PART II AUTOMATA AND FORMAL LANGUAGES
MICHAELMAS 2020-21
EXAMPLE SHEET 3

Unless explicitly asked to, you need not *prove* that any machine you construct defines the language you say it does. $*$ denotes a harder problem.

(1) 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 $name{:}\alpha.address{:}\beta.$ where $\alpha,\beta \in \{a,\ldots,z,0,\ldots,9\}^*$. Where might you use such a machine/expression, and why?

(2) Give an ϵ-NFA with the same language as that defined by each of the following regular expressions:
(a) $(\mathbf{0}+\mathbf{1})(\mathbf{01})$ (b) $(\mathbf{a}+\mathbf{bb})^*(\mathbf{ba}^*+\epsilon)$ (c) $((\mathbf{aa}^*)^*\mathbf{b})^*+\mathbf{c}$

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

(4) 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) $\mathcal{L}(R(S+T)) = \mathcal{L}(RS) \cup \mathcal{L}(RT)$

(b) $\mathcal{L}((R^*)^*) = \mathcal{L}(R^*)$

(c) $\mathcal{L}((RS)^*) = \mathcal{L}(R^*S^*)$

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

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

(5) Use the pumping lemma to show that none of the following languages are regular:

(a) $\{a^n b^n \mid n \geq 0\}$

(b) $\{a^n b^{2n} \mid n \geq 0\}$

(c) $\{ww \mid w \in \{0,1\}^*\}$

Date: November 9, 2020.

(6) For each of the following languages, determine whether or not they are regular. Justify your answers.

(a) $\{a^n b^m \mid n \neq m\}$

(b) $\{xcx \mid x \in \{a,b\}^*\}$

(c) $\{xcy \mid x, y \in \{a,b\}^*\}$

(d) $\{a^n b^m \mid n > m\}$

(e) $\{a^n b^m \mid n \geq m \text{ and } m \leq 1000\}$

(f) $\{a^n b^m \mid n \geq m \text{ and } m \geq 1000\}$

(g) $\{1^p \mid p \text{ is a prime }\}$

(7) Prove that no infinite subset of $\{0^n 1^n \mid n \geq 0\}$ is a regular language.

(8) Find minimal DFA's for each of the following languages. In each case, *prove* that your DFA is minimal.

(a) $\{a^n \mid n \geq 0,\ n \neq 3\}$

(b) $\{a^m b^n \mid m \geq 2,\ n \geq 3\}$

(c) $\{a^m b \mid m \geq 0\} \cup \{b^n a \mid n \geq 0\}$

(9) 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.

(10) Let L, M be languages over Σ. We define the *difference* $L - M$ to be the words that are in L but not M. That is, $L - M := (L \cup M) \setminus M$. Show that if L, M are both regular languages, then $L - M$ is a regular language over Σ.

(11) Find a DFA which accepts $\{w \in \{0,1\}^* \mid w \text{ is a multiple of 3 when interpreted in binary }\}$. Compute the remainder, mod 3, of the following binary number: 101101011000011010101101011. Convert your DFA to a regular expression accepting the same language.

(12) Let D be a DFA with N states. Prove the following:
(a) If D accepts at least one word, then D accepts a word of length less than N.
(b) If D accepts at least one word of length $\geq N$, then D accepts infinitely many words.

(13) Show that the language $L := \{w01^n \mid w \in \{0,1\}^*, n \in \mathbb{K}\} \cup \{1\}^*$ satisfies the pumping lemma for regular languages. Is L a regular language?

(14) Give an algorithm that, on input of a DFA D, decides if $\mathcal{L}(D) = \emptyset$ or not.

(15*) Give an algorithm that, on input of DFA's D_1, D_2, decides if $\mathcal{L}(D_1) \subseteq \mathcal{L}(D_2)$ or not. *(You may appeal to results from the lectures.)*

(16*) For any $X \subseteq \{1\}^*$, show that X^* is a regular language.