

D3 代表記号列集合によるオートマトンの適応的修正法

榎本 肇, 富田悦次, 堂下修司, 土屋 昇 (東京工業大学工学部)

1. まえがき

オートマトン M の外的特性は、それが受理する一般には無限の入力記号列の集合 $N(M)$ として表わされる。ところが、もし適当な解釈基準のもとにおいて、その特性が、受理あるいは非受理指示の代表的な有限長記号列の有限集合によって完全に表現できれば、 M の外的ふるまいは、非常に容易にチェックできることになる。ここでは、そのようなことを保証する記号列集合を M の代表記号列集合 $\Pi_C(M)$ と名付けて、それを求めるアルゴリズムを示し⁽¹⁾、さらに、この考えを用いて、与えられたオートマトン（一般に、誤りを含む）を、教師との効果的質問応答により、適応的に目的とするオートマトンへと修正してゆく方法を示した⁽²⁾。

2. 前定条件および諸定義

特に説明のない記号、用語の使い方は、文献(3)、(1)による。

ここでは、対象とするオートマトンとしては、 ϵ -入力なし、空スタック式受理の決定性プッシュダウンオートマトン (PDA) $M = (K, \Sigma, \Gamma, \delta, P_0, S_0, \phi)$ ⁽³⁾ を考える。また、その決定性 PDA は、同じ入出力特性をもつ決定性 PDA のうちで、用いられている表層状態の種類が最小であるようなものとする。さらに簡単のため、PDA のスタックが空となって受理を決定するときにおける内部状態は、1 種のみとする。以後、単に DPDA と書いても、このような決定性 PDA をさすものとする。なお、決定性有限状態オートマトンは、このような DPDA に完全に含まれる。

DPDA のある時点における内部状態とプッシュダウン記号列の対 (P, r) を計算状態と呼び、 $r = Ar'$ であるときには、 (P, A) ($r \neq \epsilon$ のときには $A \in \Gamma$ 、さもなければ $A = \epsilon$) を、表層状態と呼ぶ。

$uvw \in N(M)$ であるとき、部分記号列 w における推移過程中に用いられる M の写像規則の系列を、 $P_M(w; u, v)$ で表わす。また、 $u : (P_0, S_0) \vdash_M^* (P, r)$ であるとき、計算状態 (P, r) を $C_M(u)$ で表わす。

ここで扱う入力記号列は、その最右端に、終止記号 ($\#$) が付けられているようなもののみとする。また、各入力記号間には、アルファベット順序が定められているとする。入力記号列に対しては、短いものから長いものへと順に、また、同じ長さのものに対しては、左端からみて、初めて入力記号の異なる位置における入力記号のアルファベット順序に従って、辞書式順序を考える。ここで、記号列 x が y に対して順序が先行することを $x \leq y$ で表わす。また、記号列集合 Π のうちで、最も順序の先行する記号列を $\text{First}(\Pi)$ で表わす。

3. DPDA の代表記号列集合

3.1 有限記号列集合による DPDA の表現

受理すべきか否（非受理）かの指示の付けられた、有限長、有限個のサンプル記号列が与えられたとき、それらを正しく識別する、表層状態の種類が最小であるような DPDA を、全て求め上げることは可能である。^{(1), (4)} ここで、受理指示で与えられたどの記号列の推移過程にも関与しない余分な写像規則は設けないことにしても、このようにして得られる DPDA は、一般に複数個となる。ところが、もしそれがただ 1 個の DPDA M か、あるいは、それと同じ入出力特性をもつ DPDA の一群に限定されるならば、ここで与えられたサンプル記号列集合は、 M の外的特性を

完全に表現しているとみなすことができる。そこで、このようなことを保証する記号列集合を M の“代表記号列集合”と呼んで、 $\Pi_C(M)$ で表し、DPDA が与えられた場合、適切な基準のもとにおいて、その代表記号列集合を有効に求める方法を示す。なお記号列に対する受理あるいは非受理の指示は、それぞれ、記号列の先頭に $\circ$ 印、 $\times$ 印を付けて示すこととするが、このような $\circ$ 印は省略することもある。

3.2 基本記号列集合 $\Pi_F(M)$ ，第 0 種副次記号列集合 $\Pi_{S_0}(M)$ ，SE 規定記号列集合 $\Pi_{SE}(M)$

〔基本記号列集合 $\Pi_F(M)$ 〕 先ず、与えられた DPDA M を表現するための基本として、 M に受理される限り順序の先行する記号列で、その推移過程に、それまでの記号列の受理過程でまだ用いられていない写像規則を少なくとも 1 回は使用するようなものを逐次求め、全ての写像規則が使用し尽された時点で得られている集合を基本記号列集合 $\Pi_F(M)$ という。その各要素は、基本記号列と呼ぶ。

〔第 0 種副次記号列集合 $\Pi_{S_0}(M)$ 〕[†] 次に、基本記号列の推移中にとられる、ある計算状態 (p, r) に対して、 $u_0 = \text{First} \{ u \mid u : (p_0, S_0) \xrightarrow{*} (p, r) \}$ とおいたとき、ある入力記号 a に対して、 $(\exists av, \neg \exists v') \{ av : (p, r) \xrightarrow{*} (r, \varepsilon) \mid u_0 av' \in \Pi_F(M) \}$ (3.1)

となっているとする。このとき、 (p, r) から a による分岐推移が可能であることを規定する副次的な記号列として、 $u_0 av \in N(M) \setminus \Pi_F(M)$ なる記号列を求め、これを第 0 種副次記号列という。 M に対して、このようにして得られる記号列の全体を $\Pi_{S_0}(M)$ で表わす。

〔SE 規定記号列集合 $\Pi_{SE}(M)$ 〕 基本記号列の推移中にとられる 2 つの計算状態 $(p, A\alpha\beta)$ 、 $(p, A\alpha r\beta)$ ($A \in \Gamma$, $r \in \Gamma^+$) に関して、(3.2) 式に示されるような推移が存在するとする。

$uxwylv : (p_0, S_0) \xrightarrow{*} (p, A\alpha\beta) \xrightarrow{*} (p, A\alpha r\beta) \xrightarrow{*} (q, r\beta) \xrightarrow{*} (q, \beta) \xrightarrow{*} (r, \varepsilon)$ (3.2)
ただし、 $w : (p, A\alpha) \xrightarrow{*} (q, \varepsilon)$ 、 $xwy : (p, A\alpha) \xrightarrow{*} (q, \varepsilon)$ かつ、 $H(x) \neq H(w)$ [‡]

このとき、 M は SE (self-embedding) 構造をもつといい、この構造を規定するために、 $ux^2 w y^2 v \in N(M)$ なる記号列をつくり出し、SE 規定記号列と名付ける。 M から得られるこのような SE 規定記号列全体の集合を $\Pi_{SE}(M)$ で表わす。

3.3 DPDA の代表記号列集合 $\Pi_C(M)$

3.3.1 Affix 条件

前節に示した $\Pi_F(M)$ 、 $\Pi_{S_0}(M)$ 、 $\Pi_{SE}(M)$ は、その中の各記号列に受理指示を付けて、 M の代表記号列集合 $\Pi_C(M)$ の中に、必ず採り入れるものとする。このような前提、すなわち、 $\Pi_F(M) \subseteq \Pi_C(M)$ 、 $\Pi_{S_0}(M) \subseteq \Pi_C(M)$ のもとでは、 M において次の関係 $[A_F]$ 、 $[A_0]$ が成立している。

〔 A_F 〕 ある基本記号列の prefix u_1, u_2 に対して、 $u_1 w_1 v_1 \in \Pi_C(M)$ 、 $u_2 w_2 v_2 \in \Pi_C(M)$ であり、かつ $u_1 w_1' v_1 \in \Pi_C(M)$ 、 $u_2 w_1' v_2 \notin \Pi_C(M)$ 、 $w_1 \leq w_1'$ 、 $u_2 w_1' v_2 \leq u_1 w_1' v_1$ であるとする。このとき、 $\neg \{ u_1 w_1' v_1 \in \Pi_F(M) \wedge C_M(u_1 w_1) = C_M(u_1 w_1') \wedge u_2 w_1' v_2 \in N(M) \wedge P_M(w_1'; u_1, v_1) = P_M(w_1'; u_2, v_2) \}$ (3.3)

特別の場合として、 $v_1 = v_2 = \varepsilon$ である場合を考えると、

$u_{10} = \text{First} \{ w \mid u, w \in \Pi_C(M) \}$ 、 $u_{20} = \text{First} \{ w' \mid u_2 w' \in \Pi_C(M) \}$
とおいたとき、もし、 $w_{10} \neq w_{20}$ であるならば、 $\tau \{ u_1 w_{10} \in \Pi_F(M) \wedge u_2 w_{20} \in \Pi_F(M) \wedge C_M(u_1) = C_M(u_2) \}$ (3.4)

〔 A_0 〕 ある u に対して、もし

(脚注) [†] このような記号列集合を特に考えないような、 M の表現法も可能である⁽¹⁾

[‡] $H(x)$ は、記号列 x の先頭記号を表わす。

$$(\exists a, \neg \exists w') \{ u a w' \in \Pi_C(M) \} \text{ であるならば, } C_M(u) = (C_u) \text{ といいたとき,} \\ \neg [u = \text{First} \{ u' | u' : (p_0, S_0) \vdash_M^* (C_u) \} \wedge (\exists w) \{ u a w \in N(M) \}] \quad (3.5)$$

ところで、前記のような $\Pi_C(M)$ から、それを正しく識別する DPDA M' を再構成したとき、もし

$$\Pi_F(M') \subseteq \Pi_C(M) \quad (3.6)$$

が成立しないならば、 M' は明らかにもとの M とは異り、このような M' は、再構成されるべき DPDA としては不適格なものとして、全て除外することにする。ここで、(3.6) 式なる関係が満たされているときには、前記関係 $[A_F]$ 、 $[A_0]$ において、 $\Pi_C(M)$ における M 以外の全ての M を M' で置きかえて、これらを再構成される DPDA M' における関係とみなしたとき、これは適格な M' においても満足されるべきものであり、このような条件付を、それぞれ “Affix 条件 A_F ” “Affix 条件 A_0 ” と呼び、DPDA 構成時には、常に考慮に入れることにする。ただし、Affix 条件 A_F は、(3.6) 式が満足されれば、自動的に満たされている。

3.3.2 計算状態および写像規則系列の異同の規定

$\pi_F(M) \cup \pi_{S_0}(M) \cup \pi_{SE}(M)$ から DPDA を再構成するとき、前項で示した Affix 条件を全て満足するとの考慮をはらうだけでは、使用する表層状態の種類を最小にするとの要請のもとでは、基本記号列に沿う推移中にとられる計算状態や写像規則系列で、正しくは区別されるべきものも不適当に統合されてしまったような DPDA が得られてしまう場合が生じる。従って、このような区別を規定するために、さらに受理または非受理指示の記号列を新しく導入して、第 1 種副次記号列と名付け、そのようなものの集合を $\pi_{S_1}(M)$ で表す。ただし、計算状態の異同を規定する操作は、写像規則系列の異同を規定する操作の特別の場合として含まれるので、特に分けて述べない。

さて Affix 条件 A_F 、 A_0 を考慮したもとにおいては $u_1 w_{10} v_1 \in \Pi_F(M)$ 、 $u_2 w_{20} v_2 \in \Pi_F(M)$ で、 $P_M(w_{10}; u_1, v_1) \neq P(w_{20}; u_2, v_2)$ なる場合に、両写像規則系列の相異を陽に記号列の指示によって規定しなければならないのは、 $w_{10} = w_{20}$ ($= w$ とおく) である場合に限られ、以下、次の (a) ~ (c) のような場合分けに応じてその対処を示す。ここで、 $\{\phi_1\}$ 、 $\{\phi_2\}$ は、それぞれ (3.7)、(3.8) で示される条件を表わすものとする。また、 $P_M(w; u_1, v_1)$ 、 $P_M(w; u_2, v_2)$ において、各対応する入力記号直前にとられる表層状態の記号を全て統合してしまったような PDA を $\bar{M}$ で表す。

$$\{\phi_1\} = \{ u_1 w_1 v_1 \in N(M) \wedge u_2 u_1 v_2 \notin N(M) \wedge C_M(u_1 w) = C_M(u_1 w_1) \wedge w \leq w_1 \} \quad (3.7)$$

$$\{\phi_2\} = \{ u_2 w_2 v_2 \in N(M) \wedge u_1 u_2 v_1 \notin N(M) \wedge C_M(u_2 w) = C_M(u_2 w_2) \wedge w \leq w_2 \} \quad (3.8)$$

(a) $(\exists w_1) \{ \phi_1 \} (\exists w_2) \{ \phi_2 \}$:

(a-1) ; $(\exists w_1) \{ \phi_1 \} \wedge (\exists w_2) \{ \phi_2 \}$ であるとき、 $\circ u_1 w_1 v_1, \circ u_2 w_2 v_2$ を $\Pi_{S_1}(M)$ に加える。

(a-2) ; $(\exists w_1) \{ \phi_1 \} \wedge (\neg \exists w_2) \{ \phi_2 \}$ であるとき、 $\circ u_1 w_1 v_1, \times u_2 w_1 v_2$ を $\Pi_{S_1}(M)$ に加える。

(a-3) ; $(\neg \exists w_1) \{ \phi_1 \} \wedge (\exists w_2) \{ \phi_2 \}$ であるときは、(a-2) の場合と同様。

(b) $(\neg \exists w_1) \{ \phi_1 \} \wedge (\neg \exists w_2) \{ \phi_2 \}$ であり、

(b-1) ; $(\exists a_1 v_1 \neg \exists v') \{ u_1 a_1 v \in N(M) \wedge u_2 a_1 v' \in N(M) \} \vee (\exists a_2 v_2 \neg \exists v') \{ u_2 a_2 v \in N(M) \wedge u_1 a_2 v' \in N(M) \}$ ($a_1 a_2 \in \Sigma$) であるときには、 $\Pi_{S_0}(M)$ と Affix 条件 A_0 によって、相異の規定はなされる。

(b-2) ; (b-1) 以外の場所で、 $(\exists w' v'_1, \exists w' v'_2) \{ u_1 w' v'_1 \in N(M) \wedge u_2 w' v'_2 \in N(M) \wedge P_M(w'; u_1, v'_1) \neq P_M(w'; u_2, v'_2) \}$ であるときには、 $u_1 w' v'_1, u_2 w' v'_2$ を基本記号列と同様にみなして、 $P_M(w'; u_1, v'_1), P_M(w'; u_2, v'_2)$ の相異を、(a) ~ (b) に相当した場合に応じて、同様の手順による規定を行う。

(c) $(\neg \exists w_1) \{ \phi_1 \} \wedge (\neg \exists w_2) \{ \phi_2 \}$ であり、(b) 以外の場合で、 $(\exists u' z) \{ u' z \in N(M) \wedge u' z \notin N$

$(\overline{M})\} \vee \{u'z \notin N(M) \wedge u'z \in N(\overline{M})\} \wedge u' \neq u_1 \wedge u' \neq u_2\}$ のとき。 $u'z \in N(M)$, または、 $u'z \notin N(M)$ に応じて、 $\circ u'z$, 又は、 $\times u'z$ を $\Pi_{S_1}(M)$ に加える。(ただし、このような処置は、 u' に続く部分記号列に沿った写像規則系列と、その他の位置における写像規則系列との相異が考えられるときになされる。)

以上の手順を、基本記号列中のあらゆる可能な同じ部分記号列の対に対して行う。ただし、以上のようにして得られる第1種副次記号列には、規定の仕方が他の規定と重複しているような場合も出るのでそれを最小化するためには、 **minimum cover** の手法を適用する。

このようにして得られる第1種副次記号列集合 $\Pi_{S_1}(M)$ と、先に示した第0種副次記号列集合 $\Pi_{S_0}(M)$ との和集合を、副次記号列集合と呼び、 $\Pi_S(M)$ で表す。

このとき、次の関係 $[A_1]$ が成立する。

$$[A_1] \quad u_0 = \text{First} \{ w \mid u_1 w \in \Pi_C(M) \} = \text{First} \{ w' \mid u_2 w' \in \Pi_C(M) \} \text{ であり、 } u_2 \leq u_{-1} \text{ とするとき、もし } (\exists u_1) \{ u_1 w_1 \in \Pi_C(M) \wedge u_2 w_1 \notin \Pi_C(M) \} \text{ であるならば、}$$

$$\neg \{ u_1 w_0 \in \Pi_F(M) \wedge u_2 w_0 \in \Pi_F(M) \wedge C_M(u_1) = C_M(u_2) \} \quad (3.9)$$

この関係を、前と同様に、再構成される DPDA M' において満たされるべき条件とみなしたとき、

“Affix 条件 A_1 ” といひ、以後、この条件も DPDA 再構成時に際して考慮に入れる。

M が SE 構造を有しない場合には、以上のようにして得られる $\Pi_F(M) \cup \Pi_S(M)$ によって、 M の推移構造は完全に規定される。 M が SE 構造を有する場合には、 $\Pi_{SE}(M)$ が採り入れられ、これによって SE 構造も正しく規定される。すなわち、 $\Pi_F(M) \cup \Pi_{SE}(M) \cup \Pi_S(M)$ を $\Pi_C(M)$ とおいたとき、それから、Affix 条件考慮のもとで構成される表層状態最小の DPDA は、 M と同じ推移構造をなして、同じ入出力特性をもち、結局、この $\Pi_C(M)$ を、このような基準のもとにおいて、 M の代表記号列集合とすることができる。

3.4 例題

下記の DPDA M_D に対する、代表記号列集合 $\Pi_C(M_D)$ を求める。

$$M_D = (\{ p_0 \}, \{ a, b, c, d, \# \}, \{ S_0, A, B, C, D, Z \}, \delta, p_0, S_0, \phi)$$

ここで、 δ ;

$$1. \delta(p_0, a, S_0) = (p_0, AZ) \quad 2. \delta(p_0, b, A) = (p_0, \varepsilon) \quad 3. \delta(p_0, \#, Z) = (p_0, \varepsilon)$$

$$4. \delta(p_0, a, A) = (p_0, B) \quad 5. \delta(p_\#, c, B) = (p_0, A) \quad 6. \delta(p_0, a, B) = (p_0, ACCD)$$

$$7. \delta(p_0, b, C) = (p_0, \varepsilon) \quad 8. \delta(p_0, d, C) = (p_0, \varepsilon) \quad 9. \delta(p_0, b, D) = (p_0, \varepsilon)$$

また、 $a < b < c < d < \#$ とする。

$$\Pi_F(M_D) = \{ ab\#, aacb\#, aaa$$

$$bbbb\#, aaabbbdb\# \}$$

$$\Pi_{SE}(M_D) = \{ aaaaaabbbbbbb\# \}$$

$$\Pi_S(M_D) = \{ \circ a a a b d b b \#, \circ a a c$$

$$a c b \#, \times a a a b b b a c b \#,$$

$$\times a a a b b d a c b \# \}$$

$$\Pi_C(M_D) = \Pi_F(M_D) \cup \Pi_{SE}(M_D)$$

$$\cup \Pi_S(M_D)$$

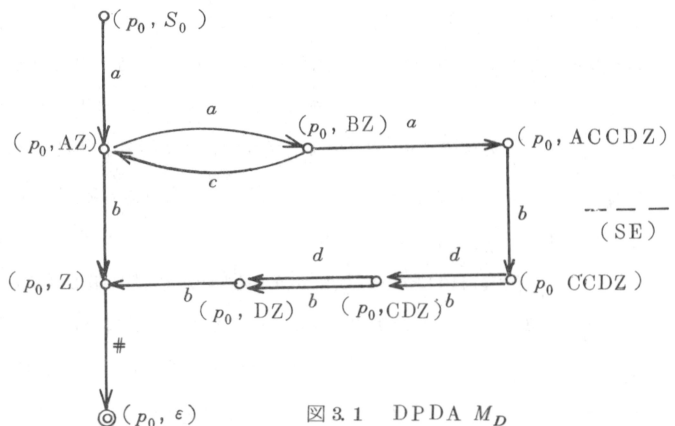


図 3.1 DPDA M_D

4 代表記号列集合による DPDA の適応的修正

4.1 修正過程の前提条件および基本方針

誤りが内在する可能性のある DPDA M_0 が与えられたときに、適切で効果的な質問を發し、その答えによって DPDA を適応的に逐次修正し、目的とする正しい DPDA、あるいはそれと同じ入出力特性をもつ DPDA へと導くことを考える。

目的とする DPDA M_D の内部構造はもちろん未知であるが、教師は、ある記号列 w が提示されたとき、それが M_D に受理されるべきものであるか否かは、完全に答えることができるものとする。

さて、前節で述べたように、 M_0 の代表記号列集合 $\Pi_C(M_0)$ は、そこで述べた解釈基準のもとにおいて、完全に M_0 を表現しているものであり、 M_0 が誤りをもっている場合には、 $\Pi_C(M_0)$ は $\Pi_C(M_D)$ とは異なったものとなっている。従って、いま述べた DPDA 修正の過程は、言い換えれば、いかにして、 $\Pi_C(M_0)$ を逐次的に $\Pi_C(M_D)$ に近づけてゆくかという問題と考えることもでき、このような方針のもとに、与えられた DPDA M_0 を適応的に修正してゆく方法について、以下に述べる。

4.2 質問応答系による、有効な記号列集合の生成

M_0 を出発点として修正をくりかえしていったある時点における DPDA を M_i とするとき、まず、それ自身を規定している、その代表記号列集合 $\Pi_C(M_i)$ を求める。続いて、それが目的とするものであるかどうかをチェックするため、 $\Pi_C(M_i)$ の要素となっている記号列の、目的とする DPDA M_D における望ましい応答を教師に質問し、もし M_i が誤った応答をしている場合には、そのような記号列の全てに対して、その先頭に付けられている受理あるいは非受理の指示を、正しい応答へと訂正し、このようにして得られる、望ましい受理、非受理の指示をもった記号列だけから成る集合を、 $\Pi_{CD}(M_i)$ で表わす)。すなわち、

$$\square x \quad \Pi_C(M_i) \ \& \ x \in N(M_D) \text{ ならば } \circ x \in \Pi_{CD}(M_i)$$

$$\square y \quad \Pi_C(M_i) \ \& \ y \notin N(M_D) \text{ ならば } \times y \in \Pi_{CD}(M_i)$$

とする（ここで、 $\square$ 印は、 $\circ$ （受理）または $\times$ （非受理）を表わす）

このような $\Pi_{CD}(M_i)$ は、 $\Pi_C(M_i)$ よりも、より $\Pi_C(M_D)$ に近いものと言える。しかし、まだ $\Pi_C(M_D) - \Pi_{CD}(M_i) \neq \emptyset$ である場合には、 $\Pi_{CD}(M_i)$ だけでは $\Pi_C(M_D)$ に不足するような記号列集合を、さらに求めてゆく必要がある。これより、このような候補となる記号列集合を効果的に質問応答系によって生成する方法を考える。

4.2.1. 付加基本記号列集合 $\Pi_F(M_i)$

$\Pi_C(M_{i-1})$ の記号列の応答を全て正しく訂正して得られる記号列集合 $\Pi_{CD}(M_{i-1})$ を完全に含み、その中の記号列に対する受理、非受理の応答指示が全て正しく付けられているような集合 $\Pi_D(M_{i-1})$ をもとにして、前章で述べた全ての **Affix** 条件考慮のもとに表層状態最小の DPDA を構成した結果として、DPDA M_i^1 が得られたとする。（なお、以下において単に DPDA 構成というときには、前章で述べたこのような構成をさす。）このとき、一般には、 $\Pi_F(M_{i-1})$ の要素は全て M_i^1 によって受理されているとは限らない。すなわち、 $\Pi_{CD}(M_{i-1})$ のうちで、 $\Pi_F(M_{i-1})$ からつくられた部分を $\Pi_{FD}(M_{i-1})$ とするとき、 $\Pi_{FD}(M_{i-1})$ の要素は必ずしも受理指示とはなっていない。しかも、 $\Pi_{FD}(M_{i-1})$ の中の受理指示の要素だけでは、一般に、 $\Pi_F(M_D)$ には不足している。ところが、先に得られている M_{i-1} が、目的とする M_D にかなり近いものである場合には、たとえ $\Pi_{FD}(M_{i-1})$ 中の非受理指示要素でも、 $\Pi_F(M_D)$ に関するかなりの情報を含んでいることが期待でき、ここでは、 $\Pi_{FD}(M_{i-1})$ の非受理指示要素からも、 M_{i-1} と M_i^1 との対比によって、 $\Pi_F(M_i^1)$ だけでは不足している基本記号列の集合をさらに付加する手順を考える。ここで、

$$\begin{aligned} u x v &\in \Pi_F(M_{i-1}), u y v \in \Pi_F(M_{i-1}), \\ x v &= \text{First} (x' v' \mid u x' v' \in \Pi_F(M_{i-1})) \quad (x = \varepsilon \text{ であってもよい}) \end{aligned} \quad (4.1)$$

であり,

$$u x v \notin N(M_i^1), u y v \in N(M_i^1) \quad (4.2)$$

かつ, $u x v$ の suffix $x v$ における推移は, M_i^1 において全く定められていないとする。また $u \leq u'$ なる prefix u' に対して,

$$u' x v \in \Pi_F(M_{i-1}), C_{M_{i-1}}(u) = C_{M_{i-1}}(u') \quad (4.3)$$

であるとする。

このような場合には, たとえ $u' y v \in \Pi_F(M_D)$ であっても, 基本記号列の作り方からわかるように, 決して, $u' y v \in \Pi_F(M_{i-1})$ とはなり得ない。従って, (4.1), (4.2) 式に示すような $x v$ と $y v$ の置換関係が, $\Pi_F(M_{i-1})$ の記号列間において成立している場合には, (4.3) に示すような記号に対して

$$u' y v \quad (4.4)$$

なる記号列を作り出し, それ自身あるいは, その部分記号列が, M_i^1 には含まれていないが, M_D としては必要な推移路を構成し得るかどうかをチェックすることは, 非常に重要なことであることがわかる。このようにして,

$\Pi_D(M_i^1)$ には含まれていない新しい受理記号列が見出されたとき, それを付加基本記号列といい, M_i^1 に対してこのような操作を逐次進めることによって得られる付加基本記号列の集合を $\Pi_F'(M_i)$ で表す。

以下に, 付加基本記号列集合 $\Pi_F'(M_i)$ を求める過程を詳しく示す。

ここで, M_{i-1} が構成された段階で, $\Pi_{FD}(M_{i-1}), \Pi_D(M_{i-1})^\dagger$ は既に求められているとする。

- (1) $\Pi_D(M_i^0) = \Pi_D(M_{i-1}), \Pi_F''(M_i) = \Pi_{FD}(M_{i-1})$ とおき $j = 0$ とする。
- (2) $j = j + 1$ とし, $\Pi_D(M_i^{j-1})$ より, 全 Affix 条件考慮のもとに最小表層状態の DPDA を構成して, それを M_i^j とおき, $\Pi_D(M_i^j)^\ddagger$ を求める。さらに, $\Pi_F'(M_i^j) = \phi$ と設定する。
- (3) $\Pi_{FD}(M_{i-1})$ を出発点としてこれまでに付加基本記号列を求めてきた過程にはまだ現われていない, (4.1), (4.2) 式に示すような置換関係をもつ記号列, および (4.3) に示すような記号列が $\Pi_F''(M_i)$ の中に存在すれば ($\Pi_{FD}(M_{i-1})$ が非受理指示の記号列を含まなければ, もちろんこのような記号列は存在しない), 次へ, さもなくば, (6) へ。
- (4) 新たに得られたあらゆる置換関係に対して, (4.4) に示したような記号列を全て求めあげ, それぞれに対して, 受理, 非受理の正しい応答指示を付して, $\Pi_F''(M_i)$ に加える。このとき受理を指示されるものはすべて付加基本記号列として採用し, $\Pi_F'(M_i^j)$ に加える。
- (5) $\Pi_F''(M_i)$ の要素であって M_i^j では非受理となる記号列

$w z = w z_1 z_2$ において, $P_{M_i^j}(w; \varepsilon, z)$ は, M_i^j において完全に定められており, suffix z 部分の推移は, 全く定められていないとする。このとき, M_i^j のとるある計算状態 (C) に対して, $\text{First} \{ z' \mid z' : (C) \mid_{M_i^j}^*(\tau, \varepsilon) \} = z_0$ とおいたとき, $w z_1 z_0 \in N(M_D)$ となるならば, $w z_1 z_0$ を付加基本記号列として採用する。以上の操作は, M_i^j のとるあらゆる計算状態に対して, および, z のあらゆる分割の仕方 $z_1 z_2$ に対して行う。さらにこの操作を, 前記のような記号列 $w z$ の全てに対して行い, ここで得られる新しい付加基本記号列の集合も $\Pi_F'(M_i^j)$ に加える。 $\Pi_F'(M_i^j) \neq \phi$ であるときには, この $\Pi_F'(M_i^j)$ も $\Pi_F'(M_i)$ に加え,

$\dagger \Pi_D(M_{i-1})$ としては, $\Pi_{CD}(M_{i-1})$ の他, 後に 4.2.2 および 4.2.3 項で述べるような $\Pi_S'(M_{i-1}), \Pi_{SE}'(M_{i-1})$ もさらに加えて, $\Pi_D(M_{i-1}) = \Pi_{CD}'(M_{i-1}) \cup \Pi_S'(M_{i-1}) \cup \Pi_{SE}'(M_{i-1})$ とする。

$$\ddagger \Pi_D(M_i^j) = \Pi_{CD}(M_i^j) \cup \Pi_S'(M_i^j) \cup \Pi_{SE}'(M_i^j)$$

さらに, $\Pi_D(M_i^j) = \Pi_D(M_i^j) \cup \Pi_F'(M_i^j)$ において, (2)へ戻る。

$\Pi_F'(M_i^j) = \emptyset$ である場合には次へ。

- (6) $\Pi_F'(M_i^1)$ に対する付加基本記号列集合の追加操作は終了し $\Pi_F'(M_i) = \bigcup_j \Pi_F'(M_i^j)$, DPDA M_i^j を改めて M_i と名付け直す。

4.2.2. 付加副次記号列集合 $\Pi_S'(M_i)$

前項の操作によって $\Pi_F(M_D)$ は全て得られたとしても, それらを正しく識別する最小表層状態の DPDA M_i が構成されたとき, そこにおいて推移中にとられる計算状態は, 一般にはまだ完全には正しく分離区別されていない。このような区別のためには, $\Pi_S(M_D)$ が中心的役割を果たすが, M_i において誤って計算状態の統合が行われてしまっている場合には, $\Pi_S(M_i)$ を単に求めただけでは, 一般にそれらの区別のための記号列は入ってこず, それに関する質問を教師に対して行うだけでは不十分であり, さらに $\Pi_S(M_D)$ に不足するような記号列集合を求める必要がある。

$$\begin{aligned} & \text{いま } u_0, v_0 \in \Pi_F(M_i), u_0, v \in \Pi_F(M_i), u, v \in \Pi_F(M_i), C_{M_i}(u_0) = C_{M_i}(u) (= (C_u) \text{ とおく}), \\ & u_0 = \text{First} \{ u \mid u; (p_0, S_0) \mid_{M_i}^* (C_u) \}, \\ & v_0 = \text{First} \{ v \mid v; (C_u) \mid_{M_i}^* (r, \epsilon) \} \quad (4.5) \end{aligned}$$

であり, $u_0, v_0, u, v \in N(M_D)$ であるが, $u, v \notin N(M_D)$ であるとする。ただし, ここで, (4.5) 式において, $u_0 \neq \epsilon$ である場合には, $u_0 = u_{0p} T(u_0)^\dagger$, $u = u_p T(u)$ とおき, $P_{M_i}(T(u_0); u_{0p}, v_0) \neq P_{M_i}(T(u); u_p, v_0)$ であるものとしておく。明らかにこの場合には, $C_{M_D}(u_0) \neq C_{M_D}(u)$ であり, このような指示は, $\times u, v \in \Pi_S(M_D)$ によってなされる。従って M_i に対しては, このような記号列 u, v も積極的につくり出してその正しい応答を質問することにより, その推移構造を知ることができ, 続いて, その正しい規定を行なうことができる。このため, (4.5) 式の関係にあるような M_i 中の計算状態 (C_u) に対しては, $u'_0 = u_{0p} \cdot T(u'_0) = \text{First} \{ u'_0 \mid C_{M_i}(u'_0) = (C_u), P_{M_i}(T(u'_0); u_{0p}, v_0) = P_{M_i}(T(u); u_p, v_0) \}$ なる prefix u'_0 を取り出して, それに suffix v を加えた記号列 $u'_0 v$ を得, その応答を質問することにする。このような suffix v としては, M_i において (C_u) を初期計算状態とみなしたときに得られる基本記号列集合に相当するものの内(ただし, $v \neq v_0$) から記号列の優先順序に従って逐次採用してゆき, はじめて非受理すべきとの指示が与えられる時点, あらういはそのような suffix の集合に対しての全てのチェックをし終るまで, 前記の質問をくり返す。

以上の手順を (4.5) 式の (C_u) に相当するような全ての推移合流点における計算状態に対して行い, このとき質問を行った新しい記号列に対して, 正しい受理, 非受理の指示を付けたものを付加副次記号列といい, その集合を $\Pi_S'(M_i)$ で表わす。4.2.1項で述べた, M_i に対する記号列集合 $\Pi_D(M_i)$ の中には $\Pi_{C_D}(M_i)$ の他に, この $\Pi_S'(M_i)$ も必ず含ませるものとする。

このような記号列の指示が加えられることによって, 計算状態の相異が正しく規定された DPDA M_{i+1} が得られると, そこにおける写像規則系列の相異は, Affix 条件考慮のもとに $\Pi_C(M_{i+1})$ ($\supset \Pi_S(M_{i+1})$) によるチェックにより, さらに, 後述する記号列集合 $\Pi_{SE}'(M_{i+1})$, $\Pi_{S_0}'(M_{i+1})$ を求める過程において調べ上げられる。

4.2.3 付加SE規定記号列集合 $\Pi_{SE}'(M_i)$

M_D が SE 構造をもっている場合には, 以上のようにして得られる $\Pi_D(M_i)$ だけでは, その構造が正しく規定されない場合があり, 従って, ここではそのために, $\Pi_{SE}(M_D)$ に相当した記号列集合をつくり出すことを考える。

ここで, $u, v, w \in \Pi_D(M_i)$, $u, v, w \in N(M_D)$; $u, x, w, y, v \in \Pi_D(M_i)$, $u, x, w, y, v \in \Pi_D(M_i)$, $H(x) \neq H$

$\dagger T(u)$ は記号列 u の最右端の記号を表わす。

$$(w), C_{Mi}(u) \neq C_{Mi}(uw), C_{Mi}(uw) = C_{Mi}(uxwy) \quad (4.6)$$

なる記号列 uwv , $uxwyv$ が存在するときには, $P_{MD}(w; u, v) = P_{MD}(w; ux, yv)$ であって, $ux^n w y^n v \in N(M_D)$ ($n = 1, 2, \dots, \infty$) なる SE 構造が M_D がもつかどうかを調べるために, M_i における写像規則系列のいかにかわからず, $uxwyv$ 中の w を xy で置換した記号列 $ux^2 wy^2 v$ をつくり出して, それに対する正しい応答を質問する。 M_i において, このような関係にある置換によって新しく得られる記号列を全て求め上げ, それぞれに正しい応答を付加して得られるような集合を $\Pi'_{SE}(M_i)$ で表わし, そのうちで特に受理指示の付けられた記号列を, 付加副次記号列と呼び, そのようなもの全体の集合を $\Pi'_{SE}(M_i)$ で表わす。この $\Pi'_{SE}(M_i)$ も, $\Pi_D(M_i)$ の中に必ず加えるものとする。なお $\Pi''_{SE}(M_i)$ 中の非受理指示の記号列は, SE 構造を否定するためのものであり, この規定は, Affix 条件 A_0 , あるいは $\Pi_S(M_i)$ を正しい応答に直したもので, または $\Pi'_S(M_i)$ によってもなされるので, 特には重要でない。

4.3 DPDA 構成時における条件の検討

これまで, ある段階で求められた, 望ましい応答指示を持つ記号列だけから成り立っている集合 $\Pi_D(M_{i-1})$ から DPDA M_i を構成するに際しては, 前章で述べた全ての Affix 条件を考慮に入れることとしていた。ところで, 先ず $\Pi_D(M_{i-1})$ の中に, $\Pi_{S_0}(M_D)$ に相当する記号列集合が全て入っているという保証はないので, このような条件付けのもとにおける DPDA 構成は, 必ずしも常に適切であるとは言えない。従って, 前記のようにして得られた M_i において, $\Pi_C(M_i) \cup \Pi'_S(M_i) \cup \Pi'_{SE}(M_i)$ に対する応答が全て正しいことがわかった場合にも, 続いて, $\Pi_D(M_i)$ から, Affix 条件 A_0 だけははずして, その他は前章で述べたのと同じ条件のもとで, $\Pi_D(M_i)$ を正しく識別し, 表層状態が最小である可態な DPDA M_i^j を全て構成し, それぞれの M_i^j に対する $\Pi_D(M_i^j)$ を求め上げる。もしこのとき, それが $\Pi_D(M_i)$ に含まれない要素を含むならば, そのうちで受理指示のものを取り出し, このような記号列を付加第 0 種副次記号列という。このようにして, 全ての M_i^j から求められる付加第 0 種副次記号列の集合を $\Pi'_{S_0}(M_i)$ で表わし, これも付け加えたものを新たな $\Pi_D(M_i)$ として, 次の DPDA M_{i+1} の構成に進むことにする。

次に, $\Pi'_{S_0}(M_i) = \emptyset$ であることがわかった場合にも, まだ $\Pi_D(M_{i-1})$ の中に, $\Pi_F(M_D)$ に相当する記号列集合が完全に含まれていたという保証はないので, DPDA M_i の構成に際して, Affix 条件 A_F を考慮に入れていたことも, 必ずしも適切であったとは言えない。従って, 続いて, 前章で述べた全ての Affix 条件をとりはずし, 従って (3.6) 式の関係も無視して, $\Pi_D(M_i)$ を正しく識別する最小状態の DPDA を全て求め上げ, 前記と同様にして, $\Pi_D(M_i)$ に含まれない新しい記号列が生じれば, そのような新しい応答を指示するもの全てを集め, そのうちで受理指示のものだけをとり出して, その集合を $\Pi_A(M_i)$ で表わす。なお, ここにおける Affix 条件を無視した DPDA 構成に際しても, 前提条件として M_D の受理する記号列は, $\{w\# \mid w \in (\Sigma - \#)^*, \# \text{ は終止記号}\}$ なる形式のものに限るとしているので, これ以外の記号列を受理するような DPDA は, 正しい DPDA としては採用しない。さらに, これまでの修正過程において, $\bigcup_k \Pi_D(M_k)$, $\bigcup_k \Pi'_F(M_k)$, $\bigcup_k \Pi''_{SE}(M_k)$ に対する応答指示は, 全て正しいものとしているので, これらの指示を満足しないような DPDA は採用しない。前記のようにして得られる集合 $\Pi_A(M_i)$ も, 新たに $\Pi_D(M_i)$ に付け加えて, 続く DPDA M_{i+1} の構成に用いる。

以上のようにして系統的に得られる $\Pi_D(M_i)$ の他に, 全く他からの条件より, ある記号列の応答を質問される場合も考え, そのような記号列の, M_i における応答が誤りであると教えられた場合には, それに正しい応答指示を付けた記号列もさらに考慮する必要がある。このような記号列全体の集合を $\Pi_E(M_i)$ で表わし, 次の DPDA の構成に先立って, $\Pi_D(M_i)$ に付け加える。

ここで, 目的とする M_D に対して, M_0 の受理する記号列の推移過程に関与する部分の推移構造抽出に関しては, 以上

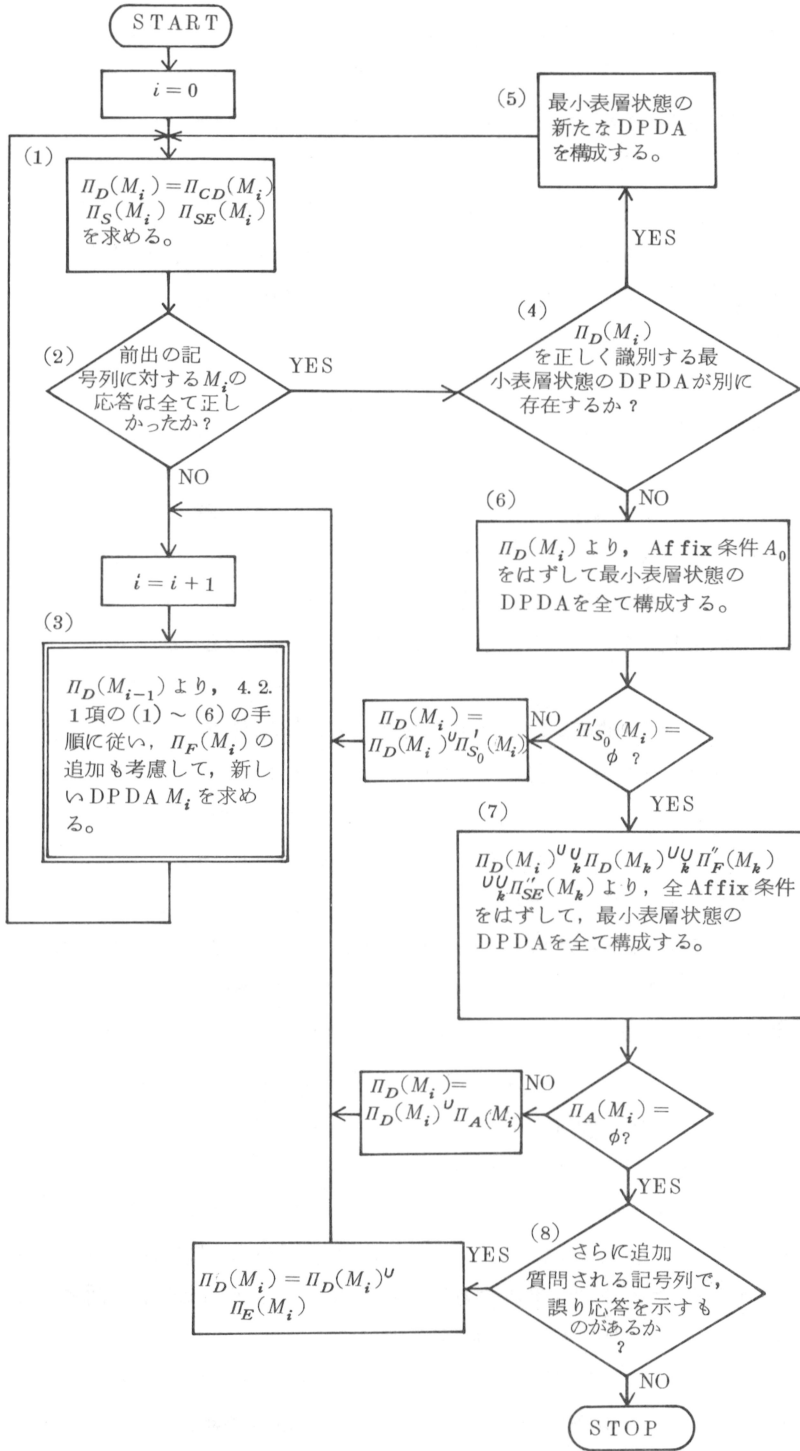


図 4.1 DPDA の適応的修正過程

のような手順によって正しく部分収束し、さらに、手順の最後において得られている $\Pi_D(M_i)$ が、 $\Pi_C(M_D)$ を完全に含むようになっていれば、それから構成される DPDA M_{i+1} は、完全に目的とするものとなって、修正過程は正しく終了する。

4.4 DPDAの適応的修正過程

前節までに述べた、DPDAの適応的修正の過程を、図 4.1 にフローチャートのかたちでまとめて示す。

4.5 例題

図 4.2 に示される DPDA M_0 が与えられたときに、これを修正してゆく過程を示す。($a \leq b < c \leq d \leq \#$ とする) 目的とする DPDA としては、前章の例題で用いた M_D あるいは、それと同じ入出力応答を示すものを想定する。

$$[0] \quad \Pi_{CD}(M_0) = \{ \times a \#, \circ a b \#, \times a a b \#, \times a a d \#, \circ a a c b \#, \times a a a b b b \# \}$$

$$\Pi'_S(M_0) = \{ \times a b b \#, \times a a b b \#, \times a a d b \#, \times a a c d \#, \times a a a b b d \# \}, \quad \Pi'_{SE}(M_0) = \phi$$

[1-1] 次に、 $\Pi_D(M_0) = \Pi_{CD}(M_0) \cup \Pi'_S(M_0)$ より図 4.3 の DPDA M_1^1 が得られる。

$$\Pi_{CD}(M_1^1) = \{ \circ a b \#, \circ a a c b \# \}, \quad \Pi'_F(M_1^1) = \{ \circ a a a b b b b \# \}$$

$$\Pi'_S(M_1^1) = \{ \circ a a c a c b \# \}, \quad \Pi'_{SE}(M_1^1) = \phi$$

[1-2] $\Pi_D(M_1^1) = \Pi_{CD}(M_1^1) \cup \Pi'_F(M_1^1) \cup \Pi'_S(M_1^1)$ より DPDA を構成すると、 M_1^1 にさらに次の 2 写像規則が追加された DPDA M_1^2 が得られる。

$$\partial(p_0, a, B) = (p_0, CCC), \quad \partial(p_0, b, C) = (p_0, \varepsilon)$$

$$\Pi_{CD}(M_1^2) = \{ \circ a b \#, \circ a a c b \#, \circ a a a b b b b \# \}, \quad \Pi'_F(M_1^2) = \{ \circ a a a b b d b \# \},$$

$$\Pi'_S(M_1^2) = \{ \circ a a c a c b \#, \times a a a b b b a c b \# \}, \quad \Pi'_{SE}(M_1^2) = \{ \circ a a a a a b b b b b b b \# \}$$

[1-3] $\Pi_D(M_1^2) = \Pi_{CD}(M_1^2) \cup \Pi'_F(M_1^2) \cup \Pi'_S(M_1^2) \cup \Pi'_{SE}(M_1^2)$ より DPDA を構成すると、図 4.4 の DPDA M_1^3 が得られる。さらに新しい付加基本記号列は得られず、 M_1^3 を M_1 と名付け直す。

[2] $\Pi_D(M_1) = \Pi_D(M_1^3) \cup \{ \times a a a b b d a c b \# \}$ からは、 M_D の 6 番目の写像規則が $\partial(p_0, a, B) = (p_0, ACCD)$ に代ったような DPDA M_2 が得られる。次は、図 4.1 のステップ(6)まで進み、 $\Pi'_{S0}(M_2) = \{ \circ a a a b d b b \# \}$ が得られる。

[3] $\Pi_D(M_2) = \Pi_D(M_1) \cup \Pi'_{S0}(M_2)$ より DPDA を構成すると、前章例題の DPDA M_D が得られる。ここで、 $\Pi_D(M_2)$ は、 $\Pi_C(M_D)$ と一致している。

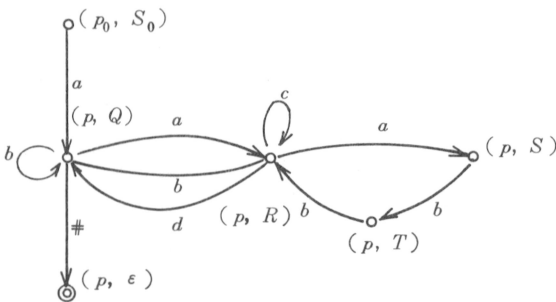


図 4.2 DPDA M_0

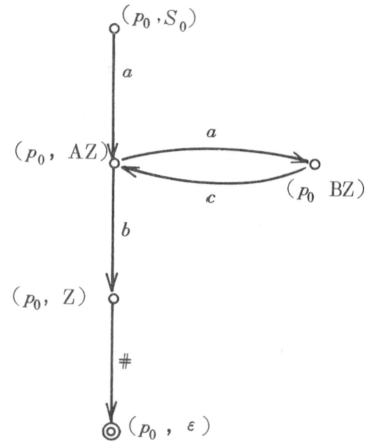


図 4.3 DPDA M_1^1

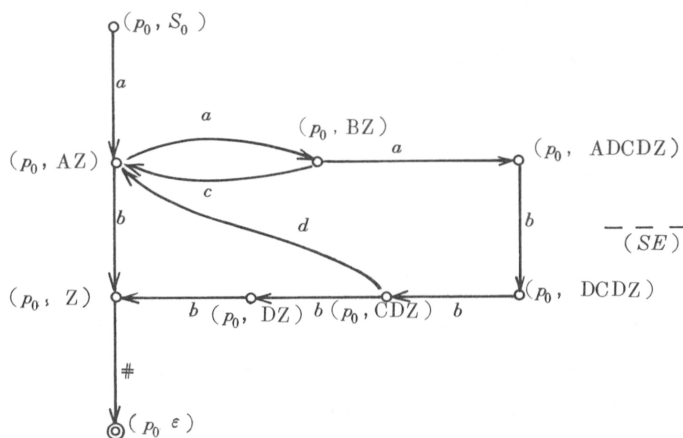


図 4.4 DPDA $M_1^3(M_1)$

5. あとがき

以上においては、入力記号間にアルファベット順が定められているものとしたが、この仮定は容易に除き得る。またここに示した方法は、さらに高いクラスのオートマトン、変換器に対しても拡張することができ、実際の記号列処理機構の適応的構成・修正に対しても供し得るようにできるものと考えられる。

最後に、本学修士課程在学中に本研究の基礎段階に対して大きい貢献をした真汐雅彦氏（現、電々公社）、および、文献(1)に従ったDPDA構成法の全体、および、本論文第3章のアルゴリズムの一部の計算機プログラム作製をそれぞれ担当した卒研生の貝塚真生君、牧野哲男君に感謝いたします。また、本研究の一部は、文部省の科学研究費の援助を受けていることを記し、謝意を表します。

参考文献

- (1) H. Enomoto, E. Tomita and S. Doshita ; "Synthesis of Automata that Recognize Given Strings and Characterization of Automata by Representative Sets of Strings," 1-st USA-Japan Computer Conference proc. pp21-27 (Oct, 1972)
- (2) 榎本, 堂下, 真汐, "オートマトンの対話型修正方法", 信学会オートマトン研資(1970-11)
- (3) J. Hopcroft and J. Ullman ; "Formal Languages and their Relation to Automata", Addison-Wesley Publishing Co. (1969)
- (4) 榎本, 堂下, 富田 ; "サンプル記号列を識別する最簡オートマトンの構成," 信学論(D), 55-D, 3, pp210-217 (昭47-03)

本 PDF ファイルは 1965 年発行の「第 6 回プログラミング・シンポジウム報告集」をスキャンし、項目ごとに整理して、情報処理学会電子図書館「情報学広場」に掲載するものです。

この出版物は情報処理学会への著作権譲渡がなされていませんが、情報処理学会公式 Web サイトの https://www.ipsj.or.jp/topics/Past_reports.html に下記「過去のプログラミング・シンポジウム報告集の利用許諾について」を掲載して、権利者の検索をおこないました。そのうえで同意をいただいたもの、お申し出のなかったものを掲載しています。

過去のプログラミング・シンポジウム報告集の利用許諾について

情報処理学会発行の出版物著作権は平成 12 年から情報処理学会著作権規程に従い、学会に帰属することになっています。

プログラミング・シンポジウムの報告集は、情報処理学会と設立の事情が異なるため、この改訂がシンポジウム内部で徹底しておらず、情報処理学会の他の出版物が情報学広場(=情報処理学会電子図書館)で公開されているにも拘らず、古い報告集には公開されていないものが少なからずありました。

プログラミング・シンポジウムは昭和 59 年に情報処理学会の一部門になりましたが、それ以前の報告集も含め、この度学会の他の出版物と同様の扱いにしたいと考えます。過去のすべての報告集の論文について、著作権者(論文を執筆された故人の相続人)を探し出して利用許諾に関する同意を頂くことは困難ですので、一定期間の権利者搜索の努力をしたうえで、著作権者が見つからない場合も論文を情報学広場に掲載させていただきたいと思います。その後、著作権者が発見され、情報学広場への掲載の継続に同意が得られなかった場合には、当該論文については、掲載を停止致します。

この措置にご意見のある方は、プログラミング・シンポジウムの辻尚史運営委員長(tsuji@math.s.chiba-u.ac.jp)までお申し出ください。

加えて、著作権者について情報をお持ちの方は事務局まで情報をお寄せくださいますようお願い申し上げます。

期間：2020 年 12 月 18 日～2021 年 3 月 19 日

掲載日：2020 年 12 月 18 日

プログラミング・シンポジウム委員会

情報処理学会著作権規程

<https://www.ipsj.or.jp/copyright/ronbun/copyright.html>