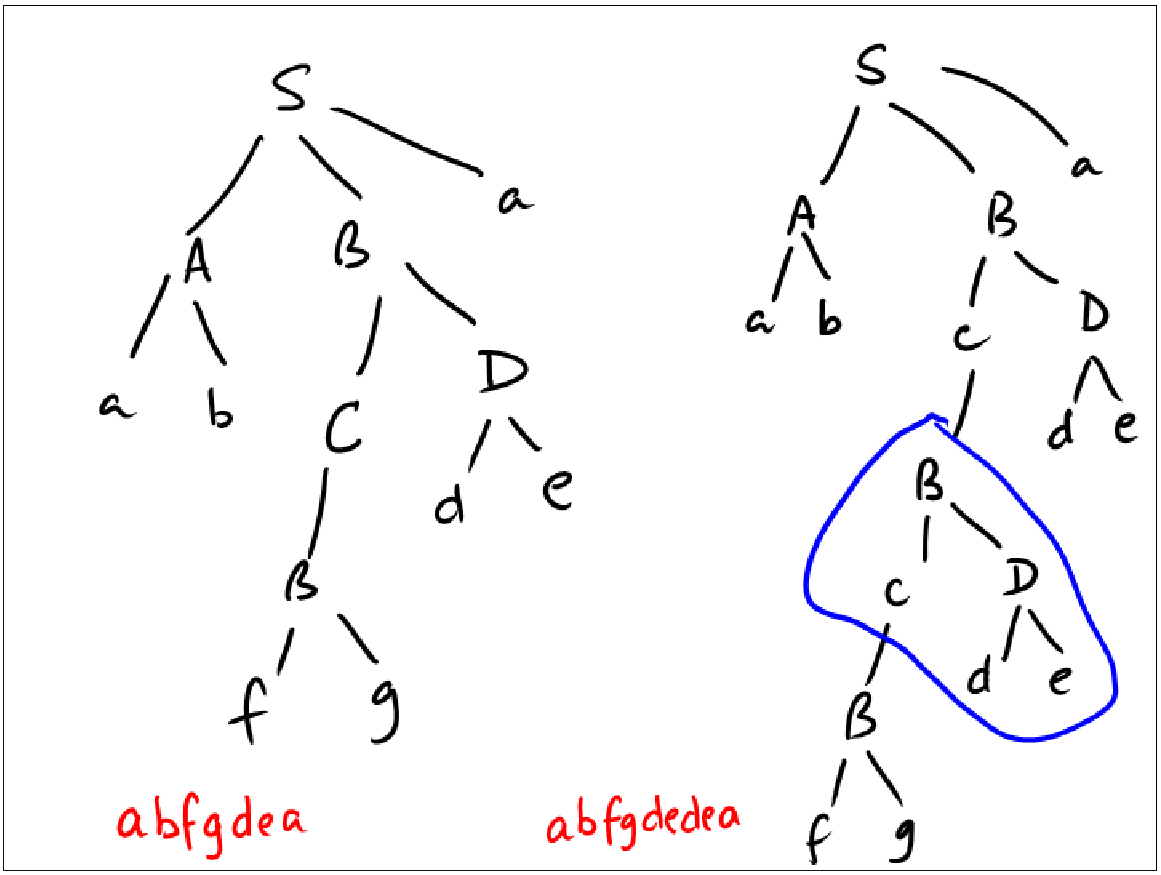
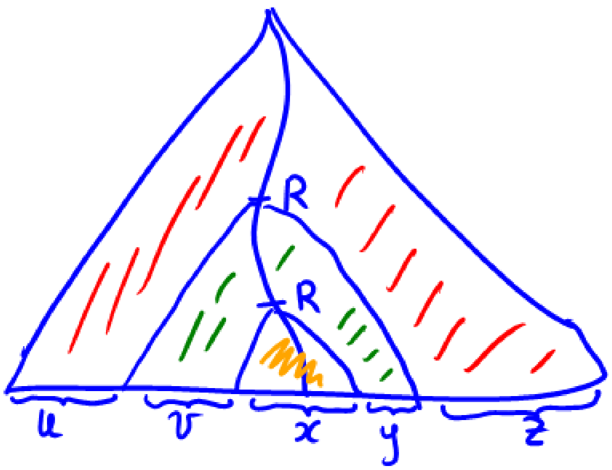
- Pumping lemma for CFLs
- Non-conter-free languages
- Non-closure of CFLs under intersection and complementation.

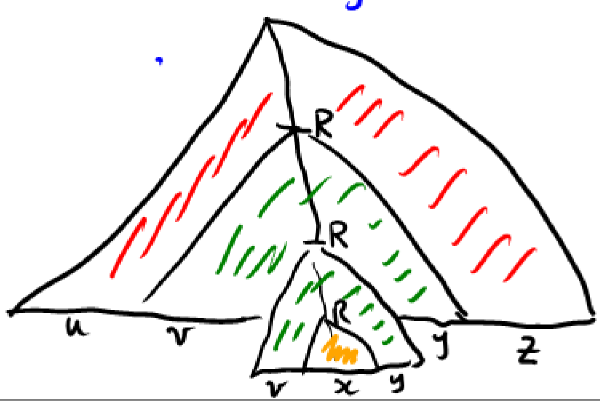

S
A
B
a
a
b
C
D
B
d
e
f
g
abfgdea
S
A
B
a
a
b
c
D
d
e
B
c
D
d
e
B
f
g
abfgdedea



If the word is long enough, the height of tree is long enough...

$R \Rightarrow^* vxy$

$R \Rightarrow^* x$



$uvvxyyz$

$= uv^2xy^2z$

So $\quad uv^3xy^3z \in L$

$uv^ixy^iz \in L$

In fact, $\quad uxz \in L$ as well.

Pumping lemma for CFLs.

If L is a CFL, then there is a $p \in \mathbb{N}$, such that for every string $s \in L$, $|s| \geq p$, there exists a split of s, $s = uvxyz$ such that:

1) $\forall i \geq 0.\ uv^ixy^iz \in L$.
2) $|vy| > 0$ (i.e. $vy \neq \epsilon$)
3) $|vxy| \leq p$

Example. $L = \{a^n c b^n \mid n \in \mathbb{N}\}$.

$\boxed{S \to aSb \mid c}$

$p = 3$

Let $s \in L$, $|s| \geq p$.

Then $s = a^n c b^n$. $(n \geq 1)$.

Eg. $s = aaacbbb = uvxyz$

$u = a$; $z = b$; $x = c$, $v = aa$; $y = bb$

But $|vxy| = 5 > 3 \longrightarrow$ $uv^ixy^iz = a(aa)^i c (bb)^i b \in L$

$u = aa$; $v = a$; $x = c$, $y = b$; $z = bb$

$|vxy| = 3 \leq p$, $|vy| > 0$. $uv^ixy^iz = aa(a)^i c (b)^i bb \in L$

$S = a^n c b^n$

Choose $u = a^{n-1}$; $v = a$; $x = c$; $y = b$; $z = b^{n-1}$

Then $u v^i x y^i z = a^{n-1} a^i c b^i b^{n-1}$

$= a^{n-1+i} c b^{n-1+i} \in L, \forall i \geq 0$

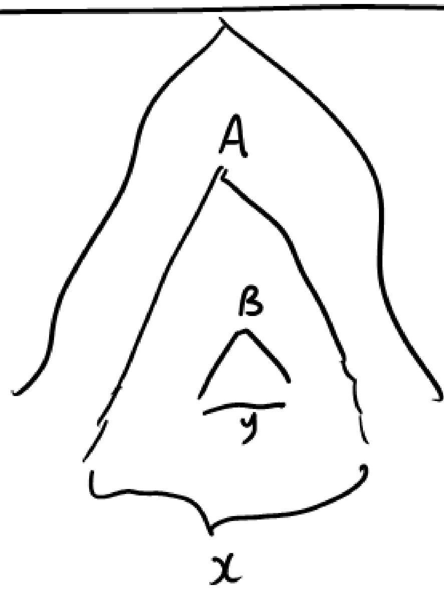
Also, $|vxy| = 3 \leq P$

$|vy| = 2 > 0$.

Hence all conditions of the PL are satisfied.

An observation of parse trees of CFGs in Chomsky Normal Form.

Let



$A \Rightarrow^* x$

$B \Rightarrow^* y$

Then $|x| > |y|$.

(Since no variables are <u>nullable</u>)

Proof of pumping lemma

Let L be a CFL.
Let $G = (V, \Sigma, R, S)$ be a CFG for L, in Chomsky Normal Form.
Let $p = 2^{|V|+1}$; we claim that for this p, the pumping lemma conditions are satisfied.
Let $s \in L$, $|s| \geq p$.
We must show that $\exists u, v, w, x, y$: $s = uvwxy$ and the three conditions in the pumping lemma are satisfied.

...

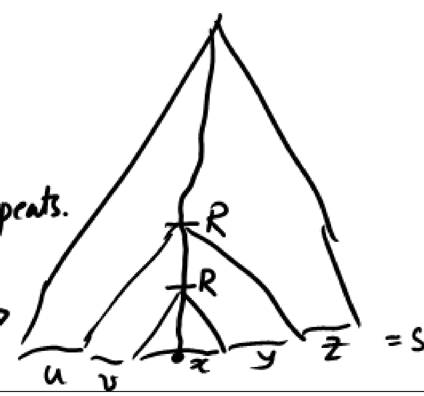
Take a parse tree for s.

Since this parse tree has $2^{|V|+1}$ leaves at least,
and since it is a binary tree, the height of
the tree must be at least $|V+1|$.

Hence there must be a path of length at least
$|V|+1$ in the tree; this path must have some
variable repeated. Pick a variable that is repeated
"lowest" in this path...

i.e. go from the leaf to the
root along the path and
pick the first variable R that repeats.

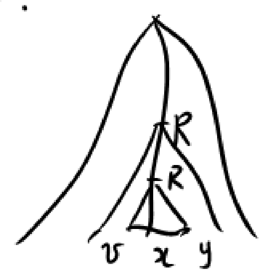
Pick u, v, w, x, y according to the figure →



Then $s = uvxyz$.

1) By previous arguments, we can pump this string and be within L.
   i.e. $\forall i.\ uv^i xy^i z \in L$.

2) $vy \neq \epsilon$.
   This is because G is in CNF.
   By earlier observation
   since the first R generates vxy
   and the second R generates x,
   $|vxy| > |x|$, i.e. $|vx| > 0$
   i.e. $vx \neq \epsilon$.



3) Finally, let's show $|vxy| \le p$.

Since we chose the lowest variable that repeated (along some path), the height of the subtree generating vxy is of height at most $|V|+1$.

Hence $|vxy| \le 2^{|V|+1} = p$.

Non-context-free languages

$L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ is not context-free.

Assume it is context-free.
Then PL holds, and there is a $p \in \mathbb{N}$ sat. conditions of PL.
Let $s = a^p b^p c^p \in L$, $|s| \geq p$.
Then by PL, there must be u, v, x, y, z st.
$s = uvxyz$, $\forall i.\ uv^ixy^iz \in L$,
$|vxy| \leq p$, $|yv| > 0$.
Case I: If v (or y) have two types of letters,
then $uv^2xy^2z \notin a^*b^*c^*$
So $uv^2xy^2z \notin L$.

<u>Case 2</u>. v has one type of character
y has one type of character

Then uv^2xy^2z
does not have equal no. of a's & b's & c's.

So $uv^2xy^2z \notin L$.

So no such partition $uvxyz$ exists.
Hence PL fails & L is not a CFL.

$L_1 = \{a^n b^n c^i \mid n, i \in \mathbb{N}\}$
$L_2 = \{a^i b^n c^n \mid n, i \in \mathbb{N}\}$ CFLs.

$L_1 \cap L_2 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ is not a CFL.

Hence the class of CFLs is not closed under intersection.