The Proposition Bank
An Annotated Corpus of Semantic Roles

TzuYi Kuo
EMLCT
Saarland University
June 14, 2010
Outline

• Introduction
• Motivation
• PropBank
  – Semantic role
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Outline

• Introduction
• Motivation
• PropBank
  – Semantic Role
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Introduction

• Represent the full meaning of sentences

• Alternation
  – Syntactic realization of semantic arguments

(1) John broke the window.
(2) The window broke.
Introduction

• Represent the full meaning of sentences

• Alternation
  – Syntactic realization of semantic arguments

(1) John broke the window.
(2) The window broke.

same underlying semantic role
Introduction

Proposition Bank

Predicate-argument information

Penn Treebank
Introduction

• Focus on
  – Argument structure of verbs
  – Provide a complete corpus annotated with semantic roles

• Goal
  – Provide a broad-coverage hand-annotated corpus for supervised automatic role labelers
  – Show how and why these syntactic alternations take place
Outline

• Introduction
• Motivation
• PropBank
  – Semantic Role
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Motivation

• Inspired by Levin (1993)
  – Research into the linking between semantic roles and syntactic realization
  – Syntactic frames are a direct reflection of the underlying semantics
  – Define verb classes
    • Based on the ability of particular verbs
    • In syntactic frames
Motivation

- **VerbNet** (Kipper et al., 2000)
  - Extend Levin’s classes
    - Adding an abstract representation of the syntactic frames for each class
    - Correspond between syntactic positions and the semantic roles they express
  
  *Ex. Break*
  
  Agent REL Patient
  Patient REL into pieces
Outline

• Introduction
• Motivation
• PropBank
  – Semantic Role
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
PropBank

• From sentences to propositions

John met Mary.

John and Mary met.

John met with Mary.

John and Mary had a meeting.

Proposition: meet(John, Mary)
Outline

• Introduction
• Motivation
• PropBank
  – Semantic Role
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Semantic Role

• Difficult to define a universal set of semantic roles covering all types of predicates

• Verb-by-verb basis
  – Arg0 → Agent
  – Arg1 → Prototypical Patient
Semantic Role

• Verb-specific numbered role

Frameset accept.01 “take willingly”

Arg0: Acceptor
Arg1: Thing accepted
Arg2: Accepted-from
Arg3: Attribute
Semantic Role

• Verb-specific numbered role

Frameset accept.01 “take willingly”

Arg0: Acceptor
Arg1: Thing accepted
Arg2: Accepted-from
Arg3: Attribute
Semantic Role

Attempt to cover the range of syntactic alternations afforded by the usage
Outline

• Introduction
• PropBank
  – Semantic Role
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Framing

• Distinguishing Framesets
  – Different numbers of arguments

(15) Frameset decline.01 “go down incrementally”

Arg1: entity going down
Arg2: amount gone down by, EXT
Arg3: start point
Arg4: end point
Ex: ... [Arg1 its net income] declining [Arg2-EXT 42%] [Arg4 to $121 million] [ArgM-TMP in the first 9 months of 1989]. (wsj_0067)

(16) Frameset decline.02 “demure, reject”

Arg0: agent
Arg1: rejected thing
Ex: [Arg0 A spokesman] declined [Arg1 *trace* to elaborate] (wsj_0038)
Framing

• Distinguishing Framesets
  – Verb-particle

Frameset cut.01 “slice”
Arg0: cutter
Arg1: thing cut
Arg2: medium, source
Arg3: instrument
Ex: [Arg0] Longer production runs]
  [Arg2] would cut [Arg1] inefficiencies
  from adjusting machinery
  between production cycles]. (wsj_0317)

Frameset cut.05 “cut back = reduce”
Arg0: cutter
Arg1: thing reduced
Arg2: amount reduced by
Arg3: start point
Arg4: end point
Ex: “Whoa,” thought John,
  “[Arg0] I’ve got [Arg0] *trace* to start
  [Arg0] *trace* cutting back
  [Arg1] my intake of chocolate].}
Framing

• Distinguishing Framesets
  – Different syntactic type

Frameset see.01 “view”

Arg0: viewer
Arg1: thing viewed

Ex1: \([\text{Arg0}, \text{John}] \text{ saw } [\text{Arg1}, \text{the President}]\) NP

Ex2: \([\text{Arg0}, \text{John}] \text{ saw } [\text{Arg1}, \text{the President collapse}]\) Clause object
Framing

• Secondary Predications

**predicative reading**
Mary called John a doctor.

(Arg0: Mary
Rel: called
Arg1: John (item being labeled)
Arg2-PRD: a doctor (attribute))

**ditransitive reading**
Mary called John a doctor.

(Arg0: Mary
Rel: called
Arg2: John (benefactive)
Arg1: a doctor (thing summoned))
Framing

- Traces
  - Empty category which known as trace
  - Coindex with other constituents in tree

[Arg0 John] tried [Arg0 *trace*] to kick [Arg1 the football], but Mary pulled it away at the last moment.
Framing

• Frames file
  – the collection of framesets for each lexeme

... [Arg1, the company] to ... *offer* [Arg2, a 15% to 20% stake] [Arg2, to the public] (wsj_0345)

... [Arg2, Sotheby’s] ... *offered* [Arg2, the Dorrance heirs] [Arg1, a money-back guarantee] (wsj_1928)

... [Arg2, an amendment] *offered* [Arg1, by Rep. Peter DeFazio] ... (wsj_0107)

... [Arg1, Subcontractors] will be *offered* [Arg1, a settlement] ... (wsj_0187)
Framing

• In Wall Street Journal
  – Over 3,300 verbs framed
  – 4,500 framesets described
  – Average polysemy of 1.36
  – Each instance of a polysemous verb is marked as to which frameset it belongs to
  – Interannotator (ITA) agreement of 94%
Outline

• Introduction
• Motivation
• PropBank
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Development Process

• Annotation
  – Rule-based argument tagger (Palmer, Rosenzweig, and Cotton 2001)
    • Class-based mappings between grammatical and semantic roles
    • 83% accuracy
    • The output is then corrected by hand
      – Examining the descriptions of the arguments and the example tagged sentences
Development Process

• Annotation
  – Kappa statistic (Siegel and Castellan, 1988)
    • Measure agreement between annotators

\[ \kappa = \frac{P(A) - P(E)}{1 - P(E)} \]

– P(A) : the probability of inter-annotator agreement
– P(E) : the agreement expected by chance
Development Process

• Annotation
  – Kappa statistic

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Interannotator agreement.</th>
<th></th>
<th></th>
<th>κ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( P(A) )</td>
<td>( P(E) )</td>
<td></td>
</tr>
<tr>
<td>Including ArgM</td>
<td>Role identification</td>
<td>.99</td>
<td>.89</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>Role classification</td>
<td>.95</td>
<td>.27</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>Combined decision</td>
<td>.99</td>
<td>.88</td>
<td>.91</td>
</tr>
<tr>
<td>Excluding ArgM</td>
<td>Role identification</td>
<td>.99</td>
<td>.91</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>Role classification</td>
<td>.98</td>
<td>.41</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>Combined decision</td>
<td>.99</td>
<td>.91</td>
<td>.93</td>
</tr>
</tbody>
</table>
Outline

• Introduction
• PropBank
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Automatic Semantic Role Labeling

- Examine the importance of syntactic information for semantic-role labeling

- Comparing the performance of
  - System based on gold-standard parses
  - Automatically generated parser output
Automatic Semantic Role Labeling

• Gildea and Jurafsky (2002)
  – Statistical system trained on FrameNet project
    • Pass sentences through an automatic parser (Collins, 1999)
    • Extract syntactic features from the parses
    • Estimate probabilities for semantic roles from the syntactic and lexical features
    • Errors introduced by the parser no doubt negatively affected the results obtained
Outline

• Introduction
• PropBank
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Automatic Semantic Role Labeling

• Features
  – **Phrase type**: the syntactic type of the phrase expressing the semantic roles
  – **Parse tree path**: the path from the predicate through the parse tree to the constituent in question.

In order to capture the syntactic relation of a constituent to the predicate.
Automatic Semantic Role Labeling

• Features
  – **Position**: indicates whether the constituent to be labeled occurs before or after the predicate
  – **Voice**: distinguishes between active and passive, direct objects of active verbs correspond to subjects of passive verbs
  – **Headword**: a lexical feature and provides information about the semantic type of the role filler
Outline

• Introduction
• PropBank
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Automatic Semantic Role Labeling

• Predict argument roles

\[
\arg\max_{r_1, \ldots, r_n} P(r_1, \ldots, r_n | F_1, \ldots, F_n, p)
\]

\(r_i\) : role of constituents \(i\) in the sentence

\(F_i = \{pt_i, path_i, pos_i, v_i, h_i\}\) : set of features at each constituent in the parse tree
Automatic Semantic Role Labeling

- Predict argument roles

\[
P(r_1 \ldots n|F_1 \ldots n, p) \approx P(\{r_1 \ldots n\}|p) \prod_i \frac{P(r_i|F_i, p)}{P(r_i|p)}
\]

- \( P(r_i|F_i, p) \): a constituent’s role given our five features for the constituent and the predicate \( p \)

- \( P(\{r_1 \ldots n\}|p) \): a set of roles appearing in a sentence given a predicate
Outline

• Introduction
• PropBank
  – Framing
  – Annotation
• Automatic Semantic-Role labeling
  – Features
  – Algorithm
  – Evaluation
• Conclusion
Automatic Semantic Role Labeling

• Data
  – PropBank (preliminary release version)
    • 72,109 predicate-argument structures
    • 190,815 individual arguments
    • examples from 2,462 lexical predicates (types)
  – Testing data: Penn Treebank Section 23
Automatic Semantic Role Labeling

• Results
  – Given the constituents which are arguments to the predicate and merely has to predict the correct role
  – Find the arguments in the sentence and label them correctly

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>PropBank</th>
<th>PropBank &gt; 10 examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic parses</td>
<td>82.0</td>
<td>79.9</td>
<td>80.9</td>
</tr>
<tr>
<td>Gold-standard parses</td>
<td>82.0</td>
<td>82.0</td>
<td>82.8</td>
</tr>
</tbody>
</table>

Accuracy of semantic-role prediction (in percentages) for known boundaries
Automatic Semantic Role Labeling

• Results
  – Adding Traces
    • Provide hints as to the semantics of individual clauses

<table>
<thead>
<tr>
<th></th>
<th>FrameNet</th>
<th></th>
<th>PropBank</th>
<th></th>
<th>PropBank &gt; 10 examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>Precision</td>
<td>Recall</td>
<td>Precision</td>
</tr>
<tr>
<td>Automatic parses</td>
<td>64.6</td>
<td>61.2</td>
<td>68.6</td>
<td>57.8</td>
<td>69.9</td>
</tr>
<tr>
<td>Gold-standard parses</td>
<td>74.3</td>
<td>66.4</td>
<td>76.0</td>
<td>69.9</td>
<td></td>
</tr>
<tr>
<td>Gold-standard with traces</td>
<td>80.6</td>
<td>71.6</td>
<td>82.0</td>
<td>74.7</td>
<td></td>
</tr>
</tbody>
</table>

Accuracy of semantic-role prediction (in percentages) for unknown boundaries (the system must identify the correct constituents as arguments and give them the correct roles)
Automatic Semantic Role Labeling

• Results

Table 13
Accuracy of semantic-role prediction for unknown boundaries (the system must identify the correct constituents as arguments and give them the correct roles).

<table>
<thead>
<tr>
<th>Role</th>
<th>Number</th>
<th>Precision</th>
<th>Labeled recall</th>
<th>Unlabeled recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg0</td>
<td>1,197</td>
<td>94.2%</td>
<td>88.9%</td>
<td>92.2%</td>
</tr>
<tr>
<td>Arg1</td>
<td>1,436</td>
<td>95.4%</td>
<td>82.5%</td>
<td>88.9%</td>
</tr>
<tr>
<td>Arg2</td>
<td>229</td>
<td>79.0%</td>
<td>64.2%</td>
<td>77.7%</td>
</tr>
<tr>
<td>Arg3</td>
<td>61</td>
<td>71.4%</td>
<td>49.2%</td>
<td>54.1%</td>
</tr>
<tr>
<td>Arg4</td>
<td>31</td>
<td>91.7%</td>
<td>71.0%</td>
<td>83.9%</td>
</tr>
<tr>
<td>ArgM</td>
<td>127</td>
<td>59.6%</td>
<td>26.8%</td>
<td>52.0%</td>
</tr>
<tr>
<td>ArgM-ADV</td>
<td>85</td>
<td>59.1%</td>
<td>30.6%</td>
<td>55.3%</td>
</tr>
<tr>
<td>ArgM-DIR</td>
<td>49</td>
<td>76.7%</td>
<td>46.9%</td>
<td>61.2%</td>
</tr>
<tr>
<td>ArgM-DIS</td>
<td>65</td>
<td>40.0%</td>
<td>18.5%</td>
<td>55.4%</td>
</tr>
<tr>
<td>ArgM-EXT</td>
<td>18</td>
<td>81.2%</td>
<td>72.2%</td>
<td>77.8%</td>
</tr>
<tr>
<td>ArgM-LOC</td>
<td>95</td>
<td>60.7%</td>
<td>38.9%</td>
<td>62.1%</td>
</tr>
<tr>
<td>ArgM-MNR</td>
<td>80</td>
<td>62.7%</td>
<td>40.0%</td>
<td>63.8%</td>
</tr>
<tr>
<td>ArgM-MOD</td>
<td>95</td>
<td>77.6%</td>
<td>40.0%</td>
<td>43.2%</td>
</tr>
<tr>
<td>ArgM-NEG</td>
<td>40</td>
<td>63.6%</td>
<td>17.5%</td>
<td>40.0%</td>
</tr>
<tr>
<td>ArgM-PRD</td>
<td>3</td>
<td>0.0%</td>
<td>0.0%</td>
<td>33.3%</td>
</tr>
<tr>
<td>ArgM-PRP</td>
<td>54</td>
<td>70.0%</td>
<td>25.9%</td>
<td>37.0%</td>
</tr>
<tr>
<td>ArgM-TMP</td>
<td>325</td>
<td>72.4%</td>
<td>45.2%</td>
<td>64.6%</td>
</tr>
</tbody>
</table>

Labeled recall: how often the semantic-role label is correctly identified.

Unlabeled recall: how often a constituent with the given role is correctly identified as being a semantic role, even if it is labeled with the wrong role.
Automatic Semantic Role Labeling

• The relation of Syntactic Parsing and Semantic-Role labeling
  – Chunks
    • Do not build a full parse tree
    • Large advantage in speed
    • Contain basic-level constituent boundaries and labels
    • No dependencies between constituents

[\text{NP Big investment banks}] [\text{VP refused to step}] [\text{ADVP up}] [\text{PP to}]
[\text{NP the plate}] [\text{VP to support}] [\text{NP the beleaguered floor traders}] [\text{PP by}]
[\text{VP buying}] [\text{NP bigblocks}] [\text{PP of}] [\text{NP stock}] [\text{NP traders}] [\text{VP say}]. (wsj_2300)
Automatic Semantic Role Labeling

• The relation of Syntactic Parsing and Semantic-Role labeling

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold parse</td>
<td>74.3%</td>
<td>66.4%</td>
</tr>
<tr>
<td>Auto parse</td>
<td>68.6</td>
<td>57.8</td>
</tr>
<tr>
<td>Chunk</td>
<td>27.6</td>
<td>22.0</td>
</tr>
<tr>
<td>Chunk, relaxed scoring</td>
<td>49.5</td>
<td>35.1</td>
</tr>
</tbody>
</table>
Conclusion

• Consistent annotation has been achieved
• One step closer to a detailed semantic representation

• WSJ too domain specific, too financial, need broader coverage genres for more general annotation
Future work

- Add more informative thematic labels based on VerbNet
- Map annotation with FrameNet to merge two annotated data sets
- Explore
  - machine-learning approaches
  - Integration of semantic-role labeling and sense tagging with the parsing process
References
