

Linguagens Formais e Autômatos

Conceitos Básicos

Prof. Anderson Belgamo

Introdução

- Teoria das Linguagens Formais
 - Originariamente desenvolvida na década de 1950.
 - Objetivo inicial: desenvolver teorias relacionadas com as linguagens naturais.
 - HOJE: importante para o estudo de linguagens artificiais, em especial para as linguagens originárias na Ciência da Computação.

Sintaxe e Semântica

- Historicamente:
 - O problema sintático foi reconhecido antes do problema semântico e são de tratamento mais simples que os semânticos.
- Conseqüência:
 - Grande ênfase à sintaxe ao ponto de levar a idéia de que questões de linguagem de programação resumiam-se às questões da sintaxe

Sintaxe e Semântica

- Consequentemente
 - Sintaxe basicamente manipula símbolos e não considera os correspondentes significados
 - Mas, para resolver qualquer problema real é necessário dar uma interpretação semântica aos símbolos
 - Exemplo: "estes símbolos representam os inteiros"

Sintaxe e Semântica

- **Sintaticamente "errado"**
 - não existe uma noção de programa "errado"
 - neste caso, simplesmente não é um programa
- **Sintaticamente "Correto"**
 - pode não ser o programa que o programador esperava escrever
- **Programa "Correto" ou "Errado"**
 - deve considerar se modela adequadamente o comportamento desejado

Abordagem para Representação

- Axiomático: associam-se regras às componentes da linguagem.
 - Exemplo: gramáticas, a qual também pode ser chamada de **formalismo gerador** pois permite verificar se um determinado elemento da linguagem é gerado.
- Operacional: autômato com estados, instruções primitivas e como cada instrução modifica da estado.
 - Conhecido como um **formalismo reconhecedor** pois analisa uma entrada para verificar se é reconhecida pela máquina.

Abordagem para Representação

- Denotacional: ou **formalismo funcional**.
 - Define-se uma função que caracteriza o conjunto de palavras admissíveis na linguagem.
 - Exemplo: expressões regulares.

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: símbolo, caractere
 - entidades abstratas básica não definida formalmente
 - Exemplo de símbolos:
 - Letras, dígitos
- Definição: alfabeto
 - conjunto finito de símbolos
 - Exemplo:
 - $\Sigma_1 = \{a, b, c\}$
 - $\Sigma_2 = \{0, 1, \dots, 9\}$
 - $\Sigma_3 = \{ \}$

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Conjunto de palavras sobre Σ_1
 - Σ^* (conjunto de todas as palavras sobre Σ)
 - $\Sigma^+ = \Sigma^* - \{\epsilon\}$
 - » Exemplo: para $\Sigma = \{a,b\}$
 - » $\Sigma^+ = \{a,b,aa,ab,ba,bb,aaa,\dots\}$
 - » $\Sigma^* = \{\epsilon, a,b,aa,ab,ba,bb,aaa,\dots\}$
- Definição: tamanho ou comprimento
 - O tamanho ou comprimento de uma palavra w , representado por $|w|$, é o número de símbolos que compõem a palavra.

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Prefixo, Sufixo, Subpalavra
 - prefixo (sufixo): qualquer seqüência de símbolos inicial (final) de uma palavra
 - subpalavra: qualquer seqüência de símbolos contígua de uma palavra
 - » Exemplo: para a palavra abcb
 - » Prefixo? Sufixo? Subpalavras?
- Definição: Linguagem Formal
 - Uma linguagem formal é um conjunto de palavras sobre um alfabeto.
 - Exemplo: suponha o alfabeto $\Sigma = \{a,b\}$. Então:
 - » Conjunto vazio: $\{ \}$
 - » Conjunto formado pela palavra vazia: $\{ \epsilon \}$
 - » Conjunto de palíndromos

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Concatenação

- É uma operação binária, definida sobre uma linguagem, a qual associa a cada par de palavras uma palavra formada pela justaposição da primeira com a segunda.
- A operação de concatenação satisfaz às seguintes propriedades (suponha v, w, t palavras):
 - a) Associatividade: $v(wt) = (vw)t$
 - b) Elemento Neutro à Esquerda e à Direita: $\varepsilon w = w = w\varepsilon$

IMPORTANTE: uma operação de concatenação definida sobre uma linguagem L não é, necessariamente, fechada sobre L , ou seja, a concatenação de duas palavras de L não é, necessariamente, uma palavra de L .

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Concatenação Sucessiva

- É uma concatenação de uma palavra (com ela mesma), representada na forma de um expoente w^n , onde w é a uma palavra e n indica o número e concatenações sucessivas.

- Exemplos:

- » $w^3 = ?$

- » $w^1 = ?$

- » $a^5 = ?$

- » $a^n = ?$

- » $w^0 = \varepsilon$

- » Se $w = \varepsilon$, então $w^n = \varepsilon$, para $n > 0$. Se $n = 0$, w^n é indefinida.

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Gramática

- É uma quádrupla ordenada $G = (V, T, P, S)$, na qual:
 - » V : conjunto finito de símbolos variáveis ou não-terminais
 - » T : conjunto finito de símbolos terminais
 - » P : conjunto finito de pares, denominados regras de produção tal que a primeira componente é a palavra de $(V \cup T)^+$ e a segunda componente é palavra de $(V \cup T)^*$
 - » S : elemento de V denominado variável inicial
- Regras de Produção: uma regra de produção (α, β) é representada por $\alpha \rightarrow \beta$.

As regras de produção definem as condições de geração das palavras da linguagem.

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Derivação

- A aplicação de uma regra de produção é denominado derivação de uma palavra.
- A aplicação sucessiva de regras de produção permite derivar as palavras da linguagem representada pela gramática.
- Em resumo:
 - Derivação é substituição de uma palavra de acordo com a regra de produção.

- Definição: Linguagem Gerada

- $G = (V, T, P, S)$
 - $V = \{S, D\}$
 - $T = \{0, 1, 2, \dots, 9\}$
 - $P = \{ S \rightarrow D, S \rightarrow DS, D \rightarrow 0 \mid 1 \mid \dots \mid 9 \}$

Alfabetos, Palavras, Linguagens e Gramáticas

- Definição: Linguagem Gerada
 - Uma derivação do número 243:
 - $S \rightarrow DS \rightarrow 2S \rightarrow 2DS \rightarrow 24S \rightarrow 24D \rightarrow 243$
 - Portanto:
 - $S \rightarrow^* 243$
 - $S \rightarrow^+ 243$
 - $S \rightarrow^6 243$
- Definição: Equivalência de Gramáticas
 - G_1 e G_2 são equivalentes se e somente se
 - $GERA(G_1) = GERA(G_2)$

Alfabetos, Palavras, Linguagens e Gramáticas

– Convenções:

- A, B, C,..., S, T símbolos variáveis
- a, b, c,..., s, t símbolos terminais
- u, v, w, x, y, z palavras de símbolos terminais
- α, β palavras de símbolos variáveis e/ou terminais

Alfabetos, Palavras, Linguagens e Gramáticas

– Exemplo:

– $\{ ww \mid w \text{ é palavra de } \{a, b\}^* \}$

– $G = (\{S, X, Y, A, B, F\}, \{a, b\}, P, S)$

– $P = \{$
 $S \rightarrow XY,$
 $X \rightarrow XaA \mid XbB \mid F$
 $Aa \rightarrow aA, Ab \rightarrow bA, AY \rightarrow Ya,$
 $Ba \rightarrow aB, Bb \rightarrow bB, BY \rightarrow Yb,$
 $Fa \rightarrow aF, Fb \rightarrow bF, FY \rightarrow \epsilon \}$

Exemplo da derivação da palavra baba

$S \rightarrow XY \rightarrow XaAY \rightarrow XaYa \rightarrow XbBaYa \rightarrow XbaBYa \rightarrow XbaYba \rightarrow FbaYba \rightarrow$
 $bFaYba \rightarrow baFYba \rightarrow ba\epsilon ba \rightarrow baba$