

## 품사태깅을 위한 어휘규칙의 자동획득†

이상주 류원호 김진동 임해창

고려대학교 컴퓨터학과  
서울시 성북구 안암동 5가 1  
우: 163-701

{zoolwhryuljinlrjm}@nlp.korea.ac.kr

### Automatic Acquisition of Lexical Rules for Part-of-Speech Tagging

Sang-Zoo Lee Won-Ho Ryu Jin-Dong Kim Hae-Chang Rim  
Department of Computer Science and Engineering  
Korea University

#### 요약

기존의 어휘규칙기반 품사태거는 품사문맥이나 어휘확률만을 사용하는 통계적 품사태거에 의해 해결되지 않는 형태론적 중의성을 어휘문맥을 참조하는 어휘규칙을 사용함으로써 효과적으로 해결할 수 있었다. 그러나 어휘규칙을 수작업으로 획득하기 때문에 규칙획득에 많은 시간이 소요되어 소량의 규칙만이 사용되었다. 본 논문에서는 품사부착말뭉치로부터 어휘규칙을 자동으로 획득하는 방법을 제안한다. 제안된 방법으로 자동획득된 어휘규칙을 사용하여 실험말뭉치의 66.1%를 98.8%의 정확률로 태깅하였다. 이로써 통계적 품사태거만을 사용할 때(95.43% 정확률)보다 어휘규칙과 결합할 때(96.12% 정확률) 통계적 품사태거의 성능이 약 15.1%(0.69% 정확률)만큼 향상되었다. 또한 제안된 방법은 영어 품사태깅에 대해서도 효과적임이 실험을 통해 증명되었다.

#### 1 서론

한국어는 교착어이면서 다양한 음운현상을 갖고 어절단위 품사태깅보다 형태소단위 품사태깅이 선호되기 때문에 영어와 같은 굴절어에 비해 형태론적 중의성이 훨씬 복잡하다. 특히 한국어에는 품사문맥으로는 구별하기 어려운 동품사 중의성[9]이 존재하며 이러한 중의성은 의미정보나 어휘문맥을 고려해야만 해결될 수 있다. 그러나 기존의 통계적 품사태거[2, 3, 4, 9]는 품사문맥 확률과 어휘확률만을 참조하므로 동품사 중의성

과 같은 형태론적 중의성을 제대로 해결할 수 없었다.

어휘문맥을 고려하는 기존의 품사태거들은 변형규칙[1], 최대엔트로피모델[5], 어휘규칙[10, 12] 등에 기반을 두고 있다. [1]에서는 변형규칙을 어휘화함으로써 어휘문맥을 고려할 수 있지만, 어휘화된 규칙보다 기존의 규칙들이 훨씬 더 높은 점수를 갖기 때문에 어휘화된 규칙에 의해 해결되는 중의성이 많지 않게 된다. [5]에서는 어휘를 참조하는 자질을 사용함으로써 어휘문맥을 고려할 수 있지만, 학습과 태깅에 많은 시간이 소요되기 때문에 제한된 어휘문맥만을 고려하였다. 더욱이 [1, 5] 등은 영어에서와 같은 어절단위 품사태깅에 적용된 모델이기 때문에 한국어에 적용하기 위해서는 형태소단위 품사태깅 모델로 적절한 변형이 필요하다는 문제점이 있다. 반면에 [10, 12]에서는 주변 어휘를 고려할 수 있는 어휘규칙을 사용하는 모델로 어절단위 또는 형태소단위 품사태깅에 모두 적용할 수 있으며 어휘문맥을 고려함으로써 높은 태깅 정확도를 보인다. 그러나 수작업으로 어휘규칙을 획득하기 때문에 대량의 규칙을 획득하기 어렵고 획득된 규칙에 일관성이 결여되거나 오류가 존재하는 문제점이 있다.

본 논문은 어휘규칙을 수동으로 획득함으로써 발생하는 문제를 해결하기 위해 품사태깅된 말뭉치로부터 어휘규칙을 자동으로 획득하는 방법을 제안한다. 어휘규칙을 말뭉치로부터 자동학습하기 위해 기존의 규칙들을 말뭉치 통계를 반영하도록 변형하였다. 제안된 방법으로 자동획득된 어휘규칙을 사용하여 학습말뭉치의 100%를 100% 정확률로 태깅하였고 실험말뭉치의 66.1%를 98.8%의 정확률로 태깅하였다. 이로써 통계적 품사태거만을 사용할 때(95.43% 정확률)보다 어휘

† 본 논문은 1997년도 한국과학재단 특정기초 연구과제 연구비 지원에 의한 것입니다.

(제 10회 한글 및 한국어 정보처리 학술대회)

규칙과 결합할 때(96.12% 정확률) 통계적 품사태거의 성능이 약 15.1%(0.69% 정확률)만큼 향상되었다. 또한 제안된 방법은 영어 품사태거에 대해서도 효과적임이 실험을 통해 증명되었다.

2. 관련연구

[10, 12]에서 제안된 기존 어휘규칙의 형태는 [그림 1]과 같다. 기존 어휘규칙은 주어진 어휘 문맥(좌우 최대 세 어절)에서 중심어가 하나의 품사태그에 대응됨을 나타낸다<sup>1)</sup>.

$$\langle P;N \rangle \langle \text{중심어} \rangle \langle \text{문맥} \rangle = \langle \text{품사태그} \rangle$$

$$\langle \text{문맥} \rangle := \{ \text{앞어절} \}_0^P * \{ \text{뒤어절} \}_0^N$$

\* : 중심어를 나타내는 표시

$$0 \leq P, N \leq 3$$

[그림 1] 기존 어휘규칙의 형태

예를 들어, '나는'이 '씩이 나는 중이다.'에서는 '나\_동사+는\_어미'로 태깅되고, '새가 나는 중이다.'에서는 '날\_동사+는\_어미'로 태깅된다. 위 두 문장은 같은 품사문맥 '명사+조사 동사+어미 명사+서술격조사+어미'를 가지기 때문에 '나는'의 형태론적 중의성은 품사문맥에 의해서 해결될 수 없다. 또한 어휘확률을 사용할 때는 두 태그 중에서 어휘확률이 더 높은 태그가 일괄적으로 선택되기 때문에 오류가 발생하게 된다. 그러나 [그림 2]의 두 어휘규칙은 두 어절 '씩이'와 '나는', '새가'와 '나는' 사이의 의미관계를 묵시적으로 표현함으로써 중의성을 해결한다.

$$[1:0] [나는] [씩이 *] = [나_동사+는_어미]$$

$$[1:0] [나는] [새가 *] = [날_동사+는_어미]$$

[그림 2] 기존 어휘규칙의 예

그러나 기존 어휘규칙은 언제나 하나의 올바른 규칙만을 결정하는 결정적 규칙이기 때문에 규칙이 적용될 수 있는 모든 가능한 용례가 고려된 후 중의성이 없을 때에 규칙이 작성돼야 한다. 그러나 규칙을 사람이 획득하기 때문에 규칙 획득 시 문맥을 충분히 고려하지 않고 중의성이 남아있는 문맥에서 규칙을 작성함으로써 오류를 야기하게 된다[10]. 예를 들어, [그림 2]의 두 번째 어휘규칙은 '새가 나는 좋다.'에서도 '나는'을 '나\_대명사+는\_조사'가 아닌 '날\_동사+는\_어미'로 잘못 태깅하게 된다.

또한 어휘규칙을 수동으로 획득하기 때문에 대량의 규칙을 획득하기 어려운 한계를 가지고 있다.

3 통계기반 어휘규칙

기존의 어휘규칙이 사람의 지식에 기반을 둔 결정적 규칙이기 때문에 발생하는 문제를 해결하기 위해 품사부착말뭉치의 통계에 기반을 둔 어휘규

칙을 제안한다. 통계기반 어휘규칙의 형태는 [그림 3]과 같다.

$$\langle \text{규칙} \rangle := \langle \text{중심어} \rangle \langle \text{문맥정보} \rangle \langle \text{태깅정보} \rangle$$

$$\langle \text{중심어} \rangle := W_0$$

$$\langle \text{문맥정보} \rangle := P \ N \ \langle \text{좌문맥} \rangle \ \langle \text{우문맥} \rangle \ C$$

$$\langle \text{좌문맥} \rangle := \{ \text{앞어절} \}_0^P := P \ W_{-P} \ \dots \ W_{-1}$$

$$\langle \text{우문맥} \rangle := \{ \text{뒤어절} \}_0^N := N \ W_1 \ \dots \ W_n$$

$$\langle \text{태깅정보} \rangle := A \ \{ \text{품사태그, 빈도} \}_0^A$$

$$:= A \ (T_1, F_1) \ \dots \ (T_A, F_A)$$

C := 용례의 수, P := 좌문맥의 크기,  
N := 우문맥의 크기, A := 중의성 정도  
W<sub>-i</sub> := 중심어의 왼쪽 i번째 어절  
W<sub>i</sub> := 중심어의 오른쪽 i번째 어절  
T<sub>i</sub> := 빈도가 i번째로 많은 품사태그  
F<sub>i</sub> := T<sub>i</sub>로 태깅된 용례의 수

[그림 3] 통계기반 어휘규칙의 형태

통계기반 어휘규칙은 학습말뭉치에서 중심어 W<sub>0</sub>이 주어진 문맥 W<sub>-P</sub> ... W<sub>-1</sub> W<sub>0</sub> W<sub>1</sub> ... W<sub>N</sub>에서 C번 출현하고 A개의 품사태그 중에서 T<sub>1</sub>로 F<sub>1</sub>번, ..., T<sub>A</sub>로 F<sub>A</sub>번 태깅되었다는 것을 나타낸다. 예를 들면 [그림 4]의 어휘규칙은 어휘문맥 '할 수도' (좌문맥 '할'과 우문맥 '')에서 중심어 '수도'가 태그 '수\_명사+도\_조사'로 태깅된 말뭉치 내의 용례 수가 14개이고 태그 '수도\_명사'로 태깅된 용례 수가 2개라는 것을 나타낸다.

$$\begin{matrix} (\text{수도 } 1 \ 0 \ \text{할 } 16 \ 2 \ (\text{수\_명사+도\_조사 } 14) \\ (\text{수도\_명사 } 2) \end{matrix}$$

[그림 4] 통계기반 어휘규칙의 예

통계기반 어휘규칙 ρ의 가장 가능성이 높은 태그 MPT(ρ)와 태깅 정확률 ACC(ρ)는 식 [1,2]와 같이 정의된다. 식 [1,2]에서 ρ(A)는 규칙 ρ의 요소 A를 나타낸다.

$$MPT(\rho) = \rho(T_1) \quad [1]$$

$$ACC(\rho) = \frac{\rho(F_1)}{\rho(A)} \quad [2]$$

예를 들어, [그림 4]의 규칙의 MPT는 '수\_명사+도\_조사'이며 ACC는 14/16이다.

3.1 말뭉치기반 학습

1. 주어진 문맥을 갖는 모든 용례를 품사부착 말뭉치로부터 추출한다.
2. 용례로부터 규칙의 중심어, 문맥정보, 태깅 정보를 설정한다.

[그림 5] 통계기반 어휘규칙의 획득

중심어와 문맥이 주어질 때 통계기반 어휘규칙은 [그림 5]와 같이 획득된다<sup>2)</sup>. 예를 들어 중심어 '결었다.'에 대한 용례가 [그림 6]과 같이 획득될 때 문맥을 참조하지 않는 규칙(가장 일반적인 규칙)은 [그림 6]의 a와 같이 획득된다. [그림

1) 기존 어휘규칙에서 사용되는 품사태그는 형태 소단위이지만 어절단위로도 사용될 수 있다.

2) 용례 추출을 위한 색인 및 탐색 방법은 [11]을 참조하였다.

(제 10회 한글 및 한국어 정보처리 학술대회)

6]에서 '/s'는 문장 경계를 나타낸다.

걸_동사+었_어미+다_어미+... 정겹게 거리를 걸었다. /s 없이 그냥 걸었다. /s 내쉬며 길을 걸었다. /s 모험주의 노선을 걸었다. /s 그를 따라 걸었다. /s 조금 떨어져 걸었다. /s 흥뻑 맞으며 걸었다. /s 앞을 바라보며 걸었다. /s 거리를 바쁘게 걸었다. /s 신혼의 밤거리를 걸었다. /s 어두운 밤거리를 걸었다. /s 우산을 쓰며 걸었다. /s 발작 앞서 걸었다. /s 다시 앞으로 걸었다. /s 웃으며 앞장서 걸었다. /s 힘 없이 걸었다. /s 앞에까지 함께 걸었다. /s 걸_동사+었_어미+다_어미+... 웃으며 말을 걸었다. /s 가서 전화를 걸었다. /s 마음으로 전화를 걸었다. /s 집에 전화를 걸었다. /s 집에 전화를 걸었다. /s 집에 전화를 걸었다. /s 집에 전화를 걸었다. /s 집으로 전화를 걸었다. /s 철규집으로 전화를 걸었다. /s 하영에게 전화를 걸었다. /s a. (걸었다. 0 0 28 2 (걸_동사+었_어미+다_어미+... 18) (걸_동사+었_어미+다_어미+... 10) )
---

[그림 6] '걸었다.'에 대한 용례와 가장 일반적인 규칙

기존 어휘규칙과 달리 통계기반 어휘규칙은 하나의 품사태그만을 가지면 결정적이고 둘 이상의 품사태그를 가지면 비결정적이다. 어떤 규칙이 비결정적이라면 더 많은 어휘정보를 참조함으로써 결정적인 규칙을 획득할 수 있다. 하나 더 적은 어휘정보를 참조하는 규칙을 일반규칙, 하나 더 많은 어휘정보를 참조하는 규칙을 세부규칙이라 한다. 세부규칙은 좌문맥 또는 우문맥 중 어느 문맥에서 어휘정보를 참조하느냐에 따라 좌세부규칙과 우세부규칙으로 구분된다. 따라서 어떤 비결정적 규칙의 세부규칙 집합은 좌세부규칙 집합과 우세부규칙 집합으로 분리될 수 있다.

$$DR(S) = \frac{\sum_{\rho \in S, R(A)=1} \rho(C)}{\sum_{\rho \in S} \rho(C)} \quad [3]$$

$$AR(S) = \frac{\sum_{\rho \in S} \rho(F_1)}{\sum_{\rho \in S} \rho(C)} \quad [4]$$

$$NR(S) = |S| \quad [5]$$

규칙집합의 중의성해결 능력은 결정률, 정확률, 크기에 의해 판단된다. 규칙집합 S의 결정률

DR(S)는 식 [3]과 같이 규칙집합 S가 적용되는 용례 수와 결정적으로 중의성이 해결되는 용례 수의 비율로 정의되고, 정확률 AR(S)는 식 [4]와 같이 규칙집합 S가 적용되는 용례 수와 정확하게 중의성이 해결되는 용례 수의 비율로 정의되며, 크기 NR(S)는 식 [5]와 같이 규칙집합 S의 규칙 수로 정의된다.

만약 두 개의 경쟁하는 규칙집합 S와 S'가 있을 때 다음과 같은 원칙으로 최소집합을 선택한다.

- 원칙1: DR(S)가 DR(S')보다 크면 S를 선택한다.
- 원칙2: DR(S)와 DR(S')가 같을 때 AR(S)가 AR(S')보다 크면 S를 선택한다.
- 원칙3: DR(S)와 DR(S')가 같고 AR(S)와 AR(S')가 같을 때 NR(S)가 NR(S')보다 크면 S를 선택하고 NR(S)와 NR(S')가 같으면 S와 S'를 모두 선택한다.

원칙1은 최대 결정주의 원칙으로서 규칙집합이 학습말뭉치 내의 모든 어절에 대해서 결정적으로 중의성을 해결할 수 있다면 학습말뭉치가 모든 실세계 언어현상을 반영할 수 있을 만큼 충분히 클 때 어떠한 실험말뭉치에 대해서도 결정적으로 중의성을 해결할 수 있다는 것을 의미한다. 원칙2는 최대 정확도의 원칙으로서 결정률이 같다면 정확률이 높은 규칙집합을 선호한다는 것이다. 원칙3은 최소 정보량의 원칙으로서 같은 성능을 갖는 규칙집합 중에서 규칙이 적은 것을 선호한다는 것을 의미한다.

비결정적규칙의 세부규칙 집합에서 위의 원칙에 의해 불필요한 집합을 제거하고 남은 집합을 세부규칙의 최소집합이라 한다. 따라서 비결정적 규칙 ρ의 세부규칙 최소집합 MS(ρ)는 식 [6]과 같이 정의된다. 식 [6]에서 ls는 좌세부규칙 집합이고, rs는 우세부규칙 집합이다.

$$MS(\rho) = \begin{cases} ls & \text{if } DR(ls) > DR(rs) \\ & \text{if } DR(ls) \equiv DR(rs), AR(ls) > AR(rs) \\ & \text{if } DR(ls) \equiv DR(rs), AR(ls) \equiv AR(rs), |ls| > |rs| \\ rs & \text{if } DR(rs) < DR(ls) \\ & \text{if } DR(rs) \equiv DR(ls), AR(rs) < AR(ls) \\ & \text{if } DR(rs) \equiv DR(ls), AR(rs) \equiv AR(ls), |rs| < |ls| \\ ls \cup rs & \text{if } DR(rs) \equiv DR(ls), AR(rs) \equiv AR(ls), |rs| \equiv |ls| \end{cases} \quad [6]$$

예를 들어 [그림 6]의 규칙 a는 비결정적이다. 규칙 a의 세부규칙 최소집합은 [그림 7]과 같이 획득된다. [그림 7]의 규칙들은 규칙 a보다 좌문맥에 하나의 어휘정보를 더 참조하고 있는데 이것은 규칙 a의 좌세부규칙 집합이 우세부규칙 집합보다 결정적으로 중의성을 해결하는 용례의 수가 많기 때문에 우세부규칙 집합이 제거되었기 때문이다.

최대 결정주의 원칙, 최대 정확도의 원칙, 최소 정보량의 원칙에 입각하여 말뭉치로부터 통계기반 어휘규칙을 학습하는 과정은 [그림 8]과 같다. 이 학습과정에 의해 말뭉치에 중의성이 존재하면 그 중의성을 해결할 수 있는 규칙이 계속 학습된다. 결국 말뭉치 내의 중의성이 학습을 주

- (걸었다. 10 거리를 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 그냥 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 길을 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 노선을 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 따라 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 떨어져 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 맞으며 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 바라보며 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 바쁘게 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 밤거리를 21 (긴 동사+었 어미+다 어미+\_ 2))
- (걸었다. 10 쓰며 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 앞서 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 앞으로 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 앞장서 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 없이 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 함께 11 (긴 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 말을 11 (걸 동사+었 어미+다 어미+\_ 1))
- (걸었다. 10 전화를 91 (걸 동사+었 어미+다 어미+\_ 9))

[그림 7] 규칙 a의 세부규칙 최소집합  
 도한다는 의미에서 이 학습방법을 중의성주도 (ambiguity-driven) 학습이라 명명한다.

1. 학습말뭉치에 있는 모든 중심어들에 대해 가장 일반적인 규칙을 획득하여 규칙베이스에 추가한다.
2. 규칙베이스에 추가되는 모든 비결정적규칙들에 대해서 세부규칙 최소집합을 획득하여 규칙베이스에 추가한다.
3. 2에서 새로운 세부규칙 최소집합이 추가되지 않을 때까지 2의 과정을 반복한다.

[그림 8] 중의성주도 학습

### 3.2 품사태깅

품사부착말뭉치로부터 학습된 규칙들은 원시 말뭉치를 태깅하기 위해 사용된다. 그러나 학습말뭉치가 실제계의 모든 언어현상을 반영할 수 있을 만큼 충분히 크지 않기 때문에 자료부족 현상이 발생한다. 따라서 충분히 신뢰할 수 있는 규칙을 사용하여 태깅해야 한다. 실험을 통해 규칙의 신뢰성은 규칙의 적용범위(즉, 규칙이 획득된 용례의 수)에 비례한다는 것을 알 수 있었다. 결국 자료부족에 의한 태깅 정확률의 저하를 최소화하기 위해서는 가능한 한 적용범위가 넓은 규칙을 선택해야 한다.  $\alpha$ 는 신뢰할 수 있는 규칙의 최소 적용범위를 나타내는 임계치이다.

신뢰할 수 있는 규칙이라 하더라도 정확률이 낮다면 소용이 없다. 일반적으로 규칙기반 태깅의 태깅 정확률은 적용되는 규칙의 정확률에 비례한다. 높은 정확도를 위해 결정적 규칙만을 태깅에 사용한다면 태깅의 적용범위가 좁게 된다. 실험을 통해 충분히 높은 정확률을 갖는 비결정적 규칙들은 전체 정확률을 크게 저하시키지 않으면서 태깅의 적용범위를 넓힐 수 있다는 것을 알았다. 태깅에 사용할 비결정적 규칙의 최소 정

확률은 임계치  $\beta$ 에 의해 결정된다.

1. 어절에 적용될 수 있는 규칙을 탐색한다.
2. 적용범위가 임계치  $\alpha$ 보다 작은 규칙을 제거한다.
3. 1과 2를 거쳐 남은 규칙이 없으면 태깅을 보류한다. ('/u'로 태깅한다.)
4. 1과 2를 거쳐 남은 규칙 중에서 가장 정확률이 높은 규칙의 정확률이 임계치  $\beta$ 보다 높으면 그 규칙의 MPT(가장 빈도가 높은 태그)로 태깅한다.
5. 그렇지 않으면 태깅을 보류하거나(1-best 태깅) 모든 가능한 태그를 부여한다(k-best 태깅).

[그림 9] 어휘규칙을 이용한 품사태깅 알고리즘

규칙기반 태깅은 임계치  $\alpha$ 와  $\beta$ 에 따라서 원시말뭉치의 각 어절을 [그림 9]와 같이 태깅한다. [그림 9]의 알고리즘에 의해 비결정적 규칙보다 결정적 규칙이 선호되고 비결정적 규칙들 중에서는 정확률이 높은 규칙이 선호된다. [그림 9]에서 가장 정확률이 높은 규칙이 충분히 높은 정확률을 갖지 않을 때, k-best 태깅 방식<sup>3)</sup>에 의해 모든 가능한 태그를 부여하는 것은 중의성을 최소화하기 위한 것이다.

특수기호 '/u'로 태깅된(태깅이 보류된) 어절들은 형태소분석이나 사전참조를 통해서 가능한 태그들을 찾아내고 통계적 태깅과 같은 적용범위가 넓은 태깅에 의해 가장 가능성이 높은 태그를 선택하는 과정을 거쳐야 한다. 물론 k개의 최선의 태그들에 대해서도 통계적 태깅 등에 의해 가장 가능성이 높은 태그를 선택하는 과정이 필요하다.

학습말뭉치가 충분히 크지 않다면 어휘규칙에 의해 결정적으로 태깅되지 않는 어절들이 많이 발생할 것이다. 그러나 비록 어휘규칙을 이용하는 규칙기반 태깅이 상대적으로 낮은 적용범위를 갖는다 하더라도, 높은 태깅 정확률을 가지기 때문에 여러 분야에서 매우 유용하게 활용될 수 있다.

### 3.3 통계적 태깅과 결합

많은 연구들에서 태깅의 정확도를 향상시키기 위해 규칙기반 태깅과 통계적 태깅을 결합하였다 [6, 7, 12]. [6]에서는 먼저 규칙기반 태깅과 통계적 태깅이 독립적으로 태깅한 뒤 태깅결과가 다를 경우 규칙기반 태깅의 결과를 선호하는 규칙 우선 병렬통합 방법을 사용하여 태깅의 정확률을 향상시켰다. [7]에서는 먼저 통계적 태깅이 신뢰도 있는 통계를 근거로 일부 어절들을 태깅하고, 나머지 태깅되지 않은 어절들을 규칙기반 태깅으로 태깅하는 통계 우선 직렬통합 방법을 사용하였다. [12]에서는 수동구축 규칙기반 태깅 [10]으로 일부 어절들의 중의성을 해결하고 통계적 태깅 [4]로 나머지 어절들의 중의성을 해결하

3) [1]에서 중의성을 최소화하기 위해 k개의 최선의 태그(k-best tags)를 사용하였다.

(제 10회 한글 및 한국어 정보처리 학술대회)

학습말뭉치		규칙		내부실험		외부실험(크기=16,665)			
크기	평균 빈도	모든 규칙	빈도 제한	TA(%)	TC(%)	모든 규칙		빈도 제한	
						TA(%)	TC(%)	TA(%)	TC(%)
16,807	1.79	11,698	4,162	100	100	97.55	45.84	98.52	34.42
50,744	2.27	29,950	13,707	100	100	98.16	57.18	98.77	46.22
84,102	2.61	44,679	23,002	100	100	98.43	61.59	98.95	51.87
117,008	2.88	59,945	37,770	100	100	98.57	64.25	99.05	55.08
150,450	3.12	73,543	47,880	100	100	<b>98.79</b>	<b>66.05</b>	<b>99.19</b>	<b>57.65</b>

[표 1] 한국어 실험 결과: 규칙학습, 내부실험, 외부실험

	$\alpha=1$		$\alpha=2$		$\alpha=3$		$\alpha=4$		$\alpha=5$	
	TA	TC	TA	TC	TA	TC	TA	TC	TA	TC
$\beta=1.0$	98.79	66.05	99.38	55.43	99.54	49.25	99.60	44.98	<b>99.60</b>	41.53
$\beta=0.9$	98.65	70.53	99.11	60.74	99.21	54.89	99.24	50.78	99.22	47.42
$\beta=0.8$	98.19	72.78	98.49	63.43	98.50	57.70	98.46	53.71	98.38	50.41
$\beta=0.7$	97.56	<b>74.63</b>	97.68	65.62	97.61	60.01	97.47	56.11	97.38	52.68

[표 2] 한국어 실험 결과: 임계치를 달리한 외부실험

는 규칙 우선 직렬통합 방법을 사용하였다. [12]에서는 [6]과 같은 규칙 우선 병렬통합 방법과 규칙 우선 직렬통합 방법을 모두 실험하여 규칙 우선 직렬통합 방법이 더 높은 정확도 향상을 보인다는 것을 보고하였다. 그러나 [12]에서 사용하고 있는 규칙기반 태거는 수동으로 구축되기 때문에 소량의 규칙만을 사용하므로 적용범위가 낮은 한계가 있다.

통계기반 어휘규칙도 통계적 태거와 결합될 수 있다. 통계기반 어휘규칙은 자동으로 획득되면서도 높은 정확도를 보이기 때문에 통계적 태거와 결합될 때 기존의 연구보다 더 좋은 결과를 얻을 수 있다. 본 논문에서는 규칙 우선 직렬통합 방법처럼 규칙기반 태거에 의해 일부 어절들을 1-best 태깅 또는 k-best 태깅으로 중의성을 최소화하고 통계적 태거에 의해 남은 중의성을 해결하는 방법을 사용하여 실험하였다.

4 실험 및 평가

본 논문에서는 통계기반 어휘규칙을 한국어뿐만 아니라 영어에 대해서도 실험하였다.

4.1 한국어태깅

실험에 사용된 한국어 품사부착 말뭉치는 형태소 태그 44개와 기호태그 17개의 품사태그 집합으로 태깅되었고 여러 장르의 문서로 구성되었으며 167,115 어절이 포함되었다. 이 중에서 150,450 어절(약 90%)은 학습말뭉치로 사용되었고, 16,665 어절(약 10%)은 성능평가를 위한 실험말뭉치로 사용되었다. 실험말뭉치는 전체말뭉치에서 10문장에 1문장씩 추출되었다.

[표 1]의 첫 번째와 두 번째 열은 학습말뭉치의 크기와 어절의 평균빈도, 세 번째와 네 번째 열은 학습된 총 규칙의 수와 중심어의 빈도가 2 이상인 규칙의 수, 다섯 번째와 여섯 번째 열은 결정적 규칙의 학습말뭉치(내부실험)에 대한 태깅 정확률 TA(정확하게 태깅된 어절 수 / 태깅된 어절 수)과 태깅 적용률 TC(태깅된 어절 수 / 총

어절 수), 일곱 번째와 여덟 번째 열은 결정적 규칙의 실험말뭉치(외부실험)에 대한 태깅 정확률과 태깅 적용률, 아홉 번째와 열 번째 열은 중심어의 빈도가 2이상인 결정적 규칙의 실험말뭉치에 대한 태깅 정확률과 태깅 적용률이다.

[표 1]에 따르면 통계기반 어휘규칙을 이용하는 규칙기반 태거는 학습말뭉치를 100% 정확률로 태깅한다. 이것은 통계기반 어휘규칙이 중의성을 해결할 수 있을 정도로 충분한 어휘문맥을 참조할 수 있기 때문으로 해석된다. 한편 실험말뭉치에 대해서 결정적 규칙을 사용할 때 98.79% 태깅 정확률과 66.05% 태깅 적용률을 보인다.

반면 학습말뭉치의 크기가 증가할수록 외부실험의 태깅 정확률과 적용률이 증가하지만 획득되는 규칙의 수도 따라서 증가한다. 그러나 이것은 그리 심각한 문제는 아닐 것 같다. 왜냐하면 필요에 따라서 사용할 어휘규칙을 선택할 수 있기 때문이다. 예를 들면, [표 1]에서 중심어의 출현 빈도가 1인 규칙을 제외하면, 규칙의 수가 약 35% 감소하지만 태깅 정확률은 오히려 늘어나고 태깅 적용률만 약간 줄어든 것을 알 수 있다.

[표 2]는 임계치를 달리한 외부실험의 결과를 보여준다. [표 2]에 따르면 태깅 정확률은 임계치  $\alpha$ 와  $\beta$ 에 비례하지만 태깅 적용률은 임계치  $\alpha$ 와  $\beta$ 에 반비례함을 알 수 있다. 따라서 높은 정확률을 필요로 하는 경우에는 임계치를 높이고 높은 적용률을 필요로 할 경우에는 임계치를 낮추어야 한다.

[표 3]은 통계적 태거와 결합할 때의 외부실험 결과를 보여준다. 첫 번째는 형태소분석기만을 사용한 경우의 k-best 태깅 정확률 KA(k-best 태그 중에서 정확한 태그가 있는 어절 수 / 태깅된 어절 수)와 평균 중의성 AA(총 태그 수 / 총 어절 수), 두 번째는 통계적 태거만을 사용한 경우, 세 번째는 어휘규칙기반 태거만을 사용한 경우( $\alpha=1, \beta=1.0$ )로서 1-best 태깅 후에 태깅이 보류된 어절에 형태소분석 결과를 부여한 경우, 네 번째는 세 번째의 1-best 태깅 대신 k-best 태깅을 사용한 경우, 다섯 번째와 여섯 번째는 세 번째와 네 번째에 통계적태거가 통합된 경우.

(제 10회 한글 및 한국어 정보처리 학술대회)

학습말뭉치		규칙 수		내부실험		외부실험(크기=17,543)			
크기	평균 빈도	총	빈도 제한	TA(%)	TC(%)	모든 규칙		빈도제한 규칙	
						TA(%)	TC(%)	TA(%)	TC(%)
17,100	3.78	7,513	4,441	100	100	97.54	71.24	98.50	63.72
52,036	5.73	18,816	13,479	100	100	98.11	77.38	98.78	71.95
86,366	7.09	29,147	22,476	100	100	98.29	79.42	98.73	74.72
120,710	8.20	38,511	30,724	100	100	98.31	80.76	98.67	76.89
155,828	9.17	47,164	38,348	100	100	<b>98.43</b>	<b>82.20</b>	<b>98.75</b>	<b>78.66</b>

[표 5] 영어 실험 결과: 규칙학습, 내부실험, 외부실험

	$\alpha=1$		$\alpha=2$		$\alpha=3$		$\alpha=4$		$\alpha=5$	
	TA	TC	TA	TC	TA	TC	TA	TC	TA	TC
$\beta=1.0$	98.43	82.20	99.20	74.58	99.49	70.65	99.61	67.84	<b>99.65</b>	65.72
$\beta=0.9$	98.25	87.17	98.95	79.55	99.21	75.62	99.30	72.81	99.34	70.68
$\beta=0.8$	97.87	89.23	98.51	81.61	98.75	77.68	98.82	74.87	98.84	72.74
$\beta=0.7$	97.42	<b>90.69</b>	98.02	83.06	98.22	79.14	98.27	76.33	98.35	74.06

[표 6] 영어 실험 결과: 임계치를 달리한 외부실험

일곱 번째부터 열 번째는 세 번째부터 여섯 번째와 임계치만이 다른 경우( $\alpha=1, \beta=0.9$ )이다. 실험에 사용한 통계적 태거는 [8]의 태거로서 학습 말뭉치 전체에 대해서 95.80% 정확률을 보였다.

	KA	AA
형태소분석	100	3.38
통계적태거	95.43	1
어휘규칙( $\alpha=1, \beta=1.0, 1\text{-best}$ )	<b>99.20</b>	2.02
어휘규칙( $\alpha=1, \beta=1.0, k\text{-best}$ )	99.10	1.63
통합( $\alpha=1, \beta=1.0, 1\text{-best}$ )	96.06	1
통합( $\alpha=1, \beta=1.0, k\text{-best}$ )	96.11	1
어휘규칙( $\alpha=1, \beta=0.9, k\text{-best}$ )	99.05	1.85
어휘규칙( $\alpha=1, \beta=0.9, 1\text{-best}$ )	98.96	<b>1.58</b>
통합( $\alpha=1, \beta=0.9, 1\text{-best}$ )	96.07	1
통합( $\alpha=1, \beta=0.9, k\text{-best}$ )	<b>96.12</b>	1

[표 3] 한국어 실험 결과: 통계적 태거와 결합

[표 3]에 따르면 어휘규칙기반 태거를 k-best 태거( $\alpha=1, \beta=0.9$ )로서 통계적 태거와 결합한 경우에 정확률이 최고로 15.1% (0.69% 정확률)만큼 향상되었다는 것을 알 수 있다. 이것은 정확률이 유사하면서 평균 중의성이 훨씬 낮은 경우, 즉 훨씬 더 넓은 적용률을 갖는 경우( $\beta=1$ 보다  $\beta=0.9, 1\text{-best}$ 보다 k-best)에 통합 모델의 성능이 더 좋다는 것을 말해준다.

통합된 결과의 정확률이 비교적 낮다 하더라도, 규칙에 의해 태깅된 어절은 높은 정확도로 태깅된 것을 보장할 수 있다. 또한 규칙에 의해 태깅된 어절을 쉽게 알 수 있기 때문에 나머지 어절들에 대해서만 오류수정을 거친다면 오류수정을 위한 비용을 최소화할 수 있을 것이다.

[표 4]는 임계치를 달리하면서 어휘규칙기반 태거(1-best)를 통계적 태거와 결합했을 때의 외부실험의 정확률을 보여준다. [표 4]에서 우리는 흥미로운 현상을 발견하였다. 즉, [표 2]에서는 가장 좋은 정확도가  $\alpha=5$ 이고  $\beta=1.0$ 일 때 얻어졌지만, [표 4]에서는  $\alpha=1$ 이고  $\beta=0.9$  또는  $\beta=0.8$ 일 때 가장 좋은 결과를 얻었다. 이로써 우리는 어휘규칙기반 태거가 통계적 태거보다 훨씬

높은 정확도를 가지기 때문에 통합 후의 성능은 규칙의 정확률보다 적용률에 더 영향을 받는다는 것을 알 수 있다.

	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$	$\alpha=5$
$\beta=1.0$	96.06	95.90	95.85	95.79	95.75
$\beta=0.9$	<b>96.07</b>	95.91	95.87	95.81	95.77
$\beta=0.8$	<b>96.07</b>	95.89	95.85	95.78	95.75
$\beta=0.7$	96.00	95.83	95.78	95.75	95.66

[표 4] 한국어 실험 결과: 임계치를 달리한 결합

## 4.2 영어태깅

Brown Corpus에서 추출된 173,371 단어(한국어와 영어의 비교를 위해 말뭉치의 크기가 비슷하도록 선정하였다)의 품사부착 말뭉치(품사집합은 동일) 중에서 155,828 단어(약 90%)는 학습말뭉치로 사용되었고 17,543 단어(약 10%)는 실험말뭉치로 사용되었다.

[표 5]는 [표 1]과 비교해 볼 때 비슷한 크기의 학습말뭉치에 대해서 한국어의 경우보다 영어의 경우가 비교적 성능이 좋다는 것을 보여준다. 그 이유는 한국어는 영어보다 훨씬 많은 어형을 가지므로 어절의 평균빈도가 낮기 때문에 신뢰도 높은 규칙을 학습하기 위해서는 더 큰 학습말뭉치가 필요하기 때문이다.

[표 6]은 [표 2]와 비교해 볼 때 한국어의 경우보다 영어의 경우에 어휘규칙의 적용률이 훨씬 높은 것을 보여준다.

[표 7]은 통계적 태거와 결합할 때의 외부실험 결과를 보여주고 있다. 실험에 사용된 통계적 태거는 품사bigram모델과 어휘확률을 사용하는 은닉마르코프모델(HMM)기반 태거이다. 실험에 사용된 통계적 태거의 성능(내부실험: 98.19% 정확률, 외부실험: 96.89% 정확률)이 다른 기존 태거들의 성능에 비해서 비교적 높은 것은 사전이 제한되었고(Brown Corpus 전체에서 추출됨) 따라서 미등록어가 없기 때문인 것으로 추측된다.

[표 7]에 따르면 어휘규칙기반 태거와 통계적 태거를 통합함으로써 성능이 최고 6.8%(0.21% 정

확률)만큼 향상된다. 이것은 [표 3]의 한국어와 비교하면 낮은 향상률이다. 이런 결과가 나오는 이유 중에 하나는 제한된 사건을 사용함으로써 통계적 태거의 성능이 높았기 때문이고 또 하나는 영어에서는 품사문맥과 어휘확률로 해결될 수 있는 중의성이 한국어보다 많으므로 어휘문맥을 고려해야 해결되는 중의성이 한국어에서보다 적기 때문으로 풀이된다.

	KA	AA
형태소분석	100	1.71
통계적태거	96.89	1
어휘규칙( $\alpha=1, \beta=1.0, 1\text{-best}$ )	98.71	1.24
어휘규칙( $\alpha=1, \beta=1.0, k\text{-best}$ )	98.68	<b>1.18</b>
통합( $\alpha=1, \beta=1.0, 1\text{-best}$ )	97.06	1
통합( $\alpha=1, \beta=1.0, k\text{-best}$ )	97.06	1
어휘규칙( $\alpha=2, \beta=1.0, 1\text{-best}$ )	<b>99.40</b>	1.33
어휘규칙( $\alpha=2, \beta=1.0, k\text{-best}$ )	99.36	1.24
통합( $\alpha=2, \beta=1.0, 1\text{-best}$ )	<b>97.10</b>	1
통합( $\alpha=2, \beta=1.0, k\text{-best}$ )	<b>97.10</b>	1

[표 7] 영어 실험 결과: 통계적 태거와 결합

또한 한국어의 경우에는 k-best 태깅 방식으로 통합하였을 경우에 약간의 정확률 향상을 얻었지만 영어의 경우에는 그렇지 못함을 알 수 있다. 이것은 영어 실험에 사용된 형태소분석기가 제한된 사건을 사용함으로써 k-best 태깅의 효과로 이미 갖고 있기 때문이다.

	$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$	$\alpha=5$
$\beta=1.0$	97.06	<b>97.10</b>	97.06	97.01	97.01
$\beta=0.9$	97.05	97.08	97.04	96.99	96.99
$\beta=0.8$	97.01	97.03	96.99	96.94	96.94
$\beta=0.7$	96.88	96.91	96.87	96.82	96.85

[표 8] 영어 실험 결과: 임계치를 달리한 결합

[표 8]은 임계치를 달리하면서 어휘규칙기반 태거(1-best)를 통계적 태거와 결합했을 때의 외부실험의 정확률을 보여준다. [표 8]은 한국어와는 다른 양상을 보여준다. 즉, [표 6]과 비교할 때, 결정적인 규칙을 사용하는 경우( $\beta=0.1$ )와  $\alpha$ 가 2인 경우가 비교적 다른 경우보다 좋은 성능을 나타냄을 알 수 있다. 이런 결과가 나오는 것은 영어 실험에서는 통계적 태거도 비교적 높은 정확도를 보이기 때문에 통합 후의 성능은 어휘규칙의 적용률보다 정확률에 더 영향을 받는 것으로 풀이된다.

#### 4.3 기존 연구와 비교

[12]는 20만 원시말뭉치에서 수동으로 추출된 8,002개의 어휘규칙을 사용하여 25,463 어절의 학습말뭉치와 20,893 어절의 실험말뭉치에 대해서 실험한 결과 어휘규칙의 정확률과 적용률은 내부실험에서 99.13%와 63.57%를 보였고 외부실험에서 99.01%와 61.89%를 보였다. 또한 통계적 품사태거와 결합하여 내부실험에서 30.48%만큼(93.70%에서 95.62%로) 정확률을 향상시켰고 외부실험에서 26.07%만큼(93.21%에서 94.98%로) 정확률을 향상시켰다. [12]에서 사용하는 통계적

태거[4]와 본 논문에서 사용하는 통계적 태거[8]의 정확률이 다소 차이가 나고 또한 말뭉치의 크기가 다르기 때문에 직접적인 비교는 될 수 없지만 본 논문의 향상률이 외부실험에서 15.1%로서 [12]에서보다는 낮은 향상률을 보인다. 그러나 본 논문의 통계기반 어휘규칙은 자동으로 획득할 수 있다는 장점을 갖는다.

#### 4.4 태깅오류 교정

일반적으로 수동 태깅된 말뭉치는 여러 태깅 전문가들의 견해의 차이나 태깅 전문가 한 사람의 시간에 따른 견해의 변화에 의해 잘못 태깅되거나 일관성이 없이 태깅되는 경향이 높다. 이러한 태깅오류가 문제가 되는 것은 대량의 말뭉치에서 태깅오류를 찾는 것이 쉽지 않기 때문이다.

통계기반 어휘규칙은 학습말뭉치에 대해서 100% 태깅 정확률을 보이는 기억태거(memorizing tagger)이기 때문에 말뭉치의 태깅오류를 감지하고 교정하는 데에 효과적으로 이용될 수 있다.

만약 어떤 어절이 잘못 태깅되거나 일관성이 없이 태깅되었다면<sup>4)</sup> 그 어절에 대해서 비결정적 규칙이 획득될 것이고 그 비결정적 규칙에는 오류태그가 존재하고 그 규칙에 대한 세부규칙들 중에도 오류가 존재한다. 예를 들어, '나가야'가 '... 가꾸어 나가야 할 ...'에서는 옳은 태그 '나가\_본동사+아야\_연결어미'로 태깅되고 '... 놓여 나가야 할 ...'에서는 오류 태그 '나가\_보조동사+아야\_연결어미'와 같이 태깅되었다고 하자. '나가야'가 나타나는 다른 용례는 없다고 할 때 '나가야'에 대한 가장 일반적인 규칙은 [그림 10]의 b이고 규칙 b의 세부규칙들은 c와 d가 될 것이다. 규칙 b의 두 번째 태그는 오류 태그이므로 규칙 d에도 오류가 존재함을 알 수 있다.

b. (나가야 0 0 2 2 (나가_본동사+아야_연결어미 1) (나가_보조동사+아야_연결어미 1))
c. (나가야 1 0 가꾸어 1 1 (나가_본동사+아야_연결어미 1))
d. (나가야 1 0 놓여 1 1 (나가_보조동사+아야_연결어미 1))

[그림 10] '나가야'에 대한 규칙의 예

따라서 오류를 포함하는 비결정적규칙과 세부규칙들에서 오류를 교정하고 교정된 규칙을 말뭉치에 적용함으로써 말뭉치내의 오류를 교정할 수 있다. 예를 들어, [그림 10]의 규칙 b는 [그림 11]의 규칙 b'로 교정하고 규칙 c와 d는 제거한다. 그러면 규칙 b'에 의해 말뭉치의 오류가 교정된다.

b'. (나가야 0 0 2 1 (나가_본동사+아야_연결어미 2))
--------------------------------------

[그림 11] '나가야'에 대한 교정된 규칙

이렇게 함으로써 오류를 감지하기 위해 말뭉치내의 모든 어절을 검사하지 않고 상대적으로 적은 수의 비결정적 규칙들을 검사함으로써<sup>5)</sup> 효율

4) 같은 태그로 태깅돼야 할 때 다른 태그로 태깅된 경우를 말한다.

적으로 태깅오류를 감지하고 교정할 수 있다. 실제로 본 논문에서 사용한 한국어 말뭉치 전체에서 규칙을 획득한 후, 모든 비결정적 규칙 대신에 적용범위가 넓은 148개의 비결정적 규칙을 조사하여 이 중에서 오류가 있는 107개의 규칙을 수정한 뒤 다시 말뭉치에 적용함으로써 222개의 태깅오류를 교정하였다. 이로써 교정된 규칙이 적용될 수 있는 53,215 어절들에 대해서는 비일관성 태깅 오류가 없다는 것을 보장할 수 있다. 또한 형태소분석 원칙이나 사전정보가 상이하여 발생하는 비일관성 오류를 가진 5,313 어절에 대한 규칙 9,188개에서 오류를 수정하여 얻은 7,850개의 규칙과 원시어절 오류를 수정하여 얻은 23개를 수정하여 얻은 11개의 규칙을 이용하여 14,583어절에서 12,895개의 오류를 교정함으로써 최종적으로 실험에 사용할 말뭉치를 구축하였다.

### 5 결론

본 논문에서는 통계기반 어휘규칙을 말뭉치로부터 중의성주도 학습을 이용하여 자동획득하는 방법과 어휘규칙을 이용하여 통계적 품사태거의 성능을 향상시키는 방법을 제안하였다. 제안된 규칙태거는 다음과 같은 특징을 갖는다.

첫째로, 제안된 태거는 학습가능 규칙기반 태거이다. 이것은 [10, 12]의 어휘규칙이 수동으로 획득되는 것과 대조적이다.

둘째로, 제안된 태거는 어휘문맥을 고려하는 통계기반 어휘규칙을 사용하기 때문에 높은 정확률로 태깅할 수 있다. 또한 학습말뭉치의 크기가 커질수록 외부실험의 태깅 정확률과 태깅 적용률이 향상되는 것을 관찰할 수 있었다.

셋째로, 제안된 태거는 통계적 태거와 같은 다른 태거들과 쉽게 결합될 수 있다. 예를 들면, 1-best 태거나 k-best 태거로 통계적 태거와 결합되어 중의성을 최소화함으로써 비교적 정확도가 낮은 통계적 태거의 성능을 향상시킬 수 있다.

넷째로, 제안된 태거는 어절단위 띄어쓰기 언어라면 어디에든 적용될 수 있다. 한국어뿐만 아니라 영어에 대해서도 실험을 수행한 결과, 비교적 어휘문맥이 형태론적 중의성 해결에 중요한 영향을 미치는 한국어가 그렇지 않은 영어보다 제안된 태거가 더욱 효과적이라는 것을 알 수 있었다.

마지막으로, 통계기반 어휘규칙은 기존 어휘규칙과 마찬가지로 이해하기가 쉽기 때문에 규칙 수정이 용이하며 말뭉치를 그대로 반영하기 때문에 어휘규칙 수정으로 말뭉치의 태깅 오류를 교정할 수 있다는 장점을 갖는다.

반면 제안된 태거는 규칙의 수가 많다는 단점을 갖고 있다. 그러나 응용분야에 따라서 필요한 규칙을 적절히 선택할 수 있으므로 이 문제는 심각한 것 같지 않다. 예를 들면 우리는 앞선 실험

에서 학습말뭉치에서 단지 한번 출현한 중심어에 대한 규칙을 제거함으로써 규칙 수를 상당히 줄일 수 있었다. 계속해서 태깅 정확률을 저하시키지 않으면서 덜 효과적인 규칙들을 제거함으로써 규칙 수를 줄이는 방법에 대한 연구를 수행할 것이다.

### 참고문헌

- [1] Brill E., Some Advances in Transformation-Based Part of Speech Tagging, *Proc. of AAAI-94*, pp. 722-727, 1994
- [2] Church K. W. Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text, *Proc. of the 2nd Conference on ANLP*, pp. 136-143, 1988
- [3] Cutting D., Kupiec J., Pedersen J., and Sibun P. A Practical Part-of-Speech Tagger, *Proc. of the 3rd Conference on ANLP*, pp. 133-140, 1992
- [4] Kim Jin-Dong, Lim Heui-Seok, Lee Sang-Zoo and Rim Hae-Chang, Twoply Hidden Markov Model: A Korean Part-of-Speech Tagging Model Based on Morpheme-Unit with Eojeol-Unit Context," *CPOL(Computer Processing of Oriental Languages)*, Vol.11, No.3, pp. 277-289, 1997
- [5] Ratnaparkhi A., A Maximum Entropy Model for Part-of-Speech Tagging, *Proc. of EMNLP*, pp. 133-142, 1996
- [6] Tapanainen Pasi and Voutilainen Atro, Tagging Accurately - Don't Guess if You Know, *Proc. of the 4th Conference on ANLP* pp. 149-156, 1994
- [7] Zhang M., Li S. and Zhao T., Tagging Chinese Corpus Based on Statistical and Rule Techniques, *Proc. of the 17th ICCPOL*, pp. 503-506, 1997
- [8] 김진동, 이상주, 임해창, 어절 띄어쓰기를 고려한 형태소 단위 품사 태깅 모델, 제10회 한글 및 한국어정보처리 학술대회 논문집, 1998
- [9] 이상주, 임희석, 임해창, 은닉 마르코프 모델을 이용한 두단계 한국어 품사 태깅, 제6회 한글 및 한국어정보처리 학술대회 논문집, 6권, 1호, pp. 305-312, 1994
- [10] 이정규, 이상주, 임희석, 임해창, 규칙기반 한국어 품사 태깅을 위한 어휘 규칙 획득의 수작업 최소화 방안, 제24회 정보과학회 봄 학술 발표논문집, Vol.24, No.1, pp. 479-482, 1997
- [11] 이호, 언어 정보 획득을 위한 한국어 코퍼스 분석 도구, 고려대학교 석사학위논문, 1994
- [12] 임희석, 언어 지식과 통계 정보를 이용한 한국어 품사 태깅 모델, 고려대학교 박사학위 논문, 1997

5) 이 방법은 결정적 규칙에 존재하는 오류, 즉 다르게 태깅되어야 하는 것이 같게 태깅되는 경우는 발견할 수 없다.