Combining Lexical Resources:

Mapping Between PropBank and VerbNet

Edward Loper,Szu-ting Yi, Martha Palmer

September 2006

Using Lexical Information

- Many interesting tasks require…

  Information about lexical items…

  and how they relate to each other.
- E.g., question answering.

  Q: *Where are the grape arbors located?*

  A: Every path from back door to yard was covered by a grape-arbor, and every yard had fruit trees.

Lexical Resources

- Wide variety of lexical resources available
  - VerbNet, PropBank, FrameNet, WordNet, etc.
- Each resource was created with different goals and different theoretical backgrounds.
  - Each resource has a different approach to defining word senses.

SemLink:
Mapping Lexical Resources

- Different lexical resources provide us with different information.
- To make useful inferences, we need to *combine* this information.
- In particular:
  - **PropBank** -- How does a verb relate to its arguments? Includes **annotated text.**
  - **VerbNet** -- How do verbs w/ shared semantic & syntactic features (and their arguments) relate?
  - **FrameNet** -- How do verbs that describe a common scenario relate?
  - **WordNet** -- What verbs are synonymous?
  - **Cyc** -- How do verbs relate to a knowledge based ontology?

* Martha Palmer, Edward Loper, Andrew Dolbey, Derek Trumbo, Karin Kipper, Szu-Ting Yi

PropBank

- 1M words of WSJ annotated with predicate-argument structures for verbs.
  - The location & type of each verb's arguments
- Argument types are defined on a per-verb basis.
  - Consistent across uses of a single verb (sense)
- But the same tags are used (Arg0, Arg1, Arg2, …)
  - Arg0 ≈ proto-typical agent (Dowty)
  - Arg1 ≈ proto-typical patient

PropBank:
cover (smear, put over)

- Arguments:
  - Arg0 = causer of covering
  - Arg1 = thing covered
  - Arg2 = covered with
- Example:

John ***covered*** the bread with peanut butter.

PropBank: Trends in Argument Numbering

- Arg0 = proto-typical agent (*Dowty*)
  Agent (85%), Experiencer (7%), Theme (2%), …
- Arg1 = proto-typical patient (*Dowty*)
  Theme (47%),Topic (23%), Patient (11%), …
- Arg2 = Recipient (22%), Extent (15%), Predicate (14%), …
- Arg3 = Asset (33%), Theme2 (14%), Recipient (13%), …
- Arg4 = Location (89%), Beneficiary (5%), …
- Arg5 = Location (94%), Destination (6%)

PropBank: Adjunct Tags

- Variety of ArgM's (Arg#>5):
  - TMP: when?
  - LOC: where at?
  - DIR: where to?
  - MNR: how?
  - PRP: why?
  - REC: himself, themselves, each other
  - PRD: this argument refers to or modifies another
  - ADV: others

Limitations to PropBank as Training Data

- Args2-5 seriously overloaded → poor performance
  - VerbNet and FrameNet both provide more fine-grained role labels
- 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.]

Limitations to PropBank as Training Data (2)

- WSJ too domain specific & too financial.
- Need broader coverage genres for more general annotation.
  - Additional Brown corpus annotation, also GALE data
  - FrameNet has selected instances from BNC

How Can SemLink Help?

- In PropBank, Arg2-Arg5 are overloaded.
  - But in VerbNet, the same thematic roles across verbs.
- PropBank training data is too domain specific.
  - Use VerbNet as a bridge to merge PropBank w/ FrameNet

    → Expand the size and variety of the training data

VerbNet

- Organizes verbs into classes that have common syntax/semantics linking behavior
- Classes include…
  - A list of member verbs (w/ WordNet senses)
  - A set of thematic roles (w/ selectional restr.s)
  - A set of frames, which define both syntax & semantics using thematic roles.
- Classes are organized hierarchically

VerbNet Example

No Comments

contiguous_location-47.8

POST COMMENT

Members: 37, Frames: 1

MEMBERS

BESTRIDE	EDGE (WN 1)	HEAD (WN 1)
BLANKET (FN 1; WN 1, 2)	ENCIRCLE (FN 1; WN 1)	HUG (WN 1)
BORDER (WN 1, 2, 3)	ENCLOSE (WN 1, 2)	LINE (FN 1; WN 1)

ROLES

- THEME1 [+CONCRETE]
- THEME2 [+CONCRETE]

FRAMES

BASIC TRANSITIVE

EXAMPLE	"Italy borders France"
SYNTAX	THEME1 V THEME2
SEMANTICS	CONTACT(DURING(E), THEME1, THEME2) EXIST(DURING(E), THEME1) EXIST(DURING(E), THEME2)

What do mappings look like?

- 2 Types of mappings:
  - **Type mappings** describe which entries from two resources might correspond; and how their fields (e.g. arguments) relate.
    - Potentially many-to-many
    - Generated manually or semi-automatically
  - **Token mappings** tell us, for a given sentence or instance, which type mapping applies.
    - Can often be thought of as a type of classifier
      - Built from a single corpus w/ parallel annotations
    - Can also be though of as word sense disambiguation
      - Because each resource defines word senses differently!

Mapping Issues

- Mappings are often many-to-many
  - Different resources focus on different distinctions
- Incomplete coverage
  - A resource may be missing a relevant lexical item entirely.
  - A resource may have the relevant lexical item, but not in the appropriate category or w/ the appropriate sense
- Field mismatches
  - It may not be possible to map the field information for corresponding entries. (E.g., predicate arguments)
    - Extra fields
    - Missing fields
    - Mismatched fields

VerbNet↔PropBank Mapping: Type Mapping

- Verb class ↔ Frame mapped when PropBank was created.
  - Doesn't cover all verbs in the intersection of PropBank & VerbNet
    - This intersection has grown significantly since PropBank was created.
- Argument mapping created semi-automatically
- Work is underway to extend coverage of both

VerbNet↔PropBank Mapping: Token Mapping

- Built using parallel VerbNet/PropBank training data
  - Also allows direct training of VerbNet-based SRL
- VerbNet annotations generated semi-automatically
  - Two automatic methods:
    - Use WordNet as an intermediary
    - Check syntactic similarities
  - Followed by hand correction

Using SemLink: Semantic Role Labeling

- Overall goal:
  - Identify the semantic entities in a document & determine how they relate to one another.
- As a machine learning task:
  - Find the predicate words (verbs) in a text.
  - Identify the predicates' arguments.
  - Label each argument with its semantic role.
- Train & test using PropBank

Current Problems for SRL

- PropBank role labels (Arg2-5) are not consistent across different verbs.
  - If we train within verbs, data is too sparse.
  - If we train across verbs, the output tags are too heterogeneous.
- Existing systems do not generalize well to new genes.
  - Training corpus (WSJ) contains a highly specialized genre, with many domain-specific verb senses.
  - Because of the verb-dependant nature of PropBank role labels, systems are forced to learn based on verb-specific features.
  - These features do not generalize well to new genres, where verbs are used with different word senses.
  - System performance drops on the Brown corpus

Improving SRL Performance w/ SemLink

- Existing PropBank role labels are too heterogeneous
  - So subdivide them into new role label sets, based on the SemLink mapping.
- Experimental Paradigm:
  - Subdivide existing PropBank roles based on what VerbNet thematic role (Agent, Patient, etc.) it is mapped to.
  - Compare the performance of:
    - The original SRL system (trained on PropBank)
    - The mapped SRL system (trained w/ subdivided roles)

Subdividing PropBank Roles

- Subdividing based on *individual VerbNet theta roles* leads to very sparse data.
- Instead, subdivide PropBank roles based on *groups of VerbNet roles*.
- Groupings created manually, based on analysis of argument use & suggestions from Karin Kipper.
- Two groupings:
  1. Subdivide Arg1 into 6 new roles:
     $Arg1_{Group1}$, $Arg1_{Group2}$, …, $Arg1_{Group6}$
  2. Subdivide Arg2 into 5 new roles:
     $Arg2_{Group1}$, $Arg2_{Group2}$, …, $Arg2_{Group5}$
- Two test genres: Wall Street Journal & Brown Corpus

Arg1 groupings
(Total count 59,710)

Group1 (53.11%)	Group2 (23.04%)	Group3 (16%)	Group4 (4.67%)	Group5 (.20%)
Theme; Theme1; Theme2; Predicate; Stimulus; Attribute	**Topic**	**Patient; Product; Patient1; Patient2**	**Agent; Actor2; Cause; Experiencer**	**Asset**

Arg2 groupings
(Total count 11,068)

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**

Experimental Results:
What do we expect?

- By subdividing PropBank roles, we make them more coherent.
  … so they should be easier to learn.
- But by creating more role categories, we increase data sparseness.
  … so they should be harder to learn.
- Arg1 is more coherent than Arg2
  … so we expect more improvement from the Arg2 experiments.
- WSJ is the same genre that we trained on; Brown is a new genre.
  … so we expect more improvement from Brown corpus experiments.

Experimental Results:
Wall Street Journal Corpus

	Precision	Recall	F1
Arg1-Original	89.24	77.32	82.85
Arg1-Mapped	90.00	76.35	82.61
Difference	+0.76	-1.03	-0.24
Arg2-Original	73.04	57.44	64.31
Arg2-Mapped	84.11	60.55	70.41
Difference	+11.07	+3.11	+6.10

Experimental Results:

Brown Corpus

	Precision	Recall	F1
Arg1-Original	86.01	71.46	78.07
Arg1-Mapped	88.24	71.15	78.78
Difference	+2.23	-0.31	+0.71
Arg2-Original	66.74	52.22	58.59
Arg2-Mapped	81.45	58.45	68.06
Difference	+14.71	+6.23	+9.47

Conclusions

- By using more coherent semantic role labels, we can improve machine learning performance.
  - Can we use learnability to help evaluate role label sets?
- The process of mapping resources helps us improve them.
  - Helps us see what information is missing (e.g., roles).
  - Semi-automatically extend coverage.
- Mapping lexical resources allows to combine information in a single system.
  - Useful for QA, Entailment, IE, etc…