SemLink+: FrameNet, VerbNet, and Event Ontologies

Martha Palmer,
Claire Bonial, Diana McCarthy
University of Colorado

Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929 – 2014)
ACL Workshop
June 27, 2014
Outline

- Deep NLU?
- Where we are now
- Where we need to go
- More details about where we need to go
- The contributions and limitations of lexical resources to this process
Where we are now – shallow semantics

- Syntactic Structure – parse trees, Treebanks
- Semantic types – nominal entities [Person, Location, Organization], NE tagging
- Semantic roles – Agents, [PropBank FrameNet, VerbNet]
- Sense distinctions – *call me a taxi, call me an idiot*, WordNet, OntoNotes groups, FrameNet, VerbNet, vectors, etc.
- Coreference – [*President Obama: he*]
Where we are now - DETAILS

- 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

- DARPA BOLT – discussion forum, SMS
  - PropBank extensions: light verbs, function tags on core args, nominalizations, adjectives, constructions, often relying on FrameNet
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.

FrameNet and VerbNet should have the same coverage, and we (or at least VerbNet) desperately need help to do this semi-automatically!!
WordNet: - call, 28 senses, 9 groups

- WN5, WN16, WN12: Loud cry
- WN3, WN19, WN1: Label
- WN18, WN27, WN2: Challenge
- WN2, WN13, WN28: Phone/radio
- WN17, WN11, WN24: WN10, WN14, WN21, Bid
- WN15, WN26: Bird or animal cry
- WN4, WN7, WN8, WN9: Request
- WN20, WN25: Call a loan/bond
- WN6, WN23: Visit
SEMLINK-PropBank, VerbNet, FrameNet, WordNet, OntoNotes

PropBank
Frameset1*

carry

cost-54.2, ON2

fit-54.3, ON3

WN1 WN2

WN5 WN20 WN22 WN24

WN24 WN31 WN33 WN34

WN1 WN3 WN8

WN9 WN16 WN17 WN19

WN28 WN32 WN35 WN36

WN11 WN 23

WN27 WN37 WN38

ON4 – win election

carry-11.4, CARRY,-FN ,ON1

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

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

- Sense Groups (Senseval-2/OntoNotes) - ITA 89%
  Intermediate level
  (includes Verbnet/some FrameNet) – SVM, 88+% 
  *Dligach & Palmer, ACL2011*

- WordNet – ITA 73%
  fine grained distinctions, 64%
Extended VerbNet: 6,340 senses
  92% PB tokens (8114 verb senses/12,646 all)
Type-type mapping PB/VN, VN/FN, VN/WN
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; Croce, et. al., ACL2012
Run SRL, map Arg2 to VerbNet roles, Brown performance improves  Yi, Loper, Palmer, NAACL07
“Saucedo said that guerrillas in one car opened fire on police standing guard, while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building, and a third car rushed the attackers away.”

Saucedo said – reporting event, evidential
What we can do

- *that guerrillas in one car opened fire on police standing guard*

- *opened fire* = aspectual context,
  - fire(guerrillas, police)

- *standing guard* = support verb construction/aspectual?, reduced relative
  - guard(police, X)
What we can do, cont.

- while a second car **carrying** 88 pounds (40 kgs) of dynamite **parked** in front of the building

- *carrying* - reduced relative, correct head noun - pounds or dynamite?
  - carry(car2, dynamite)

- park(car2, front_of(building))
What we can do, cont.

- *and a third car rushed the attackers away*

- rush(car3, attackers, away)
“Saucedo said that guerrillas in one car opened fire on police standing guard, while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building, and a third car rushed the attackers away.

- guarding BEFORE/OVERLAP firing
- Narrative container – TimeX

  - [firing, parking, rushing] all overlap, all in the same temporal bucket?
  - [see Styler, et. al, ACL2014, Events Workshop & RED Guidelines]
Don’t mark the relations between EVENTS.

Instead, put EVENTS in temporal buckets and relate the buckets
“Saucedo said that guerrillas in one car opened fire on police standing guard, while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building, and a third car rushed the attackers away

- guarding BEFORE/OVERLAP firing
- X CONTAINS [firing, parking, rushing]
- firing BEFORE parking
- parking BEFORE rushed
Implicit arguments

- *that guerrillas in one car opened fire on police standing guard*

- *opened fire* = aspectual context,
  - fire(guerillas, police)

- *standing guard* = support verb construction or aspectual?, reduced relative
  - guard(police, X)
More compelling example
(thanks to Vivek Srikumar)

- The bomb exploded in a crowded marketplace. Five civilians were killed, including two children. Al Qaeda claimed responsibility.

- Killed by Whom?
- Responsibility for what?

- Need recovery of implicit arguments
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,
VerbNet: \textit{send-11.1} (Members: 11, Frames: 5) includes \textit{“ship”}\n
- **Roles**
  - Agent [+animate | +organization]
  - Theme [+concrete]
  - Source [+location]
  - Destination [+animate | [+location & -region]]

- **Syntactic Frame:** NP V NP PP.\textit{destination}
  - example "\textit{Nora sent the book to London.}"
  - syntax Agent V Theme \{to\} Destination
  - semantics motion(during(E), Theme)\n    location(end(E), Theme, Destination)\n    cause(Agent, E)
Recovering Implicit Arguments*

* AKA definite null complements

[ Arg0 The two companies] [ REL1 produce] [ Arg1 market pulp, containerboard and white paper]. The goods could be manufactured closer to customers, saving [ REL2 shipping] costs.

- Used VerbNet for subcategorization frames
Implicit arguments

- **SYNTAX**  
  Agent V Theme {to} Destination
  
  `[AGENT] shipped [THEME] to [DESTINATION]`

- **SEMANTICS**
  - `CAUSE(AGENT,E)`
  - `MOTION(DURING(E), THEME),`
  - `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**

- \(\text{CAUSE(Companies, E)}\)
- \(\text{MOTION(DURING(E), goods)}\),
- \(\text{LOCATION(END(E), goods, customers)}\),

*Can annotate, semi-automatically!*
Another type of Implicit Relation
Example from Daniel Marcu, GALE Wrap-up Mtg

- Between Munich and LA you need less than 11 hours by plane.

- You can fly to Los Angeles from Munchen in no more than eleven hours.

- From Munich to Los Angeles, it does not take more than eleven hours by plane.
Constructions allow us to

- Recognize a path prepositional phrase, and that it necessarily goes with a “MOTION” event – Caused-motion constructions
  - John sneezed the tissue off the table.
  - Mary blinked the snow off of her eyelashes.
- If we detect a MOTION event we can associate the *plane* with it as a vehicle
- Just the *plane* itself can suggest a motion event…
Construction Grammar

- In Construction Grammar

  - constructions are carriers of meaning
  - constructions are assigned meaning in the same way that words are – via convention rather than composition.

- Invaluable resource –
  FrameNet Constructicon, Cxn Viewer
Introduce a constructional \``layer\'' to VerbNet, which attaches orthogonally to relevant VerbNet classes.
VerbNet can also provide inferences – sometimes…

- 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)`
  `location(E,grape_arbor,path)`
Inferences can inform Coreference?

“Saucedo said that guerrillas in one car opened fire on police standing guard, while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building, and a third car rushed the attackers away.”

AMR

- guerilla – Arg0 of fire.01
- attacker – person – Arg0-of attack.01
FrameNet Attack frame

attacker.n, fire.n

Loosely connected to “fire.v” via Hostile Encounter
Needed – An Event Ontology

- That can provide appropriate levels of generalization
- DEFT - Event Ontology conference calls
  - Martha Palmer, James Pustejovsky, Annie Zaenen, Diana McCarthy, Teruko Mitamura, German Rigau, Ann Bies, Kira Griffit, Julie Fitzgerald, Claire Bonial, Derek Palmer

- Map ERE event types to FrameNet?
- Develop an upper level event ontology that ERE and FN can both be mapped to?
Conflict events

- **ERE**
  - Attack events
  - Protest/Demonstration events

- **FrameNet**
  - Attack events - See previous slide
  - Protest, not present
  - Demonstration – Reasoning frame

- **VerbNet**
  - Attack - Judgment
  - Protest - Conspire
  - Demonstrate – Transfer_message
Using protégé for Ontologies

• Will coordinate with SUMO, WN also

• Any advice?

• Could really use some help from an experienced user

• Could also REALLY use input from all of these “clustering” techniques!
Predicate Matrix - Lacalle, Laparra, Rigau, LREC 2014 -

WordNet

VerbNet

A: WN Senses (206,941)
B: WN Verb senses (25,041)
C: VN predicates (6293)
D: PB predicates (6,181)
E: FN lexical-units (10,195)
F: FN verb lexical-units (4095)

PropBank

FrameNet
Predicate Matrix

B: WN Verb senses (25,041)
C: VN predicates (6,293)
D: PB predicates (6,181)
F: FN verb lexical-units (4,095)

SemLink < UNION (B,C,D,E)
Predicate Matrix

- **First version 1.0 (GWC 2014)**
  - SemLink +
  - Monosemous verbs from VN +
  - Synonyms from WN

- **Second version 1.1 (LREC 2014)**
  - SemLink +
  - Automatic mappings between predicates +
    - WN-VN and WN-FN (**new mappings!**)
  - Project VN roles to FN roles (**complete gaps!**) +
  - Synonyms from WN
Where we need to go

- Recovery of implicit arguments
- Recovery of implicit relations
- Better Entity coreference
- Event coreference
- Temporal and Causal ordering of events
- Generalizations over event types
Lexical resources can provide

- Generalizations about subcat frames & roles
- Backoff classes for OOV items for portability
- Semantic similarities/"types” for verbs
- Event type hierarchies for inferencing
- Need to be unified and empirically validated and extended: Semlink+
  - VN & FN need PB like coverage, and techniques for extension and automatic domain adaptation

**Hybrid lexicons – symbolic and statistical lexical entries?**
Acknowledgments

- We gratefully acknowledge the support of the National Science Foundation Grants for Consistent Criteria for Word Sense Disambiguation, Robust Semantic Parsing, Richer Representations for Machine Translation, A Bayesian Approach to Dynamic Lexical Resources for Flexible Language Processing, DARPA-GALE via a subcontract from BBN, DARPA-BOLT & DEFT via a subcontract from LDC, and DTRA: SemLink+ via a subcontract from BBN, and NIH THYME.

- 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, DARPA or NIH.
And thanks to

- Postdocs: Paul Kingsbury, Dan Gildea, Nianwen Xue, Jinying Chen
- Students: Joseph Rosenzweig, Hoa Dang, Tom Morton, Karin Kipper Schuler, Jinying Chen, Szu-Ting Yi, Edward Loper, Susan Brown, Dmitriy Dligach, Jena Hwang, Will Corvey, Claire Bonial, Jinho Choi, Lee Becker, Shumin Wu, Kevin Stowe
- Collaborators: Christiane Fellbaum, Suzanne Stevenson, Annie Zaenen, Orin Hargraves, James Pustejovsky