

Programmation PYTHON

Cours 6

Nassim ZELLAL

2020/2021

Expressions régulières - 2

La mémorisation de modèle - référence arrière

Les parenthèses **()** : mémorisation (référence arrière).

- `ligne="eleve3 au poste3"`

- `if re.search("(\\b).+poste\\b1",ligne):`

`#on peut ajouter le préfixe « r » et écrire (r"(\\b).+poste\\b1",ligne):`

- `print("ok")`

- `#-----#`

- `ligne="eleve3 au poste5"`

- `if re.search("(\\b).+poste\\b1",ligne):`

`#on peut ajouter le préfixe « r » et écrire (r"(\\b).+poste\\b1",ligne):`

- `print("ok")`

La mémorisation de modèle - group()

- Les parenthèses **()** s'utilisent avec la méthode **group()** pour extraire la sous-chaîne de caractères placée entre ces parenthèses.
- `import re`
- `ligne="eleve3 au poste5"`
- `x=re.search("(.+)(\d)(.+)(po(.+))(\d)",ligne)`
- `print(x.group(6)+" "+x.group(5)+" "+x.group(4)+" "+x.group(3)+" "+x.group(2)+" "+x.group(1))`

La mémorisation de modèle - group()

```
5 ste poste au 3 eleve
```

La mémorisation de modèle - groups()

- La méthode **groups()** retourne un tuple contenant tous les sous-groupes reconnus par l'expression régulière.
- `import re`
- `m = re.match("(\\d+)\\.(\\d+)", "24.1632")`
- #il n'est pas nécessaire de déspecialiser le point (\\.) dans cet exemple
- `print(m.groups())`

La mémorisation de modèle - groups()

```
C:\Users\user\Desktop>test.py  
( '24' , '1632' )
```

Expressions régulières - regroupement

- `import re,sys`
- `var=open(sys.argv[1],'r')`
- `res=var.readlines()`
- `for i in range(len(res)):`
 - `if re.search(r"\b(homme|femme)\b",res[i]):`
 - `print(res[i],end="")`
- `var.close()`

1 Vers midi, un homme de haute taille, vêtu d'un costume défraîchi et d'une cravate, arriva
2 Ils regardaient l'homme debout sur le banc et écoutaient le texte malgré eux.
3 Dans la routine prodigieusement lassante de la salle d'attente, cet homme qui lisait à vo:
4 L'homme, de sa grosse voix de ténor aux accents faussement pathétiques, déversait ces mon
assise.
5 Et l'homme qui lisait ainsi, sur un banc vert de la gare, ne leur annonçait rien qui fût :
6 Lorsqu'elles virent l'homme en train de lire, elles échangèrent un regard et sentirent tou
7 La seconde était une femme d'action.
8 L'homme semblait ne pas avoir entendu la question.
9 La femme de ménage fit un pas en arrière, mit ses mains sur ses hanches, bomba le torse et
10 La femme de ménage traîna le banc jusqu'à son emplacement initial, dans le prolongement de
11 Elle n'avait pas l'intention d'écouter ce que l'homme lisait.
12 « Il y a un homme qui lit à haute voix dans la salle d'attente ! »
13 Dans le coin le plus reculé se tenait bel et bien un homme de grande taille, en train de :
moments les plus calmes, ses paroles parvenaient même jusqu'aux oreilles du chef de gare.
14 « Il y a un homme qui lit un livre à haute voix dans la salle d'attente.
15 – C'est cela même ! » homme avare
16 « Que faire lorsqu'un homme ne crache pas et ne se bat pas dans la salle d'attente d'une
17 Après cette entrée en matière claire et directe, il dut fournir à l'homme de loi des expl:
18 Il pensa qu'il pourrait peut-être simplement aller voir l'homme et lui dire :
19 Au fond, il n'y avait peut-être rien de gênant dans le fait qu'un homme lise dans la salle
20 « Nous avons dans notre salle d'attente un homme qui lit à haute voix.
21 – C'est un homme sérieux, il lit un livre, ajouta le surveillant.
22 Une fois seul, le chef de gare alla se poster à la porte de la salle d'attente pour obser
23 Et puis il avait une femme et des enfants.
24 Et il y eut bientôt sur le bureau toutes sortes de renseignements et de photographies, ai
25 l'homme lisait toujours.
26 Mais ce que le chef de gare ne savait pas, c'est que dans la salle d'attente était resté
lecteur n'y faisait pas des ajouts de son cru.
27 Sa journée de travail touchait à sa fin, et l'homme lisait toujours.
28 Une grande partie des gens présents écoutaient la voix de ténor de l'homme et accueillaie
qu'ils sortaient de sa bouche.
29 Et le chef ne savait toujours pas que, parmi les auditeurs, se trouvait un homme digne de

1 Vers midi, un homme de haute taille, vêtu d'un costume défraîchi et d'une cravate, arriva dans la salle d'attente de la gare.
2 Ils regardaient l'homme debout sur le banc et écoutaient le texte malgré eux.
3 Dans la routine prodigieusement lassante de la salle d'attente, cet homme qui lisait à voix haute était un événement aussi troubla
4 L'homme, de sa grosse voix de ténor aux accents faussement pathétiques, déversait ces mondes et ces passions dans le brouhaha de l'assemblée.
5 Et l'homme qui lisait ainsi, sur un banc vert de la gare, ne leur annonçait rien qui fût indispensable à leur existence.
6 Lorsqu'elles virent l'homme en train de lire, elles échangèrent un regard et sentirent toutes deux qu'il s'agissait là d'une attitude
7 La seconde était une femme d'action.
8 L'homme semblait ne pas avoir entendu la question.
9 La femme de ménage fit un pas en arrière, mit ses mains sur ses hanches, bomba le torse et adopta une pose qu'elle devait juger mauvaise
10 La femme de ménage traîna le banc jusqu'à son emplacement initial, dans le prolongement des autres, puis considéra la salle d'un air
11 Elle n'avait pas l'intention d'écouter ce que l'homme lisait.
12 « Il y a un homme qui lit à haute voix dans la salle d'attente ! »
13 Dans le coin le plus reculé se tenait bel et bien un homme de grande taille, en train de lire un livre à haute voix, le visage tous
14 moments les plus calmes, ses paroles parvenaient même jusqu'aux oreilles du chef de gare.
14 « Il y a un homme qui lit un livre à haute voix dans la salle d'attente.
15 – C'est cela même ! » homme avare
16 « Que faire lorsqu'un homme ne crache pas et ne se bat pas dans la salle d'attente d'une gare, mais lit un livre à haute voix aux
17 Après cette entrée en matière claire et directe, il dut fournir à l'homme de loi des explications longues et embrouillées.
18 Il pensa qu'il pourrait peut-être simplement aller voir l'homme et lui dire :
19 Au fond, il n'y avait peut-être rien de gênant dans le fait qu'un homme lise dans la salle d'attente, mais tout de même, il valait
20 « Nous avons dans notre salle d'attente un homme qui lit à haute voix.
21 – C'est un homme sérieux, il lit un livre, ajouta le surveillant.
22 Une fois seul, le chef de gare alla se poster à la porte de la salle d'attente pour observer l'homme en cachette.
23 Et puis il avait une femme et des enfants.
24 Et il y eut bientôt sur le bureau toutes sortes de renseignements et de photographies, ainsi qu'un exemplaire du livre que l'homme
25 l'homme lisait toujours.
26 Mais ce que le chef de gare ne savait pas, c'est que dans la salle d'attente était resté un homme qui suivait le texte livre en main
27 lecteur n'y faisait pas des ajouts de son cru.
27 Sa journée de travail touchait à sa fin, et l'homme lisait toujours.
28 Une grande partie des gens présents écoutaient la voix de ténor de l'homme et accueillaient avec un naturel parfait les idées et les
29 tels qu'ils sortaient de sa bouche.
29 Et le chef ne savait toujours pas que, parmi les auditeurs, se trouvait un homme digne de confiance qui veillait, inébranlable, se

Expressions régulières - flags (arguments optionnels)

- re.**I** : recherche insensible à la casse.
- re.**S** : traiter la chaîne comme une ligne simple.
(avec l'argument optionnel re.**S**, le caractère `.` reconnaît aussi les sauts de ligne).

Expressions régulières - flags (arguments optionnels)

Tout · le · monde · aime · le · chocolat · ! · Le · chocolat · est · une · source · de · magnésium · ! **CRLE**

Certains · sont · allergiques · au · chocolat · ! **CRLE**

Chocolat · noir. **CRLE**

CRLE

- import sys, re
 - var=open(sys.argv[1], 'r')
 - var2=open("res.txt", 'w')
 - a=1
 - for i in var.readlines():
 - x=re.search("!(. \$)", i, re.**S**)
 - if **x!=None** :#éviter les objets de type "None" (**NoneType object**) /ou bien if x:
 - var2.write("Ligne "+str(a)+" résultat : "+x.group(1))
 - a+=1
- Extraction des sauts de ligne précédés d'un point d'exclamation.
- Un objet "None" est incompatible avec la méthode group().

Expressions régulières - flags (arguments optionnels)

```
1 Ligne 1 résultat : CR LF  
2 Ligne 2 résultat : CR LF
```

Expressions régulières - flags (arguments optionnels)

Tout · le · monde · aime · le · chocolat · ! · Le · chocolat · est · une · source · de · magnésium · ! **CR LF**

Certains · sont · allergiques · au · chocolat · ! **CR LF**

Chocolat · noir. **CR LF**

CR LF

- import sys, re
- var=open(sys.argv[1], 'r')
- var2=open("res.txt", 'w')
- for i in var.readlines():
 - x=re.search("(chocolat)", i, re. **I**)
 - if x: #éviter les objets de type "None"
 - var2.write("Résultat : "+x.group(1)+"\n")

Expressions régulières - flags (arguments optionnels)

1	Résultat : chocolat
2	Résultat : chocolat
3	Résultat : Chocolat

Expressions régulières - flags (arguments optionnels)

Tout le monde aime le chocolat ! Le chocolat est une source de magnésium ! **CRLF**

Certains sont allergiques au chocolat ! **CRLF**

Chocolat noir. **CRLF**

CRLF

- import sys, re
- var=open(sys.argv[1], 'r')
- var2=open("res.txt", 'w')
- for i in var.readlines():
 - x=re.findall("chocolat", i, re.I)
 - for c in x:
 - var2.write("Résultat : "+c+"\n")

Expressions régulières - flags (arguments optionnels)

1	Résultat : chocolat
2	Résultat : chocolat
3	Résultat : chocolat
4	Résultat : Chocolat

Expressions régulières - substitution

```
import sys,re
var=open(sys.argv[1],'r')
var2=open("res.txt",'w')
for i in var.readlines():
    x=re.sub("chocolat","pain",i,flags=re.I)
    var2.write(x)
```

#Avec `sub()` on doit écrire « flags » avant un flag.


Expressions régulières - substitution

```
1 Tout le monde aime le pain ! Le pain est une source de magnésium !  
2 Certains sont allergiques au pain !  
3 pain noir.  
4
```

Expressions régulières - look-around

- **(?=expression) regarde-devant positif (look ahead)**
- **(?!expression) regarde-devant négatif (look ahead)**
- **(?<=expression) regarde-derrière positif (look behind)**
- **(?<!expression) regarde-derrière négatif (look behind)**

Expressions régulières - look-around

- `import re`
- `x = "J'ai mangé le croissantqui était sur la table"`
- `y = "I catch the red fly"`
 Il y a une tabulation ici.
- `z = "Je m'appelle Tom-cat "`
- `xx=re.search("(croissant)(?!\\s)",x)`
- `print(xx.group(1))` # reconnaît croissant
- `yy=re.search("(?<=\\t)(.+)",y)`
- `print(yy.group(1))` #reconnaît 'fly'
- `zz=re.search("(?<=-)(cat)(?=\\s)",z)`
- `print(zz.group(1))`# reconnaît 'cat' dans 'Tom-cat'

Exercice 1

- Modifier le script ci-dessous (slide 12), afin qu'il puisse extraire tous les sauts de ligne du fichier « flags.txt ». Proposez trois solutions.
- `import sys,re`
- `var=open(sys.argv[1],'r')`
- `var2=open("res.txt",'w')`
- `a=1`
- `for i in var.readlines():`
- `x=re.search("!(. $)",i,re.S)`
- `if x:`
- `var2.write("Ligne "+str(a)+" résultat :
"+x.group(1))`
- `a+=1`

Exercice 2

- Modifier le script ci-dessous (slide 16), afin qu'il puisse extraire toutes les occurrences de « chocolat » du fichier « flags.txt » sans utiliser les 2 boucles for.
- `import sys,re`
- `var=open(sys.argv[1],'r')`
- `var2=open("res.txt",'w')`
- `res=var.readlines()`
- for `i in res:`
- `x=re.findall("chocolat",i,re.I)`
- for `c in x:`
- `var2.write("Résultat : "+c+"\n")`

Exercice 3

- 1/ Faire un script qui prend en entrée "text-b.txt" encodé en UTF-8 du cours 6;
- 2/ Trier les tokens par ordre alphabétique décroissant et afficher ce tri dans un fichier;
- 3/ Afficher dans un autre fichier :
 - 3.1/ le nombre de tokens;
 - 3.2/ les phrases exclamatives;
 - 3.3/ les phrases interrogatives;
 - 3.4/ afficher toutes les occurrences du token « homme » qui sont suivies d'un verbe ;