

準 LL(2) 文法に対する構文解析表の作成アルゴリズム†

吉田 敬一†† 竹内 淑子†††

LL(k) 文法の解析表作成に関する研究は、そのほとんどが $k=1$ の場合に関するものである。 $k \geq 2$ に関する研究は全くといってよいほど行われていない。わずかに、Aho らの研究があるにすぎない。その理由は、 $k=1$ の場合は必ず強 LL(1) 文法であるが、 $k \geq 2$ の場合には、強でないものと強の種類文法が存在するためである。強 LL(k) 文法に対しては LL(1) 文法とほぼ同じ方法で作成の実現が可能であるが、強でないものに対しては文脈問題がからむので、全く異なる方法をとらなければならない。本論文では、LL(k) 文法にわずかに制限を加えた準 LL(k) 文法の、 $k=2$ の場合の表作成のアルゴリズムとその正当性を示すものである。準 LL(k) 文法にも強のもの強でないものが存在するが、その表作成のアルゴリズムは比較的平易である。具体例にもとづき、そのアルゴリズムの性能についても実験した結果、表作成の時間は Aho らの方法の 1/10 程度で、表 (解析表と生成規則表) の大きさは 1/120~1/400 程度ですむことがわかった。また、文法に依存しないコード部分は 7% 程度筆者らのほうが大きくなったが、プログラム全体に占める割合は PASCAL- の場合で約 29%、ISO PASCAL の場合で 10% 程度であり、文法が大きくなるほどこの比率は小さくなるという一般的傾向があるので、あまり問題にはならない。

1. はじめに

筆者らは文献 1) で、準 LL(k) 文法とその解析法を提案した。準 LL(k) 文法の表現能力は LL(k) 文法と強 LL(k) 文法の間位置するものである¹⁾。

いままでにいくつかの LL(1) 文法の解析表作成法とその解析法が研究され、その成果が報告されている^{2)~6)}。しかし、 $k \geq 2$ に対する研究はほとんどみられず、表駆動型のものでは文献 5) だけである。それは、解析表の作成が LL(1) のそれにくらべて数段の困難を伴うからである。その理由は、LL(1) 文法はすべて強 LL(1) であるが、 k が 2 以上になると強と強でないものが存在し、後者の表構造は文脈に依存するからである。

本論文が提案する解析表作成法は、基本的には文献 4), 6) で提案された方法に文脈の概念を加えたものである。この実現のために、 π 型生成規則番号というものを導入した。文献 2) の Aho らの方法と筆者らの方法を性能的にくらべてみると具体例に関して筆者らの方法は表作成の所要時間で約 1/10、表の部分 (解析表と生成規則表) の記憶領域は約 1/120~1/400 程度ですむことがわかった。さらに、大部分が表操作のみで作成が可能であるので、作りやすいという特長を

もつ。

本論文ではこうした特長をもつ準 LL(2) 文法に対する解析表の作成アルゴリズムを提案し、アルゴリズムの正当性、性能評価を示す。

2. Aho-Ullman の方法の概要²⁾

この方法は、LL(2) 文法を Aho-Ullman 固有の強 LL(2) 文法にかえる作業と、それをもとにして解析表を作る作業から成る。そのためには非終端記号 A に右接しうる n 個 ($n \geq 1$) の異なる文脈があるとき、A に対して新しい非終端記号 T_1, T_2, \dots, T_n を生成する必要がある。この作業は、同じ非終端記号が右辺に 2 度以上現れてもよいことを考慮すると、費やす時間はかなりの量となる。

こうして作り出される T_i の個数は、 $k=2$ の場合に対して、PASCAL- 文法⁷⁾ について実験したところ、もともと 54 しかなかった非終端記号から作られた T_i の個数は 400 を超えた。さらに、それによって生成規則の書きかえが行われるが、もとの PASCAL- 文法の生成規則が 98 本であったものが、書きかえ後は 800 本を超えた。

こうして、実質的に文法を強 LL(2) に書きかえたあと、解析表を作る作業に入るわけであるが、PASCAL- に対して解析表作成を試行したところ、解析表の一部と書きかえられた生成規則表の記憶領域だけで約 7.4 M バイトを要した。

† A Constructing Algorithm for Semi-LL(2) Grammar's Parsing-table by KEIICHI YOSHIDA (Department of Computer Science, College of Engineering, Shizuoka University) and YOSHIKO TAKEUCHI (Department of Computer Science, Hamamatu Polytechnic College).

†† 静岡大学工業短期大学部情報工学科
††† 浜松職業訓練短期大学校情報処理科

3. 基本的定義と記法

本論文の中で使用される記号について述べる。ただしここで述べない定義、記法については、文献 1) に従うものとする。

【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 $\text{FIRST}_k(\alpha)$ は以下で定義される。

$$\text{FIRST}_k(\alpha) = \{u \mid (\alpha \Rightarrow_k u\beta, \|u\| = k) \text{ または } (\alpha \Rightarrow_k u, \|u\| < k)\}$$

ただし、 $\alpha, \beta \in (N \cup \Sigma)^*$, $u \in \Sigma^*$, $\|u\|$ は u の長さを表す。また、 $\alpha \Rightarrow_k u\beta$ は生成規則を 0 回以上使用する最左導出を表す。なお、 \Rightarrow_k は生成規則を k 回使用する最左導出を表し、とくに $k=1$ のとき \Rightarrow で表す。本論文で扱う導出は、とくに断りが無い限り、すべて最左導出である。

【定義 3】 集合 $\text{END-FOLLOW}(X)$ は以下で定義される。

$$\begin{aligned} \text{END-FOLLOW}(X) &= \{A \mid (A \Rightarrow \alpha X \beta, X \in N, \beta \Rightarrow \varepsilon), \\ &\text{または } (A \Rightarrow \alpha B X, X \in \Sigma)\} \end{aligned}$$

【定義 4】²⁾ L_1, L_2 を Σ^* の部分集合とするととき、演算子 \oplus_k は以下で定義される。

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

【定義 5】 文脈自由文法 $G = (N, \Sigma, P, S)$ において $S \Rightarrow u_i A X \xi_i$ のとき、 PF (partial-FOLLOW) は以下で定義される。

$$\text{PF}_k(A, X) = \cup_i \text{FIRST}_k(X \xi_i)$$

【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とき

$$S \Rightarrow u A X v$$

なるすべての $u A X v$ に対して

$$(\text{FIRST}_k(\alpha) \oplus_k \text{PF}_k(A, X))$$

$$\cap (\text{FIRST}_k(\beta) \oplus_k \text{PF}_k(A, X)) = \phi$$

が成り立つとき、 G は準 LL(k) 文法であるという。

4. 解析表の構造

4.1 基本的な定義

解析表作成のアルゴリズムを述べるのに必要ないくつかの定義を与える。

【定義 7】 文脈自由文法 $G = (N, \Sigma, P, S)$ を用いて次の拡大文法 G' を定義する。

$$\begin{aligned} G' &= (N', \Sigma', P', S'), \\ N' &= N \cup \{S'\}, \Sigma' = \Sigma \cup \{\$, \}, \\ P' &= P \cup \{S' \rightarrow S\$\$ \} \end{aligned}$$

以下では、この拡大文法 G' を用いて議論を展開するが、便宜上、 N', Σ', P' をそれぞれ N, Σ, P と読みかえるものとする。

【記法 2】 生成規則ならびに導出

$$A \xrightarrow{p} \alpha, \alpha \Rightarrow_p \beta$$

における p 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 $\alpha \Rightarrow_p \beta$ で用いられた生成規則の番号を示す。

【定義 8】 生成規則の番号を p, q で表すとき $[] p(q)$ または $[X]p(q)$ を π 型生成規則番号という。ここで、 p, q は以下の導出を満足するものとする。

$$S' \Rightarrow u' Y \alpha' \Rightarrow_p u' \xi A \gamma \alpha' \Rightarrow_p u A \alpha \Rightarrow_p u \beta \alpha$$

ただし、 $\alpha = \gamma \alpha'$ で X は α の最左端一記号を表す。 π 型生成規則番号は解析表作成のために、途中の文脈の保存としてのみ用いられる。また、 π 型生成規則番号 $[X]p(q)$ において、 $X = \varepsilon, q = \varepsilon$ のとき、 $[X]p(q)$ は $[]p()$ を表すものとする。

【定義 9】 π 型生成規則番号 $[X]p(q)$ から (q) を取り除いた生成規則番号 $[X]p$ を、 τ 型生成規則番号という。ここでも、 $X = \varepsilon$ のとき $[X]p$ は $[]p$ を表すものとする。

【定義 10】 解析表 T は行、列がそれぞれ $(N - \{S'\}) \cup \Sigma, \Sigma$ の要素によって名付けられたマトリクスとする。解析表 T の A 行、 a 列の要素を $T(A, a)$ 、また a 行、 b 列の要素を $T(a, b)$ で表し、これら各要素には τ 型生成規則番号の集合または nil が記入される。

4.2 解析表の性質

準 LL(2) の定義より、解析表は以下の性質をもっていなければならない¹⁾。

【性質 1】

$$S' \Rightarrow u A \alpha \Rightarrow_p u \beta \alpha \Rightarrow_p u a b \gamma \alpha$$

$\longleftrightarrow []p \in T(A, a) \text{ かつ } []p \in T(a, b)$

【性質 2】

$S' \Rightarrow uAX\alpha \Rightarrow u\beta X\alpha \Rightarrow uaX\alpha \Rightarrow uab\delta$

$\longleftrightarrow []p \in T(A, a) \text{ かつ } [X]p \in T(a, b)$

【性質 3】

$S' \Rightarrow uAX\alpha \Rightarrow u\beta X\alpha \Rightarrow uX\alpha \Rightarrow uaby$

$\longleftrightarrow [X]p \in T(A, a) \text{ かつ } [X]p \in T(a, b)$

【性質 4】 解析表の要素への記入は、前記 3 つの場合以外にはない。

4.3 解析表作成用の諸表

前節の性質をもつ解析表の作成のアルゴリズムに関連する諸表の構造ならびにそれらと前出の【定義】との関連について述べる。

(1) FIRST 表: FIRST は本来は【定義 2】で定義される集合であるが、アルゴリズムの中ではそれに対応する表の名前である。その行および列がそれぞれ $NU\Sigma$, $NU\Sigma$ の要素で名付けられたマトリクスであり、その要素には π 型生成規則番号の集合または nil が記入される。FIRST 表の $N \times (NU\Sigma)$ 領域、 $\Sigma \times (NU\Sigma)$ 領域をそれぞれ FT_1 , FT_2 で表す。このとき、集合 $FIRST_2$ と対応する部分は FT_1 の $N \times \Sigma$, FT_2 の $\Sigma \times \Sigma$ の部分の和であって、その関係は次のとおりである。

$$ab \in FIRST_2(A) \longleftrightarrow [X]p(q) \in FT_1(A, a) \text{ かつ } [Y]p(r) \in FT_2(a, b)$$

ただし、 p, q, r は生成規則番号で、 $X, Y \in NU\Sigma \cup \{\epsilon\}$ 。

(2) PF 表: PF は本来は【定義 5】で定義される集合であるが、アルゴリズムの中ではそれに対応する表の名前とする。その構造は FIRST 表と同じで、その要素には π 型生成規則番号の集合または nil が記入される。PF 表の $N \times (NU\Sigma)$ 領域ならびに $\Sigma \times (NU\Sigma)$ 領域をそれぞれ PL_1 , PL_2 で表す。このとき、集合 PF_2 と対応する部分は PL_1 の $N \times \Sigma$, PL_2 の $\Sigma \times \Sigma$ の部分の和であって、その関係は次のとおりである。

$$ab \in PF_2(A, X) \longleftrightarrow [X]p(q) \in PL_1(A, a) \text{ かつ } [Y]p(r) \in PL_2(a, b)$$

ただし、 p, q, r は生成規則番号で、 $X, Y \in NU\Sigma \cup \{\epsilon\}$ 。

(3) END-FOLLOW 表: END-FOLLOW は本来は【定義 3】で定義される集合であるが、アルゴリズムの中ではそれに対応する表の名前とする。その行および列はそれぞれ $NU\Sigma$, N の要素で名付けられた

マトリクスであり、その要素には π 型生成規則番号の集合または nil が記入される。END-FOLLOW 表を EF で表すとき、集合 END-FOLLOW と EF との関係は次のとおりである。

$$A \in \text{END-FOLLOW}(Y) \longleftrightarrow [X]p(q) \in \text{EF}(Y, A)$$

ただし、 p, q は生成規則番号で、 $Y \in NU\Sigma$, $X \in NU\Sigma \cup \{\epsilon\}$ 。

解析表は、最終的な FIRST 表の、 $(NU\Sigma) \times \Sigma$ 領域として得られる。すでに述べたように、その要素には τ 型生成規則番号の集合または nil が記入される。

5. アルゴリズム

本論文で提案するアルゴリズムを便宜上、アルゴリズム θ と呼ぶことにする。まず、アルゴリズム θ の中で用いられる記号と定義について述べる。

【記法 3】 記号列 β の左から i 番目の記号を ${}^{(i)}\beta$ で表す。

【記法 4】 各表の要素を表すのに、【定義 10】で示した解析表の要素に対するものと同様な記法が使用される。

【定義 11】 集合 Q, R, Γ はそれぞれ以下で定義される。

$$Q = \{(A, p) \mid A \xrightarrow{p} \alpha \Rightarrow \epsilon\}$$

$$R = \{(A, a, p) \mid A \xrightarrow{p} \alpha \Rightarrow a\}$$

$$\Gamma = \{p \mid p \text{ は各生成規則につけた固有の番号で正整数}\}$$

次に、アルゴリズム θ の詳細を述べる。

【Step 1】 FIRST 表の初期化。つまり、 $A \xrightarrow{p} \alpha Y_i \beta$, $\alpha \Rightarrow \epsilon$ ならば $[]p(p)$ を $FT_1(A, Y_i)$ に加え、 $Y_i \in \Sigma$ ならば $FT_2(Y_i, {}^{(1)}\beta)$ にも加える。 $Y_i \xrightarrow{p_i} \gamma \Rightarrow a$ (つまり、 $(Y_i, a, p_i) \in R$) ならば $[]p(p)$ を $FT_2(a, {}^{(1)}\beta)$ に加える。

Firstly, let every entry of the FIRST-table

be nil.

begin

for each production such that $A \xrightarrow{p} Y_1 Y_2 \dots Y_n$
and $Y_1 Y_2 \dots Y_n \neq \epsilon$ do

begin

$i \leftarrow 0$;

repeat

$i \leftarrow i + 1$;

$FT_1(A, Y_i) \leftarrow FT_1(A, Y_i) \cup \{ []p(p) \}$;

/* find 2nd symbol Y_j on A^* */

if $i + 1 \leq n$ then

```

begin
  if  $Y_i \in \Sigma$  then  $X \leftarrow Y_i'$ 
  else if  $(Y_i, a, p_i) \in R$  then  $X \leftarrow a'$ 
  else goto  $\ell$ ;
   $j \leftarrow i$ ;
  repeat
     $j \leftarrow j+1$ ;
     $FT_2(X, Y_j) \leftarrow FT_2(X, Y_j) \cup \{ [ ]p(p) \}$ ;
  until  $(Y_j, p_j) \in Q$  or  $j+1 > n$ 
end;
 $\ell$ :
until  $(Y_i, p_i) \in Q$  or  $i+1 > n$ 
end
end.

```

[Step 2] FIRST 表の閉包. 前半は $A \Rightarrow_p \alpha B \beta$, $B \Rightarrow \gamma C \delta$, $\alpha, \gamma \Rightarrow \varepsilon$ ならば $[]p(p)$ を $FT_1(A, C)$ に加える. 後半は終端記号による非終端記号の置換えを行う.

```

begin
  repeat
    for each  $A, B, C \in N$  do
      if  $[ ]p(p)$  in  $FT_1(A, B)$  and  $[ ]q(q)$ 
        in  $FT_1(B, C)$  then
         $FT_1(A, C) \leftarrow FT_1(A, C) \cup \{ [ ]p(p) \}$ ;
    until no change occurs in the FIRST-table;
  repeat
    for each  $A, B \in N$ , and  $a \in \Sigma$  do
      if  $FT_1(A, B) \neq \phi$  and  $FT_1(B, a) \neq \phi$  then
        begin
          for each  $[ ]p(p) \in FT_1(A, B)$  do
            begin
               $FT_1(A, a) \leftarrow FT_1(A, a) \cup \{ [ ]p(p) \}$ ;
            /* find 2nd symbol  $X$  on  $A$  */
            for each  $X \in N \cup \Sigma$  do
              if  $FT_1(B, a) \cap FT_2(a, X) \neq \phi$  then
                 $FT_2(a, X) \leftarrow FT_2(a, X) \cup \{ [ ]p(p) \}$ 
            end
          end
        until no change occurs in the FIRST-table;
      if  $R = \{ \}$  and  $Q = \{ \}$  then skip Step 3
        through Step 7

```

end.

[Step 3] PF 表の初期化. つまり, $A \rightarrow_p \alpha Y_i \beta Y_j \gamma$, $\beta \Rightarrow \varepsilon$ ならば $[^{(1)}\xi]p(p)$ を $PL_1(Y_i, Y_j)$ に加え, $Y_j \in \Sigma$ ならば $PL_2(Y_j, ^{(1)}\gamma)$ にも加える. ただし, $\xi = \beta Y_j$.

Firstly, let every entry of the PF-table be nil.
 begin /* 0 indicates an index for an imaginary
 production $S' \rightarrow S \$ \$$ */

$PL_1(S, \$) \leftarrow [\$]0(0)'$;

$PL_2(\$, \$) \leftarrow [\$]0(0)'$;

for each production such that $A \rightarrow_p Y_1 Y_2 \dots Y_n$
 and $Y_1 Y_2 \dots Y_n \neq \varepsilon$ do

begin

 $i \leftarrow 1$;

repeat

 $j \leftarrow i+1$;if $Y_i \in N$ and $j \leq n$ then

begin

 $PL_1(Y_i, Y_j) \leftarrow PL_1(Y_i, Y_j)$ $\cup \{ [Y_j]p(p) \}$; $k \leftarrow j$;if $Y_k \in N$ then

begin

while $(Y_k, p_k) \in Q$ and $k+1 \leq n$ do

begin

 $k \leftarrow k+1$; $PL_1(Y_i, Y_k) \leftarrow PL_1(Y_i, Y_k)$ $\cup \{ [Y_j]p(p) \}$

end

end;

/* find 2nd symbol Y_m on Y_i */if $Y_k \in \Sigma$ thenif $k+1 \leq n$ then

begin

 $m \leftarrow k+1$; $PL_2(Y_k, Y_m) \leftarrow PL_2(Y_k, Y_m)$ $\cup \{ [Y_j]p(p) \}$;while $(Y_m, p_m) \in Q$ and $m+1 \leq n$ do

begin

 $m \leftarrow m+1$; $PL_2(Y_k, Y_m) \leftarrow PL_2(Y_k, Y_m)$ $\cup \{ [Y_j]p(p) \}$

```

end
end
end;
i ← i + 1
until i ≥ n
end
end.

```

[Step 4] END-FOLLOW 表の初期化. つまり, $A \xrightarrow{p} Y_1 Y_2 \dots Y_n$, $Y_1 Y_2 \dots Y_n \neq \varepsilon$ に対して $Y_n \in N$ ならば $[]p(p)$ を $EF(Y_n, A)$ に加え, $Y_n \in \Sigma$ かつ $Y_{n-1} \in N$ ならば $[Y_n]p(p)$ を加える.

Firstly, let every entry of the
END-FOLLOW-table be nil.

```

begin
for each production such that  $A \xrightarrow{p} Y_1 Y_2 \dots Y_n$ 
and  $Y_1 Y_2 \dots Y_n \neq \varepsilon$  do
if  $Y_n \in N$  then
begin
 $EF(Y_n, A) \leftarrow EF(Y_n, A) \cup [ ]p(p)$ ;
i ← n;
while  $(Y_i, p_i) \in Q$  and  $i > 1$  do
begin
if  $Y_{i-1} \in N$  then
 $EF(Y_{i-1}, A) \leftarrow EF(Y_{i-1}, A)$ 
 $\cup [Y_i]p(p)$ ;
i ← i - 1
end
end
else if  $Y_{n-1} \in N$  then
 $EF(Y_n, A) \leftarrow EF(Y_n, A) \cup [Y_n]p(p)$ 
end.

```

[Step 5] END-FOLLOW 表の閉包. つまり, $B \xrightarrow{q} \alpha_1 B_1 \beta_1$, $B_1 \xrightarrow{q} \alpha_2 B_2 \beta_2, \dots$, $B_{n-1} \xrightarrow{q} \alpha_n X \beta_n$ かつ $C \xrightarrow{q} \gamma_1 C_1 \delta_1$, $C_1 \xrightarrow{q} \gamma_2 C_2 \delta_2, \dots$, $C_{n-1} \xrightarrow{q} \gamma_n B \delta_n$ で, $\beta_n \beta_{n-1} \dots \beta_1 \Rightarrow \varepsilon$, $\delta_n \delta_{n-1} \dots \delta_1 \Rightarrow \varepsilon$ のとき, $X \in N$ ならば $[V]p(h)$ を $EF(X, C)$ に加え, $X = a \in \Sigma$ ならば $[a]p(p)$ を $EF(a, C)$ に加える. ただし, V は $\beta_n \beta_{n-1} \dots \beta_1 \delta_n \delta_{n-1} \dots \delta_1$ の先頭一記号. また, 記号列 $\beta_n \beta_{n-1} \dots \beta_1$ および $\delta_n \delta_{n-1} \dots \delta_1$ の先頭一記号をそれぞれ Z, Y とする.

```

begin
repeat
for each  $B, C \in N$ ;  $Y, Z \in NU\{\varepsilon\}$ ,
and  $X \in N$  do

```

```

if  $[Z]p(q)$  in  $EF(X, B)$  and  $[Y]g(h)$ 
in  $EF(B, C)$  then

```

```

begin
if  $Z \neq \varepsilon$  then  $V \leftarrow 'Z'$  else  $V \leftarrow 'Y'$ ;
 $EF(X, C) \leftarrow EF(X, C) \cup \{[V]p(h)\}$ 
end

```

until no change occurs in the
END-FOLLOW-table;

for each $B, C \in N$, $Y \in NU\{\varepsilon\}$, and $X \in \Sigma$ do
if $[X]p(p)$ in $EF(X, B)$ and $[Y]g(h)$

```

in  $EF(B, C)$  then
 $EF(X, C) \leftarrow EF(X, C) \cup \{[X]p(p)\}$ 
end.

```

[Step 6] END-FOLLOW 表を用いて PF 表を補完. 前半は $C \Rightarrow \alpha_1 B \xi Y \beta$, $B \Rightarrow \gamma_1 X_1 \delta_1$, $X_1 \Rightarrow \gamma_2 X_2 \delta_2, \dots$, $X_{n-1} \Rightarrow \gamma_n X a$, $\delta_{n-1} \dots \delta_1 \Rightarrow \varepsilon$ ならば $[a]p(p)$ を $PL_2(a, Y)$ に加える. 後半は, $B \Rightarrow \alpha_1 A_1 \beta_1$, $A_1 \Rightarrow \alpha_2 A_2 \beta_2, \dots$, $A_{n-1} \Rightarrow \alpha_n A \beta_n$ かつ $C \Rightarrow \gamma_1 X_1 \xi_1 Y \delta$, $X_1 \Rightarrow \gamma_2 X_2 \xi_2, \dots$, $X_{n-1} \Rightarrow \gamma_n B \xi_n$ で, $\xi_n \xi_{n-1} \dots \xi_1 \Rightarrow \varepsilon$ ならば $[V]p(r)$ を $PL_1(A, Y)$ に加え, $Y \in \Sigma$ ならば $PL_2(Y, W)$ にも加える (付図). ただし, V, W はそれぞれ $\beta_n \beta_{n-1} \dots \beta_1 \xi_n \xi_{n-1} \dots \xi_1, \delta$ の先頭の一記号.

```

begin
for each  $B \in N$ ,  $Y, Z \in NU\Sigma$ ,
 $X \in NU\Sigma \cup \{\varepsilon\}$ , and  $a \in \Sigma$  do
if  $[a]p(p)$  in  $EF(a, B)$  and  $[Z]q(q)$ 
in  $PL_1(B, Y)$  then
 $PL_2(a, Y) \leftarrow PL_2(a, Y) \cup \{[a]p(p)\}$ ;
for each  $A, B \in N$ ,  $Y \in NU\Sigma$ 
and  $X, Z \in NU\Sigma \cup \{\varepsilon\}$  do
if  $[X]p(s)$  in  $EF(A, B)$  and  $[Z]q(r)$ 
in  $PL_1(B, Y)$  then

```

```

begin
if  $X \neq \varepsilon$  then  $V \leftarrow 'X'$  else  $V \leftarrow 'Z'$ ;
 $PL_1(A, Y) \leftarrow PL_1(A, Y) \cup \{[V]p(r)\}$ ;
/* find 2nd symbol  $W$  on  $A$  */
if  $Y \in \Sigma$  then
for each  $W \in NU\Sigma$ 
if  $PL_1(B, Y) \cap PL_2(Y, W) \neq \emptyset$  then
 $PL_2(Y, W) \leftarrow PL_2(Y, W)$ 
 $\cup \{[V]p(r)\}$ 
end

```

```

if  $R = \{ \}$  then goto Step 8
end.

```

[Step 7] $A \Rightarrow a$ に対する処理. つまり, $A \Rightarrow \alpha \Rightarrow a$, $B \Rightarrow \alpha_1 X_1 Y \beta_1 \Rightarrow \alpha_1 \dots \alpha_{n-1} X_{n-1} \beta_{n-1} \dots \beta_2 Y \beta_1 \Rightarrow \alpha_1 \alpha_2 \dots \alpha_n A \beta_n \beta_{n-1} \dots \beta_2 Y \beta_1$, $\beta_n \beta_{n-1} \dots \beta_2 \Rightarrow \varepsilon$ ならば $[Z]p(r)$ を $FT_2(a, X)$ に加える. ただし, Z は $\beta_n \beta_{n-1} \dots \beta_2$ の先頭一記号.

```
begin
  for each  $(A, a, p) \in R$  do
    for each  $X, Z \in NU\Sigma$ 
      if  $[Z]q(r)$  in  $PL_1(A, X)$  then
         $FT_2(a, X) \leftarrow FT_2(a, X) \cup \{[Z]p(r)\}$ 
      end.
```

[Step 8] $FT_2(a, A)$ の A を, 終端記号で置き換える ($A \Rightarrow \varepsilon$ のとき).

```
begin
  for each  $a \in \Sigma, A \in N, X \in NU\Sigma \cup \{\varepsilon\}$  do
    if  $[X]p(q)$  in  $FT_2(a, A)$  and  $[ ]r(r)$ 
      in  $FT_1(A, b)$  then
       $FT_2(a, b) \leftarrow FT_2(a, b) \cup \{[X]p(q)\}$ ;
    if  $Q = \{ \}$  then skip Step 9 through Step 10
  end.
```

[Step 9] $u_0 D \beta_0 \Rightarrow u_0 \alpha_1 D_1 \beta_1 \beta_0 \Rightarrow u_1 D_1 \beta_1 \beta_0 \Rightarrow u_{n-1} D_{n-1} \beta_{n-1} \beta_{n-2} \dots \beta_1 \beta_0 \Rightarrow u_{n-1} \alpha_n A \gamma B \beta_n \beta_{n-1} \dots \beta_1 \beta_0 \Rightarrow u A \gamma B \beta_n \beta_{n-1} \dots \beta_1 \beta_0 = u A \gamma B \beta_n \Rightarrow u \alpha \gamma B \beta_n \Rightarrow u \gamma B \beta_n \Rightarrow u B \beta_n \Rightarrow u \delta \beta_n \Rightarrow u a b \xi$ に対する $PL_1(A, a)$, $PL_2(a, b)$ を求める. つまり, $PL_1(A, B) \neq \phi$, $A \Rightarrow \varepsilon$ に対して, 前半は $B \Rightarrow a b \xi$ の場合の処理で, 後半は $B \Rightarrow a$, $\beta_n \Rightarrow b \xi$ の場合の処理を行う. さらに, $PL_2(a, {}^{(1)}\beta)$ の ${}^{(1)}\beta$ が非終端記号のとき ${}^{(1)}\beta$ を終端記号で置き換える.

```
begin
  for each  $(A, h) \in Q, B \in N, X \in NU\Sigma,$ 
     $Z \in NU\Sigma \cup \{\varepsilon\}$ , and  $a, b \in \Sigma$  do
    if  $PL_1(A, B) \neq \phi$  and  $FT_1(B, a) \neq \phi$  then
      begin
        if  $[ ]p(p)$  in  $FT_1(B, a)$  and  $[Z]p'(q)$ 
          in  $FT_2(a, b)$  then
          if  $p = p'$  then for each  $[Z]p(q)$ 
            if  $p = q$  then
              begin /*  $B \beta_n \Rightarrow a b \xi, \beta_n = a b \xi$  */
                for each  $[X]r(s)$ 
                  in  $PL_1(A, B)$  do
                  begin
                     $PL_1(A, a) \leftarrow PL_1(A, a)$ 
                       $\cup \{[X]r( )\}$ ;

```

```
 $PL_2(a, b) \leftarrow PL_2(a, b)$ 
   $\cup \{[X]r( )\}$ 
```

```
end
end
else
  begin /*  $B \beta_n \Rightarrow a b \xi, \beta_n = a b \xi$  */
    for each  $C \in NU\Sigma$ 
      and  $Y \in NU\Sigma$  do
      for each  $[Y]t(u)$  in  $PL_1(B, C)$ 
        and  $[X]r(s)$  in  $PL_1(A, B)$  do
        if  $s = t$  and  $u = q$  then
          for each  $[X]r(s)$ 
            in  $PL_1(A, B)$  do
            begin
               $PL_1(A, a) \leftarrow PL_1(A, a)$ 
                 $\cup \{[X]r( )\}$ ;
               $PL_2(a, b) \leftarrow PL_2(a, b)$ 
                 $\cup \{[X]r( )\}$ 
            end
          end
        end
      end
    end
  end
end;
```

```
/* Replace nonterminal 2nd symbol with
  terminal symbol */
for each  $B \in N, Z \in NU\Sigma$  and  $a, b \in \Sigma$  do
  if  $[Z]p(q)$  in  $PL_2(a, B)$ 
    and  $FT_1(B, b) \neq \phi$  then
     $PL_2(a, b) \leftarrow PL_2(a, b) \cup \{[Z]p(q)\}$ 
  end.
```

[Step 10] $A \Rightarrow \alpha \Rightarrow \varepsilon$, つまり $S' \Rightarrow u A \gamma \Rightarrow u \alpha \gamma \Rightarrow u \gamma \Rightarrow u a b \xi$ に対する $FT_1(A, a)$, $FT_2(a, b)$ を求める.

```
begin
  for each  $A$  such that  $(A, p) \in Q$  do
    for each  $a, b \in \Sigma$  do
      if  $PL_1(A, a) \cap PL_2(a, b) \neq \phi$  then
        begin
          for each  $[X]q(r)$  in  $PL_1(A, a)$ 
            for each  $X \in NU\Sigma$  do
              begin
                 $FT_1(A, a) \leftarrow FT_1(A, a)$ 
                   $\cup \{[X]p( )\}$ ;
                 $FT_2(a, b) \leftarrow FT_2(a, b)$ 
                   $\cup \{[X]p( )\}$ 
              end
            end
          end
        end
      end
    end
  end
end
```

end.

[Step 11] 解析表を作成する. つまり, π 型生成規則番号を τ 型生成規則番号に変換する.

begin

for each $Z \in NU\Sigma \cup \{\epsilon\}$, $A \in N$,

and $a, b \in \Sigma$ do

replace $[Z]p(r)$ in $FT_1(A, a)$

or $FT_2(a, b)$ by $[Z]p$

end.

[End of Algorithm]

6. 証明

前章で提案された LL(2) 文法のアルゴリズム θ で最終的に生成される FIRST 表の $N \times \Sigma$, $\Sigma \times \Sigma$ の領域をそれぞれ T_1 , T_2 で表し $T = T_1 + T_2$ とするとき, T が LL(2) 文法解析表になっていることを証明する. π 型生成規則番号は解析表作成のためののみ必要であり, 解析表作成の最終ステップ Step 11 ですべて τ 型生成規則番号にかえられる. したがって, 解析表としてみたときは, π 型生成規則番号と τ 型生成規則番号とは等価である.

ゆえに, 解析表が正しいことを示すには

$$\begin{aligned} S' &\xrightarrow{p} uAv \xrightarrow{p} uav \xrightarrow{p} uab\xi \\ \longleftrightarrow [X]p &\in T_1(A, a) \\ [Y]p &\in T_2(a, b) \end{aligned} \quad (6.1)$$

が成り立つことを示せばよい. ただし, $X, Y \in NU\Sigma \cup \{\epsilon\}$,

$X = \epsilon$ のとき, $Y = {}^{(1)}\nu$ または $Y = \epsilon$

$X \neq \epsilon$ のとき, $Y = X = {}^{(1)}\nu$

とする. また, Step i で $FT_1, FT_2, PL_1, PL_2, EF$ に追加される内容をそれぞれ $FT_1^i, FT_2^i, PL_1^i, PL_2^i, EF^i$ で表し, Step i, j のいずれかで表に追加される内容をそれぞれ $FT_1^{i/j}, FT_2^{i/j}, PL_1^{i/j}, PL_2^{i/j}, EF^{i/j}$ で表す.

さて, 文法が準 LL(2) であるから, (6.1) 式が成り立つためには, 次のことが成り立てばよい. すなわち

(i) $A \xrightarrow{p} \alpha$, $\alpha \xrightarrow{p} ab\xi$ のとき, つまり

$$\begin{aligned} S' &\xrightarrow{p} uAv \xrightarrow{p} uav \xrightarrow{p} uab\xi\nu \\ \longleftrightarrow FT_1^{i/2}(A, a) &\ni []p, \text{ かつ} \\ FT_2^{i/2/8}(a, b) &\ni []p \end{aligned}$$

(ii) $A \xrightarrow{p} \alpha$, $\alpha \xrightarrow{p} a$, $\nu \xrightarrow{p} b\xi$ のとき, つまり

$$\begin{aligned} S' &\xrightarrow{p} uAv \xrightarrow{p} uav \xrightarrow{p} uab\xi \\ \longleftrightarrow FT_1^{i/2}(A, a) &\ni []p, \text{ かつ} \end{aligned}$$

$$FT_2^{7/8}(a, b) \ni [{}^{(1)}\nu]p$$

(iii) $A \xrightarrow{p} \alpha$, $\alpha \xrightarrow{p} \epsilon$, $\nu \xrightarrow{p} ab\psi$ のとき, つまり

$$\begin{aligned} S' &\xrightarrow{p} uAv \xrightarrow{p} uav \xrightarrow{p} uv \xrightarrow{p} uab\psi \\ \longleftrightarrow FT_1^{10}(A, a) &\ni [{}^{(1)}\nu]p, \text{ かつ} \\ FT_2^{10}(a, b) &\ni [{}^{(1)}\nu]p \end{aligned}$$

前述したように, 準 LL(2) 文法の解析表作成では文脈を考慮する必要がある場合 (上記の (ii), (iii) の場合) があるので, 便宜上, 解析表の正しさを示す証明ではすべて π 型生成規則番号を用いることとする.

(i) の証明

導出

$$A \xrightarrow{p} \alpha \xrightarrow{p} ab\nu, \quad a, b \in \Sigma, \nu \in (NU\Sigma)^*$$

は文法がサイクルを含まないから, 有限回の導出である. この回数を $k_0 (= k+1)$ とすると, 上の導出は

$$A \xrightarrow{p} \alpha \xrightarrow{p} ab\nu$$

と書き表すことができる. これは一般に

$$A \xrightarrow{p} \alpha = \mu_1 X_1 \nu_1 \xrightarrow{p} X_1 \nu_1 \xrightarrow{p} ab\eta_1, \quad \eta_1 = \nu \quad (6.2)$$

ただし, $X_1 \xrightarrow{p} \epsilon$, $k_1 < k_0$

と書き直すことができるから, 文法の中に

$$A \xrightarrow{p} \mu_1 X_1 \nu_1, \quad \mu_1 \xrightarrow{p} \epsilon$$

なる生成規則が存在しなければならない. したがって, Step 1 より

$$\begin{aligned} A &\xrightarrow{p} \mu_1 X_1 \nu_1, \quad \mu_1 \xrightarrow{p} \epsilon \\ \longleftrightarrow FT_1^1(A, X_1) &\ni []p(p) \end{aligned}$$

$X_1 \in \Sigma$ ならば

$$FT_2^1(X_1, {}^{(1)}\nu_1) \ni []p(p)$$

が成り立つ. (6.2) 式の X_1 からの導出はさらに,

$$X_1 \xrightarrow{p_1} \mu_2 X_2 \nu_2 \xrightarrow{p_1} X_2 \nu_2 \xrightarrow{p_1} ab\eta_2$$

ただし, $X_2 \xrightarrow{p_1} \epsilon$, $k_2 < k_1$

と書き直すことができ, はじめの部分の導出に対して

$$X_1 \xrightarrow{p_1} \mu_2 X_2 \nu_2, \quad \mu_2 \xrightarrow{p_1} \epsilon$$

なる生成規則が存在しなければならない. したがって, Step 1 より

$$\begin{aligned} X_1 &\xrightarrow{p_1} \mu_2 X_2 \nu_2, \quad \mu_2 \xrightarrow{p_1} \epsilon \\ \longleftrightarrow FT_1^1(X_1, X_2) &\ni []p_1(p_1) \\ X_2 \in \Sigma \text{ ならば} & \\ FT_2^1(X_2, {}^{(1)}\nu_2) &\ni []p_1(p_1) \end{aligned}$$

この手順は一般に

$$X_i \xrightarrow{p_i} \mu_{i+1} X_{i+1} \nu_{i+1} \xrightarrow{p_i} X_{i+1} \nu_{i+1} \xrightarrow{p_i} ab\eta_{i+j}$$

ただし, $X_{i+1} \xrightarrow{p_i} \epsilon$, $k_{i+1} < k_i$

と書くことができ, このとき $X_0 = A$, $p_0 = p$ として

$$X_i \xrightarrow{p_i} \mu_{i+1} X_{i+1} \nu_{i+1}, \quad \mu_{i+1} \xrightarrow{p_i} \epsilon$$

ただし, $\xi_j = \beta_j \xi_{j-1} \xrightarrow{\varepsilon} \varepsilon, k_m < k_{m-1}$

と表すことができる. ゆえに

$$X_{j-1} \xrightarrow[p_{j-1}]{} \alpha_j X_j \beta_j, \beta_j \xrightarrow{\varepsilon} \varepsilon$$

$$\longleftrightarrow EF^4(X_j, X_{j-1}) \ni [^{(1)}\beta_j] p_{j-1}(p_{j-1}) \quad (6.10)$$

(6.7)と(6.8)式に対する Step 5 の操作により

$$X_i \xrightarrow[p_i]{} \alpha_{i+1} X_{i+1} \beta_{i+1} \xrightarrow{\varepsilon} u_{i+1} X_{i+1} \beta_{i+1}$$

$$\xrightarrow[p_{i+1}]{} u_{i+1} \alpha_{i+2} X_{i+2} \xi_{i+2}$$

ただし, $u_{i+1} \in \Sigma^*, \xi_{i+2} = \beta_{i+2} \beta_{i+1} \xrightarrow{\varepsilon} \varepsilon,$

$$\longleftrightarrow EF^{4/6}(X_{i+2}, X_i) \ni [^{(1)}\xi_{i+2}] p_{i+1}(p_i) \quad (6.11)$$

さらに, (6.11), (6.9)式に対する Step 5 の操作により

$$X_i \xrightarrow[p_{i+3}]{} u_{i+2} X_{i+2} \xi_{i+2} \xrightarrow{\varepsilon} u_{i+2} \alpha_{i+3} X_{i+3} \xi_{i+3}$$

ただし, $u_{i+2} \in \Sigma^*, \xi_{i+3} = \beta_{i+3} \beta_{i+2} \beta_{i+1} \xrightarrow{\varepsilon} \varepsilon$

$$\longleftrightarrow EF^{4/6}(X_{i+3}, X_i) \ni [^{(1)}\xi_{i+3}] p_{i+2}(p_i)$$

同様にして, Step 5 の操作をくりかえすことにより

$$X_i \xrightarrow[p_i]{} \alpha_{i+1} X_{i+1} \beta_{i+1} \xrightarrow{\varepsilon} u_{j-1} X_{j-1} \xi_{j-1}$$

$$\xrightarrow[p_{j-1}]{} u_{j-1} \alpha_j X_j \xi_j, \xi_j = \beta_j \xi_{j-1} = \beta_j \beta_{j-1} \cdots \beta_{i+1} \xrightarrow{\varepsilon} \varepsilon$$

$$\longleftrightarrow EF^{4/6}(X_j, X_i) \ni [^{(1)}\xi_j] p_{j-1}(p_i) \quad (6.12)$$

が成り立つ.

また, (6.7)式の生成規則において, $\beta_{i+1} = \varepsilon, X_{i+1} = a \in \Sigma, \alpha_{i+1} \in N$ のときは Step 4 より

$$X_i \xrightarrow[p_i]{} \alpha_{i+1} a \longleftrightarrow EF^4(a, X_i) \ni [a] p_i(p_i) \quad (6.13)$$

であるから, (6.12)式において, $\beta_j = \varepsilon, X_j \in \Sigma, \alpha_j \in N$ ならば, $[^{(1)}\xi_j] p_{j-1}(p_i)$ は $[X_j] p_{j-1}(p_{j-1})$ とする.

(ii-b) つぎに, (ii-a) を用いて以下の関係が成り立つことを示す.

$$A \xrightarrow[q]{} \delta_i X_i \gamma_i, X_i \xrightarrow[p_{j-1}]{} u_{j-1} X_{j-1} \gamma' \xrightarrow{\varepsilon} u_{j-1} \delta_j X_j \gamma,$$

$$\gamma = \gamma_j \gamma' = \gamma_j \gamma_{j-1} \cdots \gamma_{i+1} \xrightarrow{\varepsilon} \varepsilon$$

$$\longleftrightarrow PL_1^{3/6}(X_j, \gamma_i) \ni [^{(1)}\nu] p_{j-1}(q) \quad (6.14)$$

$\gamma_i \in \Sigma$ ならば

$$PL_2^6(\gamma_i, \gamma_i) \ni [^{(1)}\nu] p_{j-1}(q)$$

ただし, $\nu = \gamma \gamma_i, u_{j-1} \in \Sigma^*$ (図 2).

まず, Step 3 により

$$A \xrightarrow[q]{} \delta_i X_i \gamma_i$$

$$\longleftrightarrow PL_1^3(X_i, \gamma_i) \ni [^{(1)}\gamma_i] q(q) \quad (6.15)$$

$\gamma_i \in \Sigma, \gamma_i \neq \varepsilon$ ならば

$$PL_2^3(\gamma_i, \gamma_i) \ni [^{(1)}\gamma_i] q(q)$$

は明らか. (6.6), (6.15)式に対する Step 6 の操作により, (6.14)式が成り立つ.

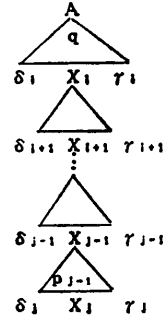


図 2 $\gamma = \gamma_j \gamma_{j-1} \cdots \gamma_{i+1} \xrightarrow{\varepsilon} \varepsilon$ のとき, $PL_1(X_j, \gamma_i)$ に $[^{(1)}\nu] p_{j-1}(q)$ を加え, $\gamma_i \in \Sigma$ ならば, $PL_2(\gamma_i, \gamma_i)$ に $[^{(1)}\nu] p_{j-1}(q)$ を加える. ただし, $\nu = \gamma \gamma_i$.

Fig. 2 In the case of Fig. 2, we can find $PL_1(X_j, \gamma_i)$ and $PL_2(\gamma_i, \gamma_i)$. That is, we add $[^{(1)}\nu] p_{j-1}(q)$ to $PL_1(X_j, \gamma_i)$ if $\gamma = \gamma_j \gamma_{j-1} \cdots \gamma_{i+1} \xrightarrow{\varepsilon} \varepsilon$. If $\gamma_i \in \Sigma$, we add $[^{(1)}\nu] p_{j-1}(q)$ to $PL_2(\gamma_i, \gamma_i)$, where $\nu = \gamma \gamma_i$.

さて, (ii-b) を用いて, (ii) のはじめに示した関係を証明する. (ii-b) より次の関係が成り立つ.

$$X \xrightarrow[q]{} \beta B \delta, B \xrightarrow[p_{j-1}]{} u_{j-1} B_{j-1} \xi' \xrightarrow{\varepsilon} u_{j-1} \beta_j A \xi, \xi \xrightarrow{\varepsilon} \varepsilon$$

$$\longleftrightarrow PL_1^{3/6}(A, \delta) \ni [^{(1)}\nu] p_{j-1}(q)$$

$\delta \in \Sigma$ ならば

$$PL_2^{3/6}(\delta, \delta) \ni [^{(1)}\nu] p_{j-1}(q)$$

ただし, $\nu = \xi \delta, u_{j-1} \in \Sigma^*$

したがって

$$S' \xrightarrow[q]{} u_i X \gamma \xrightarrow[p_{j-1}]{} u_i \beta B \delta \gamma \xrightarrow{\varepsilon} u_i B \delta \gamma$$

$$\xrightarrow[p_{j-1}]{} u_i B_{j-1} \xi' \delta \gamma \xrightarrow{\varepsilon} u_i \beta_j A \xi \delta \gamma, \xi \delta \xrightarrow{\varepsilon} \varepsilon$$

$$\longleftrightarrow PL_1^{3/6}(A, \gamma) \ni [^{(1)}\zeta] p_{j-1}(q) \quad (6.16)$$

$\gamma \in \Sigma$ ならば

$$PL_2^{3/6}(\gamma, \gamma) \ni [^{(1)}\zeta] p_{j-1}(q)$$

ただし, $\zeta = \xi \delta \gamma, u_j, u_i \in \Sigma^*$.

いま, (6.16)式において

$$A \xrightarrow[p]{} \alpha \xrightarrow{\varepsilon} a, \xi \delta \gamma \xrightarrow{\varepsilon} b \eta$$

とする. まず, $A \xrightarrow[p]{} \alpha$ に対しても有限回の導出であるから

$$A \xrightarrow[p]{} \alpha \xrightarrow{k} a \quad k \text{ は定数}$$

と書き直すことができる. Step 1, 2 より

$$A \xrightarrow[p]{} \alpha \xrightarrow{k} a \longleftrightarrow FT_1^{1/2}(A, a) \ni [] p(p) \quad (6.17)$$

ここで(6.16), (6.17)式に対する Step 7 の操作より

$$S' \xrightarrow[q]{} u_i X \gamma \xrightarrow[p_{j-1}]{} u_i \beta B \delta \gamma \xrightarrow{\varepsilon} u_i B \delta \gamma \xrightarrow{\varepsilon} u_i \beta_j A \xi \delta \gamma$$

$$\xrightarrow[p_{j-1}]{} u_i B_{j-1} \xi' \delta \gamma \xrightarrow{\varepsilon} u_i \alpha \xi \delta \gamma \xrightarrow{\varepsilon} u_i a \xi \delta \gamma, \xi \delta \xrightarrow{\varepsilon} \varepsilon$$

$$\longleftrightarrow FT_2^7(a, \gamma) \ni [^{(1)}\zeta] p(q) \quad (6.18)$$

ただし, $\zeta = \xi \delta \gamma$.

ところで, $\xi \delta \gamma \xrightarrow{\varepsilon} b \delta$ より

$$\xi\delta\gamma \Rightarrow \gamma \Rightarrow b\eta$$

であるから, Step 1, 2 より

$$\gamma \Rightarrow \gamma' \Rightarrow b\eta$$

$$\longleftrightarrow FT_1^{1/2}(A, b) \ni []r(r) \quad (6.19)$$

(6.18), (6.19)式に対する Step 8 の操作より

$$u_1a\zeta\gamma \Rightarrow u_1a\gamma \Rightarrow u_1ab\delta$$

$$\longleftrightarrow FT_2^{7/8}(a, b) \ni [^{(1)}\zeta]p(q) \quad (6.20)$$

ゆえに, (6.16), (6.17), (6.20)式より

$$S' \Rightarrow u_i X_i \gamma \Rightarrow u_i \beta B \delta \gamma \Rightarrow u_i B \delta \gamma \Rightarrow u_n A \xi \delta \gamma$$

$$\Rightarrow u_n \alpha \xi \delta \gamma \Rightarrow u_n \alpha \xi \delta \gamma \Rightarrow u_n a \gamma \Rightarrow u_n a b \eta$$

$$\longleftrightarrow FT_1^{1/2}(A, a) \ni []p(p)$$

$$FT_2^{7/8}(a, b) \ni [^{(1)}\zeta]p(q)$$

ただし, $\zeta = \xi\delta\gamma$.

したがって, 生成規則番号 q は文型に A が現れるときの最初に用いられるものであり, p は a を導出するために A に適用される最初の生成規則番号であることを考慮すると, (ii) のはじめに示した次の関係が得られる.

$$S' \Rightarrow v X_i \gamma \Rightarrow v \beta A \xi \gamma \Rightarrow v \beta A v \Rightarrow u A v \Rightarrow u a v$$

$$\Rightarrow u a v \Rightarrow u a b \zeta$$

$$\longleftrightarrow FT_1^{1/2}(A, a) \ni []p(p)$$

$$FT_2^{7/8}(a, b) \ni [^{(1)}v]p(q)$$

(iii) の証明

次に, 以下の関係が成り立つことを示す.

$$S' \Rightarrow u A v \Rightarrow u a v \Rightarrow u v \Rightarrow u a b \phi$$

$$\longleftrightarrow FT_1^{10}(A, a) \ni [^{(1)}v]p()$$

$$FT_2^{10}(a, b) \ni [^{(1)}v]p() \quad (6.21)$$

この場合は,

$$A \Rightarrow a \Rightarrow \varepsilon$$

$$v \Rightarrow a b \phi \quad (6.22)$$

であり, (6.22)式については(i)の証明もしくは(ii)の証明に帰着できる. つまり, $v = X_1 X_2 \dots X_n$ (ただし, $X_i \in NU\Sigma$) とするとき, $X_1 X_2 \dots X_{i-1} \Rightarrow \varepsilon, X_i \Rightarrow \varepsilon, v' = X_{i+1} X_{i+2} \dots X_n$ のとき, (6.22)式は次のいずれかに分類できる.

(iii-a) $X_i v' \Rightarrow \zeta_1 v' \Rightarrow a b \phi v'$

$$\longleftrightarrow FT_1^{1/2}(X_i, a) \ni []p_1(p_1)$$

$$FT_2^{1/2}(a, b) \ni []p_1(p_1)$$

(iii-b) $X_i v' \Rightarrow \zeta_2 v' \Rightarrow a v' \Rightarrow a v_2' \Rightarrow a b v''$

$$\longleftrightarrow FT_1^{1/2}(X_i, a) \ni []p_2(p_2) \quad (6.23)$$

$$FT_2^{7/8}(a, b) \ni [^{(1)}v']p_2(q)$$

(iii-c) $X_i v' \Rightarrow \zeta_3 v' \Rightarrow v' \Rightarrow v_3' \Rightarrow a b v''$

$$\longleftrightarrow FT_1^{10}(X_i, a) \ni [^{(1)}v']p_3(q)$$

$$FT_2^{10}(a, b) \ni [^{(1)}v']p_3(q)$$

(iii-a), (iii-b) についてはそれぞれ(i), (ii)に帰着できる. (iii-c) についてはさらに(iii-a), (iii-b), (iii-c)に分解でき, しかも(iii-c)の導出は有限回であるから, 分解は有限回で終了する.

また(6.21)式において, $v = X_1 X_2 \dots X_n, X_1 X_2 \dots X_{i-1} \Rightarrow \varepsilon, ^{(1)}v = X_i$, とするとき, A と X_i, X_i と X_{i+1} の PF 関係, つまり $PL_1(A, X_i), PL_2(X_i, X_{i+1})$ については(ii-b)の証明に帰着できる. つまり

$$S' \Rightarrow \zeta \Rightarrow u A v, v = X_1 X_2 \dots X_n, X_1 \dots X_{i-1} \Rightarrow \varepsilon,$$

$$\longleftrightarrow PL_1^3(A, X_i) \ni [^{(1)}v]q(q) \quad (6.24)$$

$^{(1)}v \in \Sigma$ のとき

$$PL_2^{3/6}(X_i, X_{i+1}) \ni [^{(1)}v]q(q)$$

(6.24)式において, $X_i \in N$ もしくは $X_{i+1} \in N$ のときは, Step 9 を用いてこれらを終端記号で置き換える. つまり(iii-a), (iii-b), (6.24)式に対する Step 9 の操作より

$$S' \Rightarrow \zeta \Rightarrow u A v \Rightarrow u v \Rightarrow u a b \phi$$

$$\longleftrightarrow PL_1^{3/6/9}(A, a) \ni [^{(1)}v]q(q) \quad (6.25)$$

$$PL_2^{3/6/9}(a, b) \ni [^{(1)}v]q(q)$$

さらに, $A \Rightarrow a \Rightarrow \varepsilon$ であるから(6.25)式に対する Step 10 の操作により, 次の式が得られる.

$$S' \Rightarrow \zeta \Rightarrow u A v \Rightarrow u a v \Rightarrow u v \Rightarrow u a b \phi$$

$$\longleftrightarrow FT_1^{10}(A, a) \ni [^{(1)}v]p()$$

$$FT_2^{10}(a, b) \ni [^{(1)}v]p()$$

[証明終]

7. 適用例

下記の文法例をアルゴリズム θ に適用し, 各ステップごとの表の構成を図3~図10に示す.

[例] 強でない準 LL(2) 文法

1. $S \rightarrow aAaa$
4. $A \rightarrow b$

	S	A	a	b	\$
S		[]3(3)	[]1(1) []3(3)	[]2(2) []3(3)	
A				[]4(4)	
a		[]1(1)	[]1(1)		
b		[]2(2)	[]3(3)	[]2(2)	
\$					

図3 Step 2 終了後の G の FIRST 表

* 印は, Step 2 で記入されたもので, 無印は Step 1 で記入されたもの.

Fig. 3 FIRST-table for G , constructed from Step 2 through Step 1.

The entry with * denotes one entered in Step 2, and the entries without * are entered in Step 1.

	S	A	a	b	\$
S					[\$]0(0)
A			[a]1(1) [a]3(3)	[b]2(2)	
a			[a]1(1)		
b			[b]2(2)		
\$					[\$]0(0)

図 4 Step 3 終了後の G_3 の PF 表
Fig. 4 PF-table for G_3 constructed from Step 3.

	S	A
S		
A		
a	[a]3(3)	
b		
\$		

図 5 Step 5 終了後の G_5 の END-FOLLOW 表
Fig. 5 END-FOLLOW-table for G_5 constructed from Step 5 through Step 4.

	S	A	a	b	\$
S					[\$]0(0)
A			[a]1(1) [a]3(3)	[b]2(2)	
a			[a]1(1)		[a]3(3)*
b			[b]2(2)		
\$					[\$]0(0)

図 6 Step 6 終了後の G_6 の PF 表
*印は, Step 6 で記入されたもの。
Fig. 6 PF-table for G_6 constructed from Step 6 through Step 3.
The entry with * denotes one entered in Step 6.

	S	A	a	b	\$
S		[]3(3)	[]1(1) []3(3)	[]2(2) []3(3)	
A				[]4(4)	
a		[]1(1)	[]1(1)	[]1(1)**	[\$]3(0)*
b		[]2(2)	[]3(3) [a]4(1)* [a]4(3)*	[]2(2) [b]4(2)*	
\$					

図 7 Step 8 終了後の G_8 の FIRST 表
*印は, Step 7 で記入されたもの。 **は Step 8 で記入されたもの。
Fig. 7 FIRST-table for G_8 constructed from Step 8 through Step 1.
The entries with * denote ones entered in Step 7, and the entry with ** is entered in Step 8.

- 2. $S \rightarrow bAba$ 5. $A \rightarrow \epsilon$
- 3. $S \rightarrow Aa$

この文法に対する Q, R はそれぞれ以下のとおりである。

$$Q = \{(A, 5)\}, R = \{(S, a, 3), (A, b, 4)\}$$

	S	A	a	b	\$
S					[\$]0(0)
A			[a]1(1) [a]3(3)	[b]2(2)	
a			[a]1(1)		[a]3(3)
b			[b]2(2)		
\$					[\$]0(0)

図 8 Step 9 終了後の G_9 の PF 表
Fig. 8 PF-table for G_9 constructed from Step 9 through Step 3.

	S	A	a	b	\$
S		[]3(3)	[]1(1) []3(3)	[]2(2) []3(3)	
A			[a]5(*)	[]4(4) [b]5(*)	
a		[]1(1)	[]1(1) [a]5(*)	[]1(1)	[\$]3(0) [a]5(*)
b		[]2(2)	[]3(3) [a]4(1) [a]4(3) [b]5(*)	[]2(2) [b]4(2)	
\$					

図 9 Step 10 終了後の G_{10} の FIRST 表
*印は, Step 10 で記入されたもの。
Fig. 9 FIRST-table for G_{10} constructed from Step 10 through Step 1.
The entries with * denote one entered in Step 10.

	a	b	\$
S	[]1 []3	[]2 []3	
A	[a]5	[]4 [b]5	
a	[]1 [a]5	[]1	[\$]3 [a]5
b	[]3 [a]4 [b]5	[]2 [b]4	
\$			

図 10 G_{10} に対する解析表
生成規則番号はすべて π 型になっている。
Fig. 10 Parsing-table for G_{10} .
 π type production indices are transformed to τ type.

8. 性能評価

アルゴリズム θ の性能を評価するため, Aho らの作成法²⁾と, 以下の条件の下で比較検討をした。

- 使用機種: SUN-3 モデル 60 M
- 使用 OS および言語: UNIX/PASCAL
- 対象言語: • PASCAL-²⁾ (pascal minus)
- ISO PASCAL⁸⁾

プログラムは個人差をできるだけ少なくするため, すべて同一人が作成した。PASCAL- 文法は ISO PASCAL のサブ・セットである。

表作成の速さ

PASCAL- 文法ならびに ISO PASCAL 文法を準

LL(2) 文法に書き換え, それらに対してアルゴリズム θ と Aho らの方法により, それぞれ5回ずつ解析表作成を行い, 平均の所要時間を求めたのが表 1, 2 である. いずれの場合も, 筆者らの方法では Aho らの約 1/10 以下である. ただし, 表 2 の Aho らの場合は, 記憶領域の関係で解析表がメモリ上に展開できず, 解析表の各エントリに書き込むためのデータを求めるところまでの時間である. したがって, 表 2 の Aho らの場合, 実際にはさらに表に書き込むための時間が加わる.

集合 Q, R の計算に要する時間は, 表 1, 2 に示すように全体の計算所要時間の 0.8~2.1%程度であって十分に無視できる範囲である.

また, 集合 Q を求めるアルゴリズムは文献 6) に示すアルゴリズムに準じた. 集合 R を求めるアルゴリズムは Q を求めるアルゴリズムの延長線上に容易に求めることができる.

記憶領域

記憶領域はコード部分と表部分に大きく分けられる. 前者は対象となる文法の影響を受けないが, 後者は大きく影響を受ける. 表には解析表, 生成規則表のほかに, 解析表を生成する時に必要となる一時的な諸表 (筆者らの場合でいえば, FIRST 表の一部分, PF 表, END-FOLLOW 表など) がある. 解析表の大きさは筆者らの場合, 対象となる文法の非終端記号と終端記号の個数により決定されるが, Aho らの場合には生成される新しい非終端記号, つまり T_i の個数により決定される (図 11, 12). 一例として, PASCAL- と ISO PASCAL の各々の場合の T_i , 生成規則の増え方を表 3 に示す. ともに約 8~16 倍に増えている. 表 4 は PASCAL- と ISO PASCAL の各々に対する記憶領域の実測値である. 筆者らの解析表では, 1つのエントリに複数個の τ 型生成規則番号が入る場合があるが, そのエントリに入るべき生成

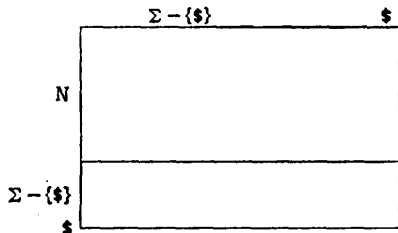


図 11 筆者らの解析表の構造

Fig. 11 The structure of authors' parsing-table.

表 1 PASCAL- に対する解析表の作成時間 (実測値) (単位: 秒)

Table 1 The experimental construction times of parsing-tables for PASCAL- (time unit: sec).

アルゴリズム θ (筆者ら) X		Aho ら Y		X: Y
集合 Q の計算	0.26	集合 FIRST の計算	0.24	
集合 R の計算		T_i の計算		
Step 1	11.61	構文則の書き換え	128.54	
Step 11		表作成		
合計	11.87	合計	128.78	1: 11

表 2 ISO PASCAL に対する解析表の作成時間 (実測値) (単位: 秒)

Table 2 The experimental construction times of parsing-tables for ISO PASCAL (time unit: sec).

アルゴリズム θ (筆者ら) X		Aho ら Y		X: Y
集合 Q の計算	1.47	集合 FIRST の計算	1.38	
集合 R の計算		T_i の計算		
Step 1	190.69	構文則の書き換え	2388.55	
Step 11		表作成		
合計	192.16	合計	2389.93	1: 12

規則番号は順次ポインタで結ぶことでプログラム上は実現している. 表 4 の I にはこうした領域も含まれている. また, I には解析表, 生成規則表だけではな

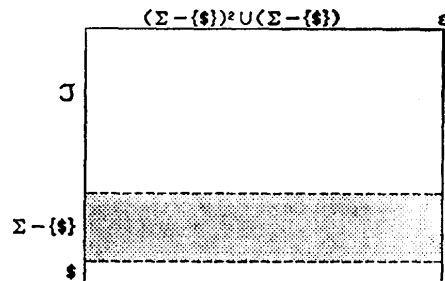


図 12 Aho らの解析表の構造

性能評価では, 記憶容量の関係で斜線部は生成しなかったが, この部分は生成規則による置換えはない (スタックからの降ろし, エラー発見, など) ので, 解析ルーチンで補うことができる. $(\Sigma-\{\$\})^2$ は長さ 2 の終端記号列の集合を表す.

Fig. 12 The structure of Aho and Ullman's parsing-table.

We can't construct the oblique portion of the Fig. 12 as an array because authors' computer does not have enough memory size, but we can code this part of tables in the parsing steps. $(\Sigma-\{\$\})^2$ denotes a set of terminal string of length 2.

表 3 生成規則数, 非終端記号数等の変化 (PASCAL- および ISO PASCAL の場合) (実測値)
 Table 3 Actual numbers of production rules, terminals, and non-terminals resulted from rewriting Wirth's production rules for applying Aho's algorithm (in the cases of PASCAL- and ISO PASCAL).

	PASCAL- 文法				ISO PASCAL 文法			
	準 LL(2) X_1	筆者ら	Aho ら Y_1	$X_1: Y_1$	準 LL(2) X_1	筆者ら	Aho ら Y_1	$X_1: Y_1$
生成規則数	98	98	804	1: 8	247	247	3997	1: 16
非終端記号数	54	54	427	1: 8	151	151	2369	1: 16
終端記号数	46	46	46		61	61	61	

表 4 PASCAL- および ISO PASCAL に対する解析表作成時の記憶領域 (実測値) (単位: バイト)
 Table 4 Actual memory capacity occupied in the construction times of parsing-tables for PASCAL- and ISO PASCAL (memory unit: byte).

	PASCAL- 文法			ISO PASCAL 文法		
	筆者ら X_1	Aho ら Y_1	$X_1: Y_1$	筆者ら X_1	Aho ら Y_1	$X_1: Y_1$
コード部分 (A)	65560	60896	1: 0.93	65560	60896	1: 0.93
表 { I. 全体* (B)	161552**	7530080		558896**	71947144	
	60228	7439496	1: 124	179732	71805620	1: 400
II. 解析表+生成規則表***						
合計 (A)+(B)	227112	7590976	1: 33	624456	72008040	1: 115

* 全体には, 解析表, 生成規則表の領域のほかに, 表作成時に必要な諸表のための領域も含まれる.

** 筆者らの解析表では, 1つのエントリに複数個の τ 型生成規則番号が入る場合がある. そうした場合に, プログラム上では各生成規則番号をポインタで結ぶことにより実現している. この値にはこうした領域も含まれている.

*** 解析時に用いられる解析表と生成規則表のみの値を示す.

* The values in the row consist of memories for the parsing-table and production-rules, and all the tables required for constructing a parsing-table.

** In authors' parsing-table, it is possible to enter multiple τ type production indices into a single entry of tables. In such a case, the linked-list technique is applied in authors' program. The values involve memories for these linked-lists.

*** The values in the row consist of only the memories for parsing-table and production-rules used for parsing.

く, 解析表を生成するのに必要なすべての諸表も含まれている. 表 4 からわかるように, 解析時に要する解析表と生成規則表のための記憶領域に限っていえば, 筆者らのものは Aho らの約 1/120~1/400 程度である. こうした傾向は他のプログラミング言語についてもいえよう.

表 4 によると, コード部分は筆者らがやや大きくなるが, 全体に占めるコード部分の割合は 10~29% 程度なのであまり問題にはならない.

Aho らの場合, 文法が大きくなると書き換えられる生成規則も多くなるので, そのための表も大きくなり, さらに図 12 からわかるように解析表も幾何級数的に大きくなる. 解析時にはこれがほぼそのまま持ち込まれる. 筆者らの場合, 解析表となって持ち込まれるのは FIRST 表の $(N \cup \Sigma) \times \Sigma$ 部分と, そこを根としたポインタで結ばれた部分と, もとの生成規則表のみである.

9. むすび

はじめに述べたように LL(k) 文法, $k \geq 2$ に対する解析表作成は多くの困難を伴っている. 1つには作成法そのものの難しさであり, もう1つは実現する上での記憶領域の問題である. 本論文では, LL(k) 文法にわずかに制限を加えた準 LL(k) 文法¹⁾ の, $k=2$ の場合につき, その解析表の作成アルゴリズムを示し, その正当性を証明した.

さらに, Aho らの方法との性能を具体例にもとづいて比較し, 提案する作成法が表作成のための所要時間において約 1/10, 解析表と生成規則表のみの記憶領域では 1/120~1/400 程度ですむことがわかった.

また, Aho らの方法ではその性質上, 生成規則を書き換える必要があり, その増え方はもとの生成規則の 8~16 倍に増えることがわかった. これらの比較は, アルゴリズムの性質から他の具体例に対しても一般的な傾向をもつものと推定できる.

Aho らの方法が準 LL(2) 文法のみならず, LL(2) 文法にも適用できるという利点もつが, アルゴリズムの実現の容易さ, LL(2) 文法のかなりのものが準 LL(2) 文法に書き換えが可能であると思われるので, そのことを考慮に入れると本論文で提案する解析表作成法は実用上も十分に有効であると考え.

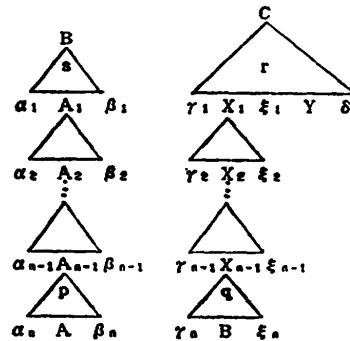
なお, 本論文で提案した方法により生成された解析表にもとづく解析法については, 文献 1) に述べられている.

謝辞 本研究をすすめるにあたり, 長年にわたりご指導いただいた東京理科大学・井上謙蔵教授に心から感謝の意を表したい. また, 査読者からの有益なご助言に謝意を表す.

参 考 文 献

- 1) 吉田, 竹内: 準 LL(2) 文法に対する解析表の構造と解析アルゴリズム (情報処理学会論文誌投稿中).
- 2) Aho, A. V. and Ullman, J. D.: *The Theory of Parsing, Translation, and Compiling*, Vol. 1, pp. 348-356, Prentice-Hall (1972).
- 3) Lewis, P. M. and Stearns, R. E.: *Compiler Design Theory*, pp. 262-276, Addison-Wesley (1976).
- 4) 吉田, 竹内: 生成規則番号を用いた LL(1) 解析テーブル作成のアルゴリズム, 情報処理学会論文誌, Vol. 27, No. 11, pp. 1095-1105 (1986).
- 5) Backhouse, R. C.: *Syntax of Programming Languages (Theory and Practice)*, pp. 92-152, Prentice-Hall (1979).
- 6) Yoshida, K. and Takeuchi, Y.: Some Properties of an Algorithm for Constructing LL(1) Parsing-Tables Using Production Indices, *J. Inf. Process.*, Vol. 11, No. 4, pp. 258-262 (1988).

- 7) Hansen, P. B.: *Brinch Hansen on Pascal Compilers*, pp. 15-16, Prentice-Hall (1985).
- 8) Jensen, K. and Wirth, N.: *PASCAL-User Manual and Report (3rd ed.)*, ISO PASCAL standard, pp. 215-220, Springer-Verlag (1975).



付図 $PL_1(A, Y)$, $PL_2(Y, W)$ を求める.

$EF(A, B) \ni [X]p(s)$, $PL_1(B, Y) \ni [Z]q(r)$ ならば, $PL_1(A, Y)$ に $[V]p(r)$ を加える. さらに, $Y \in \Sigma$ ならば, $PL_2(Y, W)$ に $[V]p(r)$ を加える. ただし, $W = {}^{(1)}\delta, \beta_n\beta_{n-1}\dots\beta_1 \ni \epsilon, \xi_n\xi_{n-1}\dots\xi_1 \ni \epsilon$. X は記号列 $\beta_n\beta_{n-1}\dots\beta_1$ の先頭一記号, Z は記号列 $\xi_n\xi_{n-1}\dots\xi_1$ の先頭一記号. V は $X \neq \epsilon$ のとき $V = X$, $X = \epsilon$ のとき $V = Z$.

Fig. appendix In the case of Fig. appendix, we can find $PL_1(A, Y)$ and $PL_2(Y, W)$. That is, if $EF(A, B) \ni [X]p(s)$ and $PL_1(B, Y) \ni [Z]q(r)$, then we add $[V]p(r)$ to $PL_1(A, Y)$. In addition, if $Y \in \Sigma$ then we add $[V]p(r)$ to $PL_2(Y, W)$, where $W = {}^{(1)}\delta, \beta_n\beta_{n-1}\dots\beta_1 \ni \epsilon$, and $\xi_n\xi_{n-1}\dots\xi_1 \ni \epsilon$. X denotes the first symbol of a string $\beta_n\beta_{n-1}\dots\beta_1$, and Z the first symbol of a string $\xi_n\xi_{n-1}\dots\xi_1$. If $X \neq \epsilon$ then $V = X$, otherwise $V = Z$.

(平成元年 10 月 23 日受付)
(平成 2 年 3 月 6 日採録)



吉田 敬一 (正会員)

昭和 13 年生. 昭和 37 年法政大学工学部電気工学科卒業. 昭和 43 年法政大学大学院工学研究科修士課程中退. 日本電気(株)を経て, 現在, 静岡大学工業短期大学部・情報工学科助教授. コンパイラの自動生成, 情報処理教育のあり方, 人工知能に関心を持っている. 著書「コンパイラの理論と応用」(共訳, 学研), 「コンピュータ・サイエンスのための言語理論入門」(共訳, 共立出版) など. 日本ソフトウェア科学会, ACM, ACS (オーストラリア) 各会員.



竹内 淑子 (正会員)

昭和 29 年生. 昭和 55 年静岡大学工業短期大学部・情報工学科卒業. (株)東京システム技研(コンパイラ開発担当)を経て, 現在, 浜松職業訓練短期大学校情報処理科教官. コンパイラの自動生成に関心を持っている. 著書「BASIC プログラミングのすべて」(共訳, 一橋出版). 日本ソフトウェア科学会, IEEE 各会員.