고려하므로 그 영역에서 필요한 객체만

<pre>object(room, room1234, {date_from, date_to, during, room_type, size}, [{date_from, 9일 }, {date_to, Date_to }, {during, During }, {room_type, 침대방 }, {size, 2인실 }, {price, Price }, {available, Available }]),</pre>	<ul style="list-style-type: none"> ▶ 의미 유형(semantic type) ▶ 객체의 고유 이름 ▶ 복사될 수 있는 자질들 ▶ 속성을 나타내는 자질들
--	---

(그림 3) '호텔 예약' 대화 영역에서 객체 '방'의 예

객체 ‘방’만을 생성한다.

객체는 객체의 의미 유형(semantic type)과 고유 이름, 복사될 수 있는 자질, 객체의 여러 속성을 나타내는 자질들로 이루어진다. 각각의 자질은 자질의 의미 유형과 값의 쌍으로 이루어진다. 그림 3은 호텔 예약 대화 영역에서 객체 ‘방’에 대한 예를 보여준다.

그림에서 객체의 의미 유형은 ‘room’이고 이 객체의 고유 이름은 ‘room1234’라는 것을 알 수 있다. 복사될 수 있는 자질은 같은 유형의 두 개의 객체가 비교될 때 복사될 수 있는 자질을 말한다. 예를 들면, 고객이 ‘5월 9일’에 ‘두 명’이 묵을 침대방의 값을 물어본 후, “온돌방은요?”와 같은 생략 어구로 온돌방의 값을 침대방과 비교하여 물어볼 때, 고객의 원하는 온돌방은 침대방과 마찬가지로 ‘5월 9일’에 ‘두 명’이 묵을 방이다.

객체의 여러 자질들은 대화 영역에서 객체를 바라보는 관점을 나타낸다. 즉, 호텔 예약 영역에서 객체 ‘방’에 관심이 있는 내용은 고객이 방에 처음 들어올 날짜(date_from), 방에서 나갈 날짜(date_to), 머무를 기간(during), 방의 유형(room_type), 방의 크기(room_size), 방의 가격(room_price), 방의 유무 여부(available) 등이다. 이러한 내용은 방에 대한 영역 지식을 나타낸다.

또한, 객체는 객체에 대해 언급한 대화의 맥락을 나타낸다. 일반적으로, 고객은 자신이 원하는 방을 하나의 발화에서 전부 언급하는 것은 아니며 여러 발화를 통해 원하는 방을 말한다. 예를 들면, 그림 3의 객체의 예에서 고객이 원하는 방은 ‘5월 9일’에 ‘두 명’이 묵을 ‘침대방’을 나타내는데, 이는 고객과 호텔 직원 사이에 고객이 원하는 ‘방’에 대해 나눈 대화의 맥락을 나타낸다.

3.2 초점화 이론을 통한 생략 어구의 의미 복원

생략 어구로부터 화자의 의도를 인식하기 위해서 생략 어구의 생략된 문장 성분의 복원이 필요한 경우가 있다. 보조사 ‘은/는’을 포함하여 새로운 정보를 얻고자 하는 생략 어구는 문제해결 수준, 영역 수준에서 새로운 행위를 추론하게 되는데, 이 경우는 생략 어구가 정확히 어떤 정보를 원하는 것인가를 알아야 하기 때문이다.

(20) 사용자 : 침대방의 가격은 얼마입니까?

(21) 시스템 : 하루에 4만 5천원입니다.

(22) 사용자 : 온돌방은요?

(23) 사용자 : 5월 9일에 침대방 좀 예약하고 싶습니다.

(24) 시스템 : 죄송합니다. 그날 침대방은 없습니다.

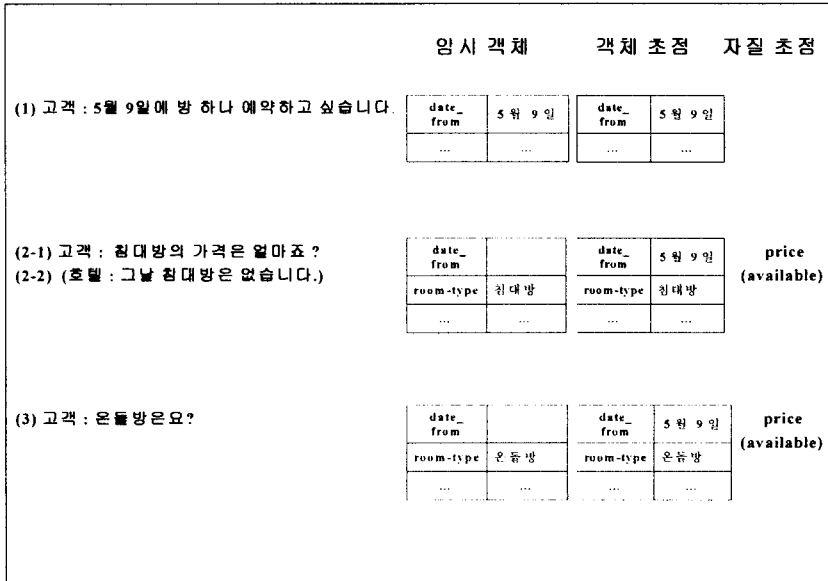
(25) 사용자 : 온돌방은요?

앞의 두 대화에서 고객의 마지막 발화 (22), (25)는 온돌방에 대한 어떤 값을 물어보는 발화이다. 전자는 온돌방의 가격을 물어보는 것이고, 후자는 온돌방이 있는지, 없는지의 여부를 물어보는 것이다. 두 경우는 모두 새로운 문제해결 행위와 영역 행위를 의도한다. 즉, 온돌방을 예약하는 영역 행위의 어떤 값을 채우는 문제해결 행위를 의도하며, 결국은 온돌방을 예약하려는 영역 행위를 나타낸다.

본 논문에서는 생략 어구의 의미 복원을 하기 위한 방법으로 초점화 이론(focusing theory)을 이용한다. 초점화(focusing)란 자신이 알고 있는 지식 중에 작은 부분에 주의 집중을 하는 것을 말한다[12]. 사람들의 지식은 아주 크기 때문에 상대방의 말을 이해하거나 자신의 말을 생성할 때, 그 많은 지식의 전부를 고려하는 것은 불가능하다. 그러므로, 상대방의 말을 이해하려면, 자신의 알고 있는 지식의 일부분에 주의 집중을 할 수 있어야 한다.

본 논문에서는 객체 초점과 자질 초점을 정의하여 생략 어구의 의미를 복원하고자 한다. ‘객체 초점’은 이제까지 생성된 여러 객체 중에서 현재 대화에서 초점이 되고 있는 객체를 말한다. ‘자질 초점’은 초점이 된 객체의 자질들 중에서 현재 초점이 되고 있는 자질을 말한다. 객체 초점과 자질 초점을 이용하여 생략 어구의 표면적 의미를 복원하는 방법은 현재 생략 어구에서 생략된 객체나 자질을 객체 초점과 자질 초점에서 구하는 것이다.

그림 4는 객체 초점과 자질 초점을 이용하여 생략 어구의 의미를 복원하는 예로서 각 발화에 의해서 생성되는 임시 객체와 객체 초점, 자질 초점을 보여준다. 임시 객체는 고객의 각 발화에 의해서 생성되는 객체로서 고객의 발화를 처리한 후 객체 초점이 될 수 있는 객체를 말한다. 그림에서는 편의상 대화에서 생략 어구의 의미 복원 과정의 설명에 불필요한 발화는 생략하였다. 그림 4에서 사용자의 (1)번 발화에 의해서 date_from이 ‘5월 9일’인 임시 객체가 생성된다. 시스템이 사용자가 방을 예약하려는 의도를 인식하면 임시 객체가 객체 초점이 된다. 다음 사용자의



(그림 4) 초점화 이론을 이용한 생략 복원의 예

하면 임시 객체가 객체 초점이 된다. 다음 사용자의 발화 (2-1)에서는 침대방의 가격을 물어본다. 이 발화에 의해 room_type이 '침대방'인 임시 객체가 생성된다. 임시 객체와 이전에 초점된 객체와 단일화(unification)를 통하여 임시 객체의 자질의 값, '침대방'이 초점된 객체에 복사된다. 발화에서 '가격'이라는 의미 표지(semantic marker)에 의해서 자질 초점이 price가 된다. 다음 (3)번 발화에 의해서 room_type의 값이 '온돌방'인 임시 객체가 생성된다. 초점된 객체와 임시 객체의 room_type의 값이 다르므로 단일화가 되지 않는다. 사용자가 새로운 객체를 추구하는 것으로 간주하고, 이전에 초점된 객체는 담화 스택(discourse stack)에 push되고 임시 객체가 객체 초점이 된다. 한편, 이전에 초점된 객체와 현재 초점된 객체는 'room'이라는 같은 유형의 객체이므로 이전에 초점된 객체의 자질들의 값이 자질 복사에 의해서 현재 초점된 객체에 복사되어 date_from이 '5월 9일'로 된다. 사용자의 (3)번 발화, "온돌방은요?"에서는 자질이 생략되었으므로 현재의 자질 초점 price에 의해 사용자가 '온돌방의 값'을 물어보는 것으로 생략 어구의 의미를 복원한다.

사용자가 "온돌방은요?"라고 발화하여 온돌방이 있는지 없는지의 여부를 묻는 경우에 생략 복원하는 방법은 가격을 묻는 경우를 처리하는 방법과 같다. 그림에서 시스템의 (2-2)발화에서 '침대방이 없다'에서

'없다'의 의미 분석 결과는 not(available)이 된다. 여기서, available은 객체 '방'의 자질이므로 자질 초점이 된다. (3)번 발화에서 사용자가 "온돌방은요?"라고 생략 어구를 발화하면 자질 초점을 이용하여 '온돌방의 유무'를 묻는 것이라고 생략 어구를 복원한다.

3.3 담화 행위의 결정 조건

본 절에서는 생략 어구가 의도하는 담화 행위의 결정 조건에 대해 살펴본다. 담화 행위의 결정 조건은 담화 행위를 정의한 담화 recipe의 제약조건이 되어 추론된 담화 행위의 조건 만족을 수행할 때 이러한 조건을 살펴보게 된다.

● 확인 요구(Ask-Confirm)

확인 요구 행위는 표층 화행 Surface-Say-Noun (_agent1, _agent2, _noun)에 의해 수행된다. 그러나, 화자의 생략 어구가 Surface-Say-Noun을 나타낸다고 해서 무조건 확인 요구 행위를 위한 것이라고 판단해서는 안된다. Surface-Say-Noun은 명확성 요구의 하위 행위도 되기 때문이다(그림 2 참조). Surface-Say-Noun이 확인 요구 행위가 되기 위한 조건은 _noun이 이전 발화의 부분이어야 한다. 이 조건은 그림 5의 Ask-Confirm recipe에 표현된다.

그림 5의 Ask-Confirm recipe의 제약조건(constraints)에서 component-of-prev-utterance(_noun)은

Action :	Ask-Confirm(_agent1, _agent2, _noun, _proposition) (<i>_agent1이 _agent2에게 _proposition의 _noun을 확인 요구한다.</i>)
Recipe-type :	Decomposition
Appl Cond :	want(_agent1, confirm(_agent1, _noun.believe(_agent2, _proposition)))
Constraints :	component-of-prev-utterance(_noun) in-focus-prop(_proposition)
Body :	Surface-Noun-Question(_agent1, _agent2, _noun)
Effects :	believe(_agent2, want(_agent1, Answer-Confirm(_agent2, _agent1, _noun, _proposition)))
Goals :	want(_agent2, Answer-Confirm(_agent2, _agent1, _noun, _proposition))

(그림 5) Ask-Confirm 담화 recipe

_noun이 이전 발화의 한 부분이 되어야 하는 것을 명시한다. 이처럼, 생략 어구와 맥락과의 연관성이 담화 행위의 결정 조건이 되며 recipe의 제약조건에서 기술된다.

● 정체 확인 요구(Ask-Identify)

정체 확인 요구 행위는 표층화행 Surface-Say-NounRN(_agent1, _agent2, _noun)이나 담화 행위 Ask-Ref에 의해서 수행된다. 결정 조건은 Ask-Confirm 행위와 마찬가지로 _noun이 이전 발화의 부분이어야 한다.

● 명확성 요구(Ask-Clarify)

명확성 요구 행위는 표층화행 Surface-Noun-Question(_agent1, _agent2, _noun)에 의해 수행된다. 이 표층화행이 명확성 요구 행위가 되기 위한 조건은 _noun이 현재 초점이 된 객체의 비워있는 한 자질의 값을 명시하는 것이다.

● 지시값 요구(Ref-Request-Noun), 예·아니오값 요구(If-Request-Noun)

표층화행 Surface-Noun-QuestionNN(_agent1, _agent2, _noun)은 지시값 요구 또는 예·아니오값 요구 행위가 될 수 있다. 예를 들면, "온돌방은요?"라는 생략 어구는 온돌방의 '가격' 뿐만 아니라 온돌방이 있는지 없는지의 '여부'도 물을 수 있는 것이다. 생략 어구의 표층화행이 Surface-Noun-QuestionNN을 나타낼 때, 이 생략 어구가 어떤 담화 행위를 의도하는가를 알려면 생략 어구의 의미 복원이 필요하다(3.2절의 생략 어구 의미 복원 참조). 생략 어구의 의미 복원은 객체 초점과 자질 초점을 통해서 이루어지

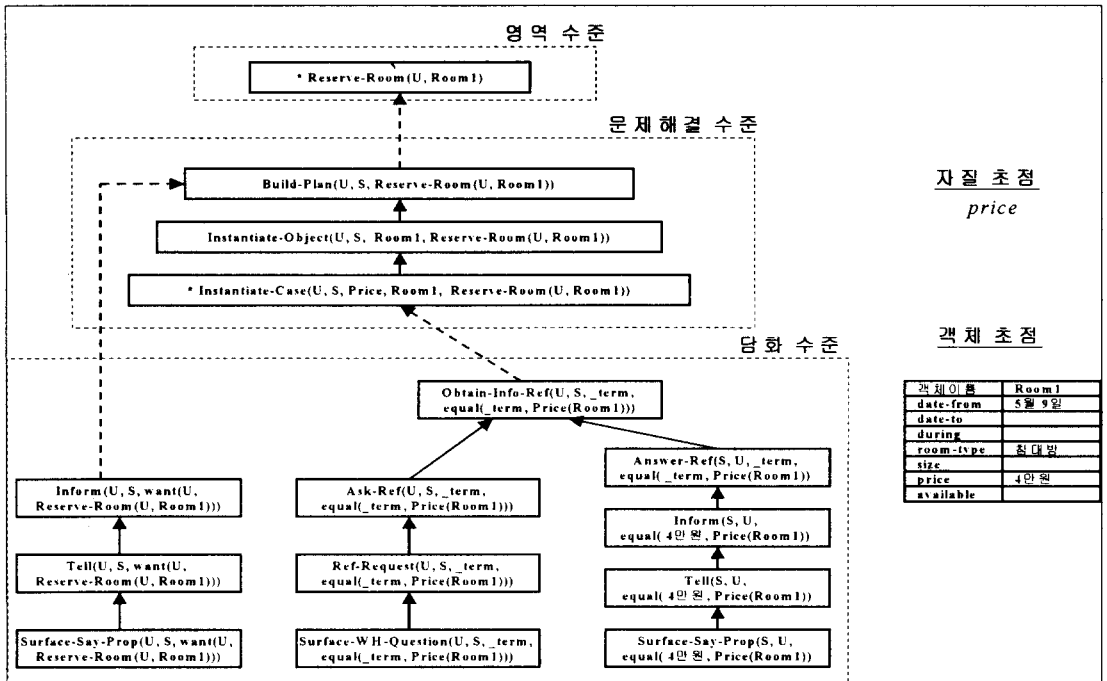
는데, 지시값 요구 혹은 예·아니오값 요구는 초점된 자질의 의미 유형(semantic type)을 통해서 결정한다. 예를 들면, 초점된 자질이 방의 유무(available)이면 이는 'proposition' 유형이므로 생략 어구가 예·아니오값 요구 행위를 나타낸다고 본다. 만약, 초점된 자질이 가격(price)이면 'proposition' 유형이 아니므로 지시값 요구 행위로 본다. 객체의 자질이 어떠한 의미 유형을 나타내는가는 'isa 계층구조(isa hierarchy)'에 의해 표현된다.

4. 대화 예제

- (26) 사용자: 5월 9일에 방 하나 예약하고 싶은데요.
 (27) 침대방 값이 얼마죠?
 (28) 시스템: 침대방의 가격은 4만원입니다.
 (29) 사용자: 2인실이요?

그림 6은 (26), (27), (28)번 발화를 처리한 후의 대화모델 및 객체 초점, 자질 초점을 보여준다. 사용자의 (26)번 발화의 표층화행은 Surface-Say-Prop(U, S, want(U, Reserve-Room(U, Room1)))이다. 이로부터 계획추론규칙을 적용하여 Inform 행위까지 추론한다¹⁾. Inform의 목적은 시스템이 사용자가 방을 예약하기를 원한다는 사실을 믿게하는 것이다. 이것은 방을 예약하는 영역 계획을 세우는 Build-Plan 문제해결 행위의 전제조건이므로 이로부터 Build-Plan 문제해결 행위를 추론한다. Build-Plan(U, S, Reserve-Room(U, Room1))의 목적은 결국 호텔방을 예약하는 영역 행위인 Reserve-Room(U, Room1)의 전제조건이므로 이로부터 영역 행위 Reserve-Room(U, Room1)을 추론한

1) 계획추론규칙과 Recipe 라이브러리를 이용하여 화자가 의도한 행위를 추론하는 자세한 방법은 [14]을 참고하기 바란다.



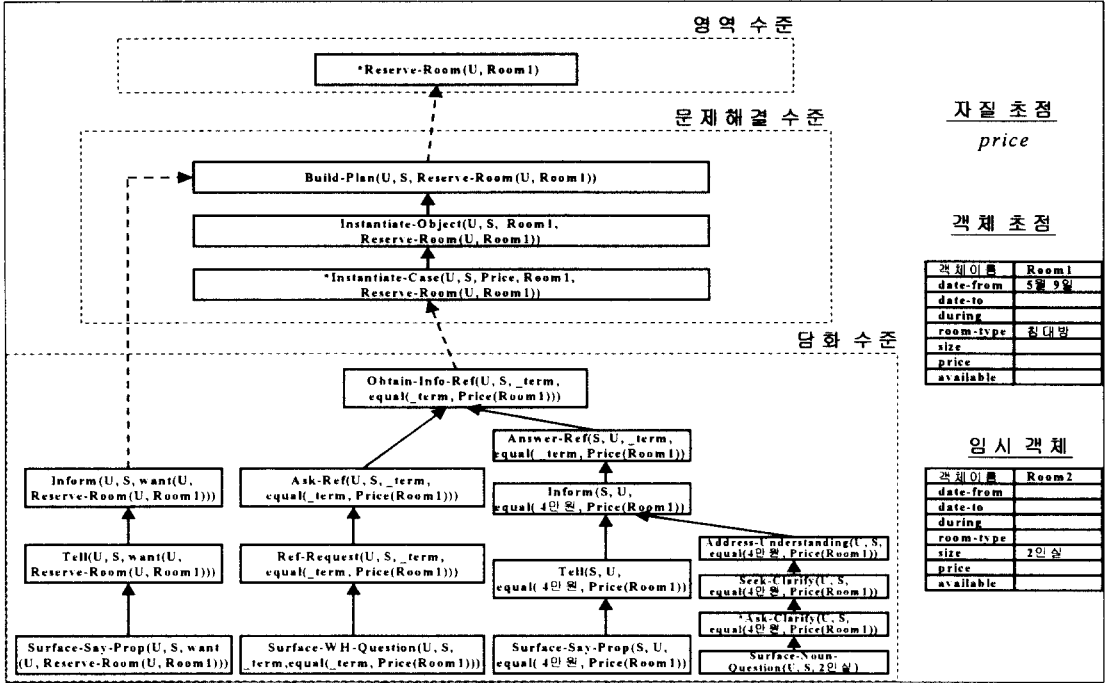
(그림 6) (26),(27),(28) 발화 처리 후의 대화모델 및 객체 초점, 자질 초점

다. 한편, 사용자의 (26)번 발화에 의해서 임시 객체 Room1이 생성되고 고객이 방을 예약하려는 의도를 인식함에 따라 임시 객체가 객체 초점이 된다.

(27)번 발화의 표층화행은 Surface-WH-Question(U, S, _term, equal(_term, Price(Room1)))이다. 이로부터 계획추론규칙을 이용하여 Obtain-Info-Ref 행위까지 추론한다. Obtain-Info-Ref(U, S, _term, equal(_term, Price(Room1)))의 목적은 knowref(U, _term, equal(_term, Price(Room1)))이고, 이는 문제해결 행위 Instantiate-Case의 전제조건이 되므로 이 문제해결 행위를 추론한다. 다시, Instantiate-Case(U, S, Price, Room1, Reserve-Room(U, Room1)) 행위로부터 초점화 휴리스틱과 계획추론규칙을 적용하여 이전 발화에 의해 인식된 Build-Plan 문제해결 행위와 가장 결속성 있는 관계를 찾아 연결한다. 이 발화에 의해 생성된 임시 객체는 그림 6에서 표현되어 있는데, 이전 객체 초점과 단일화(unification)가 되므로 새로운 객체를 생성하지 않고 단지 자질 room-type의 값 '침대방'만 객체 초점에 복사된다. 이 발화에서는 '가격'이라는 의미 표지(semantic marker)를 통해서 자질 초점이 price로 설정된다.

(28)번 시스템 발화의 표층화행은 Surface-Say-Prop(S, U, equal(4만원, Price(Room1)))이다. 이로부터 계획추론규칙과 초점화 휴리스틱을 적용하여 이전의 질문에 대답을 하는 Answer-ref 행위까지 추론하여 이전 발화와 연결한다. (28)번 발화는 시스템의 발화이므로 임시 객체를 생성하지 않으며 자질 초점은 변하지 않는다.

그림 7는 (29)번 발화를 처리한 후의 대화모델, 새로운 객체 초점, 임시 객체를 보여준다. (29)번 발화의 표층화행은 Surface-Noun-Question(U, S, 2인실)이며 이로부터 계획추론규칙을 적용하면 Ask-Confirm, Ask-Clarify 행위가 추론될 수 있다. 그러나, Ask-Confirm 담화 행위의 결정 조건은 단어 '2인실'이 이전 발화의 부분이 되어야 한다는 조건인데, 이전 발화에는 '2인실'이라는 단어가 없으므로, 이 행위는 거절된다. 다음에 Ask-Clarify 담화 행위의 결정 조건은 초점된 객체의 하나의 빈 자질을 명시하는 것인데, '2인실'은 객체 초점의 비어있는 자질 room_size의 값을 명시하므로 이 행위가 추론된다. 계속해서, Ask-Clarify 행위로부터 계획추론규칙과 초점화 휴리스틱을 이용하여 이전 발화의 한 행위인 Inform과 연결되는 Address-Understanding 행위



(그림 7) (29) 발화 처리 후의 대화모델 및 객체 초점, 자질 초점

까지 추론하여 이전 발화와 연결한다.

5. 결론 및 향후 과제

화자의 발화에 적절하고 협조적으로 대응하려면, 발화에 내재된 화자의 의도를 인식하는 것이 중요하다. 본 논문에서는 화자의 의도를 인식하는 모델로서, Lambert가 제안한 3단계 계획기반 대화이해모델을 이용하여 생략 어구로부터 화자의 의도를 인식할 수 있도록 하였다.

3단계 계획기반 대화이해모델은 완전한 문장으로부터 화자의 의도를 추론할 수 있지만, 생략 어구로부터는 화자의 의도를 인식하지 못한다. 본 논문에서는 생략 어구로부터 화자의 의도를 인식하기 위해 Lambert의 모델에 생략 어구가 의도하는 담화 행위를 담화 recipe로 정의하여 추가하였다. 기존의 계획 인식기법을 이용하려면 발화를 행위로 표현하는 것이 필요하므로 생략 어구를 표층화행으로 표현하였다. 생략 어구를 표층화행으로 표현할 때, 한국어에서 조사가 화자의 의도를 나타낸다는 특성을 이용하여 조사를 표층화행에 표현함으로써 화자의 의도 인식에 도움을 주었다. 또한, 객체와 초점화 이론을 이용하여, 화자가 생략 어구를 통하여 두 개의 계획을 동시

에 고려하는 발화를 모델링할 수 있다.

향후 과제는 다음과 같다. 우선, 본 논문에서는 처리할 생략 어구를 체언구로 제한하였는데, 다양한 형태의 생략 어구로부터 화자의 의도를 인식할 수 있도록 모델의 확장이 필요하다. 또한, 본 논문에서 고려한 조사의 다른 조사나 단서 단어 등과 의도와 의 관계를 좀 더 자세히 규명하여 이러한 언어적 단서들을 이용하여 화자의 의도를 잘 파악할 수 있도록 해야 하겠다.

그리고, 생략 어구가 의도하는 좀 더 다양한 담화 행위의 정의가 필요하다.

(30) 호텔 : 침대방의 가격은 4만 5천원입니다.

(31) 고객 : 4만 5천원이요?

위 대화의 고객의 생략 어구를 본 논문에서 제안한 모델에서는 확인 요구의 행위로 인식한다. 그러나, 맥락에 따라서 고객의 의도하는 행위는 놀라움의 표시 행위일 수 있다. 예를 들면, 고객이 일주일 전에 침대방의 값이 4만원이라고 믿었는데, 호텔측에서 현재 방값이 4만 5천원이라고 한다면 위의 대화에서 고객의 생략 어구 발화는 놀라움의 표현이거나 호텔이 말

한 사실을 못 믿겠다는 뜻을 나타내는 행위일 수도 있다. 고객이 그러한 의도를 가지고 있다면, 호텔 직원은 확인 요구 행위를 인식했을 때와는 다르게 응답 발화를 해야 된다. 실제 침대방의 가격이 4만 5천원 이라면 고객이 가지고 있는, 침대방의 가격이 4만원 이라는 믿음이 잘못되었다는 것을 알려주는 발화를 해야 할 것이다. 이와 같이 화자가 의도하는 행위에 따라서 응답이 달라지므로 이러한 다양한 담화 행위를 정의하여 다른 담화 행위와 구별하여 인식할 수 있어야 한다.

참고 문헌

- [1] 동아출판사(1994). 동아 새국어 사전. 이기문 감수, 동아출판사
- [2] 박정준(1996). 자연언어 대화 인터페이스를 위한 다영역 대화 모델의 설계. 석사학위논문, 서강대학교
- [3] 서정수(1991). 현대 한국어 문법 연구의 개관 - 제1권 주어/목적어/관련조사/기능용언, 한국문화사
- [4] 이석규(1996). '은/는'의 의미 연구, 우리말의 의미 연구, pp. 77-97. 도서출판 박이정
- [5] 이재원(1992). 대화계획 기법을 이용한 대화분석 연구. 석사학위논문, 한국과학기술원
- [6] 이현정(1996). 한국어 대화체 문장의 화행 분석. 석사학위논문, 서강대학교
- [7] James F. Allen(1983). Recognizing intentions from natural language utterances. In Michael Brady and Robert Berwick(ed.), *Computational Models of Discourse*, pp. 107-166. Cambridge, MA:MIT Press
- [8] James F. Allen(1992). *Natural Language Understanding*, 2nd ed. The Benjamin/Cummings Publishing Company, Inc.
- [9] Sandra Carberry(1989). A Pragmatics-Based Approach to Ellipsis Resolution. *Computational Linguistics*, 15(2), pp. 75-96
- [10] Sandra Carberry(1990). *Plan Recognition in Natural Language Dialogue*. Cambridge, MA:MIT Press
- [11] Robert E. Frederking(1988). *Intergrated Natural Language Dialogue*. Boston: Kluwer Academic
- [12] Barbara J. Grosz(1981). Focusing and description in natural language dialogues. In B. Webber, A. Joshi, and I. Sag(ed.), *Elements of Discourse Understanding*, pp. 85-105. Cambridge, England:Cambridge University Press
- [13] Barbara Grosz and Candace Sidner(1986). Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3), pp. 175-204
- [14] Lynn Lambert(1993). *Recognizing Complex Discourse Acts: A Tripartite Plan-Based Model of Dialogue*. Ph.D. thesis, University of Delaware
- [15] Diane Litman(1986). *Plan Recognition and Discourse Analysis : An Intergrated Approach for Understanding Dialogues*. Ph.D. thesis, Univeristy of Rochester
- [16] Diane Litman and James F. Allen(1987). A Plan Recognition Model for Subdialogues in Conversation. *Cognitive Science*, vol. 11, pp. 163-200
- [17] Martha E. Pollack(1990). Plans as Complex Mental Attitudes. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack(ed.), pp. 77-104. *Intentions in Communication*, Cambridge, MA:MIT press