Exercise 2.1
Determine the residuals of the following languages:

(a) \((aa + bb)^*\) over \(\Sigma = \{a, b\}\),
(b) \((abc)^*\) over \(\Sigma = \{a, b, c\}\),
(c) \(\{a^n b^n c^n \mid n \geq 0\}\) over \(\Sigma = \{a, b, c\}\),
(d) \(\{a^n b^{3n} \mid n \geq 0\}\) over \(\Sigma = \{a, b\}\).

Exercise 2.2
(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 2.3
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 2.4

Let $\text{msbf}: \{0,1\}^* \to \mathbb{N}$ be such that $\text{msbf}(w)$ is the number represented by $w$ in the “most significant bit first” encoding. For example, $\text{msbf}(1010) = 10$, $\text{msbf}(100) = 4$, $\text{msbf}(0011) = 3$.

For every $n \geq 2$, let us define the following language:

$$M_n = \{w \in \{0,1\}^* : \text{msbf}(w) \text{ is a multiple of } n\}.$$

(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: For every $0 \leq i < p$, consider the word $u_i$ such that $|u_i| = p - 1$ and $\text{msbf}(u_i) = i$.]

\[^1\text{Recall this type of encoding from Exercise 1.4 from the previous exercise sheet. In contrast to the function MSBF, this one (msbf) maps an encoding to its corresponding natural number.}\]
Solution 2.1

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

- For \((abc)^*\). We give the residuals as regular expressions: \((abc)^*\) (residual of \(\varepsilon\)); \(bc(abc)^*\) (residual of \(a\)); \(c(abc)^*\) (residual of \(ab\)); \(\emptyset\) (residual of \(b\)). All other residuals are equal to one of these three.

- For \(L = \{a^n b^n c^n \mid n \geq 0\}\): Every prefix of a word of the form \(a^n b^n c^n\) has a different residual. For all other words the residual is the empty set. There are infinitely many residuals:
  - \(L^c = L\),
  - for every \(i \geq 1\), we have a residual with respect to \(a^i\), which is \(L^{a^i} = \{a^{n-i}b^n c^n \mid n \geq i\}\),
  - for every \(n \geq i \geq 1\) we have a residual with respect to \(a^n b^i\), which is \(L^{a^n b^i} = \{b^n c^{n-i}\}\),
  - for every \(n \geq i \geq 1\) we have a residual with respect to \(a^n b^n c^i\), which is \(L^{a^n b^n c^i} = \{c^n\}\),
  - \(L^b = \emptyset\).

- Similarly for \(L = \{a^n b^{3n} \mid n \geq 0\}\), every prefix of a word of the form \(a^n b^{3n}\) has a different residual:
  - \(L^c = L\),
  - for every \(i \geq 1\), we have a residual with respect to \(a^i\), which is \(L^{a^i} = \{a^{n-i}b^{3n} \mid n \geq i\}\),
  - for every \(3n \geq i \geq 1\) we have a residual with respect to \(a^n b^i\), which is \(L^{a^n b^i} = \{b^{3n-i}\}\),
  - \(L^b = \emptyset\).

Solution 2.2

(a) \(L = \{w \mid w \in \Sigma^*\}\). First we prove that \(L\) has infinitely many residuals by showing that for each pair of words of the infinite set \(\{01^i \mid i \geq 0\}\) the corresponding residuals are not equal. Let \(u = 01^i, v = 01^j \in \Sigma^*\) two words with \(i < j\). Then \(L^u \neq L^v\) since \(u \in L^u\), but \(u \notin L^v\). 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 \(L^u\) 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. \(L = \{a^n \mid n \geq 0\}\). 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 power of 2. If \(d_i < d_j\) holds, we are immediately done since \(a^{d_i} \in L^a^i\) and \(a^{d_j} \notin L^a^j\). \(d_i > d_j\) 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 power of 2. These have to be unequal, since the gaps between the powers of 2 are strictly increasing and we can repeat the argument from before.

Solution 2.3

A) (a)

<table>
<thead>
<tr>
<th>Iter.</th>
<th>Block to split</th>
<th>Splitter</th>
<th>New partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>—</td>
<td>—</td>
<td>{A, B, D, E, G, H}, {C, F, I}</td>
</tr>
<tr>
<td>1</td>
<td>{A, B, D, E, G, H}</td>
<td>(b, {A, B, D, E, G, H})</td>
<td>{A, D, G}, {B, E, H}, {C, F, I}</td>
</tr>
<tr>
<td>2</td>
<td>none, partition is stable</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

The language partition is \(P_t = \{\{A, D, G\}, \{B, E, H\}, \{C, F, I\}\}\).

(b) The minimal automaton is given below:
(c) \( \Sigma\Sigma(a\Sigma\Sigma + b\Sigma)^* \).

B) (a)

<table>
<thead>
<tr>
<th>Iter.</th>
<th>Block to split</th>
<th>Splitter</th>
<th>New partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>—</td>
<td>—</td>
<td>{q_0, q_3}, {q_1, q_2, q_4}</td>
</tr>
<tr>
<td>1</td>
<td>{q_1, q_2, q_4}</td>
<td>(b, {q_1, q_2, q_4})</td>
<td>{q_0, q_3}, {q_1}, {q_2, q_4}</td>
</tr>
<tr>
<td>2</td>
<td>{q_2, q_4}</td>
<td>(a, {q_0, q_3})</td>
<td>{q_0, q_3}, {q_1}, {q_2}, {q_4}</td>
</tr>
<tr>
<td>3</td>
<td>none, partition is stable</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

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

(b)

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

Solution 2.4

(a) The following DFA accepts \( M_3 \). The states represent congruence classes w.r.t. the modulo 3 relation.

As this DFA has three states, 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 \( \varepsilon \)-residual (\( M_3^\varepsilon \)), 1-residual (\( M_3^1 \)) and 10-residual (\( M_3^{10} \)) of \( M_3 \) are distinct.

- Since \( \varepsilon \cdot \varepsilon \in M_3 \) and \( 1 \cdot \varepsilon \notin M_3 \), we know that \( \varepsilon \in M_3^\varepsilon \) but \( \varepsilon \notin M_3^1 \), and thus \( M_3^\varepsilon \neq M_3^1 \).
- Since \( \varepsilon \cdot \varepsilon \in M_3 \) and \( 10 \cdot \varepsilon \notin M_3 \), we know that \( \varepsilon \in M_3^\varepsilon \) but \( \varepsilon \notin M_3^{10} \), and thus \( M_3^\varepsilon \neq M_3^{10} \).
- Since \( 1 \cdot 1 \in M_3 \) and \( 10 \cdot 1 \notin M_3 \), we know that \( 1 \in M_3^1 \) but \( 1 \notin M_3^{10} \), and thus \( M_3^1 \neq M_3^{10} \).

(b) The following DFA accepts \( M_4 \). You can obtain in two steps: (i) construct a DFA with four states that accepts \( M_4 \), where each state represents a congruence class w.r.t. the modulo 4 relation, (ii) minimize it.

As it has three states, \( M_4 \) has at most three residuals.
(c) A DFA accepting $M_p$ can be defined as $A_p = (Q_p, \{0, 1\}, \delta_p, 0, \{0\})$ where

\[
Q_p = \{0, 1, \ldots, p - 1\},
\]

\[
\delta_p(q, b) = (2q + b) \mod p \text{ for every } q \in Q_p \text{ and } b \in \{0, 1\}.
\]

As this DFA has $p$ states, then $M_p$ 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 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 $\ell = (p - i) \mod p$. We have:

\[
\text{msbf}(u_ku_\ell) = 2^{\text{msbf}(u_\ell)} \cdot \text{msbf}(u_k) + \text{msbf}(u_\ell) \\
= 2^{p-1} \cdot k + ((p - i) \mod p) \\
\equiv k + (p - i) \mod p \quad \text{(by Fermat’s little theorem)} \\
\equiv k + p - i \\
\equiv k - i.
\]

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