

Semantics Parsing

1 Overview

{ Lexical (语义) semantics

 | Sentence Semantics

{ Implicit: Vector representation

 | Explicit: Symbolic representation

 ↳ unclear which one is best → build an accurate parser is difficult

 ↳ Blackbox

2 Lexical Semantics

 * 同义词反义判定 (仅用 lemma)

 1. 同义词真时, 同义意义不同 → independent truth conditions

 2. 同义词 → 句法行为不同 → syntactic behaviors

 3. 同义词反义判定 → antagonistic meanings

 4. 同义词 (意义词) 集合不同 → sense relations

3 Semantic Relations

Synonymy → 同义关系

Antonymy → 反义关系

(下位词)

Holonymy → 下位关系: 前者为后者的子集

(A is-a B) ⇒ <dog, mammal>

(上位词)

Hypernymy → 上位关系: 后者为前者的子集

(A contains B) ⇒ <mammal, dog>

(上位词)

Meronymy → 部分关系 (A is a part of B) <thw, ShanghaiTech> △ 同义词关系,

(整体词)

Holonymy → 整体关系 (A has a B) <ShanghaiTech, thw>

Wordnet:  Synset (同义词集合)

此部分属于多义), 这为词集间关系

Semantic distance: 通过 Synsets 通过 hypernymy / hyponymy →

最近距离

Word Sense Disambiguation (WSD)

↳ Seq labeling

* Sentence Semantics ⇒ Based on lexical semantics

paragraph semantic (句段)

* Formal meaning Representation (形式化意义表示)

Idea

Unambiguity (无歧义): 一个词 → 唯一意义

Canonical form (规范化形式): 一个意义 → 唯一语义符号

{ Expressiveness

↔ Trade off

Inference ability

{ 考虑: Database, Robot

{ 通用: Formal logic, graphs

形式语言也是形式化语义表示 (✓)

△ First order logic (FOL)

↓

○ Term: 常量或变量为一个 term

○ Formula (公式)

n元关系 RL) , 7. 公式连接 (A.V. ⇒).

量词及变量 + 公式

of Flavor O

Node: Single Word Edge: Relation

DM, PSD

Flavor 1

Node: Arbitrary part of sentence Edge: relation

EDS

Flavor 2

anchored

Node: 抽离意义符号 Edge: Relations

Abstract Meaning Representation

Parsing (句子 → 意义表示)

Syntax driven (语句驱动)

Neural approach

* Neural Semantic Parsing

Seg-to-Seg: 将句子拆分成语义表示

Parsing to Semantic graph,

Method 1: Transition-based

Method 2: Graph-based (先 node 后 edge)

* Learning

监督学习

Weak supervised

部分语义表示可运行 → 知道运行结果

→ 先算再找语义图

With latent var. reinforcement learning.

* Syntax-driven Semantic Parsing

Idea: Principle of Compositionality

⇒ 从句到词 (Constituency parse)

⇒ Synchronous context-free grammar (SCFG)

Idea: 一个语义树 →

(由语义词汇和语义标注组成) / (语义
表达式)

⇒ 遍历同义域 2 种树

⇒ 语义树与语义对应

对于自然语言: 构建语义树 → 替换为语义树

成分分析

语义化语义表示

2.3.4.5: ----

ARGM-?: 修饰语

Frame: 谓词 (含多种属性) 分属多个框架 →

一个框架 (Frame) 定义了 {谓词集} 角色

标注方法

输入序列，输出序列：
Span
label of span
role

① Seq-labeling 法： BIO 标记 span，在角色
一次标记一个词

② Graph-Based : arcs 表示 span, role

③ Seq-to-Seq

Information Extraction

Named Entity Recognition

o Nested NE Recognition

△ Entity linking → 转移到某参考集

△ Relation Extraction : 抽取实体间关系。

△ Event extraction : 抽取事件及相关信息

Methods:

△ Seq labeling : BIO 进行 NE Recognition → Ambiguous

△ Span classification → 基于 dependency arcs prediction

Decoding-based methods:

△ Seq-to-Seq

