VerbNet

Martha Palmer
University of Colorado

LING 7800/CSCI 7000-017
September 16, 2014
Outline

- Recap
- Levin’s Verb Classes
- VerbNet
- PropBank
Recap

- Fillmore – Cases
  - useful generalizations, fewer sense distinctions,

- Jackendoff – Lexical Conceptual Structure
  - Thematic roles are defined by the predicates they are arguments to

- Dowty – Proto-typical Agents and Patients
  - A bag of “agentive” entailments

- Levin – Verb classes based on syntax
  - syntactic behavior is a reflection of the underlying semantics
A Preliminary Classification of English Verbs, Beth Levin

- Based on diathesis alternations

- The range of syntactic variations for a class of verbs is a reflection of the underlying semantics
Levin classes (3100 verbs)

- 47 top level classes, 193 second and third level

- Based on pairs of syntactic frames.
  
  John broke the jar. /  Jars break easily. /  The jar broke.
  John cut the bread. /  Bread cuts easily. /  *The bread cut.
  John hit the wall. /  *Walls hit easily. /  *The wall hit.

- Reflect underlying semantic components
  contact, directed motion,
  exertion of force, change of state

- Synonyms, syntactic patterns (conative), relations
Break Levin class - Change-of-state
Confusions in Levin classes?

- Not semantically homogenous
  - \{braid, clip, file, powder, pluck, etc...\}

- Multiple class listings
  - homonymy or polysemy?

- Alternation contradictions?
  - Carry verbs disallow the Conative, but include
    \{push, pull, shove, kick, draw, yank, tug\}
  - also in Push/pull class, does take the Conative
Intersective Levin Classes

Dang, Kipper & Palmer, ACL98
Regular Sense Extensions

John pushed the chair.  +force, +contact
John pushed the chairs apart.  +ch-state
John pushed the chairs across the room.  +ch-loc
John pushed at the chair.  -ch-loc

The train whistled into the station.  +ch-loc
The truck roared past the weigh station.  +ch-loc

AMTA98,ACL98,TAG98
Intersective Levin Classes

- More syntactically and semantically coherent
  - sets of syntactic patterns
  - explicit semantic components
  - relations between senses

VERBNET

www.cis.upenn.edu/verbnet
VerbNet: Overview

- Purpose of VN is to classify English verbs based on semantic and syntactic regularities (Levin, 1993)

- Classification used for numerous NLP tasks, primarily semantic role labeling (Schuler, 2002; Shi and Mihalcea, 2005, Yi, et. al., 2007))

- In each verb class, thematic roles are used to link syntactic alternations to semantic predicates, which can serve as foundation for further inferences

\[
\begin{align*}
\text{NP} & \hspace{1em} \text{V} \hspace{1em} \text{NP} \hspace{1em} \text{THAT} \hspace{1em} \text{S} \\
\text{EXAMPLE} & \hspace{1em} "I\hspace{0.5em}promised\hspace{0.5em}him\hspace{0.5em}that\hspace{0.5em}he\hspace{0.5em}would\hspace{0.5em}arrive\hspace{0.5em}in\hspace{0.5em}time." \\
\text{SYNTAX} & \hspace{1em} \text{AGENT} \hspace{1em} \text{V} \hspace{1em} \text{RECIPIENT} \hspace{1em} \text{TOPIC} \hspace{1em} \text{THAT\_COMP} \\
\text{SEMANTICS} & \hspace{1em} \text{TRANSFER\_INFO(DURING(E), AGENT, RECIPIENT, TOPIC) CAUSE(AGENT, E)}
\end{align*}
\]
VerbNet – based on Levin, B., 93

Class entries:
- Capture generalizations about verb behavior
- Organized hierarchically
- Members have common semantic elements, semantic roles, syntactic frames, predicates

Verb entries:
- Refer to a set of classes (different senses)
- each class member linked to WN synset(s), ON groupings, PB frame files, FrameNet frames,
The Unified Verb Index

- http://verbs.colorado.edu/verb-index/
VerbNet: An in-depth example

- "Behavior of a verb . . . is to a large extent determined by its meaning" (p. 1)
  Amanda hacked the wood with an ax.
  Amanda hacked at the wood with an ax.
  Craig notched the wood with an ax.
  *Craig notched at the wood with an ax.

- Can we move from syntactic behavior back to semantics?
Hacking and Notching

- Same thematic roles: Agent, Patient, Instrument
- Some shared syntactic frames, e.g. Basic Transitive (Agent V Patient)
- **Hack: cut-21.1**

  - cause(Agent, E)
  - manner(during(E), Motion, Agent)
  - contact(during(E), ?Instrument, Patient)
  - degradation_material_integrity(result(E), Patient)
Hacking and Notching

- Same thematic roles: Agent, Patient, Instrument
- Some shared syntactic frames, e.g. Basic Transitive (Agent V Patient)
- Notch: carve-21.2

  cause(Agent, E)
  contact(during(E), ?Instrument, Patient)
  degradation_material_integrity(result(E), Patient)
  physical_form(result(E), Form, Patient)
Also Temporal Characteristics

- Needed for distinguishing between *Verbs of Assuming a Position* and *Verbs of Spatial Configuration*

- Semantic predicates are associated with an event variable, $e$, and often have an additional argument:
  - START($e$) – in force at the START of the event
  - END($e$) – in force at the END of the event
  - DURING($e$) – in force DURING the related time period for the entire event
VerbNet: send-11.1 (Members: 11, Frames: 5)
includes “ship”

- **Roles**
  - Agent [+animate | +organization]
  - Theme [+concrete]
  - Source [+location]
  - Destination [+animate | [+location & -region]]

- **Syntactic Frame:** NP V NP PP.destination
  - example  "Nora sent the book to London."
  - syntax   Agent V Theme {to} Destination
  - semantics motion(during(E), Theme)
                location(end(E), Theme, Destination)
                cause(Agent, E)
VerbNet can also provide inferences

- Every path from back door to yard was covered by a grape-arbor, and every yard had fruit trees.
- Where are the grape arbors located?
VerbNet – *cover*, *fill*-9.8 class

- **Members:** fill, …, cover, …, staff, ….

- **Thematic Roles:** Agent
  Theme
  Destination

- **Syntactic Frames with Semantic Roles**
  “The employees staffed the store"
  “The grape arbors covered every path"
  Theme V Destination

```
location(E,Theme,Destination)
llocation(E,grape_arbor,path)
```
Recovering Implicit Arguments
[Palmer, et. al., 1986, Gerber & Chai, 2010]

\[
[\text{Arg}_0 \text{ The two companies}] [\text{REL}_1 \text{ produce}] [\text{Arg}_1 \text{ market pulp, containerboard and white paper}].
\]
The goods could be manufactured closer to customers, saving [\text{REL}_2 \text{ shipping}] costs.

- Used VerbNet for subcategorization frames
Implicit arguments

**SYNTAX**  
Agent V Theme {to} Destination

\[ \text{AGENT} \text{ shipped } \text{THEME} \text{ to } \text{DESTINATION} \]

**SEMANTICS**

- \( \text{CAUSE(AGENT,E)} \)
- \( \text{MOTION(DURING(E), THEME)} \),
- \( \text{LOCATION(END(E), THEME, DESTINATION)} \),
Implicit arguments instantiated using coreference

- \([AGENT] \text{ shipped } [THEME] \text{ to } [DESTINATION]\)
- \([Companies] \text{ shipped } [goods] \text{ to } [customers]\).

**SEMANTICS**

- CAUSE(Companies, E)
- MOTION(DURING(E), goods),
- LOCATION(END(E), goods, customers),

*Can annotate, semi-automatically!*
Limitations to VerbNet as a sense inventory

- Concrete criteria for sense distinctions
  - Distinct semantic roles, but very fine-grained; leads to sparse data problems
  - Distinct frames
  - Distinct entailments

- But….

- Limited coverage of lemmas

- For each lemma, limited coverage of senses
Goal of PropBank

- Supply consistent, simple, general purpose labeling of semantic roles
- Provide consistent argument labels across different syntactic realizations
- Support the training of automatic semantic role labelers
- Semantic representation can support…
Training data supporting...

- Machine translation
- Text editing
- Text summary / evaluation
- Question and answering
The Problem

- Levin (1993) and others have demonstrated promising relationship between syntax and semantics
- Same verb with same subcategorization can assign different semantic roles
- How can we take advantage of clear relationships and empirically study how and why syntactic alternations take place?
VerbNet and Real Data

- VerbNet is based on linguistic theory – how useful is it?
- How well does it correspond to syntactic variations found in naturally occurring text?
- Use PropBank to investigate these issues
What is PropBank?

- Semantic information over the syntactically parsed (i.e. treebanked) text
- Semantic information -> predicate argument structure of a verb or a relation
- Unlike VerbNet, the predicate argument structure is specific to the verb or relation in question
- Seeks to
  1. provide consistent argument labels across different syntactic realizations of the same verb
  2. assign general functional tags to all modifiers or adjuncts to the verb
“1. PB seeks to provide consistent argument labels across different syntactic realizations”

**Jin** broke **the projector**.

Syntax: \[ \text{NP}_{\text{SUB}} \quad \text{V} \quad \text{NP}_{\text{OBJ}} \]

Thematic Roles: \[ \text{AGENT} \quad \text{REL} \quad \text{PATIENT} \]

**The projector** broke.

Syntax: \[ \text{NP}_{\text{SUB}} \quad \text{V} \]

Thematic Roles: \[ \text{PATIENT} \quad \text{REL} \]
Why numbered arguments?

- Avoids lack of consensus concerning a specific set of semantic role labels
- Numbers correspond to labels that are verb-specific
- Arg0 and Arg1 correspond to Dowty’s (1991) proto-agent and proto-patient
- Args 2-5 are highly variable
“1. PB seeks to provide consistent argument labels across different syntactic realizations”

- Uuuuuuusually…
  - Arg0 = agent
  - Arg1 = patient
  - Arg2 = benefactive / instrument / attribute / end state
  - Arg3 = start point / benefactive / instrument / attribute
  - Arg4 = end point

These correspond to VN Thematic Roles
“2. PB seeks to assign functional tags to all modifiers or adjuncts to the verb”

- Variety of ArgM’s:
  - TMP - when? yesterday, 5pm on Saturday, recently
  - LOC - where? in the living room, on the newspaper
  - DIR - where to/from? down, from Antartica
  - MNR - how? quickly, with much enthusiasm
  - PRP/CAU - why? because …, so that …
  - REC - himself, themselves, each other
  - GOL - end point of motion, transfer verbs? To the floor, to Judy
  - ADV - hodge-podge, miscellaneous, “nothing-fits!”
  - PRD - this argument refers to or modifies another: …ate the meat raw
Different verb senses…

- Have different subcategorization frames
- PropBank assigns coarse-grained senses to verbs
- PropBank “framesets,” lexical resource
- New senses, or “rolesets” are added only when the syntax and semantics of a usage are distinct
- Annotators use “frame files” to assign appropriate numbered arg structure
Propbank: sense distinctions?

- Mary left the room
- Mary left her daughter-in-law her pearls in her will

Frameset leave.01 "move away from":
Arg0: entity leaving
Arg1: place left

Frameset leave.02 "give":
Arg0: giver
Arg1: thing given
Arg2: beneficiary
WordNet: - call, 28 senses, 9 groups

- WN5, WN16, WN12
  - Loud cry

- WN3, WN19
  - Label

- WN1, WN22
  - Challenge

- WN18, WN27
  - Phone/radio

- WN2, WN13
  - WN28

- WN17, WN11
  - WN24,

- WN15, WN26
  - Bird or animal cry

- WN4, WN7, WN8, WN9
  - Request

- WN20, WN25
  - Call a loan/bond

- WN6, WN23
  - Visit

- WN10, WN14, WN21,
  - Bid
WordNet: - call, 28 senses, 9 groups, add PB

- Loud cry: WN5, WN16, WN12
- Label: WN3, WN19, WN1, WN22
- Challenge: WN18, WN27
- Phone/radio: WN2, WN13, WN28
- WN17, WN11, WN24
- Bird or animal cry: WN15, WN26
- Request: WN4, WN7, WN8, WN9
- Call a loan/bond: WN20, WN25
- Visit: WN6, WN23
- WN10, WN14, WN21, Bid
Overlap between Groups and Framesets – 95%

Frameset 1

WN1, WN2
WN6, WN7, WN8
WN11, WN12, WN13
WN19

Frameset 2

WN3, WN4
WN5, WN9, WN10
WN14
WN20

develop

Palmer, Dang & Fellbaum, NLE 2004
Sense Hierarchy
(Palmer, et al, SNLU04 - NAACL04, NLE07, Chen, et. al, NAACL06)

- PropBank Framesets – ITA >90%
  coarse grained distinctions
- 20 Senseval2 verbs w/ > 1 Frameset
- Maxent WSD system, 73.5% baseline, 90%

- Sense Groups (Senseval-2) - ITA 82%
  Intermediate level
  (includes Levin classes) – 71.7%

- WordNet – ITA 73%
  fine grained distinctions, 64%

Tagging w/groups,
ITA 90%, 200@hr;
Taggers - 86.9%
Semeval07

Chen, Dligach & Palmer, ICSC 2007
Dligach & Palmer, ACL-11, - 88%
SEMLINK

- Extended VerbNet: 5,391 senses (91% PB)
- Type-type mapping PB/VN, VN/FN
- Semi-automatic mapping of WSJ PropBank instances to VerbNet classes and thematic roles, hand-corrected. (now FrameNet also)
- VerbNet class tagging as automatic WSD
  
  Brown, Dligach, Palmer, IWCS 2011

- Run SRL, map Arg2 to VerbNet roles, Brown performance improves
  
  Yi, Loper, Palmer, NAACL07
Mapping from PropBank to VerbNet (similar mapping for PB-FrameNet)

<table>
<thead>
<tr>
<th>Frameset id = leave.02</th>
<th>Sense = give</th>
<th>VerbNet class = future-having 13.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg0</td>
<td>Giver</td>
<td>Agent/Donor*</td>
</tr>
<tr>
<td>Arg1</td>
<td>Thing given</td>
<td>Theme</td>
</tr>
<tr>
<td>Arg2</td>
<td>Benefactive</td>
<td>Recipient</td>
</tr>
</tbody>
</table>

*FrameNet Label

Baker, Fillmore, & Lowe, COLING/ACL-98
Fillmore & Baker, WordNetWKSHP, 2001
Mapping from PB to VerbNet
verbs.colorado.edu/~semlink
Generative Lexicon - VerbNet

- GL: use(Agent, Entity, Purpose)

  - use, sense 1: apply or employ something for a purpose (the most general sense)
    - Use 105
  - use, sense 2: consume or ingest, usually habitually
    - Eat 39.1-3
  - use, sense 3: expend a quantity (e.g., use up something, use something up)
    - Consume 66
Generative Lexicon - VerbNet

- GL: use(Agent, Entity, Purpose)

  - use, sense 1: apply or employ something for a purpose (the most general sense)
    - Use 105 [http://verbs.colorado.edu/vn3.2.4-test-uvi/vn/use-105.1.php](http://verbs.colorado.edu/vn3.2.4-test-uvi/vn/use-105.1.php)

  - use, sense 2: consume or ingest, usually habitually
    - Eat 39.1-3 [http://verbs.colorado.edu/vn3.2.4-test-uvi/vn/eat-39.1.php](http://verbs.colorado.edu/vn3.2.4-test-uvi/vn/eat-39.1.php)

  - use, sense 3: expend a quantity (e.g., use up something, use something up)
    - Consume 66 [http://verbs.colorado.edu/vn3.2.4-test-uvi/vn/consume-66.php](http://verbs.colorado.edu/vn3.2.4-test-uvi/vn/consume-66.php)
Additional Entailments

- Sense 1 is the most general
- Senses 2 and 3 provide additional specific entailments
  - Sense 2: Entity is ingested by an animate being, who then undergoes a change of state
  - Sense 3: in the process of using the Entity, it is depleted