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[2]、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 $A, B, C, \cdots$, Σ の要素を $a, b, c, \cdots$, $N\cup\Sigma$ の要素を X, Y, Z で表す。また、Σ^* の要素を $s, t, u, \cdots$ で表し、$(N\cup\Sigma)^*$ の要素を $\alpha, \beta, \gamma, \cdots$ 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。

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

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

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.

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

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

［定義４］文脈自由文法 $G=(N,\ \Sigma,\ P,\ S)$ において $S\overset{*}{\Rightarrow}u_i A X\xi_i$ 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き、$S\overset{*}{\Rightarrow}uA\nu$ 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き、$S\overset{*}{\Rightarrow}uAX\xi$ 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

$$A\underset{p}{\rightarrow}\alpha,\quad \alpha\underset{q}{\Rightarrow}\beta$$

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

$$S\overset{*}{\Rightarrow}uA\alpha\underset{p}{\Rightarrow}u\beta\alpha$$

における α の最左端一記号を表す。

［定理１］[3] 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

$$S\overset{*}{\Rightarrow}uAX\xi$$

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

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。

	a b
A	［X］ p
	［X］ q

図1　導出 $AX\xi \overset{*}{\Rightarrow} ab\xi$ を示す解析表の一部

図１はＡからａｂ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、$A \to \alpha$，$A \to \beta$ を相異なる生成規則とするとき、$S \overset{*}{\Rightarrow} uAX\xi$ 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
$\alpha \overset{*}{\Rightarrow} a\ \ \beta \overset{*}{\Rightarrow} \varepsilon, X \overset{*}{\Rightarrow} \varepsilon$ の場合に限られるので
(１)(２)より

(ｉ) $S \overset{*}{\Rightarrow} u_1 AX\xi_1 \Rightarrow u_1 \alpha X\xi_1 \overset{*}{\Rightarrow} u_1 aX\xi_1 \overset{*}{\Rightarrow} u_1 a\xi_1 \overset{*}{\Rightarrow} u_1 ab\xi_1'$

(ii) $S \overset{*}{\Rightarrow} u_2 AX\xi_2 \Rightarrow u_2 \beta X\xi_2 \overset{*}{\Rightarrow} u_2 X\xi_2 \overset{*}{\Rightarrow} u_2 \xi_2 \overset{*}{\Rightarrow} u_2 ab\xi_2'$

なる導出が存在する。このとき
$bc \in FIRST_2(X\xi_1), ab \in FIRST_2(X\xi_2)$,
$\{bc, ab\} \subset PF_2(A,X)$ 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。

	a b
A_1	［X］ p
A_2	［X］ q

図２　書き換え後の解析表

この表に対する導出は

(ｉ)' $S \overset{*}{\Rightarrow} u_1 A_1 X\xi_1 \Rightarrow u_1 \alpha X\xi_1 \overset{*}{\Rightarrow} u_1 aX\xi_1 \overset{*}{\Rightarrow} u_1 a\xi_1 \overset{*}{\Rightarrow} u_1 ab\xi_1'$

(ii)' $S \overset{*}{\Rightarrow} u_2 A_2 X\xi_2 \Rightarrow u_2 \beta X\xi_2 \overset{*}{\Rightarrow} u_2 X\xi_2 \overset{*}{\Rightarrow} u_2 \xi_2 \overset{*}{\Rightarrow} u_2 ab\xi_2'$

となる。このとき導出(ｉ)'(ii)'より

$bc \in FIRST_2(X\xi_1), ab \in FIRST_2(X\xi_2)$
$bc \in PF_2(A_1, X), ab \in PF_2(A_2, X)$

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のは自明である。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)