数理解析研究所講究録
第 625 巻 1987 年 48-54

ＣＦＧにおける並列性：多ヘッドＣＦＧ，部分同期ＣＦＧ，および交代ＣＦＧ

東京女子大学（文理）　守屋悦朗　（ Etsuro Moriya ）

１．はじめに

本稿ではＣＦＧ（文脈自由文法）における並列的導出法のいくつかについて考察する．それらは，（１）プロダクションの適用位置を指定する”書き換えヘッド”を複数個持つ多ヘッドＣＦＧ，（２）非終端記号を同期的なものと非同期的なものとに分け，同期的な非終端記号については，導出途中の文形式において現れる全てに同一のプロダクションを適用して書き換えを行う部分同期ＣＦＧ，（３）非終端記号を存在的なものと全称的なものに分け，alternationの概念［１，２］を導入した交代ＣＦＧ，の３つである．

本稿では［８］の結果を一部拡張し，いくつかの新しい結果を加えた．また，本稿の第４節は［９］でさらに詳細に扱われている．

２．多ヘッドＣＦＧ

$\underline{G}=(N, \Sigma, \underline{P}, S)$をＣＦＧとする．ｋを正整数，Ｐを$\underline{P}^k \times \mathbb{Z}^k$の有限部分集合とするとき，$G=(N, \Sigma, P, S)$をｋヘッドＣＦＧという．Ｐの元は

$$(*) \qquad (X_1 \rightarrow x_1, \ldots, X_k \rightarrow x_k, d_1, \ldots, d_k)$$

という形の書き換え規則であり，$(X_1 \rightarrow x_1, \ldots, X_k \rightarrow x_k)$はプロダクション部分，$(d_1, \ldots, d_k)$はヘッド移動方向指定部分である．$\underline{G}$をＧの基礎ＣＦＧという．Ｇがλ-freeであるとは$\underline{G}$がλ-freeのことであり，Ｇが１方向型であるとは，Ｐが$\underline{P}^k \times \mathbb{N}^k$の有限部分集合となっていることである．ただし，$\mathbb{N}=\{0, 1, 2, \ldots\}$である．

$\{\Gamma_1, \ldots, \Gamma_k\} \cap (N \cup \Sigma)=\phi$ とする（記号$\Gamma_1, \ldots, \Gamma_k$はＧの書き換えヘッドと呼ばれる）．Ｇにおける様相とは，各Γ_i $(1 \leq i \leq k)$を丁度１つずつ含んでいるような$(N \cup \Sigma \cup \{\Gamma_1, \ldots, \Gamma_k\})^*$の元のことをいう．様相 a，ヘッド Γ_i に対して

$$a = b\Gamma_i c$$

とする．ｃ の中の最左の非終端記号をXとするときΓ_iはXを指しているといい，ｃ が非終端記号を含まないときΓ_iは使われていないという．

多ヘッドＣＦＧにおける導出を次のように定義する．様相 a において，全てのヘッドが使わ

れており，かつ，ヘッド $\Gamma_i (1 \le i \le k)$ は非終端記号 X_i を指しているとする．もし Γ_i と Γ_j が同じ非終端記号を指しているならば，すなわち，

$$a = a'\Gamma_i a''\Gamma_j a''' \quad かつ \quad a'' \in (\Sigma \cup \{\Gamma_1, \ldots, \Gamma_k\})^*$$

ならば，(*)のプロダクション部分の第 i プロダクション $A_i \to x_i$ と第 j プロダクション $A_j \to x_j$ とが一致する場合のみプロダクション(*)は a に<u>適用でき</u>，第 i ヘッド Γ_i が指している非終端記号は(*)の第 i プロダクションによって x_i に書き換えられる．この書き換えは全てのヘッドについて同時に行われる．書き換えを行った後，第 i ヘッドは自分の右側に在る d_i 個の非終端記号を読み飛ばすように右に移動する（$d_i<0$ のときは，$-d_i$ 個左に移動する）．この結果，a から b が得られるとき

$$a \Rightarrow b$$

と書く．$\Rightarrow^*$ を $\Rightarrow$ の反射推移閉包とする．

$$L(G) = \{w \in \Sigma^* \mid \Gamma_1 \ldots \Gamma_k S \Rightarrow^* w_0 \Gamma_{i1} w_1 \Gamma_{i2} \ldots w_{k-1} \Gamma_{ik} w_k,\ w = w_0 w_1 \ldots w_k,\ かつ\ (i_1\ i_2\ \ldots\ i_k)\ は\ (1\ 2\ \ldots\ k)\ の置換\}$$

をＧが生成する<u>ｋヘッドＣＦＬ</u>と呼ぶ．

多ヘッドＣＦＧは，simple matrix 文法［６］，matrix 文法［１０］，scattered context文法［５］などの自然な拡張になっていることに注意する．（matrix文法のプロダクション $[X_1 \to x_1, \ldots, X_n \to x_n]$ においては，どの X_i も $x_1, \ldots, x_{i-1}$ の中に現れないと仮定することができるので，ヘッド移動用のプロダクションを用意しておけば，このmatrixプロダクションは n 個のヘッドでシミュレートできる．）

<u>定理１．１方向型２ヘッドＣＦＬのクラスは帰納的可算集合のクラスと一致する．実際，基礎ＣＦＧを線形文法としてこのことがいえるが，基礎ＣＦＧを正規文法とすることは出来ない．</u>

<u>証明の方針</u>　任意の帰納的可算集合は２つの線形ＣＦＬのintersectionのhomomorphic imageとして表される［３］ことを用いる．２つのＣＦＬのintersectionは，２つのGreibach標準形のＣＦＧを２つのヘッドによって同時にシミュレートする．

<u>定理２．任意のｋに対して，２方向型 λ-free ｋヘッドＣＦＬは文脈依存言語である．（２方向型 λ-free多ヘッドＣＦＧによって生成されえないような文脈依存言語が存在するか否かはわかっていない．）</u>

３．部分同期ＣＦＧ

G = (N, Σ, P, S)をＣＦＧとする．非終端記号を<u>同期的非終端記号</u>と<u>非同期的非終端記号</u>に

分ける：

$$N = N_s \cup N_a, \quad N_s \cap N_a = \phi$$

左辺が同期的非終端記号（N_sの元）であるプロダクションを<u>同期プロダクション</u>，そうでないものを<u>非同期プロダクション</u>という．$G = (N_a, N_s, \Sigma, P, S)$ を<u>部分同期ＣＦＧ</u>という．

非同期プロダクションは，通常のＣＦＧのプロダクションと同様に適用され導出が行われる．この導出（<u>非同期導出</u>）を $\Rightarrow_a$ で表す．すなわち， $a \Rightarrow_a b$ となるのは， $a = a'Xa''$, $b = a'xa''$ となる非同期プロダクション $X \rightarrow x$ が存在するときである．これに対し，同期プロダクションは次のように”同期的に”適用され，導出が行われる（<u>同期導出</u>）．

$$a = a_0 X_1 a_1 \ldots X_n a_n, \quad a_i \in (N_a \cup \Sigma)^* \qquad (0 \leq i \leq n)$$

とし，各 $X_i \rightarrow x_i$ $(1 \leq i \leq n)$ を同期プロダクションとする．

$$b = a_0 x_1 a_1 \ldots x_n a_n$$

かつ

$$X_i = X_j \quad \text{ならば} \quad x_i = x_j$$

であるとき，$a \Rightarrow_s b$ と書く．

”同期導出”は次のように定義することも考えられる．すなわち，$a = a_0 A a_1 \ldots A a_n$ で，どの a_i $(0 \leq i \leq n)$ も A を含まないとするとき，$A \rightarrow x$ が同期プロダクションで $b = a_0 x a_1 \ldots x a_n$ であるならば $a \Rightarrow_t b$ であると定義する．しかし，次の補題に示すように，$\Rightarrow_s$ と $\Rightarrow_t$ とは本質的に同じものである．

<u>補題．同期導出 $a \Rightarrow_t^* b$ が成り立つ必要十分条件は，同期導出 $a \Rightarrow_s^* t$ が成り立つことである．</u>

$a \Rightarrow_a b$ または $a \Rightarrow_s b$ であるとき

$$a \Rightarrow b$$

と書く．

$$L(G) = \{ w \in \Sigma^* \mid S \Rightarrow^* w \}$$

をＧにより生成された<u>部分同期ＣＦＬ</u>と呼ぶ．

<u>定理３．Ｌが部分同期ＣＦＬならば，Ｌ－｛λ｝は文脈依存言語である．また，部分同期ＣＦＬでない文脈依存言語が存在する．</u>

部分同期ＣＦＧにおける<u>交代導出</u>を次のように定義する．

$$(**)\quad a = a_0 \Rightarrow^+_{Q1} a_1 \Rightarrow^+_{Q2} \ldots \Rightarrow^+_{Qn} a_n = b$$

とする．ただし，どのｉについても Q_i はｓまたはａであり，かつ，$Q_i \neq Q_{i+1}$ とする．また，$\Rightarrow^+_{Qi}$ は $\Rightarrow_{Qi}$ の推移閉包である．もし，Q_1=a ならば（＊＊）を Σ_nー導出，Q_1=s ならば（＊＊）を Π_nー導出と呼ぶ．

任意の $w \in L(G)$ に対してwを生成する Σ_mー導出（$m \leq n$）が存在するならばＧは Σ_nーＰＳＣＦＧと呼ばれ，Σ_nーＰＳＣＦＧによって生成される言語を Σ_nーＰＳＣＦＬと呼ぶ．Π_nーＰＳＣＦＧ，Π_nーＰＳＣＦＬも同様に定義される．

Σ_nーＰＳＣＦＬ，Π_nーＰＳＣＦＬのクラスをそれぞれ

$$\Sigma_n^{PSCF},\quad \Pi_n^{PSCF}$$

で表す．定義から，Σ_1^{PSCF} はＣＦＬのクラスそのものである．

定理４［１１］．（１）Σ_1^{PSCF} と Π_1^{PSCF} は比較不能である．
（２）$\Sigma_1^{PSCF} \cap \Pi_1^{PSCF}$ はderivation bounded言語［４］のクラスと一致する．

$$\Sigma_n^{PSCF} \cup \Pi_n^{PSCF} \subseteq \Sigma_{n+1}^{PSCF} \cap \Pi_{n+1}^{PSCF}$$

が成り立つことは容易に証明できる．また，n=1 のときこれは真の包含関係であるが，一般のｎについてはまだ証明されていない．

$\mathfrak{L}_1$，$\mathfrak{L}_2$ を言語のクラスとするとき，

$$\mathfrak{L}_1\ \sigma\ \mathfrak{L}_2 = \{\ \tau(L_1)\ \mid\ L_1 \subseteq \Sigma^* \text{ は } \mathfrak{L}_1 \text{ の元，かつ，}$$
$$\tau \text{ はどの } a \in \Sigma \text{ についても } \tau(a) \in \mathfrak{L}_2 \text{ であるような代入}\}$$

$$\mathfrak{L}_1\ \pi\ \mathfrak{L}_2 = \{\ \mathfrak{R}(L_1)\ \mid\ L_1 \subseteq \Sigma^* \text{ は } \mathfrak{L}_1 \text{ の元，かつ，}$$
$$\mathfrak{R} \text{ はどの } a \in \Sigma \text{ についても } \mathfrak{R}(a) \in \mathfrak{L}_2 \text{ であるような準同型写像のクラス}\}$$

と定義する．ただし，$L \subseteq \Sigma^*$ に対して，

$$\mathfrak{R}(L) = \{\ h(L)\ \mid\ h \in \mathfrak{R}\ \}$$

である．とくに，言語のクラス $\mathfrak{L}$ に対して，

$$\sigma_1(\mathfrak{L}) = \mathfrak{L}\ \sigma\ \Sigma_1^{PSCF},\qquad \pi_1(\mathfrak{L}) = \mathfrak{L}\ \pi\ \Pi_1^{PSCF}$$

と定義し，$\mathfrak{L}$ を含み σ_1 および π_1 で閉じている最小のクラスを $\mathfrak{L}$ の（σ_1，π_1）閉包という．

<u>定理５．ｎを任意の自然数とする．</u>
<u>（１）ｎが偶数ならば，</u>

$$\Sigma_{n+1}^{PSCF} = \sigma_1(\Sigma_n^{PSCF}),\quad \Pi_{n+1}^{PSCF} = \pi_1(\Pi_n^{PSCF}).$$

<u>（２）ｎが奇数ならば，</u>

$$\Sigma_{n+1}^{PSCF} = \pi_1(\Sigma_n^{PSCF}),\quad \Pi_{n+1}^{PSCF} = \sigma_1(\Pi_n^{PSCF}).$$

<u>系．部分同期ＣＦＬのクラスは，有限言語のクラスの（σ_1，π_1)-閉包に一致する．</u>

４．交代ＣＦＧ

<u>交代ＣＦＧ</u>とは５つ組 $G=(N, U, \Sigma, P, S)$ のことをいう．ここに，(N, Σ, P, S) はＣＦＧであり，ＵはＮの部分集合である．Ｎの元を<u>存在的非終端記号</u>，Ｕの元を<u>全称的非終端記号</u>という．

Ｇにおける導出を次のように定義する．Ｔを各頂点に $(N \cup \Sigma)^*$ の元がラベル付けされた木とする．以下，頂点 π のラベルを $\iota(\pi)$ で表すことにする．Ｔのどの内部頂点 π についても次の（ｉ）または（ｉｉ）が成り立っているとする．

（ｉ）　$\iota(\pi)=yXz$，$X\in N-U$，$X\to x\in P$，かつ，π は丁度１つの子供 ρ を持ち，$\iota(\rho)=yxz$ である．π を<u>存在的頂点</u>と呼ぶ．

（ｉｉ）　$\iota(\pi)=yXz$，$X\in U$，P の X-プロダクション（左辺が X であるようなプロダクションのこと）は $X\to x_1 \mid x_2 \mid \ldots \mid x_k$，かつ，$\pi$ は丁度ｋ個の子供 $\rho_1, \ldots, \rho_k$ を持ち，$\iota(\rho_i)=yx_iz$ である．π を<u>全称的頂点</u>と呼ぶ．

このとき，ＴをＧにおける ι(Ｔの根)からの<u>導出</u>という．導出Ｔの各頂点 π の<u>値</u> value(π) を次のように定義する．

$$\mathrm{value}(\pi)=\begin{cases}\iota(\pi) & \pi\text{がＴの葉であるとき}\\ \mathrm{value}(\rho) & \pi\text{が存在的頂点で}\rho\text{がその子供のとき}\\ w & \pi\text{が全称的頂点で，そのどの子供}\rho\text{についても}\ \mathrm{value}(\rho)=w\text{であるとき}\\ \text{定義されない} & \text{その他}\end{cases}$$

Ｔの根の値を<u>Ｔの値</u>と定義し，やはり value(T) と表すことにする．value(T) が定義されている導出のことを<u>正当な導出</u>と呼ぶ．

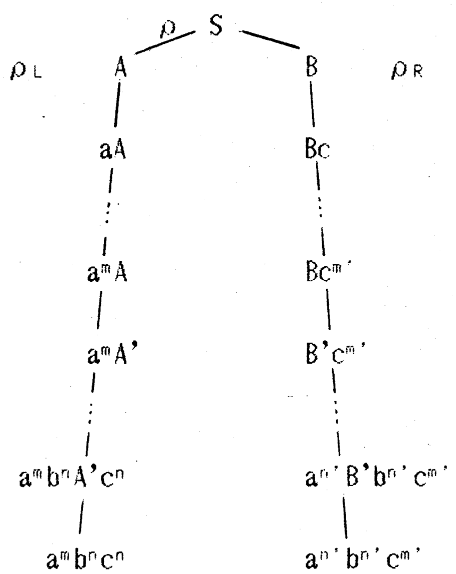
$$L(G)=\{\mathrm{value}(T)\in\Sigma^* \mid T\ は\ G\ における\ S\ からの正当な導出\}$$

をＧによって生成された<u>交代ＣＦＬ</u>という．

<u>例</u>　交代ＣＦＧ　G = ({S, A, A', B, B'}, {S}, {a, b, c}, P, S) を考える．但し，Pは次のプロダクションから成るとする：

$$S \to A \mid B$$
$$A \to aA \mid A',\ A' \to bA'c \mid \lambda$$
$$B \to Bc \mid B',\ B' \to aB'b \mid \lambda$$

Ｇにおける導出のうちで，value(T) が終端記号列となるものは必ず次の形をしている：



value(ρ_L) = $a^m b^n c^n$, value(ρ_R) = $a^{n'} b^{n'} c^{m'}$ であり，Ｓは全称的非終端記号だから，value (ρ) が定義されＴが正当な導出となるためには，$m=n=n'=m'$ となることが必要である．よって，$L(G) = \{ a^n b^n c^n \mid n \geq 0 \}$ である．

交代ＣＦＧと交代プッシュダウンオートマトン(alternating pda)とは，次のような深い関係にある．

<u>定理６．交代ＣＦＧ Ｇ に対して L=L(G)　<==>　交代プッシュダウンオートマトン Ｍ に対して L=L(M)．</u>

<u>証明の方針</u>　ＣＦＧ　<=>　ＰＤＡ　の証明をほぼなぞる．交代ＣＦＧにおける<u>最左導出</u>とは，根から葉へ至る全ての道がＣＦＧにおける導出とみたとき最左導出となっているような導出のことであると定義する．交代ＣＦＧ Ｇ においても，任意の$w \in L(G)$に対して，ｗを生成する最左導出が存在することがいえる．このことに注意すれば，ＣＦＧ　<=>　ＰＤＡ　の証明と同様に，交代ＣＦＧにおける最左導出と交代プッシュダウンオートマトンにおけるＩＤ（入力とスタックの内容）とをうまく対応させることが出来る．

交代オートマトンと同様に，交代ＣＦＧにおける導出でも，全称的導出部分と存在的導出部分との交代回数で導出を分類することが出来る．

Ｇを交代ＣＦＧ，ｋを正整数とする．Ｇにおける導出ＴがΣ_kー導出（Π_kー導出）といわれるのは，Ｔの根が存在的頂点（全称的頂点）であり，かつ，Ｔにおける根から葉へ至るどの道においても存在的頂点／全称的頂点から全称的頂点／存在的頂点への交代が高々ｋ－１回しか起こっていないときである．任意の $w \in L(G)$ に対して，ｗを生成するΣ_kー導出が存在するときＧはΣ_kー交代ＣＦＧといい，$L(G)$ をΣ_kー交代ＣＦＬという．Π_kー交代ＣＦＧ，Π_kー交代ＣＦＬも同様に定義される．Σ_kー交代ＣＦＬのクラス，Π_kー交代ＣＦＬのクラスをそれぞれ

$$\Sigma_k^{ACF}, \quad \Pi_k^{ACF}$$

で表す．次の関係が成り立つことは容易に示される：

$$\Sigma_n^{ACF} \cup \Pi_n^{ACF} \subsetneq \Sigma_{n+1}^{ACF} \cap \Pi_{n+1}^{ACF} \quad (n \geq 1).$$

定理７．

$$\begin{array}{ccccccccccc}
\Pi_1^{ACF}\subsetneq & \Sigma_2^{ACF} & \subsetneq & \Pi_3^{ACF} & \subsetneq\cdots\subsetneq & \Sigma_{2n}^{ACF} & \subsetneq & \Pi_{2n+1}^{ACF} & \subsetneq & \Sigma_{2n+2}^{ACF} & \subsetneq\cdots \\
 & \| & & \| & & \| & & \| & & \| & \\
 & \Sigma_1^{ACF} & \subsetneq & \Pi_2^{ACF} & \subsetneq\cdots\subsetneq & \Sigma_{2n-1}^{ACF} & \subsetneq & \Pi_{2n}^{ACF} & \subsetneq & \Sigma_{2n+1}^{ACF} & \subsetneq\cdots.
\end{array}$$

特に，Π_1^{ACF} は空集合およびsingletonから成るクラスである．

参考文献

1. A.K.Chandra, D.C.Kozen and L.J.Stockmeyer, Alternation, JACM 28 (1981), 114-133.
2. A.K.Chandra and L.J.Stockmeyewr, Alternation, Proc. 17th FOCS (1976), 88-108.
3. S.Ginsburg, S.Greibach and M.Harrison, One-way stack automata, JACM 14 (1967), 389-418.
4. S.Ginsburg and E.H.Spanier, Derivation bounded languages, JCSS 2 (1968), 228-250.
5. S.A.Greibach and J.E.Hopcroft, Scattered context grammars, JCSS 3 (1969), 233-247.
6. O.H.Ibarra, Simple matrix languages, Inform. Contr. 17 (1970), 359-394.
7. R.Ladner, R.J.Lipton and L.J.Stockmeyer, Alternating pushdown and stack automata, SIAM J. Comput. 13 (1984), 135-155.
8. E.Moriya, On parallelism in context-free grammars, to appear in North-Holland Mathematics Series.
9. E.Moriya, A grammatical characterization of alternating pushdown automata, in preparation.
10. A.Salomaa, Formal Languages, Academic Press, New York, 1973.
11. P.Siromoney and K.Krithivasan, Parallel context-free languages, Inform. Contr. 24 (1974), 155-162.