Laporan Praktikum Algoritma dan Pemrograman

Semester Genap 2023/2024

NIM	**<71230972>**
Nama Lengkap	**<Oktavian Christ Putranto>**
Minggu ke / Materi	**14 / Regular Expression**

SAYA MENYATAKAN BAHWA LAPORAN PRAKTIKUM INI SAYA BUAT DENGAN USAHA SENDIRI TANPA MENGGUNAKAN BANTUAN ORANG LAIN. SEMUA MATERI YANG SAYA AMBIL DARI SUMBER LAIN SUDAH SAYA CANTUMKAN SUMBERNYA DAN TELAH SAYA TULIS ULANG DENGAN BAHASA SAYA SENDIRI.

SAYA SANGGUP MENERIMA SANKSI JIKA MELAKUKAN KEGIATAN PLAGIASI, TERMASUK SANKSI TIDAK LULUS MATA KULIAH INI.

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
YOGYAKARTA
2024

BAGIAN 1: MATERI MINGGU INI (40%)

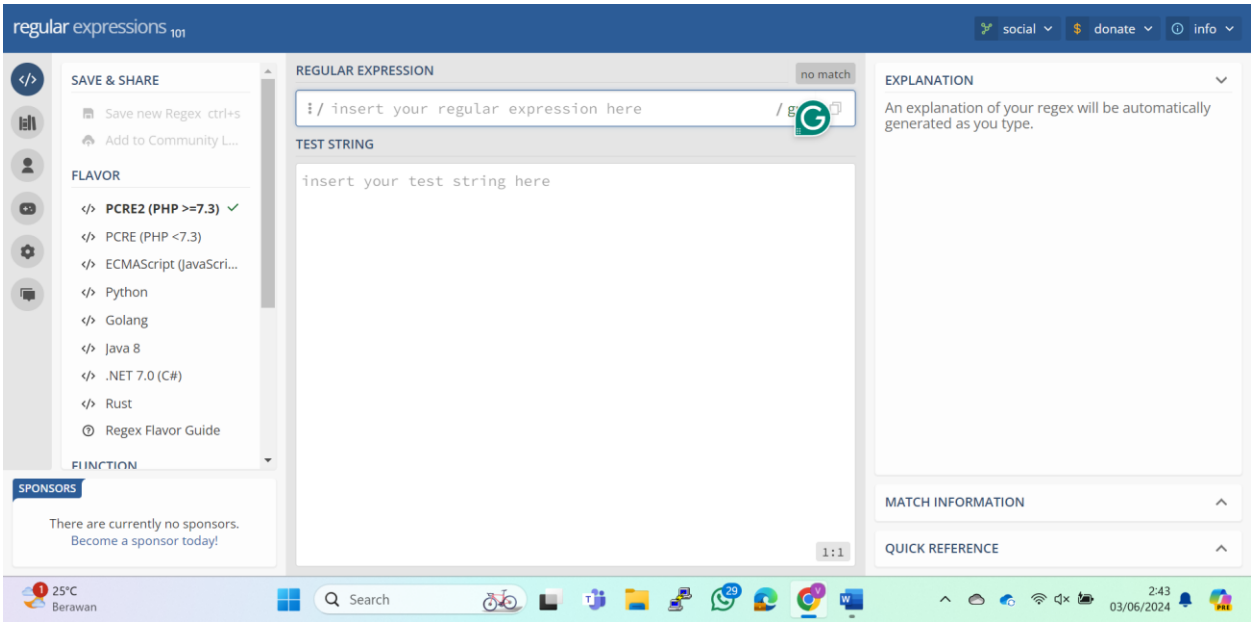
Pada bagian ini, tuliskan kembali semua materi yang telah anda pelajari minggu ini. Sesuaikan penjelasan anda dengan urutan materi yang telah diberikan di saat praktikum. Penjelasan anda harus dilengkapi dengan contoh, gambar/ilustrasi, contoh program (source code) dan outputnya. Idealnya sekitar 5-6 halaman.

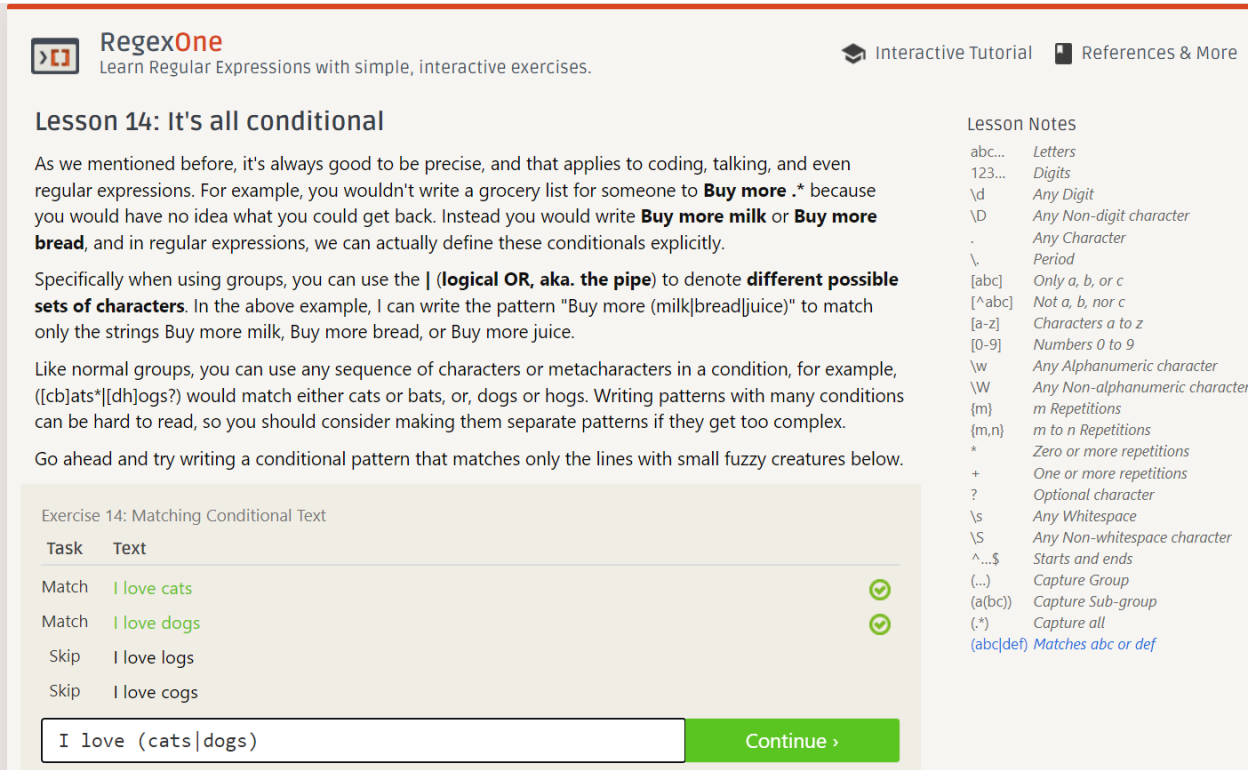
Regular Expression

Pada pengolahan string kita seringkali menemukan pola-pola dalam teksnya dan Regular Expression (Regex) sangat membantu programmer untuk mengolah string yang memiliki pola tertentu di dalamnya dan secara teori Regex akan membantu kita dalam pencarian string dengan pola tertentu, mengganti string dengan pola tertentu, dan menghapus string dengan pola tertentu. Intinya regex membantu dalam parsing string yang selama ini biasanya hanya menggunakan perintah split() dan find() saja.

Untuk dapat memakai fungsi fungsi pengolahan dalam Regex kita harus mencantumkan import re di awal code untuk dapat mengakses library nya kemudia kita dapat memakai beberapa fungsi seperti search() findall() dan lain sebaginya, Regex sendiri termasuk cukup rumit dalam penggunaan nya karena memiliki pola penulisan yang rumit dan belum tentu di support oleh semua Bahasa pemrograman.

Dalam penggunaan Regex kita harus belajar dalam melihat pola pola dalam suatu string dan untuk itu kita harus melatihnya dapat melalui beberapa web seperti Regexone dan Regex 101.





Kedua website tersebut dapat dipakai untuk melatih kemampuan kita dalam meakai Regex untuk memudahkan kita dalam mengolah data string sehingga dapat lebih mudah dan lebih beragam cara pengolahannya.

Contoh penggunaan Regex :

```
import re
handle=open('mbox-short.txt')
count = 0
for line in handle:
    line=line.rstrip()
    if re.search('From:', line):
        count += 1
        print(line)
print("Count: ",count)
```

Output :

From: stephen.marquard@uct.ac.za

From: louis@media.berkeley.edu

From: zqian@umich.edu

From: rjlowe@iupui.edu

From: zqian@umich.edu

From: rjlowe@iupui.edu

From: cwen@iupui.edu

From: cwen@iupui.edu

From: gsilver@umich.edu

From: gsilver@umich.edu

From: zqian@umich.edu

From: gsilver@umich.edu

From: wagnermr@iupui.edu

From: zqian@umich.edu

From: antranig@caret.cam.ac.uk

From: gopal.ramasammycook@gmail.com

From: david.horwitz@uct.ac.za

From: david.horwitz@uct.ac.za

From: david.horwitz@uct.ac.za

From: david.horwitz@uct.ac.za

From: stephen.marquard@uct.ac.za

From: louis@media.berkeley.edu

From: louis@media.berkeley.edu

From: ray@media.berkeley.edu

From: cwen@iupui.edu

From: cwen@iupui.edu

From: cwen@iupui.edu

Count: 27

Penjelasan :

Code tersebut akan membaca file mbox-short.txt kemudian mencari seluruh line yang berawalan ‘From’ lalu di print dan dihitung memakai counter setiap terjadi perulangan.

Contoh lain

```
txt="Sang mata-mata Sg Sing SIANG sedang g memata-matai kasus kaca mata di toko Matahari"
find=re.findall("S\w+g",txt)
search=re.search("Sg",txt)
sub=re.sub("mata","hidung", txt) #me replace
split=re.split("[ma]",txt)
print(find)
print(search)
print(sub)
print(split)
```

Output :

['Sang', 'Sing']

<re.Match object; span=(15, 17), match='Sg'>

Sang hidung-hidung Sg Sing SIANG sedang g mehidung-hidungi kasus kaca hidung di toko Matahari

['S', 'ng ', '', 't', '-', '', 't', ' Sg Sing SIANG sed', 'ng g ', 'e', '', 't', '-', '', 't', 'i k', 'sus k', 'c', ' ', '', 't', ' di toko M', 't', 'h', 'ri']

Dapat kita perhatikan dari contoh diatas terdapat beberapa penggunaan dari Regex

Contoh `find=re.findall("S\w+g",txt)`

Code tersebut berguna untuk mencari data string dari txt yang memiliki pola S huruf g dan hasil akhirnya ialah ['Sang', 'Sing']

Contoh `search=re.search("Sg",txt)`

Code akan mencari bagian maan yang memiliki pola Sg dalam txt

<re.Match object; span=(15, 17), match='Sg'>

Output tersebut keluar saat ditemukan Sg di indeks 15 sampai17 dimana Sg ditemukan

Contoh `sub=re.sub("mata","hidung", txt) #me replace`

Output ini berguna untuk mengganti kata “mata” menjadi kata “hidung” dalam txt

Contoh `split=re.split("[ma]",txt)`

Split yang tadinya hanya memakai spasi atau sebagainya jadi lebih beragam karena dapat memakai Regex dan hal ini dapat menjadikan beberapa varian untuk mengolah string

Syntax/cara penggunaan Regex

Meskipun Reges memeiliki kelebihan dalam pengolahan string yang lebih rumit dan complex yang menyebabkan penggunaan regex cukup rumit contoh nya diantaranya ialah penggunaan symbol simbolnya contoh symbol dalam Regex :

Tabel 14.1: Special Character pada Python

Karakter	Kegunaan	Contoh	Arti Contoh
[]	Kumpulan karakter	"[a-zA-Z]"	1 karakter antara a-z kecil atau A-Z besar
\{}	Karakter dengan arti khusus dan escaped character	\{}d	Angka / digit
.	Karakter apapun kecuali newline	say.n.	Tidak bisa diganti dengan karakter apapun, misal "sayang" akan valid
^	Diawali dengan	^From	Diawali dengan From
$	Dakhiri dengan	this$	Diakhiri dengan kata this
*	0 s/d tak terhingga karakter	\{}d*	ada digit minimal 0 maksimal tak terhingga
?	ada atau tidak (opsional)	\{}d?	Boleh ada atau tidak ada digit sebanyak
+	1 s/d tak terhingga karakter	\{}d+	Minimal 1 s/d tak terhingga karakter
{}	Tepat sebanyak yang ada para {}	\{}d{2}	Ada tepat 2 digit
()	Pengelompokan karakter / pola	(sayalkamu)	saya atau kamu sebagai satu kesatuan
|	atau	\{}d|\{}s	1 digit atau 1 spasi

Tabel 14.2: Escaped Character pada Regex

Special Characters	Kegunaan	Contoh
\b	Digunakan untuk mengetahui apakah suatu pola berada di awal kata atau akhir kata	"R\bin" "Rain\b"
\d	Digunakan untuk mengetahui apakah karakter adalah sebuah digit (0 s/d 9)	\d
\D	Digunakan untuk mengetahui apakah karakter yang bukan digit	\D
\s	Digunakan untuk mengetahui apakah karakter adalah whitespace (spasi, tab, enter)	\s
\S	Digunakan untuk mengetahui apakah karakter adalah BUKAN whitespace (spasi, tab, enter)	\S
\w	Digunakan untuk mengetahui apakah karakter adalah word (a-z, A-Z, 0-9, dan _)	\w
\W	Digunakan untuk mengetahui apakah karakter adalah BUKAN word (a-z, A-Z, 0-9, dan _)	\W
\A	Digunakan untuk mengetahui apakah karakter adalah berada di bagian depan dari kalimat	"\AThe"
\Z	Digunakan untuk mengetahui apakah karakter adalah berada di bagian akhir dari kalimat	"End\Z"

Lesson Notes

abc...	*Letters*
123...	*Digits*
\d	*Any Digit*
\D	*Any Non-digit character*
.	*Any Character*
\.	*Period*
[abc]	*Only a, b, or c*
[^abc]	*Not a, b, nor c*
[a-z]	*Characters a to z*
[0-9]	*Numbers 0 to 9*
\w	*Any Alphanumeric character*
\W	*Any Non-alphanumeric character*
{m}	*m Repetitions*
{m,n}	*m to n Repetitions*
*	*Zero or more repetitions*
+	*One or more repetitions*
?	*Optional character*
\s	*Any Whitespace*
\S	*Any Non-whitespace character*
^...$	*Starts and ends*
(...)	*Capture Group*
(a(bc))	*Capture Sub-group*
(.*)	*Capture all*
(abc|def)	*Matches abc or def*

Kombinasi dari beragam symbol berguna untuk memastikan pola apa yang dipakai dalam penggunaan Regexnya.

BAGIAN 2: LATIHAN MANDIRI (60%)

Pada bagian ini anda menuliskan jawaban dari soal-soal Latihan Mandiri yang ada di modul praktikum. Jawaban anda harus disertai dengan source code, penjelasan dan screenshot output.

SOAL 1

Soal : Anda diminta untuk mencari seluruh teks yang berupa tanggal dengan format YYYY-MM-DD dan kemudian seluruh tanggal tersebut diambil dan ditampilkan kembali dalam format DD-MM-YYYY ditambah dengan perhitungan selisih dengan tanggal sekarang dalam hari.

Contoh soal : Pada tanggal 1945-08-17 Indonesia merdeka. Indonesia memiliki beberapa pahlawan nasional, seperti Pangeran Diponegoro (TL: 1785-11-11), Pattimura (TL: 1783-06-08) dan Ki Hajar Dewantara (1889-05-02).

Input :

```
import re
import datetime
def cari_tanggal(isi):
    x=re.findall("\d\d\d\d-\d\d-\d\d",isi)
    saat_ini=datetime.datetime.now()

    for i in x:
        i=i.split('-')
        Tahun, Bulan, Hari = i
        tanggal=datetime.datetime(int(Tahun), int(Bulan), int(Hari))
        selisih = (saat_ini-tanggal).days
        print(f" {tanggal} {selisih} hari")


cari_tanggal("Pada tanggal 1945-08-17 Indonesia merdeka. Indonesia memiliki
beberapa pahlawan nasional, seperti Pangeran Diponegoro (TL: 1785-11-11),
Pattimura (TL: 1783-06-08) dan Ki Hajar Dewantara (1889-05-02).")
```

Output :

1945-08-17 00:00:00 28780 hari

1785-11-11 00:00:00 87132 hari

1783-06-08 00:00:00 88019 hari

1889-05-02 00:00:00 49340 hari

Penjelasan :

Code ini memakai library dari import re dan import datetime

Pertama tama saya membuat function Bernama cari_tanggal:

Setelah itu say acari memakai re.findall untuk memperoleh semua tanggal yang memiliki pola

```
x=re.findall("\d\d\d\d-\d\d-\d\d",isi)
```

pola tersebut sesuai dengan contohnya yaitu yyyy-mm-dd

setelah memperoleh semua tanggal penting maka kita mabil tanggal hari ini memekai library datetime dengan fungsi

```
saat_ini=datetime.datetime.now()
```

saat_ini akan berisi waktu saat ini kemuadian

```
for i in x:
        i=i.split('-')
        Tahun, Bulan, Hari = i
        tanggal=datetime.datetime(int(Tahun), int(Bulan), int(Hari))
        selisih = (saat_ini-tanggal).days
        print(f" {tanggal} {selisih} hari")
```

code ini berfungsi untuk memisahkan tanggal yang ada di contoh soal dan sudah dicari memakai find all tersebut dipisahkan memakai split("-") setelah itu pecahannya dimasukan dalam variable Tahun, Bulan, dan Hari yang kemudian disatukan kembari dan dijadikan sebgai tanggal memakai fungsi dari library datetime untuk mendapatkan selisih harinya maka tgl hari in dikurangi dengan tanggalnya dan di print agar mengeluarkan output yang sesuai.

SOAL 2

Anda diminta untuk mencari seluruh teks yang berupa email dan kemudian ambil semua username dari email tersebut untuk digenerate password random 8 karakter yang terdiri dari angka dan huruf.

Contoh soal :

Berikut adalah daftar email dan nama pengguna dari mailing list: anton@mail.com dimiliki oleh antonius budi@gmail.co.id dimiliki oleh budi anwari slamet@getnada.com dimiliki oleh slamet slumut matahari@tokopedia.com dimiliki oleh toko matahari

Input :

```
contoh_email= '''
Berikut adalah daftar email dan nama pengguna dari mailing list:
anton@mail.com dimiliki oleh antonius
budi@gmail.co.id dimiliki oleh budi anwari
slamet@getnada.com dimiliki oleh slamet slumut
matahari@tokopedia.com dimiliki oleh toko matahari
'''
```

```
import random
import string
import re
huruf_besar = string.ascii_uppercase
huruf_kecil = string.ascii_lowercase
digit = string.digits
daftar_pass = huruf_besar+ digit+huruf_kecil
panjang_sandi= 8


daftar_email= re.findall(r"(\S+@\S+)", contoh_email)
for email in daftar_email:
    sandi = ''
    for i in range(panjang_sandi):
        sandi += ''.join(random.choices(daftar_pass))
    user = email.split("@")[0]
    print (f"{email} username: {user}, password: {sandi}")
```

output :

anton@mail.com username: anton, password: p8FyVqGo

budi@gmail.co.id username: budi, password: CtVsbO1y

slamet@getnada.com username: slamet, password: l18Coq8X

matahari@tokopedia.com username: matahari, password: ZYOQr8lc

Penjelasan :

Dalam code ini kita perlu memakai beberapa library seperti :

```
import random
import string
import re
```

pertama tam akita buat cara pembentukan password random 8 digit yang terdiri dari digit huruf besar dan huruf kecil memakai code dibawah :

```
huruf_besar = string.ascii_uppercase
huruf_kecil = string.ascii_lowercase
digit = string.digits
daftar_pass = huruf_besar+ digit+huruf_kecil
panjang_sandi= 8
```

pertama kita pakai huruf besar, huruf kecil memakai string.ascii upper dan lower setelah itu digit memakai string.digits dan disatukan semuanya dalam daftas_pass dan kita tentukan Panjang sandinya 8.

Code tersebut nantinya akan dipakai dalam randomize passwordnya

```
daftar_email= re.findall(r"(\S+@\S+)", contoh_email)
for email in daftar_email:
    sandi = ''
    for i in range(panjang_sandi):
        sandi += ''.join(random.choices(daftar_pass))
    user = email.split("@")[0]
    print (f"{email} username: {user}, password: {sandi}")
```

kemudian pada code berikutnya kita akan mengambil email dari contohnya memakai findall yang memiliki pola email dari contoh email diatas.

Setelah itu kita buat per emailnya dan kita buat randomize sandinya dan digabungkan untuk menjadi password setelahnya username didapat dari bagian depan email yang dipisah oleh (@) bagian indekx[0] nya akan menjadi usernamenya

Terakhir kita masukan seluruh datanya seperti emailnya, user dan sandi randomnya sesuai dengan ketentuan soal.