帰納的に可算な言語の同長線形文脈自由言語による特徴付けについて

大　川　　　知*

On the Characterizations of Recursively Enumerable Languages by Even-Linear Languages

Satoshi OKAWA

Abstract

For a linear context-free grammar G, if every non-terminating rule of G is of the form $A \rightarrow \alpha B \beta$ with $|\alpha| = |\beta|$, it is called an even-linear grammar and the language generated by G, L(G), is called an even-linear language.

In this note, we give two characterization results of recursively enumerable languages using even-linear languages and homomorphisms:

i) For each recursively enumerable language L, we can find two even-linear languages L_1, L_2 and two homomorphisms h, h_1 such that $L = h(h_1(L_1) \cap L_2)$.

ii) For each recursively enumerable language L, we can find an even-linear language L_E and three homomorphisms h, h_1, h_2 such that $L = h(h_1(L_E) \cap h_2(L_E))$.

1. はじめに

言語理論において、言語のクラスの特徴付けをするという問題は重要な研究課題となっており、文法によって定まる言語のクラスと、言語の受理装置であるオートマトンによって定まる言語のクラスとの間の対応付けは、その一つの重要な結果である。また、これとは別に、これらのクラスを代数的に特徴付ける問題がある。例えば、文脈自由言語ＬをＤｙｃｋ言語Ｄと正規言語Ｒと準同形写像ｈとによって、Ｌ＝ｈ（Ｄ∩Ｒ）という形で表現する[1]、帰納的に可算な言語Ｌを２つの線形文脈自由言語L_1，L_2と準同型写像ｈ

平成元年１２月１５日受理
＊　電気工学科助教授

とによって、$L=h(L_1\cap L_2)$という形で表現する[2]などの古典的な結果がよく知られている。最近、アルファベットΣに対して、Σだけに依存してアルファベットΣ'と準同形写像$h:\Sigma'^*\to\Sigma^*$とが定まり、Σ'上のDyck言語をDとして、Σ上の任意の帰納的可算な言語L_{RE}と文脈自由言語L_{CF}とに対して、$L_{RE}=h(D\cap M_{RE}')$、$L_{CF}=h(D\cap M_{CF}')$となる極小線形言語M_{RE}'と極小線形正規言語M_{CF}'が存在する[3]、文脈依存言語に対しては、このような形の特徴付けができない[4]、任意の帰納的に可算な言語Lが本論文の後で述べる右長、左長、同長の3種の線形言語のうち、同長と同長を除く任意の2つの言語のL_1, L_2と準同形写像hとによって、$L=h(L_1\cap L_2)$と表わされる[5]などの結果が得られている。

本論文においては帰納的に加算な言語の、同長線形言語と準同形写像とによる特徴付けに関する結果を与える。すなわち、任意の帰納的に可算な言語Lは、2つの同長線形言語L_1, L_2と、2つの準同形写像h, h_1とによって$L=h(h_1(L_1)\cap L_2)$と表わすことができることを示す。さらに、任意の帰納的に可算な言語Lは、1つの同長線形言語L_Eと、3つの準同形写像h, h_1, h_2とによって、$L=h(h_1(L_E)\cap h_2(L_E))$と表わすことができることを示す。

2. 準　備

本節において、本論文に必要な文法、導出、言語等に関する定義を行なうが、詳細については、文献(6)などを参照されたい。

文法とは$G=(N, \Sigma, P, S)$である。ここで、Nは非終端記号の集合、Σは終端記号の集合($V=\Sigma\cup N$とする。)、Pは書きかえ規則の集合、SはNの元で初期記号である。Pの元pは$p:\alpha\to\beta$ ($\alpha, \beta\in V^*$, αはNの元を含む)という形をしている。

文法Gの書きかえ規則$p:\alpha\to\beta$が文字列ωに適用可能であるとは、ωが$\omega_1\alpha\omega_2$と分解できるときであり、ωにpを適用した結果は、αをβに書きかえて$\omega'=\omega_1\beta\omega_2$なる文字列である。このとき$\omega\overset{p}{\underset{G}{\Longrightarrow}}\omega'$と表わす。

文法Gによる語ωの導出とは、Gの書きかえ規則の系列p_1, p_2, ･･･, p_nによる

$$S\overset{p_1}{\underset{G}{\Longrightarrow}}\omega_1\overset{p_2}{\underset{G}{\Longrightarrow}}\omega_2\overset{p_3}{\underset{G}{\Longrightarrow}}\cdots\overset{p_n}{\underset{G}{\Longrightarrow}}\omega_n=\omega$$

のことである。このことを$S\overset{p_1p_2\cdots\cdots p_n}{\underset{G}{\Longrightarrow}}\omega$, $S\overset{*}{\underset{G}{\Longrightarrow}}\omega$と表わす。文法Gがあきらかなときには、Gを省略して、$\overset{p}{\underset{G}{\Longrightarrow}}$等の代わりに$\overset{p}{\Longrightarrow}$等と書くことがある。

文法Gによって生成される言語L(G)とは、

$$L(G)=\{\omega\mid S\overset{*}{\underset{G}{\Longrightarrow}}\omega,\ \omega\in\Sigma^*\}$$

のことである。

文法Gの書きかえ規則p：$\alpha \rightarrow \beta$について、β中にNの元が入っているときpを非終端規則、そうでないとき終端規則という。また、Gのすべての書きかえ規則p：$\alpha \rightarrow \beta$について、

1）何の制約もない

2）$|\alpha| \leq |\beta|$

3）$\alpha \in N$，すなわち、p：$A \rightarrow \beta$、$A \in N$

4）p：$A \rightarrow aB$　または、$A \rightarrow a$　$B \in N$，$a \in \Sigma$

となっているとき、その文法をそれぞれ、

1）帰納的に可算な文法

2）文脈依存文法

3）文脈自由文法

4）正規文法

といい、それらが生成する言語を、

1）帰納的に可算な言語

2）文脈依存言語

3）文脈自由言語

4）正規言語

という。

文脈自由文法Gの書きかえ規則のうち、非終端規則$A \rightarrow \beta$について、$\beta \in \Sigma^* V \Sigma^*$のとき線形規則といい、すべての非終端規則が線形である文法を線形文法といい、その文法で生成される言語を線形言語という。さらに、$\beta = \omega_1 B \omega_2$（$\omega_1$，$\omega_2 \in \Sigma^*$）が$|\omega_1| < |\omega_2|$のとき右長、$|\omega_1| = |\omega_2|$のとき同長、$|\omega_1| > |\omega_2|$のとき左長規則という。すべての非終端規則が右長（同長、左長）規則であるとき、その文法を右長（同長、左長）線形文法といい、その文法によって生成される言語を右長（同長、左長）線形言語という。

先に述べたように、帰納的に可算な言語については、次のような特徴付けが知られている。

〔定理1〕[(5)]　任意の帰納的に可算な言語Lに対し、

i）2つの右長（左長）線形言語L_1，L_2と準同形写像hとが定まり、
$L = h(L_1 \cap L_2)$と表わされる。

ii）右長（左長）線形言語L_1と同長線形言語L_2と準同形写像hが定まり、
$L = h(L_1 \cap L_2)$と表わされる。

iii）右長線形言語L_1と左長線形言語L_2と準同形写像hが定まり、
$L = h(L_1 \cap L_2)$と表わされる。

この定理を見ると、2つの同長線形言語L_1，L_2による場合がないが、

これについては次の結果が成立する。

〔定理２〕(5)　２の同長線形言語L_1，L_2と準同形写像hについて、$h(L_1 \cap L_2)$は線形言語である

３．結　果

前節の定理１と定理２とは極めて対照的な結果である。すなわち帰納的に加算な言語は、定理１と同じような形式で同長言語による特徴付けができないことが示されている。従って、どのようにしても、同長言語だけでは帰納的に可算な言語が特徴付けられないのかという点に興味がある。このことについて、本節で考察する。まず初めに、次の補題が成立することを示す。

〔補題１〕　任意の線形言語Lに対して、同長言語L_Eと準同形写像hが定まり、Lは、$L=h(L_E)$と表わすことができる。

（証　明）　Lを生成する線形文法を$G=(N, \Sigma, P, S)$とする。Pの元はn個の非終端規則$p_i : A_i \to \omega_{i1} B_i \omega_{i2}$ $(i=1, 2, \cdots, n)$、及び、$|P|-n$個の終端規則$p_j : A_j \to a_j$ $(j=n+1, \cdots, |P|)$となっているとする。このとき、同長線形文法$G_E=(N_E, \Sigma_E, P_E, S_E)$を次のように定める。

$$N_E=N$$
$$\Sigma_E=\{p_{i1}, p_{i2}, \mid i=1, 2, \cdots, n\} \cup \Sigma$$
$$P_E=\{p_j : A_i \to p_{i1} B_i p_{i2} \mid i=1, 2, \cdots, n\}$$
$$\cup \{p_j : A_j \to a_j \mid j=n+1, \cdots, |P|\}$$
$$S_E=S$$

とする。準同形写像hを$h : \Sigma_E^* \to \Sigma^*$を次のように定める。

$$h(p_{i1})=\omega_{i1}, \quad h(p_{i2})=\omega_{i2}, \quad (i=1, 2, \cdots, n)$$
$$h(a)=a \quad (a \in \Sigma)$$

とする。すると、

$$L=h(L(G_E))$$

となる。$L_E=L(G_E)$として結果を得る。

〔定理３〕　任意の帰納的に可算な言語Lに対して、２つの同長線形言語L_1，L_2と２つの準同形写像h, h_1が定まりLは$L=h(h_1(L_1) \cap L_2)$と表わされる。

（証明）　定理１のii）よりLは右長線形言語L_1'と同長線形言語L_2と準同形写像hとによって

$$L=h(L_1' \cap L_2)$$

と表わされる。補題１よりある同長線形言語L_1が存在し、$L_1'=h_1(L_1)$

とすることができるから、

$$L = h\,(h_1\,(L_1) \cap L_2)$$

となる。

この定理では、2つの同長線形言語を用いているが、準同形写像を1つ増すことによって同長線形言語は1つでよいことが示される。

〔定理4〕　任意の帰納的に可算な言語Lに対し、同長線形言語L_Eと3つの準同形写像h，h_1，h_2が定まり$L = h\,(h_1\,(L_E) \cap h_2\,(L_E))$と表わされる。

（証　明）　定理1のi)により、任意の帰納的に可算な言語Lは2つの右長線形言語L_1，L_2と準同形写像hとによって$L = h\,(L_1 \cap L_2)$と表わされる。L_1，L_2を生成する右長線形文法を$G_1 = (N_1, \Sigma, P_1, S_1)$，$G_2 = (N_2, \Sigma, P_2, S_2)$とする。一般性を失うことなく、$N_1 \cap N_2 = \phi$と仮定する。さらに、

$$P_1 = \{\pi_i : A_i \to \alpha_i B_i \beta_i \mid i = 1, 2, \cdots, n\}$$
$$\cup \{\pi_j : A_j \to a_j \mid j = n+1, \cdots, |P_1|\}$$
$$P_2 = \{\rho_i : C_i \to \gamma_i D_i \delta_i \mid i = 1, 2, \cdots, m\}$$
$$\cup \{\rho_j : C_j \to b_j \mid j = m+1, \cdots, |P_2|\}$$

となっているものとする。

このとき、$N = N_1 \cup N_2$,

$$\Sigma' = \{\pi_{i1}, \pi_{i2} \mid i = 1, 2, \cdots, n\}$$
$$\cup \{\rho_{j1}, \rho_{j2}, \mid j = 1, 2, \cdots, m\} \cup \{a' \mid a \in \Sigma\}$$
$$P = \{A \to \pi_{i1} B \pi_{i2} \mid \pi_i : A \to \alpha_i B \beta_i \quad i = 1, 2, \cdots, n\}$$
$$\cup \{A_j \to a_j' \mid \pi_j : A_j \to a_j \quad j = n+1, \cdots, |P_1|\}$$
$$\cup \{C \to \rho_{i1} D \rho_{i2} \mid \rho_i : C \to \gamma_i D \delta_i \quad i = 1, 2, \cdots, m\}$$
$$\cup \{C_j \to b_j' \mid \rho_j : C_j \to b_j \quad j = m+1, \cdots, |P_2|\}$$
$$\cup \{S \to S_1, S \to S_2\}$$

とし、$G = (N, \Sigma', P, S)$とすると、Gは同長文法である。

準同形写像h_1，h_2を

$$h_1\,(\pi_{i1}) = \alpha_i,\ h_1\,(\pi_{i2}) = \beta_i,\ i = 1, 2, \cdots, n$$
$$h_1\,(a') = a, \quad a \in \Sigma$$
$$h_1\,(\rho_{i1}) = \varepsilon,\ h_1\,(\rho_{j2}) = \varepsilon,\ i = 1, 2, \cdots, m$$
$$h_1\,(a') = \varepsilon, \quad a \in \Sigma$$

及び、

$$h_2\,(\pi_{i1}) = \varepsilon,\ h_2\,(\pi_{i2}) = \varepsilon,\ i = 1, 2, \cdots, n$$
$$h_2\,(a') = a, \quad a \in \Sigma$$
$$h_2\,(\rho_{i2}) = \gamma_i,\ h_2\,(\rho_{i2}) = \delta_i,\ i = 1, 2, \cdots, m$$
$$h_2\,(a') = a, \quad a \in \Sigma$$

と定めると、

$$h_1(L(G))=L_1,\ h_2(L(G))=L_2$$

である。従って、

$$L=h(h_1(L(G))\cap h_2(L(G)))$$

となる。$L(G)=L_E$とすることによって結果を得る。

4．むすび

本論文では、帰納的に可算な言語の特徴付けについて考察を行ない、同長線形言語による特徴付けが可能であることを示した。すなわち、任意の帰納的に可算な言語Lに対し、２つの同長線形言語L_1，L_2と２つの準同形写像h_0，h_1とによって、$L=h_0(h_1(L_1)\cap L_2)$と表わされることと、１つの同長言語L_Eと３つの準同形写像h，h_1，h_2とによって、$L=h(h_1(L_E)\cap h_2(L_E))$と表わされることを示した。この結果と定理１とを対比してみると、言語の条件が厳しくなると準同形写像の数が増えていくように見うけられる。

上に述べたことから、今後の課題として、言語に対する条件と準同形写像の関係についての検討、さらに言語のクラスを小さくして帰納的に加算な言語を特性化することができるかどうかの検討などがある。

参考文献

1）N.Chomsky,“Context-free grammars and pushdown strage”,MIT Res.Lab.Electron.Quart.Prog.Report 65 pp.187－194(1962).

2）S.Ginsburg,S.A.Greibach and M.A.Harrison,“One-way stack automata,”J.ACM 14 pp.389－418(1967).

3）S.Hirose,S.Okawa and M.Yoneda,“A homomorphic characterization of recursively enumerable languages”,Theoret.Comput.Sci.35 pp.261－269(1985).

4）S.Okawa,S.Hirose and M.Yoneda,“On the impossibility of the homomorphic characterization of context-sensitive languages”, ibid,44,pp.225－228(1986).

5）大川 知、広瀬貞樹、“線形言語の部分族と言語の特性化”、投稿中

6）本多波雄、“オートマトン・言語理論”、コロナ社(昭47).