DAY 10: SEMANTICS AND PRAGMATICS
AND FUTURE CHALLENGES IN NLP

Mark Granroth-Wilding

SYNTAX AND SEMANTICS

Parsing helps to uncover aspects of sentence structure important to interpreting meaning (semantics)

Syntax describes process mapping surface form to meaning via analysis of hidden structure Parsing

TRUTH-CONDITIONAL SEMANTICS

• Semantics of sentence ≃ conditions for it to be true
• Conditions defined in terms of entities, properties, states
  The ball is on the table
  ≡
  The table is under the ball
• Typically uses predicate logic
  \[ \exists x, y. \text{on}(x, y) \land \text{ball}(x) \land \text{table}(y) \]

SYNTAX TO SEMANTICS

• Remember:
  The ball is on the table
  \[ \exists x, y. \text{on}(x, y) \land \text{ball}(x) \land \text{table}(y) \]
  depends on correct syntactic structure of sentence
• ball is subject of is
• Formal semantic representations built by compositional semantics using parsing
  word meanings → phrase meanings → sentence meaning
• Words/phrases: logical expressions with 'holes'

FORMAL SEMANTICS

• We’ve seen:
  Parsing helps to uncover aspects of structure important to interpreting meaning (semantics)
• Also: some ways to represent lexical semantics
  WordNet, word sense disambiguation, word embeddings, …
• Not yet:
  semantic representations dependent on syntactic mapping
• Formal semantics: large, complex field
• Brief tasters here: demonstrate difficult analysis

FIRST-ORDER LOGIC

• Predicate / first-order logic:
  • objects in a domain
  • logical predicates with arguments
  • variables ranging over objects
  • quantifiers applied to variable: \( \exists \) (exists), \( \forall \) (every)
    \[ \text{table}(\text{obj}) \]
    \[ \exists x. \text{table}(x) \]
• Logical expressions permit inference
  I’d like to find a restaurant where I can get vegetarian food.
  \[ \text{Serves}(x, \text{VegetarianFood}) \]
• FOL has limitations
• but well understood and computationally tractable

FOL WITH HOLES

\[ \exists x, y. \text{on}(x, y) \land \text{ball}(x) \land \text{table}(y) \]

Detection NP

[Det]

N

VP

\[ \exists y. \text{on}(A, y) \land \text{table}(y) \]

Prep

[PP]

\[ \exists y. \text{on}(A, B) \land \text{table}(y) \]

Prep NP

The ball is on the table

WHY NLP IS HARDER THAN IT SEEMS

• Real-life NLP presents more challenges than we’ve seen
• Methods here provide good starting point
• Today: examples of harder challenges
• Includes unsolved problems, ongoing research
• Sample of a few: there are plenty more!

Main take-away:

NLP is far from solved yet!
DIFFICULT PROBLEMS IN SEMANTICS
Representing time

- Key concepts in NL: events and time
  - I arrived in New York
  - I am arriving in New York
  - I will arrive in New York
- Difficult meanings: different implications for inference
- Introduce event variables
- Use to reason about time

∃e, i. Arriving(e) ∧ Arriver(e, Speaker) ∧ Destination(e, NewYork) ∧ IntervalOf(e, i) ∧ End(i, e) ∧ Precedes(e, Now)

DIFFICULT PROBLEMS IN SEMANTICS
Aspect

- Related to temporal semantics
- Covers various distinctions for events
  - Finished vs. ongoing
  - Point in time vs. interval
  - Causing change of state vs. not
- Some examples:
  - I need the cheapest fare
  - He booked me a reservation
  - He found her gate
- Subtle distinctions, but can crucially impact inference
- Need to include in semantic representations

QUANTIFIER SCOPING
Every restaurant has a menu

- Can be distinguished by FOL representation

\[ \forall x. \text{Restaurant}(x) \Rightarrow \exists y. (\text{Menu}(y) \wedge \text{Has}(x, y)) \]
\[ \exists y. \text{Menu}(y) \wedge \forall x. (\text{Restaurant}(x) \Rightarrow \text{Has}(x, y)) \]

SCOPE AMBIGUITY

- No syntactic distinction
- Same composition must yield ambiguous readings
- Build into composition process:
  - generate multiple readings
  - generate one reading that represents ambiguity

NEGATION

- Negation poorly represented by FOL
- Meaning rarely what naive FOL composition would suggest
- Complicates quantifier rules

\[ \neg (\text{Enoch likes any farmer}) \neq \text{Enoch doesn’t like any farmer} \]

- Often depends on pragmatics

‘There is nothing wrong with the aircraft’
\[ \Rightarrow \text{worry!} \]
**QUANTIFIERS**

- Some combinations only produce subset of readings
- Hard to define rules for which readings are ‘permitted’
- Quite language specific
- Need special formal mechanisms
- But humans do it with little difficulty
- Loads of theoretical work, still not solved!

**QUANTIFIER SCOPE**

- Very brief taster here
- Plenty of literature if you want to read more

---

**PRAGMATICS**

Real human language processing not like most NLP tasks

<table>
<thead>
<tr>
<th>Newswire</th>
<th>Discourse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Following a three-week trial at Bristol Crown Court, jurors told Judge Martin Picton they could not reach a verdict.</td>
<td>A: Some of them disagree, individual utterances useless some of the time. B: Either interpret structure of dialogue, by speech acts, context...majority. B: So I didn’t know if that was just something for drama or that’s truly the way it is. A: I think it does have to be unanimous. B: So I’ll. B: But uh rather interesting.</td>
</tr>
</tbody>
</table>

**DOABLE PRAGMATICS TASKS**

- Certain pragmatics problems tackled in NLP
- Particularly in discourse processing
- Language use not isolated, unrelated sentences
- Every restaurant has a menu
- Structured, coherent sequences
- We sampled the local cuisine. Every restaurant has a menu and each one featured different specialities.
- Relatively doable subtasks:
  1. Coreference resolution
  2. Discourse structure

**COREFERENCE RESOLUTION**

- These are referring expressions
- Refer to judge in California – referent (person)
- Anaphora: referring back to previously mentioned referent
- Often uses pronouns: he, she, ...
- Some references license others:
  a judge in California is antecedent for he

---

**COREFERENCE EXAMPLE**

It belongs to Dennis Coffey, a 78-year-old who still rocks his wah-wah licks once a week. Dennis doesn’t appear in Motown: The Musical. But he played on at least 100 million-selling records, and there was a year when he was on three of the Top 10 every month.

Pronouns most common form of reference
### NOT JUST ENTITIES

References not just to entities/individuals

According to Doug, Sue just bought a 1961 Ford Falcon.

1. *But that* was a lie.
2. *But that* was false.
3. *That* struck me as a funny way to describe the situation.

### PRONOUN RESOLUTION

Some clear rules you might use (in English):

- **Number agreement**
- **Person agreement**
- **Gender agreement**
- **Syntactic constraints:** binding theory – complex

John has a cup. It is red. * John has two cups. It is red.

Alice walks fast. She is late. * I walk fast. She is late.

John has a cup. He is attractive. John has a cup. It is attractive.

John said that Bill bought him a ticket.

John said that Bill bought himself a ticket.

### PRONOUN RESOLUTION

Some softer clues you might use – features:

- **Recency:** prefer things last mentioned recently
- **Grammatical role:** prefer last mentioned as subject
- **Repeated mention:** prefer mentioned often (discourse focus)
- **Parallelism:** prefer matching grammatical roles
- **Semantic knowledge:** of verbs and referents

Leo found an old map. Mark found an even older one. It described an island.

### PRONOUN RESOLUTION

Some softer clues you might use – features:

- **Recency:** prefer things last mentioned recently
- **Grammatical role:** prefer last mentioned as subject
- **Repeated mention:** prefer mentioned often (discourse focus)
- **Parallelism:** prefer matching grammatical roles
- **Semantic knowledge:** of verbs and referents

Leo thought. He looked around. He really wanted a drink. Mark had already sat down. He ordered a beer.

### PRONOUN RESOLUTION

Some softer clues you might use – features:

- **Recency:** prefer things last mentioned recently
- **Grammatical role:** prefer last mentioned as subject
- **Repeated mention:** prefer mentioned often (discourse focus)
- **Parallelism:** prefer matching grammatical roles
- **Semantic knowledge:** of verbs and referents

Leo went with Mark to Oljenkorsi. Khalid went with him to Jano.
**PRONOUN RESOLUTION**

Some softer clues you might use – features:
- **Recency**: prefer things last mentioned recently
- **Grammatical role**: prefer last mentioned as subject
- **Repeated mention**: prefer mentioned often (discourse focus)
- **Parallelism**: prefer matching grammatical roles
- **Semantic knowledge**: of verbs and referents

Mark put *his beer glass* on the table after drinking from it.

**DIFFICULT CASES**

What makes these hard?

The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress.

The boy entered the room. The door closed automatically.

Georges Cipriani left a prison in Ensisheim on parole on Wednesday. He departed the prison in a police vehicle bound for an open prison near Strasbourg.

**DOABLE PRAGMATICS TASKS**

- Certain pragmatics problems tackled in NLP
- Language use not isolated, unrelated sentences
  
  *Every restaurant has a menu*

- **Structured, coherent** sequences
  
  *We sampled the local cuisine. Every restaurant has a menu and each one featured different specialities.*

- Relatively doable subtasks:
  1. Coreference resolution
  2. Discourse structure

**LARGER-SCALE STRUCTURE**

- Most documents have high-level structure
  
  *Sections in scientific papers*

- News articles: summary at start, background later, . . .

- Stories: (mostly) temporal order

- Different conventions

- Topics discussed by document

- Tried to model with **topic modelling**

- Explicit sections, paragraphs, . . .

- Relationships between sentences

- Seen in **NLG** for structuring output

**DISCOURSE COHERENCE**

- Longer discourse: **coherence** between sentences/utterances

  *John hid Bill’s car keys. He was drunk.*

  *John hid Bill’s car keys. He likes spinach.*

- Different types of **relations**

  *I went to the shop. I had run out of milk.*

  *I went to the shop. I was out when you called.*

- Sometimes signalled explicitly

  *I went to the shop, *because* I had run out of milk.*

  *I went to the shop, so I was out when you called.*

**AN APPROACH TO COREF**

- Similar to pronoun resolution, but more general

- Typical system:

  - Identify all possible **referring expressions**

  - Extract features from each

  - Apply classifier to all pairs: coreferent/not?

- Common approach: Names/my methods for classification

- **Pronouns**

  - Definite NPs

Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to $1.3M, 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.

**COREF SYSTEM EXAMPLE**

- **Hugging Face**

  - Similar classifier-based approach to above

  - Neural network

  - Uses **word embeddings** instead of manually encoded features

  - Designed to work with **informal dialogue**

**DISCOURSE COHERENCE**

I went to the shop. I had run out of milk.

I went to the shop. I was out when you called.

- Coherence relations often implicit

  *Max fell. John pushed him.*

- Inference is important

  *Max fell because John pushed him.*

  *Max fell and then John pushed him.*
John pushed him.

Following a three-week trial at Bristol Crown Court, jurors told Judge Martin Picton they could not reach a verdict.

Max fell because John pushed him.

Max tripped. He hit his head.

Max tripped. John stumbled.

Max fell. He tripped and took a terrible tumble.

Max fell down. He lay on the floor.

These can help, but interpretation remains a hard problem.

What else do we need to do?

COURSE FEEDBACK

Almost finished course

More to come still after break, but...

This course is brand new!

Course structure

Content, slides, materials

Assignments

We need your feedback for future versions!

During (long) break: fill in course feedback form:

Follow link in email

Log into WebOodi
OTHER HARD THINGS

- Seen some difficult problems and (partial) solutions
  - Formal semantics: quantifier scope, negation
  - Coreference
  - Discourse structure
- Now a few more difficult NLP challenges
  - Open problems
  - Topics of current research
  - Some (partial) solutions...

MULTIPLE LANGUAGES

- To some extent, same methods for other languages
- Linguistic differences easy to underestimate
- E.g. Chinese: what is a character, word, token?
- Still need data

<table>
<thead>
<tr>
<th>High-resource</th>
<th>Low-resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>Few languages</td>
<td>Very many languages</td>
</tr>
<tr>
<td>Much linguistic study</td>
<td>Typically less studied</td>
</tr>
<tr>
<td>Large corpora</td>
<td>Small or no corpora</td>
</tr>
<tr>
<td>Annotated resources</td>
<td>Limited or no annotations</td>
</tr>
<tr>
<td>Commercial interest</td>
<td>Less commercial incentive</td>
</tr>
</tbody>
</table>

MULTILINGUAL WORD EMBEDDINGS

Conneau et al. (2017). Word Translation Without Parallel Data

- Aligned word embeddings open up new possibilities for
  - multilingual methods
  - language transfer
- Use word embeddings instead of words
- Some methods easily transferred across languages
  - Bag-of-words methods
- Other things language specific: e.g.
  - Word order
  - Syntax
  - Multiword expressions

BROAD-DOMAIN NLP

Syntactic parsing is really good on news text

- But I want to parse forum posts, conversational speech, Estonian, ...
- Answer: annotate a treebank for these
- Penn Treebank took years, huge effort and lots of money
- How to produce parsers that work on all these and more?
- Major area of research
- But there are some promising directions...

MULTILINGUAL WORD EMBEDDINGS

- Recent ML method
- Context-based word embeddings language specific
- Can’t compare: different vector spaces
- Transform all vectors
- Don’t know what words should match
- If lots match something, probably good mapping
- Works amazingly well!

POSSIBLE SOLUTIONS

- Annotate data, invest in small languages
- Unsupervised, or minimally supervised, methods
  - Some useful methods: e.g. topic modelling
  - Realistic replacement for supervised in some areas
  - Often, much poorer quality
  - Hard to use output: what does it mean?
- Language transfer
  - Re-use what we know about high-resource language
  - Map language-specific parts, re-use others
  - E.g. POS tagger with word translations
  - Can work for closely related languages
- Multilingual methods...

MULTILINGUAL WORD EMBEDDINGS

- Align word embeddings open up new possibilities for
  - multilingual methods
  - language transfer
- Use word embeddings instead of words
- Some methods easily transferred across languages
  - Bag-of-words methods
- Other things language specific: e.g.
  - Word order
  - Syntax
  - Multiword expressions

BROAD-DOMAIN PARSEING

Some approaches:

- Unknown words: backoff to POS tags
- Can make good guesses from context
- Is POS tagger domain independent?
- Domain transfer
  - Build grammar/parser for one domain
  - Extend to others with unlabelled data
  - Can method discover new rules/constructions?
- Transfer/expansion on particularly domain-specific parts
  - E.g. (automatically) expand lexicon
  - Keep same rules
**SPEECH**

- Speech processing: **speech recognition** at start of pipeline
- Introduces further challenges:
  - More ambiguity
  - More noise
  - Disfluency
  - Wider variety of language (informal, dialectal)

**DISFLUENCY**

- Fillers: *ya know*, *well*, *totaaa*
- Other non-lexical sounds: *hmm*, *um*, grunts, coughs
- Repairs:
  
  *the current plan is we take – okay let’s start with the bananas*
  
  **Alice saw the dog um the duck**

All cause problems for standard tools. Proper handling essential to interpretation

---

**SPEECH REPAIRS**

- One type of disfluency
- Can be categorized:
  1. **Fresh start**
     
     *the current plan is we take – okay let’s start with the bananas* → *let’s start with the bananas*
  2. **Modification**
     
     *A flight to um Berlin I mean Munich* → *A flight to Munich*
  3. **Abridged repair**
     
     *we need to um get the bananas* → *we need to get the bananas*
- In some cases, identification permits cleaning

---

**CHALLENGES OF NLP ON SPEECH**

- **Ambiguity**:
  - Many ways to interpret speech input
  - Choosing one-best may restrict later analysis
  - E.g. can’t parse
  - Interaction between SR and later components
- **Language type** (register):
  - People don’t speak like a newspaper
  - NLP tools must be able to handle informal language
  - Same issue with many text sources
- **Disfluency**
  - Real speech is not fluent: taster here

---

**BROAD-DOMAIN PARSING**

Some approaches:

- **Train on multiple domains**
  - Maybe learn domain-general characteristics
  - Better performance on new domain
- **Or dynamically select** domain(s) when parsing
- Separate **domain-general** and **domain-specific** knowledge while learning
  - Reuse general parts
  - Adapt/ignore specific parts

Many solutions apply to same domain-specificity problem on other tasks

---

**WHAT WE’VE COVERED**

- End of course imminent!
- Brief reminder of what we’ve seen...
FINITE-STATE METHODS AND MORPHOLOGY

- FSAs and FSTs
- Use of FSTs for morphological analysis
- Applications of morphology
- Expressivity of formalisms, Chomsky Hierarchy

STATISTICAL MODELS

Help model ambiguity / uncertainty
- Confidences derived from data
- Combinatorial effects – influence of context: e.g. PCFG
- Express uncertainty in output
- Statistics over previous analyses to estimate confidences
- Probabilistic models (PCFG, HMM, …)

SYNTAX

- Syntactic structure of sentences
- Hidden structures underlying semantic interpretation
- Formal syntax: grammar formalisms
- CFGs & dependencies
- Chart parsing & dependency parsing
- Statistical parsing: PCFGs

EVALUATION

- Evaluation important for:
  - Public review
  - Internal review, quantification of improvements
- Unbiased comparison of models, data, methods
- Intrinsic vs extrinsic
- Shared tasks and common test sets
- Metrics, baselines & ceilings
- Significance testing in NLP
- Splitting test data: training, test, development, cross-val
- Error analysis: confusion matrices

LANGUAGE MODELLING

- Probability distribution over sentences
- Markov-chain LM
- N-gram models of words
- Applications in: speech recognition, predictive input devices, spelling correction, …

NLG EVALUATION

- What does ‘correct’ even mean?
- Multi-method correct are a must!

Automated metrics
  - Word-overlap → fragile
  - Assume large, representative set of refs
  - Ref production can go wrong

Human evals
  - Intrinsic: ask them what they think
  - Extrinsic: watch whether they do the ‘right’ thing

DIALOGUE SYSTEMS

Three subtasks
  - NLU (stateless)
  - Dialogue management
  - What are the user and system trying to achieve?
  - What should be done next?
  - NLG (stateless)

Different purposes
  - Task-oriented: trying to achieve something specific
  - Non-task-oriented: just chatting
INFORMATION RETRIEVAL

• IR systems
• Document-term matrices → VSMs
• Bag of words
• Word importance: tf-idf

Search query
Information need
IR system
Document collection
Result (documents)
Relevant?

INFORMATION EXTRACTION

Traditional approach
• Well-defined task
• Cascade of patterns match pieces of information
• Create manually or use bootstrapping

Modern approach
• Loosely-defined task
• Extract all possible knowledge
• Constraints on pattern shape still exist
• Classification / clustering of extracted information

ADVANCED STATISTICAL MODELS

• Classification & regression in NLP
• Sentiment analysis
• SVMs
• Topic modelling: LDA
• Neural networks
• Sequence models (RNNs)
• Application to LM
• and sentiment analysis

LEXICAL SEMANTICS

• Lexical semantics: WordNet
• Word sense disambiguation (WSD)
• Distributional semantics → word vectors
• Word embeddings
• Other embedding methods: Word2vec, . . .

INFORMATION EXTRACTION

• Information extraction (IE): unstructured information in text → structured data
• Goals:
  • present information in compact, comprehensible form
  • accumulate data in knowledge bases
  • input to other data mining techniques
• Key ideas:
  • Usually complete text understanding impossible / unnecessary
  • Define what’s needed, search text for specific information

SEMANTICS & PRAGMATICS

• Formal semantics: predicate logic
• Difficult problems in semantics:
  • Time
  • Aspect
  • Quantifier scoping
  • Negation
• Pragmatics: potentially doable tasks:
  • Coreference resolution
  • Discourse structure

OPEN NLP PROBLEMS

A few examples:
• Pragmatics in real language
• Multiple languages
• Broad-domain processing
• Speech: disfluency / repairs

NLP is not solved yet!
NLP TODAY

- Hopefully this course has convinced you that:
  - NLP is not a hopeless task
  - existing work has a lot of useful applications
- Hopefully today has convinced you that:
  - NLP is a very hard task
  - there are tonnes of unsolved problems
- We can already do a lot with it
  - Recent years: much more useful for real-world applications

Now is a good time to be an NLP expert!
And now you are one . . .

READING MATERIAL

- Formal semantics for NLP
  - J&M3 ch 14
  - Representation and Inference for Natural Language (Blackburn & Bos, 2005)
- Scope, negation, . . . :
  - Taking Scope: The Natural Semantics of Quantifiers (Steedman)
- Coreference resolution: J&M2, 21.3
- Discourse coherence: J&M2, 21.2
- Speech: J&M2, Part II

SEMINAR: NLP AND NEWS

- New seminar course this autumn
- Follows up on this: NLP background required
- Focus on methods for news
  - Topical news analysis
  - Historical news exploration
  - News report generation (NLG)
- DATA20018
- Course information to appear soon . . .

NEXT UP

After lunch: practical assignments in BK107

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:15 – 12:00</td>
<td>Lectures</td>
</tr>
<tr>
<td>12:00 – 13:15</td>
<td>Lunch</td>
</tr>
<tr>
<td>13.15 – ~13:30</td>
<td>Introduction</td>
</tr>
<tr>
<td>13:30 – 16:00</td>
<td>Practical assignments</td>
</tr>
</tbody>
</table>

- Building a bigger system
- Using NLP components
- Various suggestions: you choose!