Assignment 1 Solution

COMPSCI 3331

1. (5 marks) (a) Let Σ be any finite alphabet with at least one letter. Prove that for all languages $L_1, L_2 \subseteq \Sigma^*$, if $L_1 \subseteq L_2$, then $L_1^* \subseteq L_2^*$.
(b) If $L_1^* \subseteq L_2^*$, is it true that $L_1 \subseteq L_2$? Prove or disprove (by giving a counter-example).

Solution

(a) Let L_1, L_2 be languages with $L_1 \subseteq L_2$. Let $x \in L_1^* = \cup_{i \geq 0} L_1^i$. Then there exists $i \geq 0$ such that $x \in L_1^i$. Let $x_1, x_2, \cdots, x_i \in L_1$ be words with $x = x_1 x_2 \cdots x_i$. Then as $L_1 \subseteq L_2$, $x_j \in L_2$ for all j with $1 \leq j \leq i$. Thus, $x = x_1 x_2 \cdots x_i \in L_2^i \subseteq L_2^*$.

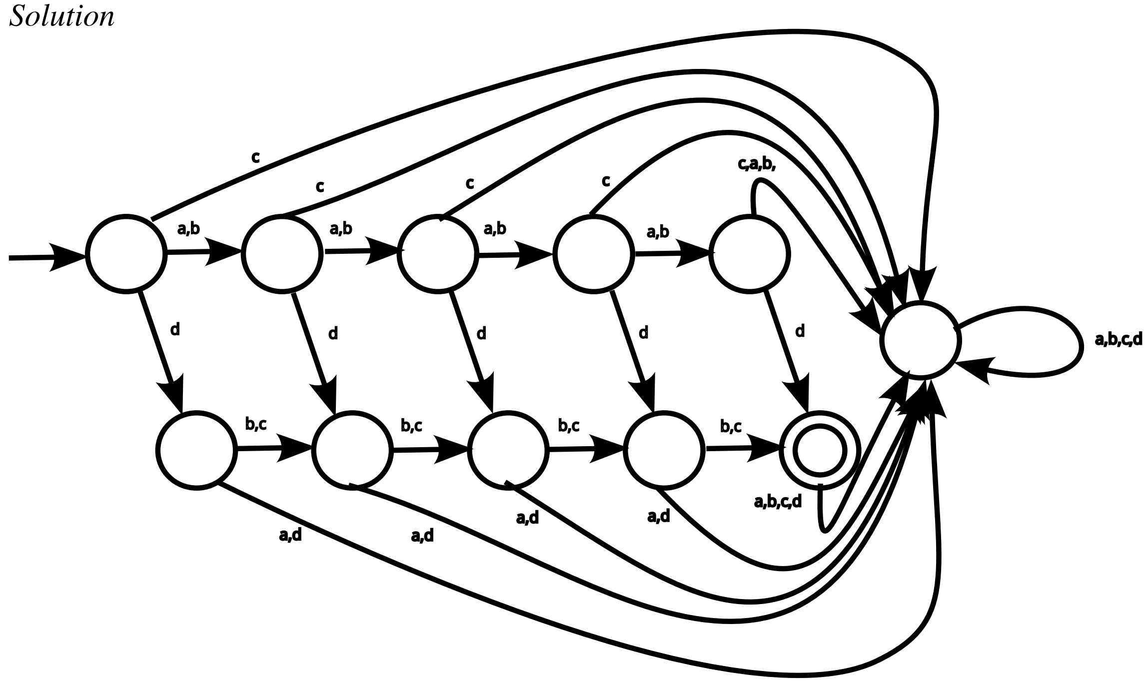
(b) The reverse implication does not hold. Let $L_1 = \{a\}^*$ and $L_2 = \{a\}$. Then $L_1^* = (\{a\}^*)^* = \{a\}^*$ and $L_2^* = \{a\}^*$. Thus, $L_1^* \subseteq L_2^*$. But $L_1 \subseteq L_2$ does not hold.

2. (5 marks) Let $\Sigma = \{a, b, c, d\}$ and $L \subseteq \Sigma^*$ be the following language

$$L = \{x_1 d x_2 \ : \ x_1 \in \{a,b\}^*, x_2 \in \{b,c\}^*, |x_1| + |x_2| = 4\}$$

Show that L is regular by giving a DFA that accepts the language. You do not need to formally prove that the language is regular, but you must include some justification that the DFA accepts the language L, such as arguing informally as to the role of the states in your DFA.

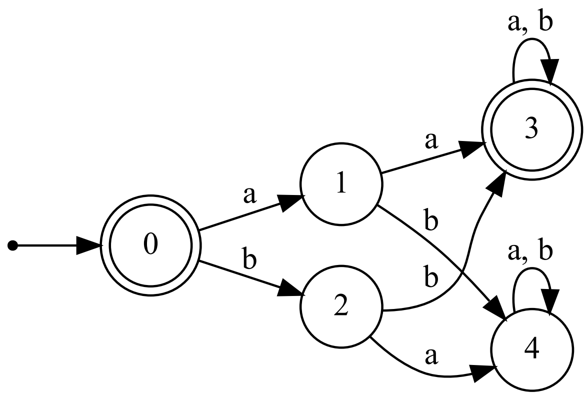
Solution



Informal justification: The DFA consists of two 'chains' - the top chain and the bottom chain. The top chain reads $x_1 \in \{a,b\}^*$ and the bottom chain reads $x_2 \in \{b,c\}^*$.

Each of the states in the top chain counts one symbol from x_1. When a d is read, then the DFA transitions from the top chain to the bottom chain, but in the same "position", and not restarting the chain. In this way, to get to the final state, the DFA must go through the length of an entire chain, but broken by the letter d, meaning that $|x_1| + |x_2| = 4$.

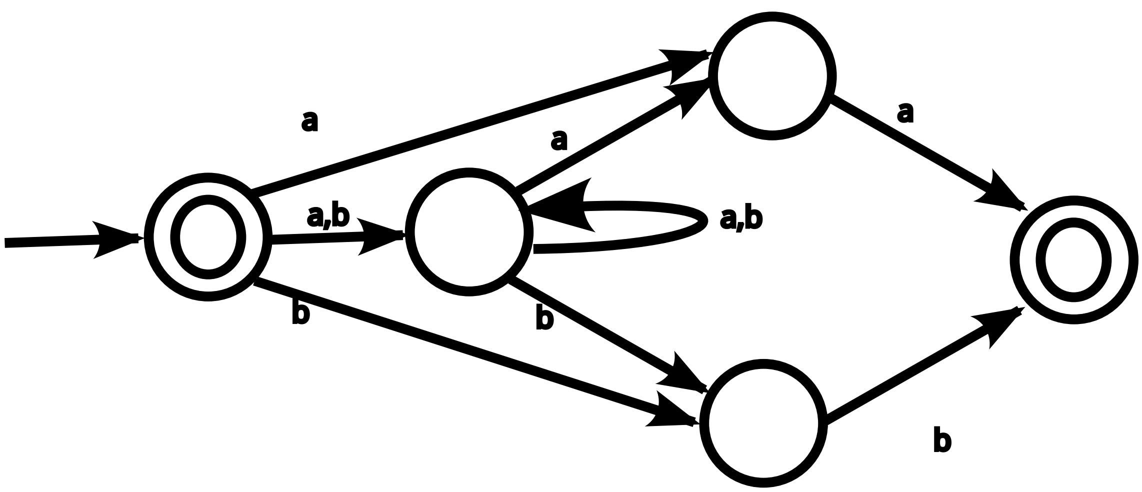
3. (5 marks) Consider the DFA M below, over the alphabet $\{a,b\}$. Give both an NFA and a DFA for $L(M)^R$. (*Hint: for the DFA it may be helpful to determine what the language $L(M)$ is, and then construct a DFA for $L(M)^R$ based on that. You are not required to perform the subset construction.*)



Solution

The language $L(M)^R$ is $\{\varepsilon\} \cup \{wz \ : \ w \in \{a,b\}^*, z \in \{aa, bb\}\}$.

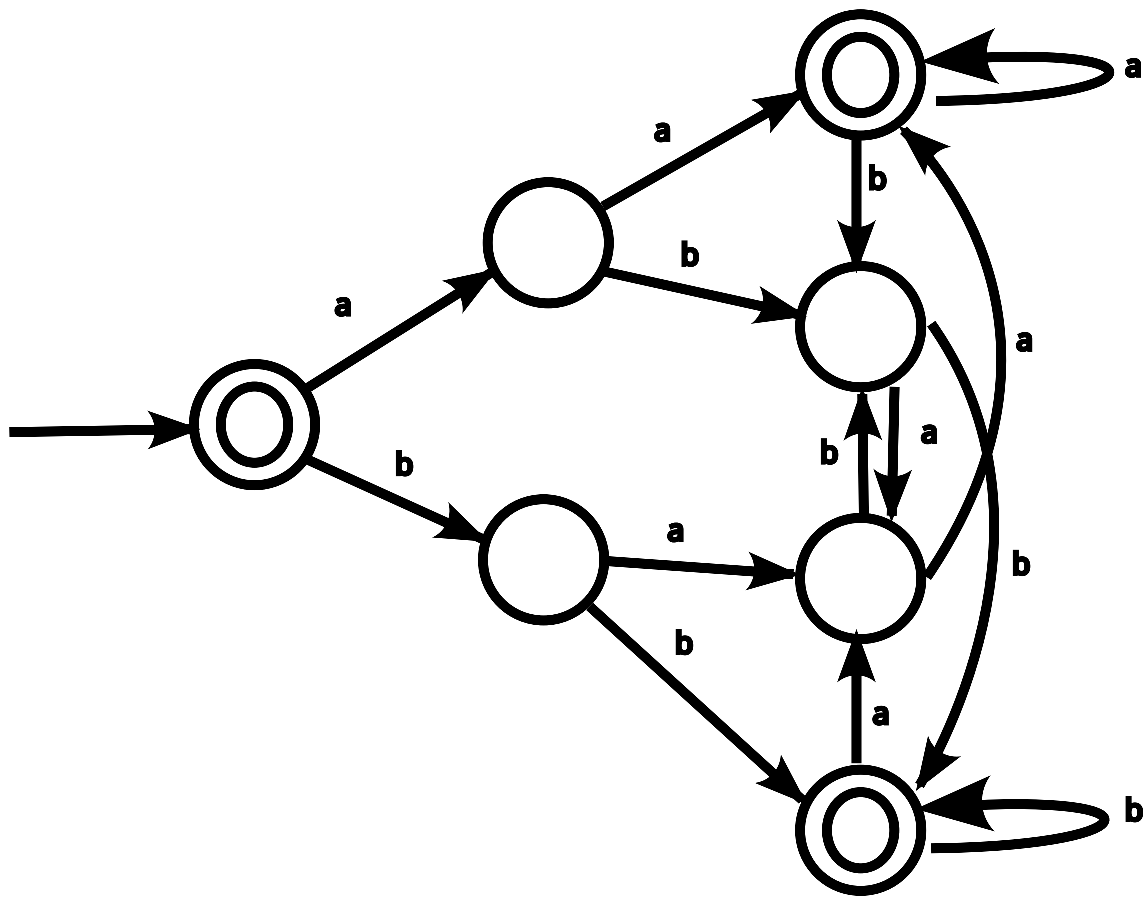
First, here an NFA for the language



There are multiple NFAs for the language. This NFA is designed intuitively as follows:

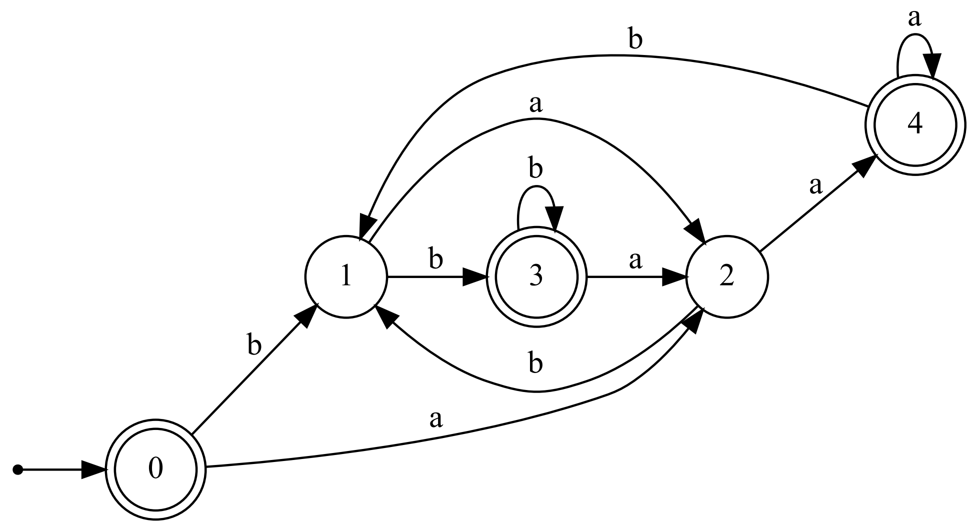
- The start state is final, since ε is in the language.
- The state with the loop allows us to read any sequence of as and bs, and then nondeterministically choose to read the final aa or bb and end up in the final state.

- The transition from the start state to the loop state means that the length of the word before *aa* or *bb* must be at least one, so we add the other transitions leaving the start state so the NFA can also accept *aa* and *bb* (with nothing before).



This DFA accepts the language by keeping track of the last two characters read at any point. After the first two letters read, we are in one of four states on the right-hand column of states. The top state corresponds to *aa* and the bottom corresponds to reading *bb*. After that, the transitions leaving this column of states reflects the new "last two letters read". For instance, in the *aa* state (top) and reading *b*, we move down one state to *ab*, reflecting that the last two characters we havfe read are now *ab*.

Note that there is a smaller DFA for this language. The DFA above is chosen to illustrate the concept in the easiest way. The DFA below accepts the same language by simplifying two states.



0
1
2
3
4
a
b