Applications
of
Regular Closure

Linz 6th. § 8.2 Closure Properties and Decision Algorithms for Context-Free Languages. [pages 1-18 here]

class13b Linz 6th, §8.1 Pumping Lemma

class13c Applications

Recall Thm 8.5: the intersection of
a context-free language and
a regular language
is a context-free language

Regular Closure

L_1 context free

L_2 regular

$\Rightarrow$ $L_1 \cap L_2$ context-free

Linz 6th, section 8.2, example 8.7, page 227

L={a^n b^n | $0 \le n$, $n \neq 100$}

is context free

An Application of Regular Closure

Prove that: $L = \{a^n b^n : \quad n \neq 100\}$

is context-free

We know: $\{a^n b^n\}$

is context-free

We also know:

$L_1 = \{a^{100}b^{100}\}$ is regular

$$\Downarrow$$

$\overline{L_1} = \{(a+b)^*\} - \{a^{100}b^{100}\}$ is regular

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

context-free

$\overline{L_1} = \{(a+b)^*\} - \{a^{100} b^{100}\}$

regular

(regular closure) $\{a^n b^n\} \cap \overline{L_1}$

is context-free

$$\{a^n b^n\} \cap \overline{L_1}$$

$$= \{a^n b^n : \quad n \neq 100\} = L$$ is context-free

Linz 6th, section 8.2, example 8.8, page 227

$L=\{w \mid \#_a(w) = \#_b(w) = \#_c(w)\}$

is not context free

Another Application of Regular Closure

Prove that: $L = \{w : \ n_a = n_b = n_c\}$

is **not** context-free

If $L=\{w: \; n_a=n_b=n_c\}$ is context-free

(regular closure)

Then $L\cap\{a^*b^*c^*\}=\{a^nb^nc^n\}$

context-free regular context-free

Impossible!!!

Therefore, L is **not** context free

Decidable Properties
of
Context-Free Languages

Linz 6th, Section 8.2, pages 227ff

Membership Question:

for context-free grammar G
find if string $w \in L(G)$

Membership Algorithms: Parsers

- Exhaustive search parser
- **CYK** parsing algorithm

Empty Language Question:

for context-free grammar G

find if $L(G) = \emptyset$

Algorithm:

1. Remove useless variables

2. Check if start variable S is useless

Infinite Language Question:

for context-free grammar G

find if $L(G)$ is infinite

Algorithm:

1. Remove useless variables

2. Remove unit and λ productions

3. Create dependency graph for variables

4. If there is a loop in the dependency graph then the language is infinite

Example: $S \to AB$

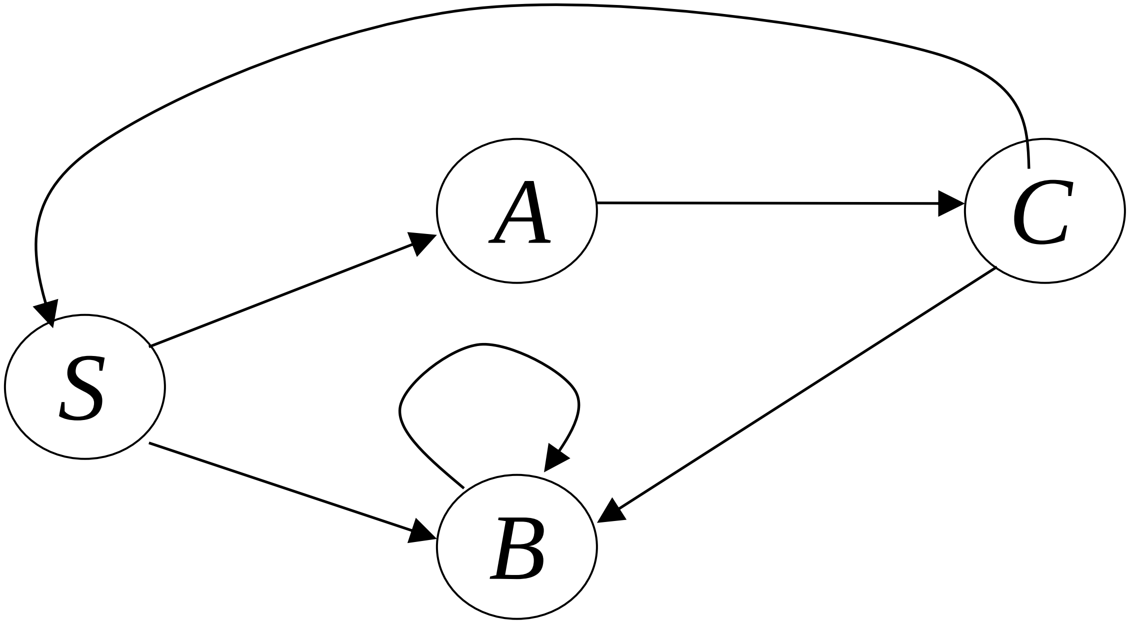
$$A \to aCb \mid a$$

$$B \to bB \mid bb$$

$$C \to cBS$$

Dependency graph

Infinite language



$$S \to AB$$

$$A \to aCb \mid a$$

$$B \to bB \mid bb$$

$$C \to cBS$$

S => AB => aB => abB => ab^{i}B => ab^{i}bb

$$S \to AB$$

$$A \to aCb \mid a$$

$$B \to bB \mid bb$$

$$C \to cBS$$

$$S \Rightarrow AB \Rightarrow aCbB \Rightarrow acBSbB \Rightarrow acbbSbbb$$

$$S \overset{*}{\Rightarrow} acbbSbbb \overset{*}{\Rightarrow} (acbb)^2 S (bbb)^2$$

$$\overset{*}{\Rightarrow} (acbb)^i S (bbb)^i$$

There is no algorithm to determine whether two context-free grammars generate the same language.

For the moment we do not have the technical machinery for defining the meaning of “there is no algorithm”.

The Pumping Lemma
for
Context-Free Languages

Linz 6th Section 8.1

Take an **infinite** context-free language

Generates an infinite number
of different strings

Example:

$$S \rightarrow AB$$

$$A \rightarrow aBb$$

$$B \rightarrow Sb$$

$$B \rightarrow b$$

$$S \to AB$$

$$A \to aBb$$

$$B \to Sb$$

$$B \to b$$

A derivation:

Variables are repeated

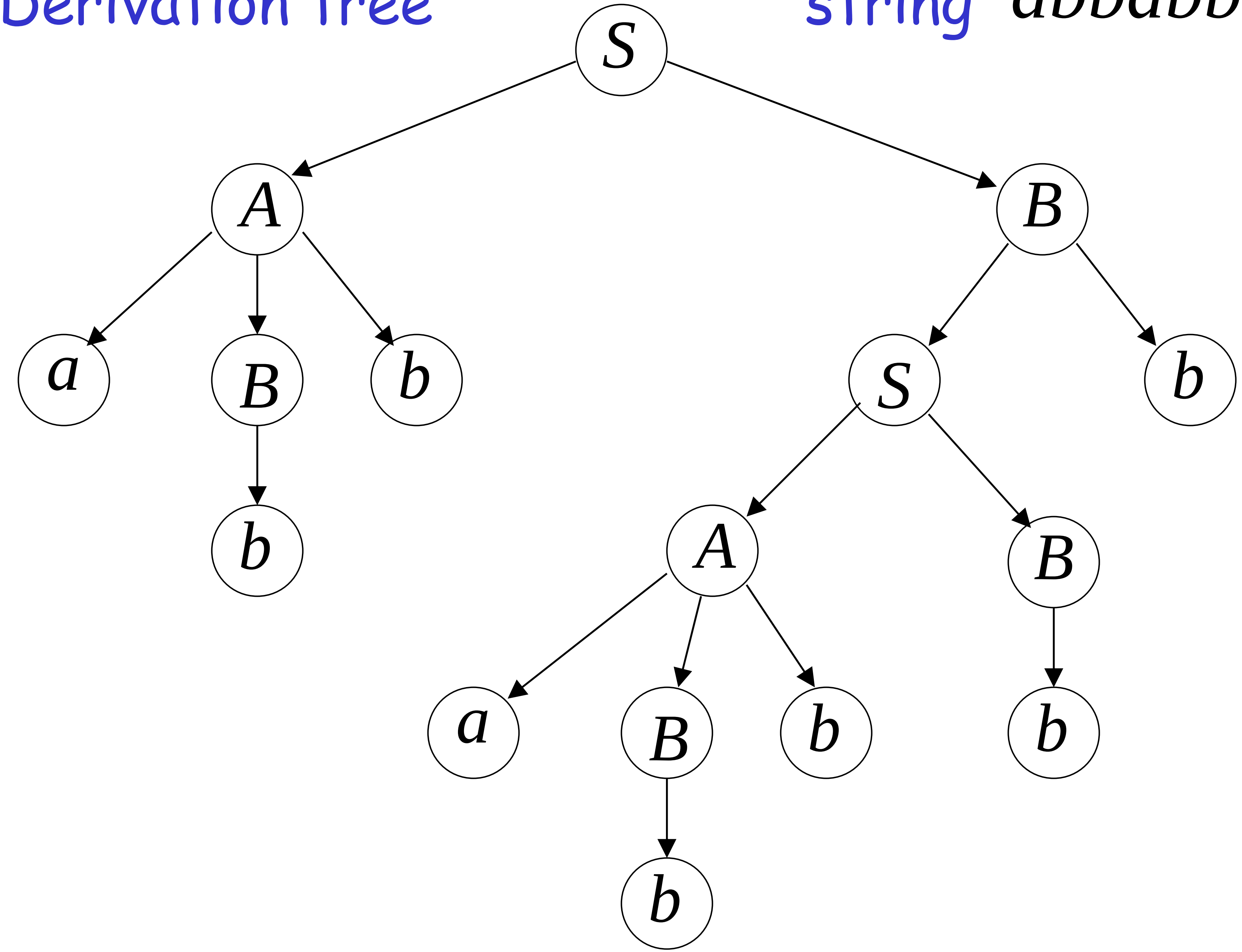
$$S \Rightarrow AB \Rightarrow aBbB \Rightarrow abb\underline{B} \Rightarrow$$

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

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

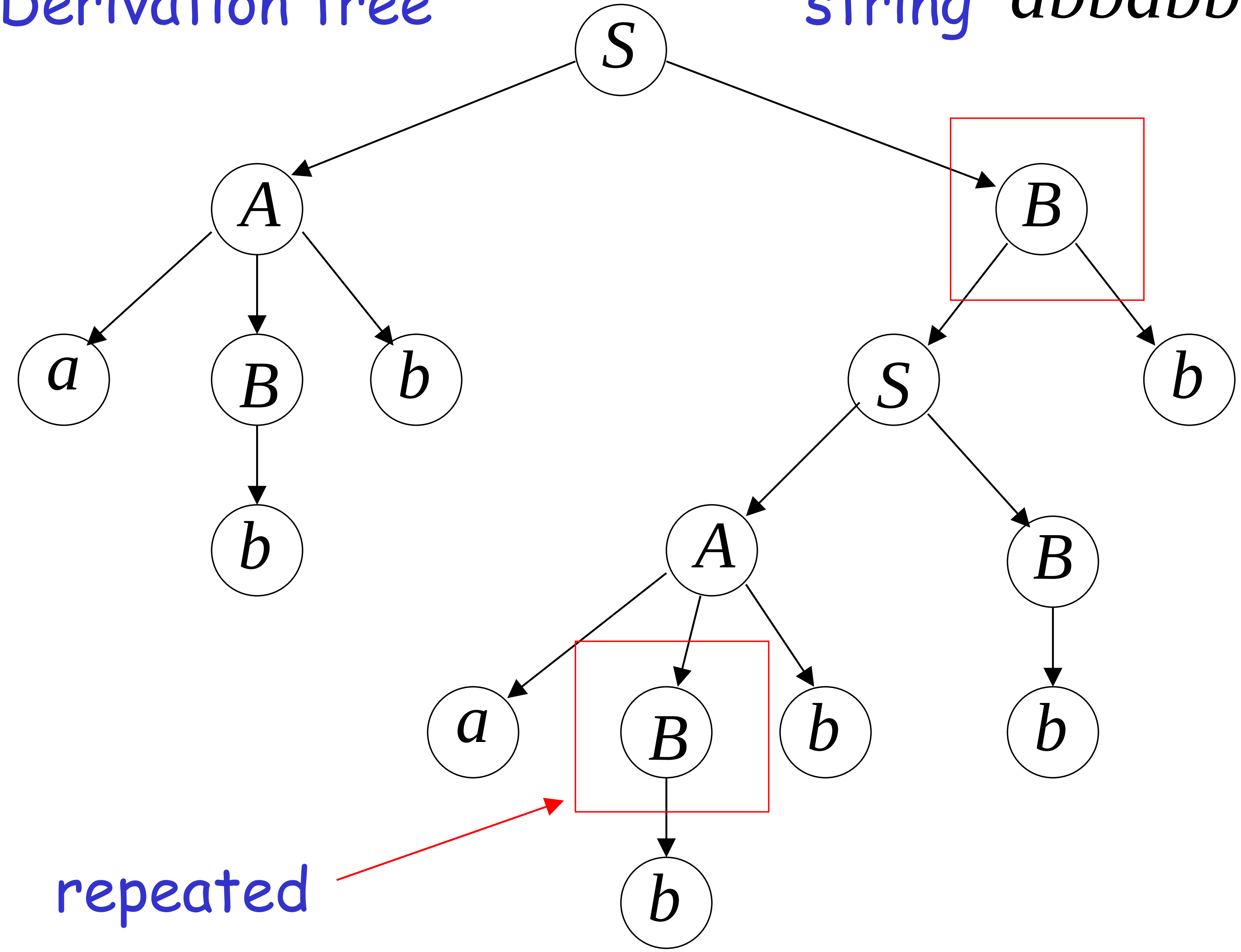
Derivation tree

string $abbabbbb$



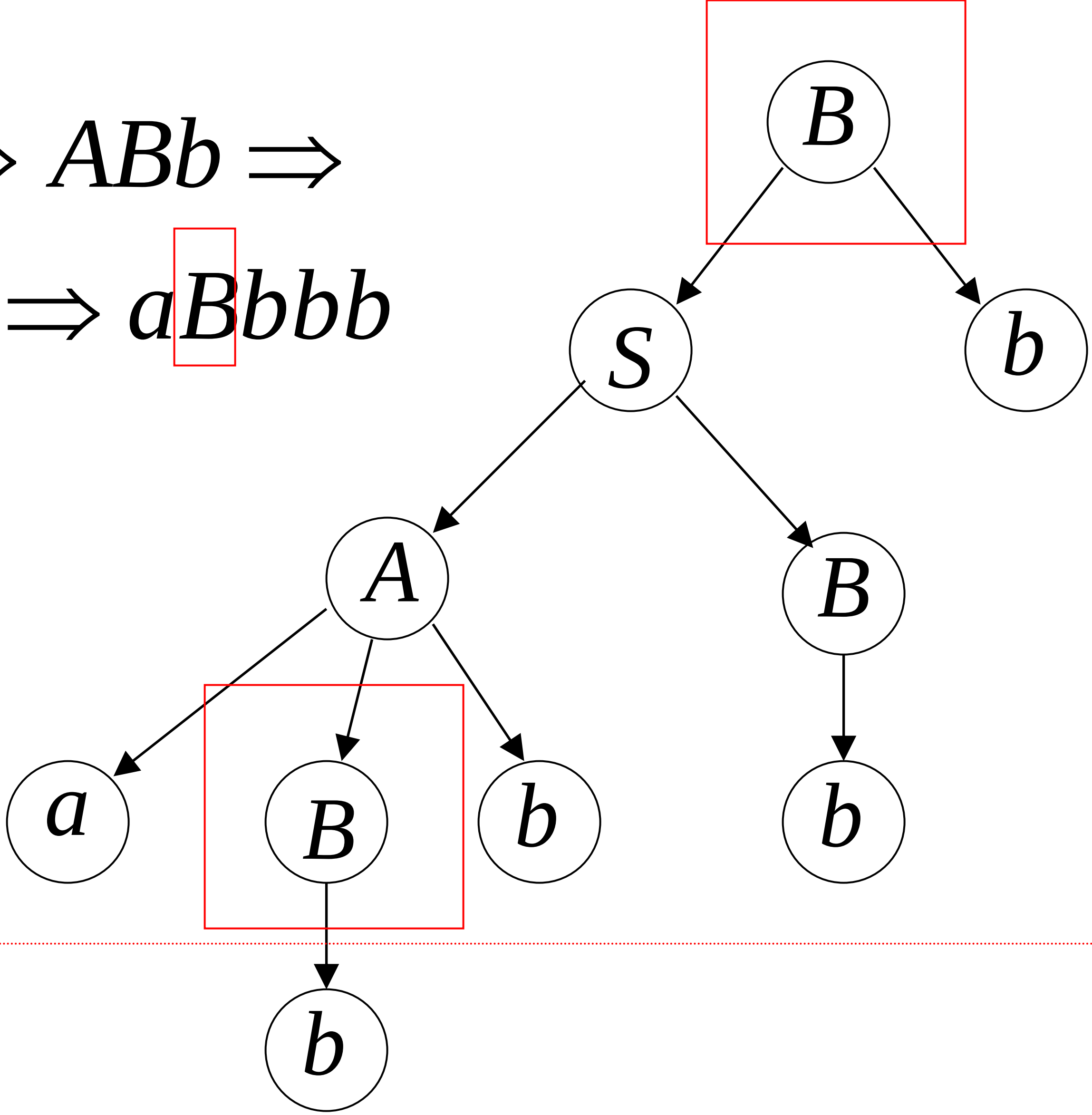
Derivation tree

string $abbabbbb$



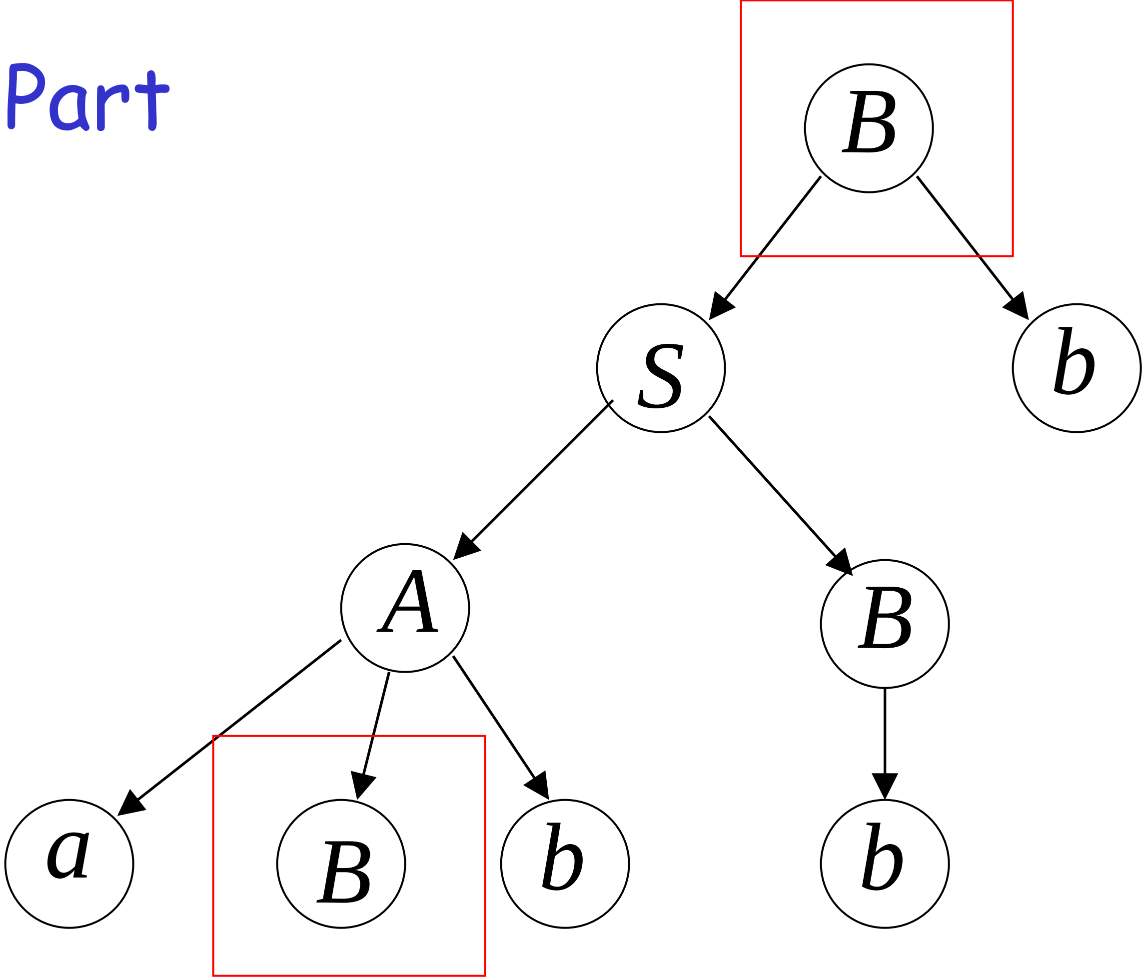
$$\boxed{B} \Rightarrow Sb \Rightarrow ABb \Rightarrow$$

$$\Rightarrow aBbBb \Rightarrow a\boxed{B}bbb$$



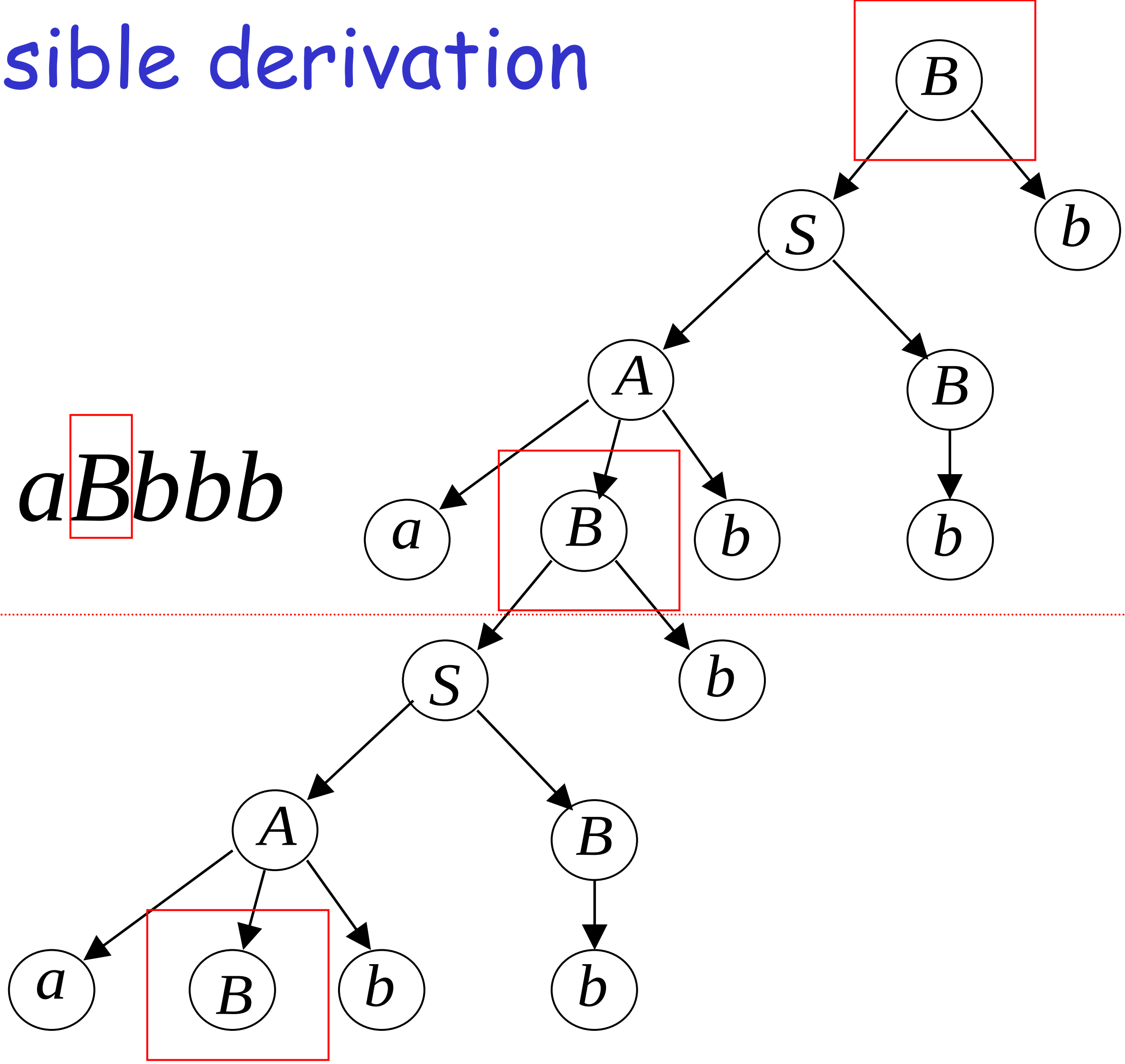
$$B \Rightarrow b$$

Repeated Part



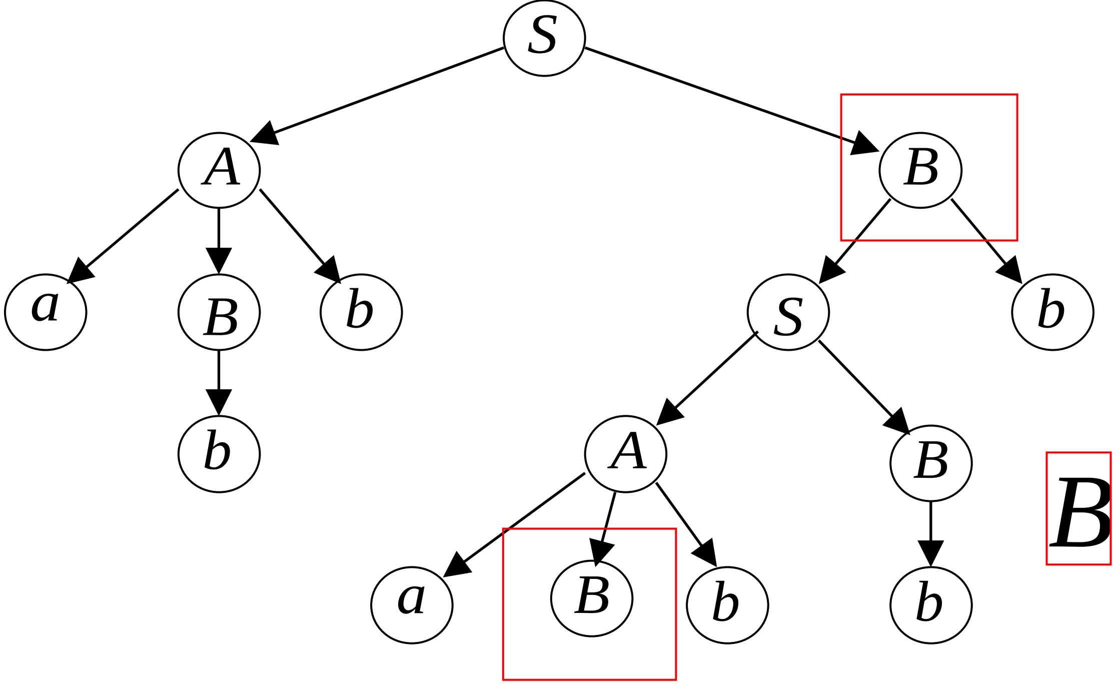
$$\boxed{B} \Rightarrow \ldots \Rightarrow a\boxed{B}bbb$$

Another possible derivation



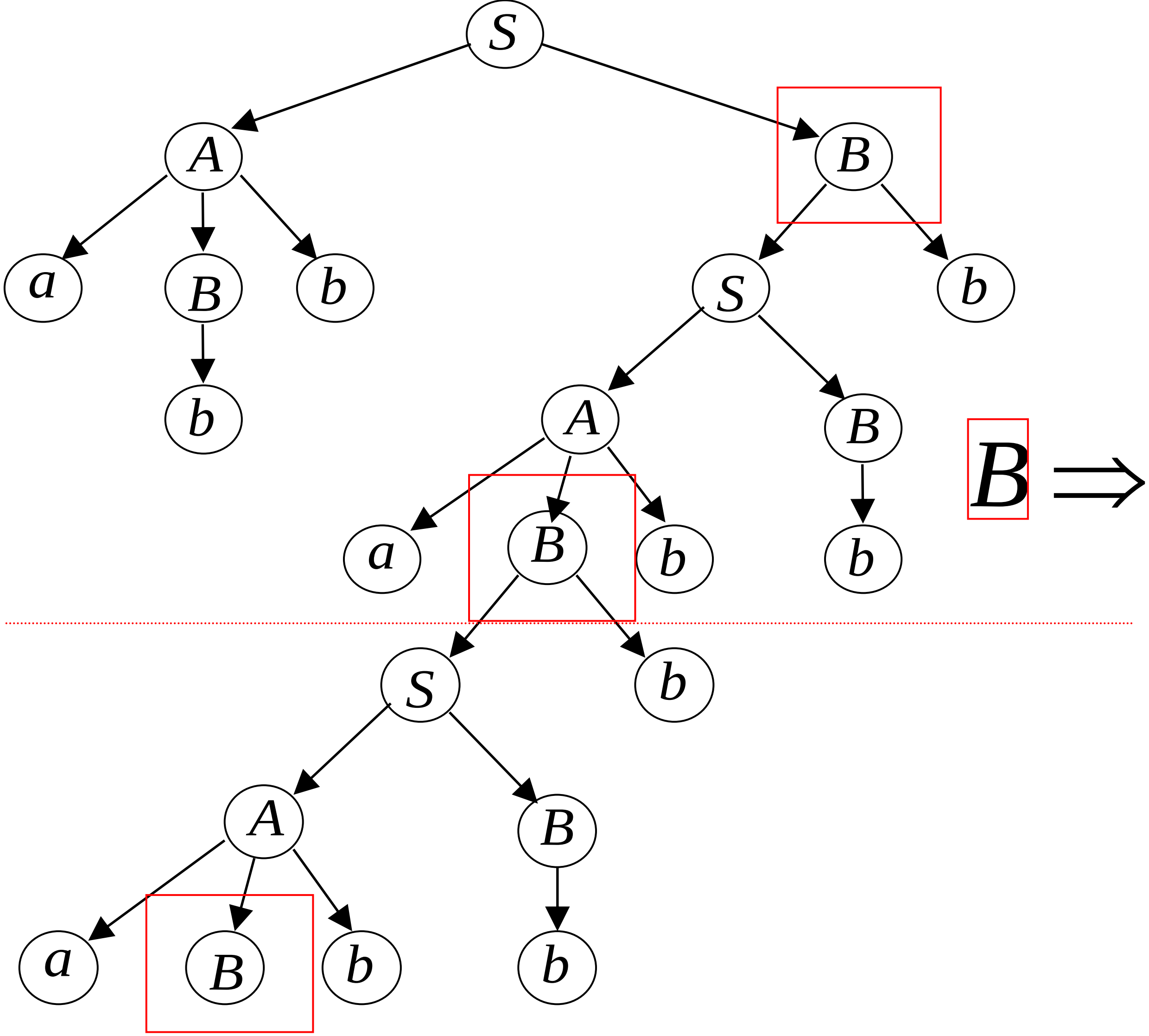
$$\boxed{B} \Rightarrow \ldots \Rightarrow a\boxed{B}bbb$$

$$\boxed{B} \Rightarrow \ldots \Rightarrow a\boxed{B}bbb \ldots \Rightarrow a\boxed{a\boxed{B}bbb}bbb$$



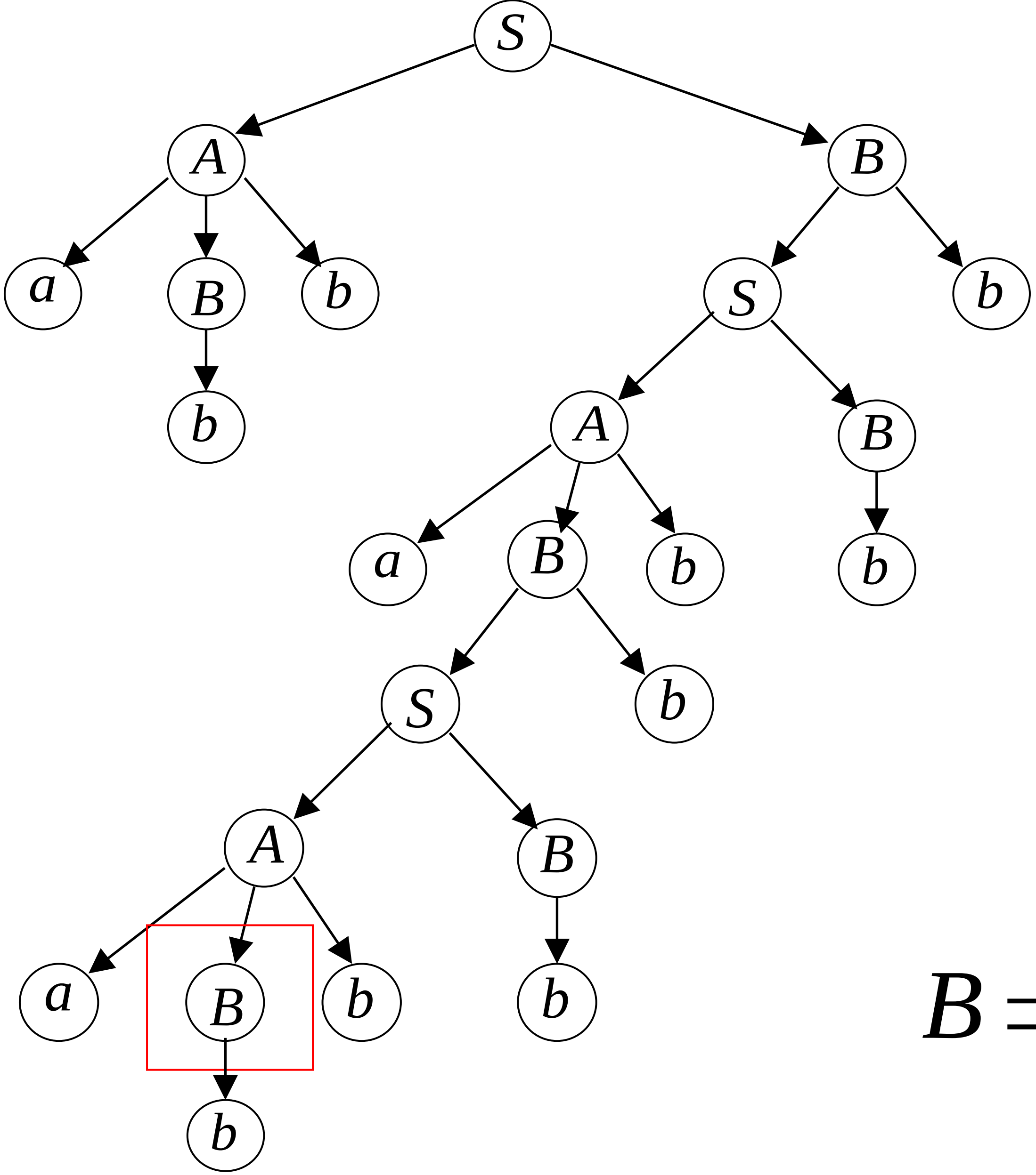
$$\boxed{B} \Rightarrow \ldots \Rightarrow a\boxed{B}bbb$$

$$S \Rightarrow \ldots \Rightarrow abba\boxed{B}bbb$$



$$\boxed{B} \Rightarrow \ldots \Rightarrow a\boxed{B}bbb$$

$$S \Rightarrow \ldots \Rightarrow abba\boxed{B}bbb \Rightarrow \ldots \Rightarrow abba\boxed{a\boxed{B}bbb}bbb$$



$$B \Rightarrow b$$

$$S \Rightarrow \ldots \Rightarrow abbaa\boxed{B}bbbbbb \Rightarrow abbaa\boxed{b}bbbbbb$$

$$S \Rightarrow \ldots \Rightarrow abbaabbbbbb$$

Therefore, the string

$$abbaabbbbbb$$

is also generated by the grammar

We know:

$$B \Rightarrow b$$

$$B \Rightarrow \ldots \Rightarrow aBbbb$$

$$S \Rightarrow \ldots \Rightarrow abbaBbbb$$

We also know this string is generated:

$$S \Rightarrow \ldots \Rightarrow abba\boxed{B}bbb \Rightarrow$$

$$\Rightarrow abbaabbbb$$

We know:

$$B \Rightarrow b$$

$$B \Rightarrow \ldots \Rightarrow aBbbb$$

$$S \Rightarrow \ldots \Rightarrow abbaBbbb$$

Therefore, this string is also generated:

$$S \Rightarrow \ldots \Rightarrow abba\boxed{B}bbb \Rightarrow$$

$$\Rightarrow abba\boxed{aBbbb}bbb \Rightarrow$$

$$\Rightarrow abbaabbbbbbb$$

We know:

$$B \Rightarrow b$$

$$B \Rightarrow \ldots \Rightarrow aBbbb$$

$$S \Rightarrow \ldots \Rightarrow abbaBbbb$$

Therefore, this string is also generated:

$$S \Rightarrow \ldots \Rightarrow abbaBbbb \Rightarrow$$

$$\Rightarrow abba(a)B(bbb)bbb$$

$$\Rightarrow abba(a)^2B(bbb)^2bbb$$

$$\Rightarrow abba(a)^2b(bbb)^2bbb$$

We know:

$$B \Rightarrow b$$

$$B \Rightarrow \ldots \Rightarrow aBbbb$$

$$S \Rightarrow \ldots \Rightarrow abbaBbbb$$

Therefore, this string is also generated:

$$S \Rightarrow \ldots \Rightarrow abbaBbbb \Rightarrow$$

$$\Rightarrow \ldots$$

$$\Rightarrow abba(a)^i B(bbb)^i bbb$$

$$\Rightarrow abba(a)^i b(bbb)^i bbb$$

Therefore, knowing that

$$abbabbbb$$

is generated by grammar G,
we also know that

$$abba(a)^i b(bbb)^i bbb$$

is generated by G

In general:

We are given an infinite context-free grammar G

Assume G has no unit-productions

no λ-productions

Take a string $w \in L(G)$

with length bigger than

$m > (\text{Number of productions}) \times (\text{Largest right side of a production})$

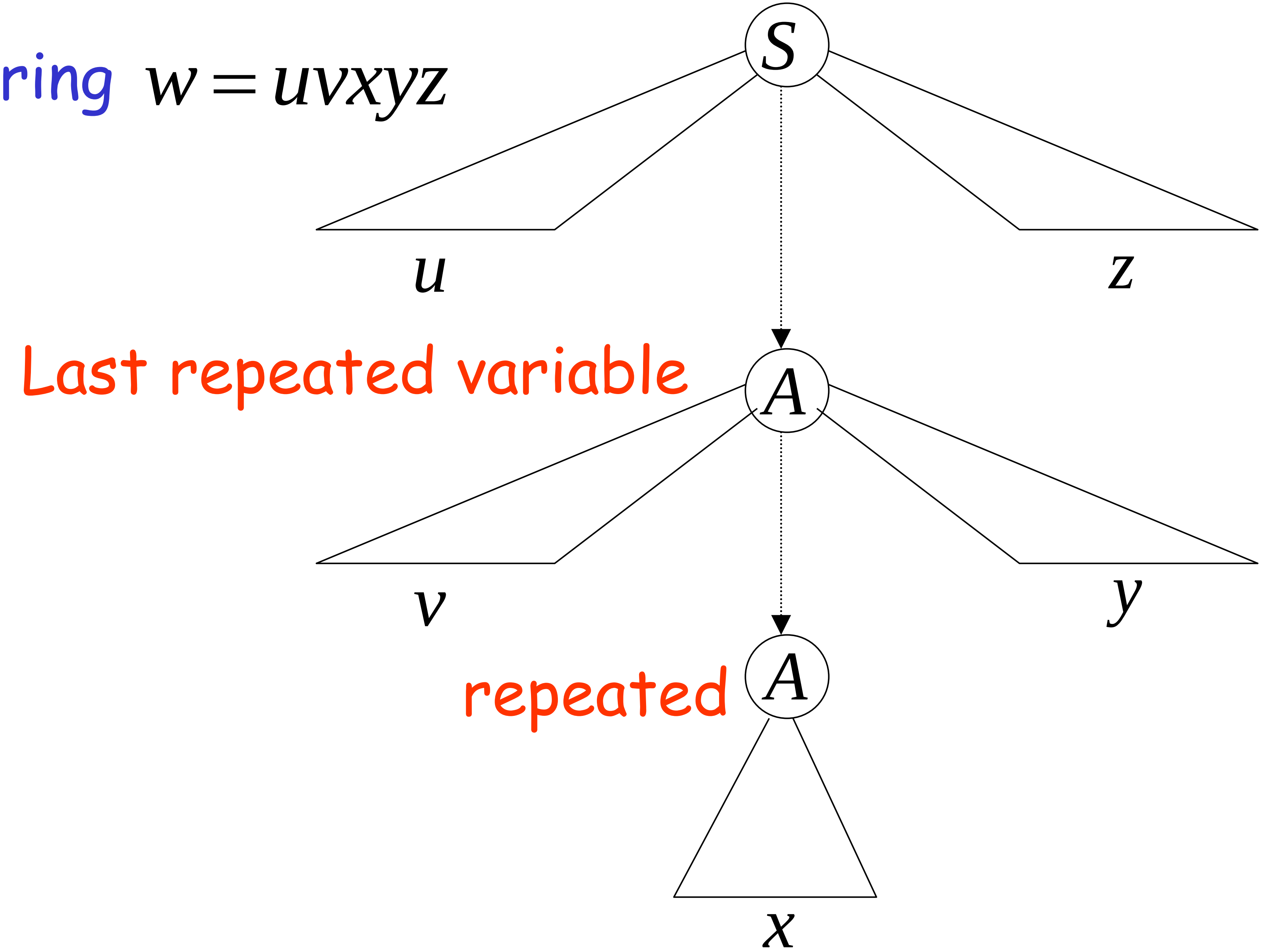
Consequence:

Some variable must be repeated

in the derivation of w

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

String $w = uvxyz$

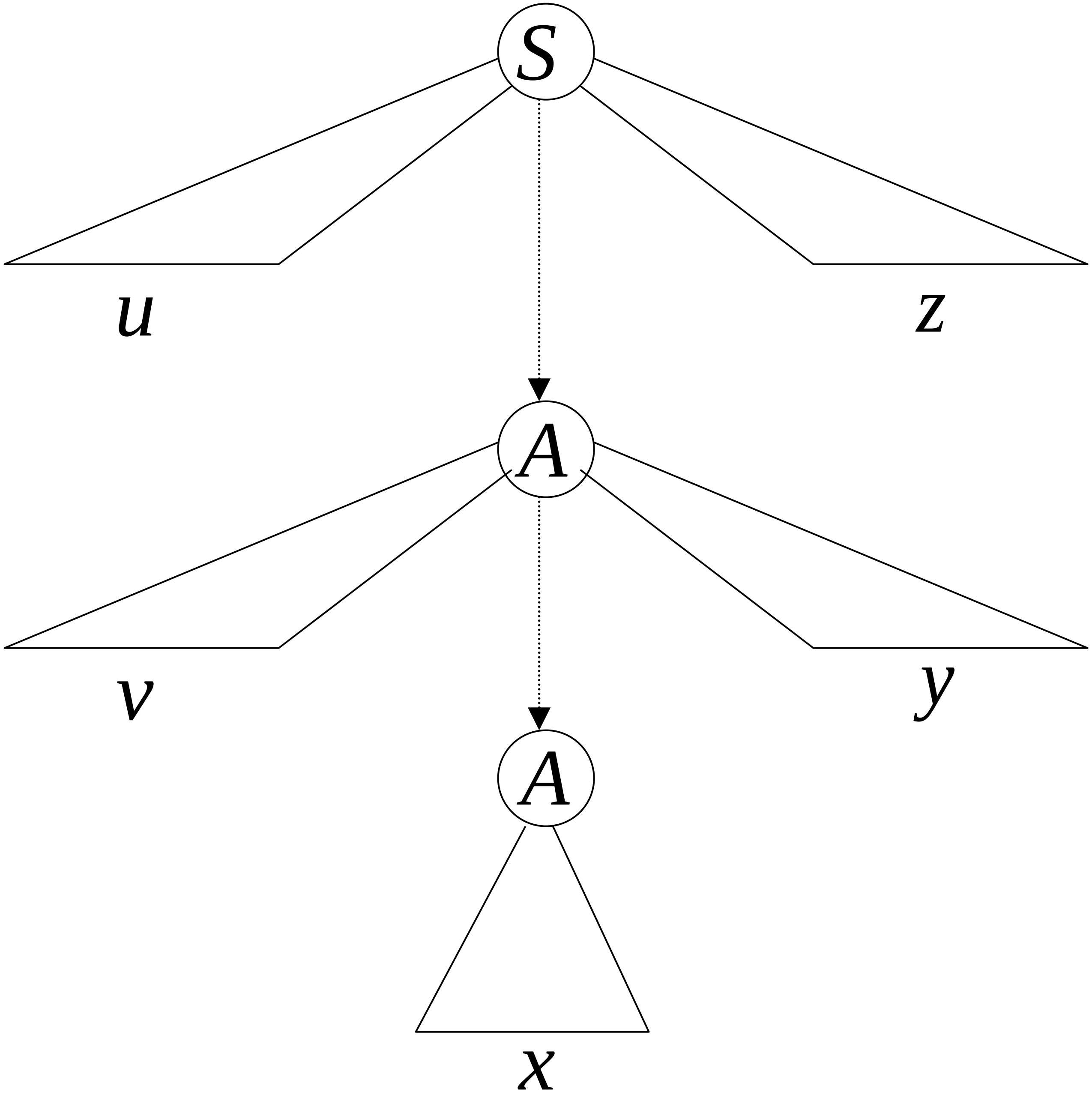


Possible derivations:

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

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

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



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:

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

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

$$\overset{*}{\Rightarrow} uvvv\cdots vAy\cdots yyyz \overset{*}{\Rightarrow}$$

$$\overset{*}{\Rightarrow} uvvv\cdots vxy\cdots yyyz$$

$$uv^ixy^iz$$

Therefore, any string of the form

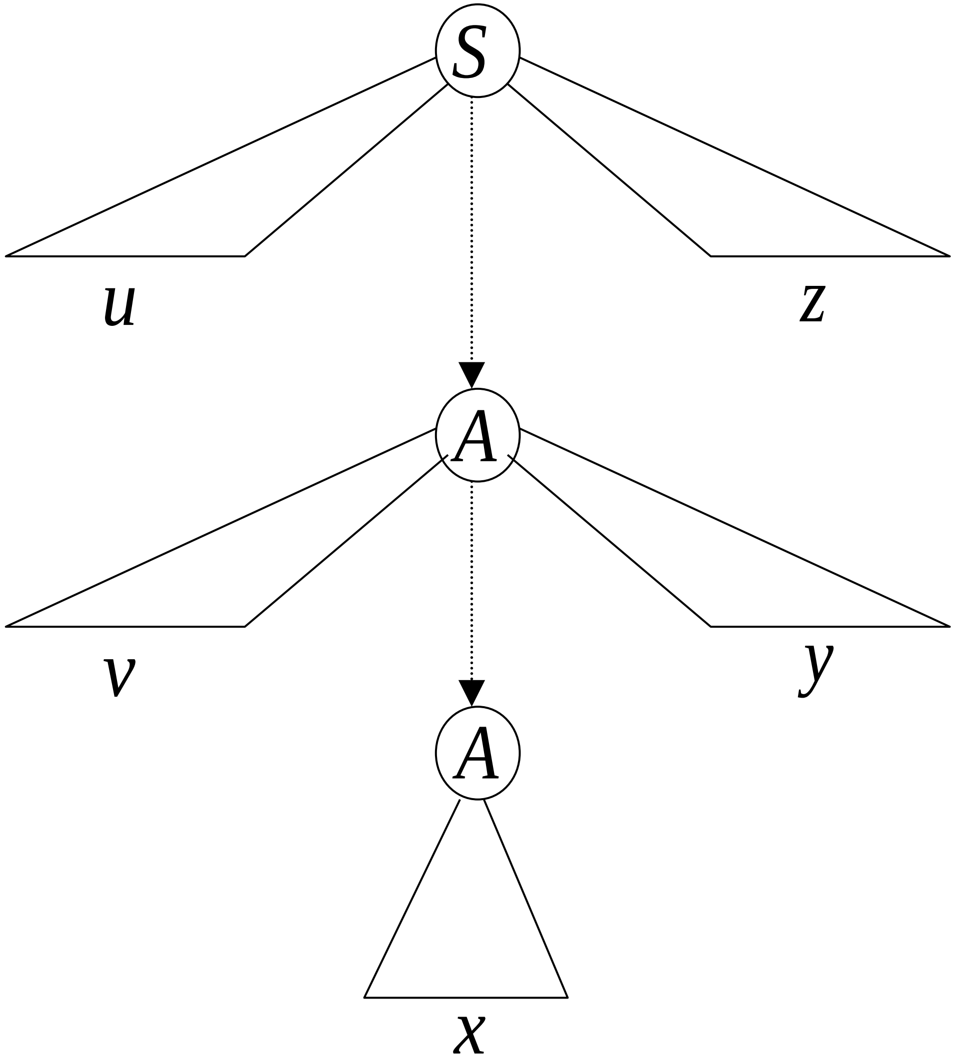
$$uv^i xy^i z \qquad i \geq 0$$

is generated by the grammar G

Therefore,

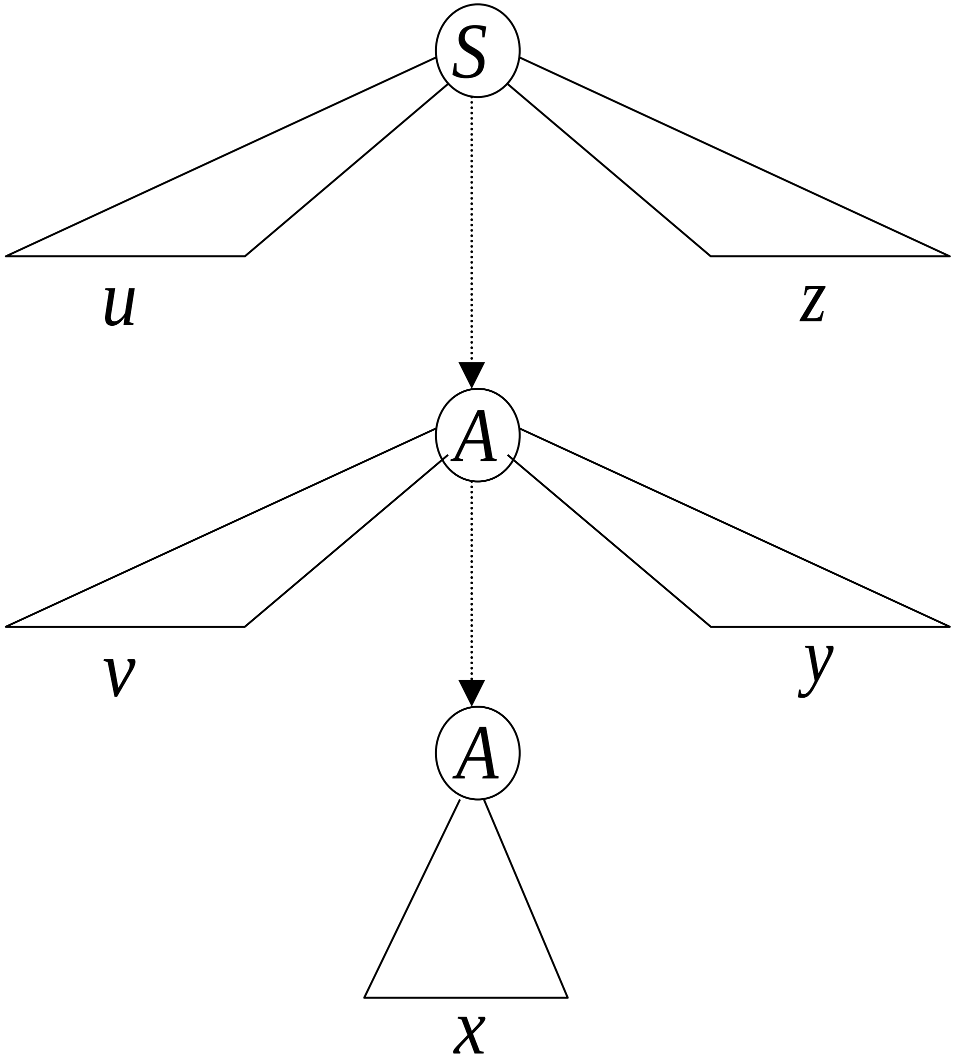
knowing that $uvxyz \in L(G)$

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



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 m such that

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

we can write $w = uvxyz$

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

and it must be:

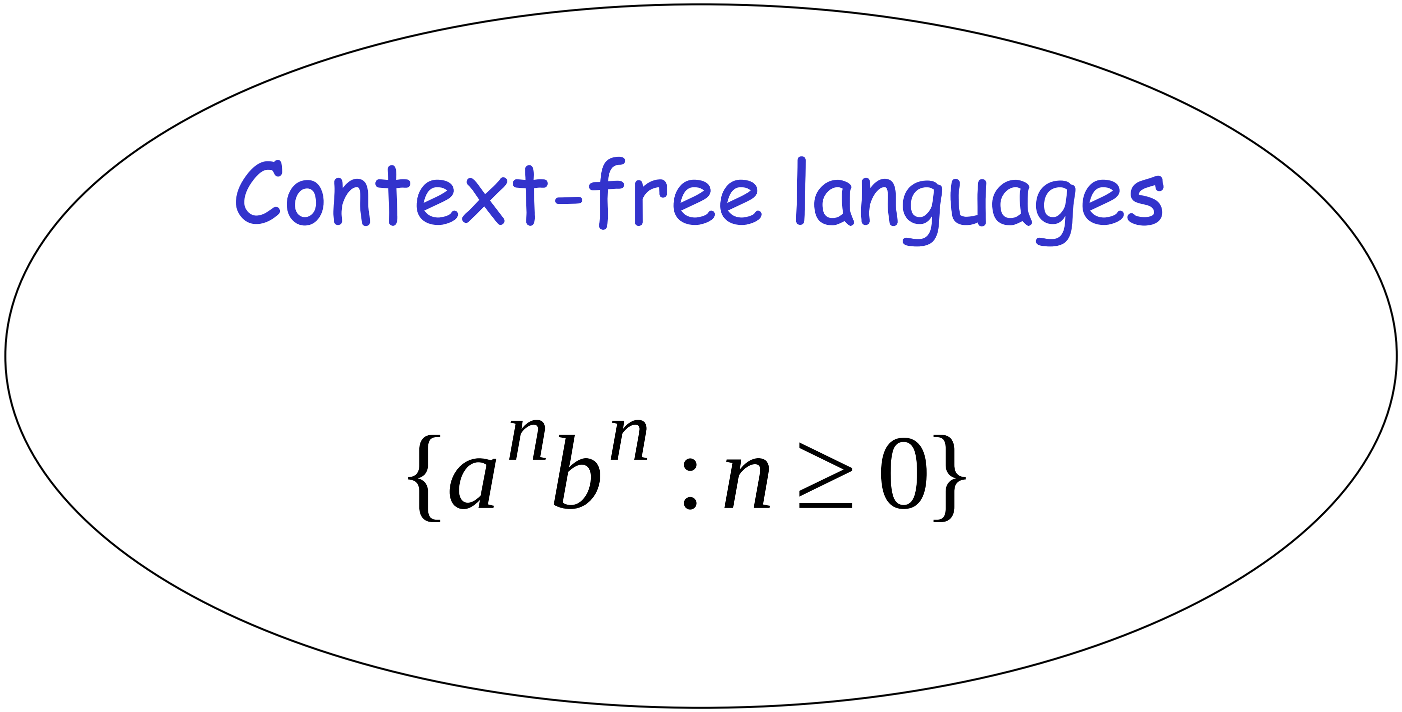
$$uv^i xy^i z \in L, \quad \text{for all } i \geq 0$$

Applications of the Pumping Lemma For Context-Free Languages

Linz 6th Section 8.1

Non-context free languages

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



Context-free languages

$$\{a^n b^n : n \geq 0\}$$

Linz 6th, section 8.1, example 8.1, page 216

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

$w = a^m b^m c^m$

Cannot cut w st vy has the same number of a's, b's and c's

Theorem: The language

$$L = \{a^n b^n c^n : n \geq 0\}$$

is **not** context free

Proof: Use the Pumping Lemma for context-free languages

$$L = \{a^n b^n c^n : n \geq 0\}$$

Assume for contradiction that L is context-free

Since L is context-free and infinite we can apply the pumping lemma

$$L = \{a^n b^n c^n : n \geq 0\}$$

Pumping Lemma gives a magic number m such that:

Pick any string $w \in L$ with length $|w| \geq m$

We pick: $w = a^m b^m c^m$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

We can write: $w = uvxyz$

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

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Pumping Lemma says:

$uv^i xy^i z \in L$ for all $i \geq 0$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

We examine **<u>all</u>** the possible locations of string vxy in w

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 1: vxy is within a^m

$$\overbrace{aaa...aaa}^{m}\,\overbrace{bbb...bbb}^{m}\,\overbrace{ccc...ccc}^{m}$$

$$\underbrace{a}_{u}\,\underbrace{aa...aa}_{vxy}\,\underbrace{a\,bbb...bbb\,ccc...ccc}_{z}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 1: v and y consist from only a

$$\overbrace{\underbrace{a}_{u}\underbrace{aa\ldots aa}_{vxy}a}^{m}\,\underbrace{\overbrace{bbb\ldots bbb}^{m}\,\overbrace{ccc\ldots ccc}^{m}}_{z}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 1: Repeating v and y

$$k \geq 1$$

$$\mathrlap{\overbrace{}^{m+k}}\underbrace{a}_{u}\underbrace{aaaaa\ldots aaaaa}_{v^2xy^2}\underbrace{a\,\overbrace{bbb\ldots bbb}^{m}\,\overbrace{ccc\ldots ccc}^{m}}_{z}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 1: From Pumping Lemma: $uv^2xy^2z \in L$

$$k \geq 1$$

$$\rlap{$\displaystyle\overbrace{}^{m+k}$}\underbrace{a}_{u}\underbrace{aaaaa...aaaa}_{v^2xy^2}\underbrace{aa\,\overbrace{bbb...bbb}^{m}\,\overbrace{ccc...ccc}^{m}}_{z}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 1: From Pumping Lemma: $uv^2xy^2z \in L$

$$k \geq 1$$

However: $uv^2xy^2z = a^{m+k} b^m c^m \notin L$

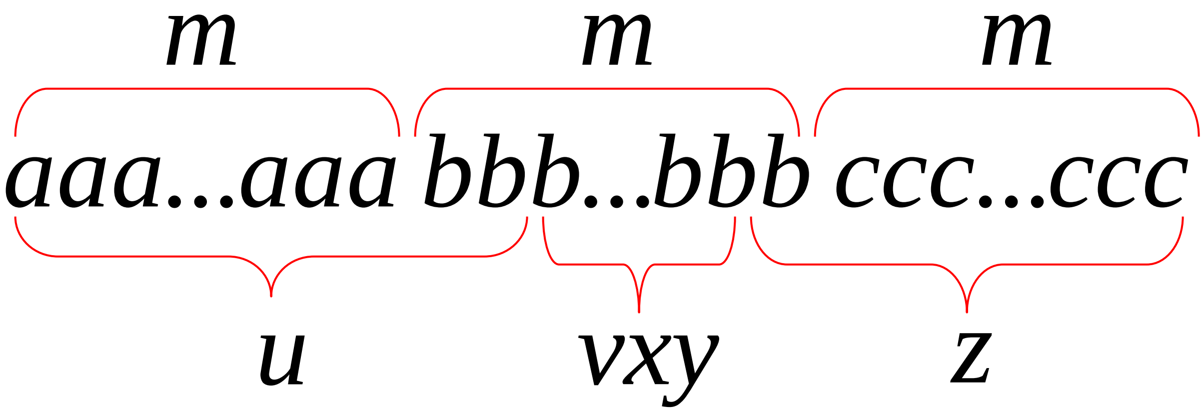
Contradiction!!!

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 2: vxy is within b^m

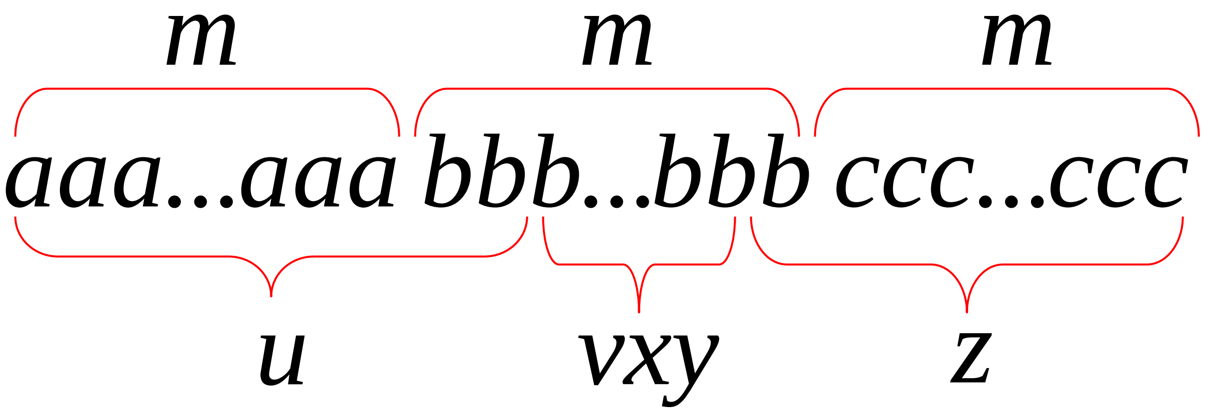


$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 2: Similar analysis with case 1



$$\overbrace{aaa...aaa}^{m}\,\overbrace{bbb...bbb}^{m}\,\overbrace{ccc...ccc}^{m}$$

$$u \qquad vxy \qquad z$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 3: vxy is within c^m

$$\underbrace{\overbrace{aaa...aaa}^{m}\,\overbrace{bbb...bbb}^{m}\,cc}_{u}\overbrace{\underbrace{c...c}_{vxy}\,\underbrace{cc}_{z}}^{m}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 3: Similar analysis with case 1

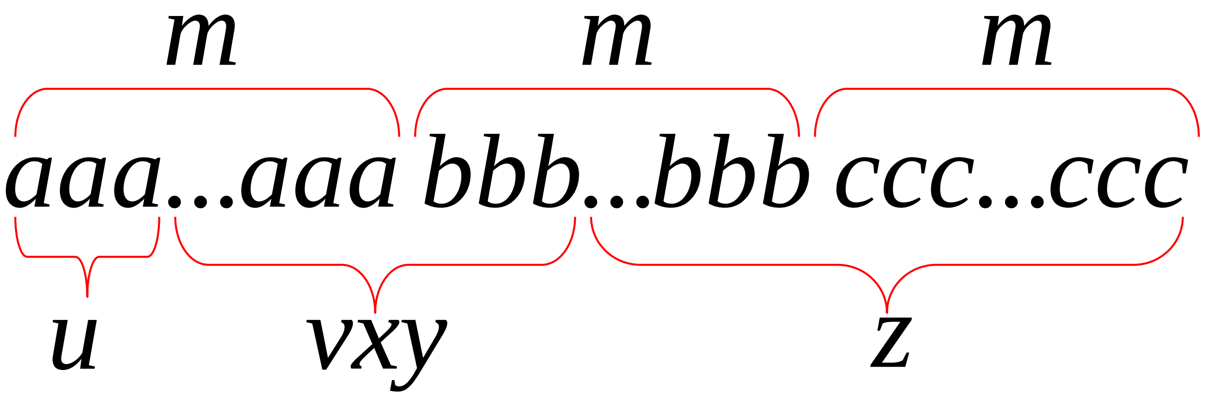
$$\underbrace{\overbrace{aaa\ldots aaa}^{m}\,\overbrace{bbb\ldots bbb}^{m}}_{u}\,\overbrace{\underbrace{ccc\ldots c}_{vxy}\,\underbrace{cc}_{z}}^{m}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: vxy overlaps a^m and b^m



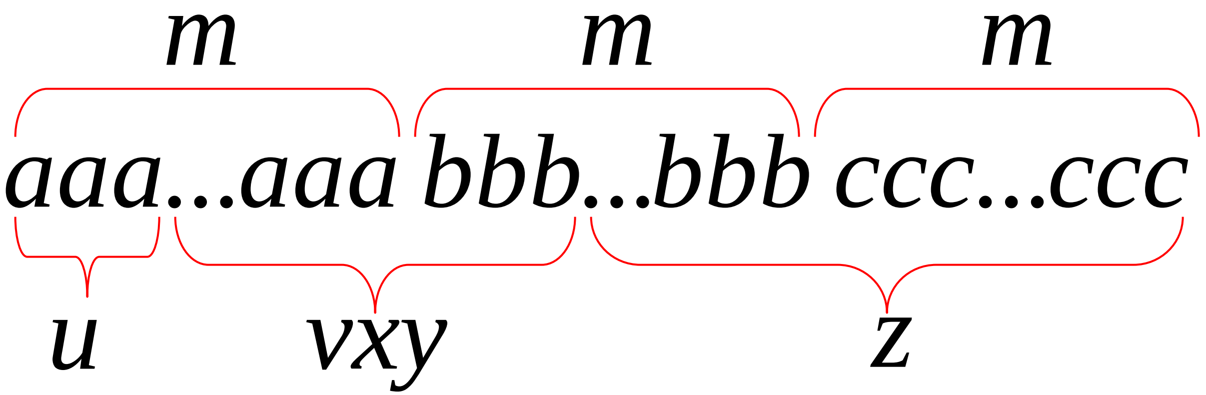
$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: Possibility 1: v contains only a

y contains only b



$$L = \{a^n b^n c^n : n \geq 0\}$$

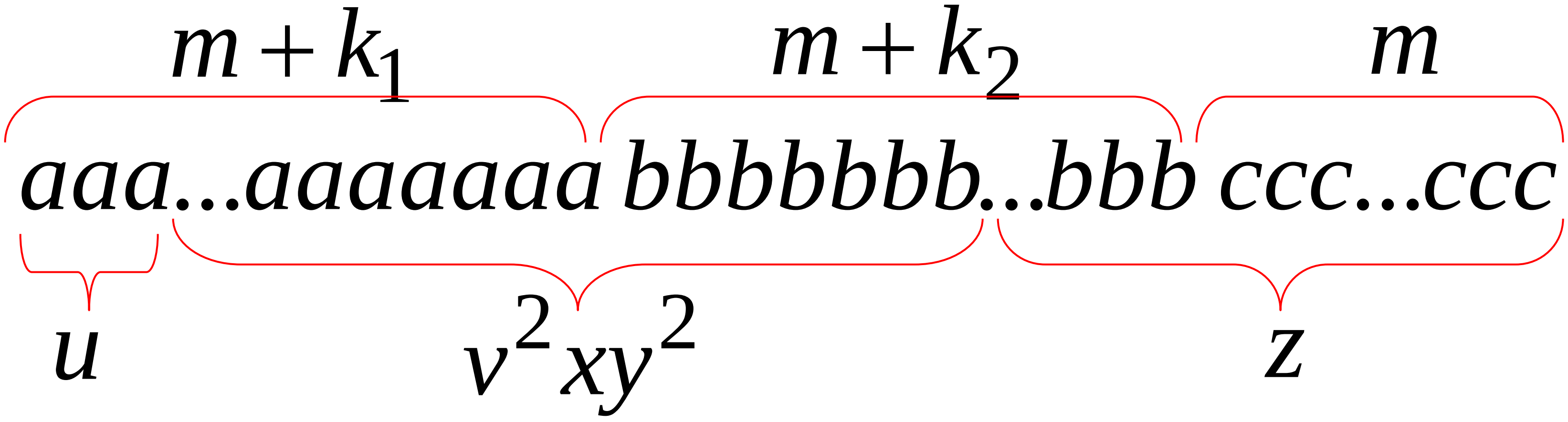
$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: Possibility 1: v contains only a

y contains only b

$$k_1 + k_2 \geq 1$$



$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: From Pumping Lemma: $uv^2xy^2z \in L$

$$k_1 + k_2 \geq 1$$

$$\overbrace{aaa\ldots aaaaaaa}^{m+k_1}\,\overbrace{bbbbbbb\ldots bbb}^{m+k_2}\,\overbrace{ccc\ldots ccc}^{m}$$

$$\underbrace{aaa}_{u}\underbrace{\ldots aaaaaaa\,bbbbbbb}_{v^2xy^2}\underbrace{\ldots bbb\,ccc\ldots ccc}_{z}$$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$w = a^m b^m c^m$

$w = uvxyz$ $\quad |vxy| \leq m$ $\quad |vy| \geq 1$

Case 4: From Pumping Lemma: $uv^2xy^2z \in L$

$k_1 + k_2 \geq 1$

However: $uv^2xy^2z = a^{m+k_1} b^{m+k_2} c^m \notin L$

Contradiction!!!

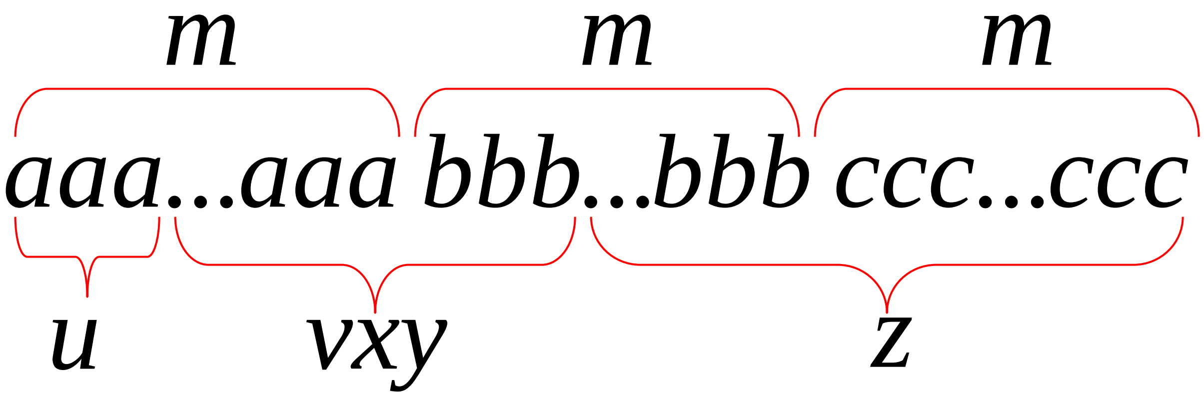
$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$w = uvxyz$ $\qquad |vxy| \leq m$ $\qquad |vy| \geq 1$

Case 4: Possibility 2: v contains a and b

y contains only b



$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: Possibility 2: v contains a and b

$k_1 + k_2 + k \geq 1$ $\qquad$ y contains only b

$$\overbrace{aaa...aaaaa}^{m}\overbrace{bb}^{k_1}\overbrace{aa}^{k_2}\overbrace{bb\ bbbbbbb...bbb}^{m+k}\ \overbrace{ccc...ccc}^{m}$$

$$\underbrace{aaa}_{u}\underbrace{...aaaaabbaabb\ bbbbbbb}_{v^2xy^2}\underbrace{...bbb\ ccc...ccc}_{z}$$

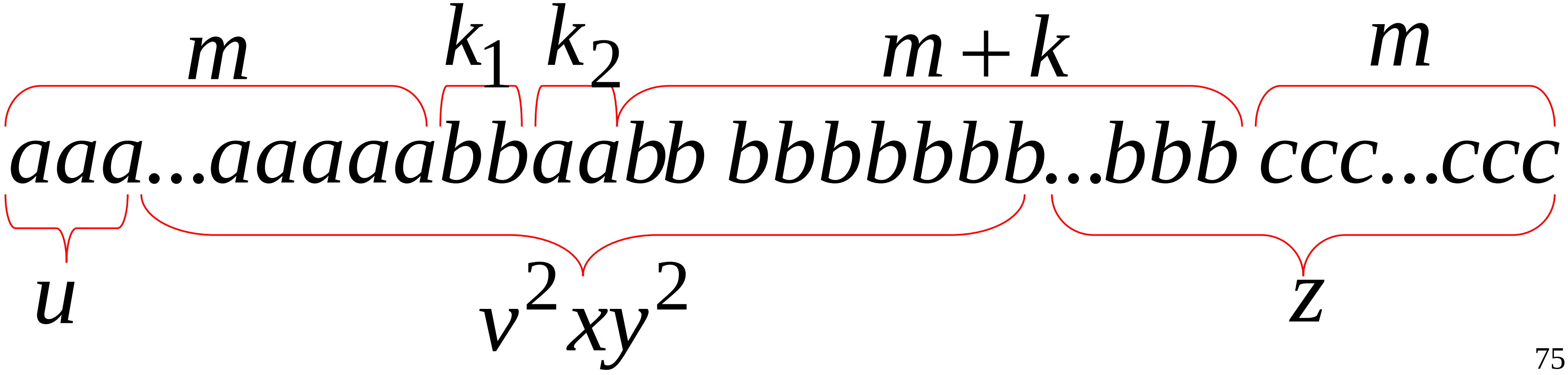
$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: From Pumping Lemma: $uv^2xy^2z \in L$

$$k_1 + k_2 + k \geq 1$$



$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: From Pumping Lemma: $uv^2xy^2z \in L$

However: $k_1 + k_2 + k \geq 1$

$$uv^2xy^2z = a^m b^{k_1} a^{k_2} b^{m+k} c^m \notin L$$

Contradiction!!!

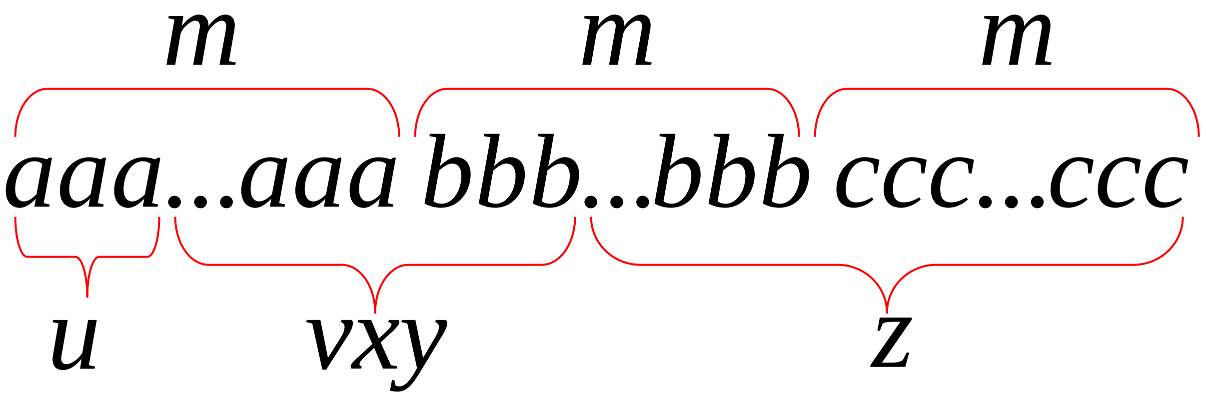
$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: Possibility 3: v contains only a

y contains a and b



$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 4: Possibility 3: v contains only a

y contains a and b

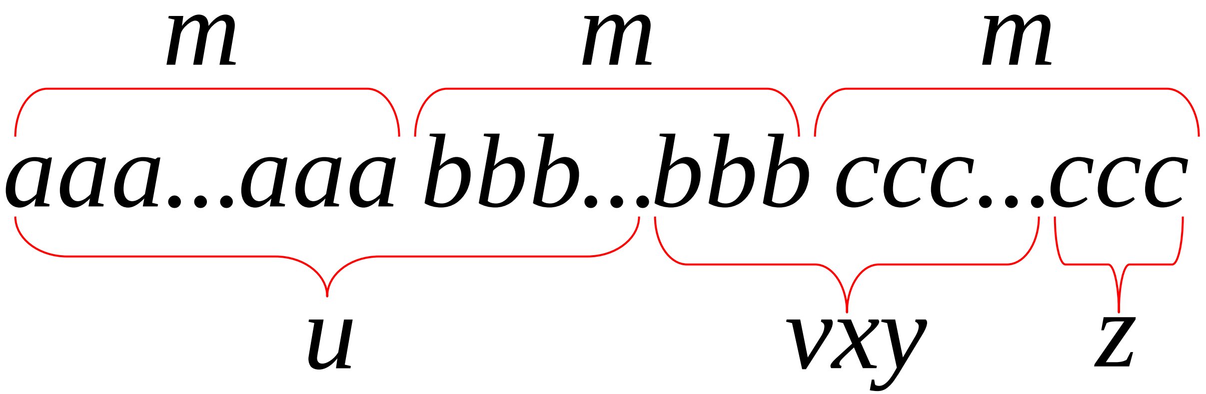
Similar analysis with Possibility 2

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 5: vxy overlaps b^m and c^m



$m \quad m \quad m$

$aaa...aaa\ bbb...bbb\ ccc...ccc$

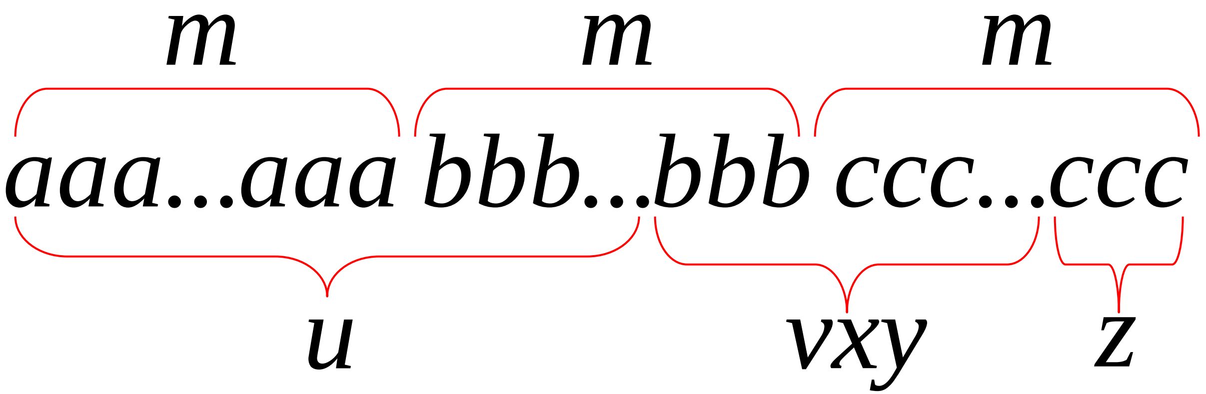
$u \quad vxy \quad z$

$$L = \{a^n b^n c^n : n \geq 0\}$$

$$w = a^m b^m c^m$$

$$w = uvxyz \qquad |vxy| \leq m \qquad |vy| \geq 1$$

Case 5: Similar analysis with case 4



There are no other cases to consider

(since $|vxy| \leq m$, string vxy cannot

overlap a^m, b^m and c^m at the same time)

In all cases we obtained a contradiction

Therefore: The original assumption that

$$L = \{a^n b^n c^n : n \geq 0\}$$

is context-free must be wrong

Conclusion: L is not context-free