Summary

- Parsing can be viewed as a search problem
- Two common architectural approaches for this search are
  - Top-down: Starting with the root S and growing trees down to the input words
  - Bottom-up: Starting with the words and growing trees up toward the root S.

- Ambiguity combined with the repeated parsing of sub-trees are a difficulty for parsing algorithms. Those algorithms use simple backtracking mechanisms.

Summary

- A sentence is structurally ambiguos if the grammar assigns it more than a possible parse.
- Common kinds of structural ambiguity include:

  - PP-attachment

  - Coordination ambiguity
- The dynamic programming parsing algorithms use a table of partial parsers to efficiently parse ambiguous sentences. The CKY, Earley and Chart-Parsing algorithms all use dynamic-programming to solve repeated parsing of subtrees problem.

Summary

- A feature structure is a set of features-value pairs, where features are atomic symbols and values are either atomic symbols or feature structure.
- Feature structures can be represented
  - Attribute-value matrices
  - Directed acyclic graphs, where features are labeled edges and values are nodes.
- Unification is the operation for
  - Combining information. Merging the information content of two features structures.
  - Comparing information. Rejecting the merger of incompatible features.

Summary

- A phrase-structure rule can be augmented with feature structures and with constraints expressing relations among the features structures of the constituents of the rule.
- Features structures can be typed. Typed feature structures place constraints on which type of values a given feature can take. They can be organized into a type hierarchy.

Summary

- Partial parsing and chunking are methods for identitying shallow syntactic constituents in a text.

- High accuracy partial parsing can be achieved either trough rule-based or machine learning-based methods.

Summary

- Probabilistic grammars assign a probability to a sentence or string of words. Usually they capture more general syntactic information than the N-gram grammars.

- A probabilistic context-free grammar (PCFG) is a context-free grammar in which every rule is annotated with the probability of choosing that rule. Each PCFG rule is treated as if it were conditionally independent; thus the probability of a sentence is computed by multiplying the probabilities of each rule in the parse of the sentence.

Summary

- There are probabilistic versions of parsers like the CYK and the Earley algorithm.
- Probabilistic lexicalized context-free grammars are another solution where each rule in the PCFG is augmented with a lexical head.The probability of a rule can then be conditioned on the lexical head or nearby heads.

- Parsers are evaluated considering three metrics: labeled recall, labeled precision and cross-brackets.