

# Talen en Automaten

Test 1, Mon 7<sup>th</sup> Dec, 2015

15h45 – 17h30

This test consists of **four** exercises over **5 pages**. Explain your approach, and **write your answer to each exercise on a separate page**. You can score a maximum of 100 points, and each question indicates how many points it is worth. The test is closed book. You are NOT allowed to use a calculator, a computer or a mobile phone. You may answer in Dutch or in English. Please write clearly, and do not forget to put on each page: your name and your student number.

**Notation** Throughout the test, we denote for any alphabet  $A$ ,  $w \in A^*$  and  $a \in A$  by  $|w|_a$  the number of  $a$ 's in  $w$ , as it was introduced in the lecture. Moreover, recall that  $v$  is a *subword* of  $w$  if  $w = xvy$  for some words  $x, y$ .

## 1 Induction

Let  $A$  and  $B$  be finite alphabets and  $f : A \rightarrow B^*$  a map from  $A$  to words over  $B$ .

- a) Define by induction a map  $\bar{f} : A^* \rightarrow B^*$  that replaces in a word  $w \in A^*$  all letters  $a$  by  $f(a)$ . **(5pt)**

**Solution:** .....

We define the required map by

$$\begin{aligned}\bar{f}(\lambda) &= \lambda \\ \bar{f}(aw) &= f(a)\bar{f}(w)\end{aligned}$$

□

- b) Let  $A = \{a, b\}$  and  $f : A \rightarrow A^*$  be given by  $f(a) = b$  and  $f(b) = abb$ .

- i) Give a word  $w \in A^*$  such that  $\bar{f}(w) = babbbb$ . **(5pt)**

**Solution:** .....

$w = abaa$  does the job.

□

- ii) Show by induction that  $|\bar{f}(w)|_b = |w|_a + 2|w|_b$ . **(10pt)**

**Solution:** .....

We show by induction in  $w$  that  $|\bar{f}(w)|_b = |w|_a + 2|w|_b$ .

- IB:  $w = \lambda$ . Here, we get that

$$|\bar{f}(\lambda)|_b = |\lambda|_b = 0 = |\lambda|_a + 2|\lambda|_b$$

- IH: For  $w \in A^*$ ,  $|\bar{f}(w)|_b = |w|_a + 2|w|_b$ .
- IS: Let  $x \in A$ . We need to distinguish two cases.

1. If  $x = a$ , then we have

$$\begin{aligned}|\bar{f}(aw)|_b &= |f(a)\bar{f}(w)|_b = |b\bar{f}(w)|_b = |\bar{f}(w)|_b + 1 \\ &\stackrel{\text{IH}}{=} |w|_a + 2|w|_b + 1 = |aw|_a + 2|aw|_b\end{aligned}$$

as required.

2. If, on the other hand,  $x = b$ , then

$$\begin{aligned} |\overline{f}(bw)|_b &= |f(b)\overline{f}(w)|_b = |abb\overline{f}(w)|_b = |\overline{f}(w)|_b + 2 \\ &\stackrel{\text{IH}}{=} |w|_a + 2|w|_b + 2 = |bw|_a + 2|bw|_b. \end{aligned}$$

So in both cases  $|\overline{f}(xw)|_b = |xw|_a + 2|xw|_b$ .

This induction proves the desired equation. □

## 2 Regular Languages [Write your answers on a separate page]

a) Let  $A = \{a, b\}$  and

$$\begin{aligned} L_1 &= \{w \in A^* \mid |w|_a \text{ is even}\} \\ L_2 &= \mathcal{L}((a+b)^*abba(a+b)^*) \\ L_3 &= \{w \in A^* \mid w \text{ does not contain the subword } abb\}. \end{aligned}$$

Explain for each  $i = 1, 2, 3$  whether and why  $L_i^* = L_i$ . (10pt)

**Solution:** .....

- $L_1^* = L_1$ , since
  1.  $|\lambda|$  is even and for all  $w, v \in L_1$  we have  $|wv|_a = |w|_a + |v|_a$  is even; thus  $L_1^* \subseteq L_1$ , see 5b) on exercise sheet 1.
  2.  $L_1 \subseteq L_1^*$  holds trivially.
- $L_2^* \neq L_2$ , since  $\lambda \in L_2^*$  but not in  $L_2$ .
- $L_3^* \neq L_3$ , since  $ab \in L_3$  and  $b \in L_3$ , hence  $abb \in L_3^*$ .

□

b) Let  $A = \{a, b\}$  and

$$L = \{w \in A^* \mid aba \text{ occurs twice as subword in } w\}.$$

Give a regular expression  $e$ , such that  $\mathcal{L}(e) = L$ . Explain your answer. (10pt)

**Solution:** .....

We put  $e = (a+b)^*(aba(a+b)^*aba+ababa)(a+b)^*$ . That  $\mathcal{L}(e) = L$  can be seen as follows.

- To show the inclusion  $L \subseteq \mathcal{L}(e)$ , let  $w \in L$  so that there are  $x_1, y_1, x_2, y_2 \in A^*$  with  $w = x_1abay_1$  and  $x_1 \neq x_2$  and  $y_1 \neq y_2$ . WLOG, we assume that  $x_1 = x_2z$  for some non-empty word  $z$ . Then either
  1.  $z = abau, u \in A^*$ . Then  $w = x_2aba u aba y_1$ , hence  $w \in \mathcal{L}(e)$ .
  2.  $z = ab$ . Then  $w = x_2abba y_1$ , hence  $w \in \mathcal{L}(e)$ .
- The other direction can be seen as follows. Let  $w \in \mathcal{L}(e)$ , then either  $w = xabayabaz$  or  $w = xababay$ . Thus we can split in both cases  $w$  twice into  $uabav$  with  $u = x, v = yabaz$  and  $u = xabay, v = z$  in the first case, and  $u = x, v = bay$  and  $u = xab, v = y$  in the second case. In other words,  $w \in L$ .

This shows that  $L = \mathcal{L}(e)$ . □

### 3 Deterministic Finite Automata [Write your answers on a separate page]

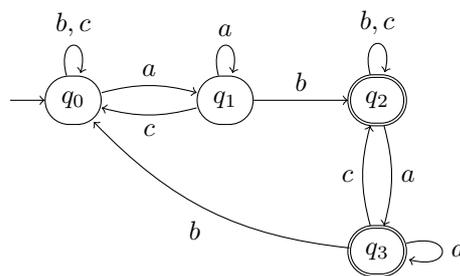
a) Let  $A = \{a, b, c\}$  and let

$$L = \{w \in A^* \mid ab \text{ occurs an odd number of times as subword in } w\}.$$

i) Give a DFA  $M$  with  $\mathcal{L}(M) = L$ . Explain your answer. (10pt)

**Solution:** .....

Define  $M$  to be



Explanation: We have the following invariants for the states.

- In  $q_0$  we have read an even number of  $ab$ .
- In  $q_1$  we are awaiting a  $b$  after an  $a$ , which would give us an odd number of  $ab$  read.
- In  $q_2$  we have read an odd number of  $ab$
- In  $q_3$  we are awaiting a  $b$  after an  $a$ , which would give us an even number of  $ab$  read.

The transitions then clearly preserve these invariants, and since only  $q_2$  and  $q_3$  are accepting, the automaton accepts exactly  $L$ . □

ii) Show that  $caba$  is accepted, and that  $abbab$  is not accepted. (5pt)

**Solution:** .....

- We have

$$q_0, caba \vdash q_0, aba \vdash q_1, ba \vdash q_2, a \vdash q_3, \lambda,$$

thus  $caba \in \mathcal{L}(M)$ .

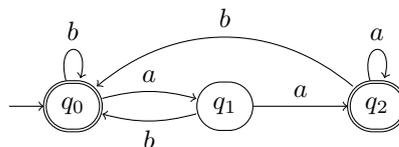
- We have

$$q_0, abbab \vdash q_1, bbab \vdash q_2, bab \vdash q_2, ab \vdash q_3, b \vdash q_0, \lambda,$$

thus  $abbab$  is not accepted.

□

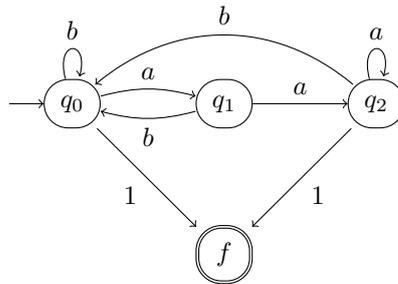
b) Let  $A = \{a, b\}$  and the DFA  $M$  over  $A$  be given by



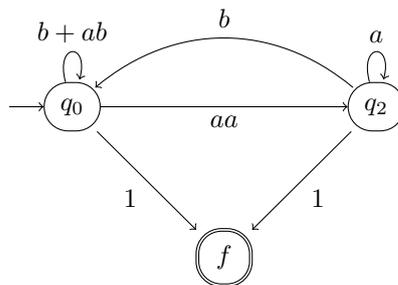
Use the procedure from the lecture to construct a regular expression  $e$  with **(10pt)**  
 $\mathcal{L}(e) = \mathcal{L}(M)$ .

**Solution:** .....

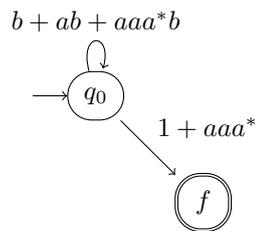
First, we need to transform  $M$  to only have one final state:



Then we eliminate  $q_1$ :



Next, we eliminate  $q_2$ :



Thus  $e$  is given by  $(b + ab + aaa^*b)^*(1 + aaa^*)$ . □

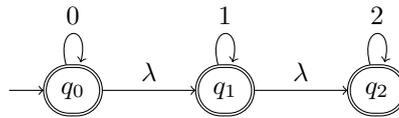
#### 4 Non-Deterministic Finite Automata [Write your answers on a separate page]

a) Let  $A = \{0, 1, 2\}$  and let  $L$  be the language of words in which the digits occur only in increasing order, i.e.,

$$L = \{x_1 \cdots x_n \mid n \in \mathbb{N}, \forall i. x_i \in A \text{ and } \forall i \leq j. x_i \leq x_j\}.$$

i) Show that  $L$  is regular by constructing an NFA- $\lambda$  that accepts  $L$ . **(10pt)**

**Solution:** .....



□

ii) Show that your automaton accepts 002 and rejects 21. (5pt)

**Solution:** .....

- We have

$$q_0, 002 \vdash q_0, 02 \vdash q_0, 2 \vdash q_1, 2 \vdash q_2, 2 \vdash q_2, \lambda,$$

hence 002 is accepted.

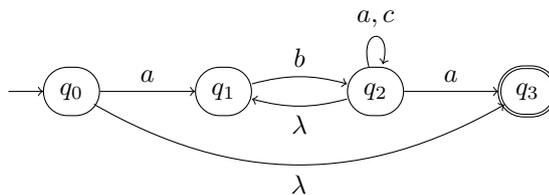
- We have

$$\begin{aligned} \delta^*(q_0, 21) &= \bigcup_{q \in \lambda\text{-closure}(q_0)} \bigcup_{p \in \delta(q, 2)} \delta^*(p, 1) \\ &= \emptyset \cup \emptyset \cup \delta^*(q_2, 1) \\ &= \emptyset, \end{aligned}$$

hence 21 is not accepted.

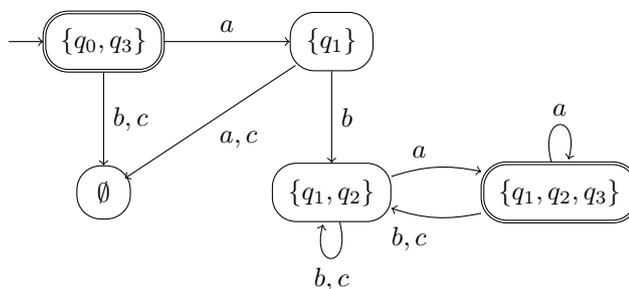
□

b) Let  $A = \{a, b, c\}$  and the NFA- $\lambda$   $M$  over  $A$  be given by



Use the procedure from the lecture to construct a DFA  $D$  with  $\mathcal{L}(D) = \mathcal{L}(M)$ . (10pt)  
Indicate clearly from which subset of states of  $M$  a state in  $D$  originates.

**Solution:** .....



□

c) Let  $e$  be the regular expression  $a + (b + 1)^*$ .

Use the procedure from the lecture to construct an NFA- $\lambda$   $M$  with  $\mathcal{L}(M) = \mathcal{L}(e)$ . (10pt)

# Talen en Automaten

Test 2, Thu 21<sup>st</sup> Jan, 2016

This test consists of **four** exercises over **4 pages**. Explain your approach. You can score a maximum of 100 points, and each question indicates how many points it is worth. The test is closed book. You are NOT allowed to use a calculator, a computer or a mobile phone. You may answer in Dutch or in English. Please write clearly, and do not forget to put on each page: your name and your student number.

**Notation** Throughout the test, we denote for any alphabet  $A$  and  $a \in A$  by  $|w|_a$  the number of  $a$ 's in the word  $w \in A^*$ , as it was introduced in the lecture.

**Write the answer to each exercise on a separate sheet!**

## 1 Non-Regular Languages

**Write your answers on a separate sheet**

Let  $A = \{a, b\}$ .

- a) We define the language  $L$  to be

$$L = \{wb^n \mid w \in A^*, |w| = n\}.$$

Show that  $L$  is not regular.

**(5pt)**

**Solution:** .....

Assume that  $L$  is regular and let  $p > 0$  be the pumping length which we get from the pumping lemma (PL). Take  $w = a^p b^p \in L$ , which can be divided by the PL into  $w = xyz$  with  $|y| \geq 1$ ,  $|xy| \leq p$  such that  $xy^i z \in L$  for all  $i \in \mathbb{N}$ . By the above constraints we immediately get that  $y = a^k$  with  $k > 0$ . Take  $i = 0$ . So  $xy^0 z = a^{p-k} b^p \in L$ . But  $p - k \neq p$ , so  $xy^0 z \notin L$ . Contradiction, hence  $L$  is not regular.

Students may use the following argument " $L \cap \mathcal{L}(a^* b^*) = \{a^n b^n \mid n \in \mathbb{N}\}$ , which is not regular. Thus  $L$  cannot be regular." However:  $L \cap \mathcal{L}(a^* b^*) = \{a^m b^n \mid m \geq n\}$ , so this argument is incorrect.  $\square$

- b) Show that the language  $L = \{w \in A^* \mid |w|_a = |w|_b\}$  is not regular, using the Pumping Lemma. **(10pt)**

**Solution:** .....

Assume that  $L$  is regular and let  $p > 0$  be the pumping length which we get from the pumping lemma (PL). Take  $w = a^p b^p \in L$ , which can be divided by the PL into  $w = xyz$  with  $|y| \geq 1$ ,  $|xy| \leq p$  such that  $xy^i z \in L$  for all  $i \in \mathbb{N}$ . By the above constraints we immediately get that  $y = a^k$  with  $k > 0$ . Take  $i = 0$  (or  $i = 2$  also works). So  $xy^0 z = a^{p-k} b^p \in L$ . But  $p - k \neq p$ , so  $xy^0 z \notin L$ . Contradiction, hence  $L$  is not regular.  $\square$

## 2 Context Free Grammars

### Write your answers on a separate sheet

Fix  $A = \{a, b\}$  for this exercise.

a) Let  $L$  be the language over  $A$  given by  $L = \{a^n b^k a^m \mid k = n + m\}$ .

i) Construct a CFG  $G$  such that  $\mathcal{L}(G) = L$ . (10pt)

**Solution:** .....

$G$  is given by the productions

$$\begin{aligned} S &\longrightarrow LR \\ L &\longrightarrow aLb \mid \lambda \\ R &\longrightarrow bRa \mid \lambda \end{aligned}$$

having non-terminals  $\{S, L, R\}$  and start symbol  $S$ . □

ii) Give a derivation for the word  $aabbba \in L$ . (5pt)

**Solution:** .....

$$\begin{aligned} S &\Rightarrow LR \Rightarrow aLbR \Rightarrow aaLbbR \Rightarrow aabbR \\ &\Rightarrow aabbbRa \Rightarrow aabbba \end{aligned}$$

□

iii) Show that the word  $aba$  is not generated. (5pt)

**Solution:** .....

The only possible way to get the first  $ab$  is by

$$S \Rightarrow LR \Rightarrow aLbR \Rightarrow abR$$

but then from  $R$  we can only get to  $\lambda$  or to  $b\dots$ , so we cannot generate  $aba$ .

**Alternative: Full Case Exploration** Starting from  $S$ , we have the following possible derivations.

$$S \Rightarrow LR \Rightarrow \{aLbR, LbRa, R, L\}$$

For these, in turn, we have

•

$$aLbR \Rightarrow \{abR, aaLbbR, aLb, aLbbRa\}$$

Thus in each case either too little a's or too many b's are produced.

•

$$LbRa \Rightarrow \{bRa, aLbbRa, Lba, LbbRaa\}$$

Similar

- $R$  cannot produce a in front of b
- $L$  cannot produce b in front of a

Thus in neither of these cases we can derive  $aba$ , so it is not in the language generated by the grammar. □

b) Let  $G$  be the following CFG over  $A$ .

$$\begin{aligned} S &\longrightarrow US \mid \lambda \\ U &\longrightarrow aa \mid ab \mid bb \mid ba \end{aligned}$$

i) Give a precise description of  $\mathcal{L}(G)$  using set notation. (5pt)

**Solution:** .....

$$\begin{aligned} \mathcal{L}(G) &= \{w \in A^* \mid |w| \text{ even}\} \\ &= \{w_1 \cdots w_n \mid n \in \mathbb{N}, w_i \in \{aa, ab, bb, ba\}\} \\ &= \mathcal{L}((aa + ab + bb + ba)^*) \end{aligned}$$

□

ii) Is  $\mathcal{L}(G)$  a regular language? Explain your answer by either giving a reason why it is not or by giving a regular grammar for  $\mathcal{L}(G)$ . (10pt)

**Solution:** .....

$\mathcal{L}(G)$  is regular. We can substitute  $U$  in the first production of  $S$  to obtain

$$S \longrightarrow aaS \mid abS \mid bbS \mid baS \mid \lambda.$$

This can be refined into the following regular grammar.

$$\begin{aligned} S &\longrightarrow aA \mid aB \mid bB \mid bA \mid \lambda \\ A &\longrightarrow aS \\ B &\longrightarrow bS \end{aligned}$$

**More compact solution**

$$\begin{aligned} S &\longrightarrow aT \mid bT \mid \lambda \\ T &\longrightarrow aS \mid bS \end{aligned}$$

□

### 3 Push Down Automata I

**Write your answers on a separate sheet**

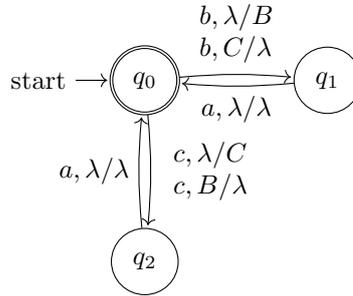
Let  $M$  be the PDA with

$$\begin{aligned} Q &= \{q_0, q_1, q_2\} & \delta(q_0, b, \lambda) &= \{(q_1, B)\} \\ \Sigma &= \{a, b, c\} & \delta(q_0, b, C) &= \{(q_1, \lambda)\} \\ \Gamma &= \{B, C\} & \delta(q_0, c, \lambda) &= \{(q_2, C)\} \\ F &= \{q_0\} & \delta(q_0, c, B) &= \{(q_2, \lambda)\} \\ & & \delta(q_1, a, \lambda) &= \{(q_0, \lambda)\} \\ & & \delta(q_2, a, \lambda) &= \{(q_0, \lambda)\} \end{aligned}$$

a) Draw a state diagram for  $M$ . (5pt)

**Solution:** .....

$M$  can be drawn as



□

- b) Check which of the following words is in  $\mathcal{L}(M)$  and explain your answer:  $abcb$  (5pt) and  $baca$ .

**Solution:** .....

- There is no transition from  $q_0$  that reads an  $a$ , so  $abcb \notin \mathcal{L}(M)$ .
- $\langle q_0, baca, \lambda \rangle \rightarrow \langle q_1, aca, B \rangle \rightarrow \langle q_0, ca, \lambda \rangle \rightarrow \langle q_2, a, C \rangle \rightarrow \langle q_0, \lambda, \lambda \rangle$ . We end in an accepting state with an empty stack, so  $baca \in \mathcal{L}(M)$ .

□

- c) Is  $\mathcal{L}((ca)^*(ba)^*) \subseteq \mathcal{L}(M)$ ? Explain your answer. (5pt)

**Solution:** .....

No,  $ca \in \mathcal{L}((ca)^*(ba)^*)$ , but  $ca \notin \mathcal{L}(M)$ . To see the latter, note that when reading a  $c$ , either a  $B$  must be popped of the stack, or a  $C$  is pushed on the stack. The first is not possible with an empty stack, and after the second the stack is not empty.

Alternatively: if  $w \in \mathcal{L}(M)$  then the number of  $a$ 's and  $b$ 's in  $w$  is equal, following the same argument as above. This does not have to be the case for every  $w \in \mathcal{L}((ca)^*(ba)^*)$ . □

- d) Give a precise description of  $\mathcal{L}(M)$  using set notation. (5pt)

**Solution:** .....

$$\mathcal{L}(M) = \{w \in \mathcal{L}((ac+bc)^*) \mid |w|_a = |w|_b\}.$$

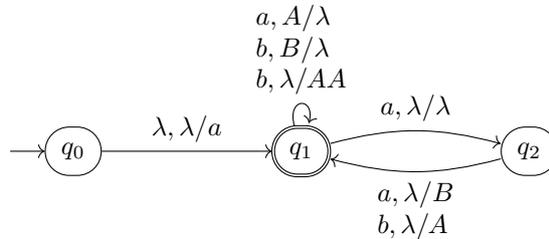
That is, the words with an equal number of  $a$ 's and  $b$ 's, where every  $a$  and every  $b$  is followed by a  $c$ , and every  $c$  is preceded by an  $a$  or a  $b$ . □

## 4 Push Down Automata II

Write your answers on a separate sheet

- a) i) Let  $A = \{a, b\}$  and let  $L$  be the language  $L = \{w \in A^* \mid |w|_a = 2|w|_b + 1\}$ . Show that  $L$  is context free by giving a PDA that accepts it. (10pt)

**Solution:** .....



□

- ii) Show that  $aaba$  and  $baaa$  are accepted, by giving the accepting computations. (5pt)

**Solution:** .....

$$\begin{aligned}
 q_0, aaba, \lambda &\rightarrow q_1, aba, \lambda \rightarrow q_2, ba, \lambda \rightarrow q_1, a, A \rightarrow q_1, \lambda, \lambda \\
 q_0, baaa, \lambda &\rightarrow q_0, aaa, AA \rightarrow q_1, aa, AA \rightarrow q_1, a, A \rightarrow q_1, \lambda, \lambda
 \end{aligned}$$

□

- iii) Show that  $aab$  is not accepted by your PDA. (5pt)

**Solution:** .....

We have the following possible computations

$$\begin{aligned}
 (q_0, aab, \lambda) &\rightarrow (q_1, aab, a) \\
 &\rightarrow \{(q_1, ab, \lambda), (q_2, ab, a)\} \\
 &\rightarrow \{(q_2, b, \lambda), (q_1, b, ba)\} \\
 &\rightarrow \{(q_1, \lambda, a), (q_1, \lambda, a)\},
 \end{aligned}$$

neither of which ends with an empty stack. Thus  $aab$  is not accepted. □

- b) Let  $G$  be the grammar on the alphabet  $\{a, b\}$  given as follows.

$$\begin{aligned}
 S &\rightarrow \lambda \mid aX \mid bY \\
 X &\rightarrow bYb \mid bb \\
 Y &\rightarrow aXa \mid aa
 \end{aligned}$$

Construct a PDA that accepts  $\mathcal{L}(G)$ , using the procedure given in the lecture. (10pt)

**Solution:** .....

First we transform the grammar into the form necessary construct a PDA.

$$S \rightarrow \lambda \mid aX \mid bY$$

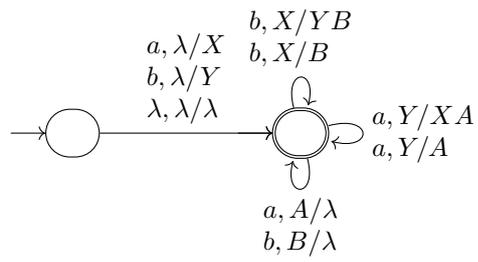
$$X \rightarrow bYB \mid bB$$

$$Y \rightarrow aXA \mid aA$$

$$A \rightarrow a$$

$$B \rightarrow b$$

The resulting PDA is then.



□