

準 LL(2) 文法の構文解析表の性質の応用について

3H-10

松浦 千賀子^{*} 吉田 敬一^{*}
静岡大学大学院理工学研究科^{*}

1. はじめに

LL(2) 文法にわずかな制限を加えた準 LL(2) 文法とその構文解析表作成のアルゴリズムが提案されているが¹⁾、解析表のもつ性質については十分に解明されていない。

本研究では、準 LL(2) 文法に対する解析表がもつ性質の1つを利用して、「準 LL(2) 文法でない LL(2) 文法はすべて準 LL(2) 文法に書き換えられる」ことを示した。

2. 基本的定義と記法

[定義1] 文脈自由文法 G を

$$G = (N, \Sigma, P, S)$$

とする。ここに、 N, Σ はそれぞれ文法 G の非終端記号の集合ならびに終端記号の集合、 P は生成規則の集合であり、 S は出発記号である。

[記法1] N の要素を A, B, C, \dots , Σ の要素を a, b, c, \dots , $N \cup \Sigma$ の要素を X, Y, Z で表す。また、 Σ^* の要素を s, t, u, \dots で表し、 $(N \cup \Sigma)^*$ の要素を $\alpha, \beta, \gamma, \dots$ で表す。とくに ϵ, ϕ はそれぞれ空列、空集合を表す。 ϵ, ϕ 以外のこれらの記号は添字をつけて用いることもある。

[定義2] 集合 $FIRST_k(\alpha)$ は以下で定義される。

$$FIRST_k(\alpha) = \{ u \mid (\alpha \Rightarrow u\beta, |u| = k) \text{ または } (\alpha \Rightarrow u, |u| < k) \}$$

ただし、 $\alpha, \beta \in (N \cup \Sigma)^*$, $u \in \Sigma^*$, $|u|$ は u の長さを表す。また、 $\alpha \Rightarrow u\beta$ は 0 回以上使用する最左導出を表す。

[定義3] L_1, L_2 を Σ^* の部分集合とすると、演算子 \oplus_k は以下で定義される。

$$L_1 \oplus_k L_2 = \{ w \mid \text{ある } x \in L_1, y \in L_2 \text{ に対して、} \\ |xy| \leq k \text{ ならば } w = xy, |xy| > k \text{ ならば } w = u, \text{ただし } xy = uv, |u| = k \}$$

[定義4] 文脈自由文法 $G = (N, \Sigma, P, S)$ において $S \Rightarrow u, AX\xi$ のとき、 PF (partial-FOLLOW) は以下で定義される。

$$PF_k(A, X) = \cup_i FIRST_k(X\xi_i)$$

[定義5] 文脈自由文法 $G = (N, \Sigma, P, S)$ において、 $A \rightarrow \alpha, A \rightarrow \beta$ を相異なる生成規則とすると、 $S \Rightarrow uA\nu$ に対して

$$(FIRST_k(\alpha) \oplus_k FIRST_k(\nu)) \cap (FIRST_k(\beta) \oplus_k FIRST_k(\nu)) = \phi$$

が成り立つとき、 G は LL(k) 文法であるという。

[定義6] 文脈自由文法 $G = (N, \Sigma, P, S)$ において、 $A \rightarrow \alpha, A \rightarrow \beta$ を相異なる生成規則とすると、 $S \Rightarrow uAX\xi$ に対して

$$(FIRST_k(\alpha) \oplus_k PF_k(A, X)) \cap (FIRST_k(\beta) \oplus_k PF_k(A, X)) = \phi$$

が成り立つとき、 G は準 LL(k) 文法であるという。

[記法2] 生成規則ならびに導出

$$A \rightarrow \alpha, \alpha \Rightarrow \beta$$

における p, q は、それぞれ生成規則 $A \rightarrow \alpha$ につけられた固有の番号、ならびに最左導出 $\alpha \Rightarrow \beta$ で用いられた生成規則の番号を示す。

[定義7] 生成規則の番号を p で表すとき [] p または $[X] p$ を τ 型生成規則番号という。ここで X は、最左導出

$$S \Rightarrow uA\alpha \Rightarrow u\beta\alpha$$

における α の最左端一記号を表す。

[定理1]²⁾ LL(2) 文法 $G = (N, \Sigma, P, S)$ において、 $A \rightarrow \alpha, A \rightarrow \beta (\alpha \neq \beta)$ とするとき

$$S \Rightarrow uAX\xi$$

に対して、 A が書き換えを必要とする可能性があるのは $\alpha \Rightarrow a, \beta \Rightarrow \epsilon, X \Rightarrow \epsilon$ のときに限られる。

An Application of the Property of Parsing-tables for Semi-LL(2) Grammars.

^{*} Chikako Matsuura

^{*} Keiichi Yoshida

^{*} Graduate School of Science and Engineering, Shizuoka University.

3. 解析表の性質とその応用

[性質] 準 LL(2)文法でない LL(2)文法の解析表には以下の書き込み現象が生じる。

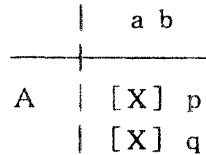


図1 導出 $AX \xi$ が $ab \xi$ を示す解析表の一部

図1はAからa bを導出するのに文脈AXのAに対して生成規則p, qのいずれもが適用可能なことを表している。準 LL(2)文法であれば、このような現象は起こりえない。

[定理2] 準 LL(2)文法でない LL(2)文法はすべて準 LL(2)文法に書き換えることができる。

[証明] LL(2)文法 $G = (N, \Sigma, P, S)$ において、 $A \rightarrow \alpha$, $A \rightarrow \beta$ を相異なる生成規則とするとき、 $S \Rightarrow u A X \xi$ に対して、準 LL(2)文法でない LL(2)文法は

- (1) $(FIRST_2(\alpha) \oplus_2 FIRST_2(X \xi)) \cap (FIRST_2(\beta) \oplus_2 FIRST_2(X \xi)) = \phi$
- (2) $(FIRST_2(\alpha) \oplus_2 PF_2(A, X)) \cap (FIRST_2(\beta) \oplus_2 PF_2(A, X)) \neq \phi$

を満足する。[定理1]より対象とするのは $\alpha \Rightarrow a$, $\beta \Rightarrow \epsilon$, $X \Rightarrow \epsilon$ の場合に限られるので(1)(2)より

- (i) $S \Rightarrow u_1 A X \xi_1 \Rightarrow u_1 \alpha X \xi_1 \Rightarrow u_1 a X \xi_1 \Rightarrow u_1 a \xi_1 \Rightarrow u_1 a b \xi_1'$
- (ii) $S \Rightarrow u_2 A X \xi_2 \Rightarrow u_2 \beta X \xi_2 \Rightarrow u_2 X \xi_2 \Rightarrow u_2 \xi_2 \Rightarrow u_2 a b \xi_2'$

なる導出が存在する。このとき

$b c \in FIRST_2(X \xi_1)$, $a b \in FIRST_2(X \xi_2)$, $\{b c, a b\} \subset PF_2(A, X)$ となる $b, c \in \Sigma$ が存在する。従って、準 LL(2)文法となるためにはこのような事象が起こらないよう、文法を書き換えればよい。つまり、図1の現象を図2のように変更してやればよい。これを実現するためには、[X] p, [X] q に対するAをそれぞれ別個の非終端記号Aに書きかえればよい。

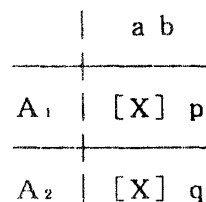


図2 書き換え後の解析表

この表に対する導出は

- (i) $S \Rightarrow u_1 A_1 X \xi_1 \Rightarrow u_1 \alpha X \xi_1 \Rightarrow u_1 a X \xi_1 \Rightarrow u_1 a \xi_1 \Rightarrow u_1 a b \xi_1'$
- (ii) $S \Rightarrow u_2 A_2 X \xi_2 \Rightarrow u_2 \beta X \xi_2 \Rightarrow u_2 X \xi_2 \Rightarrow u_2 \xi_2 \Rightarrow u_2 a b \xi_2'$

となる。このとき導出(i)(ii)より

$$b c \in FIRST_2(X \xi_1), a b \in FIRST_2(X \xi_2)$$

$$b c \in PF_2(A_1, X), a b \in PF_2(A_2, X)$$

となるから、これは準 LL(2)文法の定義を満足する。

この書き換え操作は図1に対する現象を引き起こす任意の非終端記号に対して適用可能なことは自明である。ゆえに[定理1][定理2]より、すべての準 LL(2)文法でない LL(2)文法は、準 LL(2)文法に書き換えることが可能であると言える。(証明終わり)

4. おわりに

本研究で提案した[定理]により、準 LL(2)文法でない LL(2)文法はすべて準 LL(2)文法に書き換えられることを示した。これは準 LL(2)文法に対する解析表作成のアルゴリズムを用いて、LL(2)文法に対する解析表を実質的に作成できることを意味する。今後は解析表の性質についてさらに解明していきたい。

参考文献

- 1) Aho, A.V. and Ullman, J.D.: Theory of Parsing, Translation and Compiling, Vol.1 pp334-361, Prentice-Hall (1972)
- 2) 吉田・竹内: 準 LL(2)文法に対する構文解析表の作成アルゴリズム、「情報処理学会論文誌」第31巻、第6月号(1990)
- 3) 吉田・竹内: 準 LL(2)文法に対する構文解析表の構造と解析アルゴリズム、「情報処理学会論文誌」第31巻、第9月号 (1990)
- 4) 井上謙蔵: コンパイラ「プログラム言語処理の基礎」pp77-117 丸善(1994)
- 5) 川島・吉田: テーブルを利用した LL(2)文法から準 LL(2)文法への書き換えアルゴリズム「情報処理学会第55回全国大会講演論文集」(1997)