Talen en Automaten

Assignment 1, Tue 8th Nov, 2016

Exercise teachers. The student groups are supervised by the following teachers:

Teacher	E-Mail	Room	Time
Michiel de Bondt	M.deBondt@math.ru.nl	HG00.058	10:45 – 12:30
Bas Steeg	steegbas@gmail.com	HG02.032	10:45 – 12:30
Demian Janssen	wd.janssen@student.ru.nl	HG01.058	10:45 – 12:30
Jan Martens	jan-martens@hotmail.com	HG01.139	10:45 – 12:30
Ties Robroek	ties.robroek@student.ru.nl	HG02.028	10:45 – 12:30
Sjoerd Hemels	sjoerd0707@live.nl	HG03.632	10:45 – 12:30
Rick Erkens	rjarickerkens@gmail.com	HG02.028	**13:45 – 15:30**

Postboxes are located in the Mercator building on the ground floor There will be 7 boxes labelled with *Talen en Automaten* and the corresponding group teacher's name. There will be 1 box, the *Uitleverbak*, for work that hasn't been picked up at the exercise hours.

Handing in your answers: There are two options:

1. E-mail: Send your solutions by e-mail to your exercise class teacher (see above) with subject "**T&A: assignment 1**". This e-mail should only contain a single PDF document as attachment (unless explicitly stated otherwise). Before sending an e-mail make sure:
   - the file is a PDF document
   - your name is part of the filename (for example MyName_assignment-1.pdf)
   - your name and student number are included in the document (they will be printed).
2. Post box: Put your solutions in the appropriate post box (see above). Before putting your solutions in the post box make sure:
   - your name, student number, and IC, KI or Wiskunde are written clearly on the document.

Deadline: Tue 15th Nov, 2016, 13:45 (in Nijmegen!)

Goals: After completing these exercises successfully you should be able to carry out definitions and proofs by induction on words, you should be able to write a regular expression for a simple regular language and be able to grasp what language a regular expression denotes.

There are 3 mandatory exercises, worth **10 points** in total. There is 1 more, extra hard, exercise. Be aware that this exercise is just for fun, you cannot earn any points with it.

1 Counting Letters

a) Let A be a non-empty, finite alphabet and $a \in A$ a letter.

Define by structural induction a map **(1pt)**

$$|\,.\,|_a \colon A^* \to \mathbb{N}$$

that counts the number of occurrences of the letter a in a word.

b) Show, by induction, that for any two words $w, u \in A^*$ **(2pt)**

$$|wu|_a = |w|_a + |u|_a.$$

2 Substitution

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

a) Define by induction a map $\overline{f} : A^* \to B^*$ that replaces in a word $w \in A^*$ all letters a by $f(a)$. **(1pt)**

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

i) Give a word $w \in A^*$ such that $\overline{f}(w) = babbbb$. **(1pt)**

ii) Show by induction that $\left|\overline{f}(w)\right|_b = |w|_a + 2|w|_b$. **(1pt)**

3 Regular Expressions

a) Consider the languages $L_1 = \mathcal{L}((bab)^*)$, $L_2 = \mathcal{L}((ba)^*b)$, $L_3 = \mathcal{L}(\big(b(ab)^*\big)^*)$.

For which i and j do we have $L_i \subseteq L_j$ and for which do we have $L_i \not\subseteq L_j$? Explain your answers. (There are 6 cases to consider; only a minority of the $\subseteq$ holds.) **(2pt)**

b) Let L be the language given by

$$\big\{w \in \{a, b\}^* \,\big|\, \text{every } b \text{ in } w \text{ is directly followed by an } a\big\}$$

Give a regular expression for the language L and explain your answer. **(1pt)**

c) Let L be the language given by

$$\big\{w \in \{a, b\}^* \,\big|\, |w|_b \text{ is even and } w \text{ does not end with a } b\big\}$$

Give a regular expression for the language L and explain your answer. **(1pt)**

4 Fun Exercises – Regular Expressions

a) Show that the language

$$\{w \in \{a,b\}^* \mid aa \text{ occurs exactly twice in } w\}\,.$$

is regular. [Hint: Beware of the string aaa!]

b) Show that the language

$$\{w \in \{a, b^* \mid |w|_a \text{ and } |w|_b \text{ are even}\}$$

is regular.

c) We say, that two regular expressions are equal, if they generate the same language. Symbolically: $e_1 = e_2$ iff $\mathcal{L}(e_1) = \mathcal{L}(e_2)$. There is also an order on regular expressions, given by

$$e_1 \leq e_2 \Leftrightarrow e_1 + e_2 = e_2$$

using the equality of generated languages.

Show that $e_1 \leq e_2$ iff $\mathcal{L}(e_1) \subseteq \mathcal{L}(e_2)$.

d) Show that for any regular expression e the inequality $1 + ee^* \leq e^*$ holds.