

통계에 기반한 다영역 대화형 도우미 시스템의 설계

정형일[†] 김동현[†] 장효준[†] 김학수^{††} 서정연^{†††}
서강대학교 서강대학교 서강대학교 강원대학교 서강대학교 컴퓨터학과
컴퓨터학과 컴퓨터학과 컴퓨터학과 컴퓨터정보통신공학과 바이오융합 협동과정
hijeong@sogang.ac.kr kdh2007@sogang.ac.kr hyojun77@sogang.ac.kr nlpdkim@kangwon.ac.kr seojy@sogang.ac.kr

A Design of Dialogue Interface System Based on Statistical Approach

Hyoungil Jeong[†] Donghyun Kim[†] Hyojun Jang[†] Harksoo Kim^{††} Jung-yun Seo^{†††}
Dept. of Computer Sciences, Sogang University[†]
Prog. of Computer and Communications Engineering, Kangwon National University^{††}
Dept. of Computer Science and Interdisciplinary Program of Integrated Biotechnology, University^{†††}

요 약

대화 인터페이스 시스템(dialogue interface system)은 인간과 컴퓨터의 상호작용을 위한 도구로서 자연언어(natural language)를 사용하여 정보를 공유하거나 특정 업무를 수행하는 프로그램이다. 대화 인터페이스 시스템에 대한 기존의 연구들은 영역 의존적인 스크립트나 계획 추론을 위한 계획 지식을 이용해 왔다. 스크립트 모델(script model)은 제한적인 실용시스템 개발을 위해 주로 연구되었고 계획에 기반한 모델은 대화의 원리를 이해하는 분야에서 주로 연구되어 왔다. 그러나 기존의 모델들은 시스템 확장이 매우 어려우며 예측하지 못한 사용자 발화에 대하여 대응이 어렵기 때문에 매우 제한적인 영역이나 정해진 형태의 대화만을 처리할 수 있다.

본 논문에서는 이런 단점들을 보완하기 위하여 통계에 기반한 다 영역(multi-domain) 대화 모델을 제안한다. 제안된 시스템은 각 작업들에 대하여 해당 작업에 적합한 영역 모델(domain model)을 잘 알려진 프레임 구조를 따르면서 사용자 의도 파악과 시스템 의도 생성에 통계적 방법을 사용한다. 이러한 하이브리드 형태의 구조 덕분에 제안된 시스템은 영역 확장성과 이식성이 뛰어나다는 장점을 가진다.

1. 서론

인간의 의사소통 방법들 가운데 가장 자연스러운 방법은 자연언어를 통한 대화이다. 따라서 대화를 사용하여 인간과 컴퓨터 간의 의사소통을 하려는 연구가 시도되고 있으며[1], 인간과 컴퓨터의 상호작용의 도구로서 대화를 사용하여 특정 목적을 성취할 수 있도록 도와주는 프로그램을 대화시스템이라 한다.

대화시스템은 계획에 기반한 모델[2][3], 스크립트 기반 모델[4]로 분류할 수 있다. 계획 기반 대화 모델은 복잡한 답화 현상을 처리할 수 있고 사용자의 심층적인 의도 추론이 가능하다. 하지만 계획 기반 모델은 대화 현상의 원리를 이해하는데 중점을 두었고 초기 설계가

매우 복잡하고 유지, 관리가 어렵기 때문에 실제 환경에서 활용 가능한 인간-기계 상호작용 모델에 적용하기에는 아직 많은 문제점이 있다. 스크립트 기반 모델은 일반적으로 순차적인 실행 구조를 가지며 비교적 쉬운 설계가 가능하지만 영역에 매우 의존적이며 다양한 상황에 대응이 어렵다. 대화 시스템을 실제 환경에 적용하기 위해서는 사용자의 다양한 의도에 대응할 수 있으며 영역 확장성과 시스템 관리에 용이한 구조가 필요하다.

본 논문은 기존 시스템에서 이러한 내용을 보완하기 위해 통계적 방법을 사용하여 사용자의 의도 파악과 시스템의 의도를 생성한다. 제안하는 시스템은 대화 말뭉치를 통하여 다양한 대화 상황을 추가 할 수 있고 영역 확장에 용이한 구조를 갖는다. 또한 기존의 프레임 방식[5]을 보완하여 효율적으로 다영역 대화에 대응할 수 있다. 따라서 본 논문은 기존의 대화 시스템에 통계적 접근방법을 적용한 다영역 대화 인터페이스를 제안한다.

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

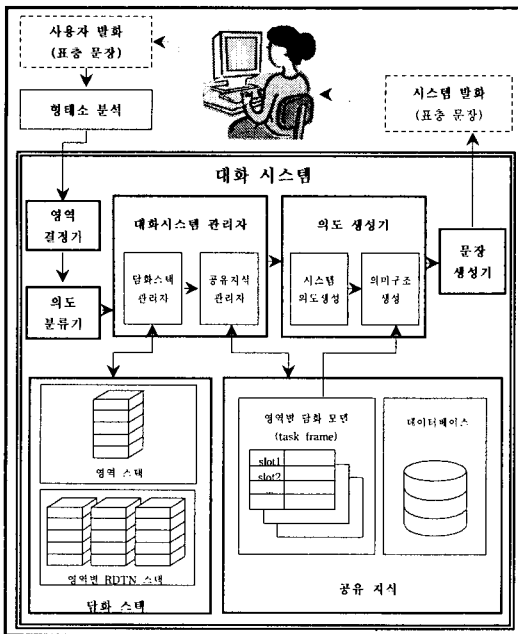
1) 계획 인식에 필요한 구조화된 지식 (e.g. recipe)

본 논문의 구성은 다음과 같다. 2장에서는 대화 인터페이스의 전체적인 구조와 각 세부 구성 요소들의 구조를 설명한다. 3장에서는 시스템 구현 진행상황과 시스템의 예비 실험으로서 사용자 의도 분류기의 성능측정 실험 결과를 보여준다. 4장에서는 결론 및 향후 과제에 대하여 설명한다.

2. 시스템 구조

본 논문에서 제안하는 통계 기반 대화 인터페이스 전체 시스템 구성은 [그림 1]과 같다. 본 절에서는 전체 시스템의 구조 설명과 각 모듈의 역할과 기능에 대하여 설명한다.

시스템은 사용자 발화의 표층 문장에 대한 형태소 분석 결과를 입력으로 사용한다. 영역 결정기 (domain detector)는 사용자 발화의 적합한 영역을 결정하고, 의도 분류기(intention classifier)는 해당 영역의 대화스택과 형태소 분석결과를 이용해 사용자의 의도를 의미 구조(semantic structure) 형태로 추출한다. 대화 시스템 관리자(dialogue system manager)는 대화의 상태를 유지하는 대화스택과, 대화에 필요한 정보들을 공유지식에 저장하고 관리하며, 의도 생성기(intention generator)는 시스템의 의도를 의미구조 형태로 생성한다. 최종적으로, 문장 생성기에서는 시스템 발화의 의미구조와 가장 유사한 후보-구(candidate phrase)를 선택하여 문장 합성기에 전달한다. 문장 합성기는 전달된 정보를 사용하여 자연스러운 시스템 응답문장을 생성한다.

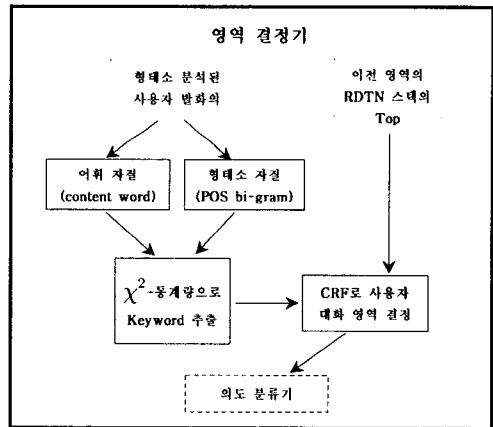


[그림 1] 통계 기반 대화 인터페이스 구성도

대화 인터페이스는 [그림 1]의 과정을 대화 종료 시까지 반복하여 사용자의 최종 목적을 이룰 수 있게 해준다.

2.1 사용자 발화의 영역 결정

영역 결정기는 사용자 발화가 정의된 여러 영역 중에 어떤 영역에 해당하는지 결정한다. 영역 결정기는 [그림 2]와 같이 형태소 분석된 사용자의 발화에서 각각의 형태소 바이그램(bi-gram)과 어휘(content word)를 자질로 삼아 X^2 -통계량[6]을 이용하여 사용자의 발화와 대화영역의 의존성을 측정한다. 영역 결정기는 측정된 결과를 통해 발화의 키워드를 추출하고, 추출된 키워드를 CRF(Condition Random Field)[7]를 이용해 영역을 결정한다. 예외 사항으로 사용자 발화가 “응, 그래”의 경우처럼 모든 영역에 빈번하게 나타나는 경우에는 영역 의존성이 없는 것으로 간주하여 영역이 변하지 않고 이전 발화의 영역과 동일한 영역으로 결정한다.



[그림 2] 영역 결정기

2.2 사용자 발화의 의도 분류

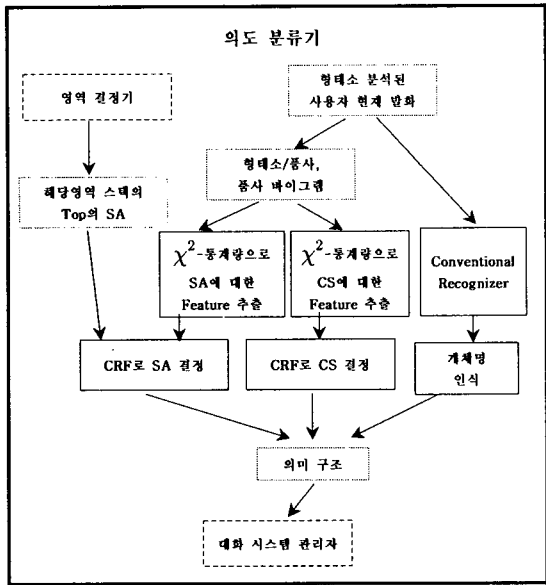
의도 분류기는 사용자의 발화에 내포되어 있는 사용자의 의도를 파악한다. 사용자의 의도를 파악하기 위해 의도 분류기는 입력된 발화의 화행(Speech Act, SA)[8]과 개념열(Concept Sequence, CS)[8], 그리고 개체명(Named Entity, NE)[2] 인식을 통하여 의미 구조에 필요한 각 단위 정보를 추출한다.

화행을 결정하기 위해서 의도 분류기는 [그림 3]과 같이 사용자 발화의 형태소/품사 쌍과 품사 바이그램 자질 중에 각 화행들과 의존성이 높은 자질들을 X^2 -통계량으로 측정하여 추출한다. 추출된 자질들과 이전 화행을 CRF의 입력 자질로 사용하여 현재 발화의 화행을 결정한다.

2) 인명, 지명, 기관명 등을 의미하는 단어 또는 단어들.

개념열을 분석하기 위해서 의도 분류기는 화행 분류와 동일한 방법을 적용하여 현재 발화에 대한 개념열을 결정한다.

의도 분류기는 Conventional Recognizer[9]를 이용하여 개체명 분석을 하고 화행, 개념열, 개체명 들을 하나의 의미구조로 구성한 후 대화 시스템 관리자에게 전달한다.



[그림 3] 의도 분류기

2.3 대화시스템 관리자

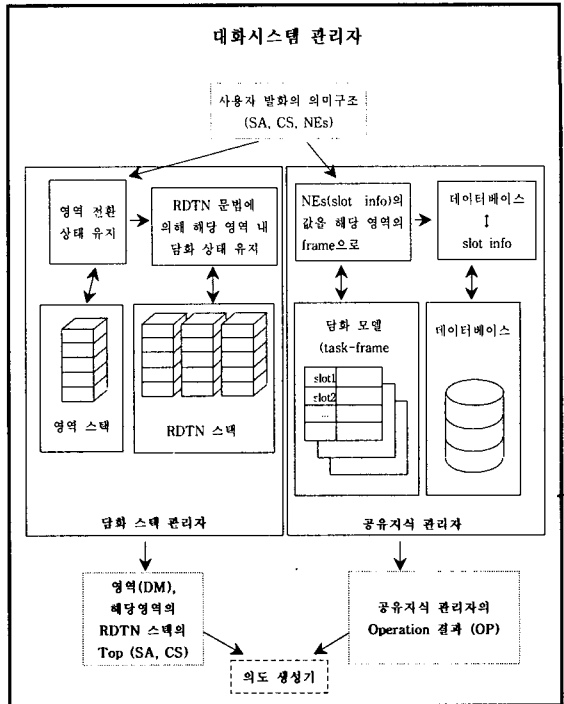
대화 시스템 관리자는 의도 결정기를 통하여 결정된 사용자 발화의 의미구조에 따라 시스템 내부의 정보들을 변화시키거나 필요에 따라 내부의 정보를 사용자에게 전달하는 역할을 한다[그림 4].

대화시스템 관리자는 담화스택 관리자와 공유지식 관리자로 구성된다. 담화스택 관리자는 영역별로 대화의 진행 상태를 유지하는 영역스택과 각 영역별 담화의 진행 상태를 저장하고 있는 RDTN(Recursive Dialogue Transition Network)[10]스택을 관리한다.

공유지식 관리자는 사용자와 시스템의 공유하는 지인 담화 모델을 프레임(frame)형태로 유지하고, 대화가 진행됨에 해당 담화 모델을 갱신하는 역할을 하며 그 결과를 의도 생성기에 전달한다.

2.3.1 담화 스택 관리자

담화 스택 관리자는 담화스택을 운용하여 사용자 발화에 따라 영역전환 상태와 담화 진행 상태를 저장 및 유지하는 역할을 한다. 담화 스택 관리자는 사용자 발



[그림 4] 대화시스템 관리자

화의 의미구조를 입력으로 받아 영역스택과 RDTN스택을 관리한다.

영역 스택의 운용방법은 새로운 영역의 대화가 시작되면 그 영역은 push되고, push된 영역의 대화가 종료되면 해당 영역은 pop되는 과정이며 이 과정을 통하여 영역 전환 상태를 유지한다.

RDTN 스택은 인접 쌍(adjacency pairs)³⁾의 원리를 사용하여 운용된다. RDTN은 대화의 전이 상태를 네트워크로 표현한 것이며 전이 규칙은 인접 쌍을 기반으로 작성한다. 예를 들어 request와 response가 인접쌍일 때, request type의 화행을 갖는 발화가 입력되면 그 발화는 push되고, 그 다음에 response type의 화행을 갖는 발화가 입력되면 request - response 인접쌍이 성립되므로 request - response에 해당하는 두 개의 발화는 pop된다. RDTN 스택은 각각의 영역별 대화를 독립적으로 유지한다. 예외 사항으로 두 개의 질문 발화에 하나의 발화로 답하거나 앞선 발화에 대답을 하지 않는 경우와 같이 인접 쌍의 원리에 해당하지 않는 경우에는 이를 처리하기 위해 스택의 정보를 강제로 제거 하는 예외 사항 처리 규칙을 추가하였다.

2.3.2. 공유지식 관리자

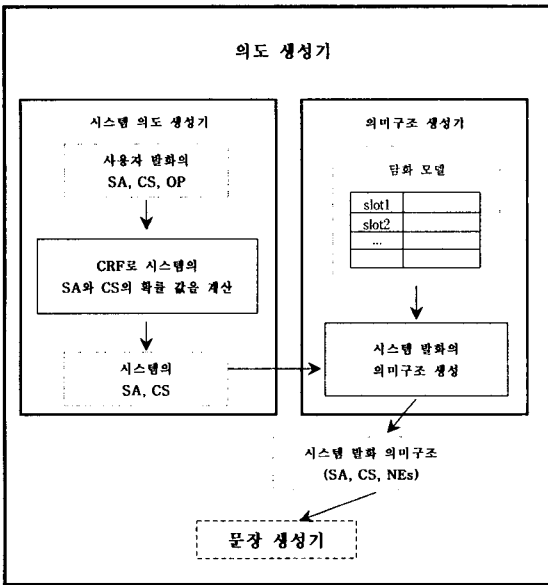
3) 언어적 기능수행에 있어서 쌍을 이루는 인접된 발화.

공유지식 관리자는 담화 모델을 통해 시스템과 사용자가 현재 진행되는 대화의 지식을 공유할 수 있도록 하는 역할을 한다.

담화 모델은 영역별로 현재 작업의 진행 상태를 유지한다. 입력에 따라 현재 담화 모델을 갱신하고 현재 상태에 따라 사용자로부터 작업을 수행하는 데 필요한 정보를 얻거나 요구한다. 예를 들어, 현재 영역은 '전화번호 검색'이고, '전화번호 검색'에 해당하는 담화모델의 slot1은 '이름', slot2 '전화번호'라 가정하자. 입력된 발화에 의해 slot1에 '홍길동'이 채워지고 개념열이 'inputel-num'인 경우 시스템은 slot2를 채우기 위하여 사용자에게 '전화번호'에 대한 정보를 요구하는 응답을 생성할 것이다.

대화시스템 관리자의 수행이 종료되면 대화시스템 관리자는 해당영역의 RDTN스택의 가장 상위에 있는 화행과 개념열, 그리고 공유지식 관리자의 처리 결과(Form Operation, OP)를 의도 생성기로 전달한다.

2.4. 시스템 의도 생성



[그림 5] 의도 생성기

의도 생성기는 시스템 의도 생성기와 의미구조 생성기로 구성되며 시스템 응답을 위한 시스템의 의도를 의미구조의 형태로 생성한다[그림 5].

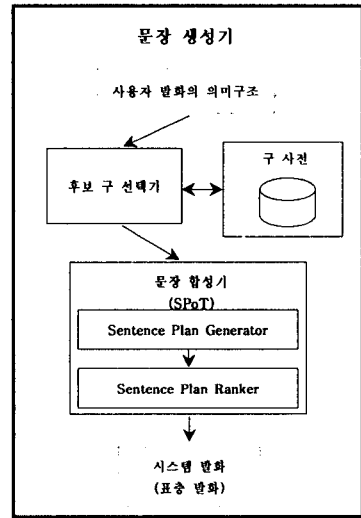
시스템 의도 생성기는 사용자 발화의 화행, 개념열, OP를 자질로 사용하고 CRF를 통해 가장 높은 확률값을 갖는 시스템의 화행과 개념열을 결정하여 의미구조 생성기로 전달한다.

의미구조 생성기는 담화모델의 개체명과 시스템 의도 생성기에서 전달된 화행과 개념열을 조합하여 시스템

발화 의미 구조를 생성한다.

2.5. 시스템 발화의 표층 문장 생성

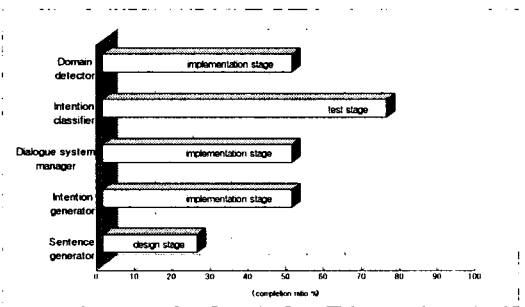
문장 생성기는 후보-구 선택기와 문장 합성기로 구성된다. 후보-구 선택기는 구 사전에서 시스템 발화 의미 구조의 성분인 화행, 개념열, 그리고 개체명과 각각 일치하는 구들을 결정한다. 결정된 구들은 문장 합성기로 전달되어 완전한 문장으로 생성된다. 문장 합성기는 SPoT[11]을 사용하였다.



[그림 6] 문장 생성기

3. 예비 실험

3.1. 구현 현황



[표 1] 구현 현황

본 논문에서 제안한 통계 기반 대화 모델의 현재 구현 현황은 [표 1]과 같다. 의도 분류기는 구현이 완료되어 테스트 과정에 있다. 영역 결정기와 대화 시스템 관리자, 그리고 의도생성기는 구현 단계에 있으며 문장

생성기는 현재 디자인 단계에 있다.

[표1]에서와 같이 우리는 의도 결정기 모듈에 대하여만 구현을 마쳤다. 따라서 본 논문에는 의도 결정기에 대한 성능 측정 결과만을 작성하였다.

3.2. 실험 데이터

의도 분류 실험을 위하여 일정관리 영역에서 Wizard-Of-Oz 방식으로 시뮬레이션한 대화 말뭉치를 수집한 후, 수동으로 화행과 개념열을 부착하였다. 수집된 말뭉치는 일정 추가, 삭제, 변경과 관련된 내용을 포함하고 있다. 말뭉치의 구성은 956개의 대화(21,336개의 발화)로 구성되며, 대화 당 평균 발화의 수는 22.32개이다. 훈련 데이터와 실험 데이터의 비율은 4:1로 나누었으며, 5배 교차검증(5-fold cross validation)을 수행하였다. CRF의 내부 설정 인자로 추정알고리즘은 L-BFGS[12]를 이용하였으며, 희소 데이터 문제를 위한 평탄화 요소는 Gaussian Prior[13]를 이용하였다. Gaussian Prior의 값은 10으로 설정하였으며, 훈련 반복 회수는 30으로 설정하였다.

3.3. 실험 결과

우리는 화행 분류와 개념열 분류를 동일한 말뭉치를 사용하여 실험하였다. 표 1은 각 분류기의 정확도를 나타낸다.

(단위: %)

| 모 델 | 화행 분류의 정확도 평균 | 개념열 분류의 정확도 평균 |
|--------|---------------|----------------|
| 김(03') | 85.5 | - |
| 최(05') | 88.6 | - |
| 제안 모델 | 93.0 | 90.2 |

[표 2] 각 분류기의 성능 평가

제안 모델과 김(03')[14], 최(05')[15]는 서로 다른 입력 자질을 사용하기 때문에 직접적으로 비교하는 것은 힘들다. 그러나 [표 2]에서 보는 것과 같이 화행 분류에 대한 실험에서 이전의 두 모델보다 월등히 좋은 결과를 보였다. 또한 제안 시스템은 개념열 분류도 동일한 구조 내에서 수행할 수 있다는 장점이 있다.

4. 결론

본 논문에서는 통계 기반 다 영역 대화 모델을 제안하였다. 제안된 모델은 영역 간 전이를 모델링하기 위하여 영역별로 답화스택을 유지하는 방법을 사용하였다. 그리고 사용자 의도 분석과 시스템 의도 생성을 통계적 방법으로 접근하여 높은 확장성과 이식성을 갖추어서 시스템 유지, 관리에 효율적인 구조를 갖는다.

5. 참고 문헌

- [1] George Ferguson, James Allen, "TRIPS: An Intelligent Integrated Problem-Solving Assistant," *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, Madison, WI, 26-30 July 1998, pp. 567-573.
- [2] Chu-Carroll and S. Carberry, "Conflict Resolution in Collaborative Planning Dialogues," *International Journal of Human-Computer Studies*, Vol. 53(6), pp. 969-1015, 2000.
- [3] 윤철진, 서정연, "제한된 영역의 대화에서 체인구 형태의 발화 이해를 위한 계획기반 생략 처리", *한국인지과학회 논문지*, 제11권 1호, pp81-92, 2000
- [4] 박정준, *자연언어 대화 인터페이스를 위한 다영역 대화 모델의 설계*, 석사학위논문, 서강대학교, 1996.
- [5] 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.
- [6] Yiming Yang and Jan O. Pedersen, "A Comparative Study on Feature Selection in Text Categorization," *Proceedings of the 14th International Conference ICML-97 on Machine Learning*, 1997
- [7] John Lafferty, Andrew McCallum and Fernando Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data", *International Conference on Machine Learning*, 2001
- [8] Austin, J. L., *How to do Things with Words*. Oxford, 1962
- [9] Choong-Nyoung Seon, Youngjoong Ko, Jeong-Seok Kim and Jungyun Seo, "Named Entity Recognition using Machine Learning Methods and Pattern- Selection Rules", *Proceeding of NLPRS 2001*, pp 229-236, Tokyo, Japan, Nov. 2001
- [10] Ron Cole, Joseph Mariani, Hans Uszkoreit, Giovanni Batista Varile, Annie Zaenen, Antonio Zampolli, and Victor Zue (eds.), *Survey of the State of the Art in Human Language Technology*, Cambridge University Press and Giardini, 1997
- [11] Marilyn A. Walker, Owen Rambow and Monica Rogati, "SPoT: A Trainable Sentence Planner", *Proceedings of ACL*, 2001.
- [12] Tamara Gibson (Kolda), Dianne P. O'Leary, and Larry Nazareth, "BFGS with Update Skipping and Varying

Memory", Technical Report CS-TR-3663, Department of Computer Science, University of Maryland College Park, 1996.

[13] Stanley F. Chen, Ronald Rosenfeld, "A Gaussian Prior for Smoothing Maximum Entropy Models", Technical Report CMU-CS-99-108, Computer Science Department, Carnegie Mellon University, 1999.

[14] Kyungsun Kim, Harksoo Kim and Jungyun Seo, "A neural

network model using a feature selection method for a speech act analysis in Korean," *International Journal of Neural Systems*, Vol. 14, No. 6, pp. 1-8, 2004.

[15] Won Seug Choi, Harksoo Kim and Jungyun Seo, "An integrated dialogue analysis model for determining speech acts and discourse structures," *IEICE Transactions on Information and Systems*, Vol. E88-D, No 1, pp150-157, Jan. 2005