Joint Learning of Preposition Senses and Semantic Roles of Prepositional Phrases

Daniel Dahlmeier\textsuperscript{1}  Hwee Tou Ng\textsuperscript{1,2}  Tanja Schultz\textsuperscript{3}

\textsuperscript{1}NUS Graduate School for Integrative Sciences and Engineering
\textsuperscript{2}Department of Computer Science
National University of Singapore
\textsuperscript{3}Cognitive Systems Lab
University of Karlsruhe

Conference on Empirical Methods in Natural Language Processing
The Semantics of Prepositional Phrases

Why prepositions are interesting

- Prepositions and prepositional phrases (PPs) convey important information, e.g.,:
  1. Daniel ate a sandwich in the morning
  2. Daniel ate a sandwich in the kitchen.

- Useful for tasks like information extraction, question answering, etc.

- Prepositions are among the most frequent words in English

Why prepositions are difficult

- Prepositions are highly ambiguous

- Semantics is highly complex and even difficult for humans [Chodorow et al.2007]
Two pillars of semantic processing

Word sense disambiguation – The Preposition Project (TPP) [Litkowski and Hargraves2005]
Find the correct meaning of a preposition in a particular context from a predefined list of senses, e.g.,
1. enclosed or surrounded by something else: I’m living in London.
2. expressing a period of time: They met in 1885.
3. the length of time before a future event: I’ll see you in fifteen minutes.

Semantic role labeling – PropBank [Palmer et al.2005]
Find the arguments of a (verb) predicate in a sentence and label each with a semantic role, e.g.,

[It]Arg0 [operates]REL [stores]Arg1 [in Iowa and Nebraska.]ArgM-LOC
WSD and SRL of PPs are similar

Where did Daniel eat the sandwich? → in the kitchen (ARGM-LOC)

Sense of preposition in: 1 – enclosed or surrounded by something else
WSD and SRL of PPs are similar (cont.)

▶ When did Daniel eat the sandwich? → *in the morning* (ARGM-TMP)

▶ Sense of preposition *in*: 2 – expressing a period of time
but still quite different

Different focus

- PropBank **semantic roles** focus on the **verb**
- TPP **senses** focus on the **preposition**
- Different level of granularity

No simple mapping TPP ↔ PropBank

1. She now lives with relatives [in/1 Alabama.] ArgM-LOC
2. The envelope arrives [in/1 the mail.] Arg4
3. [In/5 separate statements] ArgM-LOC the two sides said . . .

- Examples 1 and 2: same sense but different roles
- Examples 1 and 3: same role, but different senses
Our contribution

Observation

- Preposition senses (TPP) and semantic roles (PropBank) offer two different inventories of “meaning” for PPs.
- Knowing the sense could help to find the correct semantic role and vice-versa
- Question: can WSD help SRL and/or can SRL help WSD?

Contribution

A joint probabilistic model to infer the semantic role of the PP and the sense of the preposition that is the lexical head of the PP.
Models
Preposition WSD Baseline Model

Find the probability of sense $s$, given the context $c$

$$\hat{s} = \arg\max_s P(s|c) = \arg\max_s P(s|\psi(c)) \quad (1)$$

where $\psi(\cdot)$ is a feature map.

- Adapt WSD classifier from [Lee and Ng2002]
- Three knowledge sources:
  - Part of speech (POS) of surrounding words
  - Single words in surrounding context (bag of words)
  - Local collocations (uni-, bi- and trigrams)
- Learning algorithm: maximum entropy model
Daniel/NNP ate/VBD a/DT sandwich/NN in/IN the/DT kitchen/NN ./.  

| Feature                  | Values                                      |
|--------------------------|---------------------------------------------|
| POS                      | [VBD, DT, NN, IN, DT, NN]                   |
| Bag of words             | {daniel, ate, sandwich, kitchen}            |
| Local collocations       | [sandwich, the, a, kitchen,                 |
|                          | a sandwich, sandwich _ the, the kitchen     |
|                          | ate a sandwich, a sandwich _ the,           |
|                          | sandwich _ the kitchen, the kitchen .]      |

\[
\hat{s} = \arg\max_s P(s|c) = \arg\max_s P(s|\text{VBD, DT, NN, IN, DT, NN, ...})
\] (2)
SRL Baseline Model

Find the probability of a role $a$ for a given parse tree $t$, predicate $p$ and constituent node $v$.

\[
\hat{a} = \arg\max_a P(a|t, p, v) = \arg\max_a P(a|\Phi(t, p, v))
\]  

(3)

where $\Phi(\cdot)$ is a feature map.

- Learning algorithm: maximum entropy model
- Standard SRL features from the literature ([Gildea and Jurafsky2002, Xue and Palmer2004, Pradhan et al.2005, Jiang and Ng2006])
- Assume correct parse tree and semantic role identification
Joint Learning of Preposition Senses and Semantic Roles of Prepositional Phrases

Daniel Dahlmeier, Hwee Tou Ng, Tanja Schultz

Introduction
Models
Data Set
Experiments
Results
Conclusion

\[ \hat{a} = \operatorname{argmax}_{a} P(a|t, p, v) = \operatorname{argmax}_{a} P(a|\text{pred}=\text{ate}, \text{path}=\ldots) \]
Find the probability of sense $s$, given the context $c$ and the semantic role $a$.

$$\hat{s} = \arg\max_s P(s|c, a) = \arg\max_s P(s|\psi(c), a) \quad (6)$$

- **Training**: use the gold-standard semantic role
- **Testing**: semantic role is automatically predicted
SRL Pipeline Model

Find the probability of a role $a$ for a given parse tree $t$, predicate $p$ and constituent node $v$ and the sense $s$.

$$\hat{a} = \arg\max_a P(a|t, p, v, s) = \arg\max_a P(a|\Phi(t, p, v)s)$$ (7)

- Training: use the gold-standard semantic role
- Testing: semantic role is automatically predicted
Joint Model

Our joint model

Maximize the joint probability of the sense and semantic role:

$$\hat{(a, s)} = \arg \max_{(a, s)} P(a, s \mid t, p, v, c)$$  \hspace{1cm} (8)

Making some independence assumptions, we can factor the probability as follows:

$$\hat{(a, s)} = \arg \max_{(a, s)} P(a \mid t, p, v) \times P(s \mid c, a)$$  \hspace{1cm} (9)

$$= \arg \max_{(a, s)} P(a \mid \Phi(t, p, v)) \times P(s \mid \Psi(c, a))$$  \hspace{1cm} (10)

$$\hat{(a, s)} = \arg \max_{(a, s)} P(a \mid \text{pred=ate, path=...}) \times P(s \mid \text{VBD,DT,..., a})$$  \hspace{1cm} (11)
Joint Learning of Preposition Senses and Semantic Roles of Prepositional Phrases

Daniel Dahlmeier, Hwee Tou Ng, Tanja Schultz

Introduction
Models
Data Set
Experiments
Results
Conclusion

Data Set

PropBank sections 2–4 and section 23

- Manually annotate instances of top 7 prepositions [Jurafsky and Martin 2008]: *at, for, in, of, on, to, with* with TPP senses
- 3854 annotated PPs
- Inter-annotator agreement 86%

| Preposition | Total | Training | Test |
|-------------|-------|----------|------|
| at          | 404   | 260      | 144  |
| for         | 478   | 307      | 171  |
| in          | 1590  | 1083     | 507  |
| of          | 97    | 51       | 46   |
| on          | 408   | 246      | 162  |
| to          | 532   | 304      | 228  |
| with        | 345   | 211      | 134  |
| **Total**   | **3854** | **2462** | **1392** |
Experiments
The Experiments

Compare three models for each task

1. Baseline model: independent models for each task

Propbank Data \(\rightarrow\) SRL \(\rightarrow\) Semantic Roles

Propbank Data \(\rightarrow\) Preposition WSD \(\rightarrow\) Senses
The Experiments

Compare three models for each task

1. Baseline model: independent model for each task
2. Pipeline model: add 1-best prediction as an additional feature
The Experiments

Compare three models for each task

1. Baseline model: independent model for each task
2. Pipeline model: add 1-best prediction as an additional feature
The Experiments

Compare three models for each task

1. Baseline model: independent model for each task
2. Pipeline model: add 1-best prediction as an additional feature
3. Joint model: solve both tasks together

![Diagram showing the flow from Propbank Data to Joint Model, then Senses and Semantic Roles]

Daniel Dahlmeier, Hwee Tou Ng, Tanja Schultz
Results
Joint Learning of Preposition Senses and Semantic Roles of Prepositional Phrases

Daniel Dahlmeier, Hwee Tou Ng, Tanja Schultz

Introduction

Models

Data Set

Experiments

Results

Conclusion

Preposition WSD Results

Figure: Coarse-grained classification accuracy of WSD models

Table: Accuracy on the WSD task, statistically significantly improvements
SRL Results

Figure: $F_1$ measure on the SRL classification task

| Semantic role | Baseline | Pipeline | Joint   |
|---------------|----------|----------|---------|
| ArgM-LOC      | 72.88    | 71.54    | 74.27*  |
| ArgM-TMP      | 81.87    | 79.43    | 83.24*  |
| Overall (A)   | 71.71    | 69.47    | 72.14   |

Table: $F_1$ measure on the SRL task, statistically significantly improvements over the baseline are marked with (*)
Results

WSD

- Significant improvement for the pipeline and joint model
- Semantic role information helps WSD, but joint learning does not
- SRL feature adds information from predicate-argument structure which is not encoded in WSD features

SRL

- Modest improvement for the joint model
- Significant improvement for ArgM-LOC and ArgM-TMP
- Joint model more robust than pipeline model
Why the Joint Model can help

“He camped out at a high-tech nerve center on the floor of the big board, . . .”

- Correct sense: at_1: expressing location or arrival
- Correct semantic role: ArgM-LOC
Joint Learning of Preposition Senses and Semantic Roles of Prepositional Phrases

Daniel Dahlmeier, Hwee Tou Ng, Tanja Schultz

Introduction

Models

Data Set

Experiments

Results

Conclusion

“He/PRP camped/VBD out/RP at/IN a/DT high-tech/JJ nerve/NN center/NN”

WSD

- Local collocations unobserved (except ‘a’)
- Single words ‘camp’, ‘high-tech’ ‘nerve’, and ‘center’ unobserved
- POS not very discriminative
- Most frequent sense: at_3 – denoting a point or level on a scale (e.g., prices start at £18,500)

SRL

- Predicate ‘camped’ unobserved
- Last word ‘center’ suggests Arg1 (theme, e.g., Japan needs a business center)

|               | Gold  | Baseline |
|---------------|-------|----------|
| Sense         | at_1  | at_3     |
| Semantic role | ArgM-LOC | Arg1  |
“He camped out \text{[at/3 a high-tech nerve center]}_{\text{Arg1}}”

**WSD**
- WSD pipeline model’s prediction unchanged from the baseline

**SRL**
- The (wrong) sense feature ’at_3’ makes the SRL pipeline model change its guess to ArgM-MNR
- Nice try, but wrong again...

|                | Gold  | Baseline | Pipeline      |
|----------------|-------|----------|---------------|
| Sense          | at_1  | at_3     | at_3          |
| Semantic role  | ArgM-LOC | Arg1    | ArgM-MNR     |
“He camped out \([at/? a high-tech nerve center]?)\)”

**Joint**

- ARG1 is the *theme* of the action (*John broke the window.*)
- at\_3 denotes a point or level on a scale (*drive at 50 mph*)
- Unlikely combination (only 1 example in training data)
- Joint model “abandons” the most likely assignment for each *individual* task and finds the more likely (and correct) *joint* assignment.
- Hurray!

|               | Gold  | Baseline | Pipeline | Joint       |
|---------------|-------|----------|----------|-------------|
| **Sense**     | at\_1 | at\_3    | at\_3    | at\_1       |
| **Semantic role** | ArgM-LOC | Arg1     | ArgM-MNR  | ArgM-LOC    |
Conclusion

- We proposed a probabilistic model to jointly classify preposition senses and the semantic role of the dominating PP
- Our model showed a significant improvement over a WSD baseline model and a modest (but significant) improvement over an SRL baseline model for spatial and temporal roles
- To the best of our knowledge, this is the first joint model for preposition senses and semantic roles
Thank you!
Backup Slides
Why SRL can help WSD

“Mario Gabelli, for instance, holds cash positions well above 20% in several of his funds.”
“Mario Gabelli, for instance, holds cash positions well above 20% in several of his funds.”

- Baseline: *in* sense 2 – period of time
- ‘%’ is a strong hint for sense 2, e.g., “prices rose by 0.2% in October”.
- Governing verb ‘holds’ outside the local context window
- SRL correctly predicts ArgM-LOC which helps to find the correct sense: sense 1 – surrounded by or enclosed in
Why WSD can help SRL

“Rightly or wrongly, many giant institutional investors appear to be fighting the latest war by applying the lessons they learned in the October 1987 crash. . .”
“Rightly or wrongly, many giant institutional investors appear to be fighting the latest war by applying the lessons they learned in the October 1987 crash…”

- SRL baseline: ArgM–LOC instead of ArgM-TMP
- Predominant sense for ‘in... crash’ indeed ArgM-LOC, e.g., “died in an airplane crash”.
- The temporal clue ’October’ is not included in the SRL features.
- WSD correctly predicts sense 2 –period of time which helps to find the correct SRL role ArgM-TMP.
References I

Martin Chodorow, Joel Tetreault, and Na-Rae Han.
2007.
Detection of grammatical errors involving prepositions.
In *Proceedings of the Fourth ACL-SIGSEM Workshop on Prepositions*, pages 25–30.

Daniel Gildea and Daniel Jurafsky.
2002.
Automatic Labeling of Semantic Roles.
*Computational Linguistics*, 28(3):245–288.

Zheng Ping Jiang and Hwee Tou Ng.
2006.
Semantic Role Labeling of NomBank: A Maximum Entropy Approach.
In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, pages 138–145.

Daniel Jurafsky and James H. Martin.
2008.
*Speech and Language Processing*.
Prentice-Hall, Inc. Upper Saddle River, NJ, USA.

Yoong Keok Lee and Hwee Tou Ng.
2002.
An Empirical Evaluation of Knowledge Sources and Learning Algorithms for Word Sense Disambiguation.
In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, pages 41–48.
References II

Kenneth C. Litkowski and Orin Hargraves.
2005.
The Preposition Project.
In *Proceedings of the 2nd ACL-SIGSEM Workshop on “The Linguistic Dimensions of Prepositions and Their Use in Computational Linguistic Formalisms and Applications”*, pages 171–179.

Martha Palmer, Paul Kingsbury, and Daniel Gildea.
2005.
The Proposition Bank: An Annotated Corpus of Semantic Roles.
*Computational Linguistics*, 31(1):71–106.

Sameer Pradhan, Kadri Hacioglu, Valerie Krugler, Wayne Ward, James H. Martin, and Daniel Jurafsky.
2005.
Support Vector Learning for Semantic Argument Classification.
*Machine Learning*, 60(1):11–39.

Nianwen Xue and Martha Palmer.
2004.
Calibrating Features for Semantic Role Labeling.
In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)*, pages 88–94.