

Trabalho teórico-prático - Análise sintática descendente - 2ª avaliação

Introdução

A biblioteca Pandas do Python é usada para manipular um conjunto de dados (dataset) organizados em colunas. Um exemplo é mostrado a seguir:

	Date	State	District	Confirmed	Recovered	Deceased	Other	Tested
0	2020-04-26	Andaman and Nicobar Islands	Unknown	33	11	0	0	NaN
1	2020-04-26	Andhra Pradesh	Anantapur	53	14	4	0	NaN
2	2020-04-26	Andhra Pradesh	Chittoor	73	13	0	0	NaN
3	2020-04-26	Andhra Pradesh	East Godavari	88	48	0	0	NaN

Para acessar as colunas usamos o seguinte formato: *nome_da_variavel [índice(s)]*. Os formatos dos índices podem ser bastante diversos, como por exemplo:

- O número da coluna, que é um número inteiro, positivo, negativo ou zero.
 - Ex: `x[0]`, `x[10]`, `x[-2]`
- O nome da coluna entre aspas, simples ou duplas. Considere que os nomes das colunas são formados por letras maiúsculas, minúsculas ou espaços em branco.
 - Ex: `x['Date']`, `x["New Column"]`
- Um intervalo (slice) formado por dois números positivos (ou zero) ou dois números negativos, separados por ":". **Observe que os números dos índices podem ser opcionais, tanto no início do slice, quanto no fim, ou em ambos.**
 - Ex: `x[0:5]`, `x[2:2]`, `x[-1: -5]`, `x[:-1]`, `x[6:]`, `x [:]`.
- Um intervalo (slice) formado por dois nomes de colunas entre aspas, simples ou duplas. **Observe que as strings dos índices podem ser opcionais, tanto no início do slice, quanto no fim, ou em ambos.**
 - Ex: `x['Data': 'State']`, `x["District": 'Tested']`, `x[: 'Tested']`, `x ["Data" :]`, `x [:]`.
- Um índice dado pelo resultado do acesso a uma base, usando o mesmo formato.
 - Ex: `x[y[0]]`, `a[b[2:0]]`, `m[n['State' : "Data"]]`, `base [colunas [:]]`.

Para esta atividade considere que um analisador léxico foi implementado e transformou os tokens da seguinte forma:

- Os números positivos e o zero aparecem na cadeia como a letra *p*;
- Os números negativos aparecem na cadeia como a letra *n*;
- As strings aparecem na cadeia como a letra *s*;
- Os nomes de variáveis aparecem na cadeia como a letra *v*.

Por exemplo, a cadeia `v [p]`, representa uma o nome da variável com um número positivo ou zero entre colchetes. `v[s : s]` representa um acesso em que o índice é o slice com duas strings. Já a cadeia `v [v [n : n]]` equivale a um acesso onde o índice é dado por outro acesso a variável com slice de números negativos. As cadeias nesse formato é que irão aparecer como entrada de dados.

Descrição da atividade (cada item vale 1,0)

- 1) Construa uma gramática que reconheça cadeias no formato dado anteriormente, isto é, com as variáveis, números e strings transformados em letras.
- 2) Construa os conjuntos NULLABLE, FIRST e FOLLOW para a gramática. Se ela possuir ambiguidades, remova as recursões à esquerda e realize a fatoração à esquerda antes de construir os conjuntos.
- 3) Construa a tabela da análise sintática descendente LL, e diga se ela é LL(1). Teste sua tabela com as cadeias $v[v[p:p]]$ e $v[p:s]$.
- 4) Usando a linguagem de programação de sua preferência, construa um analisador descendente recursivo tomando como base a tabela da questão. Utilize as mesmas cadeias para testar seu programa.