LECTURE - 8

Regular Languages

- Sequencing / concatenation $A \cdot B$
- Branching $A + B$
- Repetition A^*

Context-free languages

All of the above, and

- Recursion

Example of a CONTEXT-FREE LANGUAGE.

finite sets

Σ - alphabet = $\{0, 1\}$
terminal

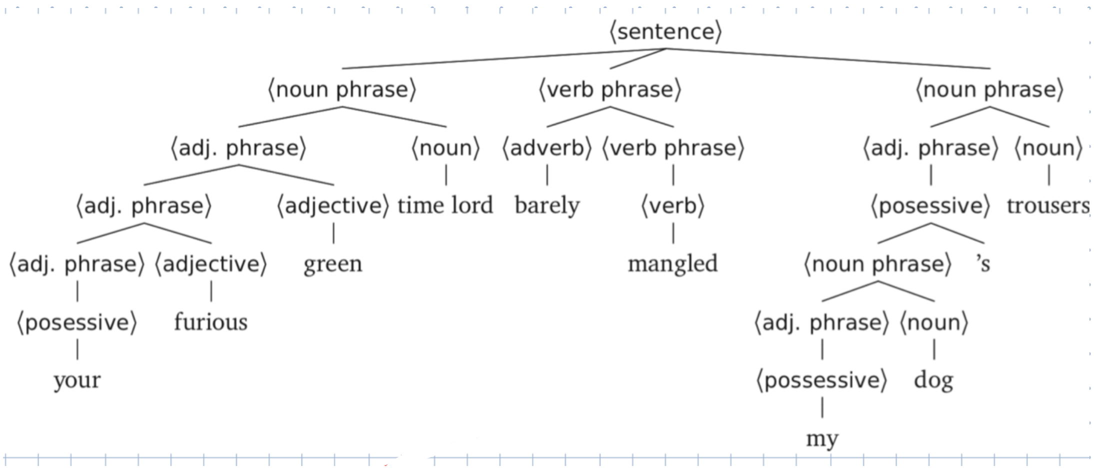
Γ - non-terminals = $\{S, A, B, C\}$

R - production rules

$$\underset{\text{non-terminal}}{A} \longrightarrow w \in (\Sigma \cup \Gamma)^*$$

S - starting non-terminal

$$G = (\Sigma, \Gamma, R, S) \qquad L(G) = L(S)$$



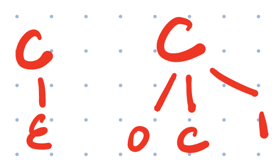
Simple example :

$\Sigma = \{0, 1\}$

$\Gamma = \{S, A, B, C\}$

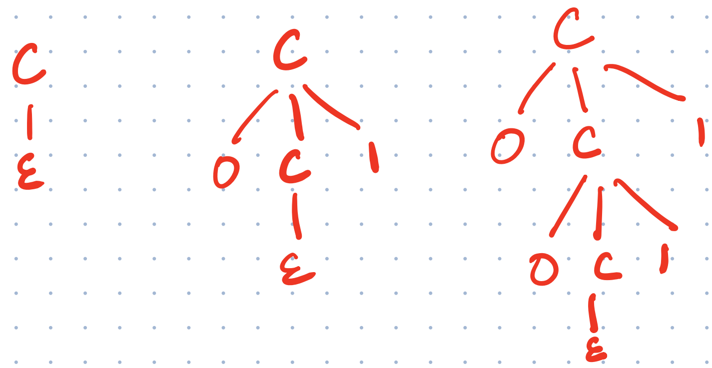
R :

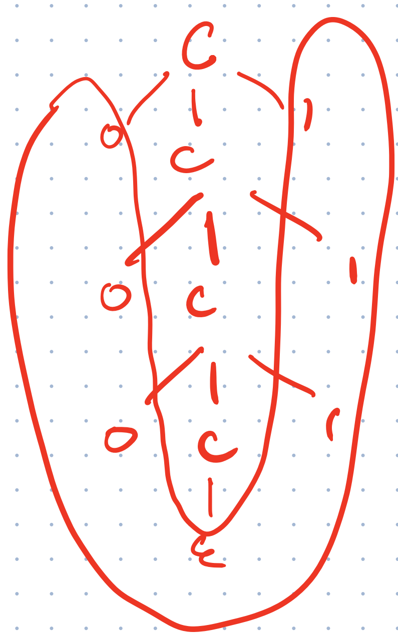
$$\boxed{\begin{array}{l} S \to A \\ S \to B \\ A \to 0A \\ A \to 0C \\ B \to B1 \\ B \to C1 \\ C \to \varepsilon \\ C \to 0C1 \end{array}}$$



$L(A)$ = set of strings in Σ^* generated by A.

$L(C) = \{\varepsilon, 01, 0011, \ldots 0^i 1^i \ldots\}$



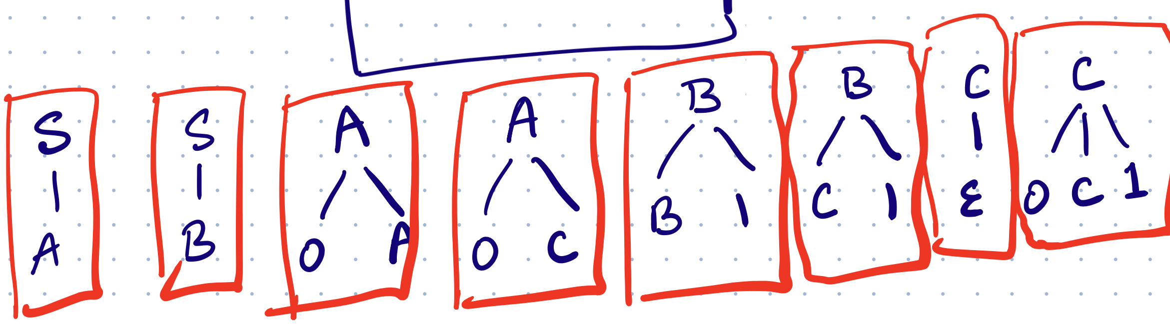


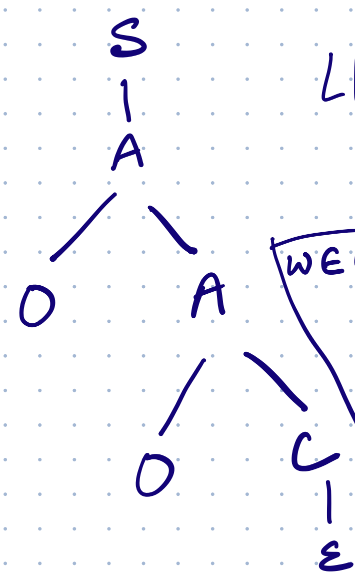
$S \to A$
$S \to B$
$A \to 0A$
$A \to 0C$
$B \to B1$
$B \to C1$
$C \to \varepsilon$
$C \to 0C1$

$S \to A \mid B$
$A \to 0A \mid 0C$
$B \to B1 \mid C1$
$C \to \varepsilon \mid 0C1$

BNF
Backus-Naur-form

$S \to A \mid B$
$A \to 0A \mid 0C$
$B \to B1 \mid C1$
$C \to \varepsilon \mid 0C1$





$L(S) = \{$ set of all strings GENERATED by $S \}$

$w \in L(S)$, As long as there is A tree rooted at S such that sequence of leaves concatenated from left to right $= w$.

what is $L(S)$ for

$$
\boxed{
\begin{aligned}
S &\to A \mid B \\
A &\to 0A \mid 0C \\
B &\to B1 \mid C1 \\
C &\to \varepsilon \mid 0C1
\end{aligned}
}
$$

$L(S) = \{0^m 1^n,\ m,n \geq 0,\ m \neq n\}.$

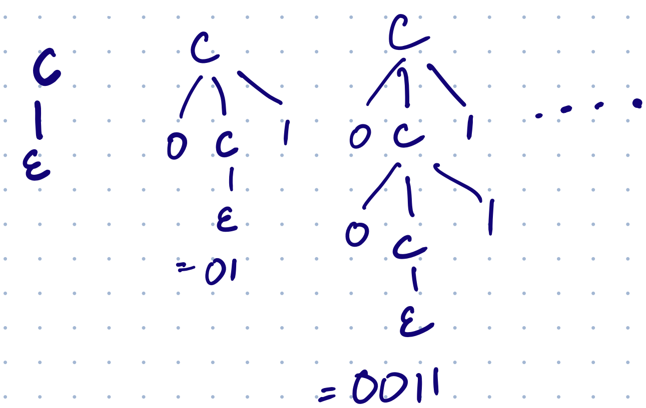
$L(A) = \{0^m 1^n,\ m,n \geq 0,\ n < m\}$

$L(B) = \{0^m 1^n,\ m,n \geq 0,\ n > m\}.$

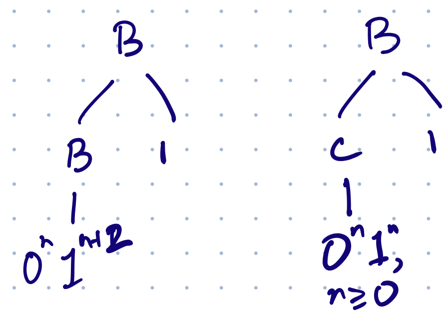
$L(C) = \{0^n 1^n,\ n \geq 0\}$

NOT REGULAR!

$L(C)$



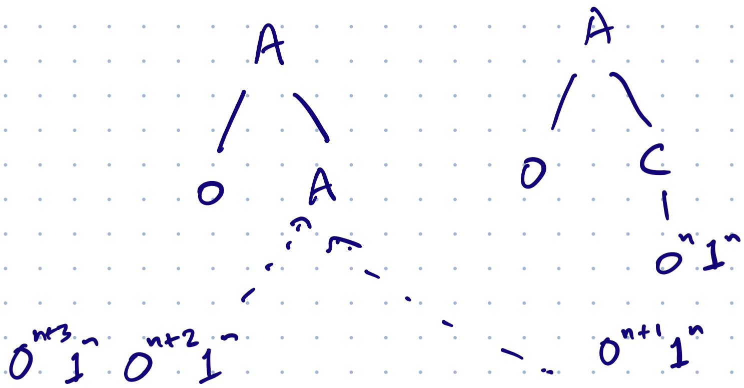
$$L(C) = \{0^i 1^i,\ i \geq 0\}$$



$L(B) =$

$\{0^n 1^{n+1},\ 0^n 1^{n+2},\ 0^n 1^{n+3}, \ldots\ldots\}$

$$L(B) = \{0^i 1^{i+j},\ i \geq 0,\ j > 0\}$$



$L(A) = \{0^{i+j}1^i,\ i \geq 0,\ j > 0\}.$

$\quad\ = \{0^m 1^n,\ m > n,\ m,n \geq 0\}$

Every regular language is also context-free languages.

Q: Is every CF language also regular?

No. 0^n1^n is CF but not regular.

$\boxed{\text{GOAL: } L(C) = \{0^n1^n \mid n \geq 0\}.}$

$C \to \varepsilon \mid 0C1 \quad = \{0^n1^n \mid n \geq 0\}.$

Lemma 1: $C \rightsquigarrow^* 0^n1^n$ for all $n \geq 0$.

$\left(\text{i.e. } \boxed{\overset{A}{0^n1^n} \in \overset{B}{L(C)}}\ \right) \qquad A \in B$

Proof: Let n be an arbitrary integer ≥ 0.

I.H. Assume $C \rightsquigarrow^* 0^m1^m$ for all $m < n,\ m \geq 0$

Two cases:

$n = 0. \quad 0^0 1^0 = \varepsilon \quad C \to \varepsilon$

$n = i$ for $i \geq 1$

$$C \to 0C1 \xrightarrow{IH} 0(0^{n-1}1^{n-1})1 = 0^n1^n$$

Therefore, $C \leadsto^{*} 0^n1^n$ for $n \geq 0$.

<u>Lemma 2</u> For all $w \in L(C)$, $w = 0^n1^n$ for some $n \geq 0$.
$\boxed{\left(L(C) \in \{0^n1^n, n \geq 0\} \right)}$. B ∈ A

<u>Proof:</u> Fix any $w \in L(C)$.

Assume $\forall x \in L(C), |x| < |w|$, $x = 0^m 1^m$ for some $m \geq 0$.

<u>Two cases:</u>

- $C \to \varepsilon \implies w = \varepsilon = 0^0 1^0 = 0^n1^n$ for $n = 0$.
- $C \to 0C1 \implies w = 0x1$ for some $x \in L(C)$
  $\overset{IH}{\implies} w = 0(0^m 1^m)1$ for some $m \geq 0$
  $= 0^{m+1} 1^{m+1}$.
  $= 0^n 1^n$ for $n = m+1$.

Thus $w = 0^n 1^n$ for some $n \geq 0$.

Lemma 1 & Lemma 2 imply

$$L(C) = \{0^n1^n \mid n \geq 0\}$$

Chomsky Normal Form

$S \to \varepsilon$

$A \to a$

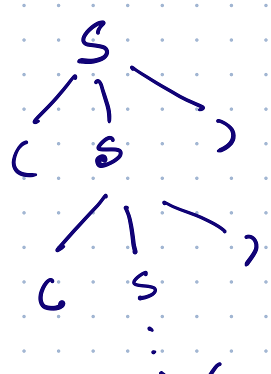
$A \to BC$

Q: Strings $w \in (0+1)^*$ with $\#(0,w) = \#(1,w)$.

$S \to \varepsilon \mid 0S1 \mid 1S0 \mid SS$ $\quad \underbrace{111000}\,\underbrace{000111}$

Q: $S \to \varepsilon \mid (S) \mid SS$

What is $L(S)$? Strings containing balanced paratheses.



$\rightsquigarrow (((\ldots))) \ldots = (^n)^n$

Q: strings w in $(\,(+)\,)^*$ where $\#((,w) = \#(),w)$

$S \to \varepsilon \mid (S) \mid)S(\mid SS$