#### INSTITUTO POLITÉCNICO NACIONAL

# ESCUELA SUPERIOR DE CÓMPUTO

# TEORÍA COMPUTACIONAL

#### PRÁCTICA 04

ALUMNO: HERNÁNDEZ CASTELLANOS CÉSAR URIEL

PROFESOR: ROSAS TRIGUEROS JORGE LUIS

FECHA DE REALIZACIÓN: 7 DE SEPTIEMBRE DEL 2017

FECHA DE ENTREGA: 10 DE SEPTIEMBRE DEL 2017

#### Marco teórico.

Una expresión regular es un modelo con el que el motor de expresiones regulares intenta buscar una coincidencia en el texto de entrada. Un modelo consta de uno o más literales de carácter, operadores o estructuras.

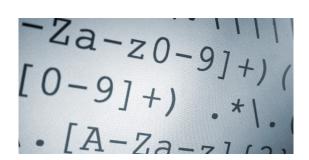


Imagen 1.0 Expresión regular.

#### **Componentes de las Expresiones Regulares**

Las expresiones regulares son un mini lenguaje en sí mismo, por lo que para poder utilizarlas eficientemente primero debemos entender los componentes de su sintaxis; ellos son:

**Literales:** Cualquier caracter se encuentra a sí mismo, a menos que se trate de un metacaracter con significado especial. Una serie de caracteres encuentra esa misma serie en el texto de entrada, por lo tanto la plantilla "raul" encontrará todas las apariciones de "raul" en el texto que procesamos.

**Secuencias de escape:** La sintaxis de las expresiones regulares nos permite utilizar las secuencias de escape que ya conocemos de otros lenguajes de programación para esos casos especiales como ser finales de línea, tabs, barras diagonales, etc.

Clases de caracteres: Se pueden especificar clases de caracteres encerrando una lista de caracteres entre corchetes [], la que que encontrará uno cualquiera de los caracteres de la lista. Si el primer símbolo después del "[" es "^", la clase encuentra cualquier caracter que no está en la lista.

**Metacaracteres:** Los metacaracteres son caracteres especiales que son la esencia de las expresiones regulares. Como son sumamente importantes para entender la sintaxis de las expresiones regulares y existen diferentes tipos, voy a dedicar una sección a explicarlos un poco más en detalle.

#### Expresiones Regulares. Módulo re

#### Módulo re

Para utilizar Expresiones Regulares, Python provee el módulo re. Importando este módulo podemos crear objetos de tipo patrón y generar objetos tipo matcher, que son los que contienen la información de la coincidencia del patrón en la cadena.

### Creando un patrón

Para crear un objeto patrón, se debe importar el módulo re y utilizamos la función compile:

```
import re

patron = re.compile('a[3-5]+') # coincide con una letra, seguida de al menos 1
dígito entre 3 y 5
```

#### Buscar el patrón en la cadena

Para buscar un patrón en una cadena, Python provee los métodos search y match. La diferencia entre ambos es que, mientras search busca en la cadena alguna ocurrencia del patrón, match devuelve Nonesi la ocurrencia no se da al principio de la cadena:

```
cadena = 'a44453'
patron.match(cadena) # <_sre.SRE_Match object at 0x02303BF0>
patron.search(cadena) # <_sre.SRE_Match object at 0x02303C28>
cadena = 'ba3455' # la coincidencia no está al principio!
patron.search(cadena) # <_sre.SRE_Match object at 0x02303BF0>
print patron.match(cadena) # None
```

### Reemplazo de cadenas

Similar a la combinación search + expand, existe el método <u>sub</u> cuya función es encontrar todas las coincidencias de un patrón y sustituirlas por una cadena. Este recibe dos parámetros: el primero es la cadena con la que se sustituirá el patrón y el segundo es la cadena sobre la que queremos aplicar la sustitución.

```
patron.sub("X", 'a455 a333b435') # sustituye todas las ocurrencias por X 'X XX'
patron.sub("LETRA(\g<1>), NUMERO(\g<2>)", 'a455 a333b435') # El reemplazo depende
de lo que se capture LETRA(a), NUMERO(455) LETRA(a), NUMERO(333)LETRA(b),
NUMERO(435)'
```

### Material y equipo.

El material utilizado en la práctica es el siguiente:

Herramientas de software:

- \* Oracle VM VirtualBox
- \* Ubuntu 17.04
- \* Python 3.6.2
- \* eric6 Web Browser (QtWebKit)
- \* Online regez texter and debugger.

Herramientas de hardware:

\* Computadora personal

#### Desarrollo de la práctica.

La práctica número cuatro de la unidad de aprendizaje teoría computacional, consistió básicamente en practicas nuestras habilidades al formar expresiones regulares.

La primera sección de la práctica se encontradaba fragmentada en cuatro ejercicios, los cuales iban aumentando gradualmente de dificultad, el propósito de los ejercicios, era el de encontrar el patrón que representaba la parte positiva, la cual debería de ser iluminada de rojo para proseguir con el siguiente nivel.

La segunda parte de la práctica consistió el elaborar diez ejercicios de expresiones regulares, con la particularidad que se usó la librería re, para poder trabajar con expresiones regulares en Python.

Las dificultades en la práctica fue principalmente lo intimidante que puede llegar a ser las expresiones regulares a primera impresión, esto se solucionó investigando en diferentes fuentes los diferentes comandos que nos brindan, para poder generar nuestros propios patrones, también una de las dificultades fue el gran periodo de tiempo invertido en los ejercicios, en los cuales algunas veces se avanzaba de forma fluida y en otras pasaban las horas sin encontrar solución alguna.

A continuación se presentan los resultados obtenidos en la práctica:

# Diagramas, gráficas y pantallas.

# **Exercise 1**

Enter a regexp that matches all the items in the first column (positive examples) but none of those in the second (negative examples). When you press "submit", you will see what matched.



Imagen 1.1 - Primer ejercicio.

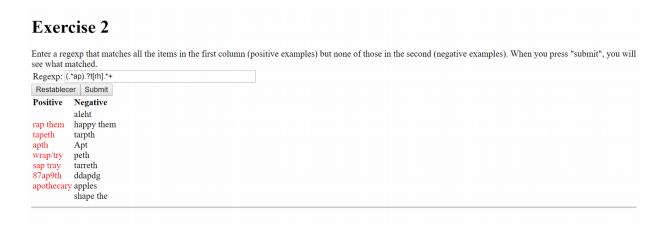


Imagen 1.2 Segundo Ejercicio.

# Exercise 3 Enter a regexp that matches all the items in the first column (positive examples) but none of those in the second (negative examples). When you press "submit", you will see what matched. Regexp: Positive Negative affgfking fgok rafgkahe a fgk bafghk affgm baffgkit affffhk affgfking fgok rafgkahe afg.K bafghk aff gm baffg kit affffhgk

Imagen 1.3 Tercer ejercicio.

#### **Exercise 4: Finding sentence breaks**

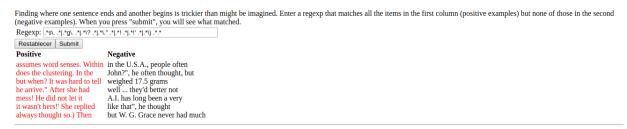


Imagen 1.4 Cuarto Ejercicio.

```
uriel@Uriel-PC:~/Escritorio$ cd teoriac
uriel@Uriel-PC:~/Escritorio/teoriac$ ls
Programa01 Programa03 Programa05 Programa07 Programa09
Programa02 Programa04 Programa06 Programa08 Programa10
uriel@Uriel-PC:~/Escritorio/teoriac$ cd Programa01
uriel@Uriel-PC:~/Escritorio/teoriac/Programa01$ python regexpy.py
['aaksdsd', 'aaoaksb', 'aalsdlskl', 'aabasasab', 'aabbbbb', 'aabsa', 'aaidjsdisjidjsijidjsijdsisdbbb']
uriel@Uriel-PC:~/Escritorio/teoriac/Programa01$
```

Imagen 1.5 Programa número uno.

- Cadenas que empiecen con dos a's y tengan cero o más b's.

```
Abrir ▼ 1.1

aaksdsd
aaoaksb
aalsdlskl
HOLA
aabasasab
aabbbbb
holita
aabsa
aaidjsdisjidjsijdsisdbbb
```

Imagen 1.6 Contenido del fichero.

```
L daksusu , daudksu , datsutskt , daudasasau , daudubu , dausa , datujsutsjtujstjustsubbu j
uriel@Uriel-PC:~/Escritorio/teoriac$ cd Programa02
uriel@Uriel-PC:~/Escritorio/teoriac/Programa02$ python regexpy.py
['aaoaksb', 'aabasasab', 'aabbbbb', 'aabsa', 'aaidjsdisjidjsijidjsijdsisdbbb', 'aabokd']
uriel@Uriel-PC:~/Escritorio/teoriac/Programa02$
```

Imagen 1.7 Programa número dos.

- Cadenas que empiecen con dos a's y tengan una o más b's.

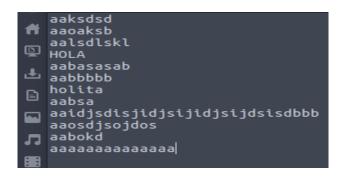


Imagen 1.8 Contenido del fichero.

```
uriel@Uriel-PC:~/Escritorio/teoriac$ cd Programa03
uriel@Uriel-PC:~/Escritorio/teoriac$ cd Programa03
uriel@Uriel-PC:~/Escritorio/teoriac/Programa03$ ls
a.txt regexpy.py
uriel@Uriel-PC:~/Escritorio/teoriac/Programa03$ python regexpy.py
['aei_oudsa', '_saij', '_eeee', '_', '_aaaa', '_ewew', '_', '_aaaa', '_eee', '_etfx', '_qqqq', '_tr', '_wqqssss', '_rrty', '_sddd', '_eeee']
uriel@Uriel-PC:~/Escritorio/teoriac/Programa03$
```

Imagen 1.9 Programa número tres.

- Cadenas en donde dos subcadenas de letras minúsculas estén unidas por un guión bajo.

```
holaAaei_oudsaA_saij_eeee_ASSSS_aaaa_ewew_A_aaa_eee_etfx_qqqq_tr_wqqssss_rrty_sddd_eeee
```

Imagen 2.0 Contenido del fichero.

```
urtelgUrtel-PC:-/Escrettor()teortac/Programa65 cd ...
urtelgurtel-PC:-/Escrettor()teortac/Programa65 ts ...
urtelgurtel-PC:-/Escrettor()teortac/Programa65 ts ...
urtelgurtel-PC:-/Escrettor()teortac/Programa65 ts ...
urtelgurtel-PC:-/Escrettor()teortac/Programa65 to ...
urtelgurtel-PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/PC:-/Escrettor()teortac/P
```

Imagen 2.1 Programa número cuatro.

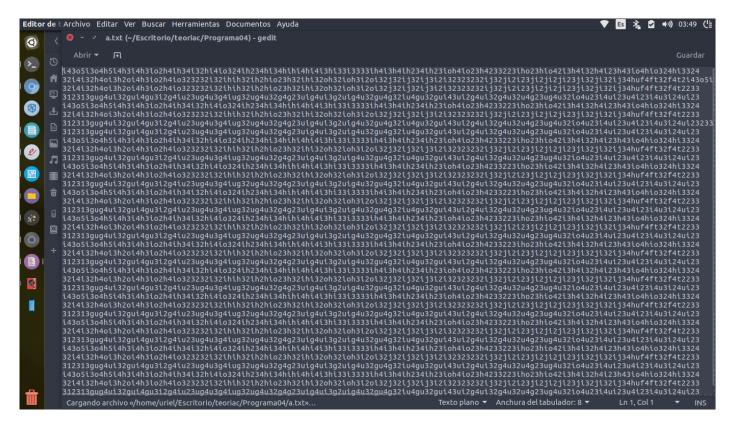


Imagen 2.2 Contenido del fichero.



Imagen 2.3 Programa número cinco.

```
por y le preguntó qué había que haccer para hacerse or. El Dr de broma, le contestó que debería de comprar un abeccdario, encargar un par de tra jes y poner un letrero que dijera "Dr Sabelotodo". Después de esto o, sólo tendría que esperar a que los cilentes fuesen llegando.

El pobre lenador no entendió la broma y vendió su mula y con el dinero compró un abecedario, dos bonos trajes y una bata blanca de Dr, además de poner en la puerta de su casa un cartel como le había indicado el Dr. Todo esto se los peco a su mujer y le dijo que era para ganar nucho diner o.

**Abian pasado pocos días cuando un señor llegó preguntando por el Dr Sabelotodo. Le dijo al leñador que le habían robado todo su dinero y este le aseguro que le explicaria la forma en que podría recuperario. El cilente le invitó a comer a su mansión y a su mujer también. Una vez en la lu josa casa, apareció un criado con el primer plato, dispuesto a servirlo y el leñador se puso tan contento que le dijo su mujer e que llegaba el primero, refritendose al primer plato, pero el criado que era un ladrón se dio por aludido y avisó a los otros, advirtiendo que el Dr ya sabía la verdad, que ellos habían robado dinero. Al aparecer el segundo criado pasó lo nismo y volvió la cocina muy asustado. Lo mismo sucedió con el tercer criado. El amo pidió al Dr sabelotodo que advivar a lo que cuarto criado traia una fuente. Y cono eran cangrejos lo que había la fuente, el criado ul lamó al falso br a la cocina y los ladrones le confesaron que eran los autores del robo y dijeron que estaban arrepentidos y pensaban devolver el dinero.

**Volvió Cangrejo al conedor y le dijo al amo que iba a leer en su abecedario donde estaba dinero robado. Sacó su libro, revisó algunas de las pág inas y por fín dijo al caballero el sitio exacto donde estaba su dinero. Como hacer todo, el Dr Sabelotodo recibió una gran recompensa, además de otra recompensa por parte de los criados y así fue como el pobre leñador se huso rico y nunca más tuvo necesidad de ir al bosque a cortar leña vivió m
```

Imagen 2.3 Programa número seis.



Imagen 2.4 Programa número siete.

# pythex

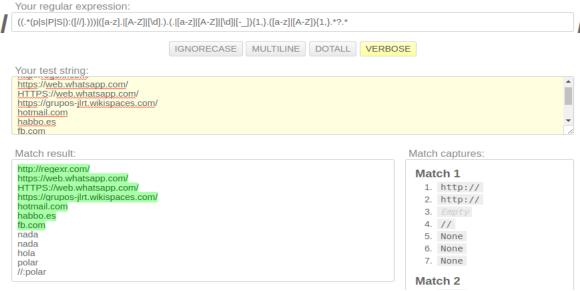


Imagen 2.5 Programa número ocho.

```
Expression share save flags

/([a-z]|[A-Z]|[0-9])*.@.([a-z]|[A-Z]|[0-9])*.*/g

Text

uriel12009@hotmail.com
uriel
mauricio@gmail.com
nunussa
holaco@gmail.com
```

Imagen 2.6 Programa número nueve.

Imagen 2.7 Pograma número diez.

# **Expresiones regulares usadas:**

```
1-. aa.*b{0,}?.*
2-. aa.*b{1,}?.*
3-. [a-z]*_[a-z]*
4-. [0-9]{1,3}
5-. (p[^p\s]*){2}
6-. Doctor
7-. [\s,.]
8-. ((.*(p|s|P|S|):([//].)))|([a-z].|[A-Z]|[\d].).(.|[a-z]|[A-Z]|[\d]|[-_]){1,}.([a-z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]|[A-Z]
```

Nota: El código de los programas se anexó en el comprimido de la práctica.

## **Conclusiones** y recomendaciones.

Las expresiones regulares resultan una herramienta poderosa para un programador, ya que reduce el tiempo de desarrollo de manera considerable, además de contar con un código más compacto.

La resolución de las expresiones regulares no es única, lo que pude percatarme que para dar resolución a estos problemas la mejor técnica a emplear es la divide y vencerás.

#### Referencias

- [1]"Lenguaje de expresiones regulares Referencia rápida", *Msdn.microsoft.com*, 2017. [Online]. Available: https://msdn.microsoft.com/es-es/library/az24scfc(v=vs.110).aspx. [Accessed: 10- Sep- 2017].
- [2]"Regular expressions | Sketch Engine", *Sketchengine.co.uk*, 2017. [Online]. Available: https://www.sketchengine.co.uk/user-guide/user-manual/concordance-introduction/regular-expressions/. [Accessed: 10- Sep- 2017].
- [3]"Guía de expresiones regulares en Python", *platzi.com*, 2017. [Online]. Available: https://platzi.com/blog/expresiones-regulares-python/. [Accessed: 10- Sep- 2017].
- [4]S. Programacion en Castellano, "Expresiones regulares en Python", Programación en Castellano., 2017. [Online]. Available: http://programacion.net/articulo/expresiones\_regulares\_en\_python\_1436. [Accessed: 10-Sep- 2017].
- [5]"Expresiones regulares en python ChuWiki", Chuwiki.chuidiang.org, 2017. [Online]. Available: http://chuwiki.chuidiang.org/index.php? title=Expresiones\_regulares\_en\_python. [Accessed: 10- Sep- 2017].
- [6]"Regex mediante ejemplos", *Pybonacci*, 2017. [Online]. Available: https://pybonacci.es/2013/02/21/regex-mediante-ejemplos/. [Accessed: 10- Sep- 2017].