Semantic Role Labeling Tutorial
NAACL, June 9, 2013

Part 1: Martha Palmer, University of Colorado
Part 2: Shumin Wu, University of Colorado
Part 3: Ivan Titov, Universität des Saarlandes
Outline

- Part 1
  Linguistic Background, Resources, Annotation
  Martha Palmer, University of Colorado

- Part 2
  Supervised Semantic Role Labeling and Leveraging Parallel PropBanks
  Shumin Wu, University of Colorado

- Part 3
  Semi- , unsupervised and cross-lingual approaches
  Ivan Titov, Universität des Saarlandes, Universteit van Amsterdam
Motivation: From Sentences to Propositions

Who did what to whom, when, where and how?

- Powell met Zhu Rongji
- Powell and Zhu Rongji met
- Powell met with Zhu Rongji
- Powell and Zhu Rongji had a meeting

Proposition: meet(Powell, Zhu Rongji)

When Powell met Zhu Rongji on Thursday they discussed the return of the spy plane.

- meet(Powell, Zhu)
- discuss([Powell, Zhu], return(X, plane))

...
Capturing semantic roles

SUBJ
- Dan broke [ the laser pointer.]

SUBJ
- [ The windows] were broken by the hurricane.

SUBJ
- [ The vase] broke into pieces when it toppled over.
Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.
Analysts have been expecting a GM-Jaguar pact that would give the US car maker an eventual 30% stake in the British company.
SRL Questions

- Why Arg0 and Arg1?
- What about nouns and adjectives?
- What about other languages?
- How does PropBank relate to VerbNet and FrameNet?
- Will we ever get past the WSJ?
- How do SRL systems get trained?
- Can this be done without training data?
- …..
Why Arg0? Linguistic Background and Resources

- Fillmore – Cases
  - Useful generalizations, fewer sense distinctions,
- Dowty – Proto-typical Agents and Patients
  - A bag of “agentive” entailments
- PropBank

- Levin – Verb classes based on syntax
  - Syntactic behavior is a reflection of the underlying semantics
- VerbNet

- Back to Fillmore and FrameNet
- SemLink
- PropBank ➔ AMR
Linguistic Background: Case Theory, The Case for Case, Charles J. Fillmore

- Case relations occur in deep-structure
  - Surface-structure cases are derived

- A sentence is a verb + one or more NPs
  - Each NP has a deep-structure case
    - A(gentive)
    - I(nstrumental)
    - D(ative) - recipient
    - F(activite) – result
    - L(ocative)
    - O(bjective) – affected object, theme

- Subject is no more important than Object
  - Subject/Object are surface structure
Case Theory Benefits - Generalizations

- **Fewer tokens**
  - Fewer verb senses
  - E.g. *cook/bake* [ ___O(A)] covers
    - Mother is cooking/baking the potatoes
    - The potatoes are cooking/baking.
    - Mother is cooking/baking.

- **Fewer types**
  - “Different” verbs may be the same semantically, but with different subject selection preferences
  - E.g. *like* and *please* are both [ ___O+D]

- Great, let’s do it!
Oops, problems with Cases/Thematic Roles

- How many and what are they?
- Fragmentation: 4 Agent subtypes? (Cruse, 1973)
  - The sun melted the ice./This clothes dryer doesn’t dry clothes well
  - Argument/adjunct distinctions – Extent?
  - The kitten licked my fingers. – Patient or Theme?


\[
\begin{align*}
\text{Agent (or Source) } & \quad \text{Esau} \ \text{sold} \ \text{Theme} \ \text{his birthright} \\
\text{Goal} & \quad \text{to Jacob} \ \text{for a bowl of porridge}.
\end{align*}
\]

\[
\begin{align*}
\text{Goal } & \quad \text{Esau} \ \text{sold} \ \text{his birthright} \\
\text{Source} & \quad \text{to Jacob} \ \text{for a Theme bowl of porridge}.
\end{align*}
\]

Jackendoff
Role definitions have to be determined verb by verb, and with respect to the other roles

- Event-dependent Proto-roles introduced
  - Proto-Agent
  - Proto-Patient
- Prototypes based on shared entailments
Proto-Agent- the *mother*

- **Properties**
  - Volitional involvement in event or state
  - Sentience (and/or perception)
  - Causes an event or change of state in another participant
  - Movement (relative to position of another participant)
  - (exists independently of event named)
    *may be discourse pragmatic*
Proto-Patient – the *cake*

- **Properties:**
  - Undergoes change of state
  - Incremental theme
  - Causally affected by another participant
  - Stationary relative to movement of another participant
  - (does not exist independently of the event, or at all)
  - *may be discourse pragmatic*
Argument Selection Principle

- For 2 or 3 place predicates
- Based on empirical count (total # of entailments for each role).
  - Greatest number of Proto-Agent entailments $\rightarrow$ Subject;
  - greatest number of Proto-Patient entailments $\rightarrow$ Direct Object.
- Alternation predicted if number of entailments for each role similar (non-discreteness).

[Mother AGENT] baked a cake.
[The cake PATIENT] baked.
PropBank Semantic Role Labels – based on Dowty’s Proto-roles

- PropBank Frame for *break*:

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

- Why numbered arguments?
  - Lack of consensus concerning semantic role labels
  - Numbers correspond to verb-specific labels
  - Arg0 – Proto-Agent, and Arg1 – Proto-Patient, (Dowty, 1991)
  - Args 2-5 are highly variable and overloaded – poor performance
PropBank seeks to provide consistent argument labels across different syntactic realizations

- **Uuuuuuusuually…**
  - Arg0 = agent, experiencer
  - Arg1 = patient, theme
  - Arg2 = benefactive / instrument / attribute / end state
  - Arg3 = start point / benefactive / instrument / attribute
  - Arg4 = end point
PropBank 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*
Statistics also revealed that Taiwanese business investment is tending to increase.
Statistics also revealed that Taiwanese business investment is tending to increase.
Statistics also revealed that Taiwanese business investment is tending to increase.
Statistics also revealed that Taiwanese business investment is tending to increase.
Why do we need Frameset ID’s?

PropBank Frames Files: tend.01, care for

Roles:
- Arg0: tender
- Arg1: thing tended (to)

Example: John tends to the needs of his patrons.
- Arg0: John
- REL: tend
- Arg1: the needs of his patrons
Sense distinctions in PropBank – coarse-grained

PropBank - Frames Files: tend.02, have a tendency

Roles:
- Arg1: Theme
- Arg2: Attribute

Example: The cost, or premium, tends to get fat in times of crisis.
- Arg1: The cost, or premium
- REL: tend
- Arg2: to get fat in times of crisis.
Visual Example: traces BASED on Jubilee

Choi, et. al., NAACL-10 Demo
Statistics also revealed that Taiwanese business investment is tending to increase.
Actual data for *leave*

Leave .01 “move away from” Arg0 rel Arg1 Arg3
Leave .02 “give” Arg0 rel Arg1 Arg2

sub-ARG0 obj-ARG1 44
sub-ARG0 20
sub-ARG0 NP-ARG1-with obj-ARG2 17
sub-ARG0 sub-ARG2 ADJP-ARG3-PRD 10
sub-ARG0 sub-ARG1 ADJP-ARG3-PRD 6
sub-ARG0 sub-ARG1 VP-ARG3-PRD 5
NP-ARG1-with obj-ARG2 4
obj-ARG1 3
sub-ARG0 sub-ARG2 VP-ARG3-PRD 3
Annotation procedure, WSJ PropBank

PTB II - Extraction of all sentences with given verb
Create Frame File for that verb  
Paul Kingsbury
(3100+ lemmas, 4400 framesets, 118K predicates)
Over 300 created automatically via VerbNet

First pass: Automatic tagging  
(Joseph Rosenzweig)
http://www.cis.upenn.edu/~josephr/TIDES/index.html#lexicon

Second pass: Double blind hand correction  
Paul Kingsbury

Tagging tool highlights discrepancies  
Scott Cotton

Third pass: Solomonization (adjudication)
Betsy Klipple, Olga Babko-Malaya
Annotator accuracy – ITA 84%

Annotator Accuracy-Primary labels only

# of annotations (log scale)

accuracy

0.86 0.87 0.88 0.89 0.9 0.91 0.92 0.93 0.94 0.95 0.96

defiler
forbesk
istas
stran
wia mstr
kingsbur
ksledge
nryant
jaywang
malayao
ptepper
delikan
SRL Questions

- Why Arg0 and Arg1?
- What about nouns and adjectives?
- What about other languages?
- How does PropBank relate to VerbNet and FrameNet?
- Will we ever get past the WSJ?
- How do SRL systems get trained?
- Can this be done without training data?
- …..
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
  - 47 top level classes, 193 second and third level, 3100 verbs

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

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

- Synonyms, syntactic patterns (conative), relations
Break Levin class - Change-of-state

- break
  - chip
  - crack
  - crash
  - crush
  - fracture
  - rip
  - tear
  - split
  - splinter
  - snap
  - smash
  - shatter
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
Intersective Levin Classes

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

VERBNET
verbs.colorado.edu/verb-index/index.php
VerbNet – Karin Kipper Schuler

- Class entries:
  - Capture generalizations about verb behavior
  - Organized hierarchically
  - Members have common semantic elements, *semantic roles* (28) and syntactic frames

- Verb entries:
  - Refer to a set of classes (different senses)
  - each class member linked to WN synset(s) and FrameNet frames

- Currently 6300 verbs

- Adding Constructions
  
  *Hwang, et.al, NAACL-HLT Construction Workshop, 2010*
  
  *Bonial, et. al., ACL RELMS Workshop, 2011*
VerbNet example – *Pour-9.5*

How does VerbNet relate to FrameNet?
The **lexical unit** (Cruse 1986), – a pairing of a word with a sense (or a FrameNet frame.)

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

  - *Our family observes* the Jewish dietary laws.
  - *You have to observe* the rules or you’ll be penalized.
The FrameNet database constitutes

- a set of **frame descriptions**
  - **Frames, Frame Elements, Valence Possibilities**
- 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.
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, …
Our family observes the dietary laws. The light switches in this room are in full conformity with the building code.
words, frames, lexical units

2 lexical units sharing same form:
Compliance.\textit{observe},
Perception.\textit{observe}
Mapping from PB to VerbNet - SemLink

http://verbs.colorado.edu/semlink
Mapping from PropBank to VerbNet
(similar mapping for PB-FrameNet) - SemLink

<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
PropBank/FrameNet - SemLink

Buy
Arg0: buyer
Arg1: goods
Arg2: seller
Arg3: rate
Arg4: payment

Sell
Arg0: seller
Arg1: goods
Arg2: buyer
Arg3: rate
Arg4: payment

More generic, more neutral – maps readily to VN, TR

Rambow, et al, PMLB03
Can SemLink improve Generalization?

- After PropBank, SRL 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
- Example
  - Rudolph Agnew,…, was named [ARG2 {Predicate} a nonexecutive director of this British industrial conglomerate.]
  - ….the latest results appear in today’s New England Journal of Medicine, a forum likely to bring new attention [ARG2 {Destination} to the problem.]
- Use VerbNet as a bridge to merge PB and FN and expand the Size and Variety of the Training
VerbNet - Arg2 groupings; (Total count 11068)

<table>
<thead>
<tr>
<th>Group1 (43.93%)</th>
<th>Group2 (14.74%)</th>
<th>Group3 (32.13%)</th>
<th>Group4 (6.81%)</th>
<th>Group5 (2.39%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recipient;</td>
<td>Extent; Asset</td>
<td>Predicate;</td>
<td>Patient2;</td>
<td>Instrument;</td>
</tr>
<tr>
<td>Destination;</td>
<td></td>
<td>Attribute;</td>
<td>Product</td>
<td>Actor2;</td>
</tr>
<tr>
<td>Location;</td>
<td></td>
<td>Theme;</td>
<td></td>
<td>Cause;</td>
</tr>
<tr>
<td>Source;</td>
<td></td>
<td>Theme2;</td>
<td></td>
<td>Experiencer</td>
</tr>
<tr>
<td>Material;</td>
<td></td>
<td>Theme1;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beneficiary</td>
<td></td>
<td>Topic</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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?
Complementary resources

Redundancy is harmless, may even be useful

PropBank provides the best training data

VerbNet provides the clearest links between syntax and semantics

FrameNet provides the richest semantics

Together they give us the most comprehensive coverage

SemLink - [http://verbs.colorado.edu/semlink/](http://verbs.colorado.edu/semlink/)

- WSJ, sense tags and SRL, mappings to VN and FN
Pierre Vinken, 61 years old, will join the board as a nonexecutive director Nov. 29.
Annotated Data – Current PropBank Status

Pradhan, et.al., IJSC 2007, Albright, et. al., JAMIA, 2013, Palmer, et. al., ICON-09

- DARPA-GALE, OntoNotes 5.0
  - BBN, Brandeis, Colorado, Penn
  - Multilayer structure: NE, TB, PB, WS, Coref
  - Three languages: English, Arabic, Chinese
  - Several Genres (@ ≥ 200K): NW, BN, BC, WT
    - Close to 2M words @ language (less PB for Arabic)
  - Parallel data, E/C, E/A
  - PropBank frame coverage for rare verbs
  - Recent PropBank extensions

- Clinical Notes – 400K available, goal is 700K
- Hindi/Urdu PropBank, 400K Hindi, 200K Urdu
- BOLT – discussion forums, SMS, email, Egyptian
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 – Current Count of Senses (lexical units)

<table>
<thead>
<tr>
<th>Language</th>
<th>Final Count</th>
<th>Estimated Coverage in Running Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>10,615*</td>
<td>99%</td>
</tr>
<tr>
<td>Chinese</td>
<td>24,642</td>
<td>98%</td>
</tr>
<tr>
<td>Arabic</td>
<td>7,015</td>
<td>99%</td>
</tr>
</tbody>
</table>

- Only 111 English adjectives
Included in OntoNotes 5.1: Extensions to PropBank

- Original annotation coverage:
  - PropBank: verbs; past participle adjectival modifiers
  - NomBank: relational and eventive nouns.

- Substantial gap – now bridging
  - Uniform treatment of light verbs,
  - Additional predicative adjectives,
  - Eventive nouns
Gaps in proposition coverage

- Event Coreference chains include nominalizations with and without light verbs
  - “China has threatened to slap sanctions on American companies that sell arms to its rival Taiwan as part of a range of punitive actions Beijing is taking to protest the deal...‘China will make further judgments as appropriate,’ Xinhua reported.”
  - Light verb/nominalization examples:
    - slap sanctions, taking actions, make judgments

- PropBank structures for eventive nouns
  - sanction(China, US companies),
  - act(China),
  - judge(China, US companies)
English Noun and LVC annotation

- **Example Noun: Decision**
  - Roleset: \texttt{Arg0: decider, Arg1: decision…}
  
    - “…[\texttt{your_{ARG0}}] \texttt{decision_{REL}}
        [to say look I don't want to go through this anymore_{ARG1}]”

- **Example within an LVC: Make a decision**
  - “…[\texttt{the President_{ARG0}}] \texttt{made_{REL-LVB}}
    the \texttt{fundamentally correct_{ARGM-ADJ}}
    \texttt{decision_{REL}}  \texttt{to get on offense_{ARG1}}”
China will make further judgments as appropriate.

**Verb - REL: [make]**
- Arg0: China,
- ArgPRX: further judgments as appropriate.

**Noun -- RELPRX: [judgment]**
- Arg0: China
- ArgM-PRD: as appropriate

**Merged -- REL: RELPRX: [make] [judgment]**
- Arg0: China
- ArgM-PRD: as appropriate
Abstract Meaning Representations – AMR, Maximal Use of PropBank Frame Files,

*Knight, et. al., LAW-2013*

He was not aware of research on smokers of the Kent cigarettes.

(r / realize-01
 :polarity -
 :ARG0 (h / he)
 :ARG1 (r3 / research-01
      :ARG1 (p4 / person
           :ARG0-of (s / smoke-02
                    :ARG1 (c2 / cigarette
                         :name (k / name
                              :op1 "Kent")))))))

To get to canonical concept, we stem to English verbs,
where PropBank arguments are best described.

General direction of stemming:
adverb → adjective → noun → verb
SRL Questions

- Why Arg0 and Arg1?
- What about nouns and adjectives?
- What about other languages?
- How does PropBank relate to VerbNet and FrameNet?
- Will we ever get past the WSJ?
- How do SRL systems get trained?
- Can this be done without training data?
- …
Acknowledgments

- We gratefully acknowledge the support of the National Science Foundation Grants for Robust Semantic Parsing, Richer Representations for Machine Translation, A Bayesian Approach to Dynamic Lexical Resources for Flexible Language Processing and DARPA-GALE via a subcontract from BBN and DARPA-BOLT via a subcontract from LDC.

- Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.
And thanks to

- **Postdocs:** Paul Kingsbury, Dan Gildea, Chunhye Han, Nianwen Xue,

- **Students:** Joseph Rosenzweig, Hoa Dang, Tom Morton, Karin Kipper Schuler, Jinying Chen, Na-Rae Han, Szu-Ting Yi, Edward Loper, Susan Brown, Dmitriy Dligach, Kevin Cohen, Jinho Choi, Lee Becker, Jena Hwang, Will Corvey, Claire Bonial, Shumin Wu, Wei-Te Chen, Kevin Stowe,

- **Collaborators:** Suzanne Stevenson, Annie Zaenen, Orin Hargraves
References

- Albright, Daniel; Lanfranchi, Arrick; Fredriksen, Anwen; Styler, William; Warner, Collin; Hwang, Jena; Choi, Jinho; Dligach, Dmitriy; Nielsen, Rodney; Martin, James; Ward, Wayne; Palmer, Martha; Savova, Guergana. 2013. Towards syntactic and semantic annotations of the clinical narrative. *Journal of the American Medical Informatics Association.*, 2013;0:1-9. doi:10.1136/amiajnl-2012-001317


- Bonial, Claire; Susan Windisch Brown; Jena D. Hwang; Christopher Parisien; Martha Palmer; Suzanne Stevenson, Incorporating Coercive Constructions into a Verb Lexicon, *In the RELMS Workshop, held in conjunction with the Association of Computational Linguistics Meeting*, Portland, Oregon, June, 2011.


References, cont.


References, cont.


