

자연어를 이용한 사용자 접속에 관한 연구

이 동애, 장 덕성
 계명대학교 공과대학 전자계산학과

요 약

MS-DOS 명령을 대신하는 자연어 인터페이스를 연구하였다. 자연어로 입력되는 한국어 문장을 형태소분석, 구문분석, 의미분석, 개념분석을 통해 대응되는 일련의 MS-DOS 명령을 생성한다. 형태소 분석에서는 Tabular Parsing법을 사용하였고, 구문분석에서는 문법적인 수식-피수식 관계를 확대하여 의미상의 수식-피수식 관계를 설정하고 이에 따라 문장을 몇개의 단위로 나눈다. 의미분석에서는 동사와 이들 단위들간의 관계와 단위를 구성하는 어절들간의 관계를 격관계로 설정하여, 개념망(semantic network)으로 문장의 의미를 표현한다. 이 개념망으로부터 MS-DOS 명령을 생성한다.

1 서론

인공지능의 일부인 자연어 처리는 인간의 사고, 학습 그리고 의사전달 과정의 기본 수단이 되는 자연어를 컴퓨터에 구현하는 것으로, 기계번역과 자연어 인터페이스를 중심으로 연구가 되어 왔다. 기존의 사람-기계 인터페이스는 형식언어나 메뉴 방식을 사용해 왔는데, 이러한 인터페이스는 사람의 입장보다는 기계의 입장에서 보다 처리가 용이한 방식이다. 그러나 자연어를 이용한 사람-기계 인터페이스는 컴퓨터를 이용하는 사용자에게 보다 자연스러우며, 편리하고 능동적인 입장을 제공한다. 이러한 장점에도 불구하고 자연어 인터페이스는 형식언어나 메뉴 방식의 인터페이스보다 실용적이지 못하다. 이는 자연어가 갖는 특정한 언어 현상들과 애매성, 그리고 절대적인 grammar rule이 존재하지 않아서 형식언어에 비해 훨씬 처리가 어렵기 때문이다[7]. 자연어 인터페이스는 전문가 시스템, CAI(computer aided instruction), 데이터 베이스 등의 인터페이스로 이용되는데 이러한 시스템들은 아직까지는 초기 단계이다.

본 연구에서는 한국어로 MS-DOS 명령을 대신할 수 있는 자연어 인터페이스 시스템을 구축하고자 한다. 이를 위해서 자연어로서의 '한국어가 갖는 특징 [6,7,9,12]을 파악하고, 문제 영역의 한국어의 범위에 따른 사전을 구성[8,10]한다. Tabular Parsing법을 이

용해 형태소 분석을 하고, 격문법을 이용해서 구문 분석과 의미 분석을 한다. 이의 결과로 생성되는 의미 구조는 개념망을 이용해 표현하고, 개념분석을 통해 개념망으로부터 MS-DOS 명령을 생성한다.

2 한국어 범위와 사건의 설정

연구의 입력 대상이 되는 문장은 성립상으로는 명명문과 의문문이다. 그외 서술문, 김만문, 칭유문은 문제의 성격상 부적합하므로 사용하지 않는다. 구조상으로는 단문, 중문, 복문까지를 허용키로 한다. 컴퓨터를 사용하고자 하는 사용자의 모든 요구를 파악하기 위하여 설문조사하고, 그 대부분을 입력대상으로 하였다.

MS-DOS 명령어가 제공하는 기능에 관련되는 단어를 중심으로 사전을 구축하되 차후에 추가·변경·삭제가 용이하도록 구축한다. 아래에는 MS-DOS 명령어가 제공하는 기능을 동사 중심으로 분류하고 이들 기능에 관련된 MS-DOS 명령을 나열한다.

- (1) DELETE
 - 지우다, 없애다, 삭제하다, etc.
 - attrib, cls, del, rmdir
- (2) CHANGE

- 바꾸다, 교환하다, 변경하다, etc.
- attrib, cd, ctty, date, exe2bin, label, prompt, replace, subst, ver

(3) OUT

- 인쇄하다, 출력하다, 인자하다, etc.
- append, attrib, break, cd, date, dir, join, more, path, print, path, print, set, time, tree, type, ver, verify, vol

(4) MOVE

- 옮기다, 백업하다, 저장하다, etc.
- backup, copy, diskcopy, store, sys, xcopy

(5) SET

- 설정하다, 할당하다, 지정하다, etc.
- append, assign, attrib, break, date, label, path, prompt, set, subst, time, verify

(6) CHECK

- 검사하다, 확인하다, 조사하다, etc.
- chkdisk, diskcomp

(7) OTHERS

- 실행하다, 초기화하다, 복구하다, etc.
- command, exit, find, format, more, recover, share, sort

사전에 저장되어야 하는 지식은 문법이론이나 채택될 모델에 따라 다르다. 본 연구에서는 기계번역시에 필요한 변환이나 생성 과정이 필요 없으므로 사전 엔트리에 저장되어야 할 지식은 entry word, pos vector, semantic이다. entry word는 형태소 단위로 저장되는 표제어에 해당한다. pos vector는 entry word의 품사와, 품사가 용언인 경우에 해당 불규칙 활용에 대한 정보를 갖고 semantic은 각 entry word의 의미를 저장한다. pos vector 각 위치의 품사와 값의 범위는 아래와 같다.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

품사 : 동사 형용사 명사 대명사 수사 부사 관형사
 값 : 0~8 0~8 0~1 0~1 0~1 0~1 0~1

entry word의 해당 품사 위치에는 1이상의 값, 나머지는 0으로 표시한다. 동사나 형용사인 경우에는 불규칙 활용에 대한 정보를 위해 그 값과 같이 구분한다.

값	1	2	3	4	5	6	7	8
불규칙 형태	#	s	u	e	#	u	e	규칙

다음 표 2.1은 사전 엔트리의 예를 보인 것이다.

[표 2.1] 사전 엔트리의 예

품사	entry word	pos vector	semantic
규칙 동사	변경하	8000000	change
으 불규칙 형용사	쓰	0300000	write
명사	디렉토리	0010000	directory
대명사	그	0001000	that
수사	하나	0000100	one
부사	다시	0000010	again
관형사	모든	0000001	all

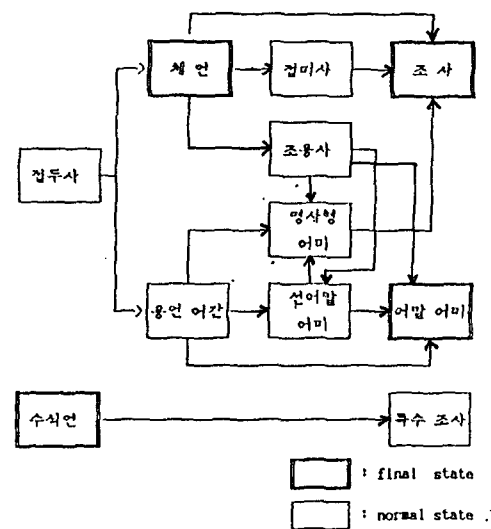
3 형태소 분석

자연어 입력 문장을 대상으로 형태소 규칙과 사전을 이용하여 어절을 최소의 의미 단위인 형태소들로 분리하고 그 정보를 추출하는 과정을 말한다. 즉 어절을 구성하는 형태소들이 어떻게 결합되어 있는가를 분석해서 가능한 결합 형태를 모두 도출하는 과정이다.

형태소 분석시 고려되어야 할 사항들은 아래와 같다.

(1) 어절의 구성

일정한 소리에 일정한 뜻이 결합되어 있는 말의 가장 작은 단위 즉, 최소 의미 단위(minimum meaningful unit)를 형태소라 하며 어절은 형태소들이 결합되어 있는 형태이다. 형태소 분석에서는 어절을 형태소 단위로 분리한다. 형태소들은 그림 3.1과 같이 결합되어 어절을 구성한다.



[그림 3.1] 어절의 형태

조용사는 명사, 의성어, 의태어에 붙어서 용언화시키는 역할을 한다. 가령 '하다'라는 접미사가 순수 우리말이나 한자어에 붙어서 용언이 되는 경우는 전체 용언의 1/3이 넘는다[3]. 그래서 '하다'를 조용사로 분류하고 독립적인 형태소로 인식하면 사전의 크기를 상당량 줄일 수 있게 된다. 또한 '이다'라는 서술격 조사도 조용사에 포함시킬 수 있는데 이는 체언에 붙어 조사로서는 유일하게 용언적 기능을 하기 때문이다.

(2) 용언의 불규칙 현상

체언은 어형이 고정되어 있어서 문법적 관계를 나타낼 수 없기 때문에 조사가 붙어 문법적 관계를 나타낸다. 반면 용언은 어미가 여러가지 형태로 변화하여 문법적 관계를 나타내게 되는데 이러한 어미 변화를 활용이라 한다. 용언의 활용에서는 용언의 어형이 변화를 받게 된다. 특히 어간과 어미의 형태 변화가 일정한 규칙 용언과는 달리 어간과 어미의 형태가 불규칙적으로 변하는 불규칙 용언인 경우에 이를 처리할 수 있어야 한다.

(3) 매개모음 '이' 와 '으' 처리

다음과 자음의 충돌을 피하기 위해서 매개모음이 첨가되어 발음을 매끄럽게 하는데 이에는 '이'와 '으'가 있다. '이'나 '으'에 접속가능한 형태소는 아래와 같다.

- (으)로, (으)며, (으)면, (으)로서
- (이)든, (이)든지, (이)나마

매개모음을 하나의 독립적인 형태소로 취급하거나, 매개모음과 매개모음에 접속가능한 형태소가 접속된 것을 하나의 형태소로 취급하여 처리할 수 있다.

(4) 축약(contraction) 현상

입력되는 어절들은 형태소 분석 단계에서 사전을 참고하게 되는데, 사전에는 각 entry word가 축약된 형태가 아닌 원형으로 저장되어 있으므로 형태소 분석의 전단계로 축약된 형태를 원형으로 변환해야 한다.

축약은 음운 현상중의 하나로 연속된 두 모음을 특별히 분리해서 발음할 필요가 없을 때, 말의 속도를 빨리 하고 노력을 경감하기 위해서 두소리의 사잇소리로 단음화하거나, 두 음절의 모음이

줄여져 한음절의 복모음으로 되는 현상을 말한다.

◦ 간음으로 축약되는 예 :

사이 -> 새, 오아 -> 와

◦ 복모음으로 축약되는 예 :

가지어 -> 가지, 하이라 -> 해라

(5) 합성어 처리

한국어에서는 합성어의 길이와 수에 제한이 없으므로 모든 합성어를 사전에는 등록할 수 없고 이러한 합성어의 인식이 필요하다.

(6) 미등록어 처리

사전에 등록되지 않은 단어에 대해서는 다른 형태소들의 정보로부터 미등록어의 짐사추정이 가능하다.

한국어의 형태소 분석에 많이 사용되는 방법으로는 최장일치법(longest match strategy), 최단일치법(shortest match strategy), Head-Tail 구분법, Tabular Parsing법이 있다. 가장 간단한 방법은 최장일치법과 최단일치법이다. 최장일치법이란 어절내에서 본절이 가능한 여러 형태소 가운데 가장 긴 형태소를 선택하는 것이며, 최단일치법이란 가장 짧은 형태소를 선택하는 방법이다. Head-Tail 구분법은 어절을 변형되지 않는 부분(Head)과 변형되는 부분(Tail)으로 구분하여 테이블을 구성하고, 이 테이블에서 접속가능한 모든 형태소들의 조합을 결과로 생성하는 방법이다. Tabular Parsing법은 입력어절로부터 추출될 수 있는 모든 형태소 정보를 테이블의 각 엔트리에 기록하는 bottom-up 방식이다.

일례로 '감기는'에 대한 각 형태소 분석법의 결과를 비교해본다.

[표 3.1] 형태소 분석 결과의 비교

최장일치법	최단일치법	Head-Tail구분법	Tabular Parsing법
감기(n)+는(z)	감(v)+기는(e)	감(v)+기(n)+는(z) 감기(n)+는(z)	감(v)+기는(e) 감(v)+기(n)+는(z) 감기(v)+는(e) 감기(n)+는(z)

e : 어미, n : 명사
 ■ : 필수형어미
 v : 동사, z : 조사

표 3.1에서 보는 것과 같이 최단일치법, 최장일치법, Head-Tail 구분법은 접속가능한 모든 형태소들의 조합을 생성해내지 못하는 반면 Tabular Parsing법은 접속가능한 형태소들의 모든 조합을 생성한다.

이러한 장점 때문에 본 논문에서는 Tabular Parsing 방법을 사용한다.

4 구문 분석

구문 분석 단계에서는 전단계인 형태소 분석의 결과를 이용하여 문장의 구조를 파악하고 문장을 구성하고 있는 각 어절들의 문장 내에서의 역할을 인식한다.

4.1 수식어와 피수식어의 패턴

알타이 어계의 공통된 특징으로서, 수식어는 피수식어 앞에서 그 기능을 한다. 수식어에는 용언을 수식하는 부사, 체언을 수식하는 관형사가 있으나 본 연구물 위해서는 수식어의 구분을 좀더 세분화할 필요가 있다. 수식어와 피수식어의 관계를 단순히 표면적으로 드러나는 문법적인 관계가 아닌 의미상의 수식과 피수식 관계로 인식하고자 하는데, 이러한 접근법은 의미를 강조하는 격문법의 취지[1,2,5]와도 일치한다.

의미상으로 인식되는 수식과 피수식의 관계는 아래와 같이 아홉가지 패턴으로 구분될 수 있다.

- 패턴 1) 관형사 체언/용언의 명사형
예 : 꽃 잎
- 패턴 2) 지시형/부사 체언/용언의 명사형
예 : 그 디렉토리
- 패턴 3) 용언+관형형 어미 체언/용언의 명사형
예 : 존재하는 파일
- 패턴 4) (체언/용언의 명사형)+관형격 조사 체언/용언의 명사형
예 : 파일의 속성
- 패턴 5) 체언 체언
예 : MS-DOS 버전
- 패턴 6) 부사 용언
예 : 각각 복사하고
- 패턴 7) 본동사 조동사
예 : 출력하지 말고 바꿔야 하는
- 패턴 8) (체언/용언의 명사형)+부사격조사 용언
예 : 디렉토리에 있는 이름으로 바꾸어
- 패턴 9) (체언/용언의 명사형)+주격조사 용언
예 : 파일이 있으면

패턴 5와 패턴 7은 수식어와 피수식어 사이에 어떤 어절도 끼어들 수 없다. 그래서 현재 처리되는 수

식어는 바로 다음 어절인 피수식어를 수식한다. 그러나 패턴 5와 패턴 7을 제외한 나머지 일곱가지 패턴은 수식어와 피수식어 사이에 다른 어절이 끼어들 수 있으므로 현재 처리되고 있는 수식어 어절의 피수식어는 우측으로 가장 가까운 해당 품사의 어절을 찾는 것을 원칙으로 한다. 아래에는 그 예를 보여주고 있는데 숫자는 위에서 구분된 해당 패턴의 번호이다.

예1: A 디렉토리에 파일이 있으면 파일들을 모두 지워라.
패턴8

예2: A 디렉토리의 모든 파일을 B 디렉토리에 복사하라.
패턴4

예2에서는 '디렉토리의'가 패턴 4에 해당하는 수식어이므로 피수식어의 자격이 있는 어절은 체언이거나 용언의 명사형인 '파일들', 'B', '디렉토리에'이다. 이 중에서 '디렉토리의'에 우측으로 가장 가까운 어절은 '파일들'이므로 '파일들'이 '디렉토리의'에 대한 피수식어가 된다.

4.2 수식-피수식 관계로 구분되는 개념 단위

문장 전체의 서술어는 수식-피수식 관계설정에서 제외한다. 이유는 문장의 골격을 이루는 서술어와 피수식어간의 관계를 파악하기 위해서이다. 즉 문장 전체의 서술어를 제외한 문장의 나머지를 몇개의 단위로 자르고, 이들 단위와 서술어와의 관계를 파악하기 위해서이다.

예1: A 디렉토리에 있는 모든 파일들 B 디렉토리에 옮겨라.
=> 단위 1 : A 디렉토리에 있는 모든 파일들
단위 2 : B 디렉토리에
단위 3 : 옮겨라

예2: 현재 디렉토리의 이름을 프린터로 출력하라.
=> 단위 1 : 현재 디렉토리의 이름
단위 2 : 프린터로
단위 3 : 출력하라

예3: *.bak 파일이 있으면 그 파일들을 지워라.
=> 단위 1 : *.bak 파일이 있으면
단위 2 : 그 파일들을
단위 3 : 지워라

각 숫자는 4.1절에서 구분한 패턴 번호이며 문장 전체의 서술어 앞에 표시된 '|'는 수식-피수식 설정에서 동사는 제외됨을 표시한다.

수식-피수식 관계로 어절들을 연결하게 되면 문장

이 몇개의 단위로 나뉘어지는 것을 알 수 있다. 위의 예에서는 모두 3개의 단위로 문장이 구분되었다.

5 의미 해석과 의미구조의 표현

격론법은 문장론적 해석과 어의론적 해석을 합친 방식이다. 한국어에서의 격(case)은 조사의 역할인데, 격론법에서의 격은 문장 표면적인 역할로서의 격, 즉 조사로 결정되는 표면격(surface case)이 아니라 의미상의 역할로서의 격, 즉 심층격(deep case)인 것이다.

수식어는 문장에서 삭제되더라도 원래의 문장 구조를 변형시키지 않는다. 그래서 수식어를 뺀 피수식어와 동사로 문장의 골격을 구축하고, 각 피수식어는 동사에 대한 어떤 관계를 유지하게 된다. 어떤 관계란 격론법에서의 격(case)으로 대응될 수 있다. 일단 삭제된 수식어는 나중에 피수식어와 관계를 맺도록 해야 하는데 이것 역시 격으로 대응될 수 있다. 이렇게 함으로써 문장내의 모든 어절에 대한 관계가 결정되며 이것이 문장의 의미로 받아 들여진다. 격은 문제의 형태에 따라 차이가 있고, 일반화된 격설정은 마련되어 있지 않으며, 본 연구에서는 필요한 격을 아래와 같이 설정하였다.

(1) 심층격 관계

동사와 피수식어간의, 피수식어와 수식어간의 관계를 설정한다.

- changed(변성격)
- dative(여격)
- from(시발격)
- location(장소격)
- object(목적격)
- partner(공동격)
- role(자격격)
- subject(주격)
- to(종착격)
- count(계수격)
- name(명칭격)
- value(속성격)

(2) 논리적 연결

문장간의 논리적 관계를 설정한다.

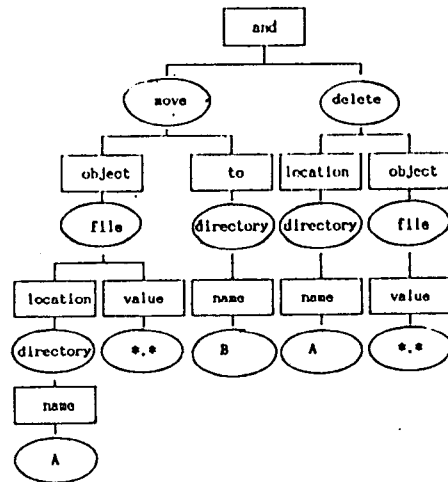
- and
- if
- not
- or

의미구조의 내부표현은 일반적으로 격론법을 기반으로 하는 개념망을 널리 사용한다. 연합회모(associative memory)라고도 불리는 개념망은 지식의 구조적 표현 방법중 가장 기본적이고 대표적인 것이다. 개념망은 개념기호를 표시하는 노드와 개념기호의 쌍으로 표현하는 edge로 구성되는 그래프이다. 개념 기호는 말의 의미를 나타내는 의도적 기호(intensional symbol)이다[4]. 개념기호는 개념어(conceptual primitive)와 개념관계(conceptual relation)로 나뉜다. 개념어는 다시 어떤 대상을 나타내는 대상 개념어와 그 대상을 묘사하는 서술개념어로 나뉜다. 예를 들면 DIRECTORY, FILE 등은 대상 개념어이고, COPY, DELETE PRINT 등은 서술개념어이다.

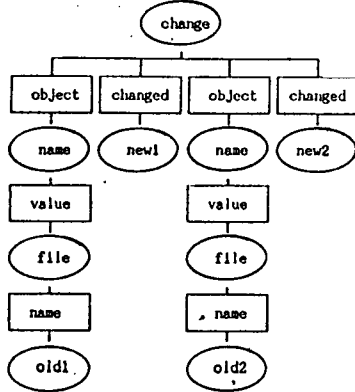
개념관계는 5.1절에서 설정한 격으로 대응된다. 그림 5.1은 각 문장에 대해서 4.1절에서 정의된 수식-피수식 패턴에 따라 구분된 단위로 개념망을 보인 것이다. 단위들은 밑줄로써 구분하기로 한다.

문장중에 지시 형용사나 대명사가 존재하면 이에 해당하는 어절을 발견할 때 전방향으로 찾는 것을 원칙으로 한다[11]. 그림 5.1의 (a)에서는 이러한 예를 보이고 있다. 그리고 생략된 용언은 후방향으로 찾는 것을 원칙으로 하는데[11] 이의 예는 그림 5.1의 (b)에서 보이고 있다.

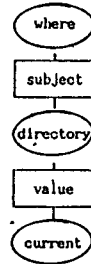
(a) B 디렉토리에는 모든 파일이 B 디렉토리에 속하고, A 디렉토리에서는 그 파일들을 지닌다.



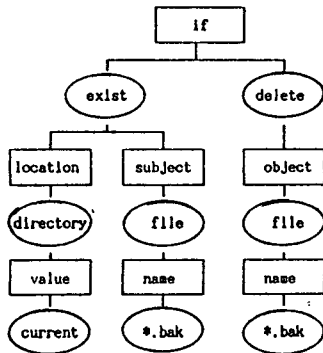
(b) old1 파일의 이름은 new1으로, old2 파일의 이름은 new2로 바뀐다.



(c) 현재의 디렉토리는 어디인가?



(d) 현재 디렉토리에 *.bak 파일이 있으면 이 파일들을 지워라.



[그림 5.1] 개념망의 예

6 개념 분석

의미 분석의 결과로 생성된 개념망으로부터 MS-DOS 명령어를 도출해내기 위해서는 MS-DOS 명령에 대한 지식을 기반으로, 개념망의 개념을 MS-DOS 명령으로 추출·대응시켜야 한다.

MS-DOS 명령은 아래와 같은 형식으로 구성된다.

명령 [옵션...]

[]은 해당 항목이 옵션임을 의미한다. 위의 옵션에 해당하는 것은 아래와 같은 것들이 있다.

- 드라이브: - 디스크의 드라이브명을 의미
- 파일명 - 파일의 이름을 의미
- 경로명 - [<디렉토리> \ [<디렉토리> ...] \ [<파일명>]
- 스위치 - '/'로 시작하여 명령을 제어
- 인수 - MS-DOS 명령에 더 많은 정보를 제공

MS-DOS 명령은 옵션의 수와 종류에 따라 아래와 같이 구분된다.

- 명령
 - cls, exit, more, ver, share
 - 명령 [ON/OFF]
 - break, verify
 - 명령 [<드라이브>]
 - format, recover, tree, vol
 - 명령 [<드라이브>] <경로명>
 - attrib, del, mkdir
 - 명령 [<드라이브>] [<경로명>]
 - dir, chkdsk, print, sort, subst
 - 명령 [<드라이브>] [<드라이브>]
 - diskcomp, diskcopy
 - 명령 [<드라이브>] <경로명> [<드라이브>]
 - [<경로명>]
 - copy, exe2bin, replace
- <>는 그 항목에 대한 텍스트를 의미한다.

그외 각기 다른 수와 형태의 옵션을 가지는 명령에 대한 지식은 물론 명령의 입·출력의 방향을 바꾸는 '<', '>', ">>"와 파이프 처리를 위한 '|'도 포함한다.

의미분석의 결과로 생성된 개념망으로부터 MS-DOS 명령어를 도출해내기 위해서 먼저 명령을 찾고, 다음에 명령에 해당하는 옵션을 찾는다. 해당 명령은 동사와 object에 해당하는 개념어를 통해 추출이 가능하다. 2.2절에서 구분된 동사와 object에 해당하는 개념어를 서로 관련지으면 다음과 같은 MS-DOS 명령을 추출해낼 수 있다. 그중 DELETE와 CHANGE에 대해서만 아래에 요약한다. DELETE의 경우 object에 해당되는 개념어는 화면, 파일, 파일의 속성, 디렉토리가 있고 각각은 MS-DOS 명령 cls, del, attrib, rmdir에 대응된다.

(1) DELETE(지우다, 없애다, 삭제하다, etc.)

- 화면

... cls

- 화일 ... del
- 화일의 속성 ... attrib
- 디렉토리 ... rmdir

(2) CHANGE(바꾸다, 교환하다, 변경하다, etc.)

- 화일의 이름 ... ren
- 화일의 속성 ... attrib
- 작업 디렉토리 ... cd
- 명령어의 입출력을 위한 디바이스 ... cty
- 날짜 ... date
- 실행 화일 ... exe2bin
- 디스크의 레이블 ... label
- 프롬프트 ... prompt
- 화일의 버전 ... replace
- 경로명 ... subst

MS-DOS 명령이 결정되면 해당 명령이 필요로 하는 옵션에 대한 지식을 이용하여 옵션들을 개념망의 정보를 통해 완성시킨다.

그림 5.1로 표현된 경우에 MS-DOS 명령은 아래와 같다.

- (a) COPY A*. * B
DEL A*. *
- (b) REN OLD1 NEW1
REN OLD2 NEW2
- (c) CD
- (d) IF EXIST *.BAK DEL *.BAK

본 논문은 C로 구현했으며, 개념 분석의 결과로 생성되는 MS-DOS 명령은 프로그램내에서 C 라이브러리 system을 이용하여 바로 수행이 된다.

7 결론

형태소 분석은, 컴파일러의 front-end와 같이 결과는 고정적이며, 거의가 비슷한 형태를 띤다. 구문 분석과 의미 분석은 컴파일러의 back-end처럼 문제의 성격에 따라 달라지며, 결과적으로 의미구조 역시 달라진다. 그래서 형태소 해석기는 공개하고 구문분석과 의미분석은 문제에 따라 설계하도록 하여 자연어 처리 연구에 박차를 가했으면 한다.

또한 자연언어의 인터페이스로 MS-DOS 명령을 대신하기에는 더 많은 연구가 필요하다. 본 논문에서는 MS-DOS가 제공하는 명령의 기능을 미리 안다는 가정하

에서 입력되는 문장을 처리하는, 제한된 범위의 한국어를 처리한다. 그러므로 전혀 MS-DOS 명령을 모르는 사용자의 요구를 처리하기 위해서는 처리대상이 되는 한국어의 범위는 상당부분 확대되어야 하며, 이에 따라 사전의 설정, 형태소 분석, 구문 분석, 의미 분석, 개념분석 역시 보충 확대되어야 한다. 이러한 면에서 본다면 사용자에게 최대한의 용용성을 제공하는 자연어 인터페이스는 사실상 자연언어 전체를 처리할 수 있는 시스템을 말하며, 이를 위해서는 아직도 더 많은 연구가 있어야 할 것으로 본다.

참고문헌

1. Bar, A. & Feigenbaum, E. A., "The Handbook of Artificial Intelligence", Vol. 1, William Kaufmann, Inc., 1981
2. Harris, M.D. "Introduction to Natural Language Processing"; Reston Publishing Company Inc., 1985
3. 김성용, 최기선, 김길창, "Tabular Parsing 방법과 접속 정보를 이용한 한국어 형태소 분석기", 한국어를 중심으로 한 기계번역, 한국과학기술원, 1990
4. 김재훈, 김철호, 최기선, 김길창, "PIVOT 방식의 기계번역에서 한국어 격구조 설정과 중간언어로부터 조사 생성", 정보과학회지 Vol.14, No.2, 1987
5. 김재희, "인공지능의 기법과 응용", 교학사, 1989
6. 박병수, "기계번역에서 본 한국어의 특징", 정보과학회지, 제7권 제6호, 1989
7. 박세영, "기계번역을 위한 자연언어의 syntax 처리", 정보과학회지, 제7권 제6호, 1989
8. 이상조, "기계번역 시스템을 위한 사전 구성", 정보과학회지, 제7권 제6호, 1989
9. 조규빈, "하이라이프 국문법", 지학사, 1982
10. 최재혁, "한영 기계 번역을 위한 조사 사전 구성에 관한 연구", 한글 및 한국어정보처리 학술발표논문집, 1989
11. 한광복, "한국어 문장으로부터 개념단위의 추출과 지식베이스의 구축", 한글 및 한국어정보처리 학술발표논문집, 1989
12. 한국과학기술원, "한국어를 중심으로 한 기계번역", 1990
13. 한국전자통신연구소, "한국어 파서 및 사전 시스템 개발", 과학기술처, 1989
14. 한양대학교, "의미 네트워크를 이용한 언어이해과정의 구현에 관한 연구", 한국전자통신연구소, 1989
15. "한글/한자 MS-DOS 사용자 참고 설명서", Qnix,