

Processamento de Linguagens

Engenharia Informática (3º ano)

Teste (época normal)

15 de Junho de 2024 (9h00)

Dispõe de **2:00 horas** para realizar este teste.

Questão 1: Expressões Regulares (4v = 1+1+1+1)

A seguir são-te apresentados quatro matrizes com palavras cruzadas em que o enunciado são expressões regulares. Resolve cada um dos enigmas preenchendo a grelha com os caracteres devidos, 1 carater em cada quadrícula (na tua folha de resposta, coloca a grelha preenchida). Obs: No enigma 3 tens de resolver as interseções das expressões regulares nos extremos de cada linha e coluna.

	[[^] SPEAK] ⁺		EP IP EF	
HE LL O ⁺				
[PLEASE] ⁺				

Enigma 1

	(A B C)\1		(AB OE SK)	
. [*] M?O. [*]				
(AN FE BE)				

Enigma 2

	(MA LM)		[[^] MESH] ⁺	
[LINE] ⁺				
[LAM] ⁺				
	[ISLE] ⁺		[MALE] ⁺	

Enigma 3

	(ENT NTE NET) [*]			[WEAR] [*]		[R-Z E [*] M-R]	
(T E N) [*]							
(.) [*] W ⁺ \1							
[LENT] [*]							

Enigma 4

Questão 2: filtros (4v = 1.5+1+1.5)

Considere o seguinte multipoema, usado para pseudo geração de poemas:

```
BANDO DEL SE
DESIGNIO DE UN VERSO
SINO DE UN SENTIR
ENTE PARA SENTIR FORJADO
```

```
Certero el abrazo del destino,
Cruel el impulso del desatino,
Umbrío el caminar, solo, vespertino,
Sombra, laguna anónima del sino,
```

```
cuando sentir a veces es morir,
cuando nos pesa también existir,
si no sabemos nuestra alma sentir,
cuando no sabemos ni que decir,
```

```
es tan duro algunas veces vivir,
puede el aire a nuestro ser confundir,
sueños nos ayudan a revivir,
barco en el puerto sin poder partir,
```

```
ilusión a la que mi ser confino.
desazón que quiebra nuestro camino.
esperanza que se torna espino.
tras encrucijada de férreo afino.
```

- a) Escreva uma expressão regular capaz de encontrar um parágrafo formado por quatro linhas, capturando cada uma dessas 4 linhas, por exemplo, usando grupos de captura;
- b) Especifique uma função que usando o **re.sub** substitua cada ocorrência de tais parágrafos, pela sua primeira linha. Resultado esperado desta função:

```
BANDO DEL SE
```

```
Certero el abrazo del destino,
cuando sentir a veces es morir,
es tan duro algunas veces vivir,
ilusión a la que mi ser confino.
```

- c) Especifique uma função que usando o **re.sub** substitua cada ocorrência de tais parágrafos, por uma das linhas escolhidas por sorteio (`random.choice(listaDeLinhas)`). Exemplo de um resultado possível:

```
SINO DE UN SENTIR
```

```
Cruel el impulso del desatino,
si no sabemos nuestra alma sentir,
es tan duro algunas veces vivir,
tras encrucijada de férreo afino.
```

Questão 3: Gramáticas (5v = 3+2)

Ao visitar um país, um viajante tem de visitar várias localidades e em cada uma delas tem de visitar vários pontos de interesse cultural para ver um monumento, uma exposição, um espetáculo, um panorama, etc. Essas visitas devem ser feitas por uma determinada ordem. Precisa também de registrar o custo de cada atividade a realizar nesses vários pontos.

Neste exercício pede-se que:

- a) Imagine uma linguagem interessante para ajudar o viajante a fazer o seu plano de viagem por um ou mais países. Para definir essa linguagem, escreva, em BNF-puro, uma Gramática Independente de Contexto (GIC) apropriada;
- b) Identifique os símbolos terminais que usou na GIC da alínea anterior e escreva o respetivo analisador léxico associando a todos esses símbolos terminais as respetivas Expressões Regulares.

Questão 4: Parser ($5v = 1+1+1+1+1$)

Considere o Terminal variável **pal** (sequência não nula de letras e hífens) e a seguinte Gramática Independente de Contexto (G) em que o axioma é **LingPaises** e "&" representa a string nula.

```
NT = { LingPaises, Pares, Par, Lingua, Paises, Pais }
T  = { '.', ':', ',', DIC, CID, ORIG, pal }
P  = {
    p1: LingPaises : DIC Pares CID
    p2: Pares      : Pares Par '.'
    p3:            | &
    p4: Par        : Lingua ':' Paises
    p5: Lingua     : pal ORIG pal
    p6: Paises     : Pais ',' Paises
    p7:           | Pais
    p8: Pais      : pal
}
```

Neste contexto e após analisar a G dada e a Linguagem L por ela definida, responda às alíneas seguintes:

a) Considere a frase F :

```
DIC
PT ORIG latim: Portugal, Brasil.
FR ORIG latim: França.
CID
```

Apresentando a respetiva *árvore de derivação* mostre que F pertence a L .

Mostre ainda a diferença estrutural entre uma lista definida com recursividade à esquerda e uma lista definida com recursividade à direita.

b) Os **Conflitos LL(1)** em G podem surgir nas produções p2/p3 ou nas produções p6/p7. Recorrendo ao *Lookahead* dessas produções, diga justificando se G é uma GIC do tipo **LL(1)**.

c) Após estender a G dada, construa o respetivo *Autômato LR(0)*.

Diga, justificando, se em algum desses estados ocorrem **Conflitos shift-reduce**.

d) Suponha que G tinha sido modificada, mantendo L , da seguinte forma

```
P  = {
    ...
    p6: Paises      : Pais OutrosP
    p7: OutrosP    : ',' Pais OutrosP
    p7':           | &
    p8: Pais       : pal
}
```

Escreva, então, 2 funções de um parser RD (recursivo-descendente), para reconhecer os 2 símbolos não-terminais **Paises** e **OutrosP**.

e) Acrescente Ações Semânticas às produções da gramática para calcular o número total de línguas diferentes registadas e outras Ações Semânticas para escrever o total de países em que se fala cada uma das línguas.