T-61.231 Principles of Pattern Recognition

Exercise 7: 12.11.2001

1. Form a grammar to describe a train, which has both passenger coaches and luggage vans. First there is a locomotive, then at least one luggage van, then at least one passenger coach. The passenger coach nearest to the dining car is 1. class coach. It is also possible for the dining car to be the only form of passenger car on the train, it is interpreted as a special case of a passenger car. Form a BNF notation, a parse tree and a syntax diagram with your grammar to the train, whose car order is the following:
$\{locomotive, luggage\ van, dining\ car, 1.\ class\ coach, passenger\ coach, passenger\ coach\}$.

2. Does the grammar in question 1 accept the following train:
$\{locomotive, luggage\ van, luggage\ van, dining\ car, passenger\ coach, 1.\ class\ coach\}$?

3. Let us try to produce squares by using a grammar whose terminal symbols are line of a unit length to right o, line of a unit length downwards a, negation $\neg$ and concatenation $+$. For example a unit square is $o + a + \neg o + \neg a$. Nonterminals are $Square$, $Side1$, $Side2$, $Side3$, $Side4$ and production rules are

$$Square \rightarrow Side1 + Side2 + Side3 + Side4$$
$$Side1 \rightarrow o \mid Side1 + o \qquad Side2 \rightarrow a \mid Side2 + a$$
$$Side3 \rightarrow \neg o \mid Side3 + \neg o \quad Side4 \rightarrow \neg a \mid Side4 + \neg a$$

   a) Is the grammar able to produce squares?

   b) Show that the grammar produces also something else than just squares.

   c) How should the grammar be changed in order to make it produce only squares?

4. The electrical activity of human heart muscle simulation can be monitored by sensors whose output produces an electrocardiogram (ECG). An example of an regular pulse is *prbtbbbprbt*. The smaller of the hearts muscles, the atria, are stimulated with the first p pulse; the subsequent stimulation of the larger muscles, the ventricles, produces a larger r pulse followed by a t pulse as the ventricles re-polarize after the pumping action.

   The regularity of this normal ECG waveform can be described by a regular grammar. The waveform primitives are p, r and t for the pulses and b for the quiescent times. A normal ECG consists of concatenations of substrings *prbtb*, *prbtbb* and *prbtbbb*, in which the variable number of b's allows for variations in basically health heart rates. A regular grammar generating concatenations f these three substrings as a language is $G = (\{S, A, B, C, D, E, H\}, \{p, r, t, b\}, P, S\}$ with productions

$$S \rightarrow pA \quad A \rightarrow rB \quad B \rightarrow bC \quad C \rightarrow tD \quad D \rightarrow b \quad D \rightarrow bE$$
$$E \rightarrow b \quad E \rightarrow bH \quad E \rightarrow pA \quad H \rightarrow b \quad H \rightarrow bS \quad H \rightarrow pA$$

   a) Form the deterministic finite automaton corresponding to the language $L(G)$

   b) Does the symbol sequence *prbtbprbtbbb* ... cause an alarm?

   (Gonzales, Thomason: Syntatic PR, AnIntroduction (Addison Wesley 1998), p. 118)