Corpus Navigator Status

Jonathan Clark

Thursday, January 10, 2008

the link

1 Classes

1. NLP Lab: I would like to code a Chart Parser with Feature Constraints that reads in the same rule format as the Tomita Parser. This could be used in G&L plus I want to know more about feature constraints.
2. I would like to audit the MT Seminar since it might give me perspective when writing papers (1 hour per week, no outside obligations)

2 Available Demos

1. PDF of Navigator Oracle Output
2. PDF of Ruleless Feature Detection Output

3 Recently Accomplished

1. Integrated ATAVI Support
2. Syntax projection code (probably lacks several features of Vamshi's method)
3. Navigator "Oracle" that uses syntax projection and rules to do a greedy search for novel rules
4. Implemented Ruleless Feature detection
5. Created Latex output for ruleless feature detection

4 Current Tasks

1. Clean elicitation corpus
2. Debug ruleless feature detection
3. Discover "who" each feature is marked on
4. Implement the Navigator Heuristic Search
5. Pull components of the Corpus Navigator together

5 Planned Papers

1. ACL Short paper (Feature Clustering) due Mar 14
2. EMNLP Full paper (Corpus Navigation) due Mar 30