

高々スター次数2の拡張正規表現

Extended Regular Expressions of Star Degree At Most Two

劉 信根

橋口 攻三郎

Heekeun YOO and Kosaburo HASHIGUCHI

豊橋技術科学大学

Toyohashi University of Technology

1. まえがき

正規表現は理論的な分野ばかりでなく、応用的な分野においても重要である。特に、Kleene-閉包は重要な役割を果たす。文脈自由型言語を表現する非終端記号や複雑な演算子を持たない表現が存在すれば、便利であろう。

本研究では、次数2のスター演算の概念を提案し、和集合、連接、Kleene-閉包、次数2のスター演算に関する有限言語族の閉包、すなわち、アルファベット Σ 上の高々スター次数2の拡張正規言語のクラス($ERL(2, \Sigma)$)のいくつかの性質を明らかにする。さらに、 Σ 上の completely linearly nested context-free language のクラス($CLNCFL(\Sigma)$)を紹介し、これら2つの言語のクラスが等しいことを示す。また、 $ERL(2, \Sigma)$ のその他の性質について述べる。例えば、(1) $ERL(2, \Sigma)$ と線形言語の族との間には包含関係は成り立たない、(2) $ERL(2, \Sigma)$ と決定性文脈自由型言語の族との間には包含関係は成り立たない等を示す。

本論文の2節では、準備として、基本的な定義を述べる。3節では、 $ERL(2, \Sigma)$ と $CLNCFL(\Sigma)$ が等しいことを証明する。4節では、 $ERL(2, \Sigma)$ の基本性質について述べる。5節では、高々スター次数2の λ -free拡張正規表現の一つの重要な性質について述べる。6節では、 $ERL(2, \Sigma)$ と文脈自由型言語族の他の部分族との間の包含問題について述べる。

2. 準備

Σ は空でない有限アルファベットの集合である。 λ は空系列である。 ϕ は空集合である。任意の $w \in \Sigma^*$ に対して $l(w)$ は w の長さで、 w^R は w の逆系列である。 $L \subset \Sigma^*$ に対して $L^R = \{w^R \mid w \in L\}$ である。任意のアルファベット V, V' と $w \in V^*$ に対して $V'(w)$ は w に現われる V' の記号の集合である。任意の $A \in V$ に対して $\#_A(w)$ は w に現われる A の個数である。集合 B に対して $\#B$ は B の濃度である。

[定義 2.1]

$ERE(2, \Sigma)$ は Σ 上の高々スター次数2の拡張正規表現のクラスであり、次

のように帰納的に定義される。

(1) $\lambda, \phi, a \in \text{ERE}(2, \Sigma)$ (但し $a \in \Sigma$) .

(2) もし $n \geq 1$ で, $E_0, E_{11}, E_{12}, \dots, E_{1n}, E_{21}, E_{22}, \dots, E_{2n} \in \text{ERE}(2, \Sigma)$ であれば,
 $E_{11} \cup E_{12}, E_{11}E_{12}, (E_0)^*, (E_{11}, \dots, E_{1n}, E_0, E_{21}, \dots, E_{2n})^{**} \in \text{ERE}(2, \Sigma)$.

[注意 2.1]

表現を明瞭にするため $(E_{11}, \dots, E_{1n}, E_0, E_{21}, \dots, E_{2n})^{**}$ の代わりにしばしば

$\left(\left[\begin{array}{c} E_{11} \\ \vdots \\ E_{1n} \end{array} \right], E_0, \left[\begin{array}{c} E_{21} \\ \vdots \\ E_{2n} \end{array} \right] \right)^{**}$ の表現が導入される。

[定義 2.2]

$E \in \text{ERE}(2, \Sigma)$ が表わす各々の言語 $|E|$ は次のように帰納的に定義される。

(1) $|\lambda| = \{\lambda\}, |\phi| = \phi, |a| = \{a\}$ (但し $a \in \Sigma$) .

(2) $|E_1 \cup E_2| = |E_1| \cup |E_2|,$

$|E_1 E_2| = |E_1| |E_2| = \{vw \mid v \in |E_1|, w \in |E_2|\},$

$|(E)^*| = |E|^* = \bigcup_{i=0}^{\infty} |E|^i,$

$\left| \left(\left[\begin{array}{c} E_{11} \\ \vdots \\ E_{1n} \end{array} \right], E_0, \left[\begin{array}{c} E_{21} \\ \vdots \\ E_{2n} \end{array} \right] \right)^{**} \right|$

$= |E_0| \cup \left(\bigcup_{i=1}^{\infty} \left(\bigcup_{1 \leq j_1, \dots, j_i \leq n} |E_{1j_1}| \dots |E_{1j_i}| |E_0| |E_{2j_1}| \dots |E_{2j_i}| \right) \right).$

$E \in \text{ERE}(2, \Sigma)$ に対して $|E|$ は Σ 上の高々スター次数 2 の拡張正規言語と呼ばれる。

[定義 2.3]

$\text{ERL}(2, \Sigma) = \{|E| \mid E \in \text{ERE}(2, \Sigma)\}.$

[例 2.1]

(1) $|(a, \lambda, b)^{**}| = \{a^n b^n \mid n \geq 0\}$

(2) $\left| \left(\left[\begin{array}{c} a \\ b \end{array} \right], c, \left[\begin{array}{c} a \\ b \end{array} \right] \right)^{**} \right| = \{wcw^R \mid w \in \{a, b\}^*\}.$

文脈自由文法 (CFG) G は (V, Σ, P, S) の四つの組で示される。 V は変数の有限集合, Σ は終端記号, P は生成規則の有限集合, そして S は開始記号である。文法 G における生成規則は一般に $A \rightarrow W$ (但し $A \in V, W \in (V \cup \Sigma)^*$) の形をしている。 $L(G)$ は文法 G によって生成される文脈自由型言語 (CFL) である。

[注意 2.2]

$1 \leq i \leq m$ に対して $P = P_1 \cup \dots \cup P_m$ かつ $P_i = \{A_i \rightarrow W_{ij} \mid 1 \leq j \leq k_i\}$ の時しばしば次のように書く ;

$P: A_1 \rightarrow W_{11} | W_{12} | \dots | W_{1k_1}$

\vdots

$A_m \rightarrow W_{m1} | W_{m2} | \dots | W_{mk_m}$

[定義 2.4]

完全線形入れ子文脈自由文法 (clncfg) は次の (1)-(2) を満たす cfg $G=(V, \Sigma, P, S)$ である。

- (1) V は線形順序をもつ。すなわち, $V=\{A_1, \dots, A_n\}$ かつ $A_1 > A_2 > \dots > A_n$,
 (2) 任意の i ($1 \leq i \leq n$), $W \in (V \cup \Sigma)^*$ に対して, $A_i \rightarrow W \in P$ であれば,
 $\#_{A_i}(W) \leq 1$ かつ $V(W) \subset \{A_i, A_{i+1}, \dots, A_n\}$ である。

任意の clncfg G に対して $L(G)$ は完全線形入れ子文脈自由型言語 (clncfl) と呼ばれる。CLNCFG(Σ) は Σ 上の clncfg の集合を表わす。

[定義 2.5]

CLNCFL(Σ) = $\{L(G) \subset \Sigma^* \mid G \in \text{CLNCFG}(\Sigma)\}$.

[命題 2.1] 任意の $E \in \text{ERE}(2, \Sigma)$ に対して

$$|E|^* = |(E, \lambda, \lambda)^{**}| = |(\lambda, \lambda, E)^{**}|.$$

(証明) 明らか。

[注意 2.3]

ERE($2, \Sigma$) の定義の中で Kleene-閉包は省略できるが, 表記の便宜のため Kleene-閉包を導入する。

3. ERL($2, \Sigma$)=CLNCFL(Σ) の証明

次の定理は本研究の主な結果の一つである。

[定理 3.1]

ERL($2, \Sigma$)=CLNCFL(Σ).

定理 3.1 は次の二つの補題による。

[補題 3.1]

任意の $E \in \text{ERE}(2, \Sigma)$ に対して, $|E| \in \text{CLNCFL}(\Sigma)$ である。

(証明) 証明は E に現れる演算子の個数に関する帰納法による。

基底段階: $|\lambda|, |\phi|, |a| \in \text{CLNCFL}(\Sigma)$ (但し $a \in \Sigma$) .

帰納段階: $E \in \text{ERE}(2, \Sigma)$ は次の 4 つのいずれかの形をしている。

$$(1) E=E_1 \cup E_2 \quad (2) E=E_1 E_2 \quad (3) E=(E_1)^*$$

$$(4) E = \left[\begin{array}{c} (E_{11}) \\ \vdots \\ (E_{1n}) \end{array} \right], E_0, \left[\begin{array}{c} (E_{21}) \\ \vdots \\ (E_{2n}) \end{array} \right]^{**}$$

(1) $E=E_1 \cup E_2$ の場合

帰納法の仮定により $|E_1|$ と $|E_2|$ を生成する clncfg $G_1=(V_1, \Sigma, P_1, S_1)$ と $G_2=(V_2, \Sigma, P_2, S_2)$ が存在する。そのとき $V_1 \cap V_2 = \phi$ としてよい。次の $G=(V, \Sigma, P, S)$ は E に対する条件を満たす。

$$(1.1) V=V_1 \cup V_2 \cup \{S\}$$

$$(1.2) P=P_1 \cup P_2 \cup S \rightarrow S_1 | S_2$$

(2) $E=E_1 E_2$ の場合

帰納法の仮定により $|E_1|$ と $|E_2|$ を生成する clncfg $G_1=(V_1, \Sigma, P_1, S_1)$ と

$G_2=(V_2, \Sigma, P_2, S_2)$ が存在する。そのとき $V_1 \cap V_2 = \emptyset$ としてよい。次の $G=(V, \Sigma, P, S)$ は E に対する条件を満たす。

$$(2.1) V=V_1 \cup V_2 \cup \{S\}$$

$$(2.2) P=P_1 \cup P_2 \cup S \rightarrow S_1 S_2$$

(3) $E=(E_1)^*$ の場合

帰納法の仮定により $|E_1|$ を生成する clncfg $G_1=(V_1, \Sigma, P_1, S_1)$ が存在する。次の $G=(V, \Sigma, P, S)$ は E に対する条件を満たす。

$$(3.1) V=V_1 \cup \{S\}$$

$$(3.2) P=P_1 \cup S \rightarrow \lambda |S_1| S_1 S$$

$$(4) E = \left(\begin{array}{c} [E_{11}] \\ \vdots \\ [E_{1n}] \end{array} \right), E_0, \left(\begin{array}{c} [E_{21}] \\ \vdots \\ [E_{2n}] \end{array} \right)^{**} \text{ の場合}$$

帰納法の仮定により $|E_{1i}|, |E_0|, |E_{2i}|$ を生成する clncfg $G_{1i}=(V_{1i}, \Sigma, P_{1i}, S_{1i}), G_0=(V_0, \Sigma, P_0, S_0), G_{2i}=(V_{2i}, \Sigma, P_{2i}, S_{2i})$ が存在する (但し $1 \leq i \leq n$)。次の $G=(V, \Sigma, P, S)$ は E に対する条件を満たす。

$$(4.1) V=V_0 \cup \{S\} \cup \left(\bigcup_{i=1}^n (V_{1i} \cup V_{2i}) \right)$$

$$(4.2) P=P' \cup P_0 \cup \left(\bigcup_{i=1}^n (P_{1i} \cup P_{2i}) \right),$$

ここで, $P'=S \rightarrow S_0 |S_{11} S S_{21}| S_{12} S S_{22}| \cdots |S_{1n} S S_{2n}| \quad \square$

[補題 3.2]

任意の $L \in \text{CLNCFL}(\Sigma)$ に対して $|E|=L$ である $E \in \text{ERE}(2, \Sigma)$ が存在する。

(証明) $L \in \text{CLNCFL}(\Sigma)$ とする。 $G=(V, \Sigma, P, S)$ は L を生成する clncfg とする。証明は $\#V$ に関する帰納法による。

基底段階: $\#V=1$ 。 P の形は;

$$P=S \rightarrow a_1 |a_2| \cdots |a_m |x_1 S y_1 |x_2 S y_2| \cdots |x_n S y_n| \quad (\text{但し } 1 \leq i \leq m, 1 \leq j \leq n \text{ に対して } m, n \geq 0 \text{ かつ } a_i, x_j, y_j \in \Sigma^*)$$

次の E は条件を満たす;

$$E = \left(\begin{array}{c} [X_1] \\ \vdots \\ [X_n] \end{array} \right), a_1 \cup a_2 \cup \cdots \cup a_m, \left(\begin{array}{c} [y_1] \\ \vdots \\ [y_n] \end{array} \right)^{**}$$

帰納段階: $V=\{A_1, \dots, A_n\}$, $A_1 > \cdots > A_n$ とする。 $A_1=S$ と仮定する。

P の形は; $P=A_1 \rightarrow v_{11} | \cdots | v_{1p}$

$$|v_{11} A_1 v_{21}| \cdots |v_{1q} A_1 v_{2q}|$$

$$|w_{11}| \cdots |w_{1r}|$$

$$\cup P'$$

(但し $1 \leq i \leq p, 1 \leq j \leq q, 1 \leq k \leq r$ に対して $v_{1i} \in \Sigma^*, v_{1j} v_{2j}, w_{1k} \in (\{A_2, \dots, A_n\} \cup \Sigma)^*$ であり, 任意の $A_t \rightarrow w \in P'$ に対して $2 \leq t \leq n$ である。

帰納法の仮定によって, 任意の $w \in (\{A_2, \dots, A_n\} \cup \Sigma)^*$ に対して $E(w) \in \text{ERE}(2, \Sigma)$ を次のように定義する。

(2.1) もし $w \in \Sigma^*$ なら, $E(w)=w$ である。

(2.2) $x_i \in \Sigma^*$ ($1 \leq i \leq u+1$), $B_j \in \{A_2, \dots, A_n\}$ ($1 \leq j \leq u$) に対して
 $W = x_1 B_1 x_2 B_2 \dots x_u B_u x_{u+1}$ なら, $E(W) = x_1 E(B_1) x_2 E(B_2) \dots x_u E(B_u) x_{u+1}$ である。

(但し $1 \leq j \leq u$ に対して $E(B_j) \in \text{ERE}(2, \Sigma)$, $|E(B_j)| = L(V, \Sigma, P, B_j)$)

次の E は条件を満たす;

$$E = \left(\left[\begin{array}{c} E(V_{11}) \\ \vdots \\ E(V_{1q}) \end{array} \right], v_{11} \cup \dots \cup v_{1p} \cup E(W_{11}) \cup \dots \cup E(W_{1r}), \left[\begin{array}{c} E(V_{21}) \\ \vdots \\ E(V_{2q}) \end{array} \right] \right)^{**}. \quad \square$$

4. $\text{ERL}(2, \Sigma)$ の基本性質

[命題 4.1]

- (1) $\text{ERL}(2, \Sigma)$ の族は逆系列の演算について閉じている。
- (2) $\text{ERL}(2, \Sigma)$ の族は $\text{ERL}(2, \Sigma)$ 代入について閉じている。
- (3) $\text{ERL}(2, \Sigma)$ の族は $\#\Sigma \geq 2$ のとき共通集合の演算について閉じていない。
- (4) $\text{ERL}(2, \Sigma)$ の族は $\#\Sigma \geq 2$ のとき補集合の演算について閉じていない。

(証明) (1) 明らか。(2) もまた明らか, $\text{ERL}(2, \Sigma)$ 代入は各々の $a \in \Sigma$ に対して $f(a) \in \text{ERL}(2, \Delta)$ であるような代入 $f: \Sigma^* \rightarrow \Delta^*$ である。(3) $\text{ERL}(2, \Sigma)$ の族 $|E_1| = (a, \lambda, b)^{**} a^*$ と $|E_2| = a^*(b, \lambda, a)^{**}$ (但し $\Sigma = \{a, b\}$) の共通集合は $|E_1| \cap |E_2| = \{a^n b^n a^n \mid n \geq 0\}$ であり, 文脈自由型言語でない [1, 2].

(4) $\text{ERL}(2, \Sigma)$ の族が補集合の演算について閉じていると仮定する。一般に, $L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$ なる関係が成り立つ。ここで $L_1, L_2 \in \text{ERL}(2, \Sigma)$ とすると, 上式の左辺は $\text{ERL}(2, \Sigma)$ となる。これは (3) に矛盾する。 \square

[注意 4.1]

$\text{CFL}(\Sigma) \supset \text{ERL}(2, \Sigma) \supset \text{RL}(\Sigma)$ であり, $\#\Sigma = 1$ のとき $\text{CFL}(\Sigma) = \text{RL}(\Sigma)$ であるから, $\#\Sigma = 1$ のとき $\text{ERL}(2, \Sigma) = \text{RL}(\Sigma)$ である ($\text{RL}(\Sigma)$ は Σ 上の正規言語)。

[定義 4.1]

$\text{ERL}(2, \Sigma)$ に対する共通集合問題は任意の与えられた $L_1, L_2 \in \text{ERL}(2, \Sigma)$ に対して $L_1 \cap L_2 = \emptyset$ であるか否かを決定する問題である。

[定理 4.1]

$\text{ERL}(2, \Sigma)$ に対する共通集合問題は決定不能である。

(証明) 証明は [1] の命題 8.3.1 と補題 8.4.1 の証明のように行う。 $\Sigma_0 = \{a, b\}$, $\Sigma = \{a, b, c\}$ とする。任意の $n \geq 1$, $\overline{x} = (x_1, \dots, x_n) \in ((\Sigma_0)^+)^n$ に対して $L(\overline{x})$ を次のように定義する;

$$L(\overline{x}) = \{ba^{i_1} b \dots ba^{i_1} c x_{i_1} \dots x_{i_k} \mid k \geq 1, 1 \leq i_j \leq n, 1 \leq j \leq k\}.$$

そのとき $L(\overline{x})$ を示す $E(\overline{x})$ は;

$$E(\overline{x}) = \left(\left[\begin{array}{c} ba \\ ba^2 \\ \vdots \\ ba^n \end{array} \right], \bigcup_{i=1}^n ba^{i_1} c x_{i_1}, \left[\begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right] \right)^{**}.$$

であるから $L(\bar{x}) \in \text{ERL}(2, \Sigma)$ である。

任意の $\bar{y} = (y_1, \dots, y_n) \in ((\Sigma_0)^+)^n$ に対して $L(\bar{x}, \bar{y})$ を次のように定義する：
 $L(\bar{x}, \bar{y}) = L(\bar{x}) \{c\} (L(\bar{y}))^R$. そのとき $L(\bar{x}, \bar{y})$ を示す $E(\bar{x}, \bar{y})$ は：

$E(\bar{x}, \bar{y}) = E(\bar{x})c(E(\bar{y}))^R$ であるから $L(\bar{x}, \bar{y}) \in \text{ERL}(2, \Sigma)$ である。但し $(E(\bar{y}))^R$ は $E(\bar{y})$ の逆表現である。 $L_S = \{w_1 c w_2 c w_2^R c w_1^R \mid w_1, w_2 \in (\Sigma_0)^+\}$ とする。そのとき L_S を示す $E(L_S)$ は：

$$E(L_S) = \left[\begin{array}{c} a \\ b \end{array} \right], acE_0ca \cup bcE_0cb, \left[\begin{array}{c} a \\ b \end{array} \right]^{**}$$

$$\text{但し } E_0 = \left[\begin{array}{c} a \\ b \end{array} \right], aca \cup bcb, \left[\begin{array}{c} a \\ b \end{array} \right]^{**}$$

ゆえに、 $L(\bar{x}, \bar{y}) \cap L_S = \{t_1 c t_2 c t_2^R c t_1^R \mid t_1 = b a^{i_1} \dots b a^{i_k}, t_2 = x_{i_1} \dots x_{i_k} = y_{i_1} \dots y_{i_k}\}$.

従って $L(\bar{x}, \bar{y}) \cap L_S \neq \emptyset$ は (\bar{x}, \bar{y}) の Post 対応問題の解が存在する時そのときに限る。Post 対応問題は決定不能であるから $\text{ERL}(2, \Sigma)$ に対する共通集合問題は決定不能である (但し $\#\Sigma \geq 2$) 。

5. 高々スター次数 2 の λ -free 拡張正規表現

[定義 5.1]

$E \in \text{ERE}(2, \Sigma)$ の表現の中に λ が現われないなら、 E を λ -free という。

[定理 5.1]

任意の $E \in \text{ERE}(2, \Sigma)$ に対して $|E'| = |E| - \{\lambda\}$ であるような λ -free $E' \in \text{ERE}(2, \Sigma)$ が存在する。

(証明) 証明は E に現れる演算子の個数に関する帰納法による。

基底段階： $E = \lambda, \phi, a$ (但し $a \in \Sigma$) のとき、明かである。

帰納段階：次の四つの場合を考える。

(1) $E = E_1 \cup E_2$ の場合、

帰納法の仮定により $|E_1'| = |E_1| - \{\lambda\}$, $|E_2'| = |E_2| - \{\lambda\}$ であるような λ -free $E_1', E_2' \in \text{ERE}(2, \Sigma)$ が存在する。このとき $E' = E_1' \cup E_2'$ とおけばよい。

(2) $E = E_1 E_2$ の場合、

帰納法の仮定により $|E_1'| = |E_1| - \{\lambda\}$, $|E_2'| = |E_2| - \{\lambda\}$ であるような λ -free $E_1', E_2' \in \text{ERE}(2, \Sigma)$ が存在する。次の四つの場合を考える。

(2.1) $\lambda \in |E_1| \cap |E_2|$ のとき、 $E' = E_1' \cup E_2' \cup E_1' E_2'$ とおく。

(2.2) $\lambda \in |E_1|$ かつ $\lambda \notin |E_2|$ のとき、 $E' = E_2' \cup E_1' E_2'$ とおく。

(2.3) $\lambda \notin |E_1|$ かつ $\lambda \in |E_2|$ のとき、 $E' = E_1' \cup E_1' E_2'$ とおく。

(2.4) $\lambda \notin |E_1| \cup |E_2|$ のとき、 $E' = E_1' E_2'$ とおく。

明かに $|E'| = |E| - \{\lambda\}$ である。

(3) $E = (E_1)^*$ の場合、

帰納法の仮定により $|E_1'| = |E_1| - \{\lambda\}$ であるような λ -free $E_1' \in \text{ERE}(2, \Sigma)$ が存在する。このとき $E' = E_1' (E_1')^*$ とおけばよい。

$$(4) \quad E = \left[\left[\begin{array}{c} E_{11} \\ \vdots \\ E_{1n} \end{array} \right], E_0, \left[\begin{array}{c} E_{21} \\ \vdots \\ E_{2n} \end{array} \right]^{**} \right]$$

次の五つの場合を考える。

(4.1) $E_{11} = \lambda$ のとき、もし $n=1$ であれば、 $|E| = |E_0(E_{21})^*|$ であり、(2)と(3)によって E' が存在する。もし $n \geq 2$ であれば、

$$|E| = \left[\left[\begin{array}{c} E_{12} \\ \vdots \\ E_{1n} \end{array} \right], E_0(E_{21})^*, \left[\begin{array}{c} E_{22}(E_{21})^* \\ \vdots \\ E_{2n}(E_{21})^* \end{array} \right]^{**} \right]$$

であり、 $n-1$ の場合に還元される。

(4.2) $E_{21} = \lambda$ のとき、(4.1)と同様にして、 $|(E_{11})^*E_0|$ または $n-1$ の場合に還元される。

(4.3) ある i ($2 \leq i \leq n$) に対して、 $E_{1i} = \lambda$ のとき、

$$|E| = \left[\left[\begin{array}{c} E_{11} \\ E_{11} \\ \vdots \\ E_{1i-1} \\ E_{1i+1} \\ \vdots \\ E_{1n} \end{array} \right], E_0, \left[\begin{array}{c} E_{21} \\ E_{22} \\ \vdots \\ E_{2i-1} \\ E_{2i+1} \\ \vdots \\ E_{2n} \end{array} \right]^{**} \right]$$

となり、(4.1)に還元される。

(4.4) ある i ($2 \leq i \leq n$) に対して、 $E_{2i} = \lambda$ のとき、(4.2)に還元される。

(4.5) すべての i ($1 \leq i \leq n$) に対して、 $E_{1i}, E_{2i} \neq \lambda$ のとき、次の三つの場合を考える。

(4.5.1) $\lambda \in |E_{11}|$ のとき、 $|E_{11}'| = |E_{11}| - \{\lambda\}$ であるような λ -free $E_{11}' \in \text{ERE}(2, \Sigma)$ が存在して、

$$|E| = \left[\left[\begin{array}{c} E_{11}' \\ E_{12} \\ \vdots \\ E_{1n} \end{array} \right], E_0(E_{21})^*, \left[\begin{array}{c} E_{21}(E_{21})^* \\ E_{22}(E_{21})^* \\ \vdots \\ E_{2n}(E_{21})^* \end{array} \right]^{**} \right]$$

である。

さらに $\lambda \in |E_{21}(E_{21})^*|$ であれば、(2)と(3)によって $|E_{21}'| = |E_{21}(E_{21})^*| - \{\lambda\}$ であるような λ -free $E_{21}' \in \text{ERE}(2, \Sigma)$ が存在して、

$$E = \left[\left[\begin{array}{c} E_{11}' \\ E_{12} \\ \vdots \\ E_{1n} \end{array} \right], E_0(E_{21})^*, \left[\begin{array}{c} E_{21}'(E_{21})^* \\ E_{22}(E_{21})^* \\ \vdots \\ E_{2n}(E_{21})^* \end{array} \right]^{**} \right]$$

である。

(4.5.2) $\lambda \in |E_{21}|$ のとき、(4.5.1)と同様に考える。

(4.5.3) i ($2 \leq i \leq n$) に対して、(4.5.1)と(4.5.2)を繰り返して、 E_{1i} と E_{2i} はすべての i ($1 \leq i \leq n$) に対して λ -freeと仮定する。帰納法の仮定によって $|E_0'| = |E_0| - \{\lambda\}$ であるような λ -free $E_0' \in \text{ERE}(2, \Sigma)$ が存在する。このとき $\lambda \notin |E_0|$ であれば、

$$E' = \left[\begin{array}{c} [E_{11}] \\ \vdots \\ [E_{1n}] \end{array} , E_0', \begin{array}{c} [E_{21}] \\ \vdots \\ [E_{2n}] \end{array} \right]^{**} \text{ とおく.}$$

$E_0 = \lambda$ であれば,

$$|E| = \left| \lambda \cup \left(\bigcup_{i=1}^n E_{1i} E_{2i} \right) \cup \left(\bigcup_{1 \leq i_1, i_2 \leq n} E_{1i_1} E_{1i_2} E_{2i_2} E_{2i_1} \right) \cup \dots \right|$$

$$= \left| \lambda \cup \left[\begin{array}{c} [E_{11}] \\ \vdots \\ [E_{1n}] \end{array} , \bigcup_{i=1}^n E_{1i} E_{2i}, \begin{array}{c} [E_{21}] \\ \vdots \\ [E_{2n}] \end{array} \right]^{**} \right| \text{ である.}$$

最後に, $|E_0| = |\lambda \cup E_0'|$ かつ $|E_0'| \neq \phi$ であれば,

$$|E| = \left| \lambda \cup \left[\begin{array}{c} [E_{11}] \\ \vdots \\ [E_{1n}] \end{array} , E_0' \cup \left(\bigcup_{i=1}^n E_{1i} E_{2i} \right), \begin{array}{c} [E_{21}] \\ \vdots \\ [E_{2n}] \end{array} \right]^{**} \right| \text{ である.}$$

以上によって定理は成り立つ.

6. ERL(2,Σ)と文脈自由型言語族の他の部分族との間の包含問題

LCFL(Σ)をΣ上の線形言語の族, ULCFL(Σ)をΣ上の超線形言語の族, DCFL(Σ)をΣ上の決定性文脈自由型言語の族, UACFL(Σ)をΣ上のあいまいでない文脈自由型言語の族とする(詳細は[1, 2]).

[定義 6. 1]

- (1) ERL(2,Σ)-LCFL(Σ) ≠ φ; (2) ERL(2,Σ)-ULCFL(Σ) ≠ φ;
- (3) ERL(2,Σ)-DCFL(Σ) ≠ φ; (4) ERL(2,Σ)-UACFL(Σ) ≠ φ;
- (5) LCFL(Σ)-ERL(2,Σ) ≠ φ; (6) ULCFL(Σ)-ERL(2,Σ) ≠ φ;
- (7) DCFL(Σ)-ERL(2,Σ) ≠ φ; (8) UACFL(Σ)-ERL(2,Σ) ≠ φ;

ここで, #Σ ≥ 4 である.

(証明) (1) LCFL(Σ) ⊂ ULCFL(Σ) であるから (2) より成り立つ.

(2) $L_1 = \{(a^n c b^n \mid n \geq 1) d\}^* \in \text{ULCFL}(\Sigma)$ は山崎によって証明された [3]. 他方, L_1 を示す E_1 は: $E_1 = \{(a, acb, b)^{**} d\}^*$.

(3) $L_2 = \{ww^R \mid w \in \{a, b\}^*\} \in \text{DCFL}(\Sigma)$ は知られている. 他方, L_2 を示す E_2 は:

$$E_2 = \left[\begin{array}{c} [a] \\ [b] \end{array} , \lambda, \begin{array}{c} [a] \\ [b] \end{array} \right]^{**}.$$

(4) $L_3 = \{a^n b^n c^m d^m \mid n \geq 1, m \geq 1\} \cup \{a^n b^m c^m d^n \mid n \geq 1, m \geq 1\} \in \text{UACFL}(\Sigma)$ は知られている [1, 2]. 他方, L_3 を示す E_3 は:

$$E_3 = (a, ab, b)^{**} (c, cd, d)^{**} \cup (a, a(b, bc, c)^{**} d, d)^{**}.$$

(5) 次の線形文法 $G = (V, \Sigma, P, S)$ を考える. ここで, $\Sigma = \{a, b, c\}$, $V = \{S, A\}$,

$P: S \rightarrow aba \mid aSa \mid cAc, \quad A \rightarrow bAb \mid cSc.$

$L(G) \in \text{ERL}(2, \Sigma)$ と仮定する. そのとき $L(G)$ を示す λ -free $E \in \text{ERE}(2, \Sigma)$ が存

在する。まず次の系を示す。

(5.1) 任意の $w \in L(G)$ に対し, $w = vabav^R$ であるような $v \in \Sigma^*$ が存在する。

(5.2) $L(G) \cap \Sigma^*aba\Sigma^*aba\Sigma^* = L(G) \cap \Sigma^*abb\Sigma^* = L(G) \cap \Sigma^*bba\Sigma^* = \emptyset$ 。

(5.3) 任意の $u, v, w, x, y \in \Sigma^*$ に対して, すべての $i \geq 0$ に対して $ux^ivy^iw \in L(G)$ であれば, $l(x) = l(y)$ である。

(系の証明) (5.1) と (5.2) は定義によって明らかである。

(5.3) すべての $i \geq 0$ に対して $ux^ivy^iw \in L(G)$, $l(x) \neq l(y)$ と仮定する。 $l(x) > l(y)$ と仮定する (その他は対称)。そのときある十分大きい $i > 0$ に対して $l(x^i) > 2l(uvy^iw) + 3$ が存在する。(5.1) よりある $x_0, x_1 \in \Sigma^*$ に対して $x^i = x_0abax_1$ である。しかし $ux^{2i}vy^{2i}w = ux_0abax_1x_0abax_1vy^{2i}w \in L(G)$ であり, (5.1) に矛盾する。□

定理の証明 (続き) n を E に現われる Σ に属するすべての記号の数とする。 $w = (a^{n+1}cb^{n+1}c)^{n+1}aba(cb^{n+1}ca^{n+1})^{n+1}$ を考える。明かに $w \in L(G)$ である。従って w は E の中で導かれる。 E の位置 i に, 位置 i に現われる各々の $x \in \Sigma$ を結びつける, 但し i は正の整数で, x は左から E の i 番目の位置に現われる記号を示す。従って w に対してペアの系列 $((x_1, i_1), (x_2, i_2), \dots, (x_p, i_p))$ が存在する, ここで $w = x_1x_2 \dots x_p$, $x_1 = \dots = x_{n+1} = a$, $x_{n+2} = c$, $x_{n+3} = \dots = x_{2n+3} = b$, \dots 等, そして各々の i_j は E における x_j の位置である。 n の定義によって $1 \leq j_{11} < j_{11}' \leq n+1$ かつ $i_{j_{11}} = i_{j_{11}'}$ であるような j_{11}, j_{11}' が存在する。(5.3) によって (H)* の型の E の部分表現は存在しないことに注意する。従って $i_{j_{11}}$ 番目の位置は次のように ** の中に現われる:

$$\dots \left[\left[\begin{array}{c} \vdots \\ a^k \\ \vdots \end{array} \right], \dots, \left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \right] \right]^{**}, \text{ 或は } \dots \left[\left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \right], \dots, \left[\begin{array}{c} \vdots \\ a^k \\ \vdots \end{array} \right] \right]^{**} \dots$$

同様に $n+3 \leq j_{21} < j_{21}' \leq 2n+3$ かつ $i_{j_{21}} = i_{j_{21}'}$ であるような j_{21}, j_{21}' が存在する。従って $i_{j_{21}}$ 番目の位置は次のように ** の中に現われる:

$$\dots \left[\left[\begin{array}{c} \vdots \\ b^r \\ \vdots \end{array} \right], \dots, \left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \right] \right]^{**}, \text{ 或は } \dots \left[\left[\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array} \right], \dots, \left[\begin{array}{c} \vdots \\ b^r \\ \vdots \end{array} \right] \right]^{**} \dots$$

(5.2) より $\left[\begin{array}{c} \vdots \\ a^k \\ \vdots \end{array} \right]$ と $\left[\begin{array}{c} \vdots \\ b^r \\ \vdots \end{array} \right]$ は明かに異なる。

w と E を通じて整数の系列 $j_{11}, \dots, j_{1n+1}, j_{21}, \dots, j_{2n+1}$ を得る。 n の選択によって $1 \leq e < f \leq n+1$ かつ $i_{j_{1e}} = i_{j_{1f}}$ であるような整数 e, f が存在する。従って E は次のような部分表現を持つ;

$$\dots \left[\left[\dots (\dots, \dots, \dots)^{**} \dots \right], \dots, \left[\dots \right] \right]^{**} \dots, \text{或は}$$

$$\dots \left[\left[\dots \right], \dots, \left[\dots (\dots, \dots, \dots)^{**} \dots \right] \right]^{**} \dots$$

従って、次の(5.4)または(5.5)が成り立つ。

(5.4) すべての $i, j \geq 0$ に対して $u_0(u_1 v_0^i w_0 v_1^j u_2)^j w_1 v_2^j u_3 \in |E|$ であるような $u_0, u_1, u_2, u_3 \in \Sigma^*$ かつ $v_0, w_0, v_1, w_1, v_2 \in \Sigma^+$ が存在する。

(5.5) すべての $i, j \geq 0$ に対して $u_0 v_0^i w_0 (u_1 v_1^j w_1 v_2^j u_2)^j u_3 \in |E|$ であるような $u_0, u_1, u_2, u_3 \in \Sigma^*$ かつ $v_0, v_1, v_2, w_0, w_1 \in \Sigma^+$ が存在する。

これは(5.3)に矛盾する。 \square

(6)-(8) $L(G) \in \text{ULCFL}(\Sigma) \cap \text{DCFL}(\Sigma) \cap \text{UACFL}(\Sigma)$ に注目すれば、(5)の証明より(6)-(8)は成り立つ。 \square

参考文献

- [1] M.A.Harrison, Introduction to Formal Language Theory, Addison-Wesley, 1978.
- [2] J.E.Hopcroft and J.D.Ullman, Introduction to Automata Theory, Language, and Computation, Addison-Wesley, 1979.
- [3] 山崎：文脈自由言語のある部分言語の包含関係に関する一考察, 信学論(D), J70-D, 1, pp.1-9(87/1).
- [4] S.A.Greibach, The unsolvability of the recognition of linear context-free languages. JACM, 13, 4 pp.582-587(1966).
- [5] S.Ginsburg and H.G.Rice, Two families of languages related to ALGOL, JACM, 9, pp.350-371(1962).
- [6] S.Ginsburg and E.H.Spanier, Bounded ALGOL-like languages, Trans.Am.Math.Soc., 113, pp.333-368(1964).
- [7] S.Ginsburg, The Mathematical Theory of Context-Free Languages, McGraw-Hill(1966).
- [8] E.Shamir, On sequential language, Z.Phonetik, Sprach. & Kommunikationsforschung, 18, pp.61-69(1965).
- [9] Kosaburo Hashiguchi and Heekeun Yoo, Extended regular expressions of star degree at most two, to appear in Theoretical Computer Science.