



Laporan Praktikum Algoritma dan Pemrograman

Semester Genap 2023/2024

NIM	71230970
Nama Lengkap	Gregorius Daniel Jodan Perminas
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.

Apa itu Regular Expression?

Dengan metode standar, pengolahan strings cukup rumit. Teknik pengolahan strings yang lebih baik adalah Regular Expression. Regular Expression dapat dikatakan sebagai sebuah tools yang mampu untuk menemukan pola (*pattern*) yang sama dengan pola regex dalam string yang ingin dicari. Regex dapat digunakan untuk:

1. Mencari string dengan pola tertentu
2. Mengganti string dengan pola tertentu
3. Menghapus string dengan pola tertentu

Intinya regex membantu dalam parsing string yang selama ini biasanya hanya menggunakan perintah `split()` dan `find()` saja. Sayangnya, tidak semua bahasa pemrograman mendukung regular expression library. Python merupakan salah satu bahasa yang mendukung library regex dengan cara `import re`.

CONTOH PROGRAM

```
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: rjlwe@iupui.edu
From: zqian@umich.edu
From: rjlwe@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
```

Meta Character, Escaped Character, Set of Character, dan Fungsi Regex pada Library Python

Sebelum menggunakan fungsi regex kita perlu mengetahui beberapa hal berikut ini:

META CHARACTERS

Character	Description	Example	Try it
[]	A set of characters	"[a-m]"	Try it »
\	Signals a special sequence (can also be used to escape special characters)	"\d"	Try it »
.	Any character (except newline character)	"he..o"	Try it »
^	Starts with	"^hello"	Try it »
\$	Ends with	"planet\$"	Try it »
*	Zero or more occurrences	"he.*o"	Try it »
+	One or more occurrences	"he.+o"	Try it »
?	Zero or one occurrences	"he.?o"	Try it »
{}	Exactly the specified number of occurrences	"he.{2}o"	Try it »
	Either or	"falls stays"	Try it »
()	Capture and group		

SPECIAL SEQUENCES

Character	Description	Example	Try it
\A	Returns a match if the specified characters are at the beginning of the string	"\AThe"	Try it »
\b	Returns a match where the specified characters are at the beginning or at the end of a word (the "r" in the beginning is making sure that the string is being treated as a "raw string")	r"\bain" r"ain\b"	Try it » Try it »
\B	Returns a match where the specified characters are present, but NOT at the beginning (or at the end) of a word (the "r" in the beginning is making sure that the string is being treated as a "raw string")	r"\Bain" r"ain\B"	Try it » Try it »
\d	Returns a match where the string contains digits (numbers from 0-9)	"\d"	Try it »
\D	Returns a match where the string DOES NOT contain digits	"\D"	Try it »
\s	Returns a match where the string contains a white space character	"\s"	Try it »
\S	Returns a match where the string DOES NOT contain a white space character	"\S"	Try it »
\w	Returns a match where the string contains any word characters (characters from a to Z, digits from 0-9, and the underscore _ character)	"\w"	Try it »
\W	Returns a match where the string DOES NOT contain any word characters	"\W"	Try it »
\Z	Returns a match if the specified characters are at the end of the string	"Spain\Z"	Try it »

SETS

Set	Description	Try it
[arn]	Returns a match where one of the specified characters (a , r , or n) is present	Try it »
[a-n]	Returns a match for any lower case character, alphabetically between a and n	Try it »
[^arn]	Returns a match for any character EXCEPT a , r , and n	Try it »
[0123]	Returns a match where any of the specified digits (0 , 1 , 2 , or 3) are present	Try it »
[0-9]	Returns a match for any digit between 0 and 9	Try it »
[0-5][0-9]	Returns a match for any two-digit numbers from 00 and 59	Try it »
[a-zA-Z]	Returns a match for any character alphabetically between a and z , lower case OR upper case	Try it »
[+]	In sets, + , * , . , , () , \$, {} has no special meaning, so [+] means: return a match for any + character in the string	Try it »

FUNGSI REGEX PADA PYTHON

Function	Description
<u>findall</u>	Returns a list containing all matches
<u>search</u>	Returns a <u>Match object</u> if there is a match anywhere in the string
<u>split</u>	Returns a list where the string has been split at each match
<u>sub</u>	Replaces one or many matches with a string

Penggunaan Fungsi Findall()

CONTOH PROGRAM 1

```
import re

# Return List dari semua yang match
txt = "Adik saya sedang main di pinggir pantai"
x = re.findall("ai", txt)
print(x)
```

OUTPUT 1

```
['ai', 'ai']
```

CONTOH PROGRAM 2

```
import re
```

```
# Mengecek apakah kata "Portugal" ada di dalam string:
txt = "The rain in Spain stays mainly in the plain"
x = re.findall("Portugal", txt)
print(x)
if (x):
    print("Yes, there is at least one match!")
else:
    print("No match")
```

OUTPUT 2

```
[]
No match
```

Penggunaan Fungsi Search()

CONTOH PROGRAM 1

```
import re

txt = "Oh my god, what a beautiful day!"
x = re.search("\s", txt)
print("The first white-space character is located in position:", x.start())
```

OUTPUT 1

```
The first white-space character is located in position: 2
```

CONTOH PROGRAM 2

```
import re

txt = "The rain in Spain"
x = re.search("Portugal", txt)
print(x)
```

OUTPUT 2

```
None
```

Penggunaan Fungsi Split()

CONTOH PROGRAM

```
import re

txt = "Oh my lord what a beautiful day!"
x = re.split("\s", txt)
```

```
print(x)
y = re.split("\s", txt, 1) # Split 1 kata pertama
print(y)
```

OUTPUT

```
['Oh', 'my', 'lord', 'what', 'a', 'beautiful', 'day!']
['Oh', 'my lord what a beautiful day!']
```

Penggunaan Fungsi Sub()

CONTOH PROGRAM 1

```
import re

txt = "The rain in Spain"
x = re.sub("\s", "9", txt)
print(x)
```

OUTPUT 1

The9rain9in9Spain

CONTOH PROGRAM 2

```
import re

# Mengganti spasi dengan 9 sebanyak 2 kali
txt = "The rain in Spain"
x = re.sub("\s", "9", txt, 2)
print(x)
```

OUTPUT 2

The9rain9in Spain

REFERENSI

[Python RegEx \(w3schools.com\)](https://www.w3schools.com/python/python_regex.asp)

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

PROGRAM LATIHAN MANDIRI 14.1

```
# PROGRAM MENGECEK SELISIH HARI
from datetime import datetime
import re

def selisihTanggal(info):
    tanggalDataFind = re.findall("\d{4}-\d{2}-\d{2}", data)
    tanggalSekarang = datetime.now()
    formatTanggal = tanggalSekarang.strftime("%Y-%m-%d")
    for i in tanggalDataFind:
        tanggalData = datetime.strptime(i, "%Y-%m-%d")
        selisih = tanggalSekarang - tanggalData
        print(i, "00:00:00", "selisih", selisih.days, "hari")

data = "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)"

selisihTanggal(data)
```

OUTPUT PROGRAM

```
1945-08-17 00:00:00 selisih 28778 hari
1785-11-11 00:00:00 selisih 87130 hari
1783-06-08 00:00:00 selisih 88017 hari
1889-05-02 00:00:00 selisih 49338 hari
```

PENJELASAN PROGRAM

- Import Modul dan Deklarasi Fungsi: Baris ini mengimpor modul datetime yang digunakan untuk bekerja dengan tanggal dan waktu. Modul re diimpor supaya bisa mengolah string dengan regex
- Deklarasi Fungsi selisihTanggal: Fungsi ini menerima satu parameter bernama info, yang akan berisi teks yang mengandung tanggal.
- Mengambil Tanggal dari Strings: Baris ini menggunakan regex untuk menemukan semua tanggal dalam string yang memiliki format "YYYY-MM-DD". Kemudian hasil pencariannya akan disimpan dalam variabel tanggalDataFind.
- Mendapatkan Tanggal Saat Ini: Baris ini mendapatkan tanggal saat ini menggunakan datetime.now().

- Mengubah Format Tanggal Saat Ini: Baris ini mengubah format tanggal saat ini menjadi "YYYY-MM-DD" menggunakan strftime("%Y-%m-%d").
- Perulangan untuk Setiap Tanggal dalam data:
 1. Mengubah tanggal tersebut menjadi objek datetime menggunakan datetime.strptime(i, "%Y-%m-%d").
 2. Menghitung selisih hari antara tanggal saat ini dan tanggal tersebut.
 3. Program mencetak selisih hari antara tanggal-tanggal yang ditemukan dalam string dengan tanggal saat ini.

SOAL 2

PROGRAM LATIHAN MANDIRI 14.2

```
import re
import random
import string

def buatUsername(email):
    # Membuat username dari alamat email
    username = email.split("@")[0]
    return username

def buatPassword():
    characters = string.ascii_letters + string.digits
    password = ''.join(random.choice(characters) for i in range(8))
    return password

def prosesData (data):
    # Mengambil Email dari Data
    alamatEmail = re.findall("\S+@\S+", data)
    for i in alamatEmail:
        username = buatUsername(i)
        password = buatPassword()
        print(f"{i} username: {username}, password: {password}")

data = "anton@mail.com dimiliki oleh antonius\nbudi@gmail.co.id dimiliki oleh\nbudi anwari\nslamet@getnada.com dimiliki oleh slamet\nslumut\nmatahari@tokopedia.com dimiliki oleh toko matahari"

prosesData(data)
```

OUTPUT PROGRAM

```
anton@mail.com username: anton, password: HtDwa7zQ
budi@gmail.co.id username: budi, password: JbMDgyAN
slamet@getnada.com username: slamet, password: Ac5xGXxz
matahari@tokopedia.com username: matahari, password: Mzfv4W0o
```


PENJELASAN PROGRAM

- Deklarasi Fungsi `buatUsername(email)`: Fungsi ini menerima satu parameter, yaitu email. Tujuan fungsi ini adalah untuk membuat username dari alamat email. Pertama-tama fungsi akan membagi alamat email berdasarkan karakter “@” menggunakan `split("@")`. Kemudian mengambil bagian pertama hasil pemisahan (sebelum karakter “@”) sebagai username.
- Deklarasi Fungsi `buatPassword()`: Fungsi ini tidak memiliki parameter dan tujuannya hanya membuat password acak. Fungsi akan menggabungkan karakter huruf (baik huruf besar maupun kecil) dan angka menggunakan `string.ascii_letters + string.digits`. Kemudian, fungsi memilih karakter secara acak sebanyak 8 kali menggunakan `random.choice(characters)` dan menggabungkannya menjadi password.
- Deklarasi Fungsi `prosesData(data)`: Fungsi ini memiliki satu parameter, yaitu data. Tujuan dari fungsi ini adalah untuk memproses data yang mengandung alamat email. Fungsi menggunakan regex untuk menemukan semua alamat email dalam teks data. Setelah menemukan alamat email, fungsi:
 1. Membuat username dari alamat email menggunakan fungsi `buatUsername(i)`.
 2. Membuat password acak menggunakan fungsi `buatPassword()`.
 3. Mencetak alamat email, username, dan password yang dihasilkan.