CS 3331a - Assignment 3 - Solutions

Paul Vrbik

November 17, 2010

Question 1 - marks

Prove by Pumping Lemma that $L = \left\{ x \in \{a,b\}^* \;\middle|\; |x|_a = 2|x|_b \right\}$ is not a DFA language.

Proof. Towards a contradiction assume that L is a DFA language. So, there is some n such that for any $w \in L$ with $|w| \geq n$ we have a decomposition $w = xyz$ given by the pumping lemma.

Consider the string $w = a^{2n}b^n \in L$. Pumping lemma gives $a^{2n}b^n = xyz$ with $|xy| \leq n$ and $|y| \geq 1$. This means that xy, and therefore y, is a sequence of a's with the later being a non-empty sequence.

However, pumping lemma also guarantees for $k \geq 0$ that $xy^kz \in L$ and therefore $xz \in L$ as well. As y is a nonempty string of a's we have $|xz|_a < |xyz|_a$ and $|xyz|_b = |xz|_b$ which permits us to deduce

$$|xz|_a < |xyz|_a = 2|xyz|_b = 2|xz|_b$$

implying $|xz|_z \neq 2|xz|_b$ and therefore $xz \notin L$; a contradiction. □

Question 2 - marks

Write a regular expression for the following languages over $\{a, b\}$:

(1) the set of all words that have no consecutive a's.

$$(ab + b)^*(a + \varepsilon) \text{ or } (a + \varepsilon)(b + ba)^*$$

(2) the set of all words containing at least three consecutive b's.

$$(a + b)^*(bbb)(a + b)^*$$

Question 3 - marks

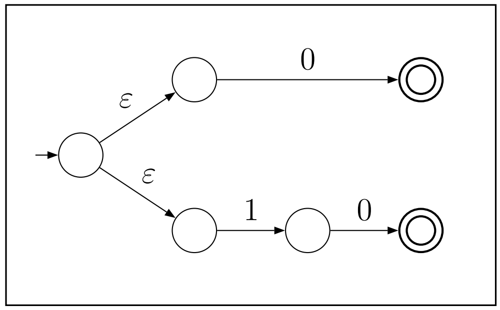
Given the following regular expression E

$$(0 + 10)^*1 + 00$$

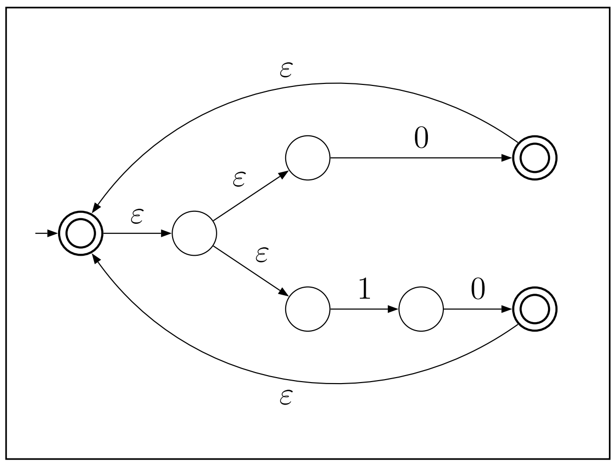
construct an $\varepsilon - NFA$ A such that $L(A) = L(E)$. (All the intermediate steps are required).

I follow the procedure of the notes and not that of the textbook.

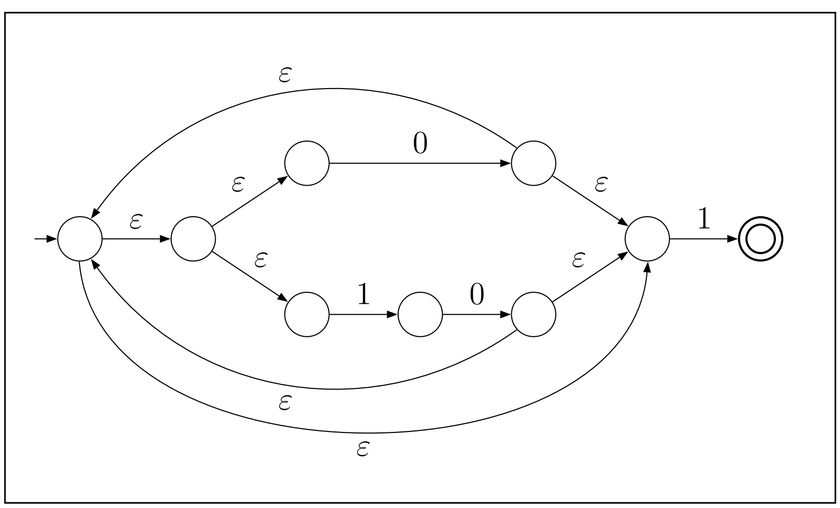
- $(0 + 10)$



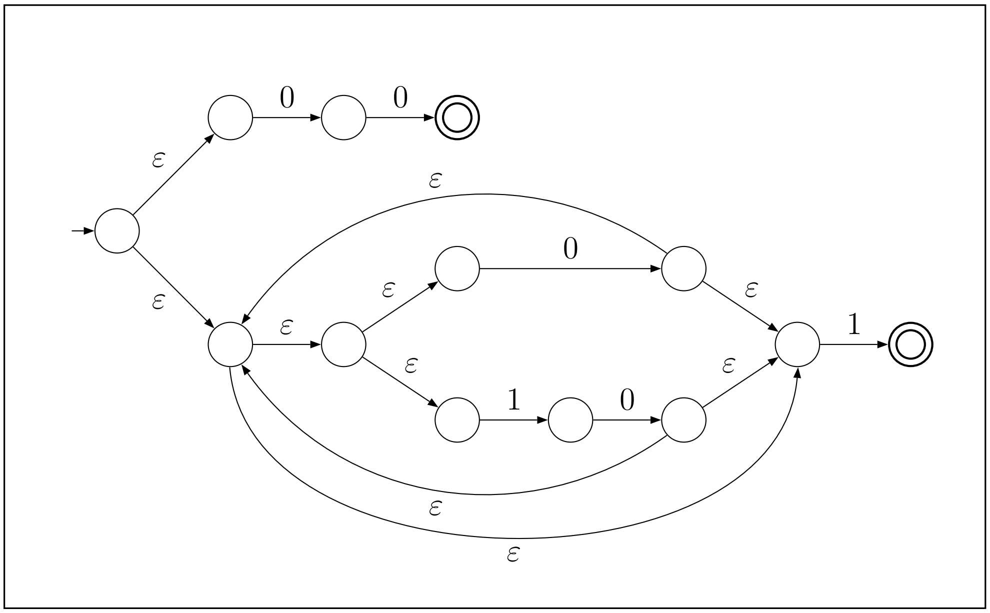
- $(0 + 10)^*$



- $(0 + 10)^*1$



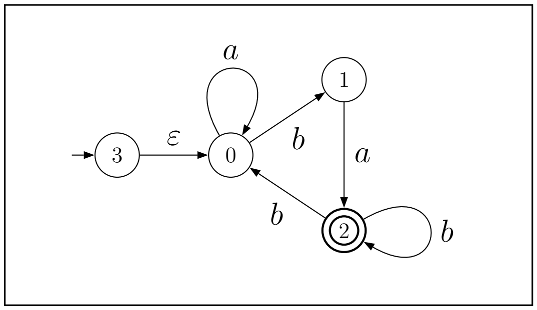
- $(0 + 10)^*1 + 00$



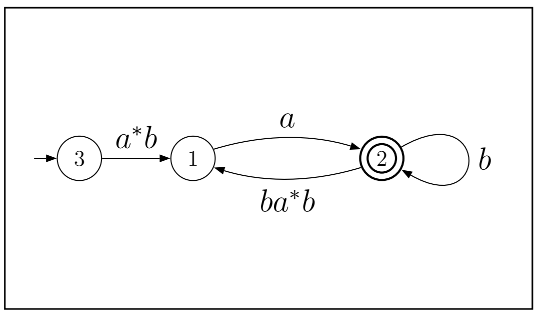
Question 4 - marks

Given the NFA A, obtain an equivalent regular expression. (All the intermediate steps are required.)

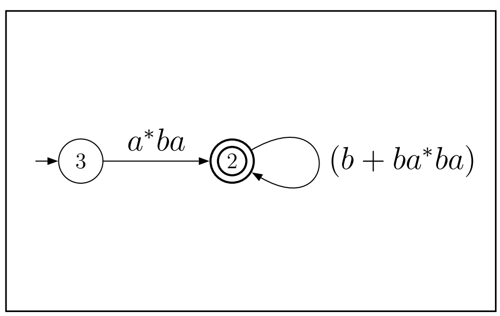
- Add a different initial state.



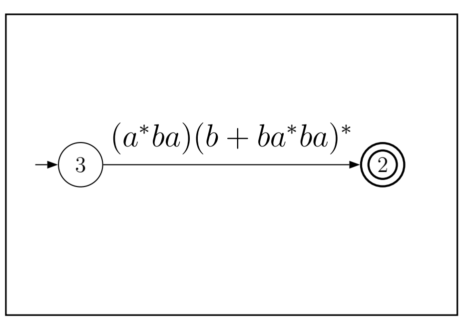
- Eliminate state 0.



- Eliminate state 1.



- Combine remaining state transition rules.



Question 5 - marks

Give a CFG for each of the following languages:

(1) $L_1 = \left\{0^i 1^{i+3} \;\middle|\; i \geq 0\right\}$

$$N = \{S\},\ \Sigma = \{0, 1\}$$
$$P : S \to 0S1 \;\big|\; 111$$

(2) $L_2 = \left\{a^i b^j \;\middle|\; 0 \leq i + 1 < j\right\}$

$$N = \{S\},\ \Sigma = \{a, b\}$$
$$P : S \to bb \;\big|\; aSb \;\big|\; Sb$$

(3) L_3 is the set of all strings over the set of symbols ')' and ')' such that '(' and ')' are well paired and nested (i.e. the set of all balanced bracketed expressions). For example, '((()))', '(()(()))', and '(())((())())()' are all words in L_3.

$$N = \{S\},\ \Sigma = \{(,)\}$$
$$P : S \to \varepsilon \;\big|\; (S) \;\big|\; SS$$

(4) $L_4 = \left\{0^m 1^n 0^{m+n} \;\middle|\; m, n \geq 0\right\}$

$$N = \{S, A\},\ \Sigma = \{0, 1\}$$
$$P : S \to \varepsilon \;\big|\; 0S0 \;\big|\; A$$
$$A \to \varepsilon \;\big|\; 1A0$$