A Variant of Extended Two-dimensional Context-free Picture Grammar

P.S. Azeezun Nisha[1], S. Hemalatha[2],
Atulya K. Nagar[3] and K.G. Subramanian[4]

[1]Department of Mathematics,
J.B.A.S. College for Women, Chennai, India
hajma29112008@gmail.com

[2]Department of Mathematics,
S.D.N.B Vaishnav College for Women, Chennai, India
hemrav2008@gmail.com

[3]Department of Mathematics and Computer Science,
Faculty of Science, Liverpool Hope University, Liverpool, UK
nagara@hope.ac.uk

[4]Department of Mathematics, Madras Christian College,
Tambaram, Chennai 600059 India
kgsmani1948@gmail.com

June 25, 2016

Abstract

Here we introduce a variant of extended two-dimensional context-free picture grammar ($E2DCFPG$), called $(l/u)E2DCFPG$ which allows rewriting only the leftmost column, or the uppermost row of variables in a picture array. Several theoretical properties of $(l/u)E2DCFPG$ are obtained and an application in generating digitized picture arrays is discussed.

AMS Subject Classification: 68Q42

Key Words and Phrases: Two-dimensional Grammar, Picture language, Context-free grammar

1 Introduction

Motivated by problems in image processing and pattern recognition, several grammar models [2] that generate two-dimensional (2D) picture languages have been proposed, generalizing concepts and techniques of formal string language theory. In [6], a simple yet expressive two-dimensional grammar model, called pure 2D context-free grammar ($P2DCFG$), was introduced and several theoretical properties were established. In [7], an extension of $P2DCFG$, called extended 2D context-free Picture Grammar ($E2DCFPG$), was considered by allowing variable symbols in addition to terminal symbols and rewriting a column or a row of variables using tables of context-free rules. $E2DCFPG$ was shown [7] to have more generative power than the $P2DCFG$. On the other hand, a variant of pure 2D context-free grammars, called $(l/u)P2DCFG$, was introduced in [3]. Unlike the $P2DCFG$, in rewriting a picture array using tables of context-free rules, the leftmost column or uppermost row of symbols in the array are rewritten in $(l/u)P2DCFG$. Here we introduce a new picture array generating 2D grammar model, which we call as $(l/u)E2DCFPG$, by considering the (l/u) mode of rewriting in the extended pure 2D context free grammar. In other words we allow rewriting all the variables (also called non terminals) in a leftmost column or an uppermost row of variables by making use of tables of context-free rules. The picture array language generated by $(l/u)E2DCFPG$ consists of arrays over a set of terminal symbols derived from axiom arrays by the application of the column / row tables of rules with an (l/u) kind of rewriting. We derive several theoretical properties by comparing the proposed model with others and also discuss the applicability of this new 2D grammar model in generating digitized picture arrays.

2 Preliminaries

We recall needed notions relating to formal string languages and two-dimensional picture languages [4, 2]. A finite alphabet Σ is a finite set of symbols. A word or a string w over Σ is a finite sequence of symbols taken from Σ. The set of all words over Σ, including the empty word λ with no symbols, is denoted by Σ^*. The length of a word w is denoted by $|w|$. Given any word $w = a_1 a_2 \dots a_n$, we denote by ${}^t w$, the word $w = a_1 a_2 \dots a_n \quad (n \geq 1)$ written vertically. A picture array (or simply, an array) p over an alphabet Σ is a rectangular $m \times n$ array of elements of Σ of the form

$$p = \begin{array}{ccc} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{array}$$

The set of all picture arrays over Σ is denoted by Σ^{**}. A picture language or a two dimensional language over Σ is a subset of Σ^{**}. The $E2DCFPG$ [7] is an extension of the $P2DCFG$ introduced in [6]. The family of picture array languages generated by extended 2D context-free picture grammars is denoted by $E2DCFPL$. When all the symbols are terminal symbols in a $E2DCFPG$ the resulting grammar is simply a $P2DCFG$ [6]. In fact the rules in a column or row table have their left and right sides consisting of only terminal symbols. The family of picture array languages generated by $P2DCFGs$ is denoted by $P2DCFL$. An $(l/u)P2DCFG$ which was introduced in [3], has all its components as in a $P2DCFG$ but involves an (l/u) mode of derivation rewriting all the symbols only in the leftmost column or the uppermost row of a picture array. The family of picture array languages generated by $(l/u)P2DCFGs$ is denoted by $(l/u)P2DCFL$.

3 E2DCFPG with (l/u) mode of derivation

We now introduce the (l/u) mode of derivation in a $E2DCFPG$.

Definition 1. Let $G = (V, \Sigma, P_1, P_2, \mathcal{M})$ be an extended 2D context-free picture grammar ($E2DCFPG$) [7]. An (l/u) mode of derivation of a picture array M_2 from M_1 in G, is a derivation in G such that only the leftmost column or the uppermost row of M_1 entirely made of variables, is rewritten using respectively, the column rule tables or the row rule tables, to yield M_2. The generated picture language consists of picture arrays over terminal symbols, each of which is derived from an axiom array of G using (l/u) mode of derivation. The family of picture languages generated by $E2DCFPGs$ under (l/u) mode of derivations is denoted by $(l/u)E2DCFPL$. We write $(l/u)E2DCFPG$ to denote $E2DCFPG$ under (l/u) mode of derivation.

We give an example of $(l/u)E2DCFPG$.

Example 2. We consider an $(l/u)E2DCFPG$ G_{tl}. We mention only the axiom array and tables of rules of G_{tl}. The axiom array is $M_0 = \begin{matrix} A & C \\ B & D \end{matrix}$. The column tables of context-free rules are $c_1 = \{A \to XAX, B \to bBb\}, c_2 = \{A \to Y, B \to a\}$, $c_3 = \{C \to ZE, E \to UE, D \to aF, F \to aF\}$, $c_4 = \{E \to U, F \to a\}$ $r_1 = \{X \to \begin{matrix} X \\ b \end{matrix}, Y \to \ Y\ , Z \to \begin{matrix} Z \\ a \end{matrix}, U \to \begin{matrix} U \\ b \end{matrix}\}, r_2 = \{X \to a, Y \to a, Z \to a, U \to b\}$. Starting from the axiom array $\begin{matrix} A & C \\ B & D \end{matrix}$, we can apply, for example, the rules of the column table c_1 two times, followed by c_2 to yield the array $\begin{matrix} X & X & Y & X & X & C \\ b & b & a & b & b & D \end{matrix}$. If now we apply c_3 two times followed by c_4 and then apply, r_1 two times followed by r_2, then we obtain the picture array M as shown in Fig. 1.

$$\begin{matrix} a & a & a & a & a & a & b & b \\ b & b & a & b & b & a & b & b \\ b & b & a & b & b & a & b & b \\ b & b & a & b & b & a & a & a \end{matrix}$$

Figure 1: Picture Array M

4 Comparison Results

We now derive certain theoretical properties that compare the picture generative power of the $(l/u)E2DCFPG$ with other picture generative models. We first show that the new picture array generating model of $(l/u)E2DCFPL$ considered here properly includes the 2D model $(l/u)P2DCFL$ [3].

Theorem 3. $(l/u)P2DCFL \subset (l/u)E2DCFPL$.

Proof. Let L be a picture language generated by a $(l/u)P2DCFG$ G which will have only one kind of symbol, namely, terminal symbol. We now describe how to construct a $(l/u)E2DCFPG$ from the given G to generate L. For every symbol $a \in G$, we create a nonterminal symbol A_a. For every rule of the form $a \to b\alpha$, a, b are terminal symbols and α is a string of terminal symbols in a column table c of G, we delete the rule $a \to b\alpha$ and add two rules to c given by $A_a \to A_b\alpha$ and $A_a \to A_b\beta$ where β is obtained from α by replacing every terminal symbol in α by its associated nonterminal symbol. The idea is that if a is the corner symbol in the leftmost column and the uppermost row of a picture array p which is rewritten by the rules of the column table c, then the rule $A_a \to A_b\beta$ (mentioned above) should be applied while the rule $A_a \to A_b\alpha$ should be applied, if a is a symbol in the leftmost column but not the corner symbol. Since the rewriting is in (l/u) mode, if the mentioned way of applying the rules of the new column table is not adopted, then nonterminals will get stuck in the array and cannot be changed into terminals. Likewise similar modifications can be done to the rules of a row table. We also add a new column table as well as a new row table each consisting of all rules of the forms $A_a \to A_a$, $A_a \to a$ for every new symbol A_a. It can be seen that the $E2DCFPG$ thus formed generates the picture language L.
To prove the proper inclusion, we consider the picture language $L_1 \cup L_2$ where L_1 and L_2 are defined [3] as follows: The picture language $L_1 \subseteq \{a, b, d\}^{**}$ is such that each $p \in L_1$ of size $(m, n), m \geq 2, n \geq 2$ has the following properties: $p(1,1) = b$; $p(1, j) = a$, for $2 \leq j \leq n$; $p(i, 1) = a$, for $2 \leq i \leq m$; $p(i, j) = d$, otherwise; The picture language $L_2 \subseteq \{a, b, e\}^{**}$ is such that each $p \in L_2$ of size $(r, s), r \geq 2, s \geq 2$ has the following properties: $p(1,1) = b$; $p(1, j) = a$, for $2 \leq j \leq s$; $p(i, 1) = a$, for $2 \leq i \leq r$; $p(i, j) = e$, otherwise. It has been proved in [3] that $L_1 \cup L_2$ cannot be generated by any $(l/u)P2DCFG$. But $L_1 \cup L_2$ is generated by the following $(l/u)E2DCFPG$: We mention the axiom arrays and the column and row tables of rules. The axiom arrays are $\begin{matrix} B & A \\ A & d \end{matrix}$ and $\begin{matrix} C & D \\ D & e \end{matrix}$. The column tables c_1, c_2, c_3, c_4 and the row tables r_1, r_2, r_3, r_4 are as given below: $c_1 = \{B \to BA, A \to Ad\}, c_2 = \{C \to CD, D \to De\}$,$c_3 = \{B \to B, A \to a\}, c_4 = \{C \to C, D \to$

$a\}$ $r_1 = \{B \to \begin{matrix} B \\ A \end{matrix}, A \to \begin{matrix} A \\ d \end{matrix}\}, r_2 = \{C \to \begin{matrix} C \\ D \end{matrix}, D \to \begin{matrix} D \\ e \end{matrix}\}$, $r_3 = \{B \to b, A \to a\}, r_4 = \{C \to b, D \to a\}$. It can be seen that the tables c_1, c_3, r_1, r_3 can be applied suitable number of times to generate the picture arrays of L_1 while c_2, c_4, r_2, r_4 can be applied suitable number of times to generate the picture arrays of L_2, thus generating $L_1 \cup L_2$. □

Theorem 4. $(l/u)E2DCFPL \setminus P2DCFL \neq \emptyset$.

Proof. We consider the picture language $L_1 \cup L_2$ where L_1 and L_2 are defined [6] as follows:$L_1 = \{X_1 \circ {}^t(c^n) \circ Y_1 \mid n \geq 1, X_1 \in \{a\}^{++}, Y_1 \in \{b\}^{++}, |X_1|_c = |Y_1|_c\}$ where $|X|_c$ stands for the number of columns of X. $L_2 = \{X_2 \circ {}^t(c^n) \circ Y_2 \mid n \geq 1, X_2 \in \{x\}^{++}, Y_2 \in \{y\}^{++}, |X_2|_c = |Y_2|_c\}$. It is known [6] that $L_1 \cup L_2$ cannot be generated by any $P2DCFG$. But $L_1 \cup L_2$ can be generated by the following $E2DCFPG$ G'. We mention the axiom arrays and tables of rules only. The Axiom arrays are ACB and XDY. The column tables are $c_1 = \{C \to ACB\}, c_2 = \{D \to XDY\}$,$c_3 = \{C \to C, E \to c\}, c_4 = \{D \to D, F \to c\}$. The row tables are $r_1 = A \to \begin{matrix} A \\ a \end{matrix}, C \to \begin{matrix} C \\ E \end{matrix}, B \to \begin{matrix} B \\ a \end{matrix}\}, r_2 = \{X \to \begin{matrix} X \\ a \end{matrix}, D \to \begin{matrix} D \\ F \end{matrix}, Y \to \begin{matrix} Y \\ a \end{matrix}$ $r_3 = \{A \to a, C \to c, B \to b\}, r_3 = \{X \to a, D \to c, Y \to b\}$ It can be seen that G' generates $L_1 \cup L_2$. □

A 2D grammar model [5], which we call here as two-dimensional tabled context-free grammar $2TCFG$ (originally called tabled context-free matrix grammar in [5]) has been extensively investigated. There are two phases of derivation in a $2TCFG$. In the first phase horizontal strings of symbols are generated using context-free string grammar rules starting from a start symbol. In the second phase, the symbols in these words are in parallel rewritten in the vertical direction using right-linear rules which are either of the form $A \to aB$ or $A \to B$ or $A \to a$ or $A \to \lambda$. But at a step all the rules applied are one of these forms. The family of picture languages generated by $2TCFGs$ is denoted by $2TCFL$. We now show that the family $(l/u)E2DCFPL$ contains picture languages that cannot be generated by any $2TCFG$.

Theorem 5. $(l/u)E2DCFPL \setminus 2TCFL \neq \emptyset$.

Proof. We consider the picture language L' consisting of picture arrays of sizes $(2m+1) \times (2n+2)$, $m, n \geq 0$, over an alphaber $\{a, b, c, d, e\}$ with a middle row of $c's$ and an equal number of rows above and below the middle row. The row above is of the form $a^{(n+1)}d^{(n+1)}$ while the row below is of the form $b^{(n+1)}e^{(n+1)}$. This picture language L' cannot be generated by any $2TCFG$ since the right-linear rules in the second phase of $2TRLG$ cannot maintain the feature of equal number of rows above and below the middle row of $c's$, although the CF rules of the first phase can handle the property present in the columns. But we construct a $E2DCFPG$ generating L'. Again we mention only the axiom array and the tables of rules. The axiom array is ${}^t(AXB)$ The row table is $r = \{X \to {}^t(AXB)\}$ while the column tables are $c_1 = \{A \to aAd, X \to cXc, B \to bBe\}$ and $c_2 = \{A \to ad, X \to cc, B \to be\}$. It can be seen that this $E2DCFPG$ generates the picture language L'. $\square$

5 Application

In the problem of character recognition, array grammars have been effectively used [1]. The $E2DCFPG$ and $(l/u)E2DCFPG$ provide grammar models for describing characters. For instance the $(l/u)E2DCFPG$ G_{tl} in the Example given in Section 3 generates picture arrays as in Fig. 1. If we interpret the symbol b as representing the blank symbol, then the array in Fig. 1 represents two alphabetic letters T and L juxtaposed, as shown in Fig. 2. This feature could prove useful in character recognition problems as the $(l/u)E2DCFPG$ provides a simple context-free kind of grammar model for description of such characters.

$$
\begin{array}{cccccccc}
a & a & a & a & a & a & & \\
 & & a & & & a & & \\
 & & a & & & a & & \\
 & & a & & & a & a & a
\end{array}
$$

Figure 2: Characters T and L juxtaposed

6 Conclusion

Here we have considered the $E2DCFPG$ in the (l/u) mode of derivation. It remains to explore whether the two grammar models $E2DCFPG$ and $(l/u)E2DCFPG$ are equivalent or not in generative power. It will be of interest to compare the generative power of $(l/u)E2DCFPG$ with other picture array generative models.

References

[1] H. Fernau, R Freund, *Bounded Parallelism in Array Grammars Used for Character Recognition.* In: Advances in Structural and Syntactical Pattern Recognition, Lecture Notes in Computer Science, Vol. **1121** (2005) 40–49.

[2] D. Giammarresi, A. Restivo, *Two-dimensional languages.* In: [7], Vol. 3, pp. 215-267 (1997).

[3] Z. Krivka, C. Martn-Vide, A. Meduna, K.G. Subramanian, *A variant of pure two-dimensional context-free grammars generating picture languages.* Lecture Notes in Computer Science, Springer, Heidelberg, vol. **8466** (2014), 123-133.

[4] G. Rozenberg, A. Salomaa (Eds.), *Handbook of Formal Languages.* Vol. **1–3**, Springer, Berlin, 1997.

[5] R. Siromoney, K.G. Subramanian, K. Rangarajan, *Parallel/Sequential rectangular arrays with tables.* Int. J. Computer Mathematics, **6** (1977), 143–158.

[6] K.G. Subramanian, R.M. Ali, M. Geethalakshmi, A.K. Nagar, *Pure 2D Picture Grammars and languages.* Discrete Applied Mathematics, **157** (2009), 3401–3411.

[7] K.G. Subramanian, M. Geethalakshmi, A.K. Nagar, S.K. Lee, *Two-dimensional picture grammar models.* Proceedings of the 2nd European Modelling Symposium (EMS2008), (2008), 263–267.