Lecture 26:
A very brief introduction to verb semantics and discourse

Julia Hockenmaier
juliahmr@illinois.edu
3324 Siebel Center

Understanding sentences

“Every chef cooks a meal”

\[ \forall x [chef(x) \rightarrow \exists y [meal(y) \land cooks(y,x)]] \]

\[ \exists y [meal(y) \land \forall x [chef(x) \rightarrow cooks(y,x)]] \]

We translate sentences into (first-order) predicate logic.

Every (declarative) sentence corresponds to a proposition, which can be true or false.

Recap: CCG derivation with semantics

\[ \text{John} \quad \text{sees} \quad \text{Mary} \]

\[ \text{NP : John} \quad \text{(S\NP)/NP : } \lambda x . \lambda y . \text{sees}(x,y) \quad \text{NP : Mary} \]

\[ \text{S\NP : } \lambda y . \text{sees}(\text{Mary},y) \]

\[ \text{S : sees(\text{Mary},\text{John})} \]
But…

… what can we do with these representations?

Being able to translate a sentence into predicate logic is not enough, unless we also know what these predicates mean.

Semantics joke (B. Partee): The meaning of life is *life*

Compositional formal semantics tells us how to fit together pieces of meaning, but doesn't have much to say about the meaning of the basic pieces (i.e. lexical semantics)

… how do we put together meaning representations of multiple sentences?

We need to consider discourse (there are approaches within formal semantics, e.g. Discourse Representation Theory)

… Do we really need a *complete* analysis of each sentence?

This is pretty brittle (it's easy to make a parsing mistake)

Can we get a more shallow analysis?

Verb semantics

Verbs describe events or states (‘eventualities’): 

- **Tom** broke the **window** with a **rock**.
- The **window** broke.
- The **window** was broken by **Tom**/by a **rock**.

**Thematic roles** refer to participants of these events:

- **Agent** (who performed the action): **Tom**
- **Patient** (who was the action performed on): **window**
- **Tool/Instrument** (what was used to perform the action): **rock**

Thematic roles are different from grammatical roles (subject or object).

Today’s lecture

**Verb semantics**:

- Thematic roles
- Semantic role labeling
- Diathesis alternations
- Verb classes

**Discourse semantics**:

- What is discourse?
- Entity-based coherence
- Coreference resolution
- Rhetorical coherence

**Thematic roles**
The inventory of thematic roles

It is difficult to give a formal definition of thematic roles that generalizes across all verbs.

Proposition Bank (PropBank):
- Arg0 = proto-agent
- Arg1 = proto-patient
- Arg2...: specific to each verb
- ArgM-TMP/LOC/...: temporal/locative/... modifiers

FrameNet:
Verbs fall into classes that define different kinds of frames (change-position-on-a-scale frame: rise, increase,...).
Each frame has its own set of frame elements.

PropBank

agree.01  Arg0: Agreer  Arg1: Proposition
Arg2: Other entity agreeing
[Arg0, The group] agreed [Arg1, it wouldn’t make an offer]
[Arg0, John] agrees with [Arg2, Mary]

fall.01  Arg1: patient/thing falling  Arg2: extent/amount fallen
Arg3: start point  Arg4: end point
[Arg1, Sales] fell [Arg4, to $251 million]
[Arg1, Junk bonds] fell [Arg4, by 5%]

Semantic role labeling: Recover the semantic roles of verbs (nowadays typically PropBank-style)
Machine learning; trained on PropBank
Syntactic parses provide useful information

Diathesis Alternations

Active/passive alternation:
- Tom broke the window with a rock. (active voice)
- The window was broken by Tom by a rock. (passive voice)

Causative alternation:
- Tom broke the window. (‘causative’; active voice)
- The window broke. (‘anticausative’/’inchoative’; active voice)

Dative alternation
- Tom gave the gift to Mary.
- Tom gave Mary the gift.

Locative alternation:
- Jessica loaded boxes into the wagon.
- Jessica loaded the wagon with boxes.

Verb classes

Verbs with similar meanings undergo the same syntactic alternations, and have the same set of thematic roles (Beth Levin, 1993)

VerbNet (verbs.colorado.edu; Kipper et al., 2008)
A large database of verbs, their thematic roles and their alternations
Discourse

What is discourse?

On Monday, John went to Einstein’s. He wanted to buy lunch. But the cafe was closed. That made him angry, so the next day he went to Green Street instead.

‘Discourse’: any linguistic unit that consists of multiple sentences

Speakers describe “some situation or state of the real or some hypothetical world” (Webber, 1983)

Speakers attempt to get the listener to construct a similar model of the situation.

What can go wrong in discourse?

Discourse 1:
John hid Bill’s car keys. He was drunk.

Discourse 2:
John hid Bill’s car keys. He likes spinach.

Coherence relations

Discourse 1:
John hid Bill’s car keys. He was drunk.

Discourse 2:
John hid Bill’s car keys. He likes spinach.

Discourse 1 is more coherent than Discourse 2 because “He(=Bill) was drunk” provides an explanation for “John hid Bill’s car keys”

What kind of relations between two consecutive utterances (=sentences, clauses, paragraphs,…) make a discourse coherent?

Rhetorical Structure Theory; also lots of recent work on discourse parsing (Penn Discourse Treebank)
Example: The Result relation

The reader can infer that the state/event described in S0 causes (or: could cause) the state/event asserted in S1:

S0: The Tin Woodman was caught in the rain.
S1: His joints rusted.

This can be rephrased as: “S0. As a result, S1”

Example: The Explanation relation

The reader can infer that the state/event in S1 provides an explanation (reason) for the state/event in S0:

S0: John hid Bill’s car keys.
S1: He was drunk.

This can be rephrased as: “S0 because S1”

Rhetorical Structure Theory (RST)

RST (Mann & Thompson, 1987) describes coherence relations between utterances.

It defines a set of rhetorical relations: Evidence, Elaboration, Attribution, Contrast, List,… Different variants of RST assume different sets of relations.

Most relations hold between a nucleus (N) and a satellite (S). Some relations (e.g. List) have multiple nuclei (and no satellite).

Every relation imposes certain constraints on its arguments (N,S), that describe the goals and beliefs of the reader R and writer W, and the effect of the utterance on the reader.

Discourse structure is hierarchical

RST website: http://www.sfu.ca/rst/
What else can go wrong in discourse?

Discourse 1:
John went to his favorite music store to buy a piano.
It was a store John had frequented for many years.
He was excited that he could finally buy a piano.
It was closing just as John arrived.

Discourse 2:
John went to his favorite music store to buy a piano.
He had frequented the store for many years.
He was excited that he could finally buy a piano.
He arrived just as the store was closing for the day.

Entity-based coherence

Discourse 1:
John went to his favorite music store to buy a piano.
It was a store John had frequented for many years.
He was excited that he could finally buy a piano.
It was closing just as John arrived.

Discourse 2:
John went to his favorite music store to buy a piano.
He had frequented the store for many years.
He was excited that he could finally buy a piano.
He arrived just as the store was closing for the day.

How we refer to entities influences how coherent a discourse is (Centering theory)

Centering Theory
Grosz, Joshi, Weinstein (1986, 1995)

A linguistic theory of entity-based coherence and salience
It predicts which entities are salient at any point during a discourse.
It also predicts whether a discourse is entity-coherent, based on its referring expressions.

Centering is about local (=within a discourse segment) coherence and salience

Centering theory itself is not a computational model
or an algorithm: many of its assumptions are not precise enough to be implemented directly. (Poesio et al. 2004)

But many algorithms have been developed based on specific instantiations of the assumptions that Centering theory makes. The textbook presents a centering-based pronoun-resolution algorithm

How can we understand discourse?

On Monday, John went to Einstein’s. He wanted to buy lunch. But the cafe was closed. That made him angry, so the next day he went to Green Street instead.

Understanding discourse requires (among other things):

1) doing coreference resolution:
‘the cafe’ and ‘Einstein’s’ refer to the same entity
He and John refer to the same person. That refers to ‘the cafe was closed’.

2) identifying discourse (‘coherence’) relations:
‘He wanted to buy lunch’ is the reason for ‘John went to Bevande.’
Discourse models

An explicit representation of:

- **the events and entities** that a discourse talks about
- **the relations** between them (and to the real world).

This representation is often written in some form of logic.

What does this logic need to capture?

Discourse models should capture...

**Physical entities**: John, Einstein’s, lunch

**Events**: On Monday, John went to Einstein’s involve entities, take place at a point in time

**States**: It was closed. involve entities and hold for a period of time

**Temporal relations**: afterwards between events and states

**Rhetorical (‘discourse’) relations**: ... so ... instead between events and states

How do we refer to entities?


‘the book’

‘this book’

‘a book’

‘my book’

‘that one’

‘the book I’m reading’
Some terminology

**Referring expressions** (‘this book’, ‘it’) refer to some entity (e.g. a book), which is called the **referent**.

**Co-reference**: two referring expressions that refer to the same entity **co-refer** (are co-referent).

*I saw a movie last night. I think you should see it too!*

The referent is **evoked** in its first mention, and **accessed** in any subsequent mention.

Indefinite NPs

- no determiner:  
  *I like walnuts.*
- the indefinite determiner:  
  *She sent her a beautiful goose*
- numerals:  
  *I saw three geese.*
- indefinite quantifiers:  
  *I ate some walnuts.*
- *(indefinite) this:*  
  *I saw this beautiful Ford Falcon today*

Indefinites usually **introduce a new discourse entity**. They can refer to a specific entity or not:  
*I’m going to buy a computer today.*

Definite NPs

- the **definite** article (**the book**),
- **demonstrative articles** (**this/that book**, **these/those books**),
- **possessives** (**my/John’s book**)
  
Definite NPs can also consist of
- personal pronouns (**I, he**)
- demonstrative pronouns (**this, that, these, those**)
- universal quantifiers (**all, every**)
- *(unmodified) proper nouns** (**John Smith, Mary, Urbana**)

Definite NPs **refer to an identifiable entity** (previously mentioned or not)

Information status

Every entity can be classified along two dimensions:

**Hearer-new vs. hearer-old**
Speaker assumes entity is (un)known to the hearer

**Hearer-old**: I will call **Sandra Thompson**.
**Hearer-new**: I will call **a colleague in California** (**Sandra Thompson**)

Special case of hearer-old: **hearer-inferrable**
*I went to the student union. The food court was really crowded.*

**Discourse-new vs. discourse-old:**
Speaker introduces new entity into the discourse, or refers to an entity that has been previously introduced.

**Discourse-old**: I will call **her/Sandra** now.  
**Discourse-new**: I will call **my friend Sandra** now.
Coreference resolution

Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to $1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.

Coreference chains:
1. {Victoria Chen, Chief Financial Officer...since 2004, her, the 37-year-old, the Denver-based financial services company’s president}
2. {Megabucks Banking Corp, Denver-based financial services company, Megabucks}
3. {her pay}
4. {rival Lotsabucks}

Coref as binary classification

Represent each NP-NP pair (+context) as a feature vector.

Training:
Learn a binary classifier to decide whether NP$_i$ is a possible antecedent of NP$_j$

Decoding (running the system on new text):
- Pass through the text from beginning to end
- For each NP$_i$:
  - Go through NP$_{i+1}$...NP$_1$ to find best antecedent NP$_j$.
  - Corefer NP$_i$ with NP$_j$.
  - If the classifier can’t identify an antecedent for NP$_i$, it’s a new entity.

Features for Coref resolution

- Do the two NPs have the same head noun? (e.g. company)
- Do they contain the same modifier? (e.g. Denver-based)?
- Does the gender and number of the NPs match?
- Does one NP contain an alias (acronym) of the other? (United States = USA, Chief Executive Office = CEO)
- Is one NP a hypernym/synonym of the other?
- Is one NP an appositive of the other? [Victoria Chen], [CFO of Megabucks]
- Are both NPs named entities of the same type? [CEO] = PERSON, Victoria Chen = PERSON

Evaluation: B-cubed F-score

The test data consists of $D$ documents $d$ with $N$ total mentions $m$ (mention boundaries are given as input)
- In the gold standard, each mention $m$ belongs to a ‘true’ cluster of mentions (=connected component) of size $t_m$
- In the system output, each mention $m$ belongs to a predicted cluster of mentions (=connected component) of size $p_m$
- For each mention $m$, the intersection of the gold standard and system output clusters defines a common cluster of mentions of size $c_m$

\[
\text{Precision } P = \frac{1}{N} \sum_{d \in D} \sum_{m \in d} \frac{c_m}{p_m} \\
\text{Recall } R = \frac{1}{N} \sum_{d \in D} \sum_{m \in d} \frac{c_m}{t_m} \\
\text{F-measure } = \frac{2PR}{P + R}
\]
**Special case: Pronoun resolution**

Task: Find the antecedent of an anaphoric pronoun in context

1. John saw a beautiful Ford Falcon at the dealership.
2. He showed it to Bob.
3. He bought it.

\[ \text{he}_2, \text{it}_2 = \text{John, Ford Falcon, or dealership?} \]
\[ \text{he}_3, \text{it}_2 = \text{John, Ford Falcon, dealership, or Bob?} \]

**Anaphoric pronouns**

Anaphoric pronouns refer back to some previously introduced entity/discourse referent:

*John showed Bob his car. He was impressed.*

John showed Bob his car. *This* took five minutes.

The antecedent of an anaphor is the previous expression that refers to the same entity.

There are number/gender/person agreement constraints: *girls* can’t be the antecedent of *he*

Usually, we need some form of inference to identify the antecedents.

**Salience/Focus**

Only some recently mentioned entities can be referred to by pronouns:

*John went to Bob’s party and parked next to a classic Ford Falcon.*
*He went inside and talked to Bob for more than an hour.*
*Bob told him that he recently got engaged.*
*He also said he bought it (???)/* the Falcon yesterday.*

Key insight (also captured in Centering Theory)
Capturing which entities are salient (in focus) reduces the amount of search (inference) necessary to interpret pronouns!