

補充資料II (Context-Free Grammar)

資科系
林偉川

CFG Important Rules

Rule 1. Substitution and elimination

CFG $G=(V,T,S,P)$, where P has the forms

$$A \rightarrow x_1 B x_2$$

$$B \rightarrow y_1 \mid y_2 \mid y_3 \mid \dots \mid y_n$$

Can be changed to

$$A \rightarrow x_1 y_1 x_2 \mid x_1 y_2 x_2 \mid \dots \mid x_1 y_n x_2$$

範例

$G = (\{A, B\}, \{a, b, c\}, A, P)$, where P are given by

$A \rightarrow a \mid aaA \mid abBc$

$B \rightarrow abbA \mid b$

can be changed to

$A \rightarrow a \mid aaA \mid abbc \mid ababbAc$

Ex.

$A \Rightarrow aaA \Rightarrow aaabBc \Rightarrow aaabbc$ (in original)

$A \Rightarrow aaA \Rightarrow aaabbc$ (in modified)

3

CFG Important Rules

Rule 2. Eliminate the Left Recursion

CFG $G = (V, T, S, P)$, where P has the forms

$A \rightarrow Ax_1 \mid Ax_2 \mid Ax_3 \mid \dots \mid Ax_n$

$A \rightarrow y_1 \mid y_2 \mid y_3 \mid \dots \mid y_m$

Can be changed to

$A \rightarrow y_i \mid y_i Z, i=1,2,\dots,m$

$Z \rightarrow x_i \mid x_i Z, i=1,2,\dots,m$

Or

$A \rightarrow y_i Z, i=1,2,\dots,m$

$Z \rightarrow x_i Z \mid \varepsilon, i=1,2,\dots,m$

4

範例

$G = (\{A, B\}, \{a, b, c\}, A, P)$, where P are given by

$A \rightarrow Aa \mid aBc \mid \lambda$

$B \rightarrow Bb \mid ba$ (assume $x_1 = a, y_1 = aBc, y_2 = \lambda$)

can be changed to

$A \rightarrow aBc \mid aBcZ \mid Z \mid \lambda$

$Z \rightarrow a \mid aZ$

$B \rightarrow Bb \mid ba$ (in original)

Continue \rightarrow

5

範例

Finally can be further changed to

$A \rightarrow aBc \mid aBcZ \mid Z \mid \lambda$

$Z \rightarrow a \mid aZ$

$B \rightarrow ba \mid baY$

$Y \rightarrow b \mid bY$

6

CFG Important Rules

Rule 3. Eliminate the useless production(Rule)

Useless Non-terminal:

- a. Cannot reach from the start symbol
- b. Cannot derive a terminal string

Given a CFG $G=(V,T,S,P)$, $\forall V_i \in V$, V_i is not useless symbol $\Rightarrow G$ is a useful grammar

7

範例

$G=({S,A,B},\{a,b\}, S, P)$, where P are given by

$S \rightarrow A$

$A \rightarrow aA \mid \lambda$

$B \rightarrow bA$ (cannot reach from the start symbol)

can be changed to

$S \rightarrow A$

$A \rightarrow aA \mid \lambda$

8

範例

$G = (\{S, A, B, C\}, \{a, b\}, S, P)$, where P are given by

$S \rightarrow aS \mid A \mid C$

$A \rightarrow a$

$B \rightarrow aa$ (cannot reach from the start symbol)

$C \rightarrow aCb$ (cannot derive a terminal string)

can be changed to

$S \rightarrow aS \mid A$

$A \rightarrow a$

9

CFG Important Rules

Rule 4. Eliminate the λ production (Rule)

λ production:

$A \rightarrow \lambda$ or $A \Rightarrow^* \lambda$

Given a CFG $G = (V, T, S, P)$ may generate a language not containing λ , the λ production can be removed

10

Example

$G = (\{S, A\}, \{a, b\}, S, P)$, where P are given by

$S \rightarrow aAb$

$A \rightarrow aAb \mid \lambda$

can be changed to

$S \rightarrow aAb \mid ab$

$A \rightarrow aAb \mid ab$

Can generate the language $L = \{a^n b^n : n \geq 1\}$

11

範例

$G = (\{S, A, B, C, D\}, \{a, b, d\}, S, P)$, where P are given by

$S \rightarrow ABaC$

$A \rightarrow BC$

$B \rightarrow b \mid \lambda$

$C \rightarrow D \mid \lambda$

$D \rightarrow d$

can be changed to

$S \rightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Aa \mid Ba \mid a$

$A \rightarrow B \mid C \mid BC$

$B \rightarrow b$

$C \rightarrow d$

12

CFG Important Rules

Rule 5. Eliminate the Unit production(Rule)

Unit production:

$$A \rightarrow B, \text{ where } A, B \in V$$

Given a CFG $G=(V,T,S,P)$ may generate a language not containing unit production, the unit production can be removed

13

範例

$G=({S,A,B},{a,b}, S, P)$, where P are given by

$$S \rightarrow Aa \mid B$$

$$A \rightarrow a \mid bc \mid B \quad (A \rightarrow B)$$

$$B \rightarrow A \mid bb \quad (B \rightarrow A)$$

can be changed to

$$S \rightarrow Aa \mid B$$

$$A \rightarrow a \mid bc \mid bb \mid A \quad (\text{by rule 1})$$

$$B \rightarrow a \mid bc \mid bb \mid B \quad (\text{by rule 1} \rightarrow A \text{ \& B are the same})$$

Finally can be changed to

$$S \rightarrow Aa \mid A$$

$$A \rightarrow a \mid bc \mid bb \mid A$$

14

CFG Important Rules usage

The use sequence of the CFG reduction should be as followed:

Rule 1. Substitution and elimination

Rule 2. Eliminate the Left Recursion

Rule 4. Eliminate the λ production(Rule)

Rule 5. Eliminate the Unit production(Rule)

Rule 3. Eliminate the useless production(Rule)

15

範例

$E \rightarrow E + T$

$E \rightarrow T$

Can be changed to

$E \rightarrow T \mid TZ$

$Z \rightarrow +T \mid +TZ$

16

BNF v.s. EBNF

BNF:

$\langle \text{expr} \rangle \rightarrow \langle \text{expr} \rangle + \langle \text{term} \rangle \mid \langle \text{expr} \rangle - \langle \text{term} \rangle \mid$
 $\langle \text{term} \rangle$

$\langle \text{term} \rangle \rightarrow \langle \text{term} \rangle * \langle \text{factor} \rangle \mid \langle \text{term} \rangle / \langle \text{factor} \rangle \mid$
 $\langle \text{factor} \rangle$

EBNF:

$\langle \text{expr} \rangle \rightarrow \langle \text{term} \rangle \{ (+ \mid -) \langle \text{term} \rangle \}$

$\langle \text{term} \rangle \rightarrow \langle \text{factor} \rangle \{ (* \mid /) \langle \text{factor} \rangle \}$

17

範例

$\langle \text{term} \rangle \rightarrow \langle \text{term} \rangle (* \mid /) \langle \text{factor} \rangle \mid \langle \text{factor} \rangle$

By rule 2. Eliminate left-recursive

$\langle \text{term} \rangle \rightarrow \langle \text{factor} \rangle \mid \langle \text{factor} \rangle \langle Z \rangle$

$\langle Z \rangle \rightarrow (* \mid /) \langle \text{factor} \rangle \mid (* \mid /) \langle \text{factor} \rangle \langle Z \rangle$

By EBNF definition

$\langle Z \rangle \rightarrow (* \mid /) \langle \text{factor} \rangle \{ \langle Z \rangle \}$

$\langle Z \rangle \rightarrow (* \mid /) \langle \text{factor} \rangle \{ (* \mid /) \langle \text{factor} \rangle \}$

$\langle Z \rangle \rightarrow \{ (* \mid /) \langle \text{factor} \rangle \}$

By rule 5. $\langle Z \rangle$ is unit production

$\langle \text{term} \rangle \rightarrow \langle \text{factor} \rangle \{ \langle Z \rangle \}$

$\langle \text{term} \rangle \rightarrow \langle \text{factor} \rangle \{ (* \mid /) \langle \text{factor} \rangle \}$

18