Spoken Human Robot Interaction

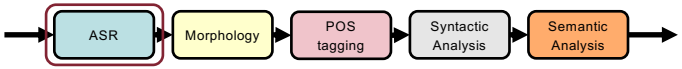
Semantic Analysis

SAPIENZA
UNIVERSITÀ DI ROMA

Andrea Vanzo, **Daniele Nardi**
lastname@dis.uniroma1.it
Department of Computer, Control, and Management Engineering
"Sapienza" University of Rome, Italy

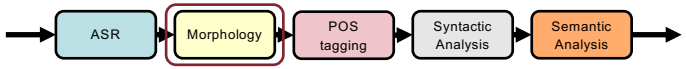
Artificial Intelligence
AY 2018/19

Our Command Interpretation pipeline



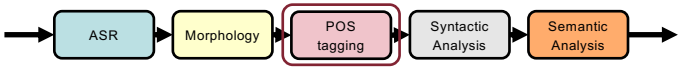
- **ASR**: process that generates a "text", starting from an audio signal

Our Command Interpretation pipeline



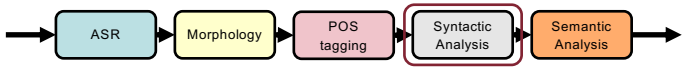
- **ASR**: process that generates a "text", starting from an audio signal
- **Morphology**: study of the words are composed starting from small informative units

Our Command Interpretation pipeline



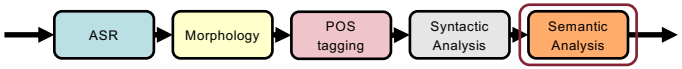
- **ASR**: process that generates a "text", starting from an audio signal
- **Morphology**: study of the words are composed starting from small informative units
- **POS tagging**: process of assigning a morpho-syntactic category to words

Our Command Interpretation pipeline



- **ASR**: process that generates a "text", starting from an audio signal
- **Morphology**: study of the words are composed starting from small informative units
- **POS tagging**: process of assigning a morpho-syntactic category to words
- **Syntactic analysis**: study of the structural relationship between words

Our Command Interpretation pipeline



- **ASR**: process that generates a "text", starting from an audio signal
- **Morphology**: study of the words are composed starting from small informative units
- **POS tagging**: process of assigning a morpho-syntactic category to words
- **Syntactic analysis**: study of the structural relationship between words
- **Semantic analysis: study of meaning of a word/sentence/text**

Outline

- Semantic Analysis
- Lexical Semantics
  - WordNet
  - Word Embeddings
- Sentence Semantics
  - Compositionality Principle
- Grounding

Semantic Analysis: what for?

- Information Extraction
  - Extract small amounts of information from large bodies of text
  - Find an answer to a question
- Text summarization
  - Sum up a text
- Information retrieval and document classification
  - Word sense disambiguation and query augmentation
- Machine translation
  - Align languages at semantic level
- Human-Robot Interaction
  - Enable Spoken Language Understanding for interacting with robots (e.g. interpreting a command,)

Why is Semantic Analysis difficult?

- Ambiguity of language

- Language understanding often requires inference to discover hidden knowledge
  - ”You need an umbrella today” → It’s a rainy day
- Language is dynamic
  - Allows defining new terms, expressions, ...
- Language grounding
  - ”Here is you yellow cab” (associate words to world objects via perception)

Why is Semantic Analysis important?

- Power of language
  - transfer thoughts from between agents
  - transfer between human and computer
- NL is a very general representation language
  - words are a powerful descriptive tool
  - NL brings about the ability to reason

What can the result of Semantic Analysis be?

- In linguistics, the output of Semantic Analysis is the meaning of a word/sentence/text…
  - ...expressed through a given representation

Example
- “bring me the mug”
- “could you please bring me the mug?”
- “take the mug to me”

All the above sentences share the same meaning (to some extent)

Lexical Semantics vs Sentence Semantics

- Lexical semantics
  - Study words’ meaning
  - Aim at
    - generating clusters of similar words (hand-crafted from statistics)
    - evaluating the semantic similarity between two words
- Sentence/text semantics
  - Study sentences/documents’ meaning
  - Based on the way words are combined to generate a meaning
  - Aim at
    - extracting the meaning of an entire sentence
    - evaluating the semantic similarity between two texts

Lexical Semantics resources: WordNet

A hierarchically organized lexical database (in English)

Example: `table` WordNet entry

Noun

- S: (n) **table**, tabular array (a set of data arranged in rows and columns) *"see table 1"*
- S: (n) **table** (a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs) *"it was a sturdy table"*
- S: (n) **table** (a piece of furniture with tableware for a meal laid out on it) *"I reserved a table at my favorite restaurant"*
- S: (n) mesa, **table** (flat tableland with steep edges) *"the tribe was relatively safe on the mesa but they had to descend into the valley for water"*
- S: (n) **table** (a company of people assembled at a table for a meal or game) *"he entertained the whole table with his witty remarks"*
- S: (n) board, **table** (food or meals in general) *"she sets a fine table"; "room and board"*

Verb

- S: (v) postpone, prorogue, hold over, put over, **table**, shelve, set back, defer, remit, put off (hold back to a later time) *"let's postpone the exam"*
- S: (v) **table**, tabularize, tabularise, tabulate (arrange or enter in tabular form)

Lexical Semantics resources

WordNet

- Sense in WordNet is defined through the ***synset***
- *Synset* is the set of near synonyms
- Example: `table` has multiple synsets
  - S1: a set of data arranged in rows and columns *(e.g. see table 1)*
  - *S2:* a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs *(e.g. it is a wooden table)*
- The above senses (S1, S2, …) share the same lemma/surface form (Homonymy)

Relations encoded in WordNet

In addition to the senses of a word, Wordnet encodes several relations:

Relation	Also called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	*breakfast*^{1} → *meal*^{1}
Hyponym	Subordinate	From concepts to subtypes	*meal*^{1} → *lunch*^{1}
Member Meronym	Has-Member	From groups to their members	*faculty*^{2} → *professor*^{1}
Has-Instance		From concepts to instances of the concept	*composer*^{1} → *Bach*^{1}
Instance		From instances to their concepts	*Austen*^{1} → *author*^{1}
Member Holonym	Member-Of	From members to their groups	*copilot*^{1} → *crew*^{1}
Part Meronym	Has-Part	From wholes to parts	*table*^{2} → *leg*^{3}
Part Holonym	Part-Of	From parts to wholes	*course*^{7} → *meal*^{1}
Antonym		Opposites	*leader*^{1} → *follower*^{1}

Wordnet can be used for:

- Word similarity
- Word-sense disambiguation

Relations between words

Word similarity: even when words are not synonyms they may be similar (e.g. *dog* and *cat*)

Word relatedness (association): connections that go beyond similarity (e.g. *coffee* and *cup*), words that belong to the same **semantic field** (e.g. hospital, restaurants)

Semantic Frames and Roles: set of words that denote perspectives of the participants in a particular type of event (e.g. transaction)

Taxonomic Relations: hyponym and hypernym (subsumption)

The Bag Of Word model (BOW)

- The starting point of each geometrical word representation
- Texts are represented as the bag (multiset) of their words
- Main idea: counting the occurrences of words into texts
- **Word-Document** Matrix:
  - Each column represents the occurrences of all the words in a doc
  - Each row represents the number of occurrences of one word in each of the docs. Typically |V| rows

The Bag Of Word model: examples

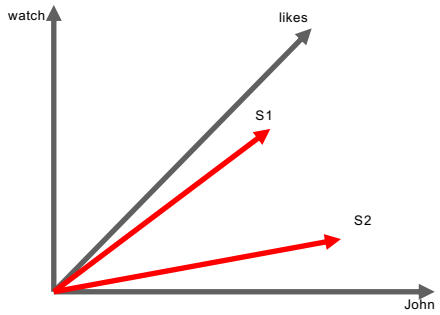
S1: John likes to watch movies. Mary likes movies too.

John	*likes*	*to*	*watch*	*movies*	*Mary*	*too*
1	2	1	1	2	1	1

S2: John also likes to watch football games.

John	*also*	*likes*	*to*	*watch*	*football*	*games*
1	1	1	1	1	1	1

Visualization in the vector space



Similarity = distance in the vector space (e.g. cosine)

From word counting to word contexts

- Looking at the words in isolation is often not enough to properly characterize the word itself
- Might be better looking at their contexts

  I love Programming. I love Math. I tolerate Biology.

	I	love	Programming	Math	tolerate	Biology	.
I	0	2	0	0	1	0	2
love	2	0	1	1	0	0	0
Programming	0	1	0	0	0	0	1
Math	0	1	0	0	0	0	1
tolerate	1	0	0	0	0	1	0
Biology	0	0	0	0	1	0	1
.	2	0	1	1	0	1	0

Co-occurrence word embeddings

A **word-word** (term-context) matrix:

- Rows contain the number of co-occurrences for a word
- Co-occurrences counted on varying-size contexts (e.g. 3)
- Vectors can again be represented in a multi-dim space

Example: *Programming* and *Math* share the same vector

– Semantically similar given this domain → cosine = 1

Limitations:

– Sparseness

– Not scalable

	I	love	Programming	Math	tolerate	Biology	.
I	0	2	0	0	1	0	2
love	2	0	1	1	0	0	0
Programming	0	1	0	0	0	0	1
Math	0	1	0	0	0	0	1
tolerate	1	0	0	0	0	1	0
Biology	0	0	0	0	1	0	1
.	2	0	1	1	0	1	0

Other word embeddings

Variants (manifold):

– Tf-idf (term frequency-inverse document frequency)

– N-grams

Approaches to reduce dimensionality (typically to 250)

– Singular Value Decomposition (SVD)

- Generalization of the eigendecomposition

– Word2vec [Mikolov et al., 2013]

- Dense vectors estimated through a 2-layered NN

```
...
with   -0.057929,0.019783,-0.021198,0.156665,0.075069,-0.033843, ...
it   -0.066015,-0.014947,-0.084252,-0.256743,-0.110134,-0.003800, ...
...
```


Summary on vector representations

Representations of meaning of words build after distributional hypothesis

Applications:

– text similarity (cosine similarity)

– Information Retrieval

– Text summarization

– Question answering

– …

– feature characterization for other NLP tasks (i.e. parsing)

Sentence Semantics

Principle of Semantic Compositionality (from Philosophy)

*The meaning of a **complex expression** is determined by the meanings of its **components***

In Word Embeddings (**meaning**) of words (**components**) composing a sentence (**complex expression**) can be combined to *"approximate"* the meaning of the sentence (e.g. for Information Retrieval)

$$\mu(S) = \mu(\overline{w}_1, \ldots, \overline{w}_n) = \sum_{i=1}^{n} \overline{w}_i$$

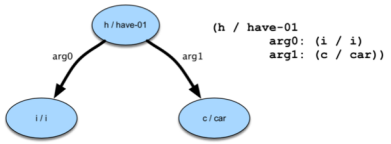
Representing semantics in NL

- Predicate Logic / Description Logics

$$\exists e, y\ Having(e) \wedge Haver(e, Speaker) \wedge HadThing(e, y) \wedge Car(y)$$

Semantic Role Labeling:

- Abstract Meaning Representation




```
(h / have-01
      arg0: (i / i)
      arg1: (c / car))
```

- Frame Semantics

```
Having:
    Haver:  Speaker
    HadThing:  Car
```


Representing semantics in Predicate Logic

Meaning representations are encoded as logic predicates

- Logical propositions enable inference
- Example

"A restaurant that serves Italian food near Princess Street"

corresponds to the meaning representation

$$\exists Restaurant(x) \wedge Serves(x, ItalianFood)$$
$$\wedge Near(LocationOf(x), LocationOf(PrincessStreet))$$

Scalability problem (large vocabulary or unrestricted domain)

Representing semantics as Abstract Meaning

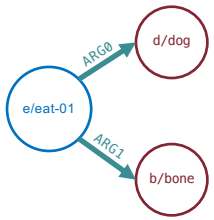
Back to Panini (Sanskrit, 400 BC)

- Human-readable semantic representation
- Focuses on *"who is doing what to whom"* in a sentence (**Role-Semantics**)
  - Who: agent
  - What: action
  - Whom: target of the action
- Whole-sentence meaning encoded in a tree-structure
- Construction: **Semantic Role Labelling**

- Pros: flexible, not linked to any semantic theory
- Cons: too flexible, too-open ended, lacking a commonly accepted def. of roles

The dog is eating a bone

```
(e/eat-01
   :ARG0 (d/dog)
   :ARG1 (b/bone))
```



Frame semantics

- Research in Empirical Semantics suggests that words represent categories of experience (*situations*)
- A *frame* is a cognitive structuring device evoked by *words* and used to support understanding (Fillmore, 1975)
  - *Lexical Units* evoke a Frame in a sentence
  - Frames are made of *elements* that express participants to the situation (*Frame Elements*)
- During communication LUs evoke the frames

Frame semantics: example from FrameNet

Bringing

Lexical Unit Index

Definition:

This frame concerns the movement of a Theme and an Agent and/or Carrier. The Agent, a person or other sentient entity, controls the shared Path by moving the Theme during the motion. In other words, the Agent has overall motion in directing the motion of the Theme. The Carrier may be a separate entity, or it may be the Agent's body. The Constant_location may be a subregion of the Agent's body or (a subregion of) a vehicle that the Agent uses.

Kari CARRIED the books across campus to the library on his head.

Kari CARRIED the books across campus to the library in his truck.

Kari CARRIED the books across campus to the library by truck.

The truck CARRIED the books across campus to the library in specially designed boxes.

The FEs include Path, Goal, and Source. Area is an area that contains the motion when the path is understood as irregular. This frame emphasizes the path of movement as opposed to the FEs Source or Goal as in Filling or Placing.

"Bring the cup on the table"

FrameNet

- Lexical resource that describes a significant portion of English in terms of precise and rich frame semantics
- The resource
  - Frame Database: a structured system of Frames and FEs
  - Lexical database: syntactic and semantic descriptions of frame-evoking words
  - Annotated Corpus: wide coverage examples

Grounding

Grounding

From Wikipedia, the free encyclopedia

Grounding or **grounded** may refer to:

- Ground (electricity), a common return path for electric current
- Grounding (punishment), restrictions placed on movement or privileges
- Grounding (metaphysics), a topic of wide philosophical interest.
- Grounding in communication, the collection of mutual knowledge, beliefs, and assumptions; "common ground"
- Intentional grounding, a rule violation in gridiron football
- Ship grounding, a type of marine accident
- Symbol grounding, a problem in cognition and artificial intelligence
- Earthing therapy, alternative health practice in which one remains in physical contact with the soil while also touching a devic

Language Grounding

Modeling semantics through symbols

- Word similarity
  - Cosine similarity among word embeddings
  - Phonetic similarity to alleviate ASR typos
  - Requires a consistent naming of objects into the semantic map

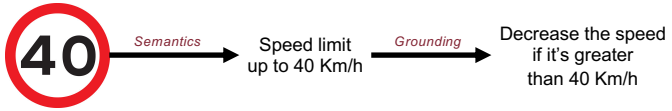
Perception-driven (e.g. on robots)

- Using vision and ML to map words to active perception
- Spatial Semantics (Pairing lexical and physical spatial relations)

Symbol Grounding

Philosophy

- *"ground a symbol meaning in something other than just more meaningless symbols"* [Harnad, 1990]

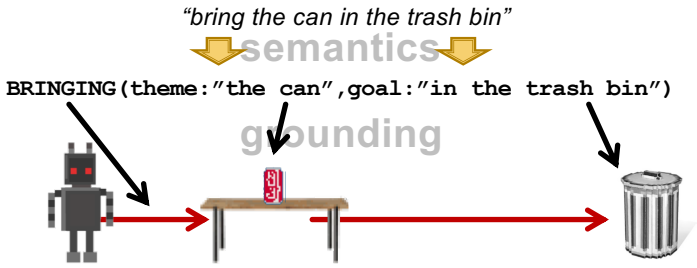


AI and Robotics

- *"the process of creating and maintaining the correspondence between symbols and sensor data that refer to the same physical objects"* [Coradeschi & Saffiotti, 2003]

Grounding language through perception

- *"map a lexical symbol (word) to a to an entity (or action) that occur into the environment"*



References

Daniel Jurafsky and James H. Martin. *Speech and Language Processing*.
https://web.stanford.edu/~jurafsky/slp3/

Chapter 6 (6.1—6.3)

Chapter 18 (18.1—18.3,18.5)