Sipser Section 2.3 Languages that are not context free

Claim $L = \{a^n b^n c^n \mid n \geq 0\}$ is not context-free

Q: can we use pumping of some kind

A: yes, but more complicated

Theorem 2.34

For every CFL L, there exists $p \in \mathbb{N}$ s.t $\forall w \in L$ with $|w| \geq p$ $\exists u, v, x, y, z \in \Sigma^*$ s.t $w = uvxyz$ and

(1) $uv^ixy^iz \in L$ $\forall i \geq 0$

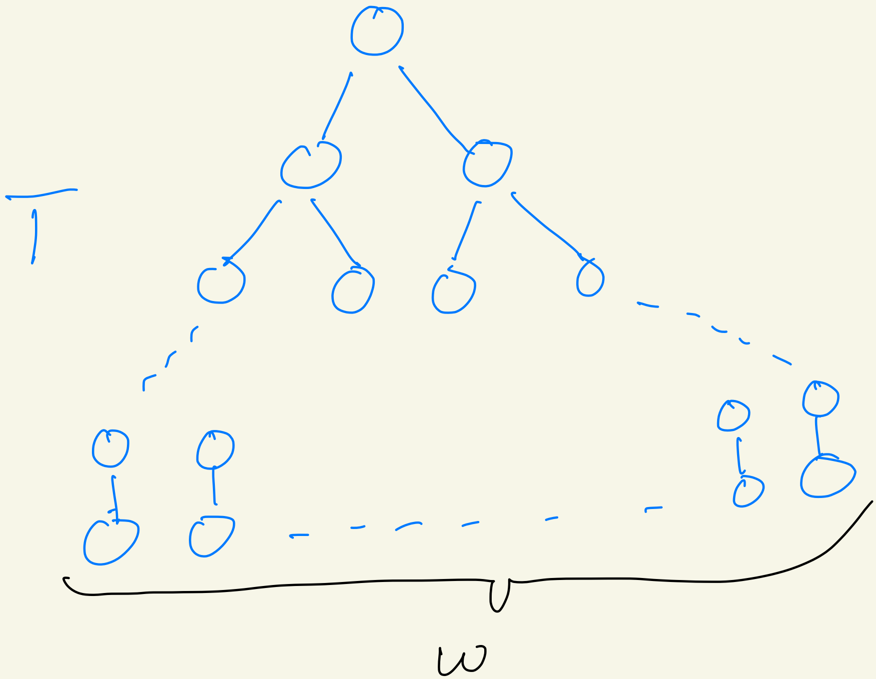
(2) $|vy| > 0$

(3) $|vxy| \leq p$

P: As L is a CFL then exists a CFG $G = (V, \Sigma, R, S)$ in Chomsky normal form such that $L = L(G)$

Let $p = 2^{|V|+1}$ (pumping length)

Suppose $|w| \geq p$ and look at a parse tree for a derivation of w in G



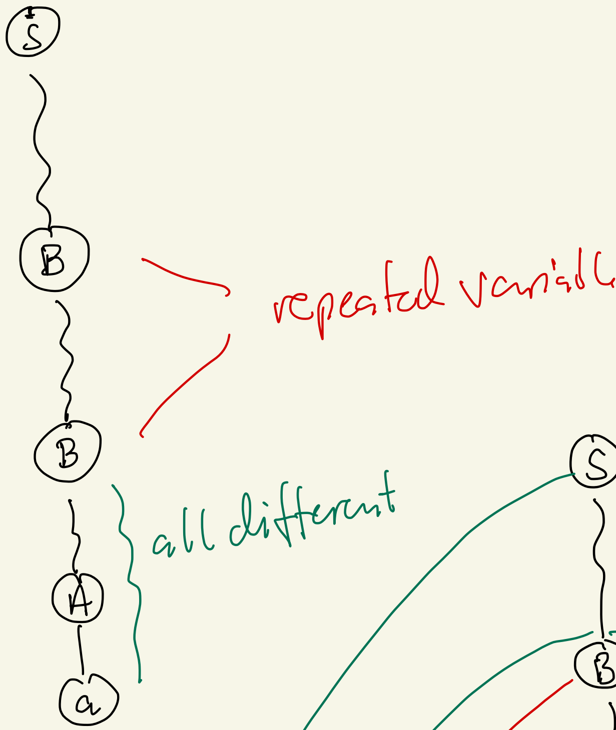
so $\geq 2^{|V|+1}$ leaves in T

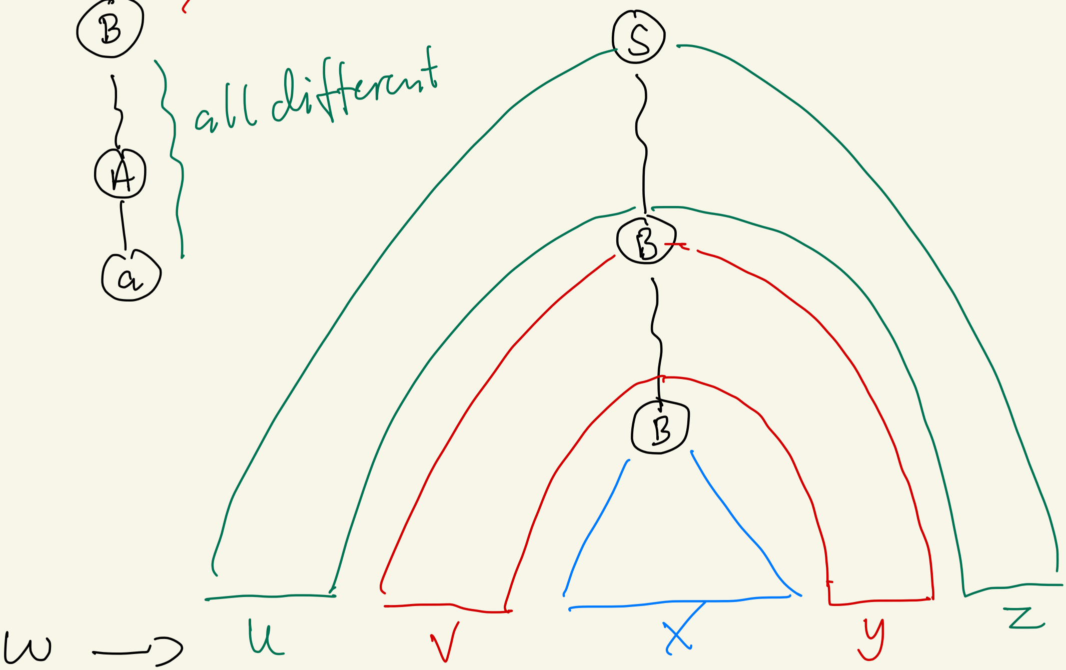
Thus the height of T is at least $|V|+2$ as the last step is of the form $A_i \to a$

Let $h(T)$ be the height of T

so $h(T) \geq |V|+2$

look at a path of length $h(T)$





$uv^ixy^iz \in L$

$|vy| > 0$

$|vxy| \leq p$

Example 2.36

$B = \{a^n b^n c^n \mid n \geq 0\}$ is not a CFL

P: suppose $L = L(G)$ for a Chomsky grammar $G = (V, \Sigma, R, S)$ and let $p = 2^{|V|+1}$

- We take $w = a^p b^p c^p$
- adversary shows $w = uvxyz$ with $|vy| > 0$ and $|vxy| \leq p$
- we claim that $uv^0xy^0z = uxz \notin L$

$aa \cdots abb \cdots bcc \cdots c$

$|vxy| \leq p \Rightarrow$ cannot have both a and c in vxy

- Case 1 $vxy \in a^p b^p$

  Then $uxz \notin L$ is it has more c's than either a's or b

Case 2 $vxy \in b^p c^p$

$\Downarrow$

$uxz \notin L$ (as in case 1)

Conclusion: no matter what partition the adversary tries we find that $uxz \notin L$

so G cannot exist, implying that L is not context-free

Example 2.37 $C = \{a^i b^j c^k \mid 0 \leq i \leq j \leq k\}$ is not a CFL

P: suppose $C = L(G)$ for some Chomsky grammar $G = (V, \Sigma, R, S)$ and let $p = 2^{|V|+1}$

We choose $s = a^p b^p c^p$

adversary: $s = uvxyz$ s.t $|vy| > 0$ and $|vxy| \leq p$

Observation: neither v nor y can contain two different symbols

$$\underbrace{aa\cdots a\,b}_{v}b\cdots b\,cc\cdots c \Rightarrow uv^2xy^2z \notin L$$

Case 1 no a in the string vy

$$aa\cdots abb\cdots \underbrace{b\,c}_{vxy}c\cdots c \Rightarrow uxz \notin L$$

Case 2 no b in vy

$$\underbrace{aa\cdots}_{vxy}abb\cdots bcc\cdots c \Rightarrow uv^2xy^2z \notin L$$

So the string vy contains both an a and a b

$$aa\cdots \underbrace{a}_{v}\underbrace{bb\cdots}_{x}\underbrace{b}_{y}cc\cdots c \Rightarrow uv^2xy^2z \notin L$$

Example 2.38

$D = \{ww \mid w \in \{0,1\}^*\}$

Suppose $D = L(G)$ for some Chomsky CFG $G = (V, \Sigma, R, S)$

We take $s = \underbrace{0^p 1^p}_{w} \underbrace{0^p 1^p}_{w}$ when $p = 2^{|V|+1}$

adversary gives us $s = uvxyz$ st $|vy| > 0$ and $|vxy| \le p$

If $\overset{w}{0 \cdots \underbrace{0 1 \cdots 1}_{vxy}} \overset{w}{0 \cdots 0 1 \cdots 1}$ or $\overset{w}{0 \cdots 0 1 \cdots 1} \overset{w}{0 \cdots \underbrace{0 1 \cdots 1}_{vxy}}$

Then $uv^2xy^2z \notin L$ So $0 0 \cdots 0 \underbrace{1 \cdots 1 \, 0 \cdots 0}_{vxy} 1 \cdots 1$

Now $uxz = 0^p 1^i 0^j 1^p$ for some i, j with

$1 \le i + j < 2p$

Hence $uxz \notin L$ ↯