582206 Models of Computation (Autumn 2007)
Exercise 9 (13 October–16 November)

1. Convert the following context-free grammar into Chomsky normal form:

$$\begin{aligned} S &\to ASB \mid B \\ C &\to \mathrm{c}C \mid \varepsilon \\ A &\to C\mathrm{a}AC \mid \mathrm{a} \\ B &\to \mathrm{abc} \mid A \quad . \end{aligned}$$

2. Call a variable A of a context-free grammar

   - *unreachable* if from the start symbol of the grammar it is impossible to derive any string containing A, and
   - *unproductive* if from A it is impossible to derive any string consisting entirely of terminals.

   As an example, in the grammar

$$\begin{aligned} S &\to A \mid BC \\ A &\to \mathrm{a}A \mid \varepsilon \\ B &\to \mathrm{b}B \mid \varepsilon \\ D &\to \mathrm{ab} \\ E &\to C \end{aligned}$$

   the variable C is unproductive, D unreachable and E both unproductive and unreachable.

   (a) Give an algorithm for finding all unreachable variables in a grammar.

   (b) Give an algorithm for finding all unproductive variables in a grammar.

   (c) Suppose we try to remove all "useless" variables by first removing all unreachable variables and rules involving them, and then, from what remains, further removing all the unproductive variables and rules involving them. Does this lead to the desired result? Why?

3. Use the CYK algorithm to decide whether the string baaab in generated by the grammar

$$\begin{aligned} S &\to AB \mid BC \\ A &\to BA \mid \mathrm{a} \\ B &\to CC \mid \mathrm{b} \\ C &\to AB \mid \mathrm{a} \quad . \end{aligned}$$

4. [Sipser Problem 2.25] For a language A over the alphabet Σ, define its set of suffixes

$$\textsc{Suffix}(A) = \{\, v \in \Sigma^* \mid uv \in A \text{ for some } u \in \Sigma^* \,\} .$$

   Prove that if A in context free, then so is $\textsc{Suffix}(A)$.