Automata & Formal Languages
Michaelmas Term 2024
Part II of the Mathematical Tripos
University of Cambridge
Prof. Dr. B. Löwe

Example Sheet #2

(14) If $v \in \mathbb{W}$, we call u a *subword* of v if there are $x, y \in \mathbb{W}$ such that $v = xuy$. In that case, we call the word xy the *result of removing the subword u from v*. For $e, w \in \mathbb{W}$, we say that e is an *excerpt* of w if it is the result of removing finitely many subwords from w, i.e., $w = x_0y_0x_1y_1 \ldots x_ny_nx_{n+1}$ and $e = x_0 \ldots x_{n+1}$ for some $x_0, \ldots, x_{n+1}, y_0, \ldots, y_n \in \mathbb{W}$; we say that it is a *proper excerpt* if $e \neq w$.

(i) Suppose that $L := \mathcal{L}(D)$ for a deterministic automaton D with $|Q| = n$ and that $w = xvy \in L$ for $x, v, y \in \mathbb{W}$ with $|v| \geq n$. Prove that there is a proper excerpt e of v such that $xey \in L$.

(ii) Show that the following language is context-free, not regular, but satisfies the regular pumping lemma with pumping number $n = 2$:

$$L = \{w\mathbf{2}\mathbf{0}^n\mathbf{1}^n\,;\, w \in \mathbb{W}, n > 0\} \cup \{\mathbf{0}, \mathbf{1}\}^+.$$

(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:

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

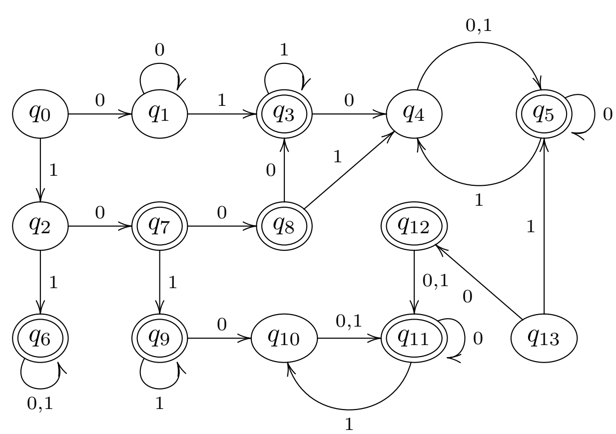
(ii) if D is a deterministic automaton with fewer than 2^n states, then $\mathcal{L}(D) \neq L_n$.

(16) Let $D = (\Sigma, Q, \delta, q_0, F)$ be a deterministic automaton and define $Q^* := Q \backslash F \cup \{F\}$ and the nondeterministic automaton $N = (\Sigma, Q^*, \Delta, F, \{q_0\})$ with

$$\begin{aligned} q' \in \Delta(q, a) &\text{ if and only if } \delta(q', a) = q, \\ q' \in \Delta(F, a) &\text{ if and only if } \delta(q', a) \in F, \\ F \in \Delta(q, a) &\text{ if and only if } \delta(q', a) = q \text{ for some } q' \in F, \text{ and} \\ F \in \Delta(F, a) &\text{ if and only if } \delta(q', a) \in F \text{ for some } q' \in F \end{aligned}$$

(where $q, q' \notin F$). Describe the relationship between $\mathcal{L}(D)$ and $\mathcal{L}(N)$ and prove your claim.

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



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

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

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

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

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

(20) Give regular expressions that describe the languages given by the grammars $G = (\Sigma, V, P, S)$

(i) where $P = \{S \to \mathbf{0}S, S \to \mathbf{0}\}$ and

(ii) where $P = \{S \to \mathbf{2}S, S \to \mathbf{2}A, A \to \mathbf{0}A, A \to \mathbf{0}B, B \to \mathbf{1}B, B \to \mathbf{1}\}$.

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

(22) Show that both the classes of regular and context-free grammars are closed under the Kleene plus operation, i.e., if L is regular or context-free, then so is L^+. Deduce that the classes of essentially regular and essentially context-free grammars are closed under the Kleene star operation, i.e., if L is regular or context-free, then so is L^*. Explain why this implies that every language associated with a regular expression is essentially regular.

(23) Let G be the context-free grammar given by $S \to ABS, S \to AB, A \to \mathbf{0}A, A \to \mathbf{0}, B \to \mathbf{1}A$. For each of the words $\mathbf{001001}$, $\mathbf{000010}$, $\mathbf{001100}$, $\mathbf{010010}$, 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) Convert the context-free grammar given by the following production rules to Chomsky normal form, explaining the transformation steps in detail: $S \to \mathbf{0}S\mathbf{11}, S \to T, T \to \mathbf{1}T\mathbf{00}, T \to S$.

(25) Give a context-free grammar in Chomsky normal form for the languages $\{\mathbf{0}^m\mathbf{1}^{2m}\mathbf{2}^k\,;\, m, k \geq 1\}$ and $\{\mathbf{0}^m\mathbf{1}^k\mathbf{0}^m\,;\, m, k \geq 1\}$. Justify your answers.

(26) 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?

(27) Suppose $\Sigma = \{\mathbf{0}\}$ and that L is a language over Σ that satisfies the context free pumping lemma. Show that it satisfies the regular pumping lemma.

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

(i) $\{\mathbf{0}^n\mathbf{1}^m\,;\, n \neq m, n + m > 0\}$;

(ii) $\{\mathbf{0}^m\mathbf{1}^n\mathbf{2}^m\mathbf{3}^n\,;\, m, n \geq 1\}$;

(iii) $\{\mathbf{0}^n\mathbf{1}^m\mathbf{2}^k\mathbf{3}^\ell\,;\, 2n = 3m \text{ and } 5k = 7\ell\}$;

(iv) $\{\mathbf{0}^n\mathbf{1}^m\mathbf{2}^k\mathbf{3}^\ell\,;\, 2n = 3k \text{ and } 5m = 7\ell\}$;

(v) $\{\mathbf{0}^n\mathbf{1}^m\mathbf{2}^k\mathbf{3}^\ell\,;\, 2n = 3k \text{ or } 5m = 7\ell\}$;

(vi) $\{\mathbf{0}^p\,;\, p \text{ is prime}\}$;

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

(viii) $\{ww\,;\, w \in \{\mathbf{0}, \mathbf{1}\}^+\}$;

(ix) $\{\mathbf{0}, \mathbf{1}\}^+ \setminus \{ww\,;\, w \in \{\mathbf{0}, \mathbf{1}\}^+\}$;

(x) $\{\mathbf{1010}^2\mathbf{10}^3 ... \mathbf{10}^{n-1}\mathbf{10}^n\,;\, n > 0\}$.