TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY

Programmeerimise süvendatud algkursus ITI0140

2014

Teema

- Regulaaravaldised (ingl *regular expressions, regex*)

Dokumentatsioon

https://docs.python.org/3.4/library/re.html

https://developers.google.com/edu/python/regular-expressions

re

```
import re

match = re.search("abc", "aabca")
print (match.group())
```

Süntaks

	r"abc"	"aabc"	"ababac"
'.'	r"a.bc"	"a2bc"	"abc"
'^'	r"^abc"	"abcd"	"aabc"
'$'	r"abc$"	"aaabc"	"abca"
'?'	r"ab?c"	"aacc" "aabcc"	"aabbcc"
'*'	r"ab*c"	"aaca" "aabca" "aabbbca"	"abab" "abbba"
'+'	r"ab+c"	"aabca" "aabbbca"	"aaca" "abbba"

Süntaks

{m}	r"ab{2}c"	"aabbcc"	"aacc", "aabcc", "aabbbcc"
{m,n}	r"ab{2,3}c"	"aabbcc", "aabbbcc"	"aacc", "aabcc", "aabbbbcc"
'\'	r"a\.bc"	"a.bc"	"a2bc"
[]	r"[ab]cd"	"acd", "bcd"	"ccd"
	r"[a-c]de"	"ade", "bde", "cde"	"dde"
	r"[a\-b]cd"	"acd", "bcd", "-cd"	"ccd"
(...)	r"(ab)?cd"	"cde" "abcde"	"acde" "bcde"
'|'	r"(ab|cd)?cd"	"acde", "abcde", "cdcde"	"ade"

Süntaks

\w - täht

\s - tühik, taandrida, reavahetus jne

\t \n \r - taandrida, uus rida, reavahetus

\d - number

Grupid

```
match = re.search("(ab)?(cd)", "abcd")
print (match.group())          "abcd"
print (match.group(0))         "abcd"
print (match.group(1))           "ab"
print (match.group(2))           "cd"
```

Grupid

```
match = re.search("(ab)?(cd)", "acd")
print (match.group())          "cd"
print (match.group(0))         "cd"
print (match.group(1))         None
print (match.group(2))         "cd"
```

Grupid

```
match = re.search("(?:ab)?(cd)","abcd")
print (match.group())          "abcd"
print (match.group(0))         "abcd"
print (match.group(1))           "cd"
```

re. Funktsioonid

compile - 1 kord avaldis valmis, N korda kasutada

search - otsib vastet kogu sõnest

match - otsib vastet sõne algusest

split - avaldise järgi sõne tükeldamine

re. Funktsioonid

group() - sama, mis group(0)
group([group1, ...]) - annab vastava indeksi või nimega sulgude vaste sõnest
groups() - kõik vasted sõnest
start - vaste algus sõnes (argumen
end - vaste lõpp sõnes

R4F

http://rise4fun.com/Rex

nt “(ab|cd)?cd”

Ülesanne

https://cs.uwaterloo.ca/~dtompkin/teaching/08a/lab7/

ülesanded 1-5 & 9-12 & valida 3 ül 13-19 seast