Learning and Predicting Verb Argument Binding

Corpora
I. OntoNotes-4.0 (CoNLL 2011 subset)
   - 10 Mbyte of text
   - multi-layer annotation
   - gold-standard annotation of i.a. coreference
   - 200 pairs
II. GigaWord 5th Ed.
   - 26 Gbyte of Text
   - automatic annotation (MATE parser)

Verb Pairs
I. Tremper & Frank (2013)
   - verbs in semantic relation: entailment, presupposition, antonymy, temporal inclusion
   - 200 pairs
II. Chambers & Jurafsky (2009)
   - ‘narrative chains’
   - 8,069 pairs

Motivation
- coreferring arguments in pairs of co-occurring verbs can be arbitrarily linked: Mary, visited the new art gallery with Susan. She loves art.
- meaning of the involved verbs can make one argument binding more probable/obligatory: Ed, shouted at Tim, because he, crashed the car. Ed, shouted at Tim, because he, was angry. (Rahman & Ng 2012)
- analysis of semantic properties of verbs reveal patterns
  - certain verbs can stand in a particular (semantic) relation
  - how can we squeeze profit from these insights?

Goals
- are there predictable argument binding (coreference) patterns?
- what are the possibly involved factors?
- does big data help finding reliable patterns?
- can we reliably predict such patterns on type/token-level?

Feature Extraction
I. pattern-based (association measures):
   - corefential probability
   - in relation to all coref. patterns of verb x:
     \[ C(\text{coref}(\text{arg}_x, \text{arg}_y)) \]
     \[ C(\text{coref}(\text{arg}_x, \text{arg}_y)) \]
   - in relation to all patterns:
     \[ C(\text{coref}(\text{arg}_x, \text{arg}_y)) \]
     \[ C(\text{coref}(\text{arg}_x, \text{arg}_y)) \]
II. argument-based:
   - e.g. same word, string match, proper name
III. verb-based:
   - e.g. tempus, voice, verb distance

Classification (WEKA J48)
number of argument patterns
- training (GigaWord)
  - WPB (1.3%): 43,726
  - XIN (25%): 838,751
  - AFP (68.6%): 2,303,020
- testing
  - OntoNotes: 5,205

Results

Comparison of performances:
- coref
  - SIEVE presented
  - Precision 0.79
  - Recall 0.58
- non-coref
  - SIEVE presented
  - Precision 0.67
  - Recall 0.85

Problems
- entire automatic annotation of GigaWord
- correct treatment of verbs containing prepositions/adverbs
- not realized arguments (evaluation):
  - not recognized by SIEVE yet
  - recognized by presented system
- gold-standard coreference information of OntoNotes: displaced word no.

Conclusions
- performance decreases on larger training sets
- insufficient ability of features to generalize for coreference, though precise prediction:
  - hence better performance on rare patterns (i.e. inObj, subj) than on frequent (i.e. subj, obj)
  - thus overfitting on non-coreferent patterns, especially for cases
    1. in which one argument is not realized
    2. in which tokens of both arguments are identical
- data is inconsistent/widespread: big data reduces determination of pattern strength and requires more robustness
- argument binding patterns can help improve coreference resolution systems (important features: include proper names, identical tokens and realization)
- to be examined:
  - suspicion of different performances of certain features (e.g. ‘ProperName’) at different argument patterns
  - precision recommends an implementation in elaborated coreference resolution systems
  - prior probability (resp. association) of coreferent argument bindings

References

www.cl.uni-heidelberg.de