

Computational Linguistics I, Winter 2006. Marcus Kracht

To be submitted: Friday, March 17, 2006.

These exercises will lead up to the construction of a general purpose chart parser. The parser takes two inputs: a grammar  $G$  in 2-standard form and a list  $l$ , and returns a set of nonterminals for  $l$ . I recommend the following: a rule is a pair  $(i, p)$ , where  $i$  is a string and  $p$  a pair of strings. A grammar is a hashtable  $H$  from pairs of strings to strings.  $H$  associates with a pair  $(a, b)$  the string  $i$  just in case the grammar contains the rule  $i \rightarrow ab$ . The chart consists in an array of arrays of sets of strings.

- [A 8.1] Create the data types necessary for the parser: sets of strings, and hashtables.
- [A 8.2] Define a function that sets up an array of arrays of sets of strings from the input list  $l$ .
- [A 8.3] (This is the chart parsing phase.) Define a function that fills the chart stepwise.
- [A 8.4] Finally, write a main loop that takes the grammar and the list as input and returns "true" if the grammar accepts.

*Remarks.* Nonterminals and terminals each are of type string. It is not necessary to distinguish terminals from nonterminals, so we ignore the distinction. There is no need to declare the set of terminals or nonterminals; they are implicitly defined by the rule system. It is however necessary to say what the start symbol is. You must therefore make that explicit. Moreover, please do not forget to supply me with evidence that you have tested the program.