

# 한국어 보조동사의 인식 및 개념그래프에 의한 표현

## The Recognition of Korean Auxiliary Verb and its Description Based on Conceptual Graph

이 병 희\*  
Byeong-Hee Lee

### 요 약

한국어 보조동사는 그 수가 적음에도 불구하고 한국어 문장에서 자주 쓰인다. 그러나 보조동사의 개념을 정확하게 처리하지 않을 경우에 기계번역의 질이 좋지 않다. 이런 보조동사처리의 문제들을 해결하기 위해 본 논문은 개념그래프에 기반을 둔 보조동사의 기술을 제안한다.

보조동사의 개념그래프 기술을 위해, 각종 논문과 국어사전에서 한국어 보조동사 40개와 여러 예문들을 수집한 후, 완료, 진행, 봉사, 시도, 강세, 원망, 지속, 추정의 8가지 의미로 분류하고 이들을 개념그래프로 기술한다. 실험에서는 한국어 보조동사가 포함된 문장을 입력 받아 개념그래프로 변환하는 프로그램을 구현하고 실험 결과를 기술한다.

### Abstract

Korean auxiliary verbs are often used in Korean sentences in spite of the small number of the auxiliary verbs. However, the incorrect processing of the verbs concept leads to the poor translation quality. To solve the problems of the auxiliary verb processing, the paper proposes a description of the auxiliary verbs based on Conceptual Graph (CG).

For the description of the auxiliary verbs within CG, we first collect 40 Korean auxiliary verbs and example sentences from papers and a Korean dictionary. Next, we perform the analysis of the Korean auxiliary verbs through a classification: perfective, progressive, benefactive, attemptive, emphatic, desirable, retentive, and presumptive. Then we depict the eight meanings based on CG. In the experiment, the paper implements the program that translates sentences included in the auxiliary verbs into CG and explains the experimental results.

## 1. 서 론

언어는 의사소통을 위한 도구로서 인간이 사회 생활을 유지하고 문화의 창조, 발전, 계승을 가능하게 하는 기본 수단이다. 최근에 와서 학교, 회사 등의 기관뿐 아니라 일반 가정에까지 보급된 컴퓨터와 인터넷(Internet)이 정보의 교환 및 습득에 큰 영향을 미치고 있다. 이에 따라 컴퓨터 발명 초기의 수치와 통계 처리에서 최근에는 자연언어로 표현된 문장의 구조와 의미를 처리함으로써 언어 지식적이고 추상적이며 차원 높은 정보를 획득하려는 자연언어 처리의 연구가 주목을

받고 있다[1].

자연언어 처리에서 기계번역(machine translation)은 원시언어 문장의 의미를 전달하기 위해 원시언어 문장에 대한 분석과 번역을 위한 정보를 이용, 문법적으로 타당한 목표 언어 문장을 생성하는 것이다. 기계번역 시스템의 전체적인 성능을 향상시키려면 원시언어의 분석구조에 대한 이해를 바탕으로 여러 가지 통사적(syntactic) 지식과 의미적(semantic) 지식 등 포괄적인 지식을 목표언어의 생성을 위한 정보로 전달, 이를 바탕으로 정확한 목표언어를 생성하도록 해야 한다[2,3].

국내에서 한/영 또는 영/한 기계번역에 대한 연구가 여러 대학과 연구소를 중심으로 80년대부터 활발히 진행되어 이제는 상용 기계번역 시스템들

\* 정희원 : 충남대학교 부설연구소 연구원  
bhlee@ce.cnu.ac.kr

이 속속 발표되고 있는 실정이다. 그럼에도 불구하고 아직까지 사람에게 의한 번역 수준에는 도달하지 못하고 있으며 잘못 번역되는 수가 상당히 많다. 특히 기계번역에서 보조동사 처리에 관한 연구가 미흡, 엉뚱한 문장으로 번역되는 문제점이 빈번하게 도출되고 있다. 본 논문의 목적은 이렇듯 기계번역 시스템을 이용한 번역에서 자주 틀리는 한국어 보조동사를 개념 그래프를 이용해 표현하고 이들을 기계번역에서 이용할 수 있도록 하는 데 있다.

개념 그래프는 여러 의미망(semantic networks)을 통합한 지식표현 언어(knowledge representation language)이며, 개념 도식을 이용해 인간이 쉽게 읽을 수 있도록 논리적으로 간결하면서도 자연어 수준의 풍부한 표현력이 있으며, 컴퓨터에 의한 자연언어 처리 등에서도 쉽게 이용할 수 있다[4].

화자나 주체의 심리 상태를 나타내는 한국어 보조동사는 진행, 완료, 봉사, 시행, 강세, 원망, 지속, 추정 등의 다양한 의미를 가지는데 반해, 이를 정확하고 체계적으로 기술하는 연구는 부족한 형편이다. 이미 영어에서는 양상(modality)이나 상(aspect)이 화자나 주체의 심리 상태를 나타낸다는 측면에서 다양한 연구가 있었다. 그러나 이는 영어에 국한된 경우이기 때문에 한국어 보조동사의 의미와 다른 부분이 많다.

이런 견지에서, 본 논문은 전통적으로 표시의 어려움을 안겨 주었던 한국어 보조동사의 의미를 분석하고, 한국어 보조동사의 구조를 개념그래프로 기술하며, 한국어 보조동사가 포함된 문장을 입력받아 개념 그래프로 변환하는 프로그램을 구현하고 그 결과를 기술한다.

이러한 목적을 위해서 본 논문에서는, 1장 서론에 이어, 2장 연구 배경에서 보조동사의 개요 및 특성과 한/영 기계번역 시스템에서 보조동사 처리의 문제점을 살펴보고, 3장 보조동사의 개념그래프 기술에서는 영어에서의 개념그래프를 살펴본 후 보조동사의 유형과 의미를 기술하고 보조동사를 개념그래프로 기술하며, 4장 실험 및 결과

에서는 보조동사 처리기의 구현과 실험결과를 알아보고, 5장 결론에서 본 논문의 연구결과를 정리하고 한국어 보조동사에 관하여 의미론적으로 시사되는 바를 알아본다.

## 2. 관련연구

### 2.1 보조동사의 개요와 특성

한국어 문장에서 서술어는 다양하게 나타난다. 한국어는 서술어의 특성에 따라 문장의 통사와 의미는 물론 유형이 달라진다. 한국어의 문장 서술어는 동사 중심으로 이루어지는 경우가 많다. 특히 <V1+V2> 구조의 서술어 구조는 한국어 문장에서 보편적으로 존재하는 문법 현상이다[5].

한국어의 <V1+V2>를 통사/의미적 관련에 따라 표 1과 같이 3개의 구성으로 나눌 수 있다.

(표 1) 한국어 동사연속 구성

	구조	의미적 관계	예
접속동사	[V1]+[V2]	병렬관계	구워먹다
합성동사	[V1+V2]v	융합관계	돌아가시다
보조동사	[[V1]+V2]	주종관계	읽어 가다

접속동사는 둘 이상의 문장이 접속하여 이루어지는 것으로 구워먹다는 굽다와 먹다가 병렬적으로 합쳐진 구성이다. 합성동사는 [V1]와 [V2]가 결합하여 본래의 의미는 희석되고 본래 의미 바탕 위에 또 다른 동사 의미로 서술어 기능을 한다. 다시 말해 돌아가시다는 돌다와 가시다가 결합하여 죽다의 의미를 갖는다. 보조동사는 V2가 의미와 문법적 기능을 갖고 통사적 관계를 이루면서 V1을 수식 내지는 서술하여 보조해 준다는 뜻이다. 여기에서 V2를 보조동사라고 하며 V2는 문장 또는 선행동사에 대하여 진행, 종결, 봉사, 지속, 사동, 강세 등의 양상적 의미로 문법적 기능을 한다[6].

(1)	a. 철수는 고구마를 구워먹었다. b. 철수는 고구마를 구웠다. c. 철수는 고구마를 먹었다.
(2)	a. 철수는 그 동화책을 거의 읽어 간다. b. 철수는 그 동화책을 거의 읽었다. *c. 철수는 그 동화책을 거의 간다.

다음으로, 보조동사의 통사/의미론적 특성을 살펴보기로 한다. 첫째, 보조동사는 홀로 설 수 없다는 의존성과 비분리성의 특성이 있다[7]. 이것은 접속동사와의 커다란 차이이다. (1)은 접속동사로, (1a)는 (1b)와 (1c)로 나누어 볼 수 있으나, (2)는 보조동사로 (2a)를 (2b)와 (2c)로 나누어 보면 (2c)가 틀린 문장이 된다.

(2)에서 읽다라는 선행동사, 즉 본동사의 의미에는 변화가 없으나, 가다라는 보조동사는 그 본래 의미를 상실하고 진행이라는 화자나 주체의 심리적 상태를 나타내고 있다. 또한 보조동사 구성의 서술어를 갖는 문장에서 보조동사는 문장의 논항(argument)을 할당받지 못하고 또한 관여하지 못하기 때문에 논항의 수와는 관련이 없다.

둘째, 보조동사 구문은 선행동사와 후행동사 사이에 다른 문장 성분을 삽입시킬 수 없다. 접속동사 구문에서는 삽입이 가능하다.

셋째, 보조동사 구문은 대형(pro-form)이라는 기준을 적용하여 접속동사 구문과 합성동사 구문과 구별해 낼 수 있다.

(3)은 접속동사 구문인데 선행동사와 후행동사가 모두 대형을 취할 수 있으며, (4) 합성동사 구문은 선행동사와 후행동사가 모두 대형을 취할

(3)	a. 철수가 옷을 빨아 입었다. b. 철수가 옷을 그렇게 해 입었다. c. 철수가 옷을 빨아 그렇게 했다
(4)	a. 나도 방으로 들어갔다. *b. 나도 방으로 그래갔다. *c. 나도 방으로 들어그랬다.
(5)	a. 나도 가 버렸다. b. 나도 그래 버렸다. *c. 나도 가 그랬다.

수 없다. 이에 반해 (5) 보조동사 구문에서는 선행동사인 본동사는 대형이 가능하나, 후행동사인 보조동사 자체는 대형이 불가능하다.

지금까지의 보조동사의 특성을 요약하면, 한국어의 보조동사는 통사/의미적으로 화자나 주체의 심리상태를 나타내는데, 자립성이 없고 항상 본동사 뒤에 위치하며, 그 사이에는 다른 문장성분이나 -서가 삽입될 수 없으며, 또한 문장 내에서 통사적 기능 변화를 주지 못하고, 대형을 취할 수도 없는 동사이다.

근래의 보조동사에 관한 연구를 보면, 보조동사를 내포문을 갖고 있는 상위문의 본동사로 보는 견해가 있으며[8], 최근 연속동사 구성(serial verb construction)이라는 이름 하에 일련의 연구가 있으며[9], 보조동사 연구에서 개별적인 의미 관계를 고찰한 논문들도 발표되고 있다[10].

## 2.2 한/영 기계번역 시스템에서 보조동사 처리

본 논문에서는 자동번역 시스템의 경우 보조동사가 포함된 문장이 어떻게 처리되는가를 살펴보기 위하여 상용 자동번역 시스템을 가지고 실험해 보았다. 그림 1은 H사의 한영 자동번역시스템에서의 한국어 보조동사 포함된 원문과 자동번역된 결과의 예이다.

보조동사 처리 측면에서 보면 그림 1에서 보듯이 보조동사의 처리가 미흡하여 잘못 번역되는

인수는 뽕뽕에 나가 버렸다. Undertaking had gone to field. 영희가 편지를 쓰고 있다. Yeonghui is writing. 철수는 조카에게 종이배를 만들어 주었다. Disengagement made paper ship to nephew. 영희는 옷을 읽어 보았다. Yeonghui dressed. 그는 문을 열어 찾았다. He turned over opening the door. 나는 밥을 먹고 싶다. I want to eat bread. 나는 공책을 책상에 얹어 놓았다. I put notebook to the desk. 철수가 영희를 좋아하는 것처럼 보인다. Chul-su is looked as like Yeonghui.
--

(그림 1) 상용 자동 번역시스템에서의 보조동사처리의 예

것을 알 수 있다. 본 논문은 이렇게 보조동사의 의미를 간과하여 잘못 번역되고 있는 시스템들에서 이용할 수 있도록 개념 그래프를 이용하여 그 의미를 표시하기 위함이다.

### 3. 보조동사의 개념그래프 기술

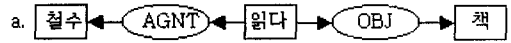
#### 3.1 영어와 개념그래프

개념그래프는 철학자이며 논리학자인 1800년 후반 C. S. Peirce가 기호논리학에 사용하기 위해서 개발한 그래픽 표기인 존재그래프(existential graph)에 바탕을 두고 있으며, 개념(concept)을 표현하는 개념노드(concept node)와 개념간의 관계를 표현하는 관계노드(relation node)로 연결되는 이분 그래프(bipartite graph)이다.

개념그래프는 1984년 Sowa의 생각들이 책으로 나오면서 활발하게 연구가 진행되어 왔으며, 근래에 들어 국제적 표준(ISO/IEC 14481 on Conceptual Schema Modeling Facilities)으로 자리잡고 있다. 또한 인공지능에서 자주 쓰이는 술어논리(predicate calculus) 또는 LISP 언어와 유사하며 이기종 시스템간의 교환 포맷이기도 한 KIF(Knowledge Interchange Format)와도 상호호환성이 있다[11].

다음 그림 2a와 그림 2b는 철수가 책을 읽다의 개념그래프로 그림 2의 a, b, c, d가 모두 같은 뜻이며, 그림 2a는 도형적 형태(display form)의 개념 그래프, 그림 2b는 표기가 편한 선조적 표시(linear notation)의 개념그래프, 그림 2c는 술어논리 표시, 그림 2d는 KIF이다.

영어권에서는 Sowa의 연구가 중심이 되어 여러 연구가 있으며 국제적 학회도 활동 중이다. 국내에서도 전산학적으로는 연산의 효율성을 위해서 개념그래프를 개념 트리로 표시하는 연구[12]와 어휘 의미정보를 이용하여 구문 의존구조로부터 개념그래프를 생성하는 연구[13], 한국어의 구문적 특성을 반영하기 위해 의미변환 규칙을 이용하는 연구[14], 개념 그래프를 이용하여 문자 인



b. [철수]←(AGNT)←[읽다]→(OBJ)→[책]

c. (∃x:철수)(∃y:읽다)(∃z:책)agt(y,x)∧obj(y,z)

d. (exist (?x 철수) (exist (?y 읽다) (exist (?z 책) (and (agt ?y ?x) (obj ?y ?z) ) ) ) )

(그림 2) 철수가 책을 읽다의 개념그래프

식 시 발생하는 오인식을 교정하자는 제안[15], 그밖에도 구현과 응용에 관한 연구가 많이 있다. 언어학적으로는 관계구문, 복합명사, 시제와 상, 준동사구문, 조응구문, AND 등의 개념 그래프에 관한 연구가 있다[16].

다음으로 본 논문의 주제와 밀접한 관련이 있는 양상과 상에 대해 알아보자. 화자가 어떤 상황에 관하여 발화할 때 의미 표현 방법은 크게 둘로 나눌 수 있다. 첫째는 화자가 그 상황에서 실제로 일어난 사실만을 기술하는 경우이다. 둘째는 화자가 상황에 대한 단순한 사실의 표현에 머무르지 않고 문장의 내용에 대한 화자의 심적 태도를 나타내는 경우이다. 화자의 심적 태도란 문장이 서술하는 내용에 대한 화자의 의견, 태도, 판단을 말한다.

둘째의 경우를 영어로 modality라고 하며, 우리 말로는 양상 또는 서법이라고 하는데 양상은 의미론적이며, 서법은 문법적인 표현이어서 본 논문에서는 의미적 관점의 양상이라는 표현을 쓰도록 하겠다.

현대 영어에서 상(aspect)의 뚜렷한 형태는 완료형과 진행형뿐이다. 고전적인 전통 문법에서 상은 동사가 의미하는 동작의 상태, 특질, 시점의 개념 등으로 정의된다.

본 논문은 전통적으로 표시의 어려움을 안겨주었던 한국어 보조동사의 의미를 영어의 양상과 상에 관한 연구에 비추어 한국어 보조동사의 구

조를 개념그래프로 기술하고 구현하고자 한다.

영어의 양상에 관한 개념그래프 연구[17]에서는 양상을 나타내기 위해 주로 사용되는 양상조동사, 양상동사, 과거시제를 중심으로 하여 이것들을 포함하는 양상표현의 의미를 분석하고, 그 구조를 10여 개의 개념관계를 이용하여 개념그래프로 기술하고 있으며, 영어의 상에 관한 개념그래프 연구[18]에서는 Sowa의 시제와 상 기술이 한 정적이고 시제의 다양한 지시 의미와 상의 함축 의미를 충분히 나타내지 못함을 밝히고 대안을 제시하였다.

### 3.2 보조동사의 유형과 의미

본 절에서는 2.1절 보조동사의 개요와 특성에서 나타난 보조동사의 통사/의미론적 특성에 의거하여 선행동사 V1의 어미유형에 따라 나누어 보

고, 일상생활에서 자주 쓰이는 보조동사들을 수집하여 이들의 의미를 중심으로 하여 보조동사의 유형을 살펴보고자 한다.

보조동사가 선행동사 V1과 결합될 때에는 V1 뒤에 보조적으로 연결해 주는 어미가 온다. 즉 <V1-연결어미 보조동사(Aux)>가 되는데 이때 연결어미로 올 수 있는 어미 형태로 부사형, 명사형, 접속형, 종결형이 있다.

전통 문법에서는 부사형에 아/어, 게, 지, 고가 있다고 일컬어 왔다. 하지만 종래 보조동사로 인정되던 것들 중 -게와 연결되는 하다, 되다, 만들다의 경우 하다와 만들다는 사동 보조동사, 되다는 피동 보조동사로 각각 처리되어 왔으나 2.1절의 보조동사의 특성에 비추어 보아 하다와 만들다는 통사적 기능변화를 하며 하다는 대동사의 역할을 하고 있어 일반 보조동사와는 다른 성격을 보인다. 또한 만들다는 그 의미에 있어서 본동

(표 2) 한국어 보조동사와 의미

보조동사	연결어미	의미	보조동사	연결어미	의미
가다	아/어	진행	버리다	아/어	완료
가지다	아/어	지속	보다	아/어, 고	시도
계시다	고	진행	보이다	아/어	추정
나가다	아/어	진행	부치다	아/어	강세
나다	아/어, 고	완료	붙이다	아/어	강세
나오다	아/어	완료	빠지다	아/어	강세
내다	아/어	완료	싶다	고	원망
내려가다	아/어	강세	싫어하다	고	원망
놓다	아/어	지속	쌓다	아/어	강세
달다	아/어	봉사	오다	아/어	진행
대다	아/어	강세	있다	고	진행
두다	아/어	지속	자빠지다	아/어	강세
드리다	아/어	봉사	젖히다	아/어	강세
들다	고	강세	제끼다	아/어	강세
들어가다	아/어	강세	제치다	아/어	강세
말다	고	완료	주다	아/어	봉사
먹다	아/어	완료	지다	고	원망
바라다	기	원망	치우다	아/어	완료
바치다	아/어	봉사	터지다	아/어	강세
버릇하다	아/어	지속	하다	기(는)	시도

사로 쓰였을 때와 차이가 없다는 주장이 있다. 또한, -지와 연결되는 말다, 아니하다, 못하다의 경우 자립성이 있어 일반 보조동사와는 다른 성격을 보인다.

양상에 관한 영어의 개념그래프에서도 사동, 피동, 부정의 경우는 없다. 그리하여 본 연구에서는 영어와의 일관성을 위해 -게와 연결되는 하다, 되다, 만들다 그리고 -지와 연결되는 말다, 아니하다, 못하다를 보조동사에서 제외하기로 한다.

이들에 의거하여 본 논문에서는 각종 논문과 국어사전에서 보조동사들을 수집하고, 양상적 의미를 분석하였다. 표 2는 이렇게 수집된 40개의 보조동사와 선행 연결어미, 의미를 보인다.

표 2에서 의미를 보면 완료, 진행, 봉사, 시도, 강세, 원망, 지속, 추정의 8개가 존재하고 있다.

하지만 아직까지 개념그래프를 이용하여 한국어의 의미를 기술하려고 하는 연구는 거의 없는 실정이다. 본 논문은 개념그래프를 이용하여 한국어 보조동사를 형식적이고 명확하게 기술하고 이것들을 한/영 기계번역 시스템에서 이용하여 번역의 품질을 향상시키는데 목적이 있다.

### 3.3 보조동사의 개념그래프 기술

본 절에서는 3.2절 보조동사의 유형 및 의미에서 보인 완료, 진행, 봉사, 시도, 강세, 원망, 지속, 추정의 의미유형을 영어에서의 개념그래프를 참고하여 기술해 보고자 한다.

#### (1) 완료의 개념 그래프

Reichenbach는 영어에서 다양한 시제와 상에 관련된 시간 해석을 사건시(point of the event), 발화시(point of speech), 준거시(point of reference)로 설명한다[19]. 발화시는 주어진 문장이 발화되는 시점이고, 사건시는 문장과 관련된 사건이나 상태가 발생하는 시점이며, 준거시는 문장에 의해 나타난 시간으로 발화시와 같지 않을 수도 있다.



(그림 3) Reichenbach의 시제와 상

전통적으로 시제의 기능은 발화시와 사건시간의 관계라고 주장되어 왔으나, 근래에 들어 시제는 상을 통하여 간접적으로 규정되어진다는 주장이 받아들여지고 있다. 즉 상은 준거시와 사건시의 관계를 규정하며 시제는 발화시와 준거시의 관계라는 것이다. 그림 3은 준거시와 사건시의 상과 발화와 준거시의 시제의 관계를 보여 준다.

또한 시제와 상을 발화시(tS), 준거시(tR), 사건시(tE) 관계로 다음 표 3과 표 4에서와 같이 나타낼 수 있다. 표 3은 그림 3의 발화시와 준거시에 따른 시제의 제약조건이다. 여기서 <은 시간의 전후의 관계로 <의 좌측이 우측보다 이전이라는 뜻이며, =는 전후의 관계가 없이 같다는 뜻이며, >는 <의 반대이다.

(표 3) 시제의 제약조건

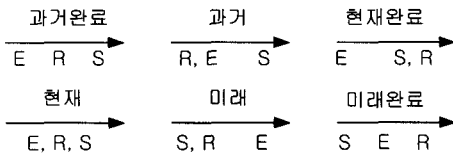
시제	제약 조건
과거(past)	tR<tS
현재(present)	tR=tS
미래(future)	tR>tS

표 4는 그림 3의 사건시와 준거시에 따른 상의 제약조건이다.

(표 4) 상의 제약조건

상	제약 조건
비완료(imperfective)	tE=tR
완료(perfective)	tE<tR

그리하여 영어에서의 과거완료, 과거, 현재완료, 현재, 미래, 미래완료를 그림으로 나타내면 다음 그림 4와 같이 나타낼 수 있다.



(그림 4) 영어의 시제와 상

그림 4에서 E는 사건시, R은 준거시, S는 발화시를 나타낸다. 시간의 방향은 화살표의 좌측 즉 과거에서 우측 즉 미래로 진행된다. 과거완료의 경우는 시간상으로 E가 나온 후 R이 나오고 다음에 S가 나온다는 뜻이며, 현재의 경우는 E, R, S가 모두 같다는 것, 미래의 경우는 S와 R이 같고 이후에 E가 나온다는 것을 나타내고 있다.

지금까지의 Reichenbach의 시제와 상에 관한 논의를 가지고 완료상을 개념그래프로 나타내면 (6)과 같다. (SUCC)은 전후의 개념인 영어의 successor의 의미이며 (PTIM)은 시점(point in time)을 의미한다.

(6)	PERFECT=[ [SITUATION: *x]->(PTIM) →[TIME:#e-time]->(SUCC) →[TIME:#r-time] ]
-----	---

표 2에서 보듯 완료(perfective)나 완수의 의미를 갖는 보조동사에는 나다, 나오다, 내다, -고 말다, 먹다, 버리다, 치우다 등이 있다. 다음 (7)은 완료의 의미인 나다의 예문과 개념그래프이다.

(2) 진행의 개념그래프

표 2에서 보듯 진행(progressive)의 의미를 갖는 보조동사에는 가다, 계시다, 나가다, 오다, -고 있다 등이 있다. 영어에서 진행은 상으로 표현하며 시제와 상의 의미를 둘 다 가지고 있어 시제의 지시 의미와 상의 내적인 함연 즉 implication (IMPL) 의미를 파악해야 한다. 앞으로 시제 표현은 간결성을 위해서 생략하도록 하겠다. 진행상은 어떤 상황 내부에 일시성(TEMPORARINESS), 순간성(INSTANTANEOUSNESS), 미완결성(INCOMPLETE-

TION), 과정(PROCESS)과 같은 여러가지 의미자질 특성 중 어느 하나를 가지고 있으며, 다음 (8)과 같이 진행상에 대해 개념그래프를 그릴 수 있다. 다음 (9)은 나가다의 개념그래프이다.

(3) 봉사(의)의 개념그래프

표 2에서 보듯 봉사(benefactive)의 의미를 갖는 보조동사에는 달다, 드리다, 바치다, 주다 등이 있다. 일반적으로 개념관계 beneficiary(BENF)를 이용하고 있으며 내용을 연결하는 개념관계 theme (THME)의 관계를 갖는다. 다음 (10)은 드리다의 개념 그래프이다.

(4) 시도의 개념그래프

표 2에서 보듯 시도(attemptive)나 시행의 의미를 갖는 보조동사에는 보다, -기(는) 하다 등이 있다. 영어의 TRY와 유사한 점이 많으며 일반적인 내용을 연결하는 개념관계 theme(THME)의 관계를 갖는다. 다음 (11)은 보다의 그에 대한 개념 그래프이다.

(5) 강세의 개념그래프

표 2에서 보듯 강세(emphasis)나 강조의 의미를 갖는 보조동사에는 내려가다, 대다, 들다, 들어가다, 부치다, 붙이다, 빠지다, 쌓다, 자빠지다, 짓히다, 제끼다, 제치다, 터지다 등이 있다. 일반적으로 어떤 문장을 강조하는 경우 개념 emphasis(EMPH)를 써서 표현할 수 있다.

다음 (12)은 강세의 의미인 자빠지다의 예문과 개념 그래프이다. 여기서 (STML)은 자극격(stimulus)에 대하여 갖는 감정을 기술한다.

(6) 원망의 개념그래프

표 2에서 보듯 원망(desirable)의 의미를 갖는 보조동사에는 바라다, -고 싶다, -고 싶어하다, -고 지다 등이 있다. 영어의 WANTING 유형의 양상과 유사한 점이 많다. 이런 동사는 보문에 의해서 기술되는 사건이나 상태에 대한 주어의 태도를

(7)	a.	나다: 철수는 밥을 먹고 나서 가게에 갔다.
	b.	[ [ [PERSON:철수 *x]←(AGNT)←[먹다]→(OBJ)→[밥] ]→(SUCC)→ [ [*x]←(AGNT)←[가다]→(LOC)→[가게] ] ]→(PTIM)→ [TIME:#e-time]→(SUCC)→[TIME:#r-time]
(8)		PROG=[ [SITUATION: *x]→(IMPL)→[TEMPORARINESS  INSTANTANEOUSNESS  INCOMPLETION  PROCESS ] ]
(9)	a.	나가다: 철수가 그 일을 수행해 나간다.
	b.	[[PERSON:철수]←(AGNT)←[수행하다]→(OBJ)→[일:#] ] →(IMPL)→[PROCESS].
(10)	a.	드리다: 철수는 선생님께 원고를 정서해 드렸다.
	b.	(PAST)→[[PERSON:철수 *x]←(AGNT)←[SERVE]→(THME)→ [PERSON:선생님]←(BENF)← [ [*x]←(AGNT)←[정서하다]→(OBJ)→[원고]]]
(11)	a.	보다: 영희는 옷을 입어 보았다.
	b.	(PAST)→[ [PERSON:영희 *x]←(AGNT)←[TRY]→(THME)→[ [*x]←(AGNT)←[입다]→(OBJ)→[옷]]]
(12)	a.	자빠지다: 그녀는 그 말에 놀라 자빠졌다.
	b.	(PAST)→[[PERSON:#그녀]←(EXPR)←[놀라다]→(STML)→[말:#]] →(IMPL)→[(EMPH)]
(13)	a.	-고 싶다: 나는 빵을 먹고 싶다.
	b.	[PERSON:#나 *x]←(EXPR)←[WANT]→(THME)→ [[*x]←(AGNT)←[먹다]→(OBJ)→[빵]]
(14)	a.	버릇하다: 그는 손가락을 물어 버릇한다.
	b.	[[PERSON:#그]←(AGNT)←[물다]→(OBJ)→[손가락]]→(IMPL)→ [ITERATIVITY]
(15)	a.	보이다: 철수가 영희를 좋아하는 것처럼 보인다.
	b.	[ [(PRBL)→[[PERSON:철수]←(EXPR)←[좋아하다]→(OBJ)→[PERSON:영희] ] ] →(DSTC)→[SHORT]→(FROM)→[CERTAINTY] ] ]→(CRTR)→[ASSUMPTION]

기술하기 때문에, 개념그래프에서 개념관계 experiencer (EXPR)와 theme(THME)을 요구한다.

다음 (13)은 원망에 관한 한국어 보조동사의 고 싶다는 예문과 그에 대한 개념 그래프이다.

(7) 지속의 개념 그래프  
표 2에서 보듯 지속(retention)이나 보유의 의미

를 갖는 보조동사에는 가지다, 놓다, 두다, 버릇하다 등이 있다. 진행과 달리 지속은 어떤 상태의 지속성(DURATION)과 어떤 상황 내부의 반복성(ITERATIVITY)과 같은 의미자질 특성 중 어느 하나를 가지고 있다.

다음 (14)은 지속의 의미를 갖는 버릇하다의 예문과 개념그래프이다.



(8) 추정의 개념그래프

표 2에서 보듯 추정(presumptive)의 의미를 갖는 보조동사에는 보이다와 ㄴ/는가, ㄹ가, 나, 구나, 다, 지, 면 뒤에 붙는 보다와 싶다 등이 있다.

인식적 양상은 화자의 지식과 믿음에 근거하기 때문에 사실판단이 아니다. 인식적 양상의 의미의 특성은 개연성에 있다. 개연성을 나타내기 위해서 [PROPOSITION]에 연결되는 개념관계 probable (PRBL)을 이용한다.

개연성의 정도는 현실성 즉 인식적 세계에서의 확실성으로부터 떨어진 거리에 의해 표시될 수 있다. 따라서, 여기에서는 개념관계 distance (DSTC)와 from(FROM)을 이용한다. 먼저, (DSTC)는 [PROPOSITION]을 [ATTRIBUTE]와 연결한다. 이를테면, 최고 유형에 속하는 것은 어떤 명제가 확실성으로부터 가장 짧은(shortest) 거리에, 고유형에 속하는 것은 짧은(short) 거리에, 중유형에 속하는 것은 중간(mid) 거리에 있다는 것을 표현한다. 그리고, (FROM)은 [ATTRIBUTE]와 이것의 기점 내지 기반이 되는 [ENTITY]를 연결한다. 인식적 양상에서는 확실성이 기반을 이루기 때문에, (FROM)은 [CERTAINTY]와 연결된다.

또한 추론의 근거를 표시하기 위해 일반적으로 개념관계 criteria(CRTR)을 이용한다. (CRTR)는 [-: \*x]를 [-: \*y]와 연결하는 개념관계로서, 이 때 \*y는 \*x의 근거가 된다.

다음 (15)은 추정에 관한 한국어 보조동사 보이다의 예문과 그에 대한 개념그래프이다

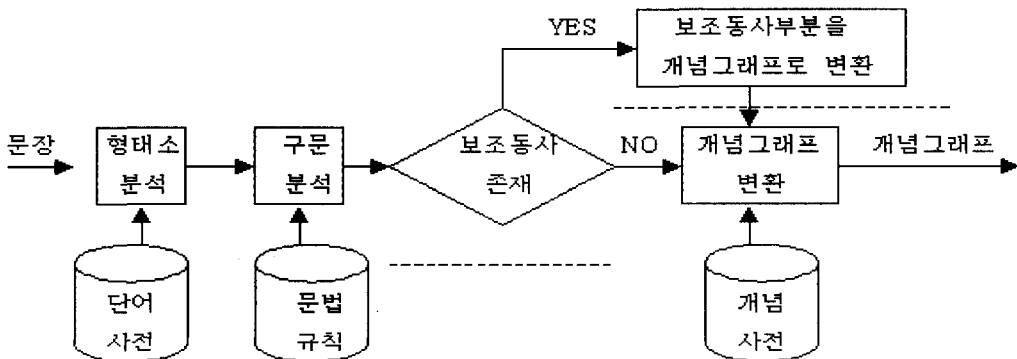
이상에서와 같이 개념그래프는 자연언어의 의미를 다루기 위해서 필요한 모든 정보를 통합하고 기술할 수 있는 지식표현 방법이며 의미의 속성이나 의미간의 논리적 관계는 물론 사용의미까지 표시할 수 있는 강력한 수단이기 때문에 한국어 보조동사의 다양한 의미를 명시적이고 체계적으로 기술할 수 있다.

4. 실험 및 결과

4.1 보조동사 처리기의 구현

영어권에서는 개념그래프의 구현과 응용에 관한 연구가 상당히 진행되고 있다[20]. 하지만 국내의 경우는 아직 구현과 응용에 관한 연구가 초보적인 수준이다. 본 논문에서는 C++언어를 이용하여 입력받은 문장에 대해서 보조동사인지를 검사하여 보조동사이면 개념그래프로 출력하는 프로그램을 구현하였다.

한국어 문장을 개념그래프로 변환하는 전체적인 시스템은 그림 5와 같으며 점선으로 된 부분이 본 논문의 주요 내용이다. 한국어 문장이 입력되면 형태소 분석과 구문 분석 단계를 거쳐 보조동사 부분을 식별하여 이 부분만을 개념그래프로



(그림 5) 한국어 문장을 개념그래프로 변환하는 단계

변환하여 일반적인 개념그래프로 변환하는 단계로 넘겨준다. 물론 어떤 문장이 형태소분석이나 구문분석에 실패하면 형태소 정보와 구문 정보를 이용하는 개념그래프 변환 단계도 실패하게 된다.

입력된 한국어 문장이 보조동사를 포함하고 있는지를 검사하는 방법은 2.1절의 보조동사의 특성을 이용하며 그 개략적인 알고리즘은 다음과 같다.

(단계 1) : 한국어 보조동사의 리스트와 각각의 의미(유형)가 수록된 테이블을 준비한다.

(단계 2) : <V1+V2>의 형태에서 보조동사가 되는 V2를 보조동사 테이블을 이용하여 검색하고 V1이 동사이면 연결어미를 검사한다. 검사하여 <V1+V2>이면 V2는 보조동사로 간주한다.

## 4.2 실험 및 결과 분석

본 실험에서 사용된 보조동사 테이블의 구조는 표 1을 따랐고, 표 2에서 보인 40개에 대해 보조동사와 그의 유형을 C++의 구조체로 선언하였다. 보조동사의 타입은 완료, 진행, 봉사, 시도, 강제, 원망, 지속, 추정인 8개를 갖는다.

본 논문에서 수집된 40개의 보조동사 중에서 각종 서적과 신문, 잡지, 교과서 등에서 조사해 본 결과 다음의 19개가 자주 사용되고 있었다. 표 5는 자주 사용되는 보조동사 19개와 [V-연결어미 Aux] 형태의 예문이다.

표 5에서와 같은 [V-연결어미 Aux] 형태의 처리 시에 본 논문에서 구현된 처리부에서 다음과 같은 문제점이 발견되었다.

- ① 보조동사가 한 문장에서 연이어 2개 이상

(표 5) 자주 사용되는 보조동사와 예문

보 조 동 사	예 문
가다	읽어 가다, 끝나치어 가다, 먹어 가다, 익어 가다, 쉬어 가다, 쉬어 가다
나가다	읽어 나가다, 풀어 나가다, 써 나가다
나다	자라 나다, 불어 나다, 깨어 나다, 살아 나다, 피어 나다, 자고 나다, 먹고 나다
내다	이기어 내다, 해 내다, 견디어 내다, 찾아 내다
놓다	갈겨써 놓다, 갈아 놓다, 받아 놓다, 붙들어 놓다
대다	먹어 대다, 울어 대다 불러 대다, 떠들어 대다, 졸라 대다
드리다	대접 해 드리다, 거들어 드리다, 보이어 드리다, 알리어 드리다
말다	떠나고 말다, 성공하고 말다, 해내고 말다, 망가지고 말다, 실패하고 말다
먹다	깨어 먹다, 잊어 먹다, 써 먹다, 해 먹다, 놀러 먹다, 읽어 먹다, 들어 먹다
버리다	먹어 버리다, 죽어 버리다, 잊어 버리다, 가 버리다
보다	먹어 보다, 가 보다, 해 보다, 들고 보다, 읽고 보다, 알고 보다
보이다	늘어 보이다, 되어 보이다
빠지다	썩어 빠지다, 낡아 빠지다, 말라 빠지다, 식어 빠지다, 늪어 빠지다
싫어하다	살고 싫어하다, h고 싫어 하다, 먹고 싫어 하다, 만나고 싫어 하다
오다	해 오다, 지키어 오다
있다	쓰고 있다, 하고 있다
주다	거들어 주다, 먹거 주다, 가 주다, 칭찬해 주다
터지다	불어 터지다
하다	만나 보기는 하다, 먹기는 하다

나타날 때, 예를 들면 ‘깨어 나고 있다.’는 보조동사 예문의 처리 시에 ‘깨(V)-어(연결어미) 나(Aux)-고(연결어미) 있다.’로 분석이 되는 데 이 때 깨어 나다의 완료의 의미와 -고 있다의 진행의 의미가 각각 나온다. 이 같은 문장은 완료와 진행의 두 의미 중에서 진행의 의미가 강하다.

- ② 형태소 분석시 축약과 불규칙 처리가 정확해야 한다. 씨 주다에서 씨를 쓰+어의 형태로 처리되어야 하며 가 버리다도 가(V) 버리다(Aux)로 잘못 처리되기 쉽다.

다음으로, 한국어 보조동사가 포함된 문장을 3.3절에서와 같은 개념그래프로 변환해 주는 개략적인 알고리즘은 다음과 같다.

(단계 1) : V2가 보조동사이면 그 유형을 반환 받는다.

(단계 2) : 리턴 받은 유형에 맞게 각각에 해당하는 개념 그래프를 출력한다.

예를 들어, 철수는 책을 읽어 내려갔다라는 문장이 입력되어 형태소분석을 거치면 철수+는 책+을 읽(다)+어 내려가+았+다이 되고, 구문 분석을 거치면 구문트리가 생성되며, 형태소분석 된 문장에서 보조동사를 찾게 된다. 보조동사 부분인 내려가다가 강세의 의미를 가지고 있으므로 [철수+는 책+을 읽다](IMPL)[EMPH]가 되며, 철수+는 책+을 읽다는 일반적인 개념그래프 변환단계에서 [철수](AGNT)[읽다](OBJ)[책]로 변환된다.

그림 6은 본 논문에서 구현한 입력된 한국어 문장이 보조동사가 포함된 경우 보조동사의 의미를 보여주는 보조동사 처리기의 결과이다. 그림 1과 같이 보조동사가 포함된 문장이 입력되면 보조동사 처리기는 원문을 해석하여 그림 6과 같이 입력 문장의 의미를 보여 준다. 원망의 의미를 갖는 (13)의 ‘나는 빵을 먹고 싶다.’는 문장이 입력

되면 위의 알고리즘을 적용하여 원망의 의미를 명확히 알려주고 있다.

보조동사 처리기는 한국어에서 보조동사의 수가 40개정도 밖에 되지 않아 관리 및 검색이 용이하며 메모리 상에 올리기에 부담이 적다. 본 논문의 보조동사 처리기를 통하여 얻는 장점은 다음과 같다.

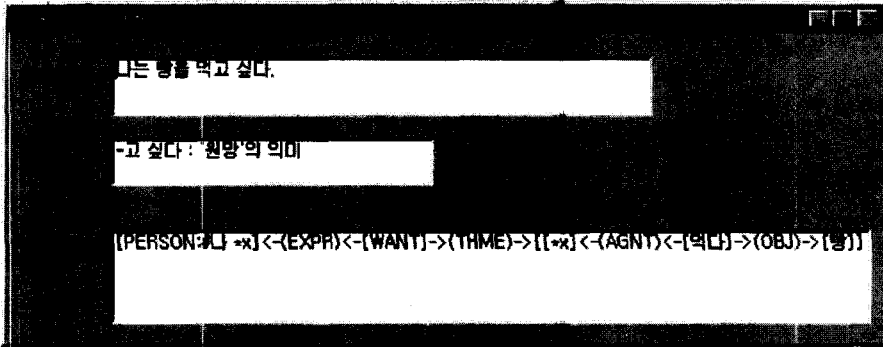
- ① 분석 단계에서 보조동사를 인식함으로써 분석은 물론 번역의 효율도 향상된다.
- ② 분석 후 복잡한 구조 변환 과정이 필요치 않아 진다.
- ③ 번역된 문장의 표현이 정확하고 자연스럽다.

이와 같은 장점들은 한/영 기계번역에서 번역의 정확성 향상에 기여할 수 있다.

## 5. 결 론

본 논문의 목적은 기계번역 시스템을 이용한 번역에서 자주 틀리는 한국어 보조동사를 개념그래프를 이용해 표현하고 이들을 기계번역에서 이용할 수 있도록 하는 데 있다. 화자나 주체의 심리 상태를 나타내는 한국어 보조동사를 각종 논문과 국어사전에서 40개를 수집하여 그 예문을 들고, 완료, 진행, 봉사, 시행, 강세, 원망, 지속, 추정의 8가지 의미로 분류하고 이들을 개념그래프로 기술해 보았다.

또한 실험에서는 이렇게 분류된 한국어 보조동사의 의미를 C++를 이용하여 한국어 보조동사가 포함된 문장을 입력받아 개념 그래프로 변환하는 프로그램을 구현하였다. 그 수가 40개 정도밖에 안되지만 한국어 문장에서 자주 쓰이는 보조동사를 처리하는 보조동사 처리기는 관리 및 검색이 용이하며 메모리 상에 올리기에 부담을 주지 않으면서도 기계번역에서 쓰일 경우 번역의 질과 효율을 높일 수 있다.



(그림 6) 본 논문에서 구현된 시스템의 결과의 예

마지막으로 개념 그래프는 자연언어의 의미를 다루기 위해서 필요한 모든 정보를 통합하고 기술할 수 있는 지식표현 방법이며 의미의 속성이나 의미간의 논리적 관계는 물론 사용 의미까지 표시할 수 있는 강력한 수단이기 때문에 한국어 보조동사의 다양한 의미를 명시적이고 체계적으로 기술할 수 있었다.

### 참고 문헌

- [1] 김나리, 패턴정보를 이용한 한국어 구문분석, 서울대 컴퓨터공학과 박사학위논문, 1997.
- [2] 이경순, 영/한 기계번역에서 분석/변환 규칙의 자동 확장, 한국과학기술원 전산학과 석사학위논문, 1997.
- [3] 윤성희, 영어/한국어 기계번역을 위한 속어 기반의 효율적 문장분석, 서울대 컴퓨터공학과 박사학위 논문, 1993.
- [4] John F. Sowa, *Conceptual Structures: Information Processing in Mind and Machine*, Addison-Wesley, 1984.
- [5] 강홍구, 국어 보조동사의 통사의미론적 연구, 충남대 국어국문학과 박사학위논문, 1999.
- [6] 남기심, 고영근, 표준 국어문법론, 탑출판사, 1993.
- [7] 이관규, 국어 보조동사 연구, 고려대 국어국문학과 석사학위논문, 1987.
- [8] 이선웅, 현대 국어의 보조용언 연구, 서울대 국어국문학과 석사학위논문, 1995.
- [9] 남미혜, 국어의 연속동사구성의 연구, 서울대 국어국문학과 박사학위논문, 1996.
- [10] 손세모들, 국어 보조용언 연구, 한국문화사, 1996.
- [11] John F. Sowa, *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks/Cole, 2000.
- [12] 배우정, 박인철, 이용석, 개념 구조의 효율적인 표현을 위한 개념 트리, 한국정보과학회 논문지(B), 제26권 제6호, pp.822-832, 1999.
- [13] 이휘봉, 이종혁, 이근배, 구문의존구조에서 개념그래프 생성을 위한 한국어의 의미분석, 한국정보과학회 봄 학술발표논문집, Vol.24, No.1, pp.463-466, 1997.
- [14] 박인철, 배우정, 안동연, 이용석, 한국어 문형에 기반한 개념 그래프의 생성, 한국정보과학회논문지(B), 제25권 제3호, pp.596-607, 1998.
- [15] 양기철, 개념그래프 기반 문서인식 후처리 기법, 한글 및 한국어정보처리 학술대회, pp. 211-214, 1994.
- [16] 서상욱, 영어개념구조연구, 형설출판사, 1996.
- [17]곽선연, 영어 양상표현의 개념구조, 충남대 영어영문학과 박사학위논문, 1997.

[18] 이창학, 영어 시제/상의 개념구조적 기술, 충남대 영어영문학과 박사학위논문, 1995.  
[19] T. Ogihara, Tense, Attitudes, and Scope, Kluwer Academic Publishers, 1996.

[20] John F. Sowa and Eileen C. Way, Implementing a Semantic Interpreter Using Conceptual Graphs, IBM Journal of Research and Development, Vol.30, No.1, 1986.

### ● 저 자 소 개 ●



#### 이 병 희

1992년 충남대학교 컴퓨터공학과 졸업(학사)

1994년 충남대학교 대학원 컴퓨터공학과 졸업(석사)

1997년 충남대학교 대학원 컴퓨터공학과 박사수료

1997~현재 충남대학교 부설연구소 연구원(병역특례)

관심분야 : 자연어처리, 한국어정보처리, 인공지능, 패턴인식, 프로그래밍언어

E-mail : bhlee@ce.cnu.ac.kr