

1. Dati i linguaggi A e B , lo *shuffle perfetto* di A e B è il linguaggio

$$\{w \mid w = a_1 b_1 a_2 b_2 \dots a_k b_k, \text{ dove } a_1 a_2 \dots a_k \in A \text{ e } b_1 b_2 \dots b_k \in B, \text{ ogni } a_i, b_i \in \Sigma\}$$

Mostrare che la classe dei linguaggi regolari è chiusa rispetto allo shuffle perfetto, cioè che se A e B sono linguaggi regolari allora anche il loro shuffle perfetto è un linguaggio regolare.

A, B sono regolari \rightarrow Shuffle è regolare

$a = |1|0|1|$

$b = |0|0|1|$

$w = |10|00|11$

$\Sigma^* = (0+1)^*$ indica tutte le stringhe (compresa la vuota) ottenibili da linguaggio

Sappiamo quindi che esistono due automi del tipo:

$$D_A = \{Q_A, \Sigma, \delta_A, q_A, F_A\}$$

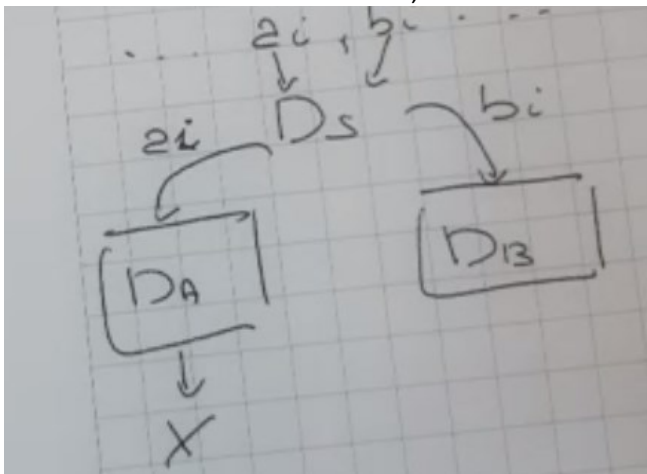
$$D_B = \{Q_B, \Sigma, \delta_B, q_B, F_B\}$$

Vogliamo produrre quindi:

$$D_S = \{Q_S, \Sigma, \delta_S, q_S, F_S\}$$

L'automata quindi sa l'ordine e che gli elementi sono invertiti, conseguentemente si va comunque almeno ad uno stato accettante.

L'idea informale è di costruire un automa che prende la posizione generica, richiama la procedura e poi non vedendo il funzionamento interno, richiamo i simboli avendo una dimostrazione generica:



Dim. formale

$$D_S = \{Q_S, \Sigma, \delta_S, q_S, F_S\}$$

$$\text{Generica } \delta(q_x, a) \rightarrow q$$

Ad esempio metto come input il simbolo "a":

$$\delta((x, y, A), a) \rightarrow (\delta_A(x, a), y, B)$$

L'idea quindi è che gli input si alternano e mi resta da aggiornare lo stato in cui si trova A , usando le sue funzioni di transizione.

x stato corrente D_A

y stato corrente D_B

A flag

$$\delta((x, y, B), b) \rightarrow (x, \delta_B(y, b), A)$$

$$\delta(q, x) \rightarrow q$$

L'idea è che la tupla rappresenti gli stati dell'automa, ottenendo gli stati nuovamente e riorganizzandoli, costruendo automaticamente la funzione di transizione di δ_s .

Quindi l'idea è di avere un aggiornamento degli stati tali da avere le transizioni ricombinate, rimanendo con gli input uguali

- Q_s

$$Q_A \times Q_B \times \{A, B\}$$

Usiamo quindi "x" come prodotto cartesiano, tipo avendo

$$\{a, b\} \times \{c, d\} \quad \text{eseguo il prodotto tra insiemi (prodotto cartesiano)}$$

$$= \{(a, c), (a, d), (b, c), (b, d)\}$$

$$\{q_x, q_y\} \times \{q_z\} \times \{A, B\}$$

$$\{(q_x, q_z), (q_y, q_z)\} \times \{A, B\}$$

$$(q_x, q_z, A) \dots$$

La risposta è: creare gli stati come prodotto (producendo tutto "brutalmente" con tutte le possibili combinazioni).

Stato iniziale q_s

$$q_s = (q_A, q_B, A)$$

$$\delta[(q_A, q_B, A), a_1] \rightarrow (\delta(q_A, a_1))$$

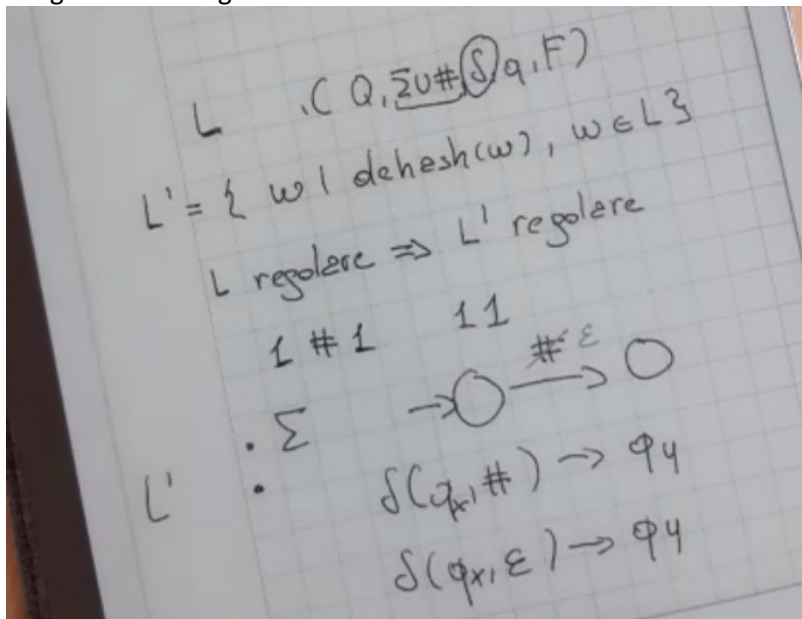
$$F_A \times F_B \times$$

$$b_K$$

$$\delta((q_x, q_y, B), b_K)$$

$$L' = \{w \mid \text{dehash}(w), w \in L\}$$

$$L \text{ regolare} \rightarrow L' \text{ regolare}$$



L linguaggio regolare su Σ

$$L' = \{y \mid xy \in L \text{ per } x \in \Sigma^*\}$$

$$w = 1011, \quad w \in L$$

11
011
1011
 ε
(Q, δ , q, F, Σ)

Essendo A e B linguaggi regolari, esistono due DFA che li riconoscono. Essenzialmente essendo ai, bi simboli appartenenti all'alfabeto, si costruisce un automa E che contiene tutti gli stati degli automi EA, EB che accettano i linguaggi A e B, e che ogni input dispari (le ai) si sposta nello stato corrispondente di EA, e ad ogni input pari (le bi) si sposta nello stato corrispondente di B. L'automa accetta quando attraversa sia lo stato finale di Ea che lo stato finale di EB.

2. Sia A un linguaggio, e sia $DROPOUT(A)$ come il linguaggio contenente tutte le stringhe che possono essere ottenute togliendo un simbolo da una stringa di A:

$$DROPOUT(A) = \{xz \mid xyz \in A \text{ dove } x, y \in \Sigma^* \text{ e } y \in \Sigma\}.$$

Mostrare che la classe dei linguaggi regolari è chiusa rispetto all'operazione $DROPOUT$, cioè che se A è un linguaggio regolare allora $DROPOUT(A)$ è un linguaggio regolare.

Quindi:

$$\{xz \mid xyz \in A, x, z \in \Sigma^*, y \in \Sigma\}$$

Sia $M=(Q, \Sigma, \delta, q_0, F)$ un DFA che riconosce A.

$$M'=(Q', \Sigma, \delta', q_0, F')$$

In pratica si saltano a random tutti gli elementi, andando in uno, un altro o un altro ancora, senza un ordine definito. Per ognuno deve ricordarsi che ha saltato quel carattere (non ne salta più di uno):

$$w_1w_2 \dots w_r$$

$$q_0q_1 \dots q_n$$

Si introduce quindi un simbolo tale da capire se si è saltato o meno, eseguendo poi la transizione:

$$Q'=Q \times \{0,1\} \quad (q, 0) \quad \text{"non ho saltato"}$$

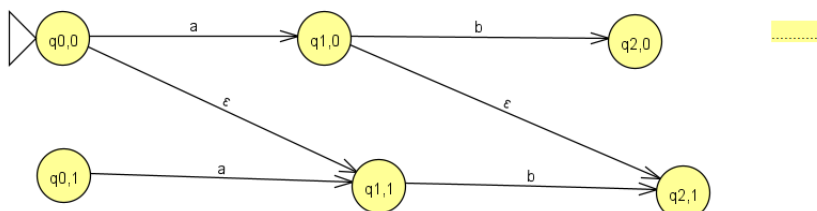
$$q' = (q_0, 0) \quad (q, 1) \quad \text{"ho saltato"}$$

Qui ragiono che per qualche $a \in \Sigma$

$$\delta'((q,0)a) = \{\delta(q_0,a),0\} \quad (\text{quindi ho un simbolo e vado avanti con 0})$$

$$\delta((q,0)\varepsilon) = \{(p,1) \mid \delta(q,a)=p\} \quad (\text{salto un simbolo oppure rimango sulla transizione stessa perché vuota})$$

$$\delta'((q,1)a) = \{\delta(q,a),1\} \quad (\text{quindi ho un simbolo e vado avanti con 1})$$



$$F'=\{(q1,1) \mid q_0 \in F\}$$

Linguaggio regolare significa che esiste un DFA che lo riconosce. Tutte le parole di L sono di lunghezza > 1 (altrimenti non si potrebbe togliere una lettera). Togliendo una lettera ad una parola composta da una sola lettera otteniamo un set vuoto, che per definizione è un linguaggio regolare. Nel caso l'automa sia composto da due o più simboli dell'alfabeto dobbiamo semplicemente costruire un automa che sostituisce la transizione che accettava y e con una epsilon transizione allo stato successivo.

3. Per una stringa $w = w_1w_2 \dots w_n$, l'inversa di w è la stringa $w^R = w_n \dots w_2w_1$. Per ogni linguaggio A , sia $A^R = \{w^R \mid w \in A\}$. Mostrare che se A è regolare allora lo è anche A^R .

Questo esercizio è diverso da quelli visti in precedenza; dato un linguaggio regolare, ci viene chiesto di dimostrare che con una modifica possiamo ottenere un altro linguaggio regolare. Dobbiamo perciò ragionare in modo più generico, senza produrre un automa specifico. Una soluzione informale è la seguente

1. Dato il linguaggio A , produco un NFA_1 che lo riconosce A .
2. Per produrre l'automa che riconosce il linguaggio A^R , posso modificare NFA_1 in NFA_2
 1. Inverto la direzione delle transizioni
 2. Lo stato iniziale di NFA_1 diventa accettante
 3. Lo stato accettante di NFA_1 diventa iniziale
3. Se NFA_1 ha più di uno stato accettante, aggiungo uno nuovo stato iniziale e una epsilon transizione verso tutti gli ex-stati accettanti.

Per questo esercizio è necessario partire da un NFA in modo da avere un solo percorso da stato iniziale ad accettante e poter seguire l'automa in senso inverso.

Dato il linguaggio A regolare formato da tutte le stringhe $w = w_1w_2w_3 \dots w_n$ si può dimostrare che il suo inverso (quello formato da tutte le stringhe $w^R = w_n \dots w_3w_2w_1$) è ancora regolare in quanto è possibile partendo dal DFA che riconosce A costruire il DFA che riconosce A^R

Procedimento:

- Si girano tutte le transizioni del DFA che riconosce A (cambiare verso delle frecce)
- Lo stato iniziale diventa il nuovo stato finale (tutti gli altri stati diventano non finali)
- Creo il nuovo stato iniziale e da questo faccio partire tante epsilon transizioni che finiscono dove c'erano i vecchi stati finali

4. Sia $A/b = \{w \mid wb \in A\}$. Mostrare che se A è un linguaggio regolare e $b \in \Sigma$, allora A/b è regolare.

Per dimostrare che A/b è regolare, possiamo costruire un automa a stati finiti, in particolare un ϵ -NFA che, partendo dall'automa a stati finiti $A = (Q \cup \{q_0\}', \Sigma, q_0, \delta, F)$ che ha lo stesso insieme di stati, stesso stato iniziale e gli stessi stati finali. Essendo che si deve sempre essere la produzione di " w ", lo stato iniziale conterrà, con l'aggiunta di una ϵ -transizione, un automa con tutti gli stati finali raggiungibili dal vecchio stato iniziale, con l'aggiunta della nuova stringa. Gli stati finali sono inoltre gli stessi.

oppure:

Il linguaggio A/b riconosce tutte le stringhe che hanno la proprietà che se concatenano b ad una qualsiasi stringa ottengo una stringa che appartiene al linguaggio regolare A . Dato che i linguaggi regolari sono chiusi rispetto alla concatenazione, per ottenere una stringa appartenente ad A concatenando b a w significa che anche w deve appartenere ad un linguaggio regolare quindi A/b è un linguaggio regolare. Si dimostra che i linguaggi regolari sono chiusi per concatenazione collegando lo stato finale di un DFA che riconosce un linguaggio regolare con un epsilon transizione allo stato iniziale di un DFA che riconosce un altro linguaggio regolare in modo da concatenare le due stringhe riconosciute da ciascun DFA

5. Sia $A/B = \{w \mid wx \in A \text{ per qualche } x \in B\}$. Mostrare che se A è un linguaggio regolare e B un linguaggio qualsiasi, allora A/B è regolare.

Per dimostrare che A/B è regolare, possiamo costruire un automa a stati finiti, in particolare un ϵ -NFA che, partendo dall'automa a stati finiti $A = (Q \cup \{q_0\}', \Sigma, q_0, \delta, F)$ che ha lo stesso insieme di stati, stesso stato iniziale e gli stessi stati finali. Essendo che si deve sempre essere una stringa in più, lo stato iniziale conterrà, con l'aggiunta di una ϵ -transizione, un automa con tutti gli stati finali raggiungibili dal vecchio stato iniziale, con l'aggiunta della stringa prefissa. Gli stati finali sono inoltre gli stessi.

oppure:

Il linguaggio A/B è formato dalle parole w ottenute a partire da una parola wx appartenente al linguaggio regolare A , alle quali viene rimossa la parte x , che è una parola del linguaggio B (linguaggio qualsiasi). $A/B = \{w \mid wx \in A \text{ per qualche } x \in B\}$

w deve quindi essere un prefisso di A e per esserlo deve essere regolare. Si può dimostrare che è regolare creando un DFA che accetta un prefisso di A . Il DFA che riconosce un prefisso di A riceve w in input e rifiuta non appena un possibile simbolo in input non fa match con quello che il DFA si aspetta. Nel caso in cui venga consumato tutto l'input (w) allora il DFA accetta andando in uno stato finale; se non fa match o se w è più lunga della stringa più lunga del linguaggio A allora rifiuta.

Il pumping lemma afferma che ogni linguaggio regolare ha una lunghezza del pumping p , tale che ogni stringa del linguaggio può essere iterata se ha lunghezza maggiore o uguale a p . La *lunghezza minima del pumping* per un linguaggio A è il più piccolo p che è una lunghezza del pumping per A . Per ognuno dei seguenti linguaggi, dare la lunghezza minima del pumping e giustificare la risposta.

- | | | |
|----------------------------|---|-------------------|
| (a) 110^* | (g) $10(11^*0)^*0$ | (m) ε |
| (b) $1^*0^*1^*$ | (h) 101101 | (n) $1^*01^*01^*$ |
| (c) $0^*1^*0^*1^* + 10^*1$ | (i) $\{w \in \Sigma^* \mid w \neq 101101\}$ | (o) 1011 |
| (d) $(01)^*$ | (j) 0001^* | (p) Σ^* |
| (e) \emptyset | (k) 0^*1^* | |
| (f) $0^*01^*01^*$ | (l) $001 + 0^*1^*$ | |

- (a) 3: si necessita di avere almeno tre caratteri per poter cominciare ad eseguire il pumping. Infatti parole come 00, 11 o simili non sono accettate.
- (b) 3: parole come 11 (cioè xy^0z) non sono accettate per costruzione, prendendo ad esempio 101 come xyz si nota che devono esserci entrambi per poter cominciare a pompare.
- (c) 3: L'unione c'è ma non ci interessa, la proprietà vale comunque. Inoltre si nota che devo avere almeno due 1 ma l'idea è la stessa dell'esercizio precedente.
- (d) 1: l'insieme vuoto deve avere almeno una stringa con cui operare e quindi pompare all'infinito una parola.
- (e) 2: abbiamo bisogno di entrambi i caratteri per pompare.
- (f) 3: le parole devono essere divise e definite in 3 pezzi anche qui per formare una parola valida e pompare.
- (g) 4: in pratica la lunghezza deve essere 4 in quanto potrebbe esserci una suddivisione che non sbilancia la stringa avendo soli tre caratteri, avendo stringhe che pompate non sarebbero nel linguaggio. Sapendo che $(11^*0)^*$ è l'espressione minima, avremo quantomeno bisogno di 0 oppure 1 per cominciare a pompare.
- (h) 4: si vede infatti che ipotizzando 3 come lunghezza minima, potremmo avere una stringa del tipo 111 che rimane sempre regolare.
- (i) 3: se sappiamo (penso io) che la stringa precedente non fa parte del linguaggio, significa considerando l'alfabeto precedente che necessitiamo di almeno tre caratteri per poter avere una stringa pompabile, considerando il caso complementare a quello descritto sopra.
- (j) 4: potremmo banalmente avere la stringa 000 che non sarebbe pompabile.
- (k) 2: anche qui, scegliendo 1 non sarebbe pompabile; con 2 almeno avremmo il possibile caso 01, regolarmente pompabile.
- (l) 3: lasciando stare il caso dell'unione, comunque si nota che la stringa 001 non può avere 2 come lunghezza minima pompabile.
- (m) 1: la stringa vuota deve necessitare di almeno un carattere qualsiasi.
- (n) 3: si considera infatti che la minima stringa effettivamente pompabile sia 001
- (o) 5: di fatto 1011 potrebbe non essere accettata dal linguaggio come pompata, perché già integralmente parte del linguaggio stesso.
- (p) 2: considerando che la stringa vuota necessita di almeno un carattere e l'alfabeto la comprende di sicuro essendo star, potrebbe bastare per un generico alfabeto avere un solo altro carattere.

7. Sia $\Sigma = \{0, 1\}$, e considerate il linguaggio

$$D = \{w \mid w \text{ contiene un ugual numero di occorrenze di } 01 \text{ e di } 10\}$$

Mostrare che D è un linguaggio regolare.

L'obiettivo come al solito è di costruire un automa in grado di riconoscere questo linguaggio presentato. L'alfabeto $\Sigma = \{0, 1\}$ e lo stato iniziale. Partiamo quindi dall'automa $A=(Q, \Sigma, q_0, \delta, F)$ con una funzione di transizione costruita come:

- $\delta(x, y, 01, a) = (\delta_1(x, a), y, 01)$
- $\delta(x, y, 10, a) = (\delta_1(x, a), y, 10)$

perché devo avere ugual numero di occorrenze di 01/10.

L'insieme degli stati finali sarà costituito da:

$$F = F_A \times F_B \times \{0, 1\}$$

Avendo come stato iniziale (partendo da A):

$$q = (q_A, q_B, A)$$

e avendo come stato accettante

$$F = F_A \times F_B \times \{A\}$$

Let $\Sigma = \{0, 1\}$ and let

$D = \{w \mid w \text{ contains an equal number of occurrences of the substrings } 01 \text{ and } 10\}$.

Thus $101 \in D$ because 101 contains a single 01 and a single 10, but $1010 \notin D$ because 1010 contains two 10s and only one 01. Show that D is a regular language.

Solution:

This language is regular because it can be described by a regular expression and a FA (NFA):

Regular Expression

$$(1^+ 0^+ 1^+)^* + (0^+ 1^+ 0^+)^*$$

8. Sia $\Sigma = \{0, 1\}$.

- Mostrare che il linguaggio $A = \{0^k u 0^k \mid k \geq 1 \text{ e } u \in \Sigma^*\}$ è regolare.
- Mostrare che il linguaggio $B = \{0^k 1 u 0^k \mid k \geq 1 \text{ e } u \in \Sigma^*\}$ non è regolare.

1) Assumiamo per assurdo il linguaggio sia regolare.

Ricordiamo che "u" è un simbolo della grammatica, quindi può essere sia 0 che 1

Allora esisterà una suddivisione $w=xyz$ tale da avere $x \neq \epsilon$, $|xy| \leq k$ con

$w=0^p u 0^p$ avendo almeno xy completamente piena di 0.

Il pumping xy^iz non dimostra lo sbilanciamento, infatti semplicemente la stringa

rimane composta di egual numero di 0 e sempre la stringa u in mezzo.

Un qualsiasi pumping per $k > 0$, ad esempio.

$w=0^k u 0^{p-k}$ comunque non dimostra alcunché,

infatti, abbiamo un numero di u sempre compreso tra un egual numero di zeri ed il linguaggio rimane sempre regolare.

2) Assumiamo per assurdo il linguaggio sia regolare.

Allora esisterà una suddivisione $w=xyz$ tale da avere $x \neq \epsilon$, $|xy| \leq k$ con

$u=0$ e $w=0^p 1 0^q$

In questo caso, siccome $x=0^k$ e $y=0^q$ allora avremo in "z" la rimanente parte della stringa, quindi:

$xy^0z = xz = 0^k 1 0^{k-p-q}$, quindi $0^{2k-p-q} 1$

Si vede quindi che il numero di 0 è sbilanciato rispetto al numero di 1 e il linguaggio quindi non può essere regolare.

9. Dimostrare che i seguenti linguaggi non sono regolari.

- (a) $\{0^n 1^m 0^n \mid m, n \geq 0\}$
- (b) $\{0^n 1^m \mid n \neq m\}$
- (c) $\{w \in \{0, 1\}^* \mid w \text{ non è palindroma}\}$
- (d) $\{wtw \mid w, t \in \{0, 1\}^+, \text{dove } \{0, 1\}^+ \text{ è l'insieme di tutte le stringhe binarie di lunghezza maggiore o uguale a 1}\}$

a) Assumiamo per assurdo il linguaggio sia regolare.

Allora esisterà una suddivisione $w=xyz$ tale da avere $x \neq \epsilon$, $|xy| \leq k$ con

$w=0^k 1^p 0^k$

e con pumping $i=0$

avremo $0^{k-p} 1^k 0^k$ che chiaramente non appartiene al linguaggio e quindi non è regolare.

b) Assumiamo per assurdo il linguaggio sia regolare.

Allora esisterà una suddivisione $w=xyz$ tale da avere $x \neq \epsilon$, $|xy| \leq k$ con

$w=0^p 1^q$ con $p > q$ o $p < q$, entrambi > 0 e $< k$

In questo caso è semplice perché basta letteralmente prendere $x=\epsilon$, $y=0^p$ e $z=1^q$

tale che pompando i avremo sempre un numero > 0 rispetto agli 1.

Banalmente la cosa è dimostrata anche nel caso xy^0z con $n \neq m$ e quindi

si dimostra che non è regolare.

c) Assumiamo per assurdo il linguaggio sia regolare.

Allora esisterà una suddivisione $w=xyz$ tale da avere $x \neq \epsilon$, $|xy| \leq k$ con

$w=0^p 1^q 0^r$

In questo modo, eseguendo un pumping di $i=0$, avremo

$0^{p-r} 1^q 0^p$ che rivela la parola non sia nel linguaggio e dunque non è regolare.

d) Assumiamo per assurdo il linguaggio sia regolare.

Allora esisterà una suddivisione $w=xyz$ tale da avere $x \neq \epsilon$, $|xy| \leq k$ con

$0^k 1^p 0^k$. Eseguendo il pumping, ad esempio con $i=2$

e scegliendo $y=0^k 1^p$ avremmo una cosa del tipo $0^{2k+2p} 0^k$ che non rispetta $|xy| \leq k$

nel numero di 0 ed il linguaggio non è regolare.

oppure:

$|xy| \leq k$, $|y| > 0$ $|w| \geq k$

(a) $\{0^n 1$

$m 0$

$m \text{ con } m, n \geq 0\}$

Ponendo $n=0$ ottengo $w = 00 1$

$k 0$

k

. $|w| > k$

$x = 1a$

, $y = 1b$

, $z = 1c0$

k e sappiamo che $a+b+c = k$.

Se noi pompiamo y con $i \neq 1$ $a + ib + c \neq k$, allora w non appartiene più al linguaggio, che non è regolare.

(b) $\{0^n 1$

$m \mid n \neq m\}$

$w = 0^{k-1} 0 1^{p+1}$, $0 = k$, $1 = k+1$

(c) $\{w \in \{0, 1\}^* \mid w \text{ è palindroma}\}$

$w = 0^k 111 0^k$

$x = 0a$

, $y = 0b$

, $z = 0c\ 111\ 0k$

Se io pompo y per $i = 2$ otteniamo che $a+bi+c > k$ il che dimostra che $0a+bi+c1110k$ non è più palindroma, quindi non è regolare.

(d) $\{wtw \mid w, t \in \{0,1\}^+\}$

$G = wtw = 0k1\ 11\ 0k1\ w = 0k1\ t = 11$

$x = 0a$

, $y = 0b$

, $c = 0c1\ 11\ 0k1$

Se io pompo y per $i = 2$ otteniamo che $a+bi+c > k$ rendendo la parola originariamente nella forma wtw nella forma $0a+bi+c1\ 11\ 0k1$ che non rispetta il linguaggio, quindi non è regolare.

10. Per ogni linguaggio A , sia $SUFFIX(A) = \{v \mid uv \in A \text{ per qualche stringa } u\}$. Mostrare che la classe dei linguaggi context-free è chiusa rispetto all'operazione di $SUFFIX$.

Sappiamo che L è context-free. Esiste quindi una grammatica G in forma normale del tipo::

$A \rightarrow BC$

$D \rightarrow d$

Se $A \rightarrow BC$ allora possiamo scrivere una parola $w=bc$

Il suffisso che è parte della parola potrà essere:

- caso 1 (prefisso nullo) BC
- caso 2 (prefisso che è una parte di B): $B'C$
- caso 3 (prefisso che è esattamente B): C
- caso 4 (prefisso che è B più una parte di C): C'
- caso 5 (ho solo prefisso): ϵ

Ci sarà quindi la regola:

$A \rightarrow BC \mid B'C \mid C \mid C' \mid \epsilon$

$D' \rightarrow d \mid \epsilon$

La grammatica G' quindi avrà S' come stato iniziale e condividerà gli stessi stati per le stesse regole. Così descritta la grammatica G' è context-free in quanto in CNF.

oppure:

Dimostro che i linguaggi context-free sono chiusi rispetto all'operazione suffisso con un PDA che riconosce il suffisso di A :

Procedimento

Chiamiamo M il PDA che riconosce il suffisso di A .

Dato che A è context-free allora esiste un PDA che lo riconosce e lo chiamiamo $M1$.

Facciamo una copia di $M1$ e la chiamiamo $M2$, che ha esattamente gli stessi stati, stack, e transizioni di $M1$.

$M1$ ed $M2$ formeranno il PDA M .

Cambio le ϵ -transizioni di $M2$: se la transizione originale era $a, b \rightarrow \epsilon$ la cambiamo in $\epsilon, b \rightarrow \epsilon$.

Per ogni stato di $M2$ aggiungo una transizione $\epsilon, \epsilon \rightarrow \epsilon$ che va al corrispondente stato di $M1$.

Lo stato iniziale di $M2$ è lo stato iniziale di M .

Possiamo vedere che M costruirà lo stack appropriatamente ignorando l'input.

Ad un certo punto M non deterministicamente transita dallo stato $M2$ al corrispondente stato in $M1$, comincia a prendere il primo carattere del suffisso come input e continua le transizioni in $M1$ da questo punto in poi.

Quindi tutti i suffissi della stringa che appartengono ad A saranno accettati da M

11. Usa i linguaggi $A = \{a^m b^n c^n \mid m, n \geq 0\}$ e $B = \{a^n b^n c^m \mid m, n \geq 0\}$ per mostrare che la classe dei linguaggi context-free non è chiusa per intersezione.

L'operazione di intersezione è definita liberamente per i linguaggi regolari; significa quindi che se eseguiamo l'intersezione di due linguaggi context-free, in questo caso A e B, anch'essa deve essere CF. Considerando ad esempio due generiche parole del linguaggio $w = a^p b^q c^p$ e $w = a^q b^p c^q$

- caso base, con $p=q=0$ sono: abc & abc , dunque l'intersezione è fattibile
- caso induttivo, con generici esponenti p, q entrambi > 0 possiamo avere:

(es. $p=2, q=3$)

$a^2 b^3 c^2$ & $a^3 b^2 c^3$

avremo che l'intersezione genererà numero diverso di a , di b , e di c .

Quindi ad esempio potremmo avere stringhe del tipo:

$a^2 b^5 c^3$, $a^5 b^3 c^3$, ecc.

Eseguendo l'intersezione si nota che otterremmo stringhe diverse e, se volessimo dimostrarlo con il PL per linguaggi CF, otterremmo sempre stringhe sbilanciate. Dunque l'operazione di intersezione non può essere chiusa per il linguaggi CF.

oppure:

Se una stringa è accetta per A significa che ha lo stesso numero di b e di c - Se una stringa è accettata per B significa che ha lo stesso numero di A e b .

Se w è riconosciuto significa che $\#A = \#B = \#C$. quindi $\{a^n + b^n + c^n \mid n \geq 0\}$ è facilmente dimostrabile che non è context free.

12. Dimostrare che i seguenti linguaggi sono context-free. Salvo quando specificato diversamente, l'alfabeto è $\Sigma = \{0, 1\}$.

- $\{w \mid w \text{ contiene almeno tre simboli uguali a } 1\}$
- $\{w \mid \text{la lunghezza di } w \text{ è dispari}\}$
- $\{w \mid w \text{ inizia e termina con lo stesso simbolo}\}$
- $\{w \mid \text{la lunghezza di } w \text{ è dispari e il suo simbolo centrale è } 0\}$
- $\{w \mid w = w^R, \text{ cioè } w \text{ è palindroma}\}$
- $\{w \mid w \text{ contiene un numero maggiore di } 0 \text{ che di } 1\}$
- Il complemento di $\{0^n 1^n \mid n \geq 0\}$
- Sull'alfabeto $\Sigma = \{0, 1, \#\}$, $\{w\#x \mid w^R \text{ è una sottostringa di } x \text{ e } w, x \in \{0, 1\}^*\}$
- $\{x\#y \mid x, y \in \{0, 1\}^* \text{ e } x \neq y\}$
- $\{xy \mid x, y \in \{0, 1\}^* \text{ e } |x| = |y| \text{ ma } x \neq y\}$
- $\{a^i b^j \mid i \neq j \text{ e } 2i \neq j\}$

Per dimostrare che i linguaggi *sono* context-free dobbiamo necessariamente fare delle derivazioni.

Se dovessimo dimostrare che *non* sono context-free allora dovremmo usare il PL per CFL.

- $S \rightarrow 1X1X1X1X$
 $X \rightarrow 0X \mid 1X \mid \epsilon$
- $S \rightarrow 0A \mid 1A$
 $A \rightarrow 0S \mid 1S \mid \epsilon$
- $S \rightarrow 0A \mid 1B$
 $A \rightarrow 0 \mid \epsilon$
 $B \rightarrow 1 \mid \epsilon$
- $S \rightarrow 0S0 \mid 1S1 \mid 0 \mid 0S1 \mid 1S0$
- $S \rightarrow 0 \mid 1 \mid 1S1 \mid 0S0 \mid \epsilon$
- $S \rightarrow T0T$
 $T \rightarrow 1T0 \mid 0T1 \mid 0T0 \mid \epsilon$
- $S \rightarrow 1A \mid 0A$
 $A \rightarrow A0 \mid$

- h) $S \rightarrow A$
 $A \rightarrow \#B|0A0|1A1$
 $B \rightarrow 0B|1B|\epsilon$
- i) $S \rightarrow A\#B|B\#A|\epsilon$
 $A \rightarrow TAT|0$
 $B \rightarrow TBT|1$
 $T \rightarrow 0|1$
- j) $S \rightarrow AB|BA$
 $A \rightarrow 0|0A0|0A1|1A0|1A1$
 $B \rightarrow 1|0B0|0B1|1B0|1B1$
- k) $S \rightarrow S_1|S_2$
 $S_1 \rightarrow aA$
 $A \rightarrow bAb|aA|\epsilon$
 $S_2 \rightarrow Bb|aBb$
 $B \rightarrow Bb|aaBb|aBb|\epsilon$

13. Se A e B sono linguaggi, definiamo $A \circ B = \{xy \mid x \in A, y \in B \text{ e } |x| = |y|\}$. Mostrare che se A e B sono linguaggi regolari, allora $A \circ B$ è un linguaggio context-free.

Se A e B sono linguaggi regolari, allora sono descrivibili mediante le medesime operazioni dei linguaggi regolari e successivamente almeno da un automa DFA/NFA.

Sia $M=(Q, \Sigma, \delta, q_0, F)$ un DFA che riconosce A .

$M'=(Q', \Sigma, \delta', q_0, F')$

Descriviamo quindi A , composto evidentemente da stringhe "x", poi B , composto da stringhe di tipo "y"

Significa quindi che per questi dettagliamo:

$(r_A, L) \rightarrow (s_A, M)$ se $\delta_L(r_A, x)=s_A$

$(r_B, L) \rightarrow (s_B, M)$ se $\delta_L(r_B, y)=s_A$

Similmente è definita $(r_A, L, M) \rightarrow (s_B, M, L)$ se $\delta_L(r_A, x)=\delta_L(r_B, y)$.

A queste condizioni notiamo che il linguaggio context-free necessariamente richiede che gli stati finali corrispondano tra i due linguaggi.

Articoliamo quindi come idea di stato finale $F(\delta_0, x), F(\delta_1, y) = F(X \bullet Y)$

Necessariamente l'idea è che, nel caso non esista uno dei due stati, lo stato finale deve essere:

$F(X \bullet Y) = F(\delta_0, x)$

oppure

$F(X \bullet Y) = F(\delta_1, y)$

per la caratterizzazione di termini di cui sopra. Dunque, il linguaggio è CF.

Il linguaggio contiene tutte le stringhe che sono la concatenazione di una stringa di A con una stringa di B della stessa lunghezza dove A e B sono regolari.

I linguaggi regolari sono chiusi per concatenazione quindi $A \circ B$ è ancora un linguaggio regolare.

I linguaggi regolari sono anche linguaggi context free perché è possibile scrivere una grammatica che li genera o un PDA che li riconosce.

Ad esempio, per riconoscere un linguaggio regolare con un PDA basta creare un automa a pila che non utilizza la pila nelle sue transizioni e si ottiene l'equivalente DFA che riconosce il linguaggio regolare

14. Dimostrare che se G è una CFG in forma normale di Chomsky, allora per ogni stringa $w \in L(G)$ di lunghezza $n \geq 1$, ogni derivazione di w richiede esattamente $2n - 1$ passi.

Sapendo che G è in forma normale di Chomsky, quindi nella forma

$A \rightarrow BC$

$A \rightarrow a$

Prendiamo l'esempio di una grammatica di un es. fatto in questo file:

$S' \rightarrow AX|YB|a|AS|SA$

$S \rightarrow AX|YB|a|AS|SA$

$A \rightarrow b|AX|YB|a|AS|SA$

$B \rightarrow b$

$X \rightarrow SA$

$Y \rightarrow a$

Consideriamo una stringa di lunghezza 1, possibilmente composta da una situazione del tipo

$A \rightarrow BC$

Sono tutti simboli terminali e quindi, ciascuno è in forma normale di Chomsky, non avendo regole unitarie/transizioni vuote/più di tre transizioni per regola.

Induttivamente, nel caso di $n > 1$, avremo una situazione in cui, per poter applicare Chomsky abbiamo bisogno di almeno 2 regole, una terminale e non terminale ed eventuali altre regole.

Quindi:

$A \rightarrow BC$

$B \rightarrow b$

$C \rightarrow b|c|d$

Ogni singola derivazione, per Chomsky, richiede esattamente l'applicazione di una sostituzione in altra regola e ogni regola può essere composta fino ad un massimo di 2 regole.

Essendoci almeno una regola terminale, per ipotesi, il numero di stringhe necessarie a comporre tale situazione sarà $2n$.

Data appunto la presenza di almeno una regola terminale, il numero di stringhe necessarie a mostrare tale situazione in una derivazione si compone di $2n - 1$ passi.

In una situazione reale come la grammatica ottenuta sopra, si nota che per ogni simbolo di partenza ne corrispondono almeno 2 derivati, con l'aggiunta della regola terminale.

Consideriamo quindi $S' \rightarrow AB$ ed $S \rightarrow XY$. Tale situazione comporta due regole non terminali ed un'altra possibilmente terminale. Dunque generalmente Chomsky richiede questo tipo di derivazione.

15. Un *automa a coda* è simile ad un *automa a pila* con la differenza che la pila viene sostituita da una coda. Una *coda* è un nastro che permette di scrivere solo all'estremità sinistra del nastro e di leggere solo all'estremità destra. Ogni operazione di scrittura (*push*) aggiunge un simbolo all'estremità sinistra della coda e ogni operazione di lettura (*pull*) legge e rimuove un simbolo all'estremità destra. Come per un PDA, l'input è posizionato su un nastro a sola lettura separato, e la testina sul nastro di lettura può muoversi solo da sinistra a destra. Il nastro di input contiene una cella con un blank che segue l'input, in modo da poter rilevare la fine dell'input. Un *automa a coda* accetta l'input entrando in un particolare stato di accettazione in qualsiasi momento. Mostra che un linguaggio può essere riconosciuto da un *automa deterministico a coda* se e solo se è Turing-riconoscibile.

L'equivalenza tra un *automa a coda* ed una TM deve essere mostrata; quindi, che il linguaggio può essere riconosciuto dall'*automa* e dunque dalla TM. Si dimostra creando una simulazione che fa capire che l'*automa* Q si comporti esattamente come la TM M .

La simulazione di Q come M funziona nel modo seguente:

si consideri l'intero nastro come una coda. Ciascun simbolo viene modificato e il movimento del nastro avviene a destra. Quando più di un numero viene pushato nella coda, si shiftano tutti i valori

a destra. Se si raggiunge la fine del nastro, allora il simbolo più a sinistra viene considerato e si capisce di essere arrivati alla fine.

La simulazione di M come Q funziona nel modo seguente:

l'alfabeto della macchina M viene espanso aggiungendo un simbolo extra e un simbolo marcatore a sinistra viene inserito nella coda. I simboli sono pushati a sinistra e letti (pop) a destra.

In modo più esteso e formale:

Il linguaggio dell'automa a coda deve essere Turing-riconoscibile.

Per fare ciò deve esistere una macchina TM che simula e riconosce questo linguaggio come una coda.

Immaginiamo quindi di avere un nastro ed una serie di simboli:

- ad ogni operazione di lettura, la macchina scorre e rimuove un simbolo a destra; per fare ciò lo marca e realizza uno shift di ogni elemento a destra (con un simbolo apposito)
- ad ogni operazione di scrittura, si deve capire se:
 - o ci si muove a sinistra, allora in questo caso si scrive il simbolo sul nastro e teniamo il simbolo di separazione, che sta a metà e ci aiuta a mantenere la separazione tra una parte e l'altra;
 - o attraverso questo, siamo capaci di riconoscere che, quando è il momento di fare il pop, si è mantenuto il simbolo corrispondente all'inizio del nastro
 - o ci si muove a destra, allora in questo caso scriviamo il simbolo sul nastro e si fa il push di # per mantenere traccia del corrispondente input

Ad ogni operazione, come descritto meglio per lo spostamento a sinistra, il # serve a mantenere traccia dell'operazione di inserimento e successivamente si fa il pop del simbolo corrispondente all'inizio del nastro (logica di FIFO). Una volta raggiunta la fine del nastro, capiamo se al simbolo più a sinistra corrisponde il #; se ciò avviene siamo in presenza di pila vuota ed effettivamente accetta in qualsiasi momento abbia svuotato tutta la pila, con o senza pila vuota (quindi anche se la testina non ha, alla fine #). Se all'inizio la pila non dovesse contenere #, sappiamo che il simbolo non sarà quello sotto la testina e cominciamo ad eseguire pop/push; comunque ad un certo punto la pila accetta.

Similmente, ammettiamo che ogni macchina a nastro singolo accetti l'automa a coda.

In questo caso, semplicemente, si aggiunge un simbolo che indica la posizione della testina e un simbolo marcatore; il nastro si muove a destra/sinistra come indicato tra prima e seconda metà di nastro per eseguire pop/push. In entrambi i casi abbiamo un linguaggio Turing-riconoscibile.

16. Una Macchina di Turing a sola scrittura è una TM a nastro singolo che può modificare ogni cella del nastro al più una volta (inclusa la parte di input del nastro). Mostrare che questa variante di macchina di Turing è equivalente alla macchina di Turing standard.

Per prima cosa simuliamo una normale macchina di Turing con una macchina di Turing che scrive due volte. La macchina a scrittura doppia simula un singolo passaggio della macchina originale copiando l'intero nastro su una parte nuova del nastro sul lato destro della porzione attualmente utilizzata. La procedura di copiatura opera carattere per carattere, marcando un carattere mentre viene copiato. Questa procedura altera ogni quadrato del nastro due volte: una volta per scrivere il nastro per la prima volta, e ancora per sottolineare che lo è stato copiato. La posizione della testina del nastro della macchina di Turing originale è contrassegnata sul nastro.

Quando si copiano le celle in corrispondenza o vicino alla posizione contrassegnata, il contenuto del nastro è aggiornato secondo le regole della macchina di Turing originale. Per eseguire la simulazione con una macchina scrivibile una volta, si opera come prima, eccetto che ogni cella del nastro precedente è ora rappresentata da due celle. La prima di queste contiene il simbolo del nastro della macchina originale e il secondo è per il simbolo utilizzato nella procedura di copiatura. L'input non viene presentato alla macchina

nel formato con due celle per simbolo; quindi, la prima volta che il nastro viene copiato, i segni di copiatura sono messi direttamente sopra il simbolo originale

Più semplicemente:

Partendo dalla macchina originale, si copia l'intero nastro tra le transizioni. In questo nuovo nastro viene aggiunto un simbolo delimitatore (caso TM a nastro semi-infinito) ed un alfabeto espanso per registrare la posizione del nastro quando è stato copiato. In questo caso si hanno due scritture, la prima usata quando si copia il simbolo sul nuovo nastro, la seconda per indicare che sia stato copiato. Avendo le transizioni uguali, la TM a sola scrittura è equivalente a quelle standard.

17. Chiamiamo k -PDA un automa a pila dotato di k pile.

- (a) Mostrare che i 2-PDA sono più potenti degli 1-PDA
- (b) Mostrare che i 2-PDA riconoscono esattamente la classe dei linguaggi Turing-riconoscibili
- (c) Mostrare che i 3-PDA *non* sono più potenti dei 2-PDA

a) Si considera un linguaggio CF, che viene riconosciuto da un PDA.

Ad esempio vediamo $L = \{a^n b^n c^n \mid n \geq 0\}$

Usiamo il PL per linguaggi CF dimostrando che non si tratta di CFL.

Ad esempio, consideriamo, per $s=uvxyz$, una stringa $s = a^k b^k c^k$, $k > 0$

Si può dimostrare che eseguendo dei pumping, a causa delle due porzioni non vuote del linguaggio, esso rimane sempre sbilanciato nel numero di 0/1.

Questo non è un CFL e quindi lo 1-PDA non lo riconosce.

Tuttavia, si sfrutta l'idea della macchina multinastro, equivalente a quella a nastro singolo, con le singole pile. Prendendo quindi l'input:

- il PDA si trova nello stato iniziale e, quando viene letta una "a", viene fatto il push sulla prima pila e si avanza nello stato successivo. Se viene letta una b o una c, si rifiuta, perché si deve essere nel format "abc".
- dallo stato dopo, che chiamiamo q_2 se viene letta una "b", si esegue il pop di una "a" e si fa il push di "b" sulla seconda pila. Il numero deve corrispondere; se non rimangono "a" rispetto alle "b" rifiuta. Se si legge ancora "a", siccome qui dovremmo leggere "b", rifiuta
- allo stato dopo, q_3 , si legge l'input "c". Qui si fa come prima; se leggiamo "a" o "b", siccome qui dovremmo leggere "c", rifiuta, altrimenti leggiamo "c" e facciamo il pop di "b" dalla seconda pila
- accettiamo se entrambe le pile sono vuote altrimenti si rifiuta

Siccome il PDA a pila singola rifiuta mentre il 2-PDA accetta, allora i 2-PDA sono più potenti degli 1-PDA.

b) I 2-PDA devono essere descritti da una TM, ciò sarà fattibile con una TM multinastro. Infatti, semplicemente, andremo ad emulare le due pile su due nastri di una TM M.

M = Su input x:

- 1) usa due pile L ed R, che rappresentano il contenuto del nastro. L'idea è di separare queste due pile con un simbolo separatore, simulando l'avanzamento sul nastro.
- 2) le due pile rappresentano l'andamento della testina; da una parte si fa pop e dall'altra si fa push, in particolare rappresentano a sinistra la parte prima della testina e sulla parte destra la parte successiva.
- 3) quando riceviamo un input, dobbiamo spostarci sulla seconda pila, facendo push. L'idea è che, essendo una TM deterministica, esegua il push di tutti gli n simboli che ha, eseguendo poi pop. Se è possibile eseguire ciò, la computazione avanza, altrimenti rifiuta
- 4) dopo aver inserito tutti gli input, l'altro nastro ragiona al contrario, partendo da destra e verificando che i simboli inseriti siano nell'ordine contrario (comincerà così subito ad eliminare, perché la pila toglie il carattere in fondo).

La macchina comunque è libera di spostarsi tra la prima e la seconda pila; tuttavia, sappiamo che se raggiungiamo la pila vuota o al massimo un solo simbolo rimasto (blank), allora accetta, altrimenti rifiuta. In questo modo i 2-PDA sono effettivamente Turing-riconoscibili.

- c) L'idea è che entrambe siano multinastro e dunque sono della stessa potenza. Abbiamo appena visto che i 2-PDA sono simulabili da una TM come descritto sopra. L'idea è che il movimento del nastro sia equamente descritto in push (prima pila) e pop (seconda pila). Sapendo che abbiamo tre pile, semplicemente, l'idea sopra si estende ad una terza pila. Su due nastri implementiamo il movimento a destra e a sinistra; similmente, una macchina a 3 pile dovrà quantomeno implementare 2 pile che agiscono in questo modo. La TM considera quindi un nastro di input e altri 3 corrispondenti a 3 pile. Il push richiede che il simbolo venga pushato da un input e tutte e tre le pile si aggiornino con il corrispondente simbolo; come detto prima, una farà il push, una farà il pop, l'altra aggiungerà il simbolo stringa ed un blank, per indicare che è avvenuta una transizione della pila. Qualora venga fatta una delle due operazioni, uno dei nastri viene implementato come copia delle altre, seguendo le transizioni della pila. Siccome la TM agisce in parallelo, essendo qui nondeterministica (dato che si può estendere fino a k), possiamo avere k nastri, eseguendo un push, un pop ed uno scorrimento del nastro in un ordine casuale. Non siamo limitati dal numero del nastro; siamo noi che decidiamo l'ordine delle singole azioni. Dunque, si può intuire che i 3-PDA non siano più potenti dei 2-PDA.

18. Sia $ALL_{DFA} = \{\langle A \rangle \mid A \text{ è un DFA e } L(A) = \Sigma^*\}$. Mostrare che ALL_{DFA} è decidibile.

L'idea è di definire una TM M tale che accetti la stringa A e stia nel linguaggio. Essendo che ha un solo stato, semplicemente, esso sarà contemporaneamente lo stato iniziale e finale.

Immaginando quindi di avere una TM la simulazione procede in questo modo:

- Simula A sul nastro segnando lo stato iniziale
- Dato che è l'unico stato, la macchina avanza e si aspetta di non segnare nessun altro stato ad eccezione di quelli che vanno verso lo stato iniziale, che saranno marcati
- Se la macchina non ha segnato altri stati, *accetta*, altrimenti *rifiuta*

Poste queste condizioni, il linguaggio può dirsi *decidibile*.

19. Sia $A_{\epsilon CFG} = \{\langle G \rangle \mid G \text{ è una CFG che genera } \epsilon\}$. Mostrare che $A_{\epsilon CFG}$ è decidibile.

In questo caso abbiamo una CFG che produce una stringa vuota. Optiamo per considerare invece una grammatica che appunto deriva ϵ :

- se la CFG deriva ϵ correttamente allora *accetta*
- se la CFG non deriva ϵ rifiuta

Definiamo M su input $\langle G \rangle$, derivando quindi una codifica di una CFG che da un punto di vista costruttivo, sarà G' ed è in forma normale di Chomsky. Questa grammatica accetta tutte le regole di G , come tale se G possiede una regola del tipo $S \rightarrow \epsilon$, allora anche G' avrà una regola del tipo $S' \rightarrow \epsilon$.

Dal punto di vista di una TM M che riconosce questa grammatica:

- marca lo stato iniziale
- si muove sul nastro seguendo le transizioni della CFG
- se essa arriva con le derivazioni ad una stringa vuota (quindi contiene una regola $S \rightarrow \epsilon$), allora *accetta*

Poste queste condizioni, il linguaggio può dirsi *decidibile*.

20. Sia $S_{REX} = \{\langle R, S \rangle \mid R, S \text{ sono espressioni regolari tali che } L(R) \subseteq L(S)\}$. Mostrare che S_{REX} è decidibile.

Sono entrambe due espressioni regolari tali che una contenga l'altra come sottoinsieme.

A livello pratico significa che una porzione di linguaggio di R è contenuta in S ; quindi, entrambe le espressioni hanno almeno una parte dello stesso linguaggio.

Si può quindi idealmente ragionare con una TM N che decide S_{REX} in questo modo:

- trasforma R ed S in due ϵ -NFA equivalenti B e C
- si esegue N sugli input di R ed S
- se scorrendo tutto il linguaggio di R si arriva, marcando tutti i suoi stati, a coprire almeno una parte di stati di S , anche essi marcati allora *accetta*, altrimenti *rifiuta*

Conviene quindi avere N multinastro, tale da avere gli input da una parte, controllando R su un nastro e vedere sul risultante terzo nastro dove starebbe S se vi è un match del pattern precedente.

Poste queste condizioni, il linguaggio può dirsi *decidibile*.

21. Sia $X = \{\langle M, w \rangle \mid M \text{ è una TM a nastro singolo che non modifica la porzione di nastro che contiene l'input } w\}$. X è decidibile? Dimostrare la vostra risposta.

Supponiamo che X sia decidibile e che esista una riduzione computabile.

Per esempio, immaginiamo una funzione di riduzione f che sulla seguente TM usa A_{TM} come riducibile:

$F = \text{"Su input } \langle M, w \rangle \text{ dove } M \text{ è una TM e } w \text{ è una stringa}$

1) Si costruisce la seguente TM M :

$M = \text{"Su input } x\text{"}$:

- Se $x = w$, allora modifica la porzione di nastro che lo contiene ed esegue M
- Se $x \neq w$, allora non modifica la porzione di nastro e va in loop
- Se M accetta, allora *accetta*
- Se M rifiuta, allora *rifiuta*

2) Si restituisce $\langle M' \rangle$

Dimostriamo quindi che f è una funzione di riduzione da A_{TM} a X :

- Se $\langle M, w \rangle \in A_{TM}$ allora la computazione di M modifica il nastro come voluto dalla funzione complemento. Ciò però rappresenta l'opposto di ciò che stiamo cercando; pertanto $f(\langle M, w \rangle) = \langle M' \rangle$ non appartiene ad X , perché modifica il nastro
- Se $\langle M, w \rangle$ non appartiene ad A_{TM} allora la macchina non modifica il nastro come vorrebbe correttamente A_{TM} , tuttavia ciò non va più bene per A_{TM} , che abbiamo detto accetta solo ma modifica del nastro.

Pertanto, si ha che $A_{TM} \leq_m X$ e sappiamo che A_{TM} è indecidibile, pertanto anche X sarà indecidibile.

22. Sia $E_{TM} = \{\langle M \rangle \mid G \text{ è una TM tale che } L(M) = \emptyset\}$. Mostrare che $\overline{E_{TM}}$, il complemento di E_{TM} , è Turing-riconoscibile.

Assumiamo che G sia una TM che ammette un linguaggio Turing-riconoscibile.

A queste condizioni, notiamo anche che la macchina ammette un linguaggio vuoto, simile al test del vuoto, ma complementata. Ciò significa che accetterà tutte le stringhe presenti possibili presenti nel linguaggio, purché siano nell'ordine del linguaggio iniziale e siano complementate.

Se la macchina G però accetta almeno una stringa, la condizione richiesta non vale più, pertanto essa continuerà a scorrere il nastro marcando ogni simbolo terminale, fino a quando non vengono marcate nuove variabili; come per il test del vuoto, se la variabile iniziale non è marcata *accetta*, altrimenti *rifiuta*.

A questo punto si pensa di costruire M tale che agendo sull'input $\langle M \rangle$:

- segua l'ordine delle stringhe dell'altra macchina
- per garantire l'ordine, dopo la prima stringa non vengono più marcate stringhe
- sequenzialmente si scorrono tutte le stringhe della lista della TM M
- in questa macchina se si accetta almeno una stringa allora *accetta*, perché la variabile iniziata è stata marcata, complementalmente all'altra macchina, altrimenti *rifiuta*

Poste queste condizioni, il linguaggio può dirsi *decidibile*.

23. Mostrare che se A è Turing-riconoscibile e $A \leq_m \overline{A}$, allora A è decidibile.

Qui dobbiamo mostrare due cose:

- A è Turing-riconoscibile
- $A \leq_m \overline{A}$

Partiamo col mostrare che A è ridotto da \overline{A} ; questo perché, se ciò accade, \overline{A} è Turing-riconoscibile, allora anche A lo è. Essendo che sia il complementare che il normale linguaggio sono Turing-riconoscibili, allora A è decidibile.

Verificando il funzionamento di A , intesa con una TM M :

$M = \text{Su input } w$, dove w è una stringa:

- Eseguiamo M su x . Se tale stringa appartiene ad A , *accetta*
- Se tale stringa non appartiene ad A , *rifiuta*

Abbiamo quindi un decisore per A ; vogliamo quindi dimostrare che anche \bar{A} ha un decisore e che quindi accetti le stringhe opposte.

Come tale, costruiamo una TM W su input $\langle M, w \rangle$:

- esegue $\langle M, w \rangle$ sul nastro
- se trova un input \underline{w} , accetta
- altrimenti rifiuta

Ora siccome evidentemente: $n \in A \iff f(n) \in A'$

equivalentemente: $n \notin A \iff f(n) \notin A'$

Avendo dimostrato l'esistenza dei decisori per il complemento e anche per le stringhe normali, vediamo che se esiste una stringa appartenente ad A , equivalentemente la stringa non viene accettata da \bar{A} e vale anche il contrario. Pertanto, per ogni n , A è decidibile.

24. Sia A un linguaggio. Dimostrare che A è Turing-riconoscibile se e solo se esiste un linguaggio decidibile B tale che $A = \{x \mid \text{esiste } y \text{ tale che } \langle x, y \rangle \in B\}$.

A è un linguaggio definito da B , ma questo viene definito solo tramite alcune condizioni su A .

Partendo da A dobbiamo dimostrare che presenta una stringa x ; se questa stringa è presente nelle transizioni che la TM di riferimento scorrendo trova, allora accetta, altrimenti rifiuta.

B invece viene descritto partendo da una descrizione simile di A , ma ammette anche una stringa " y ".

Quindi oltre ad ammettere una TM di riferimento partendo dalle transizioni percorse per " x ", similmente avremo n passi per " y ". La " y " esiste a condizione di " x ", pertanto se la TM accetta entrambe le stringhe, accetterà il linguaggio. Essendo A un riconoscitore, di sicuro accetterà almeno " x "; la condizione di accettazione su " y " poco cambia per definizione rispetto alla TM precedente. Se $x \in B$ allora accetterà " x " con un numero sufficientemente lungo di passi tali da poter accettare anche " y ", pertanto se $x \notin B$ allora $y \notin B$ per tutte le " y ".

25. $A \leq_m B$ e B è un linguaggio regolare implica che A è un linguaggio regolare? Perché sì o perché no?

Se $A \leq_m B$ e B è un linguaggio regolare, dobbiamo verificare se A possa essere anch'esso un linguaggio regolare. Però semplicemente si potrebbe avere il caso in cui B comprende già tutte le stringhe di A ed A sia non regolare. A tale scopo facciamo l'esempio di avere $A = \{0^n 1^n \mid n \geq 0\}$ e $B = \{1\}$. Dato che sono entrambi decidibili, si ha semplicemente la costruzione per entrambi di una TM che simula sul nastro sulla stringa w ed accetta qualora riconosca esattamente il linguaggio posto da A e da B . Notiamo inoltre che B riconosce tutte le stringhe di A e B è regolare; A invece non lo è. Si dimostra quindi che B risolve tutti i problemi di A , ma ciò non implica che A sia necessariamente regolare.

(Il 26 è uguale al 23, si veda quello per riferimento)

27. Sia $J = \{w \mid w = 0x \text{ per qualche } x \in A_{TM} \text{ oppure } w = 1y \text{ per qualche } y \in \bar{A}_{TM}\}$. Mostrare che sia J che \bar{J} non sono Turing-riconoscibili.

Immaginiamo di avere una TM che riceve in input una stringa w tale che sia un simbolo generico oppure $0x$. A queste condizioni possiamo immaginare una macchina $\langle M, w \rangle$ che, attraverso una simulazione, scorre tutto il nastro e trova w , in modo tale che accetti quando la trova, rifiuta altrimenti. A livello di output, si constata che $\langle M, w \rangle \in A_{TM}$, sempre appartenente al linguaggio.

Se è definita la stringa w sul linguaggio A_{TM} similmente è definito il complemento; è facile dimostrare che con una simulazione che agisce esattamente all'opposto avremmo che $\langle M, w \rangle \in \bar{A}_{TM}$, sempre appartenente al linguaggio.

L'altro caso invece considera una TM che scrive 1 seguita da una stringa $\in \langle M, w \rangle$. L'idea è totalmente uguale a prima, come descritto, pertanto la funzione di input considera sempre $\langle M, w \rangle \in A_{TM} \rightarrow \text{output di } R \in J$. Dato che $A_{TM} \leq_m J$, similmente abbiamo $\bar{A}_{TM} \leq_m \bar{J}$.

Essi non possono essere Turing riconoscibili in quanto già A_{TM} non è riconoscitore (solamente in un caso accetterà esattamente la stringa w , che deve essere $0x$ e si ferma. Ciò è molto limitato da un punto di vista applicativo; il complemento ragiona nella maniera contrapposta). Tanto ci basta per indicare che J e \bar{J} non sono Turing-riconoscibili.

(Il 28 è uguale a questo appena fatto, si veda questo per riconoscimento)

29. Un circuito Hamiltoniano in un grafo G è un ciclo che attraversa ogni vertice di G esattamente una volta. Stabilire se un grafo contiene un circuito Hamiltoniano è un problema NP-completo.

- (a) Un *circuito Toniano* in un grafo G è un ciclo che attraversa *almeno la metà* dei vertici del grafo (senza ripetere vertici). Il *problema del circuito Toniano* è il problema di stabilire se un grafo contiene un circuito quasi Hamiltoniano. Dimostrare che il problema è NP-completo.
- (b) Un *circuito quasi Hamiltoniano* in un grafo G è un ciclo che attraversa esattamente una volta tutti i vertici del grafo *tranne uno*. Il *problema del circuito quasi Hamiltoniano* è il problema di stabilire se un grafo contiene un circuito quasi Hamiltoniano. Dimostrare che il problema del circuito quasi Hamiltoniano è NP-completo.

a) Per mostrare che il problema del circuito Toniano è un problema NP-Completo, dobbiamo dimostrare che è NP e tutti i problemi NP sono riducibili in tempo polinomiale a questo problema.

Partiamo in modo semplice: la costruzione del certificato.

Per questo, ci serviranno i vertici, perché abbiamo bisogno di almeno la metà dei vertici, e:

$N = \text{Su input } \langle G, s, t \rangle$ dove G è un grafo quasi-hamiltoniano, " s " è un insieme di nodi e t è il certificato:

- prendiamo una lista di numeri, che rappresenta i nodi del grafo
- si controlla per le ripetizioni nella lista dei vertici, se ne viene trovata una si rifiuta
- si considera che i nodi s attraversino almeno $(p_m - p_i)/2$ vertici, in quanto vogliamo sapere se copriamo almeno la metà di questi
- per ciascun arco tra 1 ed $m/2$ si controlla se $\langle p_i, p_{i+1} \rangle$ induttivamente sia un arco di G . Se ciò accade, tutti i test sono passati

La verifica sui vertici viene fatta in tempo polinomiale, l'assegnazione dei vertici può potenzialmente impiegare un tempo più che polinomiale. Dunque, il problema è in NP.

Per completare la prova, dobbiamo dimostrare che è NP-Hard, usando la riduzione di un altro problema NP-completo, in questo caso, sfruttando la riduzione per mezzo del problema del circuito Hamiltoniano. Partendo da G grafo arbitrario, ci creiamo un grafo H che contiene due copie di G , senza archi in mezzo, chiamate G_1 e G_2 .

Voglio dimostrare che G ha un circuito Hamiltoniano se e solo se H ha un circuito Toniano.

- Supponendo G abbia un ciclo hamiltoniano. A queste condizioni G_1 avrà un ciclo contenente almeno la metà dei vertici, pertanto si ha un circuito toniano.
- Nei due sottografi G_1 e G_2 , avendo un circuito toniano per ciascuno, si sa che il circuito hamiltoniano coinvolgerà tutti i vertici del sottografo, tale che ciascuno contenga la metà esatta di tutti i vertici presenti. In poche parole, si ha un circuito Toniano solo a condizione che ve ne sia uno hamiltoniano. Dato inoltre che G_1 e G_2 sono copie dei vertici originali, semplicemente, anche G è circuito Hamiltoniano

Questa volta, si parla di vertici e cicli. Esiste quindi un algoritmo in tempo polinomiale che lo riduce, in questo caso, provando tutte le combinazioni per accoppiare tutti i vertici nel modo descritto (brute force). Dato quindi l'insieme di vertici di partenza, la riduzione è corretta, avendo che l'assegnazione di vertici e archi cresce di un fattore costante, nonostante la loro ricerca sia impiegata in tempo lineare.

A queste condizioni, il problema del circuito Toniano è NP-Completo.

b) Abbiamo dimostrato in precedenza che il circuito Hamiltoniano è in NP. Dato che in uno quasi-Hamiltoniano è la stessa cosa meno un vertice, riscriviamo estensivamente la dimostrazione articolando:

$N = \text{Su input } \langle G, s, t \rangle$ dove G è un grafo quasi-hamiltoniano, " s " è un insieme di nodi e t è il certificato:

- prendiamo una lista di numeri, che rappresenta i nodi del grafo
- si controlla per le ripetizioni nella lista dei vertici, se ne viene trovata una si rifiuta

- si considera che s parta dal primo numero p_1 e t vada fino a $p_m - 1$, dato che non considero l'ultimo vertice nell'attraversamento degli archi
- per ciascun arco tra 1 ed $m - 2$ si controlla se $\langle p_i, p_{i+1} \rangle$ induttivamente sia un arco di G . Se ciò accade, tutti i test sono passati

Il ciclo che si genera non include nessun nuovo vertice, infatti ognuno avrà almeno grado 1.

Non sappiamo per certo in quanto tempo polinomiale accada; alcuni sono confronti lineari e sappiamo viene verificato in un tempo al più polinomiale; la selezione stessa potenzialmente è un problema non deterministico. Pertanto, il problema è in NP.

Per completare la prova, dobbiamo dimostrare che è NP-Hard, usando la riduzione di un altro problema NP-completo, in questo caso, sfruttando la riduzione per mezzo del problema del circuito Hamiltoniano. Dobbiamo, come per il caso precedente, partire da un grafo che chiamiamo G .

Per ogni insieme di vertici, si aggiunge almeno un vertice per ogni iterazione in G

Voglio dimostrare che G ha un circuito hamiltoniano se e solo se H ha un circuito quasi-hamiltoniano.

- Supponiamo che G abbia un circuito hamiltoniano. Aggiungo un vertice senza collegargli nessun arco. Il sottografo certamente contiene un circuito hamiltoniano e raggiungiamo tutti i vertici meno uno. Questo vertice verrà aggiunto successivamente in un ciclo da H , tale che esso colleghi tutti i vertici e sia considerabile sottoinsieme del precedente (quindi H contiene un ciclo con tutti i vertici considerando anche il vertice tralasciato da G).
- A queste condizioni, G contiene per forza tutti i vertici ed H contiene a sua volta tutta una serie di vertici più quello che non fa parte del ciclo di G . Le due condizioni, come richiesto, vanno di pari passo, affinché G sia hamiltoniano se e solo se H è quasi hamiltoniano.

Dato inoltre che H è praticamente una copia di G , ma solo dei vertici utili (compreso il vertice mancante per G), si ha la correttezza della prova. Inoltre, affermiamo che dato l'insieme di vertici di partenza, la riduzione è corretta, avendo che l'assegnazione di vertici e archi cresce di un fattore costante, nonostante la loro ricerca sia impiegata in tempo lineare. Quindi, anche la riduzione è corretta, dato che il vertice è isolato. Alla luce di tutti questi fatti, anche il problema del circuito quasi-Hamiltoniano è NP-Completo.

30. Considera i seguenti problemi:

$\text{SETPARTITIONING} = \{ \langle S \rangle \mid S \text{ è un insieme di numeri interi che può essere suddiviso in due sottoinsiemi disgiunti } S_1, S_2 \text{ tali che la somma dei numeri in } S_1 \text{ è uguale alla somma dei numeri in } S_2 \}$

$\text{SUBSETSUM} = \{ \langle S, t \rangle \mid S \text{ è un insieme di numeri interi, ed esiste } S' \subseteq S \text{ tale che la somma dei numeri in } S' \text{ è uguale a } t \}$

- Mostra che entrambi i problemi sono in NP.
- Mostra che SETPARTITIONING è NP-Hard usando SUBSETSUM come problema di riferimento.
- Mostra che SUBSETSUM è NP-Hard usando SETPARTITIONING come problema di riferimento.

a) La dimostrazione di un verificatore V per SUBSET-SUM è articolata in questo modo:

Il certificato è il sottoinsieme stesso, dato che abbiamo un insieme di numeri e dobbiamo poi capire se appartengono all'insieme dei numeri completo.

L'idea di verificatore in tempo polinomiale è la seguente per V :

$V = \text{Su input } \langle \langle S, t \rangle, c \rangle$, tale che c è una collezione di numeri (certificato), S è l'insieme di tutti i numeri, c è il certificato:

- Controlla se c è una collezione di numeri la cui somma è t
- Controlla se la somma S contenga tutti i numeri
- Se entrambi passano, *accetta*, altrimenti *rifiuta*

Non sappiamo per certo in quanto tempo polinomiale accada; la selezione stessa sugli insiemi e la determinazione dei due sottoinsiemi verificatori potenzialmente è un problema non deterministico.

Quindi SUBSET-SUM è un problema in NP.

La dimostrazione di un verificatore V per SETPARTITIONING è articolata in questo modo:

Il certificato è il sottoinsieme stesso, dato che abbiamo un insieme di numeri e dobbiamo poi capire se appartengono all'insieme dei numeri interi S .

L'idea di verificatore in tempo polinomiale è la seguente per V :

$V = \text{Su input } \langle \langle S \rangle, c \rangle$, tale che c è una collezione di numeri (certificato), S è l'insieme di tutti i numeri:

- Controlla se c è una collezione di numeri la cui somma è t
- Controlla se la somma S contenga tutti i numeri di S_1 ed S_2
- Se entrambi passano, *accetta*, altrimenti *rifiuta*

La verifica è operata in tempo polinomiale, quindi SETPARTITIONING è un problema in NP.

b) Per rappresentare la struttura di SETPARTITIONING/SP come problema NP-Hard, dobbiamo provare che tutti i linguaggi in NP sono riducibili in tempo polinomiale con SUBSET-SUM/SS. Dobbiamo quindi sfruttare una struttura che permetta di rappresentare il problema. Verificando il problema in sé, SS rappresenta una singola istanza di SP, che invece considera due sottoinsiemi (SS ne considera uno solo), tali che la somma dei numeri in S_1 sia uguale alla somma dei numeri in S_2 .

Considerando che abbiamo due insiemi, necessariamente S_1 compone una metà di questa somma; dunque per ottenere la precedente basterà sfruttare come numero $t=1/2$ tale che tutti i numeri " x " che sono in S sommino tutti allo stesso valore.

Voglio quindi dimostrare che $SP \leq_p SS$

$\langle S \rangle \rightarrow \langle S', t \rangle$

Se esiste quindi un insieme di numeri in S che somma a t , allora i rimanenti numeri in S sommano a " $s - t$ ".

Pertanto, esiste una partizione in grado di sommare ad " $s - t$ ".

Sapendo che S_1 rappresenta $1/2$ di questa partizione, prendiamo due insiemi tali che la somma sia " $s - t$ ".

Pertanto, uno dei due insiemi in SP contiene almeno $s - 2t$ numeri. Rimuovendo questo numero, otterremo un set di numeri la cui somma è t e sono tutti in S (quindi una partizione è almeno la metà della seconda e si arriva allo stesso numero). Matematicamente avremo:

$$\text{Somma}(S_1) + \text{Somma}(S_2) = \text{Somma}(S)$$

che significa che

$$\text{Somma}(S_1) + \text{Somma}(S_2) = (1/2) * \text{Somma}(S)$$

L'idea quindi è che una partizione sia metà dell'altra, infatti in SP ci sono due insiemi uguali la cui somma è t , si può dire che $2S - w \in S$. Ad ogni nuovo elemento aggiunto, solo uno di questi due insiemi conterrà (dato che una è la metà dell'altra), la nuova stringa; poi, dato che nell'altro non vengono aggiunti valori, anche per SP la somma sarà sempre uguale a t .

Se ciò accade, quindi, si genera una partizione metà dell'altra e a questa aggiungo numeri che rimangono sempre nell'insieme per somma; sulla base di questa condizione si ritorna SI, altrimenti si ritorna NO. La riduzione è corretta, in quanto si considerano per istanze buone, SP rappresenta un insieme solo di SS e andrà bene per forza. Dato che la ricerca dell'insieme impiega potenzialmente un tempo più che polinomiale per ricerca continua (bruteforce), siamo in NP ed il problema è NP-Completo.

c) Per rappresentare la struttura di SUBSET-SUM/SS come problema NP-Hard, dobbiamo provare che tutti i linguaggi in NP sono riducibili in tempo polinomiale con SETPARTITIONING/SP.

Quindi, abbiamo a che fare:

- per SS un unico insieme tale che la somma sia t
- per SP due insiemi tali che la somma complessiva del primo sia uguale a quella del secondo

Per rappresentare SS prendiamo quindi due insiemi e affermiamo che per entrambi, la somma deve essere uguale a t . Definiamo quindi per l'insieme S_1 un insieme S_{new} a $2t - s$, dove s rappresenta la somma degli elementi di S_1 . In questo modo, S_1 e S_{new} rappresentano due partizioni che sommano a t .

Usando la riduzione, se S_{new} può essere di volta in volta partizionato in una serie di insiemi che contengono $2t - s$, significa che avremo sempre che S_{new} è la somma di tutti i numeri interi delle due partizioni che contiene tutti i numeri delle due sottopartizioni.

Quindi, le istanze buone, intese come somme delle sottopartizioni, vengono sommate ad una sola (SS), ma è tale anche che la somma delle due sia spezzabile in due partizioni separate (SS), che è quello che vogliamo realizzare in tempo P.

31. Considerate la seguente variante del problema SETPARTITIONING, che chiameremo QUASIPARTITIONING: dato un insieme di numeri interi S , stabilire se può essere suddiviso in due sottoinsiemi disgiunti S_1 e S_2 tali che la somma dei numeri in S_1 è uguale alla somma dei numeri in S_2 meno 1. Dimostrare che il problema QUASIPARTITIONING è NP-completo.

- (a) Dimostrare che il problema QUASIPARTITIONING è in NP fornendo un certificato per il S_i che si può verificare in tempo polinomiale.

Il certificato è dato da una coppia di insiemi di numeri interi S_1, S_2 . Per verificarlo occorre controllare che rispetti le seguenti condizioni:

- i due insiemi S_1, S_2 devono essere una partizione dell'insieme S ;
- la somma degli elementi in S_1 deve essere uguale alla somma degli elementi in S_2 meno 1.

Entrambe le condizioni si possono verificare in tempo polinomiale.

Dimostrare che il problema QUASIPARTITIONING è NP-hard, mostrando come si può risolvere il problema SETPARTITIONING usando il problema QUASIPARTITIONING come sottoprocedura.

Dimostrare che QUASIPARTITIONING è NP-hard, usando SETPARTITIONING come problema di riferimento richiede diversi passaggi:

1. Descrivere un algoritmo per risolvere SETPARTITIONING usando QUASIPARTITIONING come subroutine. Questo algoritmo avrà la seguente forma: data un'istanza di SETPARTITIONING, trasformala in un'istanza di QUASIPARTITIONING, quindi chiama l'algoritmo magico black-box per QUASIPARTITIONING.
2. Dimostrare che la riduzione è corretta. Ciò richiede sempre due passaggi separati, che di solito hanno la seguente forma:
 - Dimostrare che l'algoritmo trasforma istanze "buone" di SETPARTITIONING in istanze "buone" di QUASIPARTITIONING.
 - Dimostrare che se la trasformazione produce un'istanza "buona" di QUASIPARTITIONING, allora era partita da un'istanza "buona" di SETPARTITIONING.

3. Mostrare che la riduzione funziona in tempo polinomiale, a meno della chiamata (o delle chiamate) all'algoritmo magico black-box per QUASIPARTITIONING. (Questo di solito è banale.)

Una istanza di SETPARTITIONING è data da un insieme S di numeri interi da suddividere in due. Una istanza di QUASIPARTITIONING è data anch'essa da un insieme di numeri interi S' . Quindi la riduzione deve trasformare un insieme di numeri S input di SETPARTITIONING in un altro insieme di numeri S' che diventerà l'input per la black-box che risolve QUASIPARTITIONING.

Come primo tentativo usiamo una riduzione che crea S' aggiungendo un nuovo elemento a S di valore 1, e proviamo a dimostrare che la riduzione è corretta:

- ⇒ sia S un'istanza buona di SETPARTITIONING. Allora è possibile partizionare S in due sottoinsiemi S_1 ed S_2 tali che la somma dei numeri in S_1 è uguale alla somma dei numeri in S_2 . Se aggiungiamo il nuovo valore 1 ad S_1 otteniamo una soluzione per QUASIPARTITIONING, e abbiamo dimostrato che S' è una istanza buona di QUASIPARTITIONING.
- ⇐ sia S' un'istanza buona di QUASIPARTITIONING. Allora è possibile partizionare S' in due sottoinsiemi S_1 ed S_2 tali che la somma dei numeri in S_1 è uguale alla somma dei numeri in S_2 meno 1. Controlliamo quale dei due sottoinsiemi contiene il nuovo elemento 1 aggiunto dalla riduzione:
- se $1 \in S_1$, allora se tolgo 1 da S_1 la somma degli elementi S_1 diventa uguale alla somma dei numeri in S_2 . Abbiamo trovato una soluzione per SETPARTITIONING con input S .
 - se $1 \in S_2$, allora se tolgo 1 da S_2 quello che succede è che la somma degli elementi S_1 diventa uguale alla somma dei numeri in S_2 meno 2. In questo caso abbiamo un

problema perché quello che otteniamo *non è una soluzione di SETPARTITIONING!*

Quindi il primo tentativo di riduzione non funziona: ci sono dei casi in cui istanze cattive di SETPARTITIONING diventano istanze buone di QUASIPARTITIONING: per esempio, l'insieme $S = \{2, 4\}$, che non ha soluzione per SETPARTITIONING, diventa $S' = \{2, 4, 1\}$ dopo l'aggiunta dell'1, che ha soluzione per QUASIPARTITIONING: basta dividerlo in $S_1 = \{4\}$ e $S_2 = \{2, 1\}$.

Dobbiamo quindi trovare un modo per “forzare” l'elemento 1 aggiuntivo ad appartenere ad S_1 nella soluzione di QUASIPARTITIONING. Per far questo basta modificare la riduzione in modo che S' contenga tutti gli elementi di S *moltiplicati per 3*, oltre all'1 aggiuntivo. Formalmente:

$$S' = \{3x \mid x \in S\} \cup \{1\}.$$

Proviamo a dimostrare che la nuova riduzione è corretta:

\Rightarrow sia S un'istanza buona di SETPARTITIONING. Allora è possibile partizionare S in due sottoinsiemi S_1 ed S_2 tali che la somma dei numeri in S_1 è uguale alla somma dei numeri in S_2 . Se moltiplichiamo per 3 gli elementi di S_1 ed S_2 , ed aggiungiamo il nuovo valore 1 ad S_1 otteniamo una soluzione per QUASIPARTITIONING, e abbiamo dimostrato che S' è una istanza buona di QUASIPARTITIONING.

\Leftarrow sia S' un'istanza buona di QUASIPARTITIONING. Allora è possibile partizionare S' in due sottoinsiemi S_1 ed S_2 tali che la somma dei numeri in S_1 è uguale alla somma dei numeri in S_2 meno 1. Vediamo adesso che, a differenza della riduzione precedente, non è possibile che $1 \in S_2$: se così fosse, allora se tolgo 1 da S_2 quello che succede è che la somma degli elementi S_1 diventa uguale alla somma dei numeri in S_2 meno 2. Tuttavia, gli elementi che stanno in S_1 ed S_2 sono tutti quanti multipli di 3 (tranne l'1 aggiuntivo). Non è possibile che due insiemi che contengono solo multipli di 3 abbiano differenza 2. Quindi l'1 aggiuntivo non può appartenere a S_2 e deve appartenere per forza a S_1 . Come visto prima, se tolgo 1 da S_1 la somma degli elementi S_1 diventa uguale alla somma dei numeri in S_2 . Abbiamo trovato una soluzione per SETPARTITIONING con input S .

In questo caso la riduzione è corretta. Per completare la dimostrazione basta osservare che per costruire S' dobbiamo moltiplicare per 3 gli n elementi di S ed aggiungere un nuovo elemento. Tutte operazioni che si fanno in tempo polinomiale.

32. “Colorare” i vertici di un grafo significa assegnare etichette, tradizionalmente chiamate “colori”, ai vertici del grafo in modo tale che nessuna coppia di vertici adiacenti condivida lo stesso colore. Il problema k -COLOR è il problema di trovare una colorazione di un grafo non orientato usando k colori diversi.

- (a) Mostrare che il problema k -COLOR è in NP per ogni valore di k
- (b) Mostrare che 2-COLOR è in P
- (c) Mostrare che $3\text{-COLOR} \leq_P k\text{-COLOR}$ per ogni $k > 3$
- (d) Considerate la seguente variante del problema, che chiameremo ALMOST3COLOR: dato un grafo non orientato G con n vertici, stabilire se è possibile colorare i vertici di G con tre colori diversi, in modo che ci siano al più $n/2$ coppie di vertici adiacenti dello stesso colore. Dimostrare che il problema ALMOST3COLOR è NP-completo.

a) Per dimostrare che il problema è in NP va dimostrato che è possibile avere un verificatore del problema che lavora in tempo polinomiale. Questo è possibile perché dato un grafo di dimensioni finite, basterà passare tutti i nodi del grafo e comparare il loro colore a quello dei nodi adiacenti. Se per nessun nodo il colore sarà lo stesso dei nodi adiacenti, allora il verificatore accetterà, altrimenti rifiuterà. Quindi, avendo un grafo $G(V,E)$ con una serie di colori $\{c_1, c_2, \dots, c_k\}$ con ciascuno assegnato ad un rispettivo colore, per ogni arco $\{u, v\}$ nel grafico G si verifica che il colore $c(u) \neq c(v)$. Quindi il tempo di esecuzione sarà dato da $(V + E) \cdot k$ il quale sarà k come upper bound, perché si aggiunge sempre un vertice a quelli originali. Quindi, il problema si estende ad un limite superiore di 2 (2-Color), 3 (3-Color) e seguenti facilmente.

b) Creiamo un algoritmo che esegue in tempo polinomiale.
 Supponiamo che i nostri due colori siano il rosso e il blu. L'idea chiave è quella di colorare, se possibile, ogni componente connesso del grafico. Scegliamo un vertice v_0 nel primo componente connesso e lo coloriamo di rosso. Se v_0 è collegato a v_1, v_2, \dots, v_k , si colora questi altri vertici di blu e si ripete. Una volta colorato un vertice, si ha una sola opzione per il colore dei suoi vicini, basta solo che abbia un colore diverso; quindi o si colorano tutti i vertici diversamente in tempo polinomiale 2 a 2 oppure si ha un conflitto, quindi due colori uguali; se ciò accade, l'output è "NO". Se l'algoritmo colora correttamente tutti i componenti collegati, l'output sarà "Sì". Partendo dal grafico di input $G(V,E)$, dimostriamo che ci si impiega un tempo quadratico, come detto a due a due per i vertici, e quindi il tempo impiegato è $O(|V|^2)$.

c) Si vuole estendere il problema 3-Color a k -color. A queste condizioni, partendo da G si aggiunge una serie di vertici ad H , tale che tutti gli archi vengano così collegati.

Vogliamo quindi provare che G sia 3-colorabile se e solo se H è k -colorabile.

- Supponendo che G sia 3-colorabile, decidiamo di assegnare ad ogni vertice dei colori arbitrari, ad esempio, “rosso”, “giallo” e “blu”. Ora, abbiamo per ogni nuovo vertice, progressivamente aggiunto, un nuovo colore, possibile nel collegamento dei vertici, tale da arrivare a k .

L'idea di base dice che:

- se abbiamo due vertici, u e v con colori diversi, allora li colleghiamo
- se non hanno due colori diversi, si ha in conflitto

Dato che ogni vertice avrà quantomeno un nuovo colore (perché come detto dobbiamo arrivare a k), allora H sarà H -colorabile, dato che si aggiunge un nuovo vertice con un colore ogni volta per G .

Dimostriamo anche che se H è k -colorabile, allora G è 3-colorabile.

Abbiamo una k -colorazione, quindi ogni arco composto da due vertici contiene una coppia di colori diversi; ogni vertice a loro adiacente, a loro volta, conterrà almeno un colore diverso. Pertanto, partendo da k , possiamo progressivamente eliminare un vertice e, così facendo, verrà tolto almeno un colore. Così si procede fino ad arrivare ai 3 colori di partenza.

Dato sempre l'esame dei vertici condotto in tempo polinomiale partendo da G ed arrivando ad H , si ha un'applicazione brute-force, quindi applicata in tempo polinomiale.

d) Per dimostrare che ALMOST3COLOR è NP-completo, si deve dimostrare che è sia NP che NP-HARD. Partendo da $N = \text{Su input } \langle G, s, t \rangle$ dove G è un grafo quasi-hamiltoniano, " s " è un insieme di nodi e t è il certificato:

- si verifica che ogni insieme di nodi sia collegato ai rispettivi archi
- si verifica che tra le coppie di vertici collegate ve ne siano dello stesso colore
- se abbiamo almeno $n/2$ coppie di colori sul problema, allora si ha in output *SI*, altrimenti *NO*.

Dato che la verifica viene fatta in tempo polinomiale, il problema è in NP.

Vogliamo ora dimostrare che si tratta di un problema NP-Hard, usando ALMOST3COLOR come istanza di 3COLOR, tramite una riduzione.

Sapremo quindi che 3COLOR possiede un numero n di vertici coperti da almeno $n/2$ vertici di ALMOST3COLOR; semplicemente, supponendo che per ogni vertice di ALMOST3COLOR si abbia un collegamento, progressivamente, si assegnano tutti i colori ad ogni coppia di vertici separata da un arco. Così facendo abbiamo colorato tutti i vertici di ALMOST3COLOR.

Semplicemente si nota che 3COLOR rappresenta esattamente il doppio dei vertici dell'altro; dunque, eseguendo tutta l'operazione fino a tutti i vertici raggiungibili, è facile dimostrare come 3COLOR risolva ALMOST3COLOR.

Avendo tutte le coppie collegate, sappiamo che 3COLOR contiene le sole istanze buone del precedente, in quanto sono tutte le coppie di vertici collegati, che rispetto a 3COLOR sarebbero correttamente $n/2$ coppie.

Il tempo richiesto è polinomiale e correttamente viene risolto il problema.

oppure

Per dimostrare che ALMOST3COLOR è NP-completo, si deve dimostrare che è sia NP che NP-HARD.

È NP perché esiste un certificato del sì che si può verificare in tempo polinomiale.

Il certificato è dato da una coppia di insiemi di vertici e di archi.

Per verificarlo, occorre controllare che:

1. Il numero di archi, i cui vertici sono di colore uguale, sia al più 8939
2. Il numero di colori usati in tutto il grafo sia 3

Entrambe le condizioni si possono verificare in tempo polinomiale.

È NP-HARD, perché si può mostrare come si può risolvere il problema 3COLOR usando il problema ALMOST3COLOR come sotto procedura.

1- ALGORITMO

Dato un grafo G (che è istanza di 3COLOR), costruiamo in tempo polinomiale un nuovo grafo, che è istanza di ALMOST3COLOR, aggiungendo $8939 \times (2 \text{ vertici dello stesso colore e } 1 \text{ arco che li collega})$. Il nuovo grafo contiene 8939 coppie di vertici uguali.

2- DIMOSTRAZIONE

=> Sia S un'istanza buona di 3COLOR.

Allora, è possibile aggiungere 8939 coppie di vertici uguali, collegati da un arco (2 a 2). La parte del grafo senza colori uguali adiacenti rimane intatta (S). Vengono soddisfatti 3COLOR e ALMOST3COLOR.

<= Sia S' un'istanza buona di ALMOST3COLOR. Una volta individuate tutte le coppie di vertici, si devono contare quali sono i colori adiacenti uguali. Dato che le coppie con colori uguali sono state aggiunte a quelle diverse, quelle rimanenti sono 3 COLOR.

3- TEMPI POLINOMIALI

Dato che la costruzione del nuovo grafo implica l'aggiunta di 8939 coppie di vertici (collegate da un arco), il tempo richiesto è polinomiale

Esercizi da vecchi file di seconda parte in preparazione (anni scorsi):

15. Sia A il linguaggio che contiene solo ed unicamente la stringa s ,

$$s = \begin{cases} 0 & \text{se la vita non sarà mai trovata su Marte} \\ 1 & \text{se un giorno la vita sarà trovata su Marte} \end{cases}$$

A è un linguaggio decidibile? Giustificare la risposta. Ai fini del problema, assumere che la questione se la vita sarà trovata su Marte ammetta una risposta non ambigua Sì/No.

Avendo un linguaggio che contiene una sola stringa, quindi 0/1.

Ciò è semplice da dimostrare con una TM; infatti la macchina o prende 0 oppure 1, altrimenti non accetta il proprio linguaggio. Per uno dei due è possibile costruire un decisore, che accetti il primo o il secondo, ovviamente non entrambi. Per quanto logicamente non si sappia la risposta certa, è comunque ammesso un decisore che accetterebbe due linguaggi finiti $\{0\}$ oppure $\{1\}$.

Dunque, A è decidibile.