Online statistics for a unification-based dialogue parser

Micha Elsner, Mary Swift, James Allen, Daniel Gildea (melsner, swift, james, gildea)@cs.rochester.edu
Department of Computer Science, University of Rochester, Rochester, NY 14627

Deep Parsing with Unification

Advantages

*More complex trees
*Word sense disambiguation
*Theta roles labeled

Disadvantages

*Lots of ambiguity
***Efficiency** problems
***Parse selection**
*Evaluation of semantics?
*Hand-written grammars take **lots of development time and effort**.

Can we solve these problems and still retain the advantages?

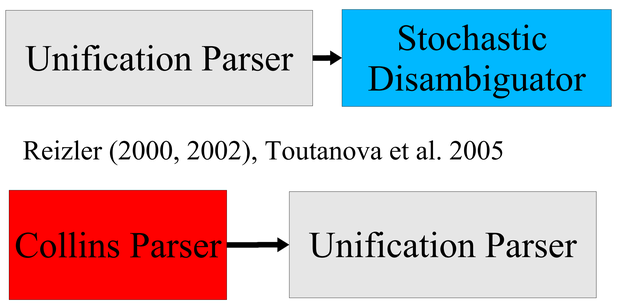
Corpus Statistics

Popular modern parsers use statistics:
*Pruning techniques for higher **speed**.
*Ranking for better **accuracy**.
*Easier to adapt to new domains where treebanks exist.

Disadvantage: making these models requires **lots of labeled training data**.

Intutition: combine statistical models with unification parsing.

Previous Uses of Statistics

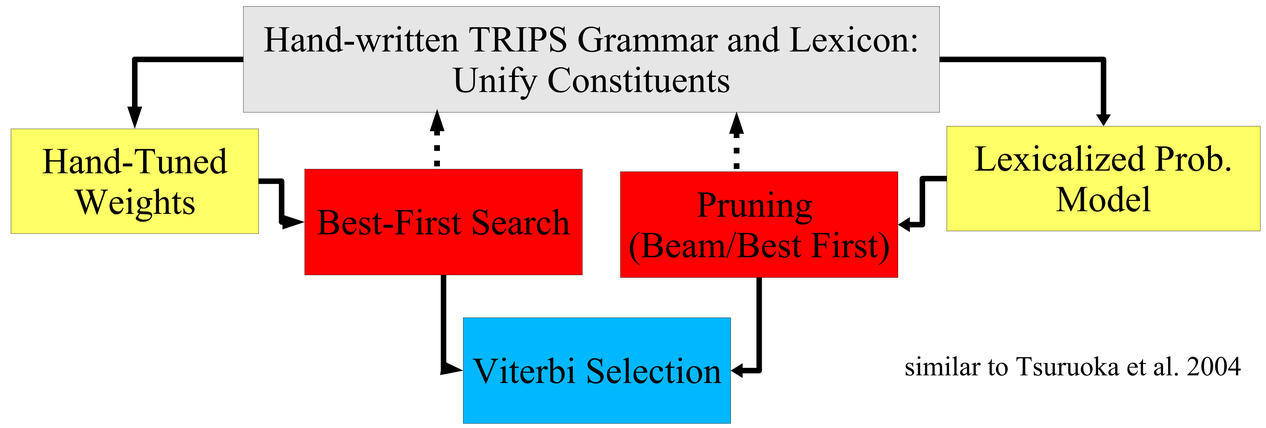


TRIPS Dialogue Parser

(Allen et al. 1996)

Augmented TRIPS

(model based on Bikel '00)



similar to Tsuruoka et al. 2004

Bracketing Accuracy: Correct Syntactic Structure

100 Random s.	Bikel	TRIPS	Aug. TRIPS
Precision	79.40%	78.20%	77.08%
Recall	79.40%	76.77%	76.09%
Complete Match	42	65	46
Constits. Fewer	-	0.00%	**36.96%**

Trained on **9282** sentences of collaborative emergency response dialogues.
Aug-Trips **reduces the search space by 40%** and is nearly as **accurate**.
Better on longer (> 7 words) sentences, **worse** on idioms which are rare in training data.
Bikel (purely statistical) does best at bracketing since it does not need to be semantically coherent.

Interpretation Accuracy: Correct Semantic Features

75 Random Sents.	TRIPS	Aug. TRIPS
Good	**26**	21
54 TRIPS Errors		
Good	0	**8**

Trained on ~**400** sentences of simulated video projector sales.
A "Good" parse requires a perfect tree, correct word senses and thematic roles.
Aug-TRIPS is not as accurate overall, but can correct some TRIPS errors.

Conclusions

*Statistics can **speed up** a real-time dialogue parser based on unification without losing too much **accuracy**, even gaining some on longer sentences.
*The augmented parser can still produce **correct semantic analyses** (word senses/roles).
*Statistical augmentation helps even **without a large training corpus**.

Selected References

James F. Allen, Bradford W. Miller, Eric K. Ringger, and Teresa Sikorski. 1996. A robust system for natural spoken dialogue. (ACL '96).
Dan Bikel. 2002. Design of a multi-lingual, parallel-processing statistical parsing engine. (HLT '02).

Stefan Riezler, Detlef Prescher, Jonas Kuhn, and Mark Johnson. 2000. Lexicalized stochastic modeling of constraint-based grammars using log-linear measures and EM training. (ACL '00).

Stefan Riezler, Tracy H. King, Richard Crouch, and John T. Maxwell. 2002. Parsing the wall street journal using a lexical-functional grammar and discriminative estimation. (ACL '02).

Mary Swift, James Allen, and Daniel Gildea. 2004. Skeletons in the parser: Using a shallow parser to improve deep parsing. (COLING-04).

Kristina Toutanova, Christopher D. Manning, Dan Flickinger, and Stephan Oepen. 2005. Stochastic HPSG parse disambiguation using the Redwoods corpus. Journal of Logic and Computation.

Yoshimasa Tsuruoka, Yusuke Miyao, and Jun'ichi Tsujii. 2004. Towards efficient probabilistic HPSG parsing: Integrating semantic and syntactic preference to guide the parsing. In Proceedings of IJCNLP-04 Workshop.