Lecture #23

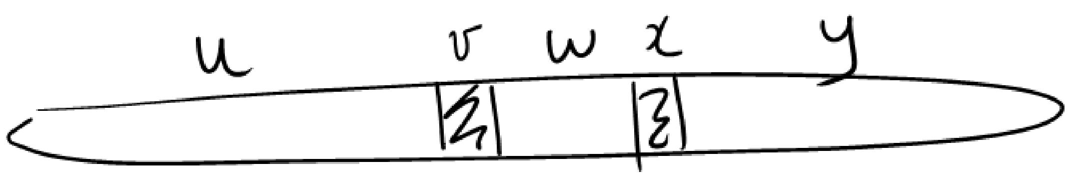
Non-context-free languages

Pumping lemma

Closure properties of CFLs.

If L is a CFL

$\exists n$



$w \in L$

$u\underline{\underline{v}}^i w \underline{\underline{x}}^i y \in L \quad \forall i.$

G - CFG is CNF

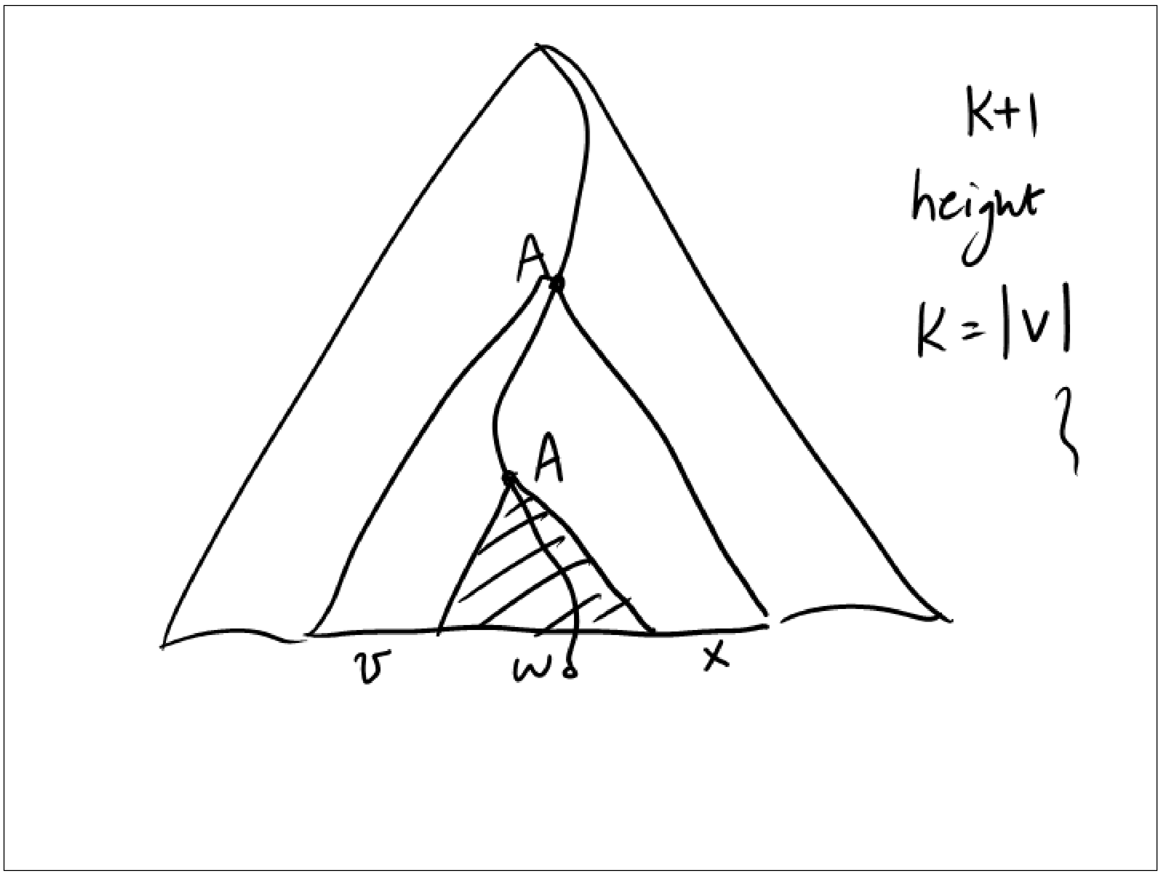
$A \to BC$

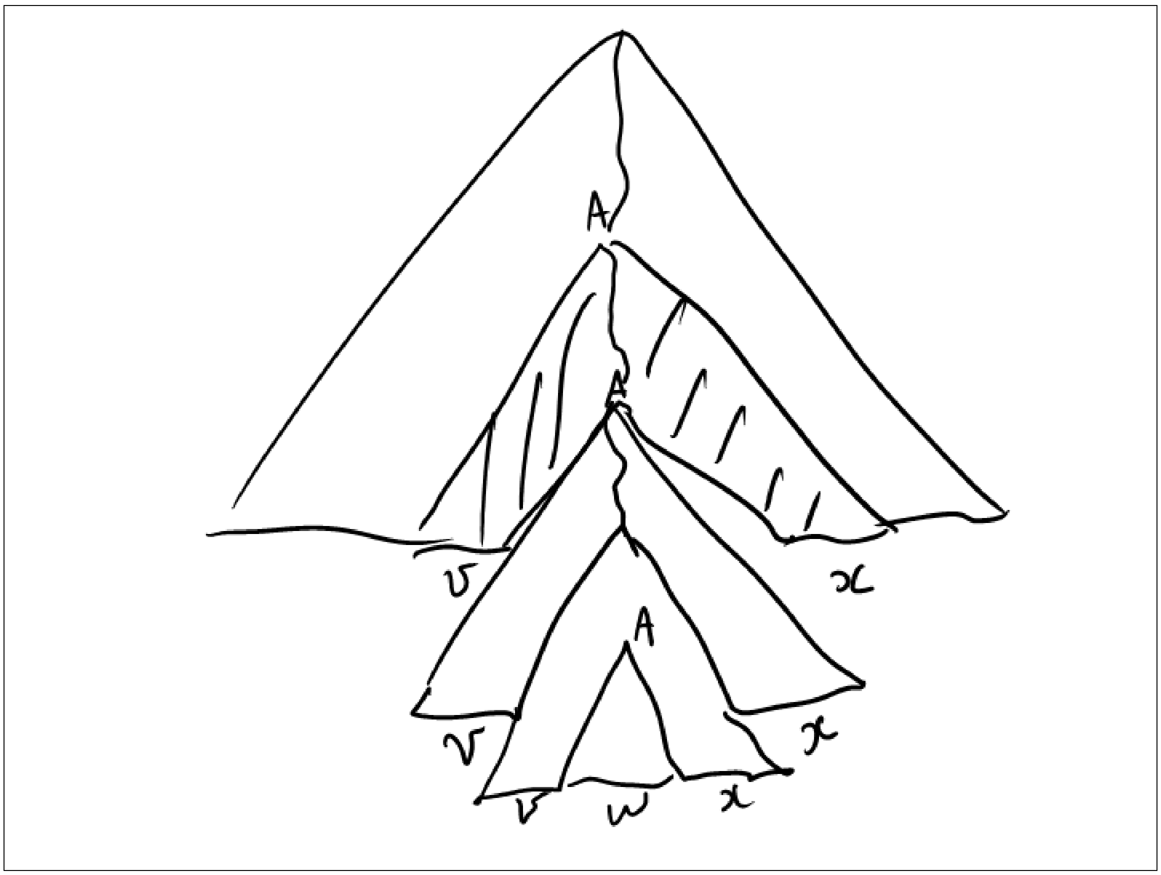
$A \to a$

w is large enough

then parse tree for w

will be tall enough


K+1
height
K=|V|
A
A
v
w
x


A
A
A
v
x
v
x
v
w
x

Lemma Let L be a CFL

Then $\exists n.$ s.t.

$\forall z \in L$, $|z| \geq n$

$\exists u, v, w, x, y \in \Sigma^*.$ $z = uvwxy$

- $|vx| \geq 1$
- $|vwx| \leq n$
- $\forall i.$ $uv^iwx^iy \in L$

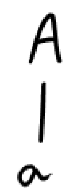
Let G be a CFG for L
in CNF.

"If $w \in L$ is long enough, then
parse tree for w will have
a long path".

<u>Claim</u>. If the parse tree for w has
no path longer than i, then
$|w| \leq 2^{i-1}$.

<u>Induction</u> We will show that if parse tree from
any var A derives w and has
no path longer than i, then $|w| \leq 2^{i-1}$
by induction on i.

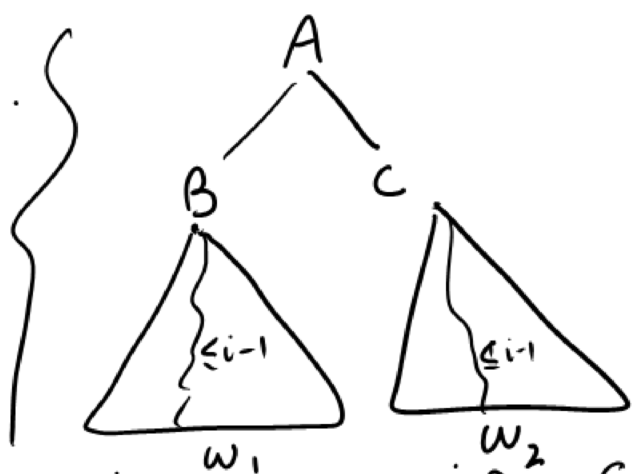
$\underline{i = 1}$



$|w| = 1$

$2^{i-1} = 1$

$\underline{i \geq 1}$



$A \rightarrow BC$

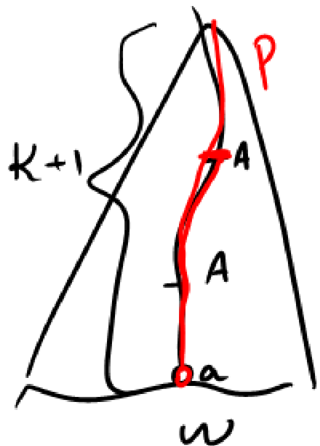
$w = w_1 w_2$

$|w_1| \leq 2^{i-2}$ $\quad |w_2| \leq 2^{i-2}$ $\quad$ So $|w| \leq 2^{i-2} \cdot 2 = 2^{i-1}$

Let $n = 2^K$

Let $w \in L$, $|w| \geq n$

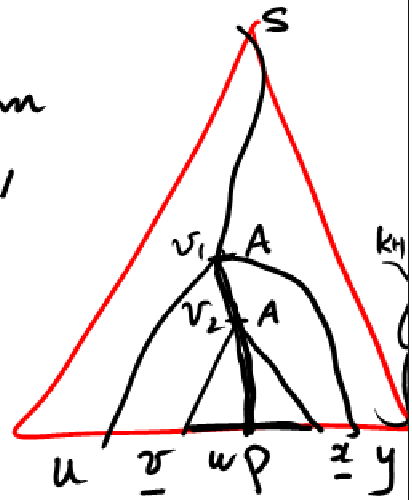
Any parse tree for w must have a path of length at least $K+1$



So there are $(K+1)$ nonterminals labeling this path — hence some non-terminal must repeat.

- Let P be (a) longest path in the tree
- Walk up this path till I find variables that repeat

Length of $\underline{\text{longest}}$ path from v_1 is at most $k+1$



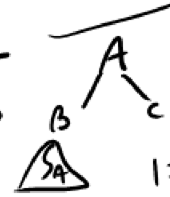
Let $S \Rightarrow^* uAy$

$\Rightarrow^* uvAxy$

$\Rightarrow uvwxy$

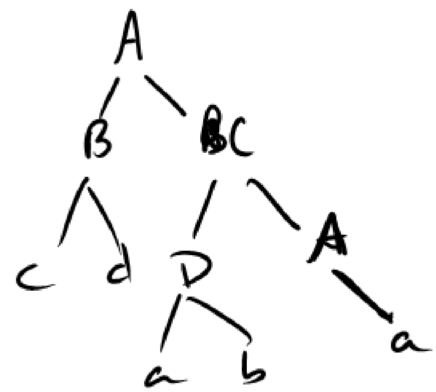
$|vwx| \leq 2^k = n$

To show $|vx| \geq 1$



$|x| \geq 1$

$|v| \geq 1$

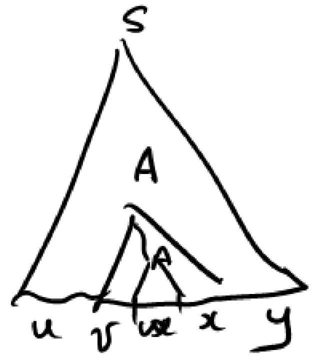

A
B
BC
c
d
D
A
a
b
a

$S \Rightarrow^* uAy$

$A \Rightarrow^* vAx$

$A \Rightarrow^* w$

uv^iwx^iy is also in L.
(for any $i \geq 0$)



$S \Rightarrow^* uAy$
$\Rightarrow^* uvAxy$
$\Rightarrow^* uvvAxxy$
$\dots\dots \Rightarrow^* \dots \Rightarrow^* uv^iAx^iy$
$\Rightarrow^* uv^iwx^iy$

COROLLARY. Let L be a CFL

Then $\exists n \in \mathbb{N}$ s.t.

$\forall z \in L$, $|z| \geq n$

$\exists u, t, z$ $\quad z = utz$ and $|t| \leq n$,

such that $\exists t'$,

t' is a contiguous substring of t

$|t'| < |t|$

and $ut'z \in L$.

Application $L = \{a^n b^n c^n \mid n \geq 0\}$ is <u>not</u> a CFL.

~~Pf~~ Assume L is a CFL.
Then PL-Cor holds, say for some n

$$w = a^{n+1} b^{n+1} c^{n+1}$$

By PL-cor, $\exists\, u, t, y$: $w = uty$, $|t| \leq n$ and $\exists t'$ which a substring of t and $ut'y \in L$.

Since t cannot have a's & b's and c's (since $|t| \leq n$) contracting t will remove some # of a's and not remove some # of b's etc.

$\cancel{a^n b^n}$

Case 1 $t \in a^*$

Then $ut'y$ has less a's than b's.

and so $ut'y \notin L$

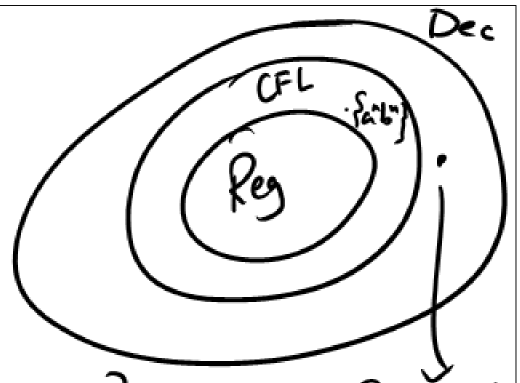
Case 2 & 3 $t \in b^*$, $t \in c^*$ ...

Case 4: $t \in \{a+b\}^*$

Then $ut'y$ will have more c's than # of a's or # of b's.

Case 5&6 : $t \in \{a+c\}^*$ $t \in \{b+c\}^*$ ...

CFLs are <u>not</u> closed under intersection.



$\{a^n b^n c^n \mid n \geq 0\}$

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

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

$$L_1 \cap L_2 = \{a^n b^n c^n \mid n \geq 0\} \rightarrow \text{not a CFL}$$

$S \rightarrow AB$
$A \rightarrow aAb$
$A \rightarrow \epsilon$
$B \rightarrow cB$
$B \rightarrow \epsilon$

Diagonalization

	w_1	w_2	w_3	$\cdots$
G_1				
G_2				
G_3				
$\vdots$				

$$L = \{ w_i \mid G_i \text{ does not generate } w_i \}$$

$$G_1 = (V_1, \Sigma, P_1, S_1)$$

Union

$$G_2 = (V_2, \Sigma, P_2, S_2)$$

$$G = \left(V_1 \cup V_2 \cup \{S\}, \Sigma, P_1 \cup P_2 \cup \begin{Bmatrix} S \to S_1, \\ S \to S_2 \end{Bmatrix}, S\right)$$

Complement.

$$L_1 \cap L_2 = \overline{\left(\bar{L_1} \cup \bar{L_2}\right)}$$

CFLs not closed under Complement.

	$\cup$	$\cap$	$-$	$\cdot$	$*$	rev	hom
CFL	✓	X	X	✓	✓	✓	✓

Membership : Decidable

Emptiness : Decidable.

$L(G) = \Sigma^*$ Undecidable

$L(G_1) \subseteq L(G_2)$ Undecidable.