Linguaggi Formali e Traduttori

2.6 Pumping lemma per i linguaggi regolari

Sommario

Per dimostrare che un linguaggio **è regolare**, basta esibire un automa a stati finiti (DFA, NFA o **ε**-NFA) che lo riconosce, oppure una espressione regolare che lo genera. L'incapacità di trovare siffatto automa o siffatta espressione non è una dimostrazione del fatto che il linguaggio non è regolare.

In questa lezione rispondiamo alle seguenti domande:

1. Esistono linguaggi **non** regolari?
2. Se sì, come dimostro che un linguaggio **non** è regolare?

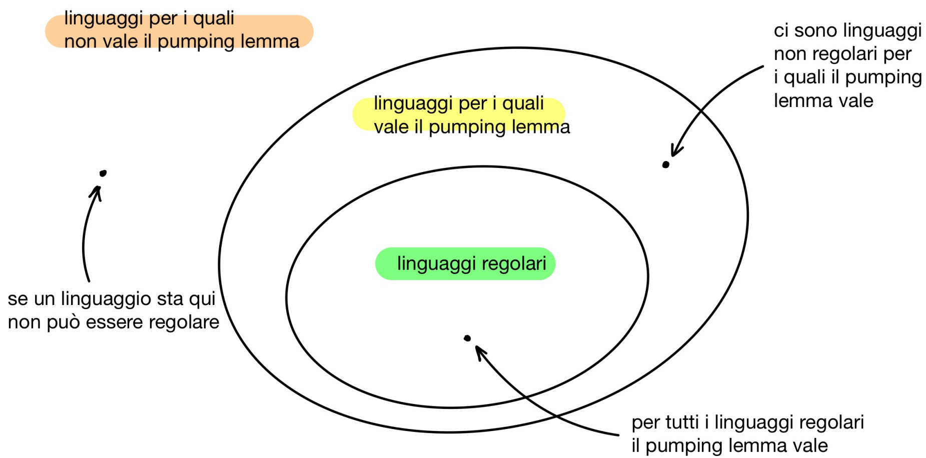
Linguaggi **non** regolari

- Cerchiamo una proprietà P soddisfatta da tutti i linguaggi regolari:

$$L \textbf{ regolare} \Rightarrow L \textbf{ soddisfa } P$$

- Se troviamo un linguaggio L che **non soddisfa** P, allora per **contrapposizione** possiamo concludere che L **non è regolare**:

$$L \textbf{ non soddisfa } P \Rightarrow L \textbf{ non è regolare}$$



Pumping lemma per linguaggi regolari

Teorema

Per ogni linguaggio regolare L esiste $n \in \mathbb{N}$ tale che, per ogni $w \in L$ con $|w| \geq n$, esistono x, y e z tali che $w = xyz$ e inoltre:

1. $y \neq \varepsilon$
2. $|xy| \leq n$
3. $xy^k z \in L$ per ogni $k \geq 0$.

In prosa

- Ogni stringa w "sufficientemente lunga" ($|w| \geq n$) di un linguaggio regolare L ...
- ... contiene una sottostringa non vuota ($y \neq \varepsilon$) ...
- ... e "non troppo distante" dall'inizio di w ($|xy| \leq n$) ...
- ... che può essere eliminata ($k = 0$) o replicata a piacere ($k > 0$) ...
- ... consentendoci di trovare altre stringhe di L ($xy^k z \in L$)

Esempio: $a^k b^k$ non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Esempio: $a^k b^k$ non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Esempio: $a^k b^k$ non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Esempio: $a^k b^k$ non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Esempio: $a^k b^k$ non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Dalla condizione 2 sappiamo che x e y sono composte di sole a.

Esempio: a^k^b^k non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Dalla condizione 2 sappiamo che x e y sono composte di sole a.

Dalla condizione 1 sappiamo che y contiene almeno una a.

Esempio: a^{k}b^k non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Dalla condizione 2 sappiamo che x e y sono composte di sole a.

Dalla condizione 1 sappiamo che y contiene almeno una a.

Dalla condizione 3 sappiamo che $xz \in L$.

Esempio: a^{k}b^k non è regolare

Dimostriamo che $L = \{a^k b^k \mid k \geq 0\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Dalla condizione 2 sappiamo che x e y sono composte di sole a.

Dalla condizione 1 sappiamo che y contiene almeno una a.

Dalla condizione 3 sappiamo che $xz \in L$.

Ma ora in xz ci sono più b che a, il che contraddice la definizione di L.

Pumping lemma: dimostrazione (1/3)

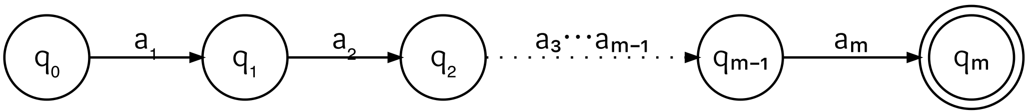
Sia L un linguaggio regolare.

Dunque esiste un DFA $A = (Q, \Sigma, \delta, q_0, F)$ tale che $L = L(A)$.

Poniamo $n = |Q|$, ovvero n è il numero di stati di un DFA che riconosce L.

Prendiamo $w \in L$ tale che $|w| \geq n$. Deve essere $w = a_1 a_2 \cdots a_m$ con $m \geq n$.

Se rappresentiamo il cammino fatto da A per riconoscere w come segue

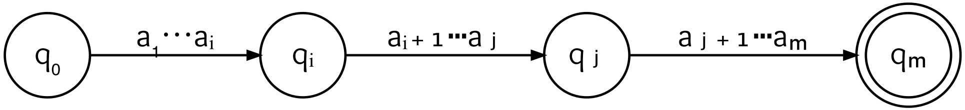


notiamo che questo cammino passa attraverso $m + 1$ stati.

Siccome $m \geq n$ abbiamo $m + 1 > n$. Ovvero, gli stati attraversati non possono essere tutti distinti, perché l'automa ne ha solo n.

Pumping lemma: dimostrazione (2/3)

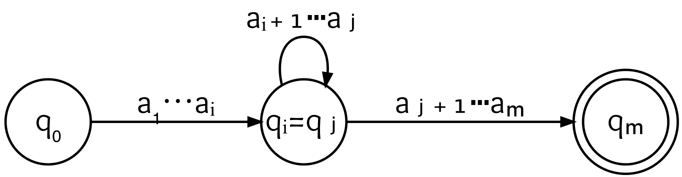
Deduciamo che il cammino fatto da A può essere rappresentato così



dove $q_i = q_j$ e $i < j$.

Possiamo supporre, senza perdere in generalità, che $q_i = q_j$ sia il **primo** stato che si ripete in questo cammino (ce ne possono essere tanti).

Il cammino fatto da A può allora essere rappresentato anche così:



Pumping lemma: dimostrazione (3/3)

Ora definiamo le stringhe x, y e z come segue:

- $x = a_1 a_2 \cdots a_i$
- $y = a_{i+1} a_{i+2} \cdots a_j$
- $z = a_{j+1} a_{j+2} \cdots a_m$

Notiamo che

1. $y \neq \varepsilon$, in quanto $i < j$ dunque in y c'è almeno un simbolo (l'automa è deterministico e non ha ε-transizioni)
2. $|xy| \leq n$ in quanto $q_i = q_j$ è il primo stato che si ripete e quindi gli stati da q_0 a q_j sono al massimo $n + 1$, attraversati leggendo al massimo n simboli di w
3. $xy^k z \in L$ per ogni $k \geq 0$ in quanto tutti i cammini etichettati con $xy^k z$ portano l'automa da q_0 (lo stato inziale) a q_m (uno stato finale)

e questo conclude la dimostrazione.

Esempio: $a^k b^m$ con $k \leq m$ non è regolare

Dimostriamo che $L = \{a^k b^m \mid 0 \leq k \leq m\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Considero la stringa $w = a^n b^n$, che è in L e ha la proprietà $|w| = 2n \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Dalla condizione 2 sappiamo che x e y sono composte di sole a.

Dalla condizione 1 sappiamo che y contiene almeno una a.

Dalla condizione 3 sappiamo che $xyyz \in L$.

Ma ora in $xyyz$ ci sono più a che b, il che contraddice la definizione di L.

Esempio: a^k con k primo non è regolare

Dimostriamo che $L = \{a^k \mid k \textbf{ primo}\}$ non è regolare facendo vedere che per L il pumping lemma non vale.

Supponiamo, per assurdo, che esista n con le proprietà enunciate nella slide 4.

Consideriamo la stringa $w = a^p$ dove p è un numero primo $p \geq n + 2$. Siamo sempre in grado di trovare p con questa proprietà in quanto esistono infiniti numeri primi. Inoltre, la stringa w è in L e ha la proprietà $|w| \geq n$.

Devono esistere x, y e z tali che $w = xyz$ e che soddisfano le condizioni 1–3 della slide 4.

Definiamo $m = |y|$, da cui segue che $|xz| = p - m$. Dalle condizioni 1 e 2 sappiamo che $1 \leq m \leq n$. Dalla condizione 3 sappiamo che $xy^{p-m}z \in L$. Tuttavia

$$|xy^{p-m}z| = |xz| + (p - m)|y| = p - m + (p - m)m = (p - m)(m + 1)$$

e ora concludiamo che $|xy^{p-m}z|$ **non è primo** in quanto:

- da $1 \leq m$ deduco $2 \leq m + 1$, e
- da $m \leq n$ e $p \geq n + 2$ deduco $p - m \geq 2$.

Esercizi e quesiti

Linguaggi non regolari

Dimostrare che i seguenti linguaggi non sono regolari:

1. $\{0^{k^2} \mid k \geq 1\}$
2. $\{0^k10^k \mid k \geq 1\}$
3. $\{0^k1^{2k} \mid k \geq 1\}$
4. $\{ww \mid w \in \{0,1\}^*\}$
5. $\{ww^R \mid w \in \{0,1\}^*\}$
6. $\{0^k1^m \mid k \neq m\}$ (suggerimento: usare una proprietà di chiusura)

Proprietà di linguaggi

1. Se L_i con $i \in \mathbb{N}$ è una famiglia **infinita** di linguaggi regolari, cosa si può dire di $\bigcup_{i\in\mathbb{N}} L_i$? È sempre un linguaggio regolare?
2. Se L è un linguaggio **finito**, mostrare come scegliere la costante n del pumping lemma senza sapere come è fatto l'automa che riconosce L.