# Positive Properties of Context-Free languages

# Union

Context-free languages are closed under: Union

$$L_1$$
 is context free 
$$L_1 \cup L_2$$
 
$$L_2$$
 is context free is context-free

# Example

# Language

$$L_1 = \{a^n b^n\}$$

$$S_1 \rightarrow aS_1b \mid \lambda$$

$$L_2 = \{ww^R\}$$

$$S_2 \rightarrow aS_2a \mid bS_2b \mid \lambda$$

# **Union**

$$L = \{a^n b^n\} \cup \{ww^R\}$$

$$S \rightarrow S_1 \mid S_2$$

# Concatenation

Context-free languages are closed under: Concatenation

 $L_1$  is context free  $L_1L_2$   $L_2$  is context free is context-free

# Example

# Language

$$L_1 = \{a^n b^n\}$$

$$S_1 \rightarrow aS_1b \mid \lambda$$

$$L_2 = \{ww^R\}$$

$$S_2 \rightarrow aS_2a \mid bS_2b \mid \lambda$$

# Concatenation

$$L = \{a^n b^n\} \{ww^R\}$$

$$S \rightarrow S_1 S_2$$

# Star Operation

Context-free languages are closed under: Star-operation

L is context free  $\stackrel{*}{\Longrightarrow}$   $L^*$  is context-free

# Example

# Language

# Grammar

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

$$S \rightarrow aSb \mid \lambda$$

# Star Operation

$$L = \{a^n b^n\}^*$$

$$S_1 \rightarrow SS_1 \mid \lambda$$

# Negative Properties of Context-Free Languages

# Intersection

Context-free languages are <u>not</u> closed under:

intersection

 $L_1$  is context free  $L_1 \cap L_2$   $L_2$  is context free  $\underbrace{ \text{not necessarily context-free} }$ 

# Example

$$L_1 = \{a^n b^n c^m\}$$

$$L_2 = \{a^n b^m c^m\}$$

# Context-free:

$$S \rightarrow AC$$

$$S \rightarrow AB$$

$$A \rightarrow aAb \mid \lambda$$

$$A \rightarrow aA \mid \lambda$$

$$C \rightarrow cC \mid \lambda$$

$$B \rightarrow bBc \mid \lambda$$

# Intersection

$$L_1 \cap L_2 = \{a^n b^n c^n\}$$
 NOT context-free

# Complement

Context-free languages are <u>not</u> closed under:

complement

L is context free  $\longrightarrow L$ 

not necessarily
context-free

# Example

$$L_1 = \{a^n b^n c^m\}$$

$$L_2 = \{a^n b^m c^m\}$$

# Context-free:

$$S \rightarrow AC$$

$$S \rightarrow AB$$

$$A \rightarrow aAb \mid \lambda$$

$$A \rightarrow aA \mid \lambda$$

$$C \rightarrow cC \mid \lambda$$

$$B \rightarrow bBc \mid \lambda$$

# Complement

$$\overline{L_1 \cup L_2} = L_1 \cap L_2 = \{a^n b^n c^n\}$$

**NOT** context-free

\* Exception >>
The intersection of
a context-free language and
a regular language
is a context-free language

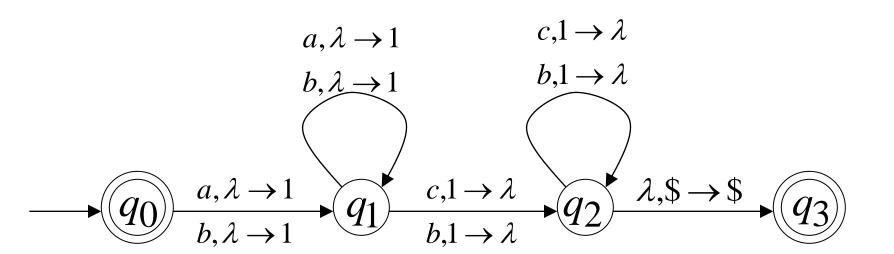
 $L_1$  context free  $L_1 \cap L_2$   $L_2$  regular context-free

# Example:

# context-free

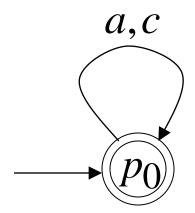
$$L_1 = \{w_1w_2 : |w_1| = |w_2|, w_1 \in \{a,b\}^*, w_2 \in \{c,d\}^*\}$$

# NPDA $M_1$



regular 
$$L_2 = \{a, c\}^*$$

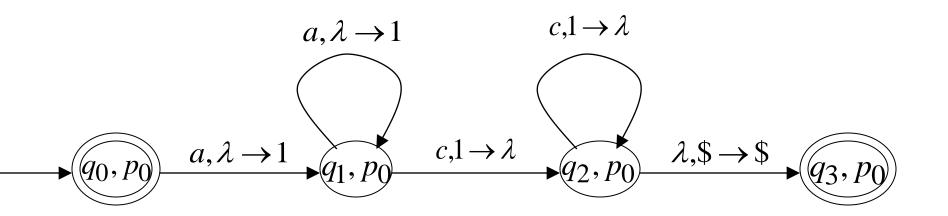
# DFA $M_2$



#### context-free

Automaton for: 
$$L_1 \cap L_2 = \{a^n c^n : n \ge 0\}$$

# NPDA M



# An Application of Regular Closure

Prove that: 
$$L = \{a^n b^n : n \neq 100, n \geq 0\}$$

is context-free

$$\{a^nb^n:n\geq 0\}$$
 is context-free

# We also know:

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

$$\{a^nb^n\}$$

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

context-free

regular





(regular closure)  $\{a^nb^n\}\cap L_1$  context-free



$$\{a^n b^n\} \cap \overline{L_1} = \{a^n b^n: n \neq 100, n \geq 0\} = L$$

is 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

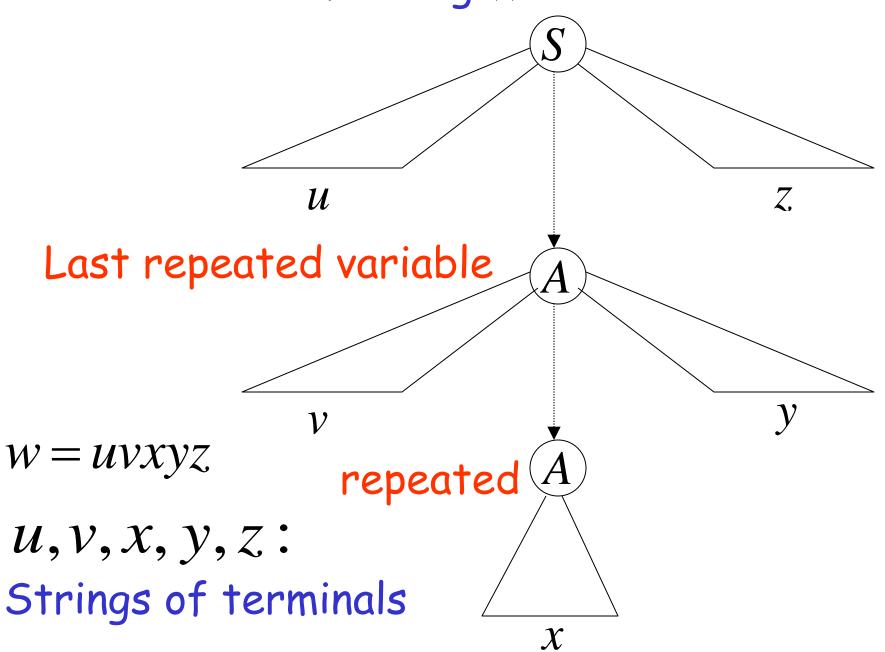
(regular closure)

Then 
$$L \cap \{a*b*c*\} = \{a^nb^nc^n\}$$
 context-free regular context-free **Impossible!!!**

Therefore, L is not context free

# The Pumping Lemma for Context-Free Languages

# Derivation tree of string W

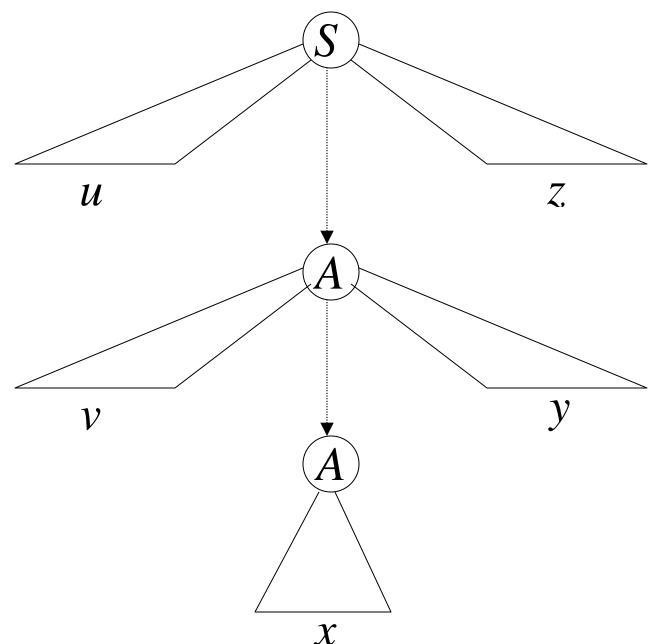


# Possible derivations:



 $A \Rightarrow vAy$ 

 $A \Longrightarrow x$ 



$$S \Longrightarrow uAz$$

$$A \Rightarrow vAy$$

$$A \Longrightarrow x$$

$$* * UAz \Rightarrow uxz$$

$$uv^0xy^0z$$

$$S \Rightarrow uAz$$

$$A \Rightarrow vAy$$

$$A \Longrightarrow x$$

$$*$$
 $S \Rightarrow uAz \Rightarrow uvAyz \Rightarrow uvxyz$ 

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

$$S \Longrightarrow uAz$$

$$A \Rightarrow vAy$$

$$A \Rightarrow x$$

$$S \Longrightarrow uAz \Longrightarrow uvAyz \Longrightarrow uvvAyyz \Longrightarrow uvvxyyz$$

$$uv^2xy^2z$$

$$S \Longrightarrow uAz \qquad \qquad * \qquad * \qquad A \Longrightarrow x$$

$$S \stackrel{*}{\Rightarrow} uAz \stackrel{*}{\Rightarrow} uvAyz \stackrel{*}{\Rightarrow} uvvAyyz \stackrel{*}{\Rightarrow} \dots$$

$$\stackrel{*}{\Rightarrow} uvvV \cdot \cdot \cdot vAy \cdot \cdot \cdot \cdot yyyz \stackrel{*}{\Rightarrow} \dots$$

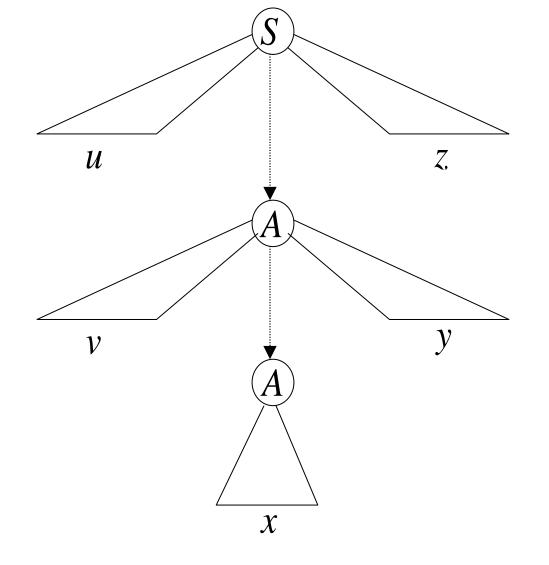
$$\stackrel{*}{\Rightarrow} uvvV \cdot \cdot \cdot vXy \cdot \cdot \cdot \cdot yyyz$$

$$uv^ixy^iz$$

# Therefore, any string of the form

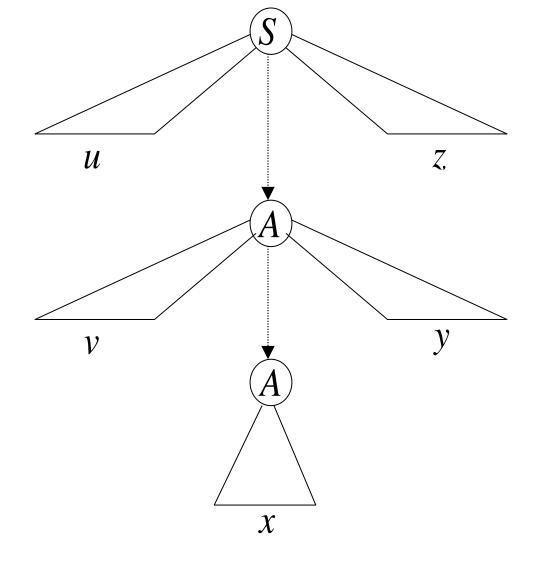
$$uv^i xy^i z$$
  $i \ge 0$ 

is generated by the grammar G



Observation:  $|vxy| \leq m$ 

m is the number of states in PDA



Observation:  $|vy| \ge 1$ 

Since repetitions are done on both v and  $y_{_{2}}$ 

# The Pumping Lemma:

For infinite context-free language L there exists an integer m such that

for any string  $w \in L$ ,  $|w| \ge m$ 

we can write w = uvxyz

with lengths  $|vxy| \le m$  and  $|vy| \ge 1$ 

and it must be:

 $uv^i x y^i z \in L$ , for all  $i \ge 0$ 

# Applications of The Pumping Lemma

# Non-context free languages

$$\{a^nb^nc^n: n \ge 0\}$$



$$\{a^nb^n: n \ge 0\}$$

# Theorem: The language

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

is **not** context free

Proof: Use the Pumping Lemma for context-free languages

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

Assume for contradiction that L is context-free

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

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

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

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

We can write: 
$$w = uvxyz$$

with lengths 
$$|vxy| \le m$$
 and  $|vy| \ge 1$ 

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

# Pumping Lemma says:

$$uv^i x y^i z \in L$$
 for all  $i \ge 0$ 

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

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 1: vxy is within  $a^m$ 

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 1: Repeating 
$$v$$
 and  $y$ 

$$k \ge 1$$

$$m+k$$
  $m$ 

aaaaaaaaaaaa bbb...bbb ccc...ccc

$$u v^2 x y^2$$

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

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

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

Contradiction!!!

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 2: vxy is within  $b^m$ 

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

# Case 2: Similar analysis with case 1

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 3: 
$$vxy$$
 is within  $c^m$ 

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

# Case 3: Similar analysis with case 1

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

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

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 4: Possibility 1: v contains only a y contains only b

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 4: Possibility 1: 
$$v$$
 contains only  $a$ 
 $k_1 + k_2 \ge 1$ 
 $y$  contains only  $b$ 
 $m + k_1$ 
 $m + k_2$ 
 $m$ 
 $aaa...aaaaaaaa bbbbbbbb...bbb ccc...ccc$ 
 $u$ 
 $v^2 x v^2$ 
 $z$ 

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 4: From Pumping Lemma:  $uv^2xy^2z \in L$   $k_1 + k_2 \ge 1$ 

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

Contradiction!!!

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 4: Possibility 2: v contains a and b y contains only b

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

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

$$w = uvxyz \qquad | v$$

$$|vxy| \le m$$
  $|vy| \ge 1$ 

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

However:

$$k_1 + k_2 + k \ge 1$$

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

Contradiction!!!

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 4: Possibility 3: v contains only a y contains a and b

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

Case 4: Possibility 3: 
$$v$$
 contains only  $a$   $y$  contains  $a$  and  $b$ 

Similar analysis with Possibility 2

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

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

$$w = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

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

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

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

$$w = uvxyz |vxy| \le m |vy| \ge 1$$

# Case 5: Similar analysis with case 4

#### There are no other cases to consider

(since  $|vxy| \le m$ , 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

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

is context-free must be wrong

Conclusion: L is not context-free

## The Pumping Lemma:



For infinite context-free language L

there exists an integer m such that

for any string 
$$w \in L$$
,  $|w| \ge m$ 

we can write w = uvxyz

with lengths  $|vxy| \le m$  and  $|vy| \ge 1$ 

and it must be:

$$uv^i x y^i z \in L$$
, for all  $i \ge 0$ 

# Non-context free languages

$$\{a^n b^n c^n : n \ge 0\}$$
  $\{vv: v \in \{a, b\}^*\}$ 

# Context-free languages

$$\{a^n b^n : n \ge 0\}$$
  $\{ww^R : w \in \{a, b\}^*\}$ 

# Theorem: The language

$$L = \{vv : v \in \{a,b\}^*\}$$

is **not** context free

Proof: Use the Pumping Lemma for context-free languages

$$L = \{vv : v \in \{a,b\}^*\}$$

Assume for contradiction that L is context-free

Pick any string of L with length at least m

we pick: 
$$a^m b^m a^m b^m \in L$$

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

We examine <u>all</u> the possible locations of string vxy in  $a^mb^ma^mb^m$ 

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 1: vxy is within the first  $a^m$ 

$$v = a^{k_1} \qquad y = a^{k_2} \qquad k_1 + k_2 \ge 1$$

$$L = \{vv : v \in \{a, b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 1: vxy is within the first  $a^m$ 

$$v = a^{k_1} \qquad y = a^{k_2} \qquad k_1 + k_2 \ge 1$$

$$L = \{vv : v \in \{a, b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 1: vxy is within the first  $a^m$ 

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

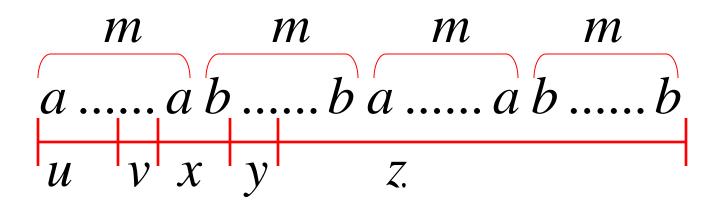
Contradiction!!! Since  $k_1 + k_2 \ge 1$ 

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 2: 
$$v$$
 is in the first  $a^m$   $y$  is in the first  $b^m$ 

$$v = a^{k_1}$$
  $y = b^{k_2}$   $k_1 + k_2 \ge 1$ 

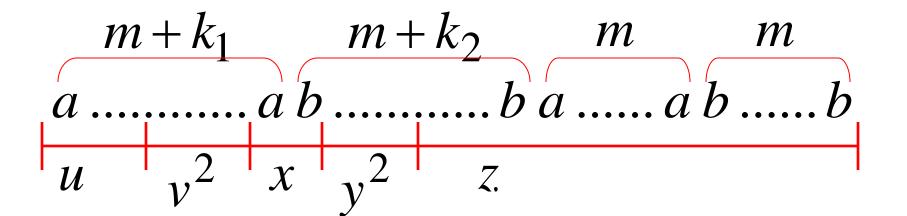


$$L = \{vv : v \in \{a, b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 2: 
$$v$$
 is in the first  $a^m$   $y$  is in the first  $b^m$ 

$$v = a^{k_1}$$
  $y = b^{k_2}$   $k_1 + k_2 \ge 1$ 



$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 2: v is in the first  $a^m$  y is in the first  $b^m$ 

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

Contradiction!!! Since  $k_1 + k_2 \ge 1$ 

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 3: 
$$v$$
 overlaps the first  $a^m b^m$   $y$  is in the first  $b^m$ 

$$v = a^{k_1} b^{k_2}$$
  $y = b^{k_3}$   $k_1, k_2 \ge 1$ 

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 3: 
$$v$$
 overlaps the first  $a^m b^m$   $y$  is in the first  $b^m$ 

$$v = a^{k_1} b^{k_2} \qquad y = b^{k_3} \qquad k_1, k_2 \ge 1$$

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 3: v overlaps the first  $a^m b^m$  y is in the first  $b^m$ 

$$a^m b^{k_2} a^{k_1} b^{m+k_3} a^m b^m = uv^2 xy^2 z \notin L$$

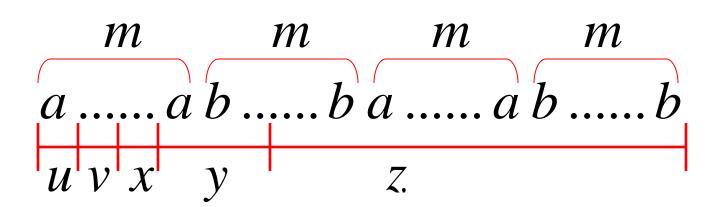
Contradiction!!!

$$L = \{vv : v \in \{a,b\}^*\}$$

$$a^m b^m a^m b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Case 4: 
$$v$$
 in the first  $a^m$   
 $y$  Overlaps the first  $a^mb^m$ 

## Analysis is similar to case 3



$$vxy$$
 is within  $a^mb^ma^mb^m$ 

or

$$a^m b^m a^m b^m$$

$$a^m b^m a^m b^m$$

### Analysis is similar to case 1:

$$a^mb^ma^mb^m$$

$$vxy$$
 overlaps  $a^mb^ma^mb^m$ 

or

$$a^m b^m a^m b^m$$

## Analysis is similar to cases 2,3,4:

$$a^m b^m a^m b^m$$

#### There are no other cases to consider

Since  $|vxy| \le m$ , it is impossible vxy to overlap:

 $a^m b^m a^m b^m$ 

nor

 $a^m b^m a^m b^m$ 

nor

 $a^m b^m a^m b^m$ 

#### In all cases we obtained a contradiction

Therefore: The original assumption that

$$L = \{vv : v \in \{a,b\}^*\}$$

is context-free must be wrong

Conclusion: L is not context-free

## Theorem: The language

$$L = \{a^{n^2}b^n : n \ge 0\}$$

is **not** context free

Proof: Use the Pumping Lemma for context-free languages

$$L = \{a^{n^2}b^n : n \ge 0\}$$

Assume for contradiction that L is context-free

Pick any string of L with length at least m

we pick: 
$$a^{m^2}b^m \in L$$

$$L = \{a^{n^2}b^n : n \ge 0\}$$

We can write: 
$$a^{m^2}b^m = uvxyz$$
 with lengths  $|vxy| \le m$  and  $|vy| \ge 1$ 

## Pumping Lemma says:

$$uv^i x y^i z \in L$$
 for all  $i \ge 0$ 

We examine all the possible locations of string vxy in  $a^{m^2}b^m$ 

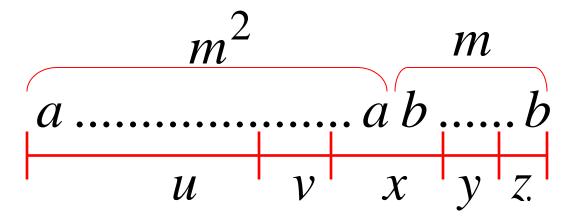
$$L = \{a^{n^2}b^n : n \ge 0\}$$

$$a^{m^2}b^m = uvxyz$$

$$|vxy| \le m \quad |vy| \ge 1$$

Most complicated case: v is in  $a^m$  y is in  $b^m$ 

$$v = a^{k_1} \qquad y = b^{k_2} \qquad 1 \le k_1 + k_2 \le m$$

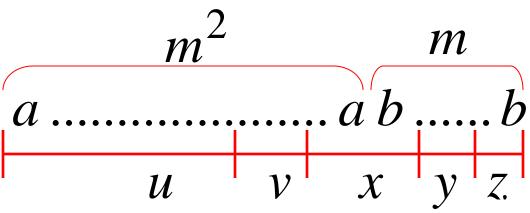


$$L = \{a^{n^2}b^n : n \ge 0\}$$

$$a^{m^2}b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

## Most complicated sub-case: $k_1 \neq 0$ and $k_2 \neq 0$

$$v = a^{k_1} \qquad y = b^{k_2} \qquad 1 \le k_1 + k_2 \le m$$



$$L = \{a^{n^2}b^n : n \ge 0\}$$

$$a^{m^2}b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Most complicated sub-case:  $k_1 \neq 0$  and  $k_2 \neq 0$ 

$$v = a^{k_1}$$
  $y = b^{k_2}$   $1 \le k_1 + k_2 \le m$ 

$$\frac{m^2 - k_1}{a \dots a b \dots b}$$

$$\frac{a \dots a b \dots b}{v^0 x y^0 z}$$

$$L = \{a^{n^2}b^n : n \ge 0\}$$

$$a^{m^2}b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

Most complicated sub-case:  $k_1 \neq 0$  and  $k_2 \neq 0$ 

$$v = a^{k_1} \qquad y = b^{k_2} \qquad 1 \le k_1 + k_2 \le m$$

$$a^{m^2 - k_1} b^{m - k_2} = u v^0 x y^0 z$$

Question: 
$$m^2 - k_1 = (m - k_2)^2$$
?

$$k_1 \neq 0 \text{ and } k_2 \neq 0 \qquad 1 \leq k_1 + k_2 \leq m$$

$$1 \le k_1 + k_2 \le m$$



$$(m-k_2)^2 \le (m-1)^2$$
  
 $\le m^2 - 2m + 1$   
 $< m^2 - k_1$ 



$$m^2 - k_1 \neq (m - k_2)^2$$

$$L = \{a^{n^2}b^n : n \ge 0\}$$

$$a^{m^2}b^m = uvxyz \qquad |vxy| \le m \qquad |vy| \ge 1$$

$$m^{2} - k_{1} \neq (m - k_{2})^{2}$$

$$\downarrow \qquad \qquad \downarrow$$

$$a^{m^{2} - k_{1}} b^{m - k_{2}} = uv^{0} xy^{0} z \neq L$$

#### Contradiction!!!

# After examining all cases, we will obtain a contradiction

Therefore: The original assumption that

$$L = \{a^{n^2}b^n : n \ge 0\}$$

is context-free must be wrong

Conclusion: L is not context-free