Automata & Formal Languages
Michaelmas Term 2022
Part II of the Mathematical Tripos
University of Cambridge
Prof. Dr. B. Löwe

Example Sheet #1

(1) Let $R = (\Omega, P)$ be a rewrite system and $\alpha, \beta, \gamma \in \Omega^*$. Show that if $\alpha \xrightarrow{R} \beta$, then $\alpha\gamma \xrightarrow{R} \beta\gamma$. Show that the converse does not hold in general. What properties of P could guarantee that the converse holds?

(2) Give an example for non-uniqueness of derivations: provide a grammar G, a word $w \in \mathcal{L}(G)$, and two different derivations of w from S in G.

(3) Give an example of two non-isomorphic, but equivalent grammars. Justify your claim.

(4) Let $\Sigma = \{a\}$, $V = \{S, A\}$, $P_8 := \{S \to aA, S \to a, A \to aS\}$, and $G_8 := (\Sigma, V, P_8, S)$. Show in detail that $\mathcal{L}(G_8) = \{a^{2n+1}\,;\, n \in \mathbb{N}\}$ and deduce that this language is regular.

(5) Let $\Sigma = \{a, b, c\}$. Show each of the following claims by producing an appropriate grammar that produces the given language. Explain why your grammar generates this language.

(a) The language consisting of words of the form $(abc)^n$ (for $n > 0$) is type 3.
(b) The language consisting of words of the form $a^n bc^n$ (for $n \in \mathbb{N}$) is type 2.
(c) The language consisting of words of the form $a^n b^n c^n$ (for $n > 0$) is type 1.
(d) The language consisting of words of the form $a^n b^m c^{n+m}$ (for $n, m > 0$) is type 0.

Are any of them of even higher type than listed?

(6) Let $G = (\Sigma, V, P, S)$ be any grammar. As in the lectures, a production rule $\alpha \to \beta$ is called *variable-based* if $\alpha \in V^*$. Suppose that $\alpha \to \beta$ is a noncontracting variable-based rule, say with $\alpha = A_1 ... A_n$ and $\beta = B_1 ... B_m$ for $A_i \in V$, $B_i \in \Omega$, and $n \leq m$. Let $X_1, ..., X_n$ be n new variables that do not occur in V and consider the following list of $2n$ rules:

$$
\begin{array}{lcl}
A_1 A_2 A_3 \;\ldots\; A_{n-2} A_{n-1} A_n & \to & X_1 A_2 A_3 \;\ldots\; A_{n-2} A_{n-1} A_n \\
X_1 A_2 A_3 \;\ldots\; A_{n-2} A_{n-1} A_n & \to & X_1 X_2 A_3 \;\ldots\; A_{n-2} A_{n-1} A_n \\
X_1 X_2 A_3 \;\ldots\; A_{n-2} A_{n-1} A_n & \to & X_1 X_2 X_3 \;\ldots\; A_{n-2} A_{n-1} A_n \\
 & \vdots & \\
X_1 X_2 X_3 \;\ldots\; X_{n-2} A_{n-1} A_n & \to & X_1 X_2 X_3 \;\ldots\; X_{n-2} X_{n-1} A_n \\
X_1 X_2 X_3 \;\ldots\; X_{n-2} X_{n-1} A_n & \to & X_1 X_2 X_3 \;\ldots\; X_{n-2} X_{n-1} X_n B_{n+1} \;\ldots\; B_m \\
X_1 X_2 X_3 \;\ldots\; X_{n-2} X_{n-1} X_n B_{n+1} \;\ldots\; B_m & \to & B_1 X_2 X_3 \;\ldots\; X_{n-2} X_{n-1} X_n B_{n+1} \;\ldots\; B_m \\
B_1 X_2 X_3 \;\ldots\; X_{n-2} X_{n-1} X_n B_{n+1} \;\ldots\; B_m & \to & B_1 B_2 X_3 \;\ldots\; X_{n-2} X_{n-1} X_n B_{n+1} \;\ldots\; B_m \\
 & \vdots & \\
B_1 B_2 B_3 \;\ldots\; B_{n-2} X_{n-1} X_n B_{n+1} \;\ldots\; B_m & \to & B_1 B_2 B_3 \;\ldots\; B_{n-2} B_{n-1} X_n B_{n+1} \;\ldots\; B_m \\
B_1 B_2 B_3 \;\ldots\; B_{n-2} B_{n-1} X_n B_{n+1} \;\ldots\; B_m & \to & B_1 B_2 B_3 \;\ldots\; B_{n-2} B_{n-1} B_n B_{n+1} \;\ldots\; B_m
\end{array}
$$

Show that each of these rules is context-sensitive and that replacing $\alpha \to \beta$ in P by this collection of $2n$ rules does not change the language produced by G. Use this to prove that a language is noncontracting if and only if it is context-sensitive.

(7) Give an example of a class of languages that is closed under unions and intersections, but not under complementation.

(8) In the lectures, we proved that the *concatenation grammar* of two grammars G and G' produces the concatenation of the two languages produced by G and G' *under the assumption that they are variable-based and do not share any variables.* Show that these assumptions are necessary by giving grammars G and G' such that the language produced by the concatenation grammar is not $\mathcal{L}(G)\mathcal{L}(G')$.

(9) Construct deterministic automata by drawing transition diagrams which accept the following languages. Explain your answers.

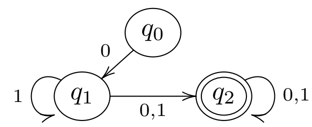
(a) $\{w \in \{0,1\}^* \,;\, |w| > 2\}$;

(b) $\{w \in \{0,1\}^* \,;\, w$ is a nonempty alternating sequence of 0s and 1s$\}$;

(c) $\{w \in \{0,1\}^* \,;\, w$ is a multiple of 3 when interpreted in binary$\}$;

(d) $\{w \in \{0,1\}^* \,;\, w$ contains 01010 as a substring$\}$.

(10) Consider the following nondeterministic automaton over the alphabet $\Sigma = \{0,1\}$:



Convert it to a deterministic automaton with $2^3 = 8$ states using the power set construction. Can you simplify the deterministic automaton without changing the accepted language?

(11) For each of the following languages $L \subseteq \{0,1\}^*$, determine whether or not they are regular. Justify your answers.

(a) $\{0^n 1^{2n} \,;\, n \geq 1\}$;

(b) $\{ww \,;\, \varepsilon \neq w \in \{0,1\}^*\}$;

(c) $\{w1w \,;\, w \in \{0\}^*\}$;

(d) $\{v1w \,;\, v, w \in \{0\}^*\}$;

(e) $\{0^n 1^m \,;\, n > m\}$;

(f) $\{0^n 1^m \,;\, n \neq m\}$;

(g) $\{0^n 1^m \,;\, n \geq m$ and $m \leq 1000\}$;

(h) $\{0^n 1^m \,;\, n \geq m$ and $m \geq 1000\}$;

(i) $\{1^p \,;\, p$ is a prime$\}$.

(12) Let $L \subseteq \{0^n 1^n \,;\, n \geq 1\}$. Show that L is regular if and only if L is finite.

(13) Suppose that $G = (\Sigma, V, P, S)$ is a regular grammar with $|V| = n$. Prove that if G produces a word of length at least 2^{n+1}, then it produces infinitely many words.

(14) In the lectures, we have seen two different constructions that prove that the class of regular languages is closed under unions: the union grammar and the product automaton construction. Take two grammars G and G' and form deterministic automata D and D', using the two mentioned constructions such that $\mathcal{L}(D) = \mathcal{L}(G) \cup \mathcal{L}(G') = \mathcal{L}(D')$. Compare the number of states that the automata D and D' have.