COMPILATORI

LINGUAGGI E COMPILATORI

Prova scritta dell'Appello del 18 aprile 2012

DODA	**21**
LO MAGRO	**23**
ROCCO	**insufficiente**

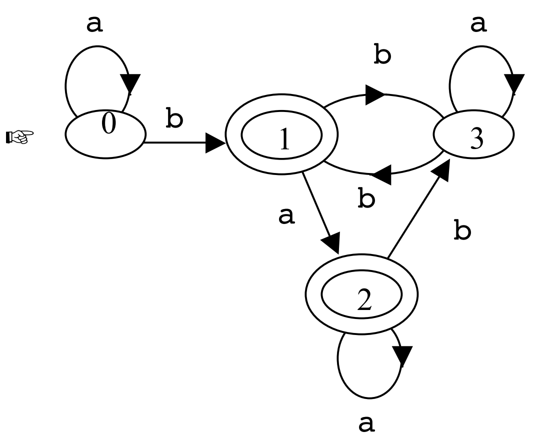
ESERCIZIO 1.
Data la grammatica acontestuale **G** con simboli nonterminali `S` e `X` e con produzioni

$\mathtt{S} \to \mathtt{aS} \mid \mathtt{X}$ $\qquad\qquad$ $\mathtt{X} \to \mathtt{bXc} \mid \varepsilon$

(1*a*) Costruire una grammatica **G′** in forma normale di Chomsky equivalente a **G**. Applicare l'algoritmo CYK a **G′** per costruire gli alberi di derivazione della frase `abc`.

(1*b*) Applicare gli algoritmi per il calcolo delle funzioni $\mathcal{I}(\alpha)$ e $\mathcal{J}(A)$ (che vanno riportati) alla grammatica **G** per decidere se **G** è o meno una grammatica *LL*(1).

ESERCIZIO 2.
Dato l'automa finito deterministico $\mathbf{D}_0$ su $\Sigma = \{\mathtt{a}, \mathtt{b}\}$



(2*a*) Trovare per via algoritmica un automa finito deterministico **D** di dimensione minima equivalente a $\mathbf{D}_0$, e applicare l'algoritmo di Moore per dimostrare che **D** e $\mathbf{D}_0$ sono equivalenti.

(2*b*) Facendo uso del Lemma di Arden trovare un'espressione regolare **E** che denoti il linguaggio accettato dall'automa finito deterministico **D**.

(2*c*) Costruire l'automa di Thomson associato all'espressione **E**.

(2*d*) Costruire una grammatica regolare che generi il linguaggio accettato dall'automa finito deterministico **D**.

Soluzione

ESERCIZIO 1.

(1*a*) La grammatica **G** non contiene simboli improduttivi.

Le variabili annullabili di **G** sono S ed X.

Eliminazione della produzioni nulla:

$S \to a \mid aS \mid X \qquad X \to bc \mid bXc$

Eliminazione della produzione unitaria:

$S \to a \mid aS \mid bc \mid bXc \qquad X \to bc \mid bXc$

Eliminazione delle produzioni multiple:

$S \to a \mid AS \mid BC \mid YC \qquad X \to BC \mid YC \qquad Y \to BX$
$A \to a \qquad B \to b \qquad C \to c$

A questo punto, per ottenere una grammatica **G*** equivalente a **G** in FNC va aggiunta la produzione nulla $S \to \varepsilon$ senonché S compare nel corpo della produzione $S \to AS$. Allora viene introdotto un nuovo simbolo iniziale, S*, cosicché **G*** = (*V**, *T*, *P**, S*) dove *V** = {S*, S, X, Y, A, B, C} e *P** contiene le produzioni:

$S^* \to S \mid \varepsilon$
$S \to a \mid AS \mid BC \mid YC \qquad X \to BC \mid YC \qquad Y \to BX$
$A \to a \qquad B \to b \qquad C \to c$

Siccome **G*** contiene una produzione unitaria, questa viene eliminata e *P** diventa

$S^* \to a \mid AS \mid BC \mid YC \mid \varepsilon$
$S \to a \mid AS \mid BC \mid YC \qquad X \to BC \mid YC \qquad Y \to BX$
$A \to a \qquad B \to b \qquad C \to c$

Vogliamo infine decidere se la stringa x = `abc` è una frase di **G***. Le sottostringhe $x_{i,j}$ di x sono di seguito riportati.

11 a	12 ab	13 abc
21 b	22 bc	-
31 c	-	-

Tabella delle sottostringhe $x_{i,j}$

Con l'algorimo CYK calcoliamo gli insiemi $N_{i,j}$:

11 S*, S, A	12 ∅	13 S*, S
21 B	22 S*, S, X	-
31 C	-	-

Tabella degli insiemi $N_{i,j}$

e, visto che il simbolo speciale `S*` appartiene ad $N_{1,3}$, possiamo concludere che la stringa `abc` è una frase di **G*** ed esiste infine un solo albero di derivazione.

(1*b*) La grammatica **G** non contiene produzioni ricorsive a sinistra. Inoltre abbiamo

$\mathcal{I}(\mathtt{aS}) = \{\mathtt{a}\}$ $\mathcal{I}(\mathtt{X}) = \{\mathtt{b}, \varepsilon\}$ $\mathcal{I}(\mathtt{bXc}) = \{\mathtt{b}\}$ $\mathcal{I}(\varepsilon) = \{\varepsilon\}$
$\mathcal{J}(\mathtt{S}) = \{\#\}$ $\mathcal{J}(\mathtt{X}) = \{\mathtt{c}, \#\}$

e, dunque, la tabella di controllo

a	b	c	#
S→aS	S→X		S→X
	X→bXc	X→ε	X→ε

Siccome nessuna cella è multipla, **G** è una grammatica *LL*(1).

ESERCIZIO 2.

(2*a*) L'automa finito deterministico $\mathbf{D}_0$ ha la seguente tabella delle transizioni

q	a	b
0	0	1
1	2	3
2	2	3
3	3	1

Si parte dalla partizione $\{0', 1'\}$ di Q dove $0' = \{0, 3\}$ e $1' = \{1, 2\}$. Si vede facilmente che partizione di Q non può essere raffinata. Pertanto, un automa finito deterministico di dimensione minima equivalente a $\mathbf{D}_0$ è l'automa finito deterministico $\mathbf{D}$ con tabella delle transizioni

q'	a	b
$0'$	$0'$	$1'$
$1'$	$1'$	$0'$

in cui $0'$ è lo stato iniziale e $1'$ è l'unico stato di accettazione. Applicando l'algoritmo di Moore si prova facilmente che $\mathbf{D}$ è equivalente a $\mathbf{D}_0$.

(2*b*) Per l'automa finito deterministico $\mathbf{D}$ abbiamo il sistema di equazioni

$$\lambda_{0'} = \{\mathtt{a}\} \cdot \lambda_{0'} + \{\mathtt{b}\} \cdot \lambda_{1'}$$
$$\lambda_{1'} = \{\mathtt{a}\} \cdot \lambda_{1'} + (\{\mathtt{b}\} \cdot \lambda_{0'} + \{\varepsilon\})$$

Il Lemma di Arden fornisce la seguente soluzione della seconda equazione:

$$\lambda_{1'} = \{\mathtt{a}\}^* \cdot (\{\mathtt{b}\} \cdot \lambda_{0'} + \{\varepsilon\}) = \{\mathtt{a}\}^* \cdot \{\mathtt{b}\} \cdot \lambda_{0'} + \{\mathtt{a}\}^*$$

che sostituita nella prima equazione dà

$$\lambda_{0'} = \{\mathtt{a}\} \cdot \lambda_{0'} + \{\mathtt{b}\} \cdot \{\mathtt{a}\}^* \cdot \{\mathtt{b}\} \cdot \lambda_{0'} + \{\mathtt{b}\} \cdot \{\mathtt{a}\}^* =$$
$$= (\{\mathtt{a}\} + \{\mathtt{b}\} \cdot \{\mathtt{a}\}^* \cdot \{\mathtt{b}\}) \cdot \lambda_{0'} + \{\mathtt{b}\} \cdot \{\mathtt{a}\}^*$$

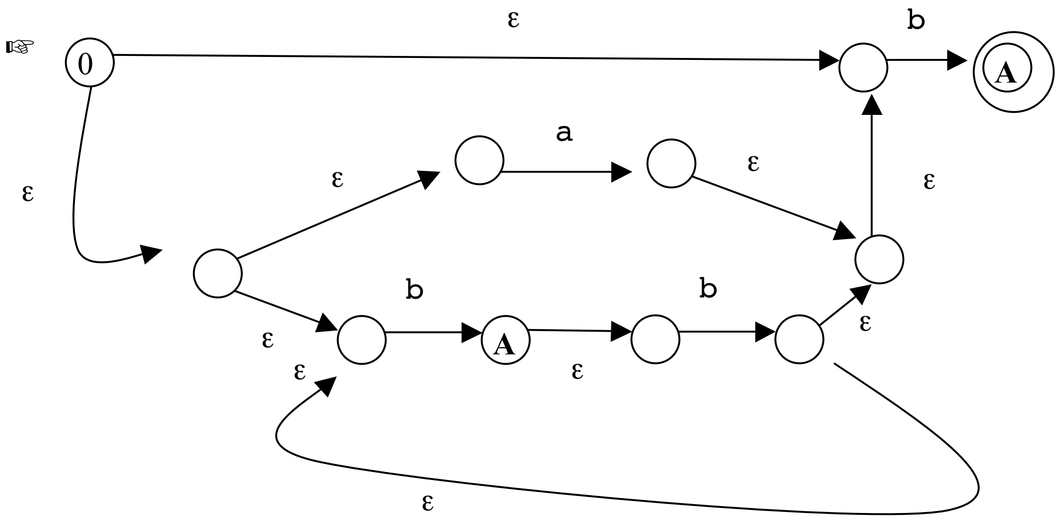
Il Lemma di Arden fornisce la seguente soluzione della prima equazione:

$$\lambda_{0'} = (\{\mathtt{a}\} + \{\mathtt{b}\} \cdot \{\mathtt{a}\}^* \cdot \{\mathtt{b}\})^* \cdot \{\mathtt{b}\} \cdot \{\mathtt{a}\}^*$$

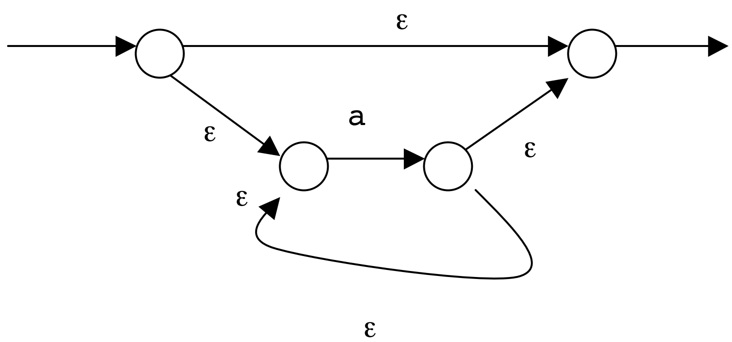
Pertanto il linguaggio accettato da $\mathbf{D}$ è denotato dall'espressione regolare

$$\mathbf{E} = (\mathtt{a}|\mathtt{ba^*b})^*\mathtt{ba^*}.$$

(2*c*) L’automa di Thompson associato ad **E** è mostrato in figura.



dove **A** è l’automa associato all’espressione `a*`.



(2*d*) **G** = (V, T, P, 0′) dove V = {0′, 1′}, T = {`a`, `b`} e P contiene le seguenti produzioni:

0′ → `a` 0′ | `b` 1′ 1′ → `a` 1′ | `b` 0′ | ε