

Exercício-programa 2 – Análise sintática

(Consultar a data de entrega no site ou no Tidia.)

Regras:

- Não é permitido utilizar códigos de outras pessoas nem da Internet. Você pode usar bibliotecas padrão para estruturas de dados básicas.
- O programa deve ser implementado **individualmente**.
- Linguagens permitidas: Java, C, C++ ou Python.
- Entregue o **código-fonte** em um único arquivo com nome `EP2.java`, `EP2.c`, `EP2.cpp` ou `EP2.py`.
- Seu programa será compilado no C99, C++11, e Java 7 (não comece o arquivo java com “package” como é o padrão do Netbeans). Para programas em Python, informar a versão utilizada (e colocar na primeira linha do script usando *shebang*).
- Programas com erro de sintaxe serão imediatamente desconsiderados.
- O programa deve seguir rigorosamente a entrada e a saída especificadas, pois serão feitos testes automatizados nos programas.

1 Objetivo

Neste exercício-programa, você deverá implementar o algoritmo CYK (Cocke-Kasami-Younger) para analisar sintaticamente um texto segundo uma gramática dada na forma normal de Chomsky. Este é o algoritmo de *parsing* que utiliza programação dinâmica e que foi visto em sala.

2 Especificação da entrada

A entrada deve ser lida da **entrada-padrão** e conterá a descrição de uma gramática na forma normal de Chomsky. Os terminais serão todos os caracteres ASCII visíveis, exceto as letras maiúsculas. Mais especificamente, os terminais são os caracteres com códigos ASCII de 33 a 64 e de 91 a 126. Cada variável será representada por uma letra maiúscula (caracteres com código de 65 a 90).

A primeira linha da entrada contém um número $n \leq 1000$ que será o número de regras, e que deve ser positivo. Cada uma das n linhas seguintes descreve uma regra. O primeiro caractere da linha de uma regra é uma variável. Se a linha tiver apenas o caractere da variável (sem contar o marcador de fim de linha `\n`), a regra codifica uma regra tipo ϵ . Se a linha tiver apenas dois caracteres, o primeiro é a variável e o segundo é um terminal. Se a linha tiver três caracteres, os dois últimos também devem ser variáveis, pois a gramática será dada na forma normal de Chomsky. A variável inicial é a primeira variável da primeira regra.

Cada linha que segue a descrição da gramática é uma sequência de terminais que deve ser analisada sintaticamente, de acordo com a gramática dada.

3 Especificação de saída

Imprima, para cada linha depois da descrição da gramática, uma linha contendo **SIM** se a palavra lida é gerada pela gramática e **NAO** (sem til) caso contrário.

4 Exemplo

```
15
SDB
SFC
SAA
SDF
SFD
S
ADB
AFC
AAA
ADF
AFD
BAF
CAD
Da
Fb
abbb
aababb
```

No exemplo acima, seu programa deve decidir se as palavras **abbb** e **aababb** são geradas pela gramática:

$$\begin{aligned} S &\rightarrow DB \mid FC \mid AA \mid DF \mid FD \mid \varepsilon \\ A &\rightarrow DB \mid FC \mid AA \mid DF \mid FD \\ B &\rightarrow AF \\ C &\rightarrow AD \\ D &\rightarrow a \\ F &\rightarrow b \end{aligned}$$

A saída correta do exemplo é:

```
NAO
SIM
```

Agradecimentos

Agradecemos à Profa. Cristiane Maria Sato que idealizou a primeira versão deste exercício-programa.