# Classifying Cellular Automata Automatically

Andrew Wuensche

SFI WORKING PAPER: 1998-02-018

# Classifying Cellular Automata Automatically

Andy Wuensche
Santa Fe Institute, 1399 Hyde Park Road,
Santa Fe, New Mexico 87501 USA, wuensch@santafe.edu

**Abstract**

"Glider" dynamics in cellular automata (CA), where coherent configurations emerge and interact, provide a stark instance of self-organization in a simple system. Such behaviour was classified as class 4 or complex (as opposed to ordered or chaotic) by Wolfram[12], and was one of the original motivations for Artificial Life[7]. Because glider dynamics is relatively rare in CA rule spaces, much study has focused on the few known complex rules. However, a more general theory would benefit from many examples, which are now available for 1d CA found by the methods described.

Cellular automata rules can be classified automatically for a spectrum of ordered, complex and chaotic dynamics, by a measure of the variance of input-entropy over time. The method allows screening out rules that display glider dynamics and related complex rules, giving an unlimited source for further study. The method also shows the distribution of rule classes in the rule-spaces of varying neighbourhood sizes. The classification produced seems to correspond to our subjective view of space-time dynamics, and to global measures on the "bushiness" of typical sub-trees in attractor basins, characterized by the distribution of in-degree sizes in their branching structure. The paper explains the methods and presents results for 1d CA.

## 1 Introduction

Cellular automata (CA) are a much studied class of discrete dynamical systems where a regular array of "cells" may take discrete values, and each cell updates in parallel according to the values in its standard local neighbourhood according to a universal rule. In this paper we consider binary, finite, CA with system size $n$, neighbourhood size $k$, and periodic boundary conditions. A state of the system is the pattern[1] of 1s and 0s across the array at a given time-step. The system traces deterministic trajectories through its state-space within which emergent patterns of 1s and 0s are observed with qualities depending on the dynamical rule[2]. The trajectories typically merge with others forming trees rooted on attractor cycles. State-space is thus linked into a set of basins of attraction[13].

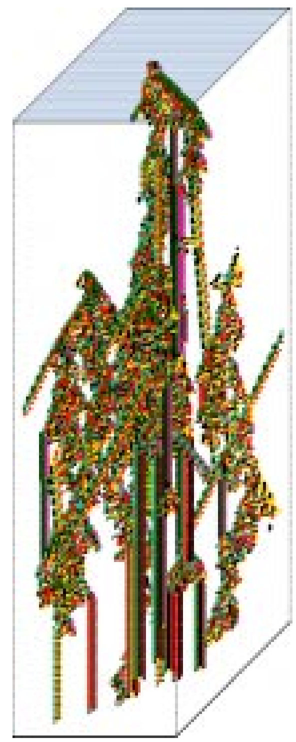


Figure 1: Right: The space-time pattern of the 2d game-of-Life on a $55 \times 55$ grid, shown in a 3d isometric projection. 2d time-steps stack below each other, and appear as if looking up at a transparent shaft. Gliders show up as diagonal structures. The initial state is →  To show up the dynamics the color of cells relate to the update neighbourhood instead of the value.
Below: A particular time-step. Dark cells are "alive", light cells "dead".

Questions that arise are how to characterise the space of all possible rules according to (1) emergent space-time patterns and (2) the attractor basin topology, and how 1 and 2 are related.

Conway's well known "game-of-Life"[2] is a 2d CA that supports coherent periodic space-time configurations that propagate and interact on a quiescent background, illustrated in figure 1.

The menagerie of configurations found have been

[1] Space-time patterns are shown as back/white (1,0). Alternatively cells are "colored" according to their update neighbourhood. If such figures are not presented in color read greyscale.

[2] CA rules in this paper are specified according to the lookup table of the outputs of $2^k$ neighbourhoods, arranged in descending order according to their binary values from left to right[10]. The outputs make a binary number with $(2^2)^k$ bits, given in hex, or decimal if $k \leq 3$. Note that odd $k$ corresponds to radius $(k-1)/2$, even $k$ is asymetric with an extra cell on the right.

grouped under various names such as gliders, glider-guns, eaters, blinkers etc. Interactions between glider streams can be contrived to achieve universal computation[2]. Langton[7] suggests that such "virtual state machines" may provide the logic for artificial life embedded in CA.

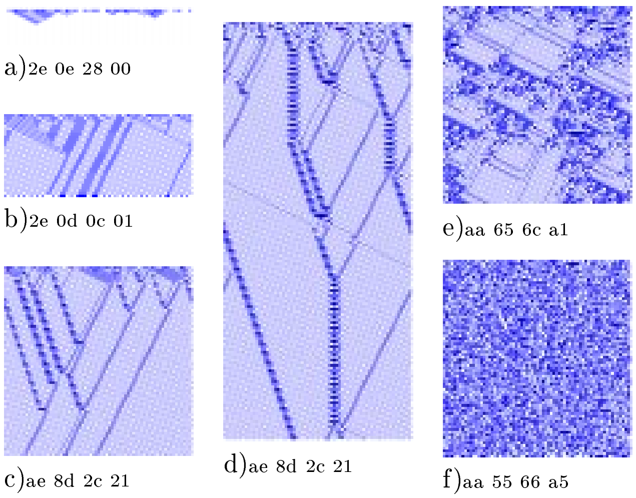
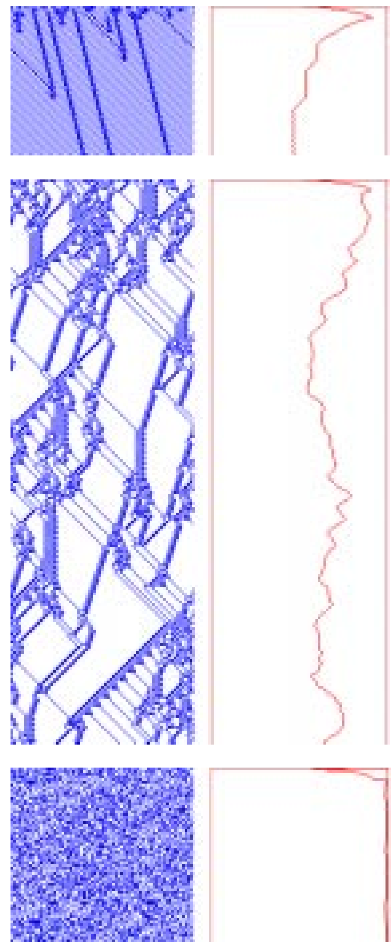



Figure 2: Typical 1d space-time patterns of a family of $k$=5 rules (shown in hex) from the same random initial state, $n$=150. Time proceeds from the top down. The rules a, b, c, d, e, f, range through ordered-complex-chaotic dynamics. a, b, d, f, correspond to Wolfram's classes 1,2,4,3, where "d" (class 4) shows sustained glider interactions. Starting with this rule, the other rules where derived by tuning the $Z$ parameter[13].

Some 1d CA rules exhibit analogous emergent structures, against backgrounds that are often periodic as well as quiescent. These coherent structures are described variously as solitary waves, gliders, virtual automata, information structures, particle-like structures and domain boundaries or defects; for simplicity they are referred to here as gliders. The emergence of gliders would in principle allow the system to be described and predictions made at a higher level, on the basis of observed glider collision rules without reference to the underlying low level CA rules. Gliders may eject or absorb a regular glider stream, or spontaneously combine to form compound gliders, which then interact at yet higher levels of description. The process could unfold without limit in large enough systems. Glider dynamics in CA provide a stark instance of self-organization in a simple system resulting from many local small scale parallel processes.

Glider dynamics can be approached from a number of perspectives, Wolfram's class 4 behaviour[12], computation[11], phase transitions between order and chaos[8], and discrete analogues of Prigogine's far-from-equilibrium dissipative structures [9]. In CA the formation, persistence and interaction of gliders can be traced at the lowest level of the system's basic components and their local interactions which are completely defined. This ability to see two levels of behaviour simultaneously, the underlying and emergent, allow insights into the mechanics of self-organization (e.g.[6]).



a) Ordered dynamics (class 2), about 100-time steps.
Rule af d1 db 47

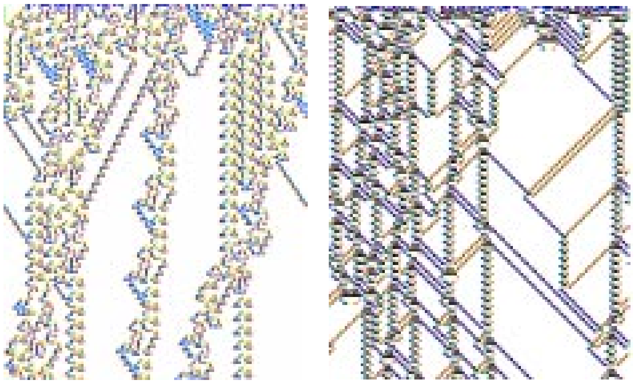
b) Complex dynamics (class 4), about 500 time-steps.
Rule 6c 1e 53 a8

c) Chaotic dynamics (class 3), about 140 time-steps.
Rule 99 4a 6a 65

Figure 3: Typical 1d CA Space-time patterns showing ordered, complex and chaotic dynamics. Time proceeds from the top down. Alongside each space-time pattern is a plot of the input-entropy described in section 3. The horizontal axis shows 0-max entropy. Note that only complex dynamics (b) exhibit high variance of this plot because glider collisions make new gliders. System size $n$=150 with periodic boundary conditions, neighbourhood size $k$=5. The same random seed was used in each case.

How simple can a CA be and yet support "interesting" glider behaviour, and what is this quality? How and why is such behaviour able to emerge? What quantitative measures can be used to identify glider dynamics? To answer these questions it would be helpful if a large sample of rules able to support gliders were available for study.

Such rules are supposed to be rare[11]. Most rules are either ordered or chaotic, though ordered rules become increasingly rare for larger $k$. In $k3$ rule-space the only two sets of glider rules that occur (rule 54 and 110, and their equivalents[13], see figure4) have been the focus of particular study (e.g. [6]).



a) $k3$ rule 110 b) $k3$ rule 54

Figure 4: Space-time patterns of the only rules in $k3$ rule-space that support interacting gliders, rule 54 and 110 (and their equivalents[13]). The rules are in decimal. $n$=150, about 200 time-steps are shown from random initial states. The color of cells relate to the update neighbourhood instead of the value. In both rules gliders are embedded in a complicated background patterns. For a clearer picture, these backgrounds have been filtered out by suppressing the relevant colors.

We will refer to glider rules as "complex" in the sense of Wolfram[10], i.e. those rules yielding localized propergating structures interacting within long transients, where the interactions are clearly "interesting". The human mind is uniquely qualified to recognize complex patterns, and to separate the interesting from the trivial, but it would be extremely useful to have measures that corresponded closely to our subjective classification. An entropy variance measure on the dynamics seems to achieve this end, and allows an unlimited source of complex rules to be found, and is further able to characterize rule-space relative to our subjective notions of order, complexity and chaos.

Figures 3(a,b and c) illustrate a range of CA behaviour, ordered, complex and chaotic. Only complex dynamics (b) exhibits high input-entropy variance (explained in sections 3-5), because glider collisions make new gliders. Ordered (a) and chaotic (c) dynamics both quickly settle to low variance. In the case of (a) because colliding gliders annihilate and the system quickly reaches its attractor at a low entropy level. In the case of (c) because the high entropy characteristic of chaos necessarily results in low variance.

This paper first discusses glider dynamics in 1d CA with examples. The automatic method for classifying rule-space is presented and the resulting rule samples are described. These local measures on trajectories are compared with global measures of the characteristic"bushiness" of sub-trees in attractor basins. Finally, a relationship between attractor basins and glider interactions is proposed.

## 2 Gliders in 1d CA

How a rule is placed within a notional order-complexity-chaos space has depended largely on our subjective appraisal of typical emergent space-time patterns. Each CA rule self-organizes its patterns in a characteristic way, and these are often recognizable given our talent for pattern recognition. For $k \leq 5$ rules a characteristic structure to the pattern is apparent even when the space-time patterns appear chaotic. This becomes less obvious for larger $k$. The characteristic pattern structure of different rules can be analyzed in formal language theory as a "regular language" with a vocabulary made up of bit sequences and a "grammar" made up of succession rules between sequences[12], and by a related "computational mechanics" approach[6].

Certain space-time patterns appear especially interesting or intriguing. Periodic sub-patterns or gliders may emerge, move across a regular background, and interact with other gliders in a particle-like manner.

Glider dynamics corresponds to Wolfram's complex class 4 behaviour in his classification of CA dynamics[11]. Wolfram orders his classes according to an alternative notion of complexity; by the increasing complexity of typical space-time patterns as measured in formal language theory[12], and draws analogies with classical continuous dynamical systems in terms of the attractors typical of each class. His classes are as follows:

| Class | CA dynamics evolves towards... | Dynamical systems analogue |
|---|---|---|
| 1. | A spatially homogeneous state... | Limit points. |
| 2. | A sequence of simple stable or periodic structures................ | Limit cycles |
| 3. | Chaotic aperiodic behaviour....... | Chaotic (strange) attractors |
| 4. | Complicated localized structures, some propagating...... | Attractors unspecified |

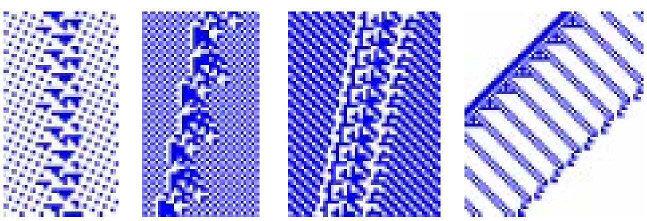
Langton and others have argued correctly that Wolfram's class 4 more naturally belongs between classes 2 and 3, at a phase transition between order and chaos, which can be traversed by tuning the $\lambda$ parameter[8], though for binary rules it has been shown that tuning the $Z$ parameter[14] gives a finer correspondence with observed behaviour. $Z$ also relates behaviour to the topology of attractor basins.

Many ordered rules have both limit points and short limit cycles (though one or the other may predominate), suggesting that class 1 and 2 may usefully be are combined. For these reasons the classification is readjusted as follows:

ordered (class 1-2) - complex (class 4) - chaotic[3] (class 3)

What are the essential features of glider behaviour? Glider dynamics occurs if a limited set of gliders emerge from random initial states, and if the interactions between gliders persist for an extended time, which requires that at least some glider collisions create new gliders. The gliders propagate at various velocities up to a maximum, the system's speed of light. At zero velocity a "glider" is static. Gliders exist against a uniform or periodic space-time background which of necessity has simultaneously emerged. This regular background (or domain) may be simple, such as a checkerboard, or a more complicated pattern (see figures 5(a-d). Backgrounds may be filtered as in figures 4 and 8 to show up gliders more clearly[4].

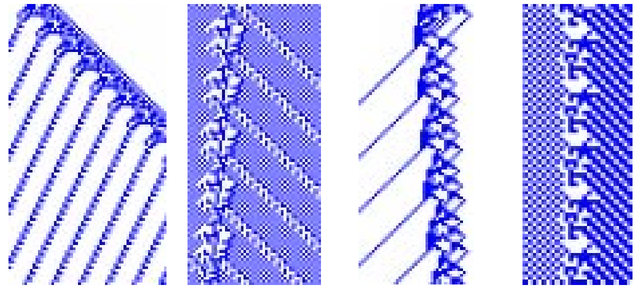
Although there are borderline cases, space-time patterns with "interesting" glider interactions are generally easily recognized in contrast to patterns that stabilize rapidly to fixed points or short periods on the one hand, or where chaotic patterns persist on the other. The borderline cases verge either on ordered or chaotic behaviour. Chaotic behaviour may also contain distinct chaotic backgrounds or domains [3], where filtering is required to uncover domain walls analogous to gliders.



a)e0 89 78 01 b)7e 86 96 de c)ad 9c 72 32 d)1c 2a 47 98

Figure 5: Gliders with various velocities and backgrounds. The $k$=5 rule numbers are shown in hex.

Gliders may be regarded as solitary waves within a background. Gliders may have the special property of solitons [1], preserving their shape and velocity after interacting with other solitons. For a neighbourhood of radius $r$, glider velocity varies from 0 to a maximum of $r$ cells per time step towards the left or right. A glider configuration that repeats at each time-step, i.e. with period one, is limited to velocities of 0,1,2,...,$r$ per time-step. Gliders with periods greater than one may have intermediate fractional velocities.

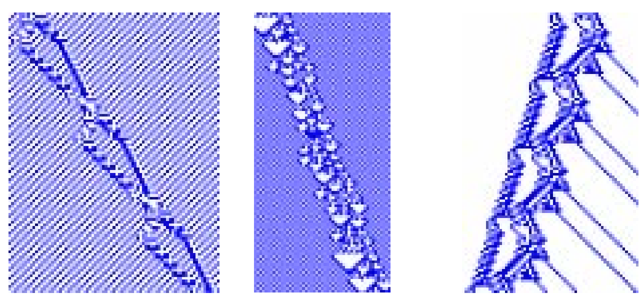


a)89 ed 71 06 b)89 ed 71 06 c)b5 1e 9c e8 d)89 ed 71 06

Figure 6: Examples of glider-guns. The $k$=5 rule numbers are shown in hex.

A glider's attributes are the background pattern and spatio-temporal period (on both sides of the glider), the glider's temporal period and velocity, its changing diameter, and the list of its repeating configurations. The same description might be applied recursively to each sub-glider component of a compound glider.

Collisions between two glider types often result in a third glider type (or more). One or both of the gliders may survive a collision with a possible shift in trajectory, or both gliders may be destroyed. In some cases a collision initially results in a chaotic interaction phase, before the final outcome emerges. The outcome of a collision is sensitive to the point of impact relative to the space-time period of each glider.



a)89 ed 71 06 b)b5 1e 9c e8 c)b5 1e 9c e8

Figure 7: A compound glider, a glider with a period of 106 time-steps, and a compound glider-gun. The $k$=5 rule numbers are shown in hex.

A glider generally represents a dislocation or defect of varying width in the background, which is often out of phase on either side of the glider, analogous to fracture planes in a crystal lattice. Alternatively, a glider may be seen as the zone that reconciles two out-of-phase backgrounds. A glider may separate two entirely different backgrounds, acting as the boundary,

[3]Note that the terms "chaos" and "chaotic" throughout this paper are used by analogy only to their meaning in continuous dynamical systems and chaos theory. Chaos in finite CA cannot conform to this strict definition although there are many common properties, for example sensitivity to initial conditions.

[4]Filtering[3, 6] in DDLab[15] is done by suppressing the display of cells that updated with reference to the most frequently occurring neighbourhoods.

as in figure 6(d). Gliders that eject a stream of sub-gliders at regular intervals, as in figure 6, and gliders that survive by absorbing a regular glider stream, as in figure 5(d), are relatively common. They are analogous to "glider-guns" and "eaters", the components for the "game-of-Life" universal computer[2]. Because a regular glider stream is essentially the same as a regular periodic background, a glider-gun creates a background, and an eater absorbs a background. Glider-guns/eaters are thus equivalent to a glider forming the boundary between two backgrounds.

Both the period and diameter of a glider may be considerable. The diameter may show a large variation within the period. Clearly gliders can only emerge in systems large enough to contain them, so that the samples described in section 5 based on $n$=150 are biased towards finding relatively small diameter gliders.

The existence of compound gliders made up of sub-gliders colliding periodically may be expected in large enough systems. Compound gliders could combine into yet higher order structures[7], and the process could unfold hierarchically without limit. The example in figure 7(a) shows a compound glider made from two independent gliders locked in a cycle of repeating collisions.

Once gliders have emerged, CA dynamics may in principle be described at a higher level, by glider collision rules as opposed to the underlying CA rules.

# 3 Input-entropy

A special case of block probability/entropy[11] is input frequency/entropy. The frequency with which each of the neighbourhoods in the rule-table is "looked up" at a given time-step can be represented by a histogram as in figure 8(right), which distributes the total of $n \times w$ lookups among the $2^k$ neighbourhoods (shown as the fraction of total lookups), where $n$=system size, $w$=the window of time steps defined ($w$=10 in figure 8), and $k$=neighbourhood size.

The Shannon entropy of the lookup frequency histogram, the "input-entropy" $S$ at time-step $t$ for one time-step ($w$=1), is given by

$$S^t = -\sum_{i=1}^{2^k} \left( \frac{Q_i^t}{n} \times log \left( \frac{Q_i^t}{n} \right) \right)$$

Where $Q_i^t$ is the lookup frequency of neighbourhood $i$ at time $t$. In practice, to smooth the measures, they are taken over a moving window of time-steps ($w$=5 in the automatic sample). Example snapshots of the changing input-frequency histogram are given in figures 8 and 9.

In a random initial state the different $k$- blocks occur with equal probability. The start entropy will be correspondingly high. Below we discuss the typical evolution of the input frequency histogram and input entropy for successive iterations of the CA, for ordered, chaotic and complex dynamics.

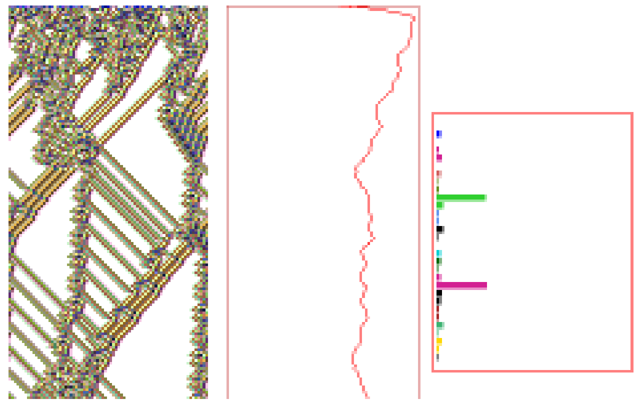
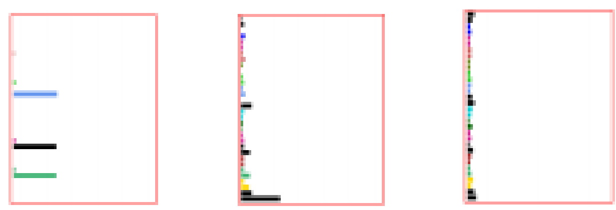


Figure 8: The complex 1d CA, $k = 5$, rule 36 0a 96 f9, $n = 150$. Left: The space-time pattern colored according to each cell's neighbourhood, with the background colors (the high bars in the histogram) filterd out.
Centre: The input-entropy plot. The horizontal axis shows 0-max entropy. The vertical axis shows time progressing from the top down.
Righ: The lookup frequency histogram. The vertical axis shows the 32 $k5$ neighbourhoods in decimal equivalent order, from 0 (bottom) to $2^5 - 1 = 31$ (top). The horizontal axis shows the look-up ($k$-block) frequency 0-max averaged over 10 time-steps



a)Ordered b)Complex c)Chaotic

Figure 9: The lookup frequency histogram for the same rules shown in figure 3. (a) ordered, (b) complex and (c) chaotic. $n$=150, $w$=10. The vertical axis shows the 32 $k5$ neighbourhoods in decimal equivalent order, from 0 (bottom) to $2^5-1 = 31$ (top). The horizontal axis shows the look-up ($k$-block) frequency 0-max averaged over 10 time-steps.

**Ordered Dynamics**
In ordered dynamics the lookup frequency histogram will rapidly become highly unbalanced, with most neighbourhoods never looked at (their lookup frequency = 0). The few remaining high frequencies settle on constant or periodic values. The entropy will settle at a low constant or periodic value, corresponding to a fixed point or short cycle attractor. Ordered behaviour produces extremely short and

bushy transient trees with a high density of states without predecessors (G-density). Ordered rules decrease disorder and entropy.

**Complex Dynamics**
In complex dynamics, the frequency histogram becomes unbalanced, with large erratic fluctuations, reflected in the entropy curve. As in ordered behaviour, a proportion of neighbourhoods are never looked at again after the initial sorting out phase. After an extended time the system generally settles onto a short attractor cycle. The high frequency neighbourhoods correspond to the emergent background. The low frequency neighbourhoods to the interacting gliders.

Glider dynamics is subject to two countervailing tendencies. On the one hand a tendency towards order because of the dominant periodic background(s), but the zones of order are mobile, their boundaries form the moving particles or gliders. When these collide there is a tendency toward chaos. The collisions may form a temporary zone of chaotic dynamics before new gliders emerge.

In systems of the order of size considered here, order or chaos may predominate at different times causing the entropy to vary. For large networks where colliding and non-colliding zones co-exist, the entropy variance will be reduced, to zero in the limit of infinite size.

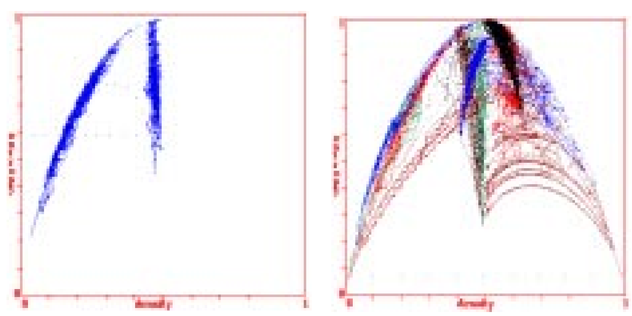
A measure of the variability of the input-entropy curve is its variance or standard deviation from the mean, high entropy variance is then a sure sign of complex space-time dynamics.

High entropy variance is not only characteristic of glider dynamics but also of other less frequent types of complex dynamics. For example two chaotic co-existing (competing) domains which are qualitatively different, having different block probabilities will produce an erratic entropy curve as one or the other domain becomes dominant. The boundary between the chaotic domains may be seen an a different type of particle lacking the regularity of a glider. Such particles may be isolated by filtering out the chaotic domains[3]. A related type of complex dynamics occurs when a chaotic domain co-exists with glider dynamics or with just a regular background.

**Chaotic Dynamics**
In chaotic dynamics, the lookup frequency histogram will fluctuate irregularly within a narrow band of low values, and the entropy will fluctuate irregularly within a narrow high band, corresponding to dynamics on very long transients or cycles, analogous to strange attractors in continuous dynamical systems. Transient trees will be sparsely branched thus will tend to be very long with relatively low G-density. Chaotic rules increase or conserve disorder and entropy.

## 4 Entropy-density signatures

Entropy-density plots for a number of complex rules are shown in figure 10. Input-entropy is plotted against the density of 1s relative to a moving window of time-steps. Each rule produces a characteristic cloud of points. The clouds lie within a parabolic envelope because high entropy is most probable at medium density, low entropy is most probable at either low or high density. For complex rules the clouds have a marked vertical extent(i.e. high variance) because the input-entropy varies significantly. Each complex rule produces a plot with its own distinctive signature. By contrast, chaotic rules will give a flat compact cloud at high entropy (at the top of the parabola).



a) $k5$ rules 6c 1e 53 a8 & 36 0a 96 f9 b) several complex $k5$ rules

Figure 10: Entropy-density scatter plots: input-entropy (vertical axis) against the density of 1s relative to a moving window of time-steps $w$=10. $k$=5, $n$=150. About 1000 time-steps from several random initial states for each rule, which of which has its own distinctive signature. Note the high variance of the input entropy.
(a) The rules shown in figures 3 and 8.
(b) A number of complex rules from the automatic sample.

For ordered rules the plot also has a large vertical extent as the entropy falls off, but there are very few data points because the system moves very rapidly to an attractor.

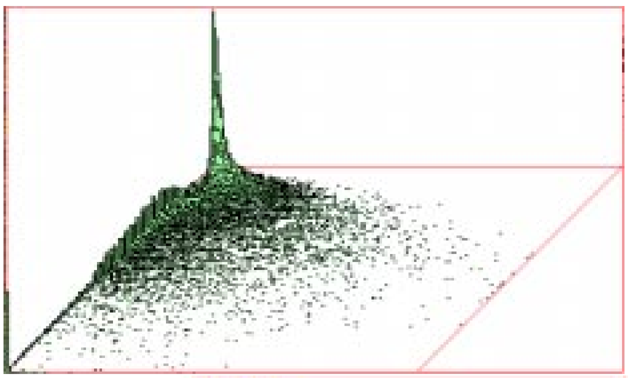
Gutowitz[5] has also shown entropy-density plots for large samples of rule-space, but his plots show a single point for each rule where the measures on that rule have settled down, whereas the plots shown here focus on the transient history of the system. These plots distinguish order, complexity and chaos by the vertical extent and density of the cloud.

# 5 Automatically classifying rule-space

To distinguish ordered, complex and chaotic rules automatically the mean input-entropy taken over a span of time steps is plotted against the standard deviation of the input entropy. The standard deviation is given by,

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

where $x_i$ = deviation of each measure from the mean, and $n$ = number of measures. The variance = $\sigma^2$.



17690 $k$=5 rules

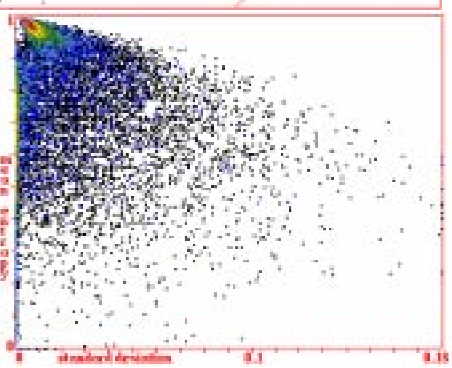



Figure 11: Bottom: A scatter plot of mean entropy (vertical axis) against standard deviation of the entropy for a random sample of 17680 $k$=5 rules, $n$=150. The color coding indicates frequency (as below). Top: The same plot laid horizontally with a third (vertical) axis showing the frequency of points falling within a 128x128 grid. Chaotic rules are concentrated in the top left "tower", ordered rules in the "ridge" close to the $y$-axis with lower mean entropy.

In the following results for $k$=5, 6 and 7 rules, the input-entropy was measured over a moving window of 5 time-steps ($w$=5). The system was run for 430 time-steps from a random initial state. The measures were only taken into account for the last 400 time-steps, the first 30 were ignored to allow the system to evolve beyond its initial sorting out phase. The mean input-entropy, and the standard deviation from this mean, were calculated relative to these 400 time-steps. This procedure was repeated 5 times from different random initial states for each rule. The measures were averaged and a point was plotted of mean input-entropy against the standard deviation of the entropy as shown in figure 11 (below).

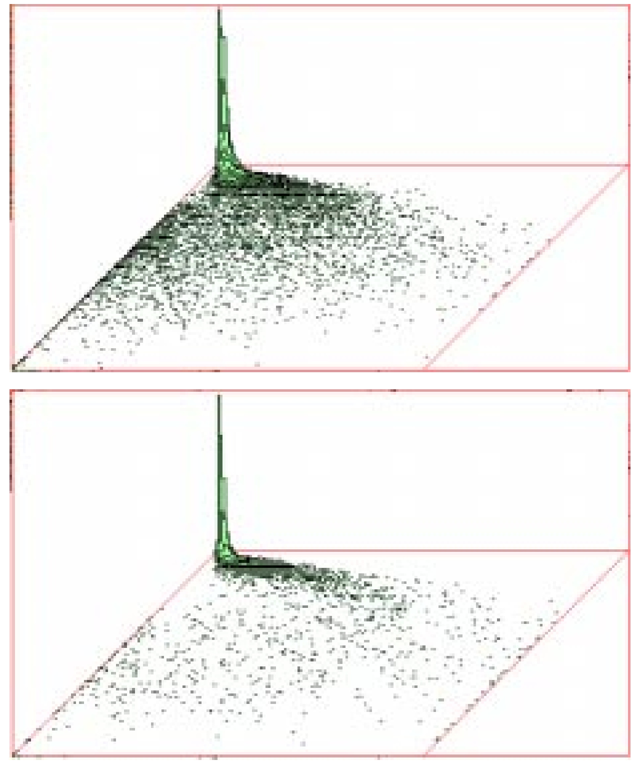


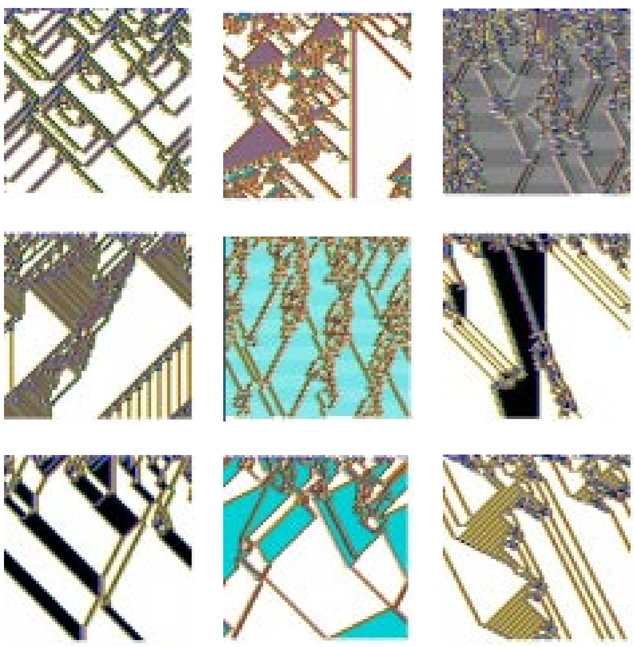
Figure 12: Mean entropy - standard deviation scatter plots as in figure 11. Top: $k$=6, sample=15425. Bottom: $k$=7, sample=14221. Note that as $k$ increases, ordered rules, and complex rules to a lesser extent, become less probable. Complex rules are spread in the area to the right with higher standard deviation.

Rules (and initial states) were selected at random by setting a 1 or 0 with equal probability for each entry in the rule's look-up table (and each bit in the initial state).

Figures 11 and 12 present the plots for $k$=5, 6 and 7 rules. The sample sizes are as follows; 17680 $k$=5, 15425 $k$=6, 14221 $k$=7. To see the distribution of rules, the plots include an extra axis, making a 2d histogram, representing the number of rules falling within blocks on a 128x128 grid overlaid over the scatter plot.

Looking at the $k$=5 2d histogram, the "tower" in the upper left represents chaotic rules with low standard deviation and high mean entropy. The ridge on the left represents ordered rules with low standard deviation and a spread of lower mean entropy. There is a low diagonal valley between the tower and the ridge representing a distinct boundary. Rules to the right of the plot, with higher standard deviation are complex rules. There is not a sharp boundary with ordered and chaotic rules; as the standard deviation decreases glider interactions become more dense with longer transients verging on chaos, or less dense with shorter transients verging on order. Any standard deviation above the maximum scale has been reset to the maximum of 1.8. The $k$=6 and $k$=7 plots show a greater frequency of chaotic rules and a declining frequency of ordered and complex rules as $k$ increases. The decrease in ordered rules is more marked.

The rule samples and measures, including each rule's $\lambda$ and Z parameters, were saved to file sorted by decreasing standard deviation, and decreasing mean entropy for each measure of standard deviation. Exmples of complex rules from the samples are shown in figure 13. More examples are presented in [16], and are all available with the DDLab software[15].



$k$=5 $k$=6 $k$=7

Figure 13: Examples of $k$=5, 6 and 7 complex space-time patterns, with high standard deviation, from the automatic samples. $n$=150, 150 times-steps from random initial states. Cells are colored according to the neighbourhood.

To check whether the expected dynamics (recognized subjectively) corresponds to the measures as plotted, the dynamics of particular rules at different positions on the plots may be examined very efficiently in DDLab, for example with a mouse click on the scatter plot. Preliminary scans indicate that the expected behaviour is indeed found, but further investigation is required to properly demarcate the space between ordered, complex and chaotic rules and to estimate the proportion of different rule classes for different $k$.

The capability to generate the automatic samples is available in DDLab[15]. This may be done for larger system size $n$ and and neighbourhood $k$, and the various other parameters for generating the samples can be adjusted.

Input entropy is a local measure on the space-time patterns of typical trajectories. The distribution of the rule samples according to these local measures may be compared with global measures on attractor basin topology, G-density and the in-degree frequency (section 6 below). A preliminary scan indicates a strong relationship between these global measures and the rule sample input-entropy plots.

# 6 Attractor Basin Topology measures

For CA rules in general, the quality of dynamical behaviour from ordered to chaotic is reflected by the topology[5] of attractor basins in the sense of the characteristic "bushiness" of subtrees, and the comparative length of transients, attractors etc. This characterizes the degree of convergence of dynamical flow seen in attractor basins, or the degree of dissipation viewing the CA as a dissipative dynamical system.

Consider a transient sub-tree with $n$ vertices linked by $n$-1 edges. In the space of all possible topologies there are two extreme cases. Maximum convergence occurs where $n$-1 edges converge in one step onto a single vertex. Here the G-density, the density of garden-of-Eden states (those without predecessors) is close to 1. Minimum convergence occurs where the edges are strung out in a chain linked by single edges.. Here the G-density is close to 0. Between these two extremes there lies a spectrum of degrees of convergence characterized by the topology of typical transient trees in terms of their "bushiness". High convergence implies short and dense trees with vertices having a high in-degree, signifying order. Low convergence implies long sparse trees, with vertices having a low in-degree, signifying chaos. Figures 14 and 15 give examples.

**G-density**

A simple measure that captures the degree of convergence is the density of garden-of-Eden states (G-density) counted in attractor basins or sub-trees, the rate of increase of G-density with $n$ is a further indication[14]. Because the average in-degree (defined

[5]The term "topology" is used here to denote how vertices are linked by edges in attractor basins, and should not be confused with its standard meaning in mathematics.

below), including in-degrees of zero, must equal one for any basin of attraction, high G-density (and a high rate of increase) signifies high convergence, low G-density (and a low rate of increase) signifies low convergence. High and low convergence in turn indicate ordered and chaotic dynamics.

**In-Degree Frequency**
A more comprehensive measure of the convergence typical of a rule, is the in-degree frequency distribution plotted as a histogram. This can be taken on attractor basins, or on a subtree[6] (or fragment thereof) for larger systems. Subtrees are portrayed as graphs showing trajectories merging onto the subtree root state, the direction of time is inwards towards the root.

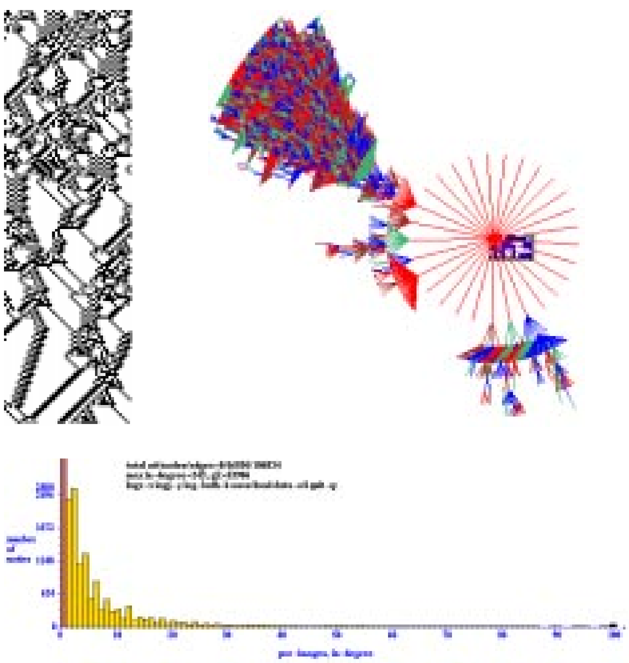


Figure 14: Top left: A transient of the complex $k5$ rule 6c 1e 53 a8, $n$=50, also shown in 3(b). Top right: A fragment of the subtree rooted on the final state in the space-time pattern. The subtree was stopped after 21 levels. Bottom: The in-degree histogram of the subtree fragment. 66896 nodes, 53906 garden-of-Eden nodes (with in-degree=0, off the scale), G-density=0.8, maximum in-degree=243 (off the scale).

The in-degree of a state (a vertex in the subtree portrait) is the number of its immediate predecessors (also called pre-images), thus the number of incoming directed edges to the vertex. Examples of in-degree histograms for particular subtrees for a complex and a chaotic rule are shown in figures 14 and 15. The horizontal axis represents in-degree size, from zero (garden-of-Eden states) upwards, the vertical axis represents the frequency of the different in-degrees. For ordered rules of the size illustrated, in-degrees become astronomical so in this case the plots can only be made for small systems. From the preliminary data gathered so far the profile of the in-degree histogram for different classes of rule is as follows:

**Ordered rules:** A bi-modal distribution, very high garden-of-Eden frequency and significant frequency of high in-degrees.

**Complex rules:** A power law distribution.

**Chaotic rules:** A relatively lower garden-of-Eden frequency compared to complex rules, and a higher frequency of low in degrees.

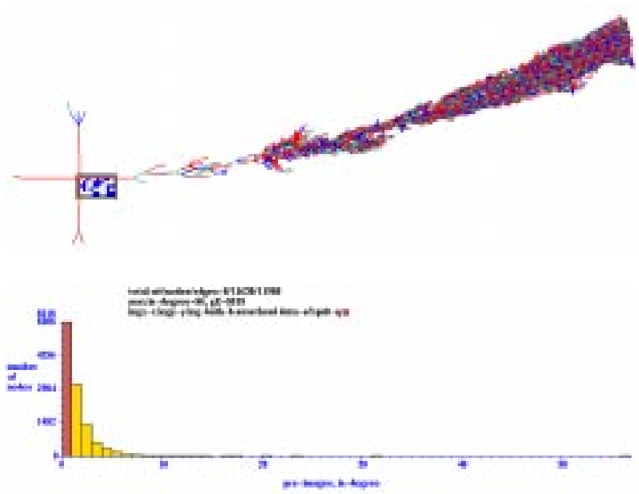


Figure 15: Top left: The chaotic $k5$ rule 99 4a 6a 65. Space-time patterns shown in figure 3(c), $n$=50. Top: A fragment of the subtree generated by running forward from a random initial state, then backwards from the state reached. The subtree was stopped after about 1000 levels. Bottom: The in-degree histogram of the subtree fragment. 11630 nodes, 5818 garden-of-Eden nodes (with in-degree=0), G-density=0.5, maximum in-degree=56.

The in-degree frequency is a finer measure of attractor basin topology and convergence than G-density alone, and the in-degree histogram gives characteristically different profiles for order, complexity and chaos.

Issues for further investigation are: a systematic look at the profiles relative to rules at various positions on the mean entropy/standard deviation scatter plots, how profiles change with system size, if a subtree fragment is representative of the dynamics as a whole, if the profile changes for subtrees deep in a basin of attraction as opposed to those close to the outer leaves, and to look at the part of subtrees close to (within a given diameter in reverse time-steps) from particular trajectories, especially in relation to glider dynamics.

[6]For more about attractor basins and subtrees, and how they are portrayed and computed see [13, 16]

# 7 Glider interactions and basins of attraction

It is possible to identify classes of configurations that make up different components of attractor basins in glider rules. In random states, configurations occur with equal probability, so the special glider-background configurations are unlikely. Non-glider/background states make up the majority of state-space, and are likely to be garden-of-Eden states, or states just a few steps forward in time from garden-of-Eden states. They occur in the initial sorting out phase of the dynamics and appear as short bushy dead-end side branches along the length of long transients, as well as at their tips.

States dominated by glider and background configurations are special cases, a small sub-category of state-space. They constitute the glider interaction phase, making up the main lines of flow within the long transients. This has also been noted by Domain [4], who described the main lines of flow as the topological skeleton of physically relevant states and the short dead end side branches from garden-of-Eden states as a skin of non-physical transitional states, comprising the bulk of the nodes in an attractor basin.

Gliders in the interaction phase can be regarded as competing sub-attractors, with the final survivors persisting in the attractor cycle. Finally, states made up solely of non-interacting gliders configurations (i.e. having equal velocity), or backgrounds free of gliders, must cycle and therefore constitute the relatively short attractors, with a period depending on the glider velocity. The attractor states are made up of gliders, compound gliders or just backgrounds, and thus form a tiny sub-category of state-space. By simply looking at the space-time patterns of a glider rule from a number of different initial states, most gliders in its glider repertoire (relative to the system size) may be identified. A complete list would allow a complete description of all the attractors in state-space, by finding all possible permutation of non-interacting gliders.

# 8 Conclusion

Glider dynamics in 1d CA is especially interesting as an example of emergent structure in a fully defined simple system. An unlimited source of complex rules that support gliders is available by an automatic procedure based on local measures, in particular input-entropy variance, which also classifies rule-space for a spectrum of ordered, complex and chaotic dynamics. Global measures, G-density and in-degree frequency, taken on the topology of attractor basins and subtrees, relate to the local measures. Though not discussed here, both local and global measures relate to the rule parameter $Z$[13, 14, 16].

Further systematic investigation of both the local and global measures, based on the automatic rule samples, and extended samples, is required for a deeper understanding of CA rule-spaces. The computer tools for such an investigation are largely in place.

**The software**

"Discrete Dynamics Lab" (DDLab)[15], was used for the computations, examples, figures and data in this paper. The software is available at:
http://www.santafe.edu/~wuensch/ddlab.html.

# References


[1] Aizawa,Y., I.Nishikawa and K,Kaneko, (1990) "Soliton Turbulence in One-Dimensional Cellular Automata", *Physica D 45*, 307-327.

[2] Conway,J.H., (1982) "What is Life?" in "*Winning ways for your mathematical plays*", Berlekamp,E, J.H.Conway and R.Guy, Vol.2, chap.25, Academic Press, New York.

[3] Crutchfield,J.P., and J.E.Hanson, (1993) "Turbulent Pattern Bases for Cellular Automata", Physica D 69: 279-301.

[4] Domain,C. and H.A.Gutowitz, (1997) "The topological skeleton of cellular automaton dynamics", Physica D, vol.103 nos 1-4, 155-168.

[5] Gutowit,H.A., and C.G.Langton, (1995) "Mean Field Theory of the Edge of Chaos", in Proceedings of ECAL3, Springer.

[6] Hanson,J.E., and J.P.Cruchfield (1997) "Computational Mechanics of Cellular Automata, An example", Pysica D, vil. 103, 169-189.

[7] Langton,C.G., (1986) "Studying Artificial Life with Cellular Automata", *Physica D 22*, 120-149.

[8] Langton,C.G., (1990) "Computation at the Edge of Chaos", *Physica D 42*, 12-37.

[9] Prigogine,I., and I.Stengers, (1984) "*Order out of Chaos", Heinemann*".

[10] Wolfram,S., ed., (1983) "Statistical mechanics of cellular automata", *Reviews of Modern Physics 55* no 3, 601-64.

[11] Wolfram,S., (1984) "Computation Theory of Cellular Automata", *Commun.Maths.Phys.96*, 15-57.

[12] Wolfram,S., (1984) "Universality and complexity in cellular automata", *Physica 10D*, 1-35.

[13] Wuensche,A., and M.J.Lesser. (1992) "*The Global Dynamics of Cellular Automata*", Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley.

[14] Wuensche,A., (1994) "Complexity in One-D Cellular Automata", Santa Fe Institute Working Paper 94-04-025.

[15] Wuensche,A.,(1996) "Discrete Dynamics Lab (DDLab)", http://www.santafe.edu/~wuensch/ddlab.html.

[16] Wuensche,A., (1997) "Attractor Basins of Discrete Networks" (DPhil thesis), Cognitive Science Research Paper 461, Univ. of Sussex.