CS 6041
Theory of Computation

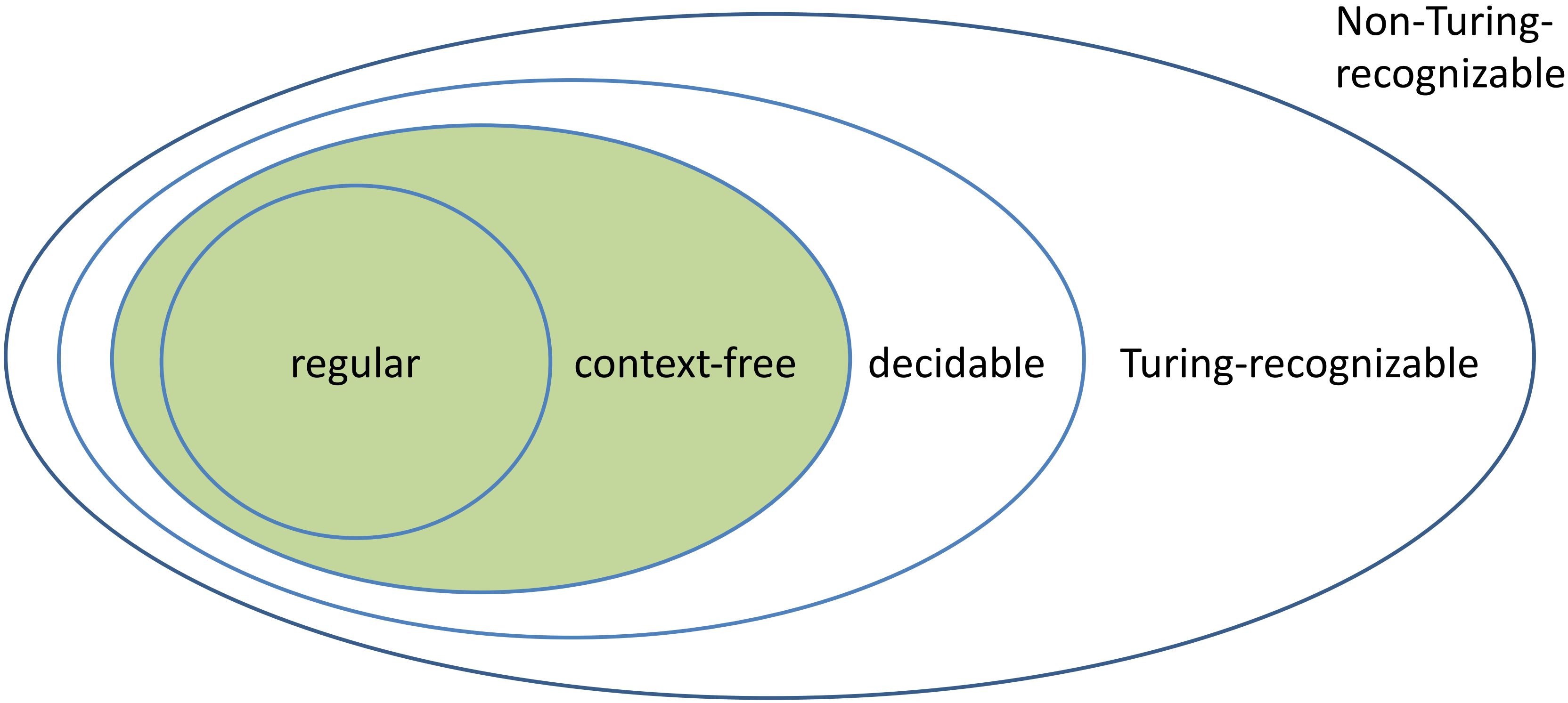
Non-context-free language

Kun Suo

Computer Science, Kennesaw State University

https://kevinsuo.github.io/

Non-context-free language



Non-context-free language

- A= { 0^n1^n | $n \geq 0$ }

  Context-free language

- Why A is context-free?
  - $G_1=(\{S\},\{0,1\}, \{S \rightarrow 0S1, S \rightarrow \varepsilon\}, S)$

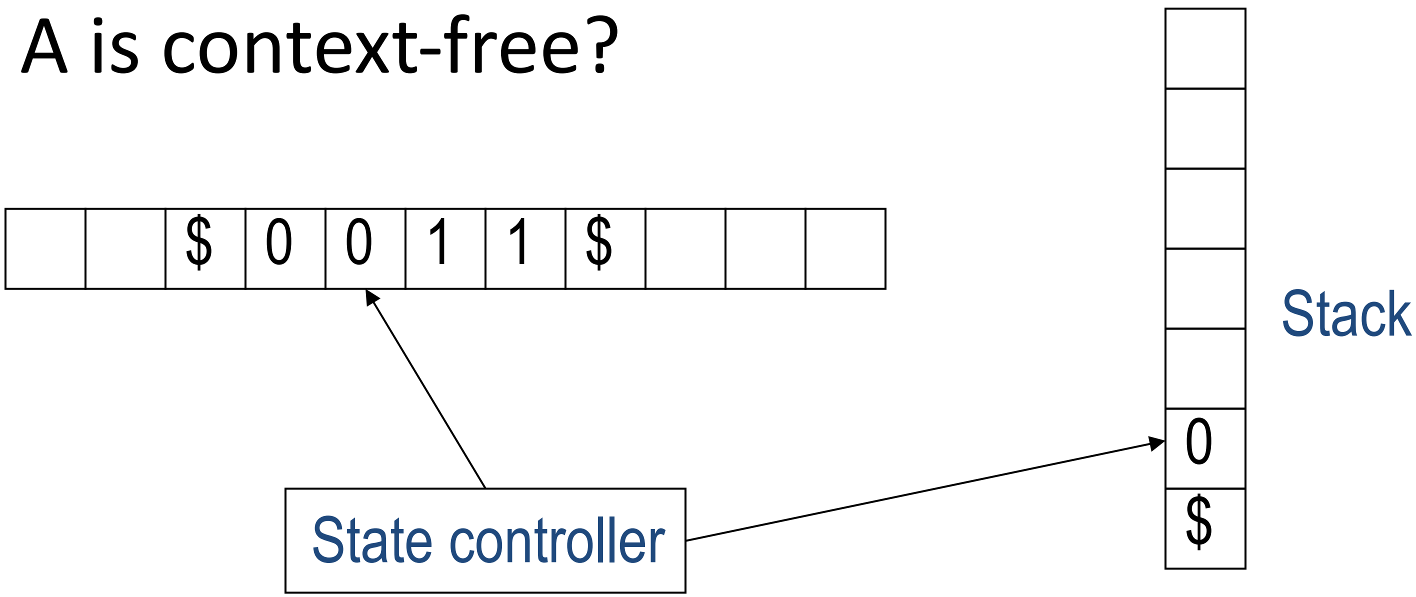
Non-context-free language

- $A = \{ 0^n1^n \mid n \geq 0 \}$

  Context-free language

  $\{ 0^n1^n \mid n \geq 0 \}$

- Why A is context-free?



Non-context-free language

- $A = \{ 0^n1^n \mid n \geq 0 \}$

  Context-free language

- $B = \{ a^nb^nc^n \mid n \geq 0 \}$

  Non-context-free language

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

  Non-context-free language

Pumping lemma

Suppose A is CFL,

then there exist a number p(the pumping length) where,

if $s \in A$ and $|s| \geq p$, then $s =$ **uvxyz**,

Satisfying the following

1) $\forall i \geq 0$, $uv^ixy^iz \in A$;

2) $|vy|>0$;

3) $|vxy| \leq p$.

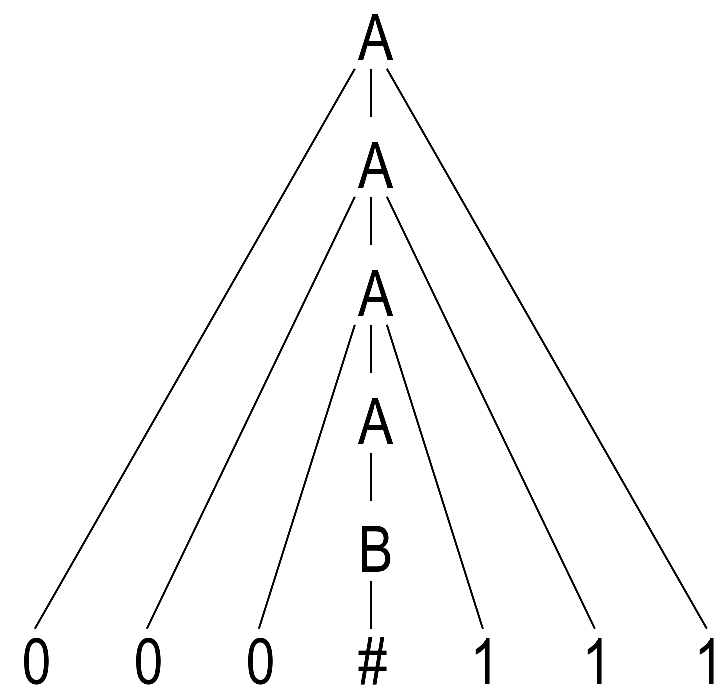
Parse tree of CFL

- Grammar G_1:

  $A \rightarrow 0A1$

  $A \rightarrow B$

  $B \rightarrow \#$



- Derivation: $A \Rightarrow 0A1 \Rightarrow 00A11$

  $\Rightarrow 000A111 \Rightarrow 000B111 \Rightarrow 000\#111$

Pumping lemma

Suppose A is CFL,

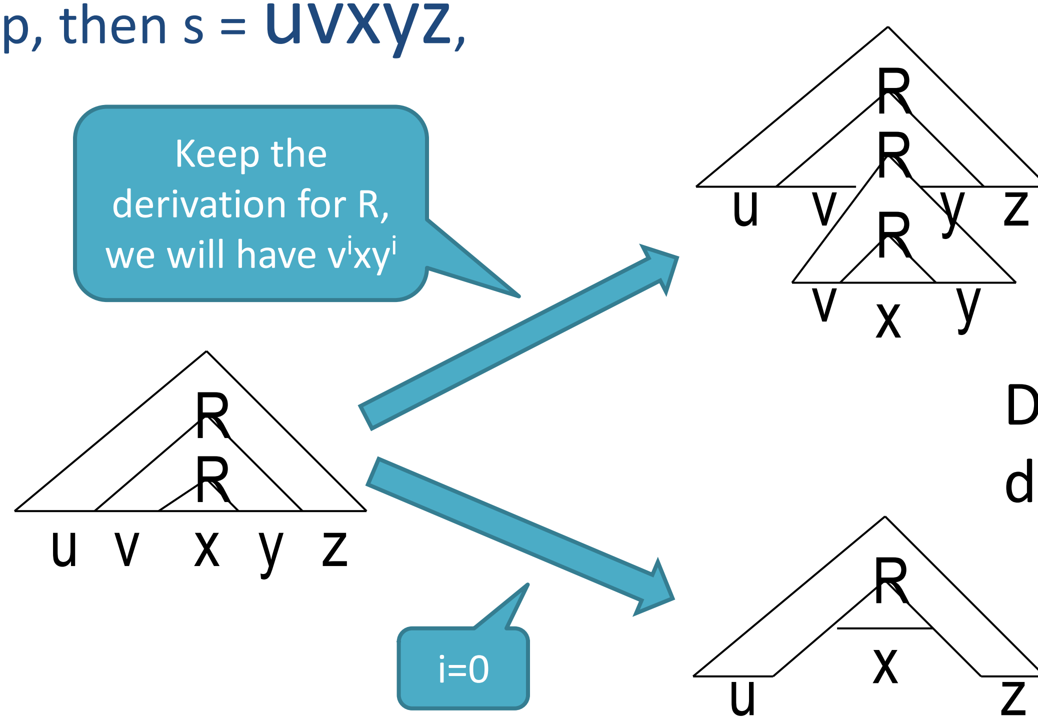
then there exist a number p(the pumping length) where,

if $s \in A$ and $|s| \geq p$, then s = **uvxyz**,

$uv^ixy^iz \in A$;

$R \rightarrow vRy$

$R \rightarrow x$



Different derivations

Pumping lemma proof

b is 3 for Grammar G_1:

$A \rightarrow 0A1$

$A \rightarrow B$

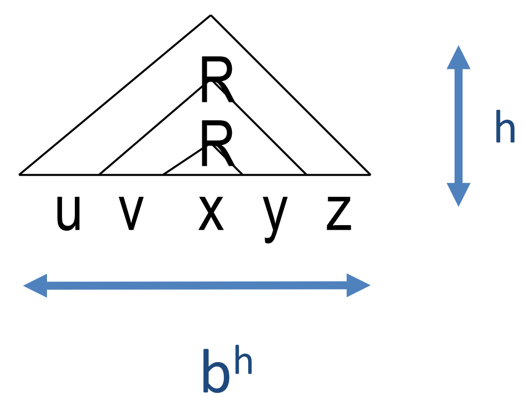
$B \rightarrow \#$

- Suppose G is A's CFG.

Let b is the longest length of right part of rule ($b \geq 2$)

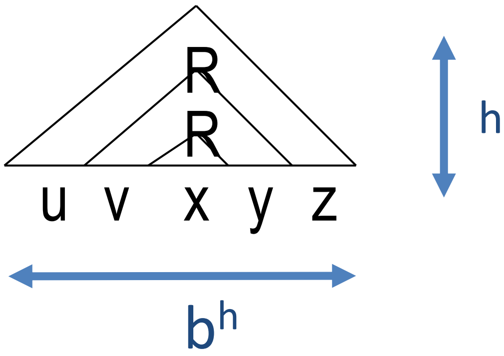
in parse tree of G, every node has at most b children.

For parse tree with h height, the length of string which it generates will not longer than b^h.

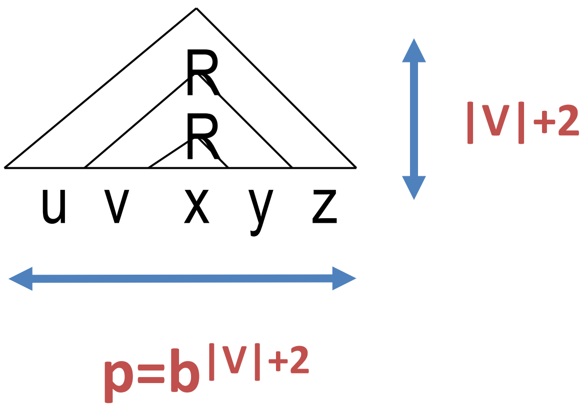


Pumping lemma proof

For parse tree with h height, the length of string which it generates will be not longer than b^h.



Suppose G has |V| variables, and let $p=b^{|V|+2}$, then for string which length is no less than p, its parse tree height is at least **|V|+2**



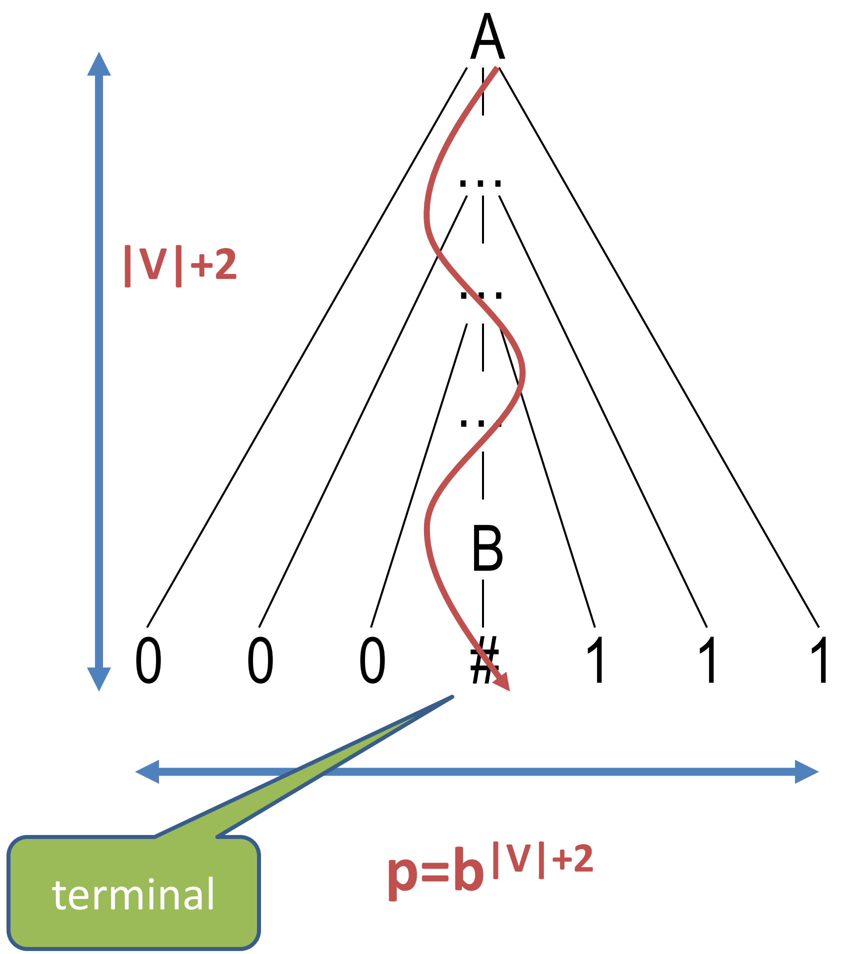
Pumping lemma proof

Suppose s is a string, $|s| \geq p$, and s has the minimum leaf nodes in all its parse tree, then

the height of parse tree for s is no less than **$|V|+2$**

As the leaf is terminal, so the variable in the path is no less than **$|V|+1$**

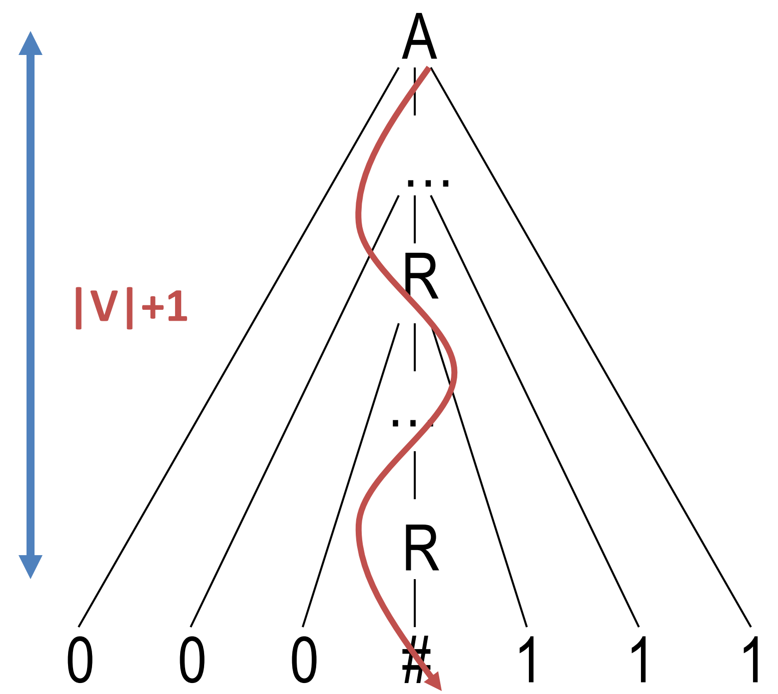
(due to $|V|+2 - 1$)



Pumping lemma proof

Based on pigeonhole principle, there must be **one variable** that appears more than once.

Suppose the last repeated variable in the path is **R**



Pumping lemma proof

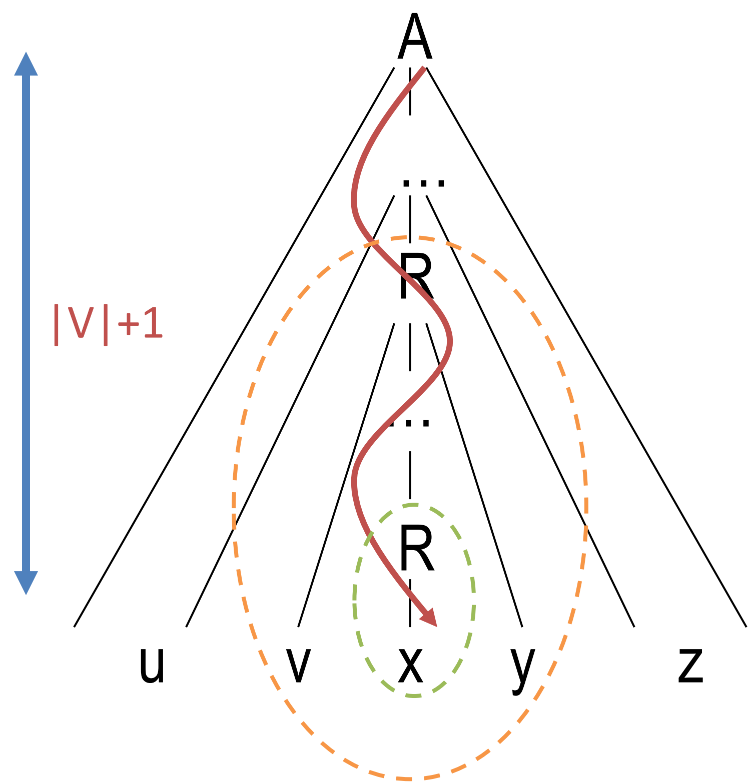
Divide s into uvxyz;

The bottom of R has smaller subtree generating x;

The top of R has larger subtree generating vxy;

As the bottom of R could have the same derivation of the top R, therefore,

$\forall i \geq 0$, $uv^ixy^iz \in A$



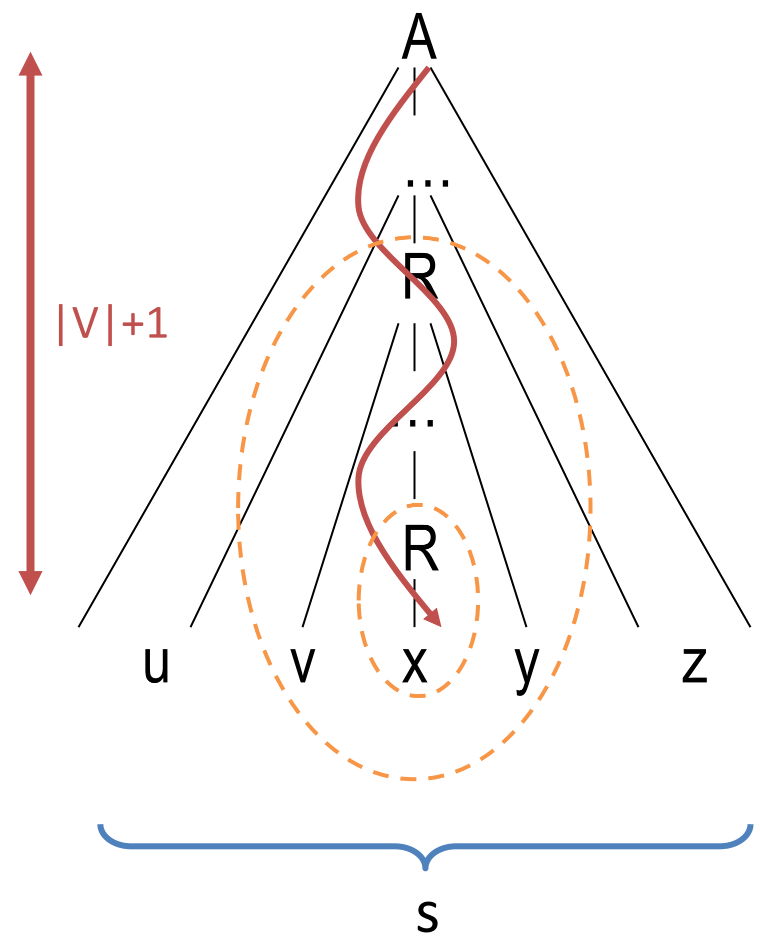
Pumping lemma proof

v and y cannot be empty string at the same time

Because if that happens, we can use the smaller subtree to replace the larger subtree to get s.

However, that is contradicted with that the parse tree has the minimum nodes. Thus,

$|vy|>0$



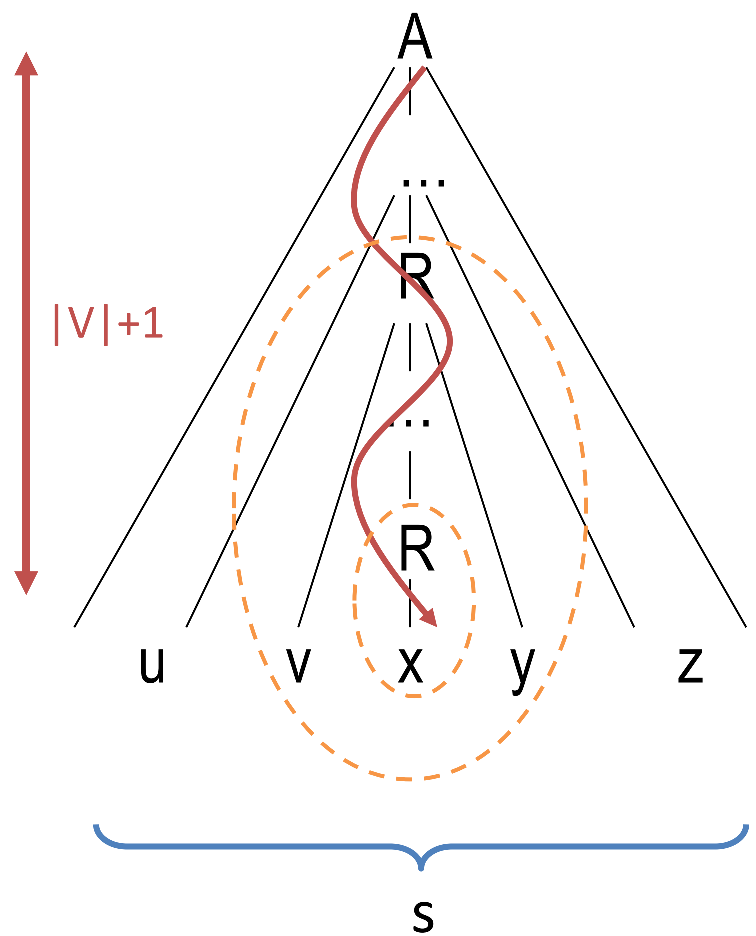
Pumping lemma proof

As the height of subtree generating vxy is no more than **|V|+2**, (R could be at most as A)

thus maximum length of string this subtree can generate is no more than **$b^{|V|+2}$=p**

$|vxy| \leq p$

#



Non-context-free language

- $A = \{ 0^n1^n \mid n \geq 0 \}$

  Context-free language

- $B = \{ a^nb^nc^n \mid n \geq 0 \}$

  Non-context-free language

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

  Non-context-free language

Example: $B=\{ a^nb^nc^n \mid n\geq 0 \}$

1) $\forall i\geq 0,\ uv^ixy^iz\in A$;
2) $|vy|>0$;
3) $|vxy|\leq p$.

- Proof:

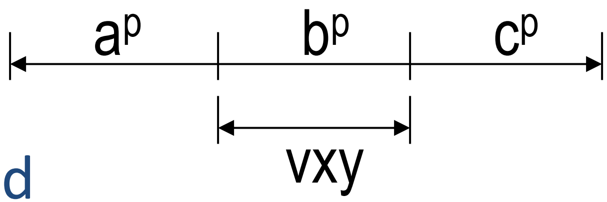
Suppose B is CFL, p is the pumping length,

let $s=a^pb^pc^p$

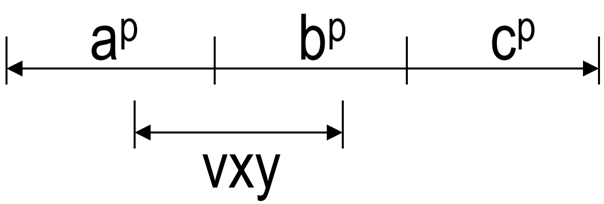
Then s = uvxyz, that

$\forall i\geq 0,\ uv^ixy^iz\in B$;

$|vy|>0$, v and y have at least one kind of symbol;



$|vxy|\leq p$, v and y have at most two kinds of symbol;

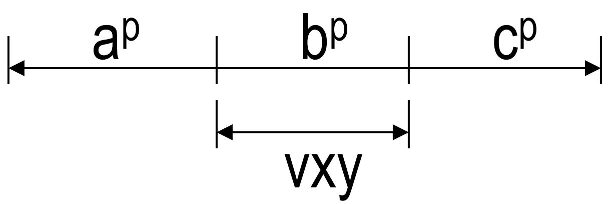


Example: $B=\{ a^n b^n c^n \mid n \geq 0 \}$

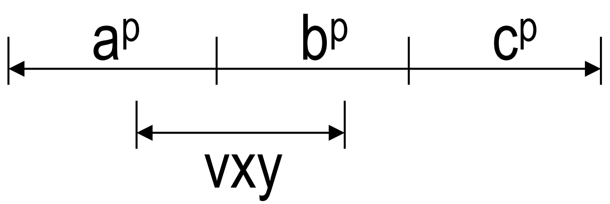
1) $\forall i \geq 0,\ uv^i xy^i z \in A$;
2) $|vy|>0$;
3) $|vxy| \leq p$.

- Proof:

If v and y have one kind of symbol, then in $uv^i xy^i z$ ($i>1$), a/b/c has different numbers;



If v and y have two kinds of symbol, then in $uv^i xy^i z$ ($i>1$), a/b/c has different numbers;



Contradiction.

Non-context-free language

- $A = \{ 0^n 1^n \mid n \geq 0 \}$

  Context-free language

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

  Non-context-free language

- $C = \{ a^i b^j c^k \mid 0 \leq i \leq j \leq k \}$

  Non-context-free language

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

  Non-context-free language

Example: C = { $a^ib^jc^k$ | $0 \le i \le j \le k$ }

1) $\forall i \ge 0$, $uv^ixy^iz \in A$;
2) $|vy|>0$;
3) $|vxy| \le p$.

- Proof:

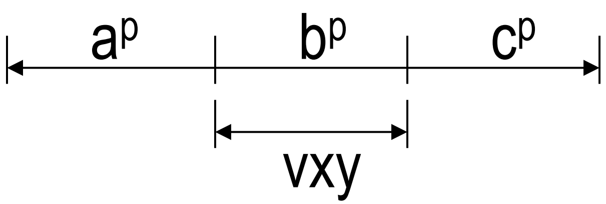
  Suppose C is CFL and p is pumping length, let **$s=a^pb^pc^p$.**

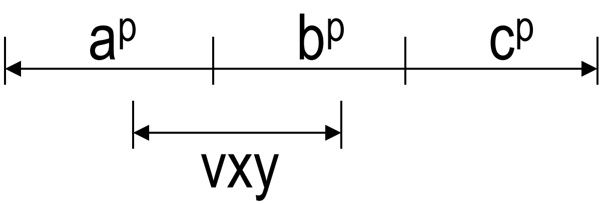
  Then **s=uvxyz,** satisfying that

  $\forall i \ge 0$, $uv^ixy^iz \in C$;

  $|vy|>0$, v and y have at least one symbol;

  $|vxy| \le p$, v and y have at most two symbols.





Example: $C = \{ a^i b^j c^k \mid 0 \le i \le j \le k \}$

1) $\forall i \ge 0,\ uv^i xy^i z \in A$;

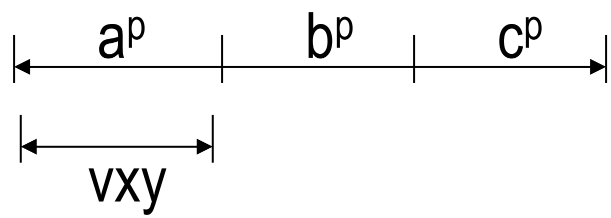
2) $|vy| > 0$;

3) $|vxy| \le p$.

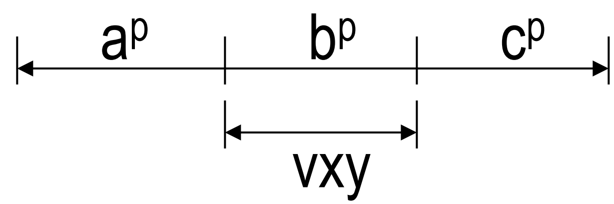
- Proof:

If v and y have one symbol, then

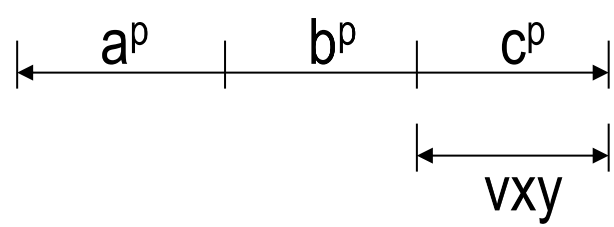
(1) v and y have a, then $i \ge 0$, $uv^i xy^i z \notin C$ because the number of a is larger than b and c;



(2) v and y have b, then $i \ge 0$, $uv^i xy^i z \notin C$ because the number of b is larger than c;



(3) v and y have c, then $uxz \notin C$ because the number of c is less than a and b;



Example: $C = \{ a^i b^j c^k \mid 0 \le i \le j \le k \}$

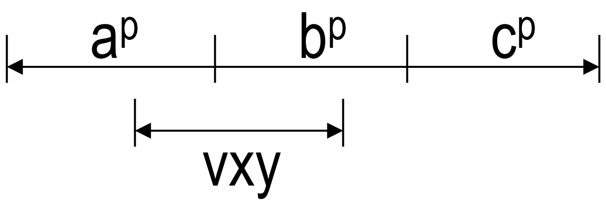
1) $\forall i \ge 0$, $uv^i xy^i z \in A$;

2) $|vy| > 0$;

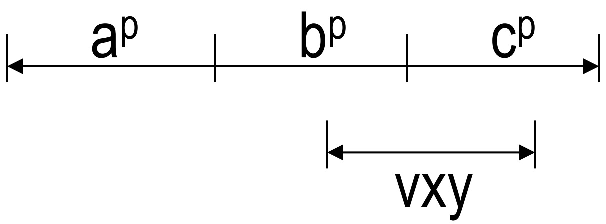
3) $|vxy| \le p$.

- Proof:

  If v and y have two symbols (a and b), then $i \ge 0$, $uv^i xy^i z \notin C$ because the number of a and b are larger than c;



  If v and y have two symbols (b and c), then $i=0$, $uxz \notin C$ because the number of c is less than a;



  Contradiction!

Example:
D={ww|w∈{0,1}*}

1) $\forall i \geq 0$, $uv^ixy^iz \in A$;

2) $|vy|>0$;

3) $|vxy| \leq p$.

- Proof:

Suppose D is CFL and p is the pumping length

Let $s = 0^p1^p0^p1^p$, then $s=uvxyz$, $|vxy| \leq p$, $uv^ixy^iz \in D$

Discuss D depends on the position of vxy

$$\overbrace{00\ldots00\ 11\ldots11}^{0^p1^p}\ \overbrace{00\ldots00\ 11\ldots11}^{0^p1^p}$$

u vxy z

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

1) $\forall i \geq 0,\ uv^ixy^iz \in A$;

2) $|vy| > 0$;

3) $|vxy| \leq p$.

- Proof:

(1) If vxy is at the first half of ww, then

in uv^2xy^2z, the second-half starts with 1 while the first-half starts with 0

uv^2xy^2z is not in form of ww. Contradiction!

$$\overbrace{00\ldots00\ 1\underbrace{1\ldots11}_{vxy}}^{0^p1^p}\ \overbrace{00\ldots00\ 11\ldots11}^{0^p1^p}$$

u z

Example:
D={ww|w∈{0,1}*}

1) $\forall i \geq 0,\ uv^ixy^iz \in A$;

2) $|vy|>0$;

3) $|vxy| \leq p$.

- Proof:

(2) If vxy is at the second half of ww, then

in uv^2xy^2z, the second-half ends with 1 while the first-half ends with 0

uv^2xy^2z is not in form of ww. Contradiction!

$$\overbrace{00\ldots00\ 11\ldots11}^{0^p1^p}\ \overbrace{\underbrace{00\ldots00}_{vxy}\ 11\ldots11}^{0^p1^p}$$

u z

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

1) $\forall i \geq 0,\ uv^ixy^iz \in A$;

2) $|vy|>0$;

3) $|vxy| \leq p$.

- Proof:

(3) If vxy is at the middle of ww containing both 1s and 0s, then

$uv^0xy^0z=uxz=0^p1^i0^j1^p$ $(i < p, j < p)$

$0^p1^i0^j1^p$ is not in form of ww. (First half has more 0 than second half)

Contradiction!

$$\overbrace{00\ldots00\ 11\ldots11}^{0^p1^p}\ \overbrace{00\ldots00\ 11\ldots11}^{0^p1^p}$$

u vxy z

CFL operation

- CFL is closure on union ($A \cup B$) operation
- Proof:

Let L_1 and L_2 be generated by the CFG, $G_1 = (V_1, T_1, P_1, S_1)$ and $G_2 = (V_2, T_2, P_2, S_2)$, respectively

Define the CFG, G, that generates $L_1 \cup L_2$ as follows:

G = ($V_1 \cup V_2 \cup \{S\}$,

$T_1 \cup T_2$,

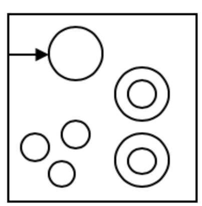
$P_1 \cup P_2 \cup \{S \rightarrow S_1 \mid S_2\}$,

S).

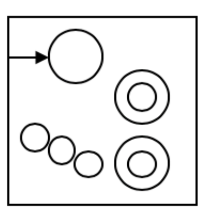
Review

- Design CFG for $\{w|w=0^n1^n \text{ or } w=1^n0^n, n\geq 0\}$
  - Design CFG for $\{w|w=0^n1^n, n\geq 0\}$
    - $G_1=(\{S_1\},\{0,1\},\{S_1\rightarrow 0S_11, S_1\rightarrow \varepsilon\},S_1)$
  - Design CFG for $\{w|w=1^n0^n, n\geq 0\}$
    - $G_2= (\{S_2\},\{0,1\},\{S_2\rightarrow 1S_20, S_2\rightarrow \varepsilon\},S_2)$
  - $G=(\{S,S_1,S_2\},\{0,1\},$
    $\{ S\rightarrow S_1,\ S\rightarrow S_2,$
    $S_1\rightarrow 0S_11,\ S_1\rightarrow \varepsilon,$
    $S_2\rightarrow 1S_20,\ S_2\rightarrow \varepsilon\},$
    $S)$

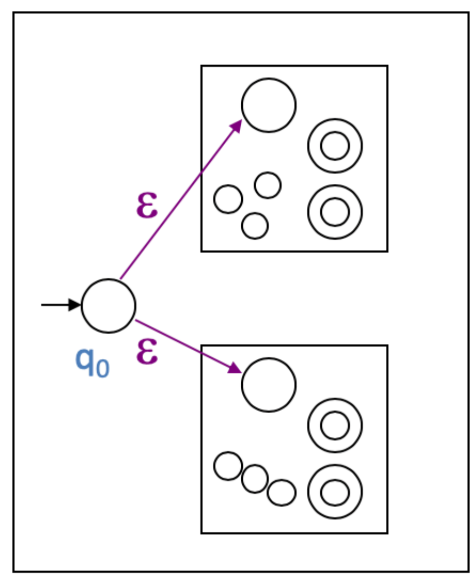
N_1



N_2



N



CFL operation

- CFL is closure on union ($A \cup B$) operation

- CFL is not closure on intersection ($A \cap B$) operation
  - $A=\{ a^n b^n c^m \mid n,m \geq 0 \}$ is CFL
  - $B=\{ a^m b^n c^n \mid n,m \geq 0 \}$ is CFL
  - $A \cap B=\{ a^n b^n c^n \mid n \geq 0 \}$ is not CFL (using pumping lemma)

- CFL is not closure on complement ($\overline{A}$) operation

CFL operation

- CFL is not closure on complement ($\bar{A}$) operation

- Proof:

  Assume the complement of CFL is also a CFL

  Let L_1 and L_2 be two CFLs

  Then $\overline{L_1}$ and $\overline{L_2}$ are also two CFLs

  Because CFL is closure on union, then $\overline{\overline{L_1} \cup \overline{L_2}} = L_1 \cap L_2$ is also a CFL, contradiction!

Operation on languages

	RL: DFA/NFA/RE	CFL: CFG/PDA	TM
Union	close	close	?
Concatenation	close	close	?
Intersection	close	not close	?
Star	close	close	?
Complement	close	not close	?
Boolean operation	close	/	?