Capítulo IV - Análise Léxica

IV.1 – Fundamentos Teóricos

- **Autômatos Finitos e Conjuntos Regulares**
  (cap. III da apostila de Linguagens Formais e Compiladores)

Geradores X Reconhecedores

Gramáticas Tipo 0	→	Máquinas de Turing
G. Sensíveis ao Contexto	→	Aut. Lim. Lineares
G. Livres de Contexto	→	Autômatos de Pilha
Gramáticas Regulares	→	Autômatos Finitos

Autômatos Finitos

- **São reconhecedores de linguagens regulares**
- **Tipos de Autômatos Finitos:**
  - Autômato Finito Determinístico (AFD)
  - Autômato Finito Não Determinístico(AFND)

IV.1.1 – Autômatos Finitos Determinísticos (AFD)

Definição formal : **$M = (K, \Sigma, \delta, qo, F)$**, onde:

K → É um conjunto finito não vazio de **Estados**;

Σ → É um **Alfabeto**, finito, de entrada;

δ → **Função de Mapeamento** (ou de transição) definida em: **$K \times \Sigma \rightarrow K$**

qo → $\in K$, é o **Estado Inicial**

F → $\subseteq K$, é o conjunto de **Estados Finais**

Exemplo: Seja **$M = (K, \Sigma, \delta, qo, F)$**, onde:

K = **{q0, q1}**

Σ = **{ a, b}**

δ = **{δ(q0, a)=q0, δ(q0, b)=q1, δ(q1, b)= q1, δ(q1,a)= - }**

qo = **q0**

F = **{q1}**

Representações de AF

- **Alem da representação formal, um AF pode também ser representado por:**
  - **Diagrama de Transição**
  - **Tabela de Transições**

Sentenças Aceitas (reconhecidas) por M:

$$\delta(qo, x) = p \mid p \in F$$

Linguagem Aceita por M:

$$T(M) = \{ x \mid \delta(qo, x) = p \wedge p \in F \}$$

IV.1.2 - A.F.N.D.

Definição: M = (K, Σ, δ, qo, F), onde:

K, Σ, qo, F → mesma definição dos AFD

δ → K x Σ = ρ(K) , onde ρ(K) ⊆ K

	Vantagem	Desvantagem
AFD	Implementação Trivial Reconhec. Eficiente	Representação menos natural de algumas L.R.
AFND	Representação mais natural de algumas LR	Implementação complexa Reconhec. Ineficiente

Exemplos: Construa um AFND M |

a) T(M) = { (a, b)*abb }

b) T(M) = { (0, 1)* (00 | 11) (0, 1)*}

c) Construa AFD ≡ AFND dos itens a) e b)

IV.1.3 – Transformação de AFND para AFD

Teorema 3.1: "Se L é um conjunto aceito por um A.F.N.D., então ∃ um A.F.D. que aceita L"

Prova: Seja M = (K, Σ, δ, qo, F) um A.F.N.D..

Construa um A.F.D. M' = (K', Σ, δ', qo', F') como segue:

1 – K' = {ρ(k)}

2 – qo' = [qo]

3 – F' = {ρ(K) | ρ(K) ∩ F ≠ φ}

4 – Para cada ρ(K) ⊂ K'

faça δ'(ρ(K),a) = ρ'(K), onde

ρ'(K) = {p | para algum q ∈ ρ(K), δ(q, a) = p};

Exemplo: Seja M um A.F.N.D. definido por:

δ	a	b
→qo	qo,q1	qo
q1	---	q2
q2	---	q3
*q3	---	---

IV.1.4 - Relação entre GR e AF

Teorema 3.2: "Se $G = (Vn, V_T, P, S)$ é uma G.R., então $\exists$ um A.F. $M = (K, \Sigma, \delta, qo, F) \mid T(M) = L(G)$".

Prova: a – Mostrar que M existe
b – Mostrar que $T(M) = L(G)$

a) **Defina M, como segue:**

1 – $K = Vn \cup \{A\}$, onde A é um símbolo novo

2 – $\Sigma = V_T$

3 – $qo = S$

4 – $F = \{A, S\}$ se $S \rightarrow \varepsilon \in P$
$\{A\}$ se $S \rightarrow \varepsilon \notin P$

5 – Construa δ de acordo com as regras a, b e c.

a) Se $B \rightarrow a \in P \Rightarrow \delta(B, a) = A$

b) Se $B \rightarrow a\,C \in P \Rightarrow \delta(B, a) = C$

c) Para todo $a \in VT$, $\delta(A, a) = -$ (indefinido)

b) **Para mostrar que T(M)=L(G), deve-se mostrar:**

1 – $L(G) \subseteq T(M)$
2 – $T(M) \subseteq L(G)$

Exemplos:

1) $S \rightarrow a\,S \mid b\,B$ $B \rightarrow b\,B \mid c$	2) $S \rightarrow b\,A \mid a\,B \mid b \mid \varepsilon$ $A \rightarrow b\,A \mid a\,B \mid b$ $B \rightarrow b\,B \mid a\,C$ $C \rightarrow b\,C \mid a\,A \mid a$

Teorema 3.3: "Se M = (K, Σ, δ, qo, F) é um A. F., então ∃ uma G.R. G = (Vn, Vt, P, S) | L(G) = T(M)"

Prova: a – Mostrar que G existe
b – Mostrar que L(G) = T(M)

a) Seja M = (K, Σ, δ, qo, F) um A.F.D..
Construa uma G.R. G=(Vn, V_T, P, S), como segue:

1 – Vn = K
2 – Vt = Σ
3 – S = qo
4 – Defina P, como segue:

a) Se δ(B, a) = C então adicione B → aC em P
b) Se δ(B, a) = C ∧ C ∈ F, adicione B → a em P
c) Se qo ∈ F,
então ε ∈ T(M).
Neste caso, L(G) = T(M) – {ε}, portanto, construa uma GR G_1 | L(G_1) = L(G) U {ε}
Senão ε ∉ T(M) e L(G) = T(M)

b) Para mostrar que L(G) = T(M), devemos mostrar que:

1 – T(M) ⊆ L(G)
2 – L(G) ⊆ T(M)

Exemplos :

δ	a	b
*→S	A	B
A	S	C
B	C	S
C	B	A

δ	a	b	b
→S	S	B	-
B	-	B	A
A	-	-	-

IV.1.5 - Minimização de Autômatos Finitos

Definição: Um AFD $M = (K, \Sigma, \delta, qo, F)$ é **mínimo** se:

1 – Não possui estados inacessíveis (inalcançáveis);

2 – Não possui estados mortos;

3 – Não possui estados equivalentes.

Alg. para Construção das Classes de Equivalência

1 – Crie, um estado ϕ para representar as indefinições;

2 – Divida K em duas CE : F e K-F;

3 – Aplique a lei de formação de CE, até que nenhuma nova CE seja formada

Algoritmo para construção do A.F. Mínimo

Entrada: Um A.F.D. $M = (K, \Sigma, \delta, qo, F)$;
Saída: Um AFD Mínimo $M' = (K', \Sigma, \delta', qo', F') \mid M' \equiv M$;
Método:

1 – Elimine os estados Inacessíveis;
2 – Elimine os estados Mortos;
3 – Construa todas as CE de M.
4 – Construa M', como segue:

a) $K' = \{ CE \}$
b) $qo' =$ CE que contiver qo;
c) $F' = \{ [q] \mid \exists\, p \in F \text{ em } [q] \}$
d) $\delta' = \delta'([p], a) = [q] \Leftrightarrow \delta(p_1, a) = q_1$ está em M $\wedge$ $p_1 \in [p]$ $\wedge$ $q_1 \in [q]$

Exemplo: **Minimize o seguinte A.F.D.**

δ	a	b
* → A	G	B
B	F	E
C	C	G
* D	A	H
E	E	A
F	B	C
* G	G	F
H	H	D

Exercícios:

1)

δ	a	b	c
* → S	A	B,F	S,F
A	S,F	C	A
B	A	-	B,S,F
C	S,F	-	A,C
*F	-	-	-

2)

δ	0	1
→ S	A , D	E
A	A , B	C , E
B	B	-
C	A , B	-
D	B , D	C
*E	E	E

3)

δ	a	b
→ q0	q1	q2
q1	q3	-
q2	-	q4
*q3	q3	q3
*q4	q4	q4

4)

δ	a	b	c
*→ S	A , C	A , D	B , C
*A	A	A	B
*B	A	A	-
*C	C	D	C
*D	C	-	C

IV.1.6 – Conjuntos e Expressões Regulares

Conjuntos Regulares (C.R.)

1 – (* definição matemática (primitiva) *)

Seja Σ um alfabeto qualquer.
Definimos um C.R. sobre Σ, como segue:

a – ϕ é um C.R. sobre Σ;

b – {ε} é um C.R. sobre Σ;

c – {a}, para todo a ∈ Σ, é um C.R. sobre Σ;

d – Se P e Q são C.R. sobre Σ, então:

1 – P ∪ Q (união),
2 – P.Q (ou PQ) (concatenação),
3 – P^* (fechamento).

Também são C.R. sobre ∑;

e – Nada mais é C.R.

2 – Linguagens geradas por Gramáticas Regulares.

3 – Linguagens reconhecidas por Autômatos Finitos.

4 – Linguagens denotados por Expressões Regulares.

Expressões Regulares (E.R.)

Definição:

1 – ϕ é uma E.R. e denota o C.R. ϕ

2 – ε é uma E.R. e denota o C.R. $\{\varepsilon\}$

3 – $\underline{a} \in \Sigma$, é uma E.R. e denota o C.R. $\{ \underline{a} \}$

4 – Se $\underline{p}$ e $\underline{q}$ são E.R. denotando P e Q, então:

a – $(p \mid q)$ é uma E.R. denotando $P \cup Q$

b – $(p.q)$ ou (pq) é uma E.R. denotando PQ

c – $(p)^*$ é uma E.R. denotando P^*

5 – Nada mais é E.R..

Observações:

1 – ordem de precedência: 1) * 2) . 3) |

2 – abreviaturas usuais:

$$p^+ = pp^*$$
$$p^? = p \mid \varepsilon$$
$$p^{£}q = p(qp)^*$$

Relação entre E.R. e C.R.

1 – Para todo C.R. $\exists$ uma E.R. que o denota

2 – Para toda E.R. é possível construir seu C.R.

3 – $E1 \equiv E2 \Leftrightarrow$ elas denotam o mesmo C.R..

IV.1.6.1 – Autômatos Finitos com ε-transições

AFND-ε : M = (K, Σ, δ, qo, F), onde:

K, Σ, qo, F → mesma definição dos A.F.D.

$\delta \rightarrow K \times \Sigma \cup \{\varepsilon\} = \rho(K)$, onde $\rho(K) \subseteq K$

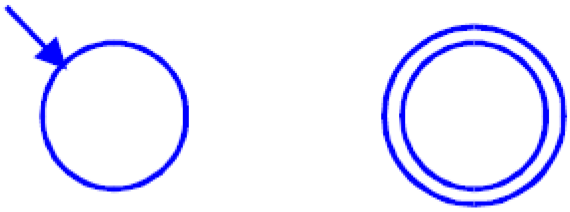
Observações:

- **ε-transições permitem movimentos independentes da entrada;**
- **O uso de ε-transições não incrementa a expressividade dos AF;**
- **Todo AFND-ε possui um AFND equivalente;**

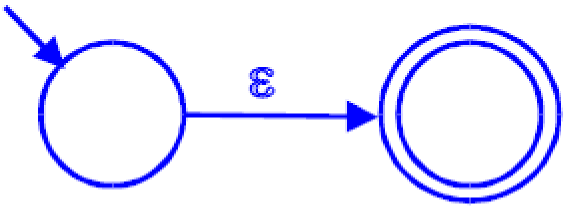
IV.1.6.2 – Correspondência entre ER e AF

Para mostrarmos que toda ER possui um AF correspondente, é suficiente mostrarmos que toda ER básica (Φ, ε, a, (α | β), (α . β) e α* - onde α, β são ERs quaisquer) possui um AF correspndente:

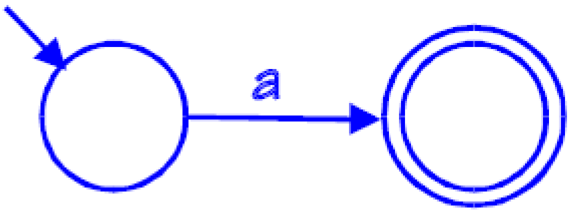
1 – AF representando a ER "ϕ" (M|T(M) = ϕ)



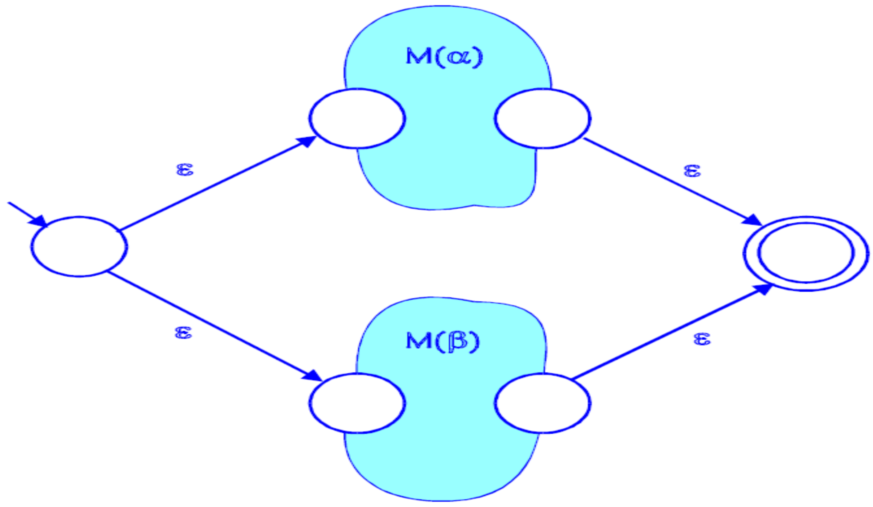
2 – AF representando a ER "ε" (M|T(M) = {ε})



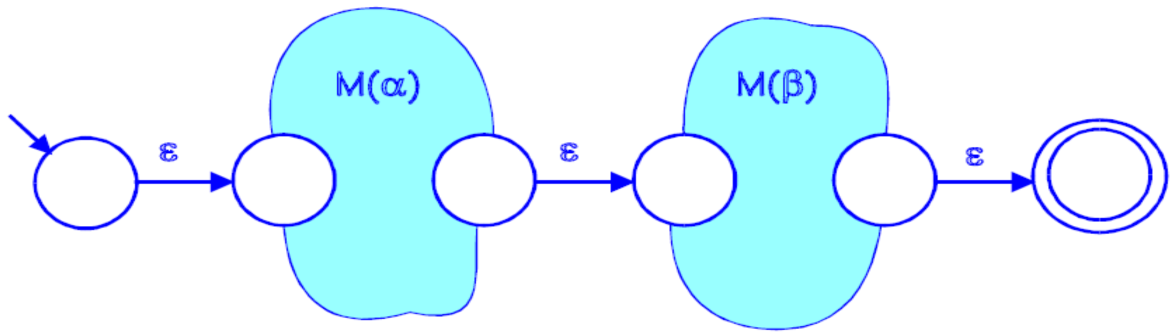
3 – AF representando a ER "a" (M|T(M) = {a})



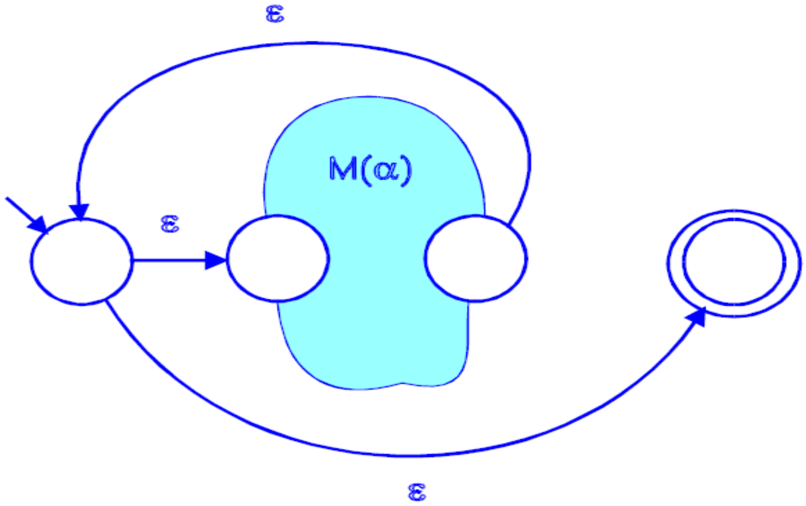
4 – AF representando a ER “α | β” (M|T(M) = {α|β})



5 – AF representando a ER “α.β” (M|T(M) = {α.β})



5 – AF representando a ER “α *” (M|T(M) = {α*})



OBS. Figuras extraídas de J.L.M.Rangel Neto (COPPE/UFRJ-PUC/RJ)

III.6.3 - Transformação de ER para AF

Diferentes métodos (estratégias):

- Método de Thompson
- Método de De Simoni (Adap. de Rennes/AHO)

III.6.3.1 - Método de Thompson

- **Consiste em:**

  1 - Decompor uma ER em suas sub-expressões constituintes;

  2 - Construir o AFNDε de cada subexpressão;

  3 - Compor o AFNDε final (usando ε-transições)

- **Exemplo:**

IV.1.7 - Implementação de Autômatos Finitos

Formas básicas para implementação de A.F.:

- **Implementação Específica**
- **Implementação Geral (ou genérica);**

IV.1.8 – Propriedades e Prob. de Decisão de CR

Propriedades Básicas de C.R.

1 – União

2 – Concatenação

3 - Fechamento

4 – Complemento: Se $L_1 \subseteq \Sigma^*$ é CR $\Rightarrow$ $\Sigma^* - L_1$ também é CR

5 – Intersecção: Se L1 e L2 são CR $\Rightarrow$ L1 $\cap$ L2 também é CR

Problemas de Decisão sobre C.R.

1 – Membership : $x \in T(M)$?

2 – Emptiness : $T(M) = \varphi$?

3 – Finiteness : T(M) é finita?

4 – Containment : $T(M1) \subseteq T(M2)$?

5 – Equivalencia : $T(M1) = T(M2)$?

6 – Intersecção Vazia : $T(M1) \cap T(M2) = \varphi$?

IV.1.9 – Aplicações de A.F. e E.R.

1 – Compiladores – Análise Léxica

2 – Editores de texto – busca/substituição

3 – Reconhecimento de padrões

4 – Outras: S.O, Redes, Hipertexto, Robótica, Criptografia, ...