

Construção de autômatos finitos determinísticos para reconhecimento de sentenças definidas por gramáticas regulares

Prof. Ivan L. M. Ricarte

Referências:	Tópico	Seção da apostila
	Gramáticas regulares	7.4
	Expressões regulares	7.5
	Definição de autômatos	8.1

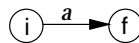
Procedimento de construção do autômato

Passo 1: Decompor a expressão regular em termos de relações elementares:

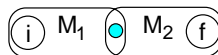
- um símbolo do alfabeto da linguagem;
- concatenação, RS ;
- alternativa, $R|S$;
- repetição, R^* .

Passo 2: Construir o autômato finito não-determinístico a partir da composição dos autômatos para reconhecer as relações elementares:

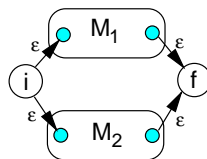
- um símbolo a do alfabeto da linguagem Σ :



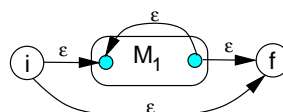
- concatenação, RS , onde M_1 reconhece R e M_2 reconhece S ; o estado final de M_1 é combinado com o estado inicial de M_2 :



- alternativa, $R|S$ com M_1, M_2 como acima; um novo estado inicial é introduzido, com transições pela *string* vazia para os estados iniciais de cada uma das máquinas M_1 e M_2 . Similarmemente, um novo estado final é atingido com transições pela *string* vazia a partir dos estados finais dessas máquinas:



- repetição, R^* , onde M_1 reconhece R ; novos estados inicial e final são introduzidos, com transições pela *string* vazia para reconhecer zero ocorrências (direto do novo estado inicial para o novo estado final) e várias ocorrências (volta do estado final de M_1 para o estado inicial de M_1):



Passo 3: Converter autômato não-determinístico (estados originais) para autômato determinístico (novos estados):

- (a) Compute a ϵ^* do estado inicial do autômato não-determinístico. A ϵ^* (ϵ -clausura) é o conjunto que inclui o estado indicado e todos os estados alcançáveis a partir dele através de transições por *strings* vazias. O resultado é um conjunto de estados que irá representar um novo estado no autômato determinístico — neste caso, um novo estado inicial, pois o conjunto resultante inclui o estado inicial original. Inclua esse novo estado em uma lista de estados não-analisados.
- (b) Para cada novo estado s na lista de estados não-analisados, obtenha o novo estado t resultante da transição através de cada símbolo α do alfabeto e inclua a transição $s \xrightarrow{\alpha} t$ no novo autômato determinístico. Para obter t , obtenha o conjunto T_α dos estados originais que podem ser alcançados através de uma transição pelo símbolo α a partir de cada estado original no conjunto s e compute $\epsilon^*(T_\alpha)$. Se t for um novo estado, inclua-o na lista de estados não-analisados. Se algum elemento de t for um estado final no autômato original, então t será um estado final no novo autômato. Este passo deve ser repetido até que a lista de estados não-analisados torne-se vazia.

Exemplo

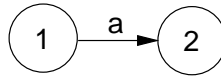
*Construir um autômato finito determinístico para reconhecer sentenças descritas pela expressão regular $R = (a|b)^*abb$.*

Passo 1

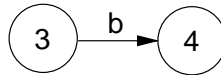
$$\begin{aligned} R_1 &= a & R_2 &= b \\ R_3 &= R_1 | R_2 & R_4 &= R_3^* \\ R_5 &= R_4 R_1 & R_6 &= R_5 R_2 \\ R &= R_6 R_2 \end{aligned}$$

Passo 2

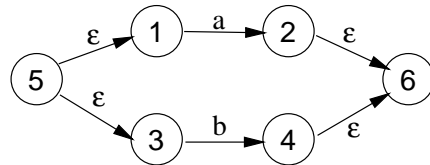
Para reconhecer R_1 ,



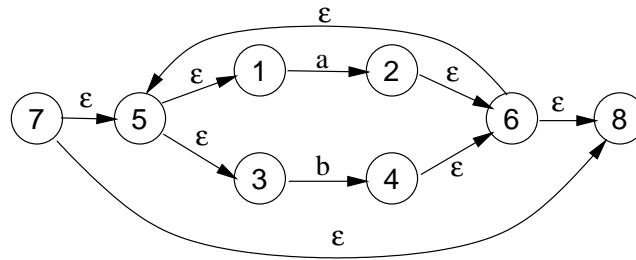
Similarmente, para reconhecer R_2 ,



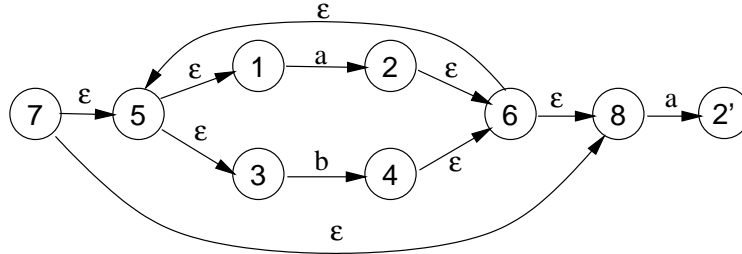
O autômato para reconhecer R_3 é construído pela combinação-alternativa de R_1 e R_2 ,



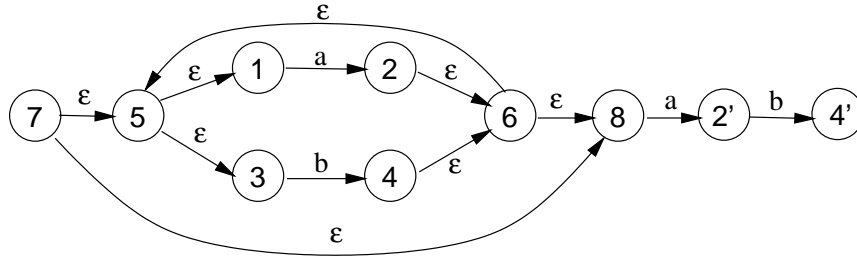
Para reconhecer R_4 , constrói-se o autômato que reconhece 0 ou mais ocorrências de R_3 :



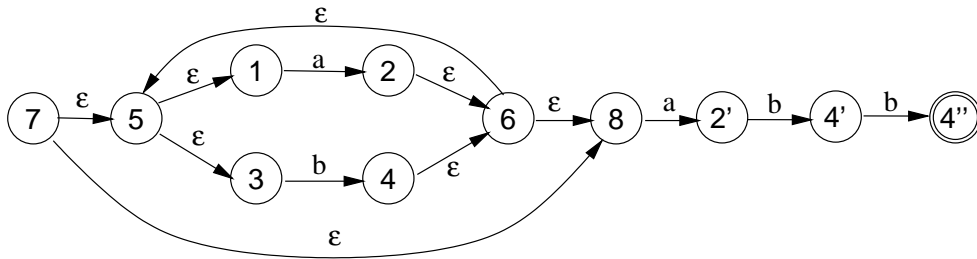
O autômato para reconhecer R_5 é formado pela combinação de concatenação de R_4 com um autômato equivalente àquele que reconhece R_1 ,



Similarmente, para reconhecer R_6 é preciso combinar o autômato que reconhece R_5 com um equivalente àquele que reconhece R_2 ,



A expressão completa é uma concatenação de R_6 com um autômato que reconhece b , equivalente à R_2 :



Passo 3

O conjunto de estados originais é $T_0 = \{7\}$; portanto, o novo estado inicial será dado por $\epsilon^*(T_0)$, ou

$$s_0 = \{1, 3, 5, 7, 8\}$$

A lista de estados não-analisados contém s_0 ; para esse estado, é preciso analisar as transições resultantes para cada um dos dois símbolos do alfabeto, a e b .

O estado s_0 contém dois estados originais a partir dos quais existem transições com o símbolo a , 1 e 8. Os estados originais atingidos por essas transições são 2 e 2'. Portanto, $T_1 = \{2, 2'\}$ e o estado atingido a partir de s_0 pela transição através do símbolo a será $s_1 = \epsilon^*(T_1)$, ou

$$s_1 = \{1, 2, 2', 3, 5, 6, 8\}$$

Similarmente, a partir de s_0 pelo símbolo b atinge-se o conjunto de estados $T_2 = \{4\}$ e, portanto, $s_2 = \epsilon^*(T_2)$ resulta em

$$s_2 = \{1, 3, 4, 5, 6, 8\}$$

O estado s_0 foi retirado da lista de estados não-analisados, mas dois novos estados — s_1 e s_2 — foram incluídos. Da análise de s_1 , obtém-se que a transição pelo símbolo a também levará ao estado s_1 , enquanto que a transição pelo símbolo b leva aos estados originais $T_3 = \{4, 4'\}$, resultando em um novo estado gerado por $\epsilon^*(T_3)$,

$$s_3 = \{1, 3, 4, 4', 5, 6, 8\}$$

A análise de s_2 indica que a transição pelo símbolo a leva ao estado s_1 , enquanto que a transição pelo símbolo b leva ao próprio estado s_2 . Nenhum novo estado é criado.

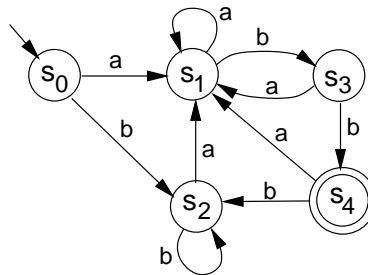
O estado s_3 permanece na lista de estados não-analisados. A transição a partir dele pelo símbolo a leva também ao estado s_1 . Pelo símbolo b , o conjunto de estados originais resultantes é $T_4 = \{4, 4''\}$; assim, um novo estado é gerado a partir do cômputo de $\epsilon^*(T_4)$,

$$s_4 = \{1, 3, 4, 4'', 5, 6, 8\}$$

Esse estado, que é incluído na lista de estados não-analisados, é um estado final, pois contém o estado original final $4''$.

Finalmente, a análise de s_4 indica que a transição pelo símbolo a leva ao estado s_1 , enquanto que a transição pelo símbolo b leva ao estado s_2 .

Após a análise de s_4 , a lista de estados não-analisados ficou vazia, indicando a conclusão do processo de conversão do autômato. O autômato determinístico resultante é:



Para fins computacionais, a representação mais adequada para autômatos é aquela na forma de tabelas de transição. Para esse autômato, a tabela correspondente é:

	s_0	s_1	s_2	s_3	s_4
a	s_1	s_1	s_1	s_1	s_1
b	s_2	s_3	s_2	s_4	s_2