2023

Theory of Computation

Kun-Ta Chuang
Department of Computer Science and Information Engineering
National Cheng Kung University

Outline

1 Two Pumping Lemmas

2 Closure Properties and Decision Algorithms for CFLs

The Pumping Lemma for Context-Free Languages

Consider now an infinite context-free language L

Let G be the grammar of $L - \{\lambda\}$

。不包含λ

Take G so that L has no unit-productions

no λ-productions

Let p = (Number of productions) x
(Largest right side of a production)

n條 production

↳ 最長的 production ⇒ f

Let $m = p + 1$ (Largest number of states in NPDA)

Example : $G \begin{cases} S \to AB \\ A \to aBb \\ B \to Sb \\ B \to b \end{cases}$

$$p = 4 \times 3 = 12$$

$$m = p + 1 = 13$$

Take a string $w \in L(G)$
with length $|w| \geq m$

↳ 找一個長度 ≥ m 且 ∈ L(G) 的 string

⇒ 必有 loop

We will show:

in the derivation of w
a ~~variable~~ (production) of G is repeated

$$S \stackrel{*}{\Rightarrow} w$$

$$v_1 \Rightarrow v_2 \Rightarrow \cdots \Rightarrow v_k \Rightarrow w$$

$$S = v_1$$

$$v_1 \Rightarrow v_2 \Rightarrow \cdots \Rightarrow v_k \Rightarrow w$$

new old

$$|v_{i+1}| \leq |v_i| + f$$

maximum right hand side of any production

↳ 右邊 production 最長的

$$|w| < k \cdot f$$

k 步

$$m \leq |w| < k \cdot f \quad \Longrightarrow \quad p < k \cdot f$$

$$v_1 \Rightarrow v_2 \Rightarrow \cdots \Rightarrow v_k \Rightarrow w$$

$$p < k \cdot f$$

$$\Downarrow$$

$$k > \frac{p}{f}$$

Number of productions in grammar

↳ num of production

$$v_1 \Rightarrow v_2 \Rightarrow \cdots \Rightarrow v_k \Rightarrow w$$

$k >$ Number of productions in grammar

$\Downarrow$

Some production must be repeated

$$v_1 \Rightarrow \cdots \Rightarrow a_1 A a_2 \Rightarrow \cdots \Rightarrow a_3 A a_4 \Rightarrow \cdots \Rightarrow w$$

Repeated variable

重覆用了rule

$$S \rightarrow r_1$$
$$A \rightarrow r_2$$
$$B \rightarrow r_2$$
$$\cdots$$

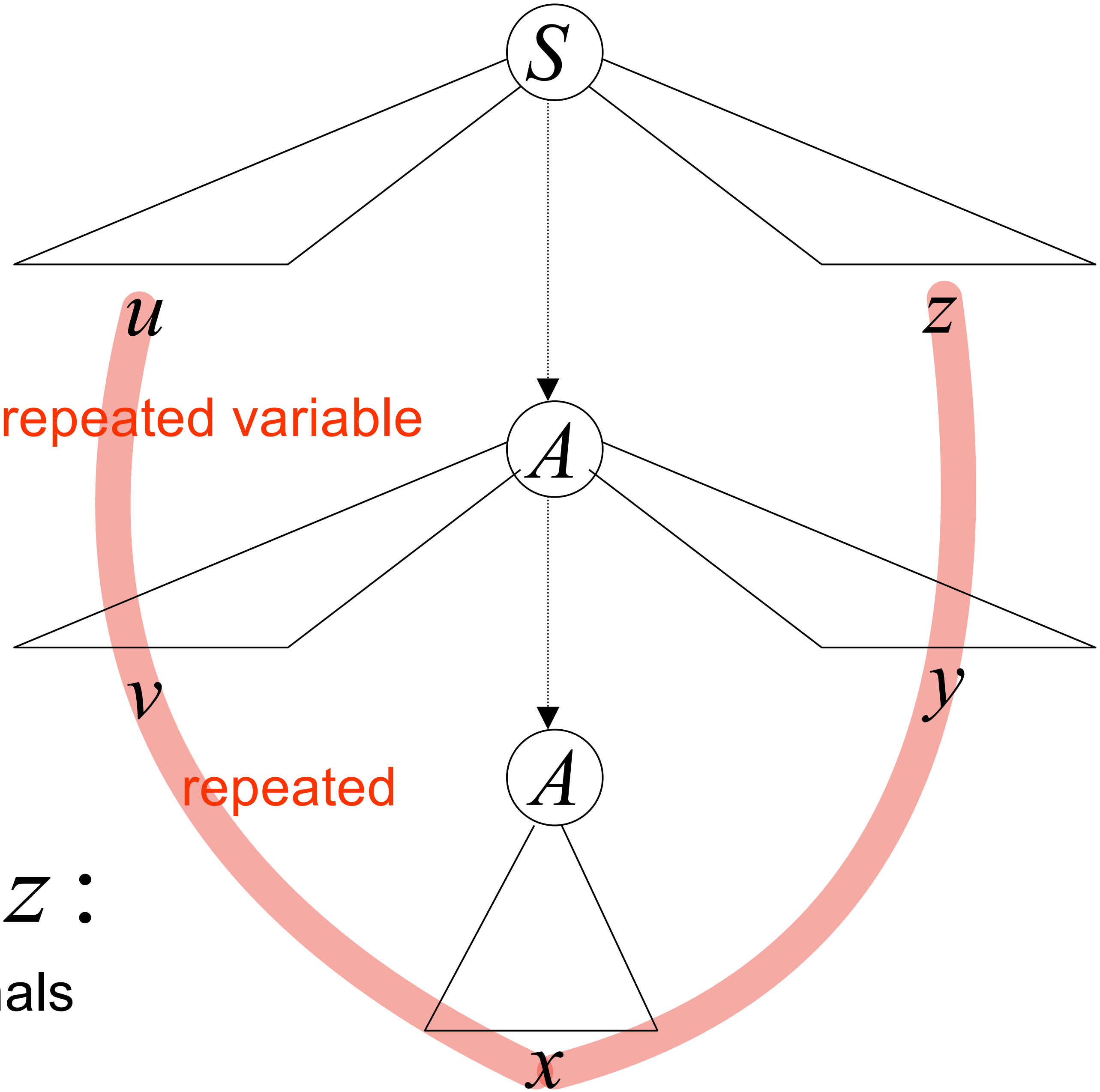
$w \in L(G)$ $|w| \geq m$

Derivation of string w

$$S \Rightarrow \cdots \Rightarrow a_1 A a_2 \Rightarrow \cdots \Rightarrow a_3 A a_4 \Rightarrow \cdots \Rightarrow w$$

Some variable is repeated

Derivation tree of string w



$$w = uvxyz$$

$$u, v, x, y, z :$$

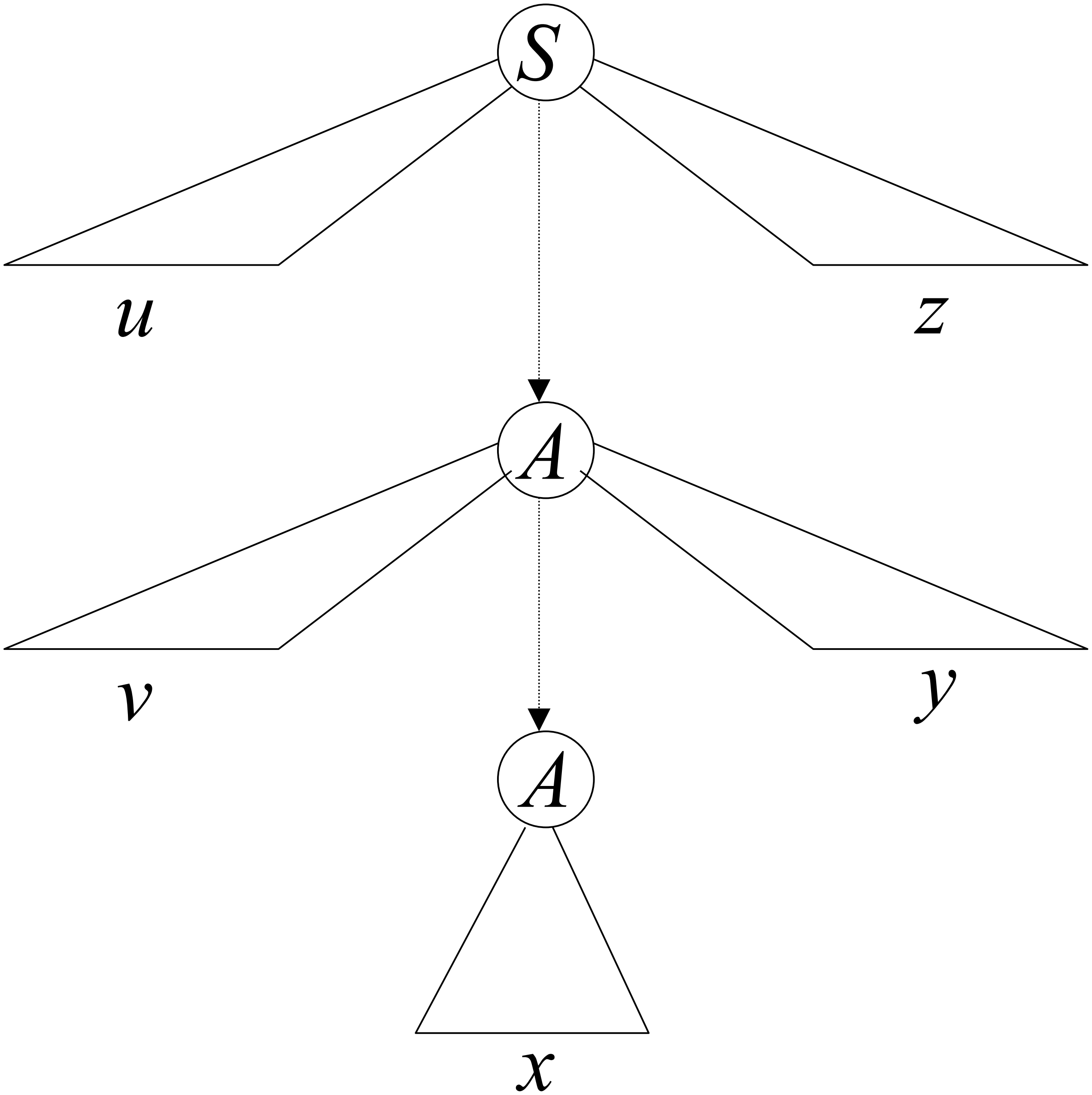
Strings of terminals

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:

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

代一次

This string is also generated:

$$S \overset{*}{\Rightarrow} uAz \overset{*}{\Rightarrow} u\underline{vAy}z \overset{*}{\Rightarrow} uvxyz$$

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

We know:

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

代2次

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:

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

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

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

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

$$uv^i xy^i z$$

Therefore, any string of the form

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

is generated by the grammar G

Therefore,

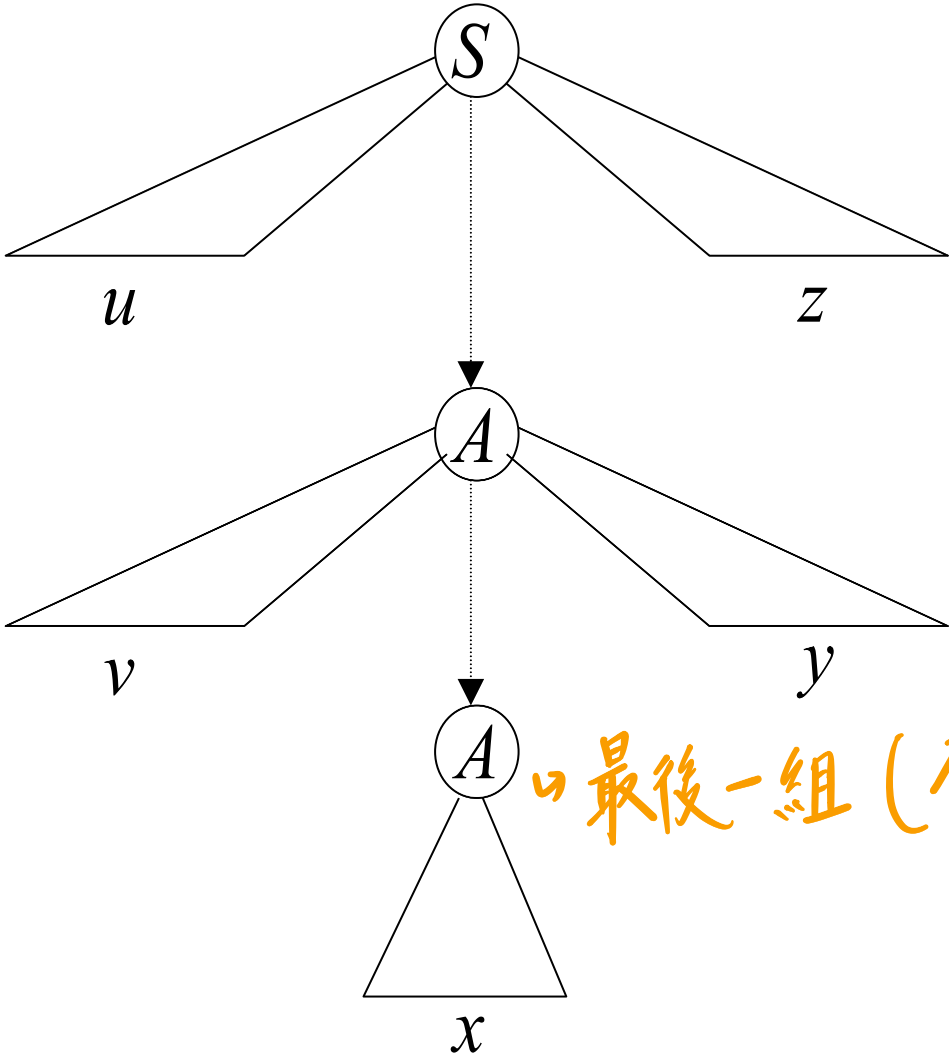
knowing that $uvxyz \in L(G)$

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

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

$$\Downarrow$$

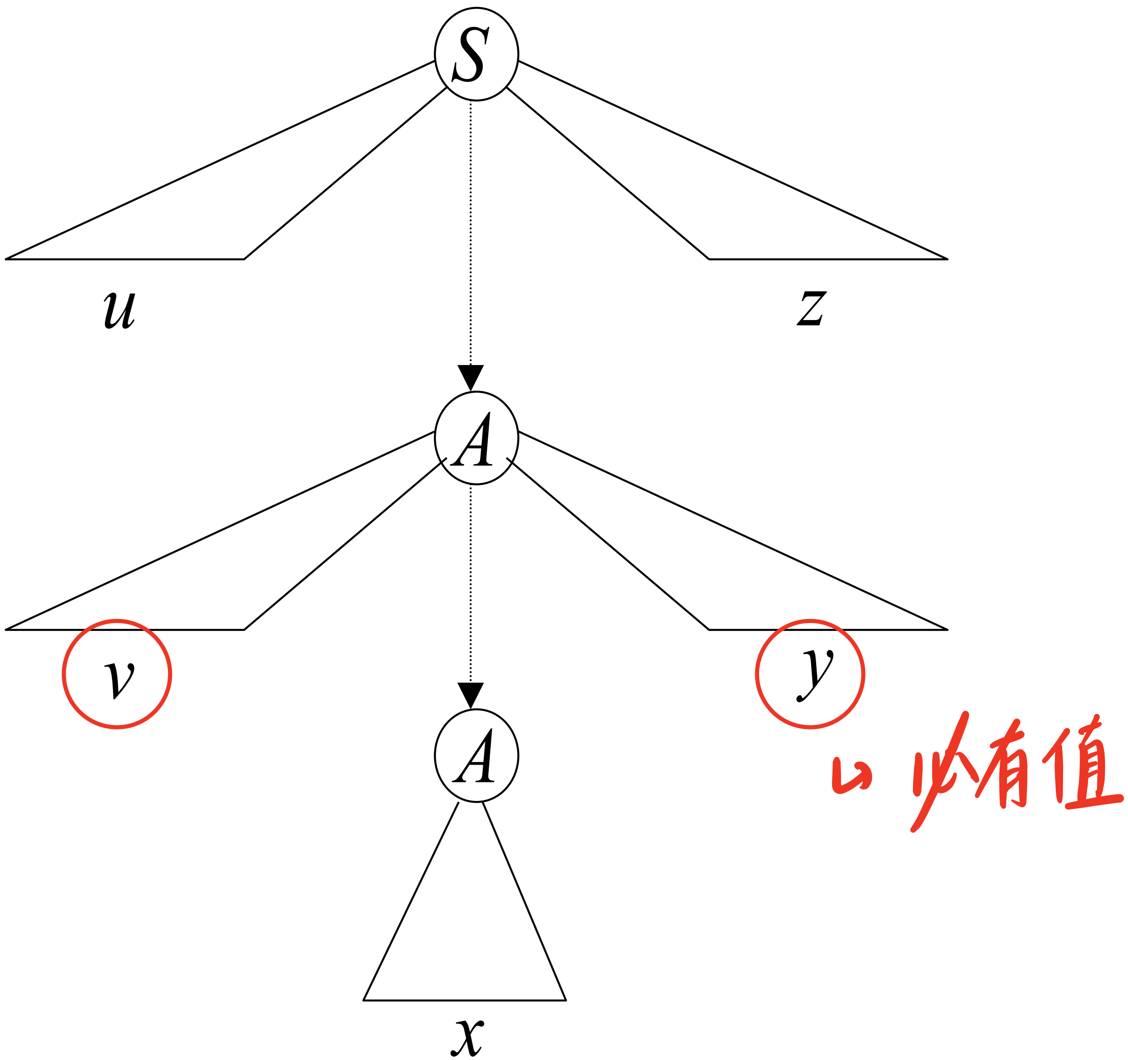
$$uv^i xy^i z \in L$$



最後一組 (不會有 unit production)
$A \to A$

Observation: $|vxy| \leq m$ 沒有重覆

Since A is the last repeated variable



Observation: $|vy| \geq 1$

Since there are no unit or λ-productions

Proof of the Pumping Lemma

- Start with a CNF grammar for $L - \{\epsilon\}$.
- Let the grammar have m variables.
- Pick $n = 2^m$.
- Let $|z| \geq n$.

  string
- We claim ("*Lemma 1* ") that a parse tree with yield z must have a path of length m+2 or more.

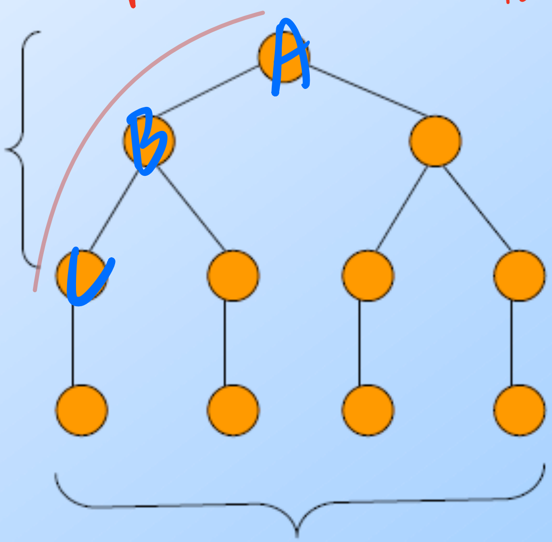
Proof of Lemma 1

- If all paths in the parse tree of a CNF grammar are of length $\leq$ m+1, then the longest yield has length 2^{m-1}, as in:

不會有variable重覆(由上到下)

m variables

m層



one terminal

2^{m-1} terminals

若terminal長度 > 2^{m-1}

則必有 variable重覆用

6

The Pumping Lemma I:

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

Non-context free languages

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

Context-free languages

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

Example 8.1: 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 ↳ 先假設為 context free，$n=0$ ($S \rightarrow \lambda$)

先不看

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:

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

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

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

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

We examine **all** the possible locations of string vxy in w ↳ proof 所有可能為 false

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

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

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

Case 1: vxy is within a^m

全在a裡面

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

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

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

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

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

Case 1: v and y consist from only a

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

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

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

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

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

Case 1: Repeating v and y

$$k \geq 1$$

多k個a

$$\underbrace{\overbrace{aaaaaa...aaaaaa}^{m+k}}\ \overbrace{bbb...bbb}^{m}\ \overbrace{ccc...ccc}^{m}$$

$$\underbrace{a}_{u}\ \underbrace{aaaaa...aaaaa}_{v^2xy^2}\ \underbrace{a\ bbb...bbb\ ccc...ccc}_{z}$$

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

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

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

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

$$k \geq 1$$

$$\overbrace{aaaaaa...aaaaaa}^{m+k}\;\overbrace{bbb...bbb}^{m}\;\overbrace{ccc...ccc}^{m}$$

$$\underbrace{aa}_{u}\underbrace{aaaa...aaaa}_{v^2xy^2}\underbrace{a\;bbb...bbb\;ccc...ccc}_{z}$$

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

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

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

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

$$k \geq 1$$

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

pump 2 次

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

↳ 在b裡面

$$\overbrace{aaa\ldots aaa}^{m}\;\overbrace{bbb\ldots bbb}^{m}\;\overbrace{ccc\ldots ccc}^{m}$$

$$\underbrace{aaa\ldots aaa\,bb}_{u}\,\underbrace{b\ldots bb}_{vxy}\,\underbrace{b\,ccc\ldots ccc}_{z}$$

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

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

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

Case 2: Similar analysis with case 1 ⇒ 和case1同理

$$\overbrace{aaa\ldots aaa}^{m}\ \overbrace{bbb\ldots bbb}^{m}\ \overbrace{ccc\ldots ccc}^{m}$$

$$\underbrace{aaa\ldots aaa\ bb}_{u}\ \underbrace{b\ldots bb}_{vxy}\ \underbrace{b\ ccc\ldots ccc}_{z}$$

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

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

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

Case 3: vxy is within c^m

↳ 全在c裡面

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

$$\underbrace{aaa...aaa\ bbb...bbb\ cc}_{u}\underbrace{c...cc}_{vxy}\underbrace{c}_{z}$$

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

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

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

Case 3: Similar analysis with case 1 ⇒ 同 case 1

$$\overbrace{aaa\ldots aaa}^{m}\;\overbrace{bbb\ldots bbb}^{m}\;\overbrace{ccc\ldots ccc}^{m}$$

$$\underbrace{aaa\ldots aaa\;bbb\ldots bbb\;c}_{u}\underbrace{cc\ldots c}_{vxy}\underbrace{cc}_{z}$$

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

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

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

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

↳ 在a,b裡面

$$\overbrace{aaa\ldots aaa}^{m}\ \overbrace{bbb\ldots bbb}^{m}\ \overbrace{ccc\ldots ccc}^{m}$$

$$\underbrace{aaa}_{u}\underbrace{\ldots aaa\ bbb}_{vxy}\underbrace{\ldots bbb\ ccc\ldots ccc}_{z}$$

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

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

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

Case 4: Possibility 1: v contains only a v全a

y contains only b y全b

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

$$\underbrace{aaa}_{u}\underbrace{\ldots aaa\,bbb}_{vxy}\underbrace{\ldots bbb\,ccc\ldots ccc}_{z}$$

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

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

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

Case 4: Possibility 1: v contains only a

y contains only b

$$k_1 + k_2 \geq 1$$

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

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

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

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

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

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

$$k_1 + k_2 \geq 1$$

$$\underbrace{aaa}_{u}\ \underbrace{\overbrace{\dots aaaaaaa}^{m+k_1}\ \overbrace{bbbbbbb\dots}^{m+k_2}}_{v^2xy^2}\ \underbrace{\dots bbb\ \overbrace{ccc\dots ccc}^{m}}_{z}$$

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

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

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

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

$$k_1 + k_2 \geq 1$$

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

↳ pump 2次

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 v有a,b

y contains only b y只有b

$$\overbrace{aaa\ldots aaa}^{m}\ \overbrace{bbb\ldots bbb}^{m}\ \overbrace{ccc\ldots ccc}^{m}$$

$$\underbrace{aaa}_{u}\underbrace{\ldots aaa\ bbb}_{vxy}\underbrace{\ldots bbb\ ccc\ldots ccc}_{z}$$

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

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

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

Case 4: Possibility 2: v contains a and b

y contains only b

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

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

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

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

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

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

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

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

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

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

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

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

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

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 v只有a

y contains a and b y有a,b

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

$$\underbrace{aaa}_{u}\underbrace{\ldots aaa\,bbb}_{vxy}\underbrace{\ldots bbb\,ccc\ldots ccc}_{z}$$

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

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

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

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

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

$$\underbrace{aaa\ldots aaa\,bbb\ldots}_{u}\underbrace{bbb\,ccc\ldots}_{vxy}\underbrace{ccc}_{z}$$

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

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

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

Case 5: Similar analysis with case 4

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

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

↳ 所有 case 都 false ⇒ L 非 context free

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

Non-context free languages

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

$$\{ww : w \in \{a,b\}\}$$

$$\{a^{n^2} b^n : n \geq 0\}$$

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

Context-free languages

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

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

The Pumping Lemma II:

For infinite linear language L

↳ 只能有一個 variable

there exists an integer m such that

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

we can write $w = uvxyz$

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

and it must be:

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

Example 8.6

- Show the following language

$L = \{w : n_a(w) = n_b(w)\}$ is not linear

$$S \to SS$$

$$S \to \lambda$$

$$S \to aSb$$

$$S \to bSa$$

O Given m

S Picks $w = a^m b^{2m} a^m$

O Picks any *uvyz* s.t. $uv=a^k$, $yz=a^l$ and $k, l \geq 1$

S Picks i = 2 → $w_2 = a^{m+k} b^{2m} a^{m+l}$ is not in L