기계 번역 의미 대역 패턴을 이용한 한국어 복합 명사 의미 결정 방법

한국전자통신연구원 음성/언어연구센터 언어처리연구팀,[1] 연세대학교 정보기술학부[2]
양성일[1] · 김영길[1] · 박상규[1] · 나동렬[2]

A Method of Word Sense Disambiguation for Korean Complex Noun Phrase Using Verb-Phrase Pattern and Predicative Noun

Seong-Il Yang,[1] Young-Kil Kim,[1] Sang-Kyu Park,[1] Dong-Yul Ra[2]
NLP Team,[1] Speech/Language Technology Research Center, ETRI, Seoul
Div. of I.T.,[2] Yonsei University, Seoul, Korea

요 약

한국어의 언어적 특성에 의해 빈번하게 등장하는 명사와 기능어의 나열은 기능어나 연결 구문의 잦은 생략현상에 의해 복합 명사의 출현을 발생 시킨다. 따라서, 한국어 분석에서 복합 명사의 처리 방법은 매우 중요한 문제로 인식되었으며 활발한 연구가 진행되어 왔다. 복합 명사의 의미 결정은 복합 명사구 내 단위 명사간의 의미적인 수식 관계를 고려하여 머리어의 선택과 의미를 함께 결정할 필요가 있다. 본 논문에서는 정보 검색의 색인어 추출 방법에서 사용되는 복합 명사구 내의 서술성 명사 처리를 이용하여 복합 명사의 의미 결정을 인접 명사의 의미 공기 정보가 아닌 구문관계에 따른 의미 공기 정보를 사용하여 분석하는 방법을 제시한다. 복합 명사구 내에서 구문적인 관계는 명사구 내에 서술성 명사가 등장하는 경우 보-술 관계에 의한 격 결정 문제로 전환 할 수 있다. 이러한 구문 구조는 명사 의미를 결정할 수 있는 추가적인 정보로 활용할 수 있으며, 이 때 구문 구조 파악을 위해 구축된 의미 제약 조건을 활용하도록 한다. 구조 분석에서 사용되는 의미 격틀 정보는 동사와 공기하는 명사의 구문 관계를 분석하기 위해 의미 정보를 제약조건으로 하여 구축된다. 이러한 의미 격틀 정보는 단문 내 명사들의 격 결정과 격을 채우는 명사 의미를 결정할 수 있는 정보로 활용된다. 본 논문에서는 현재 개발중인 한영 기계 번역 시스템 Tellus-KE의 단문 단위 대역어 선정을 위해 구축된 의미 대역 패턴인 동사구 패턴을 사용한다. 동사구 패턴에 기술된 한국어의 단문 단위 의미 격 정보를 사용하는 경우, 격 결정을 위해 사용되는 의미 제약 조건이 복합 명사의 중심어 선택과 의미 결정에 재활용 될 수 있으며, 병렬 말뭉치에 의해 반자동으로 구축되는 의미 대역 패턴을 사용하여 데이터 구축의 어려움을 개선하고자 한다.

서 론

자연어 처리에 있어 명사의 의미 모호성 해소 문제는 언어 처리의 가장 어려운 문제들 가운데 하나로 비교적 오랜 시간동안 연구되어 왔다. 한국어에서는 기능어의 발달이라는 특성에 의해 단독 품사보다 명사와 기능어가 결합된 형태가 문장 성분의 대부분을 차지한다. 하지만 기능어나 연결 구문의 잦은 생략 현상으로 문장 내 복합 명사의 사용이 빈번하여지고 반면에 자유로운 띄어쓰기나 어순의 특성은 복합 명사 처리의 어려움을 가중시킨다. 따라서 한국어 분석에서의 복합 명사 처리는 매우 중요하다고 볼 수 있으며, 최근에는 매우 활발히 연구되고 있는 분야이다.

복합 명사의 의미는 정보 검색의 색인어 추출에서부터 구조 분석을 위한 격 결정 정보, 구묶음을 위한 처리에 이르기까지 한국어 정보의 가공에서 그 비중이 매우 큰 부분이다. 특히 기계 번역의 경우, 복합 명사의 의미 결정은 대역어의 선정과 직결된 문제로 번역 시스템 전체의 질을 좌우하므로 효율적인 문제 해결이 절실한 분야이다. 한영 기

E-mail : siyang@etri.re.k
E-mail : kimyk@etri.re.kr
E-mail : parksk@etri.re.kr
E-mail : dyra@dragon.yonsei.ac.kr

계 번역의 경우 한국어 복합 명사의 대역어를 선정할 때, 일대일 대응의 대역어 선정은 곧 부자연스러운 표현으로 연결되는 경우가 많다. 따라서 가능한 추가적인 복합 명사구의 의미 관계를 파악하여 대역어를 선정하는 방법이 필요해 진다.

문장 내에서 명사의 의미 결정은 구문 분석을 수행하면서 단문 단위로 공기하는 동사에 의해 의미 격틀의 비교를 수행하여 비교적 쉽게 결정될 수 있다. 예를 들어 다음과 같은 형태의 문장에서 "새"의 의미는 동사가 취하는 의미 정보를 통해 구분할 수 있다.

새가 하늘을 날고 있다.
⇔ [조류]가 [공간]을 날!다.

단문의 분석은 동사와 공기하는 명사의 의미 정보를 격 결정을 위한 제약 조건으로 하여 문장의 구조를 파악하며, 이때 명사의 의미가 결정될 수 있다. 따라서, 복합 명사의 의미를 파악하여 기계 번역용 대역어를 산정하기 위해 가장 용이한 방법은 사전에 의미 정보와 대역어를 기술하여 복합 명사를 등록하는 방법이다. 문장에서 해당 복합 명사가 등장하는 경우, 구문 구조 분석의 의미 격틀에 의해 판단된 명사의 의미와 대역어는 비교적 쉽게 결정될 수 있다. 그러나 사실상 복합 명사의 조합수는 무한하므로 복합 명사의 머리어 선택과 의미 결정은 복합 명사구 내의 단위 명사간 의미 관계를 통해 결정되어야 한다. 따라서 복합 명사구의 내부 구조 및 의미 관계는 해당 단위 명사와 주변 어휘들과의 공기 관계를 파악하여 결정하는 방법이 연구되어 왔다.

현재까지의 연구에서는 주로 의미 말뭉치를 사용하여 추출된 인접 명사와의 의미 공기 정보나 전자 사전의 뜻풀이 용례를 사용하여 구축된 의미 정보를 이용하는 방법 등이 제안된바 있으나, 의미 정보를 기술한 데이터의 구축은 많은 비용을 요구한다는 어려움이 발생한다.

본 논문에서는 정보 검색에서의 색인어 추출을 위해 사용되는 서술성 명사의 처리 기법을 통해 복합 명사구 내 단위 명사들의 언어학적 관계를 판단하여 공기하는 단위 명사간의 의미 관계를 파악하고자 한다. 이러한 언어학적 관계는 구문 구조 분석과 동일한 방법으로 머리어의 선택과 의미 분석을 수행할 수 있도록 한다. 서술성 명사가 발견되는 경우, 서술성 명사를 용언의 역할을 수행하는 머리어로 간주할 때 두 명사간에는 보-술 관계가 성립한다고 볼 수 있다. 따라서, 명사의 의미 해소는 인접 명사의 의미 관계가 아닌 명사-동사간의 격 결정 문제로 전환하여 판단할 수 있다. 예를 들어 다음과 같은 복합 명사구의 의미 구조는 서술성 명사에 의한 구문 구조로 전환하여 파악할 수 있다.

고속도로 건설 현장 보도 내용 재분석 의뢰
⇔ [건축물]을 건설하!다
⇔ [장소]를 보도하!다
⇔ [지식]을 분석하!다

구조 분석의 격 결정을 위해 구축한 의미 격틀 정보를 사용하는 경우, 구문 분석과 동일한 분석 방법으로 복합 명사 내 단위 명사간의 의미 관계를 분석 할 수 있다. 그러나 의미 격틀의 구축은 또한 많은 구축 비용과 시간을 필요로 하게 된다. 본 논문에서는 현재 개발 중인 기계 번역 시스템 Tellus-KE에서 사용되는 단문 단위 의미 대역 패턴인 동사구 패턴을 활용한다. 동사구 패턴의 한국어 부분에서 단문 단위 의미 제약을 추출하는 경우 복합 명사 의미 해소에 사용될 수 있으며, 동시에 영문 대역어 부에 의해 대역어를 선정할 수 있는 방안을 제시하고자 한다. 복합 명사구는 관형격 조사의 사용을 포함한 명사의 나열로서 정의될 수 있지만, 본 논문에서는 서술성 명사의 쓰임에 의해 결정되는 의미 결정 방법을 효과적으로 보이기 위해 조사 사용을 배제하고 서술성 명사가 등장하는 복합 명사의 처리를 대상으로 하였다.

현재 400여 개의 의미 코드를 사용하여 구축된 의미 대역 패턴인 동사구 패턴은 약 115,000개로 일반 문장을 상대로 단문 단위 대역어를 지정하기 위해 사용된다. 동사구 패턴은 병렬 말뭉치를 사용하여 반 자동으로 구축되며, 동사구 패턴의 의미 제약 정보를 사용하여 몇 가지 의미 중의성을 갖는 명사에 대해 실험해 본 결과, 서술성 명사와 공기하는 경우 의미 모호성을 해소하는데 효과적이며, 기존 의미 모호성 해소 방법과 함께 사용될 수 있음을 보여준다.

관련연구

어휘의 의미 모호성을 해소하기 위한 방법으로는 말뭉치를 가공하여 어휘 정보와 거리 계산에 의해 결정하는 방법,[1,2] 시소러스를 사용한 방법,[3] 전자 사전의 뜻풀이 정보를 사용하는 방법,[4] 명사의 의미 정보와 접사 처리를 이용하는 방법,[5] 어휘 의미 분석 말뭉치를 이용한 방법[6] 등이 제안되고 있다. 아울러 명사 의미 결정을 2단계로 분리하여 명사구 내의 인접 명사간 의미 결정 정보와 명사구 머

리어의 명사-동사간 의미 결정 정보를 분리하여 적용하는 방법[7]을 통해 구문적 연관 관계에 따라 문맥 확률을 적용하고자 하는 시도가 이루어 지고 있다.

명사간 어휘 공기 정보와 서술성 명사를 이용하여 명사구 분석을 수행하는 방법[8]은 명사의 중심어 선택 시 보-술 관계의 격을 채우는 명사는 중심어 선택에서 제외될 수 있음을 보여준다.

아울러, 명사 의미 부착 말뭉치와 의미 체계 구축을 위한 작업으로는 세종 계획에 의한 국립 국어 연구원에서 의미 주석 말뭉치를 구축하는 작업이 있으며, 연세대학교에서는 연세 한국어 사전에 의한 의미 주석 작업을 시도하고 있다.

명사 의미 체계

주로 온톨로지(Ontology) 구축으로 대표되는 명사의 의미 분류는 개념적 기준의 판단이 모호하여 구축이 어렵다는 문제[9]와 개념 기준을 사용하는 응용 시스템의 특성에 따라 분류가 다르게 필요하다는 문제[7]가 있다. 본 논문에서는 서술성 명사와의 공기 관계를 잘 나타낼 수 있도록 한영 기계 번역 시스템인 Tellus-KE를 위해 분류된 명사 의미 체계를 사용한다. Tellus-KE에서 사용하는 기계 번역의 최소 의미 단위는 하나의 동사를 취하는 단문으로 간주되고, 최소 대역어 산정 기준은 단문으로 취급된다. 따라서 명사의 의미 코드 체계는 공기하는 동사와의 관계에 따라 분류 한다. 예를 들어 "악기"의 세부 분류는 악기를 다루는 동사와 공기하기 위해 "관악기", "현악기", "타악기" 등으로 구분한다. 정보 검색 분야에서 "서양악기", "동양악기"로 분류되는 것과 비교하면 그 특성을 알 수 있다.

Table 1은 Tellus-KE에서 사용되는 의미 코드 체계의 일부를 보여준다.

의미 코드는 총 8개 레벨 414개의 의미 노드를 사용하여 할당되며, 총 350개의 말단 노드(Leaf Node)를 갖는다. 명사의 의미 정보는 하나의 어휘 정보에 대해 여러 개의 의미 코드가 할당될 수 있으며, 형태소 사전에 등록되어 사용된다. Fig. 1은 의미 코드가 부착되어 형태소 사전에 등록된 명사의 등록 정보 예제이다.

형태소 사전에 등록된 명사 엔트리의 첫번째 숫자는 형태소 코드를 나타내는 것으로 일반 명사와 서술성 명사로 구분되어 제공된다. 의미 코드가 부착된 명사는 의미 코드에 따른 대역어와 함께 등록되며, 약 13만여 개의 엔트리가 구축되어 사용되었다.

동사구 패턴

단문이 갖는 의미는 중심어인 동사와 동사에 연결되어 격 정보를 채우는 명사가 갖는 의미로 파악할 수 있다. 한국어는 격조사와 함께 쓰이는 명사의 구문적 역할에 의해 격 정보를 결정할 수 있으며, 명사가 갖는 의미는 중심어인 동사에 해당 격 정보를 채우며 연결이 가능한지 여부를 결정하는 제약조건이 된다.

동사구 패턴은 의미 제약 조건을 기술하기 위하여 동사, 조사, 명사의 의미쌍을 사용하여 제약 조건을 기술하며, 명사-동사 쌍에 국한하지 않고 단문 단위의 제약 조건을 함께 기술한다. 이때 기술되는 동사는 서술성 명사를 취하는 용언을 포함한다(예 : 지각하다, 공부하다 …). 본 논문에서는 단문 단위의 모든 제약 조건 집합을 단문 제약 조건으로 간주하고, 단문 제약 조건에 의해 결정된 단문 단위 대역어를 기술한 대역 패턴을 동사구 패턴으로 정의한다. 동사구 패턴의 구축은 병렬 말뭉치를 사용하여 대역어에 의해 산정되는 의미 코드를 이용하여 반자동으로 구축된다. 의미제약은 4,14개의 의미 코드를 사용하며 같은 체계로 구축된 형태소 사전의 의미 코드와 비교하여 명사 의미를

```
120022 538 밤 1
{
[ ( EROOT night ) ( SEM 시간 ) ]
[ ( EROOT chestnut ) ( SEM 식품 ) ]
}
```

Fig. 1. 의미 코드 부착 사전.

Table 1. 명사 의미 체계

레벨 1	레벨 2	레벨 3	레벨 4	레벨 5	레벨 6	레벨 7	레벨 8	예
구체명사	물체	무생물	인공물	교통기관	육상교통기관	자동육상교통기관	자동차	버스, 자동차, 스포츠카
				폐기물	매연			매연, 대기오염물
		생물	동물기관	신체부분				몸, 팔, 팔꿈치, 뱃살
			식물	채소				배추, 무, 당근
추상명사	성상	크기						체온, 인구, 면적
		색깔						초록, 노랑, 회색
활동명사	경제활동	상업활동						매매, 할인, 세일
	종교활동							기도, 예배, 설교, 찬불, 숭배

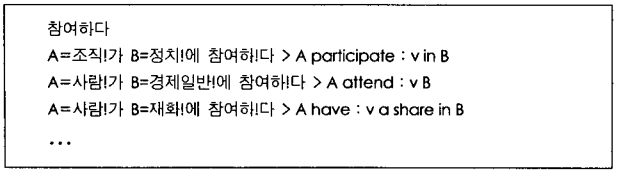

Fig. 2. 한영 동사구 패턴.

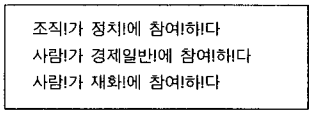

Fig. 3. 서술성 명사와 일반 명사의 의미 제약.

결정할 수 있도록 동사구 패턴에 기술하여 사용한다.

Fig. 2는 동사구 패턴의 예를 보여준다. 동사구 패턴은 크게 한국어 의미 제약 조건부와 대역어 표기부로 나뉘며, 의미 제약 조건부는 단문 내 의미구조를 표현하기 위해 동사가 취할 수 있는 격조사, 격 정보를 채우는 명사의 의미와 동사의 의미구분으로 크게 구분하여 기술한다.

명사 의미 모호성 해소를 위해 필요한 부분은 서술성 명사가 동사로 등록된 동사구 패턴에서 대역어 부분을 제외한 명사-동사 의미 공기 쌍이 되므로 필요 정보만을 가공하면 Fig. 3과 같은 의미 공기 제약 조건을 얻는다. 이때 조사는 복합 명사 의미 해소에 사용되지는 않지만, 서술성 명사의 격 결정 정보를 나타내므로 굳이 삭제하지는 않는다.

중의적인 의미를 갖는 동사는 동사가 격 정보로 취하고 있는 명사들의 의미로 결정될 수 있다. 따라서 일반 명사의 의미와 서술성 명사의 의미는 상보 관계에 의해 결정될 수 있다.

복합 명사 의미 결정

의미 전달을 위한 최소 단위를 단문으로 생각하였을 때 명사의 의미는 구조 분석 시 사용되는 의미 격틀에 의해 동사와 함께 결정하는 방법이 용이하지만, 용언과의 관계만으로 명사의 의미를 결정하기가 어려운 경우가 발생할 수 있다. 예를 들어, 대표적인 의미 중의성을 갖는 명사 "배"를 고려할 때, "배를 살펴 보았다"와 같이 의미 구분이 모호한 경우가 발생할 수 있다. 그러나, "백화점 시식용 배를 살펴 보았다"와 같이 서술성 명사와 공기하는 복합 명사를 사용하는 경우, 추가적인 의미 결정 방법이 가능해진다. 용언과 별개로 복합 명사만이 등장하는 경우도 동일하게 의미 모호성을 해결할 방법이 필요하다.

현재까지의 복합명사 의미 모호성 해결 방법은 자매 성분인 인접 명사들과의 의미 공기 정보를 분석하는 것으로

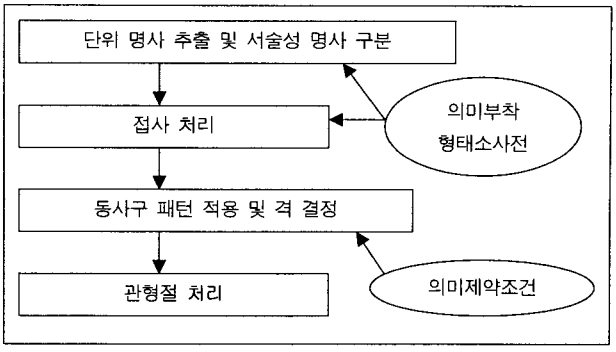

Fig. 4. 복합 명사 의미 결정 단계.

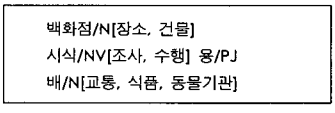

Fig. 5. 형태소 분석 결과.

의미 결정을 시도하였다. 그러나 복합 명사는 대부분 단문을 구성할 수 있는 의미적 성격으로 연결되어 생성된다는 점에 착안하여 등장하는 서술성 명사를 용언으로 간주하면 구문 분석과 유사한 방법으로 복합 명사의 의미를 결정할 수 있다. 이때 사용되는 지식은 구문 분석에서 사용하는 의미 격틀 정보를 재사용한다.

복합 명사의 의미 결정을 위해서는 크게 단위 명사 처리, 접사 처리, 의미 제약 조건 비교, 관형절 처리의 단계를 수행한다. Fig. 4는 복합 명사 의미 결정 단계를 나타낸 것이다.

1. 단위 명사 처리

분석을 수행하고자 하는 복합 명사는 우선 단위 명사로 분해되어 서술성 명사가 등장하는 여부를 판단할 수 있는 형태로 가공되어야 한다.

본 논문에서는 Tellus-KE 시스템에서 사용하는 형태소 분석기의 결과에 의해 분해된 단위 명사와 서술성 명사의 태깅 결과를 입력값으로 한다. Fig. 5는 형태소 분석에 의해 생성된 결과를 보여준다.

2. 접사 처리

복합 명사구 내에 등장하는 단위 명사들은 접사와 함께 쓰이는 것을 빈번하게 볼 수 있다. 상기 예 "시식용 배"에서 알수 있듯이 서술성 명사 "시식"을 얻기 위해서는 접사 "용"을 제거하여 비교할 수 있어야 한다. 접사는 의미를 한정하기 위한 추가적인 정보로 이용이 가능하다.

3. 의미 제약 조건 비교

단위 명사 분해와 접사 처리를 거친 복합 명사구는 서술성 명사와 일반 명사의 나열로 볼 수 있다. 서술성 명사를

용언인 v로, 서술성 명사 앞에 등장하는 단위 명사들을 의미 결정을 위한 명사 리스트 $w_{1,n}$으로 표기할 때, 이 서술성 명사에 대한 의미 확률을 최대로 하는 명사들의 의미 리스트 $s_{1,n}$을 구하는 함수 WSD는 동사와 공기하는 명사의 의미 결정 수식과 동일하게 다음 수식으로 나타낼 수 있다.

$$
\begin{aligned}
& WSD_{vp}(w_{1,n}, v) \\
&= \arg\max_{s_{1,n}} P(S_{1,n} = s_{1,n} | W_{1,n} = w_{1,n}, v) \\
&= \arg\max_{s_{1,n}} P(s_{1,n} | w_{1,n}, v) \\
&= \arg\max_{s_{1,n}} \frac{P(s_{1,n}, w_{1,n}, v)}{P(w_{1,n}, v)} \\
&\approx \arg\max_{s_{1,n}} P(s_{1,n}, w_{1,n}, v)
\end{aligned}
$$

서술성 명사와 일반 명사의 의미 결정은 병렬 말뭉치에서 적용되는 동사구 패턴의 발생 빈도에서 수집된 의미 공기 정보를 사용한다. 아울러, 의미 공기 정보는 이진 관계 정보가 아닌 단문 단위 제약 조건으로 이루어 지므로 해당 동사구 패턴과 비교될 때 단문 단위 완전 매칭인지 부분 매칭인지를 판단하여 매칭되는 논항에 가중치를 부여할 수 있다.

의미 격 정보의 비교를 수행할 때, 다른 명사에 의해 이미 채워진 격 정보는 의미 제약 검사에서 제외한다.

복합 명사 머리어의 선택에 있어 서술성 명사와 일반 명사의 공기가 없는 경우는 기본 값으로 뒤에 등장한 명사를 머리어로 하는 수식 관계를 부여할 수 있다. 예를 들어 "자주 독립 선언"의 "자주 독립"과 같이 서술성 명사 "독립하다"와 공기하는 "자주"의 의미 제약(유생물/무생물 상태)이 없는 경우 기본값으로 "독립"을 머리어로 삼아 두 명사를 하나의 일반 명사로 취급하며, 이 경우 "[상태]를 선언하!다"와 같은 동사구 패턴을 이용할 수 있다.

4. 관형절 처리

서술성 명사 뒤에 나타나는 일반 명사는 용언에 의해 꾸밈을 받는 피수식어로서 관형절의 격 성분을 채울 수 있다. 따라서 서술성 명사에 의해 결정되는 격 성분 중 필수격을 우선으로 뒤에 나타나는 일반 명사의 격 정보를 비교하여 결정하여야 한다. 예를 들어 상기 예에서 "시식"이란 서술성 명사는 뒤에 등장하는 "배"를 꾸미는 관형절의 용언으로 취급하여 "배"가 "시식하다"의 격 성분을 채우게 된다.

피수식어에 의해 관형절의 격 성분으로 채워지는 경우 서술성 명사 앞에 나오는 명사의 격 성분 검사는 채워진 필수격을 제외하고 검사하여야 한다.

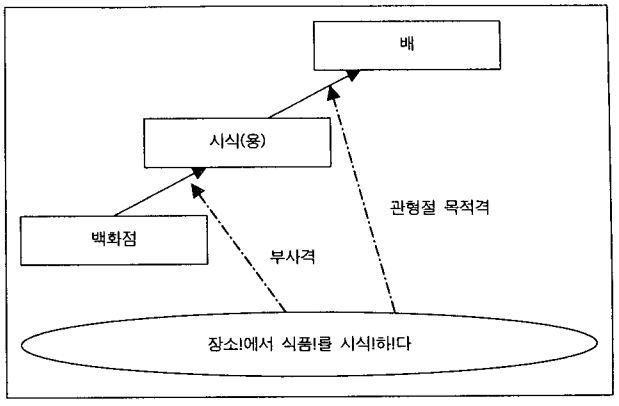

Fig. 6. 복합 명사 의미 결정.

Table 2. 중의성 해소 실험 결과

대 상	발생 빈도	정답 의미	정확률
세 제	482	408	84.65%
수 도	518	393	75.86%

최종적으로 "백화점 시식용 배"라는 복합 명사에 대한 의미 제약 정보의 적용은 "백화점에서 배를 시식하다"의 형태로 Fig. 6에서 나타낼 수 있다.

실 험

중의성을 가지는 명사가 복합 명사구에서 서술성 명사와 함께 등장하는 경우를 한정하여 의미 모호성 해소가 효과적인지 실험하여 보았다.

복합 명사가 빈번하게 발생되는 방송 뉴스 문장과 신문 말뭉치를 사용하였다. 의미적 중의성을 갖는 단어 "세제"와 "수도"가 서술성 명사와 함께 등장하는 경우 총 1000건의 복합 명사 발생 빈도에서 서술성 명사에 의한 의미 제약 조건을 만족시키는 관계를 통하여 각각 84.65%와 75.86%의 의미 분별 정확도를 보였다.

앞으로 추가적인 격틀 구축과 추가적인 실험을 통해 서술성 명사의 사용에 따른 명사 중의성 해결을 확장하고자 충분한 데이터 수집과 실험을 병행하고 있다.

결론 및 향후 연구 과제

한국어의 특성에 의해 문장 성분의 대부분을 차지하는 명사들의 사용은 빈번한 기능어의 생략 현상으로 많은 복합 명사의 출현을 야기시킨다. 구문 구조 분석에 의해 명

사의 의미는 문장에서 공기하는 동사와 동사가 취하는 격 정보 의미 제약 조건으로 분석을 수행 할 수 있다. 그러나 복합 명사구 내에 등장하는 단위 명사간의 의미 결정은 인접 명사와의 의미 공기 정보에 의존하여 파악하여 왔다.

본 논문에서는 서술성 명사와 격 결정 의미 제약 조건을 사용하여 복합 명사내의 단위 명사 간 의미 관계를 파악하는 방법을 제안하였다. 서술성 명사를 사용하는 경우 복합 명사 내 단위 명사의 의미 결정 문제는 구문 구조의 격 결정 문제로 전환되며, 의미 격틀에 의해 기술된 의미 코드는 해당 격 성분의 명사 의미를 파악할 수 있도록 한다.

향후 연구에서는 서술성 명사의 사용에 따라 다르게 나타날 수 있는 언어적 현상을 파악하여 확장하는 것이 필요하다. 본 논문에서 제안한 서술성 명사를 "하다"와 결합하는 경우에만 국한 하지 않고 "되다, 보다"등으로 확장하며, "시키다"와 같이 사/피동형에 따른 의미 제약 적용의 비교 현상을 파악하는 것도 요구된다. "연구원 근무 규정"과 같은 예에서 "연구원"의 의미가 "사람"과 "조직" 중 어떤 의미를 나타내는 지는 서술성 명사 "근무"로 구분되지 않으므로 여전히 중의성이 해소되지 못한 부분이다. 아울러, "학생용 지도"와 "학생 지도실"과 같이 서술성 명사의 태깅과 세부적인 접사 처리 방법도 연구해 나갈 예정이다.

REFERENCES

1) David Yarowsky(1995) : *"Unsupervised Word Sense Disambiguation Rivaling Supervised Method", 33rd Annual Meeting of the ACL, pp189-196*
2) 신사임, 이주호, 최용석, 최기선(2001) : "공기정보 벡터를 이용한 한국어 명사의 의미구분". 제 *13* 회 한글 및 한국어 정보처리 학술대회
3) 송영빈, 최기선(2000) : "동사의 애매성 해소를 위한 시소러스의 이용과 한계". 제 *12* 회 한글 및 한국어 정보처리 학술대회
4) 허 정, 옥철영(2001) : "사전의 뜻풀이말에서 추출한 의미 정보에 기반한 동형이의어 중의성 해결 시스템". 정보과학회 논문지 *28(9) : 688-698*
5) 강유환, 정천영, 서영훈(2002) : "명사의 의미 정보를 이용한 복합 명사 분석의 중의성 해소". 제 *14* 회 한글 및 한국어 정보처리 학술대회
6) 신지현, 최민우, 강범모(2001) : "어휘의미분석 말뭉치 구축의 절차와 문제". 제 *13* 회 한글 및 한국어 정보처리 학술대회
7) 김영길, 양성일, 박상규 외 6명(2002) : "국소 구문 관계 및 의미 공기 정보에 기반한 명사 의미 모호성 해소". 제 *14* 회 한글 및 한국어 정보처리 학술대회
8) 윤준태, 정의석, 송만석(1998) : "명사간 어휘 정보를 이용한 한국어 복합 명사 분석". 정보과학회논문지 *(B) 25(11) : 1716-1725*
9) 강범모, 박동호, 이성헌, 박진호(2001) : "한국어 명사 의미 부류 체계의 구축과 활용". 한글 및 한국어 정보처리 학술대회