Indice

Dato un linguaggio ci possiamo chiedere se questo sia CF. Ad esempio

$$L = \{a^l b^k c^j \mid k = j\} \overset{?}{\in} \text{CF}$$

Un linguaggio è CF se possiamo costruire una grammatica di tipo 2 o un automa a pila. Ad esempio per il linguaggio di sopra possiamo consumare tutte le a e controllare che il numero di b e di c sia uguale con una pila.

Prendiamo invece

$$L = \{a^i b^k c^j \mid i = j = j\} \overset{?}{\in} \text{CF}$$

L'automa per il linguaggio di prima non può essere adattato a questo linguaggio.

Prendiamo la grammatica

$$\begin{aligned} S &\to [S] \mid SS \mid T \\ T &\to (T) \mid TT \mid \epsilon \end{aligned}$$

e una derivazione

$$S \overset{*}{\Rightarrow} [()[]]$$

Ora un albero di derivazione che possiamo fare per questa stringa è sugli alberi di derivazione si possono fare operazione di sostituzione di sottoalberi, ad esempio nell'albero di sopra sostituenzo il sottoalbero 2 con il sottoalbero 1 otteniamo l'albero di derivazione per $[()[()]]$. In generale quello di sopra è un particolare albero che è rappresentato dalla derivazione

$$A \overset{*}{\Rightarrow} vAx, \qquad v, x \in \Sigma^*$$

questi sono interessanti perché possiamo ???. Ad esempio nell'albero prima abbiamo

$$S \overset{*}{\Rightarrow} [()S]$$

possiamo vedere che possiamo sia accorciare la derivazione In questo modo possiamo generare infinite stringhe.

Dal fatto che le grammatiche possono essere convertite in FNC (a patto di sacrificare la parola vuota), lavoreremo con grammatiche in FNC per semplificare la dimostrazione del pumping lemma.

Definiamo la *profondità* (o altezza) di un albero come il più lungo cammino dalla radice ad una foglia.

Lemma 1. *Sia*

$$G = \langle V, \Sigma, P, S \rangle$$

una grammatica in FNC e sia $T : A \stackrel{*}{\Rightarrow} w \in \Sigma^*$ *un albero di derivazione di altezza* h*. Allora la lunghezza di* w *è minore o uguale a* 2^{h-1}.

$$|w| \leq 2^{h-1}$$

Dimostrazione. Procediamo per induzione su h:

- per $h = 1$: per forza l'albero deve rappresentare una produzione della forma $A \to a \in \Sigma$, quindi $w = a$ e

$$|w| = 1 = 2^0 = 2^{1-1}$$

- la produzione applicata alla radice deve per forza essere della forma $A \to BC$, quindi l'albero si divide in due sottoalberi, un albero $T' : B \stackrel{*}{\Rightarrow} w'$ e un albero $T'' : C \stackrel{*}{\Rightarrow} w''$. Questi due hanno altezza minore o uguale ad $h - 1$. Ora applicando l'ipotesi induttiva

$$|w'| \leq 2^{h-2}$$
$$|w''| \leq 2^{h-2}$$

e

$$|w| = |w'| + |w''| \leq 2^{h-2} + 2^{h-2} = 2^{h-1}$$

■

Lemma 2 (Pumping lemma per CFL)**.** *Sia* L *un linguagio CF allora* $\exists N > 0$ *tale che* $\forall z \in L$ *con* $|z| \geq N$*, questa può essere scomposta*

$$z = uvwxy$$

tali che

1. $vwx| \leq N$
2. $vx \neq \epsilon$

3. $\forall i \geq 0 \mid uv^iwx^iy \in L$

Dimostrazione. Sia $G = \langle V, \Sigma, P, S\rangle$ una grammatica in FNC per $L \setminus \{\epsilon\}$. Sia $k = |V|$ e definiamo $N = 2^k$.

Sia $z \in L$ con $|z| \geq N$, allora ha un albero di derivazione $T : S \stackrel{*}{\Rightarrow} z$ visto che la lunghezza di z è maggiore di 2^k, allora dal lemma precedente abbiamo che l'altezza di T è almeno $k+1$, quindi esiste un cammino dalle foglie alle radici da $k+1$ archi, quindi $k+2$ nodi. Visto che l'ultimo nodo è un terminale, durante questo cammino incontreremo $k+1$ non terminali, e quindi almeno un non terminale si ripeterà in questo percorso. Sia A questo non terminale. Per questo non terminale A vale

$$A \stackrel{*}{\Rightarrow} w$$

$$A \stackrel{*}{\Rightarrow} vAx$$

$$S \stackrel{*}{\Rightarrow} uAy$$

è facile vedere che

$$S \stackrel{*}{\Rightarrow} uAy \stackrel{*}{\Rightarrow} uvAxy \stackrel{*}{\Rightarrow} \ldots \stackrel{*}{\Rightarrow} uv^iAx^iy \stackrel{*}{\Rightarrow} uv^iwx^iy$$

quindi abbiamo dimostrato il punto 3.

La produzione centrale di A deve essere per forza della forma $A \to BC$, supponiamo che C sia il non terminale sul percorso più lungo che genera wx, allora visto che siamo in FNC e non possiamo generare la parola vuota, allora per forza B genera qualcosa diverso da ϵ, quindi abbiamo dimostrato il punto 2.

L'altezza della parte dell'albero che genera vwx è al massimo $k+1$, cioè il numero massimo di nodi che possiamo vedere prima di trovare una ripetizione. Quindi utilizzando ancora il lemma di sopra, $|vwx| \leq N$. ■

Riprendiamo il linguaggio di prima

$$L = \{a^nb^nc^n \mid c \geq 0\}$$

mostriamo che non soddifsa il pumping lemma.
Supponiamo per assurdo che L sia CF e mostriamo che non può esistere una costante N per cui valga il pumping lemma. Sia N la costante di

L, prendiamo

$$z = a^N b^N c^N = uvwxy$$

Visto che per la prima condizione $|vwx| \leq N$, vwx potrà contenere solo due dei tre simboli, più precisamente $vwx \in a^*b^*$ o $vwx \in b^*c^*$. Supponiamo che $vwx \in a^*b^*$, prendiamo $i = 0$ e la stringa $z' = uwy$, questa per la condizione 3 dovrebbe essere in L. Calcoliamo ora le occorrenze dei simboli in z':

$$\begin{aligned} \#_c(z') &= N \\ \#_a(z') + \#_b(z') &= 2N - (\#_a(vx) + \#_b(vx)) \\ &\leq 2N \qquad \text{Per la condizione 2 } (\#_a(vx) + \#_b(vx)) \geq 1 \end{aligned}$$

e quindi $z' = a^k b^j c^N \notin L$ con $k, j < N$, quindi abbiamo un assurdo.

Prendiamo il linguaggio

$$L = \{ww \mid w \in \{a, b\}^*\}$$

questo non è CF.
Mostriamolo ancora attraverso il pumping lemma. Sia N la costante del pumping lemma, scegliamo la stringa

$$z = a^N b^N a^N b^N \in L = uvwxy$$

Utilizznado ancora la condizione 1 abbiamo due casi

- $vwx \in a^*b^*$, prendiamo la stringa $z' = uwx$, questa dovrebbe essere in L per la condizione 3. Questa è $z' = a^N b^{N'} a^{N''} b^N$, con $N' \leq N, N'' \leq N$, ora possono essere diminuite solo le a, solo le b o entrambe, ma in ogni caso $z' \notin L$.
- $vwx \in b^*a^*$, questo a sua volta dsi divide in due casi, in base al fatto che vwx sia nella prima o nella seconda parte della stringa. Prendendo ancora $i = 0$, $z' = a^{N'} b^{N''} a^N b^N$, con $N' \leq N$ e $N'' \leq N$, con ancora almeno uno tra N' e N'' minore o uguale a N.

Prendiamo il linguaggio

$$L = \{a^h b^j a^k \mid j = \max(h, k)\}$$

supponiamo sia CF e chiamiamo N la costante del pumping lemma. Prendiamo la stringa

$$z = a^N b^N a^N \in L = uvwxy$$

Anche qui ci sono due casi

- $vwx \in a^*b^*$, sappiamo che $vx \neq \epsilon$, distinguiamo tre casi
  - $vx \in a^+$, prendendo $i = 2$, otteniamo $z' = a^{N'} b^N a^N$, con $N' > N$, che non fa parte di L
  - $vx \in b^+$, prendendo $i = 0$, otteniamo $z' = a^N b^{N'} a^N$, con $N' < N$, che non fa parte di L
  - $vwx \in a^+b^+$, prendendo $i = 0$, otteniamo $z' = a^{N'} b^{N''} a^N$, con $N'' < N$, che non fa parte di L
- $vwx \in b^*a^*$, questo caso è simmetrco al precedente