文脈自由文法

離散数学・オートマトン

2020年後期

佐賀大学理工学部　只木進一

言語と文法

- 言語の構成要素
  - 語
  - 文
  - 文法
- 文法
  - 語の配置規則
  - 文の生成規則

形式文法
Formal Grammar

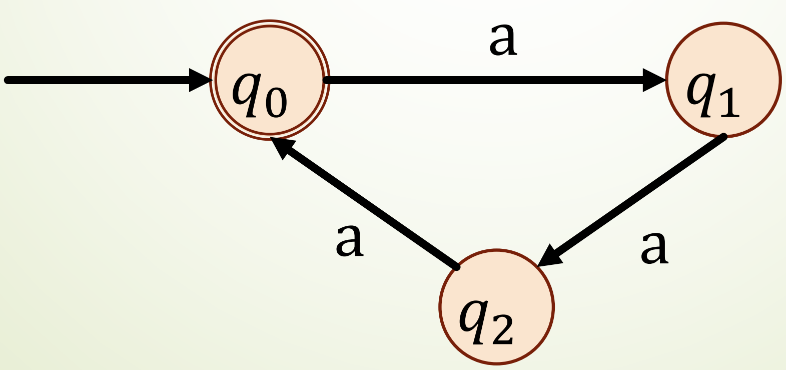
- $G = \langle N, \Sigma, P, S_0 \rangle$
  - N : 非終端アルファベット : 文法の要素に相当
  - Σ : アルファベット : 語に相当
  - P : 生成規則
  - $S_0 \in N$: 開始記号

正規文法
Regular Grammar

- 正規表現に対応した正規言語を生成
- 生成規則
  - $P: N \to \Sigma N | \Sigma$
- 例 : $G = \langle N, \Sigma, P, \mathrm{S}_0 \rangle$
  - $N = \{\mathrm{S}_0, \mathrm{S}_1, \mathrm{S}_2\}, \Sigma = \{\mathrm{a}\}$
  - $P = \{\mathrm{S}_0 \to \epsilon | \mathrm{aS}_1, \mathrm{S}_1 \to \mathrm{aS}_2, \mathrm{S}_2 \to \mathrm{aS}_0 | \mathrm{a}\}$

導出と対応するDFA

$$S_0 \Rightarrow aS_1 \Rightarrow aaS_2 \Rightarrow aaaS_0$$
$$\Rightarrow aaaaS_1 \Rightarrow aaaaaS_2 \Rightarrow aaaaaa$$



$$L = \{a^{3i} | i = N \cup \{0\}\} = (aaa)^*$$

文脈自由文法
Context Free Grammar

- 生成規則
  - $P: N \to (\Sigma \cup N)^*$
- 例: $G = \langle N, \Sigma, P, S_0 \rangle$
  - $N = \{S_0\}, \Sigma = \{\mathrm{a}, \mathrm{b}\}$
  - $P = \{S_0 \to \epsilon | \mathrm{a}S_0\mathrm{b} | \}$

$$\mathrm{S}_0 \Rightarrow \mathrm{aS}_0\mathrm{b} \Rightarrow \mathrm{aaS}_0\mathrm{bb} \Rightarrow \mathrm{aaaS}_0\mathrm{bbb} \Rightarrow \mathrm{aaabbb}$$

なぜ「文脈自由」なのか

- 生成規則の左辺は、非終端記号一つ
  - 非終端記号や終端記号との繋がり（文脈）を無視

標準形

- チョムスキー標準形 (Chomsky normal form, CNF)
  - 全生成規則が、
    - $\mathrm{A} \to \mathrm{BC}$または$\mathrm{A} \to \mathrm{a}$、$S \to \epsilon$
- グライバッハ標準形 (Greibach normal form, GNF)
  - 全生成規則が、
    - $\mathrm{A} \to \mathrm{a}\alpha, \mathrm{a} \in \Sigma, \alpha \in N^*, S \to \epsilon$も可

PDFの3種類の受理

- 入力終了時にスタックが空
  - $L_A(M) = \{w \in \Sigma^* | (q_0, w, \mathrm{Z}) \vdash^* (q, \epsilon, \epsilon)\}$
- 入力終了時に終状態
  - $L_N(M) = \{w \in \Sigma^* | (q_0, w, \mathrm{Z}) \vdash^* (q, \epsilon, \gamma), q \in F\}$
- 入力終了時にスタックが空、かつ終状態
  - $L(M) = \{w \in \Sigma^* | (q_0, w, \mathrm{Z}) \vdash^* (q, \epsilon, \epsilon), q \in F\}$

文脈自由言語Lを受理するNPDA

- $L \cap \{\epsilon\} = \emptyset$とし、$G = \langle N, \Sigma, P, S \rangle$
  - 生成規則はGNF
  - 最左導出(一番左の非終端記号から生成規則を適用)
- 等価なNPDA
  - $M = \langle \{q\}, \Sigma, N, \delta, q, S, \emptyset \rangle$
  - 入力終了時にスタックが空になる

- 最左導出

$$S \Rightarrow a_1 A_1 \gamma_1 \Rightarrow a_1 a_2 A_2 \gamma_2 \Rightarrow^* a_1 a_2 \cdots a_{n-2} A_{n-2} \Rightarrow a_1 a_2 \cdots a_{n-1}$$

- 対応する動作

$$\begin{aligned}(q, a_1 a_2 a_3 \cdots a_{n-1}, S) &\vdash (q, a_2 a_3 \cdots a_{n-1}, A_1 \gamma_1) \\ &\vdash (q, a_3 \cdots a_{n-1}, A_2 \gamma_2) \\ &\vdash^* (q, a_{n-1}, A_{n-2}) \\ &\vdash (q, \epsilon, \epsilon)\end{aligned}$$

- 遷移関数
  - 生成規則$\mathrm{A} \rightarrow \mathrm{a}\gamma$があるとき、かつその限り
    - $(q, \gamma) \in \delta(q, \mathrm{a}, \mathrm{A})$

例

- $G = \langle \{\mathrm{S}, \mathrm{A}, \mathrm{B}\}, \{\mathrm{a}, \mathrm{b}\}, P, \mathrm{S} \rangle$
  - $P = \{\mathrm{S} \to \mathrm{a}|\mathrm{b}|\mathrm{aSA}|\mathrm{bSB}, \mathrm{A} \to \mathrm{a}, \mathrm{B} \to \mathrm{b}\}$

$$
\begin{aligned}
\mathrm{S} &\Rightarrow \mathrm{aSA} \Rightarrow \mathrm{abSBA} \Rightarrow \mathrm{abaSABA} \\
&\Rightarrow \mathrm{abaaABA} \Rightarrow \mathrm{abaaaBA} \Rightarrow \mathrm{abaaabA} \Rightarrow \mathrm{abaaaba} \\
&\Rightarrow \mathrm{abaaaba}
\end{aligned}
$$

- $M = \langle \{q\}, \{\mathrm{a}, \mathrm{b}\}, \{\mathrm{S}, \mathrm{T}, \mathrm{A}, \mathrm{B}\}, \delta, S, \emptyset \rangle$

$$\delta(q, \mathrm{a}, \mathrm{S}) = \{(q, \epsilon), (q, \mathrm{SA})\},$$

$$\delta(q, \mathrm{b}, \mathrm{S}) = \{(q, \epsilon), (q, \mathrm{SB})\},$$

$$\delta(q, \mathrm{a}, \mathrm{A}) = \{(q, \epsilon)\},$$

$$\delta(q, \mathrm{b}, \mathrm{B}) = \{(q, \epsilon)\}$$

$$
\begin{aligned}
(q, \mathrm{abaaaba}, \mathrm{S}) &\vdash (q, \mathrm{baaaba}, \mathrm{SA}) \\
&\vdash (q, \mathrm{aaaba}, \mathrm{SBA}) \\
&\vdash (q, \mathrm{aaba}, \mathrm{SABA}) \\
&\vdash (q, \mathrm{aba}, \mathrm{ABA}) \\
&\vdash (q, \mathrm{ba}, \mathrm{BA}) \vdash (q, \mathrm{a}, \mathrm{A}) \\
&\vdash (q, \epsilon, \epsilon)
\end{aligned}
$$

空スタックで受理するNPDAに対応する文脈自由文法

- $M = \langle Q, \Sigma, \Gamma, \delta, q_0, \mathrm{Z}, \emptyset \rangle$
- $G = \langle N, \Sigma, P, S \rangle$
  - $q, q' \in Q, \mathrm{A} \in \Gamma$に対して$[q\mathrm{A}q'] \in N$

空スタックで受理するNPDAに対応する文脈自由文法

- $\forall q \in Q$に対して$\mathrm{S} \rightarrow [q_0 \mathrm{Z} q]$を作る
- $(q_1, \mathrm{B}_1 \cdots \mathrm{B}_k) \in \delta(q, \mathrm{a}, \mathrm{A})$に対して
  - $\forall q_2, \cdots, q_{k+1}$に対して
    - $[q\mathrm{A}q_{k+1}] \rightarrow \mathrm{a}[q_1 \mathrm{B}_1 q_2][q_2 \mathrm{B}_2 q_3] \cdots [q_k \mathrm{B}_k q_{k+1}]$を作る
  - ただし、$(q_1, \epsilon) \in \delta(q, \mathrm{a}, \mathrm{A})$に対しては、$[q\mathrm{A}q_1] \rightarrow \mathrm{a}$

例

- $M = \langle \{q_0, q_1\}, \Sigma, \Gamma, \delta, q_0, \mathrm{Z}, \emptyset \rangle$

$$\delta(q_0, \mathrm{a}, \mathrm{Z}) = \{(q_0, \mathrm{AZ})\}, \delta(q_0, \mathrm{a}, \mathrm{A}) = \{(q_0, \mathrm{AA})\},$$
$$\delta(q_0, \mathrm{b}, \mathrm{A}) = \{(q_1, \epsilon)\},$$
$$\delta(q_1, \mathrm{b}, \mathrm{A}) = \{(q_1, \epsilon)\}, \delta(q_1, \epsilon, \mathrm{Z}) = \{(q_1, \epsilon)\}$$

$$
\begin{aligned}
(q_0, \text{aaabbb}, \text{Z}) &\vdash (q_0, \text{aabbb}, \text{AZ}) \\
&\vdash (q_0, \text{abbb}, \text{AAZ}) \\
&\vdash (q_0, \text{bbb}, \text{AAAZ}) \\
&\vdash (q_1, \text{bb}, \text{AAZ}) \\
&\vdash (q_1, \text{b}, \text{AZ}) \\
&\vdash (q_1, \epsilon, \text{Z}) \vdash (q_1, \epsilon, \epsilon)
\end{aligned}
$$

- $G = \langle N, \Sigma, P, S \rangle$
- 開始記号
  - $\mathrm{S} \to [q_0 \mathrm{Z} q_0] | [q_0 \mathrm{Z} q_1]$
- $\delta(q_0, \mathrm{a}, \mathrm{Z}) = \{(q_0, \mathrm{AZ})\}$より

$$[q_0 \mathrm{Z} q_0] \to \mathrm{a}[q_0 \mathrm{A} q_0][q_0 \mathrm{Z} q_0] \mid \mathrm{a}[q_0 \mathrm{A} q_1][q_1 \mathrm{Z} q_0]$$
$$[q_0 \mathrm{Z} q_1] \to \mathrm{a}[q_0 \mathrm{A} q_0][q_0 \mathrm{Z} q_1] \mid \mathrm{a}[q_0 \mathrm{A} q_1][q_1 \mathrm{Z} q_1]$$

- $\delta(q_0, \mathrm{a}, \mathrm{A}) = \{(q_0, \mathrm{AA})\}$より

$$[q_0 \mathrm{A} q_0] \to \mathrm{a}[q_0 \mathrm{A} q_0][q_0 \mathrm{A} q_0] \mid \mathrm{a}[q_0 \mathrm{A} q_1][q_1 \mathrm{A} q_0]$$

$$[q_0 \mathrm{A} q_1] \to \mathrm{a}[q_0 \mathrm{A} q_0][q_0 \mathrm{A} q_1] \mid \mathrm{a}[q_0 \mathrm{A} q_1][q_1 \mathrm{A} q_1]$$

- $\delta(q_0, \mathrm{b}, \mathrm{A}) = \{(q_1, \epsilon)\}$より
  - $[q_0 \mathrm{A} q_1] \rightarrow \mathrm{b}$
- $\delta(q_1, \mathrm{b}, \mathrm{A}) = \{(q_1, \epsilon)\}$より
  - $[q_1 \mathrm{A} q_1] \rightarrow \mathrm{b}$
- $\delta(q_1, \epsilon, \mathrm{A}) = \{(q_1, \epsilon)\}$より
  - $[q_1 \mathrm{Z} q_1] \rightarrow \epsilon$

- まとめると

$$S \to [q_0 Z q_0] \mid [q_0 Z q_1]$$
$$[q_0 Z q_0] \to \mathrm{a}[q_0 A q_0][q_0 Z q_0] \mid \mathrm{a}[q_0 A q_1][q_1 Z q_0]$$
$$[q_0 Z q_1] \to \mathrm{a}[q_0 A q_0][q_0 Z q_1] \mid \mathrm{a}[q_0 A q_1][q_1 Z q_1]$$
$$[q_1 Z q_1] \to \epsilon$$
$$[q_0 A q_0] \to \mathrm{a}[q_0 A q_0][q_0 A q_0] \mid \mathrm{a}[q_0 A q_1][q_1 A q_0]$$
$$[q_0 A q_1] \to \mathrm{a}[q_0 A q_0][q_0 A q_1] \mid \mathrm{a}[q_0 A q_1][q_1 A q_1] \mid \mathrm{b}$$
$$[q_1 A q_1] \to \mathrm{b}$$

- 終端記号を導かないものを削除

$$S \to [q_0 Z q_0] \mid [q_0 Z q_1]$$

$$[q_0 Z q_0] \to a[q_0 A q_0][q_0 Z q_0] \mid a[q_0 A q_1][q_1 Z q_0]$$

$$[q_0 Z q_1] \to a[q_0 A q_0][q_0 Z q_1] \mid a[q_0 A q_1][q_1 Z q_1]$$

$$[q_1 Z q_1] \to \epsilon$$

$$[q_0 A q_0] \to a[q_0 A q_0][q_0 A q_0] \mid a[q_0 A q_1][q_1 A q_0]$$

$$[q_0 A q_1] \to a[q_0 A q_0][q_0 A q_1] \mid a[q_0 A q_1][q_1 A q_1] \mid b$$

$$[q_1 A q_1] \to b$$

- 生成規則

$$S \to [q_0 Z q_1]$$

$$[q_0 Z q_1] \to a[q_0 A q_1][q_1 Z q_1]$$

$$[q_1 Z q_1] \to \epsilon$$

$$[q_0 A q_1] \to a[q_0 A q_1][q_1 A q_1] \mid b$$

$$[q_1 A q_1] \to b$$

$$S \Rightarrow [q_0 Z q_1] \Rightarrow a[q_0 A q_1][q_1 Z q_1]$$
$$\Rightarrow aa[q_0 A q_1][q_1 A q_1][q_1 Z q_1]$$
$$\Rightarrow aaa[q_0 A q_1][q_1 A q_1][q_1 A q_1][q_1 Z q_1]$$
$$\Rightarrow aaab[q_1 A q_1][q_1 A q_1][q_1 Z q_1]$$
$$\Rightarrow aaabb[q_1 A q_1][q_1 Z q_1]$$
$$\Rightarrow aaabbb[q_1 Z q_1]$$
$$\Rightarrow aaabbb$$