

## 8 回目：文脈自由文法

- 文脈自由文法・言語 ( C F G / C F L )
- 導出木 ( Derivation Tree )
- 文脈自由文法の簡単化：無効記号
- 文脈自由文法の簡単化： $\epsilon$ -生成規則
- 文脈自由文法の簡単化：単位生成規則
- 文脈自由文法の標準形：チョムスキ標準形
- 文脈自由文法の標準形：グライバッハ標準形

<http://www.cit.ics.saitama-u.ac.jp/~far/Lectures/Automaton/08kaime>

## 文脈自由文法・言語 (CFG/CFL)

文脈自由文法 (Context-Free Grammar: CFG)

2型文法 (Type 2 Grammar)

$$G = \{N, \Sigma, P, S\}$$

$$A \in N; \quad a \in \Sigma;$$

$P$  の要素は以下の形式となっている :

$$P = \{ A \rightarrow \alpha, \quad | A \in N, \quad \alpha \in (\Sigma \cup N)^* \}$$

文脈自由言語 (Context-Free Language: CFL)

$$L(G) = \{w \in \Sigma^* \mid S \xRightarrow[G]{} w\}$$

$$\forall \alpha \in (\Sigma \cup N)^* \quad L(\alpha) = \{w \in \Sigma^* \mid \alpha \xRightarrow[G]{} w\}$$

## 導出木 ( Derivation Tree )

文脈自由文法による導出を見やすくする手法として導出木 ( Derivation Tree ) 図式が使われる。

- 構文木 (parse tree)
- 葉 (leaf)
- 節点 (node, vertex)
- 内部節点 (internal node)
- 枝 (edge)
- 親 (parent), 子 (offspring, son)
- 根 (root)
- 経路 (path), 経路の長さ (length)

## 文脈自由文法の簡単化：無効記号

- 定義

文脈自由文法  $G = \{N, \Sigma, P, S\}$  の非終端記号  $X$  に関して、もしいかなる  $w, x, y \in \Sigma^*$  に対しても

$$S \xRightarrow{G} wXy \xRightarrow{G} wxy$$

なる導出が存在しない場合、 $X$  は無効記号 (useless symbol) になる。

- 抽出方法

1. 生記号の抽出：

$$\exists x \in \Sigma^* \mid X \xRightarrow{G} x \quad (L(X) \neq \phi) \quad X \in N$$

2. 到達可能記号の抽出： $\exists \alpha, \beta \in (\Sigma \cup N)^* \mid S \xRightarrow{G} \alpha X \beta$

$$X \in (N \cup \Sigma)$$

注意：上記の順序での操作が必要です。

## 文脈自由文法の簡単化： $\epsilon$ -生成規則

$A \longrightarrow \epsilon$  形式の生成規則は  $\epsilon$ -生成規則 ( $\epsilon$ -rule) と呼ばれる.

- $\epsilon$ -なし文法

1.  $P$  は  $\epsilon$ -生成規則を含まない.  $\epsilon \notin L(G)$
2.  $P$  が含む  $\epsilon$ -生成規則は  $S \longrightarrow \epsilon$  である. このとき開始記号  $S$  はいかなる生成規則においてもその右辺に現れない.

- $\epsilon$ -なし文法の求め方

1. 空白化記号の検出
2.  $\epsilon$ -生成規則の除去

## 文脈自由文法の簡単化：単位生成規則

- 定義

$G = \{N, \Sigma, P, S\}$  の以下のルールは 単位生成規則 (unit production) と呼ばれる.

$$A \longrightarrow B \quad A, B \in N$$

- 単位生成規則除去方法

$A \longrightarrow N_A \{ \dots B \dots \} \longrightarrow \alpha$  が抽出されたとき  
( $B \longrightarrow \alpha$  単位生成規則ではない;  $\alpha \in P$ )

$$A \longrightarrow \alpha$$

## 文脈自由文法の標準形：チョムスキ標準形

- チョムスキ標準形 (Chomsky Normal Form: CNF) の定義

$$G = \{N, \Sigma, P, S\}$$

$$(1) \quad A \longrightarrow BC \quad (A, B, C \in N)$$

$$(2) \quad A \longrightarrow a \quad (a \in \Sigma)$$

$$(3) \quad S \longrightarrow \epsilon$$

- CNF への変換法

## 文脈自由文法の標準形：グライバッハ標準形

- グライバッハ標準形 (Greibach Normal Form: GNF) の定義

$$G = \{N, \Sigma, P, S\}$$

$$(1) \quad A \longrightarrow aB_1B_2 \dots B_m \quad (a \in \Sigma \quad B_1, B_2, \dots, B_m \in N)$$

$$(2) \quad A \longrightarrow a \quad (a \in \Sigma)$$

$$(3) \quad S \longrightarrow \epsilon$$

- GNF への変換法