Espressione regolare

dato un alfabeto Σ e un insieme di simboli $\{+, \cdot, (,), \phi\}$ tale che valga:

- $r = \phi$
- $r \in \Sigma$
- $r = (s+t)$ opp. $r = (s \cdot t)$ opp. $r = s^*$ s, t espressioni regolari

E. R	linguaggi
ϕ	Λ
a	$\{a\}$
$(s+t)$	$L(s) \cup L(t)$
$(s \cdot t)$	$L(s) \circ L(t)$
s^*	$(L(s))^*$

$Z(r)$

$(a+(b \cdot (c \cdot d)))$
$=$
$a + bcd$

$r^3 = rrr$

Esempio

Se abbiamo l'espressione regolare $(a+b)^*a$ questa rappresenta il linguaggio

$$\mathcal{L}(((a+b)^*a)) = \mathcal{L}((a+b)^*) \circ \mathcal{L}(a) = (\mathcal{L}(a+b))^* \circ \mathcal{L}(a) = (\mathcal{L}(a) \cup \mathcal{L}(b))^* \circ \mathcal{L}(a) = (\{a\} \cup \{b\})^* \circ \{a\} = (\{a,b\})^* \circ \{a\} = \{a,b\}^* \circ \{a\} = \{x \mid x \in \{a,b\}^+, x$$ termina con $a\}$

Determinare l'espressione regolare che sull'alfabeto $\{a, b\}$ definisca l'insieme delle stringhe il cui terzultimo/penultimo carattere è b

$\{a,b\}^* \circ \{b\} \circ \{a,b\} \rightarrow$ penultimo

$\underbrace{\{a,b\}^*}_{\text{generica stringa}} \circ \underbrace{\{b\}}_{\text{terzultimo}} \circ \underbrace{\{a,b\}^2}_{\text{una stringa da 2 caratteri}} \rightarrow$ terzultimo

2° Carattere

$\{a, b\} \circ \{b\} \circ \{a, b\}^*$

$C = \{0, 1, \dots 9\}$
↑
ESP. REGOLARE

$(1+2+\dots+9)((0+1+\dots+9)^* + 0) . (0+1+\dots+9)^+$

questa espressione si può usare per creare qualsiasi numero

Scrivere il linguaggio definito dall'espressione regolare

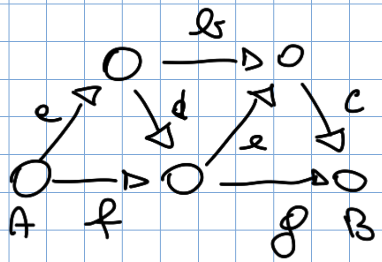
$a^*((aa)^* b + (bb)^* a) b^* = \mathcal{L}(a^*((aa)^* b + (bb)^* a) b^*)$

$= \mathcal{L}(\{a\}^* \circ (\{aa\}^* \circ \{b\} + \{bb\}^* \circ \{a\}) \circ \{b\}^*)$

$= \mathcal{L}(\{a\}^* \circ \{\{aa\}^* \circ \{b\} \cup \{bb\}^* \circ \{a\}\} \circ \{b\}^*)$

$= \{ x \mid x \in \{aa\}^* \circ \{b\} \cup \{bb\}^* \circ \{a\}\} \mid x$ inizia con a e termina con b $\}$

Descrivere una mappa stradale (con tutti i tratti a senso unico e contrassegnati dai caratteri dell'alfabeto) Fornire un'espressione regolare che colleghi percorsi da A a B (tutti)



$(ad + f) \circ (ebd)^* \circ (ec + g)$
↑
"loop"

Dato un alfabeto $\Sigma = \{a_1 \dots a_n\}$ si definisce ordinamento lessicografico delle stringhe di Σ^* l'ordinamento ($<$) ottenuto stabilendo un (qualunque) ordinamento tra i caratteri di Σ e definendo l'ordinamento di due stringhe $x, y \in \Sigma^*$ in modo tale che $x < y \iff$ una delle seguenti condizioni è verificata

1) $|x| < |y|$
2) $|x| = |y|$ ed esiste $z \in \Sigma^*$ tale che $x = z a_i u$ e $y = z a_j v$ con $u, v \in \Sigma^*$ e $i < j$

Esempio

$\Sigma = \{a, b\}$ dove $a < b$ le stringhe di Σ^* sono enumerate nel seguente modo:

$\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, aba, baa$

Def Linguaggio regolare

Un linguaggio si dice regolare se esiste un automa a stati finiti ASF che lo riconosca

---

Proprietà dei linguaggi regolari:

teorema: Dati 2 linguaggi L_1 e L_2 la loro unione $L_1 \cup L_2$ è un linguaggio regolare

Dimostrazione: Siano dati 2 automi deterministici

$A_1 = < \Sigma_1, Q_1, \delta N_1, q_{0_1}, F_1 >$

$A_2 = < \Sigma^2, Q_2, \delta N_2, q_{0_2}, F_2 >$

che accettano i linguaggi

$L_1 = L(A_1)$ e $L_2 = L(A_2)$

Vogliamo dimostrare che partendo da L_1 e L_2 e quindi da A_1 e A_2, vogliamo dimostrare che partendo da L_1 e L_2 e quindi di A_1 e A_2 costruiamo l'automa che riconosca $L(A_1) \cup L(A_2)$

$L = (A_1) \cup L(A_2)$

$A = < \Sigma, Q, \delta N, q_0, F >$

$\Sigma = \Sigma_1 \cup \Sigma_2$

$Q = Q_1 \cup Q_2 \cup \{q_0\}$

$F_2 = F_1 \cup F_2$ oppure $\underbrace{F = F_1 \cup F_2 \cup \{q_0\}}$,

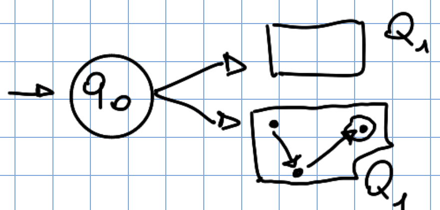
caso in cui gli automi riconoscono anche ε

$\delta_N(q, a) = \delta_{N_1}(q, a)$ se $q \in Q_1$ e $a \in \Sigma_1$

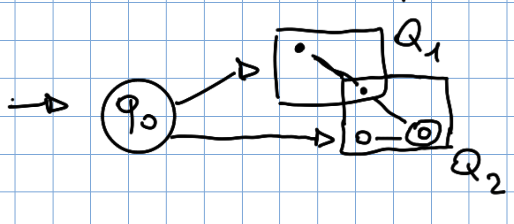
$\delta_N(q, a) = \delta_{N_2}(q, a)$ se $q \in Q_2$ e $a \in \Sigma_2$

$\delta_N(q_0, a) = \delta_{N_1}(q_0, a) \cup \delta_{N_2}(q_0, a)$ $a \in \Sigma = \Sigma_1 \cup \Sigma_2$

il fatto che A_1 e A_2 sono deterministici non implica che A sia deterministico



in questa configurazione stato finale non in comune



in questa configurazione abbiamo uno stato in comune teorema

Dato un linguaggio regolare L il suo complemento $\overline{L}$ è un linguaggio regolare

$\overline{L} = \Sigma^* \setminus \{L\}$

Dimostrazione

$A = < \Sigma, Q, \delta, q_0, F >$

↳ Definire un linguaggio riconosciuto $L = L(A)$ e possiamo costruire l'automa $\bar{A} = < \Sigma, Q, \delta, q_0, \{Q-F\} >$ che riconosce rispettivamente il linguaggio $\overline{L(A)} = \overline{L}$

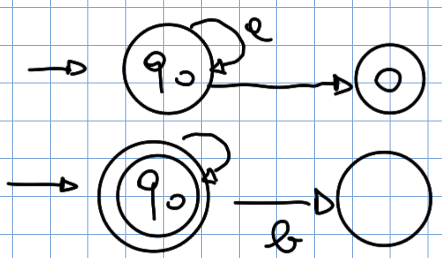
ogni stringa che porta l'automa A in uno stato finale F porta l'automa $\bar{A}$ in uno stato non finale, e viceversa

$L(\bar{A}) = \Sigma^* \setminus L(A) = \overline{L(A)}$

Dire il complemento di un automa è come dire il complemento del linguaggio

A :

$\bar{A}$:

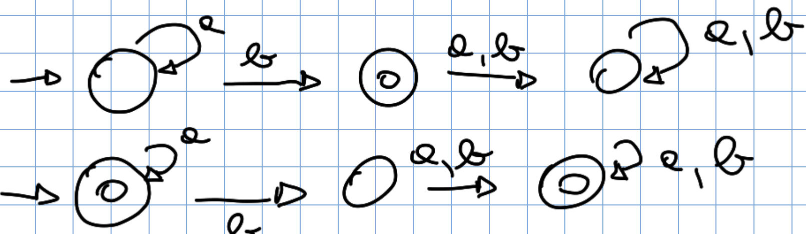


$aab \in L_1$ $ba \notin L_1$

$aab \notin L_2$ $ba \notin L_2$

A_1 :

A_2 :



Pag 95

Teorema

Dati 2 linguaggi regolari L_1 e L_2 la loro intersezione $L = L_1 \cap L_2$ è un linguaggio regolare

Dimostrazione

$L = L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$

L_1 è regolare $\Rightarrow \overline{L_1}$ è regolare
L_2 è " $\Rightarrow \overline{L_2}$ è regolare
$\Rightarrow \overline{L_1} \cup \overline{L_2}$ è regolare

$\Rightarrow \overline{L_1 \cup L_2}$ è regolare

Teorema

Dati 2 linguaggi regolari L_1 e L_2 la loro concatenazione $L = L_1 \circ L_2$ è un linguaggio regolare

Dimostrazione

ASFD $A_1 = < \Sigma_1, Q_1, \delta_1, q_{0_1}, F_1 >$
$A_2 = < \Sigma_2, Q_2, \delta_2, q_{0_2}, F_2 >$

riconoscono rispettivamente i linguaggi $L_1 = \mathcal{L}(A_1)$ e $L_2 = \mathcal{L}(A_2)$

Sia $A = < \Sigma, Q, \delta, q_0, F >$ non deterministico:

- $\Sigma = \Sigma_1 \cup \Sigma_2$
- $Q = Q_1 \cup Q_2$
- $F = \begin{cases} F_2 & \text{se } \varepsilon \notin \mathcal{L}(A_2) \\ F_1 \cup F_2 & \text{altrimenti} \end{cases}$
- $q_0 = q_{0_1}$
- $\delta_n(q, e) = \delta_1(q, e) \ \forall q \in Q_1 - F_1 \quad e \in \Sigma_1$
- $\delta_n(q, e) = \delta_2(q, e) \ \forall q \in Q_2 \quad e \in \Sigma_2$
- $\delta_n(q, e) = \delta_1(q, e) \cup \delta_2(q_{0_2}, e) \ \forall q \in Q_1 \quad e \in \Sigma = \Sigma_1 \cup \Sigma_2$

teorema

Dato un linguaggio regolare L, anche L^* è un linguaggio regolare

Dimostrazione

ASFD $A = < \Sigma, Q, \delta, q_0, F >$ che riconosce $\mathcal{L}(A) = L$

e partire da questo costruiamo $A' = < \Sigma, Q \cup \{q_0'\}, \delta', q_0', F \cup \{q_0'\} >$ che riconosce $L^* = (L(A))^*$ ponendo

$$\delta'(q, e) = \delta(q, e) \quad \forall q \in Q - F \quad e \in \Sigma$$

$$\delta'(q, e) = \delta(q, e) \cup \delta(q_0, e) \quad \forall q \in F \quad e \in \Sigma$$

$$\delta'(q, e) = \delta(q_e, e) \quad e \in \Sigma$$

il secondo automa A' riconosce stringhe in più rispetto al primo