Automi a stati finiti non deterministici

Arturo Carpi

Dipartimento di Matematica e Informatica
Università di Perugia

Corso di Linguaggi Formali e Compilatori - a.a. 2021/22

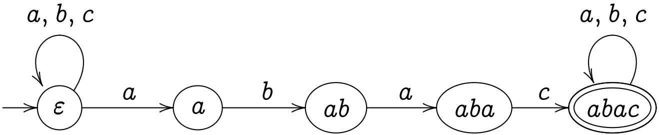
Grafo di un automa

1. Un automa è rappresentato da un grafo;
2. per ogni stato q e ogni lettera a, esiste esattamente un arco uscente da q con etichetta a;
3. le parole accettate da $\mathcal{A}$ sono le etichette dei cammini dallo stato iniziale a uno stato finale.

Se si eliminasse la **2** . . .

pro	libertà nella progettazione
contro	difficile verificare se una parola è accettata

Esempio



Automa a stati finiti non deterministico

Soluzione

- Progetto con modello non-deterministico;
- conversione automatica al modello deterministico.

Il modello non deterministico

- a ogni passo il dispositivo esegue una fra le transizioni possibili (purchè esista);
- l'input è accettato se almeno una delle possibili computazioni ha successo (termina in uno stato finale)
- l'input è rifiutato se tutte le possibili computazioni falliscono (terminano in uno stato non finale o terminano prima che il nastro sia esaurito)

Definizione

Un automa a stati finiti non deterministico è una quintupla $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F \rangle$, dove Q, Σ, q_0, F sono come nella definizione dell'automa deterministico e

$$\delta : Q \times \Sigma \rightarrow \wp(Q)$$

è la funzione di transizione.

Una parola $w = a_1 a_2 \cdots a_n$, $(a_1, a_2, \ldots, a_n \in \Sigma,\ n \geq 0)$, è accettata da $\mathcal{A}$ se esistono stati $q_1, q_2, \ldots, q_n \in Q$ tali che

$$q_i \in \delta(q_{i-1}, a_i),\ 1 \leq i \leq n, \quad q_n \in F.$$

L'insieme delle parole accettate da $\mathcal{A}$ si dice linguaggio accettato (o riconosciuto) da $\mathcal{A}$ e si denota con $L(\mathcal{A})$.

Grafo dell'automa non deterministico

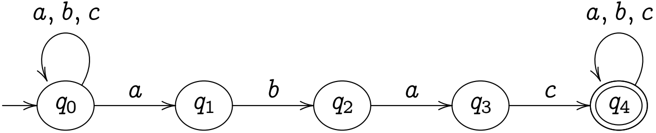
A ogni automa non deterministico $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F \rangle$ associamo un grafo diretto con frecce etichettate:

- i vertici sono gli stati;
- le frecce sono le triple (q, a, p) con $q \in Q,\ a \in \Sigma,\ p \in \delta(q, a)$;
- lo stato iniziale è denotato da una freccia entrante;
- gli stati finali sono identificati dal bordo doppio.

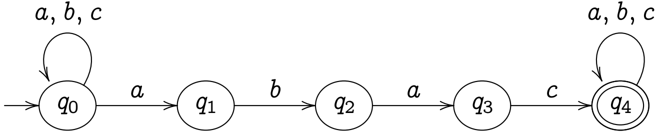
Ci si convince facilmente che

- una parola w può essere etichetta di molti cammini uscenti dallo stato iniziale, o anche nessuno;
- la parola w è accettata se e solo se c'è un cammino dallo stato iniziale a uno stato finale con etichetta w;
- questo non esclude però che vi possano essere altri cammini con etichetta w che partono dallo stato iniziale e terminano in uno stato non finale.

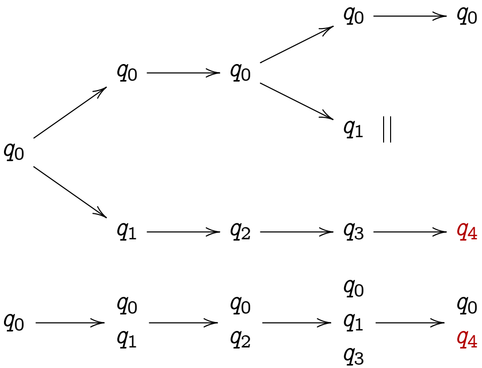
Esempio



δ	a	b	c	F
q_0	$\{q_0, q_1\}$	$\{q_0\}$	$\{q_0\}$	
q_1	$\emptyset$	$\{q_2\}$	$\emptyset$	
q_2	$\{q_3\}$	$\emptyset$	$\emptyset$	
q_3	$\emptyset$	$\emptyset$	$\{q_4\}$	
q_4	$\{q_4\}$	$\{q_4\}$	$\{q_4\}$	$\times$


a, b, c
q0
a
q1
b
q2
a
q3
c
q4
a, b, c

Input : a b a c


q0
q0
q0
q0
q0
q1
q1
q2
q3
q4
q0
q0
q1
q0
q2
q0
q1
q3
q0
q4

Determinizzazione

Teorema Sia $\mathcal{A}$ un automa a stati finiti non deterministico. Esiste effettivamente un automa a stati finiti deterministico $\mathcal{A}'$ tale che

$$L(\mathcal{A}) = L(\mathcal{A}')\,.$$

Dimostrazione

Idea: Gli stati dell'automa deterministico registreranno l'insieme gli stati raggiungibili dalle computazioni dell'automa non deterministico sul medesimo input;

Accetto se fra questi ce n'è uno finale

Sia $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F \rangle$ l'automa non deterministico.

Definisco l'automa deterministico $\mathcal{A}' = \langle Q', \Sigma, \delta', s_0, F' \rangle$ come segue:

- l'insieme degli stati è l'insieme $Q' = \wp(Q)$ costituito dai sottoinsiemi di Q,
- lo stato iniziale è $s_0 = \{q_0\}$,
- gli stati finali sono tutti i sottoinsiemi di Q che contengono almeno un elemento di F, cioè

$$F' = \{s \in \wp(Q) \mid s \cap F \neq \emptyset\},$$

- la funzione di transizione $\delta' \colon Q' \times \Sigma \to Q'$ è definita da

$$\delta'(s, a) = \bigcup_{q \in s} \delta(q, a), \quad s \in \wp(Q),\ a \in \Sigma.$$

Per ogni $w \in \Sigma^*$ risulta

$$\widehat{\delta}'(s_0, w) = \{q \in Q \mid \text{nel grafo di } \mathcal{A} \text{ c'è un cammino da } q_0 \text{ a } q \text{ con etichetta } w\}$$

(dimostrazione per induzione sulla lunghezza di w).

Pertanto sono equivalenti:

- w è accettata da $\mathcal{A}$;
- $\widehat{\delta}'(s_0, w)$ contiene uno stato di F;
- $\widehat{\delta}'(s_0, w) \in F'$;
- w è accettata da $\mathcal{A}'$.

Quindi $L(\mathcal{A}) = L(\mathcal{A}')$.

Stati inaccessibili

Osservazione

Se $\mathrm{Card}\, Q = n$, allora $\mathrm{Card}\, Q' = 2^n$.
Ma non tutti gli stati sono necessari!

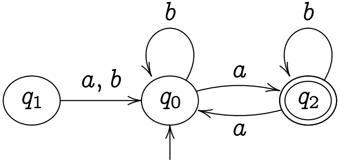
Definizione

Sia $\mathcal{A} = \langle Q, \Sigma, \delta, q_0, F \rangle$ un automa a stati finiti deterministico. Uno stato $q \in Q$ si dice accessibile se $q = \widehat{\delta}(q_0, w)$ per qualche $w \in \Sigma^*$.

Osservazione

Eliminando gli stati inaccessibili e restringendo di conseguenza la funzione di transizione, si ottiene un automa equivalente.

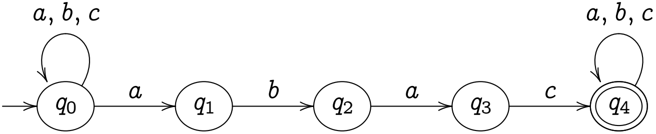
Esempio



Algoritmo di determinizzazione

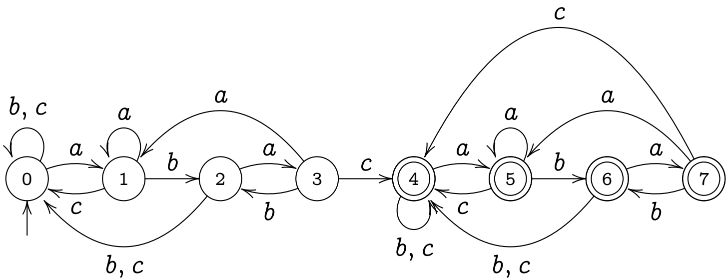
1. inserisco nella lista degli stati $s_0 = \{q_0\}$;
2. per ogni stato r della lista e ogni lettera $a \in \Sigma$
3. calcolo $s = \delta'(r, a)$;
4. se s non è nella lista degli stati allora
5. appendo s alla lista degli stati;
6. se s contiene uno stato finale di $\mathcal{A}$ aggiungo s alla lista degli stati finali;

Esempio



i	s_i	$\delta'(i, a)$	$\delta'(i, b)$	$\delta'(i, c)$	$i \in F$
0	q_0	1	0	0	
1	q_0, q_1	1	2	0	
2	q_0, q_2	3	0	0	
3	q_0, q_1, q_3	1	2	4	
4	q_0, q_4	5	4	4	$\times$
5	q_0, q_1, q_4	5	6	4	$\times$
6	q_0, q_2, q_4	7	4	4	$\times$
7	q_0, q_1, q_3, q_4	5	6	4	$\times$

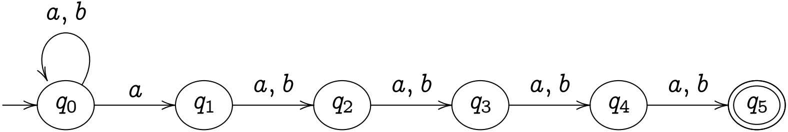
Esempio



i	s_i	$\delta'(i, a)$	$\delta'(i, b)$	$\delta'(i, c)$	$i \in F$
0	q_0	1	0	0	
1	q_0, q_1	1	2	0	
2	q_0, q_2	3	0	0	
3	q_0, q_1, q_3	1	2	4	
4	q_0, q_4	5	4	4	$\times$
5	q_0, q_1, q_4	5	6	4	$\times$
6	q_0, q_2, q_4	7	4	4	$\times$
7	q_0, q_1, q_3, q_4	5	6	4	$\times$

Un caso difficile

Le parole sull'alfabeto $\{a, b\}$ la cui 5-ultima lettera è a.



Un automa deterministico equivalente richiede almeno 2^5 stati.