論文96-33B-4-20

한국어 문장내 체언류 조응대용어의 해결방안

(A Method of the Substantives Anaphora Resolution in Korean Intra-sentential)

金 貞 海 *, 李 相 國 **, 李 相 祚 ***

(Jung Hae Kim, Sang Kook Lee, and Sang Jo Lee)

요 약

본 논문에서는 한국어 문장내에서 일어나는 대용어 문제에 대해 중심어 주도의 단방향 활성 차트 파싱을 이용하여 해결하는 방안을 제시한다. 이는 자연어 대화에서 자주 발생되는 현상이며, 실용목적의 자연어처리 시스템 구축에 있어 필수적으로 요구되는 부분이다. 대용어 해결을 위해 한국어에서의 대용어에 대한 전산학적인 정의를 내리고, 이 정의를 기본 범위로 하는 대용어와 선행어사이의 의미 분류 및 대용어 해결 과정에 필요한 처리 조건등을 설정하였다. 또한 파서내에 대용어 처리를 위해 사전내의 자질구조로 ANAPHOR, PERSON, NUM, INDEX자질을 추가하였고, 대용어 해결을 위한 알고리즘을 제안하였으며, 기존에 개발된 HPSG 파서가 처리하는 모든 문장에서 야기된 체언류 대용을 해결할 수 있도록 파서의 처리 범위를 확대하였다.

Abstract

The purpose of this paper is to show that the solutions of the problem for the anaphor occured in Korean sentence, by means of one-direction activated chart parsing leaded by a head. This is the phenomenon frequently occured in the conversation of natural language and the part necessarily required in the construction of natural language processing system for the practical use. To solve the problem of anaphor in the Korean language, we have computerized definition and the management conditions necessary in the semantic classification between the anaphor and its antecedent and the process of solution for the anaphor are established. And the features Anaphor, Person, Num and Index are added in the feature structure in lexicon. To deal with anaphor in parser and algorithm is proposed to solve the problem for anaphor. The range of management of parser is extended to solve the problem for anaphor of the indeclinable parts of speech in Korean occured in all the sentences the parser HPSG developed previously manages.

* 正會員, 上智專門大學 電算情報處理科

(Dept. of Computer Information Processing, Sangji Junior College)

** 正會員, 韓國 MJL(株) 技術研究所

(Reseach of Technical, MJL Co.)

*** 正會員, 慶北大學校 컴퓨터工學科

(Dept. of Computer Engineering, Kyungpook National Univ.)

接受日字:1995年4月26日, 수정완료일:1996年3月21日

Ⅰ. 서 론

대용어(anaphora)는 자연어 대화에서 경제성을 고려하여 자연스럽게 많이 사용되는 현상이면서도 대용의 범위가 매우 다양하고 미묘하게 발생된다는 점에서 오랫동안 이론 및 계산 언어학 분야에서 중요한 주제로 연구되어 왔다[1,2,8,9,11-15]. 또한 자연어에 있어 대용어의 빈번한 사용은 실용목적의 자연어처리 시스템 구현에서 이들을 처리할 수 있는 능력이 필수적으로 요구된다. 특히 최근에 대화형 컴퓨터 시스템에서 자연어

인터페이스에 대한 상당한 관심이 대두되었고, 이러한 자연어 인터페이스 개발의 주요 동기는 사용자가 특별한 인공언어를 배우는 대신에 자연어로 컴퓨터 시스템과 더 자연스럽고 쉽게 대화할 수 있다는 신뢰성때문이었지만, 자연어를 컴퓨터의 내부형태(internal form)로 분석하기 위한 능력은 자연스럽고 효과적인 인터페이스를 만들기에는 충분하지 못했다. 이는 컴퓨터상의 자연어 인터페이스가 자연어상에 대두되는 여러 가지 현상 특히 화용적인 현상 -대용,생략 등- 을 다루어야 하기 때문이었다. 이러한 대용어 현상의 해결에 대해 국어학 분야의 이론적인 연구[1,2,7,9]뿐만 아니라 전산학적으로는 국내에서 국어의 통사규칙과 화제 개념을 이용한 규칙기반(rule-baesd) 시스템을 구축하여 대용어의 해결을 시도한 연구[6,8]가 있었고, 기존의 자연어 처리 시스템에서는 처리 문장에서 조응대용이 있는 경우 대명사로만 처리하여 구문분석 후단계에서 처리의 애매성을 유발시키는 요인으로 남겨 둔 방법[4,5]이 있었다. 한편 국외에서는 문장에서 대용어의 유형에 따른 다양한 방법론을 제시한 연구[11,13] 및 다양한 범주의 대용어 중에 몇몇 현상을 다루는 기법이 최근에 제한된 영역 시스템(limited domain system)에 대해서 자연어 인터페이스의 개발[14,15]이 이루어지고 또한 담화이해의 일반적인 원칙을 기반으로 조응대용의 애매성을 해결하기 위해 추론적인 방법[15]을 도입하고 있다.

한국어의 대용현상은 기본적으로 하나의 문장안에서 야기되는 현상으므로 자연어 처리의 한부분인 기계번역이나 질의 응답 시스템의 구축에 있어 문장내의 대용어의 해결이 선행되어야 보다 명확한 시스템의 구축이 이루어지므로 반드시 필요한 작업이다. 따라서 본 논문에서는 한국어의 다양한 대용어 현상중에서 하나의 문장내(intra-sentential)에서 야기되는 체언류의 대용어로 조응대용을 중심으로 그 선행어(antecedent)를 찾는 것에 중점을 두며, 통합기반 문법의 한 유형이며 한국어의 특징을 잘 나타낼 수 있는 문법이론인 HPSG를 이용한 중심어 주도의 단방향 활성 차트 파서를 구현하여 대용어를 처리하였다. 따라서 본 논문의 시스템은 기존에 개발된 HPSG 파서가 처리하는 모든 문장에서 야기된 체언류 대용을 해결할 수 있도록 파서의 처리 범위를 확대하여 자연어 처리에 필수적인 화용처리에 보다 접근한 방법으로 대부분의 실용적인 국어 문장에 대해서 분석할 수 있는 파서를 구현하였다.

본 논문의 구성은 Ⅱ장에서는 대용어 해결을 위해 한국어에서의 대용어에 대한 전산학적인 정의를 내리고, 이 정의를 바탕으로 대용어의 선행어를 찾기 위한 대용어와 선행어사이의 의미 분류 및 대용어 해결 과정에 필요한 처리 조건등을 설정하였다. Ⅲ장에서는 대용어의 해결을 위해 ANAPHOR, PERSON, NUM, INDEX자질을 추가하였고 대용어의 해결을 위한 알고리즘 및 처리예문을 보였다. Ⅳ장에서는 연구 결과와 앞으로 과제를 제시하였다.

Ⅱ. 대용어의 정의

1. 한국어의 대용어 현상

대용(anaphor)이란 하나의 문장이나 문장과 문장사이에서 같은 요소가 되풀이될 때 언어 사용의 경제성(language economy)을 위하여 잉여적 표현을 제거하는 것으로 그 되풀이된 요소를 대명사와 같은 간략한 언어 표현으로 대치하여 쓰는 현상이다[1,2,9]. 이때 되풀이 된 요소를 대용어로 바꾸는 작용을 한 요소가 선행어이고 대용어로써 대치되는 요소를 대용어의 본디말(full form)이 된다. 또한 대용의 분류는 대용의 대상이 언어적 문맥속의 사실인가 아니면 담화주변의 상황이나 의식속의 사실과 같은 비언어적 사실인가에 따라 문맥대용과 상황대용으로 나눌 수 있으며[1,9], 문맥대용중에서 선행어와 대용어의 관계로써 나타나는 조응대용(anaphora)과 되풀이 관계에 있는 두 요소사이에서 앞선 요소를 대용하는 것이 아니라 '의미적으로 가리키는' 지시대용(reference)이 있다.

예문1) 어제 꽃을 샀다. 그것은 장미였다.

예문2) 철수는 머리가 좋다. 그 좋은 머리를 늘 놀음에만 쓴다.

위 두 문장에서 예문1)의 '그것'은 선행어로 '꽃'을 가지지만, 예문2)의 '그'는 선행어가 없고 앞 문장을 지시하고 있다. 따라서 본 논문의 처리는 예문1)과 같은 선행어와 대용어의 관계가 있는 문맥대용중에서 조응대용에 대해서만 처리한다. 이러한 한국어에서 조응대용 현상이 일어날 수 있는 조건을 [1,2,9]에서는 몇가지로 제시하였으나 요약하면 다음과 같다. ① 동일성 원리로 각 대용어는 그들의 선행어와 동일한 의미관계에 있다. 즉, 형태상의 동일성과 함께 의미상으로도 동일해야 함을 나타내는 데 이는 선행어와 조응어

는 지시의 동일성(identity of reference)이 아닌 의미의 동일성(identity of sense)을 나타낸다. ② 대용어가 나타난 것보다 앞선 어절에 선행어가 존재한다는 선행성원리로 순행대용만을 취급함을 나타낸다. ③ 체언이 되풀이될 때 적용되는 것으로 구정보성(old information)을 지닌다는 되풀이 관계와 구정보성이 있다.

예문3) 철수는 자기의 모자를 썼다.

예문4) 영희는 병원에 갔다가 거기서 철수를 만났다.

예문3,4)의 문장에서 나타난 '자기', '거기'는 각각 앞선 요소를 대용하고 있으나, 지금껏 자연어처리 시스템[4,5]에서는 '자기', '거기'를 대명사로만 처리하여 그대로 두었다. 이렇게 처리할 경우 구문분석기의 다음 단계 처리에서 많은 애매성을 유발시키는 요인이 될 수 있다. 따라서 말의 경제성으로 인해 문맥상에서 대용어의 사용은 자연스런 현상이므로 이러한 문장의 처리도 자연어처리의 고려의 대상이 되어야 하므로 본 논문에서는 조응대용에 대한 처리 범주로 한 문장내에서 야기되는 대용어와 체언류의 대용어로 재귀화와 대명사화만을 다루며 순행대용만으로 대용어에 대한 선행어를 찾아 보다 명확한 문장의 분석을 위한 해결 방법을 논의하였다.

2. 전산학적인 대용어의 정의

대용어의 전산학적인 정의를 위해 먼저 한국어의 대용어 구성에 대해 살펴보면, 대용어의 구성은 「지시사+대용사」로 이루어져 있으며 지시사에는 「이,그,저」와 대용사에는 「이, 너, 놈, 것, -어기, -런,...」 등이 있지만, 국어학적으로 지시사 「이,그,저」 표현이 아래와 같은 특성이 있다[7,9]. 즉, "이,그,저"표현이 본래 직시적(deictic)으로 담화장면에서 쓰이는 것이지만 전술문과 조응하여 쓰일 경우의 특징으로

① '그'는 전술된 지시체에 조응하여 화자에게는 지정(definite)한 기지(旣知)의 정보를 준다.
② 후술문에 조응할 때는 '이'계열만 쓴다.
③ '그'계열과 '이'계열이 전술문맥에 조응하는데 이 경우 '이'계열이 '그'계열보다 사람에 대한 묘사가 직접적이다.
④ '저'계열은 직시적으로만 쓰이고 전술문맥에 조응하는 용법은 없다.
⑤ 지시사'이,그,저'중에서 '그'가 담화문중에서 선행어를 언급하고 지시하는데 쓰이고 '이'와 '저'는 화시적 지시가 주된 기능을 한다.

위와 같은 특성으로 인해 본 논문에서는 체언류의 대용에 쓰이는 대용어로 지시사「그」+대용사「이,녀,남자,것,...」 등에 의해 결합된 조응대용만을 전산학적인 대용어로 정의한다.

[정의 : 전산학적인 체언류의 대용어]

대용이란 말의 경제성에 따라 반복하여 나타난 용어에 대해 대신하여 쓰는 것으로 '지시사+대용사'의 구성을 지니며, 특히 「지시사'그' + 대용사'이,녀,놈,남자,것,-어기'」로 된 경우로 언어적 문맥에서 선행어와 대용어와의 관계로써 나타나는 조응대용을 전산학적인 대용어이다.

우리말은 어휘의 특성상 대치되는 대용어의 종류를 여러가지 형태로 정의할 수 있으나 선행어의 의미에 따라 체언류 대용중에서도 사람 대용인 경우 '자기, 그, 그사람, 그분, 그녀, 그남자, 그이'가 쓰이고, 사물 대용인 경우 '그것, 그'가 쓰이며, 위치·방향 대용인 경우 '그리, 거기'가 쓰인다. 한 문장내에서 체언류 대용어의 어절의 구성은 대용어의 구조에 따라 다음과 같이 나타낼 수 있다.

- 대용어의 구조 = '지시사「그」' + '대용사'
- 체언류 대용어의 어절구성 = '대용어' + 조사

3. 대용어와 선행어의 의미분류 관계

대용어는 본래 고유한 실질적 의미가 없는 것을 특징으로 하며, 항상 선행어에 의해 그 뜻이 해석된다. 또한 대용어는 선행어가 가지는 구체적인 뜻대신에 그것을 형식적으로 가리키기만 하는 부류적 의미를 지닌 것이 특징이다. 따라서 대용어와 선행어사이의 의미관계를 다음과 같이 나타내었다.

선행어의 의미분류		대용어
체언류대용	사람	자기
	사람	그,그남자,그녀,그사람,그분,그이
	사물	그것,것
	방향,위치	그리, 거기

4. 대용어의 해결 조건

앞 절에서 나타낸 대용어의 정의와 그에 따른 대용어의 종류 및 선행어와의 의미관계에 의해 다음과 같이 대용어와 선행어의 관계를 규명하였고, 다음의 각

절에서 선행어의 선정 범위와 대용어의 처리 조건을 나타내었다.

[정리] 선행어가 대용어에 대한 가능성 있는 선행어가 되기 위한 조건[12,1,9]으로

① 최근성(recency): 대용어와 일치한 선행어보다 더 최근의 단위는 발생하지 않는다.

② 일치성(compatibility): 선행어와 대용어의 일치성이 있다.

여기서 일치성은 대용어 해결과 관련된 언어학적인 성질의 일치성(범주, 수, 성, 인칭,...)을 나타내는 것으로 본 논문에서는 의미자질(사용된 대용어와 선행어 사이의 의미값) 값의 일치성을 나타내며, 최근성은 대용어에 가장 가까운 일치성을 지닌 선행어를 의미한다.

1) 선행어의 선정 범위

대용어가 지칭하는 선행어의 범주는 '그'에 의한 조응대용에서 그 지시대상의 유일성, 특정성이 언어적 문맥에 의해서만 보증됨을 나타내므로 '그'가 붙은 명사구의 지시대상이 선행하는 언어적 문맥에 의해 특정되는 대상에서 핵심성분이 된다. 즉, '수식어+명사'구조가 대용되었을 경우 이 구조의 핵심성분인 명사가 대용어의 선행어가 된다. 따라서 Ⅲ장의 처리과정에서 보여줄 체언-사람-재귀화 예문「한 소년이 자기의 책을 읽는다」에서 대용어「자기」의 선행어는 명사구에서 핵심성분인「소년」이 된다.

2) 대용어의 처리 조건

2장 1절에서의 국어학적으로 대용어 현상이 야기되는 조건과 앞에서의 정리 및 선행어의 선정 범위에 의해서 하나의 문장이 아래와 같은 성분으로 이루어져 있을 때 대용어(NPb)에 대해 선행어(NPa)를 찾기 위한 조건을 다음과 같이 나타내었다.

여기서 S는 하나의 문장을 의미하고 문장 S의 성분으로 NPa는 선행어, NPb는 대용어, X와 Y는 각각 임의의 문장 성분요소가 된다.

S : NPa - X - NPb - Y

[대전제조건] 대용어 NPb의 왼쪽으로 선행어의 조건을 만족하는 가장 가까운 명사가 대용어 NPb의 선행어이다.

만약 NPb가 '자기'라는 대용어가 쓰인 경우 NPa가 선행어가 될 조건

[조건1] NPa는 문장의 주어로써 3인칭 단수이다.

[조건2] NPa는 사람이거나 유정명사의 의미부류에 속해야 한다.

만약 NPb이 '대명사'의 대용어가 쓰인 경우

[전제조건1] 대용어 NPb의 의미자질 값과 대상 선행어의 의미자질 값을 비교하여 동일한 값을 지니는 것을 대용어의 선행어로 한다.

[전제조건2] 대용어 NPb의 선행어로 결정되는 명사는 명사구의 핵심이어야 한다.

① NPb가 대용어 '그, 그사람, 그분, 그남자, 그녀,그이'인 경우 선행어의 조건

[조건1] 대상 선행어인 NPa는 유정명사라야 한다.

[조건2] 대상 선행어인 NPa는 3인칭 단수라야 한다.

[조건3] 대상 선행어인 NPa는 체언-사람을 지칭한다.

② NPb가 대용어 '그것, 것'인 경우 선행어의 조건

[조건1] 대상 선행어인 NPa는 무정명사라야 한다.

[조건2] 대상 선행어인 NPa는 체언-사물을 지칭한다.

③ NPb가 대용어 '그리, 거기'인 경우 선행어의 조건

[조건1] 대상 선행어인 NPa는 장소의 자질을 갖는 명사라야 한다.

[조건2] 대상 선행어인 NPa는 체언-방향 또는 체언-위치를 지칭하는 명사이다.

Ⅲ. 대용어의 해결

1. 중심어 주도의 차트 파싱

본 논문의 파서는 대용어의 해결을 위해 문법은 HPSG를 채택[10]하고 파싱기제로는 중심어 주도의 단방향 활성 차트 파서[3,4]를 이용한다. 이는 한국어의 경우 선행어는 항상 대용어 앞에 존재하기 때문에 우에서 좌로의 단방향 파서를 이용할 경우 선행어를 찾는데 유리하기 때문이다.

따라서 본 논문에서 대용어의 해결을 위해 고려한 대상은 체언 대용을 담고 있는 모든 종류의 한 문장이지만, 실제 실험 대상은 본 논문의 파서에 의해 파싱 가능한 문장으로 한다. 본 논문 파서의 파싱 범위는 보

어-중심어, 수식어-중심어, 중심어-중심어 구성의 거의 모든 실용적인 문장에 대해 분석가능하고, 파서의 입력은 형태소 분석이 이루어진 결과라고 가정하며 출력은 입력 받는 문장에 대해서 통사 및 의미를 통합적으로 처리하고 입력문장에서 추출한 정보를 파서이후의 응용 시스템에서 이용할 수 있는 내적표현으로 형식화하였다.

본 논문에서는 (그림1)과 같은 자질구조를 이용해서 언어정보를 형식화한다.

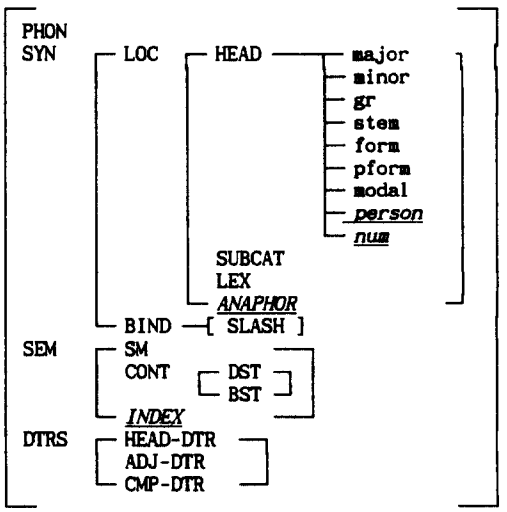


그림 1. 자질 구조
Fig. 1. Feature Structure.

여기서는 새롭게 도입한 자질에 대해서 중점적으로 기술하고 각 자질들에 대한 자세한 설명은 [3,4]에 기술되어 있다. 먼저 PERSON자질은 인칭을 나타내는 자질로서 1인칭(first), 2인칭(second), 3인칭(third)의 값을 갖는다. NUM자질은 수를 나타내는 자질로서 단수(single), 복수(couple)의 값을 갖는다. ANAPHOR 자질은 대용어의 유형을 나타내는 자질로서 1은 '자기'의 재귀화 대용, 2는 '그것'등의 사물 대용, 3은 '그리, 거기'등의 방향, 위치 대용, 4는 '그, 그사람, 그이, 그분'등의 대명사화 대용을 나타낸다. INDEX자질은 대용어의 경우 선행어 노드를 지칭하는 자질이다.

2. 알고리즘

본 논문에서는 대용어의 해결을 위한 파서로 중심어 주도의 상향식 차트 파서[4]를 사용한다. 이 파서의 대표적인 자료구조는 다음과 같다.

INPUT : 입력 edge들의 모임
PEND : 처리가 완료되지 않은 edge들의 모임
CHART : 처리가 완료된 edge들의 모임

각 edge의 구조는 시작노드 번호, 종료노드 번호, 자질구조 및 다음 노드에 대한 포인터들로 구성된 구조를 가지며, (그림2)는 HPSG의 전체 파서 알고리즘중에서 본 논문에서 사용된 대용어 해결의 흐름도만을 간략히 나타냈다.

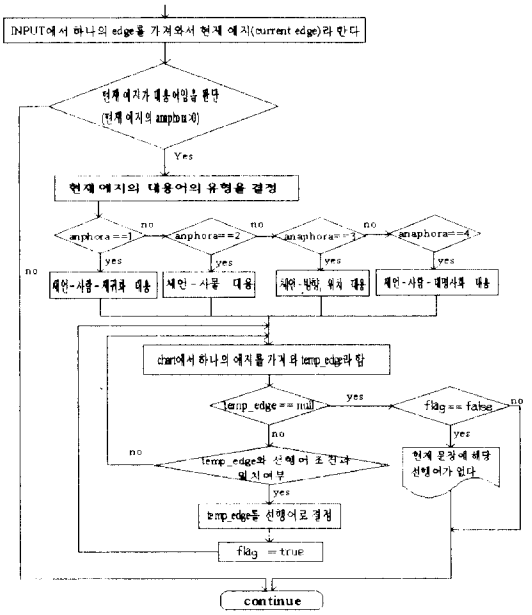


그림 2. 대용어의 해결을 위한 흐름도
Fig. 2. Flowchart for the solution of anaphora.

3. 처리 예문

1) 재언-사람-재귀화 대용

(i) 선행어의 조건

① fs->sem.sm.anim == '+' (생물)
② fs->syn.loc.head.person == third(3인칭)
③ fs->syn.loc.head.num == single(단수)
④ fs->syn.loc.head.gr == subj(주어)

(ii) 예문

[한 소년이 자기의 책을 읽는다.]
edge1 edge2 edge3 edge4 edge5

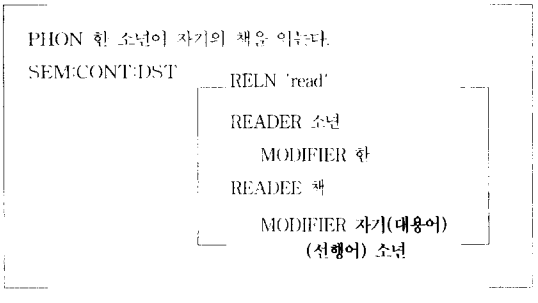
위의 예문이 입력되었을 경우 edge2까지 파싱된 결과로써 CHART에는 (edge1, edge2, edge1+edge2)가 존재한다. 그 다음 입력에서 하나의 edge를 가져와 처리를 계속할 때 edge3 [자기의]가 현재에지(current edge)가 되고, 이 에지의 anaphor자질값(an-

aphor=1)에 의해 현재 에지가 대용어임을 알 수 있다. 그래서 앞에서 제시한 대용어 처리 알고리즘에 따라 anaphor자질값이 1이므로 이 대용어가 재귀화 대용임을 결정하고 지금까지 처리가 완료된 에지들이 들어있는 CHART를 탐색해서 위의 선행어의 조건에 부합하는 에지를 이 에지의 선행어로 결정하게 된다.

이 예문의 경우 edge1과 edge2가 결합된 edge1+edge2 [한 소년] 이 선행어의 조건을 만족하므로 먼저 선행어 후보가 되고, 아직 CHART에 에지들이 남아 있으므로 탐색을 계속한다. 이때 edge2 [소년이] 가 또한 선행어의 조건을 만족하므로 선행어 후보로 등록된다. 본 논문의 대용어 선행어 선정 범위는 「수식어-중심어구조의 의미 합성」이 이루어지지 않는 단독 중심어 「LEX+」를 선행어로 선정한다. 이는 본 논문에서 사용하는 자질구조의 자질속성, LEX를 이용하여 처리할 수 있다. 즉, edge1+edge2는 fs >syn.loc.lex가 ' '값을 가지며 edge2는 '+'값을 갖는다. 따라서 edge2 [소년이] 가 선행어로 결정된다.

대용어 처리에 직접 관련되지 않는 다른 구성요소의 처리 및 본 논문의 처리 결과에 사용된 의미표현 및 의미합성 원리는 중심어 주도의 단방향 활성차트파서[4]의 파싱결과에 따르는 것이므로 본 논문에서는 생략하기로 한다.

(iii). 처리 결과



2) 체언 사람·대명사화 대용

(i) 선행어의 조건

① fs >sem.sm.anim == '+' (생물)

② fs->syn.loc.head.person == third(3인칭)

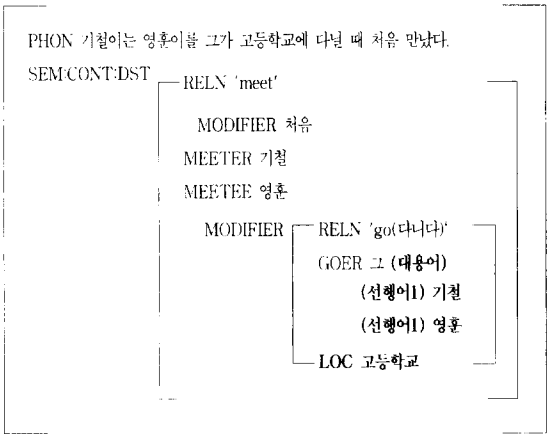
③ fs >syn.loc.head.num == single(단수)

(ii) 예문

[기철이는 영훈이를 그가 고등학교에 다닐 때 처음 만났다.]
edge1 edge2 edge3 edge4

위의 예문에서 [그] 는 [기철이] 를 대용할 수도 있고 [영훈이] 를 대용할 수도 있는 경우로서 문장 자체로서는 대용어의 선행어가 모호하다. 이러한 대용어가 대용할 선행어가 모호할 경우에는 대용어 에지 [edge3] 의 INDEX자질이 edge1 [기철이] 노드 및 edge2 [영훈이] 노드 즉, 두개의 선행어를 지칭하게 된다. 이는 선행어 조건에 의해 edge2 [영훈이] 가 먼저 선행어로 선정되지만 아직 CHART에 edge1 [기철이] 이 남아 있기 때문에 처리를 반복한다. 따라서 edge1도 선행어의 조건을 만족하므로 또한 선행어로 선정된다. edge1과 edge2는 수식어-중심어 결합구조에 참여하지 않는 단독 체언 [LEX+] 이므로 둘 다 선행어로 선정된다.

(iii) 처리 결과



3) 체언·사물 대용

(i) 선행어의 조건

① fs->sem.sm.anim == '-' (무생물)

② fs->sem.sm.huma == '-' (사물)

(ii) 예문

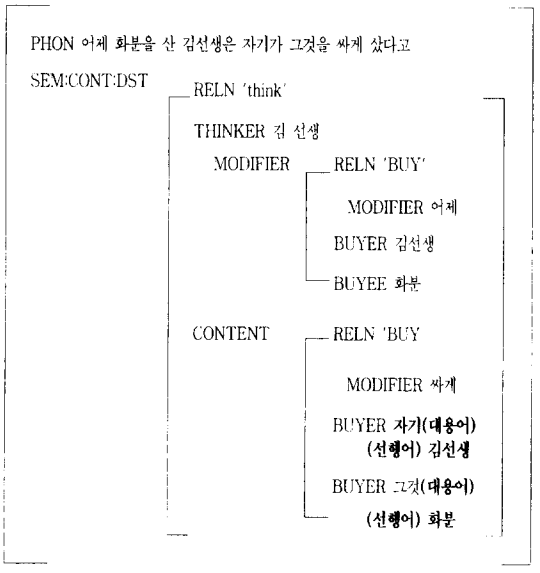
[어제 화분을 산 김선생은 자기가 그것을 싸
edge1 edge2 edge3 edge4 edge5 edge6

세 샀다고 생각했다.]

위의 예문에서 내포된 문장에 [자기가] 와 [그것을] 의 두개의 대용어가 나타나는데 edge5 [자기가] 의 대용어 처리는 앞의 Ⅲ장 3절의 체언 사람 재귀화 대용어 처리에 의해 edge4 [김선생] 가 선행어로 선정된다. edge6 [그것을] 의 처리는 위(i)의 선행어 조

건에 의해 edge2 [화분] 가 선행어임을 결정한다.

(iii) 처리 결과



Ⅳ. 결 론

본 논문에서는 중심어 주도의 단방향 활성 차트를 이용하여 자연어 사용에 있어 빈번히 발생하는 대용어 현상에 대한 해결 방안을 제시하였다. 전산학적인 대용어의 정의 및 선행어와 대용어의 관계 정립, 처리조건을 설정하였고, 이의 구현을 위해 국어와 같이 중심어 후행 언어에 대하여 효율적인 중심어 주도의 단방향 활성 차트 파싱을 이용하여 국어의 체언류 조응대용어의 해결을 구현하였다. 따라서 본 논문에서는 문장내에서 발생하는 체언류의 대용어에 대해 대용어가 지칭하는 내용인 선행어를 찾아 보다 명확한 문장의 분석을 통해 자연어 처리에 필수적인 화용처리에 보다 접근한 방법으로 대부분의 실용적인 국어 문장에 대해서 분석할 수 있는 파서를 구현하였다.

현재 한 문장내 체언류의 대용어 처리뿐만 아니라 관형어, 부사어 대용과 서술어의 처리 및 문장과 문장간의 체언류의 대용어 처리에 대해서도 계속 연구중에 있다.

참 고 문 헌

[1] 김일웅, 우리말 대용어의 연구, 부산대 박사학위논문, 1982

[2] 유현경,"국어 접속문의 통사적 특질에 대하여", 한글 제191호, pp.77-104, 1986

[3] 서영훈, 의미정보를 이용하는 중심어 주도의 한국어 파싱, 서울대 컴퓨터공학과 박사학위논문, 1991

[4] 이상국, 특수문형의 처리를 강화한 HPSG 한국어 파서의 구현, 경북대 컴퓨터공학과 석사학위 논문, 1993

[5] 이상조, 한국어 자연어 인터페이스를 위한 사전 구성에 관한 연구(Ⅱ), 한국전자통신연구소 위탁과제,경북대 전자기술연구소, 1992

[6] 이정현, 한국어 처리를 위한 구절 문법과 질의응답 시스템, 인하대학교 박사학위논문, 1988

[7] 장석진, 정보기반 한국어 문법, 도서출판 언어와 정보, 1993

[8] 허정, 국어 대용어해결 시스템의 설계 및 구현, 한국과학기술원 석사학위논문, 1985

[9] 홍순성, 국어 대명사의 조응현상에 관한 연구, 영남대 박사학위 논문, 1986

[10] C.Pollard and I.Sag, An Introduction to HPSG, 1987, Jan, Draft.

[11] James Allen, *Natural Language Understanding*, The Benjamin/Cummings Pub. Company, Inc., 1987.

[12] Joke Dorrepeal, "Discourse Anaphora", *In Proceedings of Colling 90*, Vol. II, pp. 95-99, 1990.

[13] J.G. Carbonell, R.D. Brown, "Anaphora Resolution : A Multi-Strategy Approach", *In Proceedings of Colling Budapest*, Vol. I, pp. 96-101, 1988.

[14] Philip J. Hayes, "Anaphora for Limited Domain System", *IJCAI-81*, Vol. I, pp. 416-422, 1981.

[15] T.Ukita, K.Sumita & S.Amano, "A Procedure Resolving Anaphoric Ambiguity by Finding the Most Informative Interpretation", *ICOT TR-220*, 1986, Dec..

저 자 소 개

金 貞 海(正會員) 第 31卷 B編 第3號 參照
현재 상지전문대학 전산정보처리과 조교수

李 相 祚(正會員) 第 31卷 B編 第3號 參照
현재 慶北大學校 컴퓨터공학과 교수

李 相 國(正會員) 第 31卷 B編 第11號 參照
현재 한국 MJL(주) 기술연구소 연구원