

Exercises with corrections

Finite automatas and regular languages

Exercise 1 : *First automatas.*

Let $\Sigma = \{a, b\}$. Construst a deterministic automata that accepts the following:

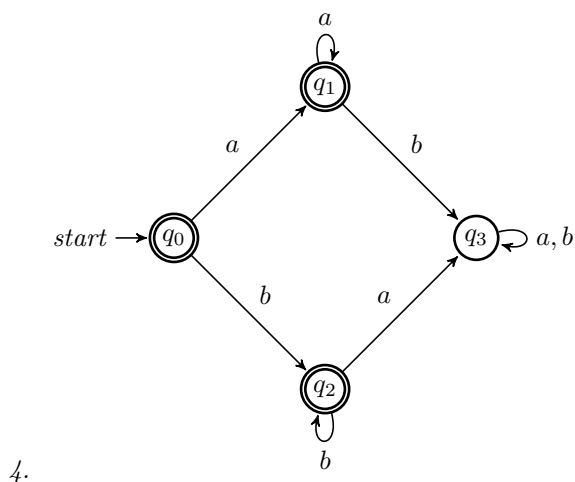
1. No word over Σ ;
2. Any word over Σ ;
3. Only the words composed of one symbol of Σ ;
4. $a^* \cup b^*$.

Answer:

1. $start \rightarrow (q) \xrightarrow{a, b}$

2. $start \rightarrow (q) \xrightarrow{a, b}$

3. $start \rightarrow (q_0) \xrightarrow{a, b} (q_1) \xrightarrow{a, b} (q_2) \xrightarrow{a, b}$



□

Exercise 2 : *Salad of automatas.*

Let $\Sigma = \{a, b\}$. Construct an automata that accepts the following languages:

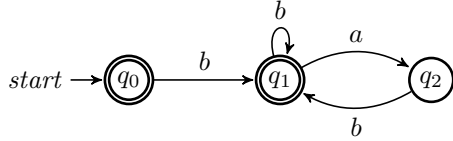
1. $L_1 = \{w \in \Sigma^* \mid \text{every } a \text{ in } w \text{ is immediately preceded and followed by a } b\}$;
2. $L_2 = \{w \in \Sigma^* \mid ab \notin w \text{ and } ba \notin w\}$;
3. $L_3 = \{w \in \Sigma^* \mid ab \in w \text{ or } ba \in w\}$;
4. $L_4 = \{w \in \Sigma^* \mid ab \in w \text{ and } ba \in w\}$;

5. $L_5 = \{w \in \Sigma^* \mid w \text{ contains an odd number of } a \text{ and an odd number of } b\}$;
6. $L_6 = \{w \in \Sigma^* \mid aaa \notin w \text{ and } w \text{ contains an odd number of } b\}$;
7. $L_7 = \{w \in \Sigma^* \mid aa \in w \text{ and the first occurrence of } aa \text{ is not preceded by } abab\}$.

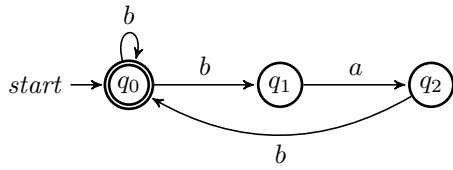
Answer:

We will not build complete automatas. If, when reading a word, the automata cannot follow a transition from its current state, we consider that it rejects the word.

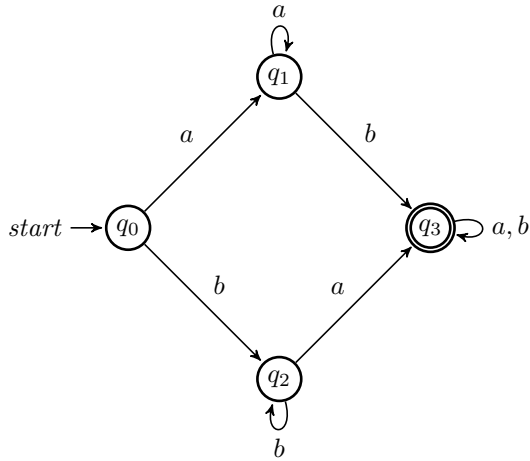
1. Deterministic solution:



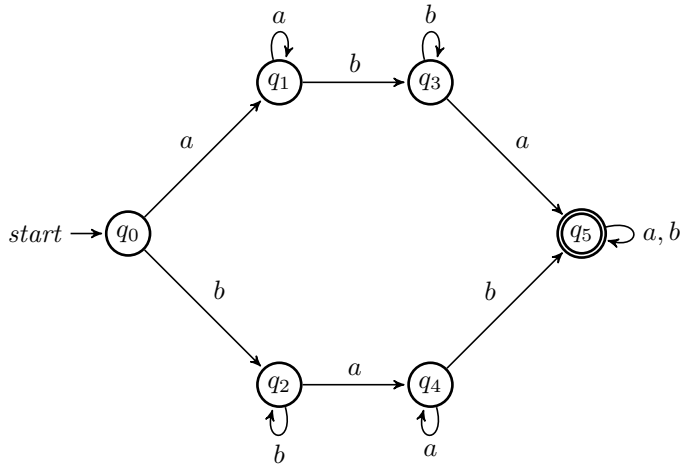
Non-deterministic solution:



2. This language is equivalent to $a^* \cup b^*$.

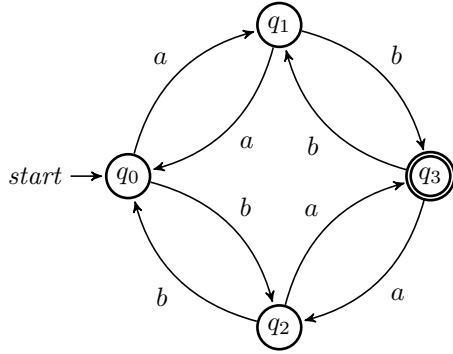


3.

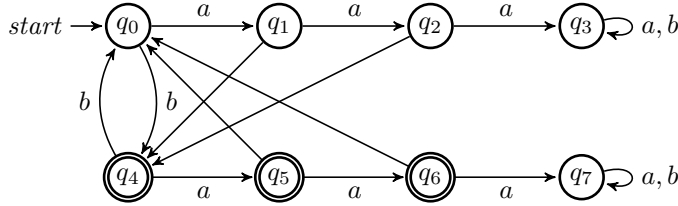


4.

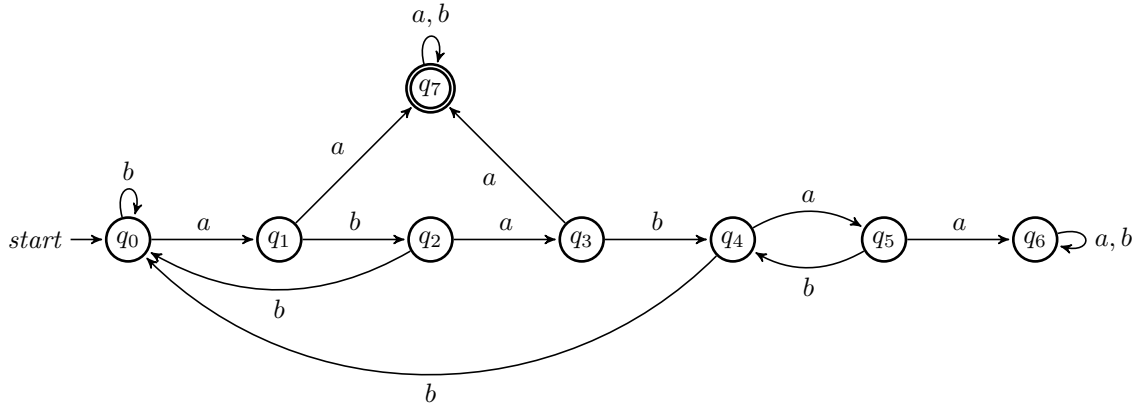
5. In the following automata, q_0 is the state where there is an even number of both a and b , q_1 is the state where there is an odd number of a and an even number of b , q_2 is the state where there is an even number of a and an odd number of b , and q_3 is the state where there is an odd number of both a and b :



6. In the following automata, the top row is when we have read an even number of b , the bottom row when we have read an odd number of b . Each column corresponds to, respectively, 0, 1, 2 and 3 consecutive a . The back edges from q_1 , q_2 , q_5 and q_6 are all labeled with b .



7. In the following automata, the state q_7 is reached when the first aa has been read and has not been preceded by $abab$, leading to us accepting the word. If the first aa is preceded by $abab$, then we reach the state q_6 , and will not accept the word.



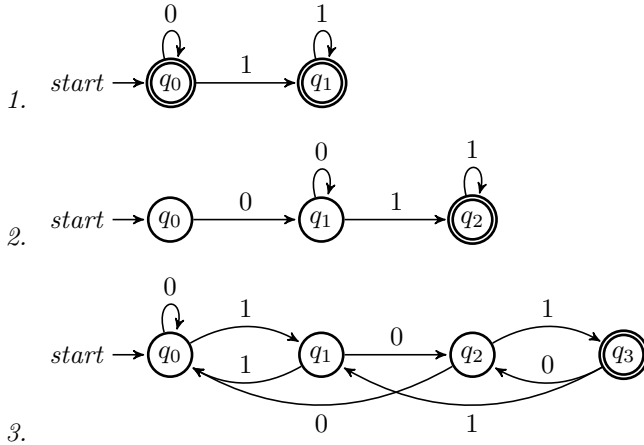
□

Exercise 3 : Binary words.

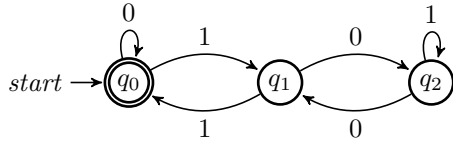
Let $\Sigma = \{0, 1\}$. Construct an automata that accepts the following languages:

1. $L_0 = \{0^m 1^n \mid m, n \geq 0\}$;
2. $L_1 = \{0^m 1^n \mid m, n > 0\}$;
3. $L_{10} = \{w \in \Sigma^* \mid w \text{ ends with } 101\}$;
4. $L_{11} = \{w \in \Sigma^* \mid w \equiv 0 \pmod{3}\}$.

Answer:



4. We will use three states. The state q_i will be reached if the partial word being read is congruent to i modulo 3. When reading a binary number a , we can deduce the class of congruence modulo 3 of $a \cdot 0$ and $a \cdot 1$ depending on the class of a . For example, if $a = 3k$, then $a \cdot 0 = 6k$ and $a \cdot 1 = 6k + 1$. The other two cases work similarly, and we can deduce the following automata:



□

Exercise 4 : Binary operations.

Let L be a language. We define L^R as the language containing all the words in L written in reverse (for example, if L contains $abab$ and $xkcd$ then L^R contains $baba$ and $dckx$).

1. Prove that if L is regular, then L^R is regular.

Let $\Sigma_2 = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. Note that any word $w \in \Sigma_2^+$ can be read as two binary numbers, one for each row. The top number is denoted by w^t and the bottom number by w^b .

2. Prove that the language $L = \{w \in \Sigma_2^+ \mid w^t > w^b\}$ is regular.

Let $\Sigma_3 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$. Again, for every word $w \in \Sigma_3^+$ we get three binary numbers. The top number is w^t , the middle number is w^m and the bottom number is w^b .

3. Prove that the language $L = \{w \in \Sigma_3^+ \mid w^b = w^t + w^m\}$ is regular.

Answer:

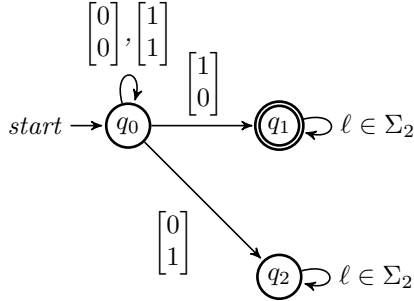
1. Let $A = (Q, \Sigma, \delta, q_0, F)$ be a finite automata that recognizes L . We construct the automata $A^R = (Q^R, \Sigma^R, \delta^R, q_0^R, F^R)$ the following way:

- $\Sigma^R = \Sigma$;
- A^R has all the states of A , that is, $Q^R = Q$;
- For all q, q', ℓ such that $\delta(q, \ell) = q'$, we have $\delta^R(q', \ell) = q$, that is, all the transitions of A are reversed in A^R ;
- $F^R = \{q_0\}$, that is, the starting state of A is an accepting state of A^R ;
- If A has only one accepting state q_f , then it becomes the starting state of A^R : $q_0^R = q_f$;

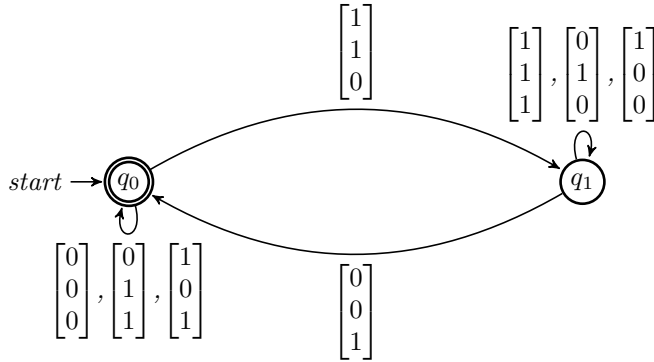
- If A has several accepting states, that is, we have $F = \{q_{f_1}, \dots, q_{f_k}\}$, then we create the starting state q_0^R of A^R , and add transitions $\delta^R(q_0^R, \epsilon) = q_{f_i}$ for all $i \in \{1, \dots, k\}$, where ϵ is the empty symbol.

It is easy to see that A^R recognizes L^R . Since A^R is a finite automata, this implies that L^R is regular (note that A^R may not be deterministic, but this is not a problem since deterministic and non-deterministic automatas recognize the same class of languages).

2. We construct a finite automata that recognizes L . In q_0 , the two numbers are still equal. We enter q_1 when the top number is bigger than the bottom number, and q_2 in the opposite case.



3. We construct an automata that recognizes L^R , which will prove by the point above that L is regular. For this, we simply do a binary addition with carry. In state q_0 , there is no carry, and in state q_1 , there is a carry. Recall that a word is rejected if it is in a state where its character being read does not have a transition arc labelled with this character.



□

Exercise 5 : The Baker Street Irregulars.

Use the pumping lemma to prove that the following languages are not regular:

1. $L_1 = \{a^n b^n \mid n \geq 0\}$;
2. $L_2 = \{a^{n^2} \mid n > 0\}$;
3. $L_3 = \{a^n b^{2n} \mid n > 0\}$;
4. $L_4 = \{(ab)^n c^n \mid n \geq 0\}$.

Use the properties of regular languages to prove that the following language is not regular:

5. $L_5 = \{a^m b^n \mid m \neq n\}$.

Prove in two different ways that the following language is not regular:

6. $L_6 = \{w \in \{a, b\}^* \mid w = w^R\}$ (that is, L is the language of all the palindromes over $\{a, b\}$).

Hint: One method uses the pumping lemma, the other one the fact that an automata accepting L is finite.

Answer: First we recall the pumping lemma:

If L is regular, then there exists an integer p such that, for all $w \in L$ with $|w| \geq p$, we can write $w = xyz$ where $|y| > 0$, $|xy| \leq p$ and $\forall n \geq 0$ $xy^n z \in L$.

1. Assume that L_1 is regular, we apply the pumping lemma. Let $w = a^p b^p = xyz$, and since $|xy| \leq p$ we have $xy = a^{|xy|}$. In particular, let $k = |y|$. Then, we know that for any integer n , $xy^n z = a^{p+(n-1)k} b^p \in L$. However, this implies that $p + (n-1)k = p$, which in turn implies $k = 0$, a contradiction since $|y| > 0$.
2. Assume that L_2 is regular, we apply the pumping lemma. Let $w = a^{p^2} = xyz$. Note that the next word in L is $w_+ = a^{(p+1)^2} = a^{p^2+2p+1}$. Furthermore, since $|xy| \leq p$, we have $|y| \leq p$. Assume that $|y| = p$, the highest possible value. But then $xy^2 z = a^{p^2+p} \notin L$ (since it is longer than w and shorter than w_+). All the other possible values for $|y|$ yield the same issue, leading to a contradiction.
3. Assume that L_3 is regular, we apply the pumping lemma. Let $w = a^p b^p = xyz$. We have $|y| = k > 0$, and as such $xy^0 z = xz = a^{p-k} b^p \notin L$, a contradiction.
4. Assume that L_4 is regular, we apply the pumping lemma. Let $w = (ab)^p c^p = xyz$. There are four possible values for y : $y_1 = (ab)^k$, $y_2 = (ba)^k$, $y_3 = b(ab)^k$ and $y_4 = a(ba)^k$. Now we only need to prove that there exists an integer n such that $xy^n z \notin L$ in each of those four cases, which will lead to a contradiction. For instance, $xy_1^2 z = (ab)^{p-k}(ab)^{2k} c^p \notin L$ since $p-k+2k = k \Rightarrow k = 0$ which is impossible.
5. Assume that L_5 is regular. Let $\overline{L_5} = \{a, b\}^* \setminus L$. Since the regular languages are closed under complementation, $\overline{L_5}$ is regular. Furthermore, let $M = a^* b^*$. It is easy to see that M is regular. Now, the regular languages are closed under intersection, and thus $\overline{L_5} \cap M$ is regular. However, $\overline{L_5} \cap M = L_1$, which is not regular as we saw previously. This contradiction proves that L_5 is not regular.
- 6.1. Using the pumping lemma. Assume that L_6 is regular. Let $w = a^p b a^p = xyz$. Then, we have $y = a^k$ and as such $xy^2 z = a^{p+k} b a^p \notin L$ since $k \neq 0$, which is a contradiction.
- 6.2 Assume by contradiction that L_6 is regular, then there exists a finite automata A that accepts every word of L_6 . Let us consider $w = a^n b a^n$. When reading the word w , A will reach the state q_n after having read $a^n b$. Since A is finite, there exists an integer $m \neq n$ such that $q_m = q_n$, that is, A reaches the state q_n after having read $a^m b$. However, after finishing reading w , A will have reached a terminating state since $w \in L$. But this implies that when A reads the word $w' = a^m b a^n$, it reaches a terminating state and as such accepts the word w' . This is a contradiction, since w' is not a palindrome over $\{a, b\}$ since $m \neq n$.

□