

DCS의 정보확장

이 창 인

베제대학교 영어영문학과

Adding New Information in DCS

Chang-in Lee

Dept. of English Language and Literature

요 약

본 논문은 DCS(Dynamic Comprehension System) 정보확장 과정을 어휘 정보의 첨가를 통해 묘사하고자한다. 즉, 기존의 사전정보의 보완작용없이, 시스템을 확장하는 과정이 이 논문에서 보여진다. 새로운 언어정보에 유연하게 대처하기 위해 화자와 청자간의 새로운 지식의 학습과정이 나무구조 형식의 보조메뉴를 통해 상호 교환방식으로 나타내진다. 본 논문은 새로운 지식의 인지과정 중 기존의 정보망(network)에 각 단위망(nection)이 첨가될 때의 과정을 구현시키고자 시도된 것이다.

1. 서론

언어의 주요기능 중 하나는 화자와 청자의 기존 정보 체제에 변화를 주어 서로간의 대화를 가능하게 하는 것이다. 이때, 인간의 인지과정이 주로 현존하는 정보망(nection)의 연결 상태를 저지, 활성화하거나 혹은 현존의 정보망에 새로운 정보를 확장시키는 작용이라면, 인간이 무엇인가를 학습(learning) 한다는 것은 주로 현 정보체제에 새로운 정보를 첨가시키는 과정이 주를 이룰 것이다. 따라서, 진정한 의미의 인지적 언어 모델이라면, 이해(Comprehension), 생성(Production) 과정이외에도 새로운 언어정보에 유연하게 대처할 수 있는 기능들을 갖추어야 할 것이다. 이러한 관점에서 비추어 볼 때 이해와 생성을 근간으로 설계된 DCS(Dynamic Comprehension System)의 정보확장이 요구되어지며, 본 논문은 DCS의 정보확장 과정과 그 과정의 정보 구조를 다루도록 한다. DCS는 언어지식과 컴퓨터 처리과정을 분리한 이해위주의 시스템으로써, 영어와 한국어의 언어지식을 토대로 시험되었다.

본 논문은 통사정보 보다는 수량적으로 대다수를 차지하는 어휘정보의 확장 과정을 다룬다. 시스템 보완은 요사이 컴퓨터 사용시 보편화 되어있는 메뉴 형식을 이용하여 사용자와의 대화를 통한 상호보완 작용에 의해 이루어진다. 논문의 전반부는 나무구조 형식의 보조 메뉴들(Help, Add New Information, Display New Information and Exit)로 이루어진 주요메뉴(Main Menu)에 관한 설명을 통해 추가입력 정보에 관한 처리과정을 다룬다.

2. 어휘 정보확장

DCS는 사전정보에 없는 새로운 단어가 입력되면 그 단어가 사용자의 타이핑 실수인지 새 정보인지의 확인작업으로 다음과 같은 질문을 하게된다.

The item ____ is not in the dictionary.
Is this a typo? (y/n)

만약 사용자가 'y'라고 답하면 시스템은 다시 다음과 같은 명령어를 보인다.

Enter the correct item:

만약 사용자가 'n'라고 답하면, DCS는 새로운 정보가 입력된 것으로 간주한다. 새 어휘를 기존의 사전정보에 추가시키기 위해선 다음의 정보들이 필요한데, 이 정보는 요사이 컴퓨터 사용시 널리 활용되는 메뉴형태를 통해 사용자와의 상호 대화를 통해 입력된다. 3장에서는 다음의 정보들이 사용자와의 대화를 통해 어떻게 저장되어 사용되어지는가를 보여준다.

- 1) 통사정보(품사정보)
- 2) 의미정보(명사의 의미속성과 동사의 의미코드)

3. 메뉴

도표-1은 DCS의 주 메뉴를 보여준다.

**** MAIN MENU ****

1. HELP
2. ADD NEW INFORMATION
3. DISPLAY NEW INFORMATION
4. EXIT

도표-1

본 3장에서는 나무구조 형태로 이루어진 각각의 보조메뉴에 대한 설명이 다루어진다.

3.1. HELP

보조메뉴 HELP는 DCS 시스템에 익숙하지 못한 사용자에게 새로운 정보 추가시 필요한 안내자 역할을 담당한다. 사용자가 HELP 메뉴를 부르게 되면 다음 도표-2의 내용이 화면위에 나타난다.

-In order to add a new item, you need to specify the number of the lexical categories and to identify the names of the lexical categories from the Menu for lexical

categories.

-You also need to choose the semantic properties that can be associated with the entity, or the semotactic code that can be applicable to the process from the Menu.

-If an appropriate semantic property or a semotactic code is not included in the Menu, you can add one by entering the command '0.' Command '0' calls up a prompt that asks you to identify the name of a new semantic property or a new semotactic code.

도표-2

3.2. ADD NEW INFORMATION

본 메뉴는 새로운 어휘 추가시 사용자가 시스템과의 직접 대화가 가능하도록 도와주는 역할을 하는 곳이다. 사용자가 위의 도표-2 주 메뉴로부터 2번 보조메뉴 'ADD NEW INFORMATION'을 택하게 되면 DCS에서 영어와 한국어를 위해 고려되어진 품사정보가 나타난다. 또한 사용자의 편의를 위하여 추가되고있는 새로운 어휘가 무엇인지도 보여진다. (도표-3)

**** Lexical Categories Menu ****

<TYPE1 Lexical Category>

1. NOUN

<TYPE2 Lexical Categories>

2. Vt 3. Vi 4. COPVERB

5. BIVERB

<TYPE3 Lexical Categories>

6. InDef 7. CONJ 8. Pl

9. IPVt

10. ADV 11. RP 12. Def

13. nPPT

14. nPPL 15. nPPDAT 16. nPPS

17. nPPIN

18. nPPQ 19. nPPDAT 20. nPPD

21. nPPAB

22. nPPINS 23. Pres 24. Past

25. Fut

26. Papl 27. Prog 28. Aux-Be

29. Aux-Have

30. Adj-C 31. Adj-D 32. Adj-CI

33. Adj-N

THE NEW ITEM =

도표-3

새로운 어휘의 다품사 (class cleavage)현상을 고려하여 DCS는 다음의 질문을 하게된다.

Lexical Category Numbers for This Item >>

여기서 사용자가 답하는 숫자는 추가 중인 새 어휘의 품사가 한 개 뿐인지 혹은 그 이상인지의 다품사 현상에 대한 정보 뿐 아니라 품사메뉴로 부터 나와 다음의 프롬프트를 읽을 수 있는 신호이기도 하다. 만약 사용자가 '1'을 답하면 새 어휘는 한 개의 품사만을 지니며 다음의 구체적인 품사정보에 대한 프롬프트를 한번만 나타내라는 신호이며, 만약 사용자가 '2'를 답하면 새 어휘는 두 개의 품사를 지니며 품사정보에 관한 프롬프트를 두 번 보이라는 신호인 것이다. 품사정보에 대한 갯수가 입력된 후에 사용자는 새 어휘에 해당되는 구체적인 품사가 무엇인지에 대한 답을 다음의 프롬프트에서 제공하게된다.

Enter the number of the lexical category >>

여기서 사용자가 명사인 '1'번이나 동사인 '2' '3' '4' '5'번 중에서 택일을 하게되면 다음의 선택사항이 주어진다.

- A. To Choose NOUN
- B. To Choose VERB
- or
- C. To Choose NOUN and VERB

3.2.1. To Choose NOUN

만약 사용자가 '나는 사과를 먹는다.'라는 입력 문장에서 '사과'를 새로운 어휘로 채택하여 기존의 사전 정보에 '사과'를 추가시키려면, 다음의 프롬프트에서 '1'(NOUN)을 답하게된다.

Enter the number of the lexical category >>

그러면, 추가될 어휘와 그어휘에 해당되는 품사 정보와 더불어 명사의 의미속성(semantic properties)에 관한 메뉴가 도표-4와 같이 주어진다.

- ** <TYPE 1 Semantic Property Menu> **
- 0. NewCategory
 - 1. HUMAN

- 2. ANIMAL
- 3. ED-LIQ
- 4. ED-SOL
- 5. CON-OBJ
- 6. PLACE
- 7. BP-IA (Body Part-InAlienable)
- 8. SPORTS
- 9. ABSTRACT
- 10. PLANT

THE NEW ITEM = SAKWA
THE LEX. CAT. = NOUN

도표-4

사용자는 다음의 프롬프트에서 새로운 어휘 '사과'에 해당되는 의미정보의 숫자를 입력시키게된다.

Enter the number of a semantic property >>

여기에서 사용자는 위의 도표-4에서 '사과'에 해당되는 의미정보 숫자를 선택하든지, 혹은 현존의 의미정보 메뉴에 새로운 의미요소를 첨가시킬 수 있는 선택사항을 갖게된다. 따라서 사용자는 주어진 입력문장('나는 사과를 먹는다.')의 새 어휘 '사과'에 대해 의미속성 'ED-SOL' (4)를 선택하거나 혹은 좀 더 구체적인 의미속성 'FRUIT'와 같은 의미속성을 추가시킬 수있다. 새로운 의미속성을 추가시키기 위해선 위의 프롬프트에서 사용자는 '0'을 입력시킨다. 그러면, DCS는 새 의미속성의 구체적 이름을 다음의 프롬프트에서 요구한다.

Enter the name of the new semantic property >>

사용자가 이 프롬프트에서 'FRUIT'를 입력시킨 후에는 현존하는 동사의 의미속성(semantic codes)에 수정을 가할 필요가 있다. 'FRUIT'라는 의미속성이 사전정보에 현존하지 않았기 때문에 '먹-'과 같은 동사의 분석을 위해선 동사의 의미속성을 확장시킬 필요가 있다. 따라서 동사 '먹-'의 현존 의미코드 ((E HB) (E ED-SOL))와 ((E ANI) (E ED-SOL))에 ((E HB) (E FRUIT))와 ((E ANI) (E FRUIT))라는 의미코드를 추가시켜야 한다. 추가 입력 과정은 다음의 동사 편에서 설명되어진다.

3.2.2. To Choose Verb

만약 사용자가 새로운 어휘 '수용하다'(나는 그를 수

용할 것이다 혹은 나는 그 아이디어를 수용했다)라는 어휘를 추가하고 싶다면, 다음의 프롬프트에서 '2'(도표-3 참조)를 입력시킴으로써 도표-5와 같은 보조메뉴의 도움을 받게된다.

** < Type 2 Semotactic Code Menu > **

<TYPE1>

1. HUMAN 2. ANIMAL 3. ED-LIQ 4. ED-SOL 5. CON-OBJ
6. PLACE 7. BP-IA 8. SPORTS 9. ABS 10. PLANT

<TYPE2>

0. New Code

1. E HUMAN E ED-SOL 2. E ANIMAL E ED-SOL
3. E HUMAN E ED-LIQ 4. E ANIMAL E ED-LIQ
5. E HUMAN E CONC-OBJ 6. E HUMAN ANY-THING ANY-THING
7. E ANIMAL ANY-THING ANY-THING
8. E HUMAN B BP-IA E HUMAN
9. E HUMAN B BP-IA E ANIMAL
10. E ANIMAL E HUMAN 11. E HUMAN E HUMAN
12. E HUMAN 13. E ANIMAL
14. E PLANT 15. E HUMAN E CONC-OBJ E HUMAN
16. E HUMAN E ANIMAL E HUMAN
17. E HUMAN E ED-LIQ E HUMAN
18. E HUMAN E ED-SOL E HUMAN
19. E HUMAN B SPORTS
20. E PLACE E PLANT

THE NEW ITEM = suwyongha-

THE LEX. CAT. = Vt

(*'B'는 'Blocking'을 'E'는 'Enabling'을 의미함)

도표-5

위의 도표에서 TYPE1인 명사에 대한 의미속성은 동사에 새로운 의미코드를 부여할 경우를 대비해 보여진다. 만약 사용자가 주어진 어휘에 대해 새로운 의미속성을 추가할 필요성이 있을 땐 명사 페와 마찬가지로 '0'를 입력시킴으로써 동사에 대한 새로운 의미코드의 추가과정이 시작된다. 어느 한 동사가 한 가지 이상의 의미코드(semotactic code)를 지닐 수 있으므로 다음의 프롬트는 주어진 동사의 요구 가능한 의미코드의 갯수에 대한 정보를 묻는다.

Enter the number of semotactic code(s) applicable to this item >>

동사 수용하-와 같은 경우, 사용자는 나는 그를 수용할

것이다 혹은 나는 그 아이디어를 수용했다와 같은 문장을 고려하여, 위의 프롬프트에서 '2'를 입력시킨다. 그러면, DCS는 구체적인 의미코드에 대한 정보를 요구하는 프롬트를 두번 보여준다.

Enter the number of a semotactic code >>

Enter the number of a semotactic code >>

이 프롬트에서 사용자는 나는 그를 수용할 것이다라는 문장에 대해선 (P1.HB P2.HB)를 나는 그 아이디어를 수용했다라는 문장에 대해선 (P1.HB P2.ABS)를 입력시킨다. 따라서 사용자는 도표-5에서 '11'(P1.HB P2.HB)를 선택하고 (P1.HB P2.ABS)를 위해선 '0'을 선택하여 새로운 의미코드를 첨가시킨다.

Enter the number of a semotactic code >> 11

Enter the number of a semotactic code >> 0

명령어 '0'은 다음의 프롬트를 부른다.

Enter a new semotactic code >>

새로운 의미코드 추가시, 사용자는 DCS에서 사용되어지는 기호(Lee 1988 참조)를 사용하여 위의 프롬트에 적절한 정보를 제공해준다. 동사 수용하-의 세 의미코드를 위해선 P1(Participant1)으로 'Human Being'을 P2로는 'Abstract'가 요구되어지므로 위의 프롬트에 사용자는 다음과 같은 정보를 제공해준다.

Enter the new semotactic code >> E HB E ABS

3.2.3. Choose NOUN and VERB

한국어의 경우엔 명사형, 동사형어미가 존재하여 한 어휘가 다품사 현상을 나타내는 경우가 없지만, 영어에서는 plant(I planted the tree or I like the plant)처럼 한 어휘가 타동사이거나 명사로 다품사 현상을 나타내는 경우가 있다. 이러한 경우, 사용자는 다음의 프롬트에서 '2'를 입력시킨다.

Enter the number of lexical categori(es) applicable to this item >> 2

다시 두개의 구체적 품사 정보에 대한 프롬트에 대해 사용자는 명사('1')와 타동사('2')를 답하면 명사에 대해선

도표-4와 같은 명사의 의미속성 메뉴가 동사에 대해선 도표-5와 같은 동사의 의미코드 메뉴가 전개되고, 각각의 처리과정은 3.2.1(명사)과 3.2.2(동사)의 설명에 준한다.

잘못되었다면 위의 프롬트에서 'n'(no)를 답하여 입력정보를 파기시키고 새 정보를 다시 입력시킨다.

3.3. Display New Information

새로운 어휘정보 처리과정을 기존의 사전정보에 저장하기 전에, 사용자의 실수여부를 확인하기 위하여 새 정보에 대한 확인작업이 필요할 경우 사용자는 'Display New Information'메뉴를 선택한다. 그러면 새 어휘에 대한 입력 정보를 보여준다. 명사 사과와 동사 수용하-의 경우 각각 아래와 같은 정보를 보여준다.

```
** <Information about a New Item> **
THE LEXICAL ITEM = SAKWA
Lexical Category No. for this Item = 1
THE LEX. CAT. = NOUN   PROP. = FRUIT
```

```
** <Information about a New Item> **
THE LEXICAL ITEM = SUYWONGHA-
Lexical Category No. for this Item = 2
THE LEX. CAT = Vt SEMOTACTIC CODES = E HB E HB,
E HB E. ABS
```

위와같이 정보가 전개된 후에 DCS는 주 메뉴로 돌아가게 된다.

3.4. Exit

새로운 정보를 확인한 후, 사용자는 도표-1 주 메뉴로 돌아가 Exit메뉴를 선택하게 된다. 만약 사용자가 화면에 나타난 정보의 정확성에 만족하게 되면 다음의 프롬트에 'y'(yes)를 답한다.

Do you want to save this item? (y/n)

그러면 새 어휘는 사전정보(Master Dictionary)에 저장되어 주어진 입력 문장에 대한 나무구조(parse tree)와 개념구조(conceptual representation)가 출력된다. 예를들어, 나는 사과를 먹었다라는 입력문장에 대해 '사과'가 사전정보에 없었다면 위의 명사 처리과정을 거쳐 '사과'가 사전정보에 추가 입력된 후에 주어진 문장에 대한 나무구조와 개념구조가 각각 출력되는 것이다.

반면에 만약 사용자가 새 정보를 입력시키는 동안 타이프 실수나 메뉴로부터의 선택오류로 인해 입력 정보가

참고문헌

1. Lamb, S. 1995. Toward a Cognitive Theory of Language. Manuscript.
2. Lee, C.I. 1988. A Dynamic Model of Comprehension with Applications to Korean and English. Hanshin Publishing Co.
3. ----. A Dynamic Comprehension System with Extended Semiotic Codes. 1994. Proceedings of the ISATED International Conference: Artificial Intelligence, Expert System, and Neural Networks. pp. 165-169.
4. Schank, R. & Childers, P. 1984. On Language, Learning and the Cognitive Computer. Addison Wesley.
5. Wilensky, R. & Arens, Y. 1980. PHRAN:a Knowledge Based Approach to Natural Language Analysis. Berkeley Electronics Research Laboratory, Memo No. UCB/ERL/M80/34.