

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, May 1, 2001
Sequoia Hall Rm. 200
(Cookies at 3:45 in 1st Floor Lounge)

Christopher Manning
Computer Science and Linguistics
Stanford University

Probabilistic Head-driven Parsing

A central problem in natural language processing is resolving sentence structure ambiguities, and this has been approached in recent years by learning probabilistic models from hand-parsed sentences. This paper discusses a particular approach of this sort, which builds on two central linguistic notions: a phrasal head is a locus of constraining information, and that around the head is an "island of certainty" – verb arguments, adjectival modifiers, and the like. A parsing model that works outwards from heads has several advantages: the model is immune to some fairly arbitrary aspects of tree geometry, it can easily capture important length and distance-based effects, and the model can more cleanly capture the intuitions of dependency-based approaches to grammar.