Talen en Automaten

Test 2, Thu 21st Jan, 2016

This test consists of **four** exercises over **6 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.

1 Non-Regular Languages

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 (incorrect) 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: $L \cap \mathcal{L}(a^* b^*) = \{a^m b^n \mid n \geq m \text{ and } n + m \text{ is even}\}$, so this argument is incorrect. □

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. □

2 Context Free Grammars

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

$$S \longrightarrow LR$$
$$L \longrightarrow aLb \mid \lambda$$
$$R \longrightarrow bRa \mid \lambda$$

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

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

Solution: ....

$$S \Rightarrow LR \Rightarrow aLbR \Rightarrow aaLbbR \Rightarrow aabbR$$
$$\Rightarrow aabbbRa \Rightarrow aabbba$$

□

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 λ or to $b\ldots$, 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 to 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 note in the language generated by the gramar. □

b) Let G be the following CFG over A.

$$S \longrightarrow US \mid \lambda$$
$$U \longrightarrow aa \mid ab \mid bb \mid ba$$

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

Solution: . . . .

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

□

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

Let M be the PDA with

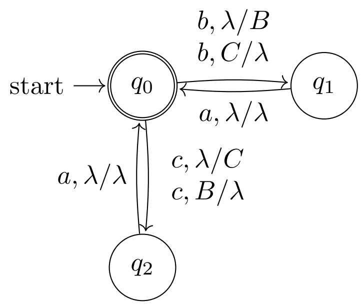
$$\begin{aligned}Q &= \{q_0, q_1, q_2\} \\ \Sigma &= \{a, b, c\} \\ \Gamma &= \{B, C\} \\ F &= \{q_0\}\end{aligned}$$

$$\begin{aligned}\delta(q_0, b, \lambda) &= \{\langle q_1, B\rangle\} \\ \delta(q_0, b, C) &= \{\langle q_1, \lambda\rangle\} \\ \delta(q_0, c, \lambda) &= \{\langle q_2, C\rangle\} \\ \delta(q_0, c, B) &= \{\langle q_2, \lambda\rangle\} \\ \delta(q_1, a, \lambda) &= \{\langle q_0, \lambda\rangle\} \\ \delta(q_2, a, \lambda) &= \{\langle q_0, \lambda\rangle\}\end{aligned}$$

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: $abcb$ and $baca$. **(5pt)**

Solution: ....

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

□

c) Is $\mathcal{L}\left((ca)^*(ba)^*\right) \subseteq \mathcal{L}(M)$? 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: ....

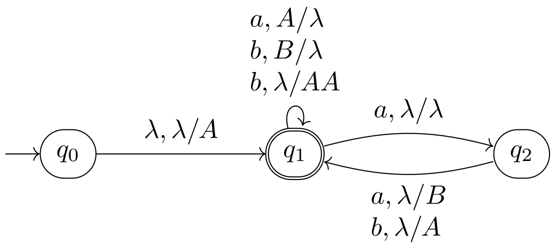
$$\mathcal{L}(M) = \{w \in \mathcal{L}((ba + ca)^*) \mid |w|_b = |w|_c\}.$$

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

4 Push Down Automata II

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: ....

$$q_0, aaba, \lambda \to q_1, aba, \lambda \to q_2, ba, \lambda \to q_1, a, A \to q_1, \lambda, \lambda$$
$$q_0, baaa, \lambda \to q_0, aaa, AA \to q_1, aa, AA \to q_1, a, A \to q_1, \lambda, \lambda$$

□

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

Solution: ....

We have the following possible computations

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

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.

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

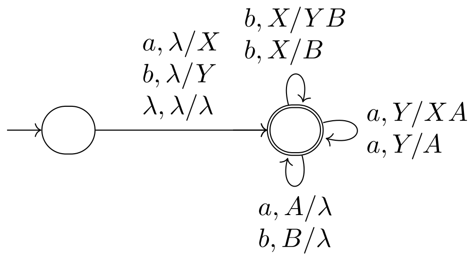
Construct a PDA that accepts $\mathcal{L}(G)$, using the procedure given in the lecture. **(10pt)**

Solution: ....

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

$$\begin{aligned}S &\to \lambda \mid aX \mid bY\\ X &\to bYB \mid bB\\ Y &\to aXA \mid aA\\ A &\to a\\ B &\to b\end{aligned}$$

The resulting PDA is then.


$a, \lambda/X$
$b, \lambda/Y$
$\lambda, \lambda/\lambda$
$b, X/YB$
$b, X/B$
$a, Y/XA$
$a, Y/A$
$a, A/\lambda$
$b, B/\lambda$

$\square$