Marco Casu

Automi, Calcolabilità e Complessità

SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Ingegneria dell'Informazione, Informatica e Statistica
Dipartimento di Informatica

Questo documento è distribuito sotto la licenza GNU, è un resoconto degli appunti (eventualmente integrati con libri di testo) tratti dalle lezioni del corso di Automi, Calcolabilità e Complessità per la laurea triennale in Informatica. Se dovessi notare errori, ti prego di segnalarmeli.

INDICE

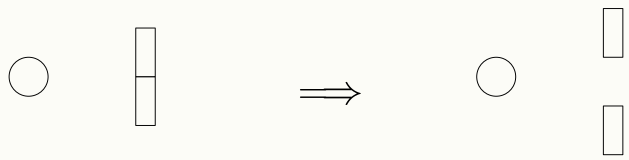
CAPITOLO

1 AUTOMI

1.1 Linguaggi Regolari

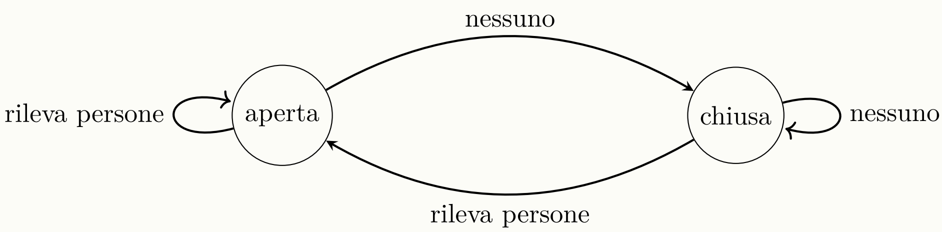
Un *automa a stati finiti* è, seppure limitato nella memoria e nella gestione dell'input, il più semplice modello di computazione. Un automa può interagire con l'input esclusivamente "scorrendolo" in maniera sequenziale.

Esempio : Si vuole modellare una semplice porta con sensore, che si apre quando qualcuno si trova nelle vicinanze.



una persona si trova davanti la porta $\Longrightarrow$ la porta si apre

Un automa che modella il problema è il seguente :



Un automa ha alcuni stati speciali, come quello iniziale, indicato con un apposita freccia, e degli stati detti *di accettazione*, ossia stati in cui deve necessariamente terminare la computazione per essere definita valida, vengono rappresentati con un doppio cerchio.

Il modello di calcolo degli automi è riconducibile al concetto di *linguaggio regolare*, che verrà formalizzato in seguito, segue ora una definizione formale di automa.

Definizione (DFA) : : Un DFA (Deterministic Finite Automa) è una 5-tupla, $(Q, \Sigma, \delta, q_0, F)$ di cui

- Q è l'insieme degli stati possibili
- Σ è l'alfabeto che compone le stringhe in input

- δ è una mappa $Q \times \Sigma \to Q$ detta *funzione di transizione*.
- $q_0 \in Q$ è lo stato iniziale.
- $F \subseteq Q$ è l'insieme degli stati di accettazione.

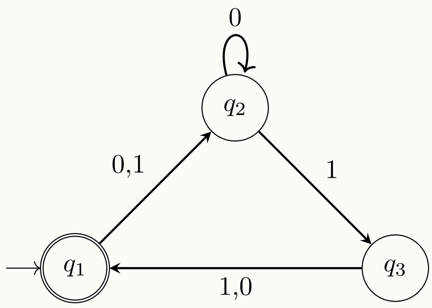


Figura 1.1: semplice automa

Nell'esempio in figura 1.1, si ha che

- $Q = \{q_1, q_2, q_3\}$
- $\Sigma = \{0, 1\}$
- $F = \{q_1\}$
- $q_0 = q_1$
- $\delta =$

	0	1
q_1	q_2	q_2
q_2	q_2	q_3
q_3	q_1	q_1

Sia D un DFA, chiamiamo **linguaggio dell'automa**, e denotiamo $L(D)$, l'insieme delle stringhe che date in input a D fanno si che D termini su uno stato di accettazione. Per definire formalmente un linguaggio

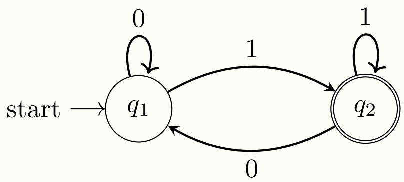


Figura 1.2: il linguaggio di tale automa risulta essere composto dalle stringhe che terminano con 1

di un automa, è necessario introdurre la **funzione di transizione estesa**:

$$\delta^*(q, \epsilon) = \delta(q, \epsilon)$$

$$\delta^*(q, ax) = \delta^*(\delta(q, a), x)$$

dove

$$a \in \Sigma, \quad x \in \Sigma^*, \quad \epsilon = \text{ stringa vuota}$$

Σ^* è l'insieme di tutte le stringhe formate dall'alfabeto Σ. Passiamo ora alla definizione di **configurazione**, essa rappresenta lo stato dell'automa ad un certo punto della computazione, essa è formata da una coppia

$$Q \times \Sigma^*$$

Rappresentante uno stato, ed una stringa di input rimanente da computare.

Un **passo della computazione** in un automa rappresenta una transizione da una configurazione ad un altra, è una relazione binaria $\vdash_D : Q \times \Sigma^*$ tale che

$$(p, ax) \vdash_D (q, x) \iff \delta(p, a) = q \quad \text{dove} \quad p, q \in Q, \quad a \in \Sigma, \quad x \in \Sigma^*$$

Si può estendere la definizione di passo di computazione, considerando la sua *chiusura transitiva* $\vdash_D^*$. Essa si ottiene aggiungendo a $\vdash_D$ tutte le coppie in $Q \times \Sigma^*$ che rendono la relazione chiusa rispetto la riflessività e rispetto la transitività.

$$(q, aby) \vdash_D (p, by) \ \wedge \ (p, by) \vdash_D (ry) \implies (q, aby) \vdash_D^* (r, y)$$

Ad esempio, nell'automa in figura 1.2, risulta chiaro che

$$\begin{cases} (q_1, 011) \vdash_D (q_1, 11) \\ (q_1, 11) \vdash_D (q_2, 1) \\ (q_2, 1) \vdash_D (q_2, \epsilon) \end{cases} \implies (q_1, 011) \vdash_D^* (q_2, \epsilon)$$

Inoltre

$$\begin{aligned} \delta^*(q_1, 011) = \\ \delta^*(q_1, 11) = \\ \delta^*(q_2, 1) = \\ \delta^*(q_2, \epsilon) = q_2 \end{aligned}$$

Se non specificato diversamente, con ϵ verrà indicata la stringa vuota. Utilizzando le precedenti definizioni, è possibile definire formalmente quali sono gli input accettati da un DFA.

Definizione : : Sia $D = (Q, \Sigma, \delta, q_0, F)$ un DFA, e sia $x \in \Sigma^*$ una stringa, essa è **accettata** da D se

$$\delta^*(q_0, x) \in F$$

Il **linguaggio riconosciuto** da D è

$$L(D) = \{x \in \Sigma^* | \delta^*(q_0, x) \in F\}$$

Definizione (Linguaggi Regolari) : : L'insieme dei linguaggi regolari, denotato REG, contiene tutti i linguaggi, tali che esiste un DFA che li ha come linguaggi riconosciuti.

$$REG = \{L \mid \exists D = (Q, \Sigma, \delta, q_0, F) \ \texttt{t.c.} \quad L \in \Sigma^* \wedge L(D) = L\}$$

Uno fra gli scopi di questo corso riguarda il capire come progettare automi, e capire se, ogni linguaggio è regolare, o ce ne sono alcuni che non possono essere riconosciuti da alcun possibile DFA.

1.1.1 Esempi di DFA

Vediamo in questa sezione alcuni semplici esempi di DFA.

Esempio 1) Si vuole progettare un DFA che accetti il seguente linguaggio

$$\{x \in \{0, 1\}^* \mid w_h(x) \geq 3\}$$

Si ricordi come

$$w_h(x) = \text{ occorrenze di 1 in } x$$

Una volta progettato il DFA, è anche importante dimostrarne la correttezza, ossia dare una prova matematico che l'automa in questione accetti il linguaggio.

- Se $x \in L(D)$ allora D accetta x
- Se D accetta x allora $w_h(x) \geq 3$

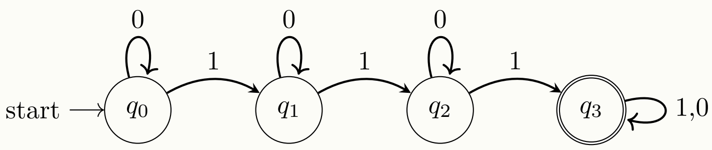


Figura 1.3: Esempio (1) di DFA

In questo, e nei seguenti casi, essendo i DFA estremamente semplici, risulta ovvio che accettino il dato linguaggio, in casi più avanzati, sarà necessario fornire una dimostrazione rigorosa.

Esempio 2) Si vuole progettare un DFA che accetti il seguente linguaggio

$$\{x \in \{0,1\}^* \mid x = 1y \land y \in \{0,1\}^*\}$$

Appunto sulla notazione : Se $a \in \Sigma^*$ e $b \in \Sigma^*$, allora con ab si denota la concatenazione di stringhe.

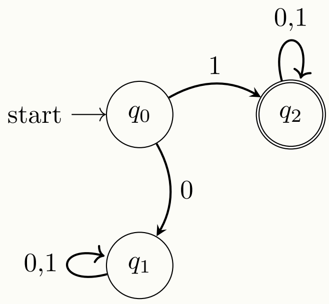


Figura 1.4: Esempio (2) di DFA

Nell'esempio (2), quando dallo stato q_0 il DFA riceve in input 0, la computazione cade su uno stato "buco nero", dalla quale non si può uscire a prescindere dall'input, l'operazione che fa cadere in questo stato è da considerarsi "non definita" in quanto non porterà mai la computazione a terminare su uno stato accettabile, è quindi comodo rimuovere tale stato dal diagramma.

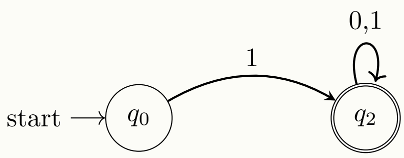


Figura 1.5: Esempio (2.1) di DFA

Anche in questo caso la dimostrazione della correttezza risulta banale.

Esempio 3) Si vuole progettare un DFA che accetti il seguente linguaggio

$$\{x \in \{0,1\}^* \mid x = 0^n 1, \quad n \in \mathbb{N}\}$$

Con $0^n 1$ si intende una stringa che sia composta esclusivamente da 0, ma con un 1 come ultimo termine, ad esempio :

0000000000000001

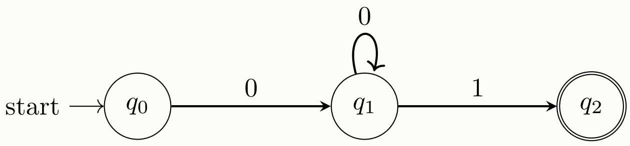


Figura 1.6: Esempio (3) di DFA

1.2 Operazioni sui Linguaggi

Lo studio delle proprietà dei linguaggi regolari può fornire opportune accortezze utili nella progettazione di automi, siccome i linguaggi sono insiemi di stringhe costruiti su un alfabeto Σ, essi godono delle operazioni insiemistiche.

Risulta utile definire formalmente la concatenazione fra stringhe, siano

$$x = a_1, a_2 \dots, a_n \qquad y = b_1, b_2 \dots, b_n$$

due stringhe, esse possono essere concatenate

$$xy = a_1, a_2 \dots, a_n, b_1, b_2 \dots, b_n$$

L'operazione di concatenazione non è commutativa, può essere definita ricorsivamente in tal modo :

$$x(ya) = (xy)a$$

dove

$$x, y \in \Sigma^* \qquad a \in \Sigma$$

Siano L_1, L_2 due linguaggi regolari in REG (per semplicità, definiti su uno stesso alfabeto Σ), e sia n un numero naturale, sono definite su di essi le seguenti operazioni :

- **unione** : $L_1 \cup L_2 = \{x \in \Sigma^* \mid x \in L_1 \lor x \in L_2\}$
- **intersezione** : $L_1 \cap L_2 = \{x \in \Sigma^* \mid x \in L_1 \land x \in L_2\}$
- **complemento** : $\neg L_1 = \{x \in \Sigma^* \mid x \notin L_1\}$
- **concatenazione** : $L_1 \circ L_2 = \{xy \mid x \in L_1 \land y \in L_2\}$
- **potenza** : $L_1^n = \underbrace{L_1 \circ L_1 \circ L_1, \cdots \circ L_1}_{n \text{ volte}}$
- **star** : $L_1^* = \{x_1, x_2 \dots, x_k \mid k \in \mathbb{Z}^+ \land x_i \in L_1\}$
  Si può definire anche diversamente
  $$L_1^* = \bigcup_{k=0}^{\infty} L_1^k$$

Esempio di concatenazione e potenza :

$$\Sigma = \{a, b\} \quad L_1 = \{a, ab, ba\} \quad L_2 = \{ab, b\} \quad L = \{a, ab, ba\}$$

$$L_1 \circ L_2 = \{aab, ab, abab, abb, baab, bab\}$$

$$L^2 = \{aa, aab, aba, abab, abba, baa, baba\}$$

Teorema (Chiusura di *REG*) : La classe dei linguaggi regolari REG, è chiusa rispetto a tutte le operazioni appena elencate, siano L_1 ed L_2 due linguaggi regolari, allora :

$$L_1 \cup L_2 \in REG \qquad L_1 \circ L_2 \in REG \qquad L_1 \cap L_2 \in REG$$

$$L_1^n \in REG \qquad \neg L_1 \in REG \qquad L_1^* \in REG$$

Dimostrazione (unione ed intersezione) : Siano L_1 ed L_2 due linguaggi regolari, considero due DFA, per semplicità, con lo stesso alfabeto

$$D_1 = (Q_1, \Sigma, \delta, q_1, F_1)$$

$$D_2 = (Q_2, \Sigma, \delta, q_2, F_2)$$

tali che

$$L(D_1) = L_1 \ \wedge \ L(D_2) = L_2$$

Si costruisce un DFA che simula contemporaneamente l'esecuzione di D_1 e D_2, in cui gli stati possibili saranno le possibili combinazioni di coppie di stati. Si definisce $D = (Q, \Sigma, \delta, q_0, F)$ tale che

- $Q = Q_1 \times Q_2 = \{(r_1, r_2) \| \ r_1 \in Q_1 \wedge r_2 \in Q_2\}$
- $\delta((r_1, r_2), a) = (\delta(r_1, a), \delta(r_2, a))$ dove $a \in \Sigma$ e $(r_1, r_2) \in Q$
- $q_0 = (q_1, q_2)$
- $F = (F_1 \times Q_2) \cup (Q_1 \times F_2) = \{(r_1, r_2) \| \ r_1 \in F_1 \vee r_2 \in F_2\}$

Nel caso si dovesse dimostrare la proprietà dell'intersezione, si avrebbe che

- $F = F_1 \times F_2$

A questo punto risulta chiaro che

(i) $x \in L_1 \cup L_2 \implies x \in L(D)$

(ii) $x \in L(D) \implies x \in L_1 \cup L_2$ ■

Dimostrazione (complemento) : : Sia L un linguaggio regolare, e D un automa che lo accetta $L(D) = L$. Si vuole dimostrare che esiste un automa che accetti $L^C = \{w \in \Sigma^* \mid w \notin L\}$. Essendo

$$D = (Q, \Sigma, \delta, q_0, F)$$

considero

$$D' = (Q, \Sigma, \delta, q_0, F^C)$$

dove $F^C = Q \backslash F$. Supponiamo che $w \in L^C$, allora sicuramente, se data come input a D', la computazione terminerà in uno stato che non è in F, dato che, per definizione, se terminasse in F, sarebbe accettato da D, ma L^C contiene tutte le stringhe che non sono accettate da D, quindi

- $w \in L^C$
- la computazione termina in uno stato $q \in F^C$
- D' accetta w
- L^C è un linguaggio regolare. ■

Per dimostrare la proprietà di concatenazione, è necessario introdurre un nuovo concetto.

1.3 Non Determinismo

Si può generalizzare il modello di DFA, in modo tale che la lettura di un input non scaturisca il passaggio da uno stato ad un'altro, ma da uno stato ad un insieme di stati, tale generalizzazione è detta *Non-Deterministic Finite Automa.*

Definizione (NFA) : : Un NFA è una tupla $N = (Q, \Sigma, \delta, q_0, F)$ tale che

- Q, Σ, q_0, F condividono la definizione con i loro corrispettivi nel DFA
- $\delta : Q \times \Sigma_\epsilon \to \mathcal{P}(Q)$

- $\Sigma_\epsilon = \Sigma \cup \{\epsilon\}$
- $\mathcal{P}(Q)$ è l'insieme delle parti di Q

Una computazione in un NFA è paragonabile ad una computazione parallela, in cui un input può risultare in diversi *rami* di computazione. Una funzione di transizione di un NFA inoltre accetta la stringa vuota ϵ, se la computazione finisce in uno stato in cui è presente un arco di questo tipo, verrà considerata anche una diramazione verso quell'arco a prescindere dall'input.

Una computazione si può rappresentare graficamente con un albero, se una delle diramazioni possibili termina in uno stato accettabile, allora la stringa in input è accettata.

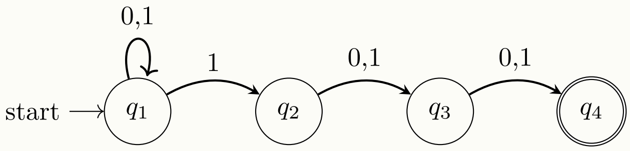


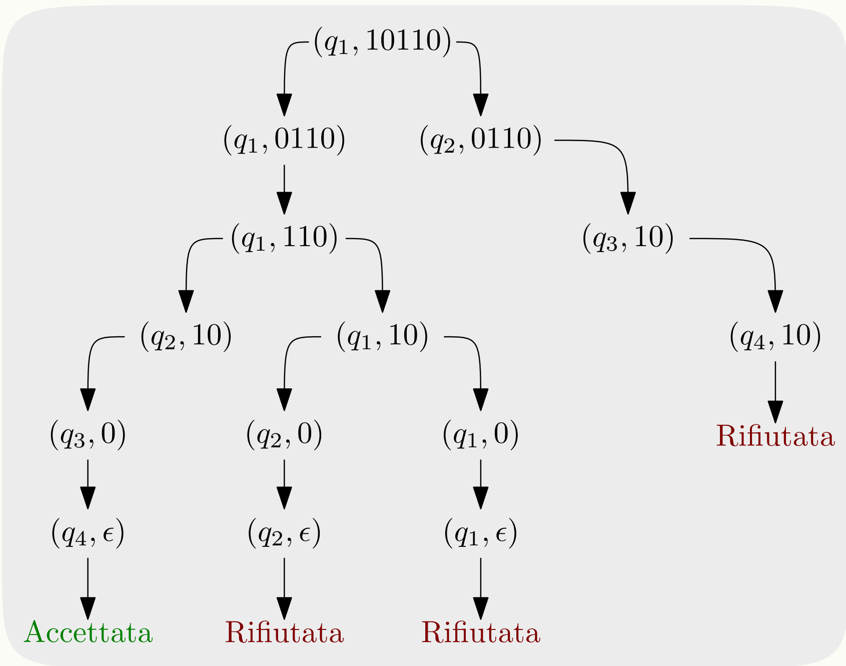
Figura 1.7: Esempio di NFA

Sia N l'NFA mostrato in figura 1.7, si ha

$$L(N) = \{x \in \{0,1\}^* \| \ x \text{ ha 1 come terzultimo valore } \}$$

Si può visualizzare il seguente albero di computazione data come input una stringa w :

$$w = 10110$$



Essendo che la traccia di sinistra accetta w, allora N accetta w.

È necessario estendere il concetto di *configurazione* per gli NFA, essa, rappresentante uno stato della computazione, sarà una coppia

$$(q, x) \in Q \times \Sigma_\epsilon$$

E diremo che

$$(p, ax) \vdash_N (q, x) \iff q \in \delta(p, a)$$

dove

$$p, q \in Q \quad a \in \Sigma_\epsilon \quad x \in \Sigma_\epsilon^*$$

Si considera ora la chiusura transitiva di $\vdash_N$, denotata $\vdash_N^*$, se w è una stringa ed N un NFA, si ha che

$$w \in L(N) \iff \exists q \in F \| \ (q_0, w) \vdash_N^* (q, \epsilon)$$

Consideriamo adesso l'unione di due NFA, risulta particolarmente semplice da definire, siano N_1 e N_2 due NFA, che per semplicità, condividono l'alfabeto

$$N_1 = \{Q_1, \Sigma_\epsilon, \delta_1, q_1, F_1\}$$

$$N_2 = \{Q_2, \Sigma_\epsilon, \delta_2, q_2, F_2\}$$

Definisco un nuovo NFA $N = (Q, \Sigma_\epsilon, \delta, q_0, F)$ tale che

- $Q = Q_1 \cup Q_2$
- $F = F_1 \cup F_2$
- Siano $q \in Q$ e $a \in \Sigma_\epsilon$:

$$\delta(q, a) = \begin{cases} \delta_1(q, a) \text{ se } q \in Q_1 \\ \delta_2(q, a) \text{ se } q \in Q_2 \\ \{q_1, q_2\} \text{ se } q = q_0 \ \wedge \ a = \epsilon \\ \emptyset \text{ se } q = q_0 \ \wedge \ a \neq \epsilon \end{cases}$$

Si avrà che $L(N) = L(N_1) \cup L(N_2)$

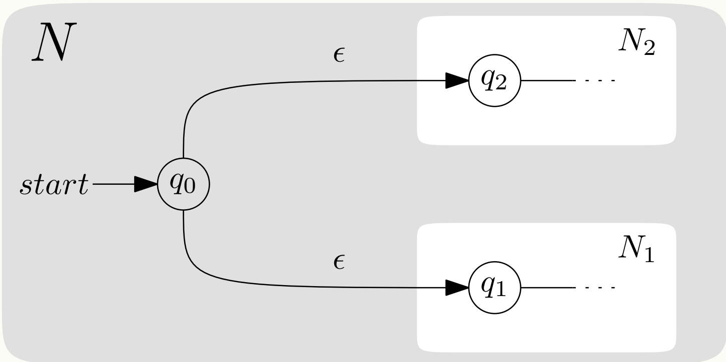


Figura 1.8: Unione di due NFA