Espressioni regolari

Un FA (i.e. automa a stati finiti) riconosce dei linguaggi detti regolari.

Vi sono varie operazioni:

(i) UNIONE: $L \cup W$

(ii) CONCATENAZIONE: $L . W$

(iii) POTENZA: L^k

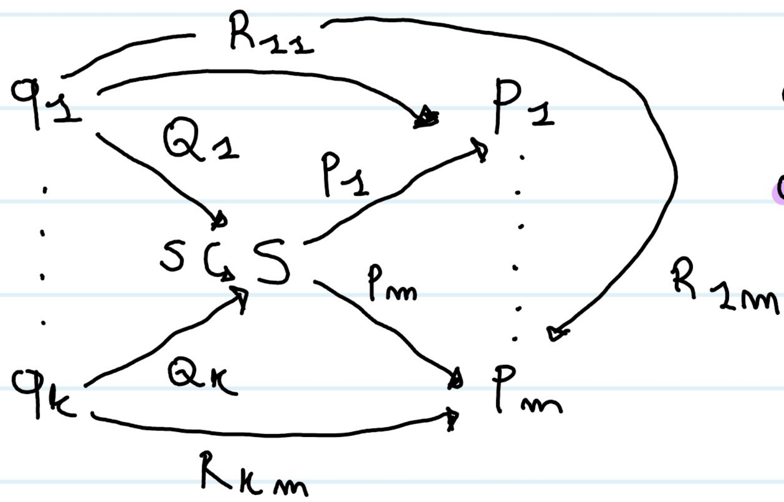
(iv) CHIUSURA DI KLENE: $L^* = \bigcup_{}^{\infty} L^i$

(v) REVERSE: L^R

Definizione induttiva

- base: ε e $\emptyset$ sono espressioni regolari, così come $a \in \Sigma$ $(L(\varepsilon) = \{\varepsilon\}, L(\emptyset) = \emptyset, L(a) = \{a\})$
- induzione: E, F espr. reg. $\Rightarrow$ $E+F$ è un espr. reg. $(L(E+F) = L(E) \cup L(F))$, EF è un espr. reg. $(L(EF) = L(E) . L(F))$, E^* è un espr. reg. $(L(E^*) = L(E)^*)$.

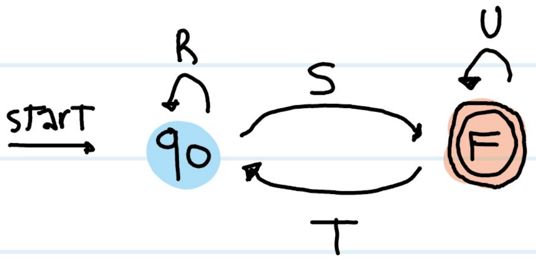
Tecnica di eliminazione degli stati



$q_1 \to p_1 : R_{11} + Q_1 S^* P_1$

$q_i \to q_j : R_{ij} + Q_i S^* P_j$

Si otterranno automi del tipo:



Leggi algebriche per i linguaggi

- $(E+S)+T = E+(S+T)$ (associatività)
- $E(S+T) = ES+ET$ (distributività)
- $E+S = S+E$ (commutatività)
- $(E^*)^* = E^*$ (nilpotenza)
- $L^+ = LL^*$, $L^* = L^+ \cup \{e\}$

Pumping lemma

Il pumping lemma è un teorema sui linguaggi regolari che illustra una proprietà fondamentale di tali linguaggi, e che permette, in innumerevoli casi, di dimostrare che taluni linguaggi non sono invece regolari.

Enunciato

Dato L linguaggio regolare, allora $\exists\, m \mid \forall\, w \in L$ t.c. $|w| \geq m$, $w = xyz$ con

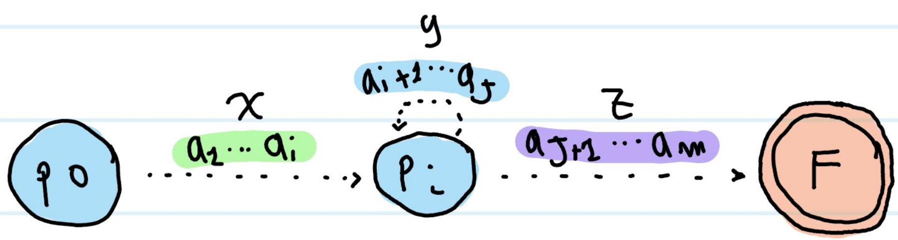
- $y \neq \varepsilon$
- $|xy| \leq m$
- $xy^kz \in L \quad \forall\, k \geq 0$

Dimostrazione

Dal momento che ogni linguaggio regolare L è accettato da almeno un DFA, chiamiamo D uno di questi. Supponiamo D abbia n stati.

Considero una generica stringa $w \in L$ t.c. $|w| \geq m$, ossia $a_1 a_2 \cdots a_m$, con $m \geq n$.

Poiché la stringa viene passata più di m-1 volte a D, per il principio dei cassetti esistono due stati su cui il cammino passa due volte.
Pertanto D è di questa forma:



- $y \neq \varepsilon$ (infatti il cammino deve passare due volte da p_1).
- x e z possono essere invece vuote.
- $|xy| \leq m$, perché $|a_1 \cdots a_{j-1}| < m$ affinché i loro stati siano distinti
- xy^kz è sempre accettata. $\boxtimes$

es. $0^n 1^n$ non è un linguaggio regolare

Se lo fosse, si supponga n sia l'indice del pumping lemma. Allora $0^n 1^n = xyz$:

(i) $y = 0^i 1^m$, allora xy^kz chiaramente non è accettato

(ii) $y = 0^i$, allora xy^kz aumenta il numero di zeri, e quindi non è accettato.

(iii) $y = 1^i$, allora xy^kz aumenta il numero di uno, e quindi non è accettato.

Non soddisfa il pumping lemma, quindi non è regolare. $\boxtimes$

es. 0^{n^2} non è un linguaggio regolare

Allora sia m l'indice del pumping lemma e quindi $0^{m^2} = xyz$ con $y \neq \varepsilon$.

Siano $x = 0^i$, $y = 0^j$ e $z = 0^k$, con $i+j+k = m^2$, $i+j < m$ e $j \geq 1$.

Per il lemma, xy^2z deve essere accettata, quindi $i+2j+k$ deve essere un quadrato m^2. Tuttavia $m^2 - m^2 = j \leq m < 2m+1$, la minima differenza tra due quadrati, $\lightning$. $\boxtimes$

Chiusura dei linguaggi rispetto alle operazioni

(i) rispetto all'unione, infatti siano E, F t.c. $L = L(E) \wedge M = L(F)$, allora $L(E+F) = L \cup M$

(ii) rispetto al complementare, sia infatti D il DFA che riconosce L, allora D con stati finali e non finali invertiti riconosce $\overline{L}$

(iii) rispetto all'intersezione: per De Morgan $L \cap M = \overline{\overline{L} \cup \overline{M}}$, che per (i) e (ii) e' regolare.

(iv) rispetto alla differenza: $L \setminus M = L \cap \overline{M}$, che per (ii) e (iii) è regolare.

Costruzione dell'automa di intersezione

Siano $A = (Q_1, \Sigma, \delta_A, q_{i_A}, F_A)$ e $B = (Q_2, \Sigma, \delta_B, q_{i_B}, F_B)$ con $L(A) = L$ e $L(B) = M$. L'automa che riconosce $L \cap M$ e' il seguente:

$$\begin{cases} A \cap M = (Q_1 \times Q_2, \Sigma, \delta_{A \cap B}, (q_{i_A}, q_{i_B}), F_A \times F_B) \\ \delta_{A \cap B}((q_1, q_2), a) = (\delta_A(q_1, a), \delta_B(q_2, a)) \end{cases}$$

Linguaggio "reverse"

Il linguaggio ottenuto ponendo al contrario ogni sua stringa si dice "reverse", e si indica con L^R. Anch'esso si dimostra essere regolare. Infatti detto $D = (Q, \Sigma, \delta, q_0, F)$ t.c. $L(D) = L$, allora

$$D^R = (Q, \Sigma, \delta^{-1*}, F, q_0) \text{ riconosce } L^R$$

* δ^{-1*} inverte tutti gli archi di D

Si può anche dimostrare induttivamente:

base: $E \in \{\emptyset, \varepsilon, a \in \Sigma\} \Rightarrow E^R = E$

passo induttivo:

$$\left.\begin{array}{ll} (i) & E = F + G \Rightarrow E^R = F^R + G^R \\ (ii) & E = FG \Rightarrow E^R = G^R F^R \\ (iii) & E = F^* \Rightarrow E^R = (F^R)^* \end{array}\right\} \text{tutti e tre sono regolari}$$