

一意解析可能アレイ文法による単連結図形及び単純閉曲線の生成
 Generation of Simply-Connected Patterns and Simple Closed Curves
 by Uniquely Parsable Array Grammars

齊金山 ・ Ruhizan Liza Ahmad Shauri ・ 森田 憲一
 Jin-Shan Qi, Ruhizan Liza Ahmad Shauri, and Kenichi Morita

広島大学工学部
 Faculty of Engineering, Hiroshima University

概要

一意解析可能アレイ文法 (uniquely parsable array grammar, UPAG) は等形アレイ文法 (isometric array grammar, IAG) の一種であり, バックトラッキングなしに構文解析を行なえるという性質を持つ. 従って, もしある図形の集合が UPAG で正確に記述できるならば, その UPAG による図形認識は効率よく (生成と同じステップ数で) 実行できる. 本論文では, 図形の位相的性質を UPAG によって記述する方法について考察し, 特に全ての単連結図形および全ての単純閉曲線をそれぞれ生成・認識できる簡潔な UPAG を与えた.

キーワード: アレイ文法, パターン生成, 一意解析可能性, 決定性解析, 単連結図形, 単純閉曲線.
 Keywords: array grammar, pattern generation, unique parsability, deterministic parsing, simply-connected patterns, simple closed curve.

1 まえがき

等形アレイ文法 (isometric array grammar, IAG) は Rosenfeld [4] によって提案された 2 次元パターン生成・認識の形式モデルである. IAG では導出と還元過程で図形の変形を避けるために書換規則の左辺と右辺が等形であるように要求されている. 等形性を保つため, 必要なときに空白記号 # で補充することが許されている. そのため空白記号 # が文脈の働きを持つことが可能となり, 2 次元文法で Chomsky 階層の最も低いサブクラスである正規アレイ文法 (regular array grammar, RAG) でさえも比較的高い生成能力を持つことになる. しかしながらこのような能力のため, IAG での 2 次元言語の解析は非常に難しくなる. 例えば RAG の認識問題は NP 完全問題となることが知られている [3]. 一方, Yamamoto and Morita [5] による一意解析可能アレイ文法 (uniquely parsable array grammar, UPAG) は書換規則にある制約を加えることにより構文解析がバックトラッキングなしに実行できるようにしたアレイ文法である. つまり還元過程がある種の合流性を持ち, これにより効率よく構文解析することが可能である. Morita and Imai [2] は図形の位相的性質を UPAG によって記述する方法を論じ, Beyer のアルゴリズム [1] に基づいて, すべての連結図形を生成・認識できる単純な UPAG を与えている. 本稿では, これらの研究を元にして, 単連結図形 (内部に穴を持たない連結図形, simply-connected pattern) と単純閉曲線 (simple closed curve) の生成と認識が可能な UPAG $G_{s\text{-connect}}$ と $G_{sc\text{-curve}}$ を与える.

2 諸定義

Σ を記号の空でない有限集合とする. Σ 上の 2 次元の語とは Σ の記号からできる任意の形状の 2 次元有限連結配列である. Σ 上の全ての 2 次元の語の集合を Σ^{2+} と記す. (但し, 空語は Σ^{2+} に含まれていな

定義 2.1 等形アレイ文法 (*isometric array grammar*, IAG) とは 5 項組

$$G = (N, T, P, S, \#),$$

である。但し, N は非終端記号の有限かつ空でない集合, T は終端記号の有限かつ空でない集合 ($N \cap T = \emptyset$), $S (\in N)$ は開始記号, $\# (\notin N \cup T)$ は特別な空白記号である。 P は $\alpha \rightarrow \beta$ の形の書換規則の有限集合である。但し, α と β は $N \cup T \cup \{\#\}$ 上の語であり, 次の条件を満たしている:

1. α と β は幾何学的に同じ形をしている。
2. α は少なくとも一つの非終端記号を含んでいる。
3. α に含まれる終端記号は書換規則 $\alpha \rightarrow \beta$ によって書き換えられることはない。
4. 書換規則 $\alpha \rightarrow \beta$ の適用はアレイの連結性に影響しない。(詳細は [4] を参照。)

η を $N \cup T$ 上の語, $\alpha \rightarrow \beta$ を書換規則とする。 η を $\#$ の 2 次元無限配列中に埋め込んだものを $\eta_{\#}$ と書く (なお, $\eta_{\#}$ は $\mathbb{Z}^2 \rightarrow N \cup T \cup \{\#\}$ なる写像として表現できることに注意 (\mathbb{Z} は整数の集合))。 α が $\eta_{\#}$ の中に部分配列として現れているとき, 書換規則 $\alpha \rightarrow \beta$ は η に適用可能と言う。 $\eta_{\#}$ に対してこの書換を実行することにより $\zeta_{\#}$ が得られるなら, η から ζ が直接的に導出されたと言い, $\eta \Rightarrow \zeta$ と書く。導出の関係 \Rightarrow は \Rightarrow の反射的推移的閉包として定義される。また, n ステップの導出を $\stackrel{n}{\Rightarrow}$ と書く。直接的導出の定義と同様に, 書換規則 $\alpha \rightarrow \beta$ の逆方向適用に基いて直接的還元 \Leftarrow の関係が定義できる。還元 \Leftarrow , および n ステップの還元 $\stackrel{n}{\Leftarrow}$ の関係も同様に定義される。 IAG G によって生成される言語は $L(G) = \{x \mid S \stackrel{*}{\Rightarrow} x \wedge x \in T^{2+}\}$ である。

$\alpha \rightarrow \beta$ を書換規則とする。 $\alpha \rightarrow \beta$ の適用によって書換えられない (つまり, 同じ記号に書換えられる) α の部分配列を文脈部分, また, 異なる記号に書換えられる部分配列を α の書換部分と言う。 β の文脈部分と書換部分もこれと同様に定義される。

定義 2.2 $G = (N, T, P, S, \#)$ を IAG とする。もし P が次の条件を満たしていれば G を一意解析可能アレイ文法 (*uniquely parsable array grammar*, UPAG) と言う。

UPAG 条件:

1. P 中の各書換規則の右辺は $\#$ と S 以外の記号を含んでいる。
2. $r_1 = \alpha_1 \rightarrow \beta_1$ と $r_2 = \alpha_2 \rightarrow \beta_2$ を P 中の任意の 2 つの書換規則とする ($r_1 = r_2$ の場合も含む)。 β_1 と β_2 を, あらゆる可能な位置関係で重ね合わせたときに重なった部分の記号がすべて一致するような各重ね合わせに対して, 次の条件が成り立つ。
 - (a) 重なった部分はそれぞれ β_1 と β_2 の文脈部分に含まれている, あるいは
 - (b) β_1 と β_2 の全体が重なっており, かつ $r_1 = r_2$ 。

例 2.1 二つの書換規則

$$aB \rightarrow ab, \quad Ca \rightarrow ca$$

は UPAG 条件を満たしている。しかし, 次ののは満たしていない。

$$\#B \rightarrow ab, \quad Ca \rightarrow ca$$

次の定理は, S から導出される任意の語 η はバックトラッキングなしに (つまり η から始まる任意順序の還元により) S まで還元されるということを示している。

定理 2.1 [一意解析可能定理] [5] $G = (N, T, P, S, \#)$ を UPAG, $\eta \in (N \cup T)^{2+}$ を $\eta \stackrel{*}{\Leftarrow} S$ (従って $S \stackrel{*}{\Rightarrow} \eta$) であるような 2 次元の語とする。このとき, $\eta \Leftarrow \zeta$ となるようなあらゆる ζ に対して, $\eta \Leftarrow \zeta \stackrel{n-1}{\Leftarrow} S$ が

3 単連結図形と単純閉曲線を生成する UPAG

ここでは、全ての単連結図形 (simply-connected pattern) を生成できる UPAG $G_{s\text{-connect}}$ と全ての単純閉曲線 (simple closed curve) を生成できる UPAG $G_{s\text{-curve}}$ を与える. (単連結図形と単純閉曲線は4隣接の関係に基づいて定義される. これらの詳細は [4] を参照.)

3.1 単連結図形を生成する UPAG

文献 [2] では、全ての連結図形 (穴を持つものと持たないものを含む) を生成する簡単な UPAG G_{connect} が提案されている. この UPAG の1つの規則を変形することにより、すべての単連結図形 (穴を持たない連結図形) を生成する UPAG が得られることを示す.

記号 “a” からできる連結図形を生成する UPAG G_{connect} は次のように定義される [2].

$$G_{\text{connect}} = (\{S, A\}, \{a\}, P_{\text{connect}}, S, \#)$$

P_{connect} に含まれる書換規則は次の通り:

$$\begin{array}{ll}
 (1) \quad \begin{array}{c} \# \\ \# S \# \\ \# \end{array} \rightarrow \begin{array}{c} \# \\ \# A \# \\ \# \end{array} & (5b) \quad \begin{array}{c} A A \\ \# A A \\ A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} A A \\ \# \# A \\ A A \# \\ \# \end{array} \\
 (2) \quad \begin{array}{c} A \\ \# \# \# \\ \# \end{array} \rightarrow \begin{array}{c} A \\ \# A \# \\ \# \end{array} & (5c) \quad \begin{array}{c} \# \\ A A A \\ A A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} \# \\ A \# A \\ A A A \# \\ \# \end{array} \\
 (3) \quad \begin{array}{c} \# \\ A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} \# \\ A A \# \\ \# \end{array} & (5d) \quad \begin{array}{c} A A \\ A A A \\ A A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} A A \\ A \# A \\ A A A \# \\ \# \end{array} \\
 (4) \quad \begin{array}{c} A A \\ A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} A A \\ A A \# \\ \# \end{array} & (6) \quad A \rightarrow a \\
 (5a) \quad \begin{array}{c} \# \\ \# A A \\ A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} \# \\ \# \# A \\ A A \# \\ \# \end{array} &
 \end{array}$$

命題 3.1 [2] UPAG G_{connect} は記号 “a” からなる連結図形を全てかつそれらだけを生成する.

すべての単連結図形を生成する $G_{s\text{-connect}}$ は G_{connect} を元にして次のように定義できる.

$$G_{s\text{-connect}} = (\{S, A\}, \{a\}, P_{s\text{-connect}}, S, \#)$$

ここで、 $P_{s\text{-connect}} = P_{\text{connect}} - \{(5d)\} \cup \{(5d')\}$, つまり P_{connect} 中の (5d) を次の (5d') に置き換えたものである.

$$(5d') \quad \begin{array}{c} \# A A \\ A A A \\ A A \# \# \\ \# \end{array} \rightarrow \begin{array}{c} \# A A \\ A \# A \\ A A A \# \\ \# \end{array}$$

$G_{s\text{-connect}}$ は UPAG である. これは各規則対に対し UPAG 条件の検査を行なうことにより確かめられる (実際にはコンピュータで検査を行なっている).

定理 3.1 UPAG $G_{s\text{-connect}}$ は記号 “ a ” からなる単連結図形を全て、かつそれらだけを生成する。

この定理を証明するために次の準備を行う。 $\eta \in \{A\}^{2+}$ とする。 $\eta_{\#}$ における **SEセル**とは、その記号が A であり、かつその南(下)と東(右)隣の記号が共に $\#$ であるようなます目を言う。SEセルは図1に示すように5つの場合に分類できる。

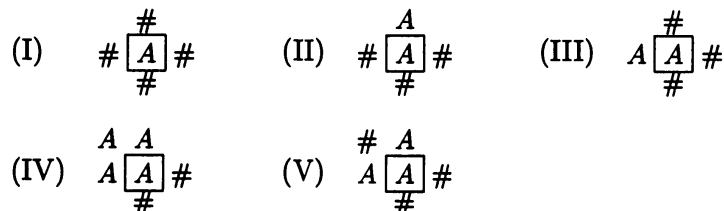


図 1: SEセル (\boxed{A} によって示す) の分類

場合 (V) はさらに図2の (a)-(j) の場合に細分できる。

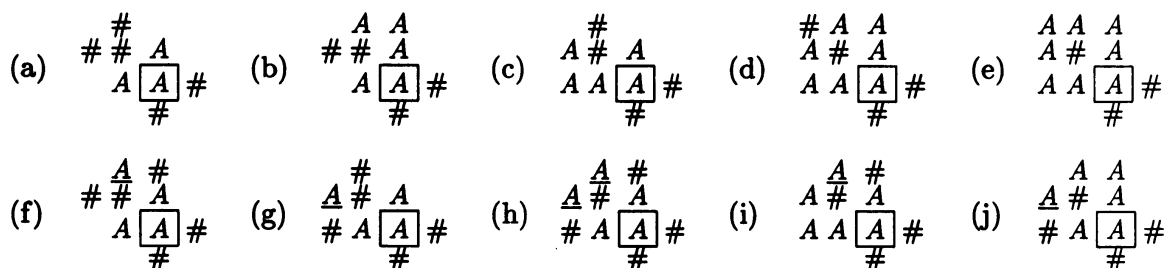


図 2: 図1における場合 (V) の細分

Σ を記号の空でない有限集合, $p: \mathbb{Z}^2 \rightarrow \Sigma \cup \{\#\}$ を無限配列とする。次のように定義される $\text{supp}(p)$ を p の台 (support) と呼ぶ: $\text{supp}(p) = \{(x, y) \mid p(x, y) \neq \#\}$. ここで, $\text{supp}(p)$ が空でない有限集合となるような p に対して x_{\min} と y_{\max} を次のように定義する: $x_{\min} = \min\{x \mid p(x, y) \neq \#\}$, $y_{\max} = \max\{y \mid p(x, y) \neq \#\}$. $\text{supp}(p)$ の各点に対し, (p の下での) インデックスを与える関数 $\text{ind}: \text{supp}(p) \rightarrow \mathbb{Z}$ を次のように定義する:

$$\text{ind}(x, y) = |x - x_{\min}| + |y - y_{\max}|.$$

さらに, p 全体に対するインデックスを次のように定義する:

$$\text{ind}^*(p) = \sum_{(x, y) \in \text{supp}(p)} \text{ind}(x, y).$$

補題 3.1 $p: \mathbb{Z}^2 \rightarrow \{\#, A\}$ を, $\text{supp}(p)$ が空でない有限集合でかつ単連結図形を形成するような任意の2次元無限配列とする。このとき p は (I)-(IV) または (a)-(d) のいずれかを部分配列として含む。

証明: $\text{supp}(p)$ が空でない有限集合だから, SEセルが少なくとも1つ存在するのは明らかである。また, $\text{supp}(p)$ が単連結図形である (穴を持たない) ことから, 部分配列 (e) は現れない。ここで p 中の全てのSEセルが (f)-(j) のいずれかであると仮定する。それらSEセルの中でインデックスの値が最も小さいものを (x_0, y_0) とする。セル (x_0, y_0) が (f) の場合を考える ((g)-(j) についても同様である)。このとき, (x_0, y_0) の左上にある \underline{A} もSEセルとなり, そのインデックスは $\text{ind}(x_0, y_0) - 3$ となり, 仮定に矛盾する。従って補題が成り立つ。 \square

補題 3.2 $p: \mathbf{Z}^2 \rightarrow \{\#, A\}$ を, $\text{supp}(p)$ が空でない有限の単連結図形を形成するような任意の2次元無限配列とする. このとき, p に対して書換規則 (1)–(4), (5a)–(5c), (5d') のいずれかが逆適用可能となる. この規則によって p を還元した結果得られる配列を p' とすると, $\text{supp}(p')$ も単連結図形を形成し, しかも $\text{ind}^*(p') < \text{ind}^*(p)$ または $\text{ind}^*(p) = \text{ind}^*(p') = 0$ が成り立つ.

証明: p に対して書換規則 (1)–(4), (5a)–(5c), (5d') のいずれかが逆適用可能となることは補題 3.1 より言える (場合 (I)–(IV) および (V) の (a)–(d) がこれらの規則の右辺に対応していることに注意). どの規則を逆適用しても図形の連結性や穴の有無が変化しないことは容易に確かめられるので, $\text{supp}(p')$ も有限の単連結図形を形成する. また, (1) 以外の規則により $\text{ind}^*(p') < \text{ind}^*(p)$ となることも, 各規則から確かめられる. 規則 (1) が逆適用できるのは $\text{supp}(p)$ が1つの点からなる集合であるときだけであり, この場合 $\text{ind}^*(p) = \text{ind}^*(p') = 0$ となる. \square

定理 3.1 の証明: UPAG $G_{S\text{-connect}}$ が記号 “a” からなる単連結図形だけを生成することは容易に分かる. なぜなら, $P_{S\text{-connect}}$ 中の規則 (1) は1つの “A” だけからなる単連結図形を生成し, また, それ以外の各規則は “A” または “a” からなる配列の連結性を保存し, かつ穴を生成することもないからである.

次に, $G_{S\text{-connect}}$ が記号 “a” からなる全ての単連結図形を生成することを示す. そのためには, 記号 “A” からなる任意の単連結図形 $w \in \{A\}^{2+}$ に対して, 還元過程 $w \xleftarrow{*} S$ が存在することを示せばよい (なぜなら, w に規則 (6) だけを繰り返し適用することにより, w と同じ形状の $x \in \{a\}^{2+}$ に対して $w \xrightarrow{*} x$ となるので, 結局 $S \xrightarrow{*} x$ が言える). $p: \mathbf{Z}^2 \rightarrow \{\#, A\}$ を, $w_{\#}$ を表す配列とする. $\text{ind}^*(p)$ の値に関する帰納法で $w \xleftarrow{*} S$ を示す. 補題 3.2 より, p (つまり $w_{\#}$) には規則 (1)–(4), (5a)–(5c), (5d') のいずれかが逆適用できる. $\text{ind}^*(p) = 0$ の場合には規則 (1) が逆適用できるので, $w \leftarrow S$ となる. $\text{ind}^*(p) = k > 0$ の場合には規則 (2)–(4), (5a)–(5c), (5d') のいずれかが逆適用でき, $\text{ind}^*(p') < k$ であるような p' で $\text{supp}(p')$ が単連結図形となるものが得られる. 従ってこの場合も $w \xleftarrow{*} S$ となる. \square

3.2 単純閉曲線を生成する UPAG

単純閉曲線は, 各要素がちょうど2つの隣接要素を持ち, 穴をちょうど1つ含むような連結図形である. なお, ここでは穴を含まないような縮退した単純閉曲線 ([4] 参照) は除外する. 記号 “a” からなる全ての単純閉曲線を生成する UPAG は次のように定義できる.

$$G_{\text{sc-curve}} = (\{S, A\}, \{a\}, P_{\text{sc-curve}}, S, \#)$$

$P_{\text{sc-curve}}$ に含まれる書換規則は次の通り:

$$(7) \begin{array}{ccc} S \# \# \# & \rightarrow & A A A \# \\ \# \# \# & & A \# A \\ \# \# \# \# & & A A A \# \\ \# \# & & \# \# \end{array}$$

$$(10) \begin{array}{ccc} \# & \rightarrow & \# \\ A A A & \rightarrow & A \# A \\ \# \# \# \# & & A A A \# \\ \# \# & & \# \# \end{array}$$

$$(8) \begin{array}{ccc} \# & \rightarrow & \# \\ \# A A & \rightarrow & \# \# A \\ A \# \# & & A A \# \\ \# & & \# \end{array}$$

$$(11) \begin{array}{ccc} \# A \# \# & \rightarrow & \# A A \# \\ A A \# & \rightarrow & A \# A \\ \# \# \# \# & & A A A \# \\ \# \# & & \# \# \end{array}$$

$$(9) \begin{array}{ccc} A \# \# & \rightarrow & A A \# \\ \# A \# & \rightarrow & \# \# A \\ A \# \# & & A A \# \\ \# & & \# \end{array}$$

$$(12) A \rightarrow a$$

$G_{\text{sc-curve}}$ が UPAG であることは, 前と同様に各規則対に対して UPAG 条件検査を行なうことにより確かめられる.

定理 3.2 UPAG $G_{sc-curve}$ は記号 “a” からできる全ての単純閉曲線を生成する.

補題 3.3 $p : Z^2 \rightarrow \{\#, A\}$ を, $\text{supp}(p)$ が単純閉曲線を形成するような任意の 2 次元無限配列とする. このとき p は図 2 の (a)–(e) のいずれかを部分配列として含む.

証明: $\text{supp}(p)$ は各点が隣接点をちょうど 2 つ持ち, かつ穴を 1 つ含む連結図形であることから, p に図 1 の部分配列 (I)–(IV) が現れることはない. また, 補題 3.1 の証明と同様にして, p 中の全ての SE セルが (f)–(j) のいずれかであると仮定すると矛盾が生じることがわかる. 従って補題が成り立つ. \square

補題 3.4 $p : Z^2 \rightarrow \{\#, A\}$ を, $\text{supp}(p)$ が単純閉曲線を形成するような任意の 2 次元無限配列とする. このとき, p に対して書換規則 (7)–(11) のいずれかが逆適用可能となる. (7) 以外の規則によって p を還元した結果得られる配列を p' とすると, $\text{supp}(p')$ も単純閉曲線を形成し, しかも $\text{ind}^*(p') < \text{ind}^*(p)$ が成り立つ.

証明: p に対して書換規則 (7)–(11) のいずれかが逆適用可能となることは補題 3.3 より言える (場合 (a)–(e) がこれらの規則の右辺に対応していることに注意). (7) 以外のどの規則を逆適用しても $\text{supp}(p')$ が再び単純閉曲線を形成することも容易にわかる. また, $\text{ind}^*(p') < \text{ind}^*(p)$ となることも, 各規則から確かめられる. \square

定理 3.2 の証明: UPAG $G_{sc-curve}$ が単純閉曲線だけを生成することは容易に分かる. なぜなら, 規則 (7) は記号 “A” だけからなる最小の単純閉曲線を生成し, また, それ以外の各規則は単純閉曲線であるという性質 (つまり, 各点が隣接点をちょうど 2 つ持ち, かつ穴を 1 つ含む) を保存する変形だからである.

次に, $G_{sc-curve}$ が記号 “a” からなる全ての単純閉曲線を生成することを示す. 定理 3.1 と同様, 与えられた任意の単純閉曲線 $w \in \{A\}^{2+}$ に対して, 還元過程 $w \stackrel{*}{\leftarrow} S$ が存在することを示せばよい. $p : Z^2 \rightarrow \{\#, A\}$ を, $w_{\#}$ を表す配列とする. $\text{ind}^*(p)$ の値に関する帰納法で $w \stackrel{*}{\leftarrow} S$ を示す. 補題 3.4 より, p (つまり $w_{\#}$) には規則 (7)–(11) のいずれかが逆適用できる. p が最小の単純閉曲線の場合には規則 (7) が逆適用できるので, $w \leftarrow S$ となる (この場合は $\text{ind}^*(p) = 16$). $\text{ind}^*(p) = k > 16$ の場合には規則 (8)–(11) のいずれかが逆適用でき, $\text{ind}^*(p') < k$ であるような p' に帰着できる. 従ってこの場合も $w \stackrel{*}{\leftarrow} S$ となる. \square

4 むすび

本稿では, 全ての単連結図形を生成・認識できる UPAG と全ての単純閉曲線を生成・認識できる UPAG を考案した. UPAG は一意解析性を持っているから, この性質を活かすと高速の解析が可能になる. これら以外の図形の位相的性質を特徴付けるような UPAG は今後の研究課題として残される.

参考文献

- [1] Beyer, W.T., Recognition of topological invariants by iterative arrays, Ph.D. Thesis, MIT (1969).
- [2] Morita, K., and Imai K., Uniquely parsable array grammars for generating and parsing connected patterns, *Pattern Recognition*, **32**, 269–276 (1999).
- [3] Morita, K., Yamamoto, Y., and Sugata, K., The complexity of some decision problems about two-dimensional array grammars, *Information Sciences*, **30**, 241–262 (1983).
- [4] Rosenfeld, A., *Picture Languages*, Academic Press, New York (1979).
- [5] Yamamoto, Y., and Morita, K., Two-dimensional uniquely parsable isometric array grammars, *Int. J. Pattern Recognition and Artificial Intelligence*, **6**, 301–313 (1992).