PART II AUTOMATA AND FORMAL LANGUAGES
MICHAELMAS 2025-26
EXAMPLE SHEET 1

(1) Construct DFAs, via transition diagrams, which accept the following languages:

(a) $\{w \in \{0,1\}^* \mid |w| > 2\}$.
(b) $\{w \in \{0,1\}^* \mid w$ is an alternating sequence of 1's and 0's $\}$.
(c) $\{w \in \{0,1\}^* \mid w$ is a multiple of 3 when interpreted in binary $\}$.
(Leading zeros are permitted and take ϵ to represent the number zero.)
(d) $\{w \in \{a,\ldots,z\}^* \mid w$ contains *ababa* as a substring $\}$.

(2) Construct NFAs, either via transition diagrams or transition tables, which accept precisely the following languages:

(a) $\{w \in \{a,\ldots,z\}^* \mid w$ contains *ababa* as a substring $\}$.
(b) $\{w \in \{a,\ldots,z\}^* \mid w$ contains *dpmms* and/or *damtp* as a substring $\}$.
(c) $\{w \in \{a,\ldots,z\}^* \mid w \in \{what, where, when\}\}$.

Use the subset construction to convert (2a) to a DFA.

(3) Construct an ϵ-NFA which accepts the *union* of the three languages from question (2), and has only *one* accept state.

(4) Given alphabets Σ, Π and a homomorphism $h : \Sigma^* \to \Pi^*$, suppose that $B \subseteq \Pi^*$ is a regular language. Show that its preimage $h^{-1}(B) \subseteq \Sigma^*$ is also regular.
If $A \subseteq \Sigma^*$ and $h(A)$ is regular then is A?

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

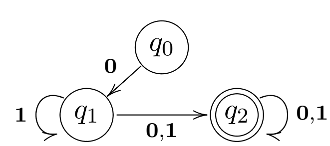
$$L_n := \{w\,;\text{ there are words } x, y \text{ such that } |y| = n-1 \text{ and } w = x1y)\},$$

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

(6) A subset U of $\mathbb{N}$ is said to be **ultimately periodic** if there exists $N \in \mathbb{N}$ and $p > 0$ such that for all $m \geq N$ we have $m \in U$ if and only if $m + p \in U$.

Let Σ be the one letter alphabet consisting of your favourite symbol. Regarding an element of Σ^* as a natural number n where n is its word length (this is sometimes called unary notation for the integers), show that a language $A \subseteq \Sigma^*$ is regular if and only if it is ultimately periodic.

(7) Consider the following nondeterministic automaton over the alphabet $0, 1$:



Jack Button, jb128@cam.ac.uk, October 16, 2025.

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?

(8) Let $D = (\Sigma, Q, \delta, q_0, F)$ be a DFA and define $R := (Q \backslash F) \cup \{\alpha\}$ for some new state α. Let the NFA N be $(\Sigma, R, \Delta, \alpha, \{q_0\})$ where Δ is given by

$$q' \in \Delta(q, a) \text{ if and only if } \delta(q', a) = q,$$
$$q' \in \Delta(\alpha, a) \text{ if and only if } \delta(q', a) \in F,$$
$$\alpha \in \Delta(q, a) \text{ if and only if } \delta(p, a) = q \text{ for some } p \in F, \text{ and}$$
$$\alpha \in \Delta(\alpha, a) \text{ if and only if } \delta(p, a) \in F \text{ for some } p \in F$$

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

(9) Construct ϵ-NFA's and regular expressions for the following regular languages:

(a) All words $w \in \{0, 1\}^*$ consisting of either the string 01 repeated some number of times (possibly none), or the string 010 repeated some number of times (possibly none).

(b) All words $w \in \{a, b, c\}^*$ consisting of some number of a's (possibly none), followed by some number of b's (at least one), followed by some number of c's (possibly none).

(c) All words $w \in \{0, 1\}^*$ which contain a 1 somewhere in the last 4 positions. If $|w| < 4$, then w must contain a 1 somewhere.

(d) All words $w \in \{a, \ldots, z, 0, \ldots, 9, ., :\}^*$ of the form *name*:α.*address*:β. where $\alpha, \beta \in \{a, \ldots, z, 0, \ldots, 9\}^*$. Where might you use such a machine/expression, and why?

(10) Convert each of the following regular expressions to ϵ-NFA's:

(a) $(\mathbf{0} + \mathbf{1})(\mathbf{01})$

(b) $(\mathbf{a} + \mathbf{bb})^*(\mathbf{ba}^* + \epsilon)$

(c) $((\mathbf{aa}^*)^*\mathbf{b})^* + \mathbf{c}$

(11) Prove that $\{w \in \{0, 1\}^* \mid w \text{ contains no more than 5 consecutive 0's}\}$ is regular.

(12) 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) $\mathcal{L}(R(S + T)) = \mathcal{L}(RS) + \mathcal{L}(RT)$
(b) $\mathcal{L}((R^*)^*) = \mathcal{L}(R^*)$
(c) $\mathcal{L}((RS)^*) = \mathcal{L}(R^*S^*)$
(d) $\mathcal{L}((R + S)^*) = \mathcal{L}(R^*) + \mathcal{L}(S^*)$
(e) $\mathcal{L}((R^*S^*)^*) = \mathcal{L}((R + S)^*)$