Universidade Católica de Pelotas
Escola de informática
053212 Linguagens Formais e Autômatos

TEXTO 2

Expressões Regulares e Gramáticas Regulares

Prof. Luiz A M Palazzo
Março de 2007

Definição de Expressão Regular

Uma expressão regular (ER) sobre um alfabeto Σ é indutivamente definida como se segue:

a) $\varnothing$ é uma ER que denota a linguagem vazia.

b) ε é uma ER que denota a linguagem contendo exclusivamente a palavra vazia, ou seja $\{\varepsilon\}$.

c) Qualquer símbolo x pertencente ao alfabeto Σ é uma ER e denota a linguagem contendo a palavra unitária x, ou seja {x}.

d) Se r e s são ERs e denotam respectivamente as linguagens R e S, respectivamente, então:

1. (r+s) é ER e denota a linguagem $R \cup S$
2. (rs) é ER e denota a linguagem $\{uv \mid u \in R \text{ e } v \in S\}$
3. (r*) é ER e denota a linguagem R*

Observações:

1. Os parênteses podem ser omitidos, respeitando-se as seguintes proridades de operações:
   - A concatenação sucessiva tem precedência sobre a concatenação e a união, e
   - A concatenação tem precedência sobre a união.
2. Uma linguagem gerada por um expressão regular r é representada por L(r) ou GERA(r).

Exemplos:

Expressão Regular	*Linguagem Representada*
aa	Somente a palavra aa.
ba*	Todas as palavras que iniciam por b, seguido de zero ou mais a.
(a+b)*	Todas as palavras sobre o alfabeto {a, b}
(a+b)*aa(a+b)*	Todas as palavras contendo aa como subpalavra.
a*ba*ba*	Todas as palavras contendo exatamente dois b
(a+b)*(aa+bb)	Todas as palavras que terminam com aa ou bb.
$(a+\varepsilon)(b+ba)^*$	Todas as palavras que não possuem dois a consecutivos

Exercícios:

Desenvolva expressões regulares que gerem as seguintes linguagens sobre $\Sigma = \{a, b\}$:

a. {w | w tem no máximo um par de a como subpalavra e no máximo um par de b como subpalavra.

b. {w | qualquer par de a antecede qualquer par de b}

c. {w | w não possui aba como subpalavra}

Expressões Regulares, Autômatos e Linguagens Regulares

Teoremas:

A classe das expressões regulares denota exatamente a classe das linguagens regulares, o que se demonstra a partir dos seguintes teoremas:

Expressão Regular → Linguagem Regular:

Se r é uma expressão regular, então GERA(r) é uma linguagem regular

Linguagem Regular → Expressão Regular

Se L é uma linguagem regular, então existe uma expressão regular r tal que GERA(r) = L.

Prova:

- Por indução sobre o número de operadores.
- Uma linguagem é regular, se e somente se é possível construir um AF (D, N ou ε) que reconheça a linguagem.
- É necessário mostrar que dada uma expressão regular r qualquer, é possível construir um autômato finito M tal que ACEITA(M) = GERA(r).

a) Base da Indução:

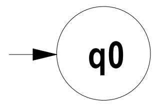
Seja r uma ER com ZERO operadores, então r só pode ser da forma:

$r = \emptyset$;
$r = \varepsilon$;

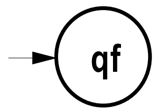
$r = x$, $(x \in \Sigma)$

Os autômatos

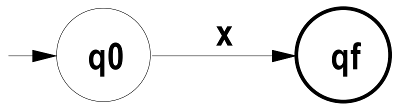
$M1 = (\emptyset, \{q0\}, \delta 1, q0, \emptyset)$



$M2 = (\emptyset, \{qf\}, \delta 2, qf, \{qf\})$



$M3 = (\{x\}, \{q0, qf\}, \delta 3, q0, \{qf\})$



aceitam as linguagens dadas na base de indução.

b) Hipótese de Indução

Seja r uma ER com até $n>0$ operadores. Suponha que é possível definir um autômato finito que aceita a linguagem gerada por r, GERA(r).

c) Passo de Indução

Seja r uma ER com n+1 operadores, então r pode ser representada por um dos seguintes casos, onde r1 e r2 possuem conjuntamente no máximo n operadores.

$r = r1 + r2$

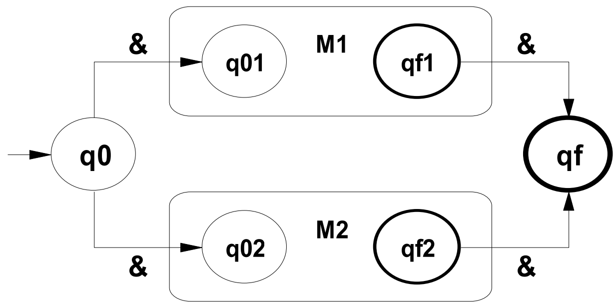
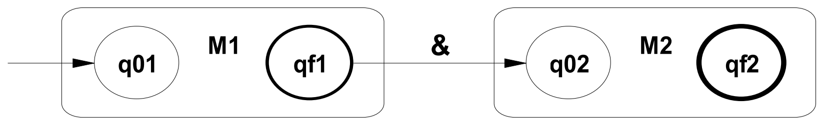
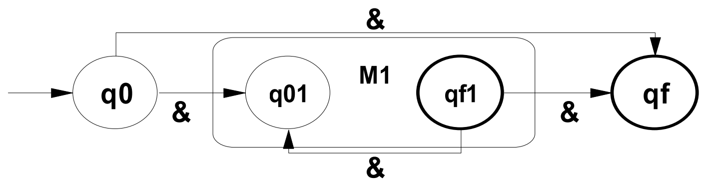
$r = r1r2$

$r = r1^*$

Portanto, por hipótese de indução, é possível construir os autômatos:

$M1 = (\Sigma 1, Q1, \delta 1, q01, \{qf1\})$, tal que ACEITA(M1) = GERA(r1)

$M2 = (\Sigma 2, Q2, \delta 2, q02, \{qf2\})$, tal que ACEITA(M2) = GERA(r2)

<table>
<tr><th>ER</th><th>Autômato Finito Correspondente</th></tr>
<tr><td>r = r1 + r2</td><td>
</td></tr>
<tr><td>r = r2r1</td><td>
</td></tr>
<tr><td>r = r1*</td><td>
</td></tr>
</table>

Para cada caso há um autômato finito com movimento vazio, cujo diagrama é dado no quadro acima, que aceita a correspondente linguagem GERA(r). As definições de tais autômatos são:

ER	Autômato Finito Correspondente
r = r1 + r2	M = (Σ1 $\cup$ Σ2, Q1 $\cup$ Q2 $\cup$ {q0, qf}, δ, q0, {qf})
r = r1r2	M = (Σ1 $\cup$ Σ2, Q1 $\cup$ Q2, δ, q01, {qf2})
r = r1*	M = (Σ1, Q1 $\cup$ {q0, qf}, δ, q0, {qf})

A construção dos módulos do AFε que reconhece a linguagem gerada pela ER a*(aa+bb) é dada a seguir:

Gramáticas Regulares

Definição:

Seja G = {V, T, P, S} uma gramática e sejam A e B elementos de V e w uma palavra de T*.
Então G é uma:

a) GLD: Se todas as regras de produção são da forma $A \rightarrow wB$ ou $A \rightarrow w$.
b) GLE: Se todas as regras de produção são da forma $A \rightarrow Bw$ ou $A \rightarrow w$.
c) GLUD: Se todas as regras de produção são como na GLD e além disso |w| =< 1.
d) GLUE: Se todas as regras de produção são como na GLE e além disso |w| =< 1.

Note-se que nas gramáticas lineares o lado direito das produções apresentam no máximo uma variável, que se existir irá sempre anteceder (LE) ou suceder (LD) qualquer subpalavra de terminais.

Teorema:
Seja L uma linguagem, então:
L é gerada por uma GLD sss
L é gerada por uma GLE sss
L é gerada por uma GLUD sss
L é gerada por uma GLUE.
Ou seja, as diversas formas de gramáticas lineares são formalismos equivalentes.

Gramática Regular:
Uma gramática regular é qualquer gramática linear.
Uma linguagem gerada por uma gramática regular G é representada por L(G) ou GERA(G).

Exemplo 1:
A linguagem a(ba)* é gerada pelas seguintes gramáticas:

a) GLD: G = ({S, A}, {a, b}, {S→aA, A→baA | &}, S).
b) GLE: G = ({S}, {a, b}, {S→Sba | a}, S).
c) GLUD: G = ({S, A, B}, {a, b}, {S→aA, A→bB | &, B→aA}, S).
d) GLUE: G = ({S, A}, {a, b}, {S→Aa | a, A→Sb}, S).

Exemplo 2:
A linguagem (a+b)*(aa + bb) é gerada pelas seguintes gramáticas:

a) GLD: G = ({S, A}, {a, b}, {S→aS | bS | A, A→aa | bb}, S).
b) GLE: G = ({S, A}, {a, b}, {S→Aaa | Abb, A→Aa | Ab | &}, S).

Teorema:
Se L é uma linguagem gerada por uma gramática regular, então L é uma linguagem regular.

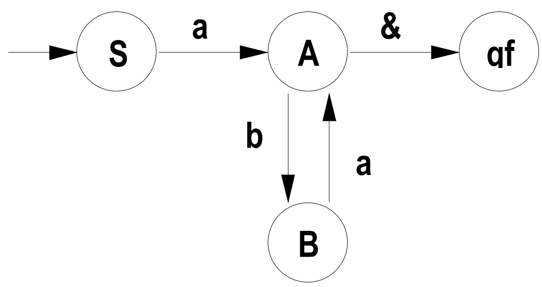
Teorema: Construção de um AFε a partir de uma Gramática Regular:
Seja G = (V, T, P, S) uma GLUD, então o AFε M = (Σ, Q, δ, q0, F) construído abaixo simula as derivações de G, ou seja, ACEITA(M) = GERA(G).

Σ = T
Q = V ∪ {qf}
F = {qf}
q0 = S
δ é construída como se segue, supondo A e B variaveis e a terminal.

Tipo da Produção	Transição Gerada
A → ε	δ(A, ε) = qf
A → a	δ(A, a) = qf
A → B	δ(A, ε) = B
A→ aB	δ(A, a) = B

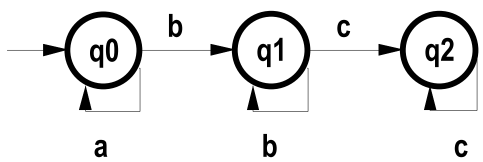
Exemplo: Construção de um AFε a partir de uma Gramática Regular
Considere a GLUD G = ({S, A, B}, {a, b}, {S→aA, A→bB | ε, B→aA}, S).
O AFε M que reconhece GERA(G) é: M = ({a, b}, {S, A, B, qf}, δ, S, {qf}), onde δ é construída como se segue:

Produção	Transição
S→aA	δ(S, a) = A
A→bB	δ(A, b) = B
A→ε	δ(A, ε) = qf
B→aA	δ(B, a) = A



Teorema:
Seja L uma linguagem regular, então existe G, uma gramática regular, que gera L

Exemplo: Construção de uma Gramática Regular a Partir de um AFD
Considere o AFD M = ({a,b,c}, {q0,q1,q2], δ, q0, {q0,q1,q2}) dado pelo diagrama:



A Gramática Regular construída a partir deste AFD é

$$G = (\{q0,q1,q2,S\}, \{a,b,c\}, S\ P),$$

Onde P é dado por:

Transição	Produção
$\delta(q0, a) = q0$ $\delta(q0, b) = q1$ $\delta(q1, b) = q1$ $\delta(q1, c) = q2$ $\delta(q2, c) = q2$	$S \rightarrow q0$ $q0 \rightarrow \varepsilon$ $q1 \rightarrow \varepsilon$ $q2 \rightarrow \varepsilon$ $q0 \rightarrow aq0$ $q0 \rightarrow bq1$ $q1 \rightarrow bq1$ $q1 \rightarrow cq2$ $q2 \rightarrow cq2$

Tradução dos Formalismos das Linguagens Regulares: