Proprietà di Chiusura

Teorema I ling. liberi sono chiusi per

1) unione
2) concatenazione
3) ripetizione (stella di Kleene)

Dim. Sia $L_1 = L(G_1)$ con $G_1 = (NT_1, T_1, R_1, S_1)$
e $L_2 = L(G_2)$ con $G_2 = (NT_2, T_2, R_2, S_2)$
(assumiamo che $NT_1 \cap NT_2 = \emptyset$)

(1) Unione

$L_1 \cup L_2 = L(G)$ dove G è

$$(NT_1 \cup NT_2 \cup \{S\},\ T_1 \cup T_2,\ S,\ R_1 \cup R_2 \cup \{S \to S_1 | S_2\})$$

(2) Concatenazione

$L_1 \circ L_2 = L(G)$ dove G è

$$(NT_1 \cup NT_2 \cup \{S\},\ T_1 \cup T_2,\ S,\ R_1 \cup R_2 \cup \{S \to S_1 S_2\})$$

(3) Stella di Kleene

$(L_1)^* = L(G)$ dove G è

$$(NT_1 \cup \{S\},\ T_1,\ S,\ R_1 \cup \{S \to \varepsilon \mid S_1 S\})$$

<u>Teorema</u> L'intersezione $L_1 \cap L_2$ di un ling. libero L_1 con un ling. regolare L_2 è un ling. libero.

per stato finale

<u>Dim</u>: $L_1 = L[N_1]$ con $N_1 = (\Sigma, Q_{N_1}, \Gamma, \delta_{N_1}, q_{N_1}, Z, F_{N_1})$

e $L_2 = L[N_2]$ con $N_2 = (\Sigma, Q_{N_2}, \delta_{N_2}, q_{N_2}, F_{N_2})$

Costruiamo il PDA $N =$

$$(\Sigma, Q_{N_1} \times Q_{N_2}, \Gamma, \delta, (q_{N_1}, q_{N_2}), Z, F_{N_1} \times F_{N_2})$$

dove $\delta((q,p), a, X)$ è l'insieme di tutte le coppie

$((r,s), \gamma)$ tali che

- $s = \delta_{N_2}(p, a)$ e
- la coppia $(r, \gamma) \in \delta_{N_1}(q, a, X)$

cioè N segue le mosse di N_1, tenendo traccia dello stato raggiunto da N_2.

Si può dimostrare che

$$(q_{N_1}, w, Z) \vdash^*_{N_1} (q, \varepsilon, \gamma) \quad \text{e} \quad \hat{\delta}_{N_2}(q_{N_2}, w) = q'$$

sse

$$((q_{N_1}, q_{N_2}), w, Z) \vdash^*_{N} ((q, q'), \varepsilon, \gamma)$$

Se $q \in F_{N_1}$ e $q' \in F_{N_2}$, allora w è riconosciuta come stringa appartenente all'intersezione $L_1 \cap L_2$.

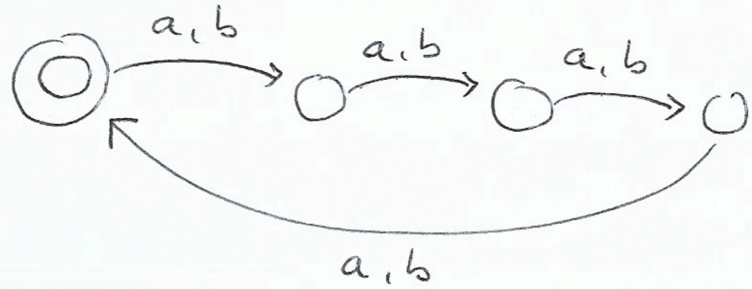
Esercizio

Mostrare che $L_1 \cap L_2$ con

$L_1 = \{a^n b^n \mid n \geq 0\}$ e $L_2 = \{w \in \{a,b\}^* \mid \exists k \in \mathbb{N}.\ |w| = 4k\}$

è un ling. libero.

- Applicando il teorema appena dimostrato, si ha che $L_1 \cap L_2$ è libero perché L_1 è libero ($S \to aSb \mid \varepsilon$) e L_2 è regolare



- $L_1 \cap L_2 = \{a^{2n} b^{2n} \mid n \geq 0\}$ ed è libero perché

$S \to aaSbb \mid \varepsilon$ lo genera

Applicazioni del teorema

$L = \{ww \mid w \in \{a,b\}^*\}$ è libero? Se lo fosse, allora anche $L \cap a^* b^* a^* b^* = \{a^n b^m a^n b^m \mid n,m \geq 0\}$ dovrebbe essere libero. Ma dimostreremo presto che questo ling. non è libero

$\Longrightarrow$ nemmeno L può essere libero

Oss: Vale che:

(1) Se $L_1 \cap L_2$ è non regolare, ma L_2 è regolare allora L_1 non regolare
(perché i ling. regolari sono chiusi per intersezione)

(2) Se $L_1 \cap L_2$ è non libero, ma L_2 è regolare allora L_1 non libero
(perché l'intersezione di un libero con un regolare dà un libero)

I ling. liberi non sono chiusi per intersezione

$L_1 = \{a^n b^n c^m \mid n, m \geq 0\}$ libero

$L_2 = \{a^n b^m c^m \mid n, m \geq 0\}$ libero

$L_1 \cap L_2 = \{a^n b^n c^n \mid n \geq 0\}$ non libero!
(lo vedremo)

I ling. liberi non sono chiusi per complementazione

$L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$ cioè se fossero chiusi per complementazione, allora dovrebbero esserlo anche per intersezione ....

$\Rightarrow$ non sono chiusi per complementazione

Pumping Theorem

Se L è libero, allora $\exists N > 0$ tale che $\forall z \in L$ con $|z| \geq N$, $\exists u, v, w, x, y$ tali che

(1) $z = uvwxy$

(2) $|vwx| \leq N$

(3) $|vx| \geq 1$

(4) $\forall k \geq 0 \quad uv^k w x^k y \in L$

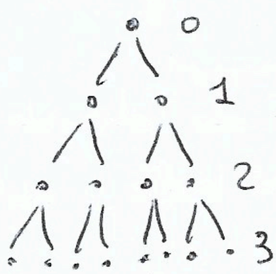
<u>Dimostrazione</u> Sia $G = (NT, T, R, S)$ una grammatica libera tale che $L = L(G)$.

- Sia b il massimo fattore di ramificazione in un albero di derivazione (ovvero il massimo numero di simboli che compaiono nella parte destra di una produzione in R)

$$b = \max \{ |\alpha| \mid A \to \alpha \in R \}$$

($\underline{\text{oss}}$: $b \geq 2$, altrimenti la grammatica sarebbe banale)

- Un albero di altezza h (con 0 la radice) e fattore di ramificazione b, ha al più b^h foglie



$h = 3$
$b = 2$
2^3 foglie

Fissiamo $N = b^{|NT|+1}$ (quindi $N > b^{|NT|}$, dato che $b \geq 2$)

Allora ogni albero di derivazione per z, con $\underline{|z| \geq N}$, <u>deve avere altezza almeno $|NT|+1$.</u>

Prendiamo una qualunque $z \in L$, con $|z| \geq N$.
Consideriamo il suo albero di derivazione
(se ne possiede più di uno, perché G è ambigua, prendiamo quello col minor numero di nodi)

Dunque,

$|z| \geq N \Rightarrow$ albero con altezza $\geq |NT|+1$

$\Rightarrow$ $\exists$ un cammino da radice S ad una foglia con almeno $|NT|+2$ nodi

$\Rightarrow$ Quel cammino attraversa $|NT|+1$ nodi interni etichettati con un nonterminale (la foglia è etichettata da un terminale)

$\Rightarrow$ almeno un nonterminale si ripete in quel cammino

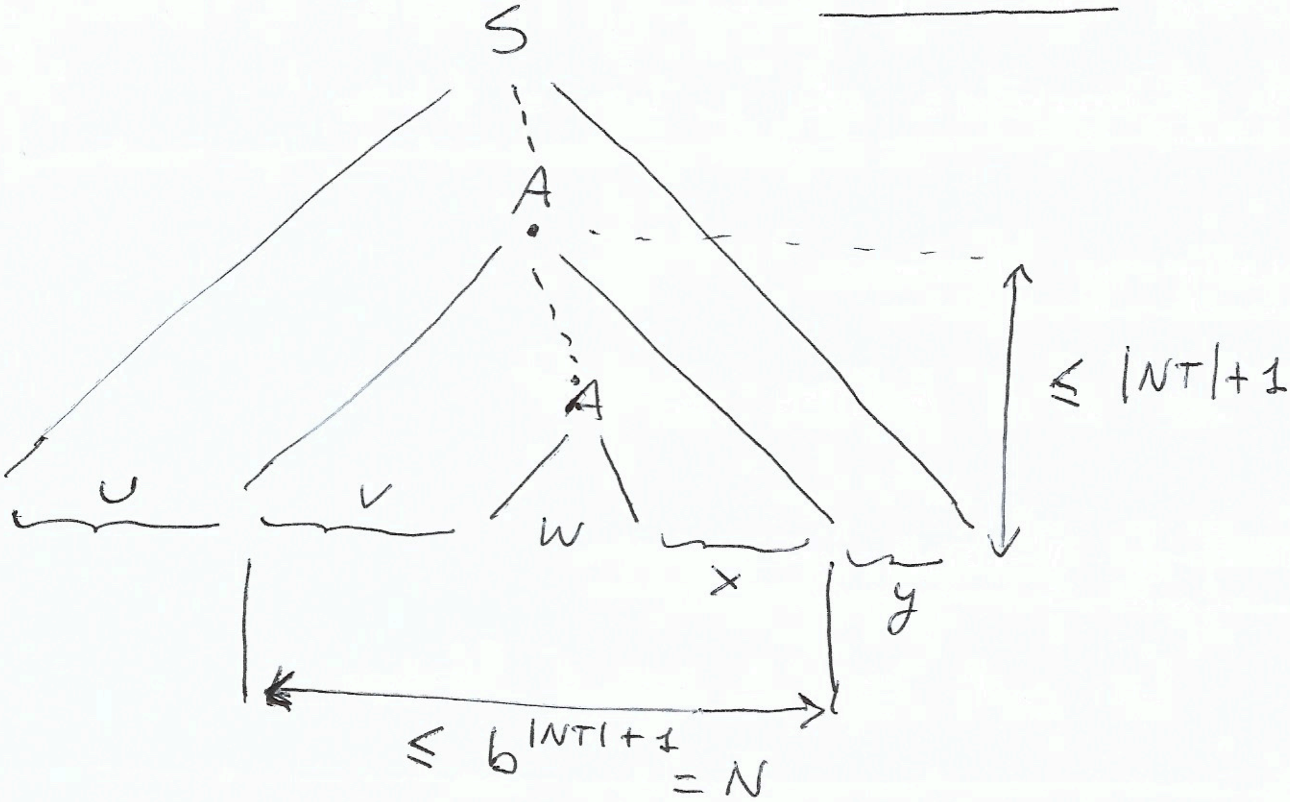
Allora $S \Rightarrow^* z$ può essere diviso come

$$\underbrace{S \Rightarrow^* uAy \Rightarrow^* uvAxy}_{} \Rightarrow^* uvwxy$$

questa parte di derivazione può essere ripetuta più volte; ad es.

$$S \Rightarrow^* uAy \Rightarrow^* uvAxy \Rightarrow^* uv^2Ax^2y \Rightarrow uv^2wx^2y$$

Come albero



Allora

$S \Rightarrow^* uAy \Rightarrow^* uwy \qquad K=0$

$S \Rightarrow^* uAy \Rightarrow^* uvAxy \Rightarrow^* uvwxy \qquad K=1$

$S \Rightarrow^* uAy \Rightarrow^* uvAxy \Rightarrow^* uv^2Ax^2y \qquad K=2$

$\vdots$

Bisogna solo verificare i vincoli:

- $|vx| \geq 1$ ovvio: se entrambe ε, allora l'albero per $K=0$ genererebbe ancora z ed avrebbe meno nodi, contraddicendo l'ipotesi di aver scelto il più piccolo albero
- $|vwx| \leq N$ ovvio: il cammino da A alla foglia è di lunghezza $\leq |NT|+1$ (cioè usa $|NT|+2$ nodi al massimo, di cui uno è il terminale foglia) $\Rightarrow$ la A in alto non può generare parole più lunghe di $b^{|NT|+1} = N$ $\blacksquare$

partendo dal basso, prendi il primo "ciclo" che si forma!

Come usare il pumping theorem "a rovescio" per dimostrare che L non è libero?

Pumping theorem

Se L è libero $\Rightarrow$ P

Pumping theorem a rovescio

Se $\sim P \Rightarrow$ L non è libero

Se $\forall N > 0 \ \exists z \in L$ con $|z| \geq N$ tale che

$\forall u, v, w, x, y$ (**se** (1) $z = uvwxy$

(2) $|vwx| \leq N$

(3) $|vx| \geq 1$

allora $\exists k \geq 0 \,.\, uv^k w x^k y \notin L$)

allora L non è libero

Dimostriamo ora che certi linguaggi non sono liberi, perché soddisfano la proprietà $\sim P$

$L = \{a^n b^n c^n \mid n \geq 0\}$ è libero? No

- Fissiamo $N > 0$ generico $(\forall N > 0)$
- Scegliamo $z = a^N b^N c^N$ $(\exists z \in L$, con $|z| \geq N)$
- Per ogni u, v, w, x, y tali che (1) $z = uvwxy$
  (2) $|vwx| \leq N$
  (3) $|vx| \geq 1$

$\Rightarrow$ le 2 estremità di vwx non possono essere "a" e "c"

- caso vwx non contiene c (i "c" stanno tutti in y)

  $\Rightarrow uv^2wx^2y$ cambia il numero di a e/o b ma non quello di c

  $\Rightarrow uv^2wx^2y \notin L$

- caso vwx non contiene a (le "a" stanno tutte in u)

  del tutto analogo.

  $\Rightarrow uv^2wx^2y \notin L$

$\Rightarrow$ L non è libero!

$L = \{ a^n b^m a^n b^m \mid n, m \geq 0 \}$

è libero? <u>NO</u>

- Fissiamo $N > 0$ generico $(\forall N > 0)$
- Scegliamo $z = a^N b^N a^N b^N$ $(\exists z \in L, |z| \geq N)$
- Per ogni u, v, w, x, y tali che (1) $z = uvwxy$

  (2) $|vwx| \leq N$

  (3) $|vx| \geq 1$

può solo essere uno dei seguenti casi

(i) $vwx \in a^*$ o $vwx \in b^*$

$\Rightarrow uv^2wx^2y \notin L$ perché ho aggiunto delle "a" (o delle "b") solo ad un pezzo

(ii) $vwx \in a^*b^*$ o $vwx \in b^*a^*$

$\Rightarrow uv^2wx^2y \notin L$ perché ho aggiunto "a" e "b" (o "b" ed "a") solo ad un blocco, ma non all'altro

In ogni caso,

$uv^2wx^2y \notin L$

$\Rightarrow L$ non è libero

$L = \{a^n \mid n$ è primo$\}$ è libero? <u>No</u>

- Fissiamo $N > 0$ generico
- Scelgo $z = a^p$ con p primo, $p \geq N+2$
- Per ogni u, v, w, x, y tali che
  (1) $z = uvwxy$
  (2) $|vwx| \leq N$
  (3) $|vx| \geq 1$

deve essere $1 \leq |vx| \leq N$. Sia $|vx| = m$.

Allora

$$|uv^0wx^0y| = |uwy| = p - m$$

Ma allora

$$|uv^{p-m}wx^{p-m}y| = |uwy| + (p-m)\cdot|vx| =$$
$$= (p-m) + (p-m)\cdot m$$
$$= (p-m)\cdot(1+m)$$

che non è primo se

$p - m > 1$ e $1 + m > 1$

ma $m \leq N$, $p \geq N+2$ $\Rightarrow p - m \geq 2$

ma $m \geq 1$ $\Rightarrow 1 + m > 1$

$$\Rightarrow uv^{p-m}wx^{p-m}y \notin L$$

$\Rightarrow$ L non è libero

Oltre i linguaggi liberi:

CLASSIFICAZIONE DI CHOMSKY

- Grammatiche regolari

$$A \to aB \quad A \to a \quad S \to \varepsilon$$

- Grammatiche libere da contesto

$$A \to \gamma \quad \text{con } \gamma \in (NT \cup T)^+ \quad S \to \varepsilon$$

- Grammatiche dipendenti dal contesto

$$\gamma A \delta \to \gamma w \delta \quad \gamma, \delta \in (NT \cup T)^* \; w \in (NT \cup T)^+$$

$$S \to \varepsilon$$

- Grammatiche monotone

$$\gamma \to \delta \quad \text{con } |\gamma| \leq |\delta|$$

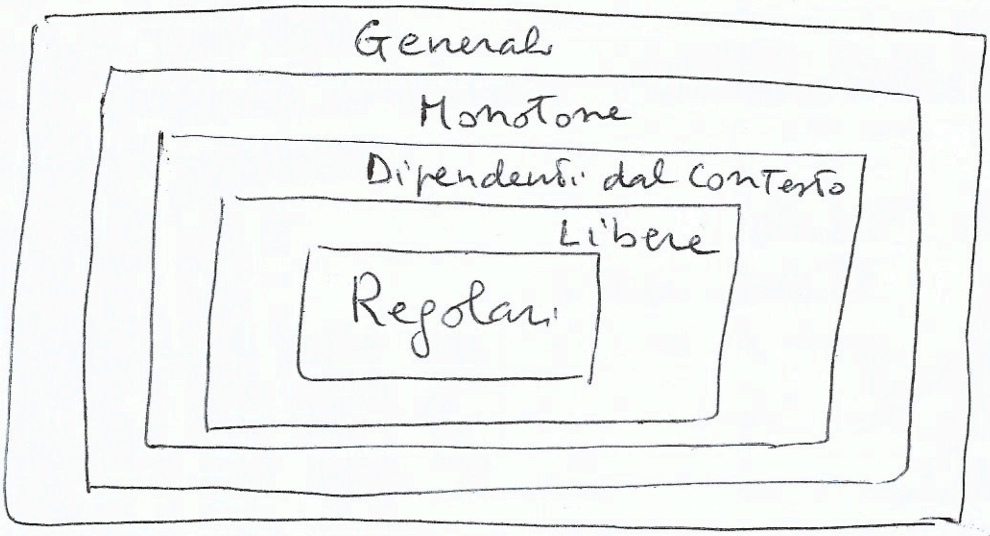
- Grammatiche generali (a struttura di frase)

$$\gamma \to \delta \quad (\text{senza alcun vincolo})$$

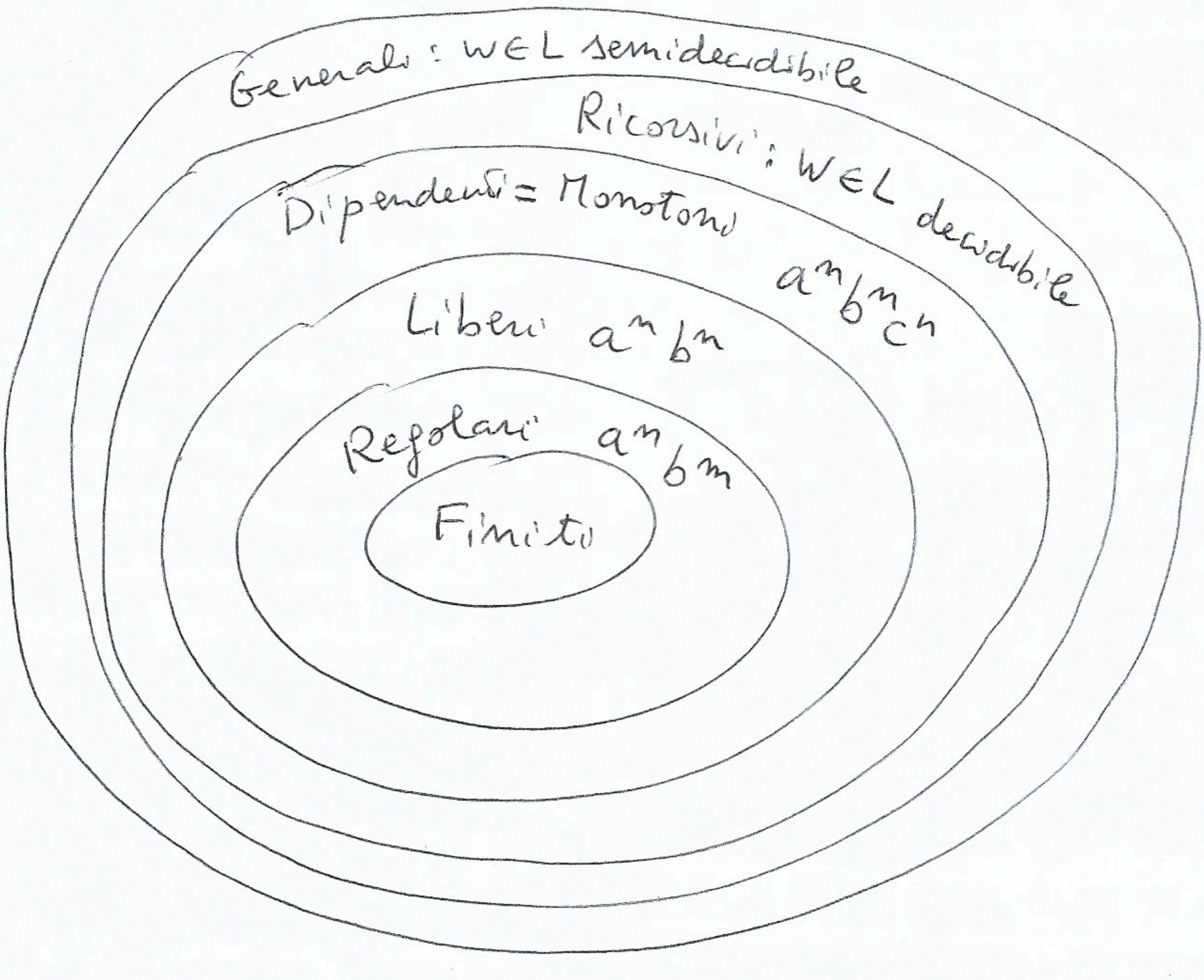
Oss1: $\left.\begin{array}{l} S \to \varepsilon \mid S' \\ S' \to ab \mid aS'b \end{array}\right]$ G libera (secondo questa def.) per $L = \{a^n b^n \mid n \geq 0\}$

non può comparire a dx in una produzione

Oss2: Teorema: Per ogni G_1 monotona, esiste G_2 dipendente dal contesto tale che $L(G_1) = L(G_2)$



Classificazione delle grammatiche

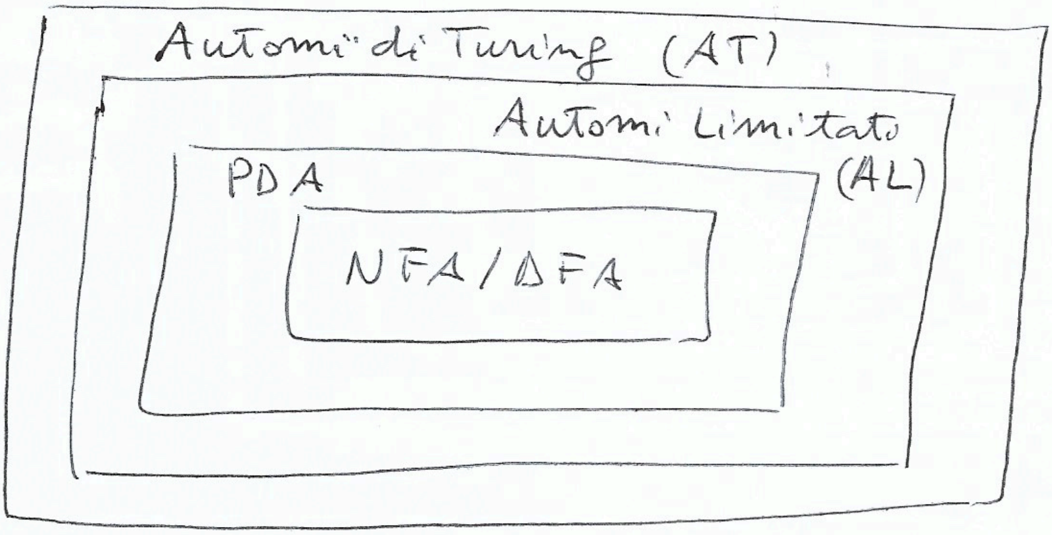


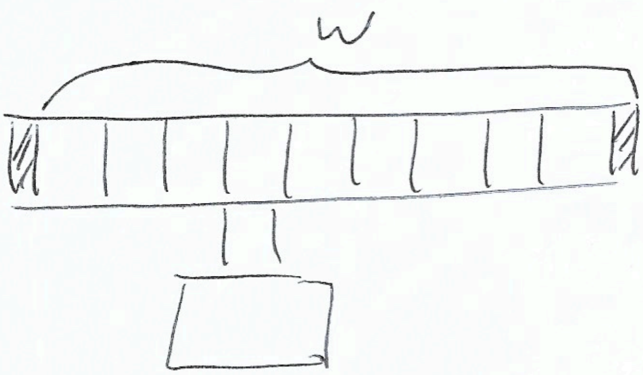
Classificazione dei linguaggi

Una grammatica monotona per $\{a^n b^n c^n \mid n \geq 0\} = L$

$$G \begin{cases} S \to \varepsilon \mid S' \\ S' \to aS'Bc \mid abc \\ cB \to Bc \\ bB \to bb \end{cases} \qquad L(G) = L$$

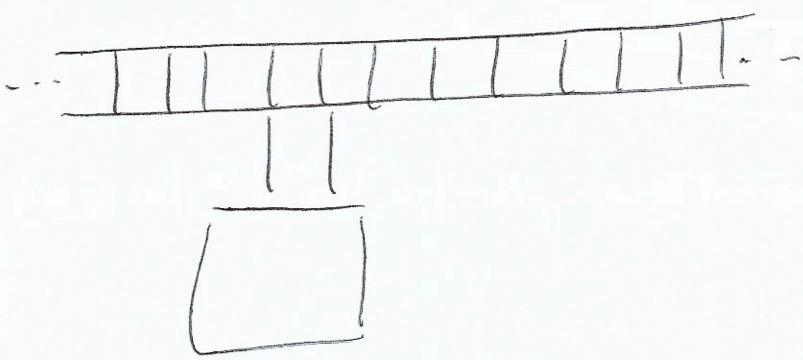
Gerarchia di Automi





AL

- legge e scrive sul nastro
- va a destra e a SX
- riconosce per stato finale
- spazio utilizzabile limitato dalla lunghezza dell'input



AT

Come prima, ma senza vincoli di spazio (nastro infinito)