Technische Universität München **Winter term 2018/19**
I7
Prof. J. Esparza / S. Sickert **31.10.2018**

Automata and Formal Languages — Homework 3

Due 06.11.2018

Exercise 3.1
Prove or disprove:

(a) A subset of a regular language is regular.

(b) A superset of a regular language is regular.

(c) If L_1 and L_1L_2 are regular, then L_2 is regular.

(d) If L_2 and L_1L_2 are regular, then L_1 is regular.

Exercise 3.2
Let $M_n = \{w \in \{0,1\}^* \mid \mathrm{msbf}(w) \text{ is a multiple of } n\}$ and let $L_{\mathrm{pal}} = \{w \in \Sigma^* \mid w \text{ is a palindrome}\}$ where Σ is some finite alphabet.

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

(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, $2^{p-1} \equiv 1 \pmod{p}$ for all prime numbers $p > 2$. [Hint:]

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

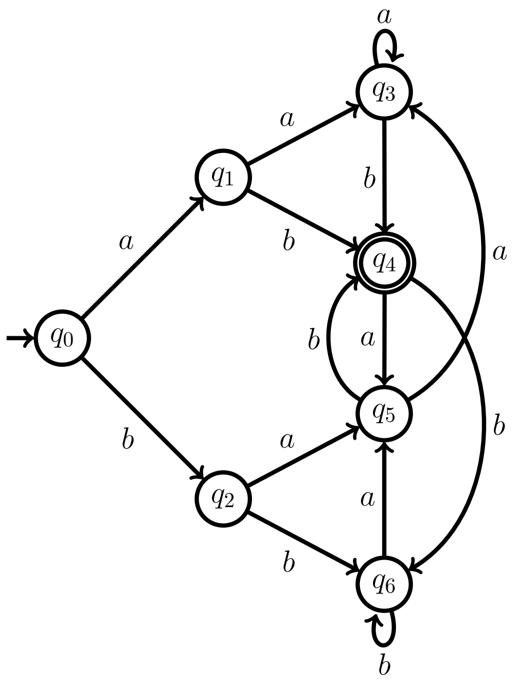
(e) Show that L_{pal} is regular for $\Sigma = \{a\}$. Is L_{pal} also regular for larger alphabets?

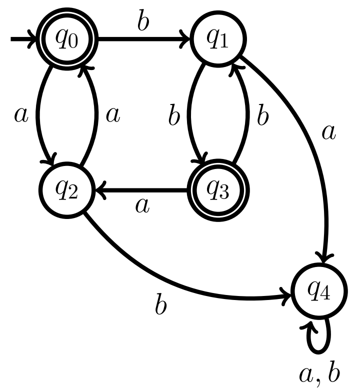
Exercise 3.3

(a) Let $\Sigma = \{0,1\}$ be an alphabet.
Find a language $L \subseteq \Sigma^*$ that has infinitely many residuals and $|L^w| > 0$ for all $w \in \Sigma^*$.

(b) Let $\Sigma = \{a\}$ be an alphabet.
Find a language $L \subseteq \Sigma^*$, such that $L^w = L^{w'} \implies w = w'$ for all words $w, w' \in \Sigma^*$.
What can you say about the residuals for such a language L? Is such a language regular?

Exercise 3.4
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 3.5

Design an efficient algorithm $Res(r, a)$, where r is a regular expression over an alphabet Σ and $a \in \Sigma$, that returns a regular expression satisfying $L\big(Res(r, a)\big) = \big(L(r)\big)^a$. Extend your approach to arbitrary words $w \in \Sigma^*$.

Solution 3.1

All statements are false. Since $\emptyset$ and Σ^* are both regular, any of the first two statements would imply that every language is regular, which is certainly not the case. For the third statement, take $L_1 = a^*$ and take for L_2 any non-regular language over $\{a\}$ (for instance, $L_2 = \{a^{n^2} \mid n \geq 0\}$). Then $L_1 L_2 = a^*$, which is regular. For the fourth statement, take $L_1 = \{a^{n^2} \mid n \geq 0\}$ and $L_2 = a^*$.

Solution 3.2

(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:

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

(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. □

(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. □

(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

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

(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. □

Solution 3.3

(a) $L = \{ww \mid w \in \Sigma^*\}$. First we prove that L has infinitely many residuals by showing that for each pair of words of the infinite set $\{0^i 1 \mid i \geq 0\}$ the corresponding residuals are not equal. Let $u = 0^i 1, v = 0^j 1 \in \Sigma^*$ two words with $i < j$. Then $L^u \neq L^v$ since $u \in L^u$, but $u \not L^v$. For the second half consider some arbitrary word w. Then $w \in L^w$, which shows the statement.

(b) We observe that for all languages satisfying that property L^w has to be non-empty for all w and thus also infinite. Furthermore all these languages are not regular, since there are infinitely many residuals.

$L = \{a^{2^n} \mid n \geq 0\}$. Let a^i and a^j two distinct words. W.l.o.g. we assume $i < j$. Let now d_i and d_j denote the distance from i and j to resp. closest larger square number. If $d_i < d_j$ holds, we are immediately done since $a^{d_i} \in L^{a^i}$ and $a^{d_i} \notin L^{a^j}$. $d_i > d_j$ is analogous. Thus assume $d_i = d_j$. Let us then define d'_i and d'_j denote the distance from i and j to resp. second closest larger square number. These have to be unequal, since the gaps between the square numbers are strictly increasing and we can repeat the argument from before.

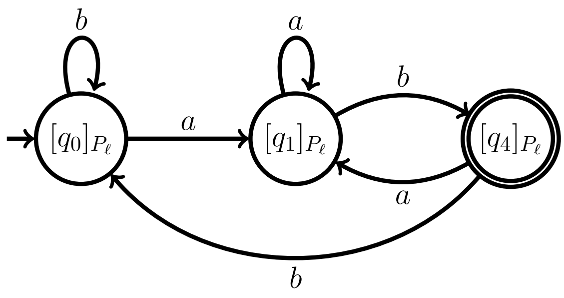
Solution 3.4

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)



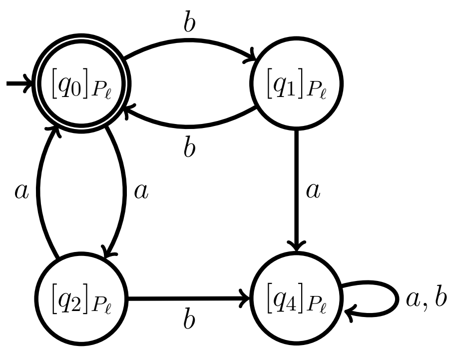
(c) $(a+b)^*ab$.

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 3.5

The solution to Exercise ... yields a linear-time algorithm to check if the language of a regular expression contains the empty word. We can easily transform it into an algorithm computing the function $E(r)$ defined by $E(r) = \varepsilon$ if $\varepsilon \in L(r)$, and $E(r) = \emptyset$ otherwise. Now we can define the function $Res(r, a)$ recursively as follows:

- $Res(\emptyset, a) = Res(\varepsilon, a) = \emptyset$;
- $Res(r_1 + r_2, a) = Res(r_1, a) + Res(r_2, a)$;
- $Res(r_1 r_2, a) = Res(r_1, a)\, r_2 + E(r_1)\, Res(r_2, a)$;
- $Res(r^*, a) = Res(r)\, r^*$.