PART II AUTOMATA AND FORMAL LANGUAGES
MICHAELMAS 2025-26
EXAMPLE SHEET 2

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

(i) $\{ww\,;\,\varepsilon \neq w \in \{0,1\}^*\}$;	(v) $\{0^n1^m\,;\,n \neq m\}$;
(ii) $\{w1w\,;\,w \in 0^*\}$;	(vi) $\{0^n1^m\,;\,n \geq m \text{ and } m \leq 1000\}$;
(iii) $\{v1w\,;\,v,w \in 0^*\}$;	(vii) $\{0^n1^m\,;\,n \geq m \text{ and } m \geq 1000\}$;
(iv) $\{0^n1^m\,;\,n > m\}$;	(viii) $\{0^p\,;\,p \text{ is a prime}\}$.

(2) Let $L \subseteq \{0^n1^n\,;\,n \geq 1\}$. Show that L is regular if and only if L is finite.

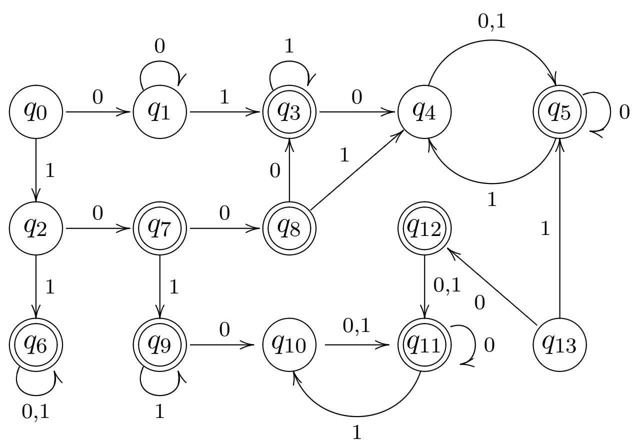
(3) Let $S = \{2^n \mid n \in \mathbb{N}\}$ be the subset of $\mathbb{N}$ consisting of the powers of 2. Show that over the alphabet $\{0,1\}$, the language consisting of the elements of S written in base 2 (with no preceding zeros) is a regular language.

However, show that over the alphabet $\{0,1,\ldots,9\}$ the language consisting of the elements of S written in base 10 (with no preceding zeros) is not a regular language.

(4) (*For those that like Groups*) Draw a nice picture of a Cayley graph of a non-abelian finite group, realised as some DFA D. What is the minimisation of D?

(5) If $D_1 = (Q,\Sigma,\delta,q_0,F)$ is a minimal DFA, and $D_2 = (Q,\Sigma,\delta,q_0,Q\setminus F)$ is a DFA for $\Sigma^* \setminus \mathcal{L}(D_1)$, then is D_2 necessarily a minimal DFA? Prove your answer.

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



Jack Button, jb128@cam.ac.uk, October 27, 2025.

(7) Let $\Sigma = \{0, 1, 2\}$. Show each of the following claims by supplying an appropriate grammar that produces the given language. Explain why your grammar generates this language.

(i) The language consisting of words of the form $(012)^n$ (for $n > 0$) is type 3.
(ii) The language consisting of words of the form 0^n12^n (for $n \in \mathbb{N}$) is type 2.
(iii) The language consisting of words of the form $0^n1^n2^n$ (for $n > 0$) is type 1.

Are any of them of even higher type than listed?

(8) Show that if L is any type 3 language then L is the accepted language of some NFA. Conversely show that if L is the accepted language of a DFA where the start state is not an accept state then L is of type 3. Use this to show that any regular language is context free.

(9) (Another proof that all regular languages are context free.) Let L, M be CFL's, and let a be any symbol. Show that the following are all CFL's: $\emptyset$, $\{\epsilon\}$, $\{a\}$, $L \cup M$, LM, L^*, and L^R (the reverse of L).
Conclude that every regular language is a CFL.

(10) For $e, w \in \Sigma^*$, we say that e is an *excerpt* of w if it is the result of removing finitely many subwords from w, i.e.,

$$w = x_1y_1x_2y_2 \dots x_ny_nx_{n+1} \text{ and } e = y_1 \dots y_n$$

for some $x_1, \dots, x_{n+1}, y_1, \dots, y_n \in \Sigma^*$; 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 $x, v, y \in \Sigma^*$ 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 over $\Sigma = \{0, 1, 2\}$ is context-free, not regular, but satisfies the regular pumping lemma with pumping number $n = 2$:

$$L = \{w20^n1^n \,;\, w \in \Sigma^*, n > 0\} \cup \{0, 1\}^*.$$

(11) Give a CFG which generates the set of regular expressions over the alphabet $\{0, 1\}$. Take as the set of terminals $\Sigma = \{\mathbf{0}, \mathbf{1}, (,), +, ^*, \emptyset, \epsilon\}$. Show that this language is not a regular language.

(12) Let $G = (N, \Sigma, P, S)$ be a CFG in CNF. Suppose we form a new CFG G' from G by adding, for each production of the form $B \to a$ in P (where $a \in \Sigma$), the production $B \to \epsilon$.

Describe the new language $\mathcal{L}(G')$ in terms of the original language $\mathcal{L}(G)$. Is this always true if G is not in CNF?

(13) Convert the following CFG to CNF:

$$S \to aSbb \mid T, \quad T \to bTaa \mid S \mid \epsilon$$

(14) Give a CFG for each of the following CFL's, and then transform each such CFG into CNF:

(a) $\{a^nb^{2n}c^k \mid k, n \geq 1\}$

(b) $\{a^nb^ka^n \mid k, n \geq 1\}$

(c) $\{a^kb^mc^n \mid k, m, n \geq 1, \ 2k \geq n\}$