T-79.1001 **Kevät 2006**
Introduction to Theoretical Computer Science (T)
Session 8
Answers to demonstration exercises

4. **Problem**: Prove that the class of context-free languages is closed under unions, concatenations, and the Kleene star operation, i.e. if the languages $L_1, L_2 \subseteq \Sigma^*$ are context-free, then so are the languages $L_1 \cup L_2$, L_1L_2 and L_1^*.

   Solution: Let L_1 and L_2 be context-free languages that are defined by grammars $G_1 = (V_1, \Sigma_1, R_1, S_1)$ and $G_2 = (V_2, \Sigma_2, R_2, S_2)$. In addition we require that $(V_1 - \Sigma_1) \cap (V_2 - \Sigma_2) = \emptyset$. That is, the grammars may not have any common nonterminals. Since the nonterminals may be renamed if necessary, this is not an essential limitation.

   Union: Let S be a new nonterminal and $G = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R_1 \cup R_2 \cup \{S \to S_1 \mid S_2\}, S\}$. Now $L(G) = L(G_1) \cup L(G_2) = L_1 \cup L_2$. This holds, since the initial symbol S may derive only S_1 or S_2, and they in turn may derive only strings that belong to the respective languages. (If the sets of nonterminals were not disjoint, this would not hold).

   Concatenation: The language L_1L_2 is defined by the following grammar:$G = (V_1 \cup V_2 \cup \{S\}, \Sigma_1 \cup \Sigma_2, R_1 \cup R_2 \cup \{S \to S_1S_2\}, S\}$

   Kleene star: The language L_1^* is defined by the following grammar: $G = (V_1 \cup \{S\}, \Sigma_1, R_1 \cup \{S \to \epsilon | SS_1\}, S\}$

5. **Problem**: Prove that the class of context-free languages is not closed under intersections and complements. (*Hint:* Represent the language $\{a^kb^kc^k \mid k \geq 0\}$ as the intersection of two context-free languages.)

   Solution: Let $L = \{a^kb^kc^k \mid k \geq 0\}$. This language has been proven to be not context-free. We can prove that context-free languages are not closed under intersection by finding two context-free languages L_1 and L_2 such that $L = L_1 \cap L_2$. Languages $L_1 = \{a^ib^kc^k \mid i, k \geq 0\}$ and $L_2 = \{a^kb^kc^i \mid i, k \geq 0\}$ fulfill this condition.

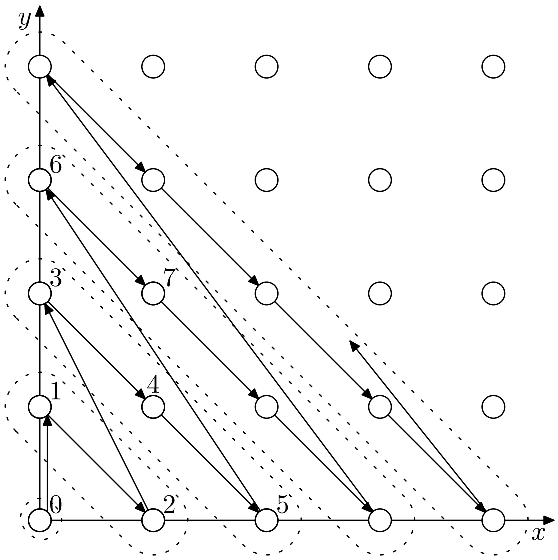
   A direct corollary is that the class of context-free languages cannot be closed under complementation, either, since they are closed under union and $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$.

   Finally, we prove that L_1 and L_2 are context-free by presenting context-free grammars that generate them. The language L_1 is generated by $G_1 = (\{S, A, B, a, b, c\}, \{a, b, c\}, P_1, S)$, where $P_1 = \{S \to AB, A \to aA \mid \varepsilon, B \to bBc \mid \varepsilon\}$. Similarily, L_2 is generated by $G_2 = (\{S, A, B, a, b, c\}, \{a, b, c\}, P_2, S)$, $P_2 = \{S \to AB, A \to aAb \mid \varepsilon, B \to cB \mid \varepsilon\}$.

6. **Problem**: Prove that the Cartesian product $\mathbb{N} \times \mathbb{N}$ is countably infinite. (*Hint:* Think of the pairs $(m, n) \in \mathbb{N} \times \mathbb{N}$ as embedded in the Euclidean (x, y) plane $\mathbb{R}^2$. Enumerate the pairs by diagonals parallel to the line $y = -x$.) Conclude from this result and the result of Problem 3 that also the set $\mathbb{Q}$ of rational numbers is countably infinite.

   Solution: A set S is countably infinite, if there exists a bijective mapping $f : \mathbb{N} \to S$. By intuition, all members of the set S can be assigned a unambiguous running number.

   The members $(x, y) \in \mathbb{N} \times \mathbb{N}$ of the set $\mathbb{N} \times \mathbb{N}$ can be assigned a number as shown in the following figure.



The idea is to arrange all pairs of numbers on diagonals parallel to the line $y = -x$ and enumerate the lines by member one at a time, starting from the shortest one. Here the enumeration can not be done parallel to the x-axis; when doing this all indices would be used to enumerate only the y-axis and no pair $(x, y), y > 0$ would ever be reached.

The enumerating scheme abowe can be defined as follows:

$$f(x, y) = x + \sum_{k=1}^{x+y} k = x + \frac{(x+y)(x+y+1)}{2}$$

For an example, $f(3, 1) = 13$, that is, the running number of pair $(3, 1)$ is 13. The function $f(x, y)$ is a bijection; for every running number there exists a unambiguous pair of numbers. Calculating a coordinate from a given index is relatively difficult, and is discussed in the appendix at the end of these solutions.

The set of positive rational numbers $\mathbb{Q}^+$ can be presented as a pair of numbers $\mathbb{N} \times \mathbb{N}$ by $(x, y) \equiv \frac{x}{y}, y \neq 0$. This is a proper subset of the countably infinite set $\mathbb{N} \times \mathbb{N}$. By Problem 3, $\mathbb{Q}^+$ is either finite or countably infinite. If $\mathbb{Q}^+$ was finite, there should exists some rational number $\frac{x}{y}, x \in \mathbb{N}, y \in \mathbb{N}, y \neq 0$, that would have the greatest running number $n < \infty$ (in the enumeration of $\mathbb{Q}$). This cannot be, because using the figure above one could always find a rational number that would have a running numberu $n' > n$. Hence, we have contradiction with the assumption that $\mathbb{Q}^+$ is finite. Therefore $\mathbb{Q}^+$ is countably infinite. By the same argument, the set $\mathbb{Q}^-$:

$$\mathbb{Q}^- = \{(-x, y) \mid (x, y) \in \mathbb{Q}^+\}$$

is countably infinite. Thus, the set $\mathbb{Q} = \mathbb{Q}^+ \cup \mathbb{Q}^-$ is the union of two countably infinite sets, and it too is countably infinite.