

## 대화 시스템을 위한 사용자 발화 문장의 감정 분류\*

강 상 우<sup>†</sup>      박 흥 민      서 정 연

서강대학교 컴퓨터학과

대화 시스템은 사용자의 의도를 파악하기 위해 발화 문장으로부터 다양한 형태론적 분석을 시도한다. 하지만 사용자는 발화 문장에 포함된 사전적 의미를 통해 의도를 전달할 뿐만 아니라 현재 감정 상태에 따라서 사전적 의미와는 다른 의도를 표현하거나 동일한 의미를 갖는 발화에서 다양한 의도를 표현한다. 따라서 대화에서 사용자의 감정을 파악하는 것은 사용자의 의도를 다양한 방향으로 분석할 수 있게 한다. 본 연구는 기계 학습 방법을 사용하여 사용자 발화 문장에 자동으로 감정 범주를 할당하는 방법을 제안한다. 일반적 감정 범주를 정의하기 위해 세부적인 감정 모델로 인정받고 있는 Plutchick의 감정 모델을 사용하여 9개 감정 범주를 재 정의하고 감정 분류를 위한 자질 집합을 문장 자질과 선행적 자질 그리고 문맥 자질로 구분하였다. 실험을 통하여 3가지 자질들의 최적 조합을 구성하고 감정의 자동 분류를 위해 SVM 분류기를 사용하였다. 실험 결과에서 제안 시스템은 비교 시스템에 비해 15% 높은 62.8%의 F1-평가치 성능을 나타냄으로서 제안된 방법이 감정 분류에 효과적임을 증명한다.

주제어 : 감정 분석, 음성 대화 시스템, 모바일, 기계 학습, 자동 분류, 로봇

---

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

† 교신저자: 강상우, 서강대학교 컴퓨터학과, 연구분야: 자연어처리

E-mail: gangsw@gmail.com

## 서론

인간은 자연어를 사용하는 대화(dialogue)를 통하여 서로간의 의도를 전달한다. 대화는 풍부한 표현력과 직관성으로 인해 인간에게 가장 자연스럽고 효과적인 의도전달 도구이다. 따라서 최근 인간과 기계가 상호작용의 도구로서 자연어(natural language)를 사용하는 대화 시스템에 대한 연구가 활발히 진행되고 있다. 대화 시스템에서는 사용자의 발화 문장에서 사용자의 의도를 파악하기 위해 언어 이해(natural language understanding)과정을 거친다. 그 중 화행 분석은 사용자 의도 파악에 매우 중요한 요소이기 때문에 다양한 방법들이 연구되었다(Choi et al. 2005; 김경선, 2006). 하지만 인간은 화행과 같이 발화에 포함된 사전적인 표현에 의해 의도를 전달할 뿐만 아니라 현재 감정 상태에 따라서 동일한 표현에서도 다양한 의도를 전달하거나 사전적 의미와는 다른 의도를 빈번하게 표현한다. 따라서 사용자의 의도를 보다 다양한 방향으로 분석하기 위해서는 발화 문장의 사전적인 의미를 분석하는 것 뿐만 아니라 대화 도중에 변화되는 사용자의 감정 파악이 필요하다. 또한 대화 시스템에서 사용자의 감정을 파악할 수 있다면 보다 다양하고 지능적인 반응이 가능할 것이다. 예를 들어 음악을 추천하는 시스템에서 사용자의 감정에 따라 추천할 수 있는 장르가 달라진다면 보다 지능적인 시스템이 될 수 있을 것이다. 이와 같이 사용자 친화적인 시스템을 구현하기 위하여 인간의 감정에 대한 접근은 매우 중요하며 최근 ‘Companions’와 같은 상업화 로봇 시스템에서도 이러한 시도가 포함되어 있다(Wilks, 2005).

최근 연구가 진행되고 있는 감정 분석(sentiment analysis) 분야에서는 사람이 작성한 문장들을 감정과 관련된 범주로 구분한다는 점에서 유사하지만 대부분 작성자의 견해나 태도(긍정, 부정 등)를 파악하는 것에 대한 연구가 주류를 이룬다(Pang et al. 2002; Turney, 2002; Wilson et al. 2005; Narayanan et al. 2009). 대화형 학습시스템에서 인간-컴퓨터 상호작용을 높이기 위해 인간의 감정을 예측하는 연구(Litman and Forbes-Riley, 2004)에서는 인식된 인간의 음성에서 4가지 감정 상태(중립, 긍정, 부정, 복합)를 파악하기 위해 적합한 음성과 어휘 자질에 관하여 논의하였다. 이 연구에서는 감정을 예측할 수 있는 음성과 어휘 자질을 찾기 위해 감정 상태를 위한 표지 부착 데이터를 활용하여 기계 학습에 적용하였다. 그러나 문맥을 고려하

지 않고 어휘 자질만을 사용하였기 때문에 학습자의 반응이 동일할 때 항상 같은 감정을 예측한다는 한계를 가진다. 또한 대화 시스템에서 사용자의 감정을 활용하기 위해서는 4가지의 감정 범주는 한계가 있으며 보다 세분화된 감정 범주의 분석이 필요하다. 한국어를 대상으로 한 연구에서는 대화체 문장에서 세분화된 감정의 범주를 정의하고 감정을 분류한 연구는 매우 드물다. 문현구(2002)에서는 한국어 대화체 문장(SMS, 메신저)에서 특정 감정을 분류하였다. 하지만 영역 의존적인 데이터(이모티콘, 신조어 등)와 범주(황당함, 씩씩함 등)를 사용함으로써 일반적인 영역에 적합하지 않다. 또한 어절만을 자질로 사용하여 형태론적 변형이 다양하게 일어나는 한국어의 특성에 따라 데이터 희소 문제(data sparseness problem)가 발생할 수 있다.

본 논문에서는 여러 영역에서 일반적으로 사용할 수 있는 세분화된 감정 범주를 결정하고 기계 학습 방법을 사용하여 자동으로 범주를 할당하는 방법을 제안한다. 감정 범주를 결정하기 위해 심리학 분야에서 접근한 연구(Plutchik et al. 1980; Plutchik and Conte, 1997; Plutchik, 2001)를 바탕으로 감정 범주를 정의하고 기계 학습을 위해 감정 분류에 필요한 자질 집합을 문장 자질, 선행적 자질 그리고 문맥 자질로 구분하여 최적의 자질 집합을 제안한다. 그리고 사용자 발화로부터 제안한 자질을 사용하여 감정을 자동 분류하기 위해 대표적 기계 학습 방법 중 하나인 SVM(support vector machine)을 사용한다(Vapnik, 1995).

본 논문의 구성은 다음과 같다. 2절에서는 관련 연구들을 설명하고 3절에서는 감정의 분류 범주를 정의하고, 감정 분류를 기계 학습에 적용하기 위하여 자질들을 문장 자질, 선행적 자질 그리고 문맥 자질로 구분하여 설명한다. 4절에서는 실험은 통해 최적화된 자질 집합을 결정하고 비교 시스템과 성능을 비교 평가한다. 5절에서는 결론 및 향후 과제에 대해 설명한다.

## 관련 연구

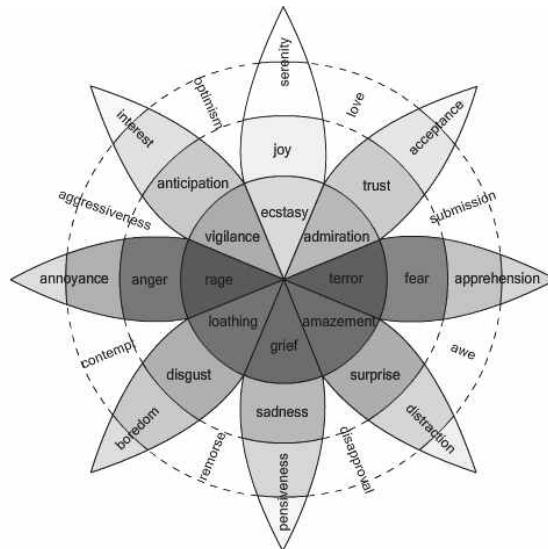
본 절에서는 감정 분류에 관한 기존 연구를 소개하고, 체계적인 감정 범주를 위한 Plutchik의 감정 모델에 대하여 설명한다.

## 기존 감정 분류 연구

감정에 관련한 기존 연구는 대부분 특정 영역을 위한 감정 분류를 위해 각 감정 범주를 위한 사전을 구성하고 문장이나 문서에서 관측되는 단어들과의 패턴 매칭 방법을 사용하였다. 그리고 감정 표현을 위해 메신저나 E-mail 전송에 이모티콘을 활용하는 시도들이 있었다. 그 중에서 ‘Expressive Real Time Communications Interface’는 감정 사전을 참조하여 각 단어의 적절한 범주를 찾고 이 과정을 통해 인식된 감정 상태를 이에 연관된 사람의 표정 이미지를 생성하여 상대방에게 전송하는 시스템이다(Zhe et al. 2002). EmpathyBuddy는 상식 정보 데이터베이스로부터 획득한 정보를 표정에 기반을 둔 6가지 감정범주에 대하여 각각 가중치를 부여함으로써 중심 감정 상태를 분석하고 그것을 이모티콘으로 나타내었다(Liu et al. 2003). Multi-modal Eliza는 기존 Eliza 시스템에 사용자의 선호와 목표 등에 대한 대화 패턴을 정의하여 대화가 지속적으로 유지되도록 하고 사용자의 질문이나 시스템의 답변에 대응되는 이모티콘을 생성하는 시스템이다(Fitriani et al. 2003). 또한 음성 대화 시스템을 위한 연구에서는 음운과 문장 그리고 문맥을 자질로 사용하였고, Boosting 알고리즘에 기반한 분류기를 사용하여 감정을 분류하였다. 감정 범주로는 7개의 범주를 제안하였지만 감정 분포가 크게 치우침에 따라 실제로는 긍정과 부정적인 감정, 2가지로 분류하였다(Liscombe et al. 2005). 한국어 대화의 감정 분류를 위한 기존 연구로는 Naïve Bayes 알고리즘과 HMM(hidden Markov model)을 결합한 모델을 이용하여 25개 세부 범주로 감정을 분류한 연구가 있다(문현구, 2002). 문현구의 연구에서는 각 어절에 대해 감정을 부여하고 이것을 Naïve Bayes 알고리즘을 이용한 통계적인 방법으로 감정을 분류하였다. 또한 감정 변화를 감정 분류에 적용하기 위하여 Naïve Bayes 알고리즘의 확률 값을 관측 확률 값으로 하는 HMM을 사용하였다. 위 감정 연구들에서 사용된 감정 범주는 데이터를 바탕으로 감정 상태를 분류함에 따라 감정 간의 구분이 모호하거나 특정 감정으로 치우치게 감정을 구분한 문제점이 있다. 따라서 인간의 감정에 대해 좀 더 체계적이고 일반적으로 사용할 수 있는 감정 범주가 필요하며 한국어의 경우 형태론적 변형이 많기 때문에 한국어의 특징을 반영한 연구가 필요하다.

Plutchik의 감정 모델

세분화된 감정 분류에 관한 연구는 Plutchik, Russell, Ekman 등 심리학의 성과에 힘입은 바 크다. 세분화된 감정 분류의 초기 연구가 기본 감정을 소수의 범주로 구분하여 이를 인식하거나 표현하고자 하였지만 초기 연구에 사용되었던 감정 범주는 너무 단순하여 감정 분류를 더욱 세분화할 필요성이 대두되었다. 이에, 최근 연구에서 세분화된 감정 분류의 주요한 모델이 된 것은 <그림 1>의 Plutchik의 감정 팽이(Wheel of Emotions)이다(Plutchik at al. 1980; Plutchik and Conte, 1997; Plutchik, 2001). Plutchik의 감정 팽이는 인간의 감정을 인간의 감정을 8개의 기본 감정군으로 분류하고 이것을 4개의 대립 쌍으로 설정하며, 강도에 따른 변화형을 제시하고 조합 감정을 기본 감정으로 환원할 수 있다. 이러한 분석적인 설명은 공학적 처리의 가능성을 높이는 데 크게 기여한다.



<그림 1> Plutchik의 감정 팽이

기본 감정들은 동서양간에도 많은 유사점을 찾을 수 있다. 동양에는 7정(七情) (희(喜), 노(怒), 애(哀), 구(懼), 애(愛), 오(惡), 욕(欲))으로 분류되는 감정 범주가 있으며 이것은 Plutchik의 기본 감정군과 매우 유사하며 미묘한 차이가 있을 뿐이다(조화태 외, 2002). 따라서 인간의 감정은 시대적 문화적 차이에도 불구하고 공통적인 특성을 지니고 있음을 알 수 있다.

한국어를 대상으로 텍스트 기반의 감정 분석에 관한 연구는 주로 Pennebaker의 모델(Pennebaker et al. 2003)을 기반으로 연구되었다. Plutchik의 감정 모델을 사용하여 한국어를 대상으로 한 연구는 감정 온톨로지를 위한 자질 분석에 대한 연구(윤애선 외, 2009)가 있으며 자동 감정 분류에 관한 연구는 아직 시도되지 않았다.

### 감정 범주 정의와 분류 자질

본 연구는 대화 시스템을 위한 감정 분석을 목적으로 하기 때문에 사용자 발화에 해당하는 문장에 대해 감정 분류를 하는 것을 목적으로 다룬다. 대부분의 대화 시스템에서 사용하는 담화 모델은 사용자의 발화에서 1개의 사용자의 의도(화행이나 행위)를 기준으로 처리하므로 감정 또한 한 문장에서 나타나는 두드러진 1개의 감정을 대상으로 한다.

감정을 분류하기 위해서는 감정의 범주와 분류 자질들을 결정해야 한다. 감정의 범주를 결정하기 위해 심리학적 접근 방법을 사용하고 감정 분류에 필요한 정보를 가지는 자질들은 3가지로 구분하여 최적의 자질 조합을 기계 학습에 적용한다. 본 절에서는 먼저 감정 범주를 결정하기 위한 배경을 설명하고 다음으로 감정을 분류하기 위하여 사용하는 각 자질들에 대하여 설명한다.

#### 감정 범주 정의

본 연구에서는 체계적이고 영역 독립적인 감정 범주를 결정하기 위해 Plutchik의 감정 모델을 이용한다. Plutchik의 연구에서는 강도에 따른 변형과 조합 감정은 기

본 감정 내에 포함되거나 기본 감정으로 설명할 수 있으므로, 일반적인 감정 범주로 적용하기 용이한 기본 감정을 기준으로 감정 범주를 결정한다. Plutchik의 연구에서 기본 감정은 8개로 구분된다(‘기쁨’(joy), ‘신뢰’(trust), ‘두려움’(fear), ‘놀람’(surprise), ‘슬픔’(sadness), ‘혐오’(disgust), ‘화남’(anger), ‘기대’(anticipation)). 하지만 실제 대화에서 ‘신뢰’의 감정이 표현되는 경우는 매우 드물며 ‘기쁨’과 ‘신뢰’의 조합감정인 ‘사랑’(love)이 ‘신뢰’를 포함하면서 빈번히 표현되기 때문에 ‘신뢰’ 대신 ‘사랑’을 감정 범주로 포함한다(<그림 1> 참조). 따라서 8개의 감정과 감정이 없는 경우를 표현하는 ‘감정 없음’(NONE)의 범주를 추가하여 총 9개의 범주를 정의한다. <표 1>은 본 연구에서 사용하는 감정 범주와 발화의 예를 보여준다.

<표 1> 감정 범주와 발화 예

감정 발화 예	발화 예
기쁨(joy)	응~ 재미있었어~
사랑(love)	사랑하는 사람과 함께인데 왜 불행해?
두려움(fear)	그러게. 내 나이도 벌써.....
놀람(surprise)	그런 것도 있어?
슬픔(sadness)	어. 속상해 죽겠어
혐오(disgust)	매우 이상하지~
화남(anger)	장난하나
기대(anticipation)	더 좋은 사람 만나겠지
감정 없음(none)	생각 해 볼게

### 자질 선택

발화 문장을 기계 학습을 통하여 각 범주로 자동 분류하기 위해서는 발화 문장으로부터 범주를 구분하기 위한 자질을 추출해야 한다. 본 절에서는 감정 분류에 사용할 자질 집합을 문장 자질, 선행적 자질, 문맥 자질로 구분하여 설명한다.

**문장 자질**

문장 자질은 현재 발화 문장으로부터 직접 추출이 가능한 자질을 말한다. 하지만 문서가 아닌 문장으로부터 자질을 추출하기 때문에 데이터 희소 문제가 발생할 수 있다. 따라서 형태소-품사의 조합과 문장 길이 등 다양한 조합 자질을 적용한다. 문장 자질을 선별하기 위해서 최근 한국어 문장 분석 분야에서 포괄적인 자질 선택 문제를 다루는 연구(김민정 외, 2008)에서 사용된 17개의 자질들을 본 연구에서 사용하는 말뭉치에 각각 한 개씩 사용하여 기여도가 있는 7개 문장 자질만을 선별하였고 실험을 통해 각 자질에 대한 감정 분류 성능과 감정 분류의 성능 향상에 어느 정도 기여할 수 있는지를 평가한다. <표 2>와 같이 형태소-품사 쌍은 문장에 출현하는 어휘들을 대표하는 자질로 사용되며, 마지막 어절, 마지막 용언, 마지막 부사 그리고 마지막 어미 자질의 경우 한국어의 경우 중심 내용이 문장의 뒤에 위치하므로 마지막으로 나온 품사의 자질들에 가중치를 부여한다. 형식 형태소 열은 전체 문장의 형태를 보여 주기 위한 자질로 사용한다.

<표 2> 문장 자질의 종류와 예

자질	자질 설명	자질의 예
형태소-품사 쌍	형태소 분석 결과로 얻어지는 어휘와 품사의 쌍	간섭-NNG, 심하-VA
마지막 어절	발화 문장의 마지막 어절	모양이냐, 비참하다
마지막 용언	마지막 용언의 어휘	좋, 모르
마지막 부사	마지막 부사의 어휘	너무, 얼마나
마지막 어미	마지막 어미의 어휘	다, 었
형식 형태소 열	의존 명사, 접미사, 조사, 감탄사, 서술격조사, 보조용언, 어미들의 어휘 열	고-싶-다
발화의 길이	어절의 수 ( S: 1개, M: 2~5개, L: 6개 이상)	Len_S, Len_M, Len_L

NNG(일반명사), VA(형용사)

**선형적 자질**

감정 표현의 경우 반 고정 표현(semi-frozen expression)과 다 어절(multi-word) 표현



이 빈번하기 때문에 단순한 문장 자질만으로는 효과적인 분류가 어렵다(윤애선 외, 2009). 대표적으로 관용적 표현의 경우 다수의 단어들끼리 서로 결합되어 감정을 표현하므로 이를 반영하기 위한 자질이 필요하다. 본 연구에서는 이러한 관용적 표현들을 반영하는 자질을 선행적 자질이라고 정의하고, 선행적 자질로 감정 어휘 열을 사용한다. 감정 어휘 열들은 길이가 정의되지 않는 n-gram으로 구성되어 있다. 그러나 문장 자질 선택 과정에서 모든 n-gram을 고려할 수 없기 때문에 별도로 구축된 감정 어휘 열 사전을 사용하는 것이 보다 효율적이다. 감정 어휘 열 사전은 실험 말뭉치와는 별도로 구성된 데이터로서 한국어 감정표현 관용어에 대한 연구(김향숙, 2001)에서 사전, 관련 논저, 신문, 문학 작품 등에서 선정된 2,456개의 감정에 대한 관용적 표현들을 수집하여 재가공한 것으로 본 연구 범위에 해당하는 1,301개의 감정 어휘 열을 선별하여 각 감정 범주의 자질로 사용하였다. 또한 교차어인 한국어의 특성에 따라 다양한 형태론적 변형이 나타나므로 정규화 과정을 추가하였다. 형태소 분석 결과의 품사에 따라 조사, 어미를 제거하였고, 감정의 강도를 주로 표현하는 관형사를 제외한 나머지 품사에 대해서 어휘 열을 생성하였다. <표 3>에서는 ‘화남’의 감정을 갖는 관용적 표현들과 이것을 정규화 과정을 거쳐 감정 어휘 열로 작성된 예를 보여준다. 감정 어휘 열은 다른 자질들처럼 분류기 학습 단계에서 데이터에 의해 학습되는 것이 아니라 다른 자질들을 학습한 후에 정규화된 1,301개의 감정 어휘 열들을 각 범주 별로 나누어 학습된 자질 집합에 포함시킨다. 분류 단계에서는 다른 자질들과 마찬가지로 이진 자질-가중치(binary feature-weighting) 기법을 사용하여 계산한다.

<표 3> 관용적 표현과 감정 어휘 열의 예

감정 표현 관용어	감정 어휘 열
열이 받다	열/NNG+받/VV
뚜껑이 열리다	뚜껑/NNG+열리/VV

### 문맥 자질

문장 자질과 선행적 자질의 경우 현재 발화 문장에서 분석할 수 있는 자질만을

고려한다. 그러나 <표 4>와 같이 대화 1과 대화 2의 마지막 발화는 동일 문장이지만 이전 문맥에 따라 현재 감정이 다르게 나타난다. 따라서 현재 발화의 정보만으로는 감정을 분류할 수 없는 현상들을 위해 이전 발화들로부터 현재 발화의 감정 분석에 필요한 문맥 자질들을 사용한다. 본 논문에서는 문맥 자질로서 사용자의 이전 감정과 이전 화행 그리고 시스템의 이전 감정을 문맥 자질로 사용한다.

<표 4> 문맥에 따른 감정 상태의 예

대화 1			대화 2		
화자 발화	감정	화행	화자 발화	감정	화행
U: 어제 새벽에 전화했었지?	두려움	ask-ref	U: 왜 이렇게 시끄러워!	화남	opinion
S: 문제가 좀 생겨서 전해드렸습니다.	-	response-ref	S: 글썄요.	-	opinion
U: 무슨 일이야?	두려움	-	U: 무슨 일이야?	화남	-

### SVM 분류기

감정 분류를 위한 확률 모델을 식 1과 같이 구성한다. 식 1에서 1개의 대화는  $n$  개의 문장으로 이루어지고 사용자와 시스템이 한 차례씩 교대로 발화한다고 가정한다.  $E(U_i)$ 는  $i$ 번째 문장의 감정 범주이며  $E_j$ 는  $j$ 번째 감정 범주이다. 그리고  $Fs_i$ 는  $i$ 번째 발화의 문장 자질이고  $Fp_i$ 는  $i$ 번째 발화의 선형적 자질이다; 선형적 자질은 감정 어휘 열을 사전으로 구성하여 학습된 자질 집합에 포함시킨다. 그리고  $Fc_{i-1}$ 은  $i-1$ 번째 문장의 문맥 자질이고  $Fc_{i-2}$ 은  $i-2$ 번째 문장의 문맥 자질이다.

식 1에서  $P(E_j|Fs_i)$ 는  $i$ 번째 발화의 문장 자질이 출현 할 때에  $j$ 번째 감정일 확률이고  $P(E_j|Fp_i)$ 는  $i$ 번째 발화의 선형적 자질이 출현 할 때에  $j$ 번째 감정일 확률이다.  $P(E_j|Fc_{i-1})$ 과  $P(E_j|Fc_{i-2})$ 는 각각  $i-1$ 번째 발화와  $i-2$ 번째 발화의 문맥 자질이 출현 할 때에  $j$ 번째 감정일 확률이다. 모든 감정 범주에 대해서 각 항들의 곱을 최대로 하는  $j$ 번째 감정 범주를  $i$ 번째 문장의 감정 범주로 할당한다.

$$E(U_i) \approx \operatorname{argmax}_{E_j} P(E_j | F s_i) P(E_j | F p_i) P(E_j | F c_{i-1}) P(E_j | F c_{i-2}) \quad \text{식 1.}$$

$$P(E_j | F) = \frac{P(E_j) P(F_1, \dots, F_n | E_j)}{P(F_1, \dots, F_n)} \quad \text{식 2.}$$

$$\begin{aligned} P(E_j) P(F_1, \dots, F_n | E_j) &= P(E_j) P(F_1 | E_j) P(F_2, \dots, F_n | E_j, F_1) & \text{식 3.} \\ &= P(E_j) P(F_1 | E_j) P(F_2 | E_j, F_1) \dots P(F_3, \dots, F_n | E_j, F_1, F_2) \\ &= P(E_j) P(F_1 | E_j) P(F_2 | E_j, F_1) \dots P(F_n | E_j, F_1, F_2, \dots, F_{n-1}) \\ &\cong P(E_j) \prod_{i=1}^n P(F_i | E_j) \end{aligned}$$

식 1을 계산하기 위해 식 1의 각 우항들을 Bayes' 정리에 의하여 식 2와 같이 전개한다. 그리고 식 2의 분자를 식 3에처럼 각각의  $F$ 에 대하여 전개하고 조건부 독립 가정을 적용하면 우항의 마지막 식을 얻을 수 있다.

본 연구에서는 자동으로 감정을 분류하기 위해 커널 기반(kernel-based) 학습 알고리즘과 자질-가중치(feature-weighting) 방법을 사용하는 SVM 분류기를 사용한다 (Vapnik, 1995). 한 발화 문장에서 각 자질들은 한 번 이상 관측되기 어렵기 때문에 이진 자질-가중치 기법은 본 연구에서 매우 효과적으로 동작하며 식 4와 같이 적용한다.

$$w_{ik} = \begin{cases} 0 & \text{noneexist} \\ 1 & \text{otherwise} \end{cases} \quad (w_{ik} \text{는 } i\text{번째 발화 문장의 } k\text{번째 자질}) \quad \text{식 4.}$$

SVM 분류기를 적용하기 위한 단계는 다음과 같다.

- 학습 단계
  1. 학습 데이터의 발화 문장들에서 형태소 분석과 정규화 과정을 거쳐 어휘 열을 생성
  2. 생성된 어휘 열로부터 각 범주마다 문장 자질들을 추출하고 감정 어휘 열 사전을 추가

3. 발화 문장에 부착된 화행과 감정 정보들을 문맥 자질로 추출, 최종 자질 집합 생성
  4. 발화 문장을 최종 자질 집합으로 변환하고 SVM 분류기를 학습
- 실행 단계
    1. 학습된 다중 SVM 분류기를 통해 각 테스트 데이터의 발화 문장의 감정 범주를 할당

#### 4. 실험 및 평가

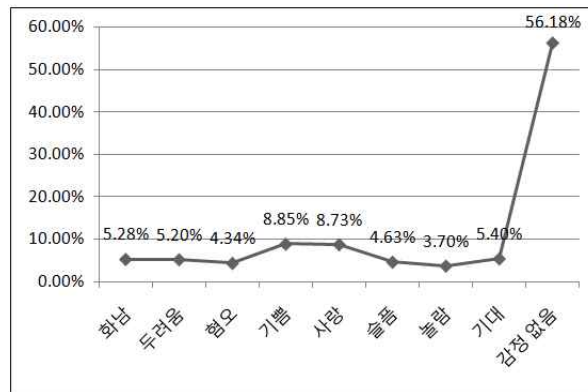
본 절에서는 제안한 방법을 평가하기 위한 실험 환경과 결과를 설명한다. 실험에 사용한 데이터 설명과 최적의 자질 집합을 위한 문장 자질, 선형적 자질 그리고 문맥 자질에 대한 실험 결과를 분석하고 마지막으로 제안시스템의 최종 자질 집합에 대한 결과를 실험을 통하여 비교 시스템과 성능을 비교 평가한다.

##### 실험 환경

본 논문에서 사용한 실험 말뭉치는 7,102개 문장으로 구성된 채팅 말뭉치를 사용하였다. 이 말뭉치는 3개의 주제(음식(3,009개 발화), 사랑(3,092개 발화), 음악(1,001개 발화))를 갖고 있으며 한 발화 당 평균 8.77개의 단어로 이루어져 있다. 이 말뭉치는 두 명의 사용자가 SMS를 이용하여 특정 주제에 관하여 대화하는 내용을 수집하였다. 이 말뭉치는 화자 정보, 채팅 영역 정보, 감정 정보, 화행 정보, 발화의 목적, 발화의 대상, 주어 정보 등을 영역 당 한 명씩의 언어 분석에 능숙한 대학원생을 할당하여 수동으로 부착되었다. 참여 대학원생들은 정보 표지 부착을 위해 미리 각 정보의 정답 예제와 설명을 위한 교육을 받았다. 그리고 부착된 정보 표지의 신뢰도를 평가하기 위해 500개의 샘플 발화로부터 Kappa value(Fleiss, 1971)를 측정하였고 모든 참여 대학원생으로부터 60.4%의 일치율(agreement rate)을 보여주었다.

<그림 2>는 전체 말뭉치의 감정 범주간의 분포를 보여준다. 전체 감정 중에서 약 56%의 발화가 감정을 가지고 있지 않으며 대부분의 감정들은 대체로 균등하게 분포되어 있다.

본 연구에서는 기계 학습을 위해 SVM을 사용한다. SVM은 이진 분류 기법을 사용하기 때문에 2개 이상의 범주를 분류하기 위해서 다중 SVM 분류기를 구성하였다(Weston at al. 1999). 실험에서 데이터가 편중되는 것을 방지하기 위해 5 fold cross validation을 사용하고 평가 지표로서 정확률<sup>1)</sup>(precision)과 재현율<sup>2)</sup>(recall)을 정의하여 사용하였으며 정확률과 재현율의 조화평균인 F1-평가치<sup>3)</sup>를 사용하였다.



<그림 2> 전체 말뭉치의 감정 분포

### 문장 자질 평가

감정 분류를 위한 문장 자질을 위해 7개의 자질(형태소-품사 쌍, 마지막 어절, 마지막 용언, 형식 형태소 열, 마지막 부사, 마지막 어미, 발화의 길이)을 실험을 통해 비교하였다. <표 5>에서는 7가지 문장 자질을 각각 단독으로 사용했을 때의

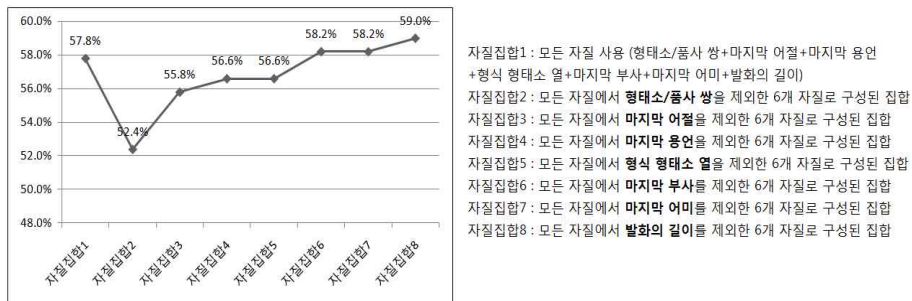
- 1) 정확률= $\frac{j\text{번째 감정범주로 선택된 문장 중 정답인 문장 수}}{j\text{번째 감정범주로 선택된 문장 수}}$
- 2) 재현율= $\frac{j\text{번째 감정범주로 선택된 문장 중 정답인 문장 수}}{j\text{번째 감정범주가 정답인 문장 수}}$
- 3) F1-평가치= $\frac{2(\text{정확률} \times \text{재현율})}{(\text{정확률} + \text{재현율})}$

성능을 보여준다. 형태소-품사 쌍과 마지막 어절의 경우 감정 분류에 가장 많은 영향을 끼치며 마지막 어미나 발화의 길이는 상대적으로 많은 영향을 주지 못함을 알 수 있다.

<표 5> 문장 자질의 성능 비교 (%)

평가자질	정확률	재현율	F1-평가치	평가자질	정확률	재현율	F1-평가치
형태소-품사 쌍	62.6	48.8	54.8	마지막 부사	28.4	17.2	21.4
마지막 어절	57.2	40.8	47.6	마지막 어미	10.2	12.8	11.4
마지막 용언	46.2	30.8	37.0	발화의 길이	9.2	13.4	10.9
형식 형태소 열	54.6	22.8	32.2				

<그림 3>은 자질들의 다양한 조합들이 감정 분류에 어떻게 기여하는지 보여준다. 자질 조합 중 형태소-품사 쌍 자질 만으로도 최종 성능에 근접한 결과를 보이므로 나머지 자질들의 추가가 얼마나 성능에 보완이 되는지를 알아보기 위해 전체 문장 자질 7개를 모두 사용하였을 때를 기본으로 하고 나머지 자질 중에 한 가지씩 제외하면서 각 6개의 자질로 실험했을 때를 비교하였다. 실험 결과로 <표 5>에서 실험한 것과 동일하게 형태소-품사 쌍이 감정 분류에 가장 큰 영향을 끼치는 것을 볼 수 있었다. 또한 형태소-품사 쌍 자질을 포함한 자질 조합들은 그 차이가



<그림 3> 다양한 문장 자질 조합의 성능 비교

미미하여 본 실험에서는 성능 하락에 영향을 미치는 자질들만을 제외하는 방법을 사용하였다. 따라서 마지막 부사와 마지막 어미, 발화의 길이의 자질의 경우에는 자질로 사용하지 않는 것이 오히려 성능이 증가하기 때문에 감정 분류를 위한 문장 자질로는 형태소-품사 쌍, 마지막 어절, 마지막 용언 그리고 형식 형태소 열을 선택하였다.

### 선택적 자질 평가

감정에 대한 관용적 표현의 정보를 나타내기 위한 선택적 자질로서 감정 어휘 열 사전을 구축하였다. 실험에 의하여 선택된 문장 자질들과 추가적으로 감정 어휘 열을 포함한 경우와 그렇지 않은 경우를 비교하여 선택적 자질에 대해 평가하였다. 감정 어휘 열을 포함하였을 경우 추가적인 자질로 인해 ‘감정 없음’ 범주에서 각 감정 범주로 분류되는 경우가 증가한다. 따라서 정확률에서 약간의 성능 하락이 있지만 <표 6>에서와 같이 재현율에서는 약 5% 성능 향상이 있고 최종적으로 F1-평가치에서는 3%의 성능 향상을 볼 수 있다. 따라서 선택적 자질이 감정 분류에 기여함을 알 수 있다.

<표 6> 문장 자질과 선택적 자질 조합의 성능 비교 (%)

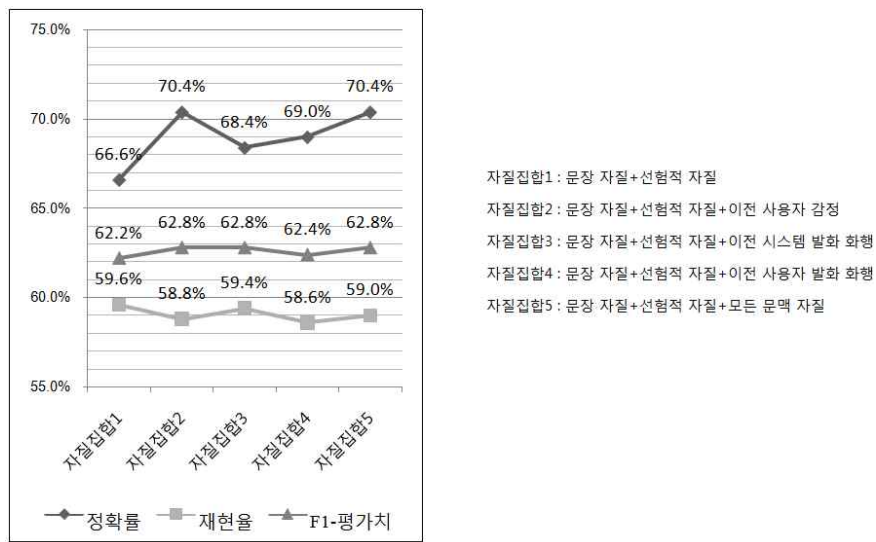
자질집합 분류	정확률	재현율	F1-평가치
선택적 자질집합*	68.8	54.4	59.2
선택적 자질집합*+감정 어휘 열	66.6	59.6	62.2

\* : 형태소-품사 쌍, 마지막 어절, 마지막 용언, 형식 형태소 열

### 문맥 자질 평가

감정 분류를 위한 문맥 자질을 위해 사용자의 이전 감정과 이전 사용자 발화의 화행 그리고 이전 시스템 발화의 화행을 문맥 자질로 선택하였다. 각 문맥 자질들을 평가하고 최적의 자질 집합을 찾기 위해 본 실험에서 선택한 문장 자질과 선택

적 자질과의 조합들을 구성하여 실험하였다. <그림 4>는 각 자질들의 다양한 조합을 실험한 결과이다. 모든 문맥 자질을 사용한 조합이 가장 높은 성능 향상을 보이며 정확률에서는 약 5%의 성능향상을 보이지만 최종적으로 F1-평가치에서 미세한 향상을 보이며 문장 자질과 선행적 자질 만을 사용했을 때에 비해 큰 차이를 보이지 못하고 있다. 이것은 실험 데이터가 본래 채팅을 위한 발화 쌍으로써 감정 표지를 부착하는 데 있어 문맥을 고려하지 않았기 때문에 문맥 자질이 감정 분류에 큰 기여를 하지 못한 것으로 추측된다. 하지만 실제 대화에서는 발화의 문장이 같더라도 문맥에 따라 다른 감정이 나타날 수 있으므로 문맥 자질은 실제 시스템에서 성능 향상에 기여 할 것으로 분석된다.



<그림 4> 문맥 자질의 조합 별 성능 비교

### 제안 시스템

<표 6>은 제안 시스템의 각 감정 별 성능 측정 결과를 보여준다. 각 분류 범주에 따른 성능을 측정한 결과 '감정 없음'을 제외하고 '기대'을 가장 잘 분류하였는



데 이것은 ‘기대’가 ‘~하고 싶다’와 같은 특정 어휘와 많이 나타나 다른 범주와는 다른 발화 형태를 보였기 때문이다. 반면에 ‘두려움’이나 ‘슬픔’의 범주의 경우 가장 낮은 성능을 보인다. ‘두려움’이나 ‘슬픔’의 경우 이전 상황에 따라 현재 사용자의 감정이 크게 의존하는데 사용하는 실험 데이터는 채팅성 발화 쌍 위주로 구성되어 있기 때문에 문맥 자질들이 크게 영향을 미치지 못한 것으로 예상된다.

<표 6> 제안 시스템의 각 범주 별 성능 (%)

분류범주	정확률	재현율	F1-평가치	분류범주	정확률	재현율	F1-평가치
기쁨	60.0	44.2	50.6	혐오	55.7	43.9	47.6
사랑	60.6	38.8	47.0	화남	75.9	44.0	55.3
두려움	67.3	49.8	56.9	기대	87.5	82.3	84.7
놀람	75.0	64.6	69.1	감정 없음	80.5	93.0	86.3
슬픔	69.6	67.5	68.1	평균 성능	70.4	59.0	62.8

<표 7>에서는 분류된 감정 범주들이 어떤 감정 범주로 오 분류 되었는지를 보여준다. 전체 오류 중에서 ‘기대’와 ‘놀람’은 다른 감정으로 잘못 분류된 비율이 가장 낮지만 ‘혐오’는 ‘화남’으로 오 분류된 경우가 빈번히 발생하게 발생하였다. ‘혐오’와 ‘화남’은 Plutchik의 감정 분류 (<그림 1>)에서 서로 인접하여 있는 감정들로서 모호한 양상을 가지고 있으며 이와 마찬가지로 각 감정들 간의 인접한 감정들 사이에 오류가 가장 빈번함을 확인할 수 있다. 데이터의 상당 부분이 ‘감정 없음’의 감정 범주에 속하기 때문에 감정이 있는 문장들을 감정이 없는 것으로 분류한 오류가 가장 많이 발생하였다. 이를 해결하기 위해서 발화 문장의 감정 분류 이전에 감정의 유/무를 미리 판단할 수 있는 분석과정을 추가한다면 성능 향상에 도움이 될 것으로 기대한다.

제안한 방법을 이용한 감정 분류 시스템의 성능을 비교 평가하기 위해한 비교 시스템(문현구, 2002)과 성능을 비교하였다. 비교 시스템의 자질과 통계 모델을 제안한 감정 범주와 동일한 실험 데이터에 적용하여 제안 시스템과 비교하였다. <표 8>에서 비교 시스템 1은 어절 자질과 Naive Bayes 분류기를 사용한 시스템이

<표 7> 각 범주 별 오류 분석 결과

정답 결과	화남	두려움	혐오	기쁨	사랑	슬픔	놀람	기대	감정 없음	합계
화남	O	18	34	4	7	25	2	0	45	135
두려움	14	O	4	0	6	32	3	2	57	118
혐오	25	9	O	4	4	9	0	1	42	94
기쁨	13	5	1	O	31	0	6	4	72	132
사랑	4	15	0	37	O	27	6	8	84	181
슬픔	24	36	6	8	23	O	0	2	32	131
놀람	3	5	2	4	0	3	O	0	24	41
기대	0	4	1	6	9	0	0	O	19	39
감정 없음	176	194	142	180	152	122	162	71	O	1199
합계	259	286	190	243	232	218	179	88	375	2070
비율	12.5%	13.8%	9.2%	11.7%	11.2%	10.5%	8.6%	4.3%	18.1%	100%

<표 8> 비교 시스템과 성능 비교 (%)

구분	정확률	재현율	F1-평가치	정확도
비교 시스템 1	45.6	53.6	47.8	60.8
비교 시스템 2	64.6	45.2	51.2	70.8
제안 시스템	70.4	59.0	62.8	77.0

고 비교 시스템 2는 단순한 분류기의 성능 차이를 알아보기 위해 어절 자질과 SVM 분류기를 사용한 비교 시스템이다. 본 논문에서는 정확률과 재현율 그리고 F1-평가치를 평가 지표로 사용하지만 비교 시스템에서는 정확도(accuracy)를 평가 지표로 사용함에 따라 이를 같이 표시하였다. 실험 결과 비교 시스템1에 비해 제안 시스템이 F1-평가치에서 15%의 성능 향상을 보인다. 또한 단순히 SVM 분류기로 인한 성능 향상을 고려한 비교 시스템 2에서도 제안 시스템은 9.6%의 성능 향상을 보였다.

## 5. 결 론

인간은 지식에 의한 판단 뿐만 아니라 현재의 감정에 영향을 받는다. 따라서 대화 시스템에서 사용자의 의도를 보다 정확하게 분석하기 위해서는 기존의 형태론적인 분석 뿐만 아니라 대화 도중에 변화되는 사용자의 감정 파악이 필요하다. 이와 같은 접근은 대화 시스템으로부터 보다 다양한 지능적인 반응을 기대할 수 있다. 본 논문에서는 대화 시스템을 위한 사용자 발화 문장에서 자동 감정 분류 방법을 제안하였다. 일반적으로 사용할 수 있는 감정 범주를 위해 심리학 분야에서 접근한 기본 범주를 대화 시스템에 적합하게 9개 범주로 재 정의하였다. 또한 감정 분류를 위한 자질 집합으로서 문장 자질과 선행적 자질 그리고 문맥 자질로 구분하여 최적 자질 집합의 조합을 실험을 통하여 구성하였다. 문장 자질로서 형태소-품사 쌍, 마지막 어절, 마지막 용언 그리고 형식형태소 열을 문장 자질로 사용하였고 선행적 자질로는 감정 분류를 위한 관용어 표현들을 이용하여 감정 어휘 열 사전을 구축하고 학습된 자질 집합에 포함하여 사용하였다. 마지막으로 문맥 자질로서 사용자의 이전 감정, 이전 사용자 발화의 화행, 이전 시스템 발화의 화행을 사용하였다. 최적 자질 집합과 SVM 분류기를 사용한 제안 시스템의 성능은 비교 시스템에 비해 15% 높은 62.8%의 F1-평가치 성능을 보임으로서 제안된 방법이 감정 분류에 보다 효과적임을 증명하였다.

## 참고문헌

- [1] Choi, W., Kim, H. and Seo, J. (2005), An integrated dialogue analysis model for determining speech acts and discourse structures, *IEICE Transactions on Information and Systems*, Vol. E88-D, No. 1, pp. 150-157.
- [2] 김경선 (2006), 개선된 자질 추출 및 가중치 부여 방법을 이용한 한국어 화행 분류 시스템, 박사학위논문, 서강대학교.
- [3] Wilks, Y. (2005), Artificial companions, *Interdisciplinary Science Reviews*, Vol. 30, pp.

145-152.

- [4] Pang, B., Lillian, L. and Vaithyanathan, S. (2002), Thumbs up? Sentiment Classification using Machine Learning Techniques, *Proceedings of Annual Meeting of the Association for Computational Linguistics and Empirical Methods in Natural Language Processing*, pp. 79-86
- [5] Turney, P. D. (2002), Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews, *Proceedings of Annual Meeting of the Association for Computational Linguistics and Empirical Methods in Natural Language Processing*, pp. 417-424
- [6] Wilson, T., Wiebe, J. and Hoffmann, P. (2005), Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis, *Proceedings of Human Language Technologies and Empirical Methods in Natural Language Processing*, pp. 347-354
- [7] Narayanan, R., Liu, B. and Choudhary, A. (2009), Sentiment Analysis of Conditional Sentences, *Proceedings of Empirical Methods in Natural Language Processing*, pp. 180-189.
- [8] Litman, D. J. and Forbes-Riley, K. (2004) Predicting Student Emotions in Computer-Human Tutoring Dialogue, *Proceedings of the International Symposium on Communication Systems, Networks and DSPs*, pp. 351-358
- [9] 문현구 (2002), *Hybrid Naïve Bayes HMM 기법을 사용한 텍스트로부터의 감정 분류*, 석사학위논문, 서울대학교
- [10] Plutchik, R. and Kellerman, H. (1980), *EMOTION Theory, Research, and Experience*, Academic press, Vol. 1
- [11] Plutchik, R. and Conte, H. R. (1997), *Circumplex models of personality and emotions*, American psychological association
- [12] Plutchik, R. (2001), The Nature of Emotions, *American Scientist*, vol. 89, pp. 344-350, 2001
- [13] Vapnik, V. (1995), *The Nature of Statistical Learning Theory*. Springer Verlag, New York.
- [14] Zhe, X. and Boucouvalas, A. (2002), Text-to-Emotion Engine for Real Time Internet Communication, *Proceedings of the International Symposium on Communication Systems,*

*Networks and DSPs*, pp. 164-168.

- [15] Liu, H., Lieberman, H. and Selker, T. (2003), A Model of Textual Affect Sensing using Real-World Knowledge, *Proceedings of International conference on User Interface*, pp. 125-132.
- [16] Fitrianie, S., Wiggers, P. and Rothkarantz, J.M. L. (2003), A Multi-modal Eliza Using Natural Language Processing and Emotion Recognition, *Proceedings of Text, Speech and Dialog (TSD), LNAI 2807*, pp. 394-399.
- [17] Liscombe, J., Riccardi, G. and Hakkani-Türk, D. (2005), Using context to improve emotion detection in spoken dialog systems, *Proceedings of Interspeech*, pp. 1845-1848.
- [18] Pennebaker, J. W., Mehl, M. R. and Niederhoffer K. G. (2003), Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, Vol. 54, No. 1, pp. 547 - 577.
- [19] 윤애선, 권혁철 (2009), 감정 온톨로지의 구축을 위한 구성요소 분석, **제 21 회 한글 및 한국어 정보처리 학술발표 논문집**, pp. 19-24
- [20] 김향숙 (2001), 한국어 감정 표현 관용어 연구, 박사학위논문, 인하대학교
- [21] 김민정, 박재현, 김상범, 임해창, 이도길 (2008), 한국어 화행 분류를 위한 최적의 자질 인식 및 조합의 비교 연구, **정보과학회논문지 소프트웨어 및 응용, 25권, 11호**, pp. 681-691
- [22] Weston, J. and Watkins, C. (1999), Support vector machines for multiclass pattern recognition, *Proceedings of European Symposium On Artificial Neural Networks*, pp. 219-224.

1 차원고접수 : 2010. 3. 22

2 차원고접수 : 2010. 12. 24

최종게재승인 : 2010. 12. 25

(*Abstract*)

## Emotion Classification of User's Utterance for a Dialogue System

Sangwoo Kang      Hongmin Park      Jungyun Seo

Department of Computer Sciences and Engineering, Sogang University

A dialogue system includes various morphological analyses for recognizing a user's intention from the user's utterances. However, a user can represent various intentions via emotional states in addition to morphological expressions. Thus, a user's emotion recognition can analyze a user's intention in various manners. This paper presents a new method to automatically recognize a user's emotion for a dialogue system. For general emotions, we define nine categories using a psychological approach. For an optimal feature set, we organize a combination of sentential, a priori, and context features. Then, we employ a support vector machine (SVM) that has been widely used in various learning tasks to automatically classify a user's emotions. The experiment results show that our method has a 62.8% F-measure, 15% higher than the reference system.

*Keywords* : *sentiment analysis, spoken dialogue system, mobile, machine learning, automatic classification, robot*