Formal Reasoning 2015
Solutions Test 3: Languages and automata
(21/10/15)

1. Give a regular expression that defines the language: (20 points)

$$L_1 := \{w \in \{a,b\}^* \mid w \text{ contains an even number of } a\text{'s}\}$$

Possible solutions are:

$$((ab^*a) \cup b)^*$$
$$b^*(ab^*ab^*)^*$$
$$(b^*ab^*a)^*b^*$$
$$(b^*ab^*ab^*) \cup b^*$$

If r is one of these expressions, then $L_1 = \mathcal{L}(r)$.

2. Give a right linear grammar G_2 that defines the language: (15 points)

$$L_2 := \{w \in \{a,b\}^* \mid w \text{ contains both an even number of } a\text{'s and an even number of } b\text{'s}\}$$

Let G_2 be the right linear grammar:

$$S \to aA \mid bC \mid \lambda$$
$$A \to aS \mid bB$$
$$B \to aC \mid bA$$
$$C \to aB \mid bS$$

Then $\mathcal{L}(G_2) = L_2$. Recall that the nonterminals represent some kind of state:

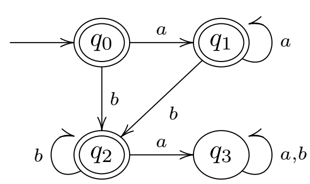
- S the number of $a's$ and the number of b's is even.
- A the number of $a's$ is odd and the number of b's is even.
- B the number of $a's$ is odd and the number of b's is odd.
- C the number of $a's$ is even and the number of b's is odd.

3. Give a finite automaton M_3 that matches the context-free grammar G_3: (20 points)

$$S \to AB$$
$$A \to aA \mid \lambda$$
$$B \to bB \mid \lambda$$

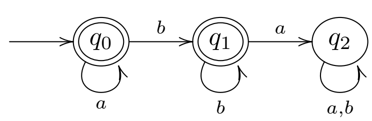
In particular $L(M_3) = \mathcal{L}(G_3)$ must hold.

Note that $\mathcal{L}(G_3) = \{a^m b^n \mid m, n \in \mathbb{N}\}$. Let M_3 be the automaton:



Then $L(M_3) = \mathcal{L}(G_3)$ holds.

An equivalent automaton, but minimal with respect to the number of states is:



4. (a) Let $|w|_a$ be the number of occurrences of the symbol a in word w, so for example $|abccb|_b = 2$, $|S|_S = 1$ and $|S|_a = 0$. Somebody claims that: (10 points)

$$P(w) := \ w \text{ contains } aa \ \text{ and/or } \ 2|w|_S + 2|w|_A + |w|_a \leq 2$$

is an invariant for the context-free grammar G_4:

$$\begin{aligned} S &\to BAB \\ A &\to aaA \mid \lambda \mid aBa \\ B &\to bB \mid \lambda \mid bBaaA \end{aligned}$$

Is this person right? Explain your answer.

Yes, this person is right: P is an invariant for G_4. Note that the 'and/or' is basically nothing else but the normal, logical inclusive 'or'. Here comes the proof of the two properties of an invariant.

- $P(S)$ holds because $2|S|_S + 2|S|_A + |S|_a = 2 + 0 + 0 \leq 2$.
- Let v be a word such that $P(v)$ holds and let v' be a word such that $v \to v'$. If $P(v)$ holds because v contains the string aa, then we are done immediately, because v' will also contain aa, because there are no rules that remove terminals and there are also no ways to put something in between two terminals. Hence we may assume that $2|v|_S + 2|v|_A + |v|_a \leq 2$. This means that there are seven possibilities for the step $v \to v'$:
  - $S \to BAB$. In this case we have that $|v'|_S = |v|_S - 1$, $|v'|_A = |v|_A + 1$ and $|v'|_a = |v|_a$. So

$$\begin{aligned} 2|v'|_S + 2|v'|_A + |v'|_a &= 2(|v|_S - 1) + 2(|v|_A + 1) + |v|_a \\ &= 2|v|_S - 2 + 2|v|_A + 2 + |v|_a \\ &= 2|v|_S + 2|v|_A + |v|_a \\ &\leq 2 \end{aligned}$$

    And hence $P(v')$ holds.
  - $A \to aaA$. In this case v' contains the string aa, which implies that $P(v')$ also holds.
  - $A \to \lambda$. In this case we have that $|v'|_S = |v|_S$, $|v'|_A = |v|_A - 1$ and $|v'|_a = |v|_a$. So

$$\begin{aligned} 2|v'|_S + 2|v'|_A + |v'|_a &= 2|v|_S + 2(|v|_A - 1) + |v|_a \\ &= 2|v|_S + 2|v|_A - 2 + |v|_a \\ &= \big(2|v|_S + 2|v|_A + |v|_a\big) - 2 \\ &\leq 2 - 2 \\ &= 0 \\ &\leq 2 \end{aligned}$$

  - $A \to aBa$. In this case we have that $|v'|_S = |v|_S$, $|v'|_A = |v|_A - 1$ en $|v'|_a = |v|_a + 2$. So

$$\begin{aligned} 2|v'|_S + 2|v'|_A + |v'|_a &= 2|v|_S + 2(|v|_A - 1) + |v|_a + 2 \\ &= 2|v|_S + 2|v|_A - 2 + |v|_a + 2 \\ &= 2|v|_S + 2|v|_A + |v|_a - 2 + 2 \\ &= 2|v|_S + 2|v|_A + |v|_a \\ &\leq 2 \end{aligned}$$

- $B \to bB$. This step doesn't change anything to the numbers of S's, A's and a's, hence $P(v')$ also holds.
- $B \to \lambda$. This step doesn't change anything to the numbers of S's, A's and a's, hence $P(v')$ also holds.
- $B \to bBaaA$. In this case v' contains the string aa, which implies that $P(v')$ also holds.

(b) Somebody else claims that: (10 points)

$$\mathcal{L}(G_4) = \{w \in \{a, b\}^* \mid w \text{ contains an even number of } a\text{'s}\}$$

Is this person right? If not, provide a word which is contained in exactly one of these two languages. Explain your answer. (Hint: have a look at $P(w)$ in Exercise 4a.)

No, this person is not correct. Consider the word $v = abababa$. This word v contains an even number of a's (namely four) and hence

$$v \in \{w \in \{a, b\}^* \mid w \text{ contains an even number of } a\text{'s}\}$$

However, $P(abababa)$ does not hold, because it doesn't contain the string aa and $2|v|_S + 2|v|_A + |v|_a = 2 \cdot 0 + 2 \cdot 0 + 4$ and $4 \not\leq 2$. So $v \notin \mathcal{L}(G_4)$.

5. (a) We define: (10 points)

$$L_5 := \{w \in \{a, b\}^* \mid w \text{ starts with an } a\}$$

Explain why:

$$L_5^* = L_5 \cup \{\lambda\}$$

The proof is split into two parts:

- $L_5 \cup \{\lambda\} \subseteq L_5^*$. For all languages L it holds that $L \subseteq L^*$, hence also for L_5. Furthermore, for all laguages L it holds that $\lambda \in L^*$, hence also for L_5. But from this it follows that $L_5 \cup \{\lambda\} \subseteq L_5^*$.
- $L_5^* \subseteq L_5 \cup \{\lambda\}$. Let $w \in L_5^*$. Then either $w = \lambda$ or $w = w_1 w_2 \ldots w_k$ for some $k \geq 1$ with $w_i \in L_5$ for all $i \in \{1, 2, \ldots, k\}$. In the first case it immediately follows that $w \in L_5 \cup \{\lambda\}$. In the second case it holds in particular that $w_1 \in L_5$. So w_1 start with an a. But if w_1 starts with an a, then $w = w_1 w_2 \ldots w_k$ also starts with an a. So $w \in L_5$ and hence also $w \in L_5 \cup \{\lambda\}$.

(b) Give two languages L and L' over alphabet $\Sigma = \{a, b\}$ such that: (5 points)

$$L \cap L' = \emptyset \quad L^* \neq \Sigma^* \quad L'^* \neq \Sigma^* \quad L \cup L' \neq \Sigma^* \quad L^* \cup L'^* = \Sigma^*$$

Explain your answer. (If you can't manage to comply to all these requirements, try to comply to as many as possible.)

Define $L_5' := \{w \in \{a, b\}^* \mid w \text{ starts with a } b\}$. Now take $L = L_5$ and $L' = L_5'$.

- $L \cap L' = \emptyset$. Because words in L always start with an a and words in L' always start with a b, the intersection $L \cap L'$ must be empty, because there are no words that start with an a and start with a b at the same time.
- $L^* \neq \Sigma^*$. In exercise 5a we have seen that $L^* = L \cup \{\lambda\}$. So if $w \in L^*$ then it holds that $w = \lambda$ or w starts with an a. But Σ^* also contains words that start with a b and therefore Σ is strictly larger than L^*.
- $L'^* \neq \Sigma^*$. Analogously, it alos holds for L' that $L'^* = L' \cup \{\lambda\}$. So if $w \in L'^*$ then it holds that $w = \lambda$ or w starts with a b. But Σ^* also contains words that start with an a and therefore Σ is strictly larger than L'^*.
- $L \cup L' \neq \Sigma^*$. We know that $\lambda \in \Sigma^*$. However, λ does not start with an a and it does not start with a b. Therefore λ is not in L and also not in L'. Hence also not in $L \cup L'$.

- $L^* \cup L'^* = \Sigma^*$. Because Σ^* contains all words over the alphabet $\{a, b\}$, it is clear that $L^* \cup L'^* \subseteq \Sigma^*$. In addition, we have already seen that $L^* = L \cup \{\lambda\}$ and $L'^* = L' \cup \{\lambda\}$. So $L^* \cup L'^* = L \cup L' \cup \{\lambda\}$. But if $w \in \Sigma^*$ then it holds that $w = \lambda$, w starts with an a or w starts with a b. And hence it follows that $w \in L \cup L' \cup \{\lambda\} = L^* \cup L'^*$.