II. Spatial Systems

A. Cellular Automata
B. Pattern Formation
C. Slime Mold
D. Excitable Media

A. Cellular Automata

Cellular Automata (CAs)

- Invented by von Neumann in 1940s to study reproduction
- He succeeded in constructing a self-reproducing CA
- Have been used as:
  - massively parallel computer architecture
  - model of physical phenomena (Fredkin, Wolfram)
- Currently being investigated as model of quantum computation (QCAs)

Structure

- Discrete space (lattice) of regular *cells*
  - 1D, 2D, 3D, …
  - rectangular, hexagonal, …
- At each unit of time a cell changes state in response to:
  - its own previous state
  - states of neighbors (within some “radius”)
- All cells obey same state update rule
  - an FSA
- Synchronous updating

Example: Conway's Game of Life

- Invented by Conway in late 1960s
- A simple CA capable of universal computation
- Structure:
  - 2D space (periodic or unbounded)
  - rectangular lattice of cells
  - binary states (alive/dead)
  - neighborhood of 8 surrounding cells (& self)
  - simple population-oriented rule

State Transition Rule

- Live cell has 2 or 3 live neighbors
  – stays as is (stasis)
- Live cell has < 2 live neighbors
  – dies (loneliness)
- Live cell has > 3 live neighbors
  – dies (overcrowding)
- Empty cell has 3 live neighbors
  – comes to life (birth)

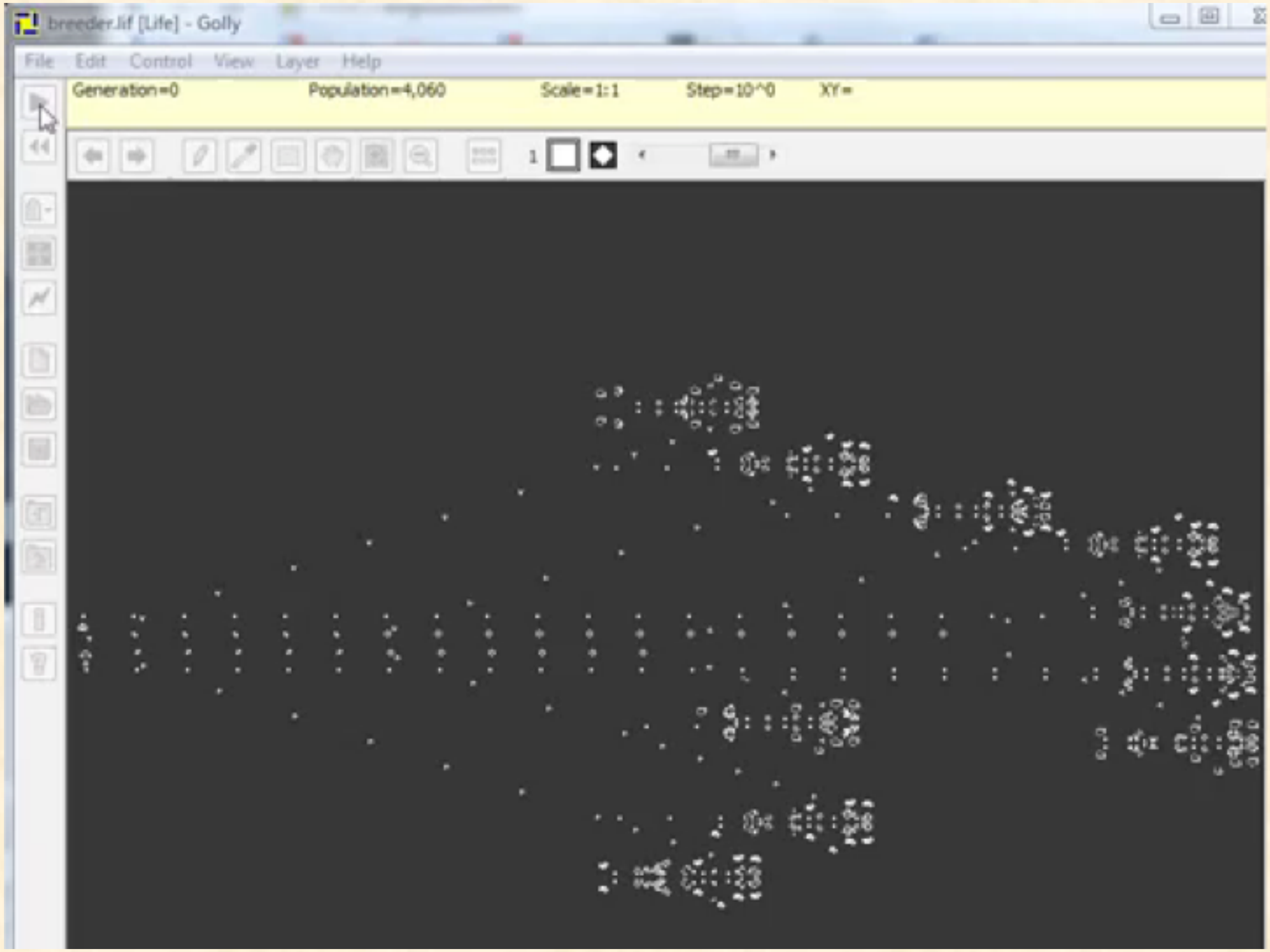
Demonstration of Life

Run NetLogo Life

or

<web.eecs.utk.edu/~mclennan/Classes/420-527/NetLogo/Life.html>

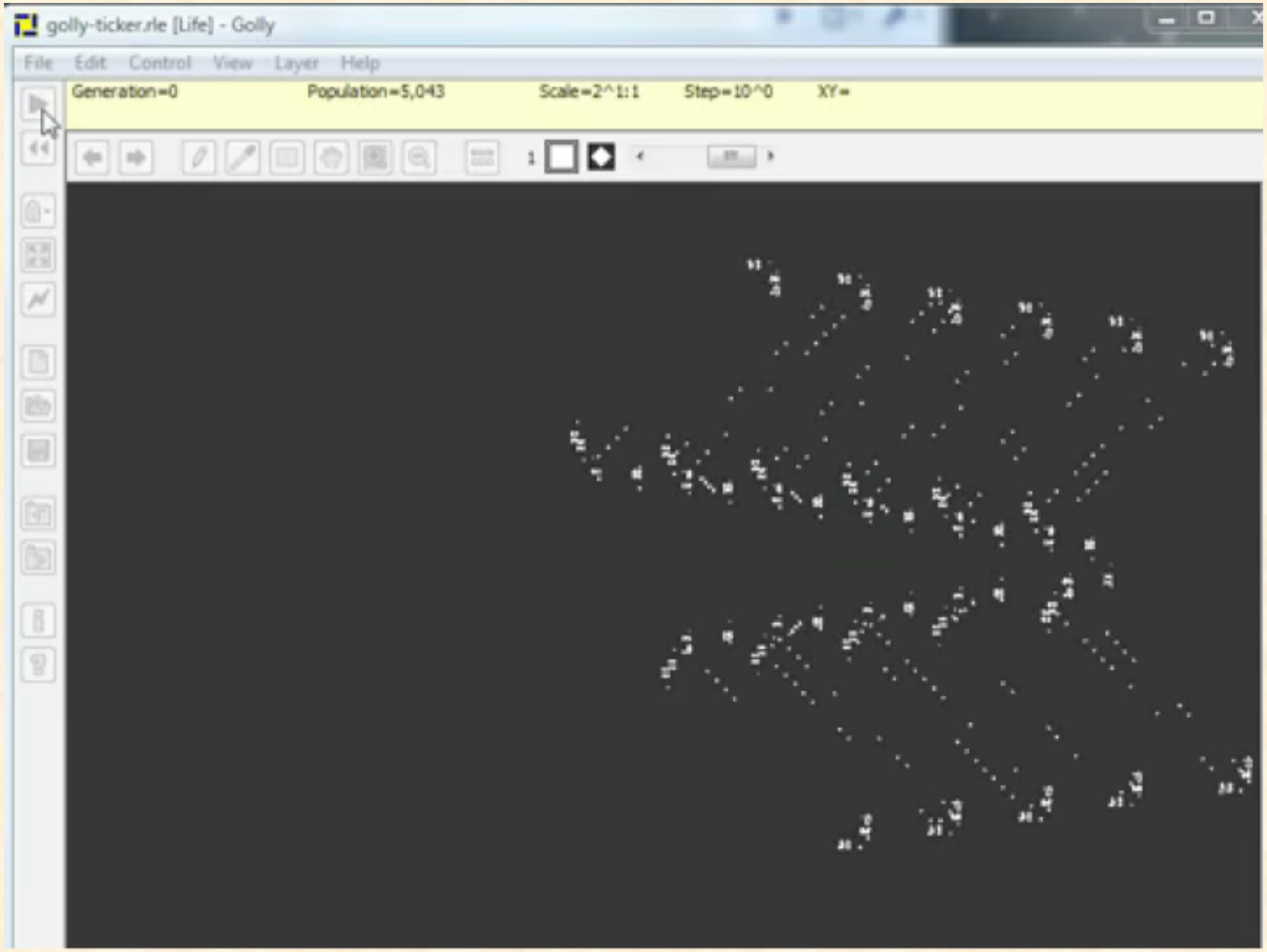
Breeder using Golly



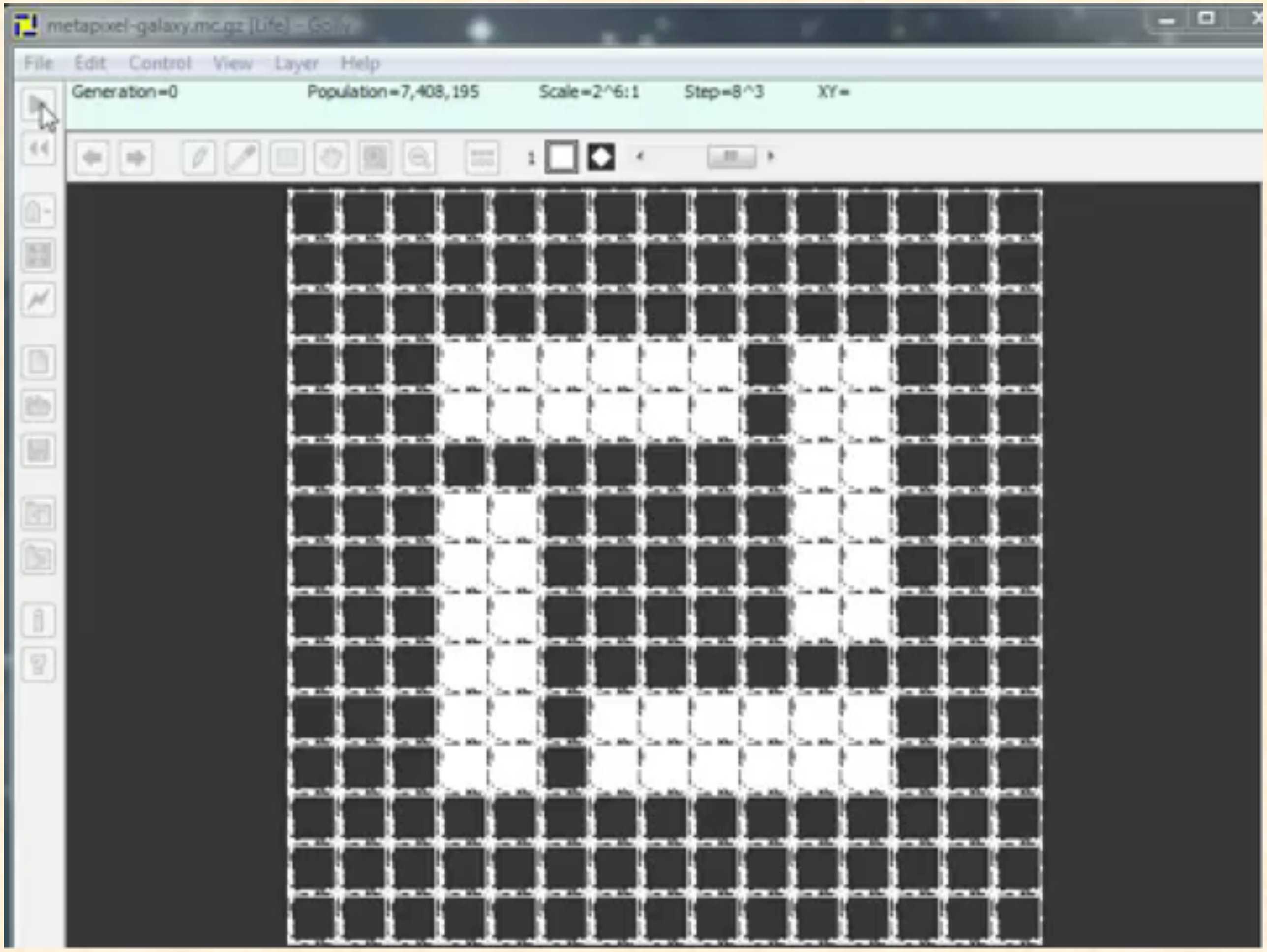
http://golly.sourceforge.net/

(videos: https://www.youtube.com/playlist?list=PLu9PfMOtQsVzSbY8zh0YLKnT4aLMY4rTh)

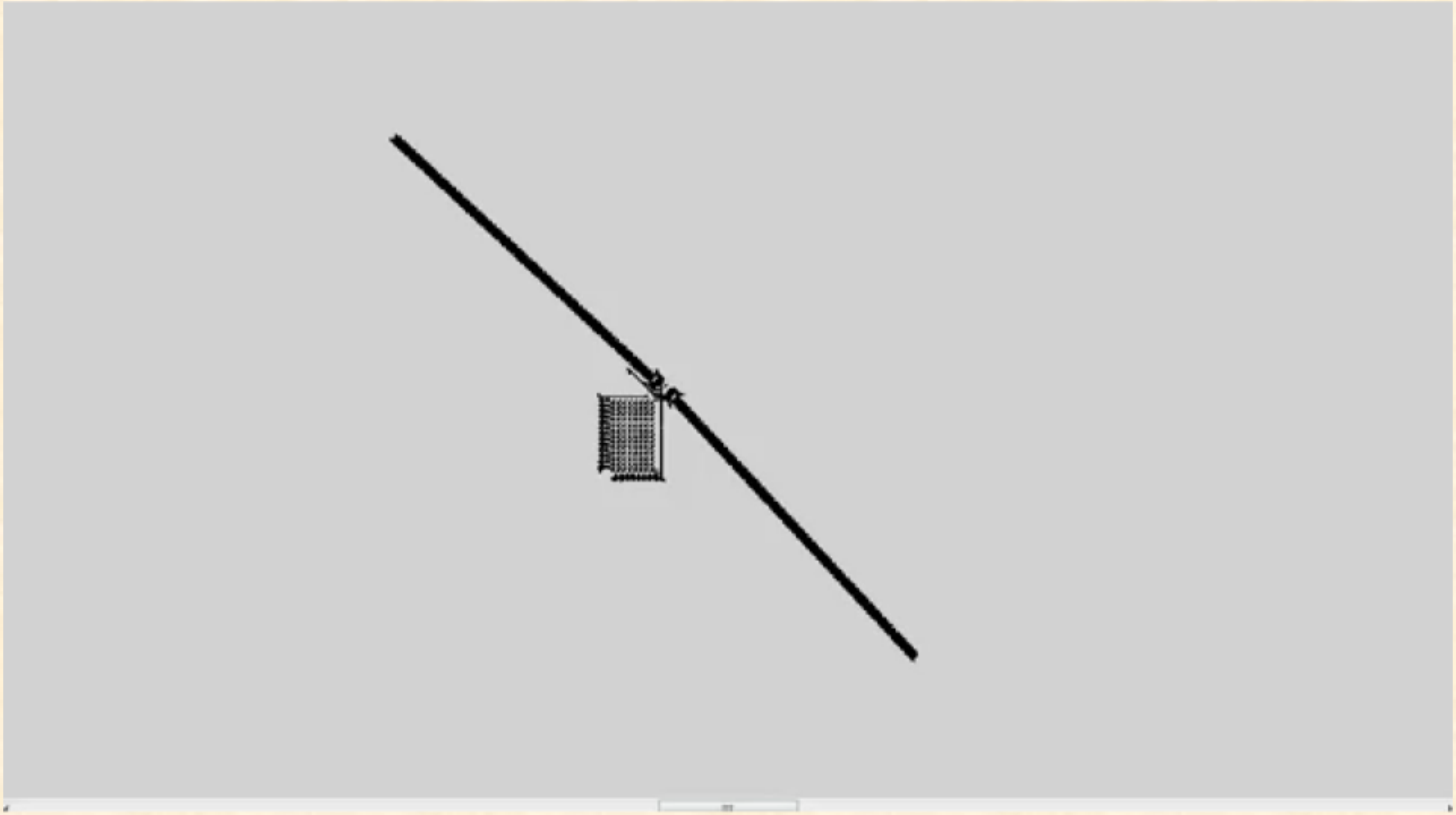
Banner


golly-ticker.rle [Life] - Golly
File Edit Control View Layer Help
Generation=0
Population=5,043
Scale=2^1:1
Step=10^0
XY=

Life Simulating Life



Universal Turing Machine



Some Observations About Life

1. Long, chaotic-looking initial transient
   – unless initial density too low or high
2. Intermediate phase
   – isolated islands of complex behavior
   – matrix of static structures & “blinkers”
   – gliders creating long-range interactions
3. Cyclic attractor
   – typically short period

From Life to CAs in General

- What gives Life this very rich behavior?
- Is there some simple, general way of characterizing CAs with rich behavior?
- It belongs to Wolfram's Class IV

The four classes of feedback behaviour

(a) Fixed points

(b) Simple periodic orbits

(c) Period-n orbit

(d) Chaos

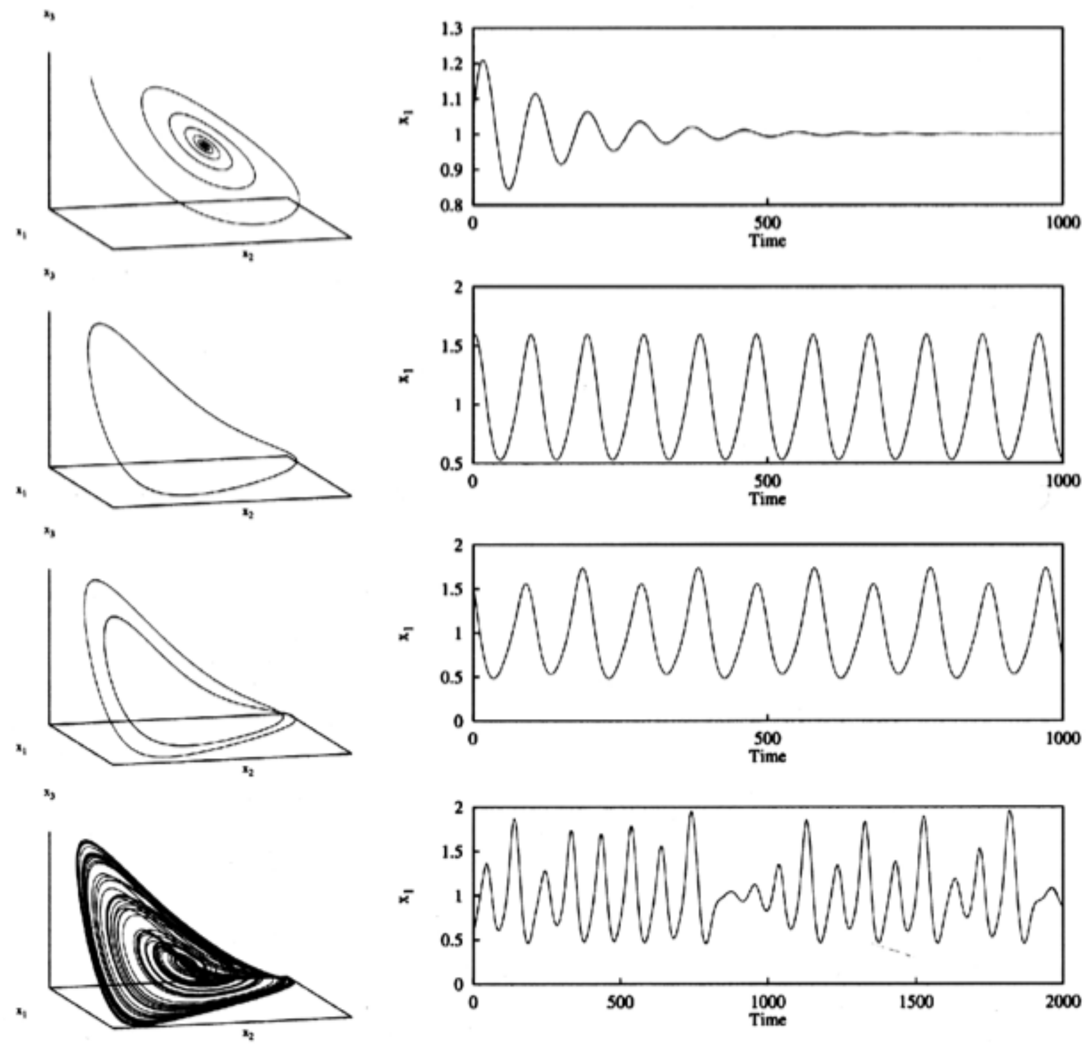


fig. from Flake via EVALife

Wolfram’s Classification

- Class I: evolve to fixed, homogeneous state
  - ~ limit point
- Class II: evolve to simple separated periodic structures
  - ~ limit cycle
- Class III: yield chaotic aperiodic patterns
  - ~ strange attractor (chaotic behavior)
- Class IV: complex patterns of localized structure
  - ~ long transients, no analog in dynamical systems

Langton's Investigation

Under what conditions can we expect a complex dynamics of information to emerge spontaneously and come to dominate the behavior of a CA?

Approach

- Investigate 1D CAs with:
  - random transition rules
  - starting in random initial states
- Systematically vary a simple parameter characterizing the rule
- Evaluate qualitative behavior (Wolfram class)

Why a Random Initial State?

- How can we characterize typical behavior of CA?
- Special initial conditions may lead to special (atypical) behavior
- Random initial condition effectively runs CA in parallel on a sample of initial states
- Addresses emergence of order from randomness

Assumptions

- Periodic boundary conditions
  - no special place
- Strong quiescence:
  - if all the states in the neighborhood are the same, then the new state will be the same
  - persistence of uniformity
- Spatial isotropy:
  - all rotations of neighborhood state result in same new state
  - no special direction
- Totalistic [not used by Langton]:
  - depend only on sum of states in neighborhood
  - implies spatial isotropy

Langton's Lambda

- Designate one state to be quiescent state
- Let K = number of states
- Let $N = 2r + 1$ = size of neighborhood
- Let $T = K^N$ = number of entries in table
- Let n_q = number mapping to quiescent state
- Then

$$\lambda = \frac{T - n_q}{T}$$

Range of Lambda Parameter

- If *all* configurations map to quiescent state:
  $\lambda = 0$
- If *no* configurations map to quiescent state:
  $\lambda = 1$
- If every state is represented *equally*:
  $\lambda = 1 - 1/K$
- A sort of measure of "excitability"

Example

- States: $K = 5$
- Radius: $r = 1$
- Initial state: random
- Transition function: random (given λ)

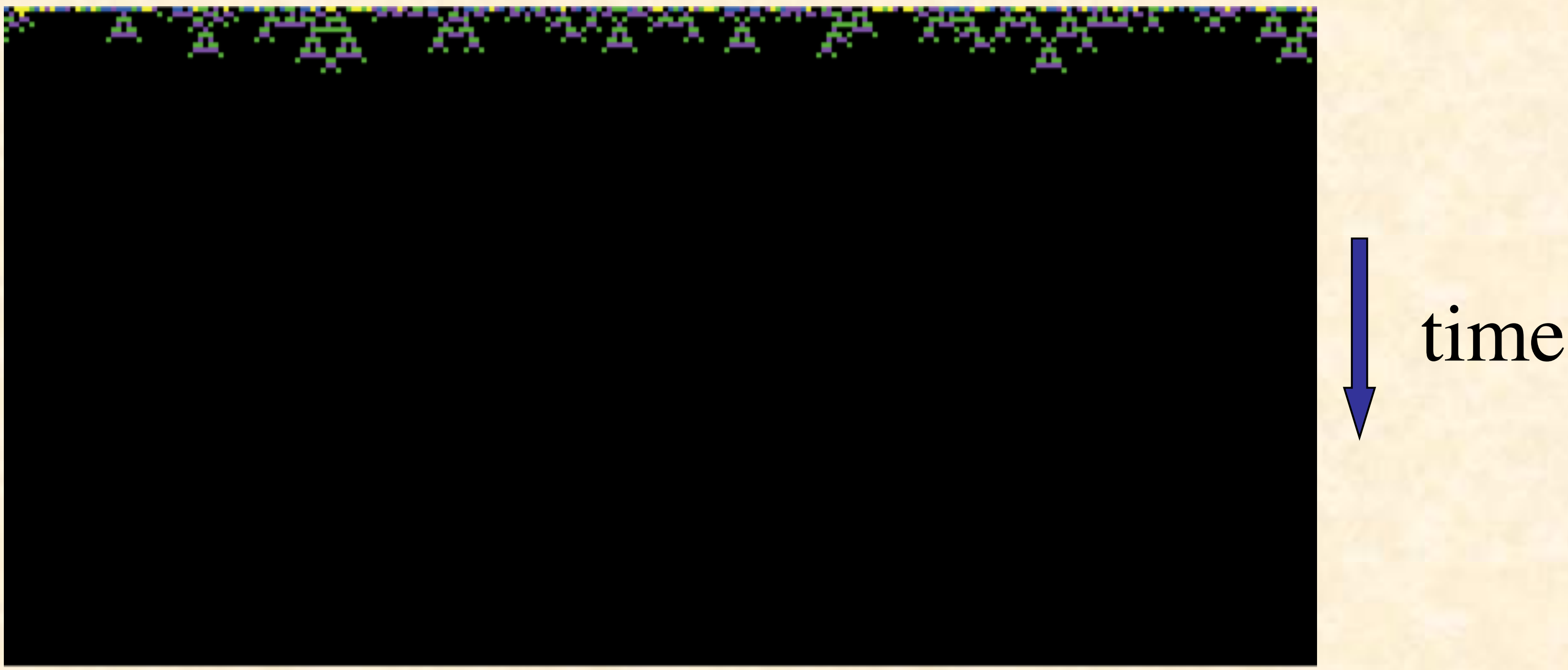
Demonstration of 1D Totalistic CA

Run NetLogo 1D CA General Totalistic

or

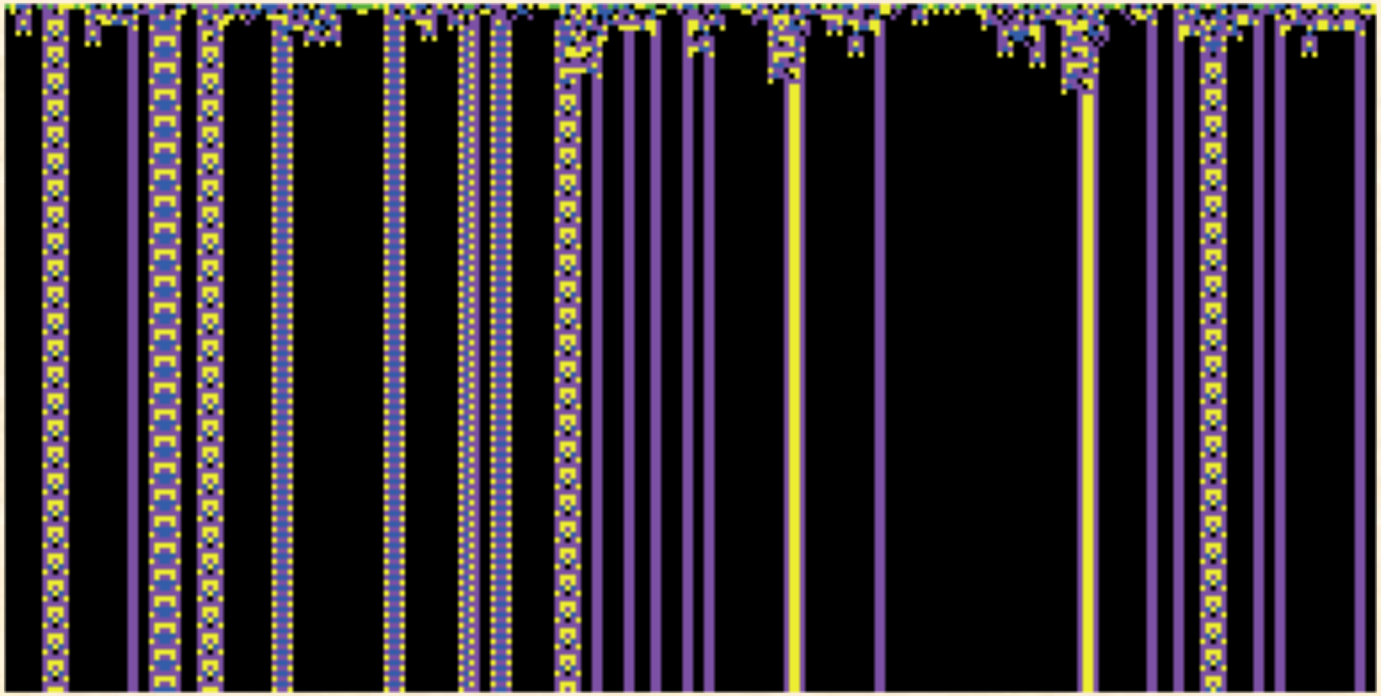
`<web.eecs.utk.edu/~mclennan/Classes/420-527/NetLogo/CA 1D General Totalistic.html>`

Class I ($\lambda = 0.3$)

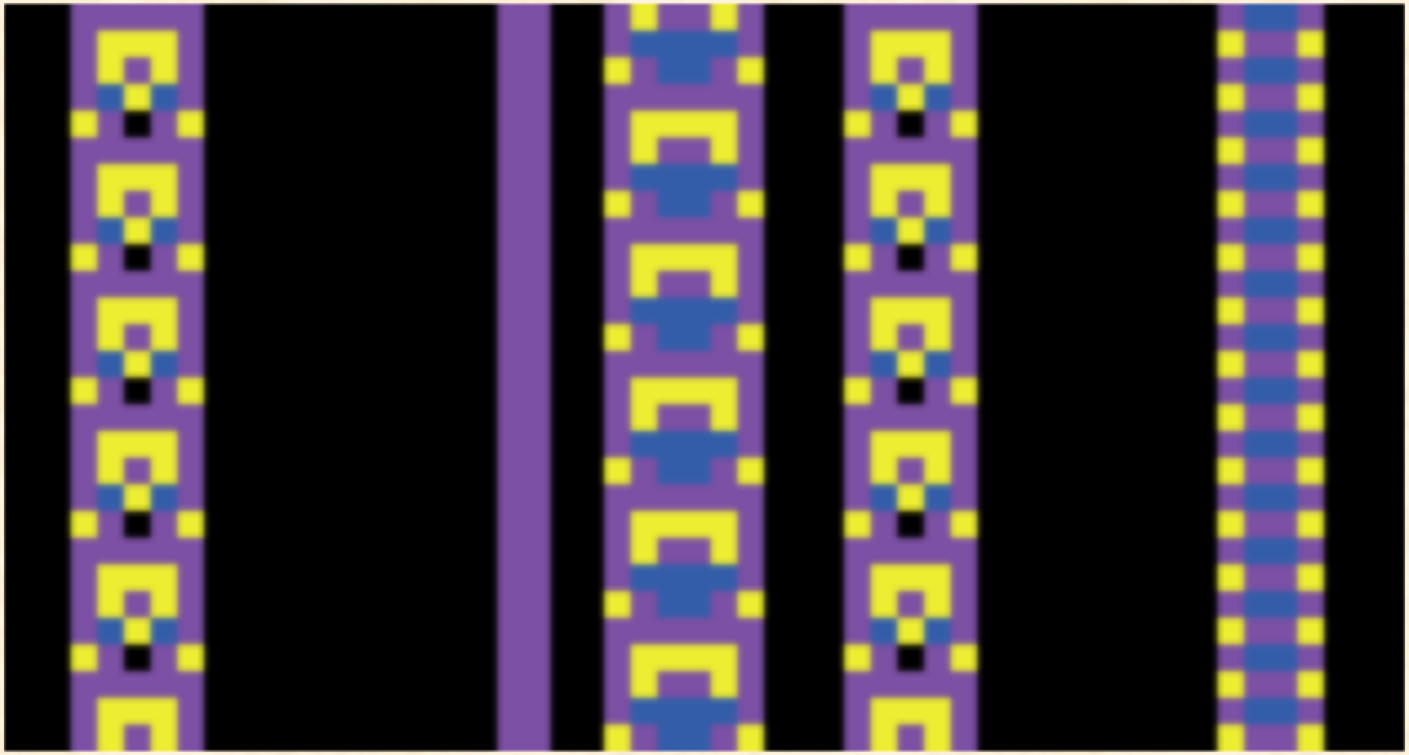


Class I ($\lambda = 0.3$) Closeup

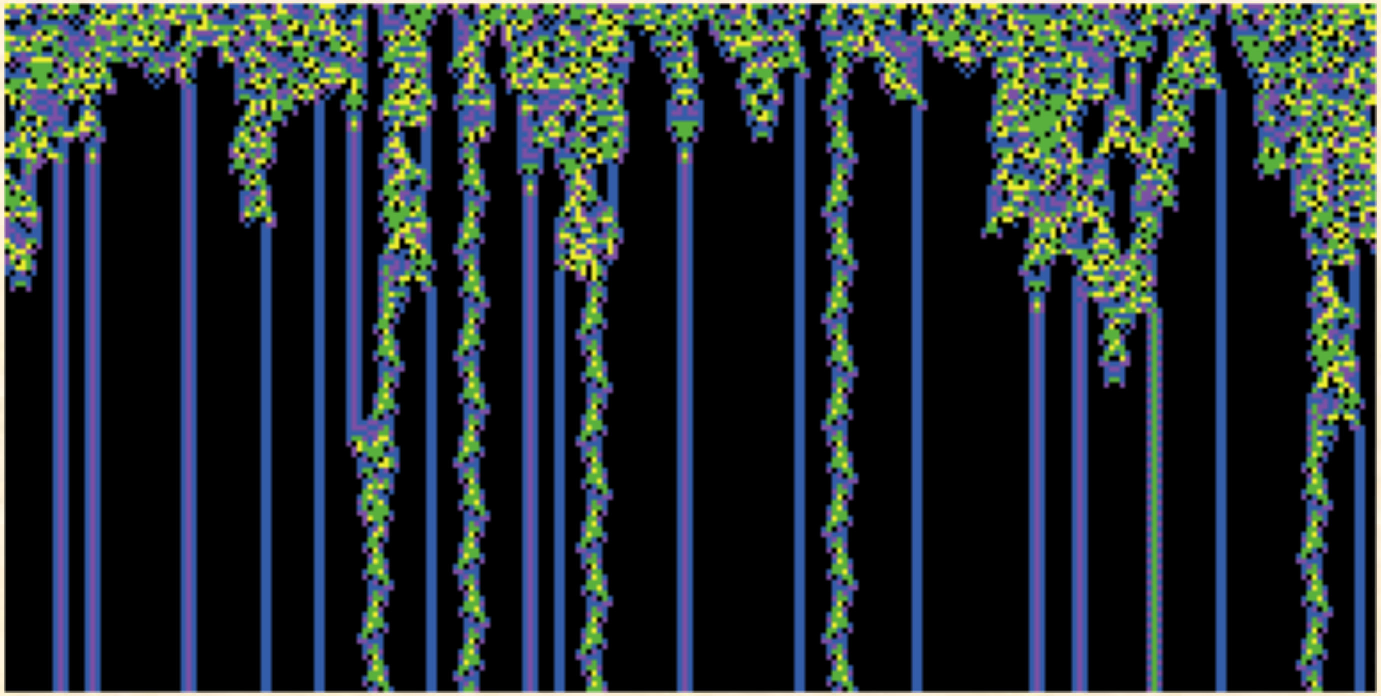
Class II ($\lambda = 0.66$)



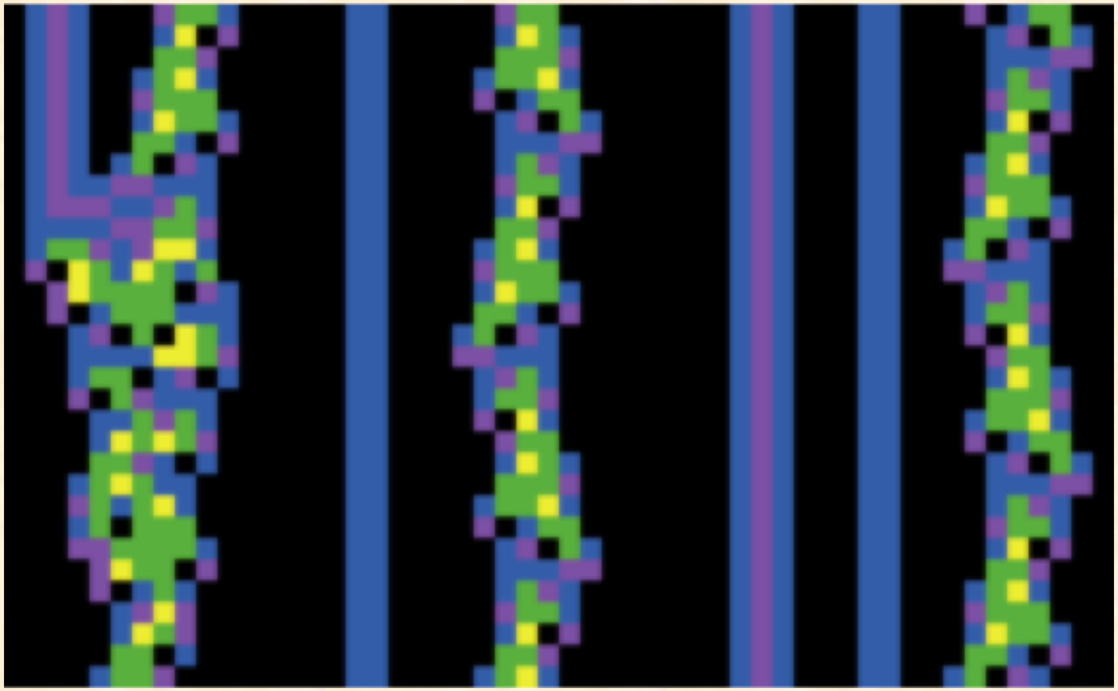
Class II ($\lambda = 0.66$) Closeup



Class II ($\lambda = 0.8$)

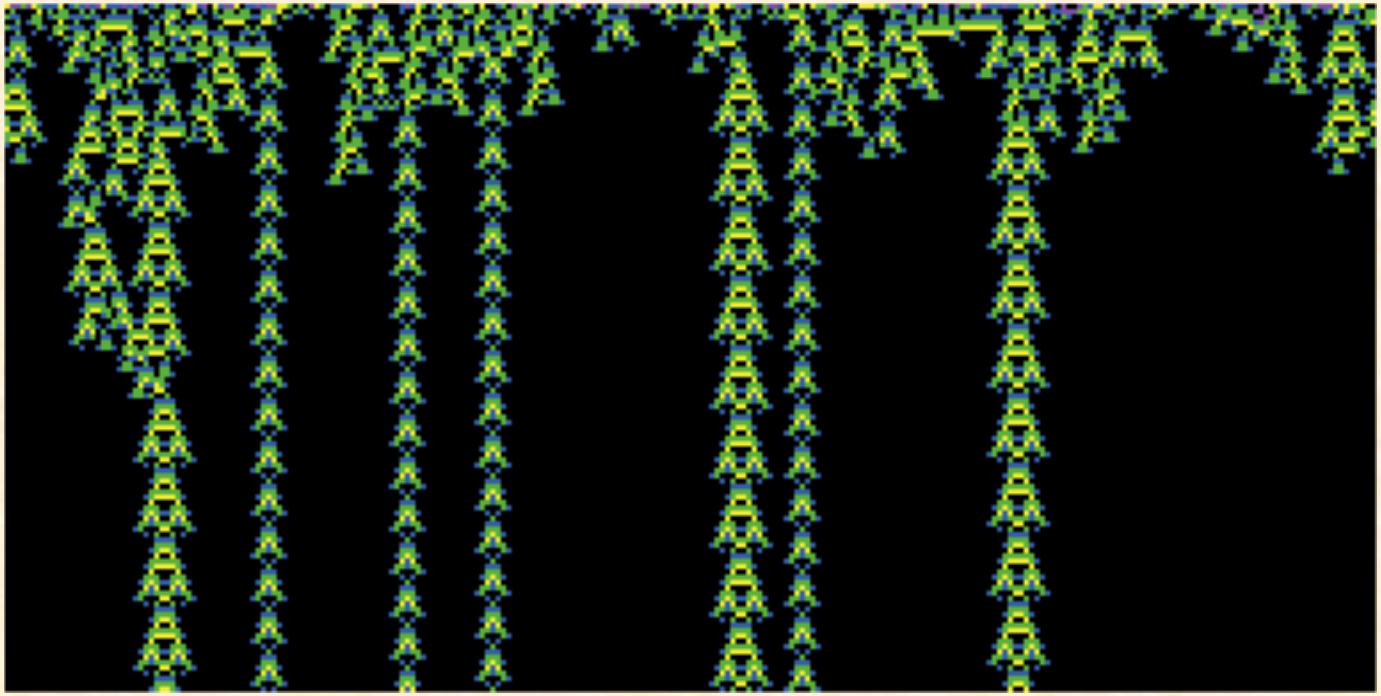


Class II ($\lambda = 0.8$) Closeup

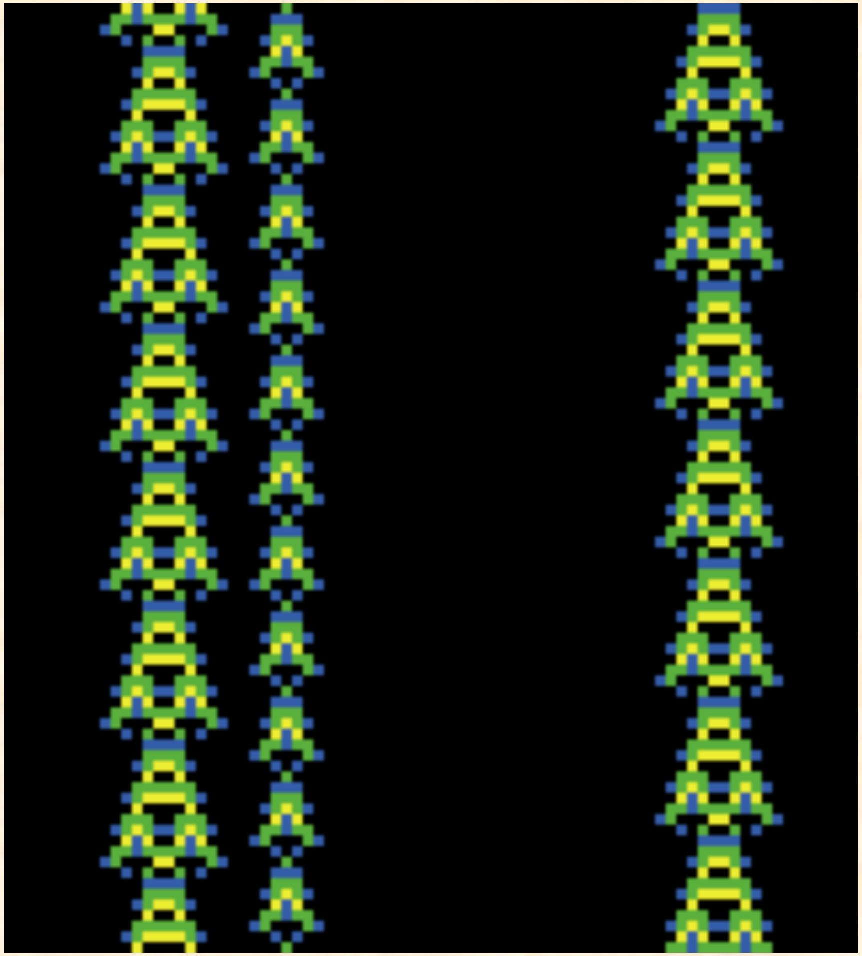


period = 20

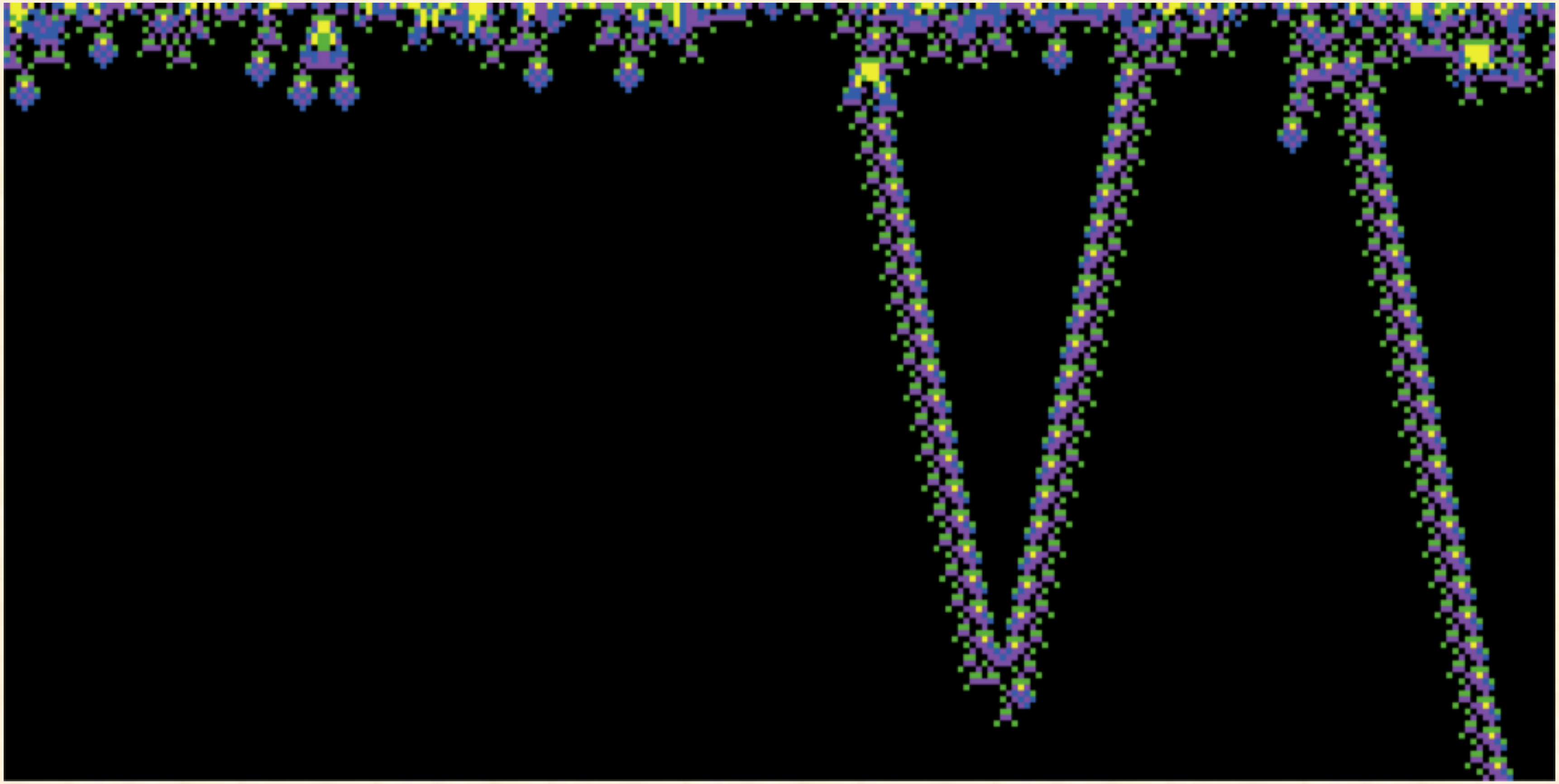
Class II ($\lambda = 0.5$)



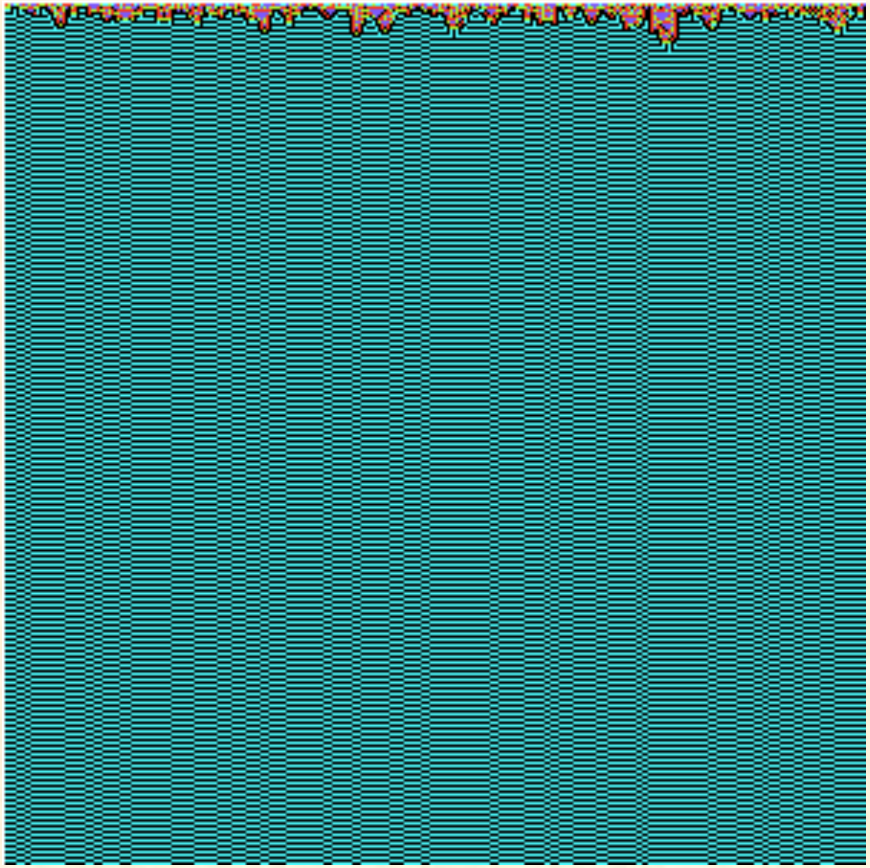
Class II ($\lambda = 0.5$) Closeup



Class II ($\lambda = 0.72$)

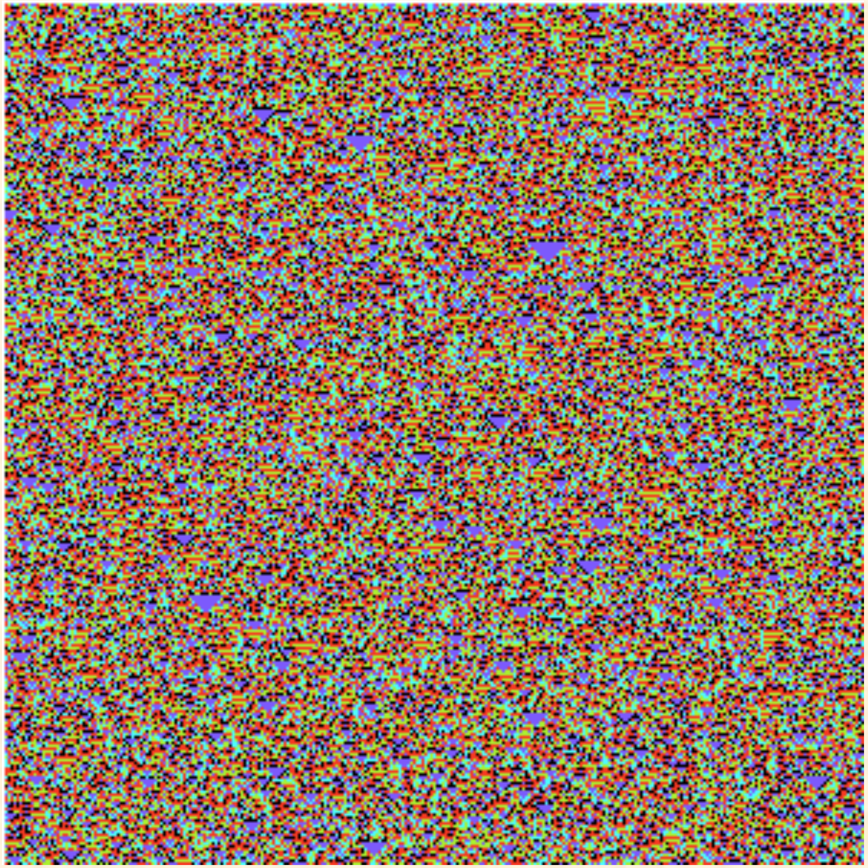


Class II ($\lambda = 0.31$)

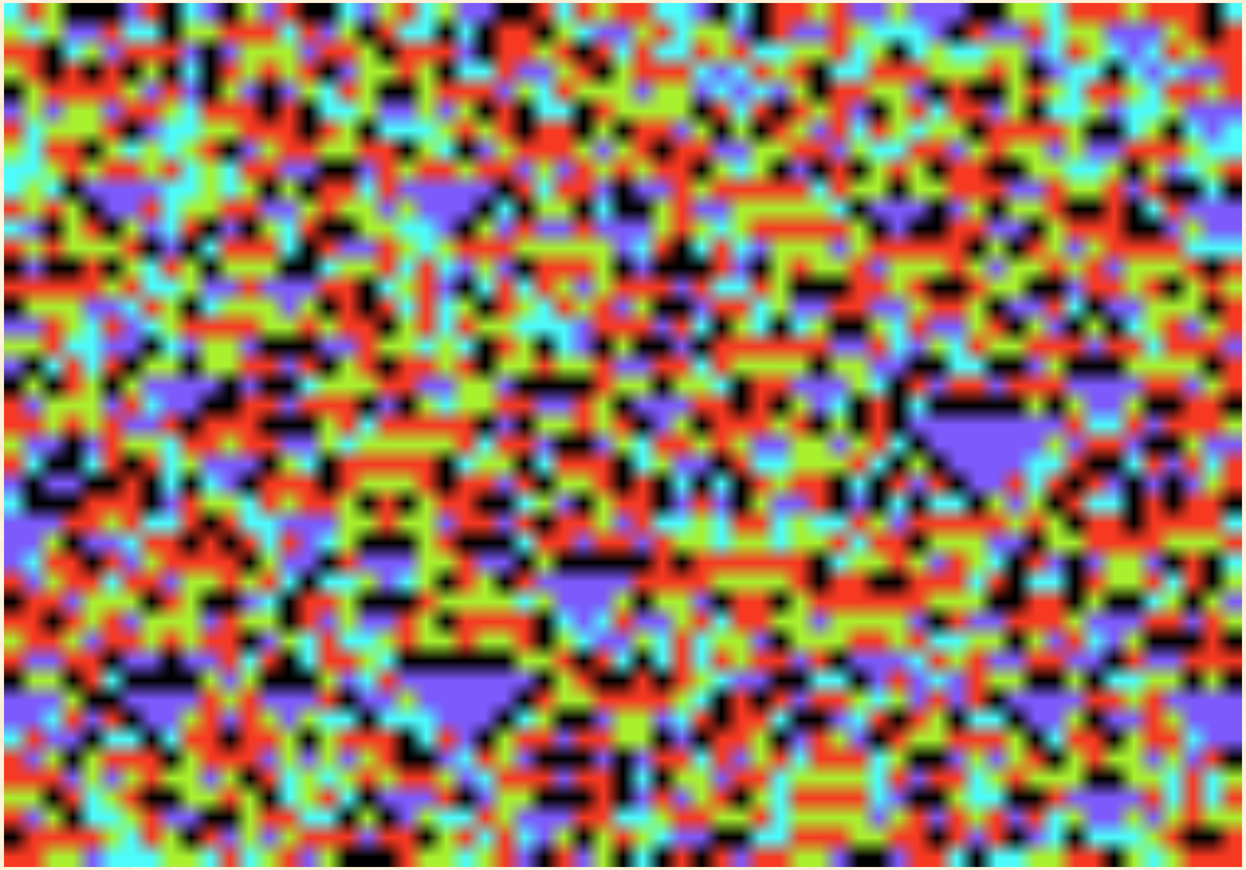


Class II ($\lambda = 0.31$) Closeup

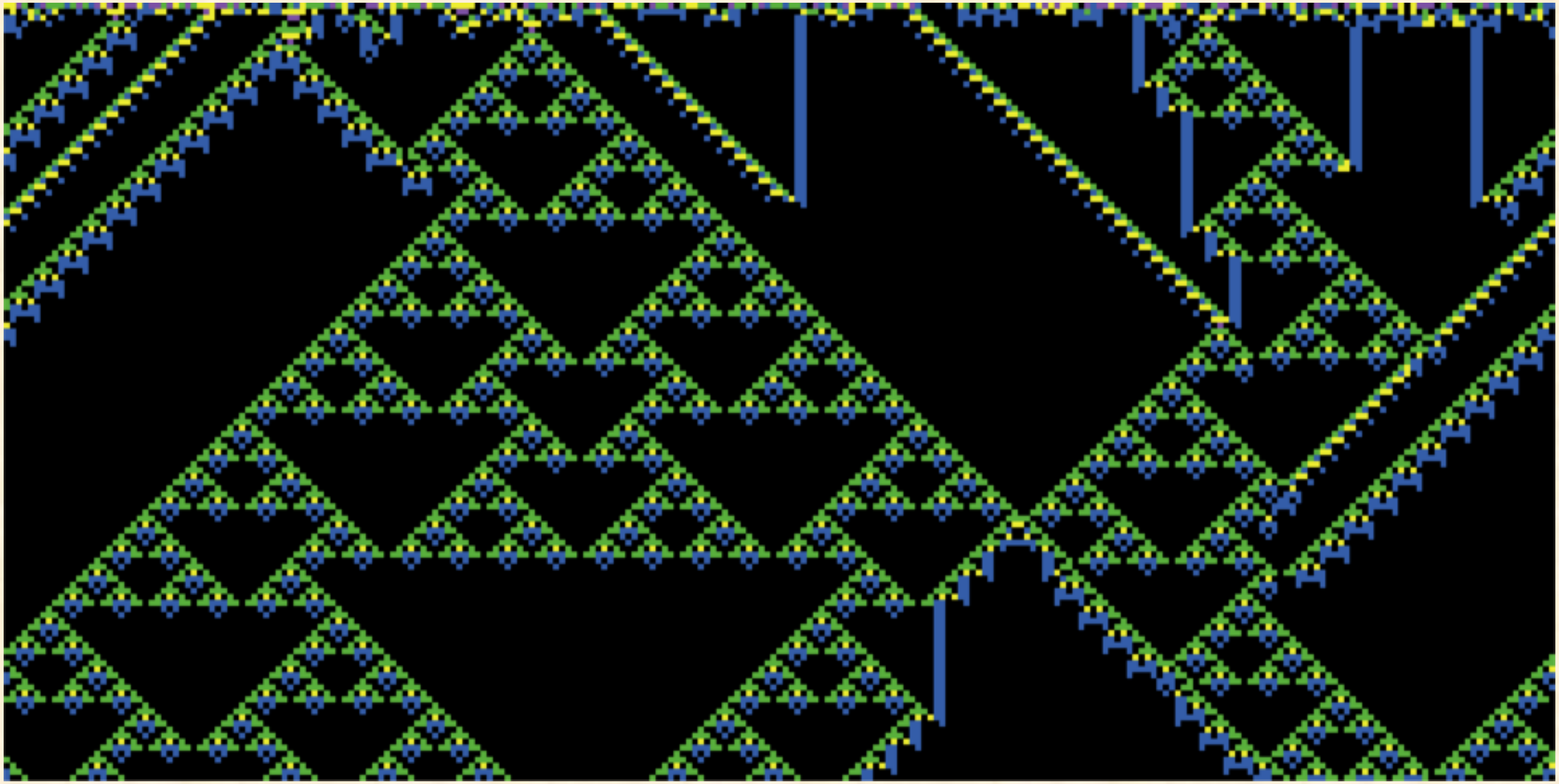
Class III ($\lambda = 0.5$)



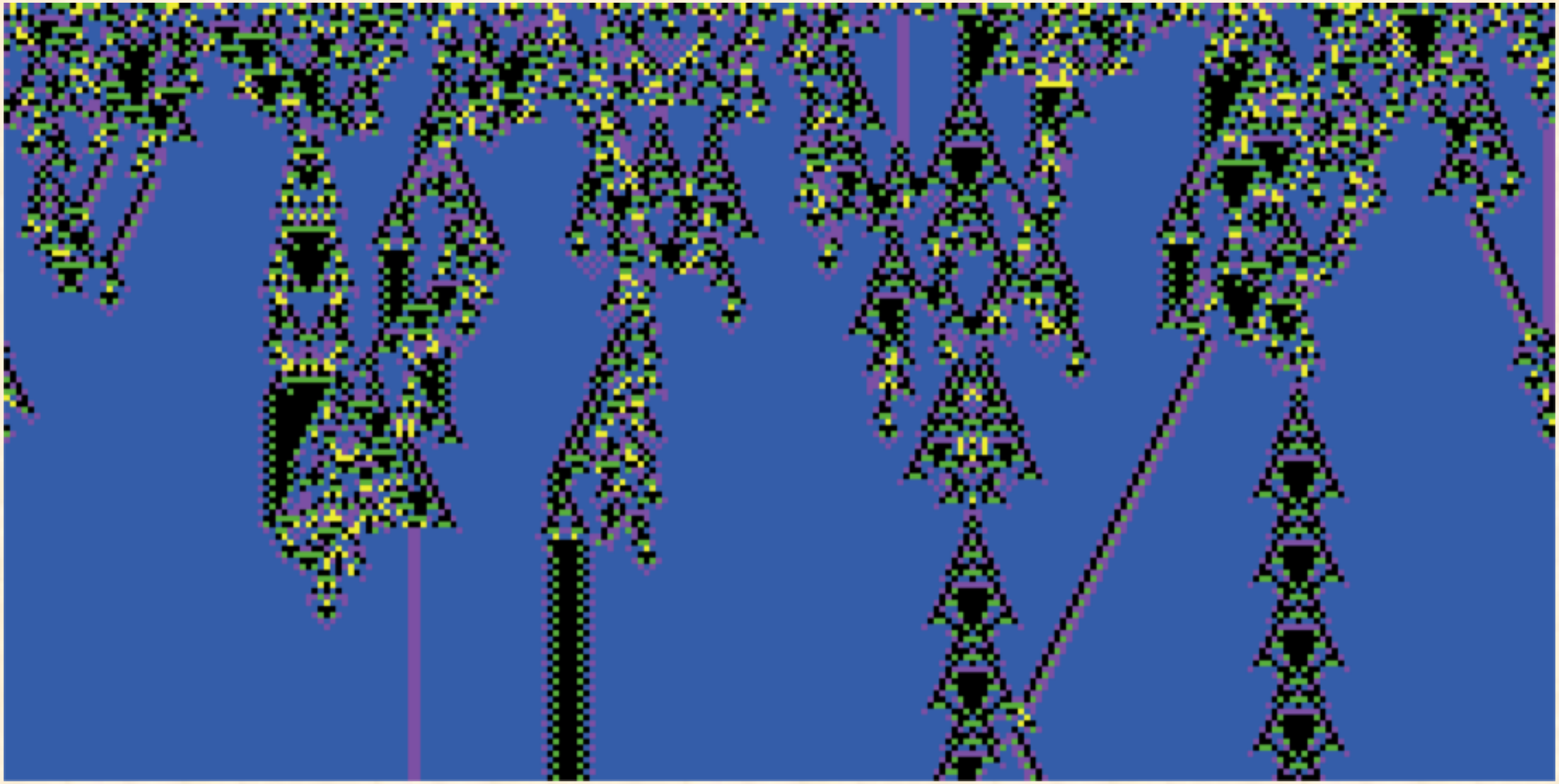
Class III ($\lambda = 0.5$) Closeup



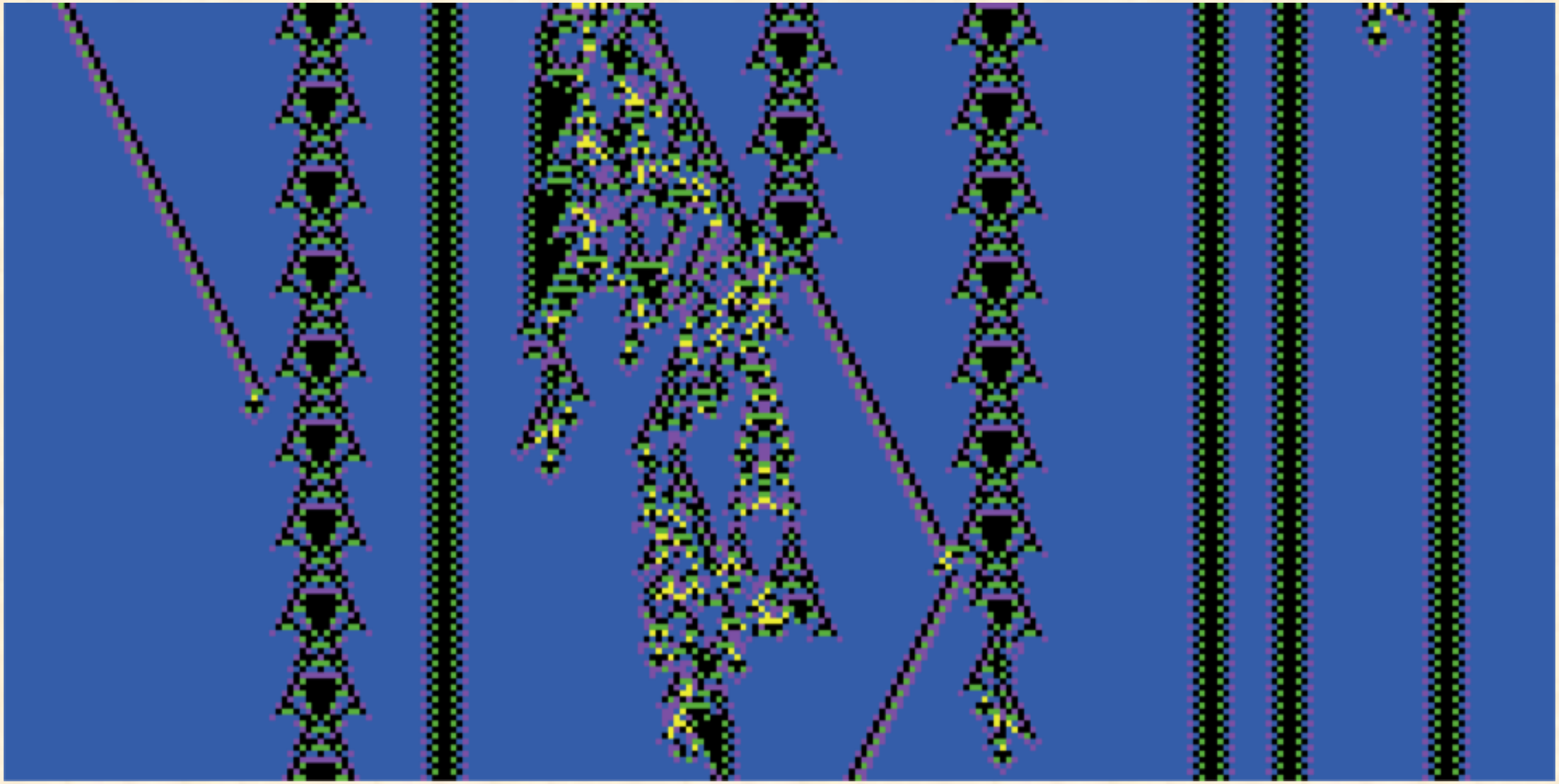
Class IV ($\lambda = 0.6$)



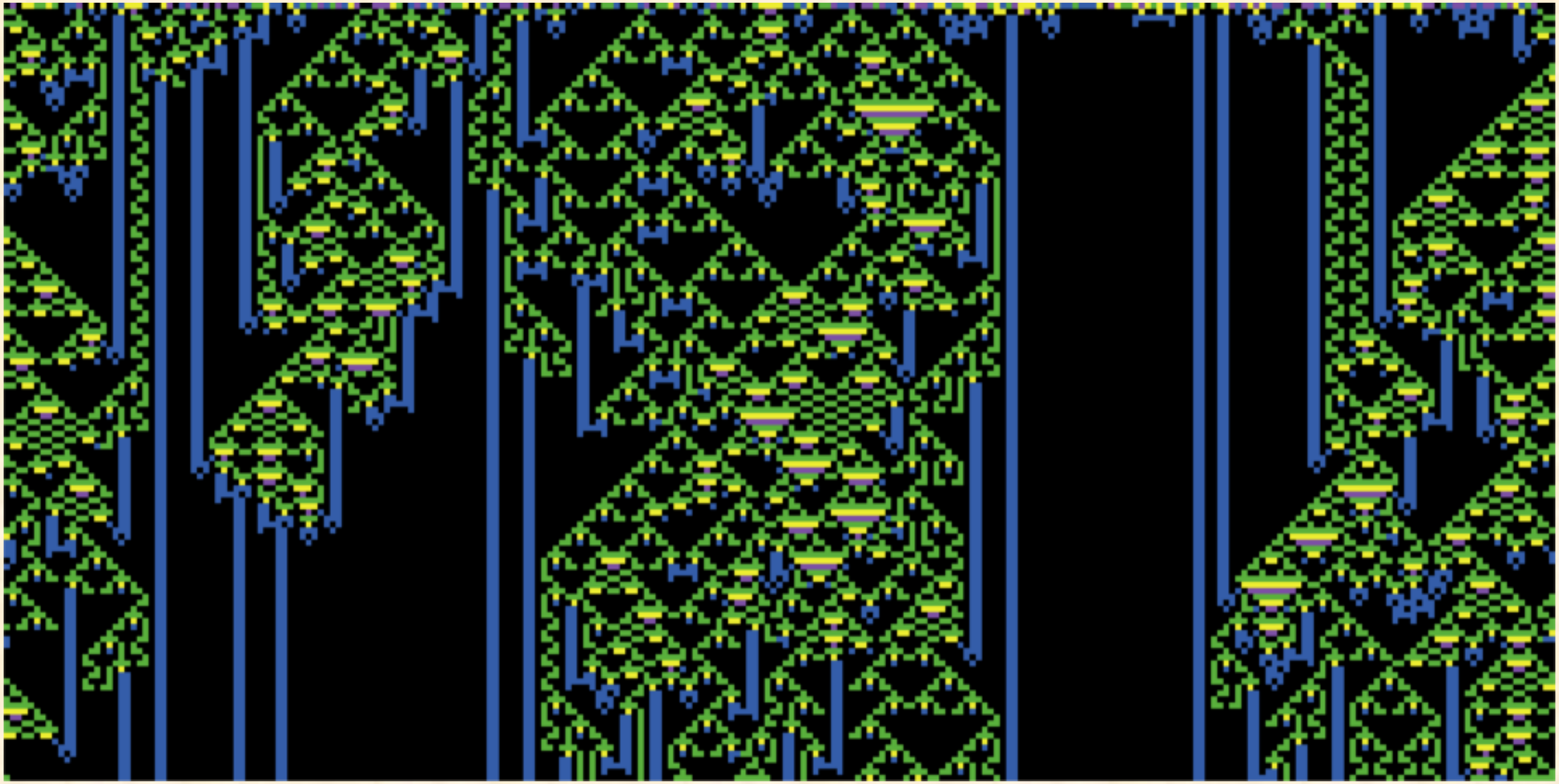
Class IV ($\lambda = 0.7$)



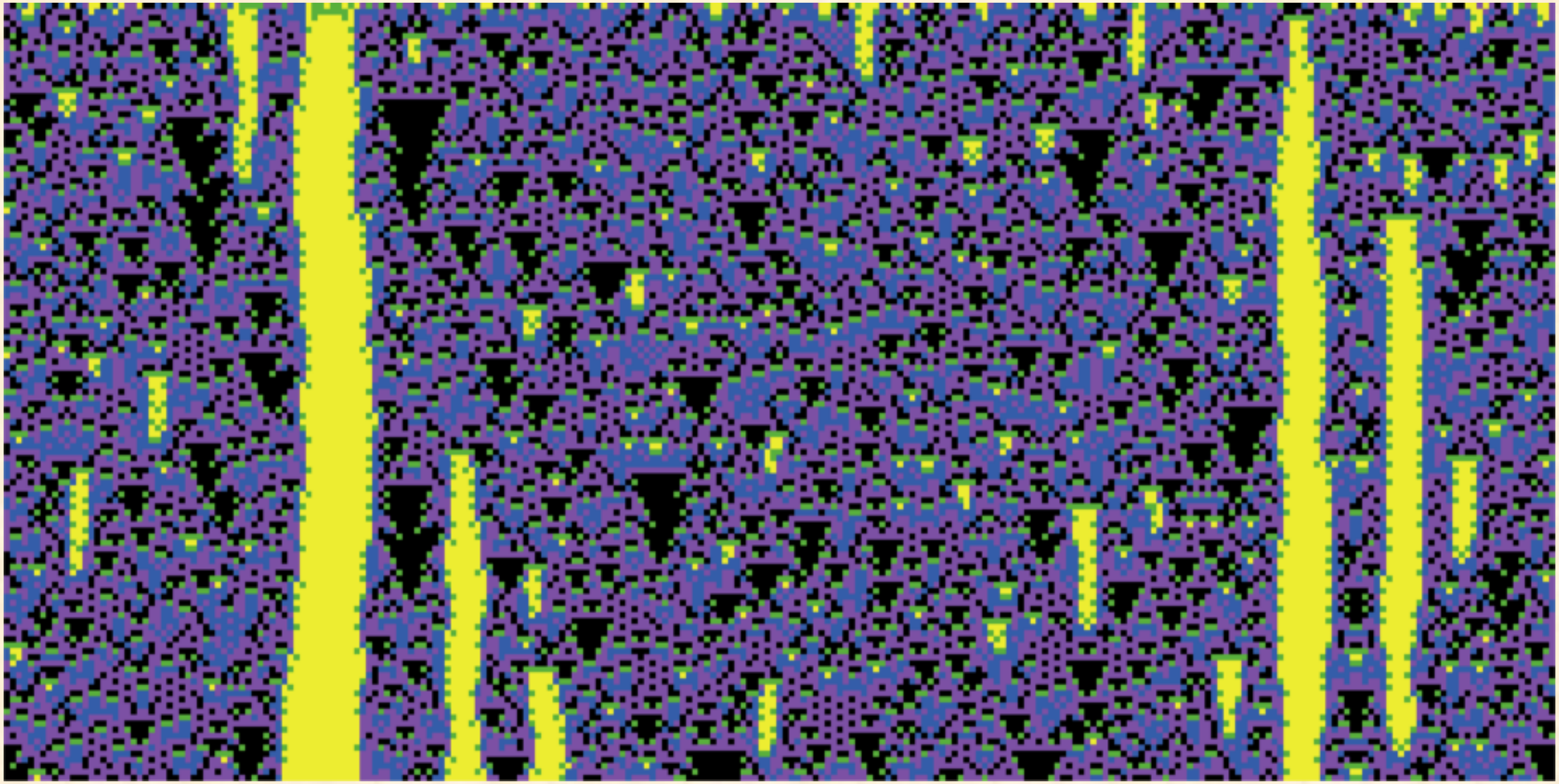
Class IV ($\lambda = 0.7$)



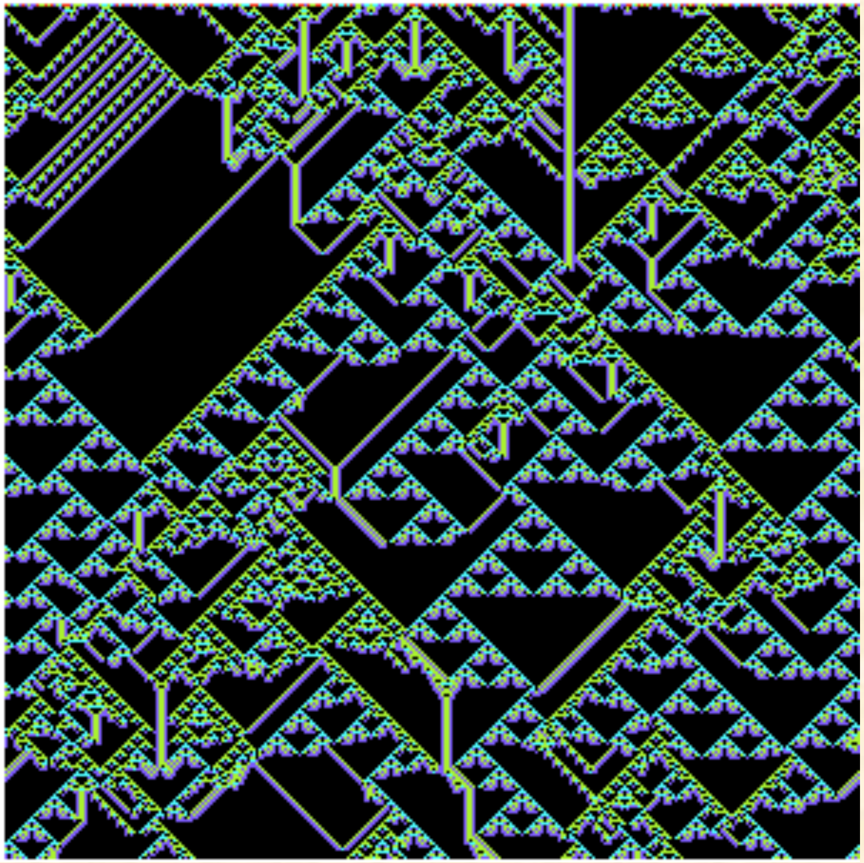
Class IV ($\lambda = 0.3$)

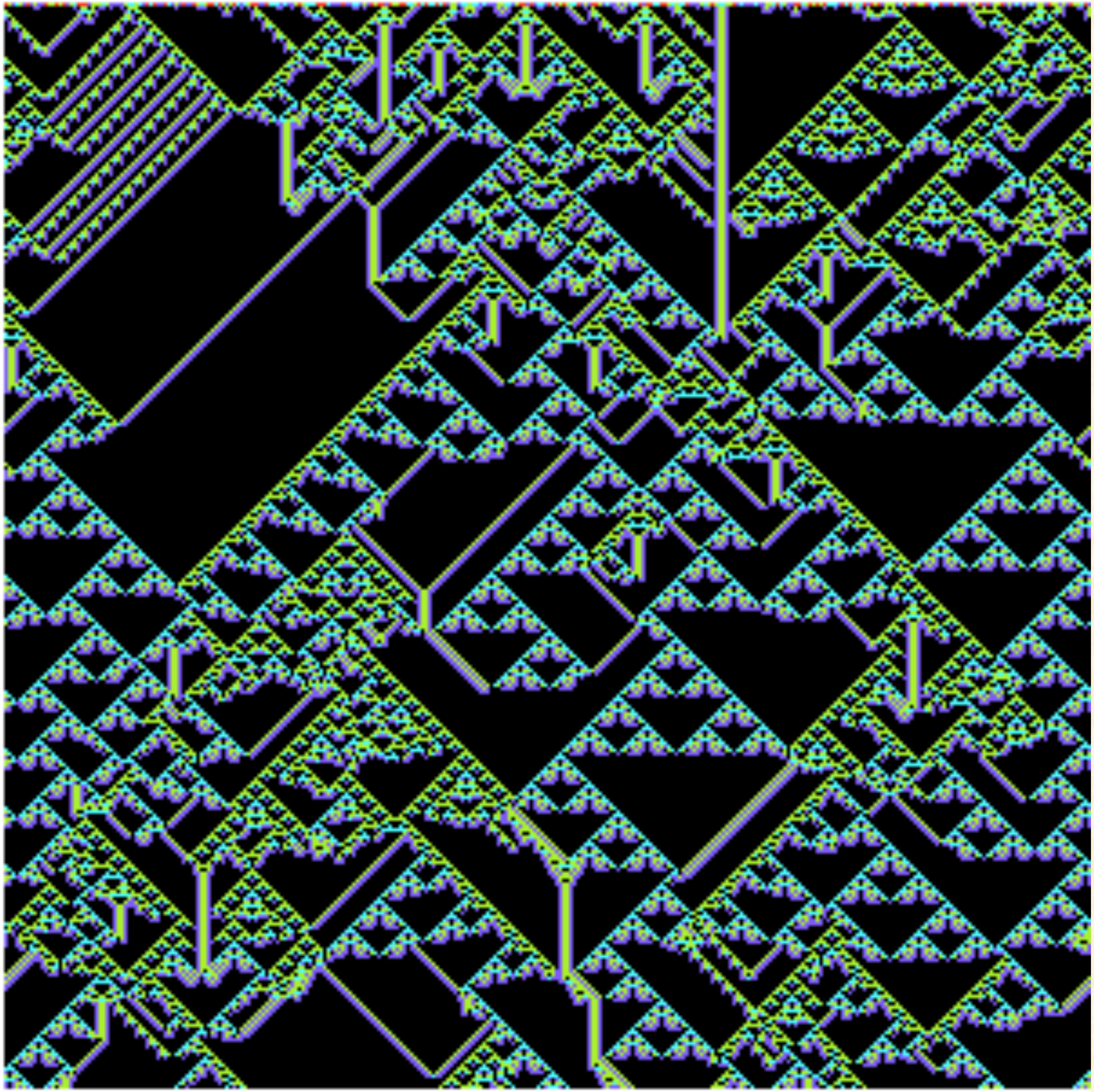


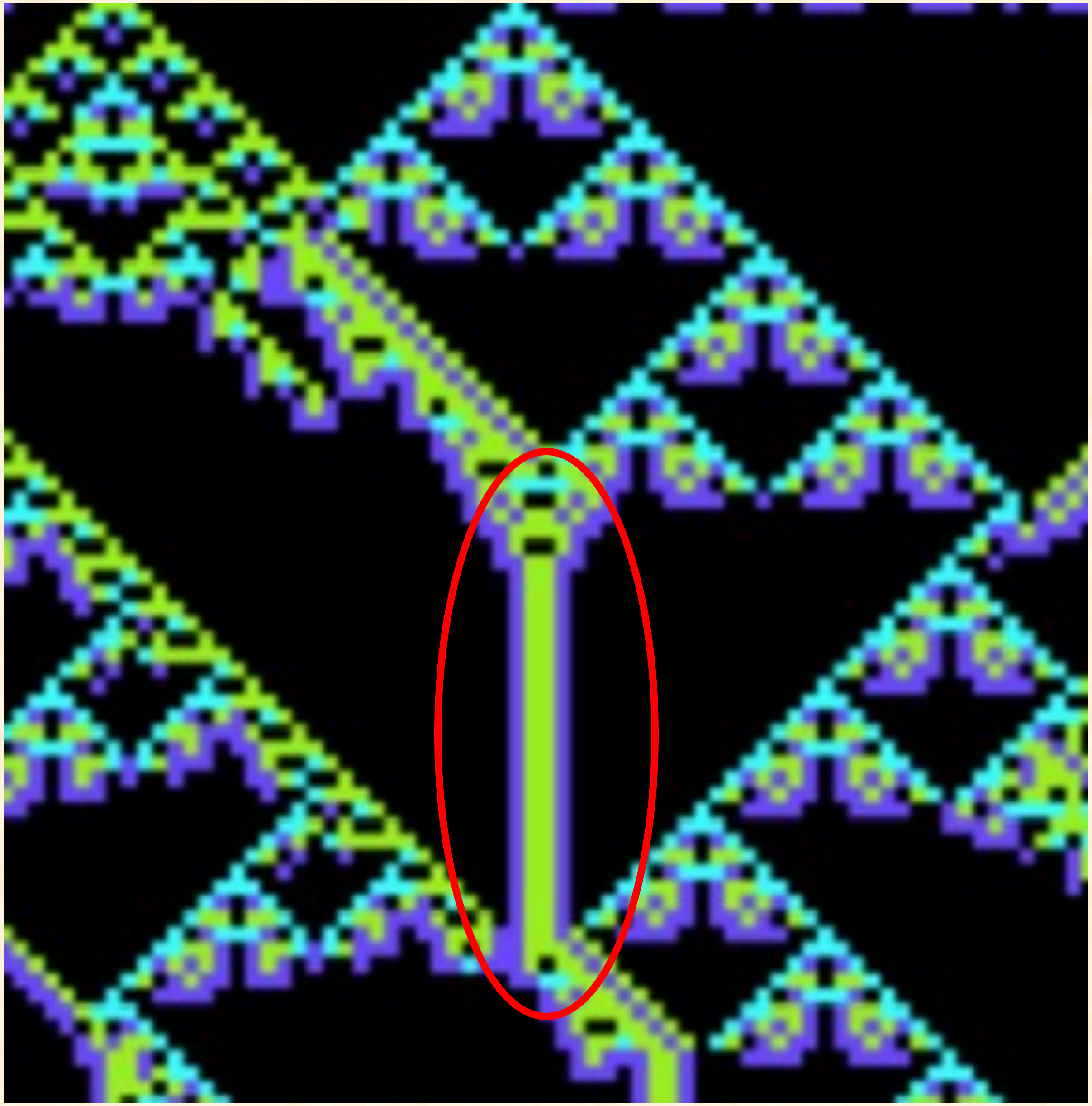
Class III–IV ($\lambda = 0.9$)

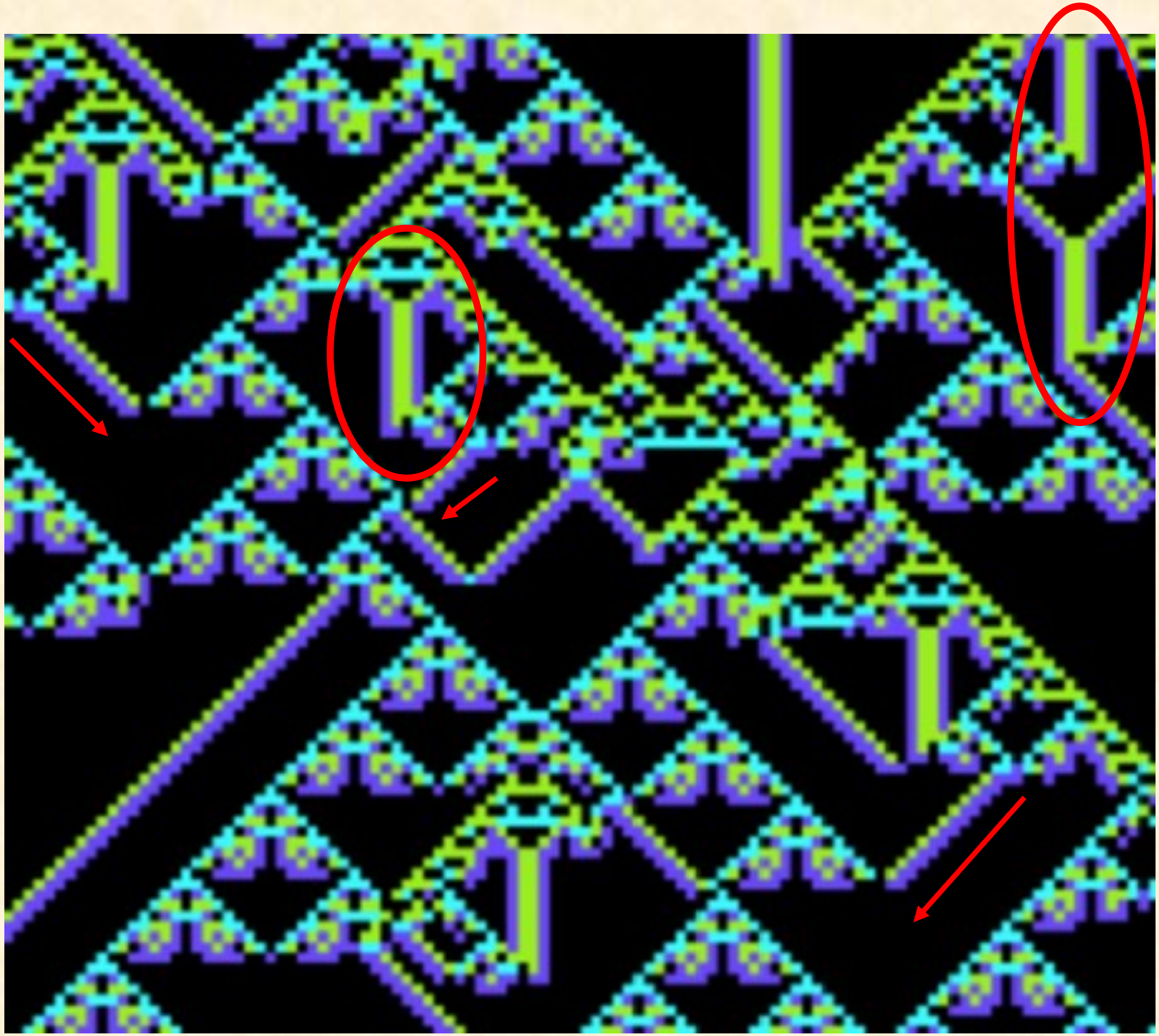


Class IV ($\lambda = 0.34$)









Class IV Shows Some of the Characteristics of Computation

- Persistent, but not perpetual storage
- Terminating cyclic activity
- Nonlocal transfer of control and information

A Computational Medium

- Storage of Information
- Transfer of Information
- Modification of Information

Class IV and Biology

- We expect biological material to exhibit Class IV behavior
- Stable
- But not too rigid
- Nonlocal coordination
- Solids, liquids, and "soft matter"

λ of Life

- For Life, $\lambda \approx 0.273$
- which is near the critical region for CAs with:

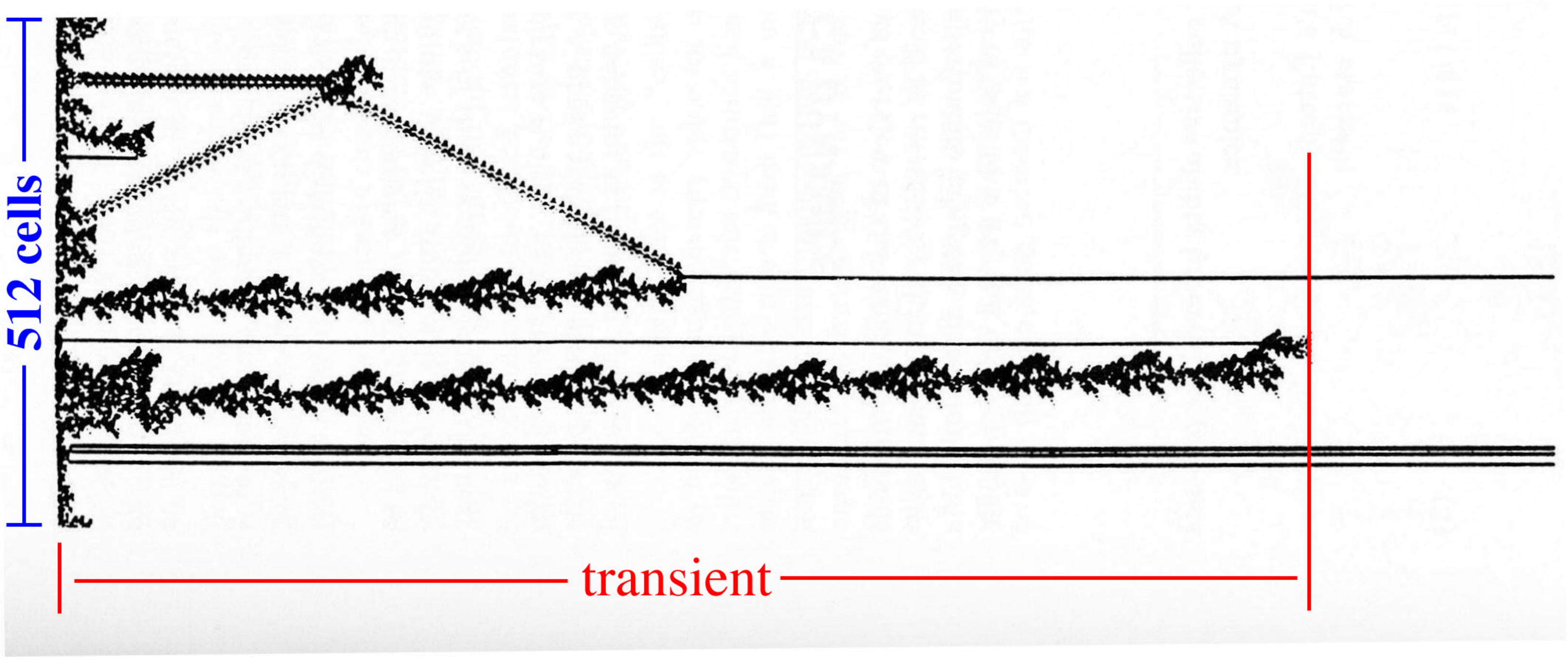
  $K = 2$

  $N = 9$

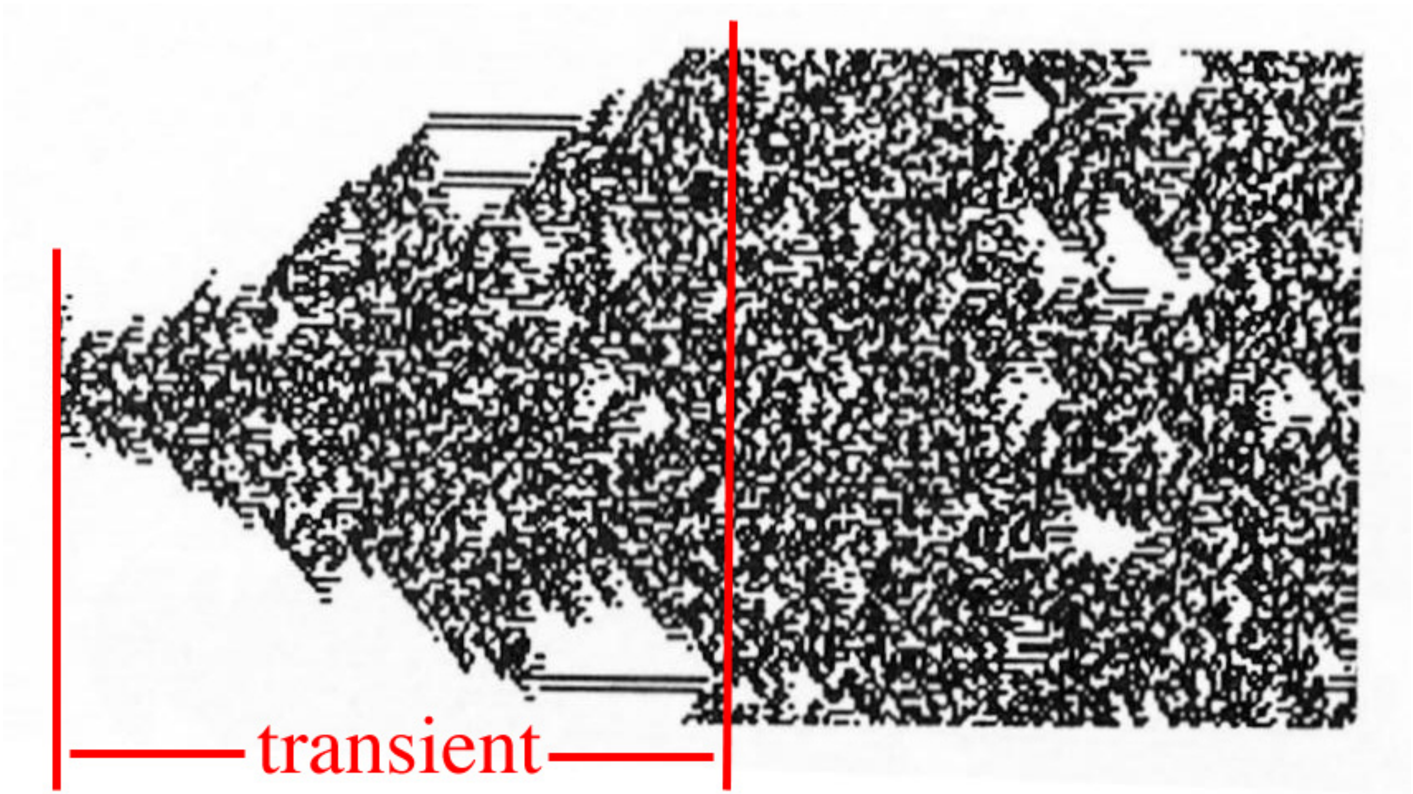
Project 1

- Investigation of relation between Wolfram classes, Langton's λ, and entropy in 1D CAs
- **Due Feb. 9**
- Information is on Canvas and course website (scroll down to "Projects / Assignments")
- Read it over and email questions or ask in class

Transient Length (I, II)



Transient Length (III)



Shannon Information
(very briefly!)

- Information varies directly with surprise
- Information varies inversely with probability
- Information is additive
- ∴The information content of a message is proportional to the negative log of its probability

$$I\{s\} = -\lg \Pr\{s\}$$

Entropy

- Suppose have source S of symbols from ensemble $\{s_1, s_2, \ldots, s_N\}$
- Average information per symbol:

$$\sum_{k=1}^{N} \Pr\{s_k\} I\{s_k\} = \sum_{k=1}^{N} \Pr\{s_k\} \left(-\lg \Pr\{s_k\}\right)$$

- This is the *entropy* of the source:

$$H\{S\} = -\sum_{k=1}^{N} \Pr\{s_k\} \lg \Pr\{s_k\}$$

Maximum and Minimum Entropy

- Maximum entropy is achieved when all signals are equally likely

  No ability to guess; maximum surprise

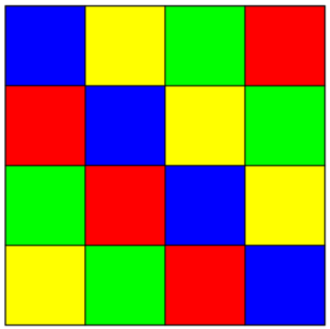
  $H_{\max} = \lg N$

- Minimum entropy occurs when one symbol is certain and the others are impossible

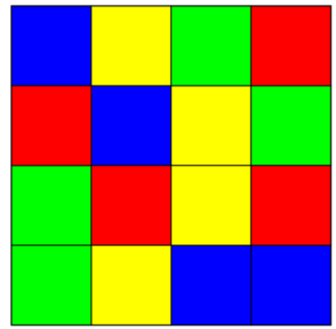
  No uncertainty; no surprise

  $H_{\min} = 0$

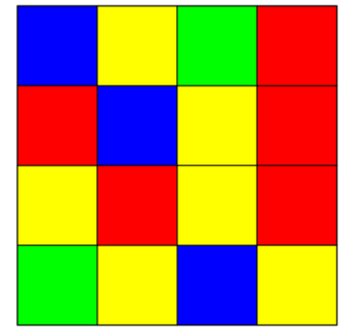
Entropy Examples



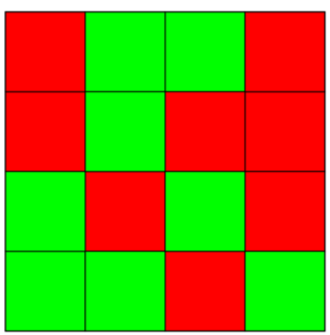
H = 2.0 bits



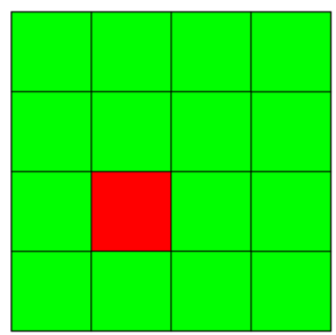
H = 2.0 bits



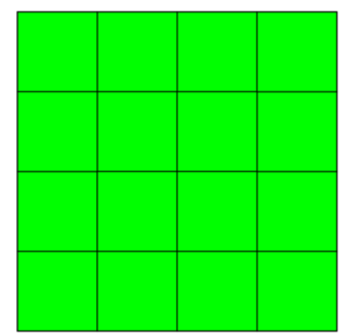
H = 1.9 bits



H = 1.0 bits



H = 0.3 bits



H = 0.0 bits

Entropy of Transition Table

- Among other things, entropy is a way to measure the uniformity of a distribution

$$H = -\sum_i p_i \lg p_i$$

- Distinction of quiescent state is arbitrary
- Let n_k = number mapping into state k
- Then $p_k = n_k / T$

$$H = \lg T - \frac{1}{T}\sum_{k=1}^{K} n_k \lg n_k$$

Entropy Range

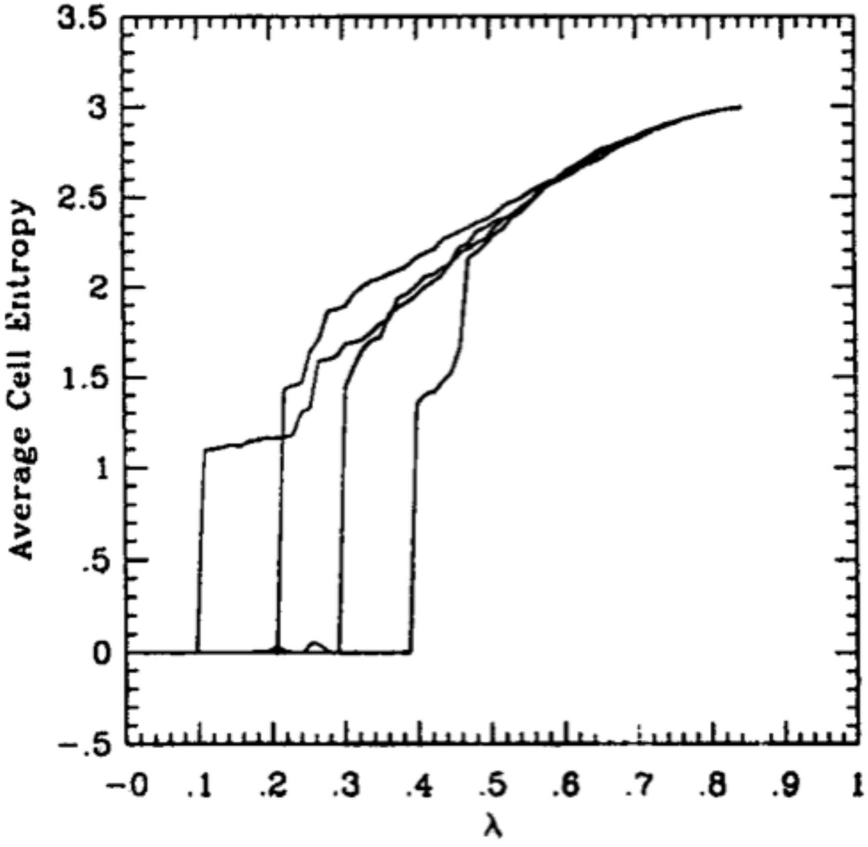
- Maximum entropy ($\lambda = 1 - 1/K$):
  uniform as possible
  all $n_k = T/K$
  $H_{\max} = \lg K$

- Minimum entropy ($\lambda = 0$ or $\lambda = 1$):
  non-uniform as possible
  one $n_s = T$
  all other $n_r = 0$ ($r \neq s$)
  $H_{\min} = 0$

Further Investigations by Langton

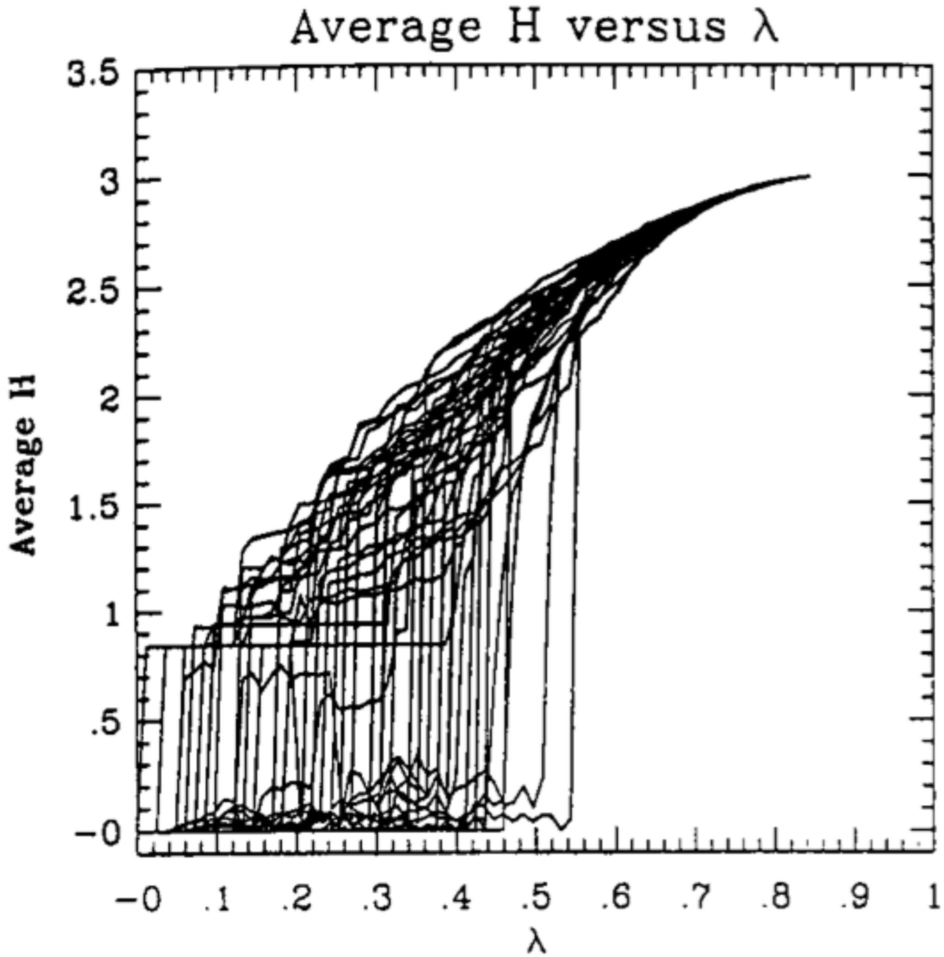
- 2-D CAs
- $K = 8$
- $N = 5$
- 64×64 lattice
- periodic boundary conditions
- measure average cell entropy
  - after 500 steps

Avg. Cell Entropy vs. λ (K=8, N=5)

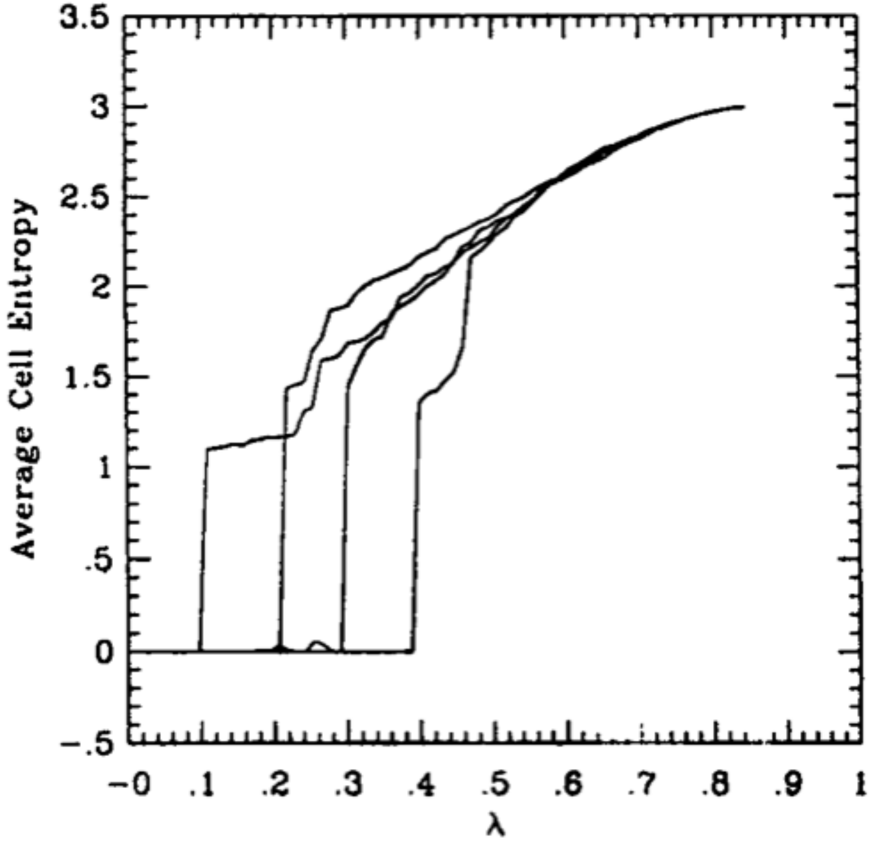


- $H = -\sum_{k=1}^{K} p_k \lg p_k$
- 1st order phase transition

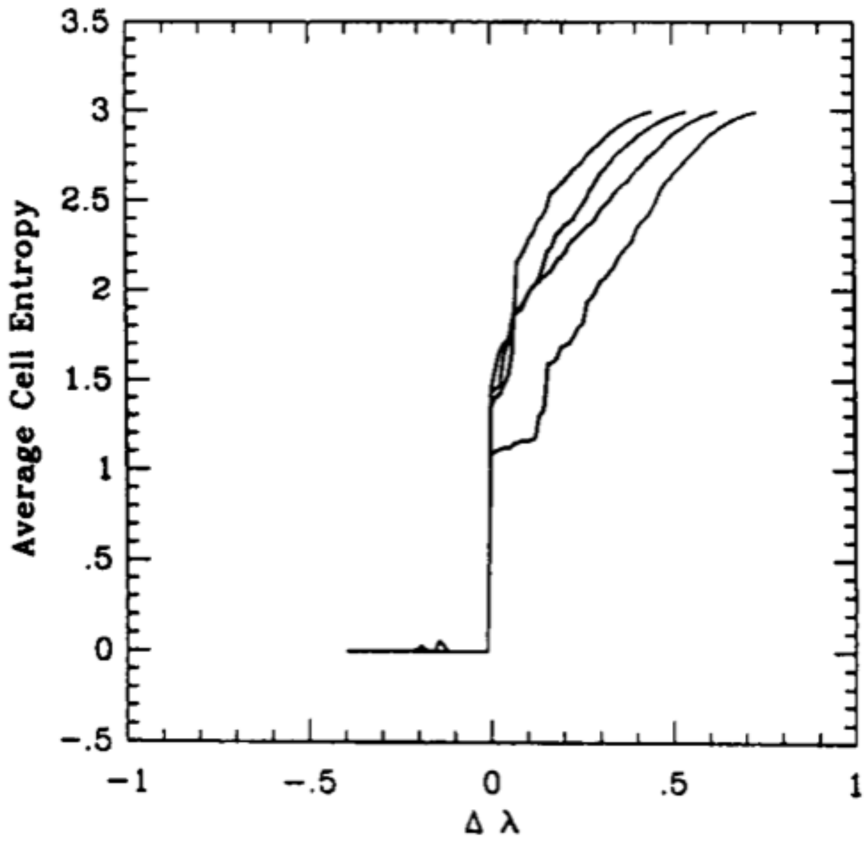
Avg. Cell Entropy vs. λ (*K*=8, *N*=5)



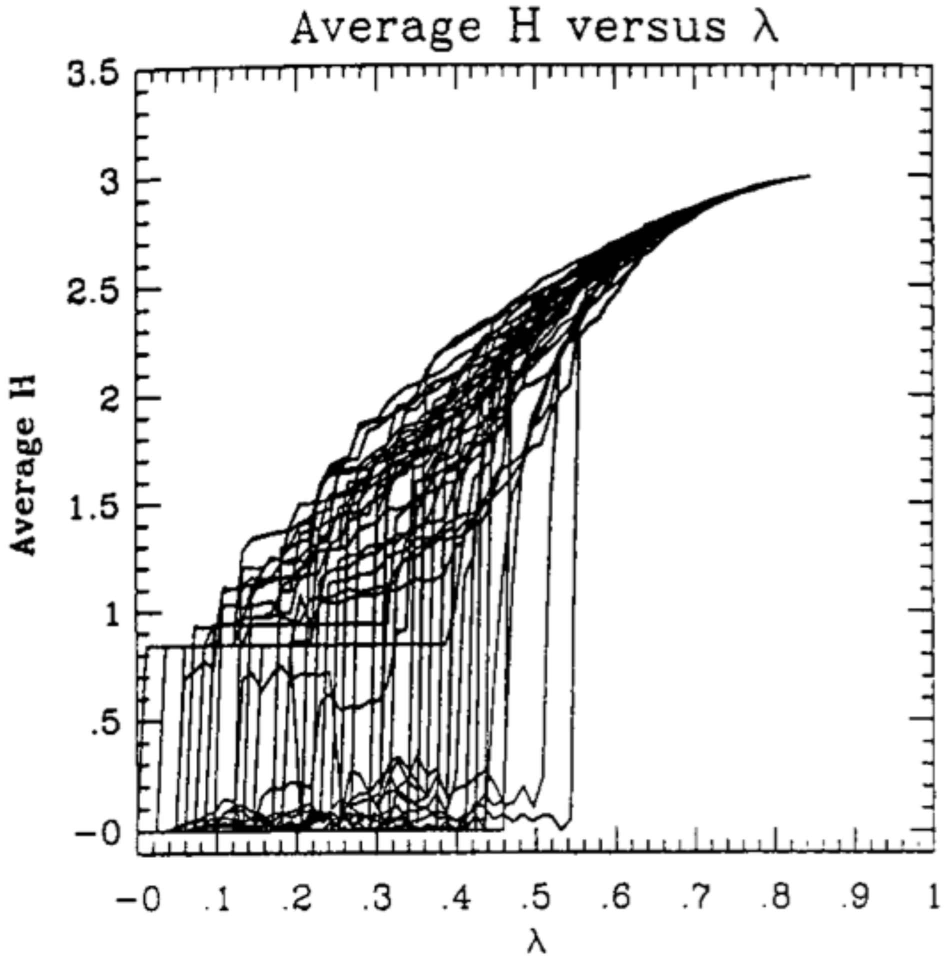
Avg. Cell Entropy vs. λ (K=8, N=5)



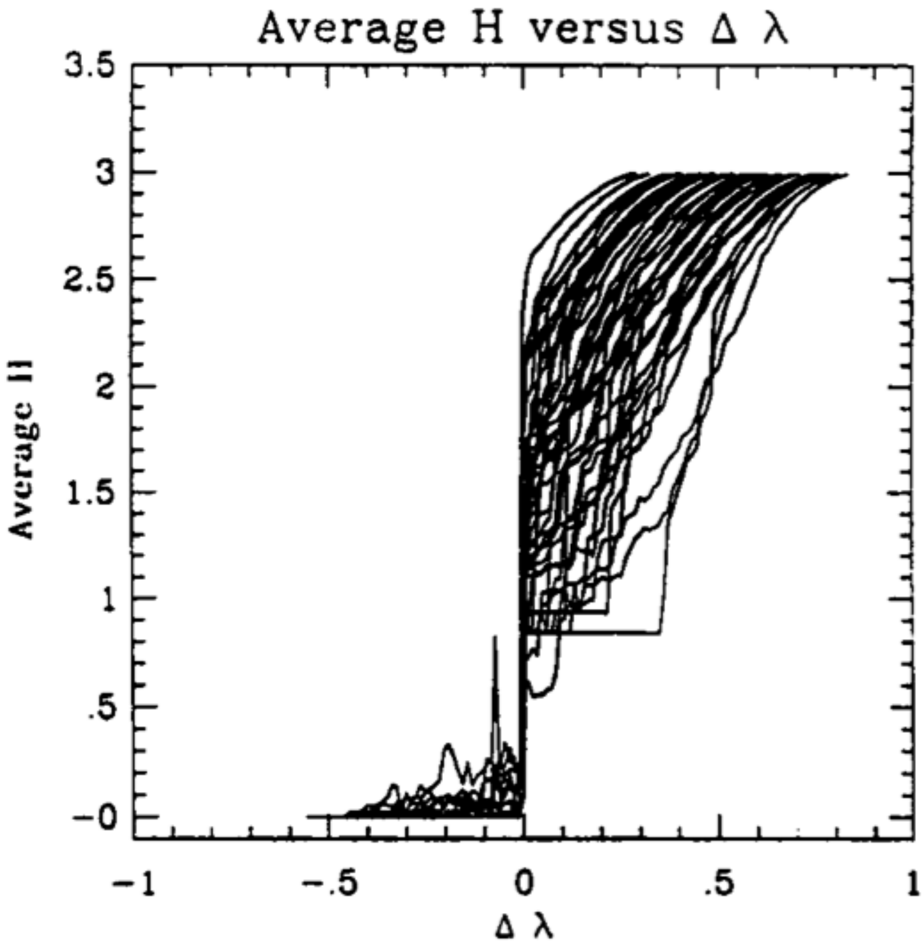
Avg. Cell Entropy vs. Δλ (K=8, N=5)



Avg. Cell Entropy vs. λ (*K*=8, *N*=5)



Avg. Cell Entropy vs. $\Delta\lambda$ (K=8, N=5)



Entropy of Independent Systems

- Suppose sources A and B are independent
- Let $p_j = \Pr\{a_j\}$ and $q_k = \Pr\{b_k\}$
- Then $\Pr\{a_j, b_k\} = \Pr\{a_j\} \Pr\{b_k\} = p_j q_k$

$$H(A,B) = -\sum_{j,k} \Pr(a_j, b_k) \lg \Pr(a_j, b_k)$$

$$= -\sum_{j,k} p_j q_k \lg(p_j q_k) = -\sum_{j,k} p_j q_k (\lg p_j + \lg q_k)$$

$$= -\sum_{j} p_j \lg p_j - \sum_{k} q_k \lg q_k = H(A) + H(B)$$

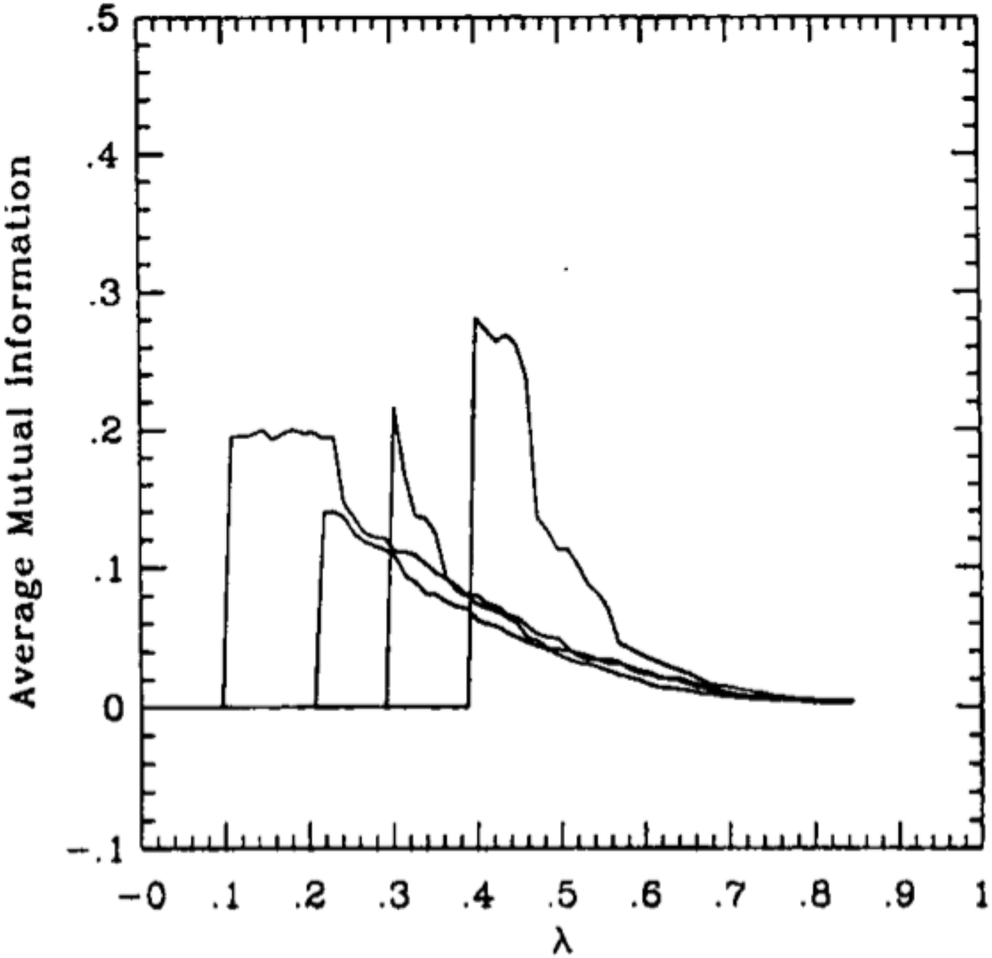
Mutual Information

- *Mutual information* measures the degree to which two sources are not independent
- A measure of their correlation

$$I(A,B) = H(A) + H(B) - H(A,B)$$

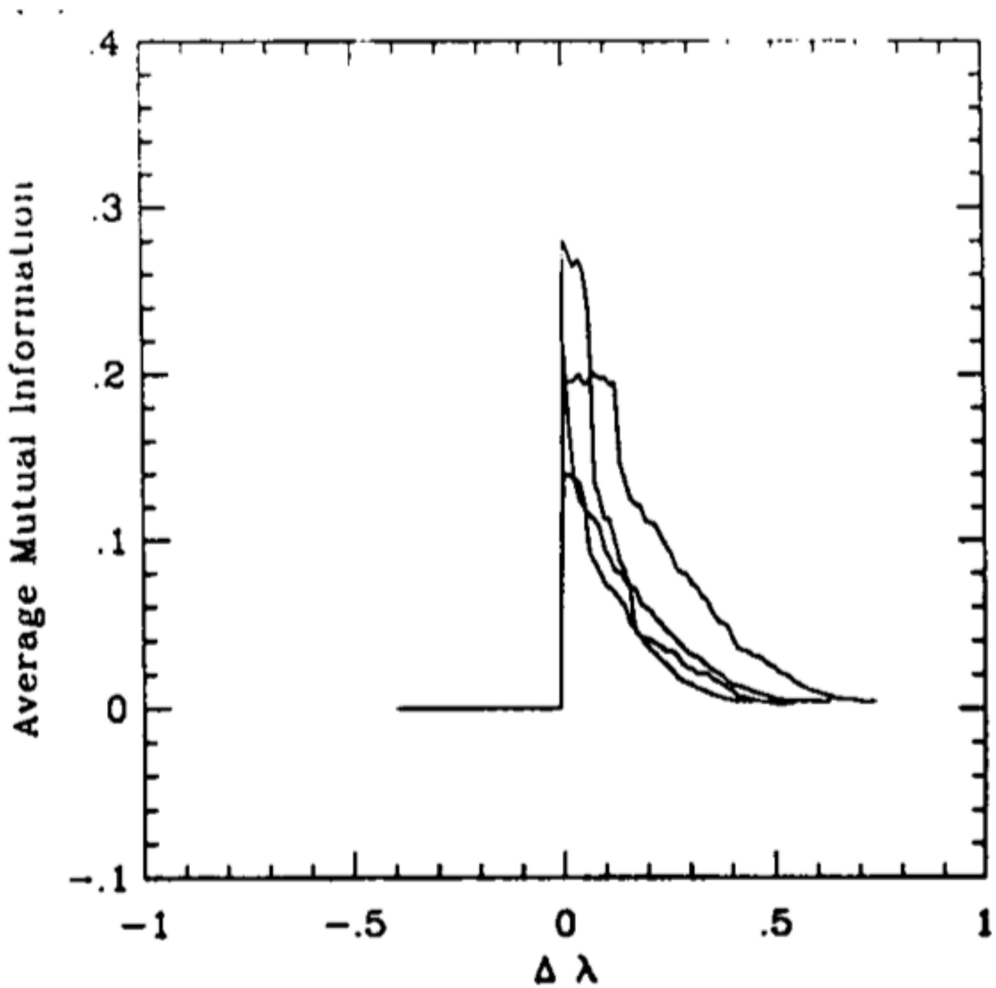
- $I(A,B) = 0$ for completely independent sources
- $I(A,B) = H(A) = H(B)$ for completely correlated sources

Avg. Mutual Info vs. λ (K=4, N=5)

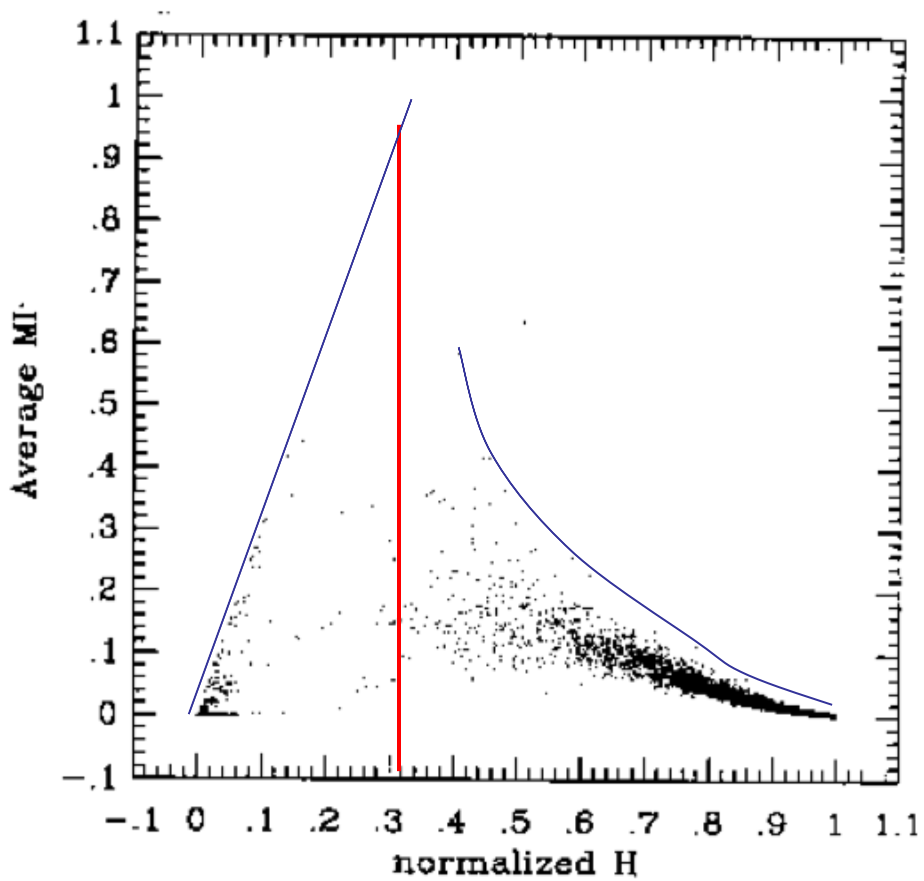


$I(A,B) =$
$H(A) + H(B)$
$- H(A,B)$

Avg. Mutual Info vs. Δ λ
(K=4, N=5)



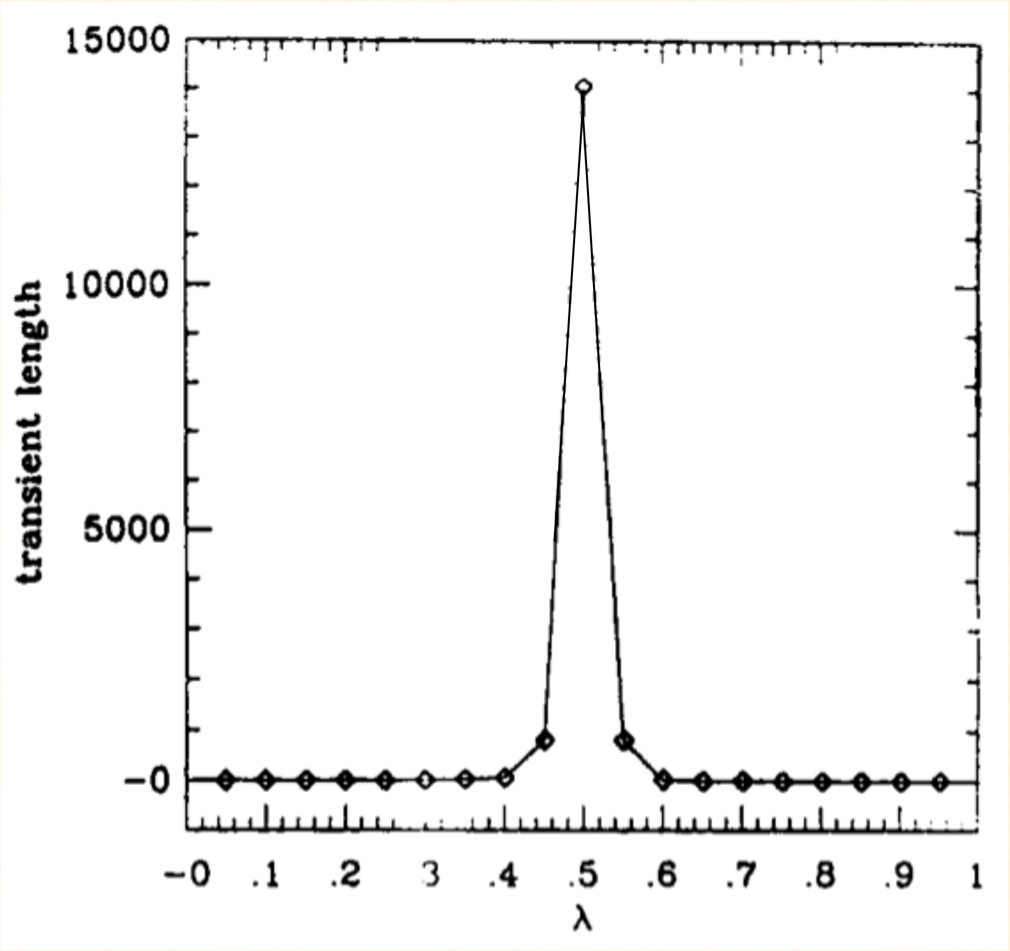
Mutual Information vs. Normalized Cell Entropy



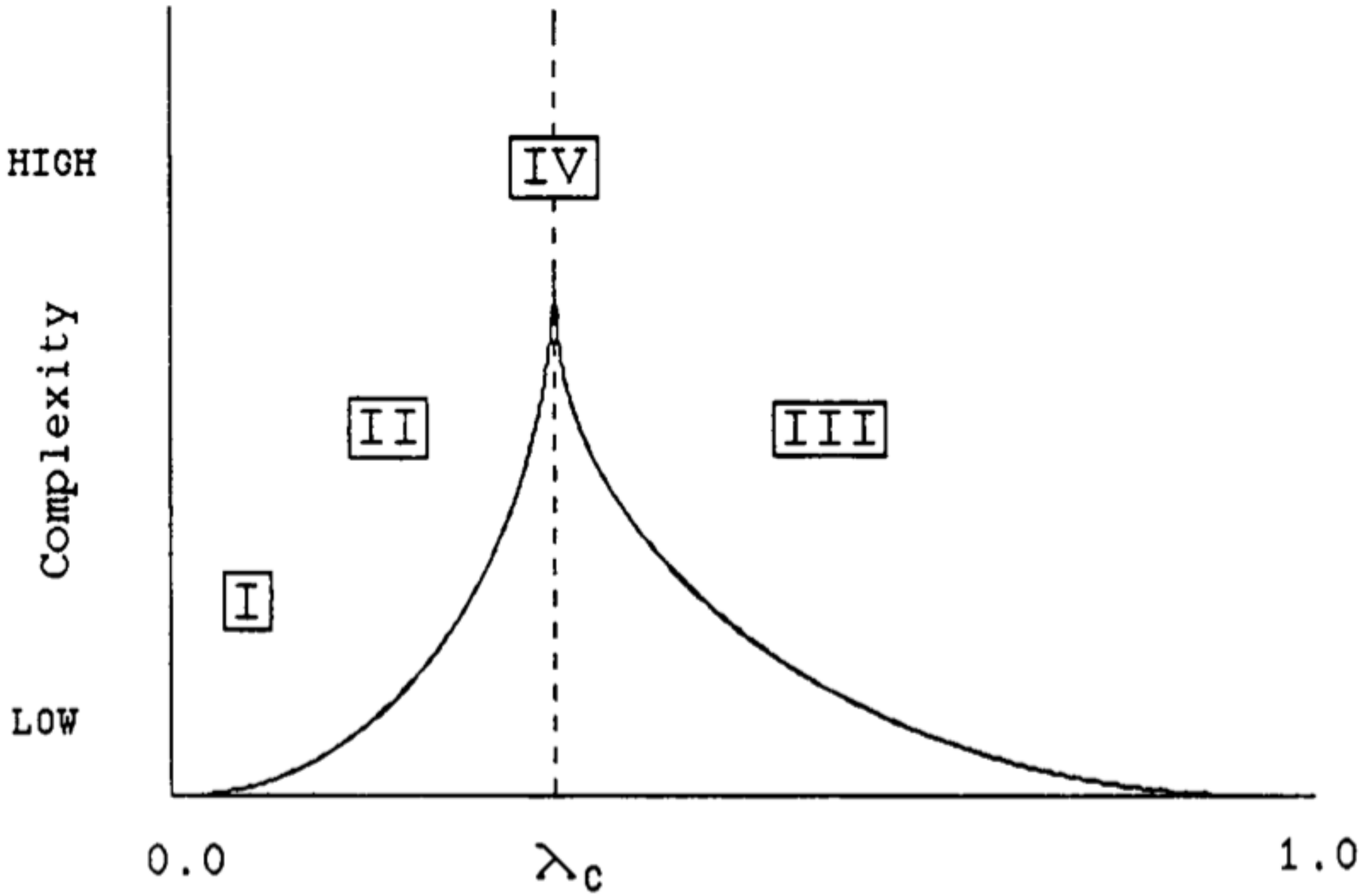
Critical Entropy Range

- Information *storage* involves *lowering* entropy
- Information *transmission* involves *raising* entropy
- Information *processing* requires a tradeoff between low and high entropy

Avg. Transient Length vs. λ (K=4, N=5)



Complexity vs. λ



Phase Transitions

- First-order phase transitions
  - Change (first derivative) is discontinuous
- Second-order (continuous) phase transitions
  - Change (first derivative) is continuous, but second derivative is discontinuous
  - Infinite correlation lengths
  - Critical slowing (long transients)
  - Statistical measure converge poorly (wide distributions)

Computation and Second-order Phase Transitions

- Long transients and long correlation lengths
- Difficulty predicting ultimate state (halting problem)
- Computation requires information storage and transmission
  - correlation too weak ⇒ independent sites ⇒ little transmission
  - correlation too strong ⇒ distant sites mimic each other

Schematic of CA Rule Space vs. λ

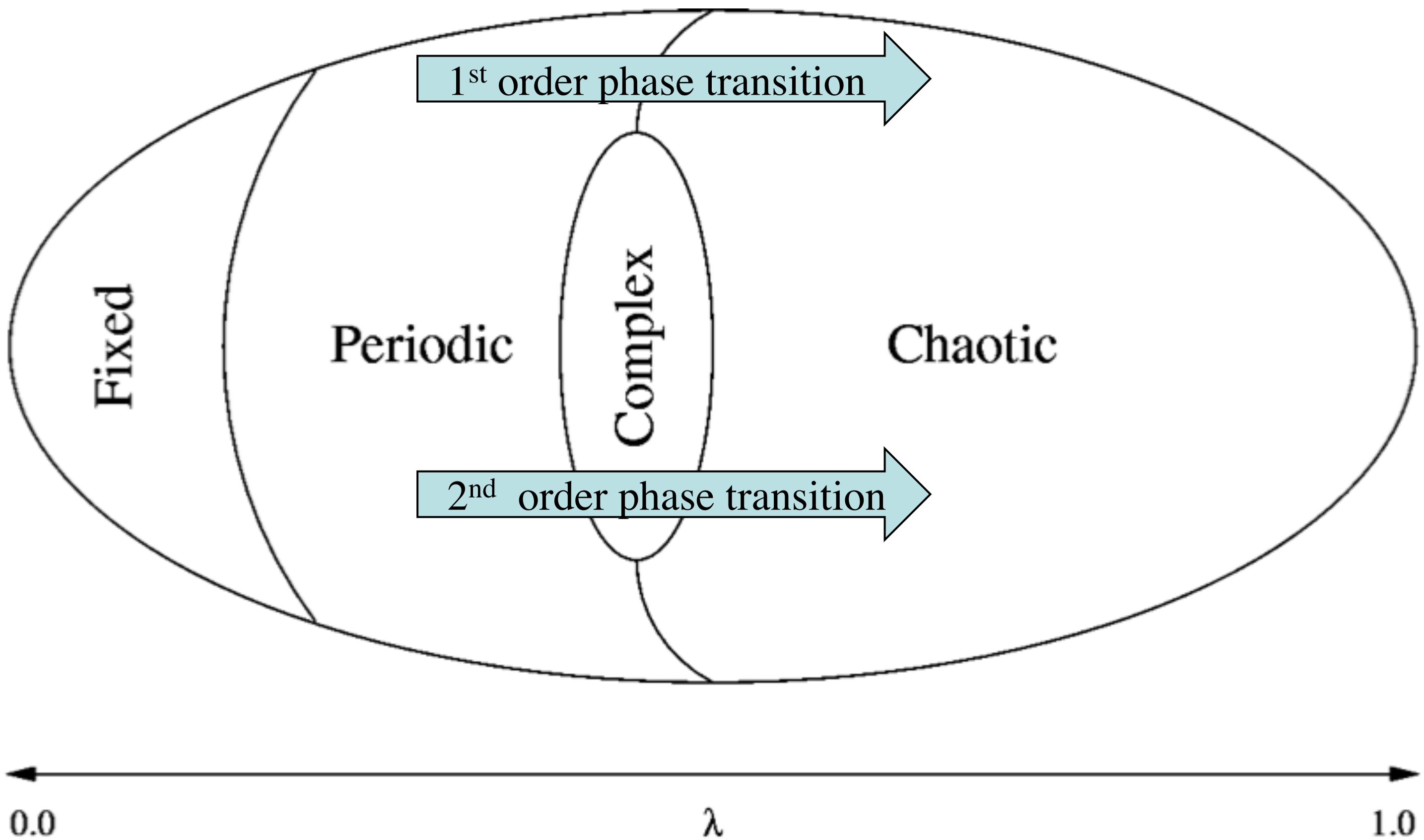


Fig. from Langton, “Life at Edge of Chaos”

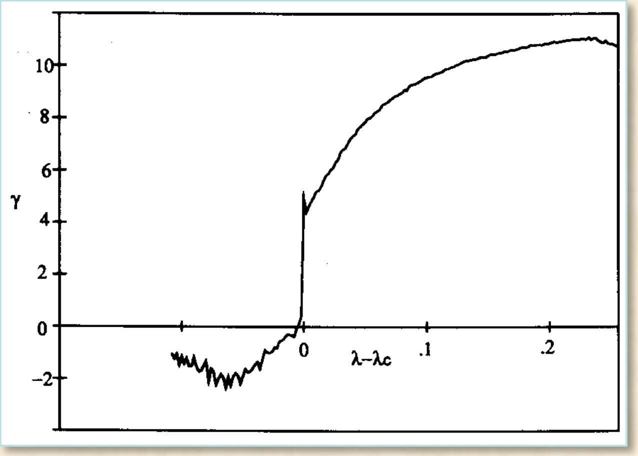
Compression-based Techniques

- Idea: lossless compression (e.g., Lempel-Ziv) approximates program-size complexity of a string
- Compare compressed and uncompressed histories
  - ✓ comp $\ll$ uncomp $\Rightarrow$ classes I or II
  - ✓ comp $\approx$ uncomp $\Rightarrow$ classes III or IV

- Hector Zenil, "Compression-Based Investigation of the Dynamical Properties of Cellular Automata and Other Systems," *Complexity*, **19** (2010).

Difference Patterns

- Sensitivity to initial conditions is characteristic of chaotic systems
- Difference patterns show the difference between evolutions from slightly different initial states
- Difference pattern spreading rate γ:
  - Little spread for small λ
  - Jumps at λ_c
  - Roughly constant rate for large λ
  - Complex behavior associated with intermediate values



Demonstration of 1D Totalistic CA Difference Patterns

Run CA 1D General Totalistic Dif

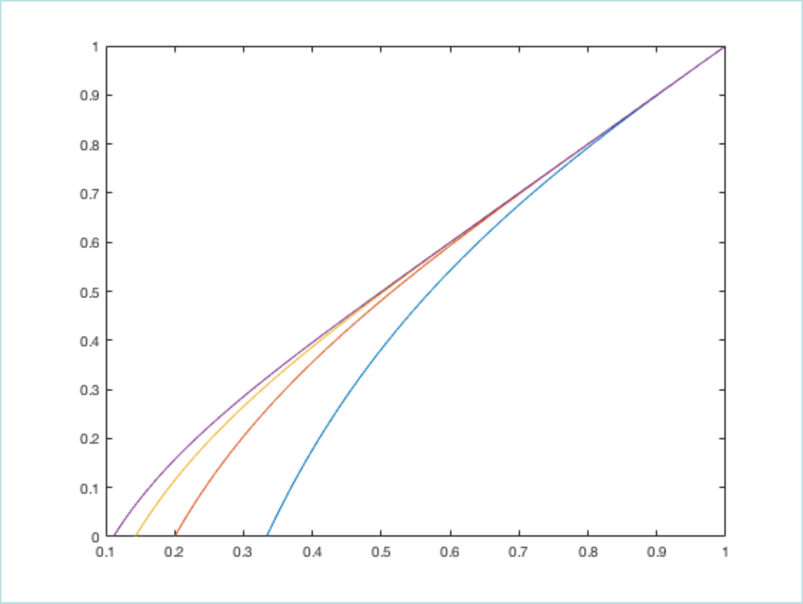
Wolfram Classes in Terms of Compressibility and Sensitivity

- Class I
  - Highly compressible
  - Insensitive to initial conditions
- Class II
  - Highly compressible
  - Sensitive to initial conditions
- Class III
  - Minimally compressible
  - Insensitive to initial conditions
- Class IV
  - Minimally compressible
  - Sensitive to initial conditions

Mean Field Approximations

- Assume states are uncorrelated
- Let q = density of quiescent states and $p = 1 - q$ be density of non-quiescent states
- Density of quiescent at next step is
$$q' = q^N + (1 - q^N)(1 - \lambda)$$
- Stationary value of q given by:
$$q = q^N + (1 - q^N)(1 - \lambda)$$
- Can solve easily for λ in terms of q (but not vice versa):
$$\lambda = \frac{1 - q}{1 - q^N}$$

Mean Field Estimates of Non-quiescence



- Estimated $p = 1 - q$ as function of λ
- $N = 3, 5, 7, 9$ ($r = 1, 2, 3, 4$ for 1D)
- Note complete quiescence below a critical λ value
- With larger neighborhoods expect non-quiescence

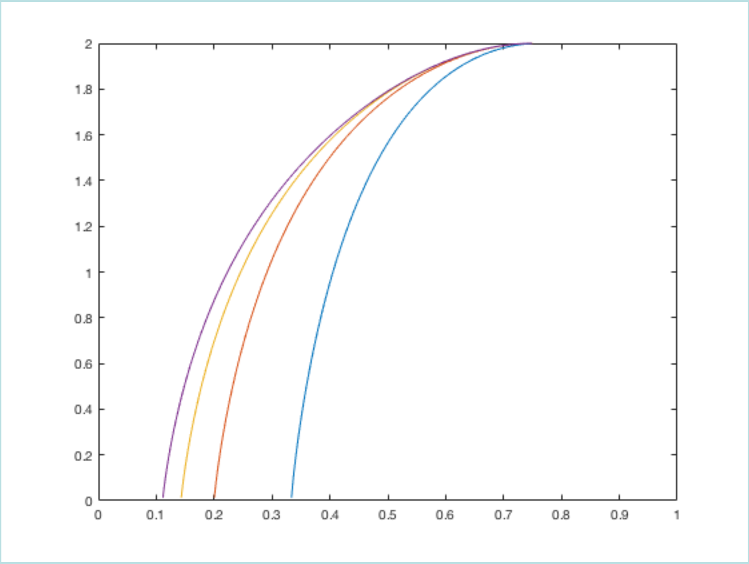
Mean Field Entropy Approximation

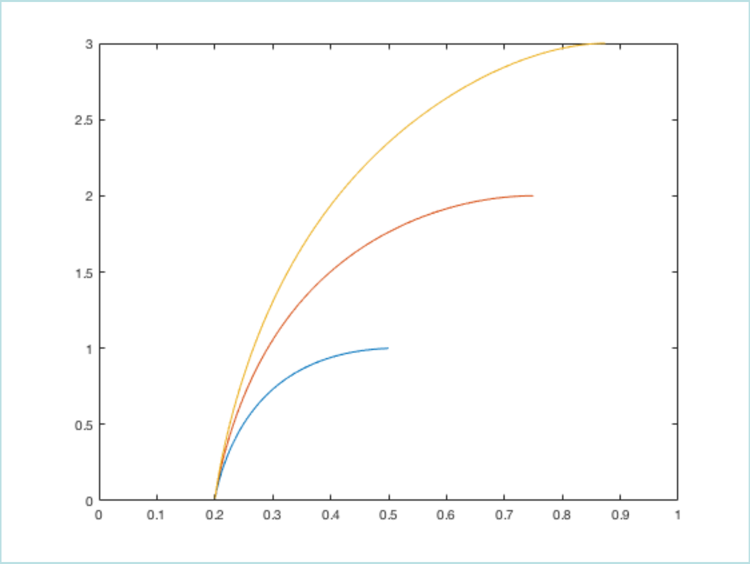
- Recall $H = \sum_k p_k \lg p_k$
- Probability of quiescent state $p_0 = q$
- Assumed equal probability on non-quiescent states: $p_k = \frac{1-q}{K-1}$
- Estimated entropy:

$$H = -\left[q \lg q + (K-1)\frac{1-q}{K-1}\lg\frac{1-q}{K-1}\right]$$

$$= -\left[q \lg q + (1-q)\lg\frac{1-q}{K-1}\right]$$

MF Estimate of Entropy vs. λ



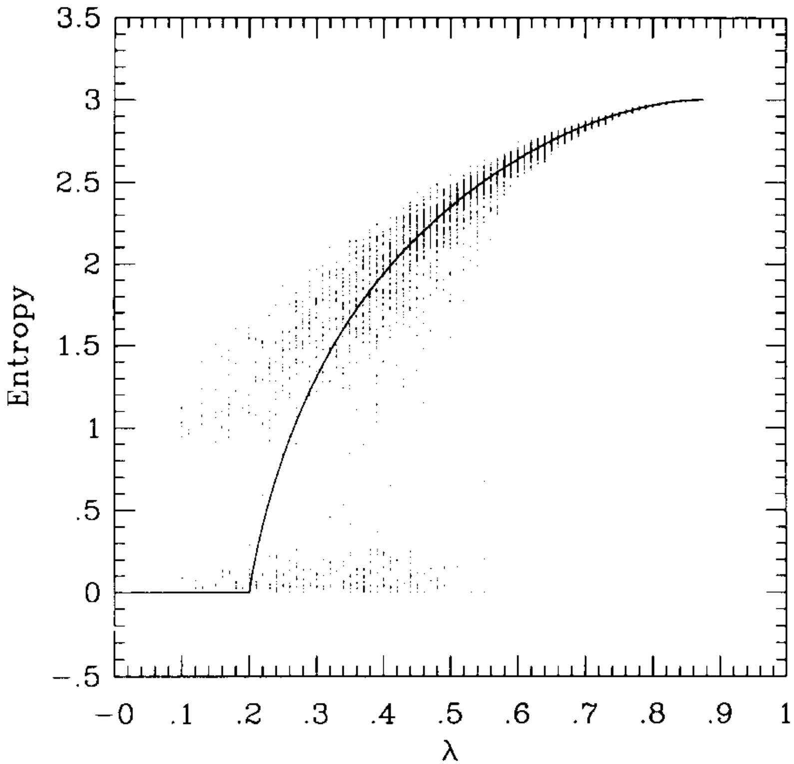


- Estimated entropy for $N = 3, 5, 7, 9$ ($K = 4$)

- Estimated entropy for $K = 2, 4, 8$ ($N = 5$)
- Note that with strong quiescence $\lambda \leq (K-1)/K$

Empirical Entropy vs. λ

- 2D CAs, 64×64
- $K = 8$
- $N = 5$ (von Neumann nbd)
- For each λ, 100 rule tables were constructed at random (rotation invariant)
- Each CA was run for 500 time steps before measuring entropy
- Entropy measured over 1000 steps
- Wooters & Langton (1990)



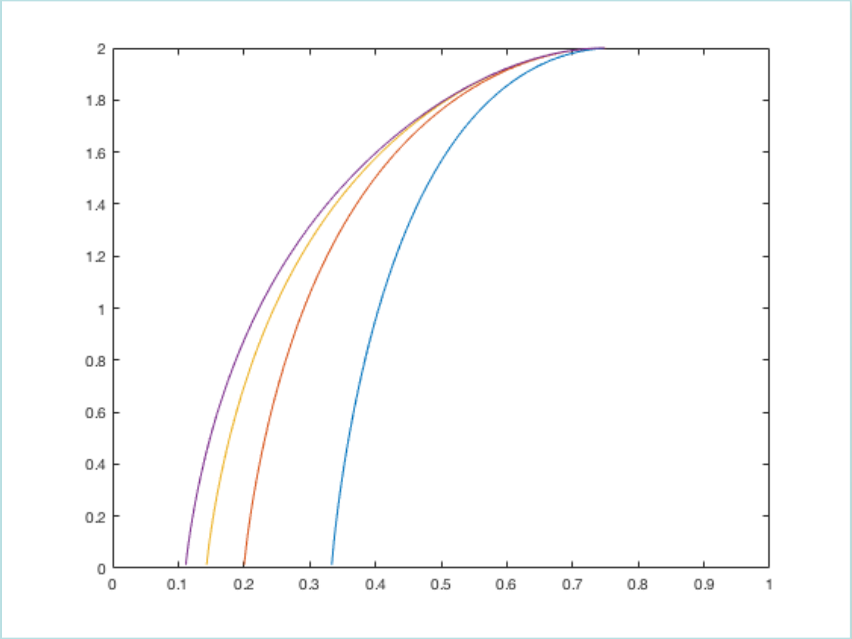
Mean Field Estimate of Spreading Rate

- Max possible rate is $\gamma_{\max} = 2r$
- For binary CA, probability two block map to same state:
$$p = \lambda^2 + (1-\lambda)^2$$
- Probability they map to different:
$$1 - p = 2\lambda(1-\lambda)$$
- Can show that average spread rate is:
$$\gamma = 2\left(r - \frac{p}{1-p}\right) = \frac{2(r+1)\lambda - 2(r+1)\lambda^2 - 1}{\lambda(1-\lambda)}$$
- Set $\gamma = 0$ and to find when becomes positive:
$$\lambda_c^\gamma = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2}{r+1}}$$

Li, Packard & Langton (1990)

MF Estimates of λ_c

MF estimated entropy for $\boldsymbol{N} = 3, 5, 7, 9$ ($\boldsymbol{K} = 4$)



Critical lambda from MF estimated spread rate ($\boldsymbol{K} = 2$)

- For $N = 3, r = 1, \gamma \; = \; 0.5$
- For $N = 5, r = 2, \gamma \; = \; 0.211$
- For $N = 7, r = 3, \gamma \; = \; 0.146$
- For $N = 9, r = 4, \gamma \; = \; 0.113$

Reversible CAs (RCAs)

- Most CAs are irreversible
  - By approaching an attractor they lose information about the initial state
  - They mix information from remote sites
- The fundamental laws of nature are time-reversible
  - Newtonian mechanics
  - Quantum mechanics
- Reversible computation is required by:
  - Ultralow power computation (below Landauer limit)
  - Quantum computation

Some Definitions and Results

- Global state vs. local (or neighborhood) state/configuration
- Global transition function F vs. local transition function f (or rule)
- A CA is *injective* if F is 1-1
  – Every global state has exactly one predecessor
- A CA is *reversible* (or *invertible*) if there is a CA with the global function F^{-1}
- Thm: A CA is reversible iff it is injective
- If a reversible CA has the same rule as its inverse, then we call it a *reversible rule*
  – The rule is *time-reversal invariant* like the laws of physics
- Thm: Reversibility of 1D CAs is decidable
- Thm: Reversibility for 2 or higher dimensional CAs is undecidable

It is hard to find RCAs, but we can make them

Second-order RCAs

- Idea: Let the neighborhood determine a reversible *change* from the preceding cell state:
  $$s_i(t+1) = f\left(S_{[i]}(t)\right) \ominus s_i(t-1)$$
  where s_i is the state of cell *i*
  and $S_{[i]}$ is the state of the nbd of cell *i*
  and $\ominus$ is mod *K* subtraction
- Note time-reversal invariance:
  $$s_i(t-1) = f\left(S_{[i]}(t)\right) \ominus s_i(t+1)$$
- First-order CAs determine new state from previous state
- Second-order CAs determine new state from previous *two* states

Demonstration of Second-order RCA

[Run CA 1D Reversible H]

Generalization of 2nd-order RCA

- The formula $s_i(t+1) = f\left(S_{[i]}(t)\right) \ominus s_i(t-1)$ uses $f\left(S_{[i]}(t)\right)$ to select a *rotation* of the state space $[0, 1, \ldots, K-1]$
- A rotation is a special case of a permutation
- Idea: Since permutations are invertible, use $f\left(S_{[i]}(t)\right)$ to choose a permutation on the state space.
- The inverse rule applies the inverse permutation $\left[f\left(S_{[i]}(t)\right)\right]^{-1}$

Reducing 2nd-order to 1st-order

- A second order RCA can be reduced to a first order RCA
- Just expand the state space from K to K^2 to include a record of the previous state:

$$s_i(t+1) = [s_i'(t+1), s_i''(t+1)]$$
$$= [s_i'(t+1), s_i'(t)]$$

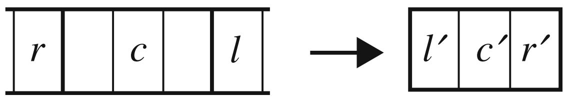
- First-order update equation:

$$s_i(t+1) = \left[f\left(S'_{[i]}(t)\right) \ominus s_i''(t), s_i'(t)\right]$$

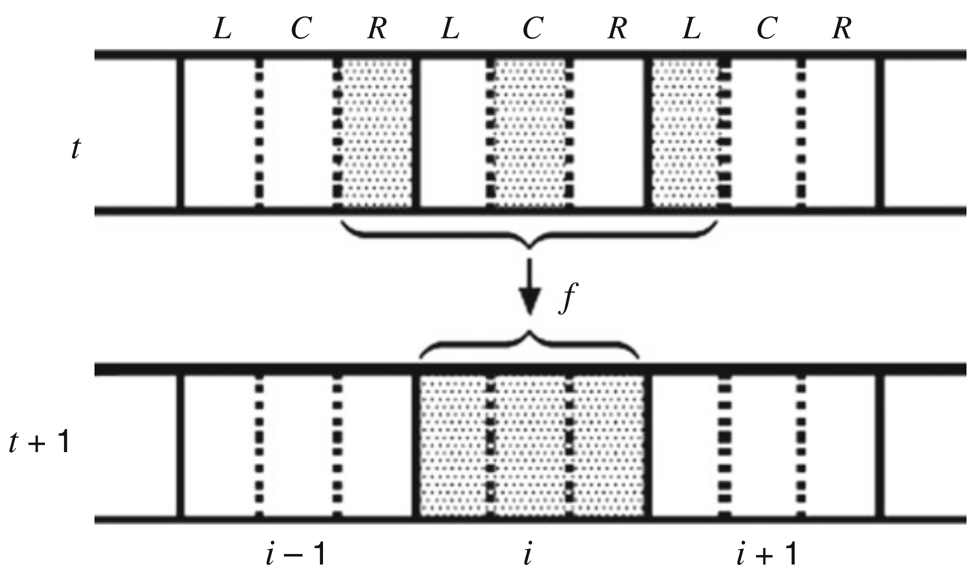
- To reverse RCA, exchange components

Partitioned CA (PCA)

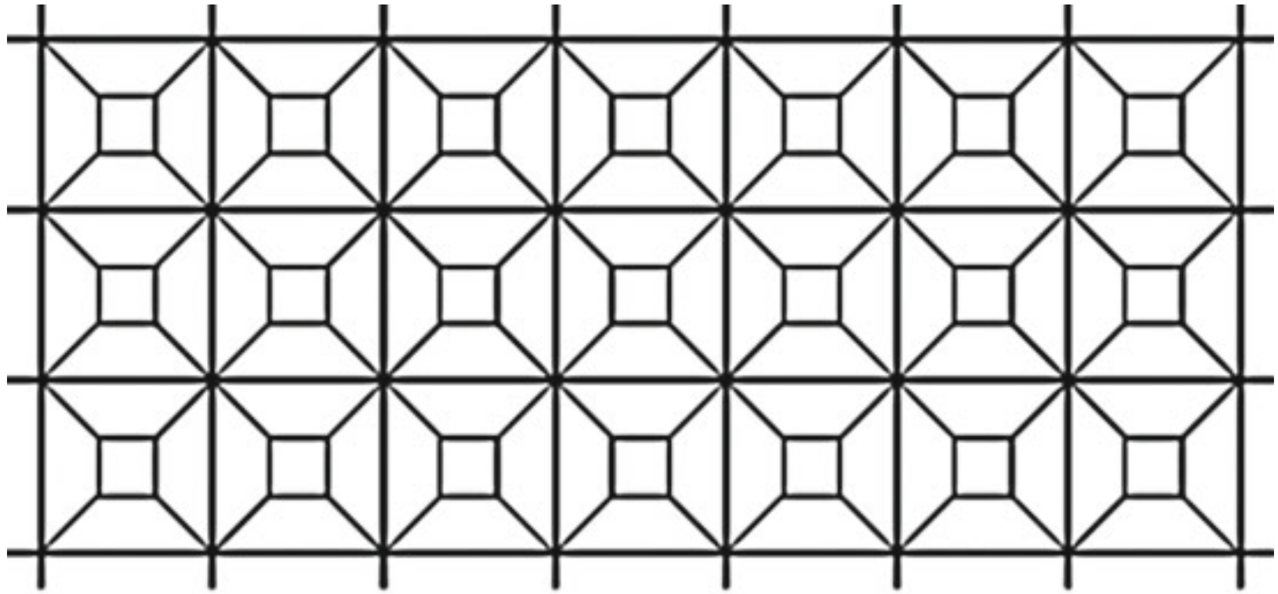
- Each cell divided into several compartments
- Extra compartments save info needed for reversibility
- 1D case: next state determined by center part and nearest parts of neighbors:
  $[l', c', r'] = f(r, c, l)$

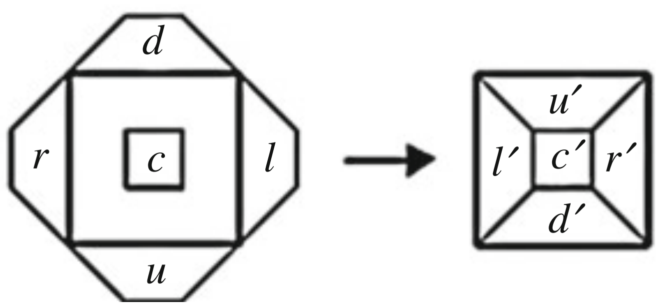


- Global function is injective iff local function is injective
- Can be simulated by ordinary CA on extended state space



2-dimensional PCA



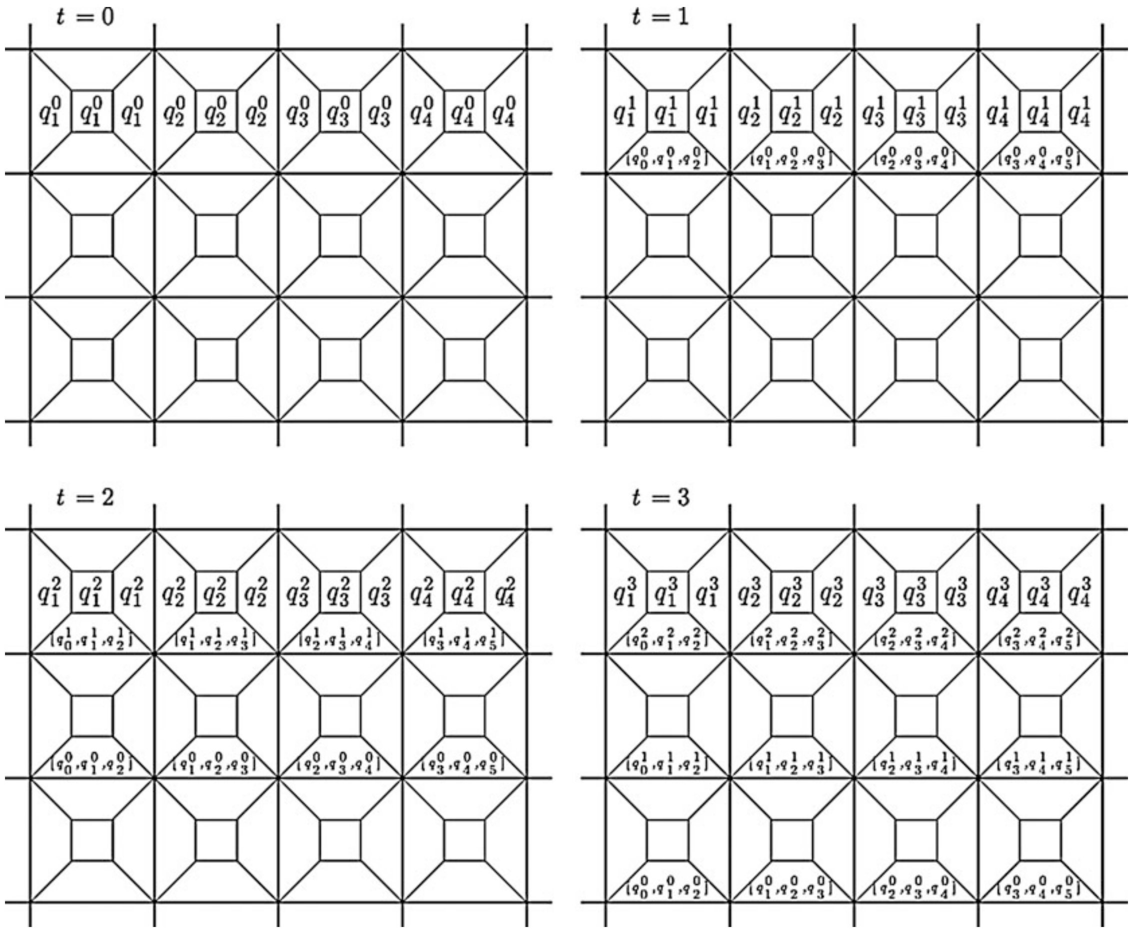


Simulation of CAs by RCAs

- It's easy to construct RCAs
- Every RCA can be simulated by an ordinary CA (on bigger state space)
- Any *d* dimensional CA can be simulated in real time by a *d*+1 dimensional RCA

Example: Simulation of 1D CA by 2D PCA

$t = 0$	q_1^0	q_2^0	q_3^0	q_4^0
$t = 1$	q_1^1	q_2^1	q_3^1	q_4^1
$t = 2$	q_1^2	q_2^2	q_3^2	q_4^2
$t = 3$	q_1^3	q_2^3	q_3^3	q_4^3



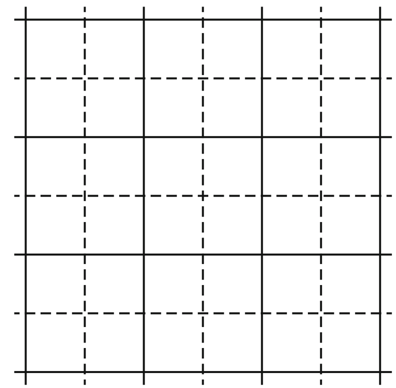
- Extra dimension holds history
- One row simulates CA and pushes nbd state down
- History rows keep shifting it down

Reversible Universal Computation

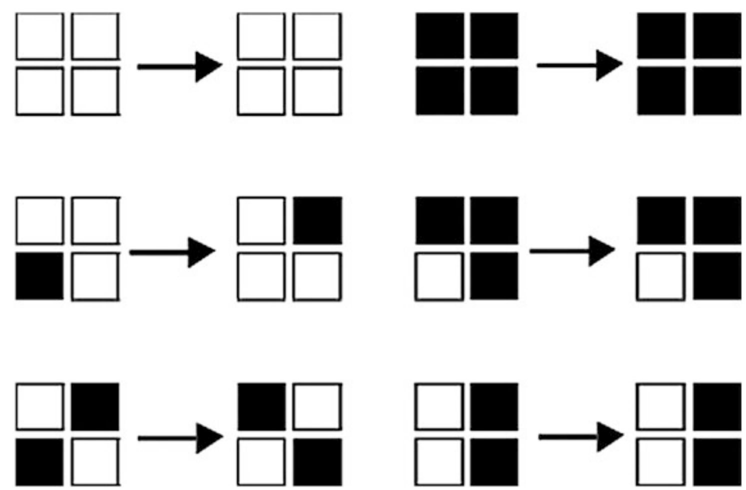
- Any TM can be simulated by a CA
  - In particular, by a 1D CA
- Any CA can be simulated by an RCA
  - In particular, a 1D CA can be simulated in real-time by a 2D RCA
- Therefore, any TM can be simulated by a 2D RCA
- Therefore, RCAs are capable of universal computation

Example of 2D RCA Capable of Universal Computation

- Two-state block CA
  - 1-1 rules apply in overlapping blocks
- Rotation symmetry assumed
- Can simulate:
  - reversible logic gates
  - billiard ball model
- For more, take COSC 494/594 Unconventional Computation



Margolus neighborhood



Suitable Media for Computation

- How can we identify/synthesize novel computational media?
  - especially nanostructured materials for massively parallel computation
- Seek materials/systems exhibiting Class IV behavior
  - may be identifiable via entropy, mut. info., etc.
- Find physical properties (such as λ) that can be controlled to put into Class IV

Some of the Work in this Area

- Wolfram: *A New Kind of Science*
  - www.wolframscience.com/nksonline/toc.html
- Langton: Computation/life at the edge of chaos
- Crutchfield: Computational mechanics
- Mitchell: Evolving CAs
- and many others…

Some Other Simple Computational Systems Exhibiting the Same Behavioral Classes

- CAs (1D, 2D, 3D, totalistic, etc.)
- Mobile Automata
- Turing Machines
- Substitution Systems
- Tag Systems
- Cyclic Tag Systems
- Symbolic Systems (combinatory logic, lambda calculus)
- Continuous CAs (coupled map lattices)
- PDEs
- Probabilistic CAs
- Multiway Systems

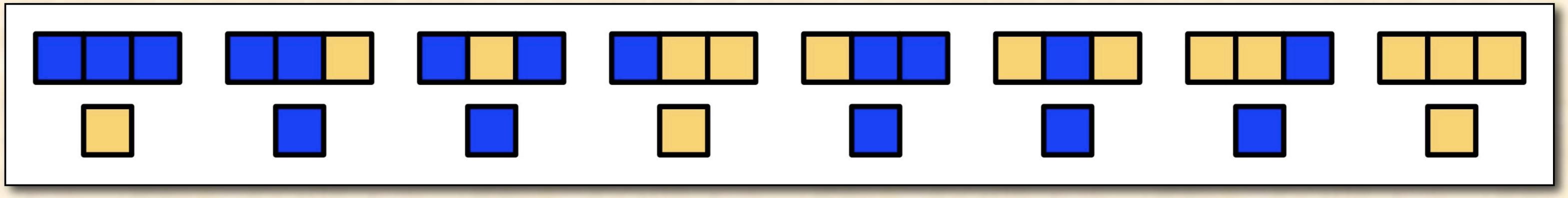
Universality

- A system is *computationally universal* if it can compute anything a Turing machine (or digital computer) can compute
- The Game of Life is universal
- Several 1D CAs have been proved to be universal
- Are all complex (Class IV) systems universal?
- Is universality rare or common?

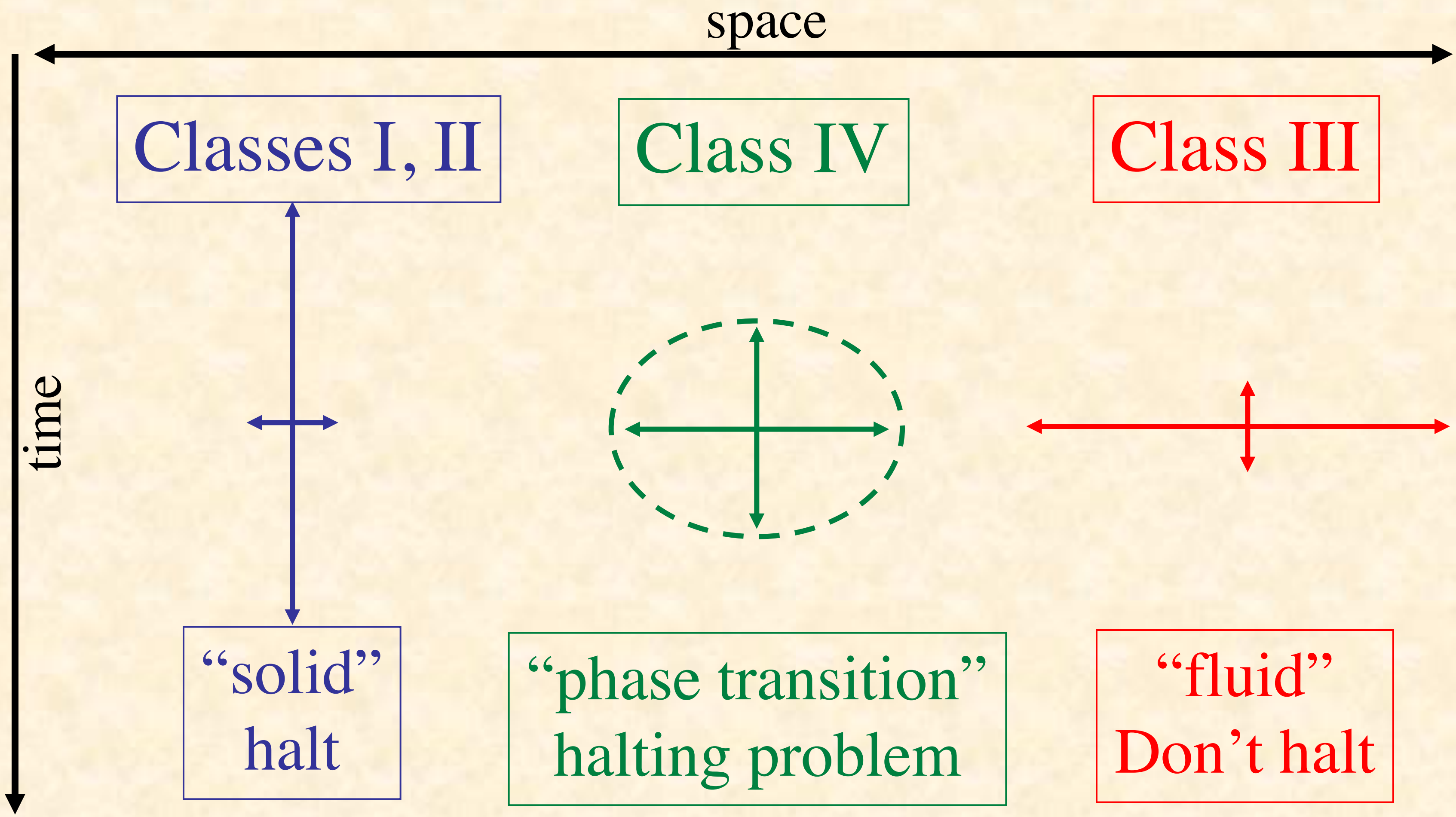
Rule 110: A Universal 1D CA

- $K = 2, N = 3$
- New state = $\neg(p \wedge q \wedge r) \wedge (q \vee r)$

  where p, q, r are neighborhood states
- Proved by Wolfram



Fundamental Universality Classes of Dynamical Behavior



Wolfram's Principle of Computational Equivalence

- "a fundamental unity exists across a vast range of processes in nature and elsewhere: despite all their detailed differences every process can be viewed as corresponding to a computation that is ultimately equivalent in its sophistication" (*NKS* 719)
- Conjecture: "among all possible systems with behavior that is not obviously simple an overwhelming fraction are universal" (*NKS* 721)

Computational Irreducibility

- "systems one uses to make predictions cannot be expected to do computations that are any more sophisticated than the computations that occur in all sorts of systems whose behavior we might try to predict" (*NKS* 741)
- "even if in principle one has all the information one needs to work out how some particular system will behave, it can still take an irreducible amount of computational work to do this" (*NKS* 739)
- That is: for Class IV systems, you can't (in general) do better than simulation.

What do CAs have to do with bio-inspired computation?

- Cellular automata were motivated by biological cells and reproduction
- Living systems display complex, organized behavior
- Yet we have seen that simple, abstract systems such as CAs display similar complexity
- Thus some of this complex behavior is not unique to living things and can appear in non-living systems as well
- CAs help us to see the essence of complex, organized behavior, so we are better able to use it in our artificial systems

Additional Bibliography

1. Langton, Christopher G. “Computation at the Edge of Chaos: Phase Transitions and Emergent Computation,” in *Emergent Computation*, ed. Stephanie Forrest. North-Holland, 1990.
2. Langton, Christopher G. “Life at the Edge of Chaos,” in *Artificial Life II*, ed. Langton et al. Addison-Wesley, 1992.
3. Emmeche, Claus. *The Garden in the Machine: The Emerging Science of Artificial Life*. Princeton, 1994.
4. Wolfram, Stephen. *A New Kind of Science*. Wolfram Media, 2002.