From Propositions to Events

Martha Palmer
University of Colorado

December 16, 2010
Clara Treecourse
Prague, The Czech Republic

Computational Language and EducAtion Research
C L E A R

Recovering implicit arguments

Gerber & Chai, ACL2010

- *[$_{Arg0}$ The two companies] [$_{REL}$ **produce**] [$_{Arg1}$ market pulp, containerboard and white paper]. The goods could be manufactured closer to customers, saving shipping costs.*

Recovering implicit arguments

Gerber & Chai, ACL2010

- *[$_{IArg0}$ The two companies] produce [$_{IArg1}$ market pulp, containerboard and white paper]. The goods could be manufactured closer to customers, saving [$_{REL}$ **shipping**] costs.*

Argument roles for *ship*

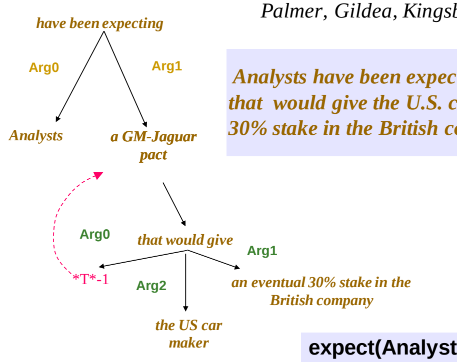
- Agent [+animate | +organization]
- Theme [+concrete]
- Source [+location]
- Destination [+animate | [+location & -region]]

Outline

- Limitations of PropBank and WordNet
- OntoNotes Groupings
- VerbNet
  - Verbs grouped in hierarchical classes
  - Explicitly described class properties
  - More informative semantic role labels
- VerbNet classifer
- Drawing inferences

PropBank – WSJ Penn Treebank

Palmer, Gildea, Kingsbury., CLJ 2005



Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.

expect(Analysts, GM-J pact)
give(GM-J pact, US car maker, 30% stake)

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?

Very, very coarse-grained.....

Limitations to PropBank

- Sense distinctions often so coarse-grained that meaningful inferences cannot be drawn
  - *Postmen carry mail/Genes carry mutations*
- Args2-4 overloaded, poor performance
  - *Rudolph Agnew,…, was named [$_{ARG2\ \{Predicate\}}$ a nonexecutive director of ….]*
  - *…. results appear in … Journal of …, … likely to bring new attention [$_{ARG2\ \{Destination\}}$ to the problem.]*
- WSJ too domain specific
  - Additional Brown corpus annotation & GALE data

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

(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%

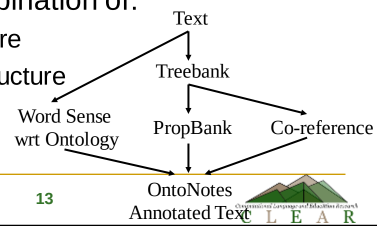
Dang & Palmer, SIGLEX02

Creation of coarse-grained resources

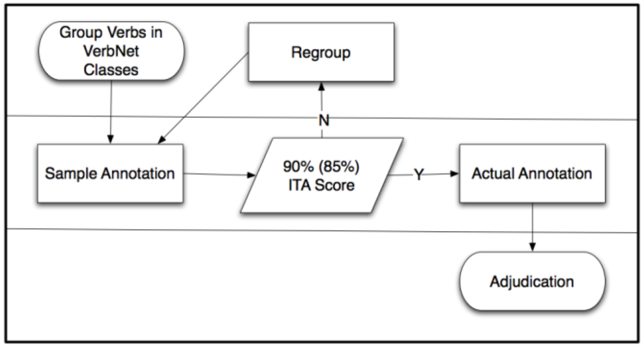
- Unsupervised clustering using rules (Mihalcea & Moldovan, 2001)
- Clustering by mapping WN senses to ODE (Navigli, 2006).
- OntoNotes - Manually grouping WN senses and annotating a corpus (Hovy et al., 2006)
- Supervised clustering WN senses using OntoNotes and another set of manually tagged data (Snow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

OntoNotes Goal: Modeling Shallow Semantics DARPA-GALE

- AGILE Team: BBN, Colorado, ISI, Penn
- Skeletal representation of literal meaning
- Synergistic combination of:
  - Syntactic structure
  - Propositional structure
  - Word sense
  - Coreference



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, JNLE07, 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

OntoNotes Status

- More than 2,500 verbs grouped
- Average ITA per verbs = 89%
- http://verbs.colorado.edu/html_groupings/
- More than 150,000 instances annotated for 2000+ verbs
- WSJ, Brown, ECTB, EBN, EBC, WebText
- Training and Testing
- *How do the groupings connect to other resources?*

Sense Hierarchy

(Palmer, et al, SNLU04 - NAACL04, NLE07, Chen, et. al, NAACL06)

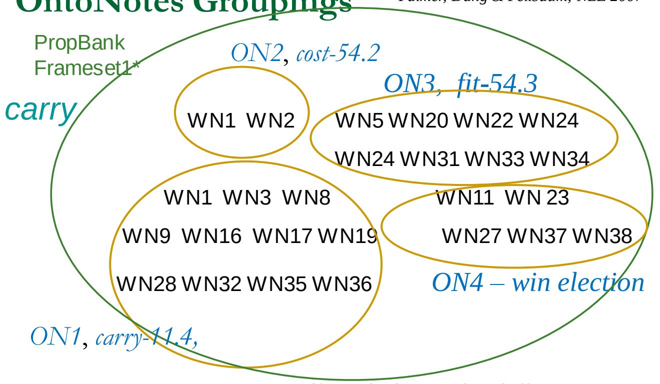
- PropBank Framesets – ITA >90%
  coarse grained distinctions
  20 Senseval2 verbs w/ > 1 Frameset
  Maxent WSD system, 73.5% baseline, 90%
  - Sense Groups (Senseval-2) - ITA 82%
    Intermediate level
    (includes Levin classes) – 71.7%
    - WordNet – ITA 73%
      fine grained distinctions, 64%

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

Chen, Dligach & Palmer, ICSC 2007

SEMLINK-PropBank, VerbNet, WordNet, OntoNotes Groupings

Palmer, Dang & Fellbaum, NLE 2007



*ON5-ON11 carry oneself,carried away/out/off, carry to term

VerbNet: Basis in Theory

- Beth Levin, *English Verb Classes and Alternations* (1993)
- 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?

Limitations to Levin Classes

Dang, Kipper & Palmer, ACL98

- 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

VerbNet – *Karin Kipper Schuler*

- 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) 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*

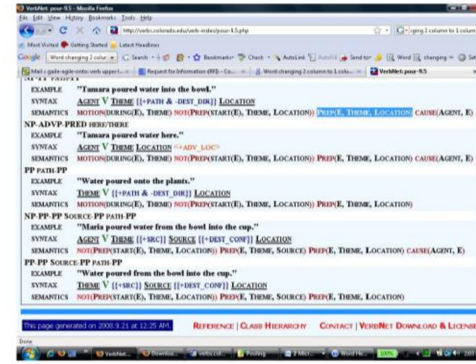
  cause(Agent, E)
  manner(during(E), Motion, Agent)
  contact(during(E), ?Instrument, Patient)
  degradation_material_integrity(result(E), Patient)

- *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 *Pour-9.5* (cont.)



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* (Tamara, water, bowl) →
  CAUSE(Tamara,E),
  MOTION(DURING(E), water),
  NOT(into(START(E), water, bowl)),
  into(E, water, bowl).

VerbNet: *send-11.1* (Members: 11, Frames: 5)

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

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

Frameset id = ***leave.02***	Sense = *give*	VerbNet class = ***future-having 13.3***
Arg0	Giver	Agent/Donor*
Arg1	Thing given	Theme
Arg2	Benefactive	Recipient

*FrameNet Label

Baker, Fillmore, & Lowe, COLING/ACL-98
Fillmore & Baker, WordNetWKSHP, 2001

MappingPropBank/VerbNet

http://verbs.colorado.edu/~mpalmer/verbnet

- Extended VerbNet 5,391 lexemes
  - (100+ new classes from *(Korhonen and Briscoe, 2004; Korhonen and Ryant, 2005))*
  - now covers 91% of PropBank tokens. *Kipper, et. al., LREC-04, LREC-06, LREJ-08, NAACL09 Tutorial*
- Semi-automatic mapping of PropBank instances to VerbNet classes and thematic roles, hand-corrected. (now FrameNet)
- VerbNet class tagging as automatic WSD
- Run SRL, map Arg2 to VerbNet roles, Brown performance improves

Yi, Loper, Palmer, NAACL07

VerbNet classifier

Brown, Dligach, Palmer, IWCS11

- Treated as a verb sense disambiguation task
- One classifier per verb
- 344 multiclass verbs
  - average 2.7 classes
  - average of 133 instances
  - Includes verbs labeled in the corpus with one VerbNet class and "No appropriate class"

Features

- Lexical
  - Neighbor words and their POS
- Syntactic
  - Passive/active
  - Types of phrases and clauses
  - Heads of phrases
- Semantic
  - Synonyms and hypernyms of arguments
  - Named entity features
  - Dynamic dependency neighbors (Dligach, 2008)

Results

- Accuracy: 88.67%
- Baseline (most frequent class): 77.78%
- Error reduction: 49%

Results

Model	Baseline (%)	Accuracy (%)	Error Reduction (%)
Lexical features only	77.78	83.07	23.81
Lexical + syntactic	77.78	84.44	29.97
Lexical + semantic	77.78	83.75	26.87
All but DDN	77.78	84.12	28.53
Lexical + syntactic + DDN	**77.78**	**84.89**	**32.00**
All features	77.78	84.65	30.92

Can SemLink improve Generalization?

- SRL Performance improved from 77% to 88%
  Automatic parses, 81% F, Brown corpus, 68%
- Overloaded Arg2-Arg5
  - PB: verb-by-verb
  - VerbNet: same thematic roles across verbs
- Use VerbNet as a bridge to merge PB and FN and expand the Size and Variety of the Training

Arg2 groupings; (Total count 11068)

Group1 (43.93%)	Group2 (14.74%)	Group3 (32.13%)	Group4 (6.81%)	Group5 (2.39%)
Recipient; Destination; Location; Source; Material; Beneficiary	**Extent; Asset**	**Predicate; Attribute; Theme; Theme2; Theme1; Topic**	**Patient2; Product**	**Instrument; Actor2; Cause; Experiencer**

Process

- Retrain the SRL tagger
  - Original: Arg[0-5,A,M]
  - ARG2 Grouping: Arg[0,2-5,A,M] Arg1-Group[1-6]
- Evaluation
  - WSJ [+6%]
  - Brown [+10%]
- More Coarse-grained or Fine-grained?
  - more specific: data more coherent, but more sparse
  - more general: consistency across verbs even for new domains?

Summary

- Reviewed limitations of PropBank and WordNet
- Described OntoNotes Groupings, VerbNet and Semlink
- VerbNet classifier will be available soon
- Hierarchical mappings of roles for PropBank/VerbNet/Framenet in progress

Acknowledgments

- We gratefully acknowledge the support of the National Science Foundation Grants for , Consistent Criteria for Word Sense Disambiguation and Robust Semantic Parsing, and DARPA-GALE via a subcontract from BBN.