

Uma linguagem regular

Considere a linguagem

- palavras que possuem um bloco de a's de tamanho ímpar

Nós sabemos que isso é uma linguagem regular.

Logo, nós podemos escrever uma expressão regular que descreve essas palavras

$$\left((a \cup b)^* b \cup e \right) a(aa)^* \left(b(a \cup b)^* \cup e \right)$$

E agora não é difícil construir a gramática.

O primeiro passo é decompor a expressão regular em 3 partes

$$\left(\underbrace{(a \cup b)^* b \cup e} \right) \underbrace{a(aa)^*} \left(\underbrace{b(a \cup b)^* \cup e} \right)$$

E daí nós podemos ter variáveis T_1 , T_2 e T_3 que produzem palavras descritas por cada parte.

As regras de T_1 são

$$\begin{aligned} T_1 &\rightarrow Qb \mid e \\ Q &\rightarrow aQ \mid bQ \mid e \end{aligned}$$

Quer dizer, a variável Q produz uma sequência qualquer de a's e b's, e essa sequência está separada do bloco ímpar de a's (na parte 2) por um símbolo b.

Abaixo nós temos as regras de T_2

$$T_2 \rightarrow aaT_2 \mid a$$

E abaixo nós temos as regras de T_3

$$T_3 \rightarrow bQ \mid e$$

onde as regras de Q já foram dadas.

A seguir, nós vamos considerar a linguagem

- palavras que possuem dois blocos de a's do mesmo tamanho

Qual é a ideia aqui?

Bom, nós precisamos gerar os dois blocos de a's que vão ter o mesmo tamanho simultaneamente — para garantir que eles vão ter o mesmo tamanho.

Isso pode ser feito com as regras

$$\begin{aligned} S &\rightarrow aS_1a \\ S_1 &\rightarrow aS_1a \end{aligned}$$

Daí, quando a construção dos blocos termina, nós podemos colocar qualquer coisa entre eles (menos um a ou a palavra vazia).

No caso mais simples, existe apenas um b entre os dois blocos, o que corresponde à regra

$$S_1 \rightarrow b$$

E, no caso geral, pode haver qualquer coisa (separada dos blocos de a's por um símbolo b), o que corresponde à regra

$$S_1 \rightarrow bQb$$

Finalmente, nós lembramos que pode aparecer qualquer coisa antes do primeiro bloco e depois do segundo bloco.

Logo, nós modificamos a primeira regra para

$$S \rightarrow T_1aS_1aT_3$$

onde T_1 e T_3 são variáveis do exemplo anterior.

Abaixo nós temos a gramática completa da linguagem

$$\begin{aligned} S &\rightarrow T_1aS_1aT_3 \\ S_1 &\rightarrow aS_1a \mid bQb \mid b \\ T_1 &\rightarrow Qb \mid e \\ T_3 &\rightarrow bQ \mid e \\ Q &\rightarrow aQ \mid bQ \mid e \end{aligned}$$