

일정 관리 영역에서의 화행 분석을 위한 효과적인 2단계 모델*

이 현 정

김 학 수[†]

서 정 연

서강대학교 컴퓨터공학과 강원대학교 컴퓨터정보통신공학전공 서강대학교 컴퓨터공학과

화행은 화자의 의도를 내포하기 때문에 지능형 대화 시스템을 구현할 때 화행 분석은 필수적이다. 본 논문에서는 효과적인 화행 분석을 위한 2단계 모델을 제안한다. 첫 번째 단계에서 기계 학습 기반의 신경망 모델과 확률 기반의 예측도 모델을 이용하여 개별적으로 화행 후보를 생성한다. 두 번째 단계에서는 신경망 모델이 출력한 화행 후보들을 예측도 모델이 출력한 화행 후보들로 필터링한 후, 남겨진 후보들 중에서 최대값을 가지는 화행을 선택한다. 일정관리 영역에서의 실험 결과, 제안한 2단계 모델링 방법이 기계 학습 모델이나 확률 모델만을 사용하는 기존의 방법들보다 좋은 성능을 보였다.

주제어 : 화행 분석, 신경망 모델, 예측도 모델

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

† 교신저자: 김학수, 강원대학교 컴퓨터정보통신공학전공, 연구세부분야: 자연어처리

E-mail: nlpdrkim@kangwon.ac.kr

서 론

대화 시스템은 인간과 컴퓨터가 자연어 대화를 통하여 상호작용할 수 있도록 인간의 언어를 이해하고 적절한 응답을 찾아 주는 지능형 소프트웨어를 말한다. 이러한 대화 시스템은 일반적으로 사용자 의도를 인식하는 부분과 인식된 의도에 따라 적절한 응답을 생성하는 부분으로 나뉜다. 사용자 의도 분석에 있어서 가장 기초적이고 핵심적인 역할을 담당하고 있는 것이 화행 분석이다. 화행 분석을 통해서 대화 시스템은 전체 대화 흐름 속에서 현재 발화를 통해서 사용자가 의도한 언어적 행위를 찾아내게 된다. 화행은 크게 표면적인 정보만을 고려한 표층 화행(surface speech act)과 문맥 정보까지 고려한 심층 화행(deep speech act)으로 나눌 수 있다. 표층 화행 분석은 현재 발화에 나타난 문형 패턴만을 고려하면 되기 때문에 비교적 단순한 패턴 매칭(matching) 방법으로 구현이 가능하다. 그러나 실제 대화 시스템을 구현하기 위해서는 전체 대화 흐름을 고려한 심층 화행 분석이 필요하며, 이를 위해서는 현재 발화에 포함된 단서 단어와 이전 발화까지의 화행 정보 등 다양한 자질을 이용한 기계 학습 방법들이 주로 이용된다. 본 논문에서는 기존의 전통적인 기계 학습 방법에 확률 기반의 범주화 모델을 결합함으로써 보다 향상된 정확률을 보이는 새로운 심층 화행 분석 모델을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 심층 화행 분석과 관련된 기존 연구를 설명하고 문제점을 지적한다. 3장에서는 신경망과 예측도를 이용한 새로운 2단계 화행 분석 모델을 제안한다. 4장에서는 실험 데이터를 설명하고 실험 결과를 분석한다. 마지막으로 5장에서 결론을 내린다.

관련 연구

화행 분석에 관한 기존의 연구들은 계획 추론 모델(plan inference model)을 위한 레서피(recipe)나 영역 의존적인 지식에 기초해 왔다[1,2,3]. Litman-1987[3]은 부대화(sub-dialogue)를 처리하기 위한 계획 인지 모델을 제시하였다. 계획 인지를 위한 기본 정보로는 REQUEST, INFORM, INFORMREF, INFORMIF라는 매우 제한적인 4개

의 화행을 이용하였으며, 간단한 격틀 형태의 의미 분석을 통해서 화행을 인식하였다. Lambert-1991[2]은 화행과 유사한 형태의 담화 목적을 정의하여 계획 인식 과정에서 나타나는 생략 문제를 해결하는데 이용하였다. 담화 목적 인식을 위해서는 Litman-1987과 유사한 방법을 사용하였다. 이러한 계획 추론 모델에 기반한 연구들은 응용 영역을 확장하거나 변경할 때마다 대량의 의미 지식을 수동으로 구축해야 한다는 단점이 있다. 최근에는 이러한 단점을 극복하기 위하여 다양한 기계 학습 모델들이 이용되고 있다[4,5,6,7]. Kim-2003[5]은 문서 분류 기법을 화행 분석에 도입하여, 현재 발화에 포함되어 있는 단서 단어 중에서 가장 정보량이 많은 것들을 찾아내는 자질 추출기법을 제안하였다. Eun-2005[6]는 SVM(support vector machine)을 이용하여 화행을 분류하였으며, Kim-2003에서 제안한 자질들에 구문 자질(구문 분석기를 이용하여 반자동으로 구축되는 문법적 역할과 관련된 자질)을 추가하였을 경우에 화행 분류 성능이 향상된다는 것을 보였다. Lee-2006[4]은 신경망을 이용하여 화행을 분류하였으며, 개념열(신경망을 통하여 발화로부터 자동으로 획득 가능한 의미 분류)이라는 자질이 화행 분류 성능에 도움이 된다는 것을 보였다. Kim-2007[7]은 예측된 다음 발화의 후보 화행을 현재 발화의 화행 분류에 적절히 사용하면 성능이 향상된다는 것을 보였다. 이러한 기계 학습 모델들은 대용량의 학습 데이터로부터 자동으로 추출된 자질들을 이용하여 화행 범주에 사용자의 발화들을 효과적으로 할당한다. 이러한 기계 학습 모델들은 입력 자질을 기반으로 화행 범주들 사이의 경계를 보다 세밀하게 나누는 것에 초점을 맞추어 왔다. 그러므로 사용자의 발화가 범주의 경계에 위치할 경우에 범주 전체의 특성이 반영되지 못한 채 경계 주변에 대한 정보만을 이용함으로써 잘못된 결과를 초래하는 경우가 종종 있었다. 이러한 약점을 보완하기 위해서 본 논문에서는 전통적인 기계 학습 모델인 신경망과 확률 기반 범주화 모델인 예측도(predictivity)를 이용하여 화행 분석의 정확률을 향상시키는 방법을 제안한다. 제안 모델은 범주의 경계 구분에 적합한 신경망 모델과 범주 전체의 특성 반영에 적합한 확률 모델로 구성되며, 각 모델은 입력 발화에 대해서 개별적으로 화행 후보들을 출력한다. 그리고 확률 모델이 출력한 화행 후보를 이용하여 신경망 모델이 출력한 화행 후보를 필터링(filtering)하는 방법으로 전통적인 기계 학습 모델이 가지는 약점을 보완한다.

2단계 화행 분석 모델

신경망을 이용한 화행 분석

n 개의 발화로 구성된 대화 $U_{1,n}$ 이 주어졌을 때, 각 발화의 화행 $SA_{1,n}$ 을 분석하는 확률 모델은 수식 (1)과 같이 정의될 수 있다.

$$SA(U_{1,n}) = \arg \max_{S_{1,n}} P(S_{1,n} | U_{1,n}) \quad (1)$$

수식 (1)에서 현재 화행은 현재 발화에 영향을 받는다는 독립가정과 바로 이전 화행에 영향을 받는다는 1차 마코프(Markov) 가정을 하면 수식 (2)와 같다.

$$SA(U_{1,n}) = \arg \max_{S_{1,n}} \prod_{i=1}^n P(S_i | U_i) P(S_i | S_{i-1}) \quad (2)$$

화자는 개인의 언어적 습관에 따라 동일한 의미의 문장을 단어의 순서를 바꾸거나 불필요한 단어를 생략하는 방법 등을 통하여 다양한 형태로 표현하므로 $P(S_i | U_i)$ 를 직접 계산하는 것은 거의 불가능하다. 이러한 문제를 해결하기 위해서 본 논문에서는 수식 (3)과 같이 발화를 문장 자질들의 집합으로 일반화한다[5].

$$SA(U_{1,n}) = \arg \max_{S_{1,n}} \prod_{i=1}^n P(S_i | F_i) P(S_i | S_{i-1}) \quad (3)$$

수식 (3)에서 F_i 는 i 번째 발화의 문장 자질 집합으로써 어휘 자질과 품사 자질로 구성된다. 어휘 자질은 품사가 부착된 어휘를 말하며, 품사 자질은 품사 바이그램(bigram)을 말한다. 예를 들어, “안녕하세요?”라는 발화의 형태소 분석 결과가 ‘안녕/ncp+하/xsp+세요/ef+ ?/sf’라고 했을 때, 어휘 자질은 ‘안녕/ncp, 하/xsp, 세요/ef, ?/sf’이고, 품사 자질은 ‘ncp-xsp, xsp-ef, ef-sf’이다. 어휘 자질과 품사 자질의 추출이 끝나면 각 자질과 화행 범주 사이의 χ^2 통계량을 계산하고, 값이 큰 상위 n 개의 자질을 선택하여 정보량이 적은 자질들을 제거한다[5]. 문장 자질 집합을 만들 때,

다양한 n -그램(n -gram) 언어모델을 사용하지 않은 이유는 자질 수의 기하급수적인 증가 때문이다. χ^2 통계량을 이용하여 상위의 자질들만을 선택한다고 하더라도 의미 있는 자질들을 포함시키려면 그 수가 적지 않다. 이러한 자질 수의 폭발적인 증가는 기계 학습 모델들의 학습 시간 증가 및 희소 데이터 문제를 일으키는 원인이 된다. 그러므로 본 논문에서는 기존의 연구들[5,6,8,9]에서 그 효용성이 입증된 품사 부착 어휘 유니그램(unigram)과 품사 바이그램을 기계 학습 모델의 자질로 사용한다. 최종적으로 완성된 화행 분석 모델인 수식 (3)을 다중 퍼셉트론(multi-layer perceptron)으로 표현하면 그림 1과 같다.

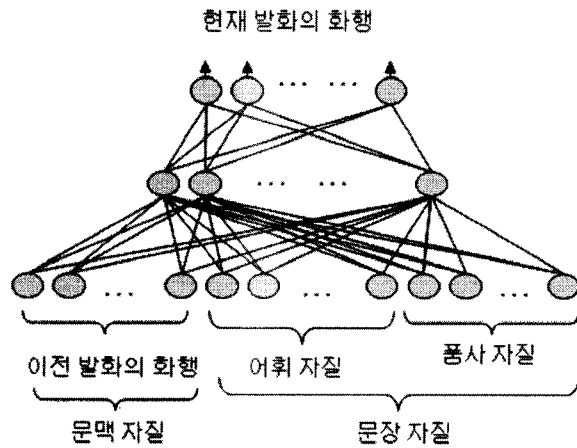


그림 1. 신경망을 이용한 화행 분석 모델

예측도를 이용한 화행 분석

발화가 특정 화행 범주에 속할 가능성을 추정하는 범주화 확률 모델은 수식 (4)와 같이 정의될 수 있다.

$$P(S_i|U) \approx P(S_i|F) \tag{4}$$

수식 (4)에서 U 는 발화를 의미하며, S_i 는 i 번째 화행 범주를 의미한다. 본 논문에서는 화행이 부착된 말뭉치에서 추출한 화행별 발화 집합을 화행 범주로 취급한다. F 는 발화를 일반화한 자질 집합으로 1에서 4 길이의 형태소 n -그램(n -gram)으로 구성된다[10]. 예를 들어, ‘안녕하세요.’라는 발화의 자질 집합은 1-그램으로 ‘안녕, 하, 세요, .’, 2-그램으로 ‘안녕;하, 하;세요, 세요.’, 3-그램으로 ‘안녕;하;세요, 하;세요.’, 4-그램으로 ‘안녕;하;세요.’가 된다. 자질 집합으로 신경망 모델이 사용한 것에 비해 다양한 언어 모델을 사용한 이유는 제안 모델이 범주 전체에 대한 단순 확률 계산을 기반으로 하기 때문에 별도의 학습 시간이 필요 없고, 희소 데이터 문제에서도 조금 더 자유로울 수 있기 때문이다. 수식 (4)에서 발화와 특정 범주 사이의 연관성을 나타내는 예측도는 수식 (5)과 같이 각 자질들이 특정 범주에 속할 가능성 중 최대값으로 표현될 수 있다[11].

$$\begin{aligned}
 P(S_i|F) &= \max_{1 \leq j \leq n} P(S_i|F^j) \\
 &= \max_{1 \leq j \leq n} \frac{\text{freq}(S_i, F^j)}{\text{freq}(F^j)}
 \end{aligned}
 \tag{5}$$

수식 (5)에서 F^j 는 발화 U 의 형태소 n -그램(n -gram) 중 j 번째 자질을 의미한다. $\text{freq}(F^j)$ 는 j 번째 형태소 n -그램이 전체 화행 범주에서 출현한 빈도수를 의미하며, $\text{freq}(S_i, F^j)$ 는 j 번째 형태소 n -그램이 i 번째 화행 범주 S_i 에서 출현한 빈도수를 의미한다.

화행 예측도를 이용한 필터링

신경망과 같이 기계 학습 기반의 화행 분석 모델들은 화행 범주 전체에 대한 특성보다는 범주 경계를 구분 짓는 특성을 잘 이용하는 것에 초점이 맞추어져 있다. 이러한 약점을 보완하기 위해서 본 논문에서는 수식 (5)에서 계산한 화행 예측도를 이용하여 신경망 모델의 출력 결과 중에서 범주 전체의 특성이 보다 잘 반영된 것을 선택하는 방법을 제안한다. 제안 모델은 먼저 그림 2와 같은 신경망을 이용하여 화행 후보들(신경망의 출력값이 0이 아닌 화행들)을 결정한다. 둘째로 수식

(5)에 따라 상위 n 개의 화행 후보를 별도로 결정한다. 그리고 신경망이 제시한 화행 후보들 중에서 화행 예측도에 따라 결정된 화행 후보에 속하지 않는 것들을 제거한다. 마지막으로 제거되지 않고 남겨진 신경망의 화행 후보 중에서 가장 큰 출력값을 가진 것을 최종 화행으로 결정한다.

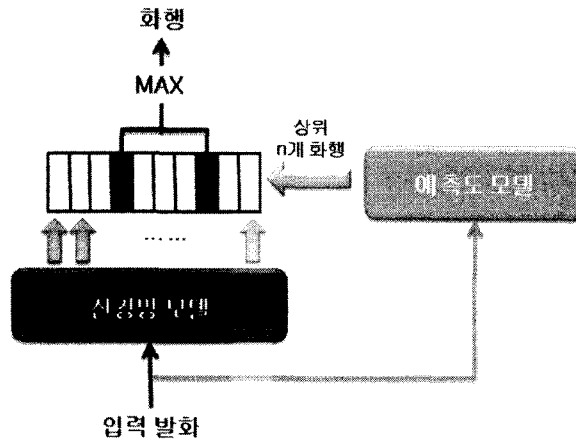


그림 2. 화행 예측도를 이용한 필터링 과정

실험 및 평가

실험 데이터

제안 모델을 실험하기 위해서 일정관리 영역에서 Wizard-Of-Oz 방식으로 시뮬레이션한 대화 말뭉치를 수집한 후, 수동으로 화행을 부착하였다. 화행 부착은 자연어처리를 전공하는 5명의 석사과정 학생이 수행하였으며, 일관성을 위하여 1명의 박사과정 학생이 최종 검토하였다. 표 1은 화행이 부착된 대화의 예를 보여주며, 표 2는 이현정-1997[12]을 참고하여 정의한 화행을 보여준다. 수집된 말뭉치는 일정 추가, 삭제, 변경과 관련된 내용을 포함하고 있다. 말뭉치의 구성은 956개의 대화(21,336개의 발화)로 구성되며, 대화당 평균 발화의 수는 22.32개이다. 말뭉치 중

856개 대화를 학습 데이터로 사용하였으며, 나머지 100개 대화를 실험 데이터로 사용하였다. 신경망의 학습은 역전파 알고리즘(back propagation algorithm)을 이용하였으며, 학습률(learning rate)은 0.2, 학습시간은 200 epoch로 설정하였다.

표 1. 화행 부착 대화의 예

발화 (U: 사용자, S:시스템)	화행
(1) U: 안녕	Greeting
(2) S: 무엇을 도와드릴까요?	Opening
(3) U: 내일 일정 좀 알려줘	Request
(4) S: 오전11시에 홍길동님과 약속이 있습니다.	Response
(5) U: 약속이 변경되었어.	Inform
(6) S: 어떤 내용이 변경되었나요?	Ask-ref
(7) U: 날짜가 바뀌었어.	Response
(8) S: 언제로 바뀌었나요?	Ask-ref
(9) U: 12월 5일이야.	Response

표 2. 화행의 종류

화행	설명
Greeting	대화 서두의 인사말
Expressive	대화 후미의 인사말
Opening	실제 대화의 시작
Ask-ref	WH-질문
Ask-if	YN-질문
Response	응답
Request	행위를 요청
Ask-confirm	이전 발화의 확인
Confirm	확인 발화의 응답
Inform	정보 제공
Accept	호응

실험 결과

표 3은 화행 예측도에 따라 결정된 상위 n 개의 화행 후보 안에 정답 화행이 존재할 확률을 보여준다.

표 3. 상위 n 개까지의 화행 분석 정확률

개수	상위 1개	상위 2개	상위 3개	상위 4개	상위 5개
정확률	85.92	95.13	98.88	98.84	99.96

표 4는 화행 예측도에 따라 결정된 화행 후보 개수별 2단계 모델의 정확률을 보여준다. 표 4에서 보듯이 화행 예측도 상위 2위까지의 화행 후보를 사용했을 때 가장 높은 정확률을 얻을 수 있었다. 화행 후보의 개수가 많아질수록 신경망 모델의 결과에 의존하게 되어 정확률이 오히려 낮아짐을 알 수 있었다.

표 4. 후보 개수별 2단계 화행 분석 모델의 정확률

개수	상위 1개	상위 2개	상위 3개	상위 4개	상위 5개
정확률	85.92	91.37	90.88	90.34	90.21

표 5는 동일한 학습 데이터와 실험 데이터를 이용하여 기존의 모델들과 정확률을 비교한 것이다. 표 5에서 Webb-2005[11]는 수식 (5)와 같은 방법으로 발화에 포함되어 있는 모든 단어 n -그램들의 화행 예측도를 계산하고, 화행 예측도가 가장 큰 범주를 최종 화행으로 결정하는 모델이다. Lee-2006[4]은 입력 자질로써 개념열이라는 의미 분류들을 추가한 신경망 모델이다. 표 5에서 보듯이 제안 모델(상위 2개의 화행 후보를 사용한 2단계 화행 분석 모델)이 화행 예측도와 신경망을 개별적으로 사용한 기존의 모델들보다 우수함을 알 수 있었다.

표 6은 대표적인 기계 학습 모델들에 2단계 모델링 방법을 적용한 후, 정확률을 측정하는 것이다.

표 5. 기존 모델과의 화행 분석 정확률 비교

모델	Webb-2005	Lee-2006	제안 모델
정확률	85.92	86.05	91.37

표 6. 2단계 모델링 방법의 성능 향상 비교

모델	SVM		MEM		NN	
	단독	복합	단독	복합	단독	복합
정확률	90.16	90.88	88.42	91.16	86.05	91.37

표 6에서 단독은 해당 기계학습 모델 단독으로 실험한 결과이고, 복합은 상위 2개의 화행 후보를 사용하여 2단계 모델링한 방법을 적용하여 실험한 결과이다. 표 6에서 보는 것과 같이 마진(margin) 기반 기계 학습 모델인 SVM, 통계 기반 기계 학습 모델인 MEM(maximum entropy model), 그리고 함수 기반 기계 학습 모델인 NN(neural network)에 제안한 2단계 모델링 방법을 적용한 결과에 따르면 모든 모델에서 성능이 향상됨을 알 수 있었다. 결론적으로 제안한 2단계 모델링 방법이 기존의 기계 학습 방법들의 성능 향상에 도움이 됨을 알 수 있었다.

표 7은 2단계 모델링 방법이 화행 분석에 긍정적인 영향을 미친 예와 그렇지 못한 예를 보여준다. 표 7에서 보는 것과 같이 2단계 모델링 방법이 화행 분석에 긍정적 영향을 미친 예는 해당 화행 범주를 대표하는(또는 화행 범주의 중심에 가까운) 단서 어절이 발화 내에 포함된 경우가 대부분이었다. 표 7의 긍정적인 예에서 해당 발화의 화행을 결정하는데 가장 중요한 단서를 제공하는 것이 ‘입력해줘’라는 어절이다. 그러나 신경망 모델의 경우에 학습 속도에 따른 자질 수의 제한 때문에 해당 어절 전체를 입력 자질로 사용할 수 없으며, ‘입력/ncp’나 ‘어주/ep’와 같이 독립된 부분 정보만을 사용하게 된다. 그러므로 형태소와 형태소 사이의 연관 관계를 통해서 형성되는 정보들에 대해서는 손실이 불가피하다. 예측도 모델에서 사용하는 형태소 n -그램이 이러한 정보의 손실을 보정해 줌으로써 정확률이 향상된 것으로 보인다. 2단계 모델링 방법이 화행 분석에 부정적 영향을 미친 예는 예측도 모델이 문맥을 사용하지 않기 때문에 생기는 오분석이 전체 시스템의 성능

표 7. 제안 모델의 대표적인 출력 결과

화행	긍정적인 예			부정적인 예		
	그 요일에 할아버지 맥에 내려간다고 일정을 입력해 줘. (정답: Request)			일요일로 검색해줘. (정답: Response)		
	신경망 결과	예측도 결과	최종 결과	신경망 결과	예측도 결과	최종 결과
Accept	0			0		
Ask-confirm	0.00001			0.00003		
Ask-if	0			0		
Ask-ref	0.86555			0.17533	2위	Ask-ref
Confirm	0.00065			0.00053		
Expressive	0.00001			0		
Greeting	0			0		
Inform	0.00002			0		
Opening	0.0013			0.00018		
Request	0.1942	1위	Request	0.00094	1위	
Response	0.0001	2위		0.95965		

에 영향을 준 경우가 대부분이었다. 즉, 표 7의 부정적인 예에서 보는 것과 같이 문맥 정보를 사용하지 않는 예측도 모델의 결과를 필터로 사용함으로써 신경망 모델의 올바른 출력 결과가 무시되는 경우가 많이 발견되었다. 예의 경우에 예측도 모델은 ‘검색해줘’라는 어절에 포함된 정보를 바탕으로 ‘request’와 ‘ask-ref’를 출력하였으며, 이것을 바탕으로 신경망 모델의 올바른 출력 결과인 ‘response’가 무시됨으로써 오분석이 초래되었다.

결 론

본 논문에서는 전통적인 기계 학습 모델이 가지는 단점을 보완하기 위한 2단계

화행 분석 모델을 제안하였다. 제안 모델은 첫 번째 단계에서 신경망과 화행 예측도를 이용하여 화행 후보를 결정한다. 그리고 두 번째 단계에서 신경망을 이용한 화행 후보를 예측도를 이용한 화행 후보로 필터링한 후, 최종 화행을 결정한다. 제안 모델은 이러한 2단계 방법을 통하여 범주 경계에 대한 정보뿐만 아니라 범주 전체의 특성을 반영할 수 있게 됨으로써 기계 학습 모델이 가지는 약점을 보완할 수 있다. 일정 관리 영역에서의 실험 결과, 제안한 2단계 모델링 방법이 전통적인 기계 학습에 기반한 화행 분석 모델들의 성능 향상에 도움이 됨을 알 수 있었다.

참고문헌

- [1] Caberry, S. (1989), *A pragmatics-based approach to ellipsis resolution*, Computational Linguistics, 15-2, 75-96.
- [2] Lambert, L. and Caberry, S. (1991), "A tripartite plan-based model of dialogue," in *Proceedings of ACL 1991*, 47-54.
- [3] Litman, D. J. and Allen, J. F. (1987), *A plan recognition model for subdialogues in conversations*, Cognitive Science, 11, 163-200.
- [4] Lee, H., Kim, H., and Seo, J. (2006), *Efficient domain action classification using neural networks*, Lecture Notes in Computer Science, vol. 4233, pp. 150-158.
- [5] 김경선, 서정연 (2003), 자질 선택 기법을 이용한 한국어 화행 결정, **정보과학회 논문지: 소프트웨어 및 응용**, 30-3, 278-284.
- [6] 은종민, 이성욱, 서정연 (2005), 지지벡터기계(support vector machines)를 이용한 한국어 화행분석, **정보처리학회 논문지**, 12B-3, 365-368.
- [7] 김세종, 이용훈, 이종혁 (2007), "이전 문장 자질과 다음 발화의 후보 화행을 이용한 한국어 화행 분석", **제19회 한글 및 한국어 정보 처리 학술대회**, 23-30.
- [8] 김학수 (2008), 능동학습법을 이용한 한국어 대화체 문장의 효율적 의미 구조 분석, **정보과학회 논문지: 소프트웨어 및 응용**, 35-5, 306-312.
- [9] Verbree, D., Rienks, R., and Heylen, D. (2006), "Dialogue-act tagging using smart

- feature selection; results on multiple corpora,” in *Proceedings of the First International IEEE Workshop on Spoken Language Technology*, 70-73.
- [10] Reithinger, N. and Klessen, M. (1997), “Dialogue act classification using language models,” in *Proceedings of EuroSpeech*, 69-78.
- [11] Webb, N., Hepple, M., and Wilks, Y. (2005), “Dialogue act classification based on intra-utterance features,” in *Proceedings of the AAAI*
- [12] 이현정 (1997), **한국어 대화체 문장의 화행 분석**, 석사학위논문, 서강대학교.

1 차원고접수 : 2008. 7. 9

2 차원고접수 : 2008. 8. 20

최종게재승인 : 2008. 9. 1

(Abstract)

An Effective Two-Step Model for Speech Act Analysis in a Schedule Management Domain

Hyunjung Lee

Harksoo Kim

Jungyun Seo

Sogang University

Kangwon National University

Sogang University

Since speech acts implies speakers' intentions, it is essential to determine speakers' speech acts if we want to implement an intelligent dialogue system. We propose a two-step model for effectively determining speakers' speech acts. In the first step, the proposed model returns speech act candidates by using a neural network model based on machine learning and a predictivity model based on statistics, respectively. In the second step, using speech act candidates which are returned by the predictivity model, the proposed model filters out speech act candidates which are returned by the neural network model. Then, the proposed model selects a speech act with maximum output value among the unremoved speech act candidates. In the experiment on a schedule management domain, the proposed two-step modeling method showed better precisions than the previous methods only using a machine learning model or a probability model.

Keywords : Speech act analysis, neural network model, predictivity model