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

Example Sheet #2

(15) Work over the alphabet $\Sigma = \{0, 1\}$ and for $n \geq 1$ consider

$$L_n := \{w\,;\text{ there are } x, y \in \mathbb{W} \text{ such that } |y| = n - 1 \text{ and } w = x1y)\},$$

i.e., the set of words that have a 1 in the nth position counted from the end of the word. Show that:

(a) There is a nondeterministic automaton N with $n + 1$ states such that $\mathcal{L}(N) = L_n$ and

(b) if D is a deterministic automaton with fewer than 2^n states, then $\mathcal{L}(D) \neq L_n$.

(16) Let R, S, T be regular expressions. For each of the following statements, either prove that it is true, or find a specific counterexample.

(a) $\mathcal{L}(R(S + T)) = \mathcal{L}(RS + RT)$;

(b) $\mathcal{L}((R + S)T) = \mathcal{L}(RT + ST)$;

(c) $\mathcal{L}(R + ST) = \mathcal{L}(RS + RT)$;

(d) $\mathcal{L}(R^*R) = \mathcal{L}(R^+)$;

(e) $\mathcal{L}((R + SR^*)^*) = \mathcal{L}((R + S)^*)$;

(f) $\mathcal{L}((R + S)^+) = \mathcal{L}(R^+ + S^+)$

(17) Let L and M be languages over an alphabet Σ and consider the set equation $X = LX \cup M$. Prove the following statements:

(a) The language L^*M is a solution for this equation.

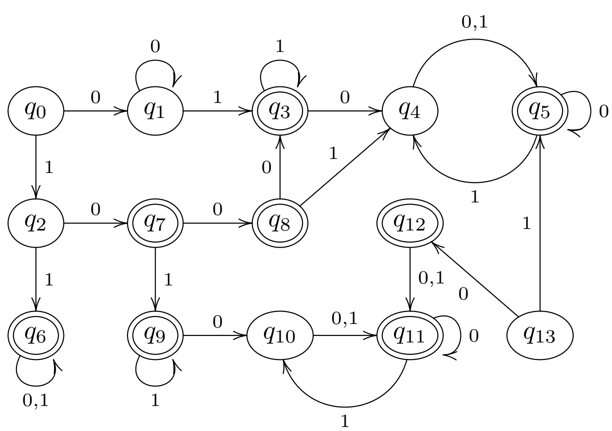
(b) If Z is a solution for this equation, then $L^*M \subseteq Z$.

(c) If $\varepsilon \notin L$, then L^*M is the only solution for this equation.

(18) Consider the regular grammar $G = (\Sigma, V, P, S)$ with $P = \{S \to aS, S \to a\}$. Prove that $\mathcal{L}(G) = \mathcal{L}(a^+)$.

(19) Consider the regular grammar $G = (\Sigma, V, P, S)$ with $P = \{S \to cS, S \to cA, A \to aA, A \to aB, B \to bB, B \to b\}$. Find a regular expression R such that $\mathcal{L}(R) = \mathcal{L}(G)$.

(20) Minimise the following automaton using the construction given in the lectures.



(21) Give the minimal deterministic automaton for the language L_n discussed in (15).

(22) A *tree graph* is an acyclic connected directed graph. A language $L \subseteq \mathbb{W}$ is called a *tree language* if it is closed under initial segments, i.e., if $w \in L$ and v is an initial segment of w, then $v \in L$. Show that every tree language L defines a tree graph by using L as the set of vertices and having $E := \{(w, wa)\,;\, w, wa \in L\}$ as the set of edges. Characterise the tree graphs that are obtained in this way from a tree language. Justify your claim.

(23) Let G be the context-free grammar given by $S \to ABS, S \to AB, A \to aA, A \to a, B \to bA$. For each of the words *aabaab*, *aaaaba*, *aabbaa*, *abaaba*, either draw the parse tree that shows that the word lies in $\mathcal{L}(G)$ or argue that it does not lie in $\mathcal{L}(G)$.

(24) Consider $\Sigma = \{0, 1, \varepsilon, (,), +, \varnothing, {}^+, {}^*\}$. Define a context-free grammar G over Σ such that $\mathcal{L}(G)$ is the set of regular expressions over the alphabet $\{0, 1\}$. Show that you cannot find a regular grammar with this property.

(25) Convert the context-free grammar given by the following production rules to Chomsky normal form, explaining the transformation steps in detail: $S \to aSbb, S \to T, T \to bTaa, T \to S$.

(26) Give a context-free grammar in Chomsky normal form for the languages $\{a^m b^{2m} c^k\,;\, m, k \geq 1\}$ and $\{a^m b^k a^m\,;\, m, k \geq 1\}$. Justify your answer.

(27) Let $G = (\Sigma, V, P, S)$ be a context-free grammar in Chomsky normal form. Let

$$P' := P \cup \{A \to \varepsilon\,;\, \text{there is some } a \in \Sigma \text{ such that } A \to a \in P)\}$$

and $G' := (\Sigma, V, P', S)$. Describe the language $\mathcal{L}(G')$ in words, giving an argument for your answer. What changes if G is not in Chomsky normal form?

(28) Suppose $\Sigma = \{a\}$ and that L is a language over Σ that satisfies the context free pumping lemma. Show that it satisfies the regular pumping lemma.

(29) For each of the following languages, decide whether it is context-free and provide an argument for your claim:

(a) $\{a^n b^m\,;\, n \neq m, n + m > 0\}$;
(b) $\{a^m b^n c^m d^n\,;\, m, n \geq 1\}$;
(c) $\{a^n b^m c^k d^\ell\,;\, 2n = 3m \text{ and } 5k = 7\ell\}$;
(d) $\{a^n b^m c^k d^\ell\,;\, 2n = 3k \text{ and } 5m = 7\ell\}$;
(e) $\{a^n b^m c^k d^\ell\,;\, 2n = 3k \text{ or } 5m = 7\ell\}$;
(f) $\{a^p\,;\, p \text{ is prime}\}$;
(g) $\{a^{2n}\,;\, n \geq 1\}$;
(h) $\{ww\,;\, w \in \{a, b\}^+\}$;
(i) $\{a, b\}^+ \backslash \{ww\,;\, w \in \{a, b\}^+\}$.

(30) Show that the class of context-free grammars is closed under the Kleene plus operation, i.e., if L is context-free, then so is L^+.