

Note: Use L^AT_EX or Word to typeset your solutions, or write them legibly by hand and submit the scan. You can use the source code of this file as a template. To draw the finite automaton, you can use TikZ or <https://madebyevan.com/fsm/>.

Problem 1 (More Closure Properties). Prove the following closure properties of regular languages.

Hint. Using the proper equivalent definitions of regular languages could significantly simplify the proofs.

1. **(10 pts).** For a string w , let w^R be the string written in the reverse order. Prove that if L is a regular language, then

$$L^R = \{w^R \mid w \in L\}$$

is also regular.

2. **(10 pts).** If L is a regular language, prove that

$$\text{sub}(L) = \{x \mid x \text{ is a subsequence of some } y \in L\}$$

is also regular. Here subsequence means not necessarily consecutive substring.

3. **(Bonus, 10 pts).** If L is a regular language over the alphabet Σ , prove that

$$L_{1/2} = \{x \mid xy \in L \text{ for some } y \in \Sigma^*, |x| = |y|\}$$

is also regular. Here $|x|$ denotes the length of the string x .

Problem 2 (NFA and DFA size). For a constant $k \in \mathbb{N}_+$, let L_k be the following language on alphabet $\Sigma = \{a, b\}$:

$$L_k = \{w \mid |w| \geq k, \text{ the } k\text{-th last symbol of } w \text{ is } a\}.$$

In other words, L_k is the language represented by regex $\Sigma^* a \Sigma^{k-1}$.

1. **(5 pts).** Describe an NFA with at most $k + 1$ states that computes L_k .
2. **(10 pts).** Prove that no DFA with less than 2^k states can compute L_k .

Problem 3 (Determine regularity). Determine whether the following languages are regular or not, and justify your answers.

1. **(5 pts).** $\{w \in \{0,1\}^* \mid w \text{ is the binary representation of } 2^k \text{ for some } k \in \mathbb{N}\}$.
2. **(5 pts).** $\{ww^R \mid w \in \{a,b\}^*\}$.
3. **(10 pts).** $\{wvw \mid v, w \in \{a,b\}^*, |w| \geq 1\}$.
4. **(10 pts).** $\{(G, k) \mid \text{undirected graph } G \text{ contains a matching of size } k\}$. Here G is encoded as an 0,1-string representing its adjacency matrix, and k is in binary.
5. **(10 pts)** $\{w \in \{a,b\}^* \mid w \text{ contains the same number of consecutive } ab\text{'s and } ba\text{'s}\}$.
6. **(10 pts)** $\{w \in \{a,b,c\}^* \mid w \text{ contains the same number of consecutive } ab\text{'s and } ba\text{'s}\}$.
7. **(Bonus, 10 pts).** $\{w \in \{0,1,\dots,9\}^* \mid w \text{ is the decimal representation of } 2^k \text{ for some } k \in \mathbb{N}\}$.

Problem 4 (Unary regular languages). In this problem, we will classify all regular languages over the unary alphabet $\Sigma = \{1\}$.

1. **(5 pts).** Given $a \in \mathbb{N}$ and finite sets $B, C \subset \mathbb{N}$, show that the unary language:

$$L_{a,B,C} = \{1^{an+b} \mid n \in \mathbb{N}, b \in B\} \cup \{1^c \mid c \in C\}$$

is regular.

2. **(10 pts).** Prove that every regular language over the unary alphabet $\Sigma = \{1\}$ is of the form $L_{a,B,C}$ for some $a \in \mathbb{N}$ and finite sets $B, C \subset \mathbb{N}$.

Hint. You can use either the pumping lemma or the Myhill-Nerode theorem.