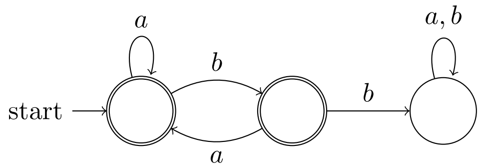
Technische Universität München **Winter term 2009/10**
I7
Prof. J. Esparza / J. Křetínský / M. Luttenberger **28. Oktober 2009**

Automata and Formal Languages – Homework 2

Due 5.11.2009.

Exercise 2.1

Let $\mathcal{A}$ be the following finite automaton:



(a) Transform the automaton $\mathcal{A}$ into an equivalent regular expression, then transform this expression into an NFA (with ε-transitions), remove the ε-transitions, determinize the automaton and then minimize it using the algorithms from the lecture. Check that the resulting automaton is isomorphic to the original one.

(b) Use JFLAP (`http://www.jflap.org/`) to perform the same transformations. Is there any difference?

(c) Let us denote $\Sigma^* = \{a, b\}$. Then $\mathcal{L}(\mathcal{A}) = \Sigma^* \setminus \Sigma^* bb \Sigma^*$. Prove that $\mathcal{A}$ is minimal by

- proving that no equivalence relation $\sim$ on Σ^* of index at most 2 can satisfy $x \sim y \Rightarrow \forall w(xw \in \mathcal{L}(\mathcal{A}) \Leftrightarrow yw \in \mathcal{L}(\mathcal{A}))$ (and thus concluding by Nerode theorem);
- computing the index of $\sim_{\mathcal{L}(\mathcal{A})}$ (and if it is 3 then concluding by Myhill-Nerode theorem);
- performing the *Bisim* algorithm for minimization (and checking the result is isomorphic to the original automaton).

Exercise 2.2

For $n \in \mathbb{N}_0$ let msbf(n) be the language of all words over $\{0, 1\}$ which represent n w.r.t. to the *most significant bit first* representation where an arbitrary number of leading zeros is allowed. For example:

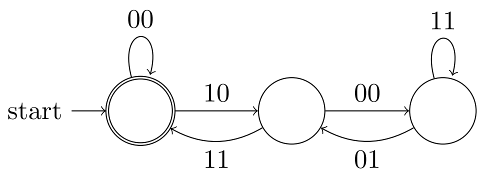
$$\text{msbf}(3) = \mathcal{L}(0^*11) \text{ and } \text{msbf}(0) = \mathcal{L}(0^*).$$

Similarly, let lsbf(n) denote the language of all *least significant bit first* representations of n with an arbitrary number of following zeros, e.g.:

$$\text{lsbf}(6) = \mathcal{L}(0110^*) \text{ and } \text{lsbf}(0) = \mathcal{L}(0^*).$$

(a) Construct and compare the minimal DFAs representing all even natural numbers $n \in \mathbb{N}_0$ w.r.t. the unary encoding (i.e., $n \mapsto a^n$), the msbf encoding, and the lsbf encoding.

(b) Consider the following FA $\mathcal{A}$ over the alphabet $\{00, 01, 10, 11\}$:



W.r.t. the msbf encoding, we may interpret any word $w \in \{00, 01, 10, 11\}^*$ as a pair of natural numbers $(X(w), Y(w)) \in \mathbb{N}_0 \times \mathbb{N}_0$. *Example*: (Underlined letters correspond to $Y(w)$.)

$$w = (00)^k 001011 \to (0\underline{0})^k 0\underline{0}1\underline{0}1\underline{1} \to (0^k 011, 0^k 001) \to (3, 1) = (X(w), Y(w))$$

- Find constants $a, b \in \mathbb{Z}$ such that $aX(w) + bY(w) = 0$ for all $w \in \mathcal{L}(\mathcal{A})$.

(c) Construct the minimal DFA representing the language $\{w \in \{0, 1\}^* \mid \text{msbf}^{-1}(w) \text{ is divisible by } 3\}$.

Exercise 2.3

In the lecture you have seen the definition of the Myhill-Nerode relation $\sim_L$ for a given language $L \subseteq \Sigma^*$:

For two words $x, y \in \Sigma^*$ we write $x \sim_L y$ iff $xv \in L \Leftrightarrow yv \in L$ holds for all $v \in \Sigma^*$.

(a) Determine the equivalence classes of $\sim_L$ w.r.t. the following languages over $\Sigma = \{a, b\}$:

- $L_1 := \mathcal{L}((ab + ba)^*)$,
- $L_2 := \mathcal{L}((aa)^*)$,
- $L_3 := \{w \in \{a, b\}^* \mid$ the number of occurrences of ab and ba in w is the same $\}$.
- $L_4 := \{a^n b^n c^n \mid n \geq 0\}$.

(b) In the definition of $\sim_L$ we compare two given words by appending all possible words. Instead of appending, we may also prepend. Consider therefore the binary relation $\sim^L$ on Σ^* defined by

For two words $x, y \in \Sigma^*$ we write $x \sim^L y$ iff $ux \in L \Leftrightarrow uy \in L$ holds for all $u \in \Sigma^*$.

- Determine the equivalence classes of $\sim^{L_1}$ and $\sim^{L_2}$.
- Show that L is regular iff $\sim^L$ has only finitely many equivalence classes.
- Is the number of equivalence classes of $\sim_L$ equal to the one of $\sim^L$?

(c) Finally, we may compare two words x, y also by appending and prepending arbitrary words, i.e., define the relation $\equiv_L$ as follows:

For two words $x, y \in \Sigma^*$ we write $x \equiv_L y$ iff $uxv \in L \Leftrightarrow uzv \in L$ holds for all $u, v \in \Sigma^*$.

For $x \in \Sigma^*$ let $[x]_L$ denote the equivalence class of x w.r.t. $\equiv_L$, i.e., $[x]_L = \{t \in \Sigma^* \mid x \equiv_L y\}$. We write $\Sigma^*/\equiv_L$ for the set of equivalence classes.

- How does $\equiv_L$ relate to $\sim_L$, resp. $\sim^L$?
- Determine the equivalence classes of $\equiv_{L_2}$.
- Show that the following multiplication on $\Sigma^*/\equiv_L$ is well-defined, associative and has $[\varepsilon]_L$ as its neutral element:

$$[w]_L \cdot [w']_L := [ww']_L.$$

Remark: $\Sigma^*/\equiv_L$ is called the syntactic monoid of L.

Exercise 2.4

Let Σ be an alphabet. The set $\mathbb{S}_\Sigma$ of *star-free* expression over Σ is inductively defined:

$$\begin{array}{lcl} S^0 & := & \Sigma \cup \{\varepsilon, \emptyset\} \\ S^{k+1} & := & \{(\phi + \psi), (\phi \cdot \psi), \overline{\phi}, (\phi \cap \psi) \mid \phi, \psi \in S^k\} \cup S^k \\ \mathbb{S}_\Sigma & := & \bigcup_{k \in \mathbb{N}} S^k. \end{array}$$

The language $\mathcal{L}(\rho)$ represented by a star-free expression $\rho \in \mathbb{S}_\Sigma$ is defined as expected:

$$\mathcal{L}(\rho) := \begin{cases} \emptyset & \text{if } \rho = \emptyset \qquad\qquad \mathcal{L}(\phi) \cup \mathcal{L}(\psi) \quad \text{if } \rho = (\phi + \psi) \\ \{\rho\} & \text{if } \rho \in \Sigma \cup \{\varepsilon\} \qquad \mathcal{L}(\phi) \cdot \mathcal{L}(\psi) \quad \text{if } \rho = (\phi \cdot \psi) \\ \Sigma^* \setminus \mathcal{L}(\phi) & \text{if } \rho = \overline{\phi} \qquad\qquad \mathcal{L}(\phi) \cap \mathcal{L}(\psi) \quad \text{if } \rho = (\phi \cap \psi). \end{cases}$$

To avoid parentheses, we assume that concatenation has the highest priority, followed by intersection, then addition. Some examples:

$$\mathcal{L}(\overline{\emptyset}) = \Sigma^* \setminus \mathcal{L}(\emptyset) = \Sigma^*, \quad \mathcal{L}(\overline{a} \cap (a + ab)) = \{ab\}.$$

(a) Show that for every star-free expression $\phi \in \mathbb{S}_\Sigma$ it holds that $\mathcal{L}(\phi)$ is a regular languages over Σ.

(b) Give a star-free expression for the regular language $\mathcal{L}((ab)^*)$ for $\Sigma = \{a, b\}$.

Remark: There is no star-free expression for the language $\mathcal{L}((aa)^*)$!