Deriving Multi-Headed Planar Dependency Parses from Link Grammar Parses

Juneki Hong
Carnegie Mellon University

Jason Eisner
Johns Hopkins University

Summary

- Multi-headed dependency corpora would allow for the development of richer syntactic formalisms.
- Link Grammar can produce projective multi-headed corpora, but Link Grammar parses are undirected.
- We want to recover this “missing information” by consistently directionally Link Grammar parses, subject to constraints such as acyclicity and reachability.
- Starting with a corpus of LG parses, we utilize ILP to find a minimum set of directionality assignments subject to these constraints.
- The resulting parses differ in style from CoNLL-style parses of the same sentences.

Multi-Headed Dependency Parsing

Relaxing single-headed constraints common in dependency parsing would allow for constructions such as Control, Relativization, and Conjunction.

ILP Link Orientation Variables

For each sentence, for each edge $i, j$, where $i \leq j$:

- Orientation of each link can be represented by variables that can either be oriented left or oriented right:
  \[ x_{ij} + x_{ji} = 1 \]
- ILP Constraints:
  \[ x_{ij} \geq 0 \]
  \[ x_{ij} + x_{ji} = 1 \]

Consistency of Directionalized Links

Links with same label type are encouraged to be oriented in the same way. We introduce variables to represent whether links with label $L$ are allowed to go left or right:

- We introduce slack variables $s_{ij}$ to allow some links to go in disallowed directions with a penalty:
  \[ x_{ij} \leq s_{ij} + L \]
  \[ s_{ij} \leq 0 \]

Data Sets

Data Sets taken from:
- CoNLL 2007 Shared Task (English)
- ACL 2013 Shared Task of Machine Translation (Russian)

1 Experiments and Results

Stability of Results

To see whether the recovered direction mapping might be unstable and sensitive to the input corpus, we compared results of increasing runs of sentences.

References