

## Semantic Role Labeling

Martha Palmer  
University of Colorado

LING 7800/CSCI 7000-017  
January 31, 2013

1



## Outline

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

2



## 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

3



## 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

4

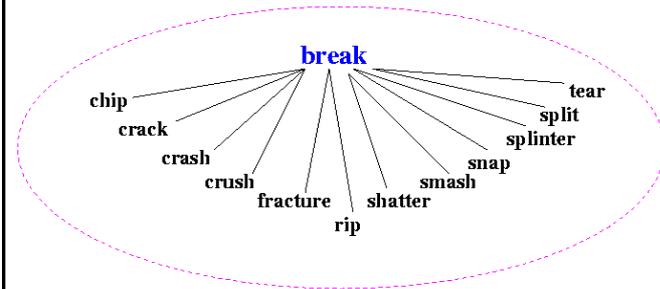


## Levin classes (3100 verbs)

- 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*

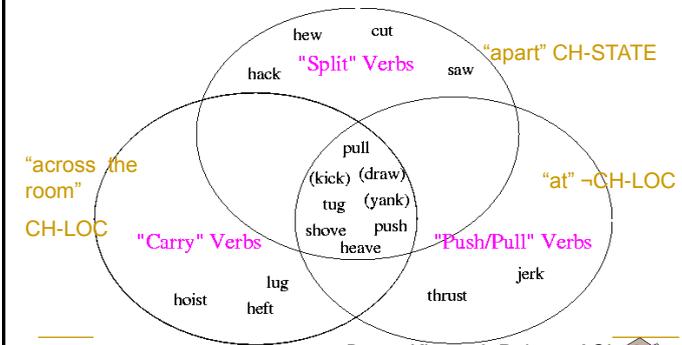


## Confusions in Levin classes?

- Not semantically homogenous
  - {*braided, 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, ACF98



## 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](http://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

NP V NP THAT S

EXAMPLE "I promised him that he would arrive in time."  
SYNTAX AGENT V RECIPIENT TOPIC <<THAT\_COMP>>  
SEMANTICS TRANSFER\_INFO(DURING(E), AGENT, RECIPIENT, TOPIC) CAUSE(AGENT, E)



## VerbNet – Karin Kipper

- Class entries:
  - Capture generalizations about verb behavior
  - Organized hierarchically
  - Members have common semantic elements, semantic roles and syntactic frames
- Verb entries:
  - Refer to a set of classes (different senses)
  - each class member linked to WN synset(s) (not all WN senses are covered)

CLEAR – Colorado



VerbNet: give-13.1 - Mozilla Firefox

http://verbs.colorado.edu/verb-index/give-13.1.php

VerbNet: give-13.1

Members: 6, Frames: 3

**MEMBERS**

GIVE (FN 1; WN 1, 3, 8, 14, 17, SELL (FN 1; WN 1, 4)  
19, 24, 29)

HOCK (WN 1)

LEASE (FN 1, 2; WN 1, 3)

PAWN (WN 1)

RENT (FN 1, 2; WN 1, 2)

**ROLES**

• ASSET

**FRAMES**

NP-PP ASSET-PP

EXAMPLE "He leased the car for \$200 a week."

SYNTAX AGENT V THEME (FOR AT) ASSET

SEMANTICS HAS POSSESSION(START(E), AGENT, THEME) HAS POSSESSION(START(E), ?RECIPIENT, ASSET) HAS POSSESSION(TRANSFER(DURING(E), THEME)

HJCL2008

## 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

CLEAR – Colorado



## 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

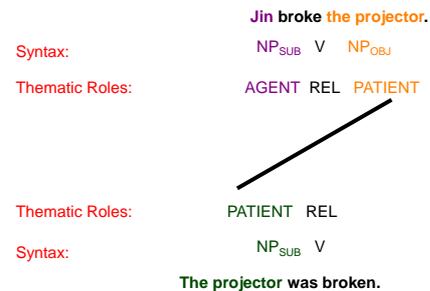


## What is PropBank?

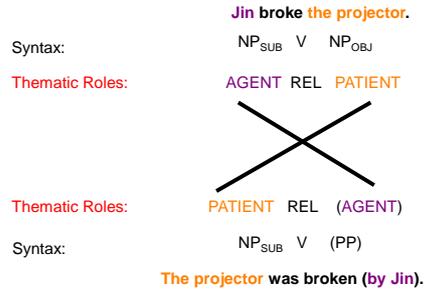
- Semantic information over the syntactically parsed (i.e. treebanked) text
- Semantic information -> predicate argument structure of a verb or a relation
- Unlike VerbNet, the predicate argument structure is specific to the verb or relation in question
- Seeks to
  1. provide consistent argument labels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



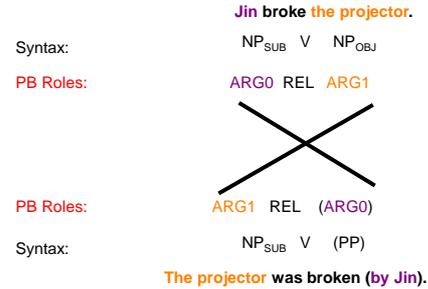
## “1. PB seeks to provide consistent argument labels across different syntactic realizations”



## “1. PB seeks to provide consistent argument labels across different syntactic realizations”



## “1. PB seeks to provide consistent argument labels across different syntactic realizations”

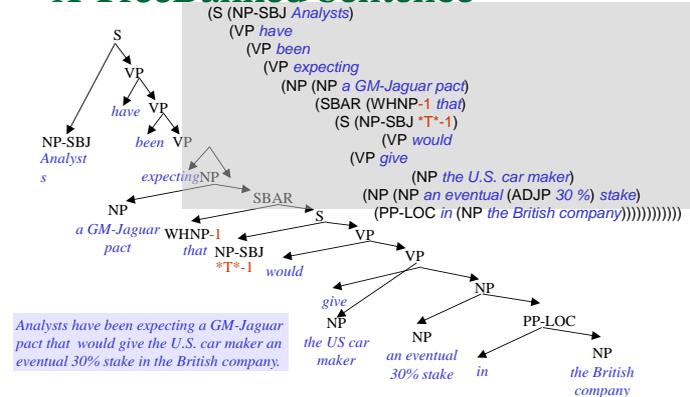


## Why numbered arguments?

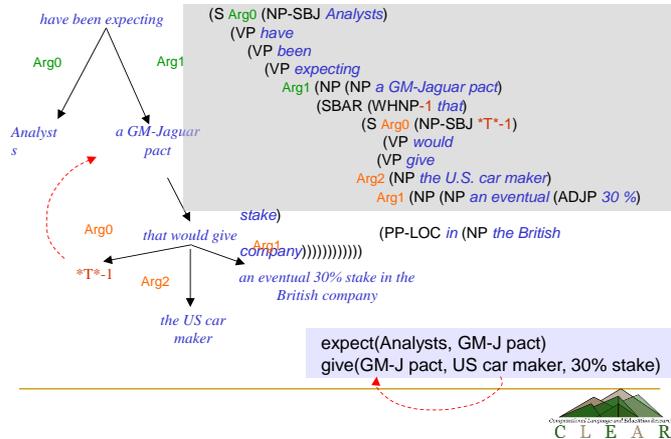
- Avoids lack of consensus concerning a specific set of semantic role labels
- Numbers correspond to labels that are verb-specific
- Arg0 and Arg1 correspond to Dowty's (1991) proto-agent and proto-patient
- Args 2-5 are highly variable



## A TreeBanked Sentence



## The same sentence, PropBanked



## Frames File Example: expect

Roles:

Arg0: expecter

Arg1: thing expected

Example: Transitive, active:

*Portfolio managers* expect *further declines in interest rates*.

Arg0:

*Portfolio managers*

REL:

*expect*

Arg1:

*further declines in interest rates*

## Frames File example: 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*

## "1. PB seeks to provide consistent argument labels across different syntactic realizations"

• Uuuuuusually...

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

## “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 Antarctica*
  - 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



## Frame files

- Argument structure is intended to be consistent across members of a VerbNet class
- Leads to certain arguments that exist in frame but never arise in usage:
  1. *double* has extent arg.
  2. *Hammer* has instrument arg.



## 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



## Returning to *Jin broke the projector*

- Frame for *break*:

Frameset **break.01** "break, cause to not be whole":  
Arg0: breaker  
Arg1: thing broken  
Arg2: instrument  
Arg3: pieces



## Full example

*Jin* broke *the projector* before class because he didn't want to hear Jena talk.

- Given relation *break*:
  - Arg0: *Jin*
  - REL: *broke*
  - Arg1: *the projector*
  - ArgM-TMP: *before class*
  - ArgM-CAU: *because he didn't want to hear Jena talk.*



## Everyday annotation...

