

2018년도
학사학위논문

한국어 구문 분석을 통한 사용자 감성 분석 어플리케이션 개발

Development of the User Sentiment
Analysis Application through the Korean
Syntax Analysis

2018년 12월 28일

순천향대학교 공과대학
컴퓨터공학과

고강문

2018년도
학사학위논문

한국어 구문 분석을 통한 사용자 감성 분석 어플리케이션 개발

Development of the User Sentiment
Analysis Application through the Korean
Syntax Analysis

2018년 12월 28일

순천향대학교 공과대학
컴퓨터공학과

고강문

한국어 구문 분석을 통한 사용자 감성 분석
어플리케이션 개발

Development of the User Sentiment
Analysis Application through the Korean
Syntax Analysis

지도교수 하 상 호

이 논문을 공학사학위 논문으로 제출함

2018년 12월 28일

순천향대학교 공과대학
컴퓨터공학과

고강문

고 강 문 의 공학사학위논문을 인준함

2018년 12월 28일

심 사 위 원 천 인 국 인
지 도 교 수 하 상 호 인

순천향대학교 공과대학

컴퓨터공학과

초록

서비스업의 발달로 인한 감정노동자의 증가는 감정노동에 따른 직무 스트레스의 증가와 자신의 감정 상태에 대한 인식 부족 현상으로 이어지고 있다. 자신의 감정 상태에 대한 인식이 부족하여 생기는 감정표현 불능증은 여러 신체장애와 정신장애를 유발하는 요인이 될 수 있기에 매우 중요한 문제 중 하나이다. 감성 분석이란 오피니언 마이닝(Opinion Mining) 분야 중 하나로 텍스트의 내용에 존재하는 감성을 긍정, 중립, 부정 중 하나로 분석하는 분야이다. 현재까지의 감성 분석 연구는 문장의 극성을 구분하기 위해 다양한 방법이 연구되었지만, 그 활용이 문장의 극성을 분류하는데 그쳐 문장의 상세한 감성 수치를 분석할 수 없다는 문제점이 존재한다.

본 논문에서는 이러한 문제점을 해결하기 위해 사용자로부터 입력받은 일기 데이터를 분석하여 사용자에게 일기 데이터 상의 감성 수치를 보여주는 어플리케이션을 제안한다. 제안한 어플리케이션을 이용하여 사용자가 일기를 작성할 경우 어플리케이션에서는 해당 일기 내용에 대한 구문 분석 과정과 감성 분석 과정을 통해 사용자가 작성한 일기 내용의 감성 수치를 계산하게 된다. 구문 분석 과정에서는 문장의 형태소 분석을 통해 생성되는 문장의 구성 형태소들을 본 논문에서 제안하는 한국어 문형 구조에 맞게 구문 트리를 생성한다. 감성 분석 과정에서는 본 논문에서 제안된 감성 사전에 등록된 감성 단어의 감성 수치를 활용한 감성 수치 연산을 실시한다. 어플리케이션을 통해 분석된 결과는 저장되어 일자에 따른 감성 수치의 변화를 사용자에게 보여줌으로써 사용자가 스스로의 감정 상태를 파악하는데 있어 도움을 주는 것을 그 목적으로 한다.

주요어 : 감정표현 불능증, 구문 분석, 감성 분석

ABSTRACT

The increase in emotional workers due to the development of the service industry leads to increased job stress and a lack of awareness of emotional state. Alexithymia caused by lack of awareness of emotional state can be a factor causing various physical and mental disorders, which is one of the key issues. The sentiment analysis is one of Opinion Mining field, this analyzes the sentiment that exists in the contents of text, either positive, neutral or negative. Until this days, studies have been researched in various ways to distinguish the polarity of sentences, but the problem is that the use only categorizes the polarity of sentences, and thus cannot analyze the detailed sentiment level of sentences.

In this paper, we propose an application that analyzes the diary data input from the user and shows the sentiment value on the diary data to the user to solve these problems. When a user writes a diary by using the application, it calculates the sentiment value of the user's diary contents through the syntax analysis and sentiment analysis of the diary content. The Syntax analysis creates a syntax tree to match the structure of the Korean sentence structure proposed in this paper. The sentiment analysis involves the calculation of sentiment value using sentiment words registered in the sentiment dictionary proposed in this paper. The analyzed results through application are stored and helps the user understand their sentiment state by showing the user the change in the sentiment value by date.

Keywords : Alexithymia, Syntax Analysis, Sentiment Analysis

차 례

제 1 장 서 론	1
제 2 장 한국어 감성 분석 기술	6
2.1 한국어 문형 구조	6
2.1.1 한국어 문형 구조 정립	7
2.2 구문 분석	9
2.2.1 형태소 분석	10
2.2.2 형태소 필터링	12
2.2.3 구문 트리 정립	16
2.2.4 품사 변환	17
2.3 감성 분석	18
2.3.1 감성 사전	18
2.3.2 한국어의 부정 표현	19
2.3.3 감성 수치 계산	20
제 3 장 시스템 설계	21
3.1 전체 시스템 구조	21
3.2 감성 분석 알고리즘	22
3.3 구문 분석	24
3.4 감성 분석	25
3.5 알고리즘 복잡도	28

제 4 장 개발 환경 및 도구.....	30
4.1 개발 환경.....	30
4.2 개발 도구.....	30
제 5 장 실험 및 결과.....	31
5.1 실험.....	31
5.2 실험 결과.....	32
5.3 어플리케이션 검증.....	34
제 6 장 결론 및 향후 과제.....	43
참고문헌.....	44
감사의 글.....	46
부록.....	47
1 꼬꼬마 형태소 분석기 품사 태그 표.....	47
2 한국어 구문 트리.....	48
3 감성 사전.....	65

표 차 례

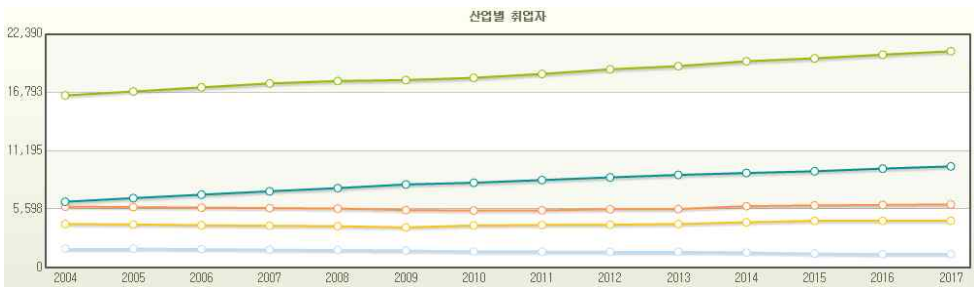
[표 1] 정립된 문형 구조.....	8
[표 2] 문형 구조에서의 명사의 형태.....	12
[표 3] 감성 분석 시 사용되는 형태소.....	14
[표 4] 특정 단어만이 구문 트리에 사용되는 형태소.....	15
[표 5] 감성 분석 어플리케이션 개발 환경.....	30
[표 6] 실험을 위해 사용한 스마트폰의 사양.....	31
[표 7] 감성 분석 결과.....	32
[표 8] 문장의 형태소 분석 결과.....	35
[표 9] 형태소 필터링 과정을 거친 문장의 형태소.....	36

그 립 차 례

[그림 1] 서비스 산업별 취업자 증감 추이	1
[그림 2] 스트레스 인지율 추이	2
[그림 3] 연도별 1일 외래환자 수	3
[그림 4] 꼬꼬마 한국어 형태소 분석기 모식도	11
[그림 5] 동사 중심 구문 트리 구조	16
[그림 6] 형용사, 명사 중심 구문 트리 구조	16
[그림 7] 전체 시스템 구조	21
[그림 8] 의사코드로 표현한 감성 분석 알고리즘	23
[그림 9] 부정 표현의 감성 수치 계산	26
[그림 10] 원래 의미에 따른 감성 수치 계산	26
[그림 11] 사용자 감성 상태 변화 추이 예시	33
[그림 12] 검증을 위해 사용된 일기 데이터	34
[그림 13] 3번 문장에 대한 구문 트리 구조	39
[그림 14] 3번 문장에 대한 감성 구문 트리	41
[그림 15] 6번째 일기 데이터에 대한 전체 일기 데이터의 감성 수치	42

제 1 장 서 론

고도성장기가 끝난 1990년대 이후 대한민국에 분 탈공업화 현상은 서비스 산업의 발전을 이끌었고, 이로 인해 서비스업이 우리 경제에서 차지하는 비중은 나날이 증가하고 있는 추세이다. 서비스업의 발달은 서비스 산업에 종사하는 사람들의 증가로도 이어졌으며 그에 따라 현재 대한민국의 서비스 산업 규모는 약 2천만 명이 종사하는 대한민국에서 가장 발달한 산업이 되었다[1].



[그림 1] 서비스 산업별 취업자 증감 추이[1]

서비스 산업 종사자가 많아짐에 따라 고객을 직접 대면하거나 일대일로 대화하면서 고객에서 편안함과 만족감을 주어 소비활동을 돕는 업무에 종사하는 감정노동자 또한 증가하였는데, 전체 임금 근로자의 38~42%에 해당하는 최소 550만 명에서 최대 740만 명까지의 노동자가 감정노동 종사자로 추정된다[2]. 하지만 감정노동자는 직업상 고객을 대할 때, 자신의 감정이 좋거나 슬프거나 화나는 상황이 있더라도 회사에서 요구하는 감정과 표현만을 고객에게 보여주는 고객응대 업무를 한다는 직업적 특징을 갖고 있다[3].

다른 직종보다 높은 직무스트레스에 시달리는 감정 노동자의 증가와 취업난은 국내에서 일상생활 중 스트레스를 느낀다고 답하는 비율인 스트레스 인지율의 수치가 2008년 불어 닥친 세계 금융위기 당시의 스트레스 인지율의 수치와 필적할 정도로 증가하는 추세에 있다[4].



[그림 2] 스트레스 인지율 추이[4]

자신의 감정을 표현할 수 없다는 감정노동자의 또 다른 직업적 특징은 감정 노동자들이 자신의 감정을 억지로 숨기려 하는 스마일 마스크 증후군에 걸릴 확률을 높게 만든다. 만일 자신의 감정을 숨김으로써 나타나는 스마일 마스크 증후군이나 가면 성 우울증을 방지할 경우, 자신의 감정을 계속하여 억누른 나머지 본인이 어떤 감정을 느끼는지도 모르는 상태가 될 수 있다[5].

이렇게 자신의 감정을 자각하지 못하는 증상은 감정표현 불능증이라고 불리는데, 불안장애, 외상 후 스트레스장애, 섭식장애, 우울장애, 성 장애, 성격장애 등 다양한 정신장애와 소화불량, 편두통, 위장장애와 같은 신체적인 질병이 나타나는 신체화 장애 현상을 유발한다[6].

이러한 감정노동자의 높은 직무 스트레스를 받지만 감정표현의 어려움을 겪는 상황은 신체장애와 정신장애의 유발 가능성을 높이게 된다. 그로인해 대한민국의 연도별 1일 외래환자 수는 날로 증가하는 추세이다[7].



[그림 3] 연도별 1일 외래환자 수[7]

오피니언 마이닝(Opinion Mining) 분야 중 하나인 감성 분석은 텍스트의 내용에 대해 감성의 극성을 긍정, 중립, 부정 중 하나로 분석하는 분야이다. 이러한 사용자의 주관이 드러난 문장은 해당 문장을 접하는 다른 사용자들에게도 영향을 미치기에, 감성 분석은 공공 기관 또는 기업체등에서 사용자의 성향을 분석하거나 마케팅에 관한 의사결정을 하는데 있어 중요한 지표로 활용되고 있다[8].

그로인해 최근 사용자의 감성을 분석하는 연구에서는 기업의 마케팅을 위한 사용자 경험의 감성 분석 등을 위해 문장의 극성을 분류하는 방법에 대한 연구가 활발히 이루어지고 있지만 그 활용이 문장의 극성을 분류하는데 그쳐 문장의 상세한 감성 수치를 분석할 수 없다는 문제점이 존재한다. 이러한 문제점은 자세한 감성 수치의 분석을 통해 이루어지는 사용자의 감성 파악 과정에서 어려움을 초래한다.

본 논문에서는 이러한 스스로에 대한 감정인식 부족으로 발생하는 다양한 신체적, 정신적 장애 유발 가능성 및 현재까지의 감성 분석 연구가 안고 있는 문제점을 해결하기 위해 사용자의 일기내용을 바탕으로 자신의 감성을 올바르게 파악할 수 있도록 하는 것을 목적으로 감성 분석을 진행하며, 사용자가 언제 어디서나 간편하게 일기를 작성하고, 그에 맞는 자신의 감성 상태를 파악하기 용이하도록 안드로이드 스마트폰 어플리케이션을 개발하였다.

문장의 감성은 문장을 구성하는 어휘들에 의해 구성되는데 이러한 문장들의 감성을 분석하는 방법에는 주로 단어를 문장의 감성 극성을 분류하는 자질인 말뭉치로 이용하는 단어 기반 시스템이 활용된다.

단어 기반 시스템을 활용한 감성 분석은 각 단어 사이의 연관성을 확률적으로 계산하여 해당 문장의 감성을 분석하는 확률론 기반 감성 분석 방법과 기존에 구축된 말뭉치를 기반으로 감성 분석을 수행하는 사전 기반 감성 분석 방법, 인공지능을 활용한 감성 분석 등이 사용된다.

확률론에 기반을 둔 한국어의 감성 분석을 위해 Turney(2002)는 두 개의 연속적인 단어를 한 쌍으로 하여 패턴을 구성하여 이를 추출하고 이러한 구문 패턴에 대해서 PMI(Pointwise Mutual Information)를 계산하여 추출한 각 패턴의 극성 강도를 계산하였다[9]. 확률 기반 감성 분석은 SO-PMI라고도 불리며, 기존에 미리 전문가에 의해 설정된 긍정 어휘와 부정 어휘들의 집합인 기준 어휘를 이용하여, 분석하고자 하는 문장과 기준 어휘와의 차이를 통해 어휘의 극성 값을 구하는 방법이다[10, 11].

사전 기반 감성분석은 구축된 감성 사전의 단어들을 감성 분석을 위한 말뭉치로써 사용하는 기법으로, 기존에 구축된 감성 사전의 내용에 따라 성능이 크게 변한다는 특징을 갖고 있다. 또한 이러한 사전 기반 감성 분석 방법은 한국어의 특징을 포함한 분석이 아닌, 형태소 분석기를 이용하여 문장의 어휘를 감성을 나타내는 말뭉치로 추출하고 이에 대한 가중치를 계산하는 방식으로 이루어져왔다. 이는 감성 단어의 빈도수를 이용하여 감성을 분석하는 과정으로, 어순이 자유롭고 어미의 변화로 극성을 다르게 나타내는 한국어의 구문적인 특징을 나타내기에는 적합하지 않다.

앞서 언급한 대로 현재까지의 한국어 감성 분석 연구들은 문장이 어떠한 극성을 나타내는지에 초점을 맞추어 연구를 진행하여 각각의 문장에 대한 상세한 감성 수치는 연구되어 있지 않다는 문제점을 보완하기 위하여, 본 논문에서는 사용자가 작성한 텍스트에 대한 극성 여부의 판단만이 아닌 해당 텍스트의 상세한 감성 수치를 파악하고 감성 수치의 변화를 보여준다는 특징이 있다.

본 논문에서 제안하는 어플리케이션은 해당 어플리케이션을 이용하는 사용자로 하여금 올바른 자신의 감정상태 파악을 통해 스스로의 감정을 깨닫고 그로 말미암아 신체화 장애를 예방하여 몸과 마음의 건강을 유지할 수 있도록 도움을 주는 것이 목적이며, 일기로 적은 한국어 일기 데이터의 구문 분석 과정과 감성 분석 과정을 통해 해당 일기의 감성 수치를 산출, 해당 일기 데이터에 적혀있는 사용자의 감성 수치를 스마트폰 화면에 보여준다.

제 2 장 한국어 감성 분석 기술

2.1 한국어 문형 구조

한국어는 SOV 형태를 갖는 자연언어이며, 자연언어는 형식언어이론에 따라 문맥 의존 문법을 따르는 언어이다. 문맥 의존 문법이란 생성규칙에 의해 치환될 기호뿐만 아니라 앞뒤의 구절, 문맥의 영향을 받는 언어를 말한다.

본 논문에서는 꼬꼬마 한국어 형태소 분석기를 통해 얻어진 문장을 구성하는 형태소들을 통한 해당 문장의 구조 해석 및 문장 구조 확정을 위해 [12]에서 연구한 조선어 문형 연구를 통해 기분석된 한국어 문형 구조를 사용하였다.

조선어 문형 연구에서는 한국어 문형 구조를 크게 동사술어기본문형, 형용사술어기본문형, 명사술어기본문형으로 나누고 이 중 동사술어기본문형은 자동사술어기본문형, 타동사술어기본문형, 양면동사술어기본문형으로 상세 분류를 하는 등의 과정을 통하여 모든 한국어 문형 구조를 41개의 문형 구조로 분류하였다.

각각의 문형 구조는 명사, 형용사, 동사, 부사와 그에 결합된 조사 또는 어미를 구성요소로 하며 각각의 구성요소의 순서에 따라 서로 다른 문형 구조가 나타난다. 이는 한국어를 포함한 자연언어의 형식언어이론에 따른 분류인 문맥 의존 문법의 특징을 따르는 문장구조 분석을 보여준다.

2.1.1 문형 구조 정립

2.1절에서 제안한 한국어 문형 구조는 동사술어기본문형을 자동사술어기본문형, 타동사술어기본문형, 양면동사술어기본문형으로 상세하게 나누었지만 양면동사술어기본문형은 특정 동사가 자동사와 타동사 모두의 의미로 사용되는 경우의 문형 구조를 의미하여 해당 문형 구조의 구조적인 설계는 자동사술어기본문형, 타동사술어기본문형에 모두 포함되는 특징이 있다.

본 논문에서 문장 트리 생성을 위해 사용되는 문형 구조는 각 문형 구조의 구조적인 구성요소를 통하여 문장트리를 생성함을 목적으로 하여 본 논문에서는 2.1절에서 제안한 한국어 문형 구조 중 타동사술어기본문형은 문형 구조 정립 시 사용하지 않았다.

또한 문형 구조를 구성하는 구성요소의 특징으로, 수의적 구성요소에 따른 문형의 확대가 일어나기에 본 논문에서 개발한 감성 분석 어플리케이션을 위한 문형 구조를 문형에 따른 수의적 구성요소를 포함하여 정립하였다.

위와 같은 과정을 통하여 본 논문에서는 20개의 동사 중심 문형 구조와 8개의 형용사 중심 문형 구조, 6개의 명사 중심 문형 구조를 정립하였다. 정립된 문형 구조를 [표 1]에서 보이고 있다.

[표 1] 정립된 문형 구조

동사 중심 문형 구조	1번 문형	N이 + V
	2번 문형	N이 + N이 + V
	3번 문형	N이 + N와 + N이 + V
	4번 문형	N이 + N에 + V
	5번 문형	N이 + N를 + N와 + N에 + V
	6번 문형	N이 + N로 + V
	7번 문형	N이 + N를 + N로 + V
	8번 문형	N이 + N와 + V
	9번 문형	N이 + N에서 + V
	10번 문형	N이 + N를 위해 + V
	11번 문형	N이 + N에 의해 + V
	12번 문형	N이 + AD + V
	13번 문형	N이 + N를 + AD + V
	14번 문형	N이 + N를 + V
	15번 문형	N이 + N에 + N를 + V
	16번 문형	N이 + N와 + N를 + V
	17번 문형	N이 + N에 + N와 + N를 + V
	18번 문형	N이 + N에서 + N를 + V
	19번 문형	N이 + N와 + N에 대해 + N를 + V
	20번 문형	N이 + N를 + N라고 + V
형용사 중심 문형 구조	1번 문형	N이 + A
	2번 문형	N이 + N이 + A
	3번 문형	N이 + N로 + N이 + A
	4번 문형	N이 + N와 + N이 + A
	5번 문형	N이 + N에 + A
	6번 문형	N이 + N와 + A
	7번 문형	N이 + N보다 + A
	8번 문형	N이 + N로 + A
명사 중심 문형 구조	1번 문형	N이 + N이다
	2번 문형	N이 + N이 + N이다
	3번 문형	N이 + N에 + N이다
	4번 문형	N이 + N와 + N이다
	5번 문형	N이 + N보다 + N이다
	6번 문형	N이 + N에 대해 + N이다

2.2 구문 분석

언어학에서 구문 분석은 문장을 문장의 구성요소로 분해한 뒤 각각의 구성요소 사이의 위계관계를 분석하여 문장의 구조를 결정하는 과정을 시행함을 말한다. 특히 한국어와 같은 자연언어의 구문 분석에서는 뜻을 가진 가장 작은 단위인 형태소는 토큰, 즉 하나의 구성요소로 분해되어 구문 분석 시 사용된다. 형태소 분석이 끝나면 해당 형태소들로 이루어진 구 묶음을 생성하고, 위와 같은 과정을 통해 작성된 구 묶음을 통해 해당 문장의 구조를 해석함과 동시에 문장의 구조를 명백히 밝히는 것이 구문 분석 이다.

한국어 구문 분석에 있어, 본 논문에서는 의존문법을 기반으로 한 파싱방법인 의존 파싱(Dependency Parsing)을 사용하였다. 의존 파싱방법의 핵심은 한 의존소에 대한 여러 지배소 후보들 중 가장 적합한 지배소를 선택하는 것인데, 한국어에서는 지배소가 의존소의 뒤에 오고 의존 관계들이 서로 교차하지 않는다는 특성을 가지고 있으며, 문장의 핵심 지배소가 문장의 제일 뒷부분에 위치한다는 특징을 가지고 있다. 이를 통해 본 논문에서는 구문 분석을 시행함에 있어 문장의 끝부터 구문 분석을 시행한다.

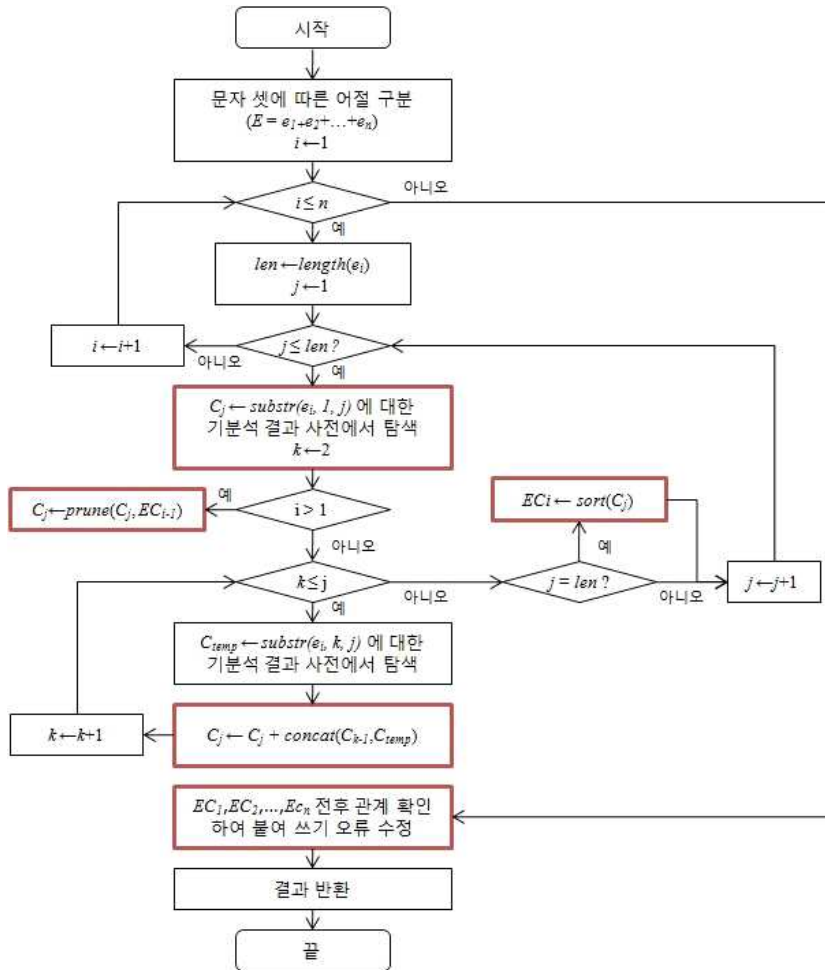
2.2.1 형태소 분석

문장에서 형태소를 분해하고 이를 분석하는 형태소 분석과정은 구문 분석의 최초 단계에서 수행된다. 본 논문에서는 형태소 분석을 위해 [13]에서 개발한 꼬꼬마 한국어 형태소 분석기를 활용하였다.

꼬꼬마 한국어 형태소 분석기는 동적 프로그래밍(Dynamic Programming)을 통해 형태소 분석을 수행하며, 기분석 사건을 이용한 인접 조건 검사에 의한 형태소 분석 과정을 통해 모든 가능한 형태소 분석 후보를 생성하고 확률 모델을 이용하여 최적 분석 후보를 선정하는 과정을 통해 형태소 분석을 수행한다. 인접 조건 검사는 품사 결합 조건, 음운 결합 조건, 형태 결합 조건의 요소를 갖고 있는 형태소 결합 조건을 통해 판단하는 과정으로, 꼬꼬마 한국어 형태소 분석기에서는 결합 선호 조건을 추가하여 형태소 분석을 수행한다.

꼬꼬마 한국어 형태소 분석기는 한글 형태소의 품사를 체언, 용언, 관형사, 부사, 감탄사, 조사, 어미, 접사, 어근, 부호, 한글 이외와 같이 나누고 각 세부 품사를 구분하여 표기한 형태소 분석 결과를 반환한다.

꼬꼬마 형태소 분석기는 [그림 4]와 같은 과정을 통해 형태소 분석을 실시하며, [부록 1]에 기술된 품사 태그 표를 바탕으로 각각의 형태소에 알맞은 형태소 품사를 분석결과에 포함시킨다.



[그림 4] 꼬꼬마 한국어 형태소 분석기 모식도[13]

2.2.2 형태소 필터링

본 논문에서 감성 분석 과정에 사용되는 꼬꼬마 한국어 형태소 분석기는 한국어 문장을 형태소 분석하는 과정을 통해 문장에 사용되는 각각의 형태소를 60 종류의 형태소 중 하나로 구분한다.

2.1.1절에서 정립한 한국어 문형 구조 목록에서 문형을 구성하는 핵심적인 구성요소는 명사, 동사, 형용사이며 동사와 형용사는 문형 구조상에서 독립적인 구성요소로 사용되지만 명사는 독립적인 구성요소로 사용되지 못하고 의존소로써 사용되어 각 명사의 지배소가 필요하다는 특징을 보인다.

2.1.1절에서 정립한 한국어 문형 구조 목록에서 문형 구조를 구성하는 지배소를 포함한 명사의 형태는 [표 2]에서 보이고 있다.

[표 2] 문형 구조에서의 명사의 형태

구성요소	의미	형태
N이	무엇이, 누가	~은, ~는, ~이, ~가
N에	무엇에, 누구에게, 어디에	~에, ~에게, ~께
N로	무엇으로, 누구로	~으로, ~로, ~이, ~가
N와	무엇과, 누구와	~과, ~와
N에서	무엇에서, 누구에서, 어디에서	~에서, ~에게
N를 위해	무엇을 위해, 누구를 위해	~을 위해, ~를 위해
N에 의해	무엇에 의해, 누구에 의해	~에 의해
N를	무엇을, 누구를	~을, ~를
N라고	무엇이라고	~라고, ~이라고, ~이라
N에 대해	무엇에 대해	~에 대해
N보다	무엇보다, 누구보다	~보다
N이다	무엇이나	~이다, ~이었다, ~이었다
N에	무엇에, 누구에게, 무엇에 대해	~에, ~에게, ~에 대해, ~한테도

이러한 문형 구조의 구성요소로서 사용되는 각각의 명사는 어근, 연결 어미, 종결 어미, 전성 어미, 선어말 어미, 조사, 보조사, 지정사를 지배소로 갖는다. 이러한 문형 구조상의 명사 지배소는 조사인 ‘~보다’처럼 하나의 지배소가 하나의 명사를 수식하기도 하며, ‘~이다’처럼 지정사인 ‘~이’와 전성 어미인 ‘~다’가 합쳐지는 여러 개의 지배소 집합이 하나의 명사를 수식하기도 한다.

문형 구조의 구성요소로 사용되는 명사의 지배소는 꼬꼬마 한국어 형태소 분석기가 구분하는 60개의 형태소 중 16개의 형태소를 포함한다.

최종적으로 감성 분석 시 사용하는 형태소에는 앞서 언급한 16개의 명사 지배소 외에도 명사 지배소 형태소의 의존소인 명사, 문형 구조상의 독립적인 구성요소인 형용사와 동사, 형태소의 품사를 변환시켜주는 파생 접미사, 부정 표현 인식을 위한 부사 등이 포함되어 꼬꼬마 한국어 형태소 분석기가 구분하는 60개의 형태소 중 30개의 형태소만을 감성 분석 시 사용되는 분석 후보군이 된다.

감성 분석 시 형태소 분석기에서 분석되는 60개의 형태소 중 30개의 형태소만을 사용한다는 의미는 형태소 분석기에서 나오는 형태소 분석 결과를 그대로 사용할 경우, 감성 분석에서 분석 후보군이 되는 형태소가 필요 이상으로 많아진다는 것을 의미한다.

감성 분석에서 분석 후보군이 되는 형태소가 많은 경우, 연산 과정이 필요 이상으로 증가하여 분석 과정에서의 소요시간이 증가하고, 보조사와 관형어, 접속조사 등 문형 구조에서 구성 요소로써 동작하지 않는 형태소들로 인하여 지배소 선정 과정에서 중의성이 발생된다는 2가지의 문제점이 발생한다.

이러한 문제점을 해결하기 위해 형태소 분석 이후 문형 구조를 통한 문형 구조 생성 이전 형태소 분석 결과를 통해 도출되는 문장을 구성하는 형태소들 중 문형 구성요소에서 사용하지 않는 형태소를 문장 트리 생성 시 사용하지 않도록 제외하는 형태소 필터링 과정을 수행하였고 이를 [표 3]에서 보이고 있다.

[표 3] 감성 분석 시 사용되는 형태소

분류	형태소 코드	형태소 설명
어근	XR	어근
연결 어미	ECD	의존적 연결 어미
	ECS	보조적 연결 어미
종결 어미	EFA	청유형 종결 어미
	EFI	감탄형 종결 어미
	EFO	명령형 종결 어미
	EFQ	의무형 종결 어미
	EFR	존칭형 종결 어미
전성 어미	ETD	관형형 전성 어미
	EFN	명사형 전성 어미
선어말 어미	EPT	시제 선어말 어미
조사	JKM	부사격 조사
	JKO	목적격 조사
	JKS	주격 조사
	JX	보조사
지정사	VCP	긍정지정사, 서술격조사
부사	MAG	부사
형용사	VA	형용사
	VXA	보조 형용사
	XSA	형용사 파생 접미사
명사	NNG	보통 명사
	NNB	일반 의존 명사
	NNM	단위 의존 명사
	NNP	고유 명사
	NP	대명사
	NR	수사
	XSN	명사 파생 접미사
동사	VV	동사
	VXV	보조 동사
	XSV	동사 파생 접미사

또한 형태소 필터링 과정에서 명사의 지배소가 포함되는 형태소 품사들 중 특정 형태소만이 명사의 지배소로 쓰여 구문 트리 구성에 사용되는 품사가 존재하는데, 이러한 품사를 [표 4]에서 보이고 있다.

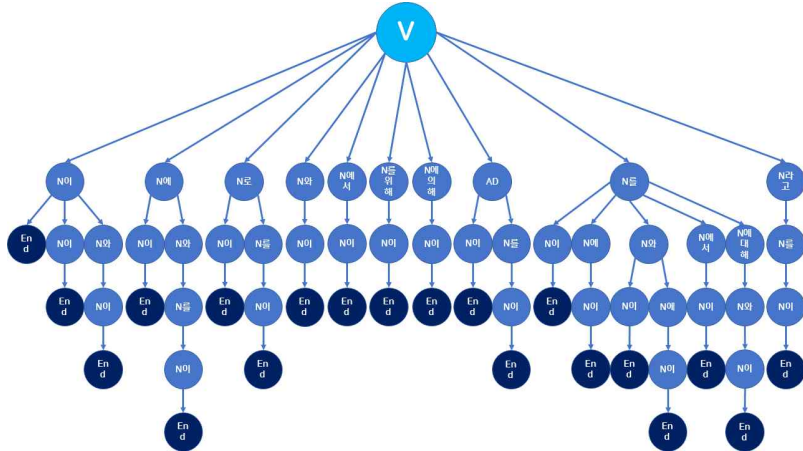
[표 4] 특정 단어만이 구문 트리에 사용되는 형태소

형태소 품사	사용 형태소
ECD	라고, 라, 지
ECS	어
ETD	ㄴ
EPT	었, 었
JKM	에, 에게, 께, 으로, 로, 과, 와, 에서, 에게, 보다
JKO	을, 를
JKS	이, 가
JX	은, 는, 라, 고, 도
VCP	이

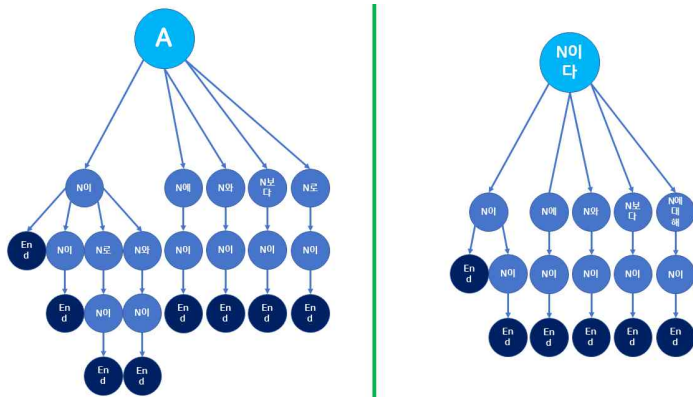
이러한 과정을 통해 구문 트리에 필수적으로 사용되는 품사들만 구문 분석에 사용되도록 형태소 필터링 과정이 수행된다.

2.2.3 구문 트리 정립

2.1.1절에서 정립된 한국어 문형 구조와 [14]에서 사용된 한국어 트리 구조를 토대로 한 구문 트리 구조를 정립하였다.



[그림 5] 동사 중심 구문 트리 구조



[그림 6] 형용사, 명사 중심 구문 트리 구조

[그림 5]와 [그림 6]은 본 논문에서 개발한 감성 분석 어플리케이션에서 실제 감성 분석에 사용되는 구문 트리 구조를 보여준다.

앞서 2.2절에서 설명한 대로 한국어는 문장의 핵심 지배소가 문장의 제일 뒷 부분에 위치한다는 특징을 갖고 있어 문장의 가장 마지막에 나오는 구성요소를 통해 문형 구조 구분을 실시하였다. 또한 문장의 끝부터 구문 분석을 시행할 수 있도록 트리의 루트 노드는 실제 한국어 문장에서 문형 구조상의 가장 마지막 구성요소이며, 트리의 앞 노드는 실제 한국어 문장에서 문형 구조상의 가장 처음 구성요소이다.

자세한 구문 트리별 구성 요소의 상세 내용은 [부록 2]에 제시되어있다.

2.2.4 품사 변환

한국어에서는 영어의 동명사처럼 형태소의 품사가 변경되어 사용될 때 파생 접미사가 사용된다. 파생접미사는 명사 파생 접미사, 동사 파생 접미사, 형용사 파생 접미사, 부사 파생 접미사 등이 있으며 이러한 파생 접미사의 사용은 ‘달리다’처럼 본디 동사의 품사를 갖는 형태소를 명사 파생 접미사인 ‘~기’를 사용하여 ‘달리기’처럼 명사의 품사를 갖는 형태소로 바꾸듯이 사용된다.

품사에 따라 문형 구조가 변화하는 한국어 문형 구조의 특성 상 형태소의 품사변환이 고려되지 않는다면 자칫 잘못된 문형 구조로의 파생과정이 일어나게 되어 사용자에게 입력받은 한국어 문장을 올바른 문형 구조로 분류할 수 없는 문제점이 생기기에 형태소의 품사변환은 문형 구조 분류에 있어 필수적이다.

파생 접미사는 해당 파생 접미사가 수식하는 형태소를 각각의 품사로 변경시켜주는 특성을 가지고 있다. 이를 통해 만일 파생 접미사가 구문 트리의 탐색에서 발생된 경우, 바로 앞에 있는 형태소를 파생 접미사의 특성에 맞게 품사변환을 해주어야 된다는 것을 알 수 있다.

본 논문에서는 문형 구조의 구성요소인 명사, 동사, 형용사의 파생 접미사인 명사 파생 접미사, 동사 파생 접미사, 형용사 파생 접미사를 형태소 필터링 과정을 통해 식별하여 감성 분석 시 사용되는 분석 후보군에 포함시켰다. 또한 앞서 말한 파생 접미사의 특성을 2.2.3절에서 정립한 구문 트리에 적용하여 루트 노드에서부터 구문 분석을 시행하고, 만일 파생 접미사가 나타날 경우 해당 파생 접미사의 문형 구조 상 구성요소의 품사를 해당 파생 접미사의 기능에 맞게 수정하였다.

2.3 감성 분석

감성 분석(Sentiment Analysis)은 오피니언 마이닝(Opinion Mining)으로 불리는 기술이며, 텍스트에 나타나는 사람의 태도, 의견, 성향과 같은 주관적인 데이터를 분석하는 자연어 처리 기술을 말한다[15].

감성 분석은 사용자 데이터에서 사용자의 감성 데이터를 추출하여 ‘좋음’과 ‘싫음’의 양 극단으로 분류하는 극성 탐지를 목표로 하는데, 본 논문에서는 사용자가 입력한 한국어 문장 분석에 사전 기반 한국어 감성 분석방법을 활용하였다. 이를 위한 감성 분석 사전은 2.3.1절에서 제안된 감성사전을 사용하였으며, 이 중 긍정·부정 값을 나타내는 수치 값인 감성 수치 값을 활용하여 감성 단어의 긍정·부정 수치를 계산하였다.

2.3.1 감성 사전

본 논문에서 감성 분석에 활용하기 위한 감성 사전은 기분석된 한국인들의 실정에 맞는 434개의 한국어 감정 어휘목록을 사용하였다. 감정사전에 등록된 여러 지표 중 감성 분석 과정에 사용된 지표는 친숙성이며, 1에서 6.99 사이의 수치 값을 가진다[16].

앞서 [16]에서 제안된 감성 사전은 형태소가 아닌 단어별 감성 수치를 나타내고, 동의어를 생략한 434개의 감성 단어들로 이루어져있다. 본 논문에서 분석 과정은 형태소단위로 이루어지기에 감성 사전에 포함된 434개의 감성 단어들을 모두 형태소 분석하여 감성 단어의 형태소를 분류하였으며, 해당 감성 단어들의 이음동의어를 감성 단어에 추가하여 443개의 감성 단어로 이루어진 감성 사전을 구성하였다[17].

본 논문에서 사용된 감성 사전은 [부록 3]에 제시되어있다.

2.3.2 한국어의 부정 표현

부정소란 어떤 문장에 덧붙여 그 명제의 진위를 정반대로 바꾸는 일을 하는 요소를 말한다. 한국어에서 대표적인 부정소로는 안, 못이 있으며 부정문이란 이러한 부정소가 들어있는 문장을 말한다.

부정문을 만들어내는 부정표현은 ‘안’부정법, ‘못’부정법, ‘말다’부정법과 특수어휘에 의한 특수부정법이 있으며 각각의 부정법은 ‘안’이나 ‘못’과 같은 단형부정을 사용한 부정표현과 ‘~지 않~’이나 ‘~지 못~’과 같은 장형부정을 사용한 부정표현으로 나뉜다.

한국어의 부정 표현에서 이러한 단형, 장형 부정표현은 문장의 구성요소 중 동사와 형용사를 수식하여 부정문을 만들어내는데 쓰인다.

한국어에서 부정문을 만드는 부정법 중 단형부정은 단정적인 부정을, 장형부정은 좀 더 완곡한 부정의 효과를 보여준다. 이를 통해 문장에서 단형부정과 장형부정이 나타날 경우 같은 부정문이지만 단형부정을 사용한 부정문에서 더 큰 의미 변화가 일어남을 알 수 있다. 본 논문에서는 1부터 6.99의 범위를 갖는 감성 단어들의 중앙값인 4를 원점으로 삼아 계산을 진행하였다[18].

본 논문에서는 단형부정이 사용된 경우 단호한 부정을 의미함으로 단형부정이 사용된 감성 단어의 긍·부정 수치를 원점인 4와의 차이만큼 뒤바꿔주고, 장형부정이 사용된 경우 완곡한 부정을 의미함으로 장형부정이 사용된 감성 단어의 긍·부정 수치를 원점인 4와의 차이의 절반만큼 뒤바꿔주었다.

2.3.3 감성 수치 계산

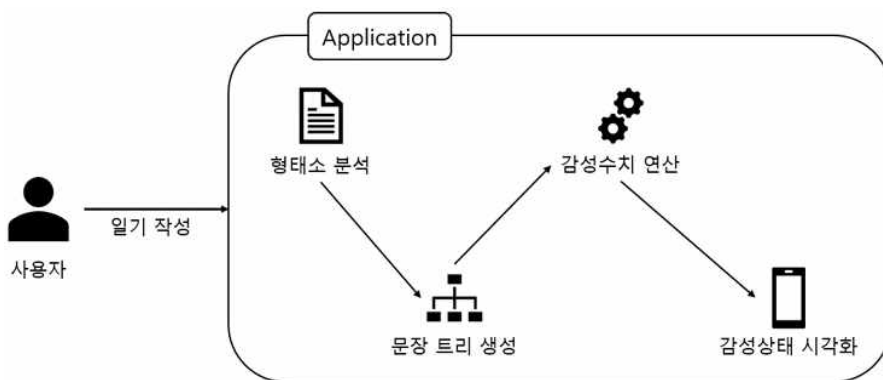
본 논문에서 제안하는 감성 수치 계산 방법은 3.2.2.1절까지 있는 감성 분석의 일련의 과정을 거치게 되면, 각 문장별로 구성된 구문 트리의 각 노드별로 감성 단어와 해당 감성 단어의 감성 수치가 저장된 감성 구문 트리가 완성이 된다. 본 논문에서 사용된 감성 수치 값이 모두 양수라는 점에 착안하여 감성 구문 트리의 각 노드에 저장되어있는 해당 감성 수치들의 수치 값 평균을 계산한다.

이러한 과정을 통해 문장 전체의 감성 수치 계산이 가능하며, 여러 문장이 존재하는 일기 내용의 감성 수치 또한 감성 수치 값이 존재하는 문장들의 수치 값 평균을 계산하는 과정을 통해 알 수 있다.

제 3 장 시스템 설계

3.1 전체 시스템 구조

본 논문에서 개발한 사용자 감성 분석 어플리케이션은 다음과 같은 일련의 과정을 통해 감성 분석을 수행한다. 먼저 사용자가 감성 분석을 수행하기 위한 일기 데이터를 어플리케이션에 작성한다. 이렇게 사용자로부터 작성된 일기 데이터는 문장단위로 나뉘지고, 각각의 문장별로 형태소 분석을 수행하게 된다. 이후 형태소 분석 과정을 통해 얻어진 형태소 토큰을 기반으로 해당 문장의 문형구조를 정의하는 문장트리를 생성한다. 문형구조가 정의된 뒤, 정의된 해당 문장의 문형구조에 사용되는 감정 단어와 부정소를 파악하여 문장 별 감성 수치를 연산하며 감성 분석 과정을 종료한다. 최종적으로, 이렇게 산출된 문장별 감성 수치를 합산하여 전체 일기 내용의 감성 수치를 파악하고 이를 사용자에게 시각화하게 된다.



[그림 7] 전체 시스템 구조

3.2 감성 분석 알고리즘

본 논문에서 설계한 알고리즘은 크게 2가지 단계로 나뉘며, 각각의 단계는 구문 분석 단계와 감성 분석 단계로 나뉜다. 구문 분석 단계는 사용자로부터 입력받은 일기 데이터를 문장 단위로 분석하여 형태소로 나누고 이후 형태소 필터링 과정을 통해 구문트리에서 사용하는 형태소들만을 분류한다. 형태소 필터링 과정을 통해 필터링 된 형태소들을 이용하여 각 문장마다의 문장 트리를 구성하는 단계이고, 감성 분석 단계는 구문 분석 과정을 통해 생성된 문장 트리를 이용하여 문장 트리를 순회하며 감성 단어가 있는 구성요소를 찾아 각 문장 트리의 구성요소에 해당 감성 단어의 감성 수치를 포함한 사용자가 작성한 문장 별 감성 문장 트리를 작성하고 이를 이용하여 문장 전체의 감성 수치를 계산함으로써 감성 분석을 실시하는 단계이다.

각각의 문장에 대하여 상기한 2가지 단계를 거치는 과정을 통해 감성 수치를 계산하고, 최종적으로 계산된 문장의 감성 수치의 총합으로 전체 일기 내용의 감성 수치를 계산하는 알고리즘을 본 논문에서 제안하였다.


```

DiaryDataSentimentAnalysis Algorithm {
// 사용자로부터 일기 데이터 입력
String DiaryData ← Input DiaryData from User
// 구문 분석 실행
// 문장별 구문 트리 생성
For FirstSentence to FinalSentence of DiaryData
    // 문장 형태소 분석
    List MorphemeList ← MorphemeAnalysis with Sentence
    // 분석된 문장 형태소 전처리
    List PrecessedMorphemeList ← Preprocessing with MorphemeList
    // 전처리된 문장 형태소를 이용하여 구문 트리 생성
    For FirstMorpheme to FinalMorpheme of PrecessedMorphemeList
        // 형태소가 구문 트리의 구성요소인 경우
        If Morpheme 가 Sentence Patterns과 Match 되면
            // 구문 트리 작성
            Than Make SyntaxTree
        End If
    End For
    // 작성이 완료된 구문 트리를 구문 트리 리스트에 저장
    TreeList SyntaxTreeList ← SyntaxTree
End For
// 감성 분석 실행
// 구문 트리별 감성 분석
For FirstTree to FinalTree of SyntaxTreeList
    // 전위 순회 방식 트리 순회
    Pre-order Tree Traversal (Tree)
    // 구문트리의 최종 감성 수치 감성 수치 리스트에 저장
    List SentimentScoreList ← SentimentScore of Tree
End For
// 모든 구문트리의 감성 수치 계산
Double Result ← Calculate Sentiment Score of SentimentScoreList
// 사용자에게 최종 감성 수치 출력
Display Result to User
}

Pre-order Tree Traversal (Tree) {
    If Tree != NULL than
        // 자식노드 순회
        Pre-order Tree Traversal (Tree Child)
        // 현재 노드의 감성 수치 계산
        SentimentScore of Tree ← Calculate Sentiment Score of Node of Tree
        // 자식 노드와 현재 노드의 감성 수치 합산
        SentimentScore of Tree ← SentimentScore + SentimentScore of Child Node
    End If
}

```

[그림 8] 의사코드로 표현한 감성 분석 알고리즘

3.3 구문 분석

형태소 필터링 과정은 2.2.3절에서 정립한 구문 트리에서 사용되는 30개의 형태소 코드를 List 변수에 저장하여 형태소 분석 과정을 통해 나온 문장의 형태소들이 List 변수에 저장된 형태소가 일치하는 경우에만 새로운 문장의 형태소 구성을 생성하도록 설계하였다.

구문 트리의 각 구성요소인 노드는 해당 구성요소의 감성 수치 값을 저장하는 Double 변수와 해당 구성요소의 형태소를 저장하는 String 변수, 구성요소가 되지 못하는 일반 명사를 저장하는 List, 해당 구성요소에 부정표현이 사용되었는지 알려주는 Boolean 변수, 사용된 부정표현의 종류를 저장하는 String 변수, 새로운 구성요소를 저장하는 노드로 구성되어 있다.

앞서 정의한 구성요소를 이용한 구문 트리는 2.2절에서 알아본 한국어의 구문적 특성을 이용하여 생성을 진행한다. 먼저 형태소 필터링 과정을 통해 생성된 문장의 형태소가 저장된 List 변수의 끝부터 구문 트리 생성을 시작한다. 제일 먼저 List 변수의 가장 마지막에 존재하는 형태소의 종류를 파악하여 중심 문형 구조를 선정한다. 중심 문형 구조가 선정되면 구문 트리 작성을 위하여 이후 List 변수에서 읽어 들이는 형태소가 선정된 중심 문형 구조의 어떠한 문형트리에 대응되는지 판단한다. 다중 Switch 구조를 활용하여 특정 구문 트리로의 분기를 구현하고, 현재 선택된 구문 트리의 식별을 위하여 식별자 변수를 사용한다. 특정 구문 트리로의 분기가 실행되면 식별자 변수를 변경하여 현재 선택된 구문 트리 루트를 파악하게 된다. 만일 List 변수에서 구문 트리의 구성요소로 사용되지 않는 일반 명사를 읽을 경우, 한국어의 지배소는 의존소의 뒤에 존재한다는 특징을 이용하여 가장 마지막으로 생성된 구문 트리 구성요소의 명사 저장 List에 저장을 하게 된다. 상기한 과정을 반복하며 구문 트리를 완성하게 되면 구문 트리만을 저장하는 List에 완성된 해당 문장의 구문트리를 저장한다.

상기한 과정을 각 문장마다 반복하며 사용자 일기 데이터의 모든 문장을 구문 트리 List에 저장하고, 이렇게 생성된 구문 트리 List를 이용하여 감성 분석 과정을 수행한다.

3.4 감성 분석

구문 분석 과정을 통해 생성된 문장의 감성 트리는 재귀적으로 순회할 수 있다. 전위순회 방식으로 순회하기에 자식노드를 먼저 방문하고 이후 작업을 진행하게 된다. 재귀함수는 반환 값으로 Double형태의 현재 노드의 감성 데이터를 반환하게 되며, 이는 부모노드가 자식노드를 참조하지 않고도 자식노드의 감성 데이터 값을 얻을 수 있게 한다. 재귀 순회 함수 안에는 감성 분석 과정 기능이 들어가게 되는데, 만일 방문하는 해당 노드가 감성 단어인 경우 해당하는 감성 분석을 진행한다.

감성 분석 과정에서는 현재 노드를 구성하는 형태소 값인 String값을 미리 구성한 감성 사전에 들어있는 감성 단어들과 비교하여 해당 형태소가 감성 사전에 등재되어 있다면 감성 사전에 정의된 감성 수치를 노드의 감성 수치 값을 저장하는 Double 변수에 저장하게 된다.

이 과정에서 부정 표현을 연산하게 되는데, 본 논문에서는 단형부정이 사용된 경우 단형부정이 사용된 감성 단어의 감성수치를 1부터 6.99의 범위를 갖는 감성 단어들의 중앙값인 4와의 차이가 같은 반대의 의미를 갖는 감성수치로 값을 조정한다. 만일 6의 값을 갖는 긍정 단어에 단형부정이 사용된다면 4와의 차이는 2가 되고, 4에서 2만큼의 차이를 갖는 부정 수치인 2로 해당 감성 수치를 변경해준다.

장형부정이 사용된 경우에는 장형부정이 단형부정에 비해 부정의 강도가 약하다는 특징을 이용하여 장형부정이 사용된 감성 단어의 감성수치를 1부터 6.99의 범위를 갖는 감성 단어들의 중앙값인 4와의 차이의 절반만큼의 감성수치로 값을 조정한다. 만일 2의 값을 갖는 부정 단어에 장형부정이 사용된다면 4와의 차이는 2가 되고, 4에서 2의 절반 값인 1만큼의 차이를 갖는 부정 수치인 5로 해당 감성 수치를 변경해준다.

ShortNegative Nation

$$Score = |(Sentiment\ Word\ Score - Origin\ Value)| * 2$$

LongNegative Nation

$$Score = |(Sentiment\ Word\ Score - Origin\ Value)| * 1.5$$

[그림 9] 부정 표현의 감성 수치 계산

부정 표현의 감성 수치 계산에 있어 부정 표현이 사용된 감성 수치가 본디 긍정적 의미의 감성 단어라면 해당 감성 수치에서 앞선 과정을 통해 계산된 Score를 빼 부정 단어로 만들어야 하지만 부정적 의미의 감성 단어라면 해당 감성 수치에서 앞선 과정을 통해 계산된 Score를 더해 긍정 단어로 만들어주어야 한다.

Positive Word

$$Sentiment\ Score = Sentiment\ Word\ Score - Score$$

Native Word

$$Sentiment\ Score = Sentiment\ Word\ Score + Score$$

[그림 10] 원래 의미에 따른 감성 수치 계산

마지막으로 해당 노드에 저장되어있는 의존소 형태의 명사들에 대한 감성 수치를 확인하고, 감성 수치가 존재하는 명사가 사용될 경우 해당 명사의 감성 수치를 노드에 포함된 Double 형태의 감성 수치 변수에 포함한다.

상기한 과정을 거치며 Double 형태의 감성 수치 변수에 저장된, 하나의 노드에서 사용된 모든 감성 수치의 합을 해당 노드에 존재하는 감성 단어의 개수로 나누어 각 노드에 사용된 모든 감성 단어의 감성 수치 평균을 구한 뒤 저장한다.

재귀 순회를 하며 2.3.3절에서 설계한 방식대로 계산된 각 노드의 감성 수치를 합산하는 과정을 통해 해당 구문 트리의 전체 감성 수치를 계산한다.

상기한 과정을 각 구문 트리마다 반복하며 사용자 일기 데이터의 모든 구문 트리의 감성 수치를 Double 형태의 변수에 저장하고, 모든 구문 트리에서 사용된 감성 수치의 합을 감성 수치가 존재하는 구문 트리의 개수로 나누어 전체 문장의 감성 수치 값을 계산한다.

3.5 알고리즘 복잡도

본 논문에서 제안하는 알고리즘의 복잡도 계산 과정에서는 형태소 분석기를 통해 문장의 분석이 완료된 뒤, 형태소 필터링 과정을 통해 필터링 되어 해당 문장의 구문 트리 작성 시 구문 트리를 구성하는 구성요소의 후보군이 되는 형태소 개수를 기준으로 하며, 이러한 형태소 개수를 n 이라 지칭한다. 또한 사용자가 작성한 일기 데이터에 존재하는 문장의 개수를 m 이라 지칭한다.

구문 분석 과정에서의 시간 복잡도는 구문 트리 작성을 위해 각 문장별로 문장에 존재하는, 형태소 필터링 과정을 통해 필터링 된 형태소의 비교를 형태소의 개수만큼 반복을 실행하므로 실행 횟수는 n 회가됨을 알 수 있다. 각각의 형태소에 대하여 해당 형태소가 구문 트리를 구성하는 형태소인지 알아보는 조건문의 실행은 반복이 시행되는 횟수만큼 실행하므로 실행 횟수는 n 회가됨을 알 수 있다. 해당하는 형태소가 구문 트리를 구성하는 형태소인 경우 구문 트리를 생성하는데, 트리 생성 시에는 대입 연산만을 수행하지만 최악의 경우 반복문이 실행될 때마다 대입연산을 수행하므로 실행 횟수는 n 회가됨을 알 수 있다. 구문 분석 과정의 마지막에 각 문장별로 생성된 구문 트리를 전체 일기 내용의 구문 트리를 저장하는 리스트에 저장하는 과정은 대입 연산만을 수행하므로 트리 생성 과정의 실행 횟수는 1회가됨을 알 수 있다. 각 문장별 구문 분석 과정에서의 실행 횟수의 총합은 $n+n+n+1$ 인 $3n+1$ 이 되며, 이러한 각 문장별 구문 분석 과정은 사용자가 작성한 문장의 개수만큼 반복되므로 모든 문장에 대한 실행 횟수는 $(3n+1) \times m$ 인 $3nm+m$ 이 된다. 최종적으로 $3nm+m$ 의 실행횟수를 갖는 구문 분석 과정의 시간 복잡도는 $O(nm)$ 이 된다.

감성 분석 과정에서의 시간 복잡도는 감성 수치 계산을 위해 구문 분석 과정에서 작성된 구문 트리를 재귀적으로 전위 순회를 시행한다. 트리의 탐색 시 실행되는 재귀 호출은 구문 트리의 각 노드를 생성하는 형태소의 개수만큼 실행되므로 재귀 호출의 실행 횟수는 n 회가됨을 알 수 있다. 재귀 호출 과정에서 실행되는 감성 수치의 계산은 단순 대입 연산만을 수행하므로 실행 횟수는 재귀 호출이 실행되는 n 회가 되며, 자식노드와 현재 노드의 감성 수치를 합산하는 과정 또한 단순 대입 연산이고 실행 횟수는 재귀 호출이 실행되는 n 회가됨을 알 수 있다. 최종적인 재귀 호출의 실행시간은 $3n$ 이 된다. 이러한 재귀 호출은 각 문장 트리의 개수만큼 실행되고 각 문장트리의 개수만큼 반복되므로 모든 문장에 대한 실행횟수는 $3n \times m$ 인 $3nm$ 이 된다. 모든 문장에 대한 감성 수치를 계산하는 과정은 모든 문장의 개수만큼 연산을 수행하므로 문장의 개수인 m 의 실행 횟수를 가지게 된다. 최종적인 감성 분석 과정에서의 실행 횟수는 재귀 호출 과정에서의 실행 횟수인 $3nm$ 과 모든 문장에 대한 감성 수치를 계산하는 과정에서의 실행 횟수인 m 의합인 $3nm + m$ 이 되므로 $3nm + m$ 의 실행 횟수를 갖는 감성 분석 과정에서의 시간 복잡도는 $O(nm)$ 이 된다.

알고리즘의 시작 시 사용자로부터 문장을 읽어드리는 연산과 알고리즘의 종료 시 계산된 감성 수치 값을 사용자에게 출력하는 연산은 각각 1회씩 시행되므로 본 논문에서 제안하는 감성 분석 알고리즘의 실행횟수 총 합은 $1 + (3nm + m) + (3nm + m) + 1$ 인 $6nm + 2m + 2$ 가 된다. $6nm + 2m + 2$ 의 실행횟수를 갖는 감성 분석 알고리즘의 시간 복잡도는 $O(nm)$ 이 된다.

제 4 장 개발 환경 및 도구

4.1 개발 환경

본 논문의 어플리케이션 개발에 사용 된 PC의 사양을 다음 [표 5]에서 보이고 있다.

[표 5] 감성 분석 어플리케이션 개발 환경

CPU	Intel(R) Core i5-6500 @3.2GHz, 4 Core
RAM	8.0GB

4.2 개발 도구

본 논문에서는 제안하는 어플리케이션의 개발을 위해 다음과 같은 개발 도구를 사용하였다. 안드로이드 스마트폰에서 사용하는 어플리케이션 개발을 위하여 안드로이드 프로그래밍에서 사용되는 프로그래밍 언어인 Java로 개발을 하였으며 4장에서 설계한 감성 분석 알고리즘의 개발 및 기능 구현은 Eclipse 통합 개발환경(IDE)을 이용하여 개발하였다. Eclipse IDE를 통해 구현한 감성 분석 알고리즘을 활용하여 사용자로부터 일기 내용을 입력받고, 감성 분석 결과를 스마트폰에 출력하는 일기장 어플리케이션은 안드로이드의 공식 IDE인 Android Studio를 이용하여 개발을 진행하였다.

제 5 장 실험 및 결과

5.1 실험

본 논문에서 개발한 어플리케이션의 실험 및 검증은 SM-G935S의 모델명을 가진 안드로이드 스마트폰 갤럭시 S7 Edge를 사용하여 실험하였다. 어플리케이션 실험에 사용된 스마트폰의 사양을 다음 [표 6]에서 보이고 있다. 어플리케이션 사용을 위해 일기 데이터는 스마트폰에 설치된 어플리케이션을 이용해 작성하고 입력되었다.

[표 6] 실험을 위해 사용한 스마트폰의 사양

CPU	Samsung Exynos 8 Octa 8890 @1.6GHz, 8 Core
RAM	4 GB
OS	Android 8.0 Oreo

일기는 21세부터 27세까지의 성인 남녀 7명으로부터 작성되었으며, 1주일동안 작성된 일기를 대상으로 실험을 진행하였다. 다양한 일기 내용의 작성을 위해 하루에 1명씩 일기를 작성하였으며, 이미 일기를 작성한 인원은 더 이상 일기를 작성하지 않는 방식으로 실험을 진행하였다.

작성된 일기는 3문장에서 4문장 사이의 단편 일기를 작성하였으며, 1주일동안 7명이 작성한 총 32문장의 일기 데이터를 대상으로 실험을 진행하였다.

5.2 실험 결과

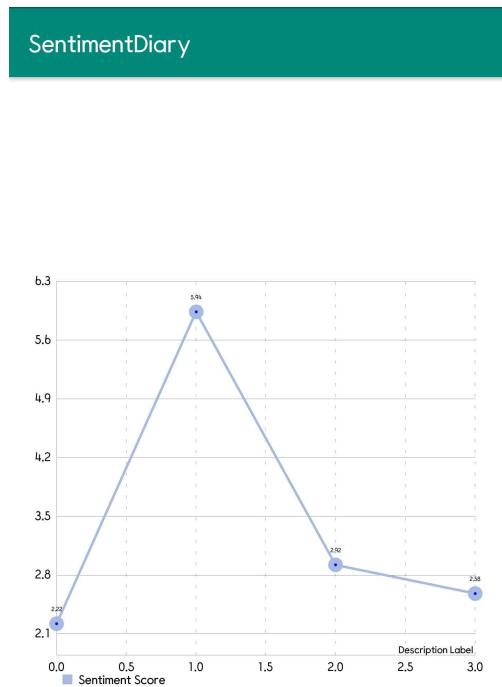
7개의 일기 데이터 중 일기 작성자가 일기 작성 시 감성을 나타내었다고 표시한 일기는 5개이며 다른 3개의 일기 내용에는 감성을 나타내지 않았다고 표시하였다. 그리고 실제 감성 분석 어플리케이션에서의 감성 분석 결과는 다음 표에서 보이고 있다.

[표 7] 감성 분석 결과

일기 번호	작성자의 감성 여부	감성 분석 결과
1	o	o (2.22)
2	o	o (5.94)
3	o	x
4	x	x
5	o	o (2.92)
6	o	o (2.58)
7	x	x

7개의 일기 내용에 대한 감성 분석 결과 감성을 나타내었다고 표시한 5개의 일기 중 4개의 일기 데이터에 대하여 감성 수치를 도출하였다. 감성 분석 결과에서 감성 여부가 도출되지 않은 3번의 일기 내용에서 작성자가 본인의 감성을 나타내었다고 표시한 ‘내가 과연 사는 건지 모르겠다.’라는 문장은 직접적인 감성 단어를 사용하지 않은 채 관용적인 표현으로 감성을 나타낸 문장이었다. 이는 정적인 사전에 정의된 감성 단어목록을 기반으로 감성 분석을 시행하는 어플리케이션의 한계를 보여준다.

또한 사용자에게 작성한 일기 내용의 추이를 보여줌으로써 본인의 감성 상태 변화를 나타내는 어플리케이션 예시 화면은 [그림 11]에서 보이고 있다. [그림 11]에서 나타내는 감성 상태 변화는 앞서 5.1절에서 언급한 21세부터 27세까지의 성인 남녀 7명으로부터 각각 작성된 일기 7편이므로 해당 그림은 감성 상태 변화 추이 기능의 예시를 들기 위해 사용되었다.



[그림 11] 사용자 감성 상태 변화 추이 예시

5.3 어플리케이션 검증

어플리케이션의 감성 분석 기능 검증을 위해 일기 내용 6번에 작성된 내용을 이용하였다. 검증과정은 해당 문장에 대한 어플리케이션 진행 과정이 4장에서 설계한 알고리즘대로 올바르게 진행되는지를 통해 검증을 시도하였다.

검증을 위해 사용된 일기 내용 6번에 작성된 일기 데이터는 [그림 12]에서 보이고 있다.

오늘 하루는 정말 슬펐다. 봉사 시간을 40시간만 채우면 되는 줄 알았는데 그게 아니라 60시간을 채워야 하는 것이었다.

매우 기분이 안 좋다. 하지만 졸업하기 위해서는 필수인 것을 피할 순 없다.

이번 방학 내내 봉사를 해서 빨리 60시간을 채워야겠다는 다짐을 했다

[그림 12] 검증을 위해 사용된 일기 데이터

해당 일기 내용에 대한 감성 분석 어플리케이션의 감성 분석 결과는 2.58이 출력되었으며, 해당 일기 내용에 4장에서 설계한 알고리즘 과정을 적용하여 도출된 결과와 감성 분석 어플리케이션에서 출력된 결과를 비교하여 검증을 하였다.

[그림 12]의 사용자 일기 데이터에 존재하는 문장의 형태소 분석 결과는 [표 8]에서 보이고 있다. 형태소 분석 과정에서는 작성된 일기 데이터를 문장단위로 나누고, 문장단위로 나누어진 일기 데이터의 형태소들을 [부록 1]에 제시된 꼬꼬마 형태소 분석기의 60여 가지 형태소 품사 태그를 활용하여 형태소 분석을 실시하게 된다.

[표 8] 문장의 형태소 분석 결과

문장	분석 결과
1	오늘/NNG, 하루/NNG, 는/JX, 정말/MAG, 슬프/VA, 었/EPT, 다/EFN, ./SF
2	내/NP, 가/JKS, 봉사/NNG, 시간/NNG, 을/JKO, 40/NR, 시간/NNG, 만/JX, 채우/VV, 면/ECE, 되/VV, 는/ETD, 줄/NNB, 알/VV, 았/EPT, 는테/ECD, 그/VV, 게/ECD, 아니/VCN, 라/ECD, 60/NR, 시간/NNG, 을/JKO, 채우/VV, 어야/ECD, 하/VX, 는/ETD, 것/NNB, 때문/NNB, 이/VCP, 었/EPT, 다/EFN, ./SF
3	매우/MAG, 기분/NNG, 이/JKS, 안/MAG, 좋/VA, 다/EFN, ./SF
4	하지만/MAC, 졸업/NNG, 하/XSV, 기/ETN, 위하/VV, 어서/ECD, 는/JX, 필수/NNG, 이/VCP, ㄴ/ETD, 것/NNB, 을/JKO, 피하/VV, ㄷ/ETD, 순/MAG, 없/VA, 다/EFN, ./SF
5	이번/NP, 방학/NNG, 내내/MAG, 봉사/NNG, 를/JKO, 하/VV, 어서/ECD, 빨리/MAG, 60/NR, 시간/NNG, 을/JKO, 채워야겠다/UN, 는/JX, 다짐/NNG, 을/JKO, 하/VV, 었/EPT, 다/EFN, ./SF

이후 형태소 분석된 결과를 바탕으로 형태소 필터링 과정을 수행하게 되는데, 형태소 필터링 과정을 거친 문장의 형태소는 [표 9]에서 보이고 있다. 형태소 필터링 과정에서는 [표 8]에 보인 문장의 형태소 분석 결과를 바탕으로 2.2.2절에서 정의한 감성 분석 과정에서 사용되는 30개의 형태소 품사만을 필터링하게 된다.

[표 9] 형태소 필터링 과정을 거친 문장의 형태소

문장	분석 결과
1	오늘/NNG, 하루/NNG, 는/JX, 정말/MAG, 슬프/VA, 었/EPT, 다/EFN
2	내/NP, 가/JKS, 봉사/NNG, 시간/NNG, 을/JKO, 40/NR, 시간/NNG, 채우/VV, 되/VV, 줄/NNB, 알/VV, 그/VV, 라/ECD, 60/NR, 시간/NNG, 을/JKO, 채우/VV, 것/NNB, 때문/NNB, 이/VCP, 었/EPT, 다/EFN
3	매우/MAG, 기분/NNG, 이/JKS, 안/MAG, 좋/VA, 다/EFN
4	졸업/NNG, 하/XSV, 위하/VV, 는/JX, 필수/NNG, 이/VCP, ㄴ/ETD, 것/NNB, 을/JKO, 피하/VV, 순/MAG, 없/VA, 다/EFN
5	이번/NP, 방학/NNG, 내내/MAG, 봉사/NNG, 를/JKO, 하/VV, 빨리/MAG, 60/NR, 시간/NNM, 을/JKO, 는/JX, 다짐/NNG, 을/JKO, 하/VV, 었/EPT, 다/EFN

형태소 필터링 과정을 통해 4장에서 작성된 알고리즘과 동일하게 감성 분석 과정에서 사용되는 30개의 형태소 품사만이 문장의 형태소로 남아있는 것을 확인 할 수 있다.

이렇게 형태소 필터링 과정이 끝난 뒤 도출되는 문장은 문장을 이루는 각 형태소들을 2.1.1절에서 정립한 한국어 문형 구조와 비교하며 문장별 구문 트리를 생성하게 된다.

첫 번째 문장을 구성하는 필수 요소는 ‘오늘’, ‘하루’라는 명사와 그 지배소인 ‘~이’라는 뜻을 가진 ‘~는’, ‘슬프다’의 형용사형 형태소인 ‘슬프’가 있다. ‘정말’은 부사이므로 문형구조에 포함되지 않는다. 이러한 ‘N이+AD’형 구조를 가진 문형 구조는 [부록 2]의 형용사 중심 1번 문형이 있다.

두 번째 문장은 안긴문장과 안은문장이 존재한다. 두 번째 문장의 구조는 ‘내가 (봉사 시간을 40 시간만 채우면 되는 줄 알았는데 그게 아니라) 60 시간을 채워야 하는 것 때문이었다.’이며, 두 번째 문장의 안긴문장은 ‘봉사 시간을 40 시간만 채우면 되는 줄 알았는데 그게 아니라’이고 안은문장은 ‘내가 60 시간을 채워야 하는 것 때문이었다.’이다. 안긴문장은 안은문장의 안에서 명사의 역할을 하므로 안은문장의 구조를 분석하면 ‘나’라는 명사와 그 지배소인 ‘~이’라는 뜻을 가진 ‘~가’, 명사절인 ‘60 시간을 채워야 하는 것 때문’과 지배소인 ‘~이다’라는 뜻을 가진 ‘~이었다’가 있다. 이러한 ‘N이+N이다’형 구조를 가진 문형 구조는 [부록 2]의 명사 중심 1번 문형이 있다.

세 번째 문장의 구성요소는 ‘기분’이라는 명사와 그 지배소인 ‘~이’라는 뜻을 가진 ‘~가’, ‘좋다’의 형용사형 형태소인 ‘좋’과 이를 수식하는 부정소인 ‘안’이 있다. 부정소는 문장구조에 포함되지 않으므로, ‘N이+AD’형 구조를 가진 문형 구조는 [부록 2]의 형용사 중심 1번 문형이 있다.

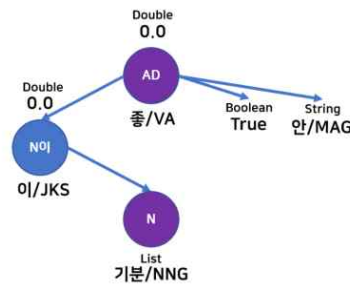
네 번째 문장의 구성요소는 ‘졸업하기 위해서’라는 명사절과 그 지배소인 ‘~이’라는 뜻을 가진 ‘~는’, ‘필수인 것’이라는 명사절과 그 지배소인 ‘~를’이라는 뜻을 가진 ‘~을’, ‘피하다’의 동사형 형태소인 ‘피하’와 그를 수식하는 부정소인 ‘없’이 있다. 이러한 ‘N는+N를+V’형 구조를 가진 문형 구조는 [부록 2]의 동사 중심 14번 문형이 있다.

다섯 번째 문장의 구성요소는 ‘이번 방학 내내 봉사를 해서’가 ‘빨리 60시간을 채워야겠다는 다짐을 했다’라는 중심 문장을 수식하므로 중심문장인 ‘빨리 60시간을 채워야겠다는 다짐을 했다’의 구성요소를 살펴보아야 한다. ‘빨리 60시간을 채워야겠다’라는 명사절과 그 지배소인 ‘~는’이라는 뜻을 가진 ‘~는’, ‘다짐’이라는 명사와 그 지배소인 ‘~를’이라는 뜻을 가진 ‘~을’, ‘하다’의 동사형 형태소인 ‘하’가 있다. 이러한 ‘N는+N를+V’형 구조를 가진 문형 구조는 [부록 2]의 동사 중심 14번 문형이 있다.

상기한 과정을 거치며 각 문장별 문형 구조를 바탕으로 한 5개의 구문 트리를 생성하는 것을 마지막으로 구문 분석 과정은 종료되게 된다.

구문 분석 과정이 종료된 뒤 생성되는 일기데이터의 구문 트리 중 형태소 분석 과정을 통해 나누어진 문장 중 3번 문장인 “매우 기분이 안 좋다.”에 대한 구문 트리 구조는 [그림 13]에서 보이고 있다.

매우 기분이 안 좋다.
(매우/MAG, 기분/NNG, 이/JKS, 안/MAG, 좋/VA, 다/EFN)
⇒ N이 + AD



[그림 13] 3번 문장에 대한 구문 트리 구조

트리는 총 2개의 노드로 구성되어있다. ‘N이’를 의미하는 노드에는 해당 구성요소의 형태소인 ‘이/JKS’와 해당 구성요소를 수식하는 일반 명사 ‘기분/NNG’가 저장되어 있고, ‘AD’를 의미하는 노드에는 해당 구성요소의 형태소인 ‘좋/VA’와 해당 구성요소에 부정표현이 사용되었음을 알려주는 ‘True’값의 Boolean 변수, 사용된 부정표현인 ‘안/MAG’를 저장하는 String 변수, 새로운 구성요소인 ‘N이’를 저장하는 노드 변수가 저장되어 있다. 각 노드에서 사용하지 않은 변수는 초기 값으로 저장되어있다.

구문 분석 과정이 종료된 뒤, 생성된 5개의 구문트리를 순회하며 감성 사전에 정의된 감성 단어를 검색한다. 생성된 5개의 구문트리의 구성요소 중 첫 번째 문장의 구문 트리 구성요소인 ‘슬프다’와 세 번째 문장의 구문 트리 구성요소인 ‘좋다’와 그의 부정소인 ‘안~’이 있다. 이를 통해 첫 번째 문장과 세 번째 문장만이 감성 수치를 갖는 감성 구문 트리임을 알 수 있다.

첫 번째 문장의 감성 분석은 ‘슬프다’라는 감성 단어에 부정형태가 사용되지 않았기에 감성 사전에 정의된 ‘슬프다’의 감성 수치를 해당 구문 트리의 감성 수치 값으로 산정하게 되는데, ‘슬프다’의 감성 수치는 2.69의 값을 가진다.

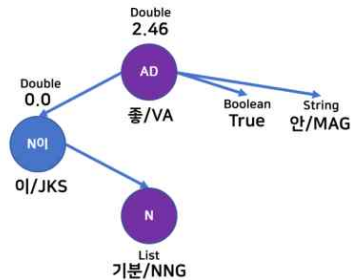
세 번째 문장의 감성 분석은 ‘좋다’라는 감성 단어를 부정소인 ‘안~’이 수식함으로 인해 감성 단어의 의미가 변하게 되고, 2.3.2절에서 정의한 알고리즘을 통해 ‘좋다’라는 긍정적인 의미의 감성 단어를 부정적인 의미의 감성 단어로 수치를 변경하여 준다. ‘좋다’의 감성 수치는 5.54의 값을 가지고 중심 값인 4와는 1.54의 차이가 존재한다. 최종적으로 긍정적인 의미의 단어인 ‘좋다’에 부정소인 ‘안~’이 수식함으로 인해 바뀌는 수치는 $4 - 1.54$ 의 값인 2.46의 값이 된다.

이후 전체 일기 데이터의 감성 수치 값 계산을 위하여 첫 번째 문장의 감정 값인 2.69와 2.46의 평균값을 구하게 되고 이는 $(2.69 + 2.46) / 2$ 의 값인 2.575가 되며, 소수점 두 자릿수까지 반올림을 하여 구하게 되므로 최종적인 전체 일기 데이터의 감성 수치 값은 2.58이 됨을 알 수 있다.

상기한 과정을 거치며 5개의 구문 트리에 감성 수치를 계산하여 대입하는 것을 마지막으로 감성 분석 과정은 종료되게 된다.

감성 분석 과정이 종료된 뒤 3번 문장인 “매우 기분이 안 좋다.”에 대한 감성 구문 트리는 [그림 14]에서 보이고 있다.

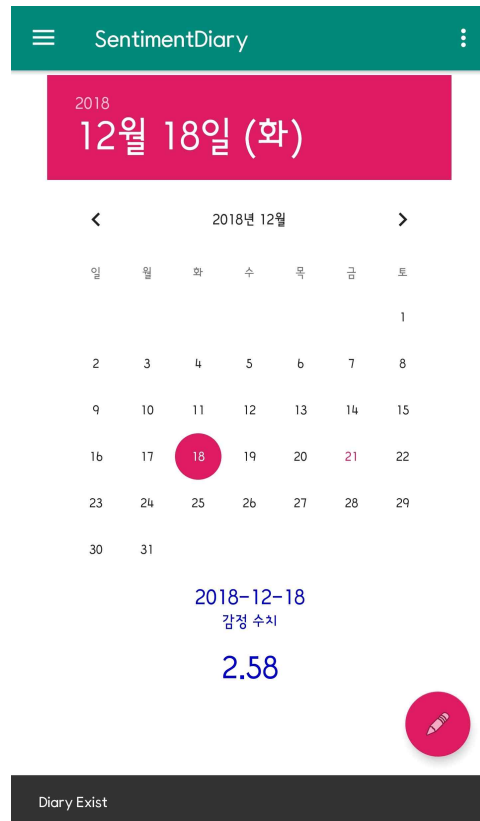
매우 기분이 안 좋다.
(매우/MAG, 기분/NNG, 이/JKS, 안/MAG, 좋/VA, 다/EFN)
⇒ N이 + AD



[그림 14] 3번 문장에 대한 감성 구문 트리

트리의 구성 요소는 [그림 13]의 구문 트리 구조와 동일하지만 감성 사전에 있는 감성 단어인 ‘좋/VA’의 감성 수치 값인 5.54의 부정형 연산을 통해 얻어진 2.46이라는 감성 수치는 해당 노드에 저장되며 감성 구문 트리를 구성한다.

어플리케이션의 검증을 위하여 본 논문에서 개발한 스마트폰의 6번째 일기 데이터에 대한 최종적인 전체 일기 데이터의 감성 수치를 확인하였다. 6번째 일기 데이터에 대한 전체 일기 데이터의 감성 수치는 [그림 12]에서 보이고 있다.



[그림 15] 6번째 일기 데이터에 대한 전체 일기 데이터의 감성 수치

앞서 계산한 전체 일기 데이터의 감성 수치 값인 2.58을 어플리케이션이 출력하는 것을 확인할 수 있다.

제 6 장 결론 및 향후 과제

본 논문에서는 기존에 연구된 한국어의 문형 구조, 단어별 감성 수치를 저장한 감성 사전을 활용하여 한국어의 문맥 의존 문법적 특징을 이용한 감성 분석 어플리케이션을 개발하였다. 감성 분석 어플리케이션은 사용자가 스마트폰을 통해 한국어 일기 데이터를 입력한 뒤 작성한 일기를 저장할 때 일기 내용에 대한 감성 분석을 실행하고 일기에 대한 감성 수치를 저장한다. 사용자는 각 날짜별로 개별적인 감성 수치의 확인이 가능하고 지정된 기간 동안의 감성 수치 추이 또한 확인할 수 있다. 접근성이 좋은 스마트폰을 이용하여 감성 분석을 진행하므로 사용자는 자신의 감성 상태를 용이하게 파악할 수 있어 사용자로 하여금 스스로에 대한 이해를 높여 건강한 삶을 유지할 수 있도록 도움을 줄 수 있는 것을 기대할 수 있다.

본 논문에서 개발한 어플리케이션은 띄어쓰기 등의 올바르지 않은 맞춤법을 기반으로 작성된 문장은 형태소 분석 결과가 정상적이지 않아 감성 분석을 수행할 수 없고 감성 사전 단어 목록의 추가가 실시간으로 이루어지지 않아 감성을 표현하는 신조어나 관용어의 사용을 감성 분석 단계에서 반영하지 못한다는 한계가 있어 학습을 통해 발전할 수 있도록 개선이 필요하다.

향후 연구로는 상기한 문제점의 해결과, 사전을 이용한 긍·부정 수치인 감성 값만을 분석하는 것이 아닌 인공지능 기술을 활용하여 여러 가지 인간의 감정을 인식하고 분석을 진행하는 기능을 추가하여야 한다.

참 고 문 헌

- [1] 국가 통계 포털, “산업별 취업자”, kosis.kr, 2018.
- [2] 안전보건공단, “감정노동 종사자”, p8-9, 2015.
- [3] 안전보건공단, “감정노동에 따른 직무스트레스 예방지침”, www.kosha.or.kr
- [4] 국가 통계 포털, “스트레스 인지율 추이”, kosis.kr, 2018
- [5] “스마일 마스크 증후군”, 지식백과,
terms.naver.com/entry.nhn?docId=3390585&cid=58345&categoryId=58345
- [6] “감정표현 불능증”, 두산백과,
terms.naver.com/entry.nhn?docId=3440270&cid=40942&categoryId=31531
- [7] 국가 통계 포털, “연도별 1일 외래환자 수”, kosis.kr, 2018
- [8] Pang, B. & Lee, L., “Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval”, 1 - 2, (2018): 1-135.
- [9] P. Turney and M. Littman. “Measuring praise and criticism: Inference of semantic orientation from association“, Proceedings of ACL-02, 40th Annual Meeting of the Association for Computational Linguistics, (2002): pages 417-424.
- [10] Sang-il Song, Dongjoo Lee, Sang-goo Lee. "Identifying Sentiment Polarity of Korean Vocabulary Using PMI." 한국정보과학회 학술발표논문집, 37.1C (2010.6): 260-265.
- [11] Myung Kyu Kim, Soo Hoan Chae, Jung Ho Kim, Myung Hoon Cha. "An Emotion Scanning System on Text Documents." Science of Emotion & Sensibility, 12.4 (2009): 433-442.

- [12] 강은국, “조선어 문형 연구”, 도서출판 박이정, 1993, 308쪽
- [13] Dongjoo Lee, Jongheum Yeon, Inbeom Hwang, Sang-goo Lee. "KKMA : A Tool for Utilizing Sejong Corpus based on Relational Database." Journal of KIISE : Computing Practices and Letters, 16.11 (2010.11): 1046-1050.
- [14] Hong Pin Im. "Parsing Strategies in Korean Syntax." Poetics & Linguistics, 6.0 (2003): 3-64.
- [15] 신수정, “감성 분석의 이해”, IDG Tech Report. 2014, 10쪽
- [16] In Jo Park, Kyung Hwan Min. "Making a List of Korean Emotion Terms and Exploring Dimensions Underlying Them." THE KOREAN JOURNAL OF SOCIAL AND PERSONALITY PSYCHOLOGY, 19.1 (2005.02): 109-129.
- [17] 국립 국어원, "한국어 기초사전", krdict.korean.go.kr/mainAction
- [18] 임홍빈, “국어 부정문의 통사와 의미”, 국어생활, (1987): 72-99.

감 사 의 글

먼저 20살의 나이로 순천향대학교에 들어와 26살에 대학교를 졸업할 때 까지 옆에서 묵묵히 응원해주시고 지지해주신 사랑하는 부모님과 가족들에게 감사의 말을 전하고 싶습니다. 그리고 매주 세미나를 통한 연구실에서의 지도로 대학 생활 동안 많은 것을 배우고 느끼게 해주시며 본 논문에 관해 아낌없는 조언을 해주신 하상호 교수님, 많은 논문 내용에 대한 피드백으로 더 나은 논문작성을 도와주신 천인국 교수님, 컴퓨터공학이라는 학문에 대해 많은 것을 알게 해주신 다른 모든 컴퓨터 공학과 교수님께도 깊은 감사의 말을 전합니다.

연구실에서 동고동락하며 많은 시간을 함께했던 객체지향 프로그래밍 연구원들인 최대영 형, 장현우 형, 이정노 형, 정기훈, 정유진, 김기태, 김재형, 이하은, 박소미, 황우녕에게도 고맙단 말을 전합니다. 때로는 짓궂게 장난도 쳤지만 언제나 중요할 때 많은 것을 도와주었던 문찬기 형, 대학생활 동안 이것저것 챙겨준 동기 손가현 등 서로를 위해주는 좋은 선후배들이 있어 대학생활을 더욱 더 즐겁게 보낼 수 있었던 것 같습니다.

끝으로 대학생활 도중 힘들 때 옆에서 많은 힘이 되어준 서유진, 박유민, 김근주와 같은 제 주변 친구들에게도 감사의 말을 전합니다.

감사합니다.

부 록

1 꼬꼬마 형태소 분석기 품사 태그 표

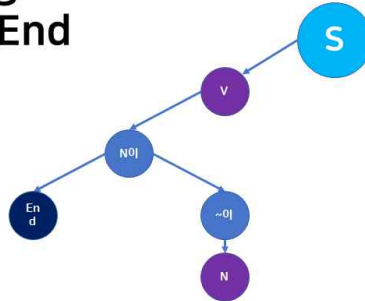
태그	설명
EC	연결 어미
ECD	의존적 연결 어미
ECE	대등 연결 어미
ECS	보조적 연결 어미
EF	종결 어미
EFA	청유형 종결 어미
EFI	감탄형 종결 어미
EFN	평서형 종결 어미
EFO	명령형 종결 어미
EFQ	의문형 종결 어미
EFR	존칭형 종결 어미
ET	전성 어미
ETD	관형형 전성 어미
ETN	명사형 전성 어미
EP	선어말 어미
EPH	존칭 선어말 어미
EPP	공손 선어말 어미
EPT	시제 선어말 어미
IC	감탄사
JK	조사
JC	접속 조사
JKC	보격 조사
JKG	관형격 조사
JKI	호격 조사
JKM	부사격 조사
JKO	목적격 조사
JKQ	인용격 조사
JKS	주격 조사
JX	보조사
MA	부사
MAG	일반 부사
MAC	접속 부사
MD	관형사

MDN	수 관형사
MDT	일반 관형사
NN	명사
NNG	보통명사
NNB	일반 의존 명사
NNM	단위 의존 명사
NNP	고유명사
NP	대명사
NR	수사
OH	한자
OL	외국어
ON	숫자
SE	줄임표
SF	마침표, 물음표, 느낌표
SO	불임표(물결, 숨김, 빠짐)
SP	쉼표, 가운뎃점, 콜론, 빗금
SS	따옴표, 괄호표, 줄표
SW	기타기호 (논리/수학기호, 화폐기호)
VA	형용사
VXA	보조 형용사
VC	지정사
VCN	부정지정사 형용사 '아니다'
VCP	긍정지정사, 서술격조사 '이다'
VV	동사
VXV	보조 동사
VX	보조 용언
XP	접두사
XPN	체언 접두사
XPV	용언 접두사
XSA	형용사 파생 접미사
XSN	명사파생 접미사
XSV	동사 파생 접미사
XR	어근
UN	명사추정범주

2. 한국어 구문 트리

동사 중심 → 1번 문형 $S \Rightarrow V + N0| + \text{End}$

N0| : 무엇이/누가 (~은/~는&-이/~가)
→ ~은: JX/~는: JX&-아: JKS/~가: JKS

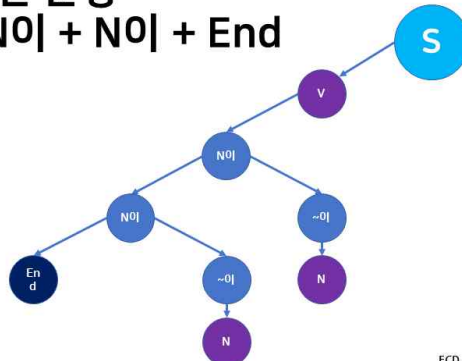


N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 2번 문형 $S \Rightarrow V + N0| + N0| + \text{End}$

N0| : 무엇이/누가 (~은/~는&-이/~가)
→ ~은: JX/~는: JX&-아: JKS/~가: JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

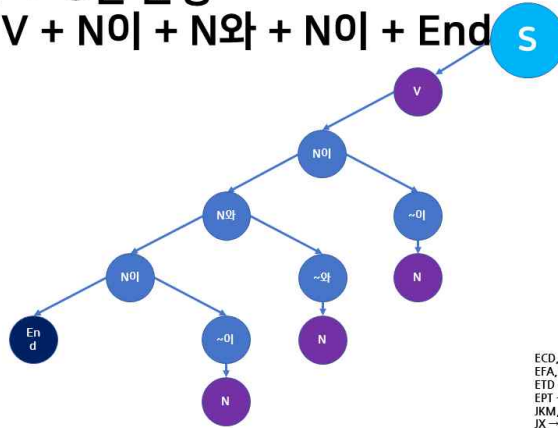
ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 3번 문형

$S \Rightarrow V + N\text{이} + N\text{와} + N\text{이} + \text{End}$

N이 : 무엇이/누가 (-은/-는&-이/-가)
→ ~은:IX/-는:IX&-아:JKS/-가:JKS

N와 : 무엇과/누구와 (-과/-와)
→ ~과:JKM/-와:JKM



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

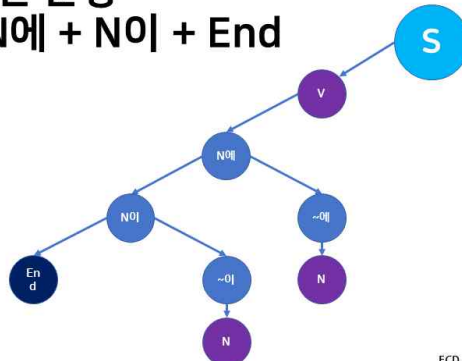
ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지정사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 4번 문형

$S \Rightarrow V + N\text{에} + N\text{이} + \text{End}$

N에 : 무엇에/누구에게/어디에 (-에/-에게/-개)
→ ~에:JKM/-에게:JKM/-개:JKM

N이 : 무엇이/누가 (-은/-는&-이/-가)
→ ~은:IX/-는:IX&-아:JKS/-가:JKS



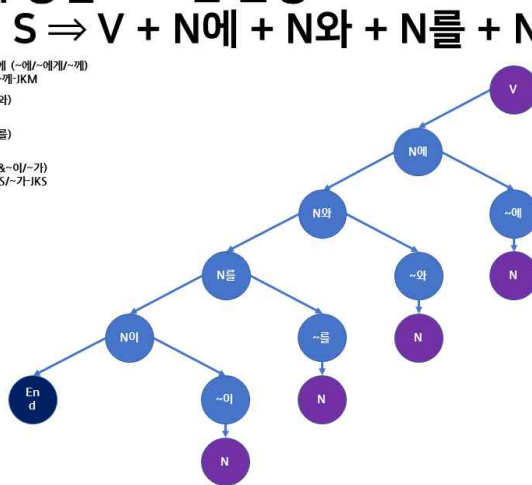
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지정사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 5번 문형

$S \Rightarrow V + N_{\text{에}} + N_{\text{와}} + N_{\text{를}} + N_{\text{이}} + \text{End}$

N_에 : 무엇에/누구에게/어디에 (∼에/∼에게/∼께)
→ ∼에-JKM/∼에게-JKM/∼께-JKM
N_와 : 무엇과/누구와 (∼과/∼와)
→ ∼과-JKM/∼와-JKM
N_를 : 무엇을/누구를 (∼을/∼를)
→ ∼을-JKO/∼를-JKO
N_이 : 무엇이/누가 (∼은/∼는&∼이/∼가)
→ ∼은-JX/∼는-JX&∼이-JKS/∼가-JKS



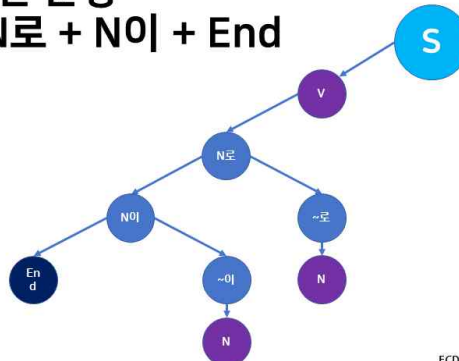
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 6번 문형

$S \Rightarrow V + N_{\text{로}} + N_{\text{이}} + \text{End}$

N_로 : 무엇으로/누구로 (∼으로/∼로&∼이/∼가)
→ ∼으로-JKM/∼로-JKM&∼이-JKS/∼가-JKS
N_이 : 무엇이/누가 (∼은/∼는&∼이/∼가)
→ ∼은-JX/∼는-JX&∼이-JKS/∼가-JKS



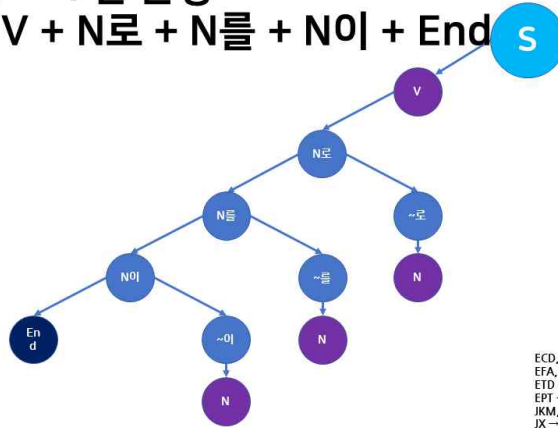
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 7번 문형

$S \Rightarrow V + N\text{로} + N\text{를} + N\text{이} + \text{End}$

N로 : 무엇으로/누구로 (~으로/~로&~이/~가)
→ ~으로 :JKM/~로 :JKM&~아 :JKS/~가 :JKS
N를 : 무엇을/누구를 (~을/~를)
→ ~을 :JKO/~를 :JKO
N이 : 무엇이/누가 (~은/~는&~이/~가)
→ ~은 :JX/~는 :JX&~아 :JKS/~가 :JKS



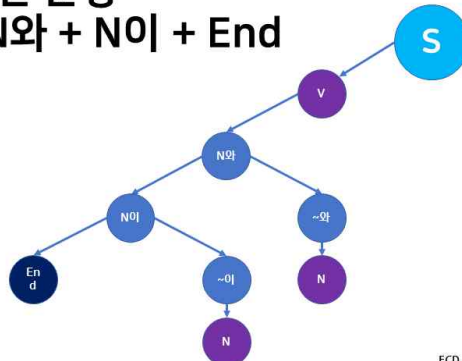
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 8번 문형

$S \Rightarrow V + N\text{와} + N\text{이} + \text{End}$

N와 : 무엇과/누구와 (~과/~와)
→ ~과 :JKM/~와 :JKM
N이 : 무엇이/누가 (~은/~는&~이/~가)
→ ~은 :JX/~는 :JX&~아 :JKS/~가 :JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

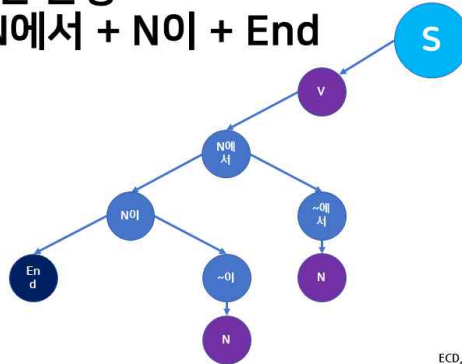
ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 9번 문형

$$S \Rightarrow V + N\text{에서} + N\text{이} + \text{End}$$

N에서 : 무엇에서/누구에서/어디에서 (~에서/~에게)
→ ~에서 :JKM/~에게:JKM

N이 : 무엇이/누가 (~은/~는&~이/~가)
→ ~은 :JX/~는 :JX&~아 :JKS/~가 :JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

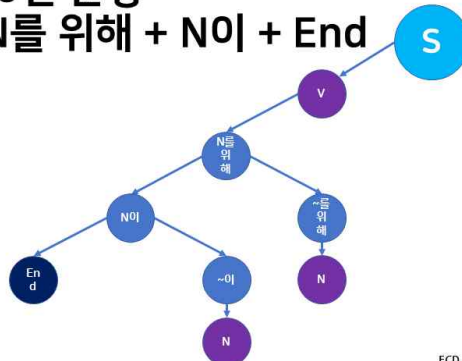
ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 10번 문형

$$S \Rightarrow V + N\text{를 위해} + N\text{이} + \text{End}$$

N를 위해 : 무엇을 위해/누구를 위해 (~을 위해/~를 위해)
→ ~을+위해+아 :JKO+VV+ECS/~를+위해+아 :JKO+VV+ECS

N이 : 무엇이/누가 (~은/~는&~이/~가)
→ ~은 :JX/~는 :JX&~아 :JKS/~가 :JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

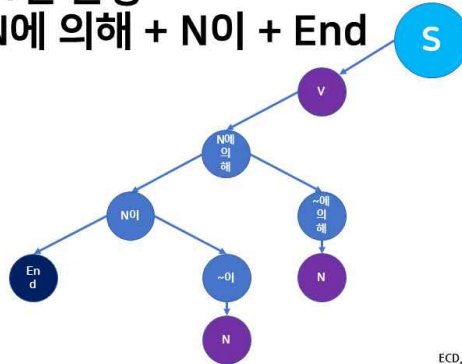
ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 11번 문형

$S \Rightarrow V + N\text{에 의해} + N\text{이} + \text{End}$

N에 의해 : 무엇에 의해/누구에 의해 (~에 의해)
→ ~에+의하+아-JKM+VV+ECS

N이 : 무엇이/누가 (~은/-는&-이/~가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



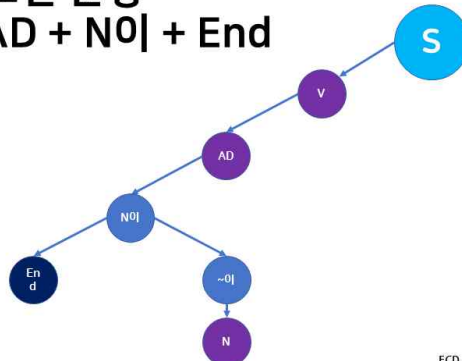
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 12번 문형

$S \Rightarrow V + AD + N\text{이} + \text{End}$

N이 : 무엇이/누가 (~은/-는&-이/~가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



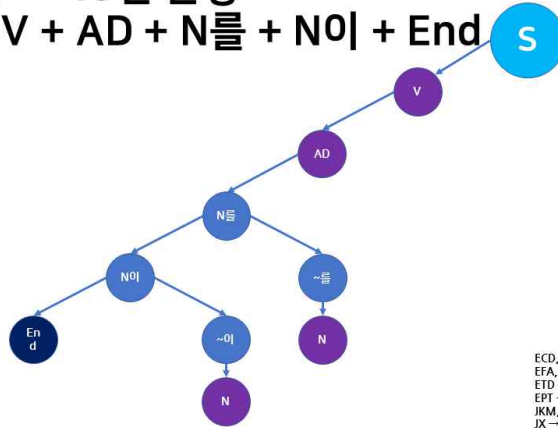
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

동사 중심 → 13번 문형

$S \Rightarrow V + AD + N\text{를} + N\text{이} + \text{End}$

N를 : 무엇을/누구를 (-을/-를)
→ ~을/~코/-를/~코
N이 : 무엇이/누가 (-은/-는&-이/-가)
→ ~은/~이/-는/~아/~스/-가/~스



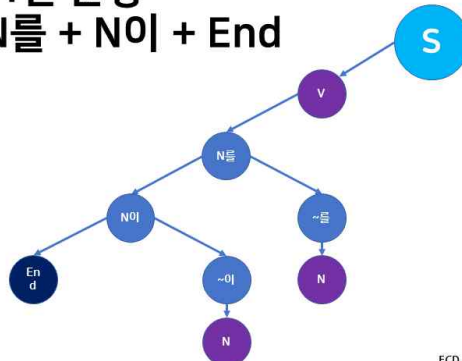
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 14번 문형

$S \Rightarrow V + N\text{를} + N\text{이} + \text{End}$

N를 : 무엇을/누구를 (-을/-를)
→ ~을/~코/-를/~코
N이 : 무엇이/누가 (-은/-는&-이/-가)
→ ~은/~이/-는/~아/~스/-가/~스



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

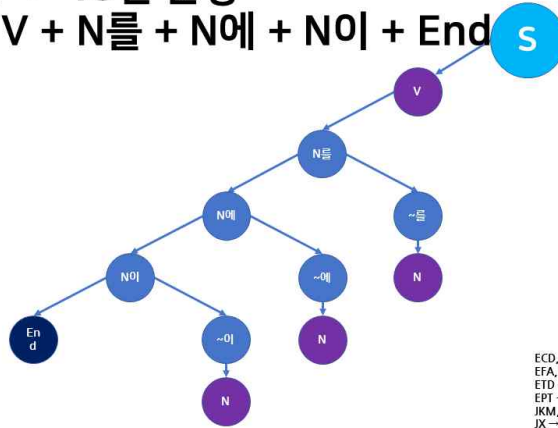
동사 중심 → 15번 문형

$S \Rightarrow V + N\text{를} + N\text{에} + N\text{이} + \text{End}$

N를: 무엇을/누구를 (~을/-를)
→ ~을-JKO/-를-JKO

N에: 무엇에/누구에게/어디에 (~에/-에게/-께)
→ ~에-JKM/-에게-JKM/-께-JKM

N이: 무엇이/누가 (~은/-는&-이/-가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

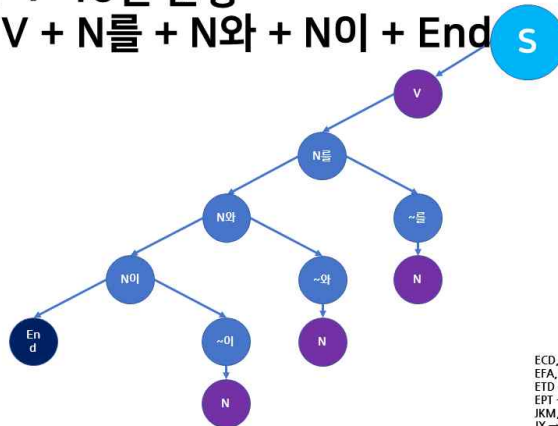
동사 중심 → 16번 문형

$S \Rightarrow V + N\text{를} + N\text{와} + N\text{이} + \text{End}$

N를: 무엇을/누구를 (~을/-를)
→ ~을-JKO/-를-JKO

N와: 무엇과/누구와 (~과/-와)
→ ~과-JKM/-와-JKM

N이: 무엇이/누가 (~은/-는&-이/-가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 17번 문형

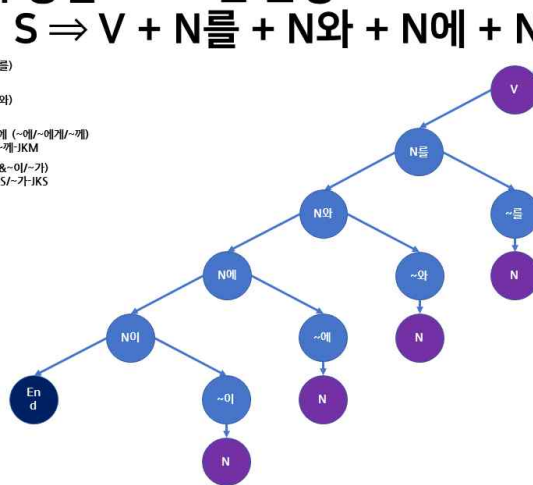
$S \Rightarrow V + N\text{를} + N\text{와} + N\text{에} + N\text{이} + \text{End}$

N를 : 무엇을/누구를 (~을/-를)
→ ~을-JKO/-를-JKO

N와 : 무엇과/누구와 (~과/-와)
→ ~과-JKM/-와-JKM

N에 : 무엇에/누구에/어디에 (~에/-에게/-에게)
→ ~에-JKM/-에게-JKM/-에게-JKM

N이 : 무엇이/누가 (~은/-는&-이/-가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN

V = VV,VXV,+XSV

AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지정사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

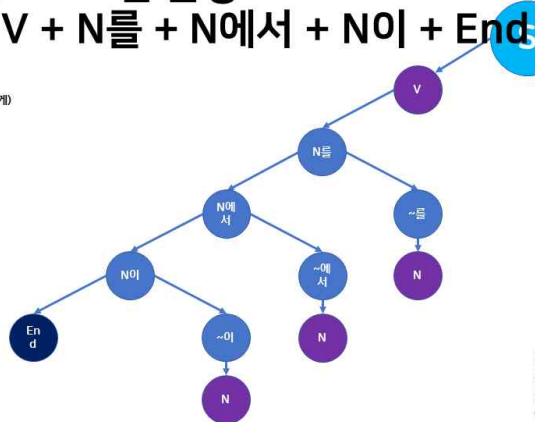
동사 중심 → 18번 문형

$S \Rightarrow V + N\text{를} + N\text{에서} + N\text{이} + \text{End}$

N를 : 무엇을/누구를 (~을/-를)
→ ~을-JKO/-를-JKO

N에서 : 무엇에서/누구에서/어디에서 (~에서/-에게)
→ ~에서-JKM/-에게-JKM

N이 : 무엇이/누가 (~은/-는&-이/-가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN

V = VV,VXV,+XSV

AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지정사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

동사 중심 → 19번 문형

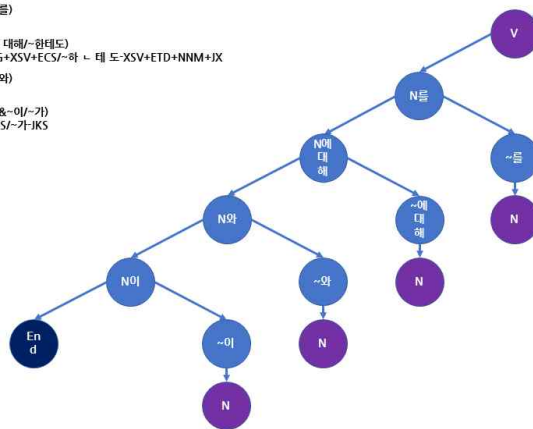
$S \Rightarrow V + N\text{를} + N\text{에 대해} + N\text{와} + N\text{이} + \text{End}$

N를 : 무엇을/누구를 (-을/-를)
→ ~을/-KO/-를/-KO

N에 대해 : 무엇에 대해 (~에 대해/~원래도)
→ ~에 대해 하 -JKM+NNG+XSV+ECS/-하 ~ 데 도-XSV+ETD+NNM+JX

N와 : 무엇과/누구와 (-과/-와)
→ ~과/-JKM/-와/-JKM

N이 : 무엇이/누가 (~은/-는&-이/~가)
→ ~은/-JX/-는 JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

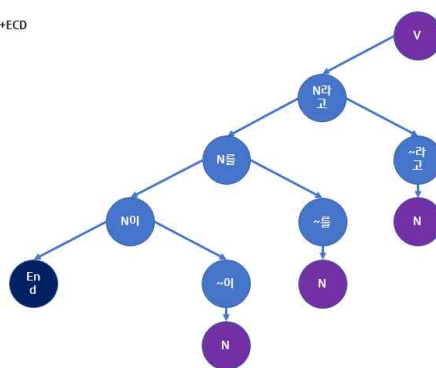
동사 중심 → 20번 문형

$S \Rightarrow V + N\text{라고} + N\text{를} + N\text{이} + \text{End}$

N라고 : 무엇이라고 (~라고/~이라고/~아래)
→ ~라고-JX/-이 라고-VCP+ECD/-이 라-VCP+ECD

N를 : 무엇을/누구를 (-을/-를)
→ ~을/-KO/-를/-KO

N이 : 무엇이/누가 (~은/-는&-이/~가)
→ ~은/-JX/-는 JX&-아-JKS/-가-JKS

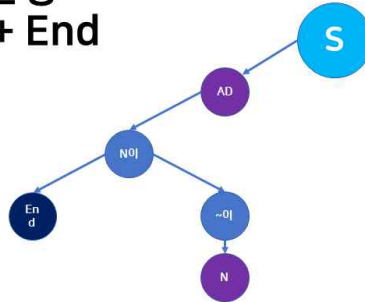


N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

형용사 중심 → 1번 문형 $S \Rightarrow AD + NOI + End$

NOI : 무엇이/누가 (~은/-는&-이/~가)
→ ~은:IX/~는:IX&-아:JKS/~가:JKS

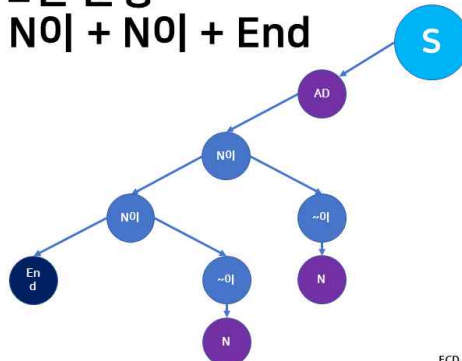


$N = NNG, NNB, NNM, NNP, NP, NR, +XSN$
 $V = VV, VXV, +XSV$
 $AD = VA, VXA, +XSA$

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

형용사 중심 → 2번 문형 $S \Rightarrow AD + NOI + NOI + End$

NOI : 무엇이/누가 (~은/-는&-이/~가)
→ ~은:IX/~는:IX&-아:JKS/~가:JKS



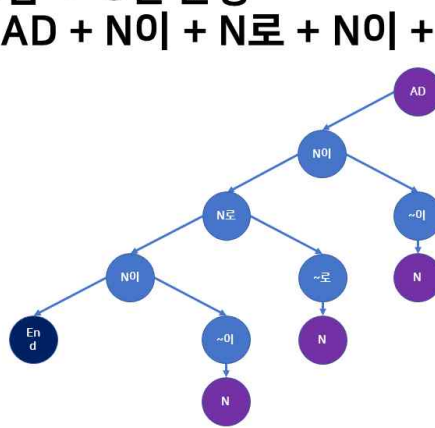
$N = NNG, NNB, NNM, NNP, NP, NR, +XSN$
 $V = VV, VXV, +XSV$
 $AD = VA, VXA, +XSA$

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 동사

형용사 중심 → 3번 문형

$S \Rightarrow AD + NOI + N로 + NOI + Ends$

NOI : 무엇이/누가 (~은/~는&-이/~가)
 → ~은:IX/~는:IX&-아:JKS/~가:JKS
 N로 : 무엇으로/누구로 (~으로/~로&-이/~가)
 → ~으로:JKM/~로:JKM&-아:JKS/~가:JKS



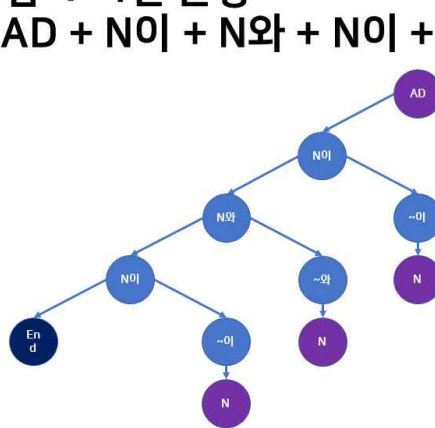
ECD, ECS → 연결 어미
 EFA, EFf, EFN, EFO, EFQ, EFR → 종결 어미
 ETD → 전성 어미
 EPT → 선어말 어미
 JKM, JKO, JKS → 조사
 JX → 보조사
 VCP → 지경사
 VA, VXA, XSA → 형용사
 NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
 VV, VXV, XSV → 등사

N = NNG,NNB,NNM,NNP,NP,NR,+XSN
 V = VV,VXV,+XSV
 AD = VA,VXA,+XSA

형용사 중심 → 4번 문형

$S \Rightarrow AD + NOI + N와 + NOI + Ends$

NOI : 무엇이/누가 (~은/~는&-이/~가)
 → ~은:IX/~는:IX&-아:JKS/~가:JKS
 N와 : 무엇과/누구와 (~과/~와)
 → ~과:JKM/~와:JKM



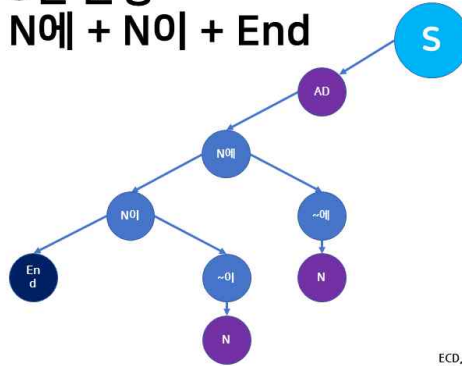
ECD, ECS → 연결 어미
 EFA, EFf, EFN, EFO, EFQ, EFR → 종결 어미
 ETD → 전성 어미
 EPT → 선어말 어미
 JKM, JKO, JKS → 조사
 JX → 보조사
 VCP → 지경사
 VA, VXA, XSA → 형용사
 NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
 VV, VXV, XSV → 등사

N = NNG,NNB,NNM,NNP,NP,NR,+XSN
 V = VV,VXV,+XSV
 AD = VA,VXA,+XSA

형용사 중심 → 5번 문형

$S \Rightarrow AD + N_{\text{에}} + N_{\text{이}} + \text{End}$

N_에 : 무엇에/누구에게/어디에 (∼에/∼에게/∼께)
 → ∼에-JKM/∼에게-JKM/∼께-JKM
 N_이 : 무엇이/누가 (∼은/∼는&∼이/∼가)
 → ∼은-IX/∼는-IX&∼아-JKS/∼가-JKS



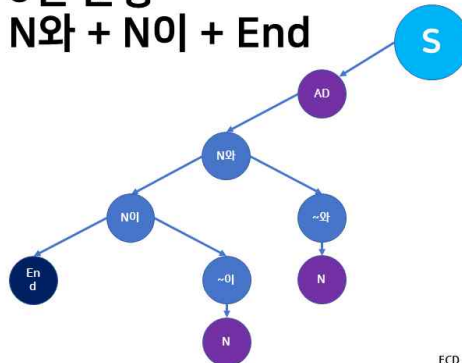
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
 V = VV,VXV,+XSV
 AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
 EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
 ETD → 전성 어미
 EPT → 선어말 어미
 JKM, JKO, JKS → 조사
 JX → 보조사
 VCP → 지경사
 VA, VXA, XSA → 형용사
 NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
 VV, VXV, XSV → 등사

형용사 중심 → 6번 문형

$S \Rightarrow AD + N_{\text{와}} + N_{\text{이}} + \text{End}$

N_와 : 무엇과/누구와 (∼과/∼와)
 → ∼과-JKM/∼와-JKM
 N_이 : 무엇이/누가 (∼은/∼는&∼이/∼가)
 → ∼은-IX/∼는-IX&∼아-JKS/∼가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
 V = VV,VXV,+XSV
 AD = VA,VXA,+XSA

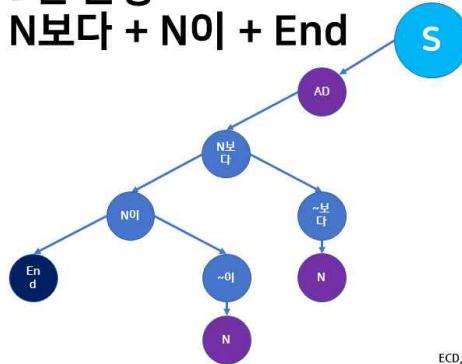
ECD, ECS → 연결 어미
 EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
 ETD → 전성 어미
 EPT → 선어말 어미
 JKM, JKO, JKS → 조사
 JX → 보조사
 VCP → 지경사
 VA, VXA, XSA → 형용사
 NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
 VV, VXV, XSV → 등사

형용사 중심 → 7번 문형

$S \Rightarrow AD + N\text{보다} + N\text{이} + \text{End}$

N보다 : 무엇보다/누구보다 (~보다)
→ ~보다-JKM

N이 : 무엇이/누가 (~은/-는&-이/~가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

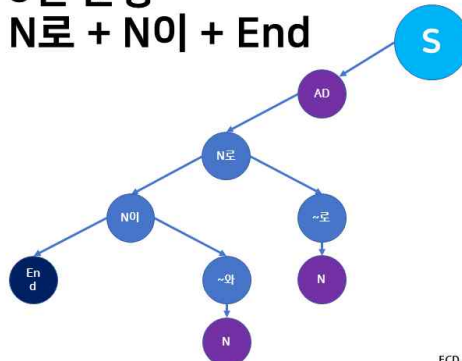
ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

형용사 중심 → 8번 문형

$S \Rightarrow AD + N\text{로} + N\text{이} + \text{End}$

N로 : 무엇으로/누구로 (~으로/~로&-이/~가)
→ ~으로-JKM/~로-JKM&-아-JKS/-가-JKS

N이 : 무엇이/누가 (~은/-는&-이/~가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS

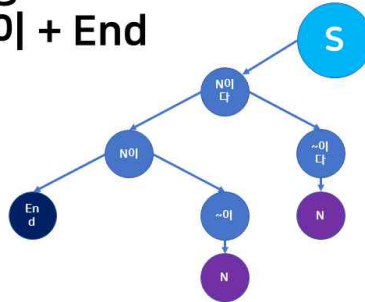


N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

명사 중심 → 1번 문형 $S \Rightarrow N\text{이다} + N\text{이} + \text{End}$

N이다 : 무엇이나 (∼이다/∼이었다/∼이었다 ∼)
→ EFA/EFV/EFN/EFO/EFQ/EFR
N이 : 무엇이/누가 (∼은/∼는&∼이/∼가)
→ ∼은:JX/∼는:JX&∼아:JKS/∼가:JKS

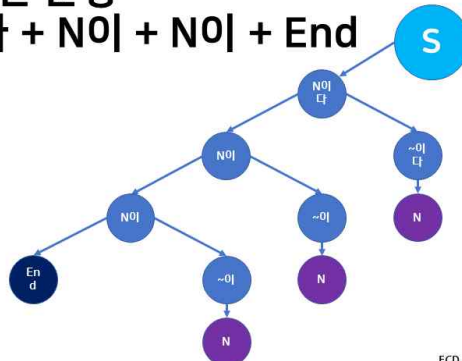


N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

명사 중심 → 2번 문형 $S \Rightarrow N\text{이다} + N\text{이} + N\text{이} + \text{End}$

N이다 : 무엇이나 (∼이다/∼이었다/∼이었다 ∼)
→ EFA/EFV/EFN/EFO/EFQ/EFR
N이 : 무엇이/누가 (∼은/∼는&∼이/∼가)
→ ∼은:JX/∼는:JX&∼아:JKS/∼가:JKS



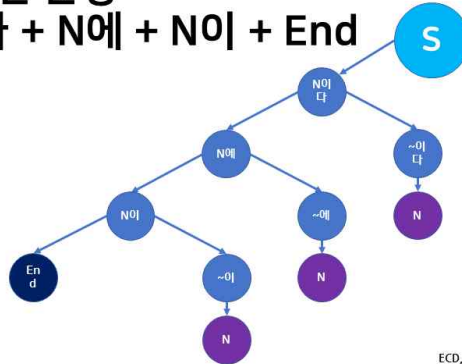
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

명사 중심 → 3번 문형

$S \Rightarrow N\text{이다} + N\text{에} + N\text{이} + \text{End}$

N이다 : 무엇이나 (~이다/~이었다/~이었다 ~)
→ EFA/EFV/EFN/EFO/EFQ/EFR
N에 : 무엇에/누구에게/어디에 (~에/~에게/~개)
→ ~에~JKM/~에게~JKM/~개~JKM
N이 : 무엇이/누가 (~은/~는&~이/~가)
→ ~은~JX/~는~JX&~아~JKS/~가~JKS



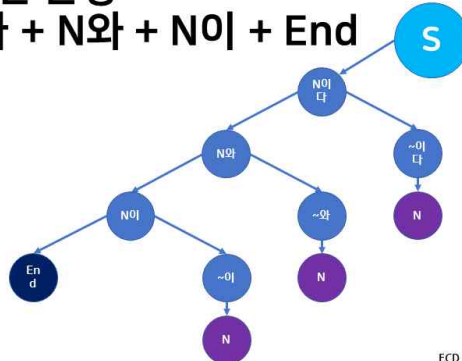
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

명사 중심 → 4번 문형

$S \Rightarrow N\text{이다} + N\text{와} + N\text{이} + \text{End}$

N이다 : 무엇이나 (~이다/~이었다/~이었다 ~)
→ EFA/EFV/EFN/EFO/EFQ/EFR
N와 : 무엇과/누구와 (~과/~와)
→ ~과~JKM/~와~JKM
N이 : 무엇이/누가 (~은/~는&~이/~가)
→ ~은~JX/~는~JX&~아~JKS/~가~JKS



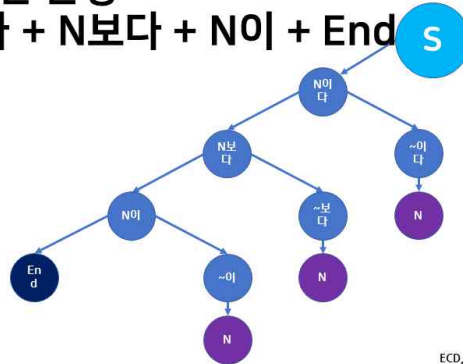
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

명사 중심 → 5번 문형

$S \Rightarrow N\text{이다} + N\text{보다} + N\text{이} + \text{End}$

N이다 : 무엇이나 (-이다/-이었다/-이었다 -)
→ EFA/EFV/EFN/EFO/EFQ/EFR
N보다 : 무엇보다/누구보다 (-보다)
→ ~보다-JKM
N이 : 무엇이/누가 (-은/-는&-이/-가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



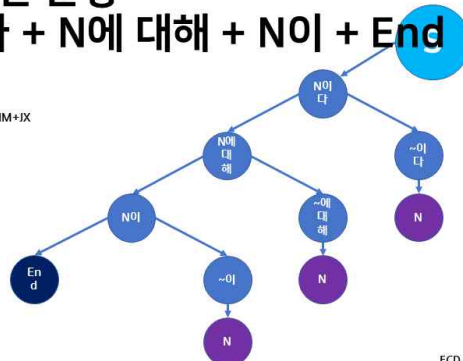
N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

명사 중심 → 6번 문형

$S \Rightarrow N\text{이다} + N\text{에 대해} + N\text{이} + \text{End}$

N이다 : 무엇이나 (-이다/-이었다/-이었다 -)
→ EFA/EFV/EFN/EFO/EFQ/EFR
N에 대해 : 무엇에 대해 (-에 대해/-원래도)
→ 에 대 하 -JKM+NNG+XSV+ECS/-하 나 테 도-XSV+ETD+NNM+JX
N이 : 무엇이/누가 (-은/-는&-이/-가)
→ ~은-JX/-는-JX&-아-JKS/-가-JKS



N = NNG,NNB,NNM,NNP,NP,NR,+XSN
V = VV,VXV,+XSV
AD = VA,VXA,+XSA

ECD, ECS → 연결 어미
EFA, EFL, EFN, EFO, EFQ, EFR → 종결 어미
ETD → 전성 어미
EPT → 선어말 어미
JKM, JKO, JKS → 조사
JX → 보조사
VCP → 지경사
VA, VXA, XSA → 형용사
NNG, NNB, NNM, NNP, NP, NR, XSN → 명사
VV, VXV, XSV → 등사

3. 감성 사전

감성 단어	감성 수치	감성 단어	감성 수치	감성 단어	감성 수치
가쁜	4.51	배신감	1.73	예쁘	5.11
가련	3.11	번뇌	2.83	온정	2.28
가소	2.48	번민	2.48	외롭	2.59
가없	3	보람차	5.69	우려	4.28
가엽	3	부끄럽	2.82	우수	3.72
가없	3	부럽	4	우습	2.49
가여	3	부아	2.28	우울	3.46
가없	3	분개	2.22	우쪼	2.37
가여	3	분노	1.84	울분	2.56
가증	1.88	분통	1.93	울적	1.87
가책	2.55	분	2.22	울화	1.29
갈등	2.59	불만	2.47	울화통	1.86
감개	4.96	불만족	2.38	원망	1.85
감개무량	4.97	불쌍	2.97	원통	1.87
감격	5.43	불안	2.51	원한	2.32
감동	5.45	불쾌	2.13	위압감	2.55
감명	5.28	불편	2.53	위축감	2.65
감미	5.54	불행	2.16	위화감	2.72
감복	4.63	비련	2.75	유감	5.58
감탄	5.11	비분강개	2.8	유쾌	2.07
감회	4.67	비애	2.71	음울	2.51
감흥	4.59	비장	3.36	이기소침	4.92
개탄	2.33	비참	1.91	이기양양	3.38
거부감	1.98	비탄	2.12	의분	2.33
거북	2.58	비통	1.8	의심	3.3
걱정	2.94	뼈아프	2.24	의아	2.72
겁나	2.61	뽀로통	2.71	의혹	5.42
격노	1.98	뿌듯	5.75	이쁘	2.27
격동	4.11	사랑	6.09	자괴	5.37
격분	2.11	사모	5.11	자궁	1.89
격양	2.33	사무	2.49	자기혐오	5.72
격정	3.57	살맛나	5.75	자랑	2.67
격하	2.7	상실감	2.29	자만심	5.02
검연쩍	3.04	상심	2.28	자부	5.28
경멸	1.74	상쾌	5.99	자신만만	5.11
경악	2.15	상큼	5.61	자족	2.24
경애	4.9	샬쭈	2.83	자책	5.72
경외	4.34	샘내	2.75	재미있	2
경쾌	5.85	서글프	2.33	적의	2.94
경탄	3.99	서러워	2.41	적적	2.7

감성 단어	감성 수치	감성 단어	감성 수치	감성 단어	감성 수치
고갯	2.42	서럽	2.51	전전긍긍	1.91
고뇌	2.71	쉴	2.51	절망	2.27
고독	2.38	서먹	2.67	절박감	2.69
고립감	2.08	서운	2.84	절절	5.56
고맙	5.5	선뚝	2.89	정감	5.56
고무	4.73	선호	4.83	정겹	5.32
고민	2.75	설레	5.21	정	3.72
고적	3.24	설워하	2.57	정욕	3.35
곤혹	2.26	섬뚝	2.17	정한	2.96
굴나	2.25	섬섬	2.47	조마조마	2.8
공감	5.07	성나	2.33	조바심	5.54
공포	2.32	성내	2.75	죃	1.87
공허	2.73	성취감	5.84	좌절	2.68
괘씸	2.11	소름	2.35	죄송	2.2
괴롭	1.93	소외감	2.16	죄	2.03
교감	5.43	속상하	2.32	죄의식	2.13
구슬프	2.8	속시원	5.43	죄책감	2.3
굴욕	1.77	속타	2.5	주눅	5.89
권태	2.67	송구	3.06	줄겁	5.57
귀찮	2.58	송연	3.32	즐기	1.75
그립	4.37	수심	2.59	증오	2.51
근심걱정	2.36	수치	2.16	지겹다	2.1
근심	2.36	속연	3.6	지긋지긋	2.64
궁지	5.74	순정	5.01	지루	2.91
기겁	2.74	스산	2.85	지리	1.97
기고만장	3.5	슬프	2.69	진노	2.07
기겁	3.37	시기	2.53	진저리나	2.45
기막히	2.67	시름겁	2.4	질겁	2.29
기쁘	5.94	시무룩	2.7	질리	2.29
기절초풍	2.64	시원섭섭	3.89	질색	2.23
기죽	2.51	시큰둥	2.88	질투	2.58
긴장	3.24	식상	2.98	짜증	2.06
꺼리	2.58	신경질	2.09	짜사랑	4.28
꺼림직하	2.45	신나다	5.96	짱	4.09
깨름직하	2.39	신명	5.61	착잡	2.83
끔찍	2.13	신물	2.2	참담	1.52
낙	5.39	신바람	5.72	창피	2.54
낙담	2.01	실망	2.19	처량	2.51
낙망	1.93	실의	2.31	처연	2.85
낙심	2.17	싫	2.24	처절	1.89
남부끄럽	2.66	싫증	2.45	처참	1.64

감성 단어	감성 수치	감성 단어	감성 수치	감성 단어	감성 수치
남부럽	3.37	심드렁	2.8	체념	2.47
낭패	2.51	심란	2.59	초조	2.74
낮간지럽	3.36	심려	2.76	충족감	5.27
낮뜨겁	2.88	심심	3.31	측은	3.07
노기	2.46	신통	3.42	치멀리	1.8
노발대발	2.1	싱송생송	3.09	치가	1.8
노심초사	2.93	쑥스럽	3.44	치욕	1.6
노엽	2.48	썰썰	2.54	친애	4.96
노파심	3.05	썩썩	2.64	침울	2.03
노하	2.16	아니꼽	2.21	침통	1.97
놀라	4.12	아련	4.03	켁기	2.59
놀랍	4.35	아리	3.36	코웃음	2.86
눈물겹	3.18	아쉽	3.07	쾌감	5.79
늪우치	3.7	아슬아슬	3.31	태평	4.55
달가워하	4.63	아연실색	2.65	토라지	2.7
담담	3.53	아연	2.9	통분	2.52
답답	2.2	아찔	2.51	통쾌	5.79
당혹	2.94	안달복달	2.75	통탄	1.83
당황	3.07	안달	2.52	통한	2.16
덤덤	3.76	안도	4.78	패배감	1.63
도취	5.1	안락	5.41	편안	5.4
동감	5.33	안심	5	편하	5.54
동경	4.81	안스럽	2.97	평안	5.63
동병상련	3.85	안정	5.11	평온	5.46
동요	3.42	안타깝	2.95	평화	5.62
동정	3.45	암담	1.79	풀죽	2.5
두렵	2.43	암울	1.98	허탈	2.67
들뜨	4.98	애	4.63	한	2.35
따분	2.77	애끓	2.94	한땀히	1.6
딱	2.99	애달프	2.95	한스럽	2.27
떠름	2.59	애달	3.14	한탄	2.02
떨떠름	2.67	애도	2.96	행복	6.16
떨	2.48	애련	3.49	향수	4.99
떳떳	5.4	애석	2.67	허망	2.11
뜨끔	2.92	애수	3.93	허무	2.28
뜨악	2.59	애잔	3.44	허전	2.91
마땅찮	2.46	애정	5.33	허허	3.07
막막	2.44	애증	2.72	혐오	1.8
만족	5.64	애지중지	4.82	호감	5.72
맘놓	4.98	애착	4.21	흔비백산	2.32
망설이	2.97	애처럽	3.01	홀가분	6.24

감성 단어	감성 수치	감성 단어	감성 수치	감성 단어	감성 수치
망연자실	2.16	애타	2.45	화나	2.17
망연	2.67	애통	2.07	환멸	1.71
매료	5.25	애틀	4.41	환희	5.72
매혹	4.84	애환	2.99	황공	4.25
맥빠지	2.5	야속	2.36	황당	3.3
머쓱	3.17	약오르	2.62	황송	4.07
멋쩍	3.38	알밋	2.28	황홀	5.72
모멸	1.34	어이없	2.89	회개	3.91
무감동	3.17	어처구니없	2.69	회오	3.63
무덤덤	3.57	억울	1.9	회한	2.49
무력감	2.19	언짢	2.5	후련	5.33
무료	2.89	여한	3.15	후회	2.45
무색	2.98	억겁	1.67	흐뭇	5.48
무섭	2.58	억정	3.02	흠모	4.94
무시무시	2.54	연모	4.45	흡족	5.77
무안	2.91	연민	3.86	흥겹	5.72
뭉클	4.48	연연	3	흥	5.5
미안	2.91	연정	4.92	흥미	5.56
민망	2.85	열광	5.52	흥미진진	5.4
밋	2.25	열등감	1.86	흥분	4.44
박진감	4.94	열애	5.07	흥취	4.71
반감	2.17	열정	5.51	희	5.21
반갑	5.91	염려	2.76	희비	3.8
반하	4.94	영광	5.99	희열	5.32
반항심	2.58	영예	5.26	희희낙락	4.71
발끈	2.57				