CS 332: *Theory of Computation* — Prof. Mark Bun
Boston University — February 1, 2024

Homework 2 – Due Thursday, February 8, 2024 at 11:59 PM

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the course staff if asked. You must also identify your collaborators and write "Collaborators: none" if you worked by yourself. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden. Collaboration is not allowed on problems marked "INDIVIDUAL."

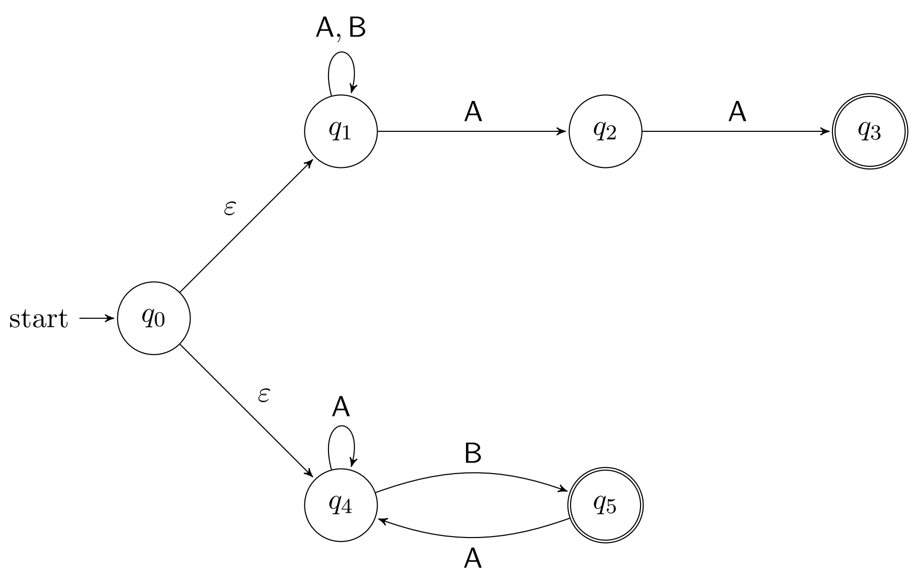
Problems There are 7 required problems and one bonus problem. Problems 1-2 and 4-7 are to be turned in via Gradescope. Problem 6 will be autograded using AutomataTutor.

1. Draw (and include in the PDF you submit to Gradescope) state diagrams of DFAs with as few states as you can recognizing the following languages. You may assume that the alphabet in each case is $\Sigma = \{0, 1\}$.

   (a) $L_1 = \{w \mid w$ is a string of the form $x_1 y_1 x_2 y_2 \dots x_n y_n$ for some natural number n such that if x is the integer with binary representation $x_1 x_2 \dots x_n$ and if y is the integer with binary representation $y_1 y_2 \dots y_n$, then $x < y\}$. Both x and y are presented starting with the most significant bit and may have leading 0s.

   (b) $L_2 = \{w \mid w$ represents a binary number that is divisible by $4\}$. The number is presented starting from the most significant bit and can have leading 0s.

2. **INDIVIDUAL (all three parts of this problem):** Consider the following state diagram of an NFA N over alphabet $\{\mathsf{A}, \mathsf{B}\}$.



   (a) Give a formal description of the machine N as a 5-tuple.

   (b) Consider running N on input BBAA. Give examples (i) of a computation path that leads N to an accept state when run on this input, (ii) a computation path that leads N to a reject state, and (iii) a computation path that leads N to fail before it's read the entire input.

(c) What is the language recognized by N?

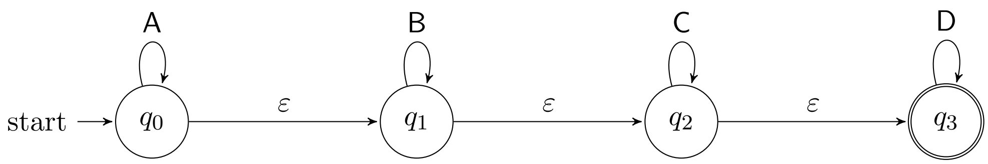
3. (Autograded on AutomataTutor – no need to hand anything in on your PDF) Give state diagrams of NFAs with as few states as you can recognizing the following languages. You may assume that the alphabet in each case is $\Sigma = \{0, 1\}$.

   (a) $L_3 = \{w \mid w \text{ ends with an odd number of 0's}\}$. For example, $101000 \in L_3$ and $11111 \notin L_3$.

   (b) $L_4 = \{w \mid \text{the length of } w \text{ is divisible by either 2 or 3}\}$.

   (c) $L_5 = \{w \mid w \text{ contains at least two 0's with exactly four symbols between them}\}$. For example, $011110 \in L_5$, while $01010 \notin L_5$.

4. Consider the following state diagram of an NFA N over alphabet $\{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}\}$.



   (a) What is the language recognized by N?

   (b) Use the subset construction to convert N into a DFA recognizing the same language. Give the state diagram of this DFA – only include states that are reachable from the start state.

5. (**NFAs can be simpler than DFAs**) Consider the language $L = \{w \in \{0, 1\}^* \mid |w| \text{ is not divisible by } 3\}$.

   (a) Give the state diagram of an NFA recognizing L that has exactly one accept state.

   (b) Prove that every DFA recognizing L requires at least two accept states.

6. (**Closure care**) On Monday, we'll show that the class of regular languages is closed under the star operation. This problem will help you investigate this property.

   (a) Let $A = \{w \in \{0, 1\}^* \mid w \text{ contains the symbol } 1\}$. Give the state diagram of a 2-state DFA D recognizing A.

   (b) Give a simple description of the language A^*.

   (c) Consider the following **failed** attempt to construct an NFA recognizing A^*: Add an ε transition from every accept state of D to the start state, and make the start state an accept state. Draw the state diagram of this NFA, and call it N.

   (d) What is $L(N)$? Give an example of a string w such that $w \in L(N)$, but $w \notin A^*$.

   (e) Give the state diagram of an NFA that *does* recognize A^*.

7. (**Set difference**) Given languages A and B over the same alphabet Σ, define their "set difference" to be the language $A \setminus B = \{w \in \Sigma^* \mid w \in A \text{ but } w \notin B\}$.

   (a) Explain how to write the set difference operation in terms of (some of the) union, concatenation, star, reverse, intersection, and complement operations. That is, describe how for arbitrary A, B one can obtain $A \setminus B$ by applying these other operations to A and B.

   (b) Use part (a) to show that the regular languages are closed under set difference.

Bonus Problems

1. Describe and analyze an algorithm that takes as input the description (e.g., the state diagram) of an NFA N and a string w, and outputs whether N accepts w. Your algorithm should run in time $O(s^2n)$ where s is the number of states of N and n is the length of the string w.