

# 술어 하위범주화 정보를 이용한 한국어 의존 파서

장 명길, 이 현아, 박 재득, 박 동인(시스템공학연구소)  
황 도삼(영남대학교)

## Korean Dependency Parser Using Subcategorization Information of Predicates

Myung-Gil Jang, Hyun-A Lee, Jae-Deuk Park, Dong-In Park(SERI)  
Dosam Hwang(Yeungnam University)

### 요 약

의존 문법을 사용한 의존 파싱에서 기본적인 단어(품사정보)들 사이의 의존 관계 검사에 의한 파싱 방법은 불필요한 의존 관계의 생성을 가져온다. 이러한 과생성을 해결하기 위하여 파싱 단계에서 보다 정교한 의존 파싱을 통해서 불필요한 의존 관계의 생성을 최소한으로 줄이는 방법에 대한 연구가 필요하다.

본 논문은 의존 파싱에서 최소한의 의존 관계를 생성하기 위하여 후보 의존소가 지배가능경로 상에서 술어 지배소와의 의존 관계 검사 시에 술어의 하위범주화 정보를 이용하는 효율적인 의존 파싱 방법을 제안한다. 이것은 의존 파싱의 다음 처리 단계인 의존 제약의 적용에 훨씬 부담을 덜어 줄 수 있다.

## I. 서론

최근의 한국어 구문 분석에 대한 연구에서 의존 문법이 널리 사용되고 있다. 의존 문법은 어형(wordform) 사이의 의존 관계에 중점을 두는 문법으로 어순의 제약을 거의 받지 않기 때문에 어순의 도치나 주요 성분의 생략이 일어나는 한국어 문장에 대한 분석에 상당히 우수한 결과를 보여준다[손광 94]. 의존 문법을 이용한 한국어 의존 파싱은 한국어 특성을 고려한 매우 견고성이 있는(robust) 파싱을 가능하게 하는 반면에 단순히 두 단어 사이의 의존 관계만을 검사하므로 불필요한 의존 관계의 생성을 가져올 수 있다.

따라서, 실제의 의존 파서 시스템의 구현에는 이러한 의존 관계의 과생성을 개선하기

위해서 초기의 의존 파싱 단계에서 최소한의 의존 관계를 생성하도록 하는 방법의 연구 [김창 93]나 이후에 의존 제약 단계에서 여러 가지 의존 제약의 적용을 통하여 불일치한 의존 관계를 제거하는 방법에 대한 연구[손광 95]가 필요하다.

[김창 93]에서는 한국어 문장에 대한 구문 분석에서 왼쪽 우선(left-to-right) 방법에 비하여 효율적인 오른쪽 우선(right-to-left) 의존 파싱 방법을 제안하였다. 또한, 의존 파싱에서 의존관계의 검사를 위해서 지배가능경로(headable path) 상의 문장 성분만을 조사함으로써 의존 트리의 수를 줄여 분석 수행 시간을 단축할 수 있음을 보여주었다. 그러나, 술어가 여러 개 나타나는 한국어 문장에서 이러한 지배가능경로 만을 고려한 의존 파싱은 불필요한 의존 관계의 생성을 어느 정도는 줄일 수 있지만 여전히 많은 의존 관계를 생성할 수 있다.

[손광 95]는 두 어절사이의 가능한 의존 관계를 구조적 의존관계 네트워크(SDRN)에 표현하고, 제약조건 만족문제(constraint satisfaction problem)라는 개념을 적용하여 애매성이 있는 문장을 자연스럽게 분석할 수 있음을 보여주고 있다.

또한, 한국어는 문장의 구조를 이루는 구성요소들 중, 술어가 가장 지배적인 위치에서 다른 구성요소들을 지배하면서 문장구조를 제약하는 기능을 수행하기[강 95] 때문에 술어를 중심으로 한 구문 분석 방법이 효과적이다. [한용 95]는 술어를 중심으로 한 한국어 문형정보를 확장된 PATR II 문법의 기술(description)에 이용하여 한국어 문장을 분석하였다. 이 방법은 한국어 문형정보가 술어의 하위범주화 정보로 완벽히 구축되어 있어야 하고, 문형정보 자체가 의미론적 특성을 반영하고 있기 때문에 명사의미체계에 대한 정립이 선결되어야 한다.

본 논문은 앞서 언급한 초기 의존 파싱 단계에서 불필요한 의존 관계를 최대한 줄이는 방법에 대한 연구에 초점을 맞추어, 의존 파싱에서 술어의 하위범주화 정보를 적절히 이용하는 보다 효율적인 의존 파싱 방법을 제안한다. 이러한 의존 파싱 방법은 지배가능경로의 개념과 오른쪽 우선 의존 파싱을 기본으로 하고, 후보 의존소의 지배가능경로 상의 술어 지배소와의 의존 관계 검사에 술어의 하위범주화 정보를 적절히 이용하도록 한다.

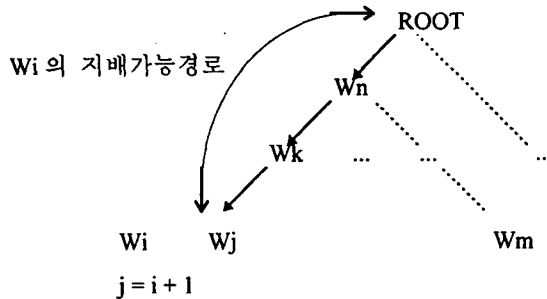
앞으로 2 장에서는 의존 파싱에서 술어 하위범주화 정보의 이용에 대하여 살펴보고 3 장에서는 예문을 가지고 의존 파싱 방법을 비교한다. 다음으로 4 장에서는 전체 시스템 구성과 실험 결과를 설명하고, 마지막으로 5 장에서 결론을 맺는다.

## II. 술어 하위범주화 정보를 이용한 의존 파싱

먼저, 의존 파싱에서의 지배가능경로의 개념에 대하여 살펴보고 술어의 하위범주화 정보를 이용한 의존 파싱에 대하여 설명한다.

## 1. 지배가능경로와 의존 파싱

[김창 93]에서 지배가능 경로를 다음과 같이 설명하고 있다. 문장 S가 문장 성분  $W_1 W_2 \dots W_n$ 으로 구성되고 문장 성분  $W_j$ 가 문장 성분  $W_k$ 에 의존할 때, 의존 관계  $W_k \Rightarrow W_j$ 를 정의할 수 있다. 이 때,  $W_j$ 의 바로 앞 문장 성분  $W_i$ 의 지배 성분은  $W_i$ 의 지배가능경로 상에 존재해야 한다. 다음 그림과 같이  $W_i$ 의 지배가능경로를 설명할 수 있다.



본 논문의 의존 파싱은 오른쪽 우선 파싱에 앞서 기술한 지배가능경로의 개념을 이용하고 의존 관계 검사는 일반적인 의존 관계 검사 루틴과 술어 하위범주화 정보를 이용한 의존 관계 검사 루틴을 둔다.

## 2. 술어 하위범주화 정보를 이용한 의존 파싱

술어 하위범주화 정보를 이용한 의존 관계 검사 루틴은 후보 의존소-지배소에 대하여 다음과 같은 ‘술어 하위범주화 정보 제약’을 적용한다.

### 술어 하위범주화 정보 제약(Subcat. Constraint)

의존소  $D$ 와 지배소 술어  $G$ 의 의존 관계 검사는 의존소의 격정보  $Case\_M$ 가 지배소 술어의 하위범주화 정보  $Subcat$ 의 한 요소를 만족해야 함.

후보 의존소-지배소의 의존 관계 검사는 위의  $Subcat. Constraint$ 를 만족하면 의존 관계를 생성할 수 있게 하고 그렇지 않으면 실패한 것으로 한다.

예를 들어, “어머니는 우리와 헤어졌다.”는 문장은 자동사 술어 ‘헤어지’가 ‘N이’+’N와’를 문형 정보로 요구하기 때문에 하위범주화 정보 값은  $Subcat\{+subj\}$ 로 표현된다. 또한, 의존소 ‘어머니’와 ‘우리와’는 각각 격정보로  $Case\_M\{subj\}$ ,  $Case\_M\{jct\}$ 임을 알 수 있다. 이때, 술어 하위범주화 정보를 이용한 의존 관계 검사에서 이들 의존소는 각각 지배소 ‘헤어지’에 대하여  $Subcat. Constraint$ 를 만족하기 때문에 ‘주격’ 관계와 ‘보격’ 관계의

의존 관계를 생성하게 된다.

문장  $S$  가 문장 성분  $W_1, W_2, \dots, W_n$  으로 구성될 때, 술어 하위범주화 정보를 이용한 의존 파싱은, 다음 두 단계로 나뉘어 수행된다.

- [1] 단계: ‘술어 하위범주화 정보 적용 범위’(Subcat\_Range)의 설정

문장  $S$  에 나타나는 술어  $W_j$  의 Subcat\_Range 는 술어  $W_j$  의 하위범주화 정보를 고려할 때 술어  $W_j$  가 지배하는 문장 성분들의 집합을 말한다. 즉,  $W_j$  와 그 앞에 나타나는 술어  $W_{j-x}$  와의 사이에 있는 문장 성분들 중, 술어  $W_j$  의 Subcat. Constraint 를 만족하는 요소이다. 그러나, 술어  $W_j$  가 관형절을 이끄는 술어인 경우는  $W_j$  의 피수식어  $W_{j+1}$  이 Subcat\_Range 에 포함될 수 있기 때문에, Subcat\_Range 는 술어  $W_j$  의 어미형태와 다음 성분  $W_{j+1}$  을 함께 고려하여 결정하여야 한다. 이러한 방법으로 얻은 Subcat\_Range 는 의존 파싱에서 문장내의 투영성(Projectivity)을 자연스럽게 따르게 한다.

- [2] 단계: Subcat\_Range 와 Subcat. Constraint 를 적용한 의존 관계 검사

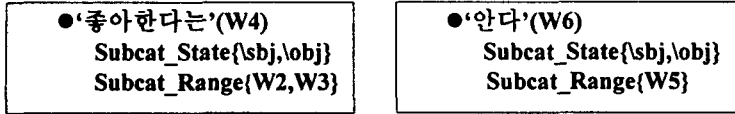
오른쪽 우선 의존 파싱에서 현재의 후보 의존소  $W_i$  와  $W_i$  의 지배가능경로 상의 술어 지배소  $W_j$  와의 의존 관계 검사에는 술어  $W_j$  의 Subcat\_Range 와 Subcat. Constraint 를 적용한다. 의존 관계 검사가 성공하면, 의존 관계를 생성하고 그 술어의 Subcat\_State 는 변경된다.

두 단계 의존 파싱의 알고리즘은 다음과 같다.

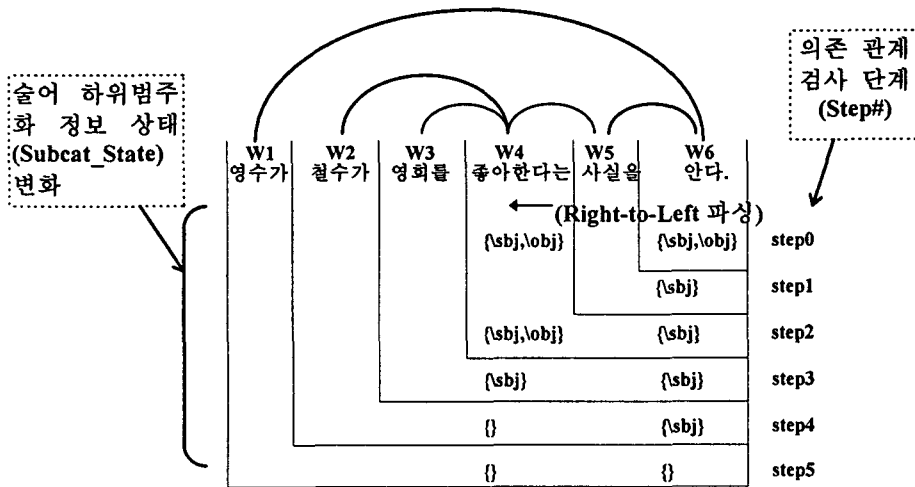
1. 문장  $S$  가  $W_1, W_2, \dots, W_n$  일 때, 문장에 나타나는 모든 술어  $W_j$  에 대해 Subcat\_Range 를 결정한다.
2. 현재의 후보 의존소  $W_i$  ( $i$  는  $n$  부터 1 까지)에 대하여  $W_i$  의 지배가능경로 상의 후보 지배소  $W_j$  와 의존 관계 검사를 할 때,
  - 2.1 if ( $W_i$  가  $W_j$  의 Subcat\_Range 와 Subcat. Constraint 를 만족하면)  
then  $W_i$  와  $W_j$  는 의존 관계를 생성하고  $W_j$  의 Subcat\_State 는 변경된다.  
또한, 더 이상 지배가능경로 상에 남아 있는 지배소와는 의존 관계 검사를 수행하지 않는다.
  - 2.2 if ( $W_i$  가 어떤  $W_j$  의 Subcat\_Range 에도 속하지 않으면)  
then if ( $W_i$  가  $W_j$  의 Subcat. Constraint 를 만족하면)  
then  $W_i$  와  $W_j$  는 의존 관계를 생성하고  $W_j$  의 Subcat\_State 는 변경된다.  
또한, 더 이상 지배가능경로 상에 남아 있는 지배소와는 의존 관계 검사를 수행하지 않는다.

예를 들어, 문장 “영수가( $W_1$ ) 철수가( $W_2$ ) 영희를( $W_3$ ) 좋아한다는( $W_4$ ) 사실을( $W_5$ ) 안다( $W_6$ ).”에 대한 의존 파싱을 설명한다.

두 단계 의존 파싱의 단계 [1]에서 술어 ‘좋아한다는’(W4)과 ‘안다’(W6)의 하위범주화 정보 초기 상태 Subcat\_State와 적용 범위 Subcat\_Range는 각각 다음과 같다.



단계 [2]의 의존 파싱은 다음 그림과 같이 의존 관계 검사가 step0부터 step5까지 수행된다.



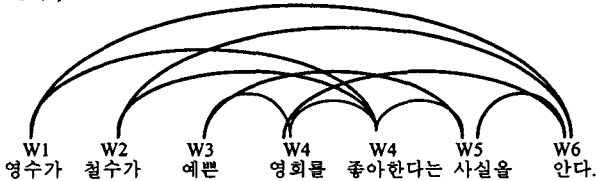
의존 관계 검사 step1에서 ‘사실은’(W5)은 술어 ‘안다’(W6)의 Subcat\_Range에 속하고 Subcat\_Constraint를 만족하기 때문에 ‘목적격’ 의존 관계가 생성되고 ‘안다’(W6)의 Subcat\_State가 {sbj}로 변경된다. Step2에서는 술어 ‘좋아한다는’(W4)과 ‘사실은’(W5) 사이에 술어 하위범주화 정보를 이용한 의존 관계 검사 루틴이 적용되지 않기 때문에 일반 의존 관계 검사 루틴에 의하여 ‘연결(동격)’ 관계의 의존 관계가 생성된다. 다음 step3에서는 ‘영희를’(W3)이 지배가능경로 상의 후보 지배소들 중, 먼저 나타나는 술어 ‘좋아한다는’(W4)와의 의존 관계 검사에서 Subcat\_Range와 Subcat\_Constraint를 만족하기 때문에, ‘목적격’ 의존 관계를 생성하면서 하위범주화 상태는 Subcat\_State{sbj}이 된다. 이때, 지배가능경로 상의 나머지 다른 지배소(술어 ‘안다’(W6) 등)와는 더 이상의 의존 관계 검사를 시도하지 않는다. 같은 방법으로 step4에서도 ‘철수가’(W2)는 술어 ‘좋아한다는’(W4)과 ‘주격’ 의존 관계를 생성한다. 마지막으로 step5에서는 지배가능경로 상의 어떤 지배소 술어의 Subcat\_Range에도 속하지 않은 ‘영수가’(W1)는 가까운 술어의 순서로 Subcat\_Constraint를 적용함으로써 어절 ‘안다’(W6)와 ‘주격’ 의존 관계를 생성한다.

### III. 의존 파싱의 비교

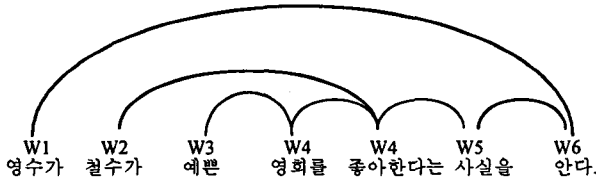
다음은 한국어 문장을 오른쪽 우선 방법으로 의존 파싱할 때, 지배가능경로만을 이용한 경우 (1)과 지배가능경로와 술어 하위범주화 정보를 함께 이용한 경우 (2)의 의존 파싱을 예문을 가지고 비교한다.

문장 “영수가(W1) 철수가(W2) 예쁜(W3) 영화를(W4) 좋아한다는(W5) 사실을(W6) 안다(W7).”에 대하여 위의 (1), (2) 각각의 경우의 의존 파싱에 따라 생성된 의존 관계를 비교하면 다음과 같다.

(1)의 경우,



(2)의 경우,



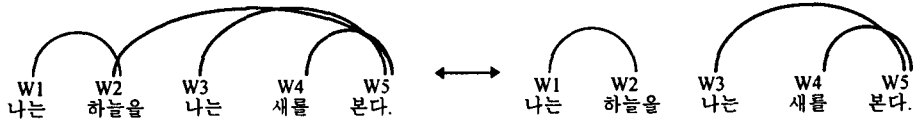
위와 같이 ‘(1)의 경우’의 의존 파싱은 생성되는 의존 관계의 수가 10 개이고, ‘(2)의 경우’의 의존 파싱은 6 개 의존 관계만이 생성됨을 알 수 있다.

다음으로 애매성이 있는 문장 “나는(W1) 하늘을(W2) 나는(W3) 새를(W4) 좋아한다(W5).”에 대한 (1), (2)의 경우의 의존 파싱을 비교한다. 이 문장은 첫번째와 세번째 어절인 ‘나는’(W1,W3)이 형태소 분석 단계에서 자동사 술어 ‘나(pv,Subcat{sbj})+ㄴ(etm2)’와 대명사 ‘나(npp)+는(jx)’ 그리고 타동사 술어 ‘날(pv,Subcat{sbj,obj})+ㄴ(etm2)’의 3 가지로 분석되기 때문에 의존 파싱에서는 모두 9 개의 후보 입력 열에 대한 의존 파싱을 수행하여야 한다. 각각의 입력 열에 대한 (1)과 (2)의 의존 파싱 결과는 다음과 같다.

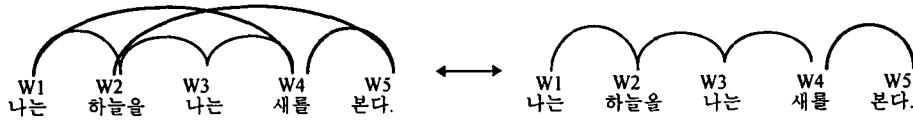
[1] W1 ‘나(pv,Subcat{sbj})’ + W3 ‘나(pv,Subcat{sbj})’



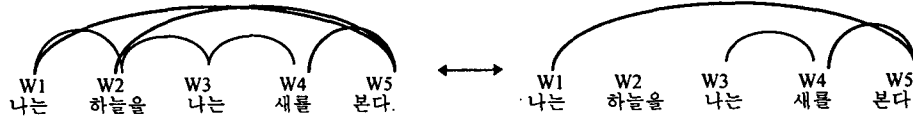
[2] W1 '나(pv,Subcat{sbj}) + W3 '나(npp)'



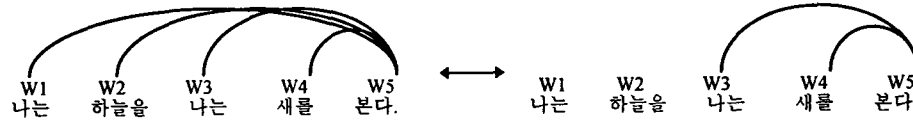
[3] W1 '나(pv,Subcat{sbj}) + W3 '날(pv,Subcat{sbj,obj})'



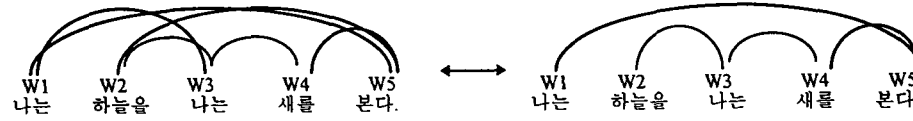
[4] W1 '나(npp)' + W3 '나(pv,Subcat{sbj})'



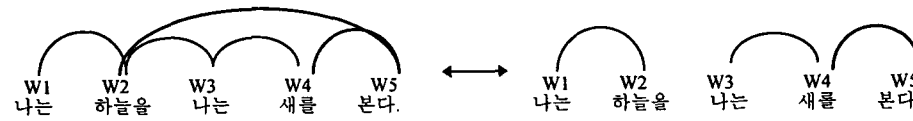
[5] W1 '나(npp)' + W3 '나(npp)'



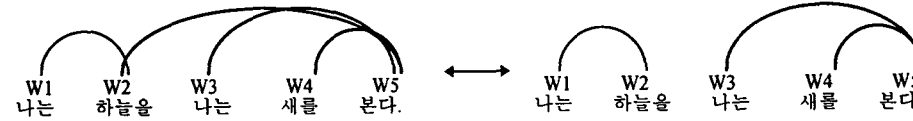
[6] W1 '나(npp)' + W3 '날(pv,Subcat{sbj,obj})'



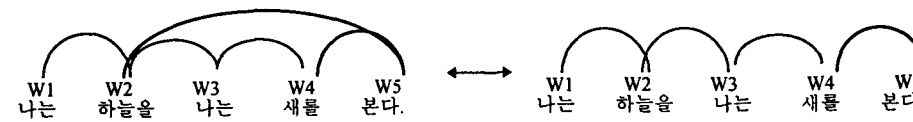
[7] W1 '날(pv,Subcat{sbj,obj}) + W3 '나(pv,Subcat{sbj})'



[8] W1 '날(pv,Subcat{sbj,obj}) + W3 '나(npp)'



[9] W1 '날(pv,Subcat{sbj,obj}) + W1 '날(pv,Subcat{sbj,obj})'



위와 같이 애매성이 있는 문장의 분석에서는 불필요한 의존 관계의 생성이 더욱 많이 발생할 여지가 있으므로 (2)의 경우에는 더욱 효과적으로 불필요한 의존 관계를 제거할 수 있게 되어, [3], [6], [9]의 단지 3개의 결과에 대해서만 의존 파싱의 다음 처리 단

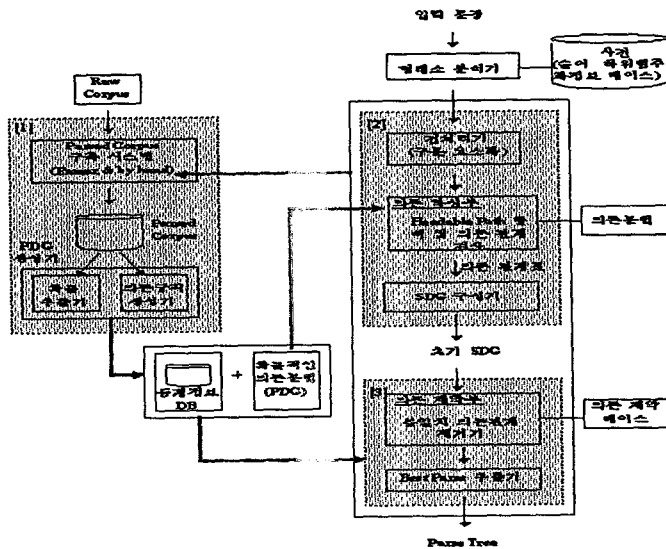
계인 의존 제약을 적용하면 된다.

#### IV. 한국어 구문 분석 시스템(KOrean Syntactic Analyzer; KOSA)

본 논문에서 구현한 한국어 구문 분석 시스템은 포항공대에서 개발하고 있는 ‘한국어 형태소 해석기’와 SERI에서 개발 중인 ‘전자사전’을 통합하여 전체 시스템을 구성하였다.

##### 1. 전체 시스템 구성

다음 그림은 한국어 구문 분석 시스템의 전체 구성도이다.



전체 시스템의 세부 모듈에 대하여 살펴 보면, 먼저, [1] 부분은 **Parsed Corpus**로부터 확률적인 의존 문법을 생성하고 통계 정보 베이스를 구축하는 부분이다. 그리고, [2] 부분은 전처리기와 의존 파싱부 그리고 SDG(Syntactic Dependency Graph) 생성기로 구성되는데, 전처리기는 구문 분석기의 처리가 효율적으로 되도록 하기 위해서 형태소 해석 결과를 구문 요소 단위로의 어절 통합을 수행한다. 즉, [김창 95]에서와 같이 본 용언 어절과 그 다음에 나오는 기능 어절간의 결합, 복합어구의 추출 등을 행한다. 그리고, 의존 파싱부는 앞에서 설명한 술어 하위범주화 정보를 이용한 의존 파싱을 수행하는 모듈이며 SDG 생성기는 의존 파싱부로부터 생성된 의존 관계표로부터 초기 SDG를 생성하는 모듈이다. 마지막으로 [3] 부분은 의존 제약부와 Best Parse 추출기로 구성되는데, 의존 제약

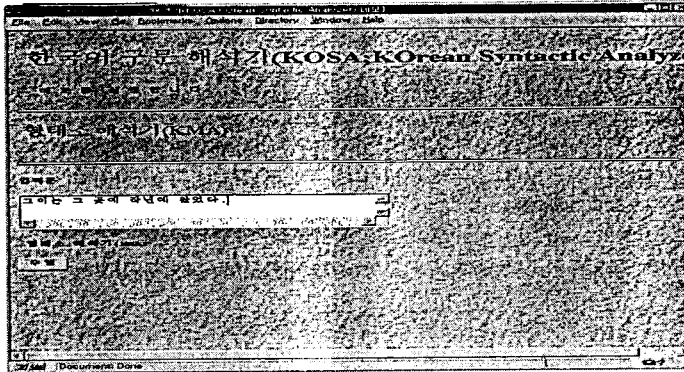


부에서는 유일 지배소 제약과 유일 범주 제약, 그리고 확률값 제약들을 사용하여 불일치하는 의존 관계를 제거하는 일을 수행한다.

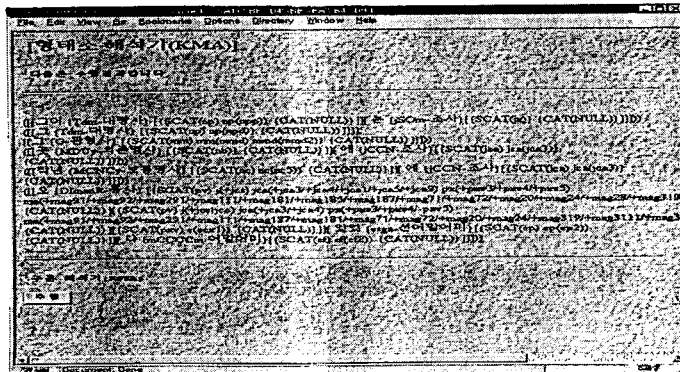
본 연구는 현재 [2]와 [3] 부분에 대한 연구를 중심으로 구현되었고 향후 연구로는 [3] 부분의 연구인 **Corpus**로부터 얻는 통계 정보를 적절히 활용하는 시스템으로 개선할 예정이다.

## 2. 응용 프로그램 인터페이스(API) 시스템

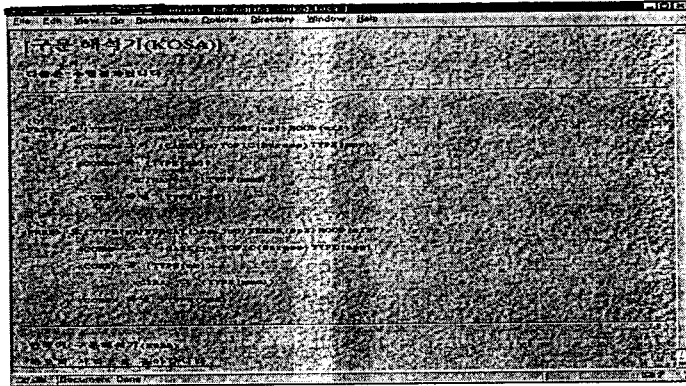
본 논문에서는 현재 다른 연구로 개발 중인 형태소 해석기와 사전을 구문 분석기에 통합하여 사용자에게 편리한 환경을 제공하는 응용 프로그램 인터페이스 시스템을 구축하였다. 이 시스템은 실제 실행 프로그램들이 수행되는 **Web 서버(Ultra O.S. W/S)**에 연결된 **Netscape** 응용 프로그램을 사용하여 사용자 인터페이스를 마련하였다. 사용자는 **Netscape** 브라우저에서 문장 입력과 형태소 해석기 수행, 그리고, 구문 해석기 수행을 차례 대로 하면서 그 결과를 보여준다. 다음은 문장 “그이는 작년에 그 곳에 왔었다.”에 대한 **Netscape** 초기 문장 입력 화면과 형태소 수행 결과, 구문 해석 수행 결과 화면이다.



<초기 문장 입력>



<형태소 해석 수행 결과>



<구문 해석 수행 결과>

### 3. 실험

본 논문은 한국어 구문 분석 시스템의 실험을 위하여 한국어 구문 분석기 성능 시험용 문장 모음인 SERISet96 을 사용하였다. SERISet96 은 [장석 95]에서 추출한 한국어 문법의 일반적인 유형을 포함한 문장들과 [강은 95]으로부터 한국어 술어 문형 기준으로 추출한 문장들, 그리고 한국어 구문 분석 평가용으로 선정한 예문들을 포함하여 모두 200 문장으로 구성되어 있다. 이들 문장은 2 어절부터 10 어절까지 한국어 기본 문형에서 나타나는 다양한 문장들로 구성되며, 평균 어절 수는 약 5 어절이다.

실험을 통한 시스템의 성능 평가는 두 가지 방법으로 수행하였다. 첫번째 방법은 SERISet96 전체 문장에 대해서 파싱 결과의 정확성에 대한 평가이고 두번째 방법은 SERISet96 에서 추출한 문장들에 대하여 생성되는 의존 관계의 수를 비교 평가한 것이다.

첫번째 평가는 파싱 결과의 정확성 여부를 다음과 기준에 따라 국문학과 졸업생에 의해 실시 되었다. 즉, 파싱 결과가 한 개이고 옳은 경우는 OO 으로 하고, 여러 개의 파싱 결과가 나오고 그 중에 맞는 것이 있을 경우는 O, 그외의 경우는 X 라고 할 때, 파싱 정확률은 전체 문장에서 OO 과 O 의 개수를 합한 수의 비율로 하였다. 평가 결과는 SERISet96 전체에 대하여 약 80%의 파싱 정확률을 나타내었으며, 5 어절 이하에 대해서는 90% 이상의 파싱 정확률을 보였다. 전체의 20%인 잘못 분석된 결과에 대한 오류의 원인을 살펴 보면, 형태소 해석기 오류(수사처리, 어미 정보 추출 오류 등), 사전의 오류(형태소 정보 미등록, 명사의미정보 누락 오류), 그리고 구문 분석 오류(전처리기에서 복합어구 처리 오류 등)가 있었다. 또한, 여러 개의 구문 분석 결과에 대한 트리 선택 규칙을 보장하는 것이 필요하다.

두번째 평가는 SERISet96 에서 추출한 술어가 1 개부터 4 개까지 나타나는 문장들에 대하여 앞의 III 장에서 언급한 의존 관계 검사에서 지배가능경로만을 이용한 경우(1)와 지배가능경로와 술어 하위범주화 정보를 함께 이용한 경우(2)의 파싱에서 생성되는 의존

관계의 수를 비교한 것으로 결과는 다음 표와 같다.

문장에서 술어의 수 (어절 수)	생성된 의존 관계의 수(평균)	
	(1) 지배가능경로	(2) 지배가능경로 + 술어 하위범주화 정보
1(4)	3	3
2(6)	8	5
3(7)	10	6
4(10)	20	9

위 결과표에서 나타난 바와 같이 문장에서 술어가 1 개 포함된 문장보다는 술어가 많이 나타나는 문장일수록 (1)의 파싱보다는 본 논문에서 제안한 (2)의 파싱이 더욱 효과적임을 알 수 있다.

## V. 결론

본 논문의 의존 파싱 방법은 후보 의존소가 지배가능경로 상의 술어 지배소와 의존 관계 검사를 수행할 때, 그 술어의 하위범주화 정보 적용 범위와 술어 하위범주화 정보 제약을 적용하여 의존 관계의 생성을 최소한으로 줄이는 효율적인 파싱 방법이다.

본 논문에서 구현한 한국어 구문 분석 시스템은 의존 파싱 과정에서 후보 의존 관계의 과생성으로 인한 비효율성을 미리 차단함으로써 의존 규칙과 지배가능경로 만에 의한 의존 파싱에 비하여 보다 정확하고 향상된 파싱 결과를 얻을 수 있었다. 그러나, 앞으로 현재의 형태소 해석기와 사전을 보완하고 구문 분석에서 의미/통계 정보를 적절히 활용하는 방법으로 시스템을 더욱 개선할 계획이다.

## 참고문헌

- [Hellwig 86] P. Hellwig, "Dependency Unification Grammar," Colling 86, pp.195-198, 1986
- [Covington 88] T. A. Covington, "Parsing Variable Word Order Languages with Unification-Based Dependency Grammar," ACMC Research Report 01-0022, Univ. Of Georgia, 1988
- [강은 95] 강은국, 조선어 문형연구, 도서출판 박이정, 1995
- [김창 95] 김창제, 정천영, 김영훈, 서영훈, "부분적인 어절 결합을 이용한 효율적인 한국어 구문 분석기," 한국정보과학회 가을 학술발표논문집, pp.597-600, 1995
- [김창 93] 김창현, 김재훈, 서정연, "지배가능경로를 이용한 오른쪽 우선 구문 분석," 제 5 회 한글 및 한국어 정보처리 학술발표 논문집, pp.35-44, 1993
- [은중 95] 은중진, "효율적인 구문 분석을 위한 전처리 구현과 복합명사의 구조 분석," 한국과학기술원 석사학위논문, 1995

[손광 95] 손광주, “신경망과 제약만족 알고리즘을 이용한 한국어 구문분석”, 포항공대 석사학위논문, 1994.12

[손광 94] 손광주, 홍영국, 이종혁, 이근배, “어절간 의존관계 해석을 위한 한국어 파서”, HCI'94, pp.351-141, 1994

[장석 95] 장석진, 정보기반 한국어 문법, 한신문화사, 1995

[한용 95] 한용기, 황이규, 이용석, “문형정보를 이용한 한국어 구문 분석”, 제 7 회 한글 및 한국어 정보처리 학술발표 논문집, pp.23-29, 1995

[홍영 93] 홍영국, 이종혁, 이근배, “의존문법에 기반을 둔 한국어 구문 분석기”, 한국정보과학회 봄 학술발표논문집, pp.781-784, 1993