DELPH-IN Grammars in CoNLL 2009 Shared Task

Yi Zhang[♠♡], Stephan Oepen[♣]

[♠]LT-Lab, DFKI, Germany

[♡]Computational Linguistics, Saarland University, Germany

[♣]Informatics, University of Oslo, Norway

DELPH-IN Summit 2009
Barcelona, Spain

Outline

CoNLL 2009 Shared Task

Syntactic and Semantic Dependencies in Multiple Languages

- Joint learning of syntactic and semantic dependencies
- Multilingual
  - Catalan (AnCora)
  - Chinese (PCTB+CPropBank)
  - Czech (PDT)
  - English (PTB+PropBank+NomBank)
  - German (Tiger+Salsa)
  - Japanese (Kyoto)
  - Spanish (AnCora)
- Different original annotation converted into a uniformed format
- In-domain and out-domain tests (for cs, en, de)

Motivation

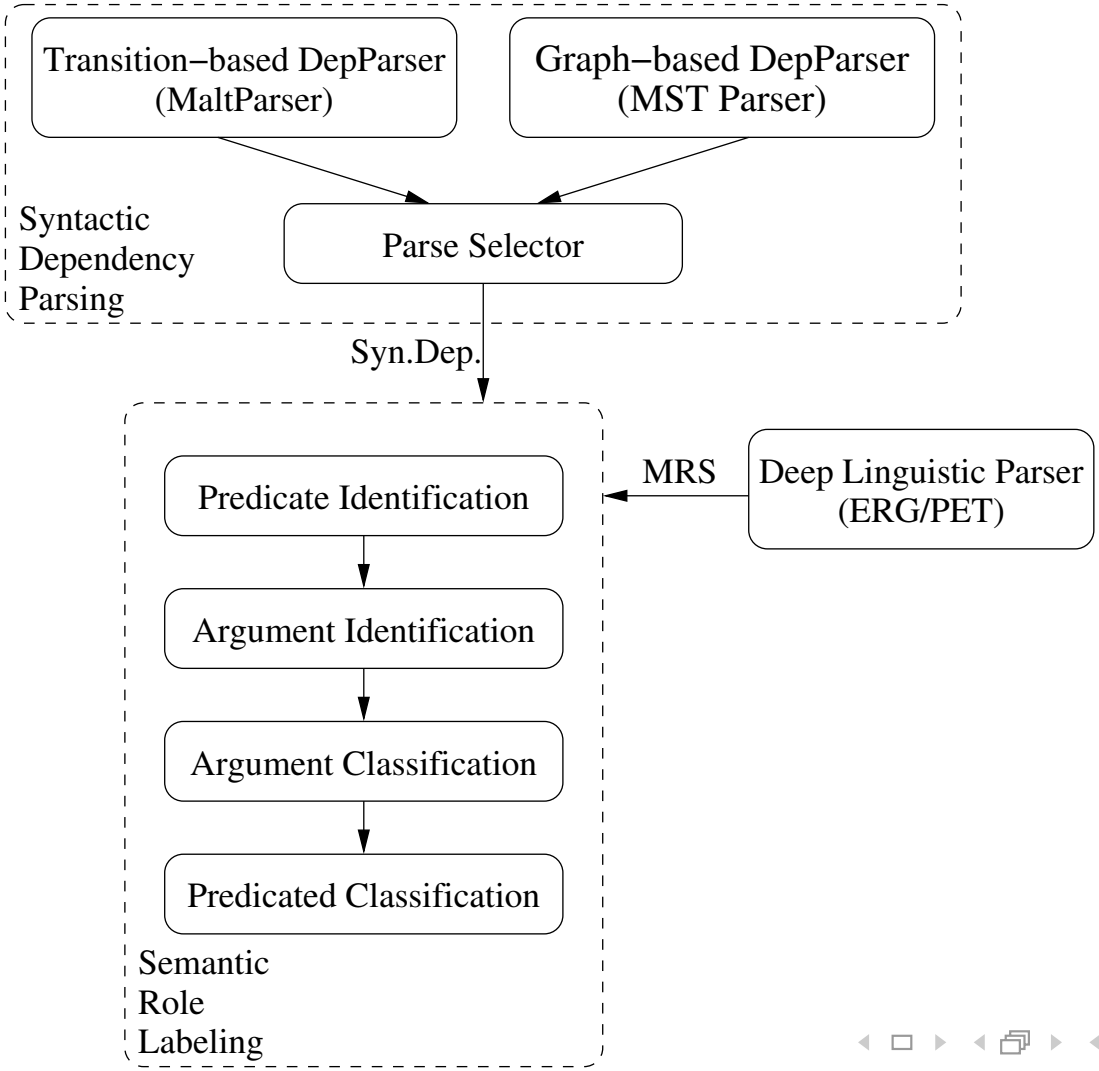
- Show benefits of applying deep parsing in learning tasks
- Road-testing grammar coverage

Previous Participation

CoNLL 2008 Shared Task

- Semantic (MRS) features help to improve the accuracy of SRL for English
- Improvement is more significant in the out-domain test

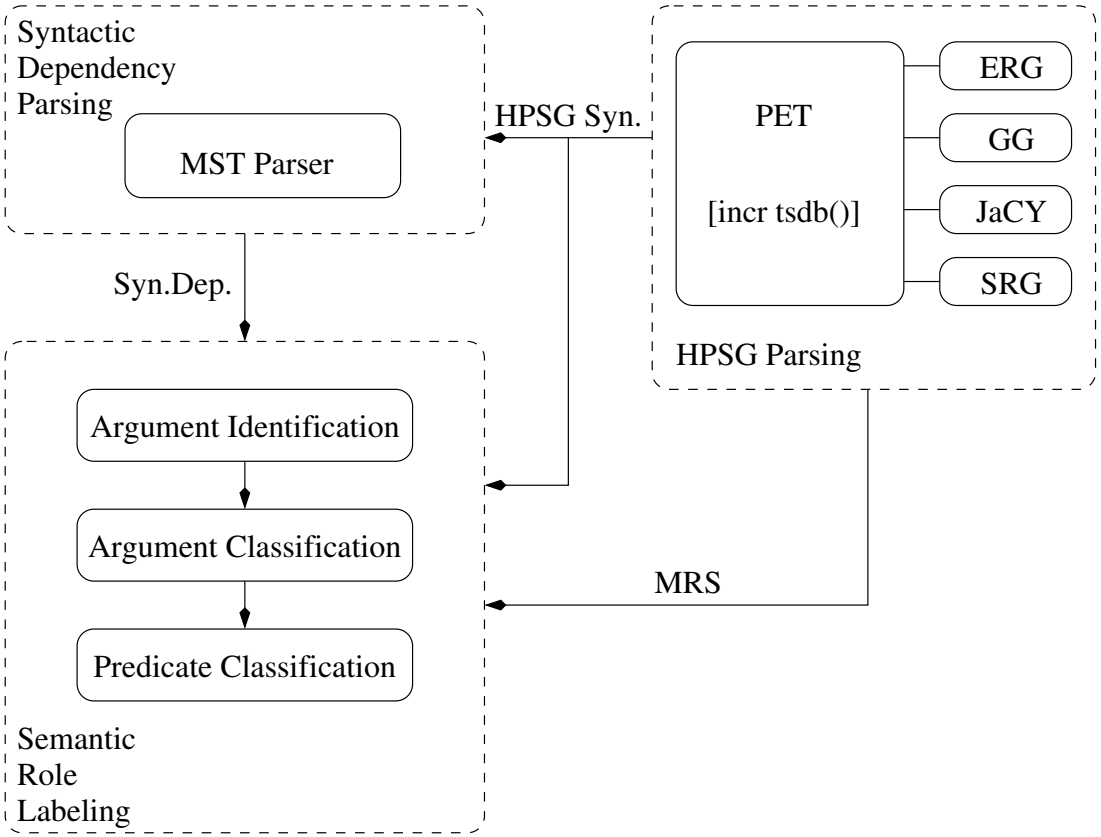
System Architecture 2008



What's New Now?

- Multilingual
- Gather both syntactic and semantic features from deep parses
- Feed deep features to both statistical parser and SRL
- Retrain parse disambiguation models with given training data

System Architecture 2009



Deep Grammar Resources

Grammar	Coverage	Speed
ERG	80.4%	10.06s
GG	28.6%	3.41s
JaCY	42.7%	2.13s
SRG	7.5%	0.80s

Parsing Setup

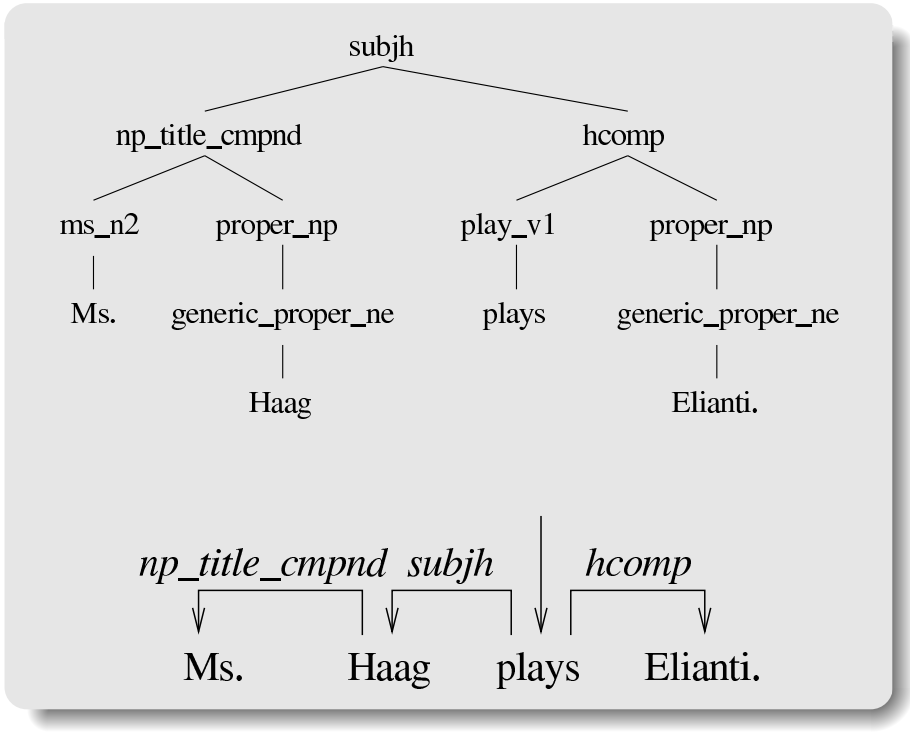
- Based on LOGON tree, avaialbe in SVN as a separate branch

  svn co http://svn.emmtee.net/snug/conll09

- Chart-mapping for preprocessing
- POS-based unknown word handling

Converting HPSG Derivations to Dependency Backbones

- Use head finding heuristics (language and grammar dependent) to identify the head word of each phrase in the derivation tree
- For binary branches, create a dependency relation from the head word of the non-head daughter to the head word of the head daughter, named after the HPSG rule



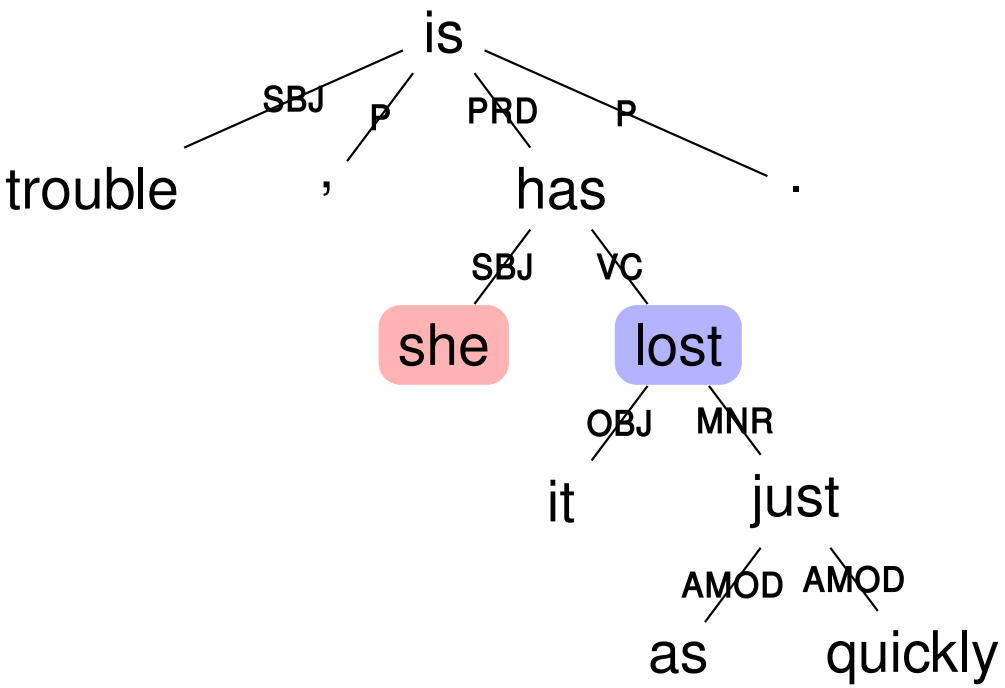
Updating Parse Disambiguation Models

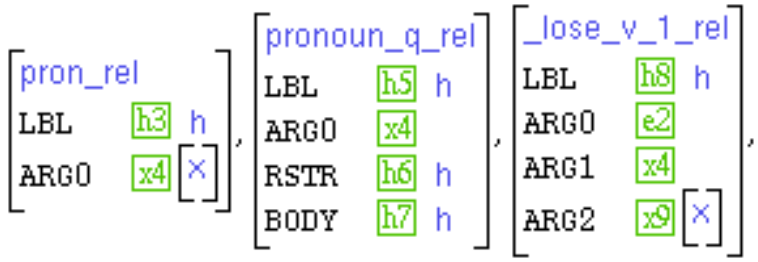
- Unlabeled dependency agreement can be calculated between the CoNLL syntactic dependencies and HPSG dependency backbone
- Parse disambiguation models are retrained to maximize the dependency agreement score
- Experiments indicate a positive correlation between the DA score and HPSG parse quality

Deep Syntactic Features

- POS of the DB parent from the predicate and/or argument
- DB label of the argument to its parent (AI/AC)
- Labeled path from predicate to argument in DB (AI/AC)
- POSes of the predicate's DB dependents

Deep Semantic (MRS) Features





- P MRS EP-name: _lose_v_1_rel
- P MRS-args labels: ARG1 ARG2
- P MRS-args POSes: PRP PRP
- A MRS EP-name: pron_rel
- A MRS-preds labels: ARG1
- A MRS-preds POSes: VBZ

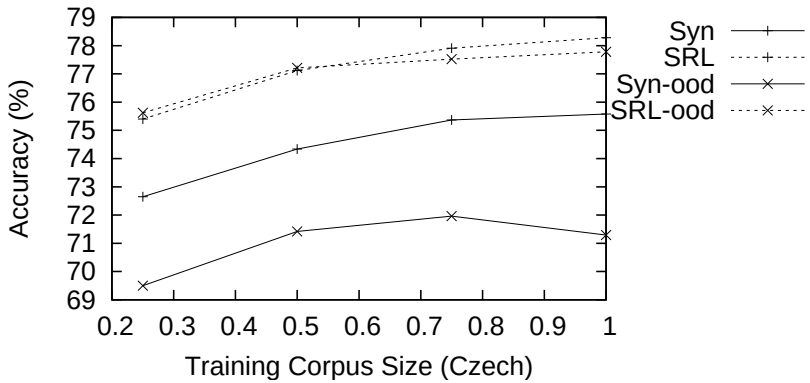
Using Deep Features for Dependency Structure Learning

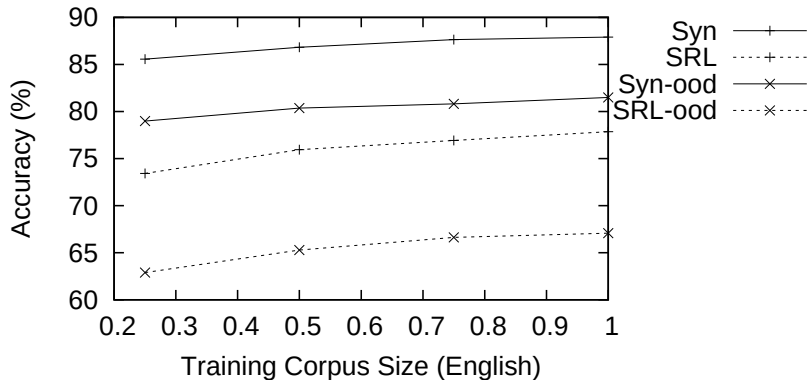
- HPSG dependency backbone features are fed back to the statistical dependency parser, and achieved better out-domain parsing performance [Zhang and Wang, 2009]
- Both HPSG dependency backbone features and MRS features are integrated into the SRL pipeline (MEM classifiers), leading to variying levels of improvements

Evaluation Results

		ca	zh	cs	en	de	ja	es
SYN	Closed	82.67	73.63	75.58	87.90	84.57	91.47	82.69
	ood	-	-	71.29	81.50	75.06	-	-
SRL	Closed	67.34	73.20	78.28	77.85	62.95	64.71	67.81
	ood	-	-	77.78	67.07	54.87	-	-
	Open	-	-	-	78.13 (↑0.28)	64.31 (↑1.36)	65.95 (↑1.24)	68.24 (↑0.43)
	ood	-	-	-	68.11 (↑1.04)	58.42 (↑3.55)	-	-

Learning Curves





Conclusion

- Conclusions from CoNLL 2008 shared task participation has been confirmed on multiple languages, even if the grammar coverage is low

References I

Zhang, Y. and Wang, R. (2009).
Cross-Domain Dependency Parsing Using a Deep Linguistic Grammar.
In *Proceedings of Proceedings of ACL-IJCNLP 2009*, Singapore.
to appear.