The Vector String Descriptor as a Tool
In the Analysis of Cellular Automata Systems

JON T. BUTLER

AND

SIMEON C. NTAFOS
Department of Electrical Engineering, The Technological Institute,
Northwestern University, Evanston, Illinois 60201

Received October 1976; revised 6 March 1977

ABSTRACT

A major impediment in the application of cellular spaces to the modeling of biological phenomena and other physical processes has been the lack of tools which relate global to local behavior. In this paper, the vector string descriptor is introduced as a tool for the analysis of cellular spaces similar to those studied by S. Ulam. With such a tool, it is quite straightforward to show the existence of global properties which would otherwise be very difficult to prove. For example, a conjecture by Ulam concerning the density of cells in a "growth" pattern is shown to be true. We also show that a self-reproduction process occurs which allows regeneration of specific patterns from "cuttings" of those patterns.

I. INTRODUCTION

The cellular space or homogeneous structure has been proposed as a model of specific biological phenomena. The origin of most of the present work in this area is the self-reproduction system of von Neumann [13]. Subsequent self-reproduction cellular spaces have been proposed by Codd [5], Banks [3], and Nourai and Kashef [14]. To Moore [12] goes the credit for the surprising discovery that certain patterns cannot be reproduced. A survey of this topic and computational aspects of cellular spaces appears in Aladyev [1].

An interesting application is the heart tissue model of Moe, Rheinbolt, and Abildskov [11]. They showed that the self-sustaining activity which occurs during atrial fibrillation can be modeled in a hexagonal array of 992 cells. Many biologically motivated numerical analyses are, in fact cellular spaces where the local transition function represents a space-time differential equation. For example, Winfree [18] has shown a cellular system which

MATHEMATICAL BIOSCIENCES 35, 55-84 (1977)
© Elsevier North-Holland, Inc., 1977
can model the rotating chemical reaction, an inorganic solution that displays a biorhythm-like periodicity. The local transition function specifies changes in local substance concentration with time as a function of reaction between local substances and of diffusion between nearby locations. Implementation of differential equations for nerve tissue has been studied by Zeeman [21].

The modeling of growth in cellular systems has received considerable attention. Ulam [17], Schrandt and Ulam [16], and Holladay and Ulam [7] have described growth patterns "intermediate in complexity between inorganic patterns like those of snow crystals and the more varied intricacies of molecules and structures". Lieblein [9] and Gajski and Yamada [6] have investigated cellular spaces in which the patterns grow to a specific size and remain forever unchanged. Growth from groupoid operations in cellular spaces has been shown by Rothstein [15] to produce rich and varied patterns. A survey emphasizing growth has been presented by Kitagawa [8].

A problem of central importance has been the characterization of global behavior in cellular spaces as a function of the local transition function. With respect to Moore's [12] result, for example, a number of researchers (Brownlee [4], Amoroso and Patt [2], and Maruoka and Kimura [10]) have shown that Garden of Eden configurations do not exist in spaces where the local transition function is "balanced". However, beyond this, few results exist which relate global parameters to local transition functions. Thus, a need exists for more analytic tools. From a mathematical point of view the cellular space is very much like a recurrence relation in which the \( n \)th number can be gotten from the \( n-1 \)st number. Some recurrences, such as the Fibonacci sequence, have the nice property that equivalent closed form solutions exist. For such cases, it is easier to obtain the \( n \)th number from an algebraic expression (i.e., a closed form solution) rather than by successive applications of the recurrence relation. In this paper, we show that an analogous situation exists in cellular spaces.

The spaces which lend themselves well to a closed form solution are growth functions similar to those studied by Ulam. In particular, it is shown that the location of cells in the pattern can be uniquely described by a "vector string descriptor". The significance of the descriptor is that it allows properties of the patterns to be established in a tractable manner. For example, we prove a conjecture by Ulam [17] concerning the density of growth pattern cells. We show also how the properties of growth patterns are related to the initial configuration. In particular, a form of self-reproduction is shown to exist, and this property allows the pattern to be regenerated from segments of the parent growth.

As with many analytical tools, the range of applicability of the vector string descriptor is limited. At one end are cellular spaces whose patterns are too complex to be described by a vector string. It is likely that
VECTOR STRING DESCRIPTOR

differential equation implementations fall into this category. At the other end are spaces whose patterns are too simple to warrant its use, such as the linear tally counters of Lieblein [9].

While the growth patterns considered do not model specific biological phenomena, they do display lifelike characteristics. For example, we see a surprising degree of complexity evolve from a relatively straightforward process. The final pattern is the result of its environment (growth function) and its origin (starting configuration). As time progresses the form of the final "organism" depends more and more on its environment and less and less on its origin.

The format of this paper is as follows. Section II is a description of the notation. Section III is a discussion of the vector string descriptor for the most basic growth rule. We then use the descriptor to determine characteristics of the pattern. In Sec. IV, we investigate patterns which result from arbitrary initial configurations and show that a self-reproduction process occurs. The vector string descriptor is shown in Sec. V to apply to whole classes of growth rules. In Sec. VI, we discuss limitations and implications of the vector string descriptor.

II. NOTATION

Let \( Z \) be the set of integers and \( Z^2 \) the set of ordered pairs \( a = (a_0, a_1) \), where \( a_i \in Z \). \( Z^2 \) can be considered the set of points with integer coordinates in the Euclidean plane. Let \( 0 = (0,0) \) denote the origin, and for \( a, b \in Z^2 \), let \( a + b = (a_0 + b_0, a_1 + b_1) \) denote the vector sum of \( a \) and \( b \).

Associated with each point \( a \in Z^2 \) is the state of \( a \), \( q_a \in \{0, 1\} \). Let a configuration \( C_t \) at time \( t \) denote an assignment of states to all points in \( Z^2 \). Thus, \( C_t : Z^2 \rightarrow \{0, 1\} \). The initial configuration \( C_0 \) is assumed to exist prior to the application of any growth function. In the primitive initial configuration the states of all points are 0, except the origin, in which case, \( q_0 = 1 \). The term cell will be used to denote a point and its state.

THE LOCAL TRANSITION FUNCTION

Configuration \( C_t \) is derived from configuration \( C_{t-1} \) by a local transition function, which specifies the state of each cell at \( t \) as a function of its state at \( t-1 \) and the states of neighboring cells at \( t-1 \). The notation used here is identical to that of Banks [3]. Figure 1 shows, for example, the local transition function described in the next section. It specifies that a cell in configuration \( C_t \) will map to the 1 state if it is in the 0 state at \( t-1 \) (as represented by the 0 surrounded by a square), and the neighboring points have the states shown. The local transition function is symmetric, in that a 90°, 180°, or 270° rotation of the pattern shown in Fig. 1 specifies a 0 to 1 state transition also. For unspecified neighborhoods, the state of a point in
$C_i$ is identical to that of $C_{i-1}$. The function in Fig. 1 is identical to that examined by Ulam [17] and will be called the \textit{u rule}.

\[
\begin{array}{c}
0 \\
0 \begin{array}{c}
0 \end{array} \begin{array}{c}
1 \\
1
\end{array}
\end{array}
\]

\textbf{FIG. 1.} Local transition function $u$.

Consider now patterns which evolve from the primitive initial configuration. Since there are no 1 to 0 transitions, a cell will remain in the 1 state once it has attained that state. Thus, under the $u$ rule, points in $\mathbb{Z}^2$ can be classified in one of two ways: those destined to be in the 1 state and those which are not. Denote the former as \textit{Ulam's pattern $U_i$}, and let $U_i$ be the set of points which undergo the 0 to 1 transition upon the $i$th application of the rule. It is appropriate to let $U_0 = \{0\}$. Thus,

$$U = \bigcup_{i=0}^{\infty} U_i$$

where $\cup$ means the union of all sets $U_i$. Figure 2 shows that part of $U$ up to the 9th application of the $u$ rule. For each $a \in U_i$ shown, the numbers in the square represent $i$.

\textbf{FIG. 2.} $U$ After 9 applications of the $u$ rule.
VECTOR STRING DESCRIPTOR

THE VECTOR STRING DESCRIPTOR

An important tool in the analysis of the pattern $U$ is the vector string descriptor. Let $I$, the fundamental vector set, be defined as follows:

$$I = \{(0,0),(1,0),(-1,0),(0,1),(0,-1)\}.$$ 

A vector string $\Gamma$ is a sequence of vectors

$$\Gamma = \gamma_m\gamma_{m-1}\cdots\gamma_0,$$

where $\gamma_i \in I$, such that for all substrings

$$\gamma_i00\cdots0\gamma_j \quad (= \gamma_i0^{j-i-1}\gamma_j)$$

of $\Gamma$, $\gamma_i \neq -\gamma_j$.\footnote{0^m = 00\cdots0 (m times).} This is called the noncancellation property. For example, $(0,1) (0,0) (-1,0) (0,-1)$ is a vector string, while $(0,1) (0,1) (0,0) (0,-1)$ is not, since the latter contains the substring $(0,1) (0,0) (0,-1)$. Note that $(0,1) (0,-1) (0,0) (0,0)$ is not a vector string either, since an empty string of 0's separates $(0,1)$ and $(0,-1)$. Let $V$ be the set of all vector strings.

The vector string $\Gamma$ is said to describe the point $a_\Gamma \in \mathbb{Z}^2$, where

$$a_\Gamma = \gamma_m2^m + \gamma_{m-1}2^{m-1} + \cdots + \gamma_02^0.$$ 

Let

$$M = \{a|a = a_\Gamma \text{ for } \Gamma \in V\}$$

denote the set of all points in $\mathbb{Z}^2$ described by a vector string, and let

$$M_t = \{a_\Gamma|a_\Gamma \in M, |\Gamma| = t\},$$

where $|\Gamma|$ is the string length of $\Gamma$:

$$|\Gamma| = |\gamma_m2^m| + |\gamma_{m-1}2^{m-1}| + \cdots + |\gamma_02^0|.$$ 

Thus,

$$M = \bigcup_{t=0}^{\infty} M_t.$$
III. PROPERTIES OF THE BASIC GROWTH PATTERN

The keystone of this section is the proof that 1-cells in Ulam's pattern are exactly those cells described by a vector string. Like the closed form solution to a recurrence relation, the vector string descriptor provides the means to establish specific characteristics. In particular, it is shown that the density of 1-cells approaches a limit, a conjecture put forth by Ulam [17], and that at generations which are a power of 2, only four cells, one on each main stem, are produced.

THE VECTOR STRING DESCRIPTOR

Before proving the main result of this section, it is convenient to establish a number of properties of points described by vector strings. In this manner, we establish (1) an algorithm to determine the vector string descriptor \( \Gamma \) from \( a_{\Gamma} \), and (2) a simple test to determine whether or not a cell \( a \) is described by some vector string. With respect to the former, it is important to know if more than one distinct vector string describes the same point. Such is not the case, as is shown below.

**LEMMA 1**

\[ a_{\Gamma} = a_{\Delta} \text{ if and only if } \Gamma = \Delta. \]

**Proof.**

If: This follows directly from the definition.

Only if: The proof here proceeds by contradiction. That is, assume

\[ a_{\Gamma} = a_{\Delta}, \]

where \( \Gamma \neq \Delta \). We have

\[ a_{\Gamma} - a_{\Delta} = 0 = \tau_0 2^1 + \tau_{s-1} 2^{s-1} + \cdots + \tau_0 2^0, \]

where \( |\tau| < 2 \). Consider the smallest \( s \) such that \( \tau_s \neq 0 \). Since \( 2^{s+1} \) divides 0 and all terms \( 2^i, \ i > s + 1 \), 2 must divide \( \tau_i \). Thus, \( \tau_i \in \{(2,0),(2,0), \ (0,2),(0,-2)\} \). Without loss of generality, we can assume \( \tau_s = (2,0) \). Thus,

\[ \frac{1}{2^{s+1}} (a_{\Gamma} - a_{\Delta}) = \tau_s 2^{s-1} + \tau_{s-1} 2^{s-2} + \cdots + [\tau_{s+1} + (1,0)] 2^0 = 0. \]

Since 2 divides 0, it must be that

\[ [\tau_{s+1} + (1,0)] \in \{(0,0),(2,0),(0,2),(0,-2)\}. \]

\(^2\) is said to divide \( a = (a_0, a_1) \) if it divides both \( a_0 \) and \( a_1 \).
VECTOR STRING DESCRIPTOR

But $\tau_{r+1} = (1,0) \neq (-2,0), (0,2), \text{ or } (0,-2)$, since equality would imply
$\tau_{r+1} = (-3,0), (-1,2), \text{ or } (-1,-2)$, respectively, and no combination of
two vectors in the fundamental vector set can produce these. Thus, $\tau_{r+1} = (1,0) \text{ or } (-1,0)$. If $\tau_{r+1} = (1,0)$, by a similar argument it can be shown that
$\tau_{r+2} = (1,0) \text{ or } (-1,0)$. If $\tau_{r+2} = (1,0)$, then $\tau_{r+3} = (1,0) \text{ or } (-1,0)$, etc. Since
the vector string is finite, there is a $j, j+1 \leq j < b$, such that $\tau_j = (-1,0)$ and
\[ a_r - a_\Delta \text{ has the form } \]
\[ a_r - a_\Delta = \tau_02^b + \cdots + (-1,0)2^j + (1,0)2^{j-1} + \cdots + (1,0)2^{j+1} + (2,0)2^j = 0. \]

However, it follows from this that either $\Gamma$ or $\Delta$ has a substring $(-1,0)$
$(0,0)^j(1,0)$, contradicting the noncancellation property. It must be that
$\tau_j = 0$, from which it follows that $\Gamma = \Delta$. Q.E.D.

The uniqueness of vector strings guarantees that an algorithm for
determining $\Gamma$ from $a_r$ produces one vector string only. The basis for such an
algorithm is established in the following lemma.

**LEMMA 2**

Let $a_r = (a_0, a_1) = \gamma_0 2^b + \gamma_1 2^{b-1} + \cdots + \gamma_0 2^0$ be an element of $M$, where
$\gamma_b \neq 0$. Then
\[ 2^b < |a_0| + |a_1| < 2^{b+1} - 1. \]

**Proof.** The upper bound is attained when
(a) $\gamma_i \neq 0$ for $0 < i < b$, 
(b) $(1,0)$ and $(-1,0)$ do not both appear in $\Gamma$, and 
(c) $(0,1)$ and $(0,-1)$ do not both appear in $\Gamma$.

The lower bound is attained when $\gamma_i = 0$ for $0 < i < b-1$. That no vector
string achieves a lower bound follows directly from the noncancellation property. Q.E.D.

Consider $a_r = (a_0, a_1) = \gamma_0 2^b + \gamma_1 2^{b-1} + \cdots + \gamma_0 2^0$. Given $a_r$, the algo-
rithm to determine $\Gamma$ proceeds by extracting $\gamma_b$ first, $\gamma_{b-1}$ second, etc. From
Lemma 2, $b$ is determined as the largest power of 2 less than $|a_0| + |a_1|$. $\gamma_b$ is
determined as follows:
(a) $\gamma_b = (0,1)$ if $|a_i| > |a_0|$ and $a_1 > 0$. 
(b) $\gamma_b = (0,-1)$ if $|a_i| > |a_0|$ and $a_1 < 0$. 
(c) $\gamma_b = (1,0)$ if $|a_0| > |a_i|$ and $a_0 > 0$. 
(d) $\gamma_b = (-1,0)$ if $|a_0| > |a_i|$ and $a_0 < 0$.

When $b$ and $\gamma_b$ are calculated, a new vector $a_r - \gamma_0 2^b = \gamma_{b-1} 2^{b-1} + \gamma_{b-2} 2^{b-2} + \cdots + \gamma_0 2^0$ is formed and $\gamma_{b-2}$ determined. The algorithm continues until
the last fundamental vector is found.
As an example, we solve for the vector string representation \( \Gamma \) of \( a_r = (4, -21) \):

\[
\begin{align*}
\mathbf{a}_r &= (4, -21): \quad |4| + |-21| = 25, \quad b = 4, \gamma_4 = (0, -1), \\
\mathbf{a}_r &= (0, -1)2^4 = (4, -5): \quad |4| + |-5| = 9, \quad b = 1 = 3, \quad \gamma_3 = (0, -1), \\
\mathbf{a}_r &= (0, -1)2^3 - (0, -1)2^2 = (4, 3): \quad |4| + |3| = 7, \quad b = 2 = 2, \quad \gamma_2 = (1, 0), \\
\mathbf{a}_r &= (0, -1)2^2 - (0, -1)2^3 - (1, 0)2^2 = (0, 3): \\
&\quad \quad |0| + |3| = 3, \quad b = 3 = 1, \quad \gamma_1 = (0, 1), \\
\mathbf{a}_r &= (0, -1)2^4 - (0, -1)2^3 - (1, 0)2^2 = (0, 1): \\
&\quad \quad |0| + |1| = 1, \quad b = 4 = 0, \quad \gamma_0 = (0, 1).
\end{align*}
\]

Thus,

\[
\Gamma = (0, -1)(0, -1)(1, 0)(0, 1)(0, 1).
\]

The algorithm fails when the original vector or an intermediate one has equal components. For this case, there is no corresponding vector string, and the cell in question is not part of \( M \). Thus, the algorithm also functions as a test to determine if a particular point is part of \( M \). However, a simpler test follows from Lemma 3.

\textbf{Lemma 3}

\textit{Given} \( \mathbf{a} = (a_0, a_1) \), \textit{then} \( \mathbf{a} \in M \) \textit{if and only if} \( (|a_0|/2^s) \mod 2 \neq (|a_1|/2^s) \mod 2 \), \textit{where} \( s \) \textit{is the largest} \( i \) \textit{such that} \( 2^i \) \textit{divides} \( \mathbf{a} \).

\textit{Proof.}

\textit{If:} Consider a cell \( \mathbf{a} = (a_0, a_1) \) such that

\[
\frac{|a_0|}{2^s} \mod 2 \neq \frac{|a_1|}{2^s} \mod 2,
\]

for \( s \) the largest \( i \) such that \( 2^s \) divides both \( a_0 \) and \( a_1 \). If \( \mathbf{a} \in M \), then \( s \) is also the largest \( i \) such that \( \gamma_i = 0 \). Apply the vector string generation algorithm. This always concludes successfully, since \( (|a_0|/2^s) \mod 2 \neq (|a_1|/2^s) \mod 2 \) guarantees that \( a_0 \neq a_1 \), for intermediate vectors \( \mathbf{a}' = (a_0', a_1') \) between the extraction of \( \gamma_0 \) and \( \gamma_s \). The extraction of \( \gamma_i \) concludes the algorithm, since the condition also guarantees \( \gamma_i = 0 \) for all \( i < s \).

\textit{Only if:} Consider \( \mathbf{a} \in M \). There exists a vector string \( \Gamma \) such that

\[
\mathbf{a} = \mathbf{a}_r = \gamma_0 2^k + \gamma_{-1} 2^{k-1} + \cdots + \gamma_0 2^0.
\]
VECTOR STRING DESCRIPTOR

Let $s$ be the smallest $i$ such that $\gamma_i \neq 0$. Thus, $2^s$ divides $a$ and

$$\frac{1}{2^s} a = \frac{1}{2^s} a_t = \gamma_0 2^{h-s} + \gamma_{h-1} 2^{h-s-1} + \cdots + \gamma_s 2^0.$$  

Considering $a_0$ and $a_1$ as a sum of powers of $2$, $\gamma_i$ contributes $1 \mod 2$ to either $|a_0|/2^s$ or $|a_1|/2^s$ while all other $\gamma_i$, $i > s$, contribute $0 \mod 2$ to either. Thus,

$$\frac{|a_0|}{2^s} \mod 2 = \frac{|a_1|}{2^s} \mod 2.$$

Q.E.D.

The test indicated by this lemma consists of finding the largest power of 2 which divides both components of a point $a = (a_0, a_1)$. If the result is one even and one odd number, then $a \in M$. Otherwise, $a \notin M$. Applying this test to (3, 3) and (6, 2), for example, shows that both points are not part of $M$ (see Fig. 2). If $a_0$ and $a_1$ are both odd, the test is trivial; the point is not part of $M$.

EQUIVALENCE OF ULM'S PATTERN AND THE VECTOR STRING REPRESENTATION

We now proceed to prove the equivalence of Ulam's pattern and the points described by vector strings. The proof is accomplished by showing for all $n$ that $U_n$, the set of cells "born" in generation $n$, is exactly $M_n$, the set of cells in $M$ with length $n$. In particular, we show that all points $a \in M_n$ are adjacent to exactly one 1-cell located at $b$, where $b \in U_{n-1}$, and that all other cells adjacent to $a$ are in the 0 state just prior to time $n$. Thus, according to the $u$ rule, the cell in question undergoes a 0 to 1 state transition. Thus, $a \in U_n$, and $M_n \subseteq U_n$. To show that $U_n \subseteq M_n$, we prove that all $a \in U_n$ have a vector string descriptor of length $n$. Formally,

**Theorem 1**

For all $n$, $M_n = U_n$.

Proof. (By induction.) For $n = 0$, we have $U_0 = M_0 = \{0\}$. Assume $M_i = U_i$ for $i < n$.

(a) $M_n \subseteq U_n$. Consider $a \in M_n$, where $\Gamma = \gamma_0 \gamma_{h-1} \cdots \gamma_0$. Let $s$ be the smallest $j$ such that $\gamma_j \neq 0$, and form $\Gamma' = \gamma_0 \gamma_{h-1} \cdots \gamma_j \gamma_{h-1} \cdots \gamma_0$, where $\gamma'_0 = 0$ and $\gamma'_{h-1} = \cdots = \gamma_0 = \gamma_s$. $\Gamma'$ describes a point $a_{\Gamma'}$, which is adjacent to $a_\Gamma$ and which is an element of $M_{\Gamma'}$. From the inductive hypothesis, $a_\Gamma \in U_{\Gamma'}$ also. Thus, $a_\Gamma$ is adjacent to at least one 1-cell. It now remains to show that the other three adjacent cells are in the 0 state. There are two cases: (1) $\gamma_0 \neq 0$ and (2) $\gamma_0 = 0$. We will consider only the first; the second proceeds analogously. If $\gamma_0 \neq 0$, $a_\Gamma$ must have one odd and one even component. $a_\Gamma$. 


because it is adjacent to $a_r$, must have either two even or two odd components. However, Lemma 3 precludes the latter. Thus, it follows that the cells adjacent to $a_r$ which share a corner with $a_r$ have two odd components. From Lemma 3, neither belong to $M$, and thus both are 0. The remaining cell, depending on $\Gamma$, is described by a vector string of length $n+1$ or is not described by any vector string. In either case, it is in the 0 state at the time the $u$ rule is applied. Thus, $a_r \in U_n$.

(b) $U_n \subseteq M_n$. Consider $a \in U_n$, $a$ must therefore be adjacent to exactly one $b \in U_{n-1}$, and from the inductive hypothesis, $b \in M_{n-1}$. Let $b$ be described by $\Gamma = \gamma_b \gamma_{b-1} \cdots \gamma_0$. There are two cases: (1) $\gamma_0 \neq 0$ and (2) $\gamma_0 = 0$. Consider (1) first. Let $\varphi$ be the smallest $j$ such that $\gamma_j \neq \gamma_0$. Either (i) $\gamma_j = 0$ or (ii) $\gamma_j \neq \gamma_0$ (from the noncancellation property, $\gamma_j \neq -\gamma_0$). The latter case is impossible, since another cell $c$ can be found that is also adjacent to $a$ and that is described by $\Gamma'$ such that $|\Gamma'| = |\Gamma|$. But this contradicts $a \in U_n$. It must be that $\gamma_0 = 0$, in which case $\Gamma = \gamma_b \gamma_{b-1} \cdots 0 \gamma_0 0$ describes $b$. If $b - a = 0$, then $\gamma_b \gamma_{b-1} \cdots 0$ describes $a$, and thus $a \in M_n$. On the other hand, if $b - a \neq 0$, then $\gamma_b \gamma_{b-1} \cdots 0 0 \gamma_0 \gamma_0 \cdots \gamma_0 (b - a)$ is a vector string of length $|\Gamma|$ which describes a point adjacent to $a$. This case is impossible, since $a \in U_n$. Trivially, $b - a = -\gamma_0$ is impossible also. For (2), $\gamma_0 = 0$, $\Gamma = \gamma_b \gamma_{b-1} \cdots \gamma_1 (b - a)$ describes $a$, and thus $a \in M_n$. Q.E.D.

From Theorem 1 it is immediately clear that for any 1-cell in Ulam's pattern there exists a path of adjacent cells from the origin 0 and that this path is described by the corresponding vector string. Cells extending in a straight line from the north, east, south, and west stems, in particular, are characterized by vector strings which contain only one type of fundamental vector. The northeast, southeast, southwest, and northwest diagonals are never occupied with 1-cells because from Lemma 3 no vector string can describe any of these.

A consequence of Theorem 1 is the following.

**COROLLARY 1**

$U = M$.

**PROPERTIES OF ULAM'S PATTERN**

The rate of growth of Ulam's pattern can be measured by the number of 1-cells $|U_n|$ added at each generation $n$. From Theorem 1 we have $|U_n| = |M_n|$, and from Lemma 1 each element of $M_n$ is described by a unique vector string. Thus, $|U_n|$ is exactly the number of vector strings of length $n$, where

$$n = |\gamma_b|^2 + |\gamma_{b-1}|^2 + \cdots + |\gamma_0|^2.$$  \hfill (1)

Since $|\gamma_0|$ is either 0 or 1, (1) is just the binary expansion of the number $n$. 

VECTOR STRING DESCRIPTOR

For example, 20 represented as a binary number is 10100, and so cells in $U_{20}$ have vector strings with $|\gamma_4|=1$, $|\gamma_3|=0$, $|\gamma_2|=1$, $|\gamma_1|=0$, and $|\gamma_0|=0$. The number of vector strings of length $n$ can be calculated in a straightforward manner as follows. For all occurrences of $|\gamma|=0$ there is only one choice for $\gamma$: $\gamma=0$. For $\gamma_b (|\gamma_b|=1)$, there are four choices: $(1,0), (-1, 0), (0, 1)$, and $(0, -1)$. For the remaining components $\gamma$ such that $|\gamma|=1$, only three choices exist, since the noncancellation property excludes as a choice the negative of the previous vector. This proves the following.

**Lemma 4**

$$|U_n| = 4 \times 3^{(n)-1}, \text{ where } t(n) \text{ is the number of 1's in the binary equivalent of } n.$$  

For example, $|U_{20}| = 12$. As one might expect, for large values of $n$ the potential for large values of $|U_n|$ is better than for small values of $n$, simply because in large binary numbers there are more places to put 1's. But asLemma 4 shows, $|U_n|$ can take on only specific values, and these are directly related to the number of 1's in the binary expansion of $n$, a somewhat unexpected result.

Lemma 4 is a generalization of a theorem cited in Ulam [17] that growth is restricted to the four main stems at generations $n = 2^m, \ m = 0, 1, 2, \ldots$. Indeed, for these values there is only one 1 in the binary expansion of $n$, and Lemma 4 yields $|U_{2^m}| = 4$. At generation $2^m - 1$ the pattern is a square bounded by cells born in generation $2^m - 1$ (viz. generations 3 and 7 in Fig. 2). All 0-cells along the sides are adjacent to two 1-cells, and so no growth occurs there. Further, no room exists within the square for additional cells. Thus, once the pattern within the square is established it remains forever unchanged. This observation leads to a calculation of the relative density of 1- and 0-cells as follows.

The area occupied by the square $a_{2^m - 1}$ is

$$a_{2^m - 1} = 2 \times 4^m - 2^{m+1} + 1. \quad (2)$$

The number of cells $n_{2^m - 1}$ in $U$ which occupy this area is obtained by summing $|U_n|$ over $n$:

$$n_{2^m - 1} = \sum_{i=0}^{2^m - 1} |U_i|. \quad (3)$$

Since there are $\binom{m}{i}$ $m$-bit binary numbers with $i$ 1's and each contributes $4 \times 3^{i-1}$ cells, $n_{2^m - 1}$ can be written as

$$n_{2^m - 1} = \sum_{i=1}^{m} \binom{m}{i} 4 \times 3^{i-1} + \binom{m}{0} 1. \quad (3)$$
or

\[ n_{2^m-1} = \frac{4^{m+1} - 1}{3}. \]  

(4)

The limiting density of Ulam's pattern, \( d_U \), is defined as

\[ d_U = \lim_{m \to \infty} \frac{n_{2^m-1}}{a_{2^m-1}}. \]  

(5)

Substituting (3) and (4) into (5) yields the following result:

**LEMMA 5**

The limiting density of the pattern generated by application of the \( u \) rule to the primitive initial configuration is \( \frac{1}{3} \).

Thus, for large patterns in areas where growth is complete, we can expect that approximately 2 out of 3 cells will be in the 1 state. Lemma 5 confirms a conjecture by Ulam [17] concerning the existence of a limiting density.

Another useful concept is fringe growth. Formally, fringe growth \( F_n \) is the set of cells

\[ F_n = \{ a | a \in U_n, |a_0| + |a_1| = n \}. \]

Considering a cell \( a \in Z^2 \), there is a period of time before which growth has not progressed to \( a \). Since a cell cannot become 1 unless it is adjacent to a cell which is already 1, the earliest time \( a \) 0 to 1 transition can occur at \( a \) is the number of cells in the path of adjacent cells 0 to \( a \). This is \( |a_0| + |a_1| \).

Thus, fringe growth represents the set of cells which become 1 at the earliest possible time.

This concept is intimately related to the growth cutoff for generations of the form \( 2^m \). At any particular time, \( t \), prospective new fringe growth is limited to cells \( a = (a_0, a_1) \) for which \( |a_0| + |a_1| = t \). These cells form the border of the square pattern at generation \( 2^m - 1 \). Consider a cell \( a \in Z^2 \) such that \( |a_0| + |a_1| = 2^m - 1 \). Since \( 2^m - 1 \) is odd, one of \( |a_0| \) and \( |a_1| \) must be odd and the other even. Thus, from Lemma 3, \( a \in M \). Since \( |a_0| + |a_1| = 2^m - 1 \), it must be that \( \gamma_i \neq 0 \) for all \( i < m - 1 \), and \( -\gamma_i \) must not be included in the vector string if \( \gamma_i \) appears. Thus, \( |\gamma_{m-1}|2^{m-1} + |\gamma_{m-2}|2^{m-2} + \cdots + |\gamma_0|2^0 = 2^m - 1 \), and \( a \in F_{2^m-1} \). Since all cells in the square of prospective fringe growth cells are indeed occupied by cells which become 1 in the same generation, cells in the next larger square are 0 because they are adjacent to two such cells. The only exception are the four 0-cells along the stems, which are adjacent to only one 1-cell. Thus, it is fringe growth which produces the cutoff effect.
VECTOR STRING DESCRIPTOR

The vector string description of a cell which is part of fringe growth is characterized by the restricted nature in the choice of fundamental vectors. A fringe growth cell \( a_t \) with vector string

\[
\Gamma = \gamma_{m-1} 2^{m-1} + \gamma_{m-2} 2^{m-2} + \cdots + \gamma_0 2^0
\]

can be specified by choosing all nonzero vectors the same (4 ways) or by choosing two orthogonal vectors (4 ways) and assigning them in any of \( 2^n - 2 \) ways to the nonzero vectors, where \( t(n) \) is the number of 1's in \( n = |\Gamma| \). Totaling the choices yields

**LEMMA 6**

The number of cells in the fringe growth cell set \( F_n \) is

\[
|F_n| = 4 \times 2^{t(n)} - 4,
\]

where \( t(n) \) is the number of 1's in the binary representation of \( n \).

Thus,

**COROLLARY 2**

\[
\lim_{t(n) \to \infty} \frac{|F_n|}{|U_n|} = 0.
\]

As \( t(n) \) increases, one expects a smaller fraction of the cells in \( U_n \) to be fringe growth cells. In comparison with \( |F_n|/|U_n| \) for small values of \( n \), we see that for \( 1 < n < 6 \), \( |F_n| = |U_n| \); the pattern consists entirely of fringe growth cells. This will, of course, be true in general when \( t(n) = 1 \) or 2.

Fringe growth, it will be shown, is also a significant factor in the determination of the properties of patterns resulting from the application of the \( u \) rule to arbitrary initial patterns.

**IV. GROWTH PATTERNS FROM FINITE INITIAL CONFIGURATIONS**

In this section we investigate patterns which result from successive applications of the \( u \) rule to arbitrary finite initial configurations. By finite we mean containing only a finite number of 1-cells. The extension of the vector string descriptor to this more general case is complicated by the fact that the state of a cell is determined, in general, by more than one 1-cell in the initial configuration. However, the fringe growth concept can be extended to provide significant information about the case at hand. For example, it is shown that the limiting density for any finite initial pattern is \( \frac{2}{3} \).
Let $R \subset \mathbb{Z}^2$ be the set of l-cells in the initial configuration, and define

$$R_{SE} = \{(a_0, a_1) \in R | a_1 - a_0 \leq b_1 - b_0, \forall (b_0, b_1) \in R\}.$$ 

$R_{SE}$ is the (initial) southeast diagonal of $R$ and represents the set of l-cells which are “southeasternmost” in the set. $R_{SW}$, $R_{NW}$, and $R_{NE}$ are defined similarly. One can think of a rectangle whose sides are oriented at 45° to the coordinate axis that just encloses $R$. The l-cells falling along the southeast edge constitute $R_{SE}$. In the description which follows, we will be concerned only with the southeast diagonal, but because of symmetry the results apply as well to the other diagonals. The significance of the four diagonals is that, because they exist on the outer edge of the initial pattern, they determine to a large extent the form of the growth pattern in the corresponding quadrant.

To make this statement more precise, consider the following. Let $r_{SE}^{NE}$ and $r_{SW}^{NE}$ denote the northernmost (easternmost) and the southernmost (westernmost) elements of $R_{SE}$ respectively. Note that $r_{SE}^{NE}$ is the source of an arbitrarily long stem which grows eastward, and $r_{SE}^{SW}$ is the source of an arbitrarily long stem growing southward. Also $r_{SE}^{NE} = r_{SE}^{SW}$ when $R_{SE}$ has only one element.

Denote the southeast diagonal length as $l_{SE}$, where $r_{SE}^{NE} - r_{SE}^{SW} = (l_{SE}, l_{SE})$, and define the $n$th-order (southeast) diagonal $D_{SE}^n$ recursively as follows:

$$D_{SE}^0 = \{r_{SE}^{SW} + (i, i), 0 \leq i < l_{SE}\}$$

and

$$D_{SE}^n = \{a | a \text{ is a south or east neighbor of } b \in D_{SE}^{n-1}\}.$$ 

$D_{SE}^0$ is a diagonal, of all points bounded by $r_{SE}^{SW}$ and $r_{SE}^{NE}$. To the southeast is $D_{SE}^1$, followed by $D_{SE}^2$, etc. Let the southeast quadrant $Q_{SE}$ be the set \{a | a \in D_{SE}, 0 < i\}.

Let the fringe growth $F^R$ of a finite configuration $R$ be the set of cells in the $n$th diagonal which undergo a 0 to 1 state transition after the $n$th application of the $u$ rule for all $n > 0$. As with the primitive initial configuration, fringe growth represents cells which become 1 at the earliest possible time. It is important to note that:

1. Fringe growth in the $n$th order diagonal depends only on fringe growth in the $n-1$th order diagonal.

2. When a fringe growth cell $c \in F^R$ in the $n$th order diagonal undergoes a 0 to 1 transition, $q_{d(n-1)}$, $q_{e+(1,0)(n-1)}$, and $q_{e+(0,1)(n-1)}$—the states of $c$ and its eastern and southern neighbor respectively at time $n-1$—are always 0.
VECTOR STRING DESCRIPTOR

These observations lead to the following.

**LEMMA 7**

Let \( q^*_c(n) = 1 \) (0) denote that cell \( c \) is (is not) in the fringe growth pattern of the \( n \)th order diagonal at time \( n \). Then

\[
q^*_c(n) = q^*_c(-1,0)(n-1) \oplus q^*_c(0,1)(n-1).
\]

(6)

Equation (6) is a recursive relation for generating the fringe growth in the southeast quadrant, and in this sense is much like a local transition function. However, (6) does not apply simultaneously to all cells in the array—for example, in any of the other three quadrants. In general, to find the fringe growth in the \( n \)th order diagonal, one must first find the fringe growth in the 1st order diagonal, the 2nd, etc. However, as we show now, for \( n \) which is a power of 2, the determination of the fringe growth is particularly simple.

**LEMMA 8**

\[
q^*_c(2^m + n) = q^*_c(-2^{m-1},0)(n) \oplus q^*_c(0,2^{m-1})(n).
\]

(7)

**Proof** (By induction on \( m \)). For \( m = 0 \), (7) follows directly from (6). Assume the hypothesis to be true for \( m - 1 \). Then we have

\[
q^*_c(2^{m-1} + (2^{m-1} + n)) = q^*_c(-2^{m-1},0)(2^{m-1} + n) \oplus q^*_c(0,2^{m-1} + n).
\]

(8)

\[
q^*_c(2^{m-1} + (2^{m-1} + n)) = q^*_c(-2^{m-1},0)(n) \oplus q^*_c(-2^{m-1},0) + (2^{m-1} + n),
\]

(9)

and

\[
q^*_c(2^{m-1} + n) = q^*_c(0,2^{m-1})(n) \oplus q^*_c(-2^{m-1},0)(n).
\]

(10)

Substituting (10) and (9) into (8) yields the derived result. Q.E.D.

The rather surprising result that a cell \( 2^m \) diagonals away from the \( n \)th diagonal depends on only two cells in that diagonal has an important effect on the fringe growth pattern. In particular, if the distance between \( c + (-2^m,0) \) and \( c + (0,2^m) \) is sufficiently large, then at least one of \( q^*_c(-2^m,0) \) and \( q^*_c(0,2^m) \) in (7) is 0 for any \( c \) in the \( 2^m + n \)th order diagonal. Since \( a = a \oplus 0 \), we have

**THEOREM 2**

For all values of \( m \) such that \( 2l^*_SE < 2^m \), the \( n \)th order (southeast) diagonal is replicated in the fringe growth of the \( 2^m + n \)th order diagonal once to the east and once to the south, where \( l^*_SE \) is the length of the \( n \)th order diagonal.
Thus, a basic form of self-reproduction exists in the fringe growth pattern. The southeast diagonal as well as all subsequent diagonals are replicated an infinite number of times. Theorem 2 shows that the u rule is closely related to Fredkin's rule (described in Winograd [19]) that operates on the neighborhood of four adjacent cells and has the specification

\[ q_e(t) = q_{e+1,0}(t-1) \oplus q_{e-1,0}(t-1) \oplus q_{e+0,1}(t-1) \oplus q_{e+0,-1}(t-1). \]

As is shown by Winograd, for generations which are a significantly large power of 2, the original pattern is replaced by four replicas, one each to the north, east, south, and west.

The fact that two replicas are reproduced is related to the following.

**COROLLARY 2**

The number of 1-cells in the fringe growth of the nth order diagonal is even for all n except possibly 0.

The result stated in Corollary 2 can be seen in the example shown in Fig. 3. The nth order diagonal, with 1 cells shown here as solid circles can be considered as a collection of groups consisting of diagonally adjacent 1 cells where each group is separated from its neighbor by at least one 0 cell (shown as a small dot). For the case at hand, there are four groups, one group of three and three groups of one. The generation of fringe growth in the n+1th diagonal (with 1-cells shown as circles), shows that each group produces exactly two offspring 1-cells.

It is important to note that all diagonals with an even number of 1-cells have a unique predecessor. The predecessor can be formed by dividing the line into pairs of 1-cells and by putting 1's in the predecessor diagonal so that the elements of each pair are "linked".

![Diagram](image-url)  
**Fig. 3.** Fringe growth along a diagonal.
VECTOR STRING DESCRIPTOR

Doing this to the \( n \)th order diagonal of Fig. 3 produces an \( n-1 \)th order diagonal with an odd number of cells. From Corollary 2, it follows that no \( n \)-2th order diagonal can exist. Thus, in this example there can only be two values for \( n \): 0 or 1. Since any diagonal \( G \) with an odd number of 1-cells can exist only initially, it is appropriate to denote it as a Garden of Eden (GOE) diagonal.

The significance of the GOE diagonal is that, although it may not exist in the initial configuration, the fringe pattern behaves as if it is, in fact, there. This can be seen as follows. Consider the southeast quadrant generated from some southeast diagonal \( R_{SE} \). If \( R_{SP} \) is not GOE, form its predecessor. If the predecessor is not GOE, form its predecessor, etc. A GOE diagonal will be formed eventually (the process cannot go on beyond a single 1-cell which is GOE). Denote this GOE diagonal by \( G_{R_{SE}} \) and consider the pattern generated in the southeast quadrant by \( G_{R_{SE}} \). Except for the diagonals between \( G_{R_{SE}} \) and \( R_{SE} \), it is identical to the quadrant generated from \( R_{SP} \). But from Theorem 2 it follows that \( G_{R_{SE}} \) is replicated once to the south and once to the east in the quadrant generated by \( G_{R_{SE}} \). It must therefore appear in the quadrant generated by \( R_{SE} \).

Associated with replication of GOE diagonals is the cutoff effect discussed in the previous section. To investigate this further, consider the following.

**Lemma 9**

\[
q_e^k(2^m - 1 + n) = q_e^k(-2^m+1,0) \oplus q_e^k(-2^m+2,1) \oplus \cdots \oplus q_e^k(0,2^m-1)(n) \tag{11}
\]

**Proof.** The proof proceeds by induction on \( m \) and is very similar to that of Lemma 8.

Lemma 9 shows that a cell \( e \) in the \( 2^m - 1 + n \)th order diagonal will be a 1-cell in the fringe growth pattern if there are an odd number of fringe growth cells in the diagonal consisting of \( e + (-2^m+1,0), e + (-2^m+2,1), \ldots, e + (0,2^m-1) \). Thus, it follows that diagonals \( 2^m - 1 \) beyond a GOE diagonal contain a line of contiguous 1-cells. Fig. 4 illustrates this. \( R_{SE} \) is the southeast diagonal, and \( G_{R_{SE}} \) is the GOE diagonal corresponding to it. The line labeled “\( 2^m - 1 \) cutoff” is the diagonal of 1-cells generated according to Lemma 9. The successor to this diagonal is a distance \( 2^m \) away from \( G_{R_{SE}} \), and so according to Theorem 2, it contains two replicas of \( G_{R_{SE}} \). Each replica, in turn, produces a fringe growth which is the same as that from \( G_{R_{SE}} \) except in the overlap region 0. The hatched area labeled \( B \) (background area) contains no fringe growth cells, because it is effectively cut off by the line of 1-cells spanning the two replicas.

However, \( B \) will be occupied by non-fringe-growth cells which stem from the two replicas. Let \( B(m) \) be the number of 1-cells in \( B \), and let \( n(m) \) be
the number of 1-cells in the sector to the northwest of $B$ when growth in those regions is complete. For the moment we are including cells to northwest of $R_{SE}$ even though they are not part of the pattern. Similarly, $n(m+1)$ denotes the number of cells in the pattern from $G_{R_{SE}}$ to the $2^{m+1}-1$ cutoff. Thus,

$$n(m+1) = 3n(m) + B(m) - O(m)$$  \hspace{1cm} (12)$$

where $O(m)$ is the number of 1 cells eliminated by competition in the overlap area $O$. If $m$ in (12) is sufficiently large, $O$ will be a sufficiently small portion of growth pattern. Thus,

$$\lim_{m \to 0} \frac{O(m)}{n(m)} = 0,$$

and $O(m)$ can be neglected. The same is not true of $B(m)$, of course. However, growth in region $B$ is identical to the corresponding region in Ulam's pattern except for a band of cells along the diagonal extending from $G_{R_{SE}}$ through the middle of the growth pattern. As $n$ increases arbitrarily, the band of discrepancy becomes a vanishingly small fraction of total area, and so we have

$$B(m) = \frac{1}{4}C_{2^{m}-1} \quad \text{for large } m.$$
VECTOR STRING DESCRIPTOR

Thus, for large $m$ (12) becomes

$$n(m + 1) = 3n(m) + \frac{1}{2} 4^m - \frac{1}{10}.$$  

Solving for the closed form solution yields

$$n(m) = A 3^m + \frac{1}{2} 4^m + \frac{1}{14},$$

where $A$ depends on initial conditions. The area occupied by $n(m)$ is approximately

$$a(m) = \frac{1}{2}(2^m - 1)^2.$$  

We have

$$\lim_{m \to \infty} \frac{n(m)}{a(m)} = \frac{2}{3}.$$  

Since the area and the number of 1-cells in the growth pattern between $G_{\text{SE}}$ and $R_{\text{SE}}$ are finite, the density of 1-cells is the same regardless of which generates the pattern. Note that $n(m)$ for large $m$ is independent of the form of the diagonal. Thus, the southwest, northwest, and northeast quadrant produce the same result. Further, since those areas not covered by any quadrant are a vanishingly small fraction of the total area, we have

**LEMMA 10**

The limiting density of 1-cells for the u rule applied to any finite initial configuration is $\frac{2}{3}$.  

Thus, although the initial configuration determines the structure of the pattern, one quantity, the density, is the same for all initial patterns.

V. OTHER GROWTH RULES

The vector string notation can be used to describe the patterns resulting from a variety of growth rules. A pattern closely related to Ulam’s pattern is obtained by applying the growth rule shown in Fig. 5 to a single 1-cell. The resulting pattern, $U_T$, is a spread-out 45° rotation of Ulam’s pattern. The fundamental vector set for $U_T$ is $I’ = \{(0, 0), (1, 1), (1, -1), (-1, 1), (-1, -1)\}$. Except for that, the vector string representations of $U_T$ and $U$ are the same.

**THEOREM 3**

$$U_T = (Z^2 - U) \cup \{0\}.$$
Proof.

(a) $U_T \subseteq (Z^2 - U) \cup \{0\}$. Let $a \in U_T$. Then there exists $b \in U$ such that

$$b = \frac{1}{2}a \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$ 

If $a = b = 0$, trivially $a \in (Z^2 - U) \cup \{0\}$. For $a \neq 0$, $b$ has the property

$$(b_0/2^k) \mod 2 \neq (b_1/2^k) \mod 2 \text{ (Lemma 3).}$$

This implies that

$$\frac{a_0}{2^k} \mod 2 = \frac{b_0 + b_1}{2^k} \mod 2 = 1 \quad \text{and} \quad \frac{a_1}{2^k} \mod 2 = \frac{b_0 - b_1}{2^k} \mod 2 = 1.$$ 

Thus $a \in (Z^2 - U)$.

(b) $U_T \equiv (Z^2 - U) \cup \{0\}$. Let $a \in (Z^2 - U) \cup \{0\}$. If $a = 0$ trivially $a \in U_T$.

For $a \neq 0$, we have $(a_0/2^k) \mod 2 = (a_1/2^k) \mod 2 = 1$, since $a \not\in U$, for $k = \max(p|2^p \text{ divides } a)$. Consider $b = (b_0, b_1) = \frac{1}{2}(a_0 + a_1, a_0 - a_1)$. We have

$$\frac{b_0}{2^k} \mod 4 = \frac{1}{2} \left( \frac{a_0 + a_1}{2^k} \mod 4 \right) \quad \text{and} \quad \frac{b_1}{2^k} \mod 4 = \frac{1}{2} \left( \frac{a_0 - a_1}{2^k} \mod 4 \right).$$

From Lemma 3, it follows that $(a_0/2^k) \mod 4 = (a_1/2^k) \mod 4 = 1, 3$. Considering all possible combinations, we obtain

$$\frac{b_0}{2^k} \mod 4 = 0 \iff \frac{b_1}{2^k} \mod 4 = 1$$

and

$$\frac{b_0}{2^k} \mod 4 = 1 \iff \frac{b_1}{2^k} \mod 4 = 0.$$ 

Thus $b \in U$ and $a \in U_T$. Q.E.D.

An immediate consequence of Theorem 3 is

**Lemma 11**

The limiting density of $U_T$ is $d_{U_T} = \frac{1}{3}$.

This result can also be obtained by following the procedure used in the determination of $d_U$.  

\[ \]
A more general class of patterns described by vector strings is obtained from the application of the extended growth rules $R^k$, $R^{k*}$, shown in Fig. 6 to the primitive initial configuration. The dots represent intervening cells whose states have no effect on the next state of the center cell. Denote the set of cells destined to go to the 1 state as $U^k$ and $U^{k*}$ for rules $R^k$ and $R^{k*}$, respectively. The resulting patterns are expansions of $U$ and $U^{*}$. The vector string notation is directly applicable. The only difference is that the fundamental vector sets are $I_k = \{(0,0),(k,0),(0,k),(-k,0),(0,-k)\}$ and $I^{k*} = \{(0,0),(k,k),(-k,k),(k,-k),(-k,-k)\}$ respectively. Cutoff again occurs in the $2^n - 1$th ($m = 1, 2, \ldots$) generation and the number of 1-cells is also $n = \frac{1}{3} - \frac{1}{3} = \frac{1}{3} + \frac{1}{3} = \frac{1}{3}$ for both patterns. The areas covered by the patterns however are effectively multiplied by $k^2$. Thus, the limiting densities become $d_{U^k} = \frac{2}{3}k^2$, $d_{U^{k*}} = \frac{1}{3}k^2$ respectively. An interesting result is obtained by considering the density of $U^k$, $d_{U^k} = \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) = \frac{1}{3}$. The pattern resulting from this rule is exactly the pattern obtained from $U$ if all cells in $U_n$ for odd $n$ are removed. If we classify the cells in $U$ according to their neighborhood, we have three distinct classes of 1-cells. They are described by

\[
\begin{array}{c}
S \\
O \\
O \\
S \\
O \\
S \\
\end{array}
\]

where $O$ is the cell under consideration, $F$ is the father cell, and $S$ represents the son cell(s) if any. It is also true that $U^1$ is exactly the class described by $\Phi$. Furthermore, by looking at $U$ we note that cells of the forms $\Phi$ and $FO$ are alternating in all stems and substems, while $FO$ cells are grown out of $\Phi$ cells. From these observations it follows that $d_{O} = d_{\Phi} + d_{O} - d_{O} - d_{\Phi} = 0$, $d_{O} = \frac{1}{3} - \frac{1}{3} - \frac{1}{3} + \frac{1}{3}$. Thus, we have the limiting densities of the three classes of 1-cells. The same ratio of 1:1:2 is characteristic of the patterns $U^k$ and $U^{k*}$ ($k = 1, 2, \ldots$), also.

![Fig. 6. Rules $R^k$ and $R^{k*}$](image)
It is interesting to note that the rules $R^k_U$ and $R^0_T$ divide the plane into $k^2$ subplanes. That is, consider an initial configuration in which the only 1-cells are located within a square of $k \times k$ points. Each 1-cell produces a pattern which is the transposition of $U^k$ or $U_T^k$. (In fact, for any local transition function, the same neighborhoods effectively divide the cellular space into $k^2$ spaces or laminal subarrays; see Yamada and Amoroso [20].) If $o$ is the number of 1's within the initial $k \times k$ square, then rules $R^k_U$ and $R^0_T$ produce densities of $2o/3k^2$ and $o/3k^2$, respectively.

Another class of patterns is obtained by modifying the growth rules $R^k_U$ and $R^k_T$, so that a 0-cell becomes a 1-cell if there exists exactly one 1-cell among the extreme cells shown in Fig. 6 or among the intervening cells (shown as dots). Denote these rules as $R^k_C$ and $R^k_C$, and the patterns which grow from the primitive initial configuration as $U^k$ and $U_T^k$, respectively. The pattern $U^T$, for example, is shown in Fig. 7. The $T$ formations are characteristic of all patterns $U^k$, $U_T^k$. The patterns can be described by the sum of two vector strings. The first vector string describes the junction cells

![Diagram](image)

**Fig. 7.** Pattern $U^T$ obtained from $R^k_C$. 

---

96 JON T. BUTLER AND SIMEON C. NTAFOS
of each $T$, and it is in fact the vector string for $U^{2k}$ ($U^k$). Then the second vector string describes, for each junction cell, the other members of the $T$ formation. That is, for $U^k$, if $a \in U^k$, $a$ is described by $a = b + c$, where $b \in U^{2k}$, and if $b = \gamma_1 2^m + \cdots + \gamma_0 2^0$ and the nonzero vector of least index is $\gamma = (2k, 0)$ (similarly for $\gamma = (0, 2k), (-2k, 0), (0, -2k)$), then $c$ is any one of the vectors $(0, x), (0, -x), (-x, 0), 1 < x < k$. The limiting densities may be obtained as follows: The density of $U^{2k}$ is $\frac{1}{3}(1/2k)^2$. This describes the junction cells of each $T$ formation. Thus each cell actually represents $3k + 1$ cells. Then we have

**Lemma 12**

$$D_{U^k} = \frac{2}{3} \left( \frac{1}{2k} \right)^2 (3k + 1), \quad D_{U^T} = \frac{1}{3} \left( \frac{1}{2k} \right)^2 (3k + 1).$$

A general property of the expanded rules considered so far is that the density goes to zero as the expansion factor is increased. An interesting question is whether or not that is a general property of expanded growth rules. Let us consider the growth rule shown in Fig. 8, where a 0 to 1 transition occurs if exactly one of the shaded cells is 1. The resulting pattern, $S$, is shown in Fig. 9. Cutoff again occurs at the $2^n - 1$ generation ($n = 1, 2, \ldots$). Also note that the cells grown in the $2^n$th generation are in some respects equivalent to the original 1-cell, that is, the growth in the direction of the main stems (diagonals) is a duplication of the previous growth. It is conjectured that the general form of the pattern is given by Fig. 10. From Figs. 9 and 10 we obtain the relations

$$q_m = 2q_{m-1} + 2p_{m-1}, \quad (13)$$

where $q_m$ is the number of 1-cells grown in $2^m - 1$ generations. Also,

$$p_m = 3p_{m-1} + p_{m-1} + 3 + m - 2,$$

where $p_{m-1} = p_{m-1} - 1$. Then

$$p_m = 4p_{m-1} + m. \quad (14)$$

![Fig. 8. Growth rule on a $3 \times 3$ neighborhood.](image-url)
Fig. 9. Southeast quadrant of a pattern produced by the growth rule of Fig. 8 on a 3 x 3 neighborhood.

Fig. 10. Block diagram of the pattern $q_{m+1}$. 
VECTOR STRING DESCRIPTOR

Initially we have $p_1 = 1$, $q_1 = 4$. Using induction on (13) and (14), we obtain the relation

$$q_m = p_m + m + 2.$$ 

Then the expression for $q_{m+1}$ becomes

$$q_{m+1} = 2q_m + 2(q_m - m - 2) = 4q_m - 2m - 4.$$ 

Solving for $q_m$ yields

$$q_m = \frac{4}{9} \times 4^m + \frac{2}{9} m + \frac{16}{9}.$$ 

The area covered by $q_m$ is $a_m = (2^m - 1)^2$.

Then if the conjecture is true, the limiting density becomes

$$d = \lim_{m \to \infty} \frac{q_m}{a_m} = \lim_{m \to \infty} \frac{\frac{4}{9} \times 4^m + \frac{2}{9} m + \frac{16}{9}}{4^m - 2 \times 2^m + 1} = \frac{4}{9}.$$ 

Although no vector string seems to describe the pattern $S$, vector strings are useful in studying its expansions. For example, consider the growth rule shown in Fig. 11. The pattern which grows from a primitive initial configuration is shown in Fig. 12. Note that the even numbered cells are exactly the pattern $U^2_t$. Also note that each even numbered cell corresponds to 17 1-cells. Then we have

$$d_{S^1} = \frac{1}{3} \left( \frac{1}{2} \right)^7 \cdot 17 = \frac{17}{48}.$$ 

A 0 to 1 transition occurs when exactly one 1 cell exists in the shaded neighborhood.

Fig. 11. Growth rule on a $5 \times 5$ neighborhood.
Fig. 12. Southeast quadrant of a pattern generated by the growth rule of Fig. 11. on a $5 \times 5$ neighborhood.

Fig. 13. The arrowhead formation.
VECTOR STRING DESCRIPTOR

The arrowhead formation is characteristic of all patterns $S^k (k = 2, 3, \ldots)$. In general, the arrowhead formation has the form shown in Fig. 13.

Thus, the corner cell that is in $U^k_1$ actually corresponds to $(2k+1)^2-(k+1)^2+1$ cells, and the density of $S^k$ is given by

$$d_{S^k} = \left( \frac{1}{3} \right) \left( \frac{1}{2k} \right)^2 [(2k+1)^2-(k+1)^2+1] \quad \text{for } k > 2.$$  

A vector string notation exists for $S^k (k \geq 2)$. Again, it is the sum of two vector strings, of which the first describes the corner cell (even numbered generations) and the second describes the arrowhead formation. The first string is exactly the string describing $U^k$. If $a \in S^k$, then $a = b + c$, where $b \in U^k$; and if $b = y_1 2^n + \cdots + y_0 2^0$ and $y_i = (2k, 2k)$ is the nonzero vector of least index, then

$$e = \begin{cases} (x, y), & 1 \leq x < k, \quad -k \leq y < k, \\ (x, y), & -k \leq x < 0, \quad 1 \leq y < k. \end{cases}$$

An analogous statement is true for the cases where $y_i = (2k, -2k)$, $(-2k, 2k)$, or $(-2k, -2k)$. If we take the limit of the expression for $d_{S^k}$ as $k \to \infty$, we have

$$\lim_{k \to \infty} d_{S^k} = \lim_{k \to \infty} \left( \frac{1}{3} \right) \left( \frac{1}{2k} \right)^2 [(2k+1)^2-(k+1)^2+1] = \frac{1}{4}.$$  

Thus $d_{S^k}$, unlike $U^k$ and $U^k_1$, goes to $\frac{1}{4}$ instead of zero as the index $k$ goes to infinity.

V. CONCLUDING REMARKS

Our interest in growth patterns is twofold. First, very little work has been done to relate the global properties of cellular array-generated patterns and the local transition functions which operate in the space. This is important, since present array “design” is heuristic, relying on the designer’s skill and luck. For example, to model periodic changes occurring in $Z$ reactions (Winfree [18]), one makes a calculated guess at the differential equation (and thus the local transition function). If the array patterns resemble the natural phenomena, it is reasonable to claim that a model has been found. However, what one wants ultimately is an algorithm to generate all local transition functions (and thus all models) which give the desired global behavior. The growth rules described in this paper are simple enough to provide a start on the problem, yet complex enough to produce interesting behavior.
The other motivation is biological. Although the rules are not complex enough to model natural growth, an interesting analogy exists with plant regeneration. Suppose we want to reproduce the pattern of growth resulting from some initial configuration by removing a part of the growth pattern and placing it in a clear region of the space. It follows from Theorem 2 that the GOE diagonal of a quadrant will be replicated at periodic intervals during its growth. Thus, a GOE diagonal (or any of its successors) can be cut and grown in an empty region. The fringe pattern and growth from the fringe pattern (backgrowth) will be identical to that of the corresponding quadrant in the original pattern. To replicate the entire pattern, four approximately spaced diagonals, one from each quadrant, are needed. The "genetic information" in this case is contained entirely within the four GOE diagonals.

The primary tool to analyze the growth patterns has been the vector string descriptor, and it is natural to inquire about its applicability to other local transition functions. Functions which have embedded in them a growth rule which is associated with a vector string appear also to be well suited—in particular, where the local transition function has a transition $q_0 \rightarrow q_i$ if all cells in the neighborhood are in state $q_0$, except one which is in state $q_i$. Rules with termination (pp. 237–243 of Schrandt and Ulam [16]), where all 1-cells older than a specific generation go to 0, fall into this category.

An interesting open question has to do with growth in hexagonal or triangular spaces. The logical extension of the u rule to these spaces produces patterns whose density is near $\frac{3}{4}$, but no suitable vector string has been found. For these cases, the fundamental vector set is apparent. However, the restrictions which must apply appear to be very complicated.

Natural biological organisms do not grow arbitrarily large. Both genetic and environmental factors limit the adult size. Gaibski and Yamada's [6] growth rules used the generation of specific patterns of states as the "key" which precluded growth beyond a certain point. For the growth patterns studied here, the cell state can be used to prevent growth beyond certain generations. For example, consider a space in which each cell has five states (0, 1, 2, 3, and 4) and the transition function shown in Fig. 14.

The pattern generated from a single 1-cell grows to a configuration which remains fixed after three applications of the local function. As long

\[
\begin{array}{c}
0 \\
0 \left[ \begin{array}{c} 0 \end{array} \right] 0 \rightarrow 2 \\
1 \\
0 \left[ \begin{array}{c} 0 \end{array} \right] 0 \rightarrow 3 \\
2 \\
0 \left[ \begin{array}{c} 0 \end{array} \right] 0 \rightarrow 3 \\
3 \\
\end{array}
\]

Fig. 14. Growth rule which produces a pattern which grows to a fixed size.
as there is a sufficient number of states, rules can by synthesized so that the pattern grows to any desired size.

This work was supported by National Science Foundation Grant MCS-7600326.

REFERENCES