Goals – Ex. Answering Questions

- Similar concepts
  - Where are the grape arbors located?
  - Every path from back door to yard was covered by a grape-arbor, and every yard had fruit trees.

Outline

- WordNet, OntoNotes Groupings, PropBank
- VerbNet
  - Verbs grouped in hierarchical classes
  - Explicitly described class properties
- FrameNet
- Links among lexical resources
  - PropBank, FrameNet, WordNet, OntoNotes groupings
- Automatic Semantic Role Labeling with PropBank/Verbnet

WordNet – Princeton
(Miller 1985, Fellbaum 1998)

On-line lexical reference (dictionary)

- Nouns, verbs, adjectives, and adverbs grouped into synonym sets
- Other relations include hypernyms (ISA), antonyms, meronyms
- Typical top nodes - 5 out of 25
  - (act, action, activity)
  - (animal, fauna)
  - (artifact)
  - (attribute, property)
  - (body, corpus)
WordNet – Princeton – *leave, n.*, *v.*.14  
(Miller 1985, Fellbaum 1998)

- Limitations as a computational lexicon
  - Contains little syntactic information
  - No explicit lists of participants
  - Sense distinctions very fine-grained,
  - Definitions often vague
- Causes problems with creating training data for supervised Machine Learning – SENSEVAL2
  - Verbs > 16 senses (including *call*)
  - Inter-annotator Agreement ITA 71%
  - Automatic Word Sense Disambiguation, WSD 64%

_Bang & Palmer, SIGLEX02_

Creation of coarse-grained resources

- Unsupervised clustering using rules (Mihalcea & Moldovan, 2001)
- Clustering by mapping WN senses to ODE (Navigli, 2006).
- OntoNotes - Manually grouping WN senses and annotating a corpus (Hovy et al., 2006)
- Supervised clustering WN senses using OntoNotes and another set of manually tagged data (Snow et al., 2007).

OntoNotes

- DARPA-GALE, OntoNotes 5.0
  - BBN, Brandeis, Colorado, Penn
  - Multilayer structure
  - Three languages: English, Arabic, Chinese
  - Several Genres (@ ≥ 200K): NW, BN, BC, WT
  - Parallel data, E/C, E/A
  - PropBank frame coverage for rare verbs
  - Recent PropBank extensions

OntoNotes: Multilayer Design

- The literal meaning of sentences
  - A frame-based representation of predicates and their arguments
  - Referring expressions and the textual phrases they refer to
  - Coarse-grained word sense tags for most polysemous verbs
- Does this lay a foundation for inference?
Empirical Validation – Human Judges

Groupings Methodology – Human Judges (w/ Dang and Fellbaum)
- Double blind groupings, adjudication
- Syntactic Criteria (VerbNet was useful)
  - Distinct subcategorization frames
    - call him an idiot
    - call him a taxi
  - Recognizable alternations – regular sense extensions:
    - play an instrument
    - play a song
    - play a melody on an instrument

SIGLEX01, SIGLEX02, INL07, Duffield, et. al., CogSci 2007

Groupings Methodology (cont.)
- Semantic Criteria
  - Differences in semantic classes of arguments
    - Abstract/concrete, human/animal, animate/inanimate, different instrument types,…
  - Differences in the number and type of arguments
    - Often reflected in subcategorization frames
    - John left the room.
    - I left my pearls to my daughter-in-law in my will.
  - Differences in entailments
    - Change of prior entity or creation of a new entity?
  - Differences in types of events
    - Abstract/concrete/mental/emotional/….
  - Specialized subject domains

PropBank Verb Frames Coverage
- The set of verbs is open
- But the distribution is highly skewed
- For English, the 1000 most frequent lemmas cover 95% of the verbs in running text.
- Graphs show counts over English Web data containing 150 M verbs.
Verb Frames Coverage By Language

<table>
<thead>
<tr>
<th>Language</th>
<th>Projected Final Count</th>
<th>Estimated Coverage in Running Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>5,100</td>
<td>99%</td>
</tr>
<tr>
<td>Chinese</td>
<td>18,200</td>
<td>96%</td>
</tr>
<tr>
<td>Arabic</td>
<td>5,250*</td>
<td>99%</td>
</tr>
</tbody>
</table>

* This covers all the verbs and most of the predicative adjectives/nouns in ATB.

How do the PropBank verb frames relate to Word Senses?

Lexical Resource - Frames Files:

**give**

Roles:
- Arg0: giver
- Arg1: thing given
- Arg2: entity given to

Example: double object

*The executives gave the chefs a standing ovation.*
- Arg0: The executives
- REL: gave
- Arg2: the chefs
- Arg1: a standing ovation

Word Senses in PropBank

- Orders to ignore word sense not feasible for 700+ verbs
  - 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

How do these relate to word senses in other resources?
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%

Power behind the throne - VerbNet
- Groupings include links to VerbNet classes
- PropBank includes links to VerbNet classes and thematic roles
- Intention to ensure that VerbNet class members are handled consistently in PropBank and Groupings
- What is VerbNet?

Levin classes (Levin, 1993)
- 3100 verbs, 47 top level classes, 193 second and third level
- Each class has a syntactic signature based on alternations.
  John broke the jar. / The jar broke. / Jars break easily.
  John cut the bread. / *The bread cut. / Bread cuts easily.
  John hit the wall. / *The wall hit. / *Walls hit easily.
Summary of semantic components

- Verb class hierarchy: 3100 verbs, 47 top level classes, 193
- Each class has a syntactic signature based on alternations.
  
  *John broke the jar. / The jar broke. / Jars break easily.*
  
  change-of-state

  *John cut the bread. / *The bread cut. / Bread cuts easily.*
  
  change-of-state, recognizable action,
  
  sharp instrument, contact, motion

  *John hit the wall. / *The wall hit. / *Walls hit easily.*
  
  contact, exertion of force, motion

  *John touched the wall. / *The wall touched. / *Walls touch easily.*
  
  contact

Common class components

markers, not distinguishers

- **cut** [CAUSE] [CHANGE] [MOTION] [CONTACT]
- **break** [CAUSE] [CHANGE]
- **touch** [CONTACT]
- **hit** [MOTION] [CONTACT]

Summary of syntactic patterning

<table>
<thead>
<tr>
<th>Touch</th>
<th>Hit</th>
<th>Cut</th>
<th>break</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conative</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Body-part ascension</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Middle</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Break Levin class - Change-of-state
Which semantic components are *grammatically* relevant?

- Pinker
  - Set of conceptually interpretable elements
  - Smaller than # of verbs, universal
  - Used by children
  - Have grammatical relevance
  - Distinguish classes that have different sets of lexical rules

Limitations to Levin Classes

- Coverage of only half of the verbs (types) in the Penn Treebank (1M words, WSJ)
- Usually only one or two basic senses are covered for each verb
- Confusing sets of alternations
  - Different classes have almost identical “syntactic signatures”
  - or worse, contradictory signatures

Multiple class listings

- Homonymy or polysemy?
  - *draw a picture, draw water from the well*
- Conflicting alternations?
  - *Carry* verbs disallow the Conative, (*she carried at the ball*), but include {push, pull, shove, kick, yank, tug}
  - also in *Push/pull* class, does take the Conative (*she kicked at the ball*)

Intersective Levin Classes

[Dang, Kipper & Palmer, ACL98](#)
VerbNet: Basis in Theory

- Verb class hierarchy: 3100 verbs, 47 top level classes, 193
- “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?

VerbNet – Karin Kipper Schuler

- 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) and FrameNet frames

Hacking and Notching

- Same thematic roles:
  - Agent, Patient, Instrument
- Some shared syntactic frames,
  - e.g. Basic Transitive (Agent V Patient)
- Different Semantic predicates

VerbNet Semantic Predicates

- **Hack: cut-21.1**
  
  cause(Agent, E)
  manner(during(E), Motion, Agent)
  contact(during(E), ?Instrument, Patient)
  degradation_material_integrity(result(E), 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)
**VerbNet example – Pour-9.5**

![VerbNet example](image1)

**VerbNet Pour-9.5 (cont.)**

![VerbNet example](image2)

**Hidden Axioms**

- **EXAMPLE:** Tamara poured water into the bowl.
- **SYNTAX:** AGENT V THEME LOCATION
- **SEMANTICS**
  - CAUSE(AGENT,E)
  - MOTION(DURING(E), THEME),
  - NOT(PREP(START(E), THEME, LOCATION)),
  - PREP(E, THEME, LOCATION)

**Hidden Axioms REVEALED!**

- **EXAMPLE:** Tamara poured water into the bowl.
- **SYNTAX:** AGENT V THEME LOCATION
- **SEMANTICS**
  - POUR.pour9.5(AGENT, THEME LOCATION) → CAUSE(AGENT,E),
  - MOTION(DURING(E), THEME),
  - NOT(PREP(START(E), THEME, LOCATION)),
  - PREP(E, THEME, LOCATION).
Hidden Axioms REVEALED!

- **EXAMPLE:** *Tamara poured water into the bowl.*
- **SYNTAX:** \text{AGENT} \text{V} \text{THEME LOCATION}
- **SEMANTICS**
  - **POUR.**\text{pour9.5} (\text{AGENT, THEME LOCATION}) \rightarrow
    
    \text{CAUSE}(Tamara,E),
    
    \text{MOTION}(\text{DURING}(E), \text{water}),
    
    \text{NOT}(\text{into}(\text{START}(E), \text{water}, \text{bowl})),
    
    \text{into}(E, \text{THEME}, \text{bowl}).

VerbNet – *cover fill-9.8*

- **WordNet Senses:** ..., cover(1,2, 22, 26),..., staff(1),
- **Thematic Roles:** Agent [+animate]
  - Theme [+concrete],
  - Destination [+location, +region]
- **Frames with Semantic Roles**
  - "The employees staffed the store"
  - "The grape arbors covered every path"
  - Theme V Destination
    
    \text{location}(E, \text{Theme}, \text{Destination})
    
    \text{location}(E, \text{grape\_arbor}, \text{path})

VerbNet as a useful NLP resource

- **Semantic role labeling**
- **Inferences**

While many of the weapons used by the insurgency are leftovers from the Iran-Iraq war, Iran is still providing deadly weapons such as EFPs -LRB- or Explosively Formed Projectiles -RRB-.

  provide(Iran, weapons, ?Recipient) \rightarrow
  
  \text{cause}(Iran, E)
  
  \text{has\_possession}(\text{start}(E), \text{Iran}, \text{weapons})
  
  \text{has\_possession}(\text{end}(E), ?\text{Recipient}, \text{weapons})
  
  \text{transfer}(\text{during}(E), \text{weapons})
Broader coverage still needed

- Only 78% of PropBank verbs included in VN
- Most classes focused on verbs with NP and PP complements
- Neglected verbs that take adverbial, adjectival, and sentential complements
- Extending VerbNet and mapping it to PropBank and FrameNet

Mapping from PropBank to VerbNet (similar mapping for PB-FrameNet)

<table>
<thead>
<tr>
<th>Frameset id =</th>
<th>Sense =</th>
<th>VerbNet class =</th>
</tr>
</thead>
<tbody>
<tr>
<td>ship.01</td>
<td>ship</td>
<td>Send -11.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arg0</th>
<th>Sender</th>
<th>Agent/Sender*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg1</td>
<td>Package</td>
<td>Theme</td>
</tr>
<tr>
<td>Arg2</td>
<td>Recipient</td>
<td>Destination/</td>
</tr>
<tr>
<td>Arg3</td>
<td>Source</td>
<td>*Goal OR Recipient</td>
</tr>
</tbody>
</table>

*FrameNet Labels*

Mapping from PB to VerbNet

http://verbs.colorado.edu/semlink

FrameNet

Introducing FrameNet
Thanks to Chuck Fillmore and Collin Baker

In one of its senses, the verb observe evokes a frame called Compliance: this frame concerns people’s responses to norms, rules or practices.

The following sentences illustrate the use of the verb in the intended sense:
- Our family observes the Jewish dietary laws.
- You have to observe the rules or you’ll be penalized.
- How do you observe Easter?
- Please observe the illuminated signs.

FrameNet

FrameNet records information about English words in the general vocabulary in terms of
1. the frames (e.g. Compliance) that they evoke,
2. the frame elements (semantic roles) that make up the components of the frames (in Compliance, Norm is one such frame element), and
3. each word’s valence possibilities, the ways in which information about the frames is provided in the linguistic structures connected to them (with observe, Norm is typically the direct object).

The FrameNet Product

The FrameNet database constitutes
- a set of frame descriptions
- a set of corpus examples annotated with respect to the frame elements of the frame evoked by each lexical unit
- lexical entries, including definitions and displays of the combinatory possibilities of each lexical unit, as automatically derived from the annotations
- a display of frame-to-frame relations, showing how some frames are elaborations of others, or are components of other frames.

Frame Elements for Compliance

The frame elements that figure in the Compliance frame are called
- Norm (the rule, practice or convention)
- Protagonist (the person[s] reacting to the Norm)
- Act (something done by the Protagonist that is evaluated in terms of the Norm)
- State_of_affairs (a situation evaluated in terms of the Norm)
- You do a whole frame for just observe?
- No. There are other Compliance words too.

V - adhere, comply, conform, follow, heed, obey, submit, ...;
AND NOT ONLY VerBS
N - adherence, compliance, conformity, obedience, observance, ...;
A - compliant, obedient, ...;
PP - in compliance with, in conformity to, ...;
AND NOT ONLY WORDS FOR POSITIVE RESPONSES TO NORMS
V - break, disobey, flout, transgress, violate, ...;
N - breach, disobedience, transgression, violation, ...;
PP - in violation of, in breach of, ...

- Are we finished with the verb observe?
- No. This verb has several other meanings too.

■ In the Perception_active frame we get the uses seen in observing children at play, observing an ant colony, sharing frame membership with watch, attend, listen to, view & pay attention.

■ In a Commenting frame, observe and observation share frame membership with remark & comment.

Tagging Compliance sentences

Protagonist
Our family
observes the dietary laws

State_of_affairs
The light switches in this room are in full conformity with the building code

Norm

Lexical Unit

Our unit of description is not the word (or “lemma”) but the lexical unit (Cruse 1986), – a pairing of a word with a sense. In our terms this is the pairing of a word with a single frame.

The lexical unit - roughly equivalent to a word in a synset - is the unit in terms of which important generalizations about lexical relations, meanings and syntactic behavior can best be formulated.
LUs and V-N relationships

- Note that the nouns based on *observe* are
  - *observance* in the **Compliance** frame,
  - *observation* in the **Perception_active** frame
- Similarly, the nouns based on *adhere* are
  - *adherence* in the **Compliance** frame,
  - *adhesion* in the **Attachment** frame.
- When we need to be precise we show the frame-specific sense of a lemma (the full name of an LU) with a dotted expression:
  - **Compliance.** *observe*, **Attachment.** *adhere*, etc.

The study of polysemy concerns membership in different frames

- 2 lexical units sharing the same form: **Compliance.** *observe*, **Perception.** *observe*
Different LU, Different Valence

**Compliance.** *observe* generally has an NP as its direct object.

**Perception.** *observe* has these patterns:
- NP: *Observe the clouds overhead.*
- NP+V-ing: *I observed the children playing.*
- wh-clause: *Observe what I’m doing.*
- that-clause: *We observed that the process terminated after an hour.*

**Comment.** *observe* occurs frequently with a quoted comment:
- “*That was brilliant,*” he observed snidely.

Lexical-units: Wrap-up

Lexical units are the entities with respect to which we define
- meanings
- grammatical behavior
- semantic relations with other entities
- morphological relations with other entities

In short, there aren’t interesting things to say about the verb *observe* in general, but only about the individual lexical units that happen to have the form *observe.*