

영한 기계 번역을 위한 혼합형 N-best 품사 태거

임희석, 권철중, 이재원, 오기은

삼성종합기술원 휴먼인터페이스 Lab.
경기도 용인시 기흥읍 농서리 산14-1
우: 449-900
limhs@nlp.korea.ac.kr

A Hybrid N-best Part-of-Speech Tagger for English-Korean Machine Translation

Heui-Seok Lim, Cheol-Joong Kwon, Jae-Won Lee, Ki-Eun Oh
HCI Lab, Samsung Advanced Institute of Technology

요약

기계 번역 시스템에서 품사 태거의 오류는 전체 번역 정확도에 결정적인 영향을 미친다. 따라서 어휘 단계의 정보만으로는 중의성 해소가 불가능한 단어에 대해서는 중의성 해소에 충분한 정보를 얻을 수 있는 구문 분석이나 의미 분석 단계까지 완전한 중의성 해소를 유보하는 N-best 품사 태거가 요구된다. 또한 N-best 품사 태거는 단어에 할당되는 평균 품사 개수를 최소화함으로써 상위 단계의 부하를 줄이는 본연의 역할을 수행하여야 한다. 본 논문은 통계 기반 품사 태깅 방법을 이용하여 N-best 후보를 선정하고, 선정된 N-best 후보에 언어 규칙을 적용하여 중의성을 감소시키거나 오류를 보정하는 혼합형 N-best 품사 태깅 방법을 제안한다. 제안된 N-best 품사 태거는 6만여 단어의 영어 코퍼스에서 실험한 결과, 단어 당 평균 1.09개의 품사를 할당할 때 0.43%의 오류율을 보인다.

제, 구문 분석 등 중의성 해소에 필요한 충분한 정보의 사용이 가능할 때 어휘 중의성을 해결하도록 함으로써 정확한 문장 분석을 지원한다. 뿐만 아니라 N-best 품사 태거를 품사 태깅 코퍼스 구축을 위한 전처리기로 사용할 경우, 적은 량의 수작업으로도 높은 정확률을 갖는 코퍼스를 구축할 수도 있다. 만약 95%의 정확률을 보이는 품사 태거로 천만단어 크기의 코퍼스를 품사 태깅하고 그 결과를 후처리하여 품사 태깅 코퍼스를 구축한다고 가정할 경우, 5%의 오류에 해당하는 50만 단어의 오류를 수정하기 위해서 천만 단어를 모두 검사하고, 수정하는 작업을 수행하여야 한다. 그러나 N-best 품사 태깅을 수행하여 99%이상의 정확률을 얻을 수 있다면, 하나 이상의 품사를 가지는 단어들만을 후처리함으로써 99%이상의 정확률 갖는 품사 태깅 코퍼스를 구축할 수 있다.

본 논문은 통계 기반 품사 태깅 모델과 규칙 기반 후처리를 이용한 혼합형 N-best 품사 태거를 제안한다.

1. 서론

품사 태거의 주요한 역할은 형태소 해석 결과 중부적절한 품사를 제거함으로써 문장 내의 각 단어에 올바른 품사를 할당하는 것이다. 그러나 자연어에는 어휘 단계의 정보만으로는 해결하기 어려운 어휘 중의성을 갖는 단어들이 존재한다. 따라서 기계 번역과 같이 품사 태거의 높은 정확률을 요구하는 응용 분야를 위한 품사 태거는 어휘 중의성 해소가 어려운 단어에 무리하게 하나의 품사를 결정하기보다 중의성 해결에 필요한 충분한 정보를 얻을 수 있는 단계까지 그 결정을 유보하는 것이 바람직하다.

품사 태깅 시 완전한 중의성 해소가 어려운 단어에 두개 이상의 품사를 할당하는 품사 태거를 N-best 품사 태거라 한다[1, 4, 6, 8, 9]. 이 때 N은 품사 태깅 이후 한 단어에 할당되는 최대 품사 개수를 의미한다. N-best 품사 태거가 갖는 장점은 다음과 같다. 첫

2. 관련 연구

N-best 품사 태깅은 유일한 품사만을 할당했을 때 오류 가능성이 있는 단어를 탐색하는 오류 탐색 단계와 그 단어에 올바른 품사일 가능성이 있는 N개의 품사를 할당하는 N-best 품사 할당 단계로 구분할 수 있다.

논문 [2]는 상태 기반 품사 태깅(state-based tagging)과 경로 기반 품사 태깅(path-based tagging)결과를 사용하여 오류를 탐색하는 다양한 방법을 제안하였다. 제안된 방법 중 가장 우수한 성능을 보인 방법은 상태 기반 품사 태깅 후, 확률값이 최대인 결과와 두 번째로 큰 결과를 이용하는 방법이었다. 이 방법을 Lob 코퍼스에 적용한 결과 전체 코퍼스 중 15% 단어에 오류 표시를 함으로써 전체 코퍼스를 정확률 99.9%로 품사 태깅할 수 있었다[2].

N-best 품사 할당 방법에 대한 대표적인 연구는

[1, 4, 6, 8, 9]를 들 수 있다. 이들 연구는 다시 특정 단어가 가질 수 있는 가능한 품사 중 현재 문맥에 부적합한 품사를 제거하거나 적합한 품사만을 선택하는 언어 규칙을 사용하는 규칙 기반 접근법[1, 6, 8]과 특정 품사 열 또는 특정 품사가 생성될 확률 값을 계산하여 임계값(threshold) 이상의 확률 값을 갖는 결과만을 N-best 품사 태깅 결과로 선택하는 통계 기반 접근법[4, 9]으로 구분할 수 있다.

기존의 N-best 품사 태깅 중 가장 우수한 성능을 보이는 품사 태깅은 ENGCG2[6]이다. ENGCG2는 규칙을 이용하여 각 단어의 부적절한 품사를 제거하여 단어의 중의성을 감소시키는 규칙 기반 N-best 품사 태깅이다. 이 시스템은 4,000여개의 규칙으로 단어 당 평균 중의성을 1.04~1.08개를 유지하며 0.3%의 오류율을 보였다[6]1). 그러나 ENGCG2는 규칙 획득 병목(knowledge acquisition bottle neck)의 문제, 규칙의 개수가 증가할수록 발생하는 규칙의 충돌 문제, 또한 시스템 확장 및 변경 시 규칙 기반 접근법이 갖는 모든 문제점을 가지고 있다.

3. N-best 품사 태깅

그림 1은 제안된 품사 태깅 방법을 도식화한 것이다. 제안된 N-best 품사 태깅 모듈은 크게 N-best 후보 선정 모듈과 규칙 기반 후처리 모듈로 구성된다. N-best 후보 선정 모듈은 하나의 품사만을 할당했을 때 오류의 가능성이 있는 단어에 기존의 통계 기반 N-best 품사 태깅 방법[4, 9]을 이용하여 올바른 품사일 가능성이 있는 후보들을 선정하는 모듈이다. 규칙 기반 후처리 모듈은 N-best 후보 선정 모듈의 결과에 휴리스틱 규칙을 적용하여 단어 당 평균 중의성이 감소시키거나 N-best 후보 선정시 발생한 오류를 보정하는 역할을 수행한다.

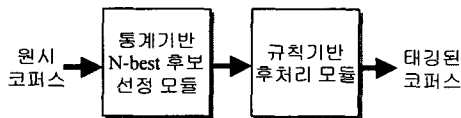


그림 1 제안된 N-best 품사 태깅 시스템

다음은 N-best 품사 태깅 모듈에서 사용할 수 있는 상태 기반 품사 태깅 방법과 경로 기반 품사 태깅 방법 그리고 이들을 이용한 N-best 후보 선정 방법을 설명한다. 품사 태깅 모델은 n개의 단어로 구성된 문장 $w^{1..n}$ 을 입력받아 올바른 품사열 $t^{1..n}$ 을 찾는 문제로 정의할 수 있다. 상태 기반 품사 태깅은 현재 단어 w^i 가 가질 수 있는 품사(k)를 통과하여 문장을 생성할 확률 값, $\gamma_i(k)$ 을 모두 계산하여 이 중 최대의 확률 값을 갖는 품사(k)를 태깅 결과로 할당하는 방법으로 $\gamma_i(k)$ 는 식 1과 같이 정의된다[5].

$$\gamma_i(k) = \frac{\alpha_i(k)\beta_i(k)}{\Pr(O|\lambda)} \quad \text{식 1}$$

식 1에서 $\Pr(O|\lambda)$ 는 $\sum_{k=1}^T \gamma_i(k)$ 를 1로 만들기 위해 사용된 정규화 요소(normalization factor)이며, T는 품사 집합 개수를 의미한다. $\alpha_i(k)$ 는 문장의 첫 단어에서부터 i번째 단어까지 관측하고 w^i 가 k번째 품사를 가질 때의 확률 값을 의미하고, $\beta_i(k)$ 는 w^i 가 k번째 품사를 가지고 i+1에서 문장 마지막 까지의 단어가 관측될 확률 값이다. 즉, 식 1의 $\gamma_i(k)$ 는 w^i 의 k번째 품사를 통과하여 문장 $w^{1..n}$ 가 생성될 확률을 의미한다. 상태 기반 N-best 품사 태깅은 각 단어에 식 1의 $\gamma_i(k)$ 의 최대값과의 비율이 임계 값 ϕ 이상인 모든 결과를 품사 태깅 결과로 생성하여 구현할 수 있다[2, 9].

경로 기반 품사 태깅은 다음 식 2와 같이 정의된다[5].

$$T(w^{1..n}) = \underset{t^{1..n}}{\operatorname{argmax}} P(t^{1..n}|w^{1..n}) \quad \text{식 2}$$

그러나 식 2의 확률 값 계산에는 많은 량의 통계 정보가 필요하며 실제로 이들 정보의 획득은 거의 불가능하다. 따라서 대부분 경로 기반 품사 태깅 모델은 단어의 어휘 정보와 근거리 문맥 정보만을 이용하여 근사한 식 3 또는 식 4로 정의된다.

$$T(w^{1..n}) = \underset{t^{1..n}}{\operatorname{argmax}} \prod_{i=1}^n P(w^i|t^i)P(t^i|t^{i-1}) \quad \text{식 3}$$

$$T(w^{1..n}) = \underset{t^{1..n}}{\operatorname{argmax}} \prod_{i=1}^n P(w^i|t^i)P(t^i|t^{i-1}t^{i-2}) \quad \text{식 4}$$

식 3과 식 4는 현재 단어의 발생은 현재 품사에만 의존하며 현재 단어의 품사는 이전 단어(또는 앞의 두 단어)의 품사에만 의존한다는 마르코프 가정을 이용하여 식 2를 근사한 것이다.

식 3과 식 4에 의한 경로 기반 품사 태깅은 viterbi 알고리즘을 이용하여 $O(NT^2)$ 시간에 계산할 수 있으며, 이를 이용한 N-best 후보 선정도 viterbi 알고리즘과 유사한 방법으로 계산할 수 있다[2, 4].

제안된 품사 태깅 방법을 위한 N-best 후보 선정 알고리즘으로는 위에서 설명한 상태 기반 N-best 품사 태깅 알고리즘이나 경로 기반 N-best 품사 태깅 알고리즘이 사용될 수 있으며 이들을 혼합한 방법도 사용될 수 있을 것이다. 상태 기반 품사 태깅과 경로 기반 품사 태깅을 혼합한 방법은 원시 문장을 각각 상태 기반 품사 태깅과 경로 기반 품사 태깅으로 품사 태깅한 후, 서로 상이하게 품사 태깅된 단어에 상태 기반 또는 경로 기반 N-best 품사를 할당하는 방법이다.

N-best 품사 태깅의 성능 측정에 사용되는 중요한 요소인 오류율과 단어 당 평균 중의성은 반비례적 관계를 갖는다. 따라서 N-best 후보 선정 알고리즘은 이 두가지 요소를 고려하여 선택되어야 하며 본 논문은 실험을 통하여 단어 당 평균 중의성이 동일할 때의

1) <http://www.conexor.fi/analysers.html>에서 ENGCG2의 성능을 평가해 볼 수 있다.

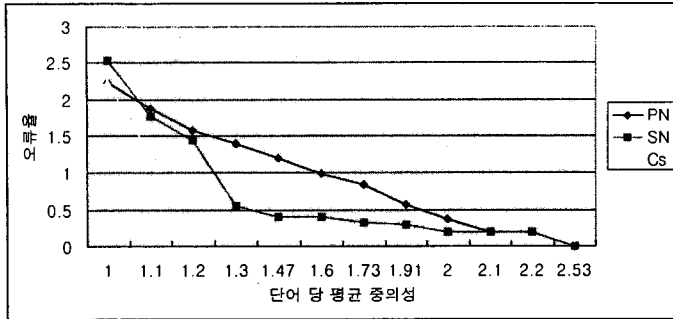


그림 2 N-best 후보 선정 알고리즘의 성능

오류율이 최소인 방법을 찾고자 한다.

그림 2는 통계 기반 N-best 후보 선정 모듈의 오류율과 중의성의 관계를 보이는 그래프이다. 본 논문에서 품사 태거의 평가 기준으로 사용하는 오류율은 아래와 같은 식에 의해서 계산한다.

$$\text{오류율} = \frac{\text{올바른 품사를 포함하고 있지 않은 어절 수}}{\text{전체 어절 수}} \times 100$$

그림 2에서 PN은 [4]의 N-best 품사 태깅 방법을 이용한 결과이고, SN은 [9]에서 사용한 상태 기반 N-best 품사 태깅의 결과이다. Cs는 경로 기반 품사 태깅과 상태 기반 품사 태깅을 혼합한 방법으로 두 방법을 이용한 품사 태깅 결과가 상이할 때 각 방법의 품사 태깅 결과뿐만 아니라 SN의 결과를 N-best 후보로 선정할 것이다. PN과 SN의 경우 단어 당 평균 중의성이 1일 때는 PN의 오류율이 낮았지만 단어 당 평균 중의성이 1.05 이상부터는 SN의 오류율이 더 낮았다. Cs 방법은 단어 당 평균 중의성이 1.1 이하까지는 경로 기반 품사 태거의 오류율을 감소시키는데 매우 효과적이었지만 그 이상부터는 오류율 감소에 영향을 미치지 못했다. Cs의 오류율이 단어 당 평균 중의성이 1.2 이하까지만 표시된 이유는 상태 기반 품사 태깅의 결과와 경로 기반 품사 태깅의 결과가 거의 유사한 결과를 보이므로 서로 상이한 품사 태깅 결과를 갖는 단어에 그 단어의 가능한 모든 품사를 할당하여도 단어 당 평균 중의성이 1.2 이하였기 때문이다. 본 논문은 그림 2의 결과를 반영하여 SN 방법을 N-best 후보 선정 알고리즘으로 사용하고자 한다.

SN 방법에 의한 N-best 후보 선정 결과를 살펴 보면 특정 문맥에서 제거할 수 있는 품사를 포함하고 있거나 특정 품사를 부여함으로써 품사 태깅 오류를 감소시킬 수 있는 단어들을 발견할 수 있다. 후처리는 이러한 단어들에 휴리스틱 규칙을 적용하여 단어 당 평균 중의성을 감소시키거나 오류율을 감소시키는 역할을 수행한다. 후처리기에서 사용하는 휴리스틱 규칙은 품사 태깅 결과를 분석하여 수작업으로 구축한 수작업 규칙과 태깅된 코퍼스를 분석하여 자동 추출한 신태그마(syntagma) 규칙으로 구성되었다. 신태그마의 사전적 의미는 발화 중에서 통합적 관계를 갖는 어구와 관용적으로 사용되는 어구를 의미하며 'all of a sudden', 'because of', 'a little' 등이 신태그마의 예이다. 신태그마 규칙은 아래와 신태그마내의 개별적인 품사 태깅 결과와 신태그마 자체의 품사를 명시한 형

태이다.

신태그마 규칙 형태 :

단어₁ 단어₂...단어_n => { 단어₁/품사 단어₂/품사,... 단어_n/품사 } {신태그마 품사 리스트}

위의 신태그마 규칙은 '단어₁ 단어₂ ..., 단어_n'가 문장에 사용되었을 경우 이들의 품사를 각각 '단어₁/품사 단어₂/품사,...단어_n/품사'로 할당하라는 규칙이다. 또한 신태그마 품사 리스트는 신태그마가 가질 수 있는 가능한 모든 품사 집합을 의미하는데, 이는 개별적인 단어의 품사보다 신태그마 자체의 품사 정보를 필요로 하는 구문 분석 단계에서 사용될 수 있는 정보이다. 신태그마 규칙 학습에는 [10]에서 관용어구 어휘 규칙 추출에 사용된 학습 알고리즘이 사용되었고 약 100만 단어 크기의 코퍼스에서 학습한 결과 500여개의 규칙이 학습되었다. 이렇게 자동 학습된 신태그마 규칙을 다시 수작업으로 편집하였고 부적절한 규칙은 제거되었으며 현재 약 400여개의 신태그마 규칙이 사용되고 있다.

수작업으로 구축한 휴리스틱 규칙은 크게 품사 제거 규칙과 품사 부여 규칙으로 나눌 수 있다. 품사 제거 규칙은 두가지 이상의 품사가 할당된 단어에서 부적절한 품사를 제거하여 중의성을 감소시키는 역할을 하며 품사 추가 규칙은 품사 태깅 오류를 자주 일으키는 특정 단어에 올바른 품사를 부여함으로써 오류율을 감소시키는 역할을 한다. 다음은 각각 품사 제거 규칙과 품사 부여 규칙의 예를 나타내고 있다.

품사 제거 규칙의 예 :

조건부: 단어 wⁱ의 품사가 pron(대명사)이고 wⁱ⁺¹의 품사가 noun(명사)과 verbp(현재형 동사)로 품사 태깅되었을 때
처리부 : wⁱ⁺¹의 품사 중 noun을 제거

품사 부여 규칙의 예 :

조건부: conj(접속사)와 prep(전치사) 품사를 가질 수 있는 단어 wⁱ가 prep로 품사 태깅되었고, w^{i-1:i-3} 단어 중 용언으로 태깅된 단어가 존재할 때
처리부 : wⁱ의 품사에 conj를 추가

수작업 규칙은 조건부-처리부 형태로 구성되며 '조

건부'는 부적절한 품사를 제거하거나 올바른 품사를 추가하기 위해 사용하는 문맥을 기술하는 부분으로 주변 단어의 어휘, 품사 그리고 단어의 중의성 유형 등을 고려하도록 되어있다. 수작업 규칙은 N-best 후보 선정 모듈의 단어 당 평균 중의성이 1.30개인 품사 태깅 결과를 이용하여 추출한 것으로 현재 약 40여개가 사용되고 있다.

4. 실험 및 비교 평가

본 논문은 제안된 N-best 품사 태거를 N-best 후보 선정 모듈의 학습에 사용되지 않은 6만여 단어 크기의 코퍼스를 이용하여 실험하였으며 형태소 분석 결과 실험 코퍼스의 단어 당 평균 중의성은 2.53개였다.

표 1 제안된 품사 태거의 성능

	중의성	오류율 (%)	
		단어 단위	문장 단위
MA	2.53	0	0
NC	1.30	0.55	6.56
NC+Rule	1.09	0.43	5.69

표 1은 실험 코퍼스에 대한 형태소 분석(MA), N-best 후보 선정 모듈(NC) 그리고 제안된 품사 태거(NC+Rule)의 단어 당 평균 중의성과 오류율을 보이고 있다. 실험 결과, 제안된 품사 태거는 한 단어에 평균 1.09개의 품사를 할당할 때 0.43%의 단어 단위의 오류율을 보였으며 5.69%의 문장 단위의 오류율을 보였다.

본 논문은 기존의 N-best 품사 태거와의 비교를 위하여 Marcken[4]과 Weischedel의 방법[9]을 제안된 시스템이 사용한 품사 집합으로 구현하였고, 동일한 실험 코퍼스에 적용하여 비교 평가하였으며 그 결과는 표 2와 같다.

표 2 기존 시스템과의 비교

단어 당 평균 중의성	각 시스템의 오류율 (%)		
	Marcken	Weischedel	제안된 방법
1.40	1.28 (12.13)	0.43 (5.63)	
1.30	1.39 (14.86)	0.55 (6.56)	
1.20	1.57 (15.72)	1.44 (10.19)	
1.10	1.88 (17.61)	1.77 (13.44)	0.44 (5.68)

비교 평가 결과, Weischedel의 방법이 Marcken의 방법보다 낮은 오류율을 보였고, 제안된 품사 태거가 단어 당 평균 중의성이 1.10일 때 0.44%의 가장 낮은 오류율을 보였다. 참고적으로 제안된 품사 태거의 이와 같은 성능은 Marcken과 Weischedel 방법으로 각각 단어 당 평균 1.81개와 1.40개의 품사를 할당하여야만 얻을 수 있는 성능이었다. 위의 평가에서 ENGCG2가 제외된 이유는 품사 태깅 규칙 구축을 위

하여 많은 수작업과 비용을 필요로 하였기 때문이다. 하지만 통계 기반 접근법과 후처리 규칙의 사용만으로도 ENGCG2와 같은 우수한 성능을 가질 수 있음을 확인하였다.

5. 결론

본 논문은 통계 정보와 휴리스틱 규칙을 이용한 N-best 품사 태거를 제안하였고, 제안된 품사 태거는 실험 결과 한 단어 당 평균 1.09개의 중의성을 유지할 때 0.43%의 오류율을 보였다. 제안된 N-best 태거는 규칙 기반 접근법과 비교하여 비교적 구현이 용이하고, 다양한 언어 현상에 대해서 견고하고, 시스템의 확장 및 변경이 용이하다는 장점을 갖는다. 또한 통계 기반 방법으로는 해결하기 어려운 언어 현상에 규칙을 적용하여 품사 태거의 성능을 향상시킬 수 있다는 장점을 갖는다.

본 논문의 기여는 다음과 같이 요약할 수 있다. 첫째, 규칙 구축에 많은 양의 비용과 노력이 필요로 하고 규칙의 제어/관리에 어려움이 있으며 시스템의 변경 및 확장이 어려운 규칙 기반 접근법을 이용하지 않고도 그와 유사한 성능을 갖는 N-best 품사 태거를 구현할 수 있음을 보였다²⁾. 둘째, N-best 후보 선정 모듈로 사용될 수 있는 기존의 통계 기반 방법을 구현 및 평가하였으며, 평가 결과 경로 기반 N-best 품사 태깅[9]의 성능이 가장 우수함을 확인하였다. 셋째, 실제 응용에 사용될 수 있는 수준의 낮은 오류율을 갖는 품사 태거 개발을 위하여 순수한 통계 기반 접근법뿐만 아니라 통계 기반 접근법의 오류를 보정할 수 있는 규칙 기반 접근법의 결합이 필요함을 보였다. 실제 본 논문에서 사용한 언어 규칙은 비교적 손쉽게 획득할 수 있는 규칙이며 품사 태거의 성능 향상에 매우 효과적이었다³⁾.

품사 태깅 후 단어 당 평균 중의성과 오류율은 반 비례적 관계를 가지므로 제안된 품사 태거의 오류율을 유지하면서 단어 당 평균 중의성을 감소시키기가 매우 어렵다. 따라서 최근에는 품사 태깅의 오류율을 유지하면서 단어 당 평균 중의성을 감소시키기 위한 방법과 품사 태거가 사용될 응용 분야에 맞게 단어 당 평균 중의성과 오류율을 조절할 수 있는 효과적인 방법에 대한 연구를 수행하고 있다.

- 2) 품사 태거의 평가는 많은 변수가 고려되어야 하는 어려운 작업이므로 논문은 제안된 품사 태깅 방법이 ENGCG2보다 우수한 성능을 갖는다고 주장하지는 않는다. 하지만 품사 태거의 개발과 유지/관리, 확장에 있어서 어려움을 겪을 수 있는 규칙 기반 접근법이 아닌 통계 기반 접근법으로도 ENGCG2와 같은 우수한 성능을 보이는 품사 태거의 구현이 가능함을 밝히고자 한다.
- 3) 실제로 ENGCG2는 본 논문이 사용한 신태그마 규칙과 같은 규칙을 사용하였으며 이 규칙은 품사 태깅 성능 향상에 중요한 영향을 미쳤다.

참고문헌

- [1] E. Brill, "Some Advances in Transformation-Based Part of Speech Tagging," *Procs. of the 12th National Conference on Artificial Intelligence*, pp. 722-727, 1994.
- [2] George F. Foster, *Statistical Lexical Disambiguation*, Master thesis, School of Computer Science, McGill Univ., 1991.
- [3] E. Macklovitch, "Where the tagger falters," *Proc. of the 4th Int. Conf. on Theoretical and Methodological Issues in Machine Translation (TMI-92)*, pp.113-126, 1992.
- [4] C. G. de Marcken, "Parsing the LOB Corpus," *Procs. 28th Annual Meeting of the Association for Computational Linguistics*, pp. 243-251, 1990.
- [5] L. R. Rabiner, B. H. Juang, "An introduction to Hidden Markov Models," *IEEE ASSP Magine*, January, pp. 4-16, 1986.
- [6] C. Samuelsson, A. Voutilainen, "Comparing a Linguistic and a Stochastic Tagger," *Procs. 35th Annual Meeting of the Association for Computational Linguistics*, pp. 248-253, 1997.
- [7] Pasi Tapanainen, Aaro Voutilainen, "Tagging accurately - Don't guess if you know," *Proc. of the 7th Conference of the European chapter of the Association for Computational Linguistics*, pp.149-156, 1994.
- [8] A. Voutilainen, J. Heikkila, "A Syntax-based Part-of-Speech Analyser," *Procs. 7thh Conference of the European Chapter of the Association for Computational Linguistics*, pp. 157-164, 1995.
- [9] R. Weischedel, M. Meter, R. Schwartz, L. Ramshaw and J. Palmucci, "Coping with Ambiguity and Unknown Words through Probabilistic Models," *Computational Linguistics*, Vol. 19, No. 2, pp. 359-382, 1993.
- [10] 임희석, 언어 지식과 통계 정보를 이용한 한국어 품사 태깅 모델, 고려대학교 컴퓨터학과 박사학위 논문, 1997.