Automata and Formal Languages — Homework 3

Due 07.11.2017

Exercise 3.1
Let $M_n = \{w \in \{0,1\}^* : \text{msbf}(w) \text{ is a multiple of } n\}$ (see Exercise #1.2) and let $L_{\text{pal}} = \{w \in \Sigma^* : w = w^R\}$ where Σ is some finite alphabet.

(a) Show that M_3 has (exactly) three residuals, i.e. show that $|\{(M_3)^w : w \in \{0,1\}^*\}| = 3$.

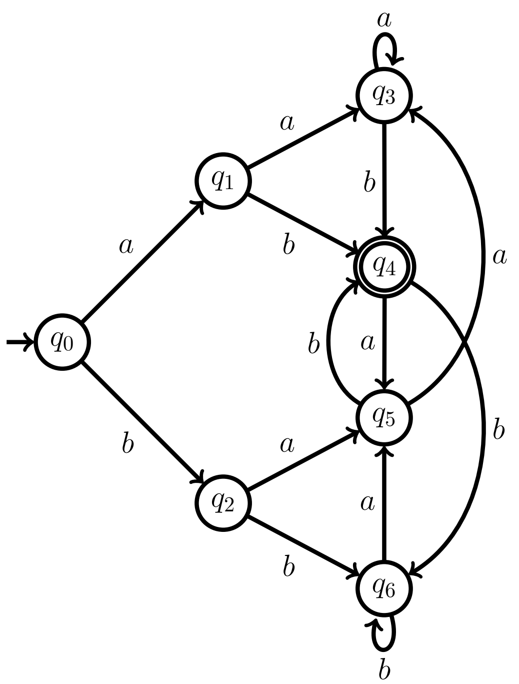
(b) Show that M_4 has less than four residuals.

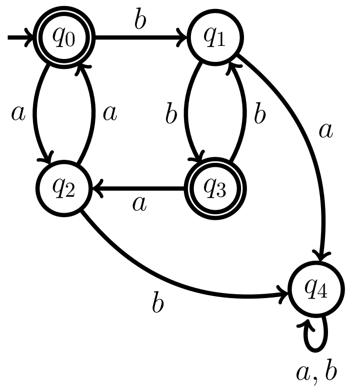
(c) ★ Show that M_p has (exactly) p residuals for every prime number p. You may use the fact that, by Fermat's little theorem, $2^{p-1} \equiv 1 \pmod{p}$. [Hint:]

(d) Show that L_{pal} has infinitely many residuals whenever $|\Sigma| \geq 2$.

(e) Show that L_{pal} is regular for $\Sigma = \{a\}$. Is L_{pal} also regular for larger alphabets?

Exercise 3.2
Let A and B be respectively the following DFAs:





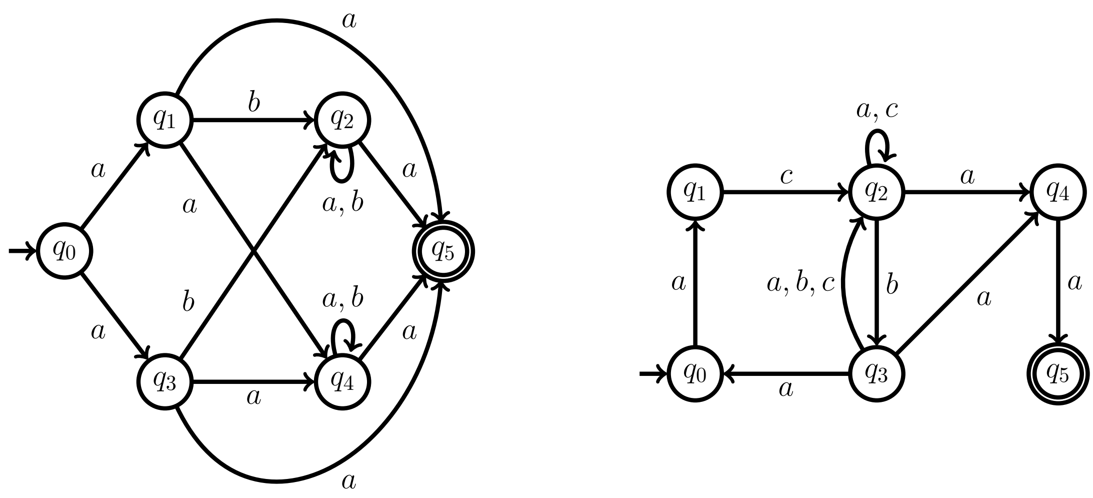
(a) Compute the language partitions of A and B.

(b) Construct the quotients of A and B with respect to their language partitions.

(c) Give regular expressions for $L(A)$ and $L(B)$.

Exercise 3.3

Let A and B be respectively the following NFAs:



(a) Compute the coarsest stable refinements (CSR) of A and B.

(b) Construct the quotients of A and B with respect to their CSRs.

(c) Show that

$$L(A) = \{w \in \{a, b\}^* : w \text{ starts and ends with } a\}$$
$$L(B) = \{w \in \{a, b\}^* : w \text{ starts with } ac \text{ and ends with } ab\}$$

(d) Are the automata obtained in (b) minimal?

Solution 3.1

(a) In exercise #1.2(c), we have seen a DFA with three states that accepts M_3. Therefore, M_3 has at most three residuals. We claim that M_3 has *at least* three residuals. To prove this claim, it suffices to show that the ε-residual, 1-residual and 10-residual of M_3 are distinct. This holds since:

$$\begin{array}{ccc} \varepsilon \cdot \varepsilon \in M_3, & \varepsilon \cdot \varepsilon \in M_3, & 1 \cdot 1 \in M_3, \\ 1 \cdot \varepsilon \notin M_3, & 10 \cdot \varepsilon \notin M_3, & 10 \cdot 1 \notin M_3. \end{array} \qquad \square$$

(b) In exercise #1.2(b), we have seen a DFA with three states that accepts M_4. Therefore, M_4 has at most three residuals. $\square$

(c) In exercise #1.2(g), we have seen a DFA with p states that accepts M_p. Therefore, M_p has at most p residuals. It remains to show that M_p has at least p residuals. For every $0 \leq i < p$, let u_i be the word such that $|u_i| = p - 1$ and $\mathrm{msbf}(u_i) = i$. Note that u_i exists since the smallest encoding of i has at most $p-1$ bits, and it can be extended to length $p-1$ by padding with zeros on the left. Let us show that the u_i-residual and u_j-residual of M_p are distinct for every $0 \leq i, j < p$ such that $i \neq j$. Let $0 \leq k < p$, and let $\ell = (p - i) \bmod p$. We have:

$$\begin{aligned} \mathrm{msbf}(u_k u_\ell) &= 2^{|u_\ell|} \cdot \mathrm{msbf}(u_k) + \mathrm{msbf}(u_\ell) \\ &= 2^{p-1} \cdot k + ((p-i) \bmod p) \\ &\equiv k + ((p-i) \bmod p) && \text{(by Fermat's little theorem)} \\ &\equiv k + p - i \\ &\equiv k - i. \end{aligned}$$

Let $0 \leq i, j < p$ be such that $i \neq j$. We have $u_i u_\ell \in M_p$ since $\mathrm{msbf}(u_i u_\ell) \equiv i - i \equiv 0$, but we have $u_j u_\ell \notin M_p$ since $\mathrm{msbf}(u_j u_\ell) \equiv j - i \not\equiv 0$. Therefore, the u_i-residual and u_j-residual of M_p are distinct. $\square$

(d) Without loss of generality, we may assume that $a, b \in \Sigma$. For every $i \in \mathbb{N}$, let $u_i = a^i b$. Let $i, j \in \mathbb{N}$ be such that $i \neq j$. We claim that the u_i-residual and the u_j-residual of L_{pal} differ. This shows that L_{pal} has infinitely many residuals. To prove the claim, observe that $u_i a^i \in L_{\text{pal}}$ and that $u_j a^i \notin L_{\text{pal}}$.

★ To see why $u_j a^i \notin L_{\text{pal}}$, assume for the sake of contradiction that $u_j a^i \in L_{\text{pal}}$. Let $w = u_j a^i$. Since w is a palindrome, it must be the case that $w_{j+1} = b = w_{|w|-(j+1)+1}$. In particular, since w contains only a single b, we must have $|w| - (j+1) + 1 = j + 1$. This yields a contradiction since

$$\begin{aligned} |w| - (j+1) + 1 &= (i + j + 1) - (j + 1) + 1 \\ &= i + 1 \\ &\neq j + 1 && \text{(by } i \neq j\text{).} \end{aligned} \qquad \square$$

(e) If $\Sigma = \{a\}$, then $L_{\text{pal}} = \Sigma^*$ since every word is trivially a palindrome. Thus, L_{pal} is accepted by a DFA with a single state. If $|\Sigma| > 1$, then by (d) we know that L_{pal} has infinitely many residuals. A language is regular if and only if it has finitely many residuals, and hence L_{pal} is not regular. $\square$

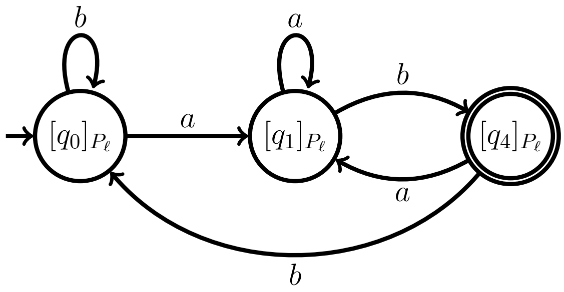
Solution 3.2

A) (a)

Iter.	**Block to split**	**Splitter**	**New partition**
0	—	—	$\{q_0, q_1, q_2, q_3, q_5, q_6\}, \{q_4\}$
1	$\{q_0, q_1, q_2, q_3, q_5, q_6\}$	$(b, \{q_4\})$	$\{q_0, q_2, q_6\}, \{q_1, q_3, q_5\}, \{q_4\}$
2	none, partition is stable	—	—

The language partition is $P_\ell = \{\{q_0, q_2, q_6\}, \{q_1, q_3, q_5\}, \{q_4\}\}$.

(b)



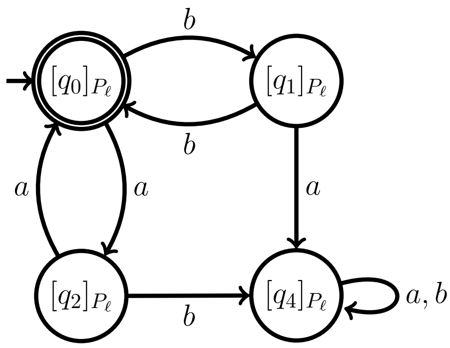
(c) $(a+b)^*ab$.

B) (a)

Iter.	**Block to split**	**Splitter**	**New partition**
0	—	—	$\{q_0, q_3\}, \{q_1, q_2, q_4\}$
1	$\{q_1, q_2, q_4\}$	$(b, \{q_1, q_2, q_4\})$	$\{q_0, q_3\}, \{q_1\}, \{q_2, q_4\}$
2	$\{q_2, q_4\}$	$(a, \{q_0, q_3\})$	$\{q_0, q_3\}, \{q_1\}, \{q_2\}, \{q_4\}$
3	none, partition is stable	—	—

The language partition is $P_\ell = \{\{q_0, q_3\}, \{q_1\}, \{q_2\}, \{q_4\}\}$.

(b)



(c) $(aa+bb)^*$ or $((aa)^*(bb)^*)^*$.

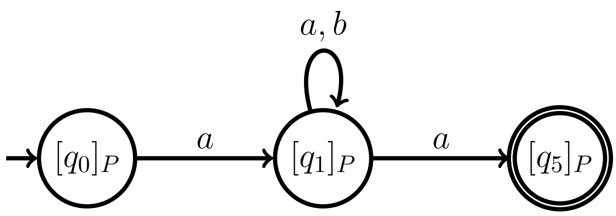
Solution 3.3

A) (a)

Iter.	**Block to split**	**Splitter**	**New partition**
0	—	—	$\{q_0, q_1, q_2, q_3, q_4\}, \{q_5\}$
1	$\{q_0, q_1, q_2, q_3, q_4\}$	$(a, \{q_5\})$	$\{q_0\}, \{q_1, q_2, q_3, q_4\}, \{q_5\}$
2	none, partition is stable	—	—

The CSR is $P = \{\{q_0\}, \{q_1, q_2, q_3, q_4\}, \{q_5\}\}$.

(b)



(c) It follows immediately from the fact that A accepts the same language as the automaton obtained in (b).

(d) Yes. By (c), the language accepted by A is $a(a+b)^*a$. An NFA with one state can only accept $\emptyset, \{\varepsilon\}, a^*, b^*$ and $\{a,b\}^*$. Suppose there exists an NFA $A' = (\{q_0, q_1\}, \{a,b\}, \delta, Q_0, F)$ accepting $L(A)$. Without loss of generality, we may assume that q_0 is initial. A' must respect the following properties:

- $q_0 \notin F$, since $\varepsilon \notin L(A)$,
- $q_1 \in F$, since $L(A) \neq \emptyset$,
- $q_1 \notin Q_0$, since $\varepsilon \notin L(A)$,
- $q_1 \in \delta(q_0, a)$, otherwise it is impossible to accept aa which is in $L(A)$.

This implies that A' accepts a, yet $a \notin L(A)$. Therefore, no two states NFA accepts $L(A)$. $\square$

B) (a)

Iter.	**Block to split**	**Splitter**	**New partition**
0	—	—	$\{q_0, q_1, q_2, q_3, q_4\}, \{q_5\}$
1	$\{q_0, q_1, q_2, q_3, q_4\}$	$(a, \{q_5\})$	$\{q_0, q_1, q_2, q_3\}, \{q_4\}, \{q_5\}$
2	$\{q_0, q_1, q_2, q_3\}$	$(a, \{q_4\})$	$\{q_0, q_1\}, \{q_2, q_3\}, \{q_4\}, \{q_5\}$
3	$\{q_0, q_1\}$	$(c, \{q_2, q_3\})$	$\{q_0\}, \{q_1\}, \{q_2, q_3\}, \{q_4\}, \{q_5\}$
4	$\{q_2, q_3\}$	$(a, \{q_0\})$	$\{q_0\}, \{q_1\}, \{q_2\}, \{q_3\}, \{q_4\}, \{q_5\}$

The CSR is $P = \{\{q_0\}, \{q_1\}, \{q_2\}, \{q_3\}, \{q_4\}, \{q_5\}\}$.

(b) The automaton remains unchanged.

(c) $\subseteq$) Let $w \in L(B)$. Every path from q_0 to q_5 first goes through q_1 and q_2 and ends up going through q_4 and q_5. This implies that $w \in L(ac(a+b+c)^*ab)$.

$\supseteq$) First note that for every $u \in \{a,b,c\}^*$, there exists $q \in \{q_2, q_3\}$ such that $q_2 \xrightarrow{u} q$. This can be shown by induction on $|u|$. Let $w \in L(ac(a+b+c)^*ab)$. There exists $u \in \{a,b,c\}^*$ such that $w = acuab$. Let $q \in \{q_2, q_3\}$ be such that $q_2 \xrightarrow{u} q$. We have $q_0 \xrightarrow{a} q_1 \xrightarrow{c} q_2 \xrightarrow{u} q \xrightarrow{a} q_4 \xrightarrow{b} q_5$. Therefore, $w \in L(B)$. $\square$

(d) No. We have seen a DFA with five states accepting the same language in Exercise #1.1.