Natural Language Processing

Lecture 11 : Semantic Parsing

Master Degree in Computer Engineering
University of Padua
Lecturer : Giorgio Satta

Semantics & meaning

Bryan Goff from Unsplash

Semantics

Semantics is the study of the meaning of linguistic expressions at the level of words, phrases, and sentences.

Semantics focuses on what expressions **conventionally** mean, rather than on what they might mean in a particular context.

The latter is the focus of pragmatics.

This distinction is generally presented as the distinction between **referential meaning** as opposed to **associative meaning**.

Lexical semantics

©Smithsonian Magazine

WordNet

A **word sense** is a discrete representation of one aspect of the meaning of a word.

We represent each sense with a superscript.

Example : For the word **bank** we highlight two word senses through in context examples

- bank^1: ... a bank can hold the investments in a custodial account ...
- bank^2: ... as agriculture burgeons on the east bank, the river ...

WordNet

WordNet is a large online database that represents

- word senses
- semantic relations between word senses

These representations are often called word taxonomies.

The set of near-synonyms for a WordNet sense is called a **synset** (for synonym set).

Example : In WordNet the word **bass** is associated, among others, to the synset $\{\text{bass}^6, \text{bass voice}^1, \text{basso}^2\}$.

Versions of WordNet have been developed for many languages.

WordNet

Semantic relations between word senses include

- **synonymy**: couch/sofa
- **antonymy**: long/short
- **hyponymy**: dog/animal

  also called IS-A relation
- **hypernymy**: the inverse relation of hyponymy

WordNet 3.0 release has 117,798 nouns, 11,529 verbs, 22,479 adjectives, and 4,481 adverbs.

Word sense disambiguation

The task of selecting the correct sense for a word in a given context is called **word sense disambiguation** (WSD).

Several algorithms for WSD, using contextual word embeddings or else symbolic collocation features.

Semantic roles

©NASA, ESA, CSA, and STScI

Semantic roles

Semantic roles, also called **thematic roles**, express the role that arguments of a predicate take in the associated event

Example :
Sasha broke the window

- Sasha is the **agent**, expressing volitional causation
- window is the **theme** or **patient**, affected in some way by the action

Roles are **invariant** with respect to syntactic constructions

- [$_{\text{Agent}}$ Sasha] broke [$_{\text{Theme}}$ the window]
- [$_{\text{Theme}}$ The window] was broken by [$_{\text{Agent}}$ Sasha]

Do not confuse semantic roles with grammatical relations such as subject and object!

Semantic roles

Commonly used semantic roles

Thematic Role	**Example**
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The city built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	He poached catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

Semantic roles

Defining a standard set of semantic roles may be very difficult, sometimes further specialization is needed.

Example :
At least two kinds of **instrument** roles: intermediary instruments and enabling instruments

- the cook opened the jar with the new gadget
- the new gadget opened the jar
- Shelly ate the sliced banana with a fork
- *The fork ate the sliced banana

Symbol * used in linguistics to denote semantically odd sentences.

Semantic roles

Each sense of a verb is associated with a set of semantic roles, called the **thematic grid**.

Alternative terminology is θ-grid, or case frame.

Thematic grids resources are at the basis of NLP tasks such as semantic role labeling (introduced later) that construct **shallow meaning** representations.

Proposition Bank

Proposition Bank, also referred to as PropBank, is a resource of syntactic analyses annotated with semantic roles.

In Proposition Bank roles are given only numbers rather than names, that are specific for an individual verb sense: ARG0, ARG1, ARG2, and so on.

Proposition Bank

Roles are assigned according to the following general conventions

- ARG0 represents (several cases of) agent
- ARG1 represents (several cases of) patient
- ARG2 is often benefactive, instrument, attribute

PropBank also has a number of non-numbered arguments called ARGMS, (ARGM-TMP, ARGM-LOC, etc.) which represent modification or adjunct meanings.

The semantics of the other roles are less consistent.

Proposition Bank

PropBank entry for one of the senses of the verb **agree**.

agree.01
ARG0: Agreer
ARG1: Proposition
ARG2: Other entity agreeing

Example :

- [$_{\text{ARG0}}$ The group] agreed [$_{\text{ARG1}}$ it wouldn't make an offer]
- [$_{\text{ARGM-TMP}}$ Usually] [$_{\text{ARG0}}$ John] agrees [$_{\text{ARG2}}$ with Mary]

FrameNet

FrameNet is another role-labeling resource/project.

Roles in FrameNet are specific to a **frame**, a knowledge structure that can be shared across several words.

This allows to make inferences across different verbs, and also between verbs and nouns.

FrameNet

Example : All of the words highlighted below share the same frame

- [ITEM Oil] **rose** [ATTRIBUTE in price] [DIFFERENCE by 2%]
- [ITEM Microsoft shares] **fell** [FINALVALUE to 7 5/8]
- A steady **increase** [INITIALVALUE from 9.5] [FINALVALUE to 14.3] [ITEM in dividends]
- [ITEM It] has **increased** [FINALSTATE to having them 1 day a month]

FrameNet

Schema for the frame **change_position_on_a_scale**, used by words such as **increase**, **fall**, **rise**, etc.

Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication.
FINAL_VALUE	The position on the scale where the ITEM ends up.
INITIAL_STATE	A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication.
INITIAL_VALUE	The initial position on the scale from which the ITEM moves away.
ITEM	The entity that has a position on the scale.
VALUE_RANGE	A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate.
Some Non-Core Roles	
DURATION	The length of time over which the change takes place.
SPEED	The rate of change of the VALUE.
GROUP	The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way.

Core roles are frame specific; non-core roles are like the Arg-M arguments in PropBank.

Semantic role labeling

©iStock/Getty Images

Semantic role labeling (SRL) is the task of automatically finding the semantic role of each argument of a predicate in an input sentence.

Reference framework can be PropBank or FrameNet.

Most popular approaches use supervised learning

- guided by phrase-structure trees
- guided by dependency trees
- exploit sequence labelling

Semantic role labeling

Example :

- [ARG0 The San Francisco Examiner] [TARGET issued] [ARG0 a special edition] [ARGM-TMP yesterday]
- [COGNIZER You] can't [TARGET blame] [EVALUEE the program] [REASON for being unable to identify it]

Semantic role labeling

A simple approach to SRL is to view the task as a sequence labeling problem, and use BIO tags.

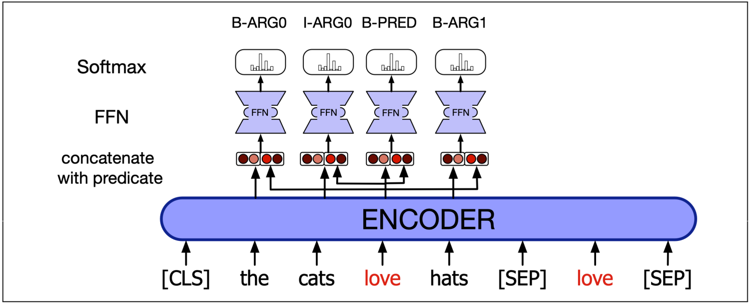
Assume

- an input sequence of words $w_{1:n}$
- the set $\mathcal{Y}(w_{1:n})$ of all tag sequences **compatible** with $w_{1:n}$

We are interested in computing the tag sequence $\widehat{y}_{1:n}$

$$\widehat{y}_{1:n} = \underset{y_{1:n} \in \mathcal{Y}(w_{1:n})}{\operatorname{argmax}} P(y_{1:n} \mid w_{1:n})$$

Semantic role labeling



Neural network algorithm

- input words mapped to embedding by a biLSTM encoder
- each token concatenated with the predicate embedding
- feedforward network with softmax maps to distribution over each SRL/BIO label

Semantic role labeling

The standard **evaluation** for SRL is to require that each argument label must be assigned to the exactly correct word sequence.

As usual for sequence labeling, precision, recall, and F-measure are computed.

Two common **datasets** used for evaluation are CoNLL-2005 and CoNLL-2012.

Argument selection

A **selectional restriction** is a semantic type constraint that a verb imposes on the concepts filling its argument roles.

Example :
The ARG0 of verb **eat** tends to be something that is edible

- I want to eat someplace nearby
- I want to eat Malaysian food

Selectional restriction can also be used to solve attachment problems in syntactic parsing.

Semantic role labeling

Selectional restrictions are associated with word senses, not lexemes.

Example :

Contrast the lexeme **serve** in the following sentences

- The restaurant serves green-lipped mussels
- Which airlines serve Denver?

Meaning

Space.com

Meaning

In past years, several tasks related to semantic analysis of sentences have been investigated in isolation

- word sense disambiguation (WSD)
- named-entity recognition (NER)
- semantic role labeling (SRL)
- coreference resolution

Recently there has been some major effort to design and process formalisms that incorporate all of the above information.

There is the hope that the smaller tasks can be better solved as a by-product of whole-sentence semantic analysis.

Meaning representations

The meaning of linguistic expressions can be captured by means of formal structures called **meaning representations**.

Meaning representations use symbols that correspond to

- objects or variables
- properties of objects
- relations among objects

Meaning representations **link** sentences to some state of affairs in some world being represented or reasoned about.

Meaning representations

Among several desiderata, we want sentences with the same semantics to be assigned the same meaning representation.

Example :

- he described her as a genius
- his description of her: genius
- she was a genius, according to his description

Meaning representations

Several semantic formalisms for meaning representations have been proposed in the literature

- abstract meaning representation (AMR)
- universal conceptual cognitive annotation
- universal decompositional semantics
- minimal recursion semantics (MRS)
- discourse representation theory (DRT)

The recent availability of annotated corpora for these formalisms has further enhanced the development and exploration of various semantic parsing models.

First-order logic

First-order logic (FOL) has been traditionally used for meaning representation

Example :
'A restaurant that serves Mexican food near ICSI'

$$\exists x : \mathit{Restaurant}(x) \wedge \mathit{Serves}(x, \mathit{MexicanFood}) \\ \wedge \mathit{Near}(\mathit{LocationOf}(x), \mathit{LocationOf}(\mathit{ICSI}))$$

'All vegetarian restaurants serve vegetarian food'

$$\forall x : \mathit{VegetarianRestaurant}(x) \\ \Rightarrow \mathit{Serves}(x, \mathit{VegetarianFood})$$

Abstract meaning representations (AMRs) are rooted, labeled, directed graphs where

- nodes denote concepts, events, properties, and states.
- arcs denote semantic relations and semantic roles

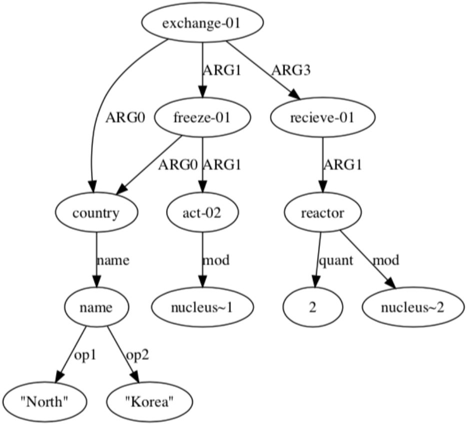
For the most, AMR are acyclic graphs.

When an entity plays multiple roles in a sentence, we employ graph **re-entrancy** (nodes with multiple parents)

AMR are thus dependency structures, but differently from the syntactic case, they show re-entrancies.

Abstract meaning representation

'North Korea froze its nuclear actions in exchange for two nuclear reactors.'



Two occurrences of modifier word 'nuclear' represented by nucleus-1 and nucleus-2.

Semantic parsing

The process of mapping sentences to meaning representations is called **semantic parsing** or semantic analysis.

At time of writing, **AMR parsing** is receiving increasing attention.

Due to the already mentioned similarity between AMR and dependency structures, transition-based parsers have been adapted to AMR parsing.

Semantic parsing

We consider a model with the following transition operators

- SHIFT: decide if and what to push on the stack after consuming a token from the buffer
- LEFTARC(ℓ): creates an edge with label ℓ between the top-most node and the second top-most node in the stack, and pops the latter
- RIGHTARC(ℓ): is the symmetric operation, but does not pop any node from the stack
- REDUCE: pops the top-most node from the stack and recovers reentrant edges with its sibling nodes

Very similar to arc-eager syntactic parser, but with operational differences for REDUCE transition.

Title: A Transition-Based Algorithm for Unrestricted AMR Parsing
Authors: David Vilares, Carlos Gómez-Rodríguez
Journal: NAACL, 2018
Content: Non-projective parsing can be useful to handle cycles and reentrancy in AMR graphs. This work introduces a greedy left-to-right non-projective transition-based parser. At each parsing configuration, an oracle decides whether to create a concept or whether to connect a pair of existing concepts. The algorithm handles reentrancy and arbitrary cycles natively.

https://aclanthology.org/N18-2023/