Proprieta' dei Linguaggi Regolari

- **Proprieta' di chiusura.** E' interessante investigare proprieta' di chiusura dei linguaggi regolari rispetto ad applicazione di *sostituzioni o omomorfismi*. Queste tecniche permettono di tradurre un linguaggio in un altro.
- **Proprieta' di decisione.** Ci interessa considerare vari problemi di decisione, comprendere quali problemi sono decidibili e quale e' il costo computazionale.

Sostituzioni

Siano Σ, Δ degli alfabeti. Una *sostituzione* e' una funzione

$$f : \Sigma \to 2^{\Delta^*}$$

La sostituzione associa, ad ogni simbolo di Σ, un insieme di stringhe di Δ, ovvero un linguaggio su Δ.
Una sostituzione $f : \Sigma \to 2^{\Delta^*}$ si applica a stringhe $w \in \Sigma^*$ in modo induttivo

$$f(\epsilon) = \epsilon$$

$$f(xa) = f(x).f(a)$$

e a linguaggi $L \subseteq \Sigma^*$,

$$f(L) = \{f(w) \mid w \in L\}$$

Esempio: traduzione

Siano $\Sigma = \{0, 1\}$ e $\Delta = \{a, b\}$ e sia $f : \Sigma \to 2^{\Delta^*}$ la sostituzione,

$$f(0) = a \quad f(1) = b^*$$

La sostituzione traduce gli 0 nel linguaggio $\{\mathrm{a}\}$ e traduce gli 1 nel linguaggio denotato dall'espressione regolare b^*.
Abbiamo

$$f(0) = a, f(1) = b^*, f(01) = ab^*$$

Consideriamo il linguaggio L denotato dall'espressione regolare $0^*(0 + 1)1^*$. L'immagine di L tramite f e' $f(L)$

$$a^*(a+b^*)(b^*)^* = a^*(a+b^*)b^* = (a^*a+a^*b^*)b^* = (a^*ab^*+a^*b^*b^*) = a^*b^*$$

Chiusura rispetto alla sostituzione

Teorema: Sia L un linguaggio regolare su Σ e $f : \Sigma \rightarrow 2^{\Delta^*}$ una sostituzione tale che, per ogni $a \in \Sigma$, $f(a)$ e' un linguaggio regolare. Allora $f(L)$ e' regolare.

Prova L'idea e' di usare le espressioni regolari come nell'esempio precedente.
Se L e' regolare allora esiste E tale che $L(E) = L$. Inoltre, per ogni carattere a $f(a)$ puo' essere descritta da un espressione regolare E_a tale che $L(E_a) = f(a)$.
Sia $f(E)$ l'espressione regolare ottenuta sostituendo ogni $a \in \Sigma$ con E_a corrispondente ad $f(a)$. Abbiamo $L(f(E)) = f(L(E)) = f(L)$.
Per induzione strutturale su E.

Chiusura rispetto alla sostituzione

Base Se E e' ϵ o $\emptyset$, allora la sostituzione non ha effetto. Quindi, $f(E) = E = f(L) = L$.
Se $E = a$, allora $f(E) = E_a$ e
$L(f(E_a)) = f(L(a)) = f(L(E)) = f(\{a\})$.

Induzione: Ci sono tre casi da considerare. Facciamo vedere il caso di $E = G + F$, gli altri sono simili. Abbiamo $f(E) = f(G) + f(F)$ per definizione di applicazione della sostituzione alle espressioni regolari.
Per ipotesi induttiva, $L(f(G)) = f(L(G))$ e $L(f(F)) = f(L(F))$.
Quindi,

$$L(f(G)+f(F)) = L(f(G)) \cup L(f(F)) = f(L(G)) \cup f(L(F)) = f(L(G) \cup L(F)).$$

Omomorfismo

Siano Σ, Δ degli alfabeti. Un *omomorfismo* e' una funzione

$$h : \Sigma \rightarrow \Delta^*$$

- Un omomorfismo e' una sostituzione che associa, ad ogni simbolo di Σ, esattamente una stringa Δ^* (invece che un linguaggio)
- Dato che una omomorfismo e' un tipo particolare di sostituzione, i linguaggi regolari sono chiusi per omomorfismo

Esempio

Siano $\Sigma = \{0, 1\}$ e $\Delta = \{a, b\}$ e sia $h : \Sigma \to \Delta^*$ l'omomorfismo,

$$h(0) = ab \quad h(1) = \epsilon$$

La sostituzione traduce gli 0 nella stringa *ab* e traduce gli 1 nella stringa vuota (quindi li cancella).
Per esempio

$$h(0110) = abab, h(11101) = ab$$

Consideriamo il linguaggio L denotato dall'espressione regolare 10^*1. Sono le stringhe di bit che hanno un numero arbitario di 0 seguito e preceduto da esattamente un 1.
Ovviamente, $h(L) = (ab)^*$.

Omomorfismo Inverso

La proprieta' interessante degli omomorfismi e' che possono essere invertiti.
Siano Σ, Δ alfabeti e $h : \Sigma \to \Delta^*$ un omomorfismo. Dato un linguaggio $L \subseteq \Delta^*$, definamo la **controimmagine** di L rispetto ad h come

$$h^{-1}(L) = \{w \mid h(w) \in L\}$$

Ovviamente, $h^{-1}(L) \subseteq \Sigma^*$ e' insieme delle stringhe che sono immagine di qualche stringa di L tramite h.

Esempio

Siano $\Sigma = \{0, 1\}$ e $\Delta = \{a, b\}$ e sia $h : \Sigma \rightarrow \Delta^*$ l'omomorfismo,

$$h(0) = aa \quad h(1) = aba$$

La sostituzione traduce gli 0 nella stringa *aa* e traduce gli 1 nella stringa *aba*. Per esempio, $h(0110) = aaabaabaaa$.
Consideriamo il linguaggio L_1 denotato dall'espressione regolare $(01)^*$. Sono le stringhe di bit che hanno un numero pari 0 e 1 alternati.
Ovviamente, $h(L_1) = (aaaba)^*$.

Esempio

Consideriamo il linguaggio L_2 denotato dall'espressione regolare $(ab + ba)^* a$.
Chi e' $h^{-1}(L_2)$? Ricordiamo che
$h^{-1}(L_2) = \{w \in \{0,1\}^* \mid h(w) \in L_2\}$.
Abbiamo $h^{-1}(L_2) = \{1\}$. E' ovvio che $h(1) = aba \in L_2$.
Bisogna pero' fare vedere che per ogni altra $w \in \{0,1\}^*$ con $w \neq 1$, $h(w) \notin L_2$.

Esempio

Per casi su w:

- se e' vuota $w = \epsilon$, allora $h(w) = \epsilon \notin L_2$;
- se inizia per 0, $w = 0y$ per qualche y. Allora $h(0y) = aah(y) \notin L_2$, dato che le stringhe iniziano con ab o ba o con a.
- se inizia per 1, $w = 1y$ per qualche $y \neq \epsilon$. Allora $h(1y) = abah(y)$. Data la forma di L_2 deve essere $h(y) = \epsilon$ ma $y \neq \epsilon$. Oppure la stringa e' per forza della forma iniziale $abab$, quindi $h(y)$ inizia per b. Tuttavia, non esiste alcun y tale che $h(y)$ inizia per b dato che $h(0) = aa$ e $h(1) = aba$.

In conclusione

Abbiamo

- $h^{-1}(L_2) = \{1\}$;
- $h(h^{-1}(L_2)) = h(\{1\}) = \{aba\} \subset L_2$.

In generale

$$h(h^{-1}(L)) \subseteq L \quad h^{-1}(h(L)) \subseteq L$$

Chiusura rispetto all'omomorfismo inverso

Teorema: Sia L un linguaggio regolare su Δ e $h : \Sigma \rightarrow \Delta^*$ un omomorfismo. Allora $h^{-1}(L)$ e' regolare.

Prova L'idea e' di usare l'automa DFA che riconosce L. Sia

$$A = (Q, \Delta, \delta, q_0, F)$$

tale che $L(A) = L$.
Costruiamo un automa B sull'alfabeto Σ che riconosce $h^{-1}(L)$.

L'automa B

Sia

$$B = (Q, \Sigma, \gamma, q_0, F)$$

- gli stessi stati di A, lo stesso stato iniziale e gli stessi stati finali
- alfabeto Σ
- $\gamma : Q \times \Sigma \to Q$ e' definita in base a $\delta : Q \times \Delta \to Q$. In particolare, $a \in \Sigma$,

$$\gamma(q, a) = \hat{\delta}(q, h(a))$$

L'automa B nello stato q, leggendo a, si muove nello stato in cui si muoverebbe A, leggendo la stringa $h(a)$.

Proprieta' della Costruzione

Per induzione sulla lunghezza di $w \in \Sigma^*$, si puo' provare che

$$\hat{\gamma}(q_0, w) = \hat{\delta}(q_0, h(w))$$

Facciamo vedere che $L(B) = h^{-1}(L(A))$. Osserviamo che $w \in L(B)$ sse $\hat{\gamma}(q_0, w) \in F$. Inoltre $\hat{\delta}(q_0, h(w)) \in F$, ovvero $h(w) \in L(A)$ e $w \in h^{-1}(L(A))$.
Viceversa e' analogo.