

감정의 표현을 위한 마크업

박성은, 이용규
동국대학교 컴퓨터공학과

Mark-up for Representing Emotion

Sung Eun Park, Yong Kyu Lee
Dept. of Computer Engineering, Dongguk University

요 약

이메일과 같은 텍스트 기반의 서비스 등이 점차 대중화되고 있지만, 이러한 텍스트 기반의 서비스에서는 메시지를 전달할 때 수신자가 필자의 감정 상태를 정확하게 파악하기 어려운 문제가 있다. 이러한 문제를 단편적으로 해결하기 위하여 감정 상태를 나타내는 이모티콘(emoji)을 사용하기도 하지만 이는 보편적이지 않아서 사용하기에 불편한 점이 있다. 따라서 본 논문에서는 이러한 문제를 해결하기 위한 방안으로 일반 텍스트 문서에 감정 태그를 삽입하여 필자의 감정을 표현할 수 있도록 새로운 마크업 언어인 EmoXML(Emotion XML)을 정의한다. 그리고 문장 내에 포함되어 있는 감정 어휘를 인식하여, 관련 감정 태그를 자동으로 생성하고 처리할 수 있는 시스템을 설계한다.

1. 서론

이메일과 같은 텍스트 기반의 서비스 등이 점차로 대중화되고 관련 기술 개발들이 활발히 진행되고 있지만, 이러한 텍스트 기반의 서비스를 통해 메시지를 전달할 때 그 내용은 정확히 전달할 수 있어도 수신자가 필자의 감정 상태를 정확하게 파악하기에는 어려운 문제가 있다. 따라서 이러한 문제를 단편적으로 해결하기 위하여 감정 상태를 나타내는 이모티콘(emoji)을 사용하기도 하지만 이는 정형화되거나 보편적이지 않아서 일반 사용자가 사용하기에는 불편한 점이 있다.

따라서 본 논문에서는 이러한 문제를 해결하기 위하여 일반 텍스트 문서에 감정 태그를 삽입하여 필자의 감정을 표현할 수 있도록 새로운 마크업 언어인 EmoXML(Emotion XML)을 정의한다. 그리고 문장 내에 포함되어 있는 감정 어휘를 인식하여, 관련 감정 태그를 자동으로 생성하고 처리할 수 있는 시스템을 설계한다.

본 논문의 구성은 다음과 같다. 먼저 2절에서는 관련 연구로 감정 인식 기반 연구와 한국어 감정 어휘 분류 연구에 대해서 살펴보고, 3절에서는 이러한 연구를 기반으로 한국어 감정 표현을 분류하며, EmoXML 태그 및 감정 태그를 정의한다. 4절에서는 이러한 감정 태그를 생성 및 변환, 검증 및 처리하는 시스템을 설계하고, 5절에서는 결론 및 향후 연구에 대해 살펴본다.

2. 관련 연구

본 절에서는 감정 인식 기반 연구와 한국어 감정 어휘 분류 연구에 대한 내용을 소개한다.

2.1 감정 인식 기반 연구

기존에 감정 인식 기반 연구 분야에서는 사람의 복잡한 감정을 처리하기 위하여 퍼지(fuzzy) 이론을 사용하거나 이를 적용한 지능적 에이전트를 개발하는 연구가 있었다[2, 4]. 또한 음성 데이터로부터 감정 정

보를 직접 추출하는 연구[3]나, 문장에 나타나는 감정을 처리하기 위하여 이모티콘(emoji)으로 문장 내의 감정을 표시하고, 이를 관련 표정을 짓는 애니메이션(animation)으로 변환하는 연구가 있었다[1]. 그리고 사람을 대신하여 에이전트가 프리젠테이션(presentation)을 할 수 있도록 MPML(Multimodal Presentation Markup Language)을 정의한 연구에서는 감정 태그를 정의함으로써 에이전트의 발표 감정을 표현하려는 연구가 있었다[5].

그러나, 이러한 연구들은 감정 정보를 추출하기 위한 장비나 대용량의 음성 샘플 데이터 구축이 필수적이다. 그리고 감정 태그를 정의한 연구에서는 감정 표현이 다양하지 않고, 발표에 적합한 감정만이 사용자에게 의해 미리 정의되어 있다. 따라서 본 논문에서는 문장에 나타나는 감정 어휘를 자동으로 인식하고 처리함으로써 필자의 감정 정도를 파악할 수 있는 방안을 연구하고자 한다.

2.2 한국어 감정 어휘 분류 연구

감정어의 색인과 검색을 위해서는 감정 어휘가 필요한데, 이는 낱양스가 미묘하여 분명한 분류 체계가 필요하다. 이에 국어학과 분류 사전의 분류 체계를 고찰하여 한국어 감정 어휘를 6가지로 분류한 연구가 있었다[8].

이 연구에서는 한국어 감정 어휘를 대표적으로 기쁨, 슬픔, 놀람, 공포, 혐오, 분노로 분류하고 있으며, 본 논문에서는 이 연구에서 분류하는 한국어 감정 체계를 따라 감정 태그들을 정의한다.

3. 감정 어휘 분류 및 EmoXML 태그 정의

본 절에서는 한국어 감정 어휘를 세부적으로 분류하고, EmoXML 태그와 감정 태그들을 정의하며, EmoXML 스키마 및 표현의 예를 살펴본다.

3.1 감정 어휘 분류

국어학과 분류 사전의 분류 체계를 고려하면 한국어 감정 어휘는 6가지로 분류할 수 있고[8], 각 감정에 해당하는 어휘의 종류는 다음 표 1과 같이 정리할 수 있다.

표 1. 한국어 감정 어휘의 분류[8]

감정	해당 어휘
기쁨	기쁘다, 웃다, 즐겁다, 미소짓다 등
슬픔	슬프다, 애잔하다, 서글프다, 울다 등
놀람	놀라다, 당황하다, 자지리지다, 질리다 등
공포	무섭다, 두렵다, 소름 돋다, 진저리나다 등
혐오	가증스럽다, 미워하다, 싫어하다 등
분노	화내다, 성내다, 노엽다, 분하다 등

또한, 감정 태그는 감정을 표현할 뿐만 아니라, 감정의 정도도 나타낼 수 있어야 하는데, EmoXML에서는 감정의 정도를 세 가지 즉, “강함, 보통, 약함”으로 구분하며 해당 어휘는 표 2와 같이 정리할 수 있다.

표 2. 감정 정도에 대한 어휘의 분류

감정 정도	해당 어휘
강함	매우, 대단히, 아주, 정말, 너무 등
보통	보통, 그러저럭 등
약함	별로, 덜, 적어도, 다소 등

3.2 EmoXML 태그 및 감정 태그 정의

앞 절에서 설명한 감정 어휘들을 EmoXML 태그들로 표 3과 같이 정의할 수 있다.

표 3. EmoXML 태그 종류

태그 종류	태그 설명	태그 종류	태그 설명
<EmoXML>	루트	<anger>	분노
<emotion>	감정 종류	<degree>	감정 정도
<joy>	기쁨		강함
<sadness>	슬픔	<normal>	보통
<surprise>	놀람	<weak>	약함
<fear>	공포	<content>	문장 내용
<hatred>	혐오		

다음의 그림 1은 표 3에서 정의한 EmoXML 태그들의 관계를 나타낸 구조도이다. 즉, 최상위 태그로는 EmoXML이 있고, 하위 태그로 emotion, degree, content 태그들을 포함할 수 있다. Emotion 태그는 각 감정을 나타내는 태그들을 포함하는데 반드시 한 개 이상 표현해야 하며, degree 태그는 감정의 정도를 나타내는 태그들을 포함하고 역시 한 개 이상 표현해야 한다. Content 태그는 해당 감정 어휘를 포함한 문장의 내용을 나타낸다.

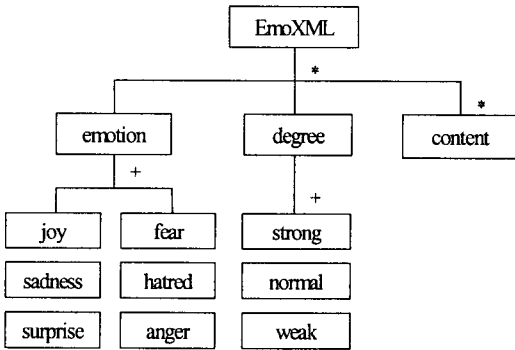


그림 1. EmoXML 태그의 구조도

```
<EmoXML>
  <emotion>
    <joy/>
  </emotion>
  <degree>
    <strong/>
  </degree>
  <content>내가 이번에 대학에 합격해서
    너무 기쁘다.</content>
</EmoXML>
```

그림 3. EmoXML 표현의 예

3.3 EmoXML 스키마 및 표현의 예

다음 그림 2는 그림 1에서 정의한 EmoXML 태그의 구조를 스키마로 표현한 것이다. 각 감정을 나타내는 joy, sadness, surprise, fear, hatred, anger 태그들과 감정의 정도를 나타내는 strong, normal, weak 태그들은 값을 갖지 않은 빈(empty)태그로 표현한다.

```
<?xml version="1.0" encoding="UTF-8"?>
<Schema name="Untitled-schema"
xmlns="urn:schemas-microsoft-com:xml-data"
xmlns:dt="urn:schemas-microsoft-com:datatypes">
  <ElementType name="EmoXML" model="closed"
content="eltOnly" order="seq">
  <ElementType name="EmoXML" model="closed"
content="eltOnly" order="seq">
  <AttributeType name="xmlns" dt:type="string"/>
  <attribute type="xmlns"/>
  <element type="emotion" minOccurs="1" maxOccurs="1"/>
  <element type="degree" minOccurs="1" maxOccurs="1"/>
  <element type="content" minOccurs="1" maxOccurs="1"/>
  </ElementType>
  <ElementType name="anger" model="closed"
content="empty"/>
  .....
  <ElementType name="surprise" model="closed"
content="empty"/>
  <ElementType name="weak" model="closed"
content="empty"/>
</Schema>
```

그림 2. EmoXML 스키마

그림 3은 그림 2에서 설명한 EmoXML 스키마에 유효한 문서의 예이다. 문장 내에 기쁨과 기쁨의 정도를 나타내는 단어를 인식하여 관련 태그들로 표현한 문서이다.

4. 감정 태그 생성 시스템 설계

본 절에서는 지금까지 살펴본 감정 태그들을 자동으로 생성하여 처리하는 시스템을 설계한다.

4.1 감정 태그 생성 절차

다음 그림 4는 감정 태그 생성 절차를 순서도로 나타낸 그림이다.

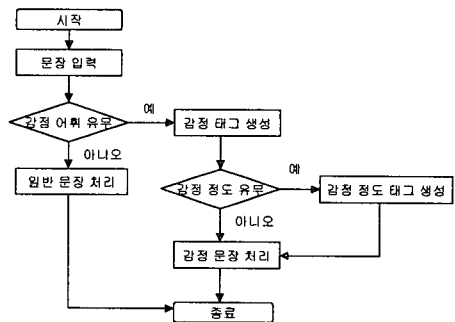


그림 4. 감정 태그 생성 절차

즉, 입력받은 문장 안에 감정 어휘가 있는지를 분석하여 감정 문장으로 처리할 것인지를 결정한 후, 감정 문장으로 분류되면 문장 안에 있는 감정 어휘에 해당하는 감정 태그를 생성한다. 이를 생성한 후에 문장 안에 감정 정도를 나타내는 단어가 있다면 감정 정도를 나타내는 태그를 생성하여 감정 문장으로 처리한다.

4.2 감정 태그 생성 시스템의 구조

앞 절에서 설명한 감정 태그 생성기를 포함한 감정 태그 생성 시스템의 구조는 다음 그림 5와 같으며, 이

는 문서 입력기, 감정 어휘 분석기, 감정 태그 생성기, 감정 태그 검증기로 구성된다.

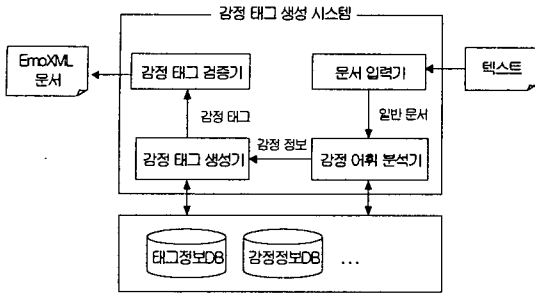


그림 5. 감정 태그 생성 시스템의 개요

먼저, 문서 입력기를 통해 일반 텍스트 문서를 받아들이면, 감정 어휘 분석기에서는 감정 정보 데이터베이스 안에 저장된 감정 정보를 기반으로 패턴 매칭 (Pattern Matching) 방법을 사용하여 문장 내에 있는 감정 어휘를 분석한다. 감정 어휘가 분석되면 감정 태그 생성기에서 관련 감정 태그들을 생성하며, 감정 태그 검증기를 통하여 태그들의 유효성을 검증한 후 최종적으로 EmoXML 문서를 생성한다.

4.3 감정 태그 생성기 모듈 설계

앞 절에서 설명한 감정 태그 생성 절차를 수행할 수 있는 감정 태그 생성기 모듈을 설계하면 다음 그림 6과 같다. 즉, 감정 태그 생성 모듈에서는 감정 정보를 입력으로 받아서 문장 내에 있는 감정 어휘를 분석한 후 감정 태그를 생성하고, 감정 정보 태그 생성 모듈에서는 감정 정보 어휘를 분석하여 감정의 정도를 나타내는 태그를 생성한다.

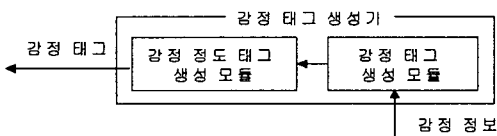


그림 6. 감정 태그 생성기 모듈 설계도

5. 결론 및 향후 연구

이미 대중화된 이메일과 같은 텍스트 기반의 서비스에서는 메시지를 전달할 경우에 그 내용은 정확하게 전달할 수 있어도 필자의 감정은 제한된 상태로 전달되므로 수신자가 필자의 감정 상태를 정확히 파

악하기에 어려운 문제가 있다.

따라서 본 논문에서는 이러한 문제를 해결하기 위하여 일반 텍스트 문서에 감정 태그를 삽입함으로써 메시지에 나타나는 필자의 감정을 자동으로 처리할 수 있는 방안을 고려했다. 이를 위해 먼저 한국어 감정 어휘들을 살펴보고, 대표적인 6가지 감정을 표현하는 태그들을 정의하였다. 그리고 이를 포함한 EmoXML의 태그들을 정의하고, 이를 자동으로 생성하고 처리할 수 있는 시스템을 설계하였다.

향후에는 감정 어휘의 다양한 어근 어미의 변화를 반영하고, 본 논문에서 설계한 감정 표현 및 처리 시스템을 실제 구현하고, 이를 적용한 응용 시스템에 대해 연구한다.

[참고문헌]

- [1] I. Albrecht, et. al. "May I talk to you? :-) - Facial Animation from Text," Proc. of the 10th Pacific Conference on Computer Graphics and Applications, pp. 77-86, 2002.
- [2] M. S. El-Nasr and J. Yen, "Agent, Emotional Intelligence and Fuzzy Logic," Proc. of the Fuzzy Information Processing Society - NAFIPS, pp. 301-305, Pensacola Beach, FL, USA, 1998.
- [3] S. Fukuda and V. Kostov, "Extracting Emotion from Voice," Proc. of the IEEE SMC '99 Conference, pp. 299-304, Tokyo, Japan, 1999.
- [4] T. Yanaru, "An Emotion Processing System Based on Fuzzy inference and Subjective Observations," Proc. of the Artificial Neural Networks and Expert Systems, pp. 15-25, Dunedin, New Zealand, 1995.
- [5] Y. Zong, H. Dohi, H. Prendinger and M. Ishizuka, "Emotion Expression Function in Multimodal Presentation," Proc. of the ICMI 2000, LNCS, vol. 1948, pp. 57-64, Beijing, China, 2000.
- [6] 강행봉, "감정 기반 컴퓨팅," 전자공학회지, 제28권 12호, pp. 53-60, 2001.
- [7] 김명관, 박영택, "감정기반 정보 검색 시스템에 관한 연구," 한국문헌정보학회지, 제32권 4호, pp. 105-115, 1998.
- [8] 장효진, "감정동사 및 감정형용사 분류에 관한 연구," 한국정보관리학회 학술대회 논문집, 제8권, pp. 29-34, 2001.