CS2333 - Summer 2025

Assignment # 2

Due: Wednesday, May 21, by 11:00 pm

Submission Instructions:

- Your answers should be submitted through the assignment dropbox on Desire2Learn. Please either submit as a single file or with one file per question.

  If you are submitting in a single file, name it `CS2333-A2.`$\langle extn \rangle$. Your answers should be in the same order as the questions.

  If you are submitting with one file per question, your files should be named `CS2333-A2Q1.`$\langle extn \rangle$, `CS2333-A2Q2.`$\langle extn \rangle$, and so on for each of the questions.

  In both cases, $\langle extn \rangle$ should be the appropriate extension for your file type. D2L supports a variety of file types for submission and marker annotation, including `.pdf`, `.docx`, `.png`, and `.jpg`.

  Contact your instructor if you have any questions.

- All answers you submit must be your own work. You may discuss general approaches to assignment problems with your classmates. However, these must be general and cannot include things such as detailed steps of an algorithm or a proof. Please see the course syllabus for more details.

- Late assignment submissions will be considered only for medical reasons or in other exceptional circumstances, and normally only if the instructor is contacted before the assignment deadline.

---

1. **(20 marks)** For each of the languages below, draw the state diagram for a **nondeterministic** finite automaton (NFA) to accept the language.

   In order to get full marks,

   - your NFA must have at most the number of states specified
   - your NFA must take advantage of nondeterminism:
     - There must be at least one case where the machine has a choice between two or more next states on some input symbol, or at least one case where the machine has no possible next state on a particular input symbol, or you must use ε-transitions in some nontrivial way.

   (a) $L_1 = \{w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \ : \ w$ starts with `abc` and ends with `bcb`$\}$

   Use no more than 7 states.

   (b) $L_2 = \{w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \ : \ w$ ends with two symbols that are different from each other, or the first two symbols of w include at least one '`a`'$\}$

   Use no more than 11 states.

   [Note: The 'or' in this question is an inclusive 'or'. A string like `cababc` satisfies both conditions and should be accepted.]

(c) $L_3 = \{w \in \{0,1\}^* \ : \ w \text{ contains both the substring 101 and the substring 00}\}$

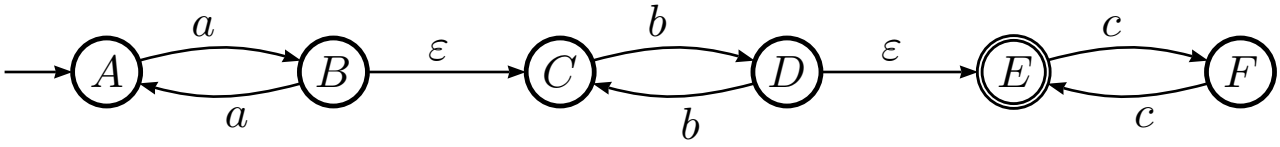
Use no more than 13 states.

(d) $L_4 = \{w \in \{\mathtt{a},\mathtt{b},\mathtt{c}\}^* \ : \ \text{the second-last symbol of } w \text{ is an `}\mathtt{a}\text{', and } |w|_{\mathtt{c}} \text{ is odd}\}$

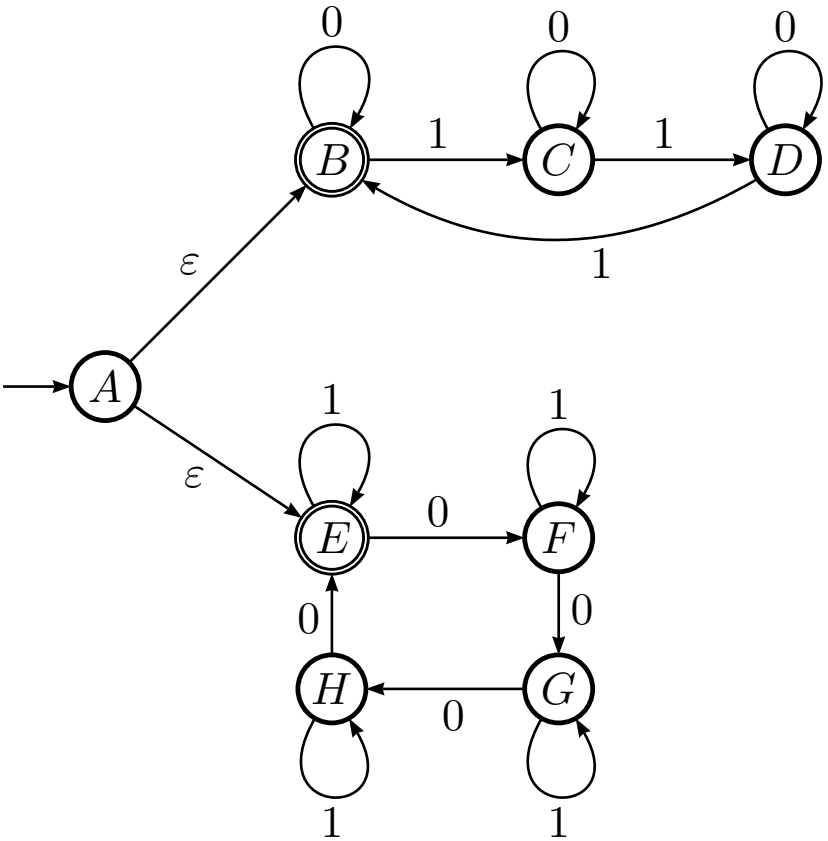
Use no more than 9 states. [Note: If you want an extra challenge, it <u>can</u> be done with only 5 states.]

2. **(6 marks)** For each nondeterministic finite automaton below, identify the language accepted. Try to describe the language as simply as you can.

(a)



(b)



3. **(4 marks)** Given the formal description below for a nondeterministic finite automaton, draw the corresponding state diagram.

   (You do not have to identify the language accepted by the finite automaton.)

   - $Q = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9, q_{10}\}$
   - $\Sigma = \{\texttt{a}, \texttt{b}, \texttt{c}\}$
   - $q = q_0$
   - $F = \{q_3, q_6, q_{10}\}$
   - $\delta : Q \times (\Sigma \cup \{\epsilon\}) \to \mathcal{P}(Q)$ is defined in the following table:

	a	b	c	ϵ
q_0	$\varnothing$	$\varnothing$	$\varnothing$	$\{q_1, q_4, q_7\}$
q_1	$\{q_2\}$	$\varnothing$	$\varnothing$	$\varnothing$
q_2	$\{q_2, q_3\}$	$\{q_2\}$	$\{q_2\}$	$\varnothing$
q_3	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
q_4	$\{q_4\}$	$\varnothing$	$\varnothing$	$\{q_5\}$
q_5	$\varnothing$	$\{q_5\}$	$\varnothing$	$\{q_6\}$
q_6	$\{q_6\}$	$\varnothing$	$\{q_6\}$	$\varnothing$
q_7	$\{q_7\}$	$\{q_7\}$	$\{q_7, q_8\}$	$\varnothing$
q_8	$\varnothing$	$\varnothing$	$\{q_9\}$	$\varnothing$
q_9	$\varnothing$	$\varnothing$	$\{q_{10}\}$	$\varnothing$
q_{10}	$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$

4. **(5 marks)** In class, we discussed the fact that if we have a DFA accepting a language L, then it is possible to obtain a DFA to accept $\overline{L}$ (the complement of L) simply by turning the accepting states into non-accepting states, and vice versa. However, we said that this approach does not necessarily work for NFAs.

   Prove that this method does not always work for NFAs by doing the following:

   - draw the state diagram for an NFA that accepts $L = \{w \in \{\texttt{a}, \texttt{b}, \texttt{c}\}^* \mid w \text{ contains the substring } \texttt{abcba}\}$
   - reverse the accepting and non-accepting states to obtain a new NFA
   - explain why the language accepted by this new NFA is **not** the complement of L

5. **(8 marks)** Consider the following languages:
   $L_1 = \{\texttt{b}, \texttt{ba}\},\ \ L_2 = \{\texttt{ab}, \texttt{aab}, \texttt{aaaa}\},\ \ L_3 = \{\texttt{a}^n\texttt{b}\ :\ n \in \mathbb{Z}^{nonneg},\ n < 5\}$

   (a) List all strings in the language $L_1 \cup L_2$.

   (b) List all strings in the language $L_2 \cup L_3$.

   (c) List all strings in the language $L_1 L_2$.

   (d) List all strings in the language $L_2 L_1$.

   (e) List all strings in the language $(L_1)^*$ that have a length less than or equal to 3.

   (f) Would the string `bbabbbbaabbbaba` be in $(L_1)^*$? Explain.

   (g) Is it possible for a string in $(L_2)^*$ to have more '`b`'s than '`a`'s? Explain.

   (h) Is it possible for a string in $(L_3)^*$ to have more '`b`'s than '`a`'s? Explain.

6. **(16 marks)** Provide a regular expression for each of the following languages.

   (a) $L_1 = \{w \in \{\texttt{a}, \texttt{b}\}^*\ :\ w$ starts with `bbb` and ends with `abba`$\}$

   (b) $L_2 = \{w \in \{0, 1\}^*\ :\ |w|_0 = 4k + 1$ for some $k \in \mathbb{Z}^{nonneg}\}$

   (c) $L_3 = \{w \in \{0, 1\}^*\ :\ |w| \geq 2$ and the first and last symbol of w are the same as each other$\}$

   (d) $L_4 = \{w \in \{\texttt{a}, \texttt{b}, \texttt{c}\}^*\ :$ every '`a`' is followed immediately by at least three consecutive '`b`'s$\}$

7. **(6 marks)** For each regular expression below, identify the language that is generated. Try to describe the language as simply as you can.

   (a) $(0 \cup 1 \cup \varepsilon)\ (0 \cup 1 \cup \varepsilon)\ (0 \cup 1 \cup \varepsilon)$

   (b) $(\texttt{aa})^*\ (\texttt{ab} \cup \texttt{ba})\ (\texttt{aa})^*$