

## 대화체 문장 처리에 관한 연구<sup>†</sup>

서강대학교 서정연\*  
한국과학기술원 이재원  
시스템공학연구소 박재득

### 1. 서 론

자연언어 처리의 궁극적인 목적은 컴퓨터가 사람의 일상의 언어를 이해하고, 생성하도록 하는 것이다. 특히, 대화는 유연성(flexibility), 명료성(succinctness), 표현력(expressive power) 면에서 뛰어난 장점을 가지고 있기 때문에 컴퓨터와 인간의 통신수단으로 유용하게 사용될 수 있다.

최근들어 자동통역기, 대화 에이전트 등에 대한 요구가 증가함에 따라 인간의 대화를 이해하고 이에 필요한 정보를 처리해 주는 기술이 많이 요구되고 있다.

대화체는 그동안 자연언어 처리 분야에서 많이 연구되어온 문어체 문장과는 다른 많은 특징들을 가지고 있으며, 이러한 특징들이 현재 대화체 처리의 주요 연구과제가 되고 있다. 대표적인 문제점들을 살펴보면 다음과 같다[17, 28].

#### (1) 불완전한(ill-formed) 문장의 분석

대화체에서는 흔히 간투사, 단어의 반복, 불완전한 종료 등을 포함하는 문장이 빈번히 발생한다. 그러므로, 대화체 문장을 처리하기 위해서는 완전한 문장뿐만 아니라 불완전한 문장을 분석하고 이해하기 위한 방법이 요구된다.

#### (2) 대용어 처리 및 생략된 문장의 분석

대용어(Anaphoric Reference) 처리란 대명

사나 지시형용사 또는 대동사와 같은 표현들이 실제 가리키는 것을 분석해 내는 작업을 말한다. 예를 들어 “그렇게 해주세요.”라는 문장의 실제 의미는 그 문장이 나타난 문맥을 분석하여야만 가능한 것이다. 이러한 문제는 대화체 문장 처리에서 매우 빈번히 나타나고 또 매우 중요한 문제이다. 이 대용어 처리 문제는 분석뿐 아니라 대화체 생성에서도 매우 중요한 문제로 알려져 있다.

대화체 문장에서는 생략이 빈번히 발생한다. 문맥적으로 분명하거나 상호 알고 있다고 여겨지는 부분은 일반적으로 생략된다. 특히, 한국어의 경우에는 주어와 동사가 함께 생략되거나 주어와 목적어가 동시에 생략되기도 한다. 그러므로, 생략된 정보를 복구할 수 있는 방법이 필수적으로 요구된다.

#### (3) 화자의 의도 분석

대화는 여러 발화들이 서로 유기적으로 관련되어 있으며, 따라서 동일한 대화체 문장일지라도 상황에 따라 서로 다른 의도를 나타내기도 한다. 그러므로, 문맥으로부터 화자의 의도를 정확하게 이해할 수 있는 방법이 요구된다.

본 논문에서는 위에서 언급한 대화체의 특징에 초점을 맞추어 이를 해결하기 위한 연구들에 대해 살펴 본다.

먼저, 대화체 문장을 분석하기 위한 연구에 대해 살펴본 다음, 3장에서는 대화를 모델링하는 방법에 대해 기술한다. 4장에서는 대화체 생성 방법에 대해 논하고, 5장에서 현재의 연구 진행 방향과 향후 연구 방향에 대해 언급한 뒤 6장에서 결론을 맺는다.

<sup>†</sup> 본 연구는 부분적으로 소프트웨어연구 “대화이해 및 생성을 위한 대화 인지모델 연구”의 결과입니다.

\*종신회원

## 2. 형태소, 구문, 의미 분석

대화체 문장은 문어체 문장과는 달리 항상 완전한 문장으로 이루어 지는 것은 아니다. 비문법적이거나 생략이 심한 문장도 허용될 수 있다. 대화체 문장에 대해 가지는 형태소 분석이나 구문 분석상의 어려움으로 다음과 같은 특징들을 들 수 있다.

- (1) 비문법적인 문장이 사용될 수 있다.
- (2) 문장 구성 성분의 생략이 심하다.
- (3) 같은 단어 혹은 문장이 반복되거나 말을 이끌어 가기 위한 간투사가 사용될 수 있다.
- (4) 대화가 중단되거나 취소됨으로써 문장이 중간에서 끝날 수 있다.
- (5) 문장의 처음과 끝을 찾기 어렵다.

한국어 대화체 문장을 분석하기 위한 연구는 그다지 많이 이루어지지 않는 않았다. [17]에서는 이러한 대화체 문장은 문법적으로 분석하기 매우 어렵다고 기술하고 있다. 그러므로, 구문 분석이나 의미분석 등이 독립되어 단계적으로 처리되는 문어체 문장 분석과는 다르게 형태소 분석, 구문 분석, 의미 분석, 때로는 그 이상의 화용 분석이 서로 밀접하게 연결되어 처리되어야 함을 제시하고 있다. 또한, 한국어 대화체 문장에서 조사 생략시 구문관계 해석에 대한 연구가 있었다[2, 5, 9]. 현재 많은 시스템들은 이러한 대화체 문장을 체계적으로 다루고 있지는 못하며, 대화체 문장의 도메인이 한정적임을 고려하여 의미 문법(semantic grammar) 등을 이용하여 부분적으로 해결하고 있을 뿐이다[4].

## 3. 대화의 이해

대화는 문 단위를 벗어나 담화상황을 배경으로 하는 발화를 기본 단위로 한다. 그러므로, 대화를 올바르게 이해하기 위해서는 문맥(context)과 관련하여 대화를 분석할 수 있는 방법이 요구된다. 본 장에서는 먼저 대화를 모델링하는 계산적인 모형들으로써 계획에 기반한 대화 모델, 대화 전이망을 이용한 모델, 통계적인 모델에 대해 살펴본다. 다음 절에서는 담화상황

과 관련한 중요한 요소들로 화행(speech act)과 초점(focus), 사용자 모델(user model)에 대해 기술한다.

### 3.1 대화 이해 모델

#### 3.1.1 계획에 기반한 대화 모델

대화를 이해하기 위한 연구로는 계획에 기반한 대화 이해 모델이 대표적이다. 대화를 계획(plan)을 이용하여 모델링한 것은 Grosz[25]의 목적지향적인 대화처리 모델에 관한 연구와 Cohen and Perrault[24]와 Allen and Perrault[21]의 화행 및 화자의 의도분석에 관한 연구가 대표적인 연구로 시작되었다. 일반적인 인공지능 분야에서는 물리적인 행위가 물리적인 세계의 변화를 목적으로 계획되어진다. 이와 마찬가지로 대화에서는 언어적인 행위(linguistic action)가 대화 참여자들의 궁극적인 목적을 이루기 위해서 계획되어 질 수 있다고 보았다. 즉, 대화에서의 각각의 발화는 대화의 궁극적인 목적을 위해 수행되는 하나의 행위에 해당되는 것이다. 예를 들어, 화자는 청자에게 어떤 행위를 하도록 요구(request)할 수 있다; 전제조건(prerequisite)으로 화자는 요구한 행위가 수행되기를 원하고 있어야 하며; 행위의 결과(effect)로 청자는 그 행위를 원하게 된다.

HEADER : REQUEST (speaker, hearer, action) PREREQ : WANT (speaker, action) DECOMP : S-REQUEST (speaker, hearer, action) EFFECT : WANT (hearer, action)
---

그림 1 요구 화행에 대한 계획 표현

Litman and Allen[30]은 다양한 형태의 부대화(subdialogue)를 처리하기 위하여 대화의 영역에 관한 지식에 해당하는 영역계획(dominan plan) 외에 대화의 일반적인 현상에 대한 지식에 해당하는 담화계획(discourse plan)을 도입함으로써 다양한 대화의 처리를 시도하여 매우 간단한 몇개의 대화를 처리할 수 있는 시스템을 처음으로 구현하였다. 이후에도 많은 연구들이 보다 다양한 대화를 처리하기 위하여 계획에 기반한 기본 모델을 지속적으로 확장시켜 왔다. Lambert and Caberry[33]은 하나

이상의 발화가 특정 의도를 가지는 복합의도 (complex intention)을 분석하기 위하여 3단계로 이루어진 계획과 이에 관련한 추론 휴리스틱을 제안하기도 하였다. 이외에 혼합주도형 대화(Mixed Initiative Dialogue)를 모델링하기 위한 연구 등이 시도되었다[18, 22]. 국내에서도 비슷한 시기에 계획 기법을 이용하여 한국어 대화처리를 위한 연구가 수행되었다 [11, 12, 13, 14, 19].

계획에 기반한 대화 이해 모델들은 인지적인 관점에서 대화의 과정을 모델링하고 있다. 그러나, 이러한 방법론은 많은 영역지식과 담화 지식을 필요로 한다. 특히, 계획지식은 쉽게 구축될 수 있는 것이 아니며, 상당한 전문가의 노력을 필요로 한다. 또한 지식에 기반하여 모든 대화를 완전하게 이해하기 위해서는 불완전한 지식(Incomplete Knowledge) 문제와 과도한 지식(Too Much Knowledge) 문제와 같은 전반적인 인공지능 연구가 가지고 있는 가장 근본적인 문제들을 해결해야만 한다. 이에 최근 들어 새로운 방법들이 시도되고 있다.

### 3.1.2 대화 전이망을 이용한 대화 모델

제한된 영역의 대화에 대해서 효율적으로 사용될 수 있는 대화 모델 방법으로 대화 전이망을 이용한 방법이 있다. 대화 전이망은 대상 도메인에서 발생할 수 있는 가능한 대화의 순서를 기술하게 된다. 이러한 대화 전이망을 이용하여 대화의 올바른 진행을 검사하거나 다음 발화에 대한 예측을 하게 된다.

이와 유사하지만 더 잘 정형화되어 있는 기법으로 재귀적 대화 흐름도(RDTN: Recursive Dialogue Transition Networks)를 이용한 방법이 있다[15, 29]. 이 모델은 재귀적 전이망문법(Recursive Transition Networks Grammar)과 비슷한 형식으로 기술되어 있으며, 스택을 이용하여 구현할 수 있다. 이 모델은 부대화에 의한 대화의 계층적인 담화구조를 효율적으로 모델링하고자 하였다.

이와 같이 대화 전이망을 이용하는 경우 대화의 상황이 미리 정의되어 있기 때문에 상황 인식의 부담이 적어 효율적으로 대화를 분석할 수 있다는 장점이 있다. 그러나 대화의 흐름이

고정되어 있기 때문에 실제 대화에서 발생할 수 있는 다양한 상황에 대해 결코하게 대처할 수 없다는 단점이 있다.

### 3.1.3 통계적인 대화 모델

지식에 기반한 대화 이해 모델의 단점을 극복하기 위한 통계적인 방법론이 Nagata와 Morimoto에 의해 시도되었다[35]. 대화에서 다음 발화의 화행(speech act)을 예측하기 위하여 통계적인 방법론을 적용하였다. 이 모델에서는 인간의 대화를 유한 심볼들의 집합으로부터 선택된 심볼들을 출력하는 일종의 정보원으로 보았다. 이때 모든 가능한 인간의 발화를 나열할 수는 없으므로 화행을 이용하여 심볼들의 집합을 근사하였다. 그러면, 대화는 다음과 같이  $n$ 개의 화행으로 근사될 수 있다.

$$D = s_1, s_2, \dots, s_n$$

제안된 모델에서는 트라이그램 확률을 이용하여 대화모델의 확률을 추정하였다. 이전 두 발화의 화행으로부터 다음에 특정 화행으로 전이할 확률을 이용하여 다음 발화의 화행을 예측하였다. 즉, 화행  $s_{i-2}, s_{i-1}$ 을 가진 문장이 발화되었을 때 다음 문장의 화행이  $s_i$ 일 확률을 구하는 것이다. 트라이그램 확률은 다음 식으로부터 구해질 수 있다.

$$P(s_i | s_{i-2}, s_{i-1}) = \frac{C(s_{i-2}, s_{i-1}, s_i)}{C(s_{i-2}, s_{i-1})}$$

이러한 모델은 자동통역 시스템인 VERBMO-BIL에서 다음 발화의 화행을 예측하는데 적용되기도 하였다[20]. 또한, 한국어 대화의 구문 패턴과 화행의 트라이그램을 이용하여 발화의 화행과 담화구조를 통계적으로 분석하고자 하는 연구가 시도되었다[29]. 구문패턴은 특정 언어 고유의 발화 패턴을 반영하기 위한 것으로 문장확률(sentential probability)을 계산하기 위해 사용되며, 화행의 트라이그램은 문맥 확률(contextual probability)로 사용된다. 이때 화행의 트라이그램은 담화의 계층적인 구조를 반영하도록 하였다.

통계적인 방법은 학습 데이터인 코퍼스로부터

터 필요한 지식을 자동으로 학습할 수 있다는 장점이 있다. 즉, 어려운 지식 구축을 요구하지 않는다는 점이다. 또한, 모호성을 처리할 수 있으며, 다른 모델과 쉽게 통합하여 사용할 수 있다는 장점이 있다. 그러나, 앞서 언급된 모델들에서 알 수 있는 바와 같이 현재의 통계적인 대화 분석에 대한 연구는 아직 초기 단계에 머무르고 있다. 따라서, 대화를 완전하게 처리할 수 있는 것은 아니며, 화행분석과 같이 부분적으로 적용되고 있는 상황이다. 이에 대한 연구는 앞으로도 계속 지속되어야 할 것이다.

## 3.2 대화 분석

### 3.2.1 화행(Speech Act)

화행은 화자가 청자에게 발화를 통하여 의도적으로 알리고 얻으려는 행위를 의미한다[16]. 화행은 발화의 사용의미, 또는 화자의 의도 등을 밝혀내는데 유용한 도구가 되고 있다. 또한, 생략이나 대응어를 처리하는데 중요한 정보가 되는 초점(focus)과도 상호 연관되어 있는 매우 중요한 담화적 요소이다.

화행에는 직접화행(direct speech act)과 간접화행(indirect speech act)이 있다. 직접화행이 문장의 표면에 나타나는 화행인 반면 간접화행은 표층적인 표현이 아니라 간접적인 표현으로 어떤 화행을 나타내는 것이다. 예를 들어 “비행기 출발 시간을 아십니까?”의 경우 표층적으로는 출발 시간을 아는지 모르는지를 묻는 질문이다. 그러나, 이 발화는 상황에 따라서는 비행기 출발 시간을 알려 달라는 간접적인 표현으로 이해될 수도 있다. 그러므로, 발화에 내포된 화자의 의도를 상황에 따라 정확히 이해하는 것이 필요하다.

화행에 대해서는 오래전부터 많은 연구가 이루어져 왔지만 어떤 표준화된 분류가 존재하지는 않으며, 흔히 주관적인 판단이나 문헌에 나타나는 분류에 크게 의존하고 있다. 화행에 대한 이론을 처음으로 정립한 Searle은 representative, directive, commissive, expressive, declarative의 5개 부류를 사용하였으며[32], Nagata는 영역에 독립적인 15개의 화행을 제시하였다[35]. Alexandersson은 회의 날짜를 결정하는 과정에서 발생하는 17개의 화행을 사

용하고 있다[20]. 한국어 대화를 대상으로 한 연구로는[15, 29] 등이 있다.

### 3.2.2 초점(Focus)

초점은 화자가 발화에서 주의를 기울이고 있는 담화객체를 의미한다. 이러한 초점은 음성적으로는 강세로 나타나며, 통사적으로는 분열문의 서술부로, 의문사 의문문에서는 의문사와 대치되는 성분들이다[16]. 예를 들어 “철수는 서울에 갔다”에서 밑줄 부분이 강세를 받고 발음될 경우, “철수는 어디에 갔다”는 것을 전제로 하여 ‘어딘가’는 서울을 나타내고 있다고 해석된다. 이와 같이 발화에 의해 새로운 정보를 기술하는 부분이 있는데 이를 초점이라고 한다.

초점은 발화에 나타나는 대응어구나 생략과 밀접한 관련을 가진다. Grosz and Sidner[26]는 대화에서 나타나는 대응어구를 이해하기 위한 방법으로 초점의 유지를 제안하였다. McKewin은 주제를 선택할 때 화자가 발화를 조직화하는 방법을 설명하기 위해 초점규칙을 제안하였다[34]. 이들이 제시한 초점 전이 규칙은 대화의 초점이 이전의 화제와 연관성이 없는 화제보다는 연관성이 있는 화제, 즉 현재 토론되고 있는 내용의 부수적인 화제로 옮겨가는 성질을 가지고 있다는 것이다. 이상에서 알 수 있는 바와 같이 대화를 이해하고 생성하는데 있어서 대화의 초점을 알아내고 이를 유지하는 것이 매우 중요하다.

### 3.2.3 사용자 모델(User Model)

질의 응답 시스템이나 에이전트 시스템 등에서 사용자에게 대한 모델을 이용하여 시스템이 사용자에게 대한 적응성을 지니도록 하여 대화의 질을 높이고자 하는 많은 연구가 수행되어 왔다[36]. 사용자 모델은 사용자의 다양한 측면에 대한 가정을 포함하는 지식을 의미한다. 사용자에게 대한 가정은 사용자의 목표(goal)와 계획(plan), 믿음(belief), 또는 사용자가 영역에 대해 가지고 있을 수 있는 잘못된 개념에 대한 정보 등이 될 수 있다. 이러한 사용자에게 대한 모델은 대화를 통해 얻어지는 다양한 정보를 통해 구축되어 진다.

이러한 사용자 모델은 시스템이 사용자의 발

화를 이해하고 사용자의 잘못된 개념을 수정하고 사용자에게 맞는 발화를 생성하는데 이용된다. 즉, 대화를 통해 얻어지는 사용자에게 대한 정보를 이용하여 사용자에게 적합한 형태의 협조적인 대화를 유도하기 위해 사용된다. 사용자 모델은 또한 협조적인 대화뿐만 아니라 지능적인 에이전트로써 사용자의 의도를 파악하거나 시스템의 발화가 사용자에게 어떠한 영향을 미치는가를 분석하기 위해서 사용되기도 한다. 그러므로, 지능적인 대화가 가능하도록 하기 위해서는 사용자 모델링이 필수적인 도구라 할 수 있다.

#### 4. 대화체 문장 생성

한국어 생성 시스템은 지금까지 주로 기계번역 시스템의 한 부분으로써 개발되어 왔다. 대표적인 시스템으로 서울대의 KSHALT와 한국과학기술원의 MATES-EK 등이 있으며, 이외에 한국어 생성에 대한 다수의 논문들이 발표되기도 하였다[7, 8, 10]. 그러나, 이들 연구는 일반적인 한국어 생성에 관한 것이며, 대화체 문장을 대상으로 한 것은 아니었다.

대화체 문장은 문어체 문장과는 다른 특징을 가지며, 대화체 생성을 위해서는 이러한 특징들을 고려하여야 한다. 대화는 효율성을 위해 결속성(coherence)과 정보성(informativity)을 가진다고 한다.

결속성이란 대화체 문장은 현재 대화의 화제와 관련성이 있는 것이어야 함을 의미하며, 정보성은 필요 이상의 정보는 제공하지 않고 필요한 정보는 빠뜨리지 않고 제공해야 한다는 것이다. 이러한 맥락에서 생략과 대응어 현상이 나타나게 된다. 그러므로, 대화체 생성에서는 이러한 현상들이 적절히 표현될 수 있어야 한다.

대화체 생성에 대한 연구는 매우 빈약한 상황이다. 이에 대한 연구로는 초점의 유지와 사용자 모델링을 이용하여 효율적인 대화를 생성하는 방법이 연구된 바 있으며[1], 문맥정보를 이용하여 생략과 대응어를 생성하는 방법이 제시되기도 하였다[23]. 완전한 문장이 대화체에서는 오히려 부자연스러운 문장이 될 수 있으

므로 이러한 연구들이 대화체 생성에서는 매우 중요하다.

#### 5. 향후 연구

최근에 다중모드 대화(multimodal dialogue)에 대한 연구가 시작되었다[3]. 단지 문자화된 언어로써의 대화뿐만 아니라, 화상 정보, 터치스크린(touch screen) 등 다양한 도구를 이용한 다중 모드 대화에 대한 필요성이 대두되고 있기 때문이다. 이러한 다중모드 대화를 처리하기 위해서는 화상에 나타나는 상대방의 제스처, 표정 등으로부터 상대방의 의도를 파악할 수 있어야 하며, 터치 스크린 등과 같은 다른 도구와 언어적인 대화 상의 대응어구 등과의 관계를 분석할 수 있어야 한다.

또한, 한편으로 학제간의 연구 필요성이 대두되고 있다. 현재 국내에서 추진되고 있는 소프트웨어에 관한 연구사업에서 언어학, 인지심리학, 인공지능 등의 여러분야 학자들이 연계하여 학제간 연구를 진행하고 있다. 이러한 학제간 연구를 통해 인간의 대화 현상에 대한 기본적인 이론을 수립하고 보다 일반성이 있는 확장된 대화처리 모델 개발을 목표로 연구 중에 있다[6].

#### 6. 결 론

이상에서는 대화체 문장의 특징을 살펴보고 이와 관련한 연구들과 진행 방향 등에 대해 살펴 보았다. 이러한 대화 분석에 대한 관심은 보다 더 편리한 컴퓨터의 개발을 위해 그 필요성에 의해 점차 커지고 있는 상황이다. 그러나, 한국어 대화체 문장의 처리에 대한 연구는 아직 초보적인 단계에 있다고 볼 수 있다. 그러므로, 앞으로는 좀 더 많은 연구가 지속적이고 체계적으로 진행되어야 할 것이다.

#### 참고문헌

- [1] 김진아, "초점의 유지와 사용자 모델링을 통한 효율적인 대화의 유도", 석사학위논문, 한국과학기술원, 1996.

- [2] 김창현, 서정연, 김길창, “한국어 대화체 문장에서 조사 생략시 구문관계의 해석”, 제2회 인공지능, 신경망 및 퍼지 시스템 종합 학술 논문집, pp. 281-290, 1992.
- [3] 김학수, 서정연, “홈쇼핑 영역에서의 다중모드 사용자 인터페이스 시스템에 대한 연구”, 한국인지과학회 춘계학술대회 논문집, pp. 74-80, 1997.
- [4] 박정준, 서정연, “다영역 대화모델”, 한국정보과학회 HCI '97 학술대회 논문집, Vol.6, No.1, pp. 64-69, 1997.
- [5] 서정연, 이종혁, 김태석, “자동통역 전화 개발을 위한 대화체 기계번역에 관한 연구”, 최종 연구 보고서, 한국통신 장기기초연구, 1996.
- [6] 서정연, 최재용, 이현호, 홍민표, “대화이해 및 생성을 위한 대화 인지모형 연구”, 한국인지과학회 춘계학술대회 논문집, pp. 24-31, 1997.
- [7] 이강천, 이상호, 서정연, “의미중심어 주도 방식에 기반한 한국어 생성 시스템”, '96 봄 정보과학회 학술대회 논문집, Vol. 23, No.1, pp. 949-952, 1996.
- [8] 이강천, 서정연, “양상에 따른 자연스러운 주격조사의 선정”, 제8회 한글 및 한국어 정보처리 학술대회 논문집, pp. 173-176, 1996.
- [9] 이공주, 권철중, 서정연, 김길창, “문법의 범위를 벗어나는 영어 문장들을 위한 견고한 구문 분석기”, 한국정보과학회 인공지능 연구회 춘계 학술발표 논문집, pp. 119-129, 1994.
- [10] 이상호, 서정연, “의미 중심어 주도 방식을 이용한 한국어 생성 시스템의 구현”, 제6회 한글 및 한국어 정보처리 학술발표 논문집, pp. 434-438, 1994.
- [11] 이재원, 서정연, 김길창, “계획에 기초한 대화 이해 시스템에 관한 연구”, 제2회 인공지능, 신경망 및 퍼지 시스템 종합 학술 논문집, pp. 281-290, 1992.
- [12] 이재원, 김재훈, 서정연, “대화모델을 이용한 대화체 기계번역”, 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 104-107, 1993.
- [13] 이재원, 서정연, 김길창, “자동통역에서의 대화체 기계번역을 위한 문맥의 구축과 이용”, 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 80-84, 1994.
- [14] 이재원, 서정연, 김길창, “인간의 대화를 이해하기 위한 대화계획의 인식 및 추적”, 한국정보과학회 HCI '95 학술대회 논문집, Vol.4, No.1, pp. 55-61, 1995.
- [15] 이현정, “한국어 대화체 문장의 화행 분석”, 석사학위논문, 서강대학교, 1997.
- [16] 장석진, “한국어 문법 -NLP를 위한 HPSG/K-”, CAIR-TR-92-33, 1992.
- [17] 전길남, “한국어 특질 및 대화체 기계번역에 관한 연구”, 최종연구보고서, 한국통신 장기기초연구, 1991.
- [18] 조영환, 서정연, 김길창, “혼합주도형 대화 시스템을 위한 대화 에이전트”, 한국정보과학회 HCI '95 학술대회 논문집, Vol.4, No.1, pp. 14-23, 1995.
- [19] 조정미, 서정연, 김길창, “한·영 대화체 기계번역 시스템”, 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 65-70, 1994.
- [20] Alexandersson, J., Maier, E., and Reithinger, N., “A Robust and Efficient Three-Layered Dialogue Component for a Speech-to-Speech Translation System”, *Proc. of the 7th European Association for Computational Linguistics*, pp. 188-193, 1994.
- [21] Allen, J. and Perrault, C.R., “Analyzing Intention in utterances”. *Artificial Intelligence*, Vol.15, pp. 143-178, 1980.
- [22] Cho, Y.H. and Kim, G.C., “MIDAS : A Mixed-Initiative Dialogue System”, *Proc. of NLPRS*, pp. 640-645, 1995.
- [23] Cho, Y.H., Kim, C.H., and Kim, G.C., “Decisions on Ellipsis and Anaphora Generation Based on Context Intimacy”, *Literary and Linguistic Computing*, Vol.12, No.2, 1997.
- [24] Cohen, P.R. and Perrault, C.R., “Ele-

ments of a Plan-Based Theory of Speech Acts”, *Cognitive Science*, Vol.3, pp. 177-212, 1979.

[25] Grosz, B., “The representation and use of focus in a system for understanding dialogues” *Proceedings of IJCAI*, pp. 67-76, 1977.

[26] Grosz, B., and Sidner, C., “Attention, Intention, and the Structure of Discourse” *Computational Linguistics*, Vol. 12, No.3, pp. 175-204, 1986.

[27] Kim, C.H., Kim, J.H., Seo, J.Y., and Kim, G.C., “Right-to-Left Chart Parsing with Headable Paths for Korean Dependency Grammar” *Computer Processing of Chinese and Oriental Language*, Vol.8 Sup., pp. 105-118, Dec.1994.

[28] Kitano, H., “Speech-To-Speech Translation: A Massively Parallel Memory Based Approach”, *Kluwer Academic Publishers*, 1994.

[29] Lee, J.W., Kim, G.C., and Seo, J.Y., “A Dialogue Analysis Model with Statistical Speech Act Processing for Dialogue Machine Translation”, *Spoken Language Translation Workshop in (E)ACL'97*, pp. 10-15, Jul.1997.

[30] Litman, D.J. and Allen, J.F. “A Plan Recognition Model for Subdialogues in Conversations”, *Cognitive Science*, Vol. 11, pp. 163-200, 1987.

[31] Seo, J.Y., “Text-Driven Multiple-Path Discourse Processing for Descriptive Texts”, *Electrical Engineering and Information Science*, Vol.1, No.2, pp. 1-8, Jun.1996.

[32] Searle, J.R., “A Taxonomy of Speech Acts”, *Expression and Meaning*, Cambridge U. Press, pp. 1-29, 1979.

[33] Lambert, L. and Caberry, S., “A Tripartite Plan-Based Model of Dialogue”. *Proc. of the 29th Association for Computational Linguistics*, pp. 47-54, 1991.

[34] McKeown, K.R., *Text Generation*, Cambridge U. Press, 1985.

[35] Nagata, M. and Morimoto, T., “First Steps towards Statistical Modeling of Dialogue to Predict the Speech Act Type of the Next Utterance”, *Speech Communication*, Vol.15, pp. 193-203, 1994.

[36] Wahlster, W. and Kobsa, A., *User Models in Dialogue Systems*, Springer-Verlag Berlin Heidelberg, 1989.

서 정 연



1981 서강대학교 수학과 학사  
 1985 미국 Univ. of Texas, Austin 전산학과 석사  
 1990 미국 Univ. of Texas, Austin 전산학과 박사  
 1990~1991 미국 Texas Austin, UniSQL Inc. Senior Researcher  
 1991 한국과학기술원 인공지능 연구센터 선임연구원  
 1991~1995 한국과학기술원 전산학과 조교수

1995~1996 서강대학교 전산학과 조교수  
 1996~현재 서강대학교 전산학과 부교수  
 관심분야: 한국어 정보처리, 자연언어처리, 기계번역, 대화처리

이 재 원



1991 서울시립대학교 전자계산학 학사  
 1993 한국과학기술원 전산학과 석사  
 1993~현재 한국과학기술원 전산학과 박사과정

박 재 득



1983 서울대학교 계산통계학과 학사  
 1985 한국과학기술원 전산학과 석사  
 1989~1995 삼성전자 멀티미디어 연구소 선임연구원  
 1994 한국과학기술원 전산학과 박사  
 1995~현재 시스템공학연구소 자연어 정보처리 연구부 실장

관심분야: 한국어정보처리, 지식표현 및 추론, 기계학습, 인 지 언어학 등