PropBank, VerbNet & SemLink

Edward Loper
PropBank

- 1M words of WSJ annotated with predicate-argument structures for verbs.
  - The location & type of each verb’s arguments
- Argument types are defined on a per-verb basis.
  - Consistent across uses of a single verb (sense)
- But the same tags are used (Arg0, Arg1, Arg2, …)
  - Arg0 ≈ proto-typical agent (Dowty)
  - Arg1 ≈ proto-typical patient
PropBank Example:

*cover (smear, put over)*

- Arguments:
  - Arg0 = causer of covering
  - Arg1 = thing covered
  - Arg2 = covered with

- Example:
  
  John *covered* the bread with *peanut butter*. 
PropBank: Trends in Argument Numbering

- **Arg0** = proto-typical agent (*Dowty*)
  
  Agent (85%), Experiencer (7%), Theme (2%), …

- **Arg1** = proto-typical patient
  
  Theme (47%), Topic (23%), Patient (11%), …

- **Arg2** = Recipient (22%), Extent (15%), Predicate (14%), …

- **Arg3** = Asset (33%), Theme2 (14%), Recipient (13%), …

- **Arg4** = Location (89%), Beneficiary (5%), …

- **Arg5** = Location (94%), Destination (6%)

(Percentages indicate how often argument instances were mapped to VerbNet roles in the PropBank corpus)
PropBank: Adjunct Tags

• Variety of ArgM’s (Arg#>5):
  – TMP: when?
  – LOC: where at?
  – DIR: where to?
  – MNR: how?
  – PRP: why?
  – REC: himself, themselves, each other
  – PRD: this argument refers to or modifies another
  – ADV: others
VerbNet

• Organizes verbs into classes that have common syntax/semantics linking behavior
• Classes include…
  – A list of member verbs (w/ WordNet senses)
  – A set of thematic roles (w/ selectional restr. s)
  – A set of frames, which define both syntax & semantics using thematic roles.
• Classes are organized hierarchically
VerbNet - **cover contiguous_location-47.8**

<table>
<thead>
<tr>
<th>Members</th>
</tr>
</thead>
<tbody>
<tr>
<td>BESTRIDE</td>
</tr>
<tr>
<td>BLANKET (FN 1; WN 1, 2)</td>
</tr>
<tr>
<td>BORDER (WN 1, 2, 3)</td>
</tr>
<tr>
<td>ROUNDED (FN 2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theme1</strong> [+CONCRETE]</td>
</tr>
<tr>
<td><strong>Theme2</strong> [+CONCRETE]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic Transitive</strong></td>
</tr>
<tr>
<td><strong>Example</strong></td>
</tr>
<tr>
<td><strong>Syntax</strong></td>
</tr>
<tr>
<td><strong>Semantics</strong></td>
</tr>
</tbody>
</table>
VerbNet Thematic Roles

- Actor
- Actor1
- Actor2
- Agent
- Asset
- Attribute
- Beneficiary
- Cause
- Destination
- Experiencer
- Extent
- Instrument
- Location
- Material
- Patient
- Patient1
- Patient2
- Predicate
- Product
- Proposition
- Recipient
- Source
- Stimulus
- Theme
- Theme1
- Theme2
- Time
- Topic
- Value
SemLink: Mapping Lexical Resources

• Different lexical resources provide us with different information.
• To make useful inferences, we need to combine this information.
• In particular:
  – PropBank -- How does a verb relate to its arguments? Includes annotated text.
  – VerbNet -- How do verbs w/ shared semantic & syntactic features (and their arguments) relate?
  – FrameNet -- How do verbs that describe a common scenario relate?
  – WordNet -- What verbs are synonymous?
  – ...
What do mappings look like?

• 2 Types of mappings:
  – *Type mappings* describe which entries from two resources might correspond; and how their fields (e.g. arguments) relate.
    • Potentially many-to-many
    • Generated manually or semi-automatically
  – *Token mappings* tell us, for a given sentence or instance, which type mapping applies.
    • Can often be thought of as a type of classifier
      – Built from a single corpus w/ parallel annotations
    • Can also be though of as word sense disambiguation
      – Because each resource defines word senses differently!
Mapping from **PB** to **VerbNet**

- **escape-51.1**
  - "move away from"
  - **leave.01** → **leave-51.2**
  - arg0 (entity leaving) → Theme
  - arg1 (place left) → Source
  - arg2 (attribute)

- **keep-15.2**

- **fulfill-13.4.1**

- **future_having-13.3**
  - "give"
  - **leave.02**
  - arg0 (giver) → Agent
  - arg1 (thing given) → Theme
  - arg2 (benefactive) → Recipient
Mapping Issues

- Mappings are often many-to-many
  - Different resources focus on different distinctions

- Incomplete coverage
  - A resource may be missing a relevant lexical item entirely.
  - A resource may have the relevant lexical item, but not in the appropriate category or w/ the appropriate sense

- Field mismatches
  - It may not be possible to map the field information for corresponding entries. (E.g., predicate arguments)
    - Extra fields
    - Missing fields
    - Mismatched fields
Mapping Issues (2)
VerbNet verbs mapped to FrameNet

- VerbNet clear-10.3
- FrameNet Classes

Diagram:
- clear
- clean
- drain
- empty
- Removing
- Emptying
Mapping Issues (3)
VerbNet verbs mapped to FrameNet

**VN Class: put 9.1**

Members: arrange*, immerse, lodge, mount, sling**

Thematic roles:
- agent (+animate)
- theme (+concrete)
- destination (+loc, -region)

Frames:
- ...

*different sense
** not in FrameNet

**FrameNet frame: place**

Frame Elements:
- Agent
- Cause
- Theme
- Goal

Examples:
- ...

*not in FrameNet*