

Reading #4

- Paper: (Collins and Duffy, 2001)
- Due date: Next Tuesday (3/2)

Questions

According to the paper, suppose you want to build a reranking for parsing using SVM:

- Q1: What does training data look like? What does the test data look like?
- Q2: how do you represent the reranker? In other words, if you store the reranker in a model file, what information should be included in the model file?
- Q3: What happens at the test time? That is, what formula(s) one needs to calculate in order to determine the correct ranking of the candidate parse trees?

- Q4: Conceptually, a parse tree is represented as a feature vector. What are the features? What are the feature values? How many features are there?
- Q5: In practice, is it necessary to represent a parse tree as a feature vector? Why or why not?