AMIT DHOMNE

PUMPING LEMMA

There are two Pumping Lemmas, which are defined for

1. **Regular Languages, and**
2. **Context - Free Languages**

APPLICATIONS OF PUMPING LEMMA

Pumping Lemma is to be applied to show that certain languages are not regular.

It should never be used to show a language is regular.

- If L is regular, it satisfies Pumping Lemma.
- If L does not satisfy Pumping Lemma, it is non-regular.

FORMAL STATEMENT

Pumping Lemma for Regular Languages

For any regular language L, there exists an integer n, such that for all $w \in L$ with

$|w| \geq n$, there exists ($x, y, z \in \Sigma^*$), such that $w = xyz$, and

(1) $|xy| \leq n$

(2) $|y| \geq 1$

(3) for all $i \geq 0$: $xy^iz \in L$

In simple terms, this means that if a string y is 'pumped', i.e., if y is inserted any number of times, the resultant string still remains in L.

FORMAL STATEMENT

Pumping Lemma for Context-free Languages (CFL)

Pumping Lemma for CFL states that for any Context Free Language L, it is possible to find two substrings that can be 'pumped' any number of times and still be in the same language. For any language L, we break its strings into five parts and pump second and fourth substring.

Pumping Lemma, here also, is used as a tool to prove that a language is not CFL. Because, if any one string does not satisfy its conditions, then the language is not CFL.

The Pumping Lemma:

For infinite context-free language L

there exists an integer n such that

for any string $w \in L, \quad |w| \geq n$

we can write $w = uvxyz$

with lengths $|vxy| \leq n$ and $|vy| \geq 1$

and it must be:

$$uv^i xy^i z \in L, \quad \text{for all } i \geq 0$$

TAKE AN INFINITE CONTEXT-FREE LANGUAGE

Generates an infinite number of different strings

Example:

$$S \to AB$$
$$A \to aBb$$
$$B \to Sb$$
$$B \to b$$

$$S \to AB$$

$$A \to$$

$$ABB \quad B$$

$$\to SB \quad B$$

$$\to B$$

In a derivation of a long string, variables are repeated

A derivation:

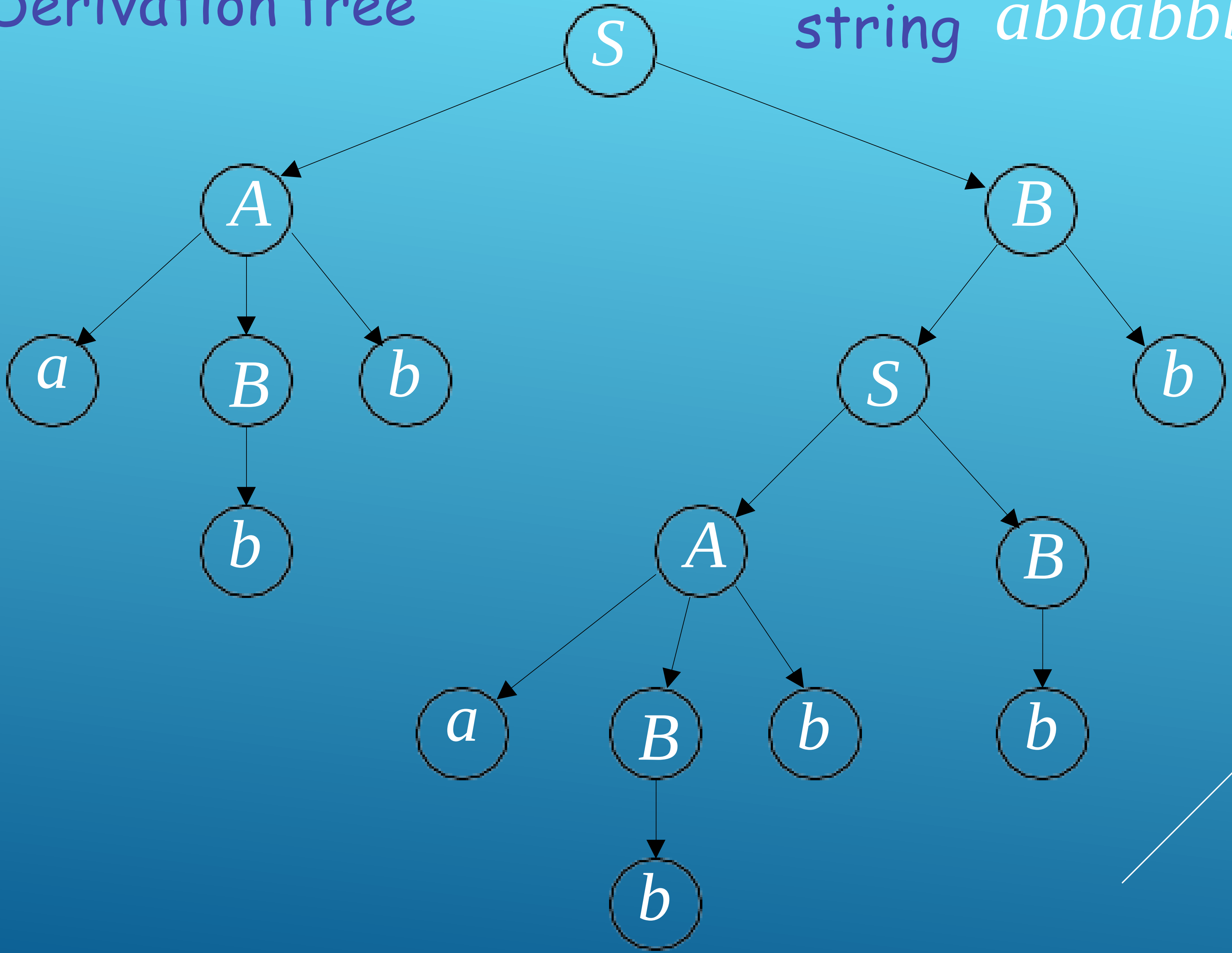
$$S \Rightarrow AB \Rightarrow aBbB \Rightarrow abb\underline{B}$$

$$\Rightarrow abbSb \Rightarrow abbABb \Rightarrow abbaBbBb \Rightarrow$$

$$\Rightarrow abbabb\underline{B}b \Rightarrow abbabbbb$$

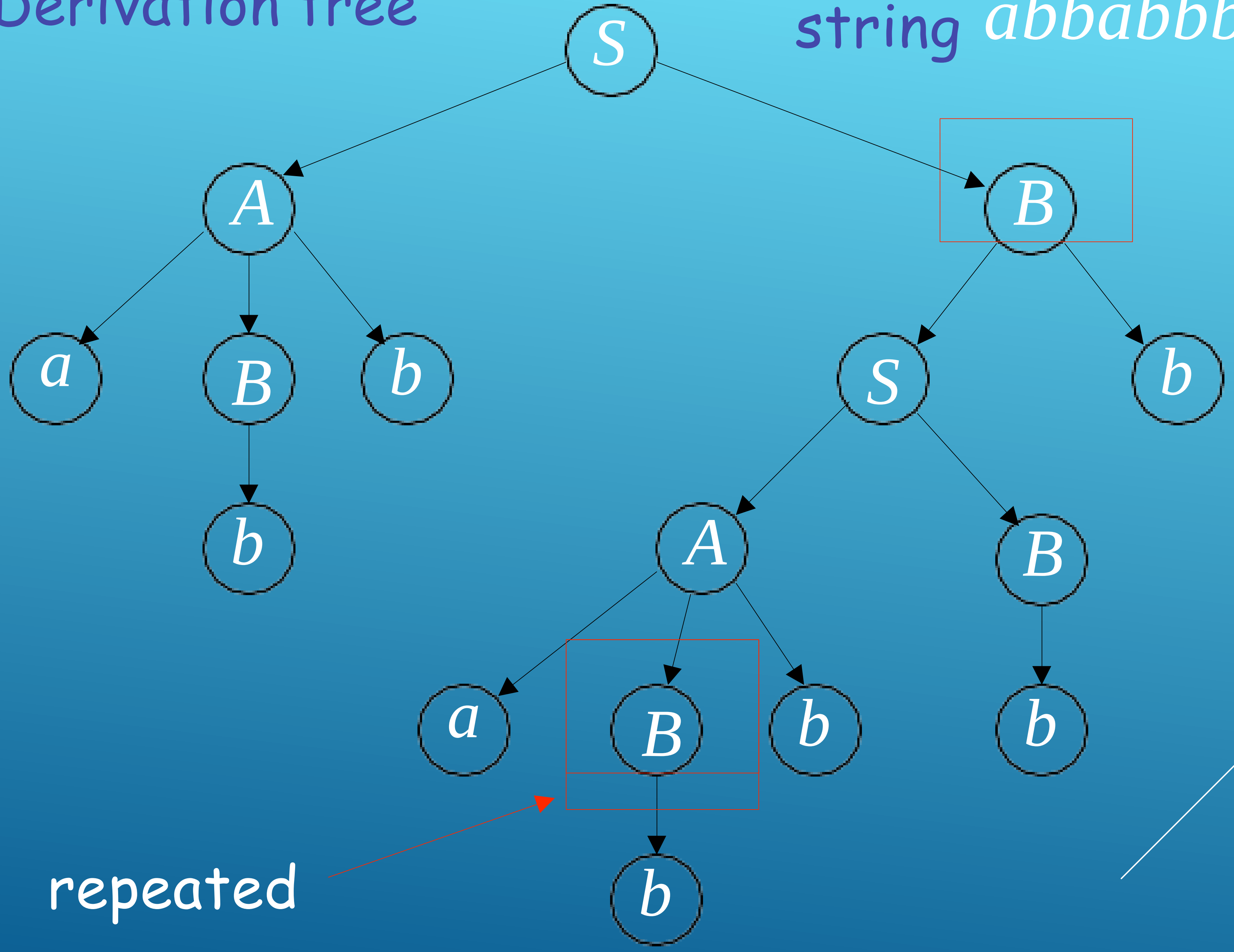
Derivation tree

string $abbabbbb$

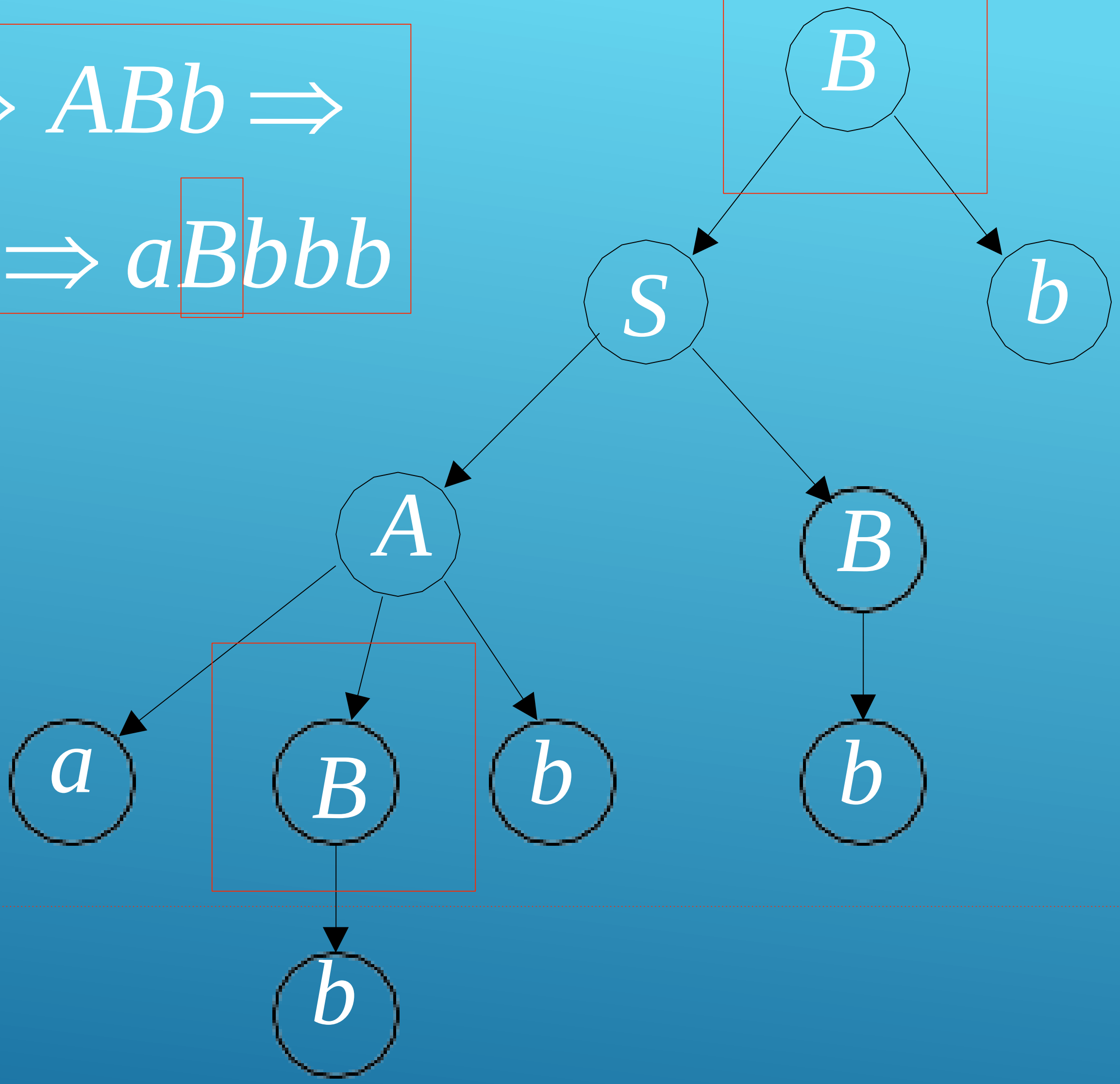


Derivation tree

string $abbabbbb$

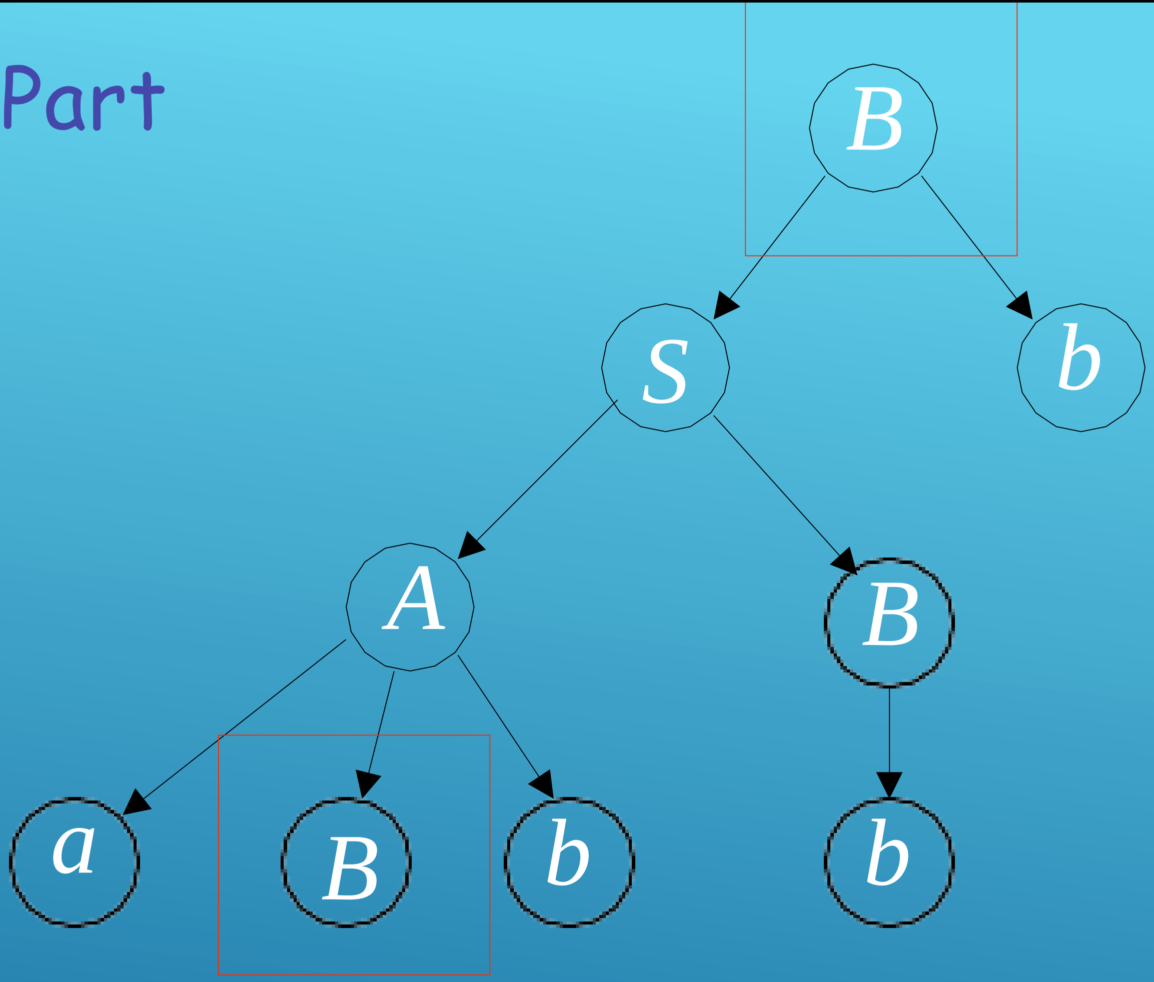


$$B \Rightarrow Sb \Rightarrow ABb \Rightarrow$$
$$\Rightarrow aBbBb \Rightarrow aBbbb$$



$$B \overset{*}{\Rightarrow} aBbbb \qquad B \Rightarrow b$$

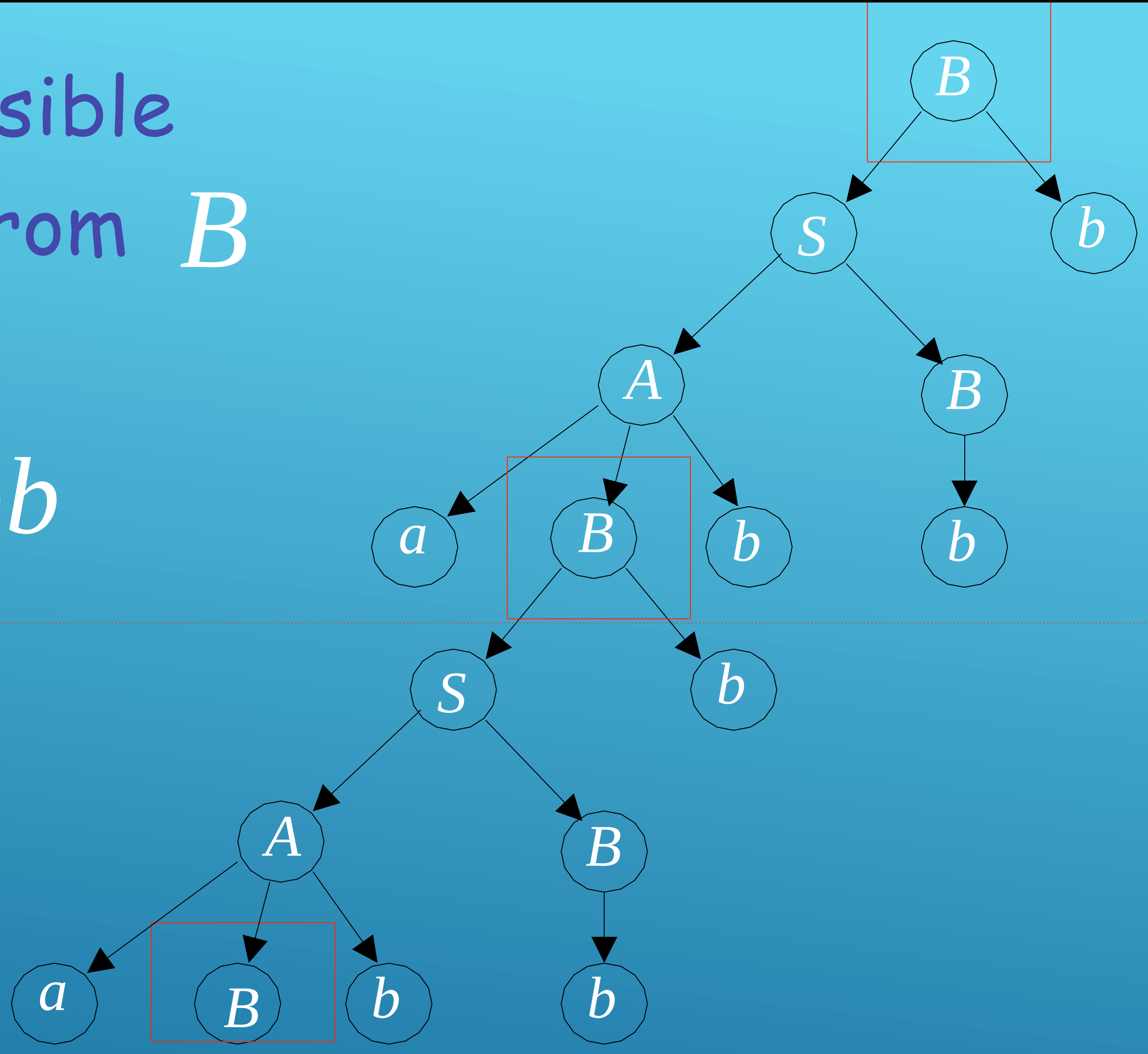
Repeated Part



$$B \overset{*}{\Rightarrow} aBbbb$$

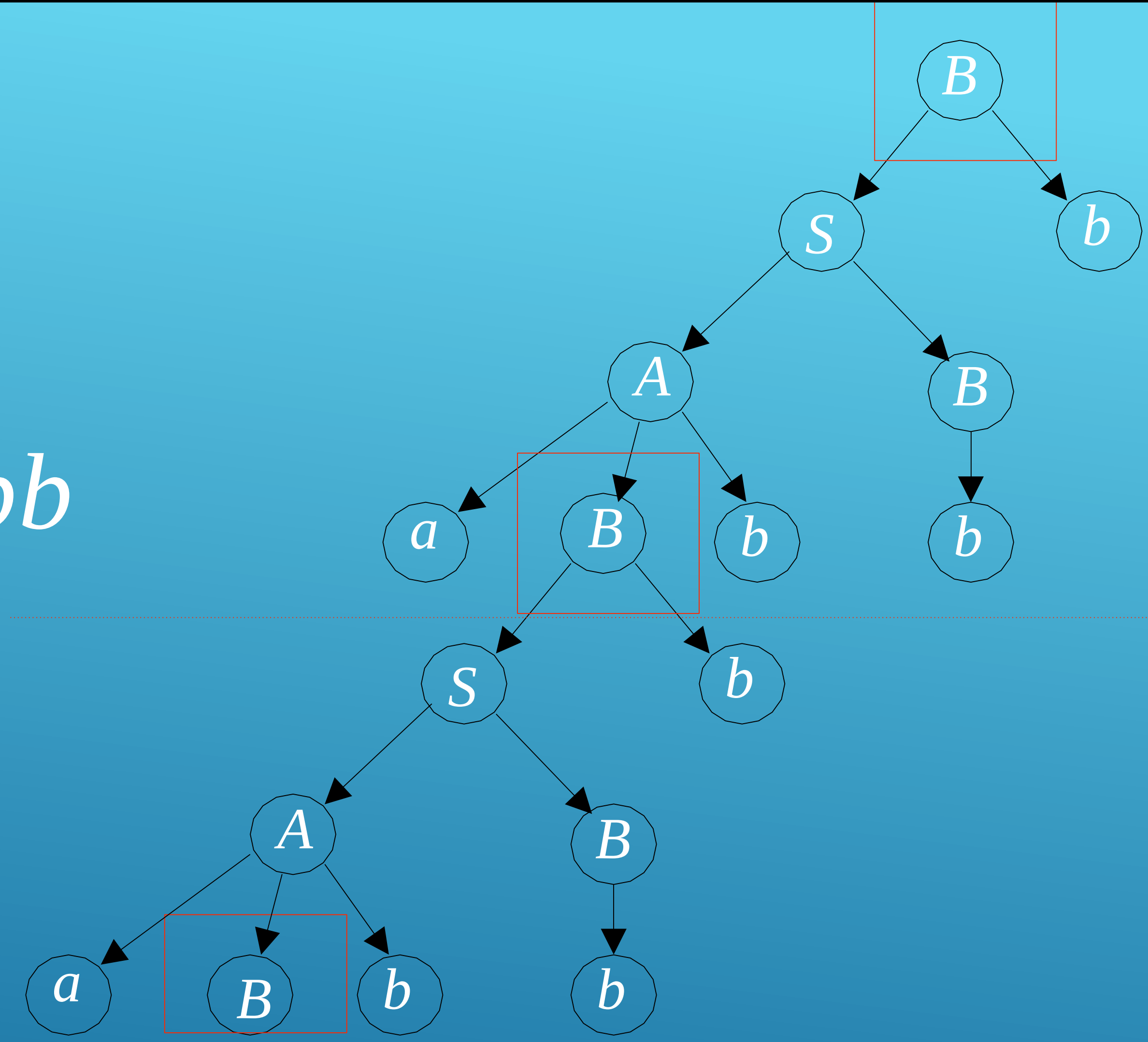
Another possible derivation from B

$$B \overset{*}{\Rightarrow} aBbbb$$



$$B \overset{*}{\Rightarrow} aBbbb \overset{*}{\Rightarrow} aaBbbbbbb$$

$$B \stackrel{*}{\Rightarrow} aBbbb$$



$$B \stackrel{*}{\Rightarrow} (a)B(bbb) \stackrel{*}{\Rightarrow} (a)^2 B(bbb)^2$$

$S \overset{*}{\Rightarrow} abbBb$ $\quad B \overset{*}{\Rightarrow} aBbbb$ $\quad B \Rightarrow b$

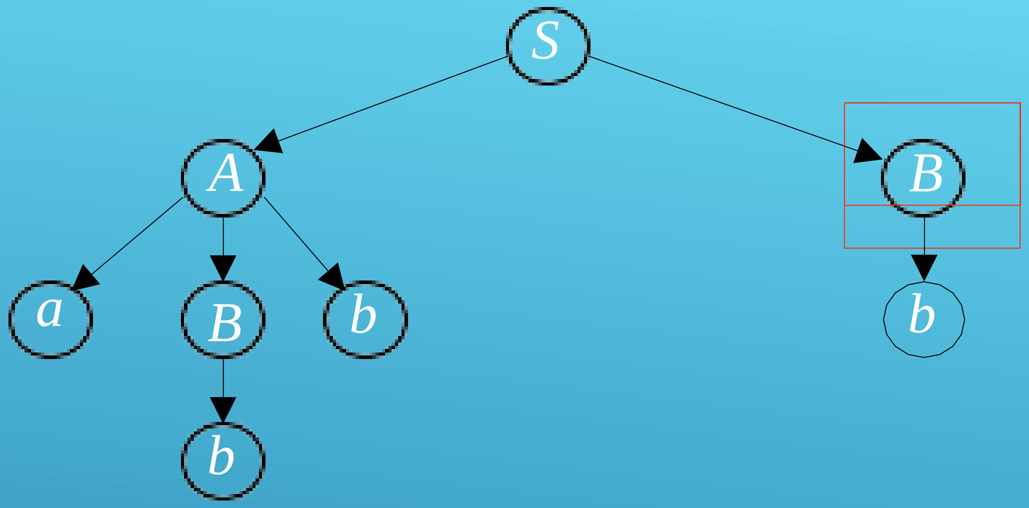
$$S \overset{*}{\Rightarrow} abb(a)^2 b(bbb)^2$$

$$abb(a)^2 b(bbb)^2 \in L(G)$$

$$S \overset{*}{\Rightarrow} abbBb$$

$$B \overset{*}{\Rightarrow} aBbbb$$

$$B \Rightarrow b$$



$$S \overset{*}{\Rightarrow} abbBb \Rightarrow abbbb \qquad = abb(a)^0 b(bbb)^0$$

$$S \overset{*}{\Rightarrow} abbBb \qquad B \overset{*}{\Rightarrow} aBbbb \qquad B \Rightarrow b$$

$$\Downarrow$$

$$S \overset{*}{\Rightarrow} abb(a)^0 b(bbb)^0$$

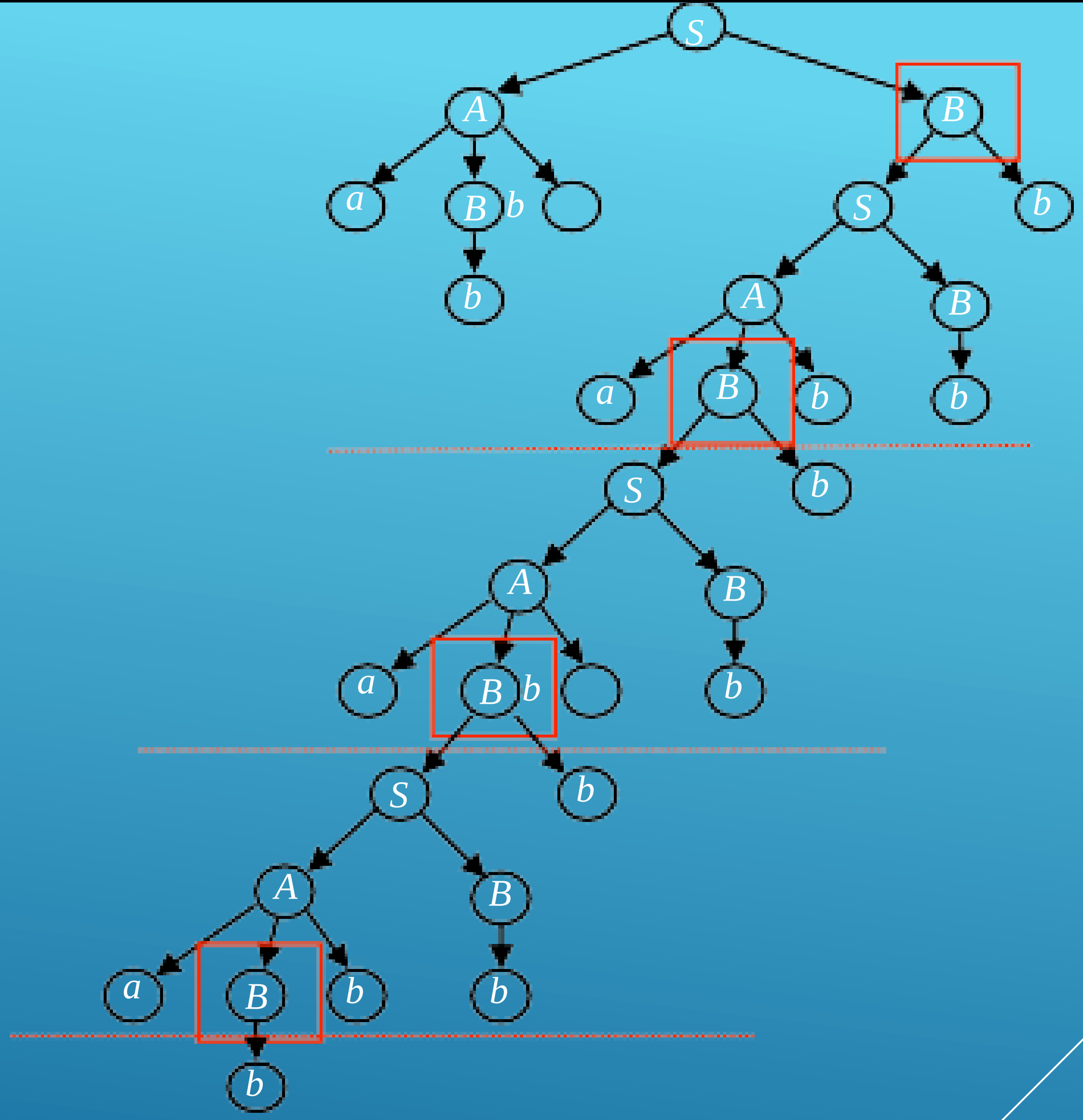
$$\Downarrow$$

$$abb(a)^0 b(bbb)^0 \in L(G)$$

$$S \overset{*}{\Rightarrow} abbBb$$

$$B \overset{*}{\Rightarrow} aBbbb$$

$$B \Rightarrow b$$



$$S \overset{*}{\Rightarrow} abb(a)^3 B(bbb)^3 \Rightarrow abb(a)^3 b(bbb)^3$$

$$S \overset{*}{\Rightarrow} abbBb \qquad B \overset{*}{\Rightarrow} aBbbb \qquad B \Rightarrow b$$

$$\Downarrow$$

$$S \overset{*}{\Rightarrow} abb(a)^3 b(bbb)^3$$

$$\Downarrow$$

$$abb(a)^3 b(bbb)^3 \in L(G)$$

In General:

$$S \overset{*}{\Rightarrow} abbBb \qquad B \overset{*}{\Rightarrow} aBbbb \qquad B \Rightarrow b$$

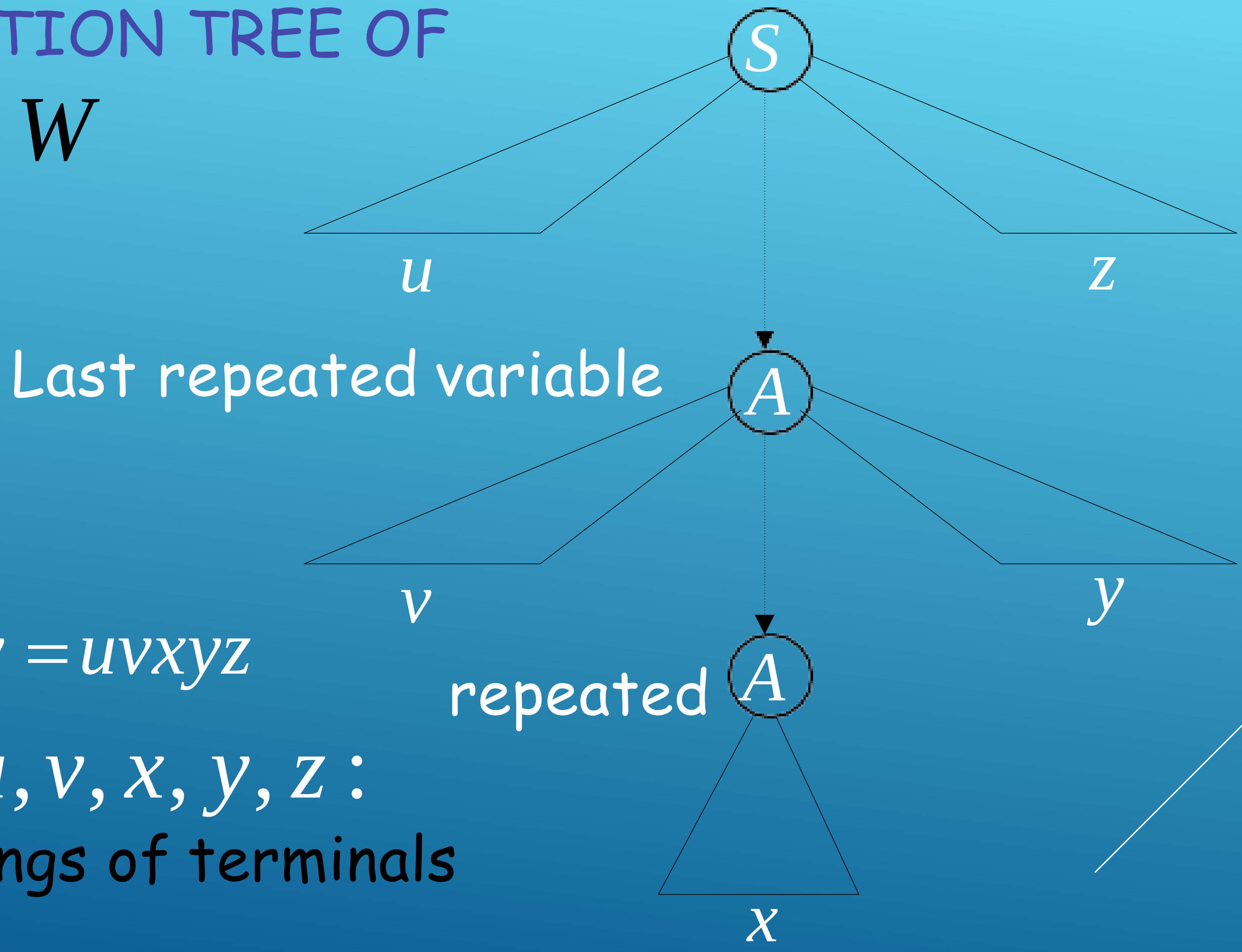
$$\Downarrow$$

$$S \overset{*}{\Rightarrow} abb(a)^i b(bbb)^i$$

$$\Downarrow$$

$$abb(a)^i b(bbb)^i \in L(G) \qquad i \geq 0$$

DERIVATION TREE OF STRING w



$w = uvxyz$

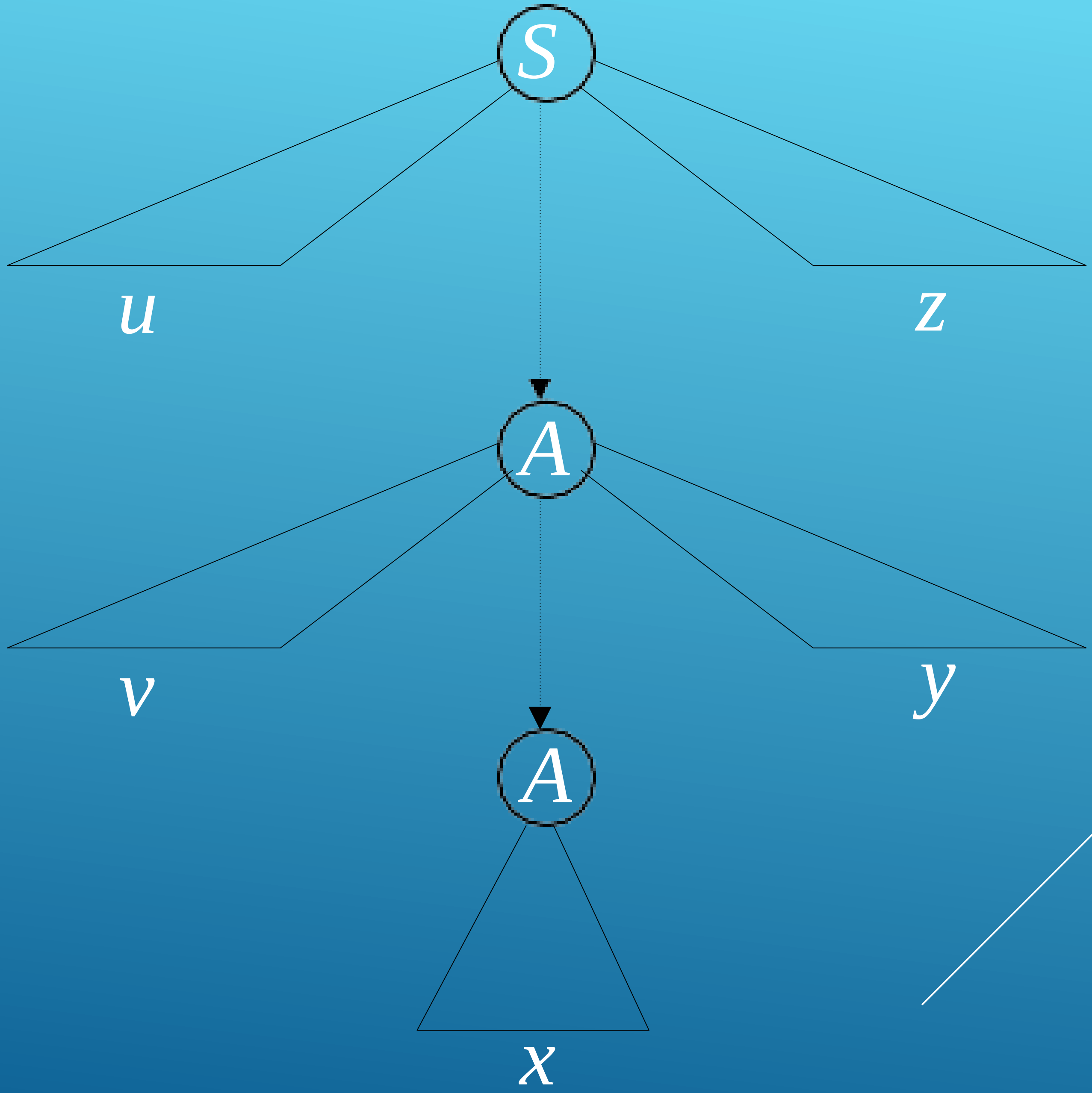
u, v, x, y, z :
Strings of terminals

Possible derivations:

$$S \overset{*}{\Rightarrow} uAz$$

$$A \overset{*}{\Rightarrow} vAy$$

$$A \overset{*}{\Rightarrow} x$$



We know:

$$S \overset{*}{\Rightarrow} uAz \qquad A \overset{*}{\Rightarrow} vAy \qquad A \overset{*}{\Rightarrow} x$$

This string is also generated:

$$S \overset{*}{\Rightarrow} uAz \overset{*}{\Rightarrow} uxz$$

$$uv^0xy^0z$$

We know:

$$S \overset{*}{\Rightarrow} uAz \qquad A \overset{*}{\Rightarrow} vAy \qquad A \overset{*}{\Rightarrow} x$$

This string is also generated:

$$S \overset{*}{\Rightarrow} uAz \overset{*}{\Rightarrow} uvAyz \overset{*}{\Rightarrow} uvxyz$$

The original $w = uv^1xy^1z$

We know:

$$S \overset{*}{\Rightarrow} uAz \qquad A \overset{*}{\Rightarrow} vAy \qquad A \overset{*}{\Rightarrow} x$$

This string is also generated:

$$S \overset{*}{\Rightarrow} uAz \overset{*}{\Rightarrow} uvAyz \overset{*}{\Rightarrow} uvvAyyz \overset{*}{\Rightarrow} uvvxyyz$$

$$uv^2xy^2z$$

We know:

$$S \overset{*}{\Rightarrow} uAz \qquad A \overset{*}{\Rightarrow} vAy \qquad A \overset{*}{\Rightarrow} x$$

This string is also generated:

$$S \overset{*}{\Rightarrow} uAz \Rightarrow uvAyz \overset{*}{\Rightarrow} uvvAyyz \overset{*}{\Rightarrow}$$

$$\overset{*}{\Rightarrow} uvvvAyyyz \overset{*}{\Rightarrow} uvvvxyyyz$$

$$uv^3xy^3z$$

We know:

$$S \overset{*}{\Rightarrow} uAz \qquad A \overset{*}{\Rightarrow} vAy \qquad A \overset{*}{\Rightarrow} x$$

This string is also generated:

$$S \Rightarrow uAz \Rightarrow uvAyz \Rightarrow uvvAyyz \Rightarrow$$

$$\Rightarrow uvvvAyyyz \Rightarrow \ldots$$

$$\Rightarrow u\,v\,v\,v \cdots v\,A\,y \cdots yyyz \Rightarrow$$

$$\Rightarrow u\,v\,v\,v \cdots v\,x\,y \cdots yyyz$$

$$uv^ixy^iz$$

Therefore, any string of the form

$$uv^i xy^i z \qquad i \geq 0$$

is generated by the grammar G

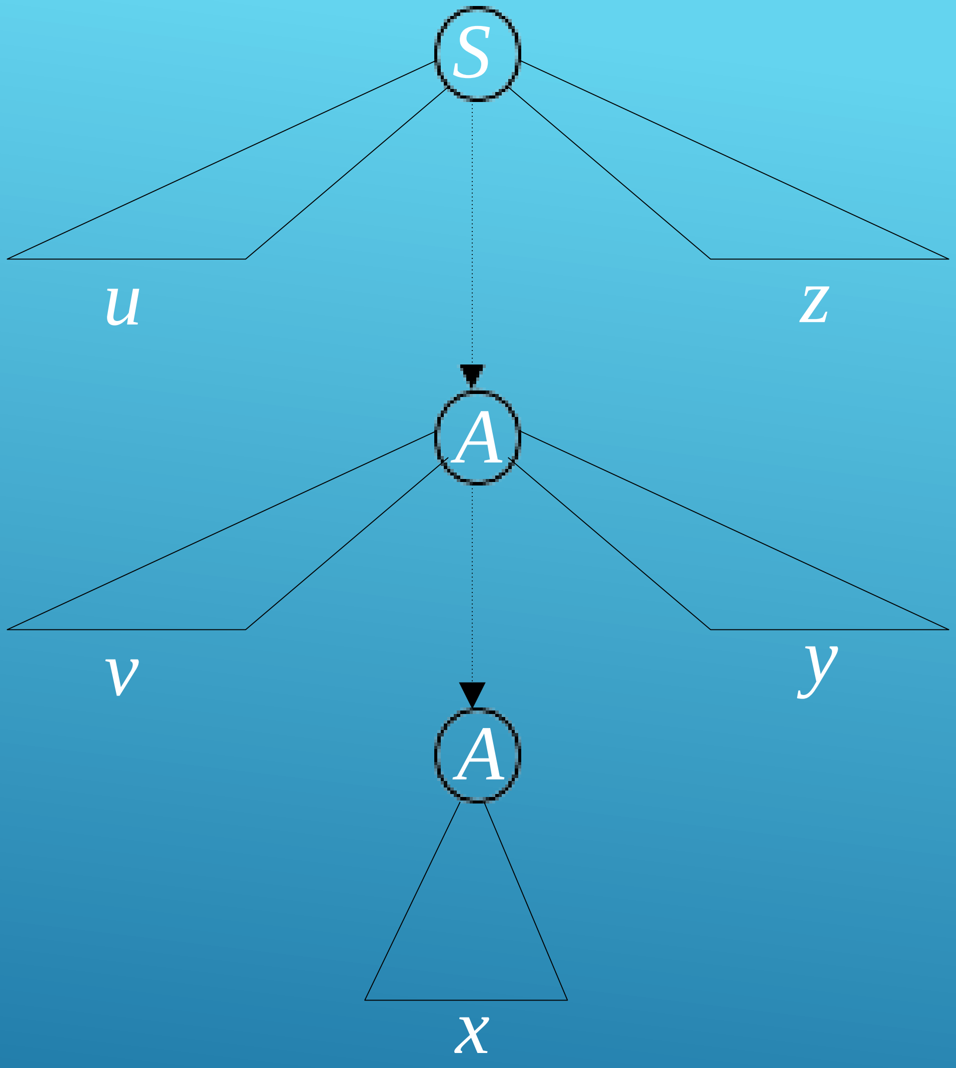
Therefore,

knowing that $uvxyz \in L(G)$

we also know that $uv^i xy^i z \in L(G)$

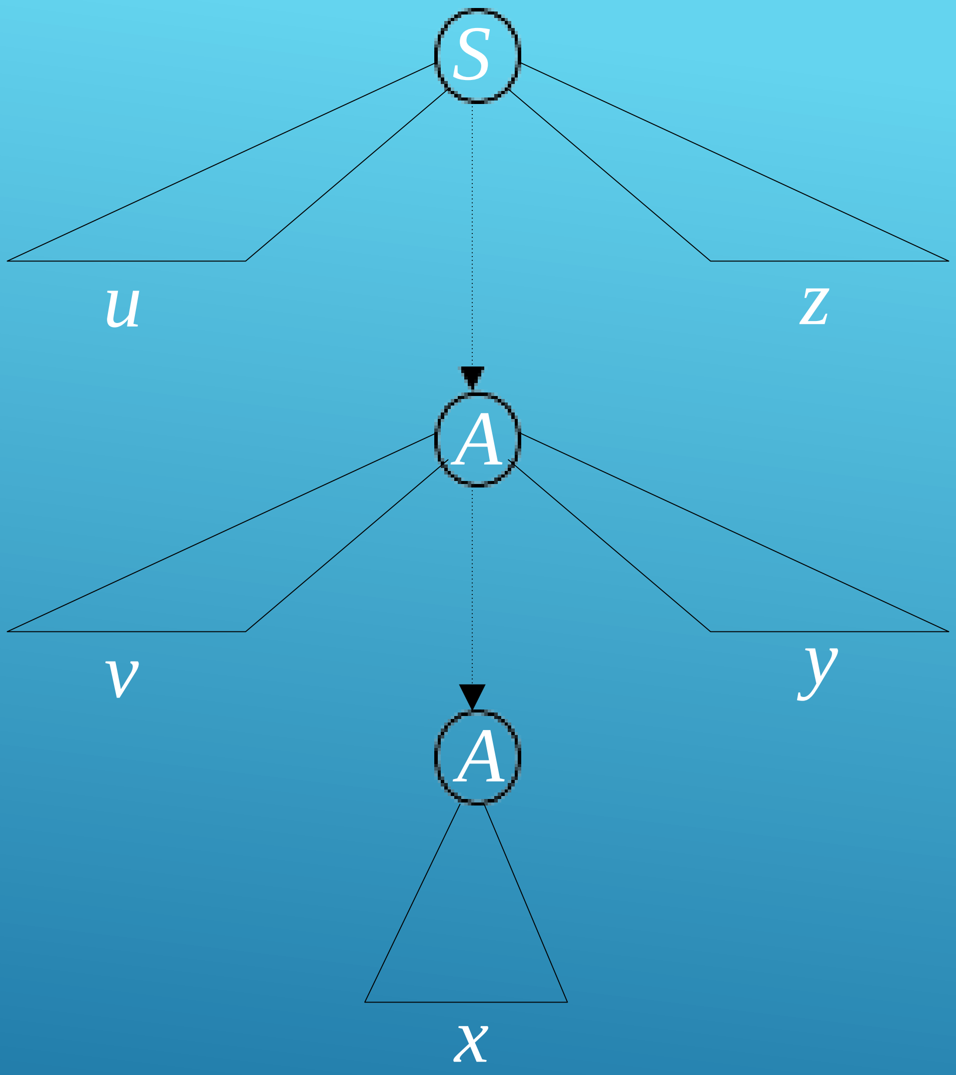
$L(G) = L - \{\lambda\}$

$uv^i xy^i z \in L$



Observation: $|vxy| \leq m$

Since A is the last repeated variable



Observation: $|vy| \geq 1$

Since there are no unit or λ-productions

The Pumping Lemma:

For infinite context-free language L

there exists an integer n such that

for any string $w \in L, \quad |w| \geq n$

we can write $w = uvxyz$

with lengths $|vxy| \leq n$ and $|vy| \geq 1$

and it must be:

$$uv^i xy^i z \in L, \quad \text{for all } i \geq 0$$