MAC 4722 - Linguagens, Autômatos e Computabilidade

Rodrigo Augusto Dias Faria - NUSP 9374992
Departamento de Ciência da Computação - IME/USP

24 de maio de 2016

Lista 5

L5.1 Construa a gramática obtida a partir do autômato a pilha da figura 2.15 usando o lema 2.27, a segunda parte do teorema 2.20. Particione o conjunto de variáveis da sua gramática de forma a exibir quais delas geram a linguagem vazia e quais geram linguagens não vazias.

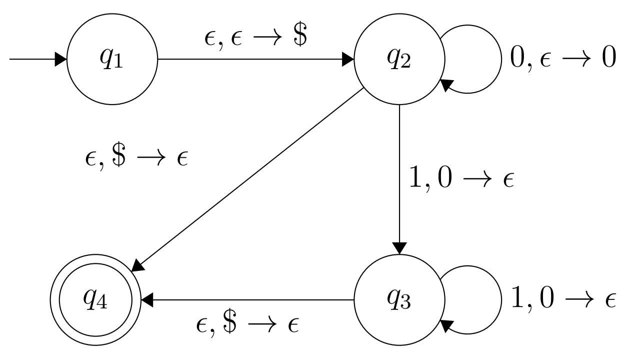
Resposta: O autômato dado na figura 2.15 não atende às três condições solicitadas no lema 2.27, a saber:

1. Deve ter um único estado de aceitação

2. A pilha deve estar vazia ao aceitar

3. Toda transição ou empilha, ou desempilha um símbolo

Claramente o autômato não atende a condição 1. Podemos adicionar uma transição $\epsilon, \$ \to \epsilon$ de q_2 a q_4 para que q_1 deixe de ser um estado de aceitação. Note que essa alteração faz com que o autômato atenda as 3 condições solicitadas.



Seja $Q = \{q_1, q_2, q_3, q_4\}$ o conjunto de estados do autômato acima modificado. Vamos denominar por G a gramática que vamos construir a partir do autômato de acordo com o procedimento dado no lema 2.27. A regra inicial é:

$S \to A_{14}$

1. Para cada $p \in Q$, insira a regra $A_{pp} \to \epsilon$ em G

$A_{11} \to \epsilon$
$A_{22} \to \epsilon$
$A_{33} \to \epsilon$
$A_{44} \to \epsilon$

2. Para cada $p, q, r \in Q$, insira a regra $A_{pq} \to A_{pr}A_{rq}$ em G

$A_{11} \to A_{11}A_{11} \mid A_{12}A_{21} \mid A_{13}A_{31} \mid A_{14}A_{41}$
$A_{12} \to A_{11}A_{12} \mid A_{12}A_{22} \mid A_{13}A_{32} \mid A_{14}A_{42}$
$A_{13} \to A_{11}A_{13} \mid A_{12}A_{23} \mid A_{13}A_{33} \mid A_{14}A_{43}$
$A_{14} \to A_{11}A_{14} \mid A_{12}A_{24} \mid A_{13}A_{34} \mid A_{14}A_{44}$

$A_{21} \to A_{21}A_{11} \mid A_{22}A_{21} \mid A_{23}A_{31} \mid A_{24}A_{41}$
$A_{22} \to A_{21}A_{12} \mid A_{22}A_{22} \mid A_{23}A_{32} \mid A_{24}A_{42}$
$A_{23} \to A_{21}A_{13} \mid A_{22}A_{23} \mid A_{23}A_{33} \mid A_{24}A_{43}$
$A_{24} \to A_{21}A_{14} \mid A_{22}A_{24} \mid A_{23}A_{34} \mid A_{24}A_{44}$

$A_{31} \to A_{31}A_{11} \mid A_{32}A_{21} \mid A_{33}A_{31} \mid A_{34}A_{41}$
$A_{32} \to A_{31}A_{12} \mid A_{32}A_{22} \mid A_{33}A_{32} \mid A_{34}A_{42}$
$A_{33} \to A_{31}A_{13} \mid A_{32}A_{23} \mid A_{33}A_{33} \mid A_{34}A_{43}$
$A_{34} \to A_{31}A_{14} \mid A_{32}A_{24} \mid A_{33}A_{34} \mid A_{34}A_{44}$

$A_{41} \to A_{41}A_{11} \mid A_{42}A_{21} \mid A_{43}A_{31} \mid A_{44}A_{41}$
$A_{42} \to A_{41}A_{12} \mid A_{42}A_{22} \mid A_{43}A_{32} \mid A_{44}A_{42}$
$A_{43} \to A_{41}A_{13} \mid A_{42}A_{23} \mid A_{43}A_{33} \mid A_{44}A_{43}$
$A_{44} \to A_{41}A_{14} \mid A_{42}A_{24} \mid A_{43}A_{34} \mid A_{44}A_{44}$

3. Por fim, para cada $p, q, r, s \in Q$, $t \in \Gamma$ e $a, b \in \Sigma_\epsilon$, insira a regra $A_{pq} \to aA_{rs}b$ em G, se $\delta(p, a, \epsilon)$ contém (r, t) e $\delta(s, b, t)$ contém (q, ϵ)

$A_{14} \to \epsilon A_{22}\epsilon \mid \epsilon A_{23}\epsilon$
$A_{23} \to 0A_{22}1 \mid 0A_{23}1$

Com isso, concluímos a construção da gramática G.

Nota: Sobre as variáveis que geram a linguagem vazia, penso naquelas do tipo $A_{11} \to A_{11}A_{11}$, porém, pelas produções em 1 e 2, estas ainda são substituíveis e não ficariam em *loop*. Pensei em eliminar regras desnecessárias com a) encontrar as variáveis que produzem cadeias somente com terminais e b) encontrar as variáveis que são atingíveis a partir da variável inicial S. Porém a) não acontece e b) é sempre verdade pois, se eu gerar o grafo de dependência, sempre consigo chegar em qualquer variável a partir de S. Não sei se entendi bem o que era pra fazer para separar o conjunto de variáveis.

L5.2 (Sipser 2.20) Seja $A/B = \{w \mid wx \in A \text{ para algum } x \in B\}$. Mostre que, se A é livre do contexto e B é regular, então A/B é livre do contexto.

Resposta: TODO

L5.3 (Sipser 2.30) Use o lema do bombeamento para mostrar que as seguintes linguagens não são livres do contexto.

a. $L_a = \{0^n 1^n 0^n 1^n \mid n \geq 0\}$

Resposta: Vamos usar o lema do bombeamento para mostrar que L_a não é livre do contexto. A prova é por contradição.

Suponha o contrário, ou seja, que L_a é livre do contexto. Seja p o comprimento de bombeamento dado pelo lema do bombeamento. Seja s a cadeia $s = 0^p 1^p 0^p 1^p$. Como $s \in L_a$ e $|s| \geq p$, o lema do bombeamento garante que s pode ser dividida em cinco partes $s = uvxyz$, onde, $\forall i \geq 0$, $uv^ixy^iz \in L_a$. Vamos mostrar que isso é impossível, analisando todas as possibilidades de particionamento de s.

Sabemos que s tem a forma:

$$\underbrace{0\ \ 0\ \ 0\ \ 0}_{p}\,\Big|\,\underbrace{1\ \ 1\ \ 1\ \ 1}_{p}\,\Big|\,\underbrace{0\ \ 0\ \ 0\ \ 0}_{p}\,\Big|\,\underbrace{1\ \ 1\ \ 1\ \ 1}_{p}$$

Além disso, a condição 3 do lema do bombeamento diz que $|vxy| \leq p$.

Caso 1: vxy não ultrapassa o limite de um segmento de s

Sem perda de generalidade, vamos considerar que vxy está no primeiro segmento de s. Pelo lema do bombeamento, podemos bombear v e y i vezes $\forall i \geq 0$. Se tomarmos $w = uv^2xy^2z$, claramente teremos mais 0s na primeira metade e, portanto, $w \notin L_a$, o que é uma contradição.

Caso 2: vxy está contido entre o primeiro e o segundo segmento de 0s e 1s em s, tal que $u = 0^a$, $v = 0^b$, $x = 0^c 1^d$, $y = 1^e$ e $z = 1^f 0^p 1^p$, onde a, b, c, d, e e $f \geq 0$, b ou $e \neq 0$, $a + b + c = p$ e $d + e + f = p$

Se tomarmos $i = 0$, temos que $w = uxz$. Logo, $a + c < p$, assim como $d + f < p$, o que provoca um deslocamento da metade de w à esquerda, desbalanceando a quantidade de 0s e 1s da primeira metade em relação à segunda e, portanto, $w \notin L_a$, o que novamente é uma contradição.

Analogamente, este caso cobre a situação em que vxy está contido entre o terceiro e o quarto segmento de 0s e 1s.

Caso 3: vxy está contido entre o segundo e o terceiro segmento de 1s e 0s, ou seja, na metade de s

Como $|vxy| \leq p$, vxy está após a primeira fronteira e antes da terceira fronteira de s. Se tomarmos $i = 0$, obtemos como resultado do bombeamento uma cadeia $w = uxz$, onde o tamanho de cada segmento de w será $p \mid\ < p \mid\ < p \mid p$, respectivamente. Logo, as ocorrências de 0s e 1s da primeira metade não correspondem às da segunda e, portanto, $w \notin L_a$, o que também é uma contradição.

b. $L_b = \{0^n \#0^{2n} \#0^{3n} \mid n \geq 0\}$ - Resposta no livro

c. $L_c = \{w\#t \mid w$ é uma subcadeia de t, onde $w, t \in \{a, b\}^*\}$ - Resposta no livro

d. $L_d = \{t_1 \# t_2 \ldots \# t_k \mid k \geq 2$, cada $t_i \in \{a, b\}^*$, e $t_i = t_j$, para algum $i \neq j\}$

Resposta: Vamos usar o lema do bombeamento para mostrar que L_d não é livre do contexto. A prova é por contradição.

Suponha o contrário, ou seja, que L_d é livre do contexto. Seja p o comprimento de bombeamento dado pelo lema do bombeamento. Seja s a cadeia $s = a^p b^p \# a^p b^p$. Como $s \in L_d$ e $|s| \geq p$, o lema do bombeamento garante que s pode ser dividida em cinco partes $s = uvxyz$, onde, para qualquer $i \geq 0$, $uv^ixy^iz \in L_d$. Vamos mostrar que isso é impossível, analisando todas as possibilidades de particionamento de s.

Caso 1: v e y contêm apenas a's da primeira metade de s, tal que $u = a^l, v = a^m, x = a^n, y = a^q$ e $z = a^r b^p \# a^p b^p$, onde $l, m, n, q, r \geq 0$, m ou $q \neq 0$ e $l+m+n+q+r = p$.

Pelo lema do bombeamento, podemos bombear v e y i vezes, para qualquer $i \geq 0$. Se tomarmos $i = 0$, temos que a cadeia uxz possui menos a's na primeira metade (antes do #) que na segunda (após o #), pois como m ou $q \neq 0$, temos que $l + n + r < p$ e, portanto, esta nova cadeia não pertence a L_d, o que é uma contradição.

Analogamente, este caso cobre a situação em que v e y contêm apenas a's da segunda metade de s.

Caso 2: v e y contêm b's da primeira metade de s e não possuem símbolos da segunda metade.

Temos duas possibilidades:

- $u = a^l, v = a^m b^n, x = b^q, y = b^r$ e $z = b^t \# a^p b^p$, onde $l, m, n, q, r, t \geq 0$, $m + n$ ou $r \neq 0$, $l + m = p$, $n + q + r + t = p$ e $|vxy| \leq p$ ou,
- $u = a^l, v = a^r, x = a^q, y = a^m b^n$ e $z = b^t \# a^p b^p$, onde $l, m, n, q, r, t \geq 0$, $m + n$ ou $r \neq 0$, $l + m + q + r = p$, $n + t = p$ e $|vxy| \leq p$.

Vamos assumir, sem perda de generalidade, que a quantidade de b's de v ou y é maior que zero, caso contrário, voltaríamos ao caso anterior. Dessa forma, para qualquer uma das duas possibilidades, se fizermos um bombeamento de v e y i vezes, para $i = 0$, temos que a cadeia uxz possui menos b's na primeira metade (antes do #) do que na segunda (após o #) e, portanto, esta nova cadeia não pertence a L_d, o que é uma contradição.

Analogamente, este caso cobre a situação em que v e y contêm b's da segunda metade de s.

Caso 3: v e y contêm b's da primeira metade de s e a's da segunda metade de s.

Pela condição 2 do lema do bombeamento, temos que v ou y possuem ao menos um símbolo. Podemos assumir, sem perda de generalidade, que v ou y possui pelo menos um b da primeira metade e pelo menos um a da segunda metade de s, caso contrário, cairíamos em um dos casos já abordados previamente. Ao bombearmos v e y i vezes, para $i = 0$, temos que a cadeia uxz possui menos b's na primeira metade do que na segunda e menos a's na segunda metade do que na primeira, o que é uma contradição, já que essa cadeia não pertence a L_d.

Vale notar que é impossível que v e y possuam símbolos iguais de partes diferentes da cadeia s pela condição 3 do lema do bombeamento.

L5.4 (Sipser 2.32) Seja $\Sigma = \{1, 2, 3, 4\}$ e $C = \{w \in \Sigma^* \mid$ em w, o número de 1s é igual ao número de 2s, e o número de 3s é igual ao número de 4s $\}$. Mostre que C não é livre do contexto.

Resposta: Suponha o contrário, ou seja, que C é livre do contexto. Seja p o comprimento de bombeamento dado pelo lema do bombeamento. Seja s a cadeia $s = 1^p 3^p 2^p 4^p$. Como $s \in C$ e $|s| \geq p$, o lema do bombeamento garante que s pode ser dividida em cinco partes $s = uvxyz$, onde, $\forall i \geq 0$, $uv^i xy^i z \in C$. Vamos mostrar que isso é impossível, analisando todas as possibilidades de particionamento de s.

Caso 1: vxy está contido em um dos segmentos de s

Isso é equivalente a dizer que vxy possui apenas um tipo de símbolo do alfabeto. Sem perda de generalidade, digamos que vxy contém apenas 1s do primeiro segmento de s. Se tomarmos $i = 2$, temos que a cadeia produzida pelo bombeamento $w = uv^2xy^2z$ tem mais 1s do que 2s e, portanto, $w \notin C$, o que nos leva a uma contradição.

Caso 2: vxy está contido exatamente na fronteira entre dois símbolos distintos

Digamos que $vxy \subseteq 1^p 3^p$. Como $|vxy| \leq p$, temos ocorrência tanto de 1s quanto 3s em vxy, caso contrário, voltaríamos ao caso anterior. Se tomarmos $i = 0$, temos que a cadeia produzida pelo bombeamento $w = uxz$ terá um número menor de 1s do que 2s, bem como um número menor de 3s do que 4s e, portanto, $w \notin C$, o que novamente nos leva a uma contradição.

Analogamente, este caso cobre a situação em que $vxy \subseteq 3^p 2^p$ ou $vxy \subseteq 2^p 4^p$.