Exercises 2
03.03.2014

Rules: The document contains a set of 5 exercises, each of them worth 2 points. You need to provide for exercises 2 to 5 a Python script and for exercise 1 a file with a single line which represents the required regex. All Python scripts together with the file for exercise 1 must be put in a ZIP archive named FirstName_LastName_ex*n*.zip, where n is the number of the exercise session (see ex_set_1.pdf). The Python scripts must be named exercise_m.py, where m is the number of the exercise. The ZIP archive must be uploaded on ILIAS until the specified deadline.
For each exercise session you will get either one point or zero. One point is given if 4 out of 5 exercises in that particular session are correctly solved.
Good luck!

Exercise 1. Write a regex (using the theoretical notation) that matches the first 4 strings below, but not the 5th.

1. abc
2. acab
3. acababababc
4. abababab

5. abcababc

Hint: there are multiple correct possibilities: they are all accepted.

Exercise 2. Write a python script that extracts and outputs to stdout the strings having at least three times the letter “a” inside, on any position. Use file input_ex2.txt as input.

Exercise 3. Write a python script that finds all strings in input_ex3.txt file having at least one group of 2 consecutive digits and does not start with the letter “b”, case insensitive.

Exercise 4. Given the list of emails in the input_ex4.txt file, output to stdout the ones that have the following format:

first_name**.**last_name**@**provider**.com**,

where the constructions in bold are fixed, first_name and last_name must begin with a capital letter, cannot contain numbers and cannot be the empty string, *provider* cannot be the empty string.

Exercise 5. Parse the text in input_ex5.txt file and count the words. A word has more than 5 letters, does not contain digits and does not start with a capital letter.