DIM. Sia $P = (Q, \Sigma, \Gamma, \delta, q_0, \{q_{acc}\})$ l'automa a pila in forma "CANONICA" (vedi sopra).

Intanto pongo $V = \{A_{pq} : p, q \in Q\}$

$S = A_{q_0 q_{acc}}$. Le regole R:

- Per ogni $p, q, r, s \in Q$, $u \in \Gamma$, $a, b \in \Sigma_\varepsilon$

se $(r, u) \in \delta(p, a, \varepsilon)$ e

$(q, \varepsilon) \in \delta(s, b, u)$

pongo $A_{pq} \to a A_{rs} b$

- Per ogni $p, q, r \in Q$, pongo $A_{pq} \to A_{pr}A_{rq}$
- Per ogni $p \in Q$, $A_{pp} \to \varepsilon$.

La prova segue dal fatto che A_{pq} genera x sse x porta P da p a q con pila vuota.

$\Longrightarrow$ S produce x sse $x \in L(P)$.

AFF. Se $A_{p,q}$ genera x, allora x porta P da p a q con pila vuota.

Per induzione su # passi della produzione di x in G.

*) BASE: # 1 passo. L'unica regola possibile è $A_{pp} \to \varepsilon$; e porta P da p a p con pila vuota.

*) PASSO INDUTTIVO: Suppongo vero per alcune: x con # passi $\leq K$ e sia $A_{pq} \overset{*}{\Rightarrow} x$ in $K+1$ passi. Ci sono due casi:

1) $A_{pq} \to a A_{rs} b$

2) $A_{pq} \to A_{pr} A_{rq}$

1) Sia $x = ayb$ così $A_{rs} \overset{*}{\Rightarrow} y$ in K passi e quindi P può andare da r a s con pila vuota.

Siccome c'è la regola $A_{pq} \to a A_{rs} b$, allora

$$(r, u) \in \delta(p, a, \varepsilon)$$
$$(q, \varepsilon) \in \delta(s, b, u) \qquad u \in \Gamma$$

quindi: P entrando in p con pila vuota, legge a, inserisce u nella pila e va in r. La stringa y porta P da r ad s mantenendo u nella pila. Infine P va da s a q rimuovendo u.

$\Rightarrow$ P va da p a q con pila vuota

2) Sia $x = yz$, $A_{pr} \overset{*}{\Rightarrow} y$ e $A_{rq} \overset{*}{\Rightarrow} z$ in $\leq k$ passi

$\Rightarrow$ y porta P da p a r con pila vuota
z porta P da r a q con pila vuota

$\Rightarrow$ $x = yz$ porta P da p a q con pila vuota.

AFF. Se x porta P da p a q con pila vuota, allora $A_{pq} \Rightarrow x$.

Per induzione su # passi di P:

– BASE: # passi $= 0$. La computazione inizia e finisce in p. In zero passi

$x = \varepsilon$. Siccome G contiene $A_{pp} \to \varepsilon$ allora $A_{pp} \Rightarrow x$.

- INDUZIONE: Suppongo vero per # passi $\leq k$, e assumo che x porta P da p a q in $k+1$ passi con pila vuote. Due casi:

1) Pila vuota solo negli stati p e q

2) Pila si svuota in mezzo.

1) In questo caso il primo simbolo inserito è lo stesso rimosso all'ultimo passo. Sia questo $u \in \Gamma$, e siano $a, b \in \Sigma_\varepsilon$ i simbolo letti nella prima e ultima mossa col

r ed s gli stati dopo la prima e prima dell'ultima mossa.

$\Rightarrow$ $\delta(p, a, \varepsilon)$ contiene (r, u)

$\delta(s, b, u)$ contiene (q, ε)

$\Rightarrow$ la grammatica G contiene la regola

$A_{pq} \to a A_{rs} b$.

Sia $x = a y b$. L'input x porta P da p a q con pila vuota e quindi y porta P da r ad s con pila vuota.

Siccome abbiamo esplicitato il primo e

l'ultimo passo della computazione, la computazione su y richiede $k+1-2 = k-1$ passi. Per induzione $A_{rs} \overset{*}{\Rightarrow} y$ ovvero $A_{pq} \overset{*}{\Rightarrow} x$.

2) Sia r lo stato in cui si svuota la pila. La computazione di P da p ad r e da r a q richiede $\leq k$ passi.
Sia $x = yz$; per induzione $A_{pr} \overset{*}{\Rightarrow} y$. $A_{rq} \overset{*}{\Rightarrow} z$; siccome $A_{pq} \to A_{pr} A_{rq}$ allora $A_{pq} \overset{*}{\Rightarrow} x$. $\blacksquare$

Pumping Lemma per CFG

Tutti i linguaggi sono CFL? No. Ad es.

$$L = \{ 0^n 1^n 2^n : n \geq 1 \}$$

TEO Se L è un CFL, esiste un numero p t.c. se $w \in L$ con $|w| \geq p$, allora si può scomporre $w = uvxyz$ in modo che:

(i) $\forall\, i \geq 0$, $uv^i x y^i z \in L$

(ii) $|vy| > 0$

(iii) $|vxy| \leq p$.

DIM. Assumiamo che la CFG associata al L sia CNF. L'albero di derivazione è binario. Inoltre:

FATTO Ogni albero di derivazione di G il cui cammino più lungo sia lungo i, genera una stringa lunga al più 2^{i-1}.

DIM. Per induzione su i. Se $i = 1$ allora $|w| = 1$ e la lunghezza è $2^{1-1} = 1$.

Sia vero per lunghezza i e vediamo che succede per $i+1$, $i \geq 1$. La prima sostituzione deve essere di tipo $S \to BC$

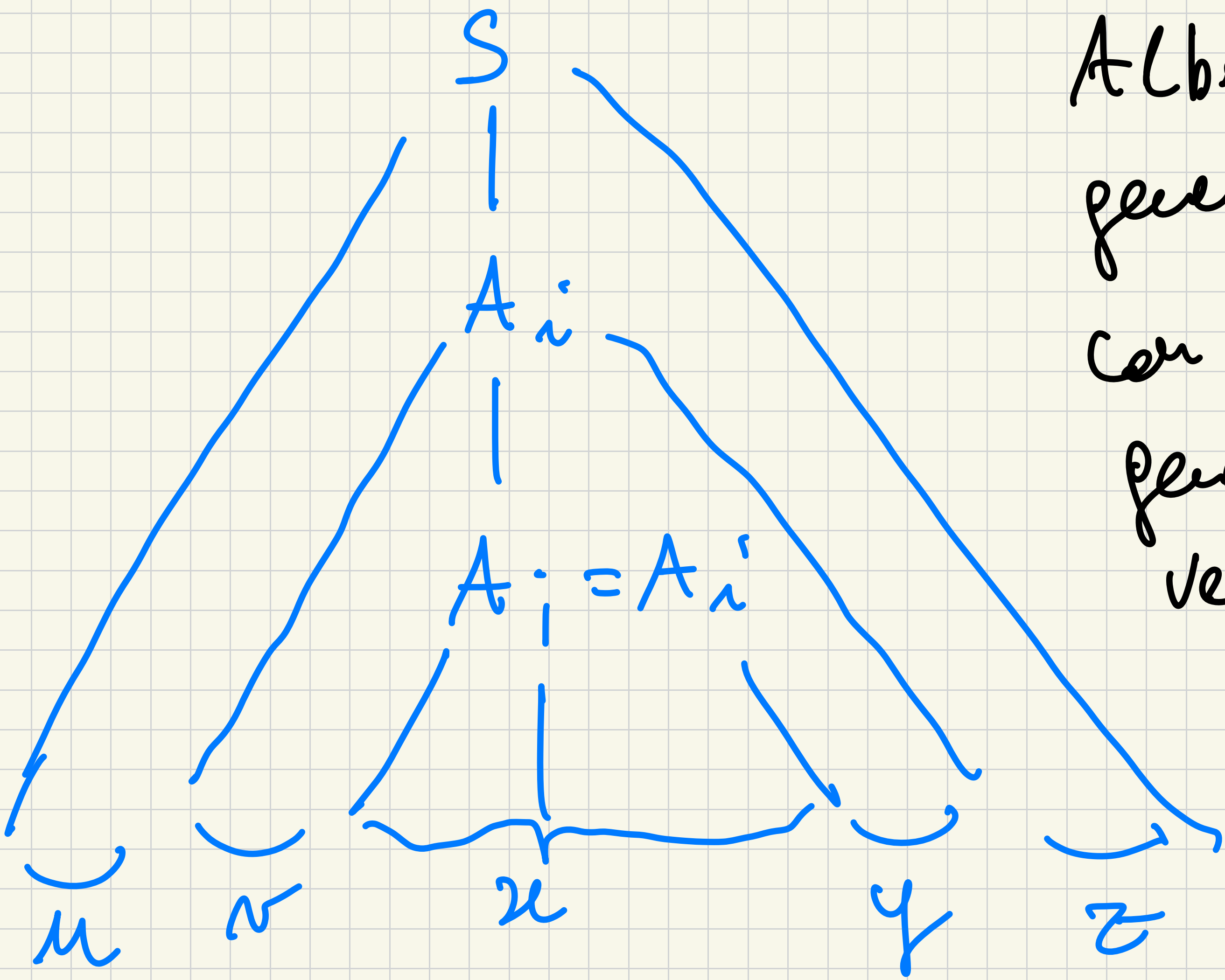
e i sottoalberi che generano B, C hanno lunghezza al più i. Per induzione generano stringhe lunghe al più 2^{i-1}

$\Rightarrow$ S genera stringhe lunghe $\le 2^{i-1} \cdot 2$

$= 2^{i}$ $\blacksquare$

Torniamo al pumping lemma. Sia $p = 2^{m}$ dove m sono le variabili. Data w con $|w| \ge p$ avrà un albero di derivazione di lunghezza $\ge m+1$. Il numero di nodi è almeno $m+2$ di cui 1 è un terminale ed $m+1$ variabili.

Siccome ci sono solo n variabili una si ripete; supponiamo che $A_i = A_j$.



Albero con radice A_i genera vxy, quello con radice $A_j = A_i$ genera x. u e z vengono generate nella derivazione di A_i da S.

Supponiamo G è in CNF, un sottoalbero consiste di simbolo terminale o 2 variabili. Quindi il sottoalbero con radice A_i è fatto di due variabili $A_i \to BC$. Una di queste genera x tranne $A_j = A_i$, l'altra genera vy.

$$\Rightarrow \quad vxy \neq x \qquad (ii)$$

$$|vy| > 0$$

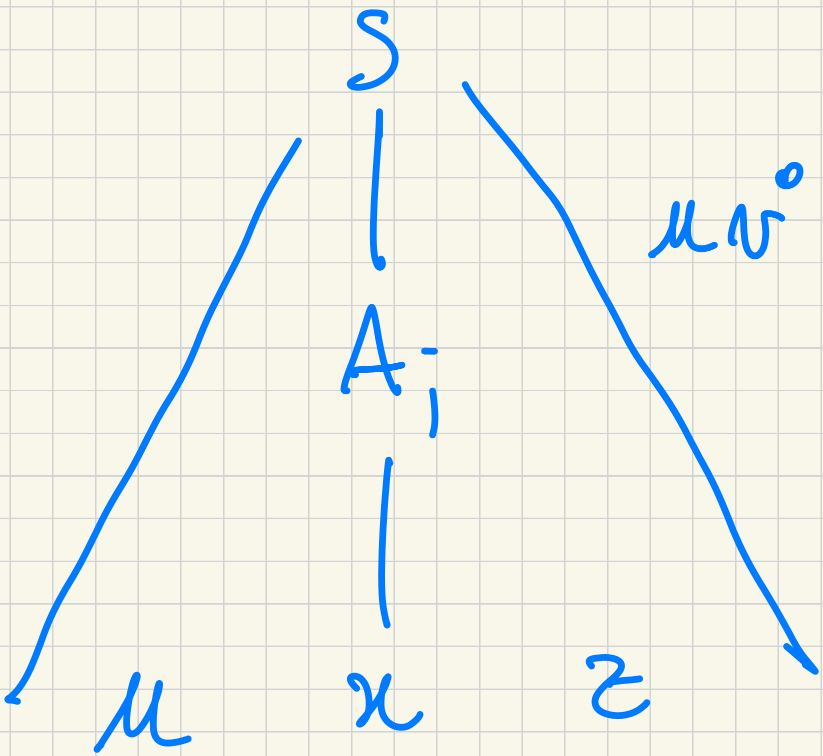
(iii) Abbiamo scelto l'albero in modo che il cammino più lungo abbia lunghezza

$n+1$ ovvero genera stringhe di lung.

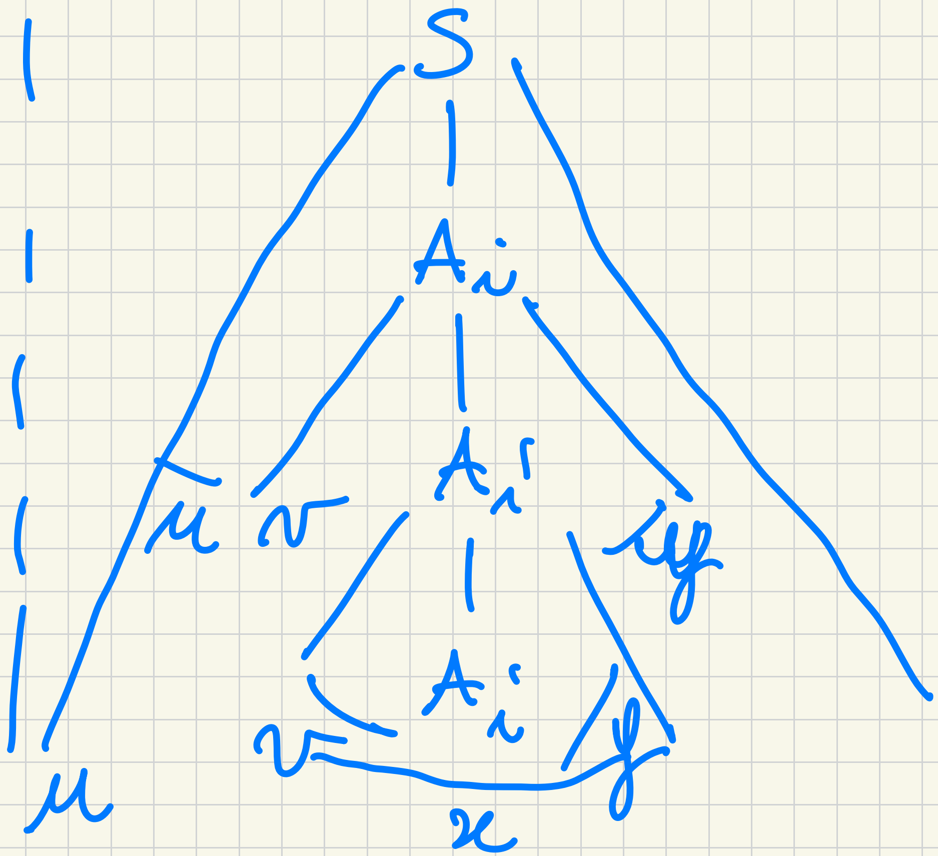
$$\leq 2^{n+1-1} = p$$

$$\Rightarrow |vxy| \leq p.$$

(i) Posso sostituire A_i con A_j e viceversa.



$uv^0xy^0z \in L$



$\Rightarrow u v^2 x y^2 z \in L$

$\forall i \geq 0 \quad u v^i x y^i z \in L$ $\blacksquare$