Exercises

Exercise 1

Obtain the position automaton of each one of the following regular expressions:

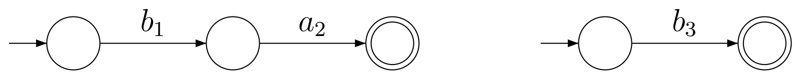
(a) $r = (ba)^*b$

Answer:

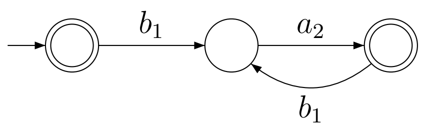
First, we obtain the linearized version of r:

$$\overline{r} = (b_1a_2)^*b_3$$

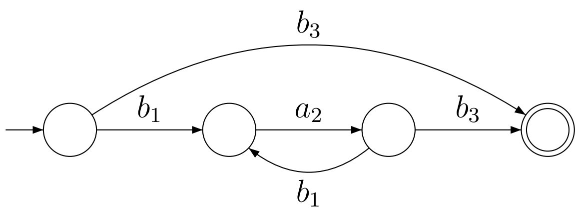
The standard local automata for the subexpressions b_1a_2 and b_3 are:



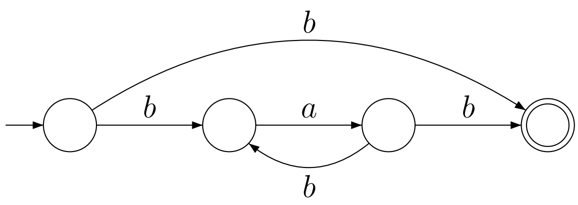
The standard local automaton for the expression $(b_1a_2)^*$ is:



and, the automaton for $\overline{r}$ is:



finally, the position automaton for r can be obtained by deleting the subindexes of the linearized alphabet:



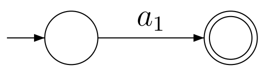
(b) $r = a(a + b)^*$

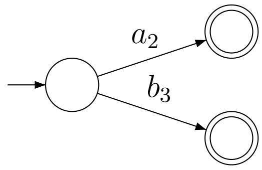
Answer:

First, we obtain the linearized version of r:

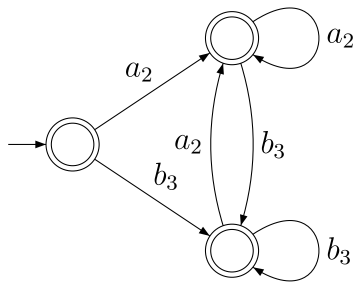
$$\overline{r} = a_1(a_2 + b_3)^*$$

The standard local automata for the subexpressions a_1 and $a_2 + b_3$ are:

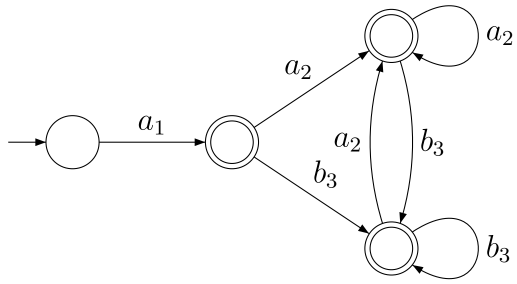




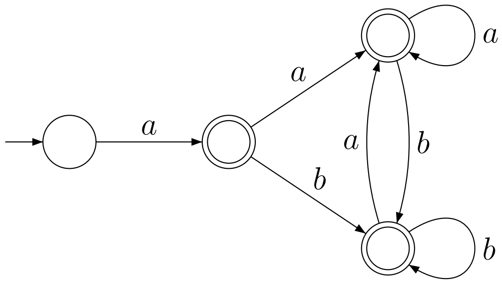
and the standard local automaton for $(a_2 + b_3)^*$ is:



Thus, the following automaton accepts $L(\overline{r})$:



and the position automaton that accepts $L(r)$ is the following one:



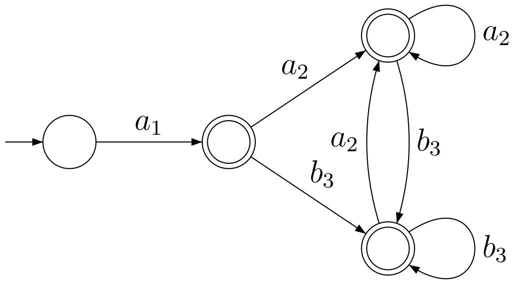
(c) $r = a(a+b)^*b$

Answer:

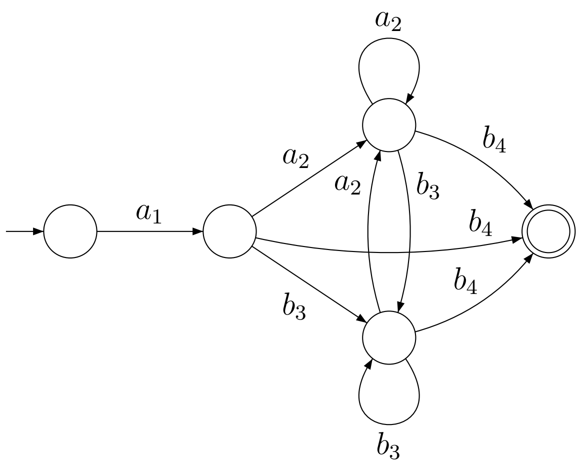
The following expression is the linearized version of r:

$$\overline{r} = a_1(a_2 + b_3)^* b_4$$

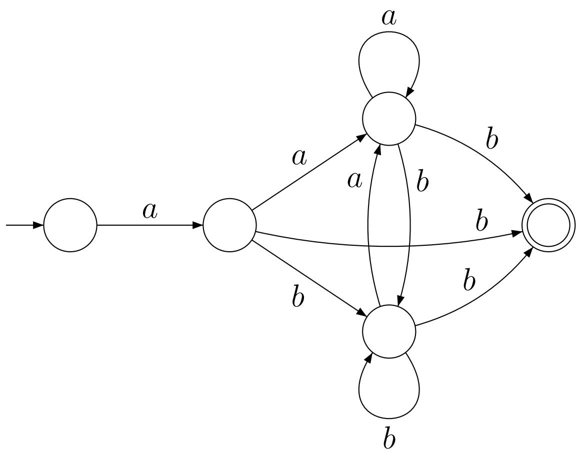
and the automaton that accepts $L(a_1(a_2 + b_3)^*)$ is the following one:



Thus, the following automaton accepts $L(\overline{r})$:



and therefore, the following automaton accepts $L(r)$.

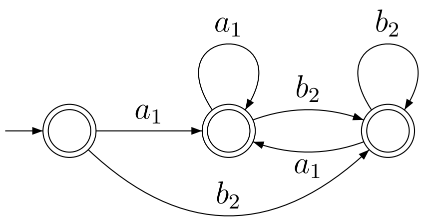


(d) $r = (a^*b^*)^* + (a+b)^*$

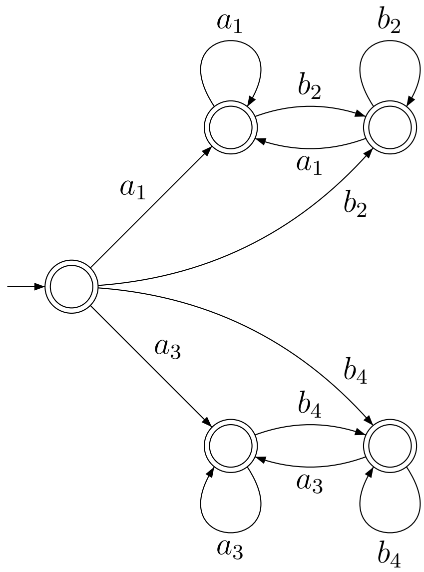
Answer:

$$\overline{r} = (a_1^*b_2^*)^* + (a_3 + b_4)^*$$

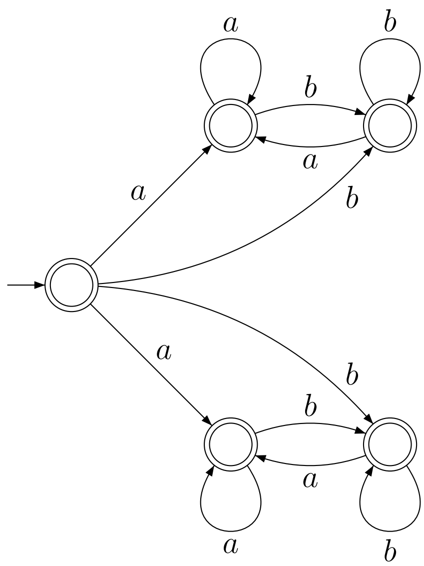
The local standard automaton for the expression $(a_1^*b_2^*)^*$ is:



and the following automaton accepts $L(\overline{r})$:



and finally the automaton that accepts $L(r)$ is:

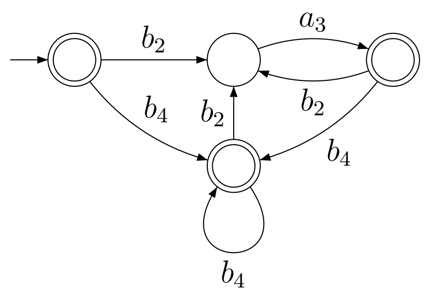


(e) $r = a(ba + b)^*$

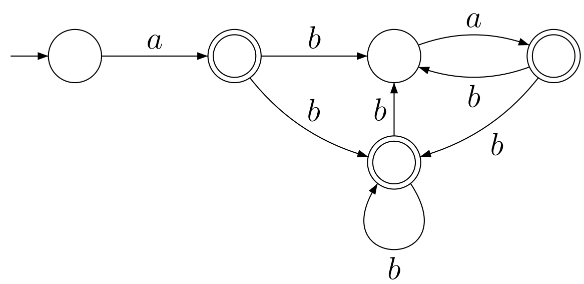
Answer:

$$\overline{r} = a_1(b_2a_3 + b_4)^*$$

The standard local automaton for $(b_2a_3 + b_4)^*$ is shown below:



as well as the position automaton for $L(r)$



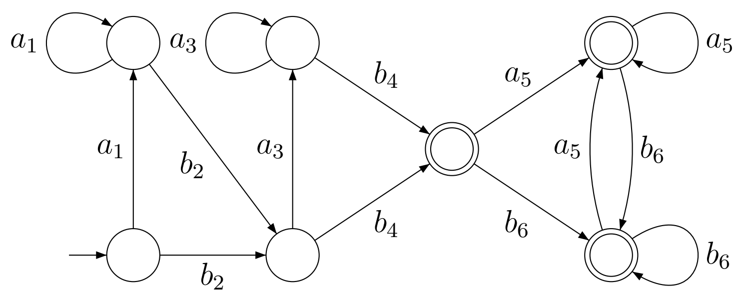
(f) $r = a^*ba^*b(a+b)^*$

Answer:

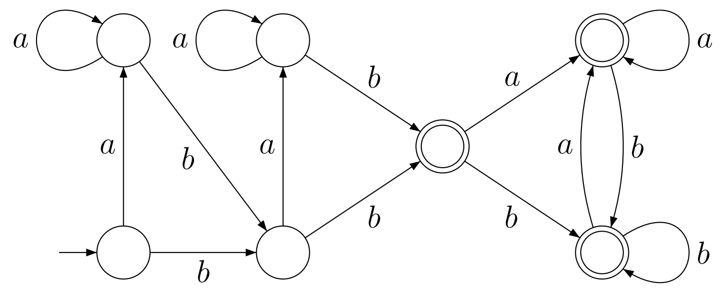
The linearized version of the input regular expression is shown:

$$\overline{r} = a_1^* b_2 a_3^* b_4 (a_5 + b_6)^*$$

as well as the standard local automaton that accepts $L(\overline{r})$:



Once the homomorphism that deletes the subindexes has been applied. the position automaton is the following:

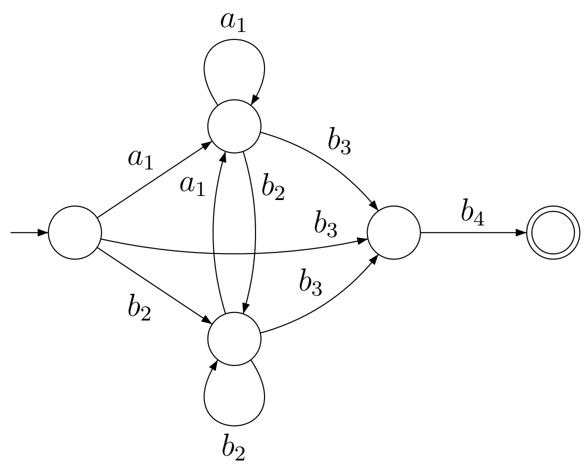


(g) $r = (a+b)^*bb + (a+b)^*a$

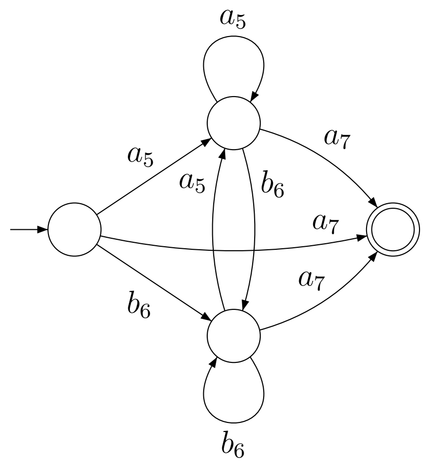
Answer:

$$\overline{r} = (a_1 + b_2)^* b_3 b_4 + (a_5 + b_6)^* a_7$$

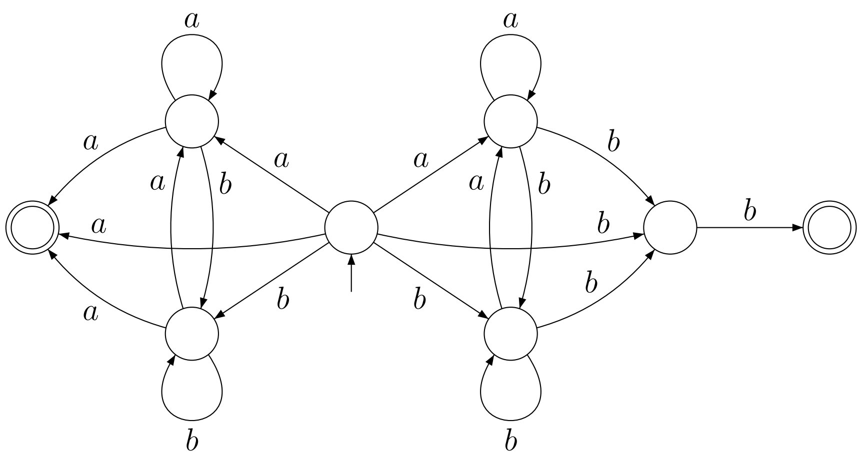
The standard local automaton that accepts $L((a_1 + b_2)^* b_3 b_4)$ is shown below:



as well as the standard local automaton that accepts the language represented by the regular expression $(a_5 + b_6)^* a_7$:



Thus, the position automaton for $L(r)$ is the following one:

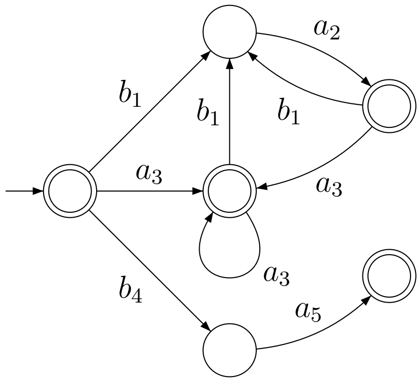


(h) $r = ((ba + a^*)^* + ba)(ab)^*$

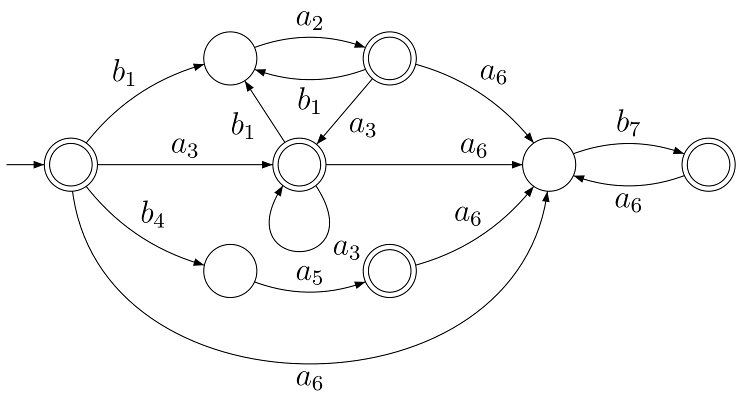
Answer:

$$\overline{r} = ((b_1a_2 + a_3^*)^* + b_4a_5)(a_6b_7)^*$$

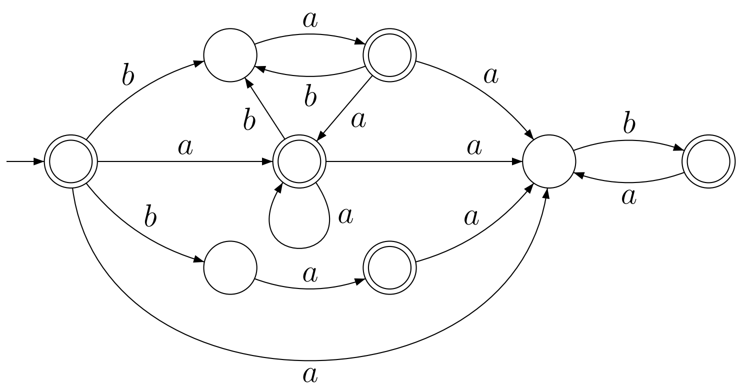
The standard local automaton for $L((b_1a_2 + a_3^*)^* + b_4a_5)$ is the following one:



and thus, the standard local automaton for $L(\overline{r})$ is the following:



therefore, the position automaton that accepts $L(r)$ is the following one:



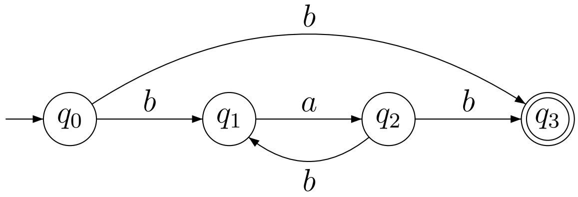
Exercise 2

Obtain the follow automaton for each one of the following regular expressions:

(a) $r = (ba)^*b$

Answer:

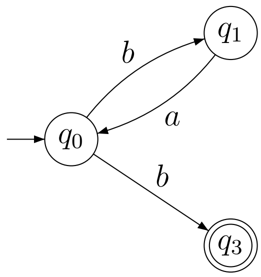
The position automaton is the following one:



The follow relation is summarized in the table below:

Q	$follow$
q_0	$\{q_1, q_3\}$
q_1	$\{q_2\}$
q_2	$\{q_1, q_3\}$
q_3	$\emptyset$

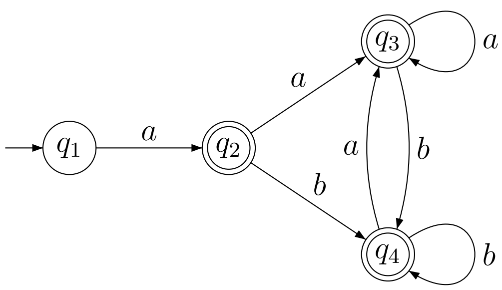
taking into account the membership of each state to the set of final states, the follow automaton is shown:



(b) $r = a(a+b)^*$

Answer:

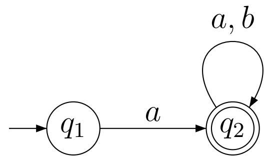
Below it is shown the position automaton that accepts $L(r)$:



the followers of each state are summarized in the following table:

Q	$follow$
q_1	$\{q_2\}$
q_2	$\{q_3, q_4\}$
q_3	$\{q_3, q_4\}$
q_4	$\{q_3, q_4\}$

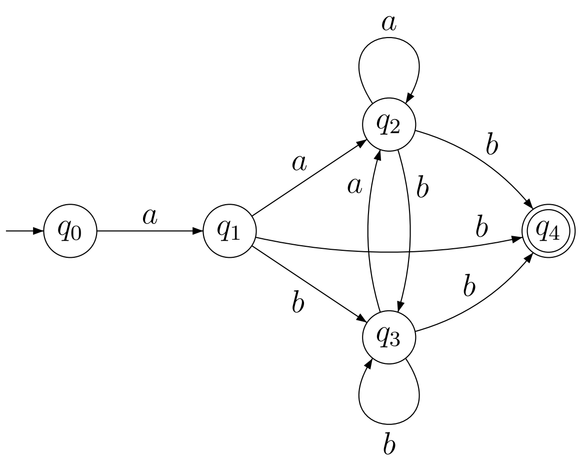
and therefore, the follow automaton for r is:



(c) $r = a(a+b)^*b$

Answer:

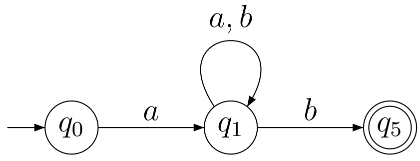
The position automaton for $L(r)$ is shown below:



Next table shows the followers of each state:

Q	$follow$
q_0	$\{q_1\}$
q_1	$\{q_2, q_3, q_4\}$
q_2	$\{q_2, q_3, q_4\}$
q_3	$\{q_2, q_3, q_4\}$
q_4	$\emptyset$

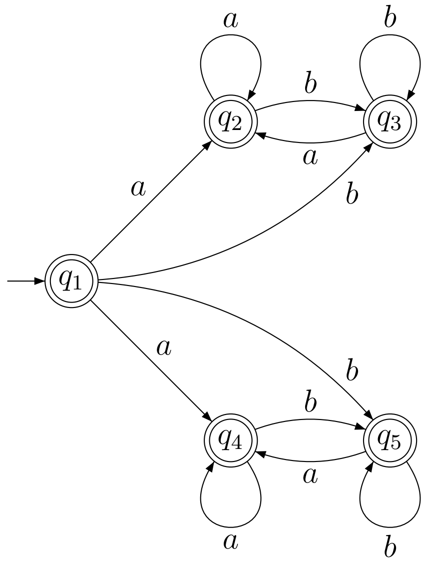
and therefore, taking into account the membership of each state to the set of final states, the follow automaton that accepts $L(r)$ is:



(d) $r = (a^*b^*)^* + (a+b)^*$

Answer:

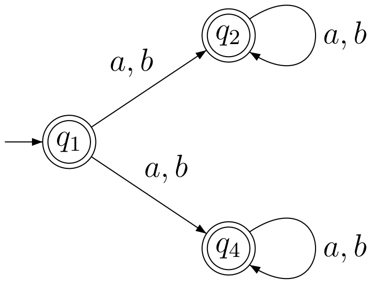
The position automaton for $L(r)$ is shown below:



Next table shows the followers of each state:

Q	$follow$
q_1	$\{q_2, q_3, q_4, q_5\}$
q_2	$\{q_2, q_3\}$
q_3	$\{q_2, q_3\}$
q_4	$\{q_4, q_5\}$
q_5	$\{q_4, q_5\}$

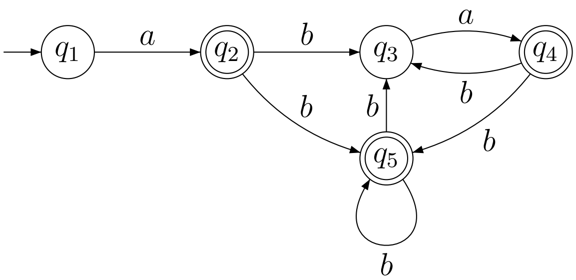
thus, the follow automaton that accepts $L(r)$ is shown.



(e) $r = a(ba + b)^*$

Answer:

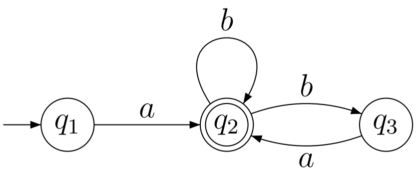
The position automaton for $L(r)$ is shown below:



taking into account the followers of each state:

Q	$follow$
q_1	$\{q_2\}$
q_2	$\{q_3, q_5\}$
q_3	$\{q_4\}$
q_4	$\{q_3, q_5\}$
q_5	$\{q_3, q_5\}$

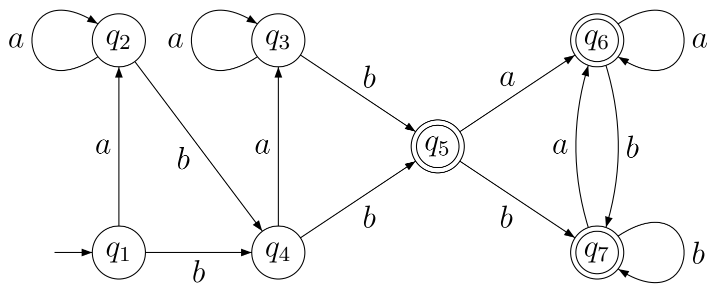
the follow automaton for $L(r)$ is shown below.



(f) $r = a^*ba^*b(a + b)^*$

Answer:

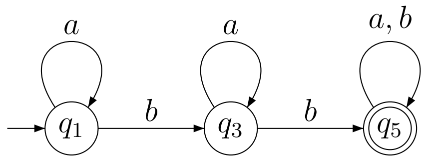
The position automaton for $L(r)$ is shown below:



Next table shows the followers of each state:

Q	$follow$
q_1	$\{q_2, q_4\}$
q_2	$\{q_2, q_4\}$
q_3	$\{q_3, q_5\}$
q_4	$\{q_3, q_5\}$
q_5	$\{q_6, q_7\}$
q_6	$\{q_6, q_7\}$
q_7	$\{q_6, q_7\}$

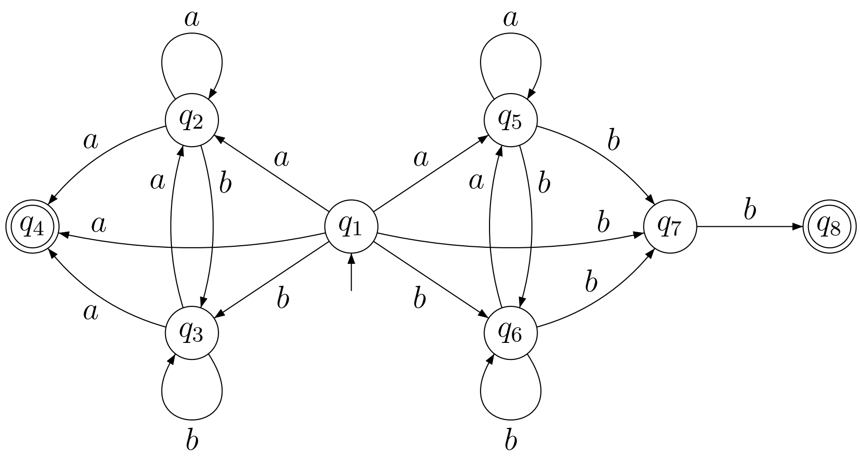
and finally the follow automaton that accepts $L(r)$ is:



(g) $r = (a+b)^*bb + (a+b)^*a$

Answer:

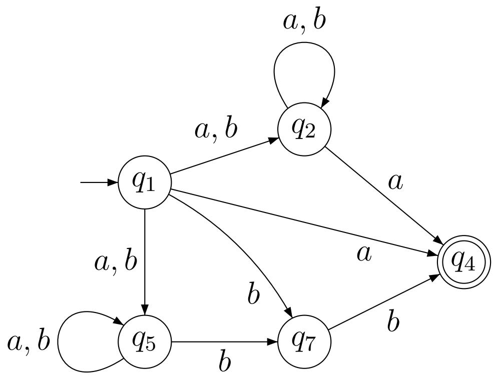
The position automaton for the regular expression is the following one:



The follow relation is summarized in the table below:

Q	$follow$
q_1	$\{q_2, q_3, q_4, q_5, q_6, q_7\}$
q_2	$\{q_2, q_3, q_4\}$
q_3	$\{q_2, q_3, q_4\}$
q_4	$\emptyset$
q_5	$\{q_5, q_6, q_7\}$
q_6	$\{q_5, q_6, q_7\}$
q_7	$\{q_8\}$
q_8	$\emptyset$

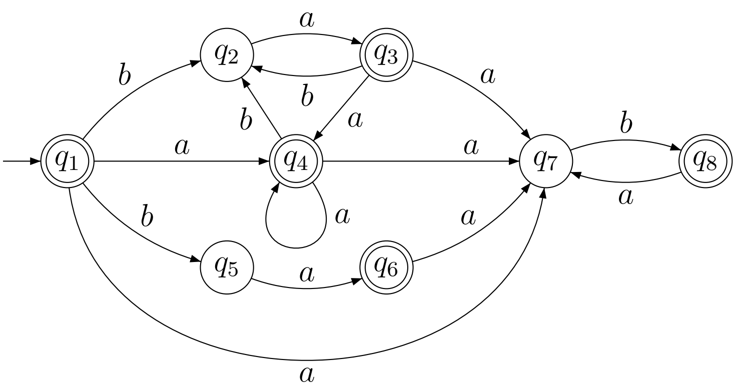
taking into account the membership of each state to the set of final states, the follow automaton is shown:



(h) $r = ((ba + a^*)^* + ba)(ab)^*$

Answer:

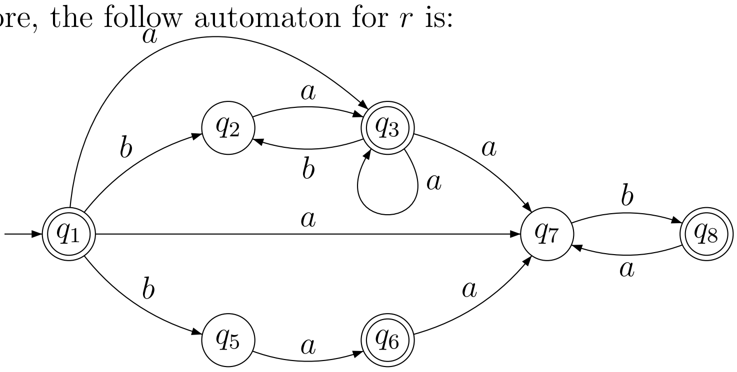
The position automaton for $L(r)$ is the following one:



the followers of each state are summarized in the following table:

Q	$follow$
q_1	$\{q_2, q_4, q_5, q_7\}$
q_2	$\{q_3\}$
q_3	$\{q_2, q_4, q_7\}$
q_4	$\{q_2, q_4, q_7\}$
q_5	$\{q_6\}$
q_6	$\{q_7\}$
q_7	$\{q_8\}$
q_8	$\{q_7\}$

and therefore, the follow automaton for r is:



Exercise 3

Consider the Brzozowski's algorithm to obtain a DFA that accepts the language represented by the following regular expressions.

(a) $r = a(ba + b)^*$

Answer:

First, we set the initial state to r. The initial state is not final because $\lambda \notin L(r)$.

The derivatives of r with respect to each symbol in the alphabet are shown:

$$\begin{aligned} a^{-1}a(ba+b)^* &= (a^{-1}a)(ba+b)^* = \\ &= \lambda(ba+b)^* = (ba+b)^* = r_1 \end{aligned}$$

$$\begin{aligned} b^{-1}a(ba+b)^* &= (b^{-1}a)(ba+b)^* = \\ &= \emptyset(ba+b)^* = \emptyset = r_2 \end{aligned}$$

both expressions denote new languages, therefore, both expressions are considered as new states (r_1 y r_2). New transitions $\delta(r,a) = r_1$ and $\delta(r,b) = r_2$ are also added to the automaton. State r_1 is also added to the set of finals because $\lambda \in L(r_1)$. The derivation process continues as follows:

$$\begin{aligned} a^{-1}r_1 &= a^{-1}(ba+b)^* = (a^{-1}(ba+b))(ba+b)^* = \\ &= (a^{-1}(ba)+a^{-1}b)(ba+b)^* = \emptyset = r_2 \end{aligned}$$

$$\begin{aligned} b^{-1}r_1 &= b^{-1}(ba+b)^* = (b^{-1}(ba+b))(ba+b)^* = \\ &= (b^{-1}(ba)+b^{-1}b)(ba+b)^* = \\ &= (a+\lambda)(ba+b)^* = r_3 \end{aligned}$$

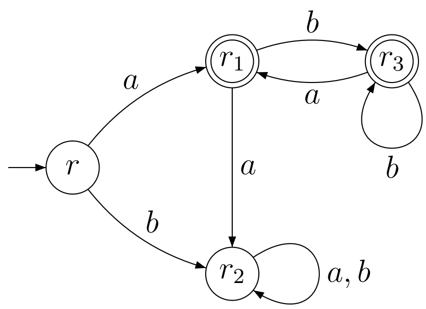
$$a^{-1}r_2 = b^{-1}r_2 = \emptyset = r_2$$

We update accordingly Q, δ and F. The derivatives of r_3 with respect to each symbol are shown below:

$$\begin{aligned} a^{-1}r_3 &= a^{-1}(a+\lambda)(ba+b)^* = (a^{-1}(a+\lambda))(ba+b)^* + (a^{-1}(ba+b)^*) = \\ &= \lambda(ba+b)^* + \emptyset = r_1 \end{aligned}$$

$$\begin{aligned} b^{-1}r_3 &= b^{-1}(a+\lambda)(ba+b)^* = (b^{-1}(a+\lambda))(ba+b)^* + (b^{-1}(ba+b)^*) = \\ &= \emptyset + (b^{-1}(ba+b)^*) = r_3 \end{aligned}$$

No new states appear, thus, the state diagram of the automaton is the following one:



(b) $r = b(ab^*a)^*b$

Answer:

The initial state is set to r. The empty string does not belong to $L(r)$, therefore, the initial state is not final. The derivatives of r with respect to each symbol are shown below.

$$a^{-1}r \;=\; a^{-1}b(ab^*a)^*b \;=\; (a^{-1}b)(ab^*a)^*b \;=\; \emptyset = r_1$$

$$b^{-1}r \;=\; b^{-1}b(ab^*a)^*b \;=\; (b^{-1}b)(ab^*a)^*b \;=\; (ab^*a)^*b = r_2$$

The automaton is updated taking into account the new states found. The set of finals is not updated. The derivation process continues as shown:

$$a^{-1}r_1 \;=\; b^{-1}r_1 \;=\; \emptyset \;=\; r_1$$

$$\begin{aligned} a^{-1}r_2 &= a^{-1}(ab^*a)^*b \;= \\ &= (a^{-1}(ab^*a)^*)b + (a^{-1}b) \;= \\ &= (a^{-1}ab^*a)(ab^*a)^*b + \emptyset \;= \\ &= b^*a(ab^*a)^*b \;=\; r_3 \end{aligned}$$

$$\begin{aligned} b^{-1}r_2 &= b^{-1}(ab^*a)^*b \;= \\ &= (b^{-1}(ab^*a)^*)b + (b^{-1}b) \;= \\ &= \emptyset + \lambda \;=\; \lambda \;=\; r_4 \end{aligned}$$

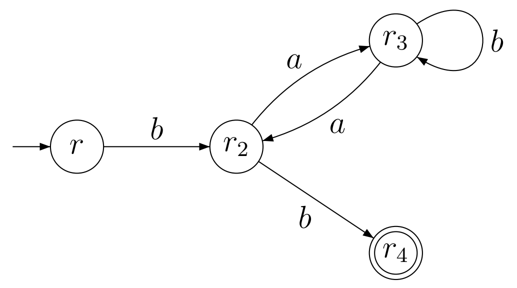
The sets Q, δ y F ($r_4 \in F$) are updated. Now the new states (regular expressions r_3 and r_4) are derived:

$$\begin{aligned} a^{-1}r_3 &= a^{-1}b^*a(ab^*a)^*b \;= \\ &= (a^{-1}b^*)a(ab^*a)^*b + (a^{-1}a(ab^*a)^*b) \;= \\ &= (a^{-1}b)b^*a(ab^*a)^*b + (ab^*a)^*b \;= \\ &= (ab^*a)^*b \;=\; r_2 \end{aligned}$$

$$\begin{aligned} b^{-1}r_3 &= b^{-1}b^*a(ab^*a)^*b \;= \\ &= (b^{-1}b^*)a(ab^*a)^*b + (b^{-1}a(ab^*a)^*b) \;= \\ &= (b^{-1}b)b^*a(ab^*a)^*b + \emptyset \;= \\ &= b^*a(ab^*a)^*b \;=\; r_3 \end{aligned}$$

$$a^{-1}r_4 \;=\; b^{-1}r_4 \;=\; \emptyset \;=\; r_1$$

As the result of this process, the following automaton is obtained.



(c) $r = (ab + b)((aa)^*(a + ba + \lambda))$

Answer:

The initial state corresponds to $\lambda^{-1}r = r$. The initial state is not final because $\lambda \notin L(r)$. The derivatives of r are shown.

$$\begin{aligned} a^{-1}(ab+b)(aa)^*(a+ba+\lambda) &= (a^{-1}(ab+b))(aa)^*(a+ba+\lambda) = \\ &= b(aa)^*(a+ba+\lambda) = r_1 \end{aligned}$$

$$\begin{aligned} b^{-1}(ab+b)(aa)^*(a+ba+\lambda) &= (b^{-1}(ab+b))(aa)^*(a+ba+\lambda) = \\ &= (aa)^*(a+ba+\lambda)) = r_2 \end{aligned}$$

The automaton is updated taking into account the new states and transitions detected. The state r_2 is included into the set of finals and the derivation process continues.

$$\begin{aligned} a^{-1}r_1 &= a^{-1}b(aa)^*(a+ba+\lambda) = (a^{-1}b)(aa)^*(a+ba+\lambda) = \\ &= \emptyset = r_3 \end{aligned}$$

$$\begin{aligned} b^{-1}r_1 &= b^{-1}b(aa)^*(a+ba+\lambda) = (b^{-1}b)(aa)^*(a+ba+\lambda) = \\ &= (aa)^*(a+ba+\lambda) = r_2 \end{aligned}$$

The sets Q, δ and F are updated. The derivatives of r_2 and r_3 are obtained:

$$\begin{aligned} a^{-1}r_2 &= a^{-1}(aa)^*(a+ba+\lambda) = \\ &= (a^{-1}(aa)^*)(a+ba+\lambda) + (a^{-1}(a+ba+\lambda) = \\ &= (a^{-1}aa)(aa)^*(a+ba+\lambda) + \lambda = \\ &= a(aa)^*(a+ba+\lambda) + \lambda = r_4 \end{aligned}$$

$$\begin{aligned} b^{-1}r_2 &= b^{-1}(aa)^*(a+ba+\lambda) = \\ &= (b^{-1}(aa)^*)(a+ba+\lambda) + (b^{-1}(a+ba+\lambda) = \\ &= (b^{-1}aa)(aa)^*(a+ba+\lambda) + a = \\ &= \emptyset + a = a = r_5 \end{aligned}$$

$$a^{-1}r_3 = b^{-1}r_3 = \emptyset = r_3$$

The sets Q, δ and F (r_4) are updated again. The process now considers the derivatives of r_4 and r_5 with respect to the symbols of the alphabet.

$$\begin{aligned} a^{-1}r_4 &= a^{-1}(a(aa)^*(a+ba+\lambda)+\lambda) = \\ &= (a^{-1}a(aa)^*(a+ba+\lambda)) + (a^{-1}\lambda) = \\ &= (aa)^*(a+ba+\lambda)+\emptyset = r_2 \end{aligned}$$

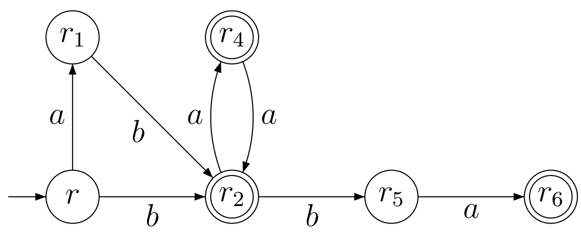
$$\begin{aligned} b^{-1}r_4 &= b^{-1}(a(aa)^*(a+ba+\lambda)+\lambda) = \\ &= (b^{-1}a(aa)^*(a+ba+\lambda)) + (b^{-1}\lambda) = \\ &= \emptyset = r_3 \end{aligned}$$

$$\begin{aligned} a^{-1}r_5 &= \lambda = r_6 \\ b^{-1}r_5 &= \emptyset = r_3 \end{aligned}$$

Finally, the derivatives of r_6 are obtained.

$$a^{-1}r_6 = b^{-1}r_6 = \emptyset = r_3$$

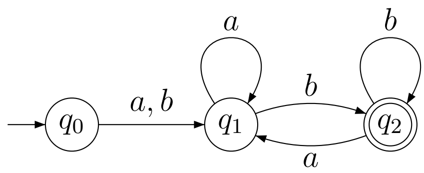
The state diagram of the automaton is shown below:



Exercise 4

Analyze the following automata to obtain a regular expression that represents the same language

(a)



Answer:

The system of equations for the automaton is shown:

$$\begin{cases} X_0 = aX_1 + bX_1 = (a+b)X_1 \\ X_1 = aX_1 + bX_2 \\ X_2 = aX_1 + bX_2 + \lambda \end{cases}$$

We apply Arden's lemma to obtain that $X_2 = b^*(aX_1 + \lambda) = b^*aX_1 + b^*$. This partial result is susbstituted in the system of equations:

$$\begin{cases} X_0 = (a+b)X_1 \\ X_1 = aX_1 + bb^*aX_1 + b^* = (a + bb^*a)X_1 + bb^* \end{cases}$$

Arden's lemma allows to obtain that $X_1 = (a + bb^*a)^*bb^*$. This is substituted in the equation of X_0 and the regular expression for the language is obtained.

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

Hint:

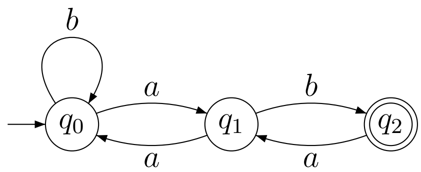
Sometimes it is interesting to try and simplify the regular expressions obtained. In this way, note that, in the exercise, the regular expression obtained for X_1 can be simplified to obtain a more reduced expression:

$$\begin{aligned} X_1 &= (a+bb^*a)^*bb^* = ((\lambda + bb^*)a)^*bb^* = \\ &= (b^*a)^*bb^* = (a+b)^*b^*b = \\ &= (a+b)^*b \end{aligned}$$

which lead to the following expression for X_0

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

(b)



Answer:

The system of equations for the automata is shown:

$$\begin{cases} X_0 = aX_1 + bX_0 \\ X_1 = aX_0 + bX_2 \\ X_2 = aX_1 + \lambda \end{cases}$$

It is possible to substitute directly the expression for X_2 to obtain the following system:

$$\begin{cases} X_0 = aX_1 + bX_0 \\ X_1 = aX_0 + b(aX_1 + \lambda) = aX_0 + baX_1 + b \end{cases}$$

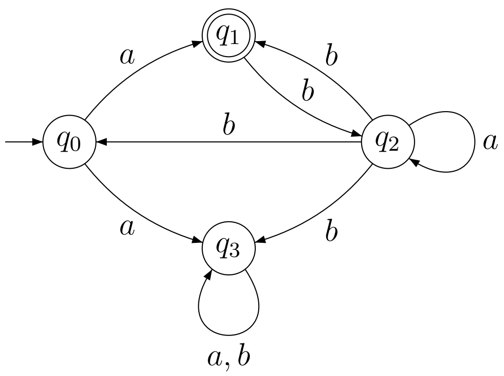
Now it is possible to apply Arden's lemma to obtain $X_1 = (ba)^*(aX_0 + b)$, once this expression is substituted, it is obtained:

$$\begin{aligned} X_0 &= a(ba)^*(aX_0 + b) + bX_0 = \\ &= a(ba)^*aX_0 + a(ba)^*b + bX_0 = \\ &= (a(ba)^*a + b)X_0 + a(ba)^*b \end{aligned}$$

Arden's lemma is applied once more time to obtain the final expression:

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

(c)



Answer:

System of equations for the automaton:

$$\begin{cases} X_0 = aX_1 + aX_3 \\ X_1 = bX_2 + \lambda \\ X_2 = bX_0 + bX_1 + aX_2 + bX_3 \\ X_3 = (a + b)X_3 \end{cases}$$

$X_3 = (a + b)^*\emptyset = \emptyset$ can be obtained by Arden's lemma, that leads to the following simplification of the system of equations:

$$\begin{cases} X_0 = aX_1 \\ X_1 = bX_2 + \lambda \\ X_2 = bX_0 + bX_1 + aX_2 \end{cases}$$

it is possible to apply Arden's lemma to obtain:

$$X_2 = a^*(bX_0 + bX_1) = a^*bX_0 + a^*bX_1$$

that can be substitured in the system:

$$\begin{cases} X_0 = aX_1 \\ X_1 = b(a^*bX_0 + a^*bX_1) + \lambda \ = \ ba^*bX_0 + ba^*bX_1 + \lambda \end{cases}$$

Arden is applied once more time:

$$X_1 \ = \ (ba^*b)^*(ba^*bX_0 + \lambda) \ = \ ba^*b(ba^*b)^*X_0 + (ba^*b)^*$$

thus:

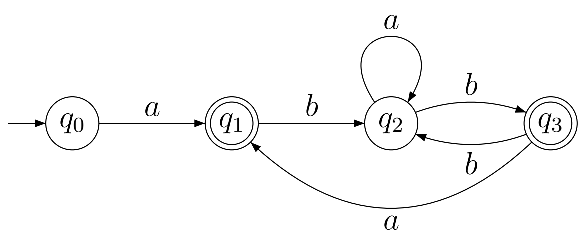
$$X_0 \ = \ aba^*b(ba^*b)^*X_0 + a(ba^*b)^*$$

and finally, by applying again Arden's lemma:

$$X_0 \ = \ (aba^*b(ba^*b)^*)^*a(ba^*b)^*$$

a expression that represent the language is obtained.

(d)



Answer:

System of equations for the automaton:

$$\begin{cases} X_0 = aX_1 \\ X_1 = bX_2 + \lambda \\ X_2 = aX_2 + bX_3 \\ X_3 = bX_2 + aX_1 + \lambda \end{cases}$$

the expression for X_3 can be substituted in the system:

$$\begin{cases} X_0 = aX_1 \\ X_1 = bX_2 + \lambda \\ X_2 = aX_2 + b(aX_1 + bX_2 + \lambda) \;=\; baX_1 + (a + bb)X_2 + b \end{cases}$$

as well as the expression for X_1:

$$\begin{cases} X_0 = a(bX_2 + \lambda) \;=\; abX_2 + a \\ X_2 = ba(bX_2 + \lambda) + (a + bb)X_2 + b \;=\; (a + bab + bb)X_2 + b + ba \end{cases}$$

Arden's lemma can be applied to obtain $X_2 = (a+bab+bb)^*(b+ba)$. Once this expression is substituted, a expression that represents the language is obtained:

$$ab(a + bab + bb)^*(b + ba) + a$$