

Homework 3 - Linguaggi non regolari, Pumping Lemma, Chiusura linguaggi regolari

Gabriel Rovesti

1. Dimostrate che i seguenti linguaggi non sono regolari:

- $\{0^n 1^m 0^n \mid m, n \geq 0\}$

Per dimostrare che il linguaggio non è regolare, possiamo utilizzare il pumping lemma per i linguaggi regolari.

Il pumping lemma afferma che per ogni linguaggio regolare L , esiste un numero intero positivo p (chiamato "pumping length") tale che per ogni stringa w appartenente a L con lunghezza maggiore o uguale a p , w può essere suddivisa in tre parti x, y e z ($w = xyz$) che soddisfano le seguenti condizioni:

- (a) y non è una stringa vuota
- (b) $|xy| \leq p$ (la lunghezza di xy è minore o uguale a p)
- (c) per ogni $i \geq 0$, la stringa $xy^i z$ appartiene a L .

Supponiamo per assurdo che questo linguaggio sia regolare. Allora, per il pumping lemma, esiste una pumping length p tale che ogni stringa w appartenente al linguaggio, con lunghezza maggiore o uguale a p , può essere suddivisa in x, y e z ($w = xyz$) soddisfacendo le condizioni.

Consideriamo la stringa $w = 0^p 1^p 0^p$, che appartiene al linguaggio. Poiché la lunghezza di w è $3p$, possiamo suddividere w in x, y e z ($w = xyz$) in modo che soddisfi le condizioni del pumping lemma. Poiché $|xy| \leq p$ (condizione 2), y può contenere al massimo p caratteri. Di conseguenza, y non può contenere contemporaneamente uno 0 e uno 1, altrimenti $|xy| > p$.

- (a) Caso 1: y contiene solo 0. In questo caso, la stringa $xy^2 z = 0^{p+|y|} 1^p 0^p$ non appartiene al linguaggio, violando la condizione 3 del pumping lemma.

- (b) Caso 2: y contiene solo 1. In questo caso, la stringa $xy^2z = 0^p 1^{p+|y|} 0^p$ non appartiene al linguaggio, violando la condizione 3 del pumping lemma.

In entrambi i casi, abbiamo una violazione delle condizioni del pumping lemma, il che contraddice l'ipotesi che il linguaggio sia regolare.

Pertanto, il linguaggio non è regolare.

- $\{0^m 1^n \mid m \neq n\}$

La parte introduttiva e dimostrativa di questo tipo di esercizi prevede sempre i primi tre paragrafi utilizzati al punto precedente. Per semplicità, li salterò, evitando di riscriverli di volta in volta.

Consideriamo la stringa $w = 0^{p+1} 1^p$, che appartiene a L poiché $(p+1) \neq p$. Poiché la lunghezza di w è $2p+1$, che è maggiore di p , possiamo suddividere w in x, y, z ($w = xyz$) in modo che soddisfi le condizioni del pumping lemma.

Poiché $|xy| \leq p$ (condizione 2), y può contenere al massimo p caratteri. Di conseguenza, y deve contenere solo 0 oppure solo 1, ma non entrambi.

- (a) Caso 1: y contiene solo 0. In questo caso, possiamo scrivere $x = 0^q e y = 0^r$, dove $q \geq 0, r > 0$ e $q + r \leq p$. Quindi, $z = 0^{p+1-q-r} 1^p$. Consideriamo la stringa $xy^2z = 0^{q+2r} 1^p$. Poiché $q + 2r = (p+1) + r$, abbiamo che $(q+2r)p$, violando la condizione 3 del pumping lemma.
- (b) Caso 2: y contiene solo 1. In questo caso, possiamo scrivere $x = 0^{p+1} e y = 1^s$, dove $s > 0$ e $s \leq p$. Quindi, $z = 1^{p-s}$. Consideriamo la stringa $xy^2z = 0^{p+1} 1^{p+s}$. Poiché $p + s \neq p$, abbiamo una violazione della condizione 3 del pumping lemma.

In entrambi i casi, abbiamo una violazione delle condizioni del pumping lemma, il che contraddice l'ipotesi che il linguaggio L sia regolare.

- $\{w \in \{0,1\}^* \mid \exists u \mid www = uu\}$

Consideriamo la stringa $w = 0^p 1^p 0^p$, che appartiene a L poiché esiste $u = 0^p 1^p$ tale che $www = uu$. Poiché la lunghezza di w è $3p$, che è maggiore di p , possiamo suddividere w in x, y e z ($w = xyz$) in modo che soddisfi le condizioni del pumping lemma.

Poiché $|xy| \leq p$ (condizione 2), y può contenere al massimo p caratteri. Di conseguenza, y non può contenere contemporaneamente uno 0 e uno 1, altrimenti $|xy| > p$.

- (a) Caso 1: y contiene solo 0. In questo caso, possiamo scrivere $x = 0^q, y = 0^r$ e $z = 0^{p-q-r}1^p0^p$, dove $q, r > 0$ e $q + r < p$. Consideriamo la stringa $xy^2z = 0^{q+2r}1^p0^p$. Poiché $q + 2r > p$, non esiste alcuna stringa u tale che $xy^2z = uu$, violando la condizione di appartenenza al linguaggio L .
- (b) Caso 2: y contiene solo 1. In questo caso, possiamo scrivere $x = 0^p1^s, y = 1^t$ e $z = 1^{p-s-t}0^p$, dove $s, t > 0$ e $s + t < p$. Consideriamo la stringa $xy^2z = 0^p1^{s+2t}0^p$. Poiché $s + 2t > p$, non esiste alcuna stringa u tale che $xy^2z = uu$, violando la condizione di appartenenza al linguaggio L .

In entrambi i casi, abbiamo una violazione delle condizioni del pumping lemma, il che contraddice l'ipotesi che L sia regolare.

- $\{w \mid w \in \{0, 1\}^* \text{ non è una stringa palindroma}\}$

Consideriamo la stringa $w = 0^p1^p0^p$, che appartiene a L poiché non è una stringa palindroma. Poiché la lunghezza di w è $3p$, che è maggiore di p , possiamo suddividere w in x, y e z in modo che soddisfi il pumping lemma.

Poiché $|xy| \leq p$ (condizione 2), y può contenere al massimo p caratteri. Di conseguenza, y non può contenere contemporaneamente uno 0 e uno 1, altrimenti $|xy| > p$.

- (a) Caso 1: y contiene solo 0. In questo caso, possiamo scrivere $x = 0^q, y = 0^r$ e $z = 0^{p-q-r}1^p0^p$, dove $q, r > 0$ e $q + r < p$. Consideriamo la stringa $xy^0z = xz = 0^p1^p0^p$, che è una stringa palindroma e quindi non appartiene al linguaggio L , violando la condizione 3 del pumping lemma.
- (b) Caso 2: y contiene solo 1. In questo caso, possiamo scrivere $x = 0^p1^s, y = 1^t$ e $z = 1^{p-s-t}0^p$, dove $s \geq 0, t > 0$ e $s + t \leq p$. Consideriamo la stringa $xy^0z = xz = 0^p1^p0^p$, che è una stringa palindroma e quindi non appartiene al linguaggio L , violando la condizione 3 del pumping lemma.

In entrambi i casi, abbiamo una violazione delle condizioni del pumping lemma, il che contraddice l'ipotesi che il linguaggio L sia regolare.

- $\{0^n \mid n \text{ è un cubo perfetto}\}$

Supponiamo per assurdo che il linguaggio L sia regolare. Consideriamo la stringa $w = 0^{p^3}$, che appartiene a L poiché p^3 è un cubo perfetto. Poiché la lunghezza di w è p^3 , che è maggiore di p , possiamo suddividere w in x, y e z in modo che soddisfi le condizioni del pumping lemma.

Poiché $|xy| \leq p$, y può contenere al massimo p caratteri. Di conseguenza, y contiene esattamente q zeri, dove $1 \leq q \leq p$.

Consideriamo la stringa $w' = xy^2z = 0^{p^3-q+2q} = 0^{p^3+q}$.

Vogliamo dimostrare che w' non appartiene a L . Per fare ciò, dobbiamo mostrare che $(p^3 + q)$ non è un cubo perfetto.

Il cubo perfetto successivo a p^3 è $(p+1)^3$, che è dato da:

$$(p+1)^3 = p^3 + 3p^2 + 3p + 1$$

Quindi, per ottenere un cubo perfetto a partire da p^3 , sarebbe necessario incrementare il numero di zeri di $3p^2 + 3p + 1$.

Tuttavia, in w' abbiamo incrementato il numero di zeri solo di q , dove $1 \leq q \leq p$. Poiché q è strettamente minore di $3p^2 + 3p + 1$, ne consegue che $(p^3 + q)$ non può essere un cubo perfetto.

Pertanto, la stringa $w' = 0^{p^3+q}$ non appartiene al linguaggio L , violando la condizione 3 del pumping lemma.

Questa violazione contraddice l'ipotesi che L sia regolare.

- $\{0^p q \mid p, q \in \mathbb{N}, p > 1, q > 1\}$

Consideriamo la stringa $w = 0^{r^2 \cdot (r+1)}$, che appartiene a L poiché possiamo scrivere $w = 0^{pq}$ con $p = r^2$ e $q = r + 1$, che soddisfano le condizioni $p > 1$ e $q > 1$. Poiché la lunghezza di w è $r^2 \cdot (r + 1)$, che è maggiore di r , possiamo suddividere w in x , y e z in modo che soddisfi le condizioni del pumping lemma.

Poiché $|xy| \leq r$ (condizione 2), y può contenere al massimo r zeri. Di conseguenza, y contiene esattamente t zeri, dove $1 \leq t \leq r$.

Consideriamo la stringa $w' = xy^2z = 0^{r^2 \cdot (r+1) - t + 2t} = 0^{r^2 \cdot (r+1) + t}$.

Vogliamo dimostrare che w' non appartiene a L . Per fare ciò, dobbiamo mostrare che $(r^2 \cdot (r + 1) + t)$ non può essere scritto come prodotto di due numeri naturali maggiori di 1.

Supponiamo per assurdo che $(r^2 \cdot (r + 1) + t)$ possa essere scritto come prodotto di due numeri naturali maggiori di 1, diciamo p' e q' . Allora:

$$r^2 \cdot (r + 1) + t = p' \cdot q'$$

Poiché $r^2 \cdot (r + 1)$ è divisibile per $r + 1$, dobbiamo avere che $q' = r + 1$ e $p' = r^2 + (t / (r + 1))$.

Tuttavia, poiché $1 \leq t \leq r$, abbiamo che $(t / (r + 1)) < 1$, il che implica che $p' < r^2 + 1$. Quindi, $p' r^2$, ma p' deve essere maggiore di 1, il che contraddice l'assunzione che p' e q' siano entrambi maggiori di 1.

Pertanto, la stringa $w' = 0^{r^2 \cdot (r+1) + t}$ non appartiene al linguaggio L , violando la condizione 3 del pumping lemma.

Questa violazione contraddice l'ipotesi che L sia regolare.

- $\{0^n 0^{2^n} \mid n, m \in \mathbb{N}\}$

Consideriamo la stringa $w = 0^p 0^{2^p}$, che appartiene a L . Poiché la lunghezza di w è $p + 2^p$, che è maggiore di p , possiamo suddividere w in x , y e z in modo che soddisfi le condizioni del pumping lemma. Poiché $|xy| \leq p$, y può contenere al massimo p zeri. Di conseguenza, y deve contenere zeri appartenenti solo alla prima parte della stringa (0^p) o solo alla seconda parte della stringa (0^{2^p}), ma non entrambe.

- Caso 1: y contiene zeri appartenenti solo alla prima parte della stringa (0^p). In questo caso, possiamo scrivere $x = 0^q$, $y = 0^r$ e $z = 0^{p-q-r} 0^{2^p}$, dove $q \geq 0$, $r > 0$ e $q + r \leq p$. Consideriamo la stringa $xy^2z = 0^{q+2r} 0^{2^p}$. Poiché $q + 2r \neq p$, la stringa xy^2z non appartiene a L , violando la condizione 3 del pumping lemma.
- Caso 2: y contiene zeri appartenenti solo alla seconda parte della stringa (0^{2^p}). In questo caso, possiamo scrivere $x = 0^p 0^s$, $y = 0^t$ e $z = 0^{2^p-s-t}$, dove $s \geq 0$, $t > 0$ e $s + t \leq 2^p$. Consideriamo la stringa $xy^2z = 0^p 0^{s+2t} 0^{2^p-s-t}$. Poiché $s + 2t \neq 2^p$, la stringa xy^2z non appartiene a L , violando la condizione 3 del pumping lemma.

In entrambi i casi, abbiamo una violazione delle condizioni del pumping lemma, il che contraddice l'ipotesi che L sia regolare.

- Per ognuno dei seguenti linguaggi, fornire la lunghezza minima di pumping e giustificare opportunamente la risposta:

- $1^* 01^* 01^*$
- $10(11^* 0)^* 0$
- 1011
- $(01)^*$

La lunghezza minima di pumping p è il più piccolo numero intero positivo che soddisfa le condizioni del pumping lemma per un dato linguaggio regolare L . Di fatto, è la più piccola lunghezza per la quale vale il pumping lemma per quel linguaggio.

Per ciascun linguaggio, possiamo quindi determinare:

- (a) La lunghezza minima di pumping per questo linguaggio è 3.
- Supponiamo che la lunghezza di pumping sia 2. Consideriamo la stringa $w = 101$ appartenente al linguaggio. Poiché $|w| = 3$, possiamo suddividere w in x, y e z ($w = xyz$) in modo che soddisfi le condizioni del pumping lemma. Tuttavia, qualunque sia la suddivisione, la condizione 3 del pumping lemma verrà violata per alcuni valori di i . Pertanto, la lunghezza di pumping non può essere 2.
- D'altra parte, se scegliamo la lunghezza di pumping come 3, possiamo suddividere qualsiasi stringa w appartenente al linguaggio in x, y e z in modo che le condizioni del pumping lemma siano soddisfatte. Quindi, la lunghezza minima di pumping per questo linguaggio è 3.
- (b) La lunghezza minima di pumping per questo linguaggio è 5.
- Supponiamo che la lunghezza di pumping sia 4. Consideriamo la stringa $w = 10110$ appartenente al linguaggio. Poiché $|w| = 5$, possiamo suddividere w in x, y e z ($w = xyz$) in modo che soddisfi le condizioni del pumping lemma. Tuttavia, qualunque sia la suddivisione, la condizione 3 del pumping lemma verrà violata per alcuni valori di i . Pertanto, la lunghezza di pumping non può essere 4.
- (c) La lunghezza minima di pumping per questo linguaggio è 5.
- Poiché il linguaggio contiene solo una stringa di lunghezza 4, non possiamo scegliere una lunghezza di pumping inferiore a 5 (per il pumping lemma, la lunghezza di pumping deve essere minore o uguale alla lunghezza della stringa più lunga del linguaggio).
- Quindi, la lunghezza minima di pumping per questo linguaggio è 5.
- (d) La lunghezza minima di pumping per questo linguaggio è 2.
- Consideriamo la stringa $w = 0101$ appartenente al linguaggio. Poiché $|w| = 4$, possiamo suddividere w in x, y e z ($w = xyz$) in modo che soddisfi le condizioni del pumping lemma con una lunghezza di pumping di 2.
- Inoltre, per qualsiasi stringa w appartenente al linguaggio, possiamo suddividere w in x, y e z in modo che le condizioni del pumping lemma siano soddisfatte con una lunghezza di pumping di 2.
- Quindi, la lunghezza minima di pumping per questo linguaggio è 2.

3. Sia $A/B = \{w \mid wx \in A \text{ per qualche } x \in B\}$. Si mostri che se A è regolare, allora A/B è regolare.

Per dimostrare che se A è regolare, allora A/B è regolare, possiamo costruire un automa a stati finiti non deterministico (NFA) che riconosce il linguaggio A/B a partire dall'NFA che riconosce il linguaggio regolare A .

Supponiamo di avere un NFA $M = (Q, \Sigma, \delta, q_0, F)$ che riconosce il linguaggio regolare A . Costruiremo un nuovo NFA $M' = (Q', \Sigma, \delta', q'_0, F')$ che riconosce il linguaggio A/B .

La costruzione di M' è la seguente:

- $Q' = Q \cup \{q'_0\}$, dove q'_0 è un nuovo stato iniziale
- Σ è l'alfabeto di M (e di conseguenza anche di M')
- $F' = F$, gli stati finali di M' sono gli stessi di M
- δ' è definita come segue:
 - $\delta'(q'_0, \epsilon) = q_0 \cup \{q \in Q \mid \exists x \in B, (\delta_0, x) = q\}$
 - $\delta'(q, a) = \delta(q, a)$ per ogni $q \in Q$ e $a \in \Sigma$

Essenzialmente, il nuovo NFA M' ha un nuovo stato iniziale q'_0 e le transizioni da q'_0 sono definite in modo tale da coprire tutte le stringhe x appartenenti a B che possono essere "saltate" per raggiungere uno degli stati di M a partire dallo stato iniziale q_0 .

Dimostriamo ora che il linguaggio riconosciuto da M' è effettivamente A/B .

(Questa parte fatta per completezza ed eventuale chiarezza sul fatto che sono equivalenti - equivale a dimostrare un "se e solo se").

- Dimostriamo che $L(M') \supseteq A/B$: Sia $w \in A/B$. Allora esiste un cammino di accettazione in M' che riconosce w . Questo cammino inizia nello stato q'_0 e, per raggiungere uno stato $q \in Q$, deve "saltare" una stringa $x \in B$. Dopo aver raggiunto q , il cammino continua a leggere w e raggiunge uno stato finale $f \in F$. Questo significa che wx è riconosciuto da M , quindi $wx \in A$. Di conseguenza, $w \in A/B$.
- Dimostriamo che $A/B \subseteq L(M')$: Sia $w \in A/B$. Allora esiste una stringa xB tale che wxA . Poiché M riconosce A , esiste un cammino di accettazione in M che riconosce wx . Questo cammino inizia nello stato q_0 , legge x e raggiunge uno stato $q \in Q$, quindi

continua a leggere w e raggiunge uno stato finale $f \in F$. Per costruzione di M' , esiste una transizione da $q0'$ a q che "salta" x , e il resto del cammino in M è valido anche in M' . Quindi, w è riconosciuto da M' . Poiché $L(M') = A/B \epsilon M'$ è un NFA, abbiamo dimostrato che se A è regolare, allora A/B è regolare.

4. Si consideri x , prefisso di una stringa y se esiste una stringa z tale che $xz = y$ e che x sia un prefisso proprio di y se in aggiunta $x \neq y$.

Si mostri che il linguaggio definito come

$NOPREFIX(A) = \{w \in A \mid \text{nessun prefisso proprio di } w \text{ è un membro di } A.\}$

(Eventuale altro riferimento per questo esercizio: Appello 26/06/2023, disponibile nel Moodle dell'anno scorso 2022-2023)

Per dimostrare che il linguaggio $NOPREFIX(A)$ è regolare, possiamo costruire un automa a stati finiti deterministico (DFA) che riconosce $NOPREFIX(A)$ a partire dal DFA che riconosce A .

Supponiamo di avere un DFA $M = (Q, \Sigma, \delta, q_0, F)$ che riconosce il linguaggio regolare A . Costruiremo un nuovo DFA $M' = (Q', \Sigma, \delta', q'_0, F')$ che riconosce il linguaggio $NOPREFIX(A)$.

La costruzione di M' è la seguente:

- $Q' = Q \times \{0, 1\}$, dove gli stati in $Q1$ rappresentano gli stati "marcati" che indicano che è stato incontrato un prefisso di una stringa accettata da M .
- Σ è l'alfabeto di M (e di conseguenza anche di M').
- $q'_0 = (q_0, 0)$, lo stato iniziale di M' è la coppia formata dallo stato iniziale di M e dal flag 0 (non marcato).
- $F' = \{(q, 0) \mid q \in F\}$, gli stati finali di M' sono le coppie $(q, 0)$ dove q è uno stato finale di M .
- La funzione di transizione δ' è definita come segue:
 - $\delta'((q, 0), a) = (\delta(q, a), 0)$ se $\delta(q, a) \notin F$
 - $\delta'((q, 0), a) = (\delta(q, a), 1)$ se $\delta(q, a) \in F$
 - $\delta'((q, 1), a) = (\delta(q, a), 1)$ per ogni $q \in Q$ e $a \in \Sigma$

Essenzialmente, il nuovo DFA M' mantiene una copia dello stato corrente di M e un flag che indica se è stato incontrato un prefisso di una stringa accettata da M . Se il flag è impostato a 1, significa che è

stato incontrato un prefisso e quindi M' non può accettare la stringa corrente.

Poiché $L(M') = NOPREFIX(A)$ e M' è un DFA, abbiamo dimostrato che se A è regolare, allora $NOPREFIX(A)$ è regolare.