To appear in *Connection Science*
Vol. 00, No. 00, Month 20XX, 1–16

Analysis of Information Gain and Kolmogorov Complexity for Structural Evaluation of Cellular Automata Configurations

Mohammad Ali Javaheri Javid[a*] and Tim Blackwell[b] and Robert Zimmer[c] and Mohammad Majid al-Rifaie[d]

Department of Computing, Goldsmiths, University of London, UK [a,b,c,d]

(*15 August 2015*)

Shannon entropy fails to discriminate structurally different patterns in two-dimensional images. We have adapted information gain measure and Kolmogorov complexity to overcome the shortcomings of entropy as a measure of image structure. The measures are customised to robustly quantify the complexity of images resulting from multi-state cellular automata. Experiments with a two-dimensional multi-state cellular automaton demonstrate that these measures are able to predict some of the structural characteristics, symmetry and orientation of cellular automata generated patterns.

Keywords: complexity, entropy, information gain, Kolmogorov complexity, computational aesthetics, cellular automata

1. Introduction

Cellular Automata (CA) were initially developed as a material independent framework to study the logic of self-reproducing behaviour of biological systems in the late 1940s by von Neumann and Ulam (Burks, 1970). In the 1960s the idea of using CA as artistic tool emerged from the works of Knowlton and Schwartz who produced "*Pixillation*", one of the early computer generated animations (Knowlton, 1964; Schwartz & Schwartz, 1992). The "*Tapestry I*" and "*Tapestry II*", still frames from "*Pixillation*" won the first prize of the *Eighth Annual Computer Art Contest* in 1970. The computer arts of Struycken (Struycken, 1976), Brown (Brown, 2001; Beddard & Dodds, 2009) and evolutionary architecture of Frazer (Frazer, 1995) are classical examples of CA based arts. Moreover, CA have been used for music composition, for example, Xenakis (Xenakis, 1992) and Miranda (Miranda, 2001).

The popularity of John Conway's Game of Life (Gardner, 1970) drew the attention of a wider community of researchers and digital artists to the unexplored potential of CA applications, especially in their capacity to generate complex behaviour from simple rules (Brown, July 1996). This fact has been noted by Wolfram, who himself produced some CA arts in the 1980s, "even a program that may have extremely simple rules will often be able to generate pictures that have striking aesthetic qualities-sometimes reminiscent of nature, but often unlike anything ever seen before" (Wolfram, 2002, p.11). He further emphasises that "one of the things I've been meaning to do is to make a bit more of a serious effort to use cellular automata in some kind of computer art" (Regis,

*Corresponding author. Email: m.javaheri@gold.ac.uk

1988, p.250).

Although classical one-dimensional CA with binary states can exhibit complex behaviours, experiments with multi-state two-dimensional (2D) CA reveal a very rich spectrum of symmetric and asymmetric patterns which are extremely challenging to generate using conventional mathematical methods (Javaheri Javid & te Boekhorst, 2006; Javaheri Javid, al Rifaie, & Zimmer, 2014).

There have been number of studies on the quantitative (Langton, 1986) and qualitative behaviour (Wolfram, 1983, 1984, 2002) of CA but they are mostly concerned with categorising the rule space and the computational properties of CA. Since CA are one of the generative tools in computer art, a means of evaluating the structure of CA generated patterns would make a substantial contribution towards further automation of CA art. There have been some interesting attempts to develop means of controlling emergence of aesthetic behaviour in CA (W. Li, 1988, 1989; Sims, 1992; Mason, 1993; Ashlock & Tsang, 2009) but with less success. This is due to lack of computational methods of human aesthetic perception.

This work follows Birkhoff's tradition in studying mathematical bases of aesthetics, especially the association of aesthetic judgement with the degree of order and complexity of a stimulus. Shannon's information theory provided an objective measure of complexity. It led to emergence of various informational theories of aesthetics. However entropy fails to take into account the spatial characteristics of 2D patterns; these characteristics are fundamental in addressing the aesthetic problem in general and of CA generated configurations in particular.

In this paper, following our earlier works (Javaheri Javid et al., 2014; Javaheri Javid, al Rifaie, & Zimmer, 2015; Javaheri Javid, Blackwell, Zimmer, & al Rifaie, 2015), we examine *information gain* and *Kolmogorov complexity* as measures of complexity in multi-state 2D CA generated configurations.

This paper is organised as follows. Section 2 provides formal definitions and establishes notations of CA. Section 3 demonstrates that entropy is an inadequate measure of discriminating multi-state 2D CA configurations. In Section 4 the potential of information gain as a structural complexity measure is discussed. Section 5 provides formal notions of Kolmogorov complexity and a method of estimating it. Section 6 gives details of experiments that test the effectiveness of information gain and its relation to Kolmogorov complexity. The paper closes with a discussion and summary of findings.

2. Definition of Cellular Automata

Definition 2.1 A cellular automaton is a regular tiling of a lattice with uniform deterministic finite state automata.

A cellular automaton $\mathcal{A}$ is specified by a quadruple $\langle L, S, N, f \rangle$ where:

- L is a finite square lattice of cells (i, j).
- $S = \{1, 2, \ldots, k\}$ is set of states. Each cell (i, j) in L has a state $s \in S$.
- N is neighbourhood, as specified by a set of lattice vectors $\{e_a\}$, $a = 1, 2, \ldots, N$. The neighbourhood of cell $r = (i, j)$ is $\{r + e_1, r + e_2, \ldots, r + e_N\}$. A a cell is considered to be in its own neighbourhood so that one of $\{e_a\}$ is the zero vector $(0, 0)$. With an economy of notation, the cells in the neighbourhood of (i, j) can be numbered from 1 to N; the neighbourhood states of (i, j) can therefore be denoted $(s_1, s_2, \ldots, s_N)$. Periodic boundary conditions are applied at the edges of the lattice so that complete neighbourhoods exist for every cell in L.

- f is the update rule. f computes the state $s_1(t+1)$ of a given cell from the states $(s_1, s_2, \ldots, s_N)$ of cells in its neighbourhood:$s_1(t+1) = f(s_1, s_2, \ldots, s_N)$. A quiescent state s_q satisfies $f(s_q, s_q, \ldots, s_q) = s_q$.

Remark 1 There are two common neighbourhoods; a five-cell von Neumann neighbourhood $\{(0,0), (\pm 1, 0), (0, \pm 1)\}$ and a nine-cell Moore neighbourhood $\{(0,0), (\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}$.

The collection of states for all cells in L is known as a *configuration* (C). The global rule F maps the whole automaton forward in time; it is the synchronous application of f to each cell. The behaviour of a particular $\mathcal{A}$ is the sequence $c^0, c^1, c^2, \ldots, c^{T-1}$, where c^0 is the initial configuration (IC) at $t = 0$.

CA behaviour is sensitive to the IC and to L, S, N and f. The behaviour is generally non-linear and sometimes very complex; no single mathematical analysis can describe, or even estimate, the behaviour of an arbitrary automaton. The vast size of the rule space, and the fact that this rule space is unstructured, mean that knowledge of the behaviour a particular cellular automaton, or even of a set of automata, gives no insight into the behaviour of any other CA. In the lack of any practical model to predict the behaviour of a cellular automaton, the only feasible method is to run a simulation.

3. Inadequacy of Entropy as a Complexity Measure for Images

Despite the dominance of entropy as a measure of order and complexity, it fails to reflect on structural characteristics of 2D patterns. The main reason for this drawback is that it measures the distribution of symbols, and not their arrangements. This is in contrast to our intuitive perception of the complexity of patterns and is problematic for the purpose of measuring the complexity of 2D CA generated images in particular.

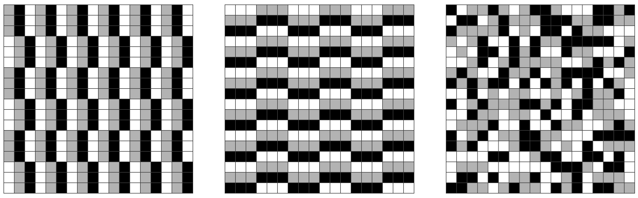
Information theory was developed in order to address reliable communication over an unreliable channel (Shannon, 1948). Entropy is the core of this theory (Cover & Thomas, 2006). Let $\mathcal{X}$ be discrete alphabet, X a discrete random variable, $x \in \mathcal{X}$ a particular value of X and $P(x)$ the probability of x. Then the entropy, $H(X)$, is:

$$H(X) = -\sum_{x \in \mathcal{X}} P(x) \log_2 P(x) \tag{1}$$

The quantity H is the average uncertainty in bits, $\log_2(\frac{1}{p})$ associated with X. Entropy can also be interpreted as the average amount of information needed to describe X. The value of entropy is always non-negative and reaches its maximum for the uniform distribution, $\log_2(|\mathcal{X}|)$:

$$0 \leqslant H \leqslant \log_2(|\mathcal{X}|). \tag{2}$$

The lower bound of relation (2) corresponds to a deterministic variable (no uncertainty) and the upper bound corresponds to a maximum uncertainty associated with a random variable. Entropy can be regarded as a measure of order and complexity. A low entropy implies low uncertainty and the message is highly predictable, ordered and less complex. High entropy implies a high uncertainty, less predictability, highly disordered and complex. Fig. 1 illustrates the measure of entropy of 2D patterns with various structural characteristics with the uniform distribution of elements. Figs. 1a-b are patterns with ordered structures and Fig. 1c is a pattern with a fairly structureless random pattern.



(a) $H = 1.58496$ (b) $H = 1.58496$ (c) $H = 1.58496$

Figure 1. Measure of H for structurally different patterns with uniform distribution of elements.

These patterns have identical entropy yet clearly have structural differences. This clearly demonstrates the failure of entropy as a structural measure. This is in contrast to our intuitive perception of complexity of patterns. For the purpose of measuring complexity of multi-state CA behaviour, it would be problematic if only entropy were to be applied.

4. Information Gain

Information gain (Bates & Shepard, 1993; Wackerbauer, Witt, Atmanspacher, Kurths, & Scheingraber, 1994; Andrienko, Yu. A., Brilliantov, N. V., & Kurths, J., 2000) has been proposed as a means of characterising the complexity of dynamical systems and of 2D patterns. It measures the amount of information gained in bits when specifying the value, x, of a random variable X given knowledge of the value, y, of another random variable Y,

$$G_{x,y} = -\log_2 P(x|y). \tag{3}$$

$P(x|y)$ is the conditional probability of a state x conditioned on the state y. Then the *mean information gain* (MIG), $\overline{G}_{X,Y}$, is the average amount of information gain from the description of the all possible states of Y:

$$\overline{G}_{X,Y} = \sum_{x,y} P(x,y) G_{x,y} = -\sum_{x,y} P(x,y) \log_2 P(x|y) \tag{4}$$

where $P(x,y)$ is the joint probability, $\text{prob}(X = x, Y = y)$. $\overline{G}$ is also known as the conditional entropy, $H(X|Y)$ (Cover & Thomas, 2006). Conditional entropy is the reduction in uncertainty of the joint distribution of X and Y given knowledge of Y, $H(X|Y) = H(X,Y) - H(Y)$. The lower and upper bounds of $\overline{G}_{X,Y}$ are

$$0 \leqslant \overline{G}_{X,Y} \leqslant \log_2 |\mathcal{X}|. \tag{5}$$

Definition 4.1 A structural complexity measure G, of a cellular automaton configuration is the sum of the mean information gains of cells having homogeneous/heterogeneous neighbouring cells over 2D lattice.

For a cellular automaton configuration, $\overline{G}$ can be calculated by considering the distribution of cell states over pairs of cells r, s,

$$\overline{G}_{r,s} = -\sum_{s_r,s_s} P(s_r, s_s) \log_2 P(s_r, s_s) \tag{6}$$

where s_r, s_s are the states at r and s. Since $|\mathcal{S}| = N$, $\overline{G}_{r,s}$ is a value in $[0, N]$.

The vertical, horizontal, primary diagonal ($\diagdown$) and secondary diagonal ($\diagup$) neighbouring pairs provide eight $\overline{G}s$; $\overline{G}_{(i,j),(i-1,j+1)} \equiv \overline{G}_{\nwarrow}$, $\overline{G}_{(i,j),(i,j+1)} \equiv \overline{G}_{\uparrow}$, $\overline{G}_{(i,j),(i+1,j+1)} \equiv \overline{G}_{\nearrow}$, $\overline{G}_{(i,j),(i-1,j)} \equiv \overline{G}_{\leftarrow}$, $\overline{G}_{(i,j),(i+1,j)} \equiv \overline{G}_{\rightarrow}$, $\overline{G}_{(i,j),(i-1,j-1)} \equiv \overline{G}_{\swarrow}$, $\overline{G}_{(i,j),(i,j-1)} \equiv \overline{G}_{\downarrow}$ and $\overline{G}_{(i,j),(i+1,j-1)} \equiv \overline{G}_{\searrow}$. The relative positions for non-edge cells are given by matrix M:

$$M = \begin{bmatrix} (i-1,j+1) & (i,j+1) & (i+1,j+1) \\ (i-1,j) & (i,j) & (i+1,j) \\ (i-1,j-1) & (i,j-1) & (i+1,j-1) \end{bmatrix}. \tag{7}$$

Correlations between cells on opposing lattice edges are not considered. The result of this edge condition is that $G_{\rightarrow}$ is not necessarily equal to $\overline{G}_{\leftarrow}$. In addition the differences between the horizontal (vertical) and two diagonal mean information rates reveal left/right (up/down), primary and secondary orientation of 2D patterns. So the sequence of generated configurations by a multi-state 2D cellular automaton can analysed by the differences between the vertical (V), horizontal (H), primary diagonal (P_d) and secondary diagonal (S_d) mean information gains by

$$\Delta\overline{G}_V = |\overline{G}_{\uparrow} - \overline{G}_{\downarrow}| \tag{8a}$$

$$\Delta\overline{G}_H = |\overline{G}_{\leftarrow} - \overline{G}_{\rightarrow}| \tag{8b}$$

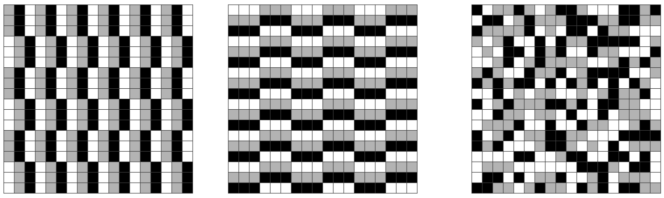
$$\Delta\overline{G}_{P_d} = |\overline{G}_{\nwarrow} - \overline{G}_{\searrow}| \tag{8c}$$

$$\Delta\overline{G}_{S_d} = |\overline{G}_{\nearrow} - \overline{G}_{\swarrow}| \tag{8d}$$

Fig. 2 demonstrates the merits of $\overline{G}$ in discriminating structurally different patterns for the sample patterns in Fig. 1. As it is evident, the measures of H are identical for structurally different patterns, however $\overline{G}$ and $\Delta\overline{G}$ differentiate spatial arrangement.

5. Kolmogorov Complexity of 2D Patterns

From the perspective of information theory, the object X is a random variable drawn according to a probability mass function $P(x)$. If X is random, then the descriptive complexity of the event $X = x$ is $\log \frac{1}{P(x)}$, because $\lceil \log \frac{1}{P(x)} \rceil$ is the number of bits required to describe x. Thus the descriptive complexity of an object depends on the probability distribution (Cover & Thomas, 2006). Kolmogorov defined the algorithmic (descriptive) complexity of an object to the minimum length of a program such that a universal computer can generate a specific sequence (Kolmogorov, 1965). Thus, the Kolmogorov



(a)	(b)	(c)
$H = 1.58496$	$H = 1.58496$	$H = 1.58496$
$\overline{G}_{\uparrow} = 1.15955$	$\overline{G}_{\uparrow} = 0$	$\overline{G}_{\uparrow} = 1.58181$
$\overline{G}_{\downarrow} = 1.15955$	$\overline{G}_{\downarrow} = 0$	$\overline{G}_{\downarrow} = 1.58209$
$\Delta\overline{G}_V = 0$	$\Delta\overline{G}_V = 0$	$\Delta\overline{G}_V = 0.00028$
$\overline{G}_{\leftarrow} = 0$	$\overline{G}_{\leftarrow} = 1.15955$	$\overline{G}_{\leftarrow} = 1.57696$
$\overline{G}_{\rightarrow} = 0$	$\overline{G}_{\rightarrow} = 1.15955$	$\overline{G}_{\rightarrow} = 1.57668$
$\Delta\overline{G}_H = 0$	$\Delta\overline{G}_H = 0$	$\Delta\overline{G}_H = 0.00028$
$\overline{G}_{\nwarrow} = 1.15843$	$\overline{G}_{\nwarrow} = 1.15843$	$\overline{G}_{\nwarrow} = 1.56727$
$\overline{G}_{\searrow} = 1.15843$	$\overline{G}_{\searrow} = 1.15843$	$\overline{G}_{\searrow} = 1.56712$
$\Delta\overline{G}_{P_d} = 0$	$\Delta\overline{G}_{P_d} = 0$	$\Delta\overline{G}_{P_d} = 0.000150$
$\overline{G}_{\nearrow} = 1.15843$	$\overline{G}_{\nearrow} = 1.15843$	$\overline{G}_{\nearrow} = 1.57688$
$\overline{G}_{\swarrow} = 1.15843$	$\overline{G}_{\swarrow} = 1.15843$	$\overline{G}_{\swarrow} = 1.57657$
$\Delta\overline{G}_{S_d} = 0$	$\Delta\overline{G}_{S_d} = 0$	$\Delta\overline{G}_{S_d} = 0.00031$

Figure 2. Measure of H, $\overline{G}s$ and $\Delta\overline{G}s$ for structurally different patterns with uniform distribution of elements.

complexity of an object is independent of the probability distribution. Kolmogorov complexity is related to entropy $H(X)$ in that the expected value of $K(x)$ for a random sequence is approximately the entropy of the source distribution for the process generating the sequence. However, Kolmogorov complexity differs from entropy in that it relates to the specific string being considered rather than the source distribution (M. Li, 1997; Cover & Thomas, 2006). Kolmogorov complexity can be described as follows, where φ represents a universal computer, p represents a program, and x represents a string and $l(p)$ represents program length,

$$K_{\varphi}(x) = \left\{ \min_{\varphi(p)=x} \quad l(p) \right\} \tag{9}$$

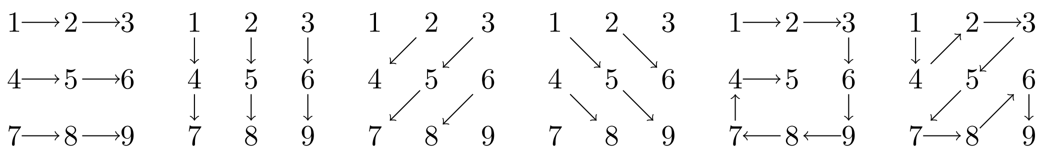
Random strings have rather high Kolmogorov complexity, of the order of their length, as patterns cannot be discerned to reduce the size of a program generating such a string. On the other hand, strings with a large amount of structure have fairly low complexity. Universal computers can be equated through programs of constant length, thus a mapping can be made between universal computers of different types. The Kolmogorov complexity of a given string on two computers differs by known or determinable constants. The Kolmogorov complexity $K(y|x)$ of a string y, given string x as input is

$$K_{\varphi}(y|x) = \left\{ \begin{array}{l} \min\limits_{\varphi(p,y)=y} \quad l(p) \\ \\ \infty, \text{if there is no } p \text{ such that } \varphi(p,x) = y \end{array} \right\} \tag{10}$$

where φ is a particular universal computer under consideration. Thus, knowledge or input of a string x may reduce the complexity or program size necessary to produce a new string y. The major difficulty with Kolmogorov complexity is its uncomputability; however, any program that produces a given string provides an upper bound.

Lempel and Ziv defined a measure of complexity for finite sequences rooted in the ability to produce strings from simple copy operations (Ziv & Lempel, 1978). This method known as $LZ78$ universal compression harnesses this principle to yield a universal compression algorithm that can approach the entropy of an infinite sequence produced by an ergodic source. Hence $LZ78$ compression can be used as an estimator for K.

In order to estimate the K value of 2D configurations generated by multi-state CA, we generate linear strings of configurations using six different templates, illustrated in Fig. 3.

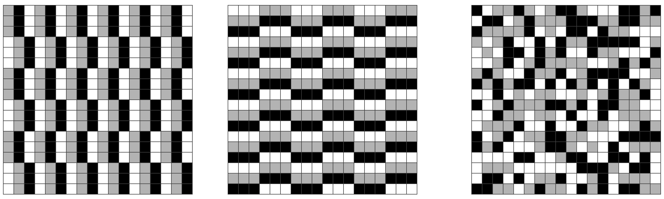


(a) (b) (c) (d) (e) (f)

Figure 3. The six different templates used to linearise 2D configurations.

Fig. 3a string horizontal $S_h = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, Fig. 3b vertical $S_v = \{1, 4, 7, 2, 5, 8, 3, 6, 9\}$, Fig. 3c diagonal $S_d = \{1, 2, 4, 3, 5, 7, 6, 8, 9\}$, Fig. 3d reverse diagonal $S_{rd} = \{3, 2, 6, 1, 5, 9, 4, 8, 7\}$, Fig. 3e spiral $S_s = \{1, 2, 3, 6, 9, 8, 7, 4, 5\}$, and Fig. 3f continuous spiral $S_{cs} = \{1, 4, 2, 3, 5, 7, 8, 6, 9\}$. Then an upper bound of K is estimated from the smallest $LZ78$ compression over these templates.

A comparison of the measurement of H, $\overline{G}s$, $\Delta\overline{G}s$ and K for structurally different patterns with uniform distribution of elements is illustrated in Fig. 4. As is evident from the measurements, K is able to discriminate the complexity of the patterns, however it fails to discriminate the spatial orientations.



(a)	(b)	(c)
$H = 1.58496$	$H = 1.58496$	$H = 1.58496$
$\overline{G}_{\uparrow} = 1.15955$	$\overline{G}_{\uparrow} = 0$	$\overline{G}_{\uparrow} = 1.58181$
$\overline{G}_{\downarrow} = 1.15955$	$\overline{G}_{\downarrow} = 0$	$\overline{G}_{\downarrow} = 1.58209$
$\Delta\overline{G}_V = 0$	$\Delta\overline{G}_V = 0$	$\Delta\overline{G}_V = 0.00028$
$\overline{G}_{\leftarrow} = 0$	$\overline{G}_{\leftarrow} = 1.15955$	$\overline{G}_{\leftarrow} = 1.57696$
$\overline{G}_{\rightarrow} = 0$	$\overline{G}_{\rightarrow} = 1.15955$	$\overline{G}_{\rightarrow} = 1.57668$
$\Delta\overline{G}_H = 0$	$\Delta\overline{G}_H = 0$	$\Delta\overline{G}_H = 0.00028$
$\overline{G}_{\nwarrow} = 1.15843$	$\overline{G}_{\nwarrow} = 1.15843$	$\overline{G}_{\nwarrow} = 1.56727$
$\overline{G}_{\searrow} = 1.15843$	$\overline{G}_{\searrow} = 1.15843$	$\overline{G}_{\searrow} = 1.56712$
$\Delta\overline{G}_{P_d} = 0$	$\Delta\overline{G}_{P_d} = 0$	$\Delta\overline{G}_{P_d} = 0.000150$
$\overline{G}_{\nearrow} = 1.15843$	$\overline{G}_{\nearrow} = 1.15843$	$\overline{G}_{\nearrow} = 1.57688$
$\overline{G}_{\swarrow} = 1.15843$	$\overline{G}_{\swarrow} = 1.15843$	$\overline{G}_{\swarrow} = 1.57657$
$\Delta\overline{G}_{S_d} = 0$	$\Delta\overline{G}_{S_d} = 0$	$\Delta\overline{G}_{S_d} = 0.00031$
$Kh = 0.13889$	$Kv = 0.13889$	$K_{d=cs} = 0.27778$

Figure 4. Measures of H, $\overline{G}s$, $\Delta\overline{G}s$ and K for structurally different patterns with uniform distribution of elements.

6. Experiments and Results

A set of experiments was designed to examine the effectiveness of $\overline{G}$ in discriminating the particular configurations that are generated by a multi-state 2D cellular automaton. The experimental rule (Table 1) maps four states, represented by *white*, *red*, *blue* and *orange*; the quiescent state is *white*.

Table 1. Update rule of experimental cellular automaton.

$L = 65 \times 65$ (4225 cells).

$S = \{0, 1, 2, 3\} \equiv \{\square, \blacksquare, \blacksquare, \blacksquare\}$

$\mathcal{N}$: Moore neighbourhood

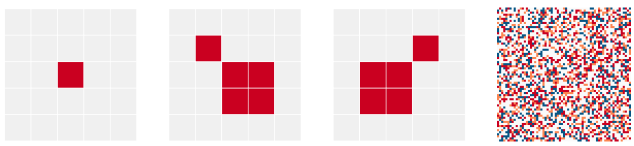
$f : S^9 \mapsto S$

$$f(s_{i,j})(t) = s_{i,j}(t+1) = \left\{ \begin{array}{lll} 1 & \text{if } s_{(i,j)}(t) = 0 & \text{and } \sigma = 1 \\ 3 & \text{if } s_{(i,j)}(t) = \text{1-3} & \text{and } \sigma = 2 \\ 2 & \text{if } s_{(i,j)}(t) = \text{1-3} & \text{and } \sigma = 3 \\ 0 & \text{otherwise} & \end{array} \right\}$$

where σ is the sum total of the neighbourhood states.

The experiments were conducted with four different ICs: (1) all quiescent states cells except for a single cell (Fig. 5a), (2) a right oriented 5 cell (Fig. 5b), (3) a left oriented 5 cell (Fig. 5c) and (4) a random configuration with 2112 *white* quiescent states cells covering $\approx 50\%$ of the lattice, 749 *red*, 682 *blue* and 682 *orange* cells (Fig. 5d).

The experimental rule has been iterated synchronously for 150 successive time steps. Figs. 6, 8, 10 and 12 illustrate a sample of time steps starting from four different ICs. Then the sequence of configurations are analysed by Eqs. 8a, 8b, 8c, 8d and K.



(a) (b) (c) (d)

Figure 5. The four different ICs used to seed experimental cellular automaton.

The behaviour of cellular automaton from the single cell IC is a sequence of symmetrical patterns (Fig. 6). This fact has been reflected on the measurements of $\Delta\overline{G}s$ (Fig. 7), where they are constant for the 150 time steps ($\Delta\overline{G}_V = \Delta\overline{G}_H = \Delta\overline{G}_{P_d} = \Delta\overline{G}_{S_d} = 0$). This is an indicator of the development of complete symmetrical patterns in four directions for each of 150 configurations generated by experimental cellular automaton. However, the measurement of entropy starts from $H_0 = 0.00319$ and reaches $H_{150} = 1.47979$ at the end of the runs (Fig. 14).

The two 5 cell ICs (5b and 5c) generate sequence of symmetrical patterns with different orientations (Fig. 8 and Fig. 10). The measurements of H for these two sequences of structurally different but symmetrical configurations are identical from $t = 0$ to $t = 150$, where $H_0^{5b} = H_0^{5c} = 0.01321$ and $H_{150}^{5b} = H_{150}^{5c} = 1.43241$ (Fig. 14). On the other hand the measurements of $\Delta\overline{G}s$ especially $\Delta\overline{G}_{P_d}$ and $\Delta\overline{G}_{S_d}$ are reflecting the differences in the orientations of symmetrical configurations (Fig. 9 and Fig. 11). This is further illustrated in Fig. 15 where the measures of H, $\overline{G}s$ and $\Delta\overline{G}s$ are compared for two configurations generated at $t = 40$ from two different 5b and 5c ICs.

The development of configurations from the random IC is a sequence of irregular

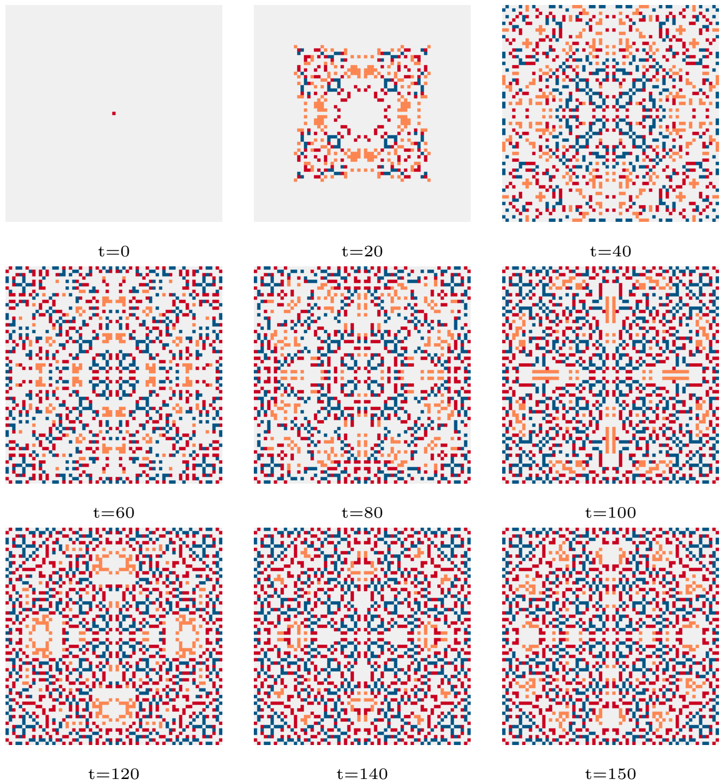


Figure 6. Space-time diagram of the experimental cellular automaton for sample time steps starting from the single cell IC (5a).

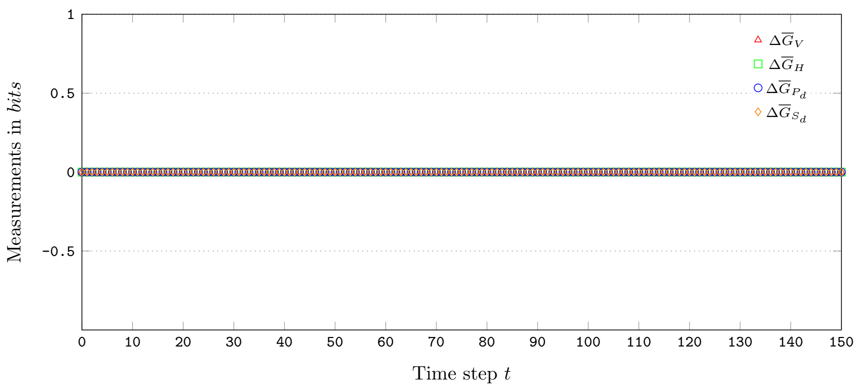


Figure 7. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 5a ICs.

structures (Fig. 12). The formation of patterns with local structures has reduced the values of $\Delta\overline{G}s$ until a stable oscillating pattern is attained (Fig. 13). This is an indicator of the development of irregular structures. However the patterns are not random since the maximum four-state value of $\log_2(4) = 2$ (Eq. 5).

These experiments demonstrate that a cellular automaton rule seeded with different ICs leads to the formation of patterns with structurally diverse characteristics. The gradient of the mean information rate along lattice axes is able to detect the structural characteristics of patterns generated by this particular multi-state 2D cellular automaton. From the comparison of H with $\Delta\overline{G}s$ in the set of experiments, it is clear that entropy fails to discriminate between the diversity of patterns that can be generated by various

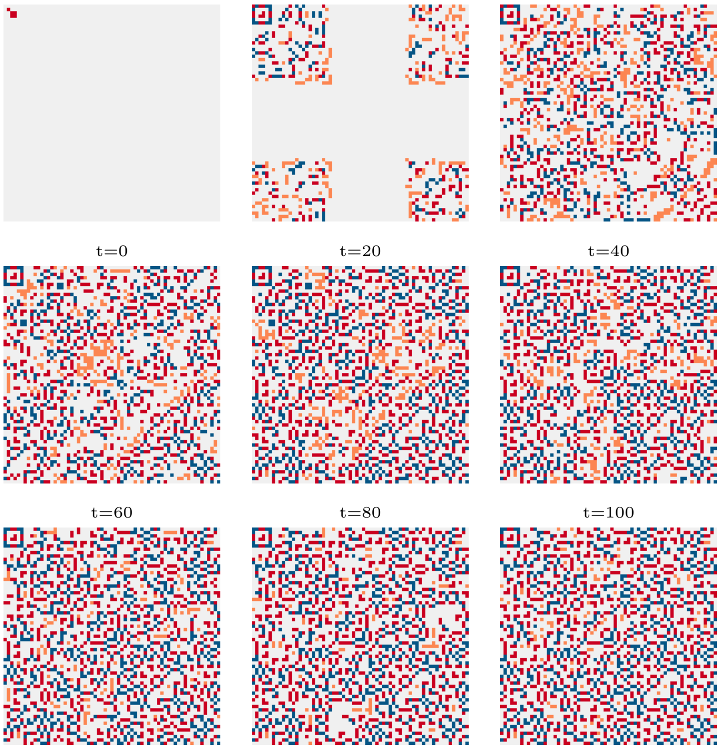


Figure 8. Space-time diagram of the experimental cellular automaton for sample time steps starting from the 5b IC.

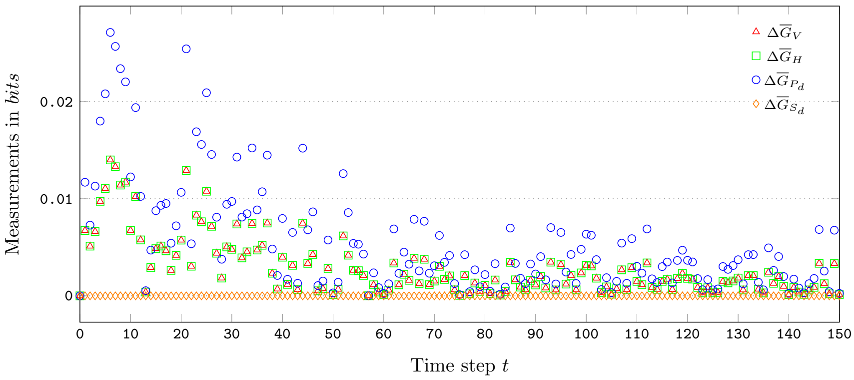


Figure 9. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 5b IC.

CA. The structured but asymmetrical patterns emerging from the random IC are clearly distinguished from the symmetrical patterns including their orientation. As it is evident from the results of experiments, the measures of H are identical for structurally different patterns, however, the measure of $\overline{G}s$ and $\Delta\overline{G}s$ are reflecting not only the complexity of patterns but their spatial arrangements (i.e. orientation of symmetries) as well.

In addition, the relationship between K and $\overline{G}s$ are examined by Eq. 11 (Pearson correlation coefficient). Table 2 illustrates the calculations of r for different directional $\overline{G}s$. Since the values of r are ≈ 0.99, so there are strong positive correlation between K and $\overline{G}s$.

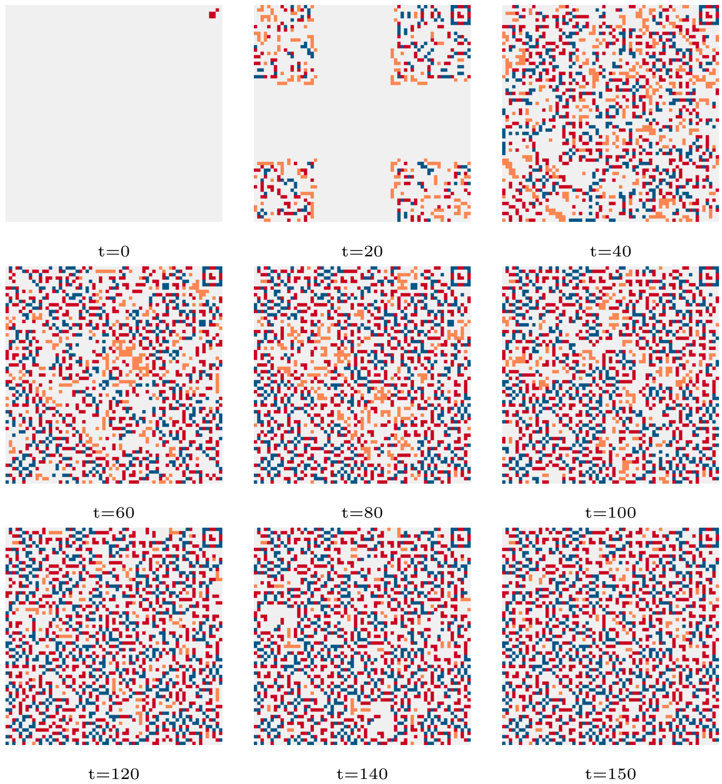


Figure 10. Space-time diagram of the experimental cellular automaton for sample time steps starting from the 5c IC.

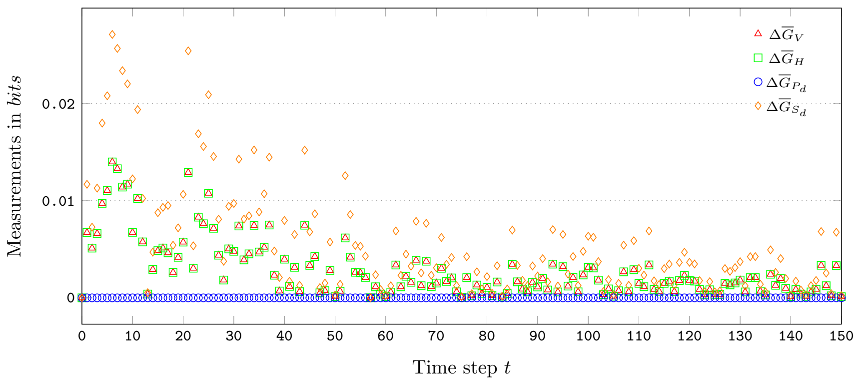


Figure 11. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 5c IC.

$$r = r_{KG} = \frac{\sum_{i=1}^{150}(K_i - \bar{K})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^{150}(K_i - \bar{K})^2}\sqrt{\sum_{i=1}^{150}(G_i - \bar{G})^2} \tag{11}$$

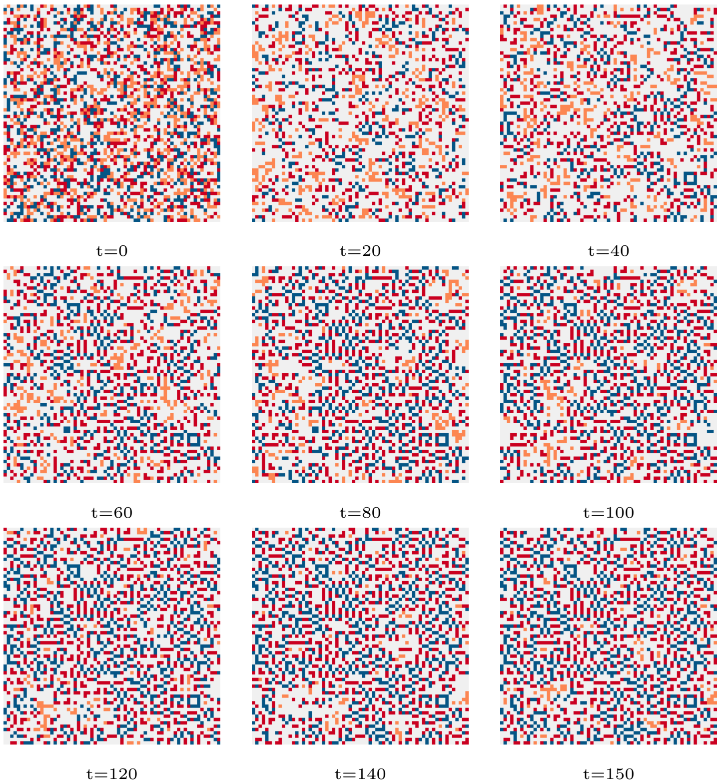


Figure 12. Space-time diagram of the experimental cellular automaton for sample time steps starting from the random (5d) IC.

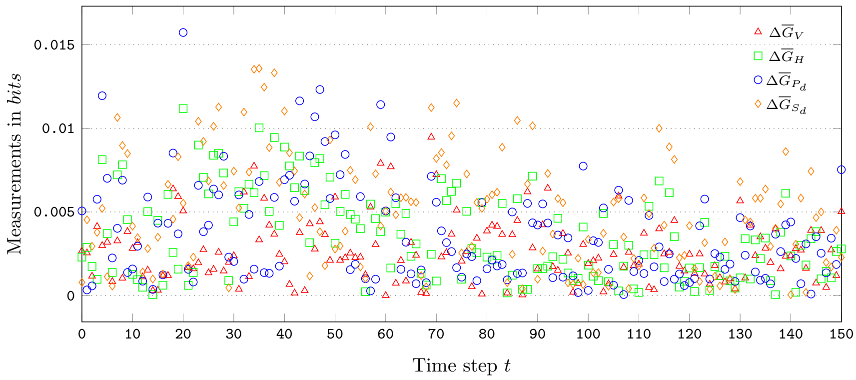


Figure 13. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from the random (5d) IC.

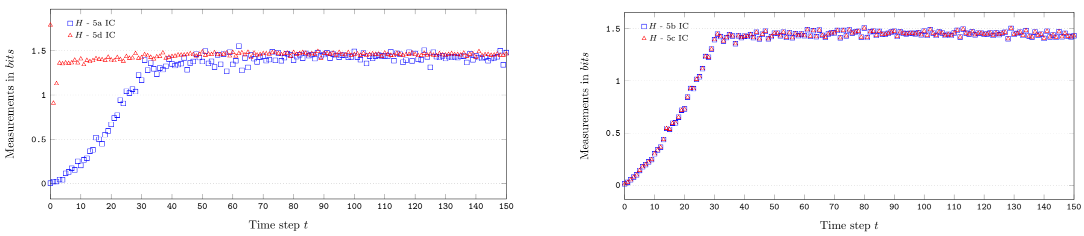
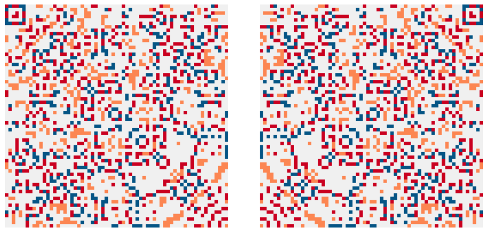


Figure 14. The measurements of H for 150 time steps starting from 5a, 5b, 5c and 5d ICs.



(a)	(b)
$H = 1.42929$	$H = 1.42929$
$\overline{G}_{\uparrow} = 1.36140$	$\overline{G}_{\uparrow} = 1.36140$
$\overline{G}_{\downarrow} = 1.36538$	$\overline{G}_{\downarrow} = 1.36538$
$\Delta\overline{G}_V = 0.00398$	$\Delta\overline{G}_V = 0.00398$
$\overline{G}_{\leftarrow} = 1.36140$	$\overline{G}_{\leftarrow} = 1.36538$
$\overline{G}_{\rightarrow} = 1.36538$	$\overline{G}_{\rightarrow} = 1.36140$
$\Delta\overline{G}_H = 0.00398$	$\Delta\overline{G}_H = 0.00398$
$\overline{G}_{\nwarrow} = 1.36634$	$\overline{G}_{\nwarrow} = 1.37148$
$\overline{G}_{\searrow} = 1.37431$	$\overline{G}_{\searrow} = 1.37148$
$\Delta\overline{G}_{P_d} = 0.00797$	$\Delta\overline{G}_{P_d} = 0$
$\overline{G}_{\swarrow} = 1.37148$	$\overline{G}_{\swarrow} = 1.37431$
$\overline{G}_{\nearrow} = 1.37148$	$\overline{G}_{\nearrow} = 1.36634$
$\Delta\overline{G}_{S_d} = 0$	$\Delta\overline{G}_{S_d} = 0.00797$
$K_s = 0.15929$	$K_v = 0.16118$

Figure 15. The comparison of H, $\overline{G}s$ and $\Delta\overline{G}s$ at $t = 40$ for 5b (a) and 5c (b) ICs.

Table 2. Calculations of r for different ICs.

$r_{K\overline{G}_{\uparrow}} = 0.9985$	$r_{K\overline{G}_{\downarrow}} = 0.9985$	$r_{K\overline{G}_{\leftarrow}} = 0.9985$	$r_{K\overline{G}_{\rightarrow}} = 0.9985$
$r_{K\overline{G}_{\nwarrow}} = 0.9975$	$r_{K\overline{G}_{\searrow}} = 0.9975$	$r_{K\overline{G}_{\nearrow}} = 0.9975$	$r_{K\overline{G}_{\swarrow}} = 0.9975$

Calculations of r for 5a IC.

$r_{K\overline{G}_{\uparrow}} = 0.9996$	$r_{K\overline{G}_{\downarrow}} = 0.9995$	$r_{K\overline{G}_{\leftarrow}} = 0.9996$	$r_{K\overline{G}_{\rightarrow}} = 0.9995$
$r_{K\overline{G}_{\nwarrow}} = 0.9996$	$r_{K\overline{G}_{\searrow}} = 0.9996$	$r_{K\overline{G}_{\nearrow}} = 0.9994$	$r_{K\overline{G}_{\swarrow}} = 0.9995$

Calculations of r for 5b IC.

$r_{K\overline{G}_{\uparrow}} = 0.9996$	$r_{K\overline{G}_{\downarrow}} = 0.9995$	$r_{K\overline{G}_{\leftarrow}} = 0.9995$	$r_{K\overline{G}_{\rightarrow}} = 0.9996$
$r_{K\overline{G}_{\nwarrow}} = 0.9995$	$r_{K\overline{G}_{\searrow}} = 0.9996$	$r_{K\overline{G}_{\nearrow}} = 0.9996$	$r_{K\overline{G}_{\swarrow}} = 0.9996$

Calculations of r for 5c IC.

$r_{K\overline{G}_{\uparrow}} = 0.9854$	$r_{K\overline{G}_{\downarrow}} = 0.9842$	$r_{K\overline{G}_{\leftarrow}} = 0.9874$	$r_{K\overline{G}_{\rightarrow}} = 0.9838$
$r_{K\overline{G}_{\nwarrow}} = 0.9885$	$r_{K\overline{G}_{\searrow}} = 0.9879$	$r_{K\overline{G}_{\nearrow}} = 0.9794$	$r_{K\overline{G}_{\swarrow}} = 0.9831$

Calculations of r for 5d IC.

7. Conclusion

CA are known for their generative capabilities and have contributed to the creation of many computer art works. Indeed, multi-state 2D CA can generate aesthetically appealing and complex patterns with various structural characteristics. Thus the means of evaluating the structure and ultimately the aesthetic qualities of CA generated patterns could have a substantial contribution towards further automation of CA art.

Entropy, one of the mostly applied measure of complexity, depends on the probability distribution of symbols, and not their arrangements. Despite the dominance of entropy as a measure of order and complexity, it fails to reflect on the structural characteristics of 2D patterns and of CA configurations.

However mean information gain takes into account conditional and joint probabilities between pairs of cells and, since it is based on correlations between cells, holds promise for patterns discrimination. Kolmogorov algorithmic complexity is another measure of complexity which can be used to estimate the complexity of 2D configurations generated by a cellular automaton.

This paper reports on a set of experiments with a cellular automaton rule seeded with four different initial conditions which lead to the formation of patterns with structurally diverse characteristics. The potential of mean information gain and Kolmogorov complexity for distinguishing multi-state 2D CA patterns is demonstrated. The measures appear to be particularly good at distinguishing different kinds of random patterns from non-random patterns. Furthermore, information gain measure is also able to discriminate the orientation of symmetries.

References

Andrienko, Yu. A., Brilliantov, N. V., & Kurths, J. (2000). Complexity of two-dimensional patterns. *Eur. Phys. J. B*, *15*(3), 539–546.

Ashlock, D., & Tsang, J. (2009). Evolved art via control of cellular automata. In *Evolutionary computation, 2009. cec'09. ieee congress on* (pp. 3338–3344).

Bates, J. E., & Shepard, H. K. (1993). Measuring complexity using information fluctuation. *Physics Letters A*, *172*(6), 416–425.

Beddard, H., & Dodds, D. (2009). *Digital Pioneers.* V&A Publishing.

Brown, P. (2001). Stepping stones in the mist. In *Creative evolutionary systems* (pp. 387–407).

Brown, P. (July 1996). Emergent behaviours: Towards computational aesthetics. *Artlink*, *16*(2 & 3).

Burks, A. W. (1970). *Essays on cellular automata.* University of Illinois Press.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory (wiley series in telecommunications and signal processing).* Wiley-Interscience.

Frazer, J. (1995). *An evolutionary architecture.* Architectural Association Publications, Themes VII.

Gardner, M. (1970, October). Mathematical games - the fantastic combinations of john conway's new solitaire game, life. *Scientific American*, 120–123.

Javaheri Javid, M. A., al Rifaie, M. M., & Zimmer, R. (2014). Detecting Symmetry in Cellular Automata Generated Patterns Using Swarm Intelligence. In A.-H. Dediu, M. Lozano, & C. Martín-Vide (Eds.), *Theory and Practice of Natural Computing* (Vol. 8890, pp. 83–94). Springer International Publishing.

Javaheri Javid, M. A., al Rifaie, M. M., & Zimmer, R. (2015). An Informational Model for Cellular Automata Aesthetic Measure. In *Aisb 2015 symposium on computational creativity.* University of Kent, Canterbury, UK.

Javaheri Javid, M. A., Blackwell, T., Zimmer, R., & al Rifaie, M. M. (2015). Spatial Complexity Measure for Characterising Cellular Automata Generated 2D Patterns. In Francisco Pereira, Penousal Machado, Ernesto Costa and Amílcar Cardoso (Ed.), *Progress in artificial intelligence, 17th portuguese conference on artificial intelligence, epia 2015, coimbra, portugal, september 8-11, 2015, proceedings.* (Vol. 9273). Springer International Publishing.

Javaheri Javid, M. A., & te Boekhorst, R. (2006). Cell Dormancy in Cellular Automata. In V. N. Alexandrov, G. D. van Albada, P. M. A. Sloot, & J. Dongarra (Eds.), *International Conference on Computational Science (3)* (Vol. 3993, pp. 367–374). Springer.

Knowlton, K. C. (1964). A computer technique for producing animated movies. In *Proceedings of the april 21-23, 1964, spring joint computer conference* (pp. 67–87).

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of information transmission*, *1*(1), 1–7.

Langton, C. G. (1986). Studying artificial life with cellular automata. *Physica D: Nonlinear Phenomena*, *22*(1), 120–149.

Li, M. (1997). *An introduction to kolmogorov complexity and its applications.* Springer.

Li, W. (1988). *Pretty pictures generated by two-state five-neighbor cellular automata* (Tech. Rep.). CCSR Technical Report No. 15.,University of Illinois.

Li, W. (1989). Complex patterns generated by next nearest neighbors cellular automata. *Computers & Graphics*, *13*(4), 531–537.

Mason, W. K. (1993). Art from cellular automata and symmetrized dot-patterns. *Computers & graphics*, *16*(4), 439–441.

Miranda, E. (2001). *Composing music with computers* (No. v. 1). Focal Press.

Regis, E. (1988). *Who got einstein's office?: eccentricity and genius at the institute for advanced study.* Basic Books.

Schwartz, L., & Schwartz, L. (1992). *The computer artist's handbook: Concepts, techniques, and applications.* W W Norton & Company Incorporated.

Shannon, C. (1948, October). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *27*, 379–423 & 623–656.

Sims, K. (1992). Interactive evolution of dynamical systems. In *Toward a practice of autonomous systems: Proceedings of the first european conference on artificial life* (pp. 171–178).

Struycken, P. (1976). Splash 1972/1974. In R. Leavitt (Ed.), *Artist and computer* (pp. 30–31). Harmony Books.

Wackerbauer, R., Witt, A., Atmanspacher, H., Kurths, J., & Scheingraber, H. (1994). A comparative classification of complexity measures. *Chaos, Solitons & Fractals*, *4*(1), 133–173.

Wolfram, S. (1983). Statistical mechanics of cellular automata. *Reviews of modern physics*, *55*(3), 601–644.

Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, *10*(1), 1–35.

Wolfram, S. (2002). *A new kind of science.* Wolfram Media Inc.

Xenakis, I. (1992). *Formalized music: thought and mathematics in composition.* Pendragon Press.

Ziv, J., & Lempel, A. (1978). Compression of individual sequences via variable-rate coding. *Information Theory, IEEE Transactions on*, *24*(5), 530–536.