Prof. Javier Esparza
Philipp Czerner

Technical University of Munich
Chair for Foundations of Software Reliability

Automata and Formal Languages

Winter Term 2023/24 – Exercise Sheet 3

Exercise 3.1.

Analyse the residuals of the following languages. If there are finitely many of them, determine them; otherwise prove that there are infinitely many of them.

(a) $(a + bbc)^*$ over $\Sigma = \{a, b, c\}$,

(b) $(aa)^*$ over $\Sigma = \{a, b\}$,

(c) $\{a^n b^{n+1} \mid n \geq 0\}$ over $\Sigma = \{a, b\}$,

(d) $\{a^{2^n} \mid n \geq 0\}$ over $\Sigma = \{a\}$.

Solution.

(a) For $(a + bbc)^*$. We give the residuals as regular expressions: $(a + bbc)^*$ (residual with respect to a); $bc(a + bbc)^*$ (residual with respect to b); $c(a + bbc)^*$ (residual with respect to bb); $\emptyset$ (residual with respect to c). All other residuals are equal to one of these four.

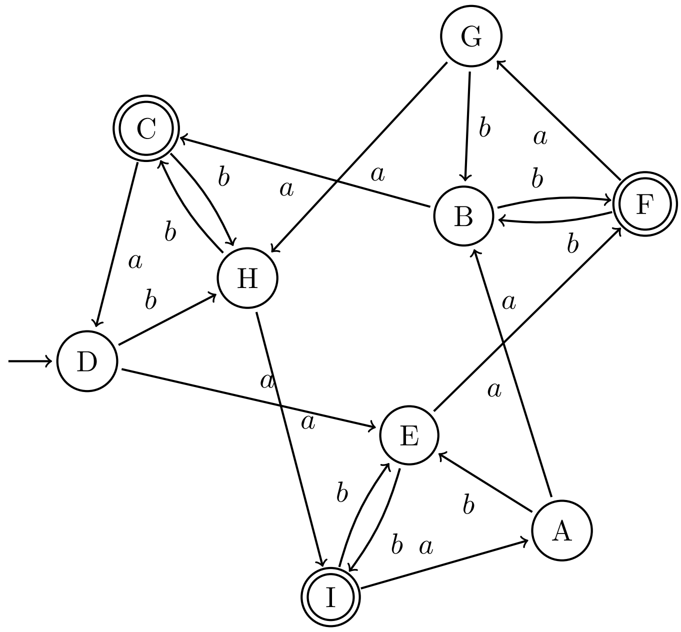
(b) For $(aa)^*$. We give the residuals as regular expressions: $(aa)^*$ (residual of ε); $a(aa)^*$ (residual of a); $\emptyset$ (residual of b). All other residuals are equal to one of these three.

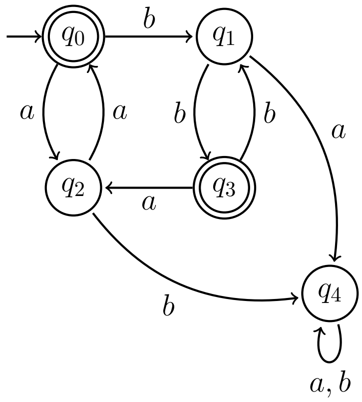
(c) For $\{a^n b^{n+1} \mid n \geq 0\}$. Note that for any $0 \leq i < j$, $a^i b^{i+1}$ belongs to the language, but $a^j b^{i+1}$ does not belong to the language. This implies that a^i and a^j have different residuals and so there are infinitely many residuals.

(d) For $\{a^{2^n} \mid n \geq 0\}$. Note that for any $0 \leq i < j$, $a^{2^i} a^{2^i}$ belongs to the language because $2^i + 2^i = 2^{i+1}$, but $a^{2^i} a^{2^j}$ does not belong to the language because $2^j < 2^i + 2^j < 2^j + 2^j = 2^{j+1}$. This implies that a^{2^i} and a^{2^j} have different residuals and so there are infinitely many residuals.

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 $L(B)$.

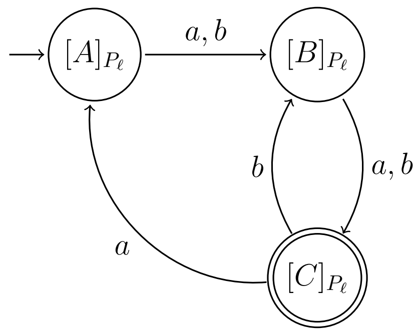
Solution.

(a) (1)

Iter.	**Block to split**	**Splitter**	**New partition**
0	—	—	$\{C, F, I\}, \{A, B, D, E, G, H\}$
1	$\{A, B, D, E, G, H\}$	$(b, \{A, B, D, E, G, H\})$	$\{C, F, I\}, \{B, E, H\}, \{A, D, G\}$
3	none, partition is stable	—	—

The language partition is $P_\ell = \{\{A, D, G\}, \{B, E, H\}, \{C, F, I\}\}$.

(2) The minimal automaton is given below:



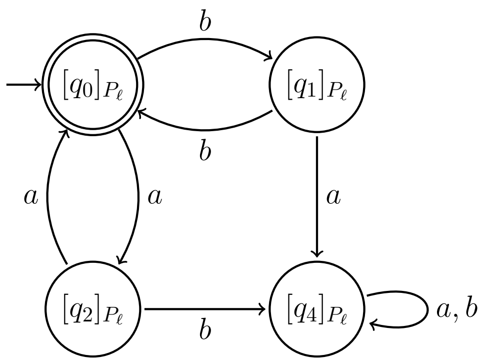
(3) $\Sigma^2(a\Sigma^2 + b\Sigma)^*$

(b) (1)

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 $P_\ell = \{\{q_0, q_3\}, \{q_1\}, \{q_2\}, \{q_4\}\}$.

(2) The minimal automaton is given below:



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

Exercise 3.3.

Given $n \in \mathbb{N}$, let $\text{MSBF}(n)$ be the set of *most-significant-bit-first* encodings of n, i.e., the words that start with an arbitrary number of leading zeros, followed by n written in binary. For example:

$$\text{MSBF}(3) = 0^*11 \quad \text{and} \quad \text{MSBF}(9) = 0^*1001 \quad \text{MSBF}(0) = 0^*$$

Similarly, let $\text{LSBF}(n)$ denote the set of *least-significant-bit-first* encodings of n, i.e., the set containing for each word $w \in \text{MSBF}(n)$ its reverse. For example:

$$\text{LSBF}(6) = 0110^* \quad \text{and} \quad \text{LSBF}(0) = 0^*$$

For any $n \geq 2$, let $M_n = \{w \in \{0,1\}^* \mid w \in \text{MSBF}(k) \text{ and } k \text{ is a multiple of } n\}$ and $L_n = \{w \in \{0,1\}^* \mid w \in \text{LSBF}(k) \text{ and } k \text{ is a multiple of } n\}$.

In the following, let $p > 2$ be any prime number.

(a) Prove that M_p and L_p have at least p many residuals.

(b) Give the minimal DFA A_p (with p states) for the language M_p.

(c) Prove that the NFA obtained by reversing the transitions of A_p and swapping the initial and final states is a DFA. Conclude that the minimal DFA for L_p has p states.

Solution.

(a) For a word $w \in \{0,1\}^*$, let $\text{msbf}(w)$ denote the number n such that $w \in \text{MSBF}(n)$. Similarly, let $\text{lsbf}(w)$ denote the number n such that $w \in \text{LSBF}(n)$. Note that the functions msbf and lsbf satisfy the following identities.

$$\text{msbf}(uv) = 2^{|v|} \cdot \text{msbf}(u) + \text{msbf}(v) \tag{1}$$

$$\text{lsbf}(uv) = \text{lsbf}(u) + 2^{|u|} \cdot \text{lsbf}(v) \tag{2}$$

First, let us show that M_p has at least p many residuals. For every $0 \leq i < p$, let u_i be a word such that $\text{msbf}(u_i) = i$ and $|u_i| = p-1$. Note that such an 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 $0 \leq k < p$, and let $\ell = (p-i) \bmod p$. We have:

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

Let $0 \leq i < j < p$. We have $u_i u_\ell \in M_p$ since $\text{msbf}(u_i u_\ell) \equiv i - i \bmod p \equiv 0 \bmod p$, but we have $u_j u_\ell \notin M_p$ since $\text{msbf}(u_j u_\ell) \equiv j - i \bmod p \not\equiv 0 \bmod p$. Therefore, the u_i-residual and u_j-residual of M_p are distinct. It follows that M_p has at least p many residuals.

To show that L_p has at least p many residuals, we use the same technique, except that we now let u_i be a word such that $\text{lsbf}(w) = i$ and $|u_i| = p-1$ and we use equation 2 instead of 1.

(b) We now give a DFA A_p for M_p with p states. By the previous subproblem, A_p has to be the minimal DFA for M_p. A_p is given by $A_p = (Q_p, \{0,1\}, \delta_p, 0, \{0\})$ where

$$\begin{aligned}
Q_p &= \{0, 1, \ldots, p-1\}, \\
\delta_p(q, b) &= (2q + b) \bmod p \quad \text{for every } q \in Q_p \text{ and } b \in \{0,1\}.
\end{aligned}$$

By using equation 1 and by induction on the length of w, we can show that $\delta_p(0, w) = q$ if and only if $\text{msbf}(w) \equiv q \bmod p$. It will then follow that A_p recognizes M_p.

(c) Let $B_p = (Q_p, \{0,1\}, \delta'_p, 0, \{0\})$ be the NFA obtained by reversing the transitions of A_p and then swapping its initial and final states. Note that $\delta'_p(q, b) = \{q' :$

$\delta_p(q', b) = q\}$. Hence, to show that B_p is a DFA, it is enough to show that for every $b \in \{0, 1\}$, the function $\delta_p^b : q \mapsto \delta_p(q, b)$ is bijective.

First, for every $b \in \{0, 1\}$, we will show that δ_p^b is injective. Fix a $b \in \{0, 1\}$. Note that $\delta_p^b(q) = (2q + b) \bmod p$. Suppose $2q_1 + b \equiv (2q_2 + b) \bmod p$ for some $q_1, q_2 \in Q_p$. Then $2(q_1 - q_2) \equiv 0 \bmod p$ and since $p > 2$ is a prime, this would imply that $q_1 = q_2$. Hence, the function δ_p^b is indeed injective.

Further, note that any injective function from a finite set to itself must also be a surjective function, i.e., the range of the function must be the entire finite set. It follows then that δ_p^b is bijective for every $b \in \{0, 1\}$ and this concludes the proof.