Esercizi grammatiche context free

a.a. 2021-2022

Corso di Fondamenti di Informatica - 1 modulo

Corso di Laurea in Informatica

Università di Roma "Tor Vergata"

Prof. Giorgio Gambosi

Definire un automa a pila che accetti il linguaggio

$$L = \{a^n b^m | 1 \leq n \leq m\}$$

per pila vuota.

1: soluzione

- L'automa legge la sequenza iniziale di a ponendo sulla pila un simbolo A per ogni simbolo letto.
- L'automa cambia stato per leggere la sequenza di b, eliminando i caratteri A dalla pila.
- Se si raggiunge il fondo della pila (il simbolo Z_0) la stringa va accettata, completando la lettura degli eventuali b mancanti ed eliminando poi Z_0.

1: soluzione

	(q_0, Z_0)	(q_0, A)	(q_1, Z_0)	(q_1, A)
a	(q_0, AZ_0)	(q_0, AA)	-	-
b	-	(q_1, ε)	(q_1, Z_0)	(q_1, ε)
ε	-	-	(q_1, ε)	-

Definire un automa a pila che accetti il linguaggio

$$L = \{w \in \{a, b\}^+ | \#_a(w) \geq \#_b(w)\}$$

Dove $\#_c(x)$ indica il numero di occorrenze del carattere c nella stringa x.

2: soluzione

- L'automa manitine traccia, sulla pila, della differenza tra il numero di caratteri a e il numero di caratteri b letti fino a ora (o vice versa, a seconda che siano stati letti più a o più b).
- La stringa è accettata se al termine della sua lettura la pila è vuota o contiene tutti simboli A.
- Per accettare per pila vuota l'automa prevede che in qualunque istante in cui il numero di a lette è almeno pari al numero di b possa entrare in uno stato q_1 di svuotamento della pila.

2: soluzione

	(q_0, Z_0)	(q_0, A)	(q_0, B)	(q_1, Z_0)	(q_1, A)
a	(q_0, AZ_0)	(q_0, AA)	(q_0, ε)	-	-
b	(q_0, BZ_0)	(q_0, ε)	(q_0, BB)	-	-
ε	(q_1, ε)	(q_1, ε)	-	(q_1, ε)	(q_1, ε)

Si definisca un automa a pila (eventualmente non deterministico) che accetti il linguaggio $L = \{a^r b^s c^t a^n c^n | s = r + t, r, t, n \geq 0\}$.

Si consideri il linguaggio $L \subset \{0, 1\}^*$ tale che $\sigma \in L$ se e solo se $\#_0(\sigma) = \#_1(\sigma)$, dove $\#_a(s)$ indica il numero di occorrenze del carattere a nella stringa s. Si definisca una grammatica CF in GNF che generi L.

4: soluzione

Una possibile grammatica è la seguente:

$$S \quad \longrightarrow \quad 0S1S|1S0S|\varepsilon$$

L'eliminazione della ε-produzione porta alla grammatica equivalente

$$S \quad \longrightarrow \quad 0S1S|1S0S|0S1|01S|1S0|10S|01|10$$

che non presenta produzioni unitarie o simboli inutili.

4: soluzione

La grammatica in CNF che ne deriva è

$$\begin{aligned} S &\rightarrow XY|YX|XU|ZY|YZ|UX|ZU|UZ \\ X &\rightarrow ZS \\ Y &\rightarrow US \\ Z &\rightarrow 0 \\ U &\rightarrow 1 \end{aligned}$$

4: soluzione

da cui deriva immediatamente la grammatica in GNF

$$\begin{aligned}
S &\longrightarrow 0SY|1SX|0SU|0Y|1SZ|1X|0U|1Z \\
X &\longrightarrow 0S \\
Y &\longrightarrow 1S \\
Z &\longrightarrow 0 \\
U &\longrightarrow 1
\end{aligned}$$

Si consideri il linguaggio

$$L = \{0^i 1^j | i \geq j\}$$

Il linguaggio è regolare o context free? Dimostrare quale delle due affermazioni è vera. Si definisca inoltre una grammatica (di tipo 3 o di tipo 2, rispettivamente) che generi tutte e sole le stringhe del linguaggio.

5: soluzione

Il linguaggio non è regolare, ma è context free. Per verificare che non è regolare si può utilizzare il pumping lemma applicato (fissato n)alla stringa $0^n1^n \in L$. Dato che per ogni $uvx = 0^n1^n$ con $|uv| \leq n$ e $|v| \geq 1$ si deve avere necessariamente che $v = 0^k$ per un qualche $k > 0$, si che $uv^0w = uv = 0^{n-k}1^k \notin L$, per cui L non è regolare.

5: soluzione

Una grammatica CF che genera L è ad esempio

$$\begin{aligned} S &\longrightarrow 0S1 | 0T1 | \varepsilon \\ T &\longrightarrow 0T | 0 \end{aligned}$$

Definire un automa a pila che accetta per stato finale il linguaggio composto dalle stringhe $w \in \{0, 1\}^+$ contenenti uno stesso numero di 0 e di 1.

6: soluzione

Un possibile automa ha 2 soli stati q_0, q_F e un alfabeto di pila Z_0, Z, U. Ad ogni istante la pila contiene, al di sopra di Z_0, una sequenza di Z di dimensione pari a $\#(0) - \#(1)$ se $\#(0) - \#(1) > 0$ o una sequenza di U di dimensione pari a $\#(1) - \#(0)$ se $\#(0) - \#(1) < 0$.

	$(q_0, 0)$	$(q_0, 1)$	(q_0, ε)
Z_0	(q_0, ZZ_0)	(q_0, UZ_0)	(q_F, ε)
Z	(q_0, ZZ)	(q_0, ε)	-
U	(q_0, ε)	(q_0, UU)	-

Definire un automa a pila che accetti il seguente linguaggio

$$L = \{a^p b^{p+2q} a^q ; p, q > 0\}$$

descrivendo e giustificando le scelte effettuate.

7: soluzione

- Nello stato q_0 vengono posti nella pila tanti simboli A quanti simboli a sono letti. Lo stato diventa q_1 al primo simbolo b letto
- Nello stato q_1, un simbolo A viene tolto dalla pila per ogni b letto, fino a giungere al fondo della pila e passare in q_2.
- In q_2, per ogni simbolo b letto viene posto sulla pila un simbolo B. L'automa passa in q_3 quando legge un nuovo simbolo a
- In q_3, per ogni simbolo a letto l'automa dovrà togliere due simboli B: per far ciò, passerà ciclicamente in q_3, in cui toglierà la prima B dalla pila avendo letto a, e in q_4, in cui toglierà la seconda B con una ε-transizione.
- Infine, se l'automa si trova in q_4, ed ha quindi tolto BB dalla pila avendo letto a, può eliminare Z_0 dalla pila con una ε. La stringa è accettata per pila vuota.

7: soluzione

	(q_0, Z_0)	(q_0, A)	(q_1, Z_0)	(q_1, A)	(q_2, B)	(q_3, B)	(q_4, Z_0)	(q_4, B)
a	(q_0, AZ_0)	(q_0, AA)	-	-	(q_3, ε)	-	-	(q_3, ε)
b	-	(q_1, ε)	(q_2, BZ_0)	(q_1, ε)	(q_2, BB)	-	-	-
ε	-	-	-	-	-	(q_4, ε)	(q_4, ε)	-

Si consideri il linguaggio

$$L = \{w\#x | w, x \in \{0, 1\}^+, w^R \text{ è suffisso di } x\}$$

Si verifichi che L è context free definendo un automa a pila che lo accetta.

8: soluzione

L'automa dapprima (nello stato q_0) legge w e la trascrive sulla pila in ordine inverso. Alla lettura del carattere # l'automa passa nello stato q_1 di lettura di x: in qualunque passo in cui il carattere letto corrisponde a quello in cima alla pila l'automa effettua una scelta non deterministica tra due opzioni:

1. assumere che w^R compaia in x a partire da questo carattere, in tal caso passa nello stato q_2 ed elimina il primo carattere dalla pila
2. assumere che w^R non compaia in x a partire da questo carattere, e continuare a leggere caratteri, nello stato q_1

Nello stato q_2, l'automa procede nella computazione fin tanto che i caratteri letti corrispondono a quelli via via estratti dalla pila. Nel caso positivo, la stringa termina con Z_0 sulla pila: questo carattere viene quindi estratto con una ε-transizione.

8: soluzione

	(q_0, Z_0)	(q_0, Z)	(q_0, U)	(q_1, Z)	(q_1, U)	(q_2, Z)	(q_2, U)	(q_2, Z_0)
0	(q_0, ZZ_0)	(q_0, ZZ)	(q_0, UZ)	$\{(q_1, Z), (q_2, \varepsilon)\}$	(q_1, U)	(q_2, ε)	-	-
1	(q_0, UZ_0)	(q_0, UZ)	(q_0, UU)	(q_1, Z)	$\{(q_1, U), (q_2, \varepsilon)\}$	-	(q_2, ε)	-
#	-	(q_1, Z)	(q_1, U)	-	-	-	-	-
ε	-	-	-	-	-	-	-	(q_2, ε)

Sia dato il linguaggio

$$L = \{a^n b^m c^k | k = |n - m|\}$$

Definire una grammatica context free che generi il linguaggio.
Discutere se la grammatica risultante è ambigua.

9: soluzione

Una possibile grammatica è la seguente:

$$\begin{array}{rcl} S & \longrightarrow & S_1 | S_3 \\ S_1 & \longrightarrow & aS_1b | S_2 \\ S_2 & \longrightarrow & aS_2c | \varepsilon \\ S_3 & \longrightarrow & S_4S_5 \\ S_4 & \longrightarrow & aS_4b | \varepsilon \\ S_5 & \longrightarrow & bS_5c | \varepsilon \end{array}$$

S_1 corrisponde al caso $n \geq m$, mentre S_3 al caso $m \geq n$.

9: soluzione

La grammatica in questo caso risulta ambigua, in quanto ad esempio la stringa $aabb$ può essere generata sia come
$S \Rightarrow S_1 \Rightarrow aS_1b \Rightarrow aaS_1bb \Rightarrow aabb$ che come
$S \Rightarrow S_3 \Rightarrow S_4S_5 \Rightarrow aS_4bS_5 \Rightarrow aaS_4bbS_5 \Rightarrow aabbS_5 \Rightarrow aabb$

Dimostrare che il seguente linguaggio.

$$L = \{w \in \{a, b\}^+ : \#_w(a) = 2\#_w(b)\}$$

è context free, dove $\#_w(x)$ indica il numero di occorrenze del carattere x nella stringa w

10: soluzione

Una possibile soluzione è quella di definire un PDA che accetta il linguaggio.

	(q_0, Z_0)	(q_0, X)	(q_0, Y)	(q_1, Z_0)	(q_1, Y)
a	(q_0, XXZ_0)	(q_0, XXX)	(q_1, ε)	-	-
b	(q_0, YZ_0)	(q_0, ε)	(q_0, YY)	-	-
ε	(q_0, ε)	-	-	(q_0, X)	(q_0, ε)

Si definisca un automa a pila (eventualmente non deterministico) che accetti il linguaggio $L = \{a^r b^s c^t | t = r - s\}$.