

한국어 화행 분류를 위한 최적의 자질 인식 및 조합의 비교 연구

(A Comparative Study on Optimal Feature Identification and Combination for Korean Dialogue Act Classification)

김민정[†] 박재현^{**} 김상범^{***} 임해창^{****} 이도길^{*****}
 (Min-Jeong Kim) (Jae-Hyun Park) (Sang-Bum Kim) (Hae-Chang Rim) (Do-Gil Lee)

요약 본 논문은 통계 기반 한국어 화행분류를 위하여 필요한 각 자질이 분류 성능에 미치는 영향과 성능 향상에 기여하는 자질 조합을 비교 평가한다. 지지벡터기계 학습 방법을 이용하여 구현한 화행 분류 시스템을 통해 실험한 결과, n-gram 자질 중 품사 바이그램은 유용하지 않으며 형태소-품사 쌍과 다른 자질들을 결합했을 때 성능이 향상됨을 알 수 있었다. 또한, 자질 선택 기법을 사용한 자질 비율에 따른 실험을 통해서 매우 적은 자질만으로도 화행 분류에 있어 어느 정도 안정된 성능을 낼 수 있었다. 아울러, 실험 결과의 분석을 통해 한국어에서 마지막 어절이 문장 전체의 화행분류에 중요한 역할을 하며, 한국어의 특징인 자유 어순이나 주어의 빈번한 생략 등이 화행 분류 실험의 성능에 영향을 미친다는 사실도 알 수 있었다.

키워드 : 화행 분류, 한국어의 특성, 자질 인식

Abstract In this paper, we have evaluated and compared each feature and feature combinations necessary for statistical Korean dialogue act classification. We have implemented a Korean dialogue act classification system by using the Support Vector Machine method. The experimental results show that the POS bigram does not work well and the morpheme-POS pair and other features can be complementary to each other. In addition, a small number of features, which are selected by a feature selection technique such as chi-square, are enough to show steady performance of dialogue act classification. We also found that the last eojeol plays an important role in classifying an entire sentence, and that Korean characteristics such as free order and frequent subject ellipsis can affect the performance of dialogue act classification.

Key words : Dialogue Act Classification, Korean Characteristics, feature identification

· 본 연구는 2단계 BK21사업과 2008년도 한국과학재단의 지원을 받아 수행된 연구임(No. R01-2006-000-11162-0)

[†] 학생회원 : 고려대학교 컴퓨터,전파통신공학과
 mjkim@nlp.korea.ac.kr

^{**} 비회원 : 고려대학교 컴퓨터,전파통신공학과
 jhpark@nlp.korea.ac.kr

^{***} 비회원 : SK Telecom 검색사업팀 매니저
 sangbumkim@sktelecom.com

^{****} 종신회원 : 고려대학교 컴퓨터,통신공학부 교수
 rim@nlp.korea.ac.kr

^{*****} 비회원 : 고려대학교 민족문화연구원 HK연구교수
 motdg@korea.ac.kr

논문접수 : 2008년 5월 20일

심사완료 : 2008년 10월 2일

Copyright©2008 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 소프트웨어 및 응용 제35권 제11호(2008.11)

1. 서론

대화를 하는 사람들은 발화를 통해 자신의 의도를 전달하며, 이렇게 발화를 통해 전달되는 의도를 화행이라고 한다[1]. 화행은 발화가 대화 내에서 수행하는 역할을 결정하기 때문에 대화에 관한 연구에서 매우 중요하다. 다음 그림 1의 대화에서 *진술(질문)* 다음에는 *진술*이 오고 *요청* 다음에는 *수락*, *거절*이 오거나 *이유*를 서술하는 *진술*이 오는 것을 볼 수 있다.

대화 시스템에서는 이와 같은 화행간의 연관성에 근거하여 현재 발화의 화행을 이용하여 다음 발화에 올 화행을 예측하고 다음 발화를 생성하는데 사용한다[2]. 또한 격틀(frame) 기반 대화 관리 시스템에서 발화를

1) 본 논문에서는 본문 설명에 나타나는 화행은 이탤릭체로 구분한다.

A: 빌 하틀리 씨 좀 바꿔 주시겠어요? (요청)
 B: 죄송하지만 지금 잠시 자리 비우셨는데요. (진술)
 A: 음..(단순감탄) 언제쯤 돌아오실까요? (진술질문)
 B: 점심 식사가 끝나면 바로 돌아오실 겁니다. (진술)
 A: 알겠습니다.(진술) 이따가 다시 전화 드리죠.(의지표명/약속)

그림 1 화행 부착 대화 말뭉치의 예

표현하기 적합한 격들을 결정할 때 화행을 이용하기도 한다[3].

화행 분류는 크게 규칙기반 방법과 통계기반 방법이 있으며, 규칙기반 방법은 규칙을 구축하는데 비용이 많이 들고 다른 도메인으로의 이식성이 좋지 않으므로 통계기반 방법을 많이 사용한다.

이전 통계기반 한국어 화행 분류 연구에서는 화행 분류에 필요한 자질을 제안하기만 했을 뿐 다른 연구에서 사용한 자질을 같이 사용하려는 시도나 각 자질에 대한 비교 평가 및 자질의 조합을 다양하게 하여 자질의 사용이 화행 분류의 성능에 미치는 영향을 분석하려는 평가가 없었다. 이에 본 논문에서는 이전 연구에서 사용한 다양한 화행 분류를 위한 자질들을 비교 평가하고 화행 분류에 적합한 자질 조합을 찾아본다.

기존의 통계기반 한국어 화행 분류에서 사용한 자질은 형태소-품사 쌍과 품사 바이그램, 구문 정보, 마지막 어절 및 특정 형태소를 이용한 자질이 있다[4-10]. 이중 구문 정보는 한국어 화행 분류를 위한 정보를 완전히 표현하지 못하기 때문에 한국어 화행 분류에 단독으로 쓰이기에 적합하지 않으며 다른 화행의 발화가 같은 구문 패턴으로 표현되어 중의성을 증가시키는 문제가 있다[8].

이에 본 논문에서는 구문 파서를 이용한 구문정보 자질은 배제하고 n-gram 유형 자질과 마지막 어절 및 특정 형태소 어휘를 이용한 자질을 대상으로 비교 실험을 수행한다. 실험은 도메인 비종속적인 화행 부착 대화 말뭉치를 대상으로 하며 지지벡터기계 학습방법을 이용한 화행 분류 시스템을 자질들과 결합하여 성능을 측정한다.

본 논문은 다음과 같은 순서로 구성되어 있다. 2장에서는 화행 분류와 관련된 기존 연구를 살펴본다. 그리고 3장에서 본 논문에서 사용할 자질을 설명하고, 4장에서는 3장에서 설명한 자질을 지지벡터기계 학습방법을 이용하여 실험 및 평가한다. 마지막으로 5장에서는 결론을 맺는다.

2. 관련 연구

화행 부착 말뭉치를 이용한 많은 통계 기반 영어 화행 분류에서 n-gram을 화행 분류의 자질로 사용하였다 [3,11-16]. 이중 Stolcke, Webb, Lan은 SWITCHBOARD 말뭉치[17]을 동일하게 사용하였다. SWITCH-

BOARD 말뭉치는 70가지 다른 주제를 다룬 전화 내용으로 구성되어 있으며 205,000말뭉치로 구축되어 있다. Stolcke는 발화 내의 단어 열과 발화간의 화행 열에 은닉 마코브 모델을 사용하였으며[14], Webb은 n-gram을 이용한 발화 내 자질만을 사용하여 Stolcke와 비슷한 성능을 보였다[15]. 이는 도메인 비종속적인 대화의 화행 분류에서는 대화의 흐름이 다양하기 때문에 발화간 자질이 발화 내 자질보다 상대적으로 성능에 영향을 미치지 못한다는 것을 보여준다. Lan은 n-gram과 함께 발화의 마지막 톤, 발화 내 동사의 유무 등의 다양한 자질을 이용하여 Stolcke보다 좋은 성능을 보였다[16]. 이러한 연구를 통해 영어 화행 분류에서 n-gram이 매우 간단한 자질이고 좋은 성능을 내지만 다른 자질과 함께 사용했을 때 더 좋은 성능을 보임을 알 수 있다.

통계 기반 한국어 화행 분석에 관한 기존 연구는 구문 분석 파서를 이용한 구문 정보만 사용한 연구, 형태소-품사 쌍과 품사 바이그램, 구문 정보를 모두 사용한 연구와 형태소-품사 쌍과 품사 바이그램만 사용한 연구, 그리고 마지막 어절 및 특정 형태소의 어휘를 이용한 연구로 나눌 수 있다.

초기의 통계 기반 한국어 화행 분석 연구에서는 구문 정보만을 이용하여 화행을 분류하였다[4,5]. 한국어 대화의 화행 분류 연구에서 사용한 구문 자질은 문장 유형, 주동사, 시제, 문장의 부정형 여부, 양상이나 보조용언, 단서 단어이다. 이러한 자질만으로 화행을 분류할 경우 제한된 정보만을 추출하게 되므로 화행 분류에 필요한 정보를 충분히 표현하기 어렵게 된다.

아래 두 예제 문장은 각기 다른 화행의 발화이다. 하지만 단서 단어를 제외한 구문 정보로 나타내면 {문장 유형=평서문, 주동사=같다, 시제=현재, 문장의 궁부정=긍정, 양상=추측}으로 같은 패턴을 갖게 되고 중의성을 남겨두게 된다.

- (1) “내 생각도 그의 생각과 같다.” (진술)
- (2) “그가 아픈 것 같다.” (추측)

따라서 다음 연구는 이러한 구문 정보와 함께 n-gram에 기반을 둔 형태소-품사 쌍과 품사 바이그램을 같이 사용하는 방향으로 진행되었다[6,7]. 하지만 구문 정보로 인한 중의성 문제는 해결되지 않았고, 이 문제를 해결하기 위해 구문 정보를 사용하지 않는 화행 분류 연구가 진행되었다[8,9].

구문 정보를 사용하지 않음으로 인해 구문 파서의 오류 전파 문제나 구문 정보가 발화를 불완전하게 표현하는 문제는 해결되었지만 다른 문제가 발생하였다. 예를 들어, 동사 ‘고맙’은 감시 화행을 결정하는데 매우 중요한 단서가 된다. 하지만 이 동사가 안긴문장이나 명사를 수식하는 구에 들어갔을 경우 화행의 결정에 영향을 미

치게 되므로 화행을 잘못 분류하게 된다. 다음 (4)번 예제 문장의 경우 ‘고맙’은 ‘사람’을 수식하며, (5)번 예제 문장의 ‘고맙’은 안긴문장의 서술어이다. 하지만 형태소-품사 쌍과 품사 바이그램만을 사용한 화행 분류에서는 위의 경우 모두 ‘고맙’이라는 동사가 들어가 있기 때문에 감시의 화행으로 분류하게 된다. 이는 형태소-품사 쌍과 품사 바이그램이 형태소 단위로 한두 개의 문맥만을 보기 때문에 생기는 문제이다.

(3) “고맙습니다.” (감사)

(4) “그는 고마운 사람이야.” (진술)

(5) “그에게 고맙다고 말해줘.” (요청)

한국어가 형식 형태소와 실질 형태소로 구성된다는 점과 화행의 결정에 동사, 어미 등의 특정 형태소가 중요한 역할을 한다는 점을 이용하여 마지막 어절과 동사, 어미 등의 특정 형태소를 내용 자질과 형식 자질로 나누어 화행 분류의 자질로 이용한 연구[10]가 있다. 이 연구에서는 어미, 보조용언과 같은 특정 형태소를 이용하여 구문 정보를 보다 유연성 있게 표현할 수 있었다. 하지만 이 논문에서 주장한 내용 자질과 형식 자질의 구분이 모호하고, 자질로 사용한 형태소에 제한을 두어 명사 등의 형태소가 화행 분류에 유용할 수 있음에도 제외하였다는 문제가 있다. 예를 들어, 다음 예제문장 (6)은 [10]의 오류 분석에서 지적한 ‘-겠’ 선어말어미를 사용한 문장이다. 이 경우 ‘예약’이라는 명사는 추측이나 회빙 발화보다는 의지표명/약속을 나타내는 발화에서 많이 발생하는 단어이다. 따라서 ‘예약’은 이 발화를 화행분류를 하는데 중요한 역할을 한다. 하지만 이 연구에서는 명사를 전혀 사용하지 않고 있으므로 이러한 정보를 사용할 수 없다.

(6) “그 방으로 예약을 하겠습니다.” (의지표명/약속)

본 논문에서는 n-gram 유형 자질과 마지막 어절 및 특정 형태소 어휘를 이용한 자질을 같이 사용할 경우 서로의 단점을 보완해줄 수 있을 것이라는 가정 하에 각 자질에 대한 비교 평가를 수행하고 최적의 자질 조합을 찾아본다.

3. 한국어 대화의 화행 분류를 위한 자질

본 논문에서는 한국어 화행 분류를 위해 다음 자질들을 비교한다. 이전의 마지막 어절 및 특정 형태소를 사용한 연구에서는 이 외에도 의문사 사용 여부와 문장부호를 자질로 사용하였다[10]. 하지만 이 두 가지 자질은 형태소-품사 쌍 자질과 겹치기 때문에 본 연구에서는 제외하였다. 본 논문에서는 자질의 특징에 따라 다섯 개의 자질 군으로 나누었으며, n-gram 기반 자질, 마지막 서술어부 자질, 형식 형태소 열/집합 자질, 표층 자질, 문맥 자질로 구성된다.

3.1 n-gram 기반 자질

본 연구에서는 이전 한국어 화행 분류 연구에서 n-gram 기반 자질로 사용한 형태소-품사 쌍과 품사 바이그램을 그대로 사용한다.

형태소-품사 쌍 형태소 분석, 품사 태깅 결과에서 얻을 수 있는 어휘-품사의 쌍이다. 본 논문에서는 세종 품사 집합[18]을 사용하였으며 자동 형태소 분석 및 품사 부착기[19]를 사용하여 형태소 분석 및 품사 부착을 수행하였다. 형태소 분석기 평가를 위해 100건의 발화에 대해 수작업으로 평가를 한 결과 어절 단위 정확도 94.9%의 성능을 보였다. 다음 (8)은 예제 문장 (7)을 형태소 분석, 품사 태깅한 결과이다. 이를 형태소-품사 쌍으로 표현하면 {나/NP, 의/JKG, 말/NNG, 을/JKO, 못/MAG, 믿/VV, 을/EF, /SF}과 같이 표현된다.

(7) “내 말을 못 믿을걸.” (추측)

(8) 나/NP+의/JKG 말/NNG+을/JKO 못/MAG 믿/VV+을/EF+./SF

품사 바이그램 형태소 분석, 품사 태깅 결과인 품사로 이어지는 품사-품사 쌍을 의미한다. 위 예제 문장 (7)을 품사 바이그램으로 표현하면 {NP-JKG, JKG-NNG, NNG-JKO, JKO-MAG, MAG-VV, VV-EF, EF-SF}과 같이 표현된다.

3.2 마지막 서술부 자질

영어는 주어와 서술어 등의 위치가 정해져 있고 주어-동사-목적어의 순서로 나오는데 비해 한국어는 비교적 어순이 자유로우며 기본적으로 주어-목적어-동사의 순서를 갖는다. 따라서 한국어는 마지막까지 들어서 그 상황을 알아야 내용을 파악할 수 있는 특성을 지닌 언어이며, 마지막 서술부가 문장 전체를 해석하는데 가장 중요한 역할을 한다. 이러한 한국어의 특징을 반영할 수 있는 다음과 같은 자질들을 마지막 서술부 자질 군으로 분류한다.

마지막 어절 한국어에서 마지막 서술부는 예제 문장 (9)에서 “감사합니다.”와 같이 마지막 어절 내에 주로 위치한다. 또한 아래 예제 문장 (10)과 같이 “응.,” “설마.,” “알았어.”와 같이 하나의 어절로 구성된 발화는 발화 자체가 화행을 분류하는데 중요한 자질이 된다. “말했다.,” “설마” 등과 같이 형태소 분석을 하지 않은 마지막 어절을 자질로 사용한다.

(9) “도와주셔서 감사합니다.” (감사)

(10) “응.” (긍정대답)

마지막 동사 하나의 발화 내에 여러 개의 동사가 나올 경우 마지막 동사를 자질로 사용한다. 예를 들어 다음 예제 문장 (11)과 (12)에서 “생각하다”와 “부탁하다”는 각각 주장, 요청 화행을 구분하는 중요한 단서가 된다.

(11) “네 말도 일리가 있다고 생각해.” (주장)

(12) “존슨씨 좀 부탁합니다.” (요청)

마지막 부사 “아마”, “좀”과 같은 부사들은 동사와 함께 화행 결정에 큰 영향을 미친다. 위 예제 문장 (12)에서와 같이 “좀”은 “부탁하다” 등의 동사와 함께 요청 화행의 발화에서 빈번하게 발생하며, 아래 문장 (13)과 같이 “아마”는 “~었나 보다”나 “인 것 같다” 등의 표현과 함께 추측 화행의 발화에 많이 발생한다. 이와 같이 부사가 서술어와 같이 나오는 경우가 많은 것을 고려하여 부사를 마지막 서술부 자질로 분류한다.

(13) “아마 감기 걸렸나 보군요.” (추측)

마지막 보조 용언 보조 용언은 양상과 연관 지어져 발화자의 의도를 보충한다. 위 예제 문장 (13)에서의 “~려나 보다”의 “보다”와 아래 예제 문장 (14)에서 “~ㄴ/ㄹ 것 같다”의 “같다”는 추측 화행을 구분하는 중요한 단서가 된다.

(14) “그가 올 것 같다.” (추측)

어미 종결어미, 연결어미, 선어말어미 등은 용언의 쓰임을 보완해준다. 이에 연결어미, 종결어미, 선어말어미 중 가장 마지막에 나온 것을 자질로 사용한다. 특히 선어말 어미는 시제나 높임 등을 나타내는데 예를 들어 다음 예제 문장 (15)에서 선어말어미 ‘-겠’은 화자의 의지표명/약속을 나타낸다.

(15) “오늘은 내가 밥값을 내겠다.” (의지표명/약속)

3.3 형식 형태소 열/집합 자질

한국어에서는 형식 형태소는 개별로도 쓰이지만 여러 형태소가 한 발화 내에서 같이 사용되어 양상 등을 나타내는 경우도 있다. 이에 본 연구에서는 한 발화 내에서 같이 쓰인 형식 형태소나 품사의 열, 혹은 집합 자질을 화행 분류의 자질로 사용한다.

형식 형태소 어휘 열 발화 내의 의존 명사, 접미사, 조사, 감탄사, 서술격조사, 보조용언, 어미들의 형태소 열은 전체 문장의 뼈대를 보여줄 수 있다. 다음 예제 문장 (16)에서 “-이-려나-보다”는 추측 화행의 발화에서 나타난다.

(16) “눈이 오려나 보다” (추측)

형식 형태소 어휘 집합 한국어는 어순이 자유롭기 때문에 발화 내에서 형식 형태소가 나타나는 순서도 자유롭기 때문에 형식 형태소 열보다 형식 형태소 집합이 나올 수 있다. 예를 들어, 예제문장 (17)과 (18)는 모두 화자의 의지를 표명하는 발화이다. 하지만 이를 형식 형태소 열로 표현할 경우 ‘에게-을-고-싶-어’와 ‘을-에게-싶-어’로 다르게 표현이 된다. 따라서 한 발화 내에 나타난 형식 형태소를 순서에 상관없이 자질로 사용할 수 있도록 형식 형태소 집합을 자질로 제안하고 실험을 통해 형식 형태서 어휘 열과 비교한다.

(17) “너에게 선물을 주고 싶어.” (희망)

(18) “선물을 너에게 주고 싶어.” (희망)

형식 형태소 품사 열, 형식 형태소 품사 집합 형식 형태소 중 주격 조사 ‘은’, ‘는’, ‘이’, ‘가’와 같이 동일한 역할을 하면서 이전에 오는 형태소의 음운에 따라 다른 형태소를 쓰는 경우가 있다. 이러한 점을 반영하여 형식 형태소 품사 열과 형식 형태소 품사 집합 또한 형식 형태소의 어휘를 사용한 자질과 비교한다.

3.4 표층 자질

주장, 요청 등의 발화의 주어는 발화의 화행을 결정하는데 있어 중요하며, 발화의 평균 길이는 화행에 따라 다르다[16]. 주어와 길이는 다른 발화 자질들과는 달리 특정 형태소 보다는 발화의 표면적인 특성에 중점을 두었으므로 [16]과 같이 따로 표층 자질 군으로 분류한다.

주어 주어가 화행 분류에 결정적인 역할을 할 경우가 있다. 예를 들어, 다음 예제 문장 (19)와 같이 주장이나 의지를 표명하는 발화에서는 1인칭 주어가 많이 쓰고 예제 문장 (20)과 같이 상대방에게 요청하거나 명령하는 발화에서는 2인칭 주어를 많이 쓴다.

(19) “나는 긍정적으로 생각해.” (주장)

(20) “긍정적으로 생각해.” (명령)

길이(어절 수), (S,M,L) 화행에 따라 평균 발화의 길이는 다르다[16]. 감탄, 긍정 대답 등의 발화의 길이는 상대적으로 짧고, 주장, 희망 등의 발화의 길이는 상대적으로 짧다. 본 논문에서는 어절 수를 그대로 자질로 사용한 것과, 어절 수를 기준으로 S, M, L 3단계로 나누는 자질을 각각 사용하여 비교한다. 3단계로 나누는 것은 학습 집합 말뭉치 내의 어절 수 단위 길이 분포를 고려하여 나누었으며, 한 단어로 된 발화, 여섯 단어 미만, 여섯 단어 이상으로 세 단계로 나누었다.²⁾

3.5 문맥 자질

화행은 대화 문맥 내에서 나타나는 화자의 의도이므로 같은 발화도 주변 문맥에 따라 다른 화행으로 분류할 수 있다. 따라서 대상 발화 자질 뿐만이 아니라 대상 발화 주변 문맥을 반영할 수 있는 문맥 자질을 같이 사용한다.

상대방 발화 화행 아래의 대화 예제에서와 같이 요청 다음에는 수락이나 거절이 오는 경우가 많고 질문 뒤에는 긍정대답 등이 오는 경우가 많다. 이와 같이 이전 발화에 의해 다음에 올 발화를 예측할 수 있는 경우가 있다[20]. 이는 이전 발화가 다음 발화에 영향을 미치는 것을 나타낸다.

자기 발화 화행 그림 2의 대화 예제에서와 같이 요

2) 학습 집합 말뭉치 9,711 발화 중 한 단어로 된 발화는 2,571개, 두 단어 이상 여섯 단어 미만인 발화는 각각 1,000개 이상 2,000개 미만, 여섯 단어 이상인 발화는 각각 1,000개 미만의 분포로 나타났다.

A: 도넛 같은 것 좀 사다 주실래요? (요청)
 B: 그렇지요. (수락)
 A: 고마워요. (감사)

그림 2 대화 예제

A: 차네 이 차 산거야? (확인질문)
 B: 그림. (긍정대답)
 B: 주문하고 두 달이나 기다리고 나서야 어제 왔어. (진술)

그림 3 대화 예제

칭이나 제인 뒤에 상대방의 수락 발화가 올 경우 그에 대한 감사가 오는 경우가 있다. 혹은 다음 그림 3의 대화 예제와 같이 턴을 상대방에게 넘기지 않고 긍정 대 답 뒤에는 그에 대한 부연 설명을 하기 위한 진술이 올 수 있다. 이는 이전 자기 발화 화행도 다음의 발화에 영 향을 미칠 수 있음을 보여준다.

발화의 위치 만남 인시 발화는 대화의 시작 부분에 많이 나타나며 헤어짐 인시 발화는 대화의 마지막 부분 에 많이 나타난다. 긍정대답, 부정대답과 같은 발화는 이전에 온 다른 발화에 대한 대답이기 때문에 대화의 맨 처음에 나올 수 없다. 이와 같이 화행에 따라 대화 내에서 발화가 나오는 위치가 연관이 있는 경우가 있기 때문에 각 화자 별로 대화의 시작(B1,B2)과 대화의 내 부(I)의 세 가지로 분류하였다.

4. 실험 및 평가

4.1 실험 말뭉치

본 논문에서는 도메인에 비종속적인 한국어 대화 말뭉치[10]를 이용하여 화행 분류 실험을 수행하였다. 이 말뭉치는 외국어 회화의 한글 해석 부분을 사용한 것으로 번역을 거친 것이지만 구어체이기 때문에 자연스러운 문장으로 이루어져 있어 한국어의 특성을 반영한 자질을 사용하는데 문제가 없었다. 1,671개의 대화, 11,207개의 발화로 구성되어 있으며 수작업을 통해 37개의 화행 분류 체계 중 하나로 화행을 부착한 말뭉치이다. 서론의 그림 1은 화행 정보가 부착된 말뭉치의 일부분이다. 본 연구에서는 이 중 9,711 발화를 학습에 이용하였으며 나머지 1,496 발화를 실험에 사용하였다.

4.2 실험 환경

본 논문에서는 위 자질들을 지지벡터기계와 결합하여 화행 분류의 성능을 실험해 보았다. 지지벡터기계의 학습을 위한 자질은 각각의 자질들의 가중치로 벡터를 구성하였다. 본 연구에서는 이진 가중치를 적용하여 해당 자질이 주어진 문장에서 출현했을 경우 1을, 출현하지 않았을 경우에는 0을 가중치로 부여하였다. 지지벡터기계는 이진 분류기이므로 각각의 화행을 위한 지지벡터기계를 따로 학습하고 실험하여 출력 값이 가장 큰 지

지벡터기에 해당하는 화행을 주어진 발화의 화행으로 최종적으로 결정한다. 본 연구에서는 SVMlight[21]를 이용하였고 선형 커널을 이용하여 학습하였다.

추가로 실행한 자질선택기법 실험에서는 여러 자질 선택 기법 중 이전 연구에서 한국어 화행 분류에 가장 높은 성능을 보인 카이제곱 통계량을 사용하였다[8]. 카이제곱 통계량은 문서 분류 분야에서 자질 선택 기법으로 많이 사용되며 본 논문에서는 자질 f 와 화행 s 간의 정보량을 측정하기 위해 다음 수식과 같이 계산하였다.

$$\chi^2(f,s) = \frac{(A+B+C+D) \times (AD-CB)^2}{(A+C) \times (B+D) \times (A+B) \times (C+D)}$$

- A: 자질 f 를 가지고 화행이 s 인 발화 수
- B: 자질 f 를 가지고 화행이 s 가 아닌 발화 수
- C: 자질 f 를 가지지 않고 화행이 s 인 발화 수
- D: 자질 f 를 가지지 않고 화행이 s 가 아닌 발화 수

각 자질에 카이제곱 통계량을 부여하는 방법은 평균 값을 이용하는 방법과 최대값을 이용하는 방법이 있다. 본 논문에서는 이전 연구에서 좋은 성능을 보인 최대값을 이용하는 방법을 사용하였다[8].

실험 결과에 대한 평가 척도로는 정확도를 사용하였으며 다음 수식과 같이 계산하였다.

$$\text{정확도(\%)} = \frac{\text{올바르게 화행분류된 발화수}}{\text{전체 실험에 사용한 발화수}} \times 100$$

4.3 실험 결과 및 분석

본 논문에서는 각 자질이 한국어의 화행 분류에 미치는 영향을 살펴보기 위해 다음과 같이 실험을 하였다. 우선, 각 자질을 개별적으로 사용하여 화행 분류를 하여 보고 화행 분류에서 가장 널리 쓰이는 n -gram 자질을 기저 성능으로 하여 다른 자질들을 추가해보는 실험을 하였다.

그리고 발화 자질만을 대상으로 모든 자질을 사용한 것을 기저 성능으로 하여 개별 자질을 제거하고 실험하여 성능을 살펴본 후 가장 좋은 자질 조합을 선정하였다. 이렇게 선정한 발화 자질 조합을 대상으로 각 문맥 자질을 추가해서 역시 최적의 성능을 보이는 자질 조합을 최종적으로 선정하였다.

마지막으로 자질 선택 기법을 사용하여 사용하는 자질의 수를 제한하였을 때의 성능 변화를 살펴보는 것으로 마무리하였다.

표 1은 각 자질을 개별적으로 사용한 화행 분류 실험 결과를 정확도를 기준으로 내림차순으로 정렬한 표이다. 자질 개수는 학습 집합 9,711 발화에서 출현한 각 자질 값의 개수이다. 본 연구에서 실험한 결과는 '지지벡터기계' 열이며 '최대 엔트로피' 열은 이전 연구[10]에서 최대 엔트로피 모형을 사용하여 같은 말뭉치에 같은 자질을 사용한 화행 분류 실험 결과이다. '-'표시가 된 부분

표 1 각 자질별 화행 분류 실험 결과 (정확도, %)

사용한 자질	지지벡터기계	최대엔트로피	자질 개수
형태소-품사 쌍	83.22	-	5,395
마지막 어절	64.17	53.63	3,706
형식 형태소 어휘 열	56.89	52.23	5,904
형식 형태소 어휘 집합	53.61	-	5,749
품사 바이그램	49.26	-	533
형식 형태소 품사 집합	43.58	-	1,561
형식 형태소 품사 열	42.65	-	2,362
이전 자기 발화 화행	34.22	33.28	37
어미	33.96	48.66	347
이전 상대방 발화 화행	27.54	33.00	35
마지막 용언	22.66	40.03	624
마지막 부사	20.59	35.70	326
주어	20.25	33.23	4
발화의 길이 (S,M,L)	11.50	33.71	3
대화 내 발화의 위치	7.09	33.26	3
마지막 보조용언	6.55	36.10	31
발화의 길이 (어절 수)	1.20	-	16

은 이전 연구에서 사용하지 않은 자질이다. 두 통계 기반 학습 방법의 결과를 통해 본 연구에서 사용하는 자질이 특정 학습 방법에 종속적이지 않음을 확인할 수 있다.

표 1을 통해 형태소-품사 쌍만을 사용했을 때도 83%가 넘는 성능을 보여 매우 많은 정보를 제공한다는 것을 알 수 있다. 또한 품사 바이그램은 그에 비해 현저하게 떨어지는 성능을 보였으며 오히려 마지막 어절을 사용했을 때 성능이 높다. 형식 형태소 열/집합 자질 군에서는 어휘를 사용한 것이 전반적으로 품사를 사용한 것보다 좋은 성능을 보이며, 길이 자질은 그룹화 하여 사용한 것이 어절 수를 그대로 사용한 것보다 성능이 월등히 좋다.

표 2는 각 자질이 다른 자질과 결합하여 얼마나 보완해주는지를 보기 위해 최근의 한국어 화행 분류 연구 [8,9]에서 사용한 형태소-품사 쌍과 품사 바이그램을 기저성능으로 하여 다른 자질들을 추가해본 화행 분류 실험 결과이다.

품사 바이그램을 형태소-품사 쌍과 함께 사용한 것이 형태소-품사 쌍을 단독으로 사용한 것과 크게 차이가 나지 않는다는 점에서 기존 한국어 화행 분류 연구에서 기본 자질로 쓰인 품사 바이그램이 형태소-품사 쌍과 같이 사용했을 때 많은 정보를 제공하지 않는다는 것을 알 수 있다.

주어 자질을 사용했을 때 기저 성능보다 정확도가 낮다. 이는 주어의 생략이 빈번한 한국어의 특징[22]이 반영된 것으로 볼 수 있다. 실제 말뭉치에서도 학습 집합 9,711 발화 중 7,793 발화(80.15%)에 주어 생략되었고

표 2 n-gram 자질과 다른 자질 결합 실험 결과

사용한 자질	정확도(%)
형태소-품사 쌍 + 품사 바이그램 (①)	83.42
①+마지막 어절	84.89
①+형식 형태소 어휘 집합	84.22
①+형식 형태소 어휘 열	84.02
①+발화의 길이(S,M,L)	83.96
①+이전 상대방 발화 화행	83.96
①+마지막 용언	83.82
①+마지막 보조용언	83.76
①+발화의 길이(어절 수)	83.76
①+어미	83.62
①+대화 내 발화의 위치	83.62
①+이전 자기 발화 화행	83.56
①+마지막 부사	83.16
①+주어	83.16
①+형식 형태소 품사 열	83.09
①+형식 형태소 품사 집합	83.09

실험 집합 1,493 발화 중 1,134 발화(75.95%)에 주어 생략되었다.

또한 위 표 1에서는 형식 형태소 어휘 열이 형식 형태소 어휘 집합보다 성능이 좋으나, 이 자질들을 n-gram 자질과 결합하였을 때는 성능이 비슷하게 되는 것을 볼 수 있다. 실제 형식 형태소 어휘 열로 표현했을 때 자질은 5,904개로 표현되었으나 형식 형태소 어휘 집합으로 표현했을 때는 722개가 적은 5,182개로 표현되었다. 이는 한국어의 자유 어순 특성 때문에 형식 형태소 열로 표현했을 때 다양하게 표현이 되지만 이것을 집합으로 표현했을 때 거의 성능의 변화가 없이 자질 수를 많이 줄일 수 있기 때문에 군집화와 같은 효과를 낼 수 있음을 보여준다. 본 논문에서는 각 자질의 평가를 하기 위해 이후 실험에서는 형식 형태소 어휘 집합과 형식 형태소 어휘 열을 따로 사용하여 실험 결과를 비교한다.

위 실험 결과를 토대로 이후 실험에서는 형식 형태소 품사 열과 형식 형태소 품사 집합, 어절 수를 그대로 사용한 발화의 길이는 배제하고 형식 형태소 어휘 열 혹은 집합과 3단계로 나눈 발화의 길이 자질을 사용한다. 표 3과 표 4는 모든 발화 자질을 사용하고 한 자질만을 제외한 화행 분류 실험 결과이다.

형식 형태소 어휘 집합과 형식 형태소 어휘 열은 사용한 형태소는 동일하며 순서를 고려하였는가의 여부만 차이가 난다. 이에 본 논문에서는 한국어의 자유 어순 특성으로 두자질 개수가 차이가 나지만 표 2에서 성능이 비슷하게 나타났으며, 표 3, 표 4를 통해 이를 다른 자질들과 함께 사용했을 때도 역시 성능이 비슷함을 알 수 있다. 또한, 표 3과 표 4에서 공통적으로 품사 바이

표 3 모든 발화 자질을 사용하고 한 자질만 제외한 화행 분류 실험 결과(형식 형태소 어휘 집합 사용 시)

사용한 자질	정확도(%)
모든 자질 사용 (②)	85.56
②-형태소-품사 쌍	75.40
②-품사 바이그램	85.76
②-마지막 어절	84.22
②-마지막 용언	84.96
②-마지막 보조용언	85.70
②-마지막 부사	85.23
②-어미	85.16
②-형식 형태소 어휘 집합	85.49
②-주어	85.36
②-발화의 길이 (S,M,L)	85.56

표 4 모든 발화 자질을 사용하고 한 자질만 제외한 화행 분류 실험 결과(형식 형태소 어휘 열 사용 시)

사용한 자질	정확도(%)
모든 자질 사용 (③)	85.23
③-형태소-품사 쌍	75.73
③-품사 바이그램	86.03
③-마지막 어절	84.43
③-마지막 용언	85.10
③-마지막 보조용언	85.76
③-마지막 부사	85.50
③-어미	85.36
③-형식 형태소 어휘 열	85.49
③-주어	85.43
③-발화의 길이 (S,M,L 그룹)	85.76

그램을 제외하였을 때 가장 좋은 성능을 보인 것을 알 수 있다. 특히 형식 형태소 어휘 열을 사용했을 때 더 높은 성능 향상을 보이는데 이는 품사 바이그램과 형식 형태소 어휘 열이 중복되어 서로 상충되는 것으로 해석할 수 있다.

기존 연구[16]에서는 최적의 자질 조합을 찾기 위해 자질을 하나씩 추가하면서 제일 좋은 자질조합을 찾아가는 방법을 사용하였다. 본 논문에서는 위 실험 결과를

표 5 문맥 자질 추가 실험 결과

사용한 자질	정확도(%)
선택한 발화 자질(④)	86.23
④+상대방 발화 화행	86.30
④+자기 발화 화행	85.70
④+발화의 위치	85.56
④+상대방 발화 화행+자기 발화 화행	86.56
④+상대방 발화 화행+발화의 위치	86.36
④+자기 발화 화행+발화의 위치	85.63
④+모든 문맥 자질	85.96

기반으로 한 개 자질을 제외하였을 때 더 높은 성능을 보인 것이 있을 경우 해당 자질을 제외하고 다시 다른 자질을 한 개씩 제외하는 추가 실험을 통해 가장 좋은 발화 자질 조합을 찾아보았다. 품사 바이그램과 마지막 보조용언, 주어 제외된 모든 발화 자질을 사용하였을 때 가장 좋은 성능을 보이며, 형식 형태소 집합을 사용하였을 때 86.16%, 형식 형태소 열을 사용하였을 때 86.23%로 성능의 차이가 미미하다. 이는 앞서 언급한 바와 같이 한국어의 자유 어순에 의한 영향으로 해석할 수 있다.

위 표 5는 앞서 선택한 대상 발화 자질 조합에 문맥 자질을 추가한 실험 결과이다. 발화 자질 조합으로 사용한 자질은 형태소-품사 쌍, 마지막 어절, 마지막 용언, 마지막 부사, 어미, 형식 형태소 어휘 열, 발화의 길이 (S,M,L)이다. 위 발화 자질 조합과 문맥 자질 중 자기 발화 화행과 상대방 발화 화행을 사용하였을 때 가장 높은 성능을 보임을 알 수 있다.

아래 표 6은 실험 말뭉치에서 10회 이상 출현한 화행들과 그것의 분류 과정에서 가장 빈번하게 발생한 오류이다. 괄호 안의 숫자는 해당 오류 화행으로 분류된 발화의 개수이다. 감시 화행의 경우 “감사합니다.”나 “고맙습니다.”와 같이 전형적인 형태로 나타나므로 오류가 발생하지 않았다. 가장 빈번하게 발생한 오류는 진술 화행으로 잘못 분류된 오류이다. 이는 진술 화행이 전체 말뭉치의 1/3을 차지하고 다른 화행이 말뭉치 내에서

표 6 화행 분류 오류 분석

화행(출현 발화 개수)	정확률(%)	가장 빈번하게 나타난 오류	화행(출현 발화 개수)	정확률(%)	가장 빈번하게 나타난 오류
감사(17)	100.00	-	명령(56)	73.21	진술(10), 사과(2)
확인질문(164)	97.56	진술질문(3)	희망(23)	69.57	진술(4), 격려(3)
부정(18)	94.44	진술(1)	추측(112)	66.07	진술(27), 의지표명/약속(8)
진술(576)	93.92	추측(9), 명령(5)	의지표명/약속(46)	65.22	추측(7), 진술(4), 명령(3)
긍정(48)	91.67	진술(2)	주장(18)	61.11	진술(7)
진술질문(192)	91.15	확인질문(16)	요청(16)	56.25	호출(2), 확인질문(2)
호출(54)	88.89	확인질문(4), 진술(2)	제안(21)	47.62	확인질문(4), 진술(4)
단순감탄(37)	86.49	찬동(2)	전언(11)	27.27	진술(6)
찬동(29)	82.76	진술(3)			

상대적으로 적게 발생하여 생기는 문제로 볼 수 있다. 이 문제를 해결하기 위해서는 많이 발생하지 않는 화행을 보다 일반적인 화행에 병합시키고 말뭉치의 양을 늘리는 것이 필요하다.

본 연구에서 사용한 자질 중 형태소-품사 쌍, 품사 바이그램, 형식 형태소 집합과 같은 자질들은 학습 집합의 크기에 따라 자질의 개수가 많아지게 된다. 자질 선택 기법을 사용하면 자질이 지나치게 많아 학습 시간이 많이 소요될 경우 적절한 양만을 사용하여 학습 시간을 줄일 수 있으며, 자질 선택 기법으로 사용하는 자질의 양을 조절하여 과적응을 방지할 수 있다. 또한 자질 선택 기법으로 부여한 점수를 이용하여 상대적으로 유용한 자질을 알아보는데 사용할 수도 있다.

이전 연구[8]에서는 자질 선택 기법 중 Information Gain, 상호정보량(Mutual Information), 카이 제곱 통계량을 한국어 화행 분류에 적용하여 보았으며, 세 가지 중 카이제곱통계량이 가장 좋은 성능을 기록하였다. 이에 본 논문에서는 카이 제곱 통계량을 이용하여 자질의 개수를 제한하였을 때의 성능을 측정하였으며 전체 자질 조합과 위 표 5에서 가장 높은 정확도를 보인 자질 조합(형태소-품사 쌍, 마지막 어절, 마지막 용언, 마지막 부사, 어미, 형식 형태소 어휘 열, 발화의 길이(S,M,L), 상대방 발화 화행, 자기 발화 화행)을 대상으로 실험하였다. 그 결과는 표 7, 그림 4와 같다. 전체 자질 중 4% 이상을 썼을 때 가장 좋은 성능의 95% 이상을 기록하였다.

기존 연구 [6]에서는 자질 개수를 적절한 수로 선택했을 때 최고의 성능을 보였으나, 본 논문에서는 모든 자질을 사용했을 때는 전체 자질을 사용하는 것보다 상위 40%의 자질을 사용하는 것이 더 성능이 좋으며, 최고 성능 자질 조합에서는 상위 50% 이상 쓸 경우 전체 자질을 사용한 것과 성능이 같다. 이는 자질 조합을 구성하면서 화행 분류에 필요없는 자질을 적절하게 삭제한 것으로 볼 수 있다. 또한, 기존 연구 [6]에서는 전체 자질의 개수가 4,000개 정도이나 본 논문에서 사용한 자질

표 7 자질 선택 기법을 사용하여 일부 자질만 사용했을 때의 성능 변화(정확도, %)

사용한 자질 비율	모든 자질 사용	최고 성능 자질 조합
상위 100%	85.70	86.56
상위 90%	85.76	86.50
상위 80%	85.90	86.36
상위 70%	85.76	86.43
상위 60%	85.90	86.50
상위 50%	85.90	86.56
상위 40%	86.10	86.50
상위 30%	85.56	86.23
상위 20%	85.76	85.90
상위 10%	85.76	85.70
상위 9%	85.29	85.23
상위 8%	85.43	84.96
상위 7%	85.10	84.96
상위 6%	84.83	84.96
상위 5%	84.76	84.76
상위 4%	83.29	82.42
상위 3%	81.15	79.55
상위 2%	76.60	75.87
상위 1%	64.57	63.37

은 16,948개(표 8)로 전체 자질의 개수가 4배가 넘고 카이 제곱 통계량에 따른 분포(표 8, 그림 5)에서 알 수 있듯 여러 자질군의 자질이 고루 상위에 분포함을 알 수 있다.

위 실험 결과를 통해 이후 보다 큰 말뭉치를 학습에 이용하여 학습 시간이 많이 소요될 경우 사용하는 자질의 양을 제한하여 학습 시간을 줄여도 성능에 크게 영향을 미치지 않을 것임을 예측할 수 있다.

표 8과 그림 5는 표 6과 같이 전체 자질에 대해 카이 제곱 통계량을 계산한 결과로 다시 각 자질 내의 비율을 계산한 결과를 나타낸 것이다. 예를 들어, 맨 처음 막대는 n-gram 자질(형태소-품사 쌍과 품사 바이그램) 중 전체 자질의 상위 5%에 속하는 것이 전체 n-gram

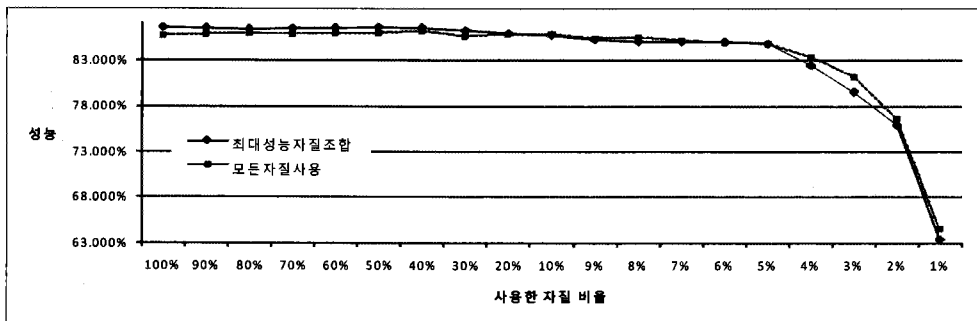


그림 4 자질 선택 기법을 이용하여 일부 자질을 사용했을 때의 성능

표 8 자질 선택 기법의 값에 따른 자질의 분포

	n-gram 자질	마지막 서술부 자질	형식 형태소 열 자질	표층 자질	문맥 자질	개수
상위 0~10%	9.2%	12.7%	8.2%	71.4%	26.7%	1694
상위 10~20%	9.0%	8.9%	11.9%	14.3%	13.3%	1695
상위 20~30%	10.6%	10.6%	8.6%	14.3%	25.3%	1695
상위 30~40%	9.8%	9.4%	10.7%	0.0%	10.7%	1695
상위 40~50%	10.6%	8.9%	10.4%	0.0%	12.0%	1695
상위 50~60%	9.0%	9.2%	11.8%	0.0%	1.3%	1694
상위 60~70%	15.8%	7.8%	6.1%	0.0%	8.0%	1695
상위 70~80%	2.5%	30.2%	0.4%	0.0%	0.0%	1695
상위 80~90%	9.6%	1.7%	17.7%	0.0%	1.3%	1695
상위 90~100%	13.8%	0.7%	14.2%	0.0%	1.3%	1695
개수	5928	5034	5904	7	75	16948

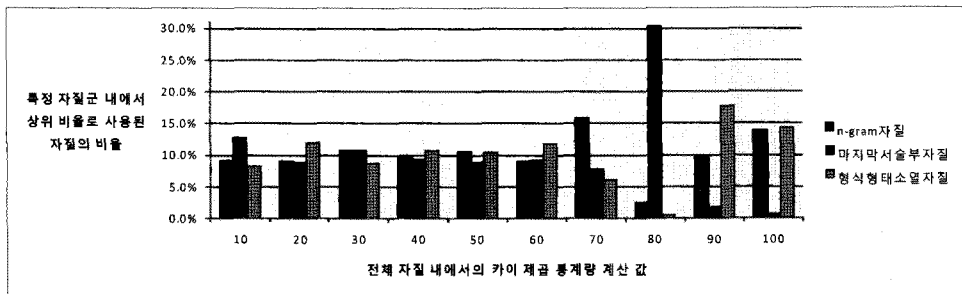


그림 5 자질 선택 기법의 값에 따른 자질의 분포

자질 중에 얼마나 되는가를 본 것으로, 상위에 많이 분포한 자질이 화행 분류에 유용한 자질이라고 말할 수 있다. 문맥 자질과 표층 자질의 경우 전체 자질의 개수가 다른 자질 군에 비해 매우 적기 때문에 그림 5에서는 자질의 개수가 5,000이 넘는 n-gram 자질, 마지막 서술부 자질, 형식 형태소 열 자질만을 표현하여 성향을 비교하였다. 그림 5를 통해 특히 마지막 서술부의 경우 대체적으로 기본자질보다 상위에 분포하는 것으로 마지막 서술부 자질이 유용하다는 것을 확인할 수 있다.

본 연구는 다음과 같은 한계를 지니고 있다. 제안하는 자질 중 마지막 서술부 자질의 경우 형태소의 위치 정보를 포함한 것이긴 하나 마지막 용언 등의 자질은 형태소-품사 쌍과 중복이 될 수 있다. 본 연구에서는 n-gram 자질을 기본으로 마지막 어절 및 특정 형태소를 이용한 자질을 추가하여 각 자질이 가지는 문제들 서로 보완해주는지 여부를 확인하고자 실험을 진행하였으므로 이 문제는 본 연구에서 다루지 않았다. 하지만 이러한 한계는 추후 연구에서 좀 더 정밀한 자질 선별 작업을 통해 개선되어야 할 것으로 판단된다.

5. 결론

본 논문은 지지벡터기계 학습기법을 이용하여 구현한 화행 분류 시스템을 통해 통계기반 한국어 대화의 화행

분류를 위한 n-gram 유형 자질과 마지막 어절 및 특정 형태소 어휘를 사용한 자질을 비교 평가하고, 한국어 화행 분류에 가장 적합한 자질 조합을 선별하였다.

개별 자질의 화행 분류 성능 실험 결과를 통해 형태소-품사 쌍, 마지막 어절, 형식 형태소 어휘 열/집합 등의 순으로 유용한 것을 알 수 있었다. 그리고 자질 결합 실험 결과, 품사 바이그램과 주어, 마지막 보조 용언, 발화의 위치를 제외한 자질을 모두 사용했을 때 가장 좋은 성능을 보였다.

또한, 이전 연구[10]에서 최대 엔트로피 모델을 사용한 실험 결과와의 비교를 통해 본 연구에서 사용한 자질이 특정 학습 방법에 종속적이지 않으며 특히 마지막 어절, 형식 형태소 어휘 열과 같은 자질은 학습 방법에 상관없이 화행 분류에 유용함을 알 수 있었다.

형태소-품사 쌍 자질은 개별 자질만으로도 83%가 넘는 높은 정확도를 보였으나 볼 수 있는 문맥이 지나치게 짧기 때문에 생기는 문제가 있다. 마지막 어절과 형식 형태소 어휘 열/집합 자질은 문장의 구문정보를 표현할 수 있으나 사용하지 않는 형태소의 정보를 보지 못하는 단점이 있다. 형태소-품사 쌍 자질과 다른 자질들을 같이 사용할 때 성능향상이 있는 것을 통해 이 자질들이 서로 보완함을 알 수 있었다.

또한, 한국어의 특성이 화행 분류에 미치는 영향을 실

험을 통해 확인할 수 있었다. 마지막 어절, 마지막 용언 등의 자질이 유용하다는 것을 통해 한국어의 화행 분류에서 마지막 서술부가 유용하다는 점을 알 수 있었으며, 형식 형태소 어휘 열 자질과 형식 형태소 어휘 집합 자질의 비교 실험을 통해 자유 어순을 고려한 자질 군집화가 가능함을 보였다. 화행 분류에서 주어와 주어는 직관과는 달리 주어 자질은 사용하지 않았을 때 오히려 성능이 좋았는데, 이것은 한국어의 잦은 주어 생략 현상이 화행 분류에 미치는 영향을 보여준다.

자질 선택 기법을 사용한 자질 비율 실험에서는 전체 자질의 약 4%만 사용하였을 때에도 최고 성능의 95% 이상의 성능을 보이는 것을 통해 매우 적은 자질만으로도 화행 분류에 있어 안정된 성능을 낼 수 있다는 점을 알 수 있었다.

제안한 자질이 형태소 분석 쌍과 중복될 수 있는 점이 있으나 이러한 부분은 기존에서 사용한 자질을 기반으로 하여 제안하는 자질을 추가하는 실험을 위주로 진행되었기 때문에 발생한 것이며, 추후 연구에서 보다 자질을 정밀히 선별하여 해결할 수 있을 것으로 판단된다.

또한, 본 연구에서는 기존에 사용한 n-gram 자질을 그대로 사용하도록 하였으나 이것을 좀 더 다양하게 시도해보는 것도 의미 있는 연구가 될 것이라고 생각한다. 따라서 향후 연구로써 n-gram 자질을 형태소 분석 쌍과 품사 바이그램뿐만이 아닌 다른 형태로 적용하는 방법과 제안하는 자질을 n-gram 자질과 중복되지 않고 적용시킬 수 있는 방법에 대해 연구하고자 한다.

본 연구에서 사용한 말뭉치는 국내 화행 분류 연구 [4-10]에서 사용한 말뭉치 중 가장 크다. 하지만 실험 집합이 1,496개의 발화로 되어 있어 2개의 발화만으로 0.1%의 성능 차이를 보이게 되므로 소수의 발화의 오답 여부가 상대적으로 성능에 크게 반영될 수 있다. 영어권에서 화행 분류 연구에 많이 쓰는 SWITCHBOARD 말뭉치[17]의 경우 205,000개의 발화로 구축되어 있는 것에 비교해 볼 때, 보다 큰 한국어 대화 관련 말뭉치의 구축이 필요하다는 것을 알 수 있다.

참고 문헌

- [1] J.A. Austin, "How to Do Things with words," Harvard University Press, 1962.
- [2] Massaki Nagata and Tsuyoshi Morimoto, "First steps towards statistical modeling of dialogue to predict the speech act type of the next utterance," *Speech Communication* Vol. 15, issue 3-4, pp. 193-203, 2004.
- [3] Maria José Castro et al., "Dialogue Act Classification in a Spoken Dialogue system," CAEPIA-TTIA 2003, LNAI 3040, pp. 260-270, 2004.
- [4] 이성욱, 서정연, "결정트리를 이용한 한국어 화행결정", 제 11회 한글 및 한국어 정보처리 학술대회 논문집, pp. 377-381, Oct. 1999
- [5] Won Seug Choi, Jeong-Mi Cho, Jungyun Seo, "Analysis System of Speech Acts and Discourse Structures Using Maximum Entropy Model," in the Proc. of 37th Annual Meeting of the ACL, pp. 230-237, Maryland, U.S.A., Jun. 1999.
- [6] 은종민, 이성욱, 서정연, "지지벡터기계(Support Vector Machines)를 이용한 한국어 화행분석", 정보처리학회논문지B, 제 12-B권 3호, pp. 365-368, 한국정보처리학회, Jun. 2005.
- [7] 김세중, 이용훈, 이종혁, "이전 문장 자질과 다음 발화의 후보 화행을 이용한 한국어 화행 분석", 제 19회 한글 및 한국어 정보처리 학술대회 발표 논문집, pp. 23-30, 2007.
- [8] 김경선, "개선된 자질 추출 및 가중치 부여 방법을 이용한 한국어 화행 분류 시스템", 서강대학교, 박사학위논문, 2005.
- [9] 이현정, 서정연, "일정관리 영역에서 신경망을 이용한 사용자 의도 파악", 한국인지과학회 춘계학술대회 논문집, pp. 87-90, Jun, 2006.
- [10] 김민정, 한경수, 박재현, 송영인, 임해창, "도메인에 비종속적인 대화에서의 화행 분류", 제 18회 한글 및 한국어 정보처리 학술대회 발표 논문집, pp. 246-253, 2006.
- [11] Massaki Nagata and Tsuyoshi Morimoto. "An experimental statistical dialogue model to predict the speech act type of the next utterance," In Proceedings of the International Symposium on Spoken Dialogues, pp. 83-86, 1993.
- [12] Norbert Reithinger and Martin Klesen, "Dialogue act classification using language models," In Proceedings of EuroSpeech-97, pp. 2235-2238, 1997.
- [13] Max Louwerse and Scott Crossley, "Dialog act classification using N-Gram algorithms," In Proceedings of the International Florida Artificial Intelligence Research Society, Menlo Park, California, pp. 758-763, 2006.
- [14] Stolcke, A. et al., "Dialogue act modeling for automatic tagging and recognition of conversational speech," In *Computational Linguistics* 26(3), pp. 339-373, 2000.
- [15] Webb, N., Hepple, M., Wilks, Y., 2005. Dialog act classification based on intra-utterance features. In Proceedings of the AAAI Workshop on Spoken Language Understanding.
- [16] Kwok Cheung Lan et al., "Dialogue Act Recognition Using Maximum Entropy," In *Journal of the American Society for Information Science and Technology*, 59(6), pp. 859-874, 2008.
- [17] Godfrey, J., Holliman, E., & McDaniel, J. SWITCHBOARD: Telephone speech corpus for research and development. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-92) Vol. 1, pp. 517-520,

1992.

- [18] 21세기 세종계획 국어 기초자료 구축 분과 연구 보고서, 2003.
- [19] 이상주, "자동 품사 부착을 위한 새로운 통계적 모형", 고려대학교, 박사학위논문, 1999.
- [20] Jacob L. Mey. Pragmatics: an introduction. Blackwell Publishers, 2nd edition, 2001.
- [21] T. Joachims, Making large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning, B. Schölkopf and C. Burges and A. Smola (ed.), MIT-Press, 1999.
- [22] 전성희, "한국어의 생략 현상 연구", 고려대학교, 석사학위논문, 2000.



이도길

1999년 고려대학교 컴퓨터학과 학사. 2001년 고려대학교 컴퓨터학과 석사. 2005년 고려대학교 컴퓨터학과 박사. 2006년 4월~2008년 2월 NHN(주) 과장. 2008년 3월~현재 고려대학교 민족문화연구원 HK연구교수. 관심분야는 한국어 정보처리, 기계학습, 정보검색



김민정

2005년 고려대학교 컴퓨터학과 학사. 2007년 고려대학교 컴퓨터학과 석사. 현재 고려대학교 컴퓨터·전파통신공학과 박사과정. 관심분야는 자연어처리, 기계번역, 기계학습



박재현

2005년 고려대학교 컴퓨터학과 학사. 2007년 고려대학교 컴퓨터학과 석사. 관심분야는 자연어처리, 언어 모델, 대화 관리



김상범

2005년 3월~2005년 10월 고려대학교 컴퓨터정보통신연구소 연구조교수. 2005년 11월~2006년 10월 동경공업대학 정밀공학연구소 JSPS 특별연구원. 2006년 11월~2007년 6월 NHN 검색센터 검색품질팀. 2007년 7월~현재 SK텔레콤 인터넷사업단 검색사업팀. 관심분야는 정보검색, 기계학습, 자연어처리

임해창

정보과학회논문지 : 소프트웨어 및 응용
제 35 권 제 10 호 참조