I113 オートマトンと形式言語
レポート3の解説
(I113 Automaton & Formal Languages Answer & Comments for Report 3)

上原 隆平(Ryuhei UEHARA)
uehara@jaist.ac.jp

レポート (3)

[問題1] 正則表現 $(1+\varepsilon)(01)^*(0+\varepsilon)$ で表現される言語を文脈自由文法で表現せよ。(ちなみに正則言語のクラスは文脈自由言語に真に含まれる)

[アイデア]

$$(1+\varepsilon)(01)^*(0+\varepsilon)$$

3つの部分に分けて生成する

Report (3)

[Problem 1] Define the language represented by a regular expression $(1+\varepsilon)(01)^*(0+\varepsilon)$ by context free grammar. (Besides, the class of regular languages is properly contained in the class of context free languages)

[Idea] $\underline{(1+\varepsilon)}\,\underline{(01)^*}\,\underline{(0+\varepsilon)}$

We partition into 3 parts.

[問題1] 正則表現（$(1+\varepsilon)(01)^*(0+\varepsilon)$）で表現される言語を文脈自由文法で表現せよ。(ちなみに正則言語のクラスは文脈自由言語に真に含まれる)

[解答例]

文法を $G=(V,T,P,S)$ とする。ここで

$V = \{ S, A, B, C \}$

$T = \{ 0, 1 \}$

$$P \begin{cases} S \rightarrow ABC \\ A \rightarrow 1 \mid \varepsilon \\ B \rightarrow 01\,B \mid \varepsilon \\ C \rightarrow 0 \mid \varepsilon \end{cases}$$

である。

$$\underbrace{(1+\varepsilon)}_{A}\underbrace{(01)^*}_{B}\underbrace{(0+\varepsilon)}_{C}$$

[Problem 1] Define the language represented by a regular expression $(1+\varepsilon)(01)^*(0+\varepsilon)$ by context free grammar. (Besides, the class of regular languages is properly contained in the class of context free languages)

[Solution]

It is represented by the grammar $G=(V,T,P,S)$, where

$V=\{S, A, B, C\}$, $T=\{0, 1\}$,

$$P \begin{cases} S \rightarrow ABC \\ A \rightarrow 1 \mid \varepsilon \\ B \rightarrow 01B \mid \varepsilon \\ C \rightarrow 0 \mid \varepsilon \end{cases}$$

$$\underbrace{(1+\varepsilon)}_{A}\underbrace{(01)^*}_{B}\underbrace{(0+\varepsilon)}_{C}$$

[問題2] $\Sigma=\{0,1\}$ 上の言語 $L=\{0^n1^n \mid n>0\}$ を文脈自由文法で表現せよ。($\varepsilon \notin L$)

[解答例] $G=(\{S\}, \{0, 1\}, S\rightarrow 01 \mid 0S1, S)$

[Problem 2] Define the language L={$0^n 1^n$ | n>0} over Σ={0,1} by context free grammar. ($\varepsilon \notin L$)

[Solution]

$$G = (\{S\}, \{0, 1\}, S \rightarrow 01 \mid 0S1, S)$$

[問題3] 問題2で構成した文脈自由文法によって、言語Lが正しく表現されていることを証明せよ。

[解答の骨子] ここでは

- 文法Gで定義される言語$L(G)$と言語$L(=\{0^n1^n\})$が一致することを示す。これは
- $L(G) \subseteq L$ かつ $L \subseteq L(G)$

を示すことで証明する。

[典型的な誤解答]:

「Gの生成する語が0^n1^nの形をしていることを示す」

⇒ これは $L(G) \subseteq L = \{0^n1^n\}$ しか示していない

⇒ もしかしたら特定の 0^m1^m は生成しないかもしれない

[Problem 3] Prove that the CFG defined in Problem 2 certainly represents the language L.

[Outline of the solution] Here, we show

- $L(G)$, defined by the grammar G, coincides with $L(=\{0^n1^n\})$.

This can be proved by showing

- $L(G) \subseteq L$ and $L \subseteq L(G)$.

[Typical wrong (or insufficient) answer]:

'All words generated by G form 0^n1^n for some n'

⇨ That only shows $L(G) \subseteq L = \{0^n1^n\}$.

⇨ Possibly, 0^m1^m for some m may not be generated.

[問題3] 問題2で構成した文脈自由文法によって、言語Lが正しく表現されていることを証明せよ。

[解答例] 文法G = ({S}, {0, 1}, S→01 | 0S1, S)で定義される言語L(G)と言語L={0^n1^n | n>0}が一致することを示す。これは (1)$L(G) \subseteq L$ かつ(2)$L \subseteq L(G)$ が成立することを示せばよい。

(1) $L(G) \subseteq L$を示す。Gの2つの規則より、

① 0と1の個数はいつでも同じ

② 0の前に1が生成されることは決してない

したがって、Gによって生成される語は必ず 0^n1^n という形式の語になり、L の元である。

[Problem 3] Prove that the CFG defined in Problem 2 certainly represents the language L.

[Solution] The language $L(G)$ defined by $G = (\{S\}, \{0,1\}, S \to 01 \mid 0S1, S)$ coincides with $L=\{0^n 1^n \mid n>0\}$. This is done by showing (1) $L(G) \subseteq L$ and (2) $L \subseteq L(G)$.

(1) We show $L(G) \subseteq L$. By two rules of G,

① The number of 0s is equal to the number of 1s.

② 1 is never generated before 0.

Hence, all words generated by G form $0^n 1^n$ for some n, and hence in L.

[解答例] 文法G = ({S}, {0, 1}, S→01｜0S1, S)で定義される言語$L(G)$と言語L={0^n1^n ｜ n>0}が一致することを示す。これは (1)$L(G) \subseteq L$ かつ(2)$L \subseteq L(G)$ が成立することを示せばよい。

(2)　$L \subseteq L(G)$ を示す。n に関する帰納法で示す。

①　n=1 のとき: 規則 S→01より、語01は生成できる。

②　n=k (k>1) のとき: 語$0^{k-1}1^{k-1}$はGで生成できると仮定する。

このとき、まず規則 S→ 0$\underline{S}$1 を適用する。次に右辺の$\underline{S}$から語$0^{k-1}1^{k-1}$を導出する。これは帰納法の仮定から、Gで生成可能である。したがって、

$$S \underset{G}{\Rightarrow} 0S1 \overset{*}{\underset{G}{\Rightarrow}} 00^{k-1}1^{k-1}1=0^k1^k$$

となる。よって語0^k1^kはGで生成できる。

[Solution] The language $L(G)$ defined by $G=(\{S\}, \{0,1\}, S\rightarrow 01 \mid 0S1, S)$ coincides with $L=\{0^n1^n \mid n>0\}$. This is done by showing (1) $L(G)\subseteq L$ and (2) $L\subseteq L(G)$.

(2) We show $L\subseteq L(G)$ by induction for n.

① $n=1$: By the rule $S\rightarrow 01$, the word 01 can be generated.

② $n=k$ ($k>1$): Suppose that the word $0^{k-1}1^{k-1}$ can be generated by G.

Then, we first apply the rule $S\rightarrow 0\underline{S}1$. Next, we derive the word $0^{k-1}1^{k-1}$ from $\underline{S}$ on the right side, which can be done by inductive hypothesis. Hence, we have

$$S \underset{G}{\Rightarrow} 0S1 \overset{*}{\underset{G}{\Rightarrow}} 00^{k-1}1^{k-1}1=0^k1^k$$

Thus the word 0^k1^k can be generated by G.