SPLN MI-EI

Exame, 28 de Junho de 2018

Questão 1 () Escreva da forma mais simples e elegante que conseguir:

a). uma função python que calcule a reunião de multi-conjuntos: $A \cup B$; Sugestão: represente os multi-conjuntos como dicionários

```
bagunion(\{a:3,c:3,e:6\},\{a:2,c:8,d:9\}) = \{a:5,c:11,e:6,d:9\}
```

b). Construa um programa python que dado um número representando uma quantia, (ex 12345678\$) produza uma saída contendo apenas as quantias, mas formatadas de modo semelhante ao que aparece nos documentos oficiais: (ex 12.345.678\$).

Questão 2 () Um dos modos habituais de distinguir entre duas variantes de línguas (ex pt_PT pt_BR) é o método das "listas negras": que envolve *arranjar* uma lista de palavras que só aparece (ou aparece muito mais vezes) numa das variantes, e contar o número de indícios para cada lado.

```
pt_BR trem
pt_PT comboio
pt_PT polaco
pt_BR galera
```

- a). Faça programas que dado um ficheiro (lista negra) contendo uma lista como a apresentada, e dado um texto, indique a que variante pertence.
- b). Faça um programa ou apresente um algoritmo que dado um par de textos grandes em PT_pt e outro em PT_br construa um ficheiro lista-negra.

Questão 3 () Considere o seguinte exemplo de um texto anotado com freeling, shallow parsing em formato conll:

```
44 o
                                     DAOMSO
                                               (sn:45(espec-ms:44(j-ms:44))
1
                                     NCFS000
                                               (grup-nom-fm:45(n-fm:45))
    45 capital
                       capital
2
                                     AQOMSOO
                                               (s-adj:46(s-a-ms:46(a-ms:46))))
    46 estrangeiro
                       estrangeiro
3
                                     VMIP3S0
                                               (grup-verb:47(verb:47))
    47 fica
4
                       ficar
    48 com
                                     SP
                                               (grup-sp:48(prep:48)
                       com
5
    49 um
                                     DIOMSO
                                               (sn:51(espec-ms:49(indef-ms:49))
                       um
6
    50 major
                       maior
                                     ADOCSOO
                                               (grup-nom-ms:51(s-a-ms:50(a-ms:50))
7
                       controlo
                                     NCMS000
                                               (grup-nom-ms:51(n-ms:51)))))
    51 controlo
8
    52 de
                                     SP
                                               (sp-de:52
                       de
    53 um
                                     DIOMSO
                                               (sn:54(espec-ms:53(indef-ms:53))
10
                       um
    54 sistema
                       sistema
                                     NCMS000
                                               (grup-nom-ms:54(n-ms:54)
11
                                               (s-a-ms:56(sadv:55)
    55 mais
                       mais
                                     R.G
12
                                     AQOCSOO
    56 vulnerável
                                               (a-ms:56)))))
13
                       vulnerável
    57 a
                                     SP
                                               (grup-sp:57(prep:57)
14
                       а
    58 prazo
                                     NCMS000
                                               (sn:58(grup-nom-ms:58(n-ms:58))))
                       prazo
1.5
```

- a). Indique a que corresponde cada coluna. Descreva a estrutura da última coluna.
- b). Escreva uma script que transforme este texto em:
- o=o+DAMS capital=capital+NCFS estrangeiro=estrangeiro+AQOMS fica=ficar+VMIP3S ...
- c). Escreva uma função que recebendo um paragrafo CONLL escreva os agrupamentos (sn o capital estrangeiro) (grupo-verb fica) (grupo-sp com um maior controlo) (sp-de de um sistema mais vulnerável) (grupo-sp a prazo)

Questão 4 () Suponha que foi construída uma função trad que recebe uma string e devolve a sua tradução. Faça uma função tradC que traduza todas as strings contidas num programa C.

Questão 5 () Para pôr ordem na casa construiu-se uma listagem dos livros com o seguinte aspecto:

```
Volfrâmio
Os Maias
Os Lusíadas
Os Maias
```

Constatou-se que há algumas repetições que pretende encontrar.

Escreva um filtro usando o Python para anote títulos de livros duplicados de modo a obter:

Notas:

- (1) Nas linha contendo um título repetido, anotou-se o número da primeira ocorrência, de acordo com o exemplo.
- (2) Não analise nem altere linhas com menos de 5 caracteres.

Questão 6 () Considere a seguinte extracto de Lark – gramática de Json simplificado:

```
g = r"""
     ?start : json
                                          s1
2
     ?json : STR
                                      -> j1
3
       | "[" lista "]"
                                      ->
                                        j2
       | "{" mapping "}"
                                      -> j3
     ?lista : json
                                      -> 11
6
       | lista "," json
                                      -> 12
     ?mapping : STR ":" json
                                      ->
                                          m1
       | mapping "," STR ":" json
q
                                      -> m2
     %import common.ESCAPED_STRING -> STR
10
     %import common.WS
11
     %ignore WS
12
13
     ....
14
     class Json2Html(InlineTransformer):
15
         def s1(self,j):
16
                 print(j)
17
         def j1 ...
18
19
         def m2(self,m,s,j):
20
     jh = Lark(g, parser='lalr', transformer=Json2Html())
21
     . . .
```

- a). Escreva a expressão regular que estará associada a ESCAPED_STRING. Pelo menos tenha em conta casos como o exemplo da alínea seguinte.
- b). Junte-lhe as ações semânticas necessárias de modo a calcular a máxima profundidade de do Json.