

## Ch7 의미 분석

### #1 중의성

- 언어 표현에서 둘 이상의 의미를 가지는 것
- 의사 전달을 하는 대상에게 혼동을 줄 수 있다.
- 정확한 분석을 위해선 중의성 표현들의 의미를 결정하는 외부 정보들을 정확하게 파악하는 것이 중요하다.
- 어휘적 중의성
  - 다의어에 의한 중의성: 하나의 단어가 주변 문맥이나 상황에 따라 다른 의미들로 해석되는 것
  - 동음어에 의한 중의성: 문장 주변의 문맥정보가 동일하지만, 상황에 따라 다른 의미로 해석되는 것
- 구조적 중의성
  - 문장을 이루고 있는 어휘 간의 관계에 따라 여러 의미로 해석되는 것
  - 수식어에 의한 중의성: 수식어가 수식하는 범위에 따라 문장의 해석이 달라지는 것(예: 부유한 철수와 영희가 명품매장을 갔다. 에서 철수만 부유한지, 철수와 영희 모두 부유한지 알 수 없다.)
  - 대칭 동사에 의한 중의성: 문장의 동사가 수식하는 범위에 따라 해석이 달라지는 것

### #2.1 단어 중의성 해소

- 문장 내 중의 성을 가지는 어휘를 사전에 정의된 의미와 매칭하여 어휘적 중의성을 해결하는 문제로 문장에 표현된 단어들을 기계가 정확하게 이해하게 하는 것

#### #2.1.2 지식기반 방법

- 문장에 등장한 단어들을 사전에 정의된 어휘지식을 활용하여 예측
- 현재 있는 어휘지식은 WordNet, BabelNet, ConceptNet, Freebase등이 있다.
- 사전 정의 기반 방법
  - 사전에 정의된 뜻풀이를 기반으로 의미를 추론
  - 대표적으론 레스크(lesk) 알고리즘이 있다. 중의성 단어의 사전 뜻풀이에 쓰인 단어들과 중의

성 단어와 함께 주변 문맥에 나타나 단어의 사전 뜻풀이에 쓰인 단어들의 문맥 패턴이나 중복된 단어 수를 보고 의미를 결정

- 뜻풀이에만 쓰인 단어와 문장에 너무 의존적임
- 그래프 기반 방법
  - 단어의 의미 관계 정보를 분석하여 의미를 추론
  - 어휘 의미망 사전에 정의된 의미 관계 정보 중, 중의성 단어와 주변에 있는 단어들의 부분 그래프를 새로 추출하고, 가장 연결성이 높은 중의성 단어의 의미를 선택하는 방법
  - EX) "She drinks some milk."에서 중의성을 가지는 'drink'와 'milk'의 의미를 선택하기 위해 'drink'가지는 모든 의미와 'milk'가 가지는 모든 의미를 특정 거리만큼 그래프를 추출하고, 연결성을 측정하여, 선택
  - 문법적 정보가 포함되어 있지 않아 정확한 의미를 분석하는데 어려움

#### #2.1.3 지도학습기반 방법

- 문장 내 단어의 의미가 레이블된 데이터를 이용하여 기계학습 모델에 하급하고 단어의 의미를 예측
- 대량의 학습데이터를 구축해야 하는 단점이 존재
- 하나의 의미로 다양한 문맥 패턴이 있으므로 학습데이터를 구축하는데 많은 시간이 필요

#### #3 의미역 분석

- 의미역: 의미를 해석하기 위해 서술어가 수식하는 대상의 의미 관계를 파악하고, 그 역할을 분류하는 것.
- 논항: 수식을 받는 대상, 문장의 구조가 바뀌어도 행위주, 피동작주는 변하지 않음

#### #3.1 의미역의 종류

- 필수적 의미역
  - 서술어의 의미를 구성하는 데 필수적으로 요구되는 것
  - 행동주, 도구, 피동주, 경험자, 수혜자, 출처, 도달점등으로 분류

행동주	어떤 행위를 하는 주체, 감정이나 의지를 가지는 사람이나 동물, 공의성이나 의도성을 가
-----	--

	짐.
도구	행위, 이동의 의미를 표현하는 동사의 수단
피동주/수동자	동사가 행위를 표현하는 경우, 그 행위에 영향을 받거나 상태 변화를 겪는 것
경험자	인지, 지각, 감정을 나타내는 용언일 때, 경험 주체가 되는 논항에 의해 내적인 상태에 영향을 받는 것
수혜자	어떤 행위가 행해졌을 때 이익을 받는 개체
출처.근원	행동의 동기나 이유, 장소
도달점/목표	행위가 목적하는 곳, 출처와 반대 개념

- 수의적 의미역