Multiword Expressions & Semantic Roles

CMSC 723 / LING 723 / INST 725

MARINE CARPUAT
marine@cs.umd.edu
• Q: what is understanding meaning?
• A: predicting relations between words (similarity, entailment, synonymy, hypernymy ...)

Approaches:
• Learn from raw text vs. thesaurus/wordnet
• Supervised vs. unsupervised
Today

- From word meaning to sentence meaning
  - Semantic Role Labeling [Textbook: 20.9]

- When minimal unit of analysis are not words
  - Multiword Expressions [Not in Textbook]
SEMANTIC ROLE LABELING

Slides Credit: William Cohen, Scott Yih, Kristina Toutanova
Yesterday, Kristina hit Scott with a baseball

Scott was hit by Kristina yesterday with a baseball

Yesterday, Scott was hit with a baseball by Kristina

With a baseball, Kristina hit Scott yesterday

Yesterday Scott was hit by Kristina with a baseball

Kristina hit Scott with a baseball yesterday
Syntactic Variations

Kristina hit Scott with a baseball yesterday

With a baseball, Kristina hit Scott yesterday
Semantic Role Labeling – Giving Semantic Labels to Phrases

• \([\text{AGENT John}] \text{ broke} \ [\text{THEME the window}]\)

• \([\text{THEME The window}] \text{ broke}\)

• \([\text{AGENT Sotheby’s]} \text{ .. offered} \ [\text{RECIPIENT the Dorrance heirs}] \ [\text{THEME a money-back guarantee}]\)

• \([\text{AGENT Sotheby’s}] \text{ offered} \ [\text{THEME a money-back guarantee}] \text{ to} \ [\text{RECIPIENT the Dorrance heirs}]\)

• \([\text{THEME a money-back guarantee}] \text{ offered} \text{ by} \ [\text{AGENT Sotheby’s}]\)

• \([\text{RECIPIENT the Dorrance heirs}] \text{ will} \ [\text{ARM-NEG not}] \text{ be offered} \ [\text{THEME a money-back guarantee}]\)
Why is SRL Important – Applications

• Question Answering
  – Q: When was Napoleon defeated?
  – Look for: [PATIENT Napoleon] [PRED defeat-synset] [ARGM-TMP *ANS*]

• Machine Translation
  English (SVO)  Farsi (SOV)
  [AGENT The little boy]  [AGENT pesar koocholo]  boy-little
  [PRED kicked]  [THEME toop germezi]  ball-red
  [THEME the red ball]  [ARGM-MNR moqtam]  hard-adverb
  [ARGM-MNR hard]  [PRED zaad-e]  hit-past

• Document Summarization
  – Predicates and Heads of Roles summarize content

• Information Extraction
  – SRL can be used to construct useful rules for IE
SRL: REPRESENTATIONS & RESOURCES
FrameNet [Fillmore et al. 01]

Frame: Hit_target
(hit, pick off, shoot)

Lexical units (LUs):
Words that evoke the frame (usually verbs)

Frame elements (FEs):
The involved semantic roles

[Agent Kristina] hit [Target Scott] [Instrument with a baseball] [Time yesterday].
Methodology for FrameNet

1. Define a frame (e.g. DRIVING)
2. Find some sentences for that frame
3. Annotate them

- Corpora
  - FrameNet I – British National Corpus only
  - FrameNet II – LDC North American Newswire corpora

- Size
  - >8,900 lexical units, >625 frames, >135,000 sentences

http://framenet.icsi.berkeley.edu
Proposition Bank (PropBank) [Palmer et al. 05]

• Transfer sentences to propositions
  – Kristina hit Scott $\rightarrow$ hit(Kristina, Scott)

• Penn TreeBank $\rightarrow$ PropBank
  – Add a semantic layer on Penn TreeBank
  – Define a set of semantic roles for each verb
  – Each verb’s roles are numbered

…[A0 the company] to … offer [A1 a 15% to 20% stake] [A2 to the public]
…[A0 Sotheby’s] … offered [A2 the Dorrance heirs] [A1 a money-back guarantee]
…[A1 an amendment] offered [A0 by Rep. Peter DeFazio] …
…[A2 Subcontractors] will be offered [A1 a settlement] …
Proposition Bank (PropBank)  
Define the Set of Semantic Roles  

• It’s difficult to define a general set of semantic roles for all types of predicates (verbs).
• PropBank defines semantic roles for each verb and sense in the frame files.
• The (core) arguments are labeled by numbers.  
  – A0 – Agent; A1 – Patient or Theme  
  – Other arguments – no consistent generalizations
• Adjunct-like arguments – *universal* to all verbs  
  – AM-LOC, TMP, EXT, CAU, DIR, PNC, ADV, MNR, NEG, MOD, DIS
Proposition Bank (PropBank) Frame Files

- **hit.01 “strike”**
  - A0: agent, hitter; A1: thing hit;
  - A2: instrument, thing hit by or with
  - \[\text{\textit{Kristina}} \text{hit} \text{\textit{Scott}} \text{\textit{with a baseball}} \text{yesterday}.\]

- **look.02 “seeming”**
  - A0: seemer; A1: seemed like; A2: seemed to
  - \[\text{\textit{It}} \text{looked} \text{\textit{to her}} \text{like} \text{\textit{he deserved this}}.\]

- **deserve.01 “deserve”**
  - A0: deserving entity; A1: thing deserved;
  - A2: in-exchange-for
  - \textit{It looked to her like } \text{\textit{he}} \text{\textit{deserved} \textit{this}}.
FrameNet vs PropBank

FrameNet Annotation:

[Buyer Chuck] bought [Goods a car] [Seller from Jerry] [Payment for $1000].

PropBank Annotation:

[Arg0 Chuck] bought [Arg1 a car] [Arg2 from Jerry] [Arg3 for $1000].

[Arg0 Jerry] sold [Arg1 a car] [Arg2 to Chuck] [Arg3 for $1000].
FrameNet vs PropBank -2

FRAMENET ANNOTATION:

[Goods A car] was bought [Buyer by Chuck].

[Goods A car] was sold [Buyer to Chuck] [Seller by Jerry].

[Buyer Chuck] was sold [Goods a car] [Seller by Jerry].

PROPBNANK ANNOTATION:

[Arg1 A car] was bought [Arg0 by Chuck].

[Arg1 A car] was sold [Arg2 to Chuck] [Arg0 by Jerry].

[Arg2 Chuck] was sold [Arg1 a car] [Arg0 by Jerry].
Kristina hit Scott with a baseball yesterday.

\[
[A_0 \text{Kristina}] \text{hit} [A_1 \text{Scott}] [A_2 \text{with a baseball}] [\text{AM-TMP yesterday}].
\]
Proposition Bank (PropBank) Statistics

• Proposition Bank I
  – Verb Lexicon: 3,324 frame files
  – Annotation: ~113,000 propositions
    http://www.cis.upenn.edu/~mpalmer/project_pages/ACE.htm

• Alternative format: CoNLL-04,05 shared task
  – Represented in table format
  – Has been used as standard data set for the shared tasks on semantic role labeling
    http://www.lsi.upc.es/~srlconll/soft.html
SRL: TASKS & SYSTEMS
Semantic Role Labeling: Subtasks

- **Identification**
  - Very hard task: to separate the argument substrings from the rest in this exponentially sized set
  - Usually only 1 to 9 (avg. 2.7) substrings have labels ARG and the rest have NONE for a predicate

- **Classification**
  - Given the set of substrings that have an ARG label, decide the exact semantic label

- **Core argument** semantic role labeling: (easier)
  - Label phrases with core argument labels only. The modifier arguments are assumed to have label NONE.
Evaluation Measures

Correct: \([_{A0} \text{The queen}] \text{ broke} \ [_{A1} \text{the window}] \ [_{AM-TMP} \text{yesterday}]\)
Guess: \([_{A0} \text{The queen}] \text{ broke the} \ [_{A1} \text{ window}] \ [_{AM-LOC} \text{yesterday}]\)

<table>
<thead>
<tr>
<th>Correct</th>
<th>Guess</th>
</tr>
</thead>
<tbody>
<tr>
<td>{The queen} \rightarrow A0</td>
<td>{The queen} \rightarrow A0</td>
</tr>
<tr>
<td>{the window} \rightarrow A1</td>
<td>{window} \rightarrow A1</td>
</tr>
<tr>
<td>{yesterday} \rightarrow AM-TMP</td>
<td>{yesterday} \rightarrow AM-LOC</td>
</tr>
<tr>
<td>all other \rightarrow NONE</td>
<td>all other \rightarrow NONE</td>
</tr>
</tbody>
</table>

- Precision, Recall, F-Measure
- Measures for subtasks
  - Identification (Precision, Recall, F-measure)
  - Classification (Accuracy)
  - Core arguments (Precision, Recall, F-measure)
What information can we use for Semantic Role Labeling?

- Syntactic Parsers

- Shallow parsers

- Semantic ontologies (WordNet, automatically derived), and named entity classes

  (v) **hit** (cause to move by striking)

  *WordNet hyponymph*:

  ![propel, impel](cause to move forward with force)
Arguments often correspond to syntactic constituents!

Most commonly, substrings that have argument labels correspond to syntactic constituents

- In Propbank, an argument phrase corresponds to exactly one parse tree constituent in the correct parse tree for 95.7% of the arguments;
- In Propbank, an argument phrase corresponds to exactly one parse tree constituent in Charniak’s automatic parse tree for approx 90.0% of the arguments.
- In FrameNet, an argument phrase corresponds to exactly one parse tree constituent in Collins’ automatic parse tree for 87% of the arguments.
Labeling Parse Tree Nodes

• Given a parse tree $t$, label the nodes (phrases) in the tree with semantic labels
• To deal with discontiguous arguments
  – In a post-processing step, join some phrases using simple rules
  – Use a more powerful labeling scheme, i.e. C-A0 for continuation of A0
Combining Identification and Classification Models

**Step 1. Pruning.**
Using a hand-specified filter.

**Step 2. Identification.**
Identification model (filters out candidates with high probability of NONE)

**Step 3. Classification.**
Classification model assigns one of the argument labels to selected nodes (or sometimes possibly NONE)
Combining Identification and Classification Models

\[
\neg P(l|c, t, p) = P_{ID}(Id(l)|\Phi(c, t, p)) \ast P_{CLS}(l|Id(l), \Phi(c, t, p)) \\
\text{or} \\
\neg P(l|c, t, p) = P(l|\Phi(c, t, p))
\]

One Step. Simultaneously identify and classify using \( P(l|c, t, p) \)
Gildea & Jurafsky (2002) Features

• Key early work
  – Future systems use these features as a baseline

• Constituent Independent
  – Target predicate (lemma)
  – Voice
  – Subcategorization

• Constituent Specific
  – Path
  – Position (left, right)
  – Phrase Type
  – Governing Category (S or VP)
  – Head Word

```
She                     broke       the      expensive  vase
PRP  VBD  DT  JJ  NN

S
NP
VP
NP

Target     broke
Voice      active
Subcategorization VP→VBD NP
Path       VBD↑VP↑S↓NP
Position    left
Phrase Type NP
Gov Cat     S
Head Word   She
```
Performance with Baseline Features using the G&J Model

• Features combined using a linear classifier

FrameNet Results

Propbank Results
Improving performance with better learning + better features

• Better Machine Learning: 67.6 → 80.8 using SVMs [Pradhan et al. 04])

• Better features
  ▪ Head Word and Content Word POS tags
  ▪ **NE labels (Organization, Location, etc.)**
  ▪ Structural/lexical context
  ▪ Head of PP Parent
  ▪ If the parent of a constituent is a PP, the identity of the preposition
Pradhan et al. (2004) Features

• More (31% error reduction from baseline due to these + Surdeanu et al. features)
Joint Scoring: Enforcing Hard Constraints

• **Constraint 1: Argument phrases do not overlap**

  By \([A_1 \text{ working } A_1 \text{ hard }], \text{ he} \text{ said}, \text{ you can achieve a lot.}\)
  
  – Pradhan et al. (04) – greedy search for a best set of non-overlapping arguments
  
  – Toutanova et al. (05) – exact search for the best set of non-overlapping arguments (dynamic programming, linear in the size of the tree)
  
  – Punyakanok et al. (05) – exact search for best non-overlapping arguments using integer linear programming

• **Other constraints** ([Punyakanok et al. 04, 05])

  – no repeated core arguments (good heuristic)
  
  – phrases do not overlap the predicate
  
  – *(more later)*
There are many statistical tendencies for the sequence of roles and their syntactic realizations
- When both are before the verb, AM-TMP is usually before A0
- Usually, there aren’t multiple temporal modifiers
- Many others which can be learned automatically
Per Argument Performance
CoNLL-05 Results on WSJ-Test

• Core Arguments (Freq. ~70%)

<table>
<thead>
<tr>
<th></th>
<th>Best ( F_1 )</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0</td>
<td>88.31</td>
<td>25.58%</td>
</tr>
<tr>
<td>A1</td>
<td>79.91</td>
<td>35.36%</td>
</tr>
<tr>
<td>A2</td>
<td>70.26</td>
<td>8.26%</td>
</tr>
<tr>
<td>A3</td>
<td>65.26</td>
<td>1.39%</td>
</tr>
<tr>
<td>A4</td>
<td>77.25</td>
<td>1.09%</td>
</tr>
</tbody>
</table>

• Adjuncts (Freq. ~30%)

<table>
<thead>
<tr>
<th></th>
<th>Best ( F_1 )</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMP</td>
<td>78.21</td>
<td>6.86%</td>
</tr>
<tr>
<td>ADV</td>
<td>59.73</td>
<td>3.46%</td>
</tr>
<tr>
<td>DIS</td>
<td>80.45</td>
<td>2.05%</td>
</tr>
<tr>
<td>MNR</td>
<td>59.22</td>
<td>2.67%</td>
</tr>
<tr>
<td>LOC</td>
<td>60.99</td>
<td>2.48%</td>
</tr>
<tr>
<td>MOD</td>
<td>98.47</td>
<td>3.83%</td>
</tr>
<tr>
<td>CAU</td>
<td>64.62</td>
<td>0.50%</td>
</tr>
<tr>
<td>NEG</td>
<td>98.91</td>
<td>1.36%</td>
</tr>
</tbody>
</table>

Data from Carreras&Màrquez’s slides (CoNLL 2005)
What are Multi Word Expressions?

– Decomposable into multiple words

– Lexically, syntactically, semantically, pragmatically and/or statistically idiosyncratic
Some examples

San Francisco
ad hoc
by and large
part of speech
take a walk
take advantage of
call (someone) up
Why do we care?

• MWEs are pervasive
  – Estimated to be equivalent in number to simplex words in mental lexicon

• MWEs are a challenge to NLP systems
MWE or not MWE?

“there is no unified phenomenon to describe but rather a complex of features that interact in various, often untidy, ways and represent a broad continuum between non-compositional (or idiomatic) and compositional groups of words.”

[Moon 1998]
Indicators of MWE-thood

• Institutionalization/conventionalization

• Lexicogrammatical fixedness:
  – Formal rigidity, preferred lexical realization, restrictions on voice, etc

Fixed MWE: kick the bucket
Non-fixed MWE: keep tabs on
Indicators of MWE-hood

• Semantic non-compositionality
  – Mismatch between semantics of the parts and the whole
    Kick the bucket (but also: At first)

• Syntactic irregularity
  – all of a sudden, the be all and end all of
  – (but also: kick the bucket, fly off the handle)
Indicators of MWE-hood

• Non-identifiability: meaning cannot be predicted from surface form
  – kick the bucket, fly off the handle
  – (but also: wide awake, plain truth)
Indicators of MWE-hood

• Situatedness: expression situated with a fixed pragmatic point
  – Good morning, all aboard
  – But also: first off

• Figuration: expression encodes some metaphor, metonymy, hyperbole
  – Figurative expressions: bull market
  – Non figurative expressions: first off
Indicators of MWE-hood

• Single-word paraphrasability: the expression has a single word paraphrase
  – Leave out = omit
  – (but also: look up)

• Informality:
  – Expression associated with more informal or colloquial registers

• Affect
  – Expression encodes a certain evaluation of affective stance toward the thing it denotes
Indicators of MWE-hood

• Substitutability: MWEs stand in opposition to anti-collocations
  – Expressions derived through synonym/word order substitution which occur with markedly lower frequency than the MWE

many thanks
*several thanks
*many gratitudes
Concept of “Multiword”

• ~ a lexeme that crosses word boundaries

• Complications
  – non-segmenting languages
  – Languages without a pre-existing writing system

• But there is fuzziness even in English
  – Houseboat vs. house boat
  – Trade off vs. trade-off vs. tradeoff
<table>
<thead>
<tr>
<th>Expression</th>
<th>MWE?</th>
</tr>
</thead>
<tbody>
<tr>
<td>library card</td>
<td></td>
</tr>
<tr>
<td>at arm’s length</td>
<td></td>
</tr>
<tr>
<td>old tree</td>
<td></td>
</tr>
<tr>
<td>foreign direct investment</td>
<td></td>
</tr>
<tr>
<td>the sun</td>
<td></td>
</tr>
<tr>
<td>at [nine] o’clock</td>
<td></td>
</tr>
<tr>
<td>to go bush</td>
<td></td>
</tr>
<tr>
<td>give a demo</td>
<td></td>
</tr>
<tr>
<td>kick the bucket</td>
<td></td>
</tr>
<tr>
<td>once upon a time</td>
<td></td>
</tr>
<tr>
<td>at home</td>
<td></td>
</tr>
<tr>
<td>in the meantime</td>
<td></td>
</tr>
<tr>
<td>to read Shakespeare</td>
<td></td>
</tr>
</tbody>
</table>
MWEs vs. Collocations

• A collocation is an arbitrary and recurrent word combination

• Tends to be compositional (e.g., strong coffee)

• Generally contiguous word sequences (often bigrams)
Brainstorming Exercise

How can we identify MWEs automatically?
Today

• From word meaning to sentence meaning
  • Semantic Role Labeling [Textbook: 20.9]

• When minimal unit of analysis are not words
  • Multiword Expressions [Not in Textbook]