GRAMMATICA A STRUTTURA DI FRASE

Formalmente è una quadrupla

$$G = (\Sigma, V, S, P)$$

Σ alfabeto del linguaggio

V insieme finito di simboli (categorie sintattiche oppure simboli non terminali)

$S \in V$ categoria sintattica iniziale (oppure simbolo distinto)

P è un insieme finito di PRODUZIONI.

Una produzione ha la forma

$$A \to \alpha$$

dove $A \in V$, $\alpha \in (\Sigma \cup V)^*$

α è una stringa che contiene
sia simboli di Σ
sia " " V

Esempio

$$G = (\{a,b\}, \{S\}, S, \{S \to ab, S \to aSb\})$$

$$\in (\{a,b\} \cup \{S\})^*$$

Nella teoria dei linguaggi formali esiste una gerarchia di grammatiche a struttura di frase, che si distinguono per la forma delle produzioni.

Se le produzioni hanno la forma

$$A \to \alpha \quad \text{dove} \quad A \in V, \ \alpha \in (\Sigma \cup V)^*$$

la grammatica si chiama LIBERA DAL CONTESTO (LIBERA)

Come può una grammatica libera definire un linguaggio.

$$G = (\{a, b\}, \{S\}, S, \{S \to ab,\ S \to aSb\})$$

$S \to ab$
$S \to aSb$

simboli dell'alfabeto sono minuscoli

le categorie sintattiche sono maiuscole

la prima produzione ha come cat. sintattica quella iniziale

$S \to ab$
$S \to aSb$
$\equiv$ $S \to ab \mid aSb$ (oppure)

le produzioni con la stessa cat. sintattica alla sinistra si possono scrivere, più compattamente in questo modo

Derivazione di una stringa $\alpha \in \Sigma^*$ a partire da una grammatica libera.

Si inizia con il simbolo distinto (cat. sintattica iniziale) e si procede sostituendo ad una categoria sintattica nella stringa ottenuta con la parte destra di una prod. per quella cat. sintattica.

$S \to ab \mid aSb$

derivazioni

$S \to ab \in \Sigma^*$

$S \to aSb \to aabb \in \Sigma^*$

$S \to aSb \to aaSbb \to aaabbb \in \Sigma^*$

i simboli di Σ si chiamano anche SIMBOLI TERMINALI perché terminano le derivazioni.

Una grammatica libera G definisce il linguaggio di tutte e sole le stringhe DERIVABILI in G.

$S \to ab \mid aSb$

S è definito in termini di S (definizione ricorsiva, produzione ricorsiva)

Qual'è il linguaggio definito? (linguaggio generato)

$S \to ab$

$S \to \ldots \to aabb$

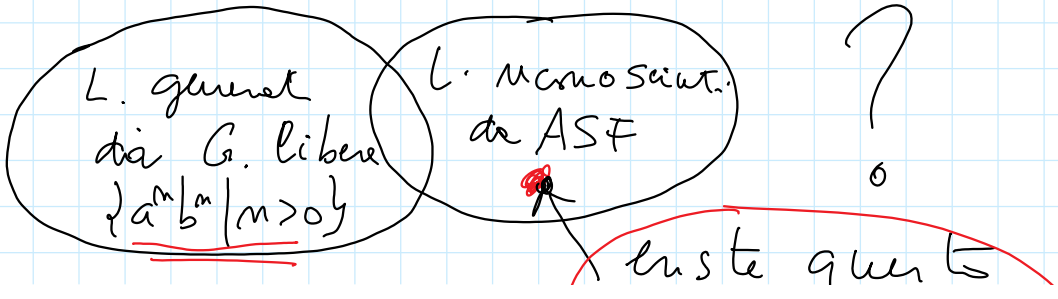
$S \to \ldots \to aaabbb$

?

$L = \{ a^n b^n \mid n > 0 \}$ linguaggio che non può essere riconosciuto da un ASF

grammatiche → strumento generativo

ASF → " riconoscitivo



$\{a^n b^m \mid n > 0\}$

esiste questo linguaggio?

Esiste un linguaggio riconosciuto da un AST che non può essere generato da una gramm. libera?

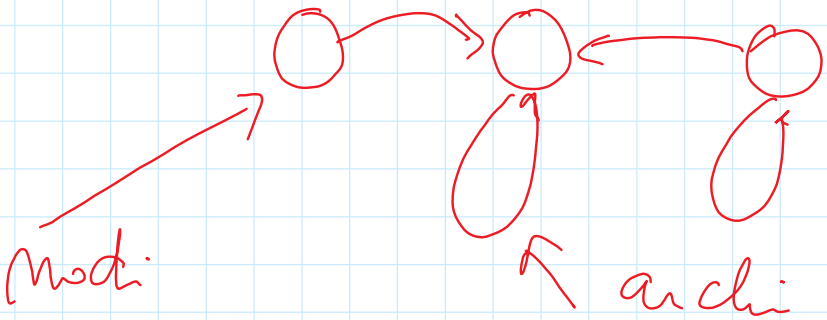
$S \to aSb \to aaSbb \to aaabbb \in \Lambda^*$

Un modo di rappresentare le derivazioni in modo compatto e dato dagli:

ALBERI DI DERIVAZIONE

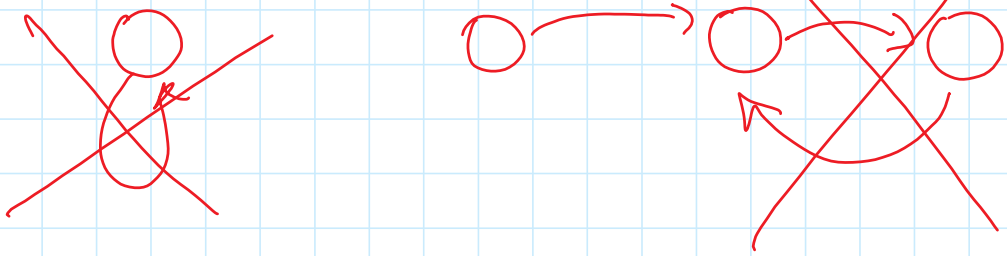
Albero

è un grafo (che abbiamo usato per rappresentare gli AST)

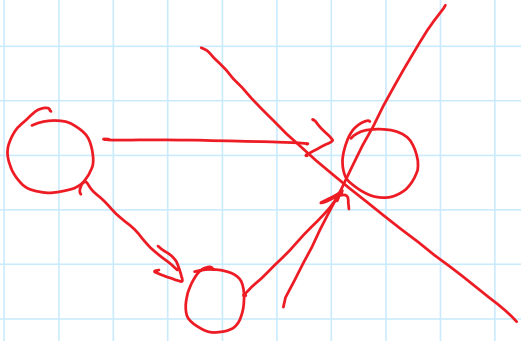


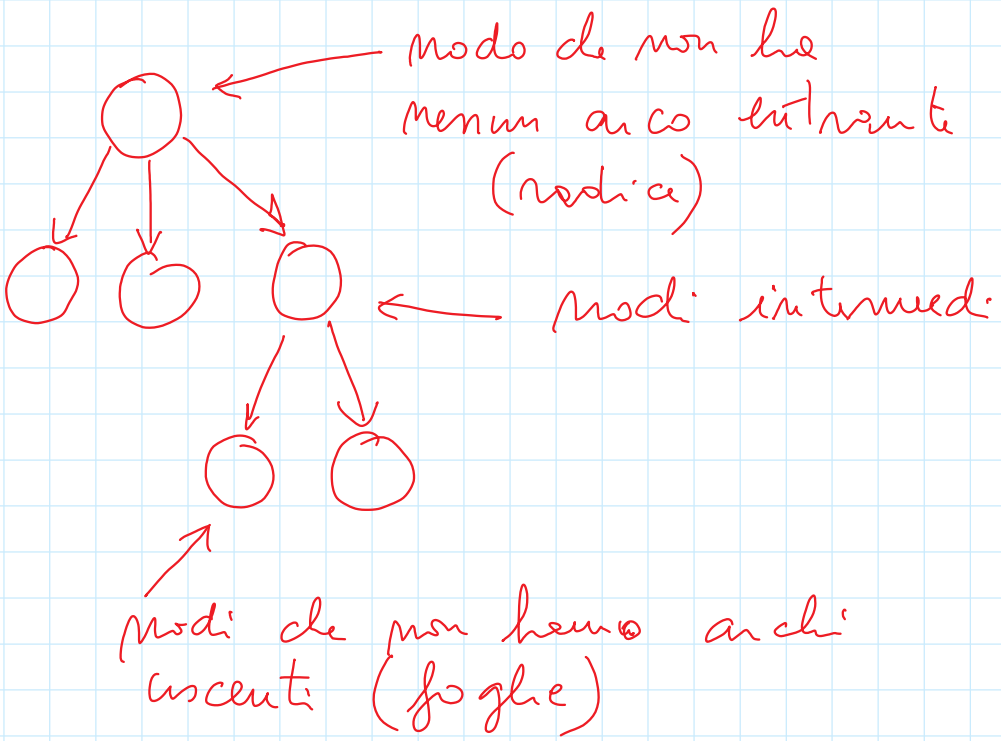
Albero è un grafo:

- aciclico (non ci sono cicli)



- ogni nodo ha, al più, un arco entrante



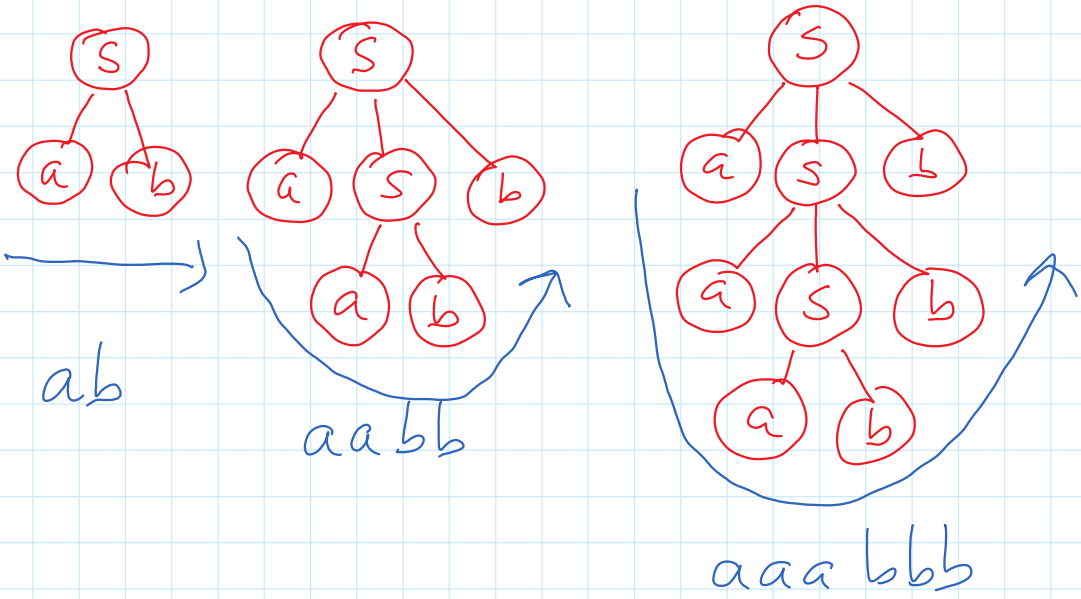


ALBERO DI DERIVAZIONE RISPETTO A UNA GRAMMATICA G:

- la radice contiene la cat. sint. iniziale
- le foglie contengono simboli di Σ
- i nodi intermedi contengono cat. sintattiche
- i figli di un nodo intermedio che contiene la cat. sintattica $A \in V$, contengono i simboli della parte destra di una produzione per A, nell'ordine.

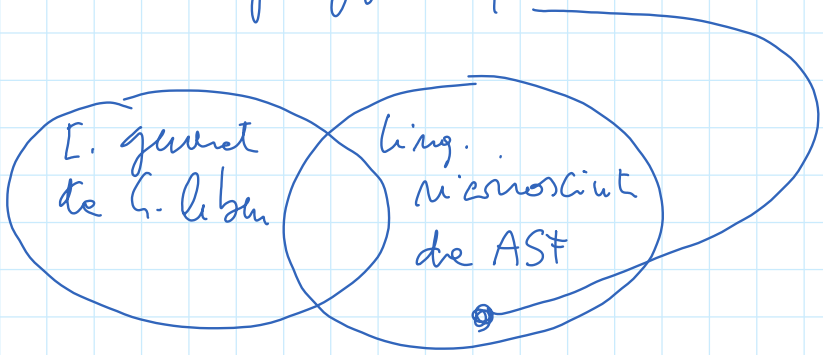
lunedì 14 ottobre 2019 09:57

$S \to ab \mid aSb$

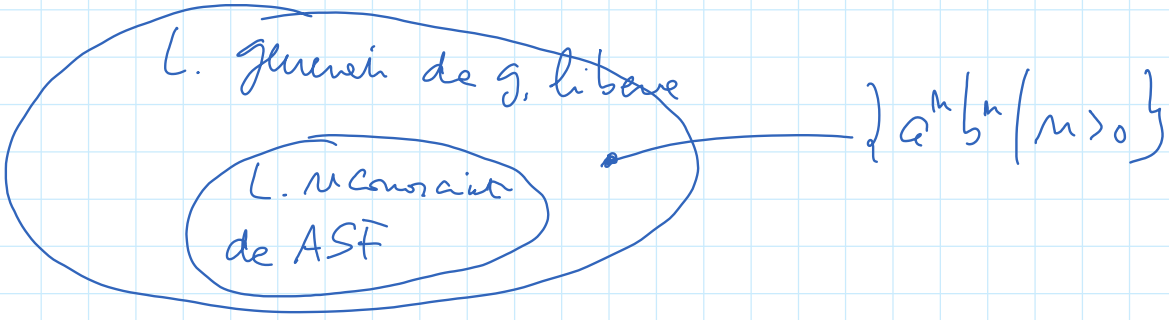


Il linguaggio generato da una grammatica libera G è l'insieme di tutte e sole le stringhe che hanno un albero di derivazione in G.

Esiste un linguaggio così?



NO! Tutti i linguaggi riconosciuti da ASF possono essere generati da una gramm. libera.



Dimostrazione costruttiva.

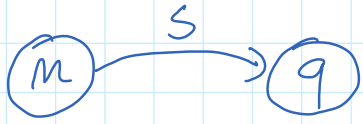
Data un qualsiasi ASF A si costruisca la grammatica equivalente (che genera il linguaggio riconosciuto da A)

lunedì 14 ottobre 2019 10:19

Dato $A = (\Sigma, Q, s_0, F, \delta)$

costruiamo la grammatica G

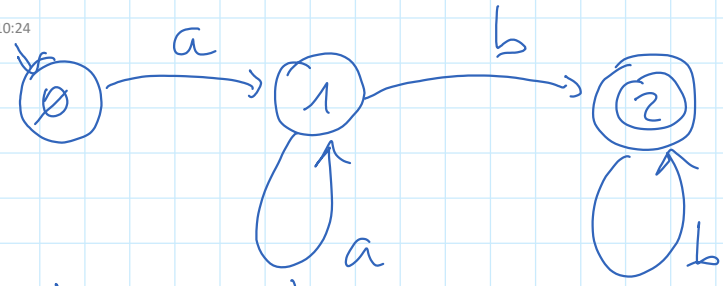
$G = (\Sigma,$
$Q,$
$s_0,$

se nell'automa ho 

allora aggiungo a G la produzione $n \to s\,q$ ← simbolo di Σ

stato dell'ASF e quindi cat. rivolte alle grammatiche

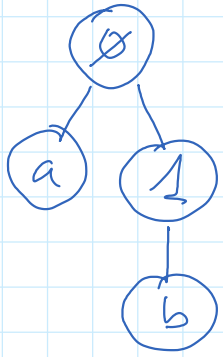
Se $q \in F$ allora aggiungo anche la produzione $n \to s$



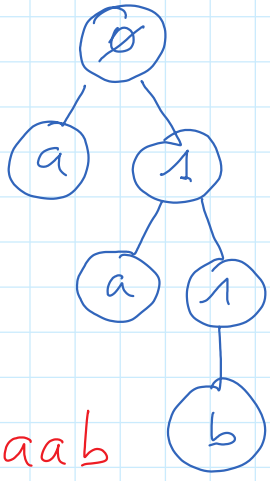
$$L = \{ a^n b^m \mid n, m > 0 \}$$

$$G = (\{a, b\}, \{0, 1, 2\}, 0,$$
$$\{ 0 \to a1,\ 1 \to a1,\ 1 \to b2,\ 1 \to b,$$
$$2 \to b2,\ 2 \to b \})$$

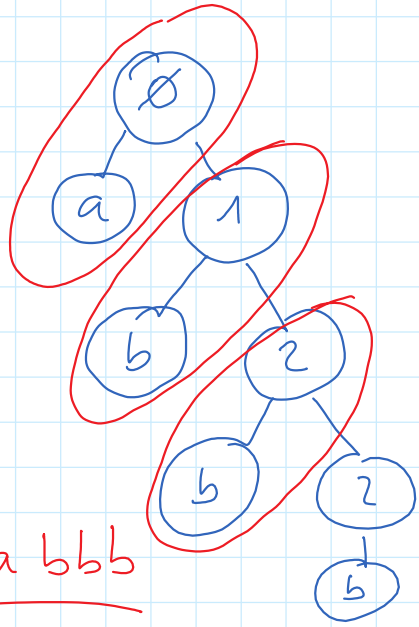
$$0 \to a1$$
$$1 \to a1 \mid b2 \mid b$$
$$2 \to b2 \mid b$$



ab



aab



abbb

Produzioni generate dalla dimostrazione costruttiva hanno la forma

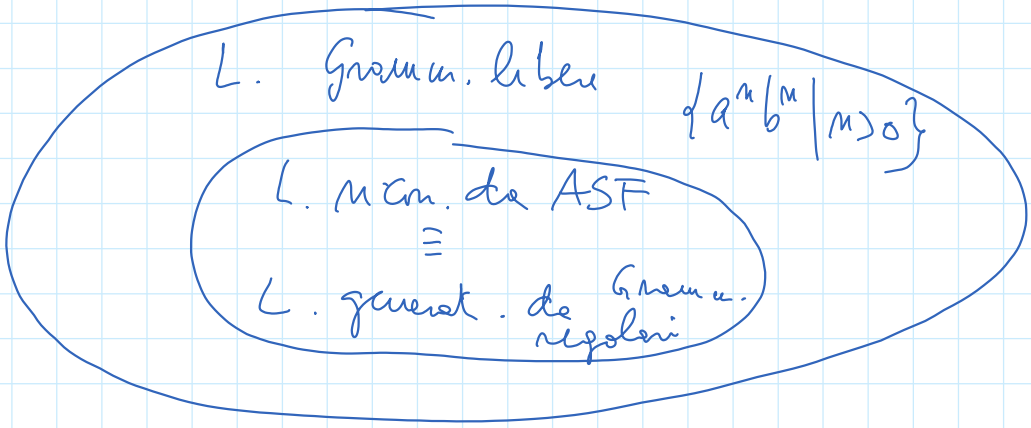
$A \to aB$ $\quad A, B \in V$, $a \in \Sigma$

oppure

$A \to a$ $\quad A \in V$, $a \in \Sigma$

Le grammatiche con questo tipo di produzioni si chiamano GRAMMATICHE REGOLARI

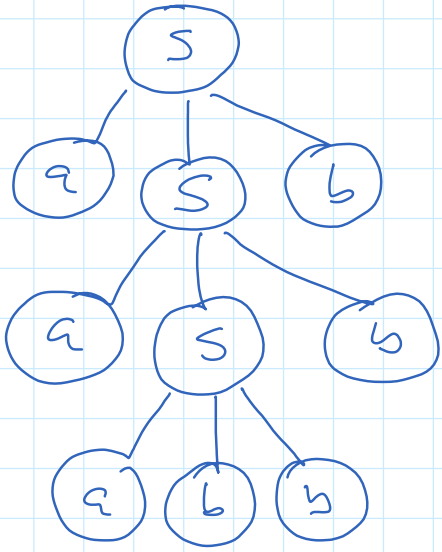
Le grammatiche regolari generano lo stesso insieme di linguaggi riconosciuti dagli ASF



lunedì 14 ottobre 2019 10:35

$L_1 = \{ a^n b^{n+1} \mid n > 0 \}$

$S \rightarrow abb \mid aSb$

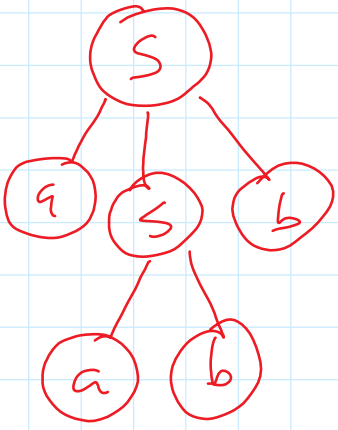


aa a bb bb

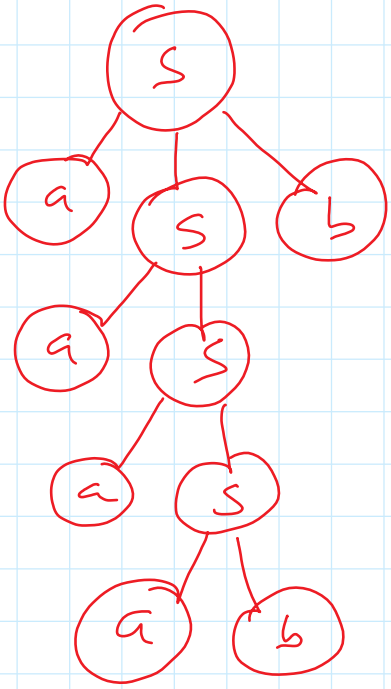
$$L_2 = \{ a^n b^m \mid n \geq m > 0 \}$$

$aabb \in L_2$ $\quad aaaab \in L_2$ $\quad abb \notin L_2$

$$S \to ab \mid aSb \mid aS$$



aabb



aaaabb

Linguaggio delle parentesi bilanciate

$((a+b)*c)+(d+e)$

$(())()$

- quando apro una parentesi dovrò anche chiuderla
- non posso mettere una parentesi chiusa se non ho la corrispondente parentesi aperta

L_p

$()()(()) \in L_p$

$()(() \notin L_p$

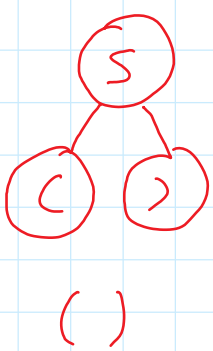
$()) \notin L_p$

$\Sigma = \{ (,) \}$

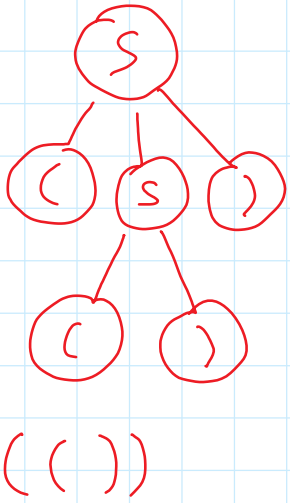
$S \rightarrow () \mid (S) \mid SS$

()()

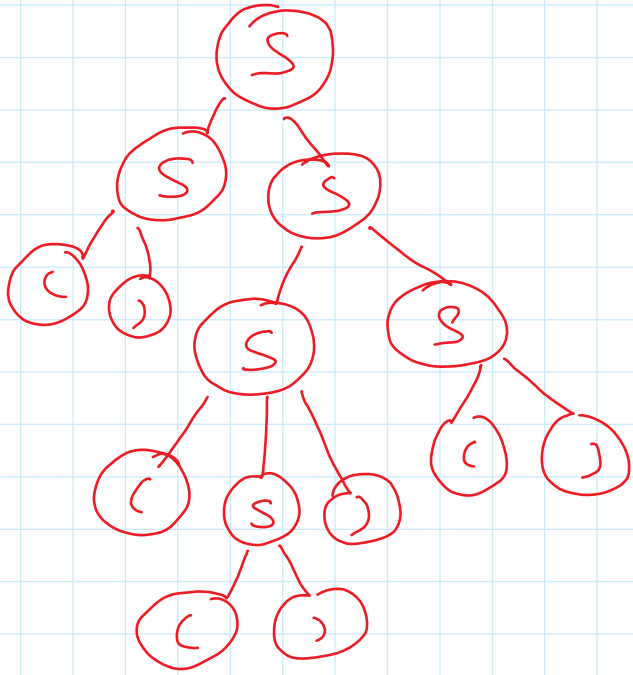
()(())()



( )



(( ))



( )(( ))( )

$\Sigma = \{0, \ldots, 9, a, \ldots, z, =, ;, if, (,), else, while, \{, \}, \}$

Cifra → 0 | ... | 9

Lettera → a | ... | z

Seq → Cifra | Lettera | Cifra Seq | Lettera Seq

Ide → Lettera | Lettera Seq

Comando → Ide = Exp ; |
if (Exp) Comando |
if (Exp) Comando else Comando |
while (Exp) Comando |
{ Dec_list Com_list }

Com_list → Comando | Comando Com_list

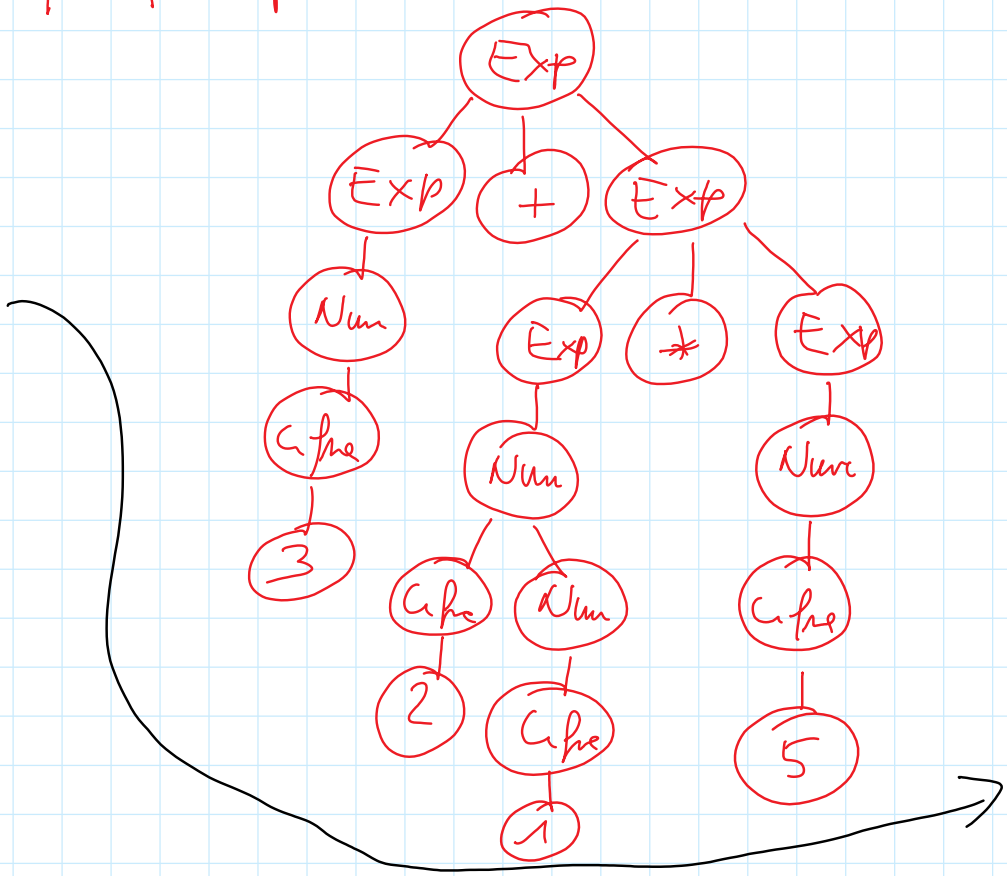
Dec_list → ....

Exp → Ide | Num | Exp + Exp | Exp * Exp

Ide → ....

Num → Cifra | Cifra Num

3+21*5



3+21*5

x+1