Lab session Regular Expressions

Group A: October 16, 2009
Group B: October 13, 2009

Work in the given groups of two. Submit your solutions to the respective assignment on Blackboard. The file name is:
`s02_s0XXXXX_s0XXXXX.tar.gz`
One of the group members commits your solution. The other(s) submit a txt-file with a confirmation. Keep an eye on the deadline (see Blackboard)!

1 Exercises

Test your regular expressions using "gsed -r".

1. Find a regular expression that matches a number at the end of a line.
2. Find a regular expression that matches filenames with a "tar.gz" extension.
3. Find a regular expression that matches all words of 4 characters long.
4. Find a regular expression that matches any number between 1 and 999.
5. Find a regular expression that matches dates of the form:
   31/08/1933
   2-03-2002
   09 4 1966
   15.12.1999
6. Find a regular expression that matches an IPv4 number (0.0.0.0 through 255.255.255.255).
7. Find a regular expression that matches hexadecimal representations of the form:
   0x2a
   0XF
   0X1111
   0x0

8. Find a regular expression that matches floating point numbers. Some examples of floating points are:

   12.245
   -234
   +.0009
   3.11 e33
   43.1E11
   2e-14

9. Find a regular expression that matches strings surrounded by square brackets. Beware of greedy evaluation! For example, the HTML string "Hello, this is emphasized." should match twice, for and for .

10. Find a Sed command that extracts HTML tags (without attributes or nested tags) from a text. A text must be converted as follows:

    <h1>This is a valid HTML tag</h1>.
    These <1>invalid <a}>tags</a}> should be ignored.

    Becomes: This is a valid HTML tag.
    These <1>invalid <a}>tags</a}> should be ignored.

2 Project

There is no project this week. You only have to submit your solutions to the exercises. There will be no feedback loop on this lab session.