RANDOM CONTEXT STRUCTURE GRAMMARS

E.M. Ehlers
S.H. von Solms
Deparment of Computer Science
Rand Afrikaans University
P.O. BOX 524 Johannesburg 2000

ABSTRACT

Random Context Structure Grammars are defined. The language generated by a Random Context Structure Grammar is a set of three-dimensional digital structures. A hierarchy of Random Context Grammars is discussed briefly. Random Context Structure Grammars are defined as an extension of the one-dimensional Random Context Grammars and two-dimensional Random Context Array Grammars. Finally an example is given of a Random Context Structure Grammar that generates three-dimensional digital structures in the form of square-based pyramids.

1. INTRODUCTION

In the last few years much work has been done on grammars generating two-dimensional figures [5], [9]. An example of two-dimensional grammars is Random Context Array Grammars (RCAG) defined by Von Solms [7], [8]. Arrays are generated by using sequential rewriting. The rewriting of each symbol is determined by a horizontal, vertical and global permitting and forbidding context.

Many disciplines use three-dimensional models, for instance chemical structures in chemistry, and the DNA-molecule in biochemistry. A formal method to generate three-dimensional structures could have interesting applications in these fields. In this paper RCAG's have been extended to Random Context Structure Grammars (RCSG) that generate three-dimensional structures. RCSG's generate three-dimensional structures using a number of permitting and forbidding context conditions in every production. The power of RCSG's is inherent in the fact that the rewriting of a single symbol in the structure may be influenced by the presence and/or absence of certain specified symbols in the structure.

This is an attempt at defining a grammar that generates three-dimensional structures and it is interesting both in a theoretical and practical sense, as mentioned above.

The discussion given here is informal and intuitive. For a more complete definition the reader is referred to [1].

2. A HIERARCHY OF RANDOM CONTEXT GRAMMARS

To enable us to understand Random Context Structure Grammars a short discussion of one-dimensional and two-dimensional Random Context Grammars will be given.

2.1 The One-Dimensional Case : Random Context Grammars

Random Context Grammars [6] generate one-dimensional strings. The class of languages generated by Random Context Grammars contains the class of context free languages and is contained in the class of context sensitive languages. Random Context languages are important in the syntactical definition of natural and programming languages.

Definition

A Random Context Grammar is an ordered four tuple $G = (V_N, V_T, P, S)$ where

a) V_N and V_T are disjunct finite non-empty sets of nonterminal and terminal symbols respectively. $S \in V_N$.

quæstiones informaticæ, 5, 2, pp 23-28, issn 0254-2757

b) P is a finite set of productions of the form

$$A \rightarrow \alpha (U; T), A \in V_N, \alpha \in (V_N \cup V_T)^+, \text{U and T are subsets of } V_N$$

If $A \rightarrow \alpha (U; T)$ is a production in P and β and γ in $(V_N \cup V_T)^*$ then we may write

$$\beta A \gamma \Rightarrow \beta \alpha \gamma$$

if every symbol in U appears in the string $\beta\gamma$ and no symbol in T appears in $\beta\gamma$. U is called the permitting context and T the forbidding context.

Example

The following example shows a Random Context Grammar that generates the language $L(G) = \{a^{n(n+1)/2} / n \geq 1\}$.

$G = (\{S, A, X, Y\}, \{a\}, P, S)$

P consists of the following productions:

$S \rightarrow A$	$A \rightarrow aY (X; \emptyset)$
$S \rightarrow A S$	$X \rightarrow a (\emptyset; A)$
$A \rightarrow X (\emptyset; X, Y, S)$	$Y \rightarrow A (\emptyset; X)$

Lets look at the derivation of aaa = $a^{2(2+1)/2} = a^3$

$A \Rightarrow AS \Rightarrow AA \Rightarrow XA \Rightarrow XaY \Rightarrow aaY \Rightarrow aaA \Rightarrow aaX \Rightarrow aaa.$

2.2 The Two-Dimensional Case : Random Context Array Grammars

In this section we briefly define the two-dimensional extention of Random Context Grammars. This extension is called Random Context Array Grammars (Von Solms [7,8]).

An RCAG is a 4-tuple $G = (V_N, V_T, P, S)$ where V_N, V_T and S have the usual Formal Language meaning. P is a finite set of productions consisting of 4 different types:

1 $A \rightarrow \alpha (U_1; T_1/U_2; T_2/U_3; T_3)$
2 $A \leftarrow \alpha (U_1; T_1/U_2; T_2/U_3; T_3)$
3 $A \uparrow \alpha (U_1; T_1/U_2; T_2/U_3; T_3)$
4 $A \downarrow \alpha (U_1; T_1/U_2; T_2/U_3; T_3)$

$A \in V_N, \alpha \in (V_N \cup V_T)^+, U_i, T_i \subseteq V_N, 1 \leq i \leq 3, U_i \cap T_i = \emptyset, 1 \leq i \leq 3, 1 \leq |\alpha| \leq 2$

A production of type 1 is interpreted as follows:

If A appears in a two-dimensional figure of symbols from $(V_N \cup V_T)$. A may be replaced horizontally to the right by α if —

(i) all symbols of U_1 and no symbols of T_1 appears in the same horizontal row as A
(ii) all symbols of U_2 and no symbols of T_2 appear in the same vertical row as A
(iii) all symbols of U_3 and no symbols of T_3 appear somewhere in the figure
(iv) $|\alpha| = 2$, then only background symbols must appear immediately to the right of A
(v) $|\alpha| = 1$, then A is simply replaced by α if (i), (ii) and (iii) hold.

U_1 and T_1 are known as the horizontal permitting and forbidding context.
U_2 and T_2 are known as the vertical permitting and forbidding context.
U_3 and T_3 are known as the global permitting and forbidding context.

For the complete definition and a more detailed explanation see Von Solms [7, 8].

2.3 The Three-Dimensional Case: Random Context Structure Grammars

Three-dimensional models are used in several disciplines, for example structures in Chemistry and the DNA-molecule in Biochemistry. It seems as if a fomal method to generate three- dimensional structures can have very interesting applications [2, 3, 4] . In this section the two-dimensional Random Context Array Grammars are extended to three dimensions and called Random Context Structure Grammars (RCSG's). RCSG's are able to generate three-dimensional structures. Once again we stress the fact that because of the volume and complexity of the productions, RCSG's will only be discussed informally. A formal and complete introduction appears in[1].

The language generated by an RCSG is a set of three-dimensional digital structures. (For a discussion of three-dimensional digital structures the reader is referred to Rosenfeld [4].

An RCSG is a 4-tuple $G = (V_N, V_T, P, S)$ where V_N, V_T, P and S have the usual Formal Language meanings. The productions of P are of the form:

$$A \rightarrow \alpha\,(U_1; T_1 /U_2; T_2 /U_3; T_3 /U_4; T_4 /U_5; T_5 /U_6; T_6 /U_7; T_7)$$

where $A \in V_N, \alpha \in (V_N \cup V_T)^+, U_i, T_i \subseteq V_N, 1 \leq i \leq 7, 1 \leq |\alpha| \leq 2, U_i \cap T_i = \emptyset, 1 \leq i \leq 7.$

The productions will now be discussed by considering figure 1 which shows a single symbol in a three-dimensional structure. The lines and planes through the point where the symbol is located, determine the different contexts which must be considered before a production is applied.

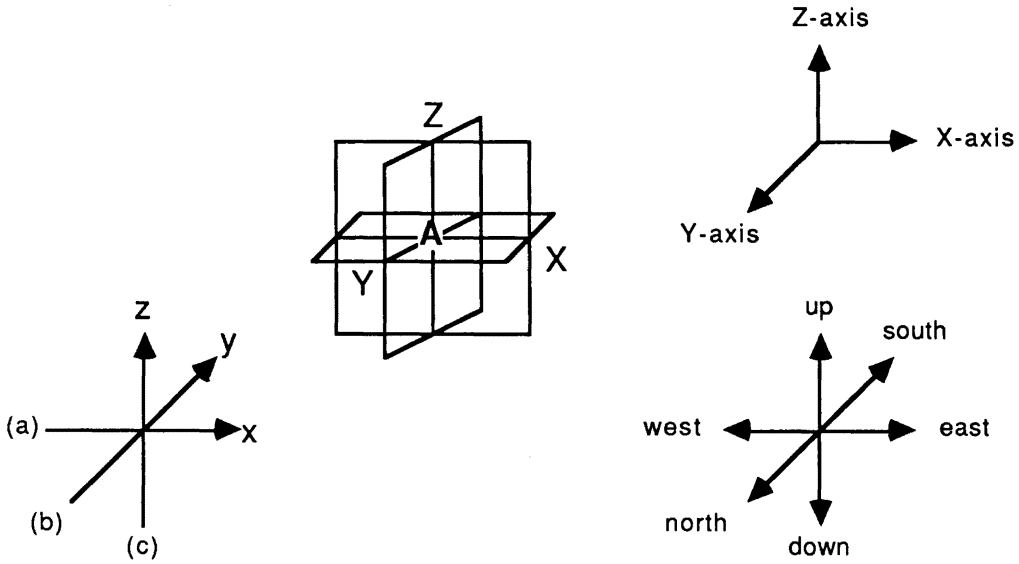

figure 1
The Context Associated with a Symbol A in a 3-Dimensional Structure

There are seven different contexts associated with the symbol A:

a)	The horizontal line context	All the symbols appearing on the line through A parallel to the X-axis.
b)	The vertical line context	All the symbols appearing on the line through A parallel to the Y-axis.
c)	The depth line context	All the symbols appearing on the line through A parallel to the Z-axis.
d)	The horizontal plane context	All the symbols appearing on the plane through A parallel to the XY-plane.
e)	The vertical plane context	All the symbols appearing on the plane through A parallel to the XZ plane.
f)	The depth plane context	All the symbols appearing on the plane through A parallel to the YZ-plane.
g)	The global context	All the symbols appearing in the structure.

Productions can cause substitution in six different directions:

$A \rightarrow \alpha$ to the east
$A \leftarrow \alpha$ to the west
$A \nearrow \alpha$ to the south
$A \swarrow \alpha$ to the north
$A \uparrow \alpha$ upwards
$A \downarrow \alpha$ downwards

A production

$$A \rightarrow \alpha\,(U_1;\, T_1\,/U_2;\, T_2\,/U_3;\, T_3\,/U_4;\, T_4/U_5;\, T_5/U_6;\, T_6/U_7;\, T_7)$$

of G can be interpreted as follows:

If A appears in a three-dimensional structure of symbols of $V_N \cup V_T$, A may be substituted by α eastwards if —

(a) all the symbols of U_1 and none of T_1 appear in the horizontal line context
(b) all the symbols of U_2 and none of T_2 appear in the vertical line context
(c) all the symbols of U_3 and none of T_3 appear in the depth line context
(d) all the symbols of U_4 and none of T_4 appear in the horizontal plane context
(e) all the symbols of U_5 and none of T_5 appear in the vertical plane context
(f) all the symbols of U_6 and none of T_6 appear in the depth plane context
(g) all the symbols of U_7 and none of T_7 appear in the global context
(h) If $|\alpha| = 2$ then only background symbols must appear directly east of A
(i) If $|\alpha| = 1$, then A is simply substituted by α if (a) — (g) holds.

The other types of productions are interpreted in a similar way.

It is easy to see that RCAG's are a special case of RCSG's. The structures generated by RCAG's can be seen as structures lying in a single plane. The power of RCSG's is inherent in the fact that the substitution of a symbol in a structure can be influenced by the presence or absence of other symbols in seven different contexts in the structure. This should have very interesting practical applications.

Example of an RCSG

The language generated by the following RCSG is the set of square-based pyramids.

$$G = (\{S, Y, Z, X_1, X_2, X_3, X_4\}, \{*\}, P, S)$$

P:

1. S → * (//////)
2. S ↓ $\overset{*}{Y}$ (//////∅; {X_1, X_2, X_3, X_4, Y, Z})
3. Y → YX_1 (//////∅; {$X_2; X_3; X_4$})
4. Y ← X_2Y (X_1; ∅//////∅; {X_3, X_4})
5. Y ↑ Y^{X_3} ({X_1, X_2}; ∅//////∅; { X_4})
6. Y ↓ X_4Y ({X_1, X_2}; ∅/{X_3}; ∅/////)
7. Y ↓ $\overset{*}{Z}$ (///{X_1, X_2, X_3, X_4}; ∅///)
8. X_1 ↓ $\overset{*}{Y}$ (///{X_2, X_3, X_4}; ∅///)
9. X_2 ↓ $\overset{*}{Y}$ (///{X_3, X_4}; {X_1} ///)
10. X_3 ↓ $\overset{*}{Y}$ (///{ X_4}; {X_1, X_2} ///)
11. X_4 ↓ $\overset{*}{Y}$ (///∅; {X_1, X_2, X_3} ///)
12. Z ↓ $\overset{*}{Z}$ (///{X_1, X_2, X_3, X_4}; ∅///)
13. Y → Z (///{X_1, X_2, X_3, X_4}; ∅///)
14. Z → * (///{X_1, X_2, X_3, X_4}; ∅///)
15. Y → * (//////{X_1, X_2, X_3, X_4}; {Z))
16. X_1 → * (///{ X_2, X_3, X_4}; ∅///∅; {Z, Y})
17. X_2 → * (///{X_3, X_4}; {X_1} ///∅; {Z, Y})
18. X_3 → * (///{ X_4}; { X_1, X_2} ///∅; {Z, Y})
19. X_4 → * (///{X_1, X_2, X_3}; ∅///∅; {Z, Y})

The elements of the language is three-dimensional digital structures in the form of square-based pyramids: (Seen from the top)

```
Level 1        Level 2        Level 3

                                  *
                  *             * * *
   *            * * *         * * * * *      etc.
                  *             * * *
                                  *
```

figure 2

A Digitized Square Based Pyramid

To illustrate the use of the productions in the generation of a structure let us consider the first few steps in the generation of a square-based pyramid.

Notation: Each of the levels of the three-dimensional structure is given seperately.
⊗ → denotes the symbols of every level that lie on a straight depth line.

Generation Production used		Resulting structure
2. $S \downarrow \overset{*}{Y}$	This production is used as none of the symbols X_1, X_2, X_3, X_4, Y or Z appear in the global context.	Level 1 : (*) Level 2 : Y

(The following productions will generate a square at level 2).

3. $Y \rightarrow YX_1$	This production is used as no X_2, X_3 or X_4 appears in the global context.	Level 1: (*) Level 2: (Y)X_1
4. $Y \leftarrow X_2Y$	This production is used as no X_3 or X_4 appears in the global context in the structure and a X_1 appears on the same horizontal line as the Y.	Level 1: (*) Level 2: X_2(Y)X_1
5. $Y \uparrow Y^{X_3}$	This production may be used as X_1 and X_2, appear on the same horizontal line as Y and there is no X_4 in the global context.	Level 1: (*) Level 2: X_3 X_2(Y)X_1
6. $Y \downarrow X_4Y$	This production is used as: X_1 and X_2 appear on the same horizontal line as Y and X_3 appears on the same vertical line as Y.	Level 1: (*) Level 2: X_3 X_2(Y)X_1 X_4

The following level of the pyramid can be generated or the structure completed by using productions 15-19.

REFERENCES

1. Ehlers and von Solms., A Hierarchy of Random Context Grammars and Automata., Submitted for publication.
2. Lest and Teaneck., [1979], Detection of Three-dimensional Patterns of Atoms in Chemical Structures., *CACM*, **22,** 4.
3. Requicha., [1980], Representations for Rigid Solids : Theory Methods and Systems., *Computing Surveys*, **12**, 4.
4. Rosenfeld A., Three-dimensional Digital Topology., *Information and Control*, **50,** pp. 119-127.
5. Rosenfeld A., [1979], *Picture Languages*, Academic Press.
6. Van der Walt APJ., Random Context Languages., *Information Processing - 71.*
7. Von Solms SH., [1979], Random Context Array Grammars., *ACM Computer Science Conference 1979.*
8. Von Solms SH., [1980], Random Context Array Grammars., *Information Processing 80.*, Proceedings of IFIP Congress 80 North-Holland Publishing Company 1980.
9. Wang., [1982], A New Hierarchy of Two-dimensional languages, *Information Processing Letters*, **15**, 5.