

Site report: Cambridge DELPH-IN activities

Ann Copestake

Natural Language and Information Processing Group
Computer Laboratory
University of Cambridge

July 2010

Forthcoming attractions

- Underspecified quantification: Aurélie Herbelot (Tuesday)
Ducks lay eggs
The reporters asked questions
All, most or some?
- DMRS: Ann Copestake (Tuesday)
 - Overview/introduction to DMRS
 - DMRS packing and comparison (Vaughan Eveleigh, MPhil)
 - Induction of relationships between DMRS (Andy MacKinlay, visiting Cambridge from Melbourne)

Other Cambridge activities

1. SciBorg — now finished. Good results with ERG on Chemistry papers after preprocessing with OSCAR. Difficulties exploiting RMRS (part of motivation for DMRS).
2. Monte Carlo Semantics (Richard Bergmair, PhD).
3. Syntactically irregular MWEs (LI Xiaohua, MPhil): automatic selection of POS tags for MWE extraction, tested on ERG lexicon.
4. Anaphora resolution (Ann Copestake): initial experiments, proposal for a discourse extension for *MRS.
5. Lexicalised compositionality (Ann Copestake and Aurélie Herbelot): using distributional concepts in place of model-theoretic predicates (very preliminary).

Parse selection via combined syntactic and distributional techniques

- Aurélie Herbelot: JHU summer workshop, 2009
- Clark and Curran CCG parser, but should be relevant to DELPH-IN grammars
- Experiments with coordination
- Report is available via
<http://www.cl.cam.ac.uk/~sc609/pubs.html>

Coordination problems (CCG)

- *Food and Drug Administration spokesman Jeff Nesbitt said the agency has turned over evidence in a criminal investigation concerning Vitarine Pharmaceutical Inc. to the US attorney's office in Baltimore.*
- *I just don't feel that the company can really stand or would want a prolonged walkout.*

ERG (top parse, web demo June 29th, 2010)

- *Food and Drug Administration spokesman Jeff Nesbitt slept.*
- *I just don't feel that the company can really stand or would want a prolonged walkout.*

ERG (top parse, web demo June 29th, 2010)

- *Food and Drug Administration spokesman Jeff Nesbitt slept.*
- *I just don't feel that the company can really stand or would want a prolonged walkout.*

ERG (top parse, web demo June 29th, 2010)

- *Food and Drug Administration spokesman Jeff Nesbitt slept.*
- *I just don't feel that the company can really stand or would want a prolonged walkout.*

JHU Experimental setup

- 1000 best parses for each sentence
- Baseline calculated on top parse
- Pick the correct coordination(s)
- Development corpus:
 - 180 short Wikipedia sentences (under 30 words) containing coordination, split into 100 sentences for training and 80 for testing
- Evaluation corpus:
 - Sections 2 to 21 of the Wall Street Journal in CCG output (CCGbank, Hockenmaier, 2003) for training
 - 300 sentences of Wikipedia data, annotated during the workshop for testing

Features for Naive Bayes classifier

- parser ranking
- distance of the second coordinate to the conjunction
- distance between coordinates
- similarity between coordinates (WordNet or distributional similarity)
- N-gram similarity (based on parts of speech)

Additional binary features for specific constructions:

- first coordinate likely object of preposition?
- second coordinate likely subject of verb?

Results

Table: Results on Wikipedia300 coordinations

	Precision	Recall	F-score	Gain
Baseline	77.4%	66.3%	71.4%	-
WordNet similarity	79.7%	69.9%	74.5%	3.1
Distributional similarity	79.2%	69.4%	74.0%	2.6
plus PP and subject features	80.3%	70.3%	75.0%	3.5

Conclusion on JHU experiment

- 3.5% improvement on Wikipedia test set
- Distributional similarity system trained on CCG-parsed Wikipedia data performs roughly at the level of WordNet
- Haven't evaluated on entire parse
- Considerable scope for further experiments

Possible future work in Cambridge

- CCG-DMRS and parser comparison
- DMRS plus distributional techniques
- Incremental processing and parse selection
- Incremental generation?

Generally, emphasis on combining compositional semantics and distributional techniques and on psycholinguistically plausible models of semantics.