582206 Models of Computation (Autumn 2009)
Exercise 8 (10–13 November)

1. Consider the grammar

$$\begin{array}{rcl} \langle\texttt{stmt}\rangle & \to & \langle\texttt{if-then-else}\rangle \mid \langle\texttt{if-then}\rangle \mid \texttt{p} \\ \langle\texttt{if-then-else}\rangle & \to & \textbf{if}\ \texttt{b}\ \textbf{then}\ \langle\texttt{stmt}\rangle\ \textbf{else}\ \langle\texttt{stmt}\rangle \\ \langle\texttt{if-then}\rangle & \to & \textbf{if}\ \texttt{b}\ \textbf{then}\ \langle\texttt{stmt}\rangle \quad . \end{array}$$

   To clarify, the set of terminal is $\{\, \textbf{if}, \textbf{then}, \textbf{else}, \texttt{b}, \texttt{p} \,\}$; for readability, some space has been added between the terminals.
   Show that the grammar is ambiguous. Desing an unambiguous grammar for the same language.

2. Convert the following grammar into Chomsky normal form:

$$\begin{array}{l} S \to BSA \mid A \\ A \to \mathrm{a}A \mid \varepsilon \\ B \to B\mathrm{ba} \mid \varepsilon \quad . \end{array}$$

3. 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

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

   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) Try to remove all "useless" variables in the above example 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? What if the opposite order is used: first remove unproductive variables and then unreachable variables?

   Hint for (a) and (b): See algorithm on page 160 in the lecture notes that computes the set NULL.

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 is context free, then so is $\textsc{Suffix}(A)$.
   Hint: Chomsky normal form.