II. Spatial Systems
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
  - 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
  $\Rightarrow$ stays as is (stasis)
- Live cell has < 2 live neighbors
  $\Rightarrow$ dies (loneliness)
- Live cell has > 3 live neighbors
  $\Rightarrow$ dies (overcrowding)
- Empty cell has 3 live neighbors
  $\Rightarrow$ comes to life (reproduction)

Demonstration of Life

Run NetLogo Life

or

<www.cs.utk.edu/~mclennan/Classes/420/NetLogo/Life.html>

Go to CBN
Online Experimentation Center

<mitpress.mit.edu/books/FLAOH/cbnhtml/java.html>

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

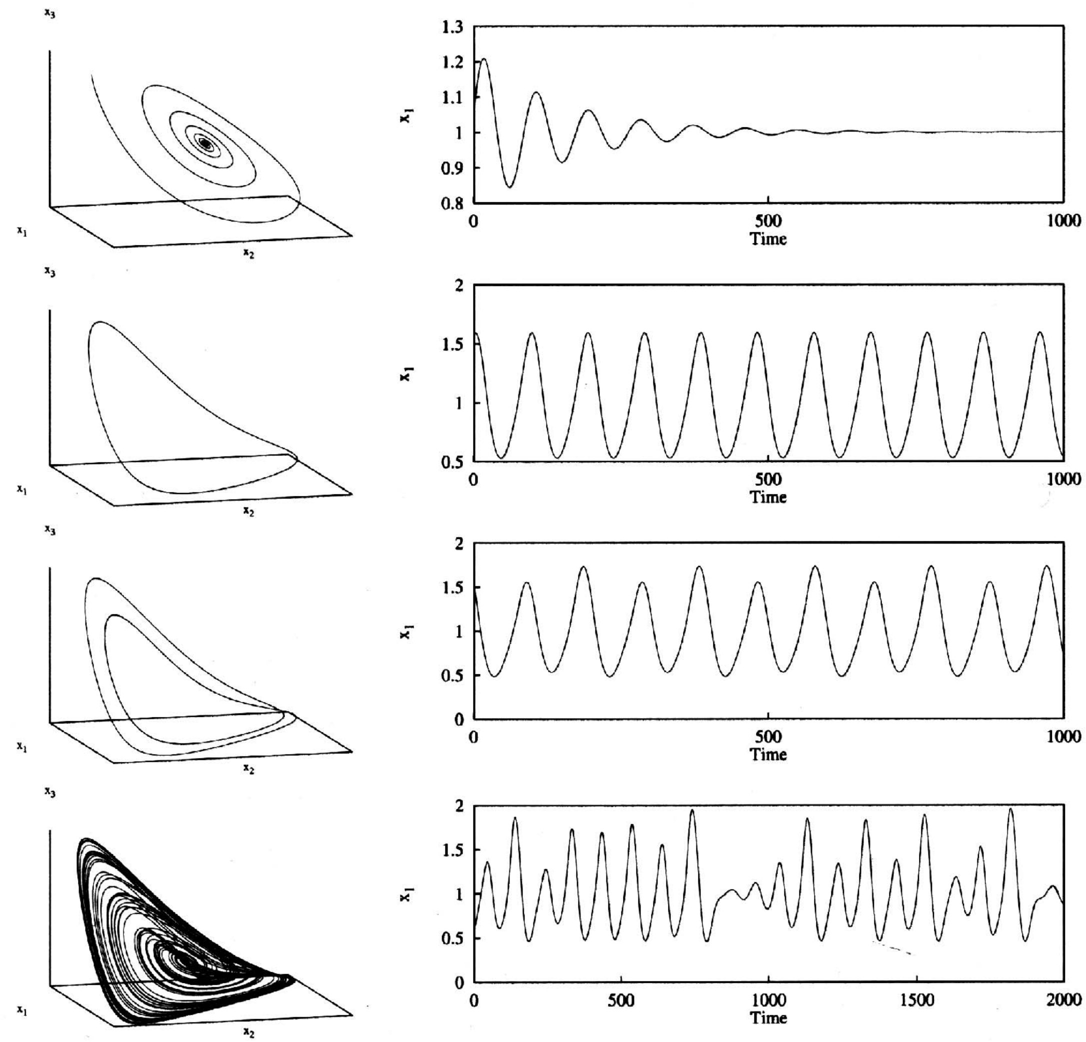


fig. from Flake via EVALife

Wolfram's Classification

Wolfram’s Classification

- Class I: evolve to fixed, homogeneous state

Wolfram's Classification

- Class I: evolve to fixed, homogeneous state
  - ~ limit point

Wolfram's Classification

- Class I: evolve to fixed, homogeneous state
  - ~ limit point
- Class II: evolve to simple separated periodic structures

Wolfram’s Classification

- Class I: evolve to fixed, homogeneous state
  ~ limit point
- Class II: evolve to simple separated periodic structures
  ~ limit cycle

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

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)

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

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

Assumptions

- Periodic boundary conditions

Assumptions

- Periodic boundary conditions
  - no special place

Assumptions

- Periodic boundary conditions
  - no special place
- Strong quiescence:

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

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

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:

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

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

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]:

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

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

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

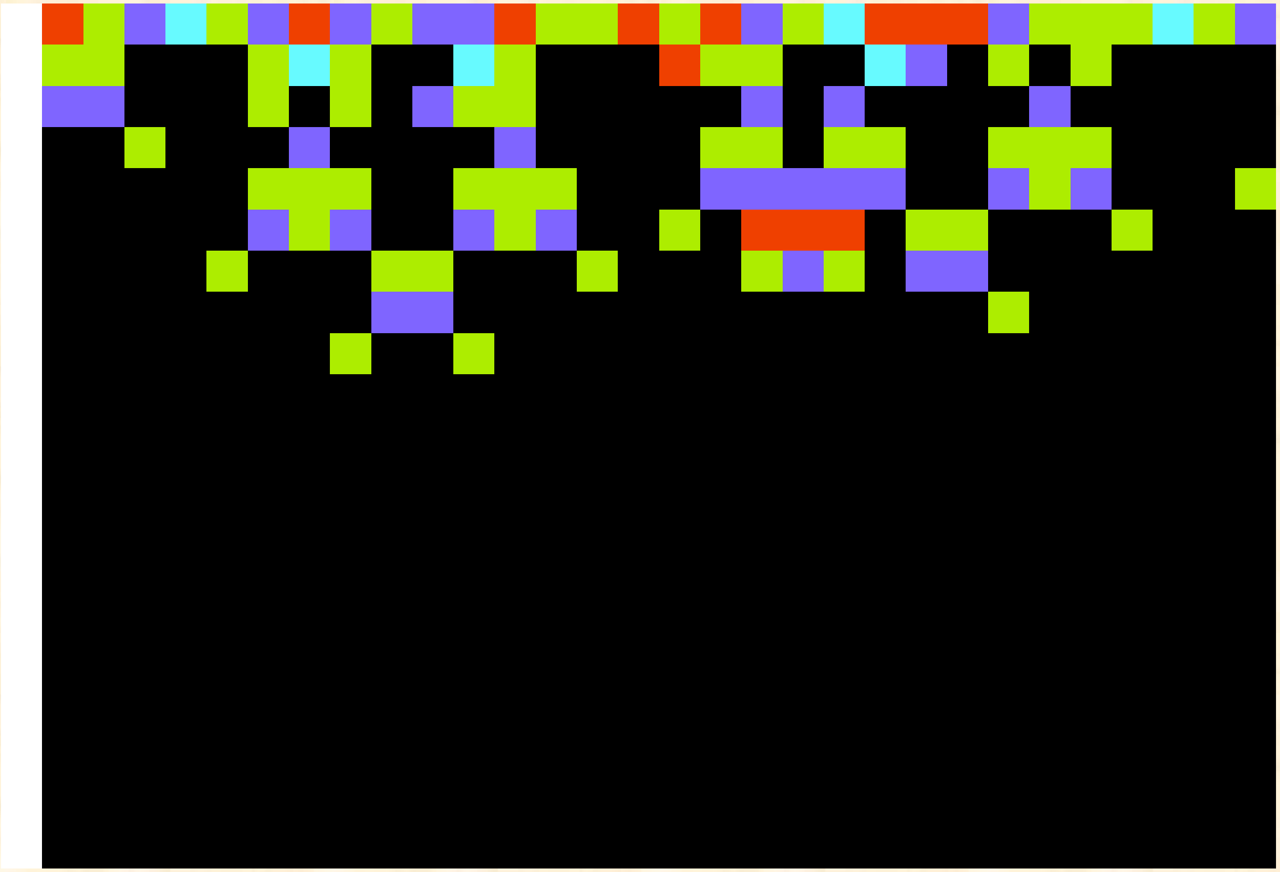
<www.cs.utk.edu/~mclennan/Classes/420/NetLogo/CA-1D-General-Totalistic.html>

Go to CBN
Online Experimentation Center

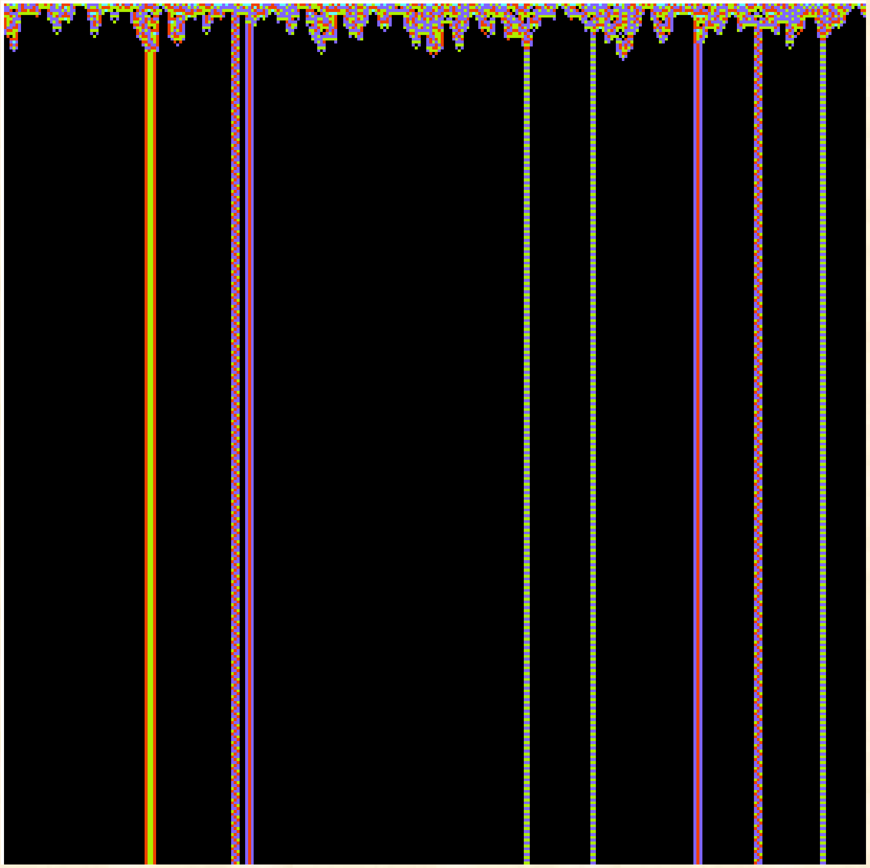
<mitpress.mit.edu/books/FLAOH/cbnhtml/java.html>

Class I ($\lambda = 0.2$)

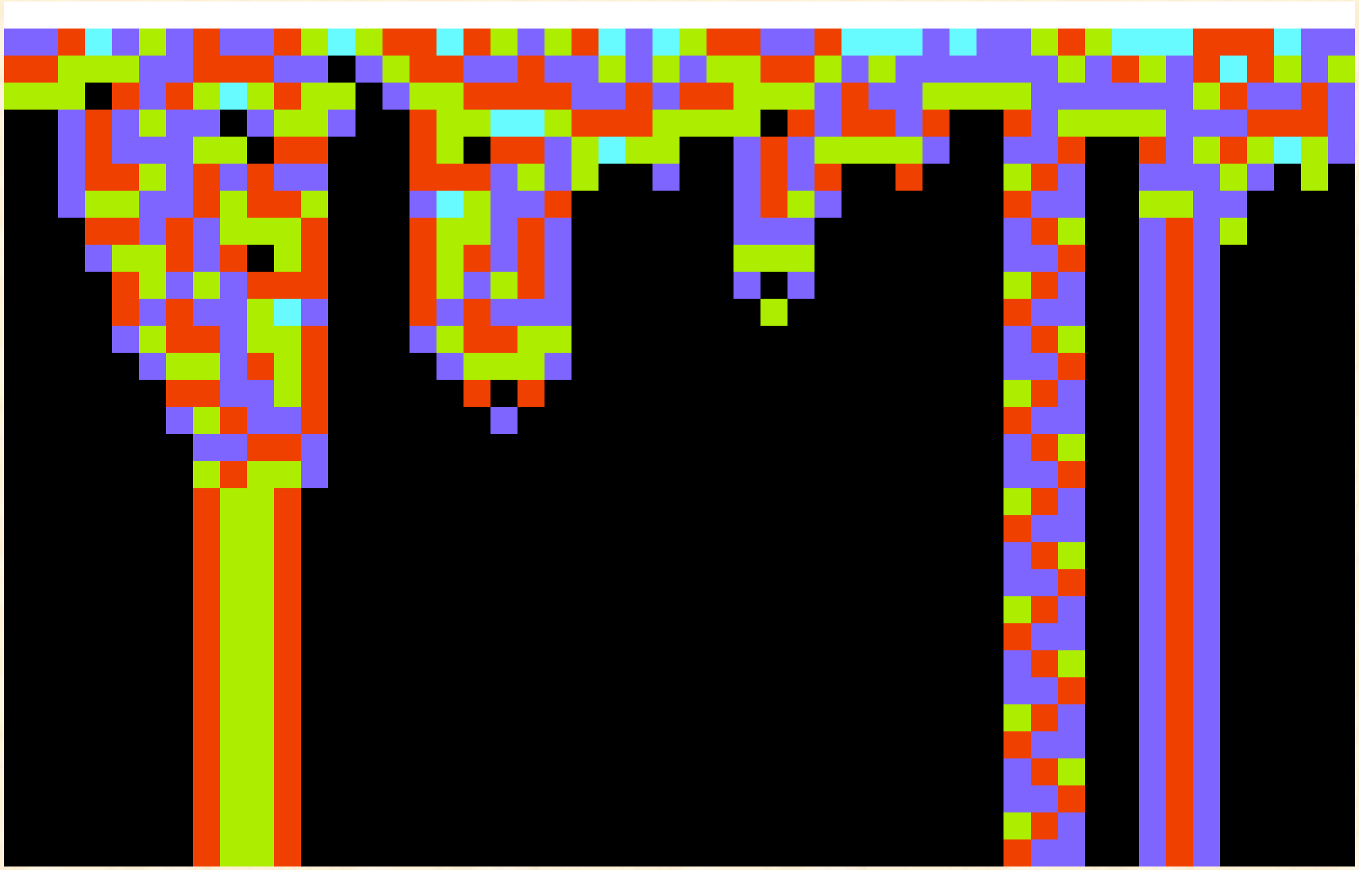
Class I ($\lambda = 0.2$) Closeup



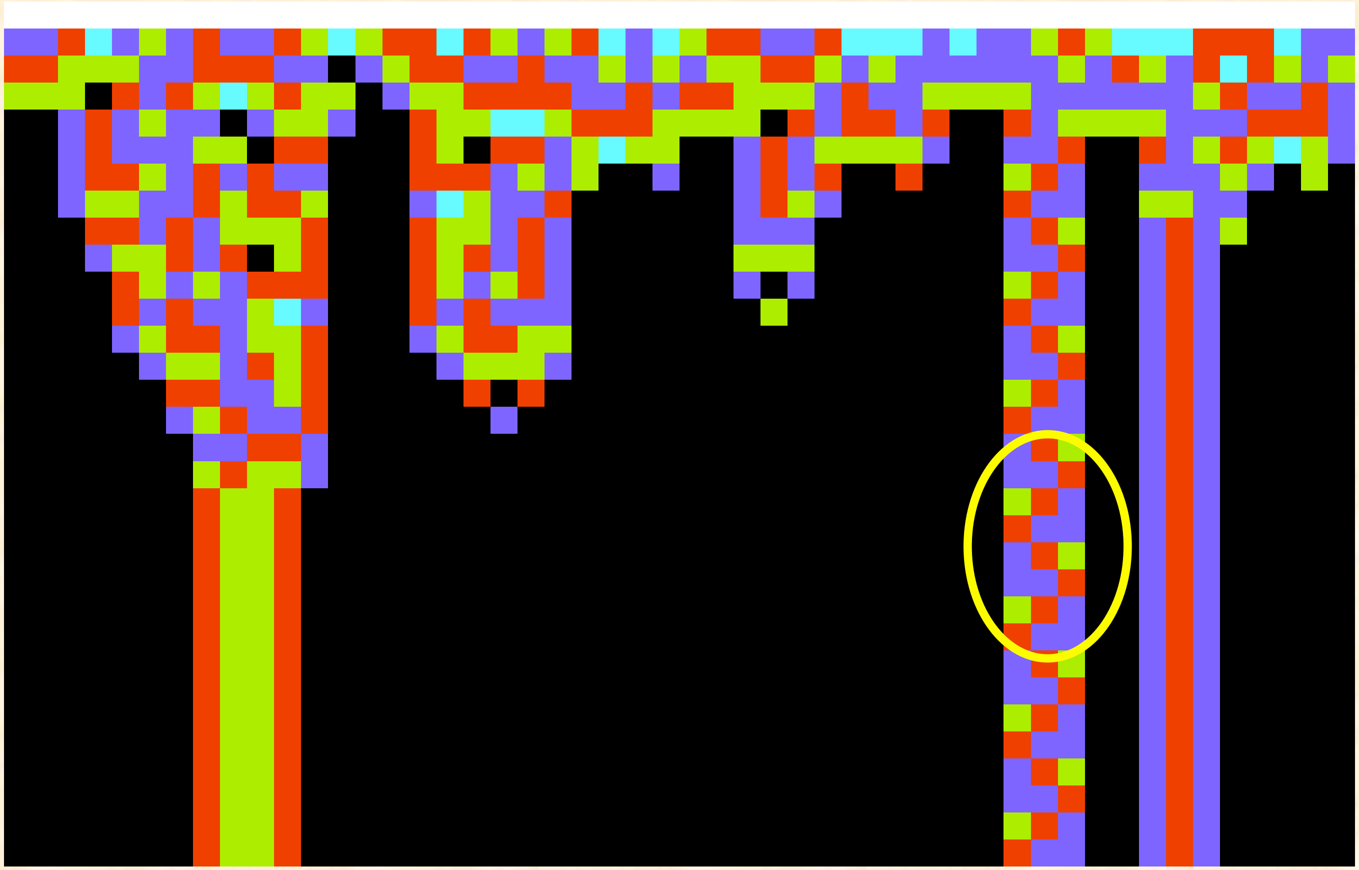
Class II ($\lambda = 0.4$)



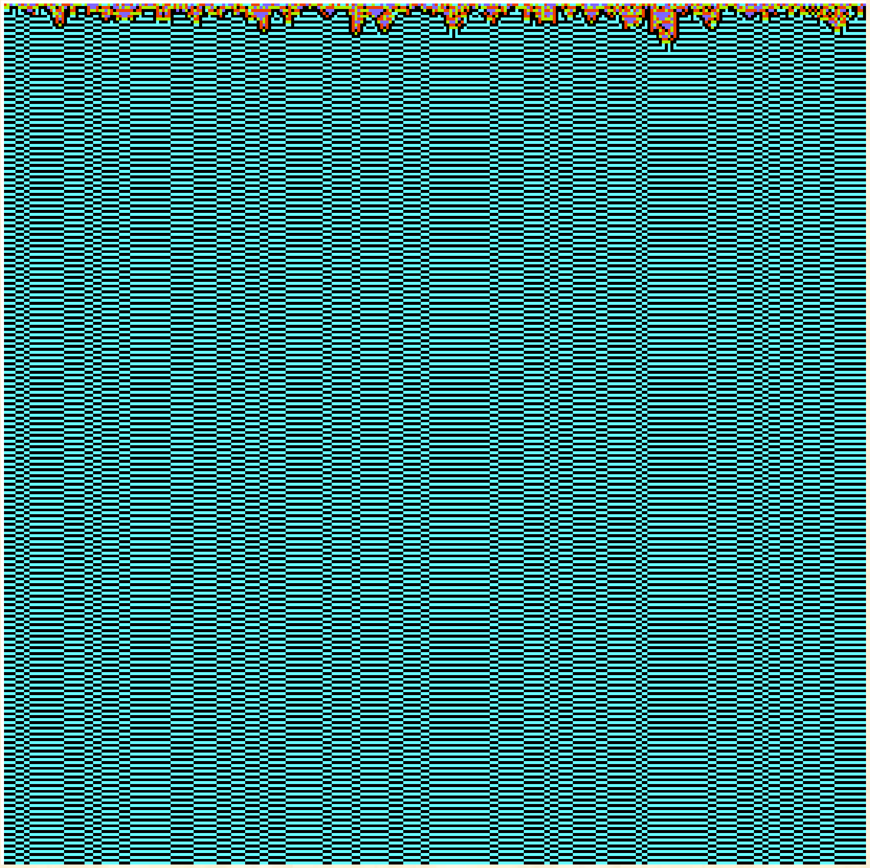
Class II ($\lambda = 0.4$) Closeup



Class II ($\lambda = 0.4$) Closeup

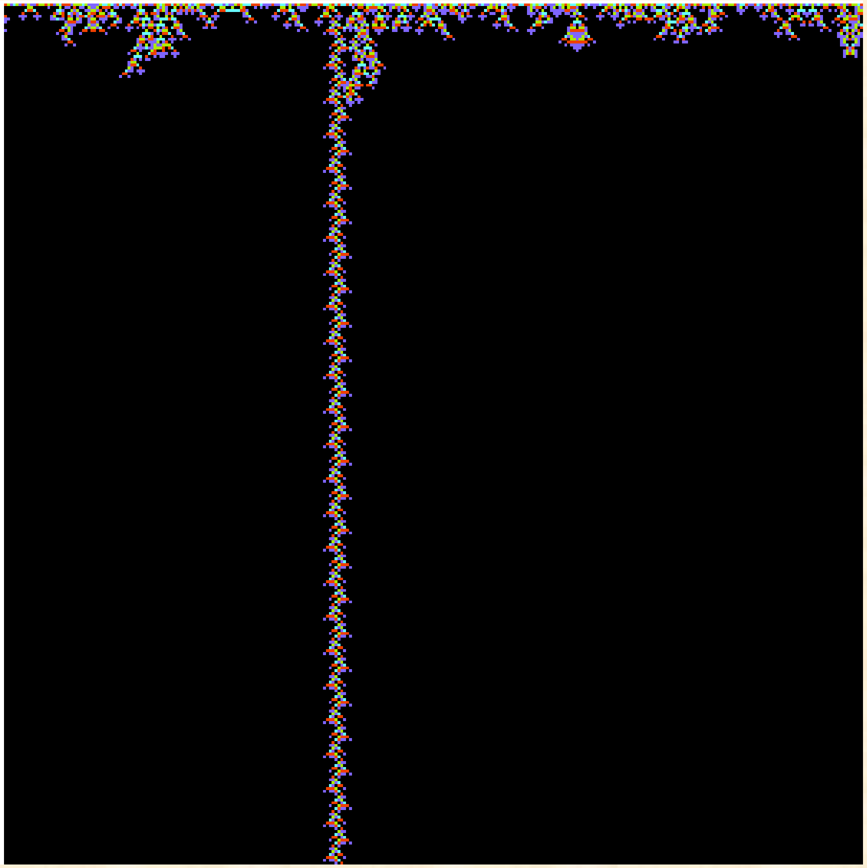


Class II ($\lambda = 0.31$)

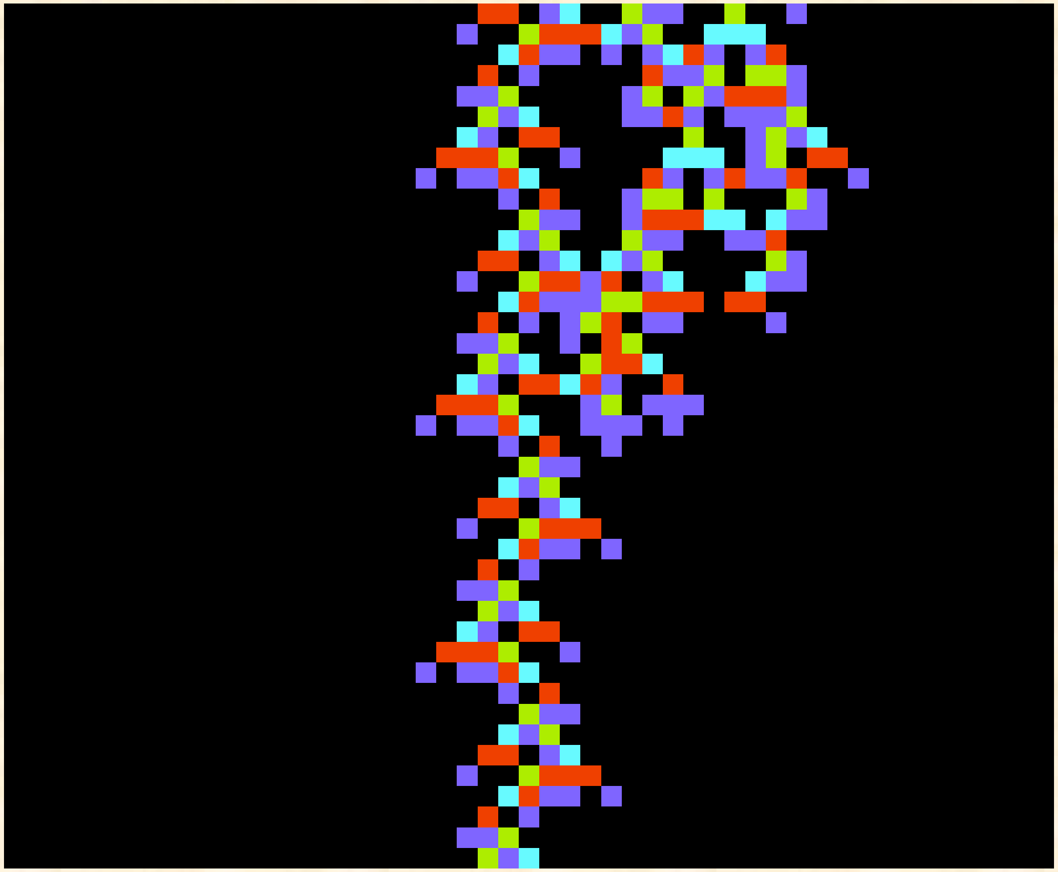


Class II ($\lambda = 0.31$) Closeup

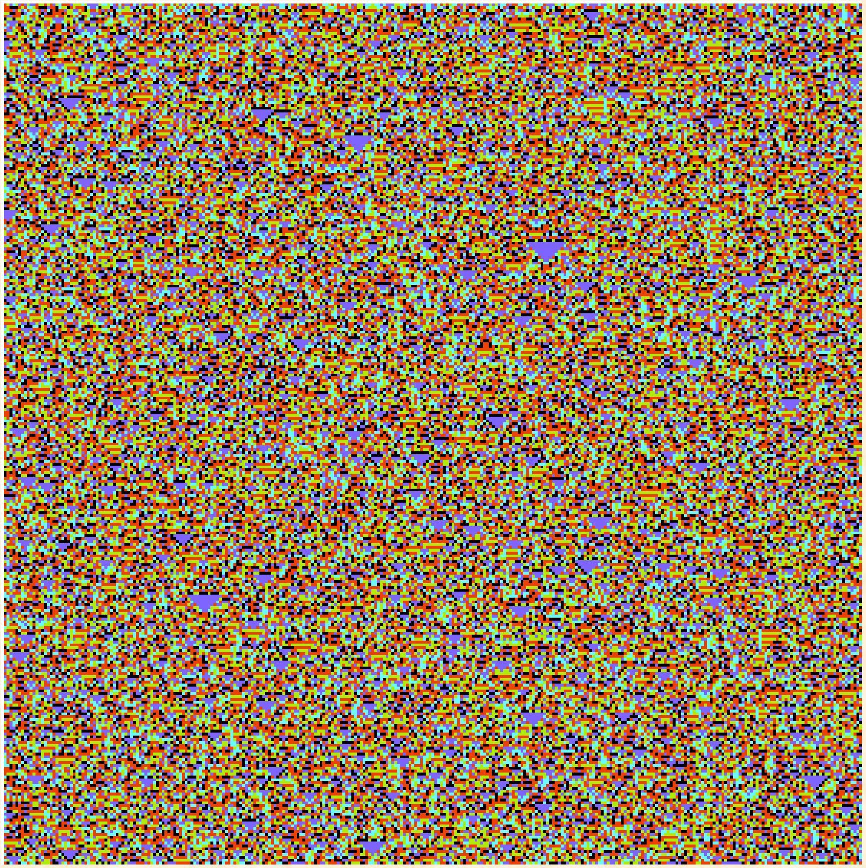
Class II ($\lambda = 0.37$)



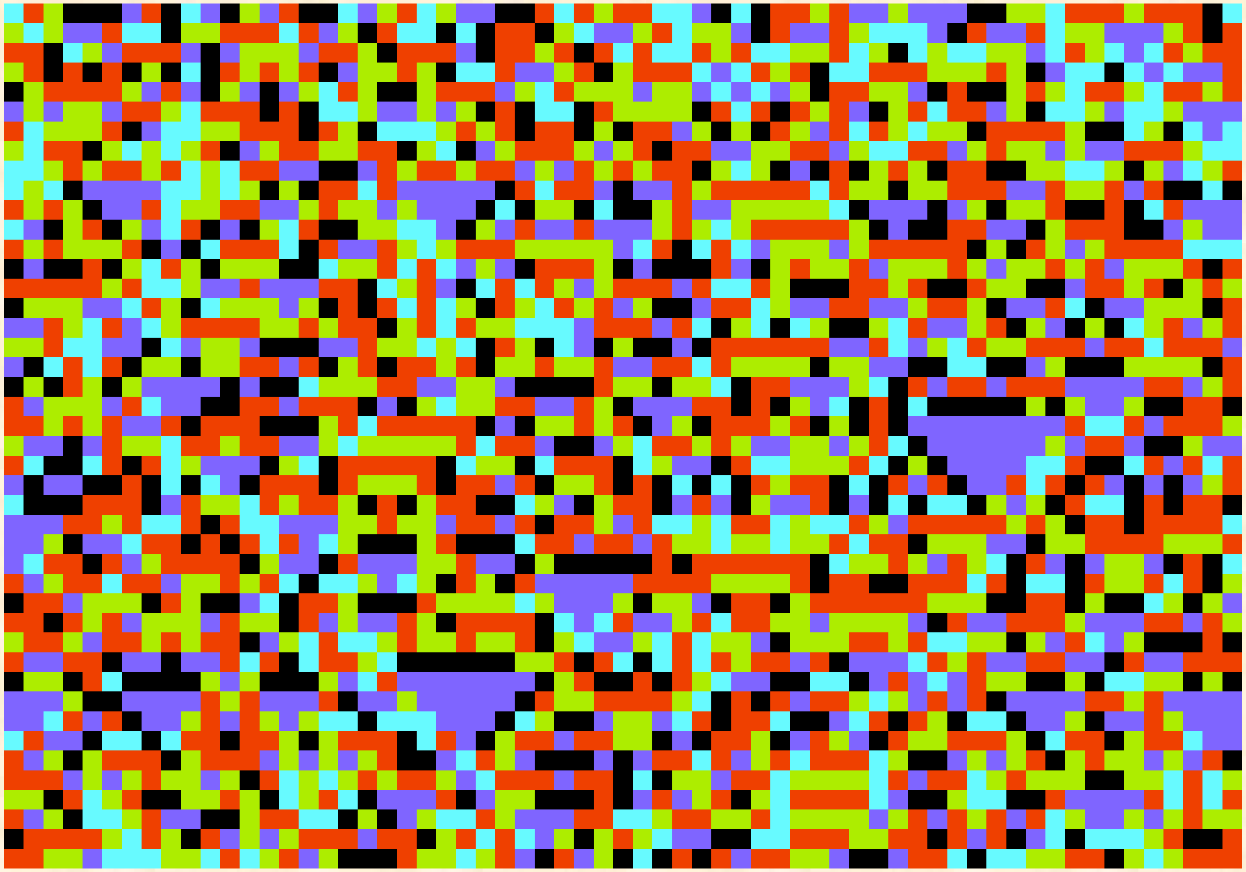
Class II ($\lambda = 0.37$) Closeup



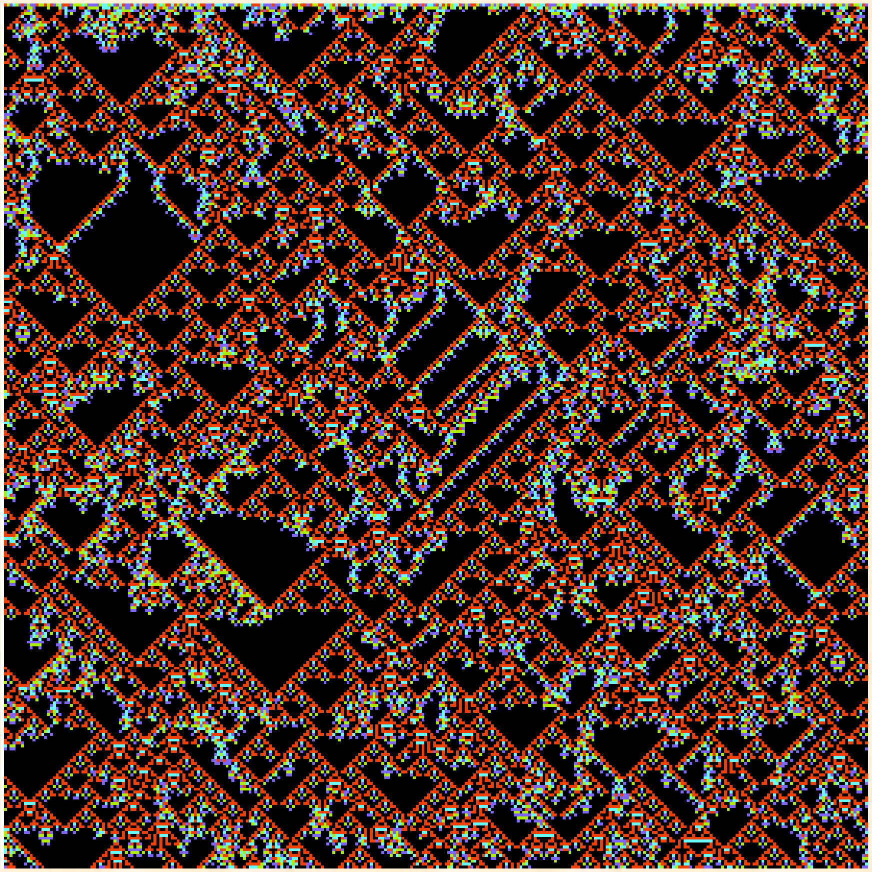
Class III ($\lambda = 0.5$)



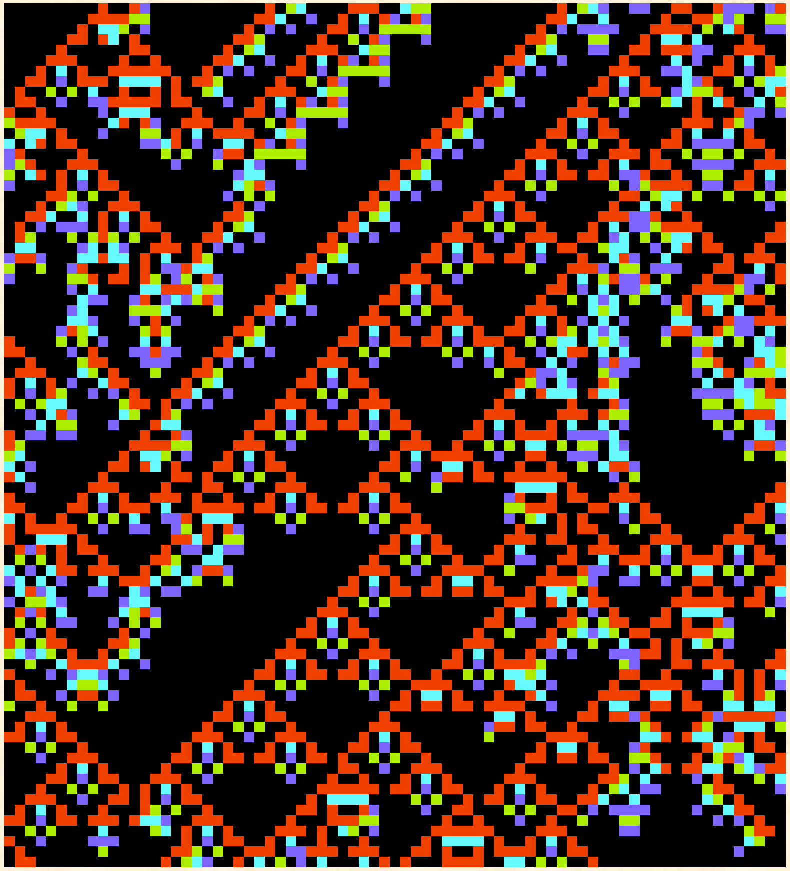
Class III ($\lambda = 0.5$) Closeup



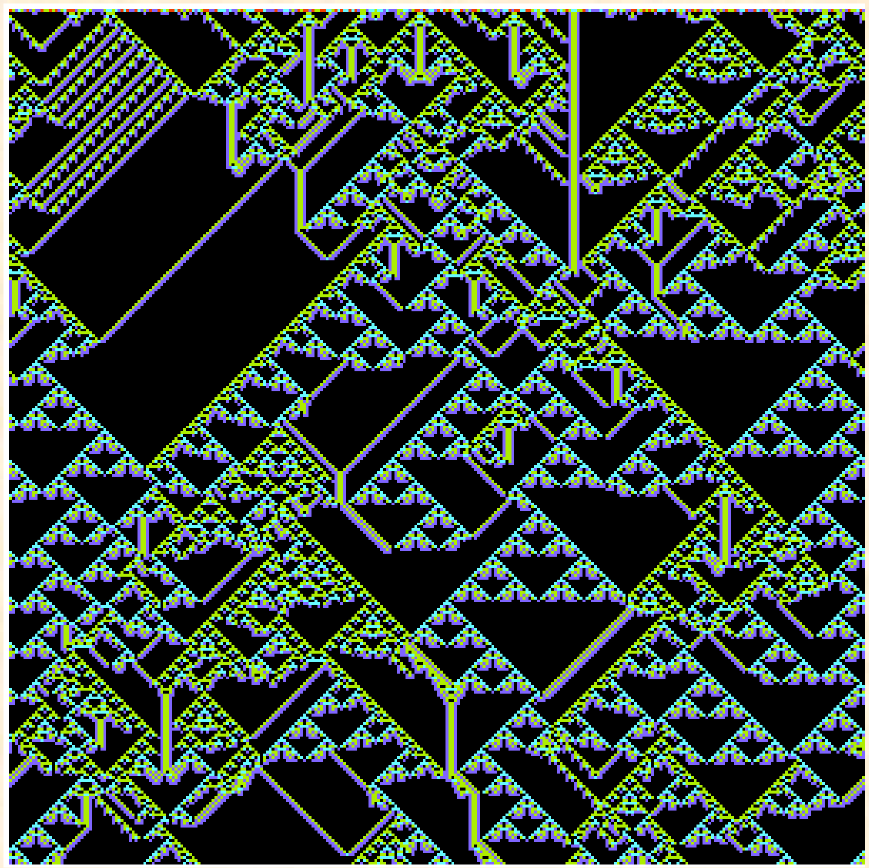
Class IV ($\lambda = 0.35$)

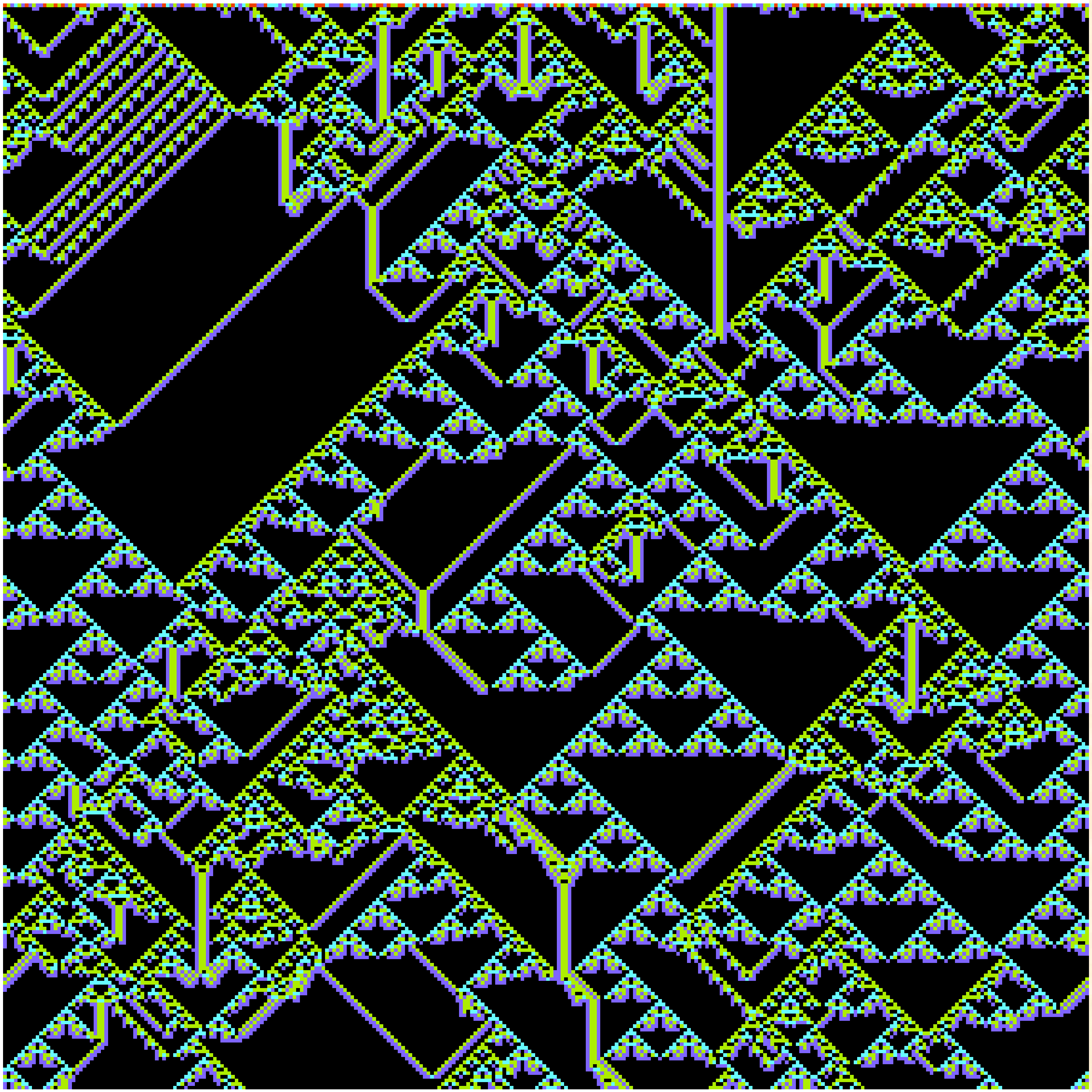


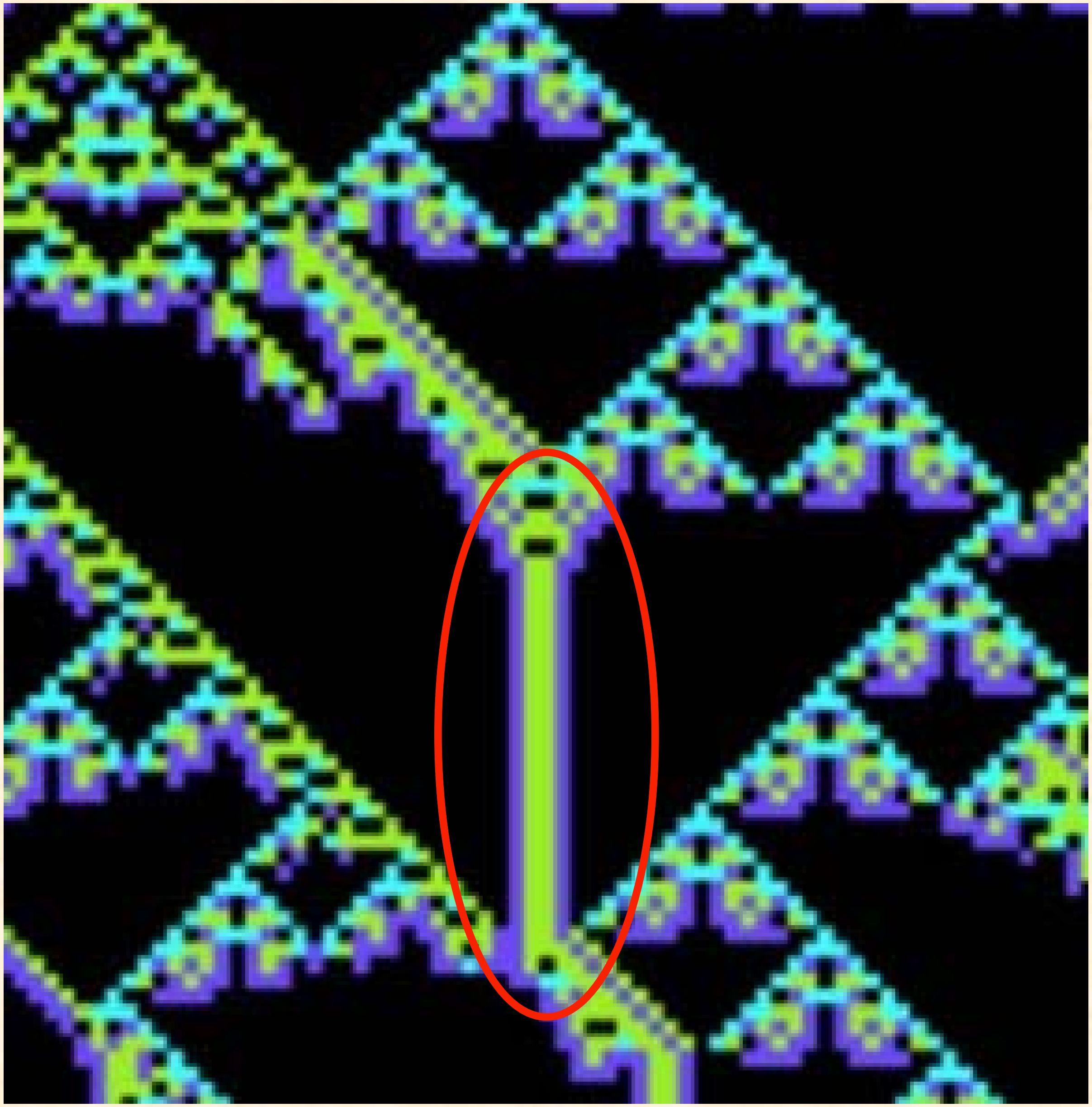
Class IV ($\lambda = 0.35$) Closeup

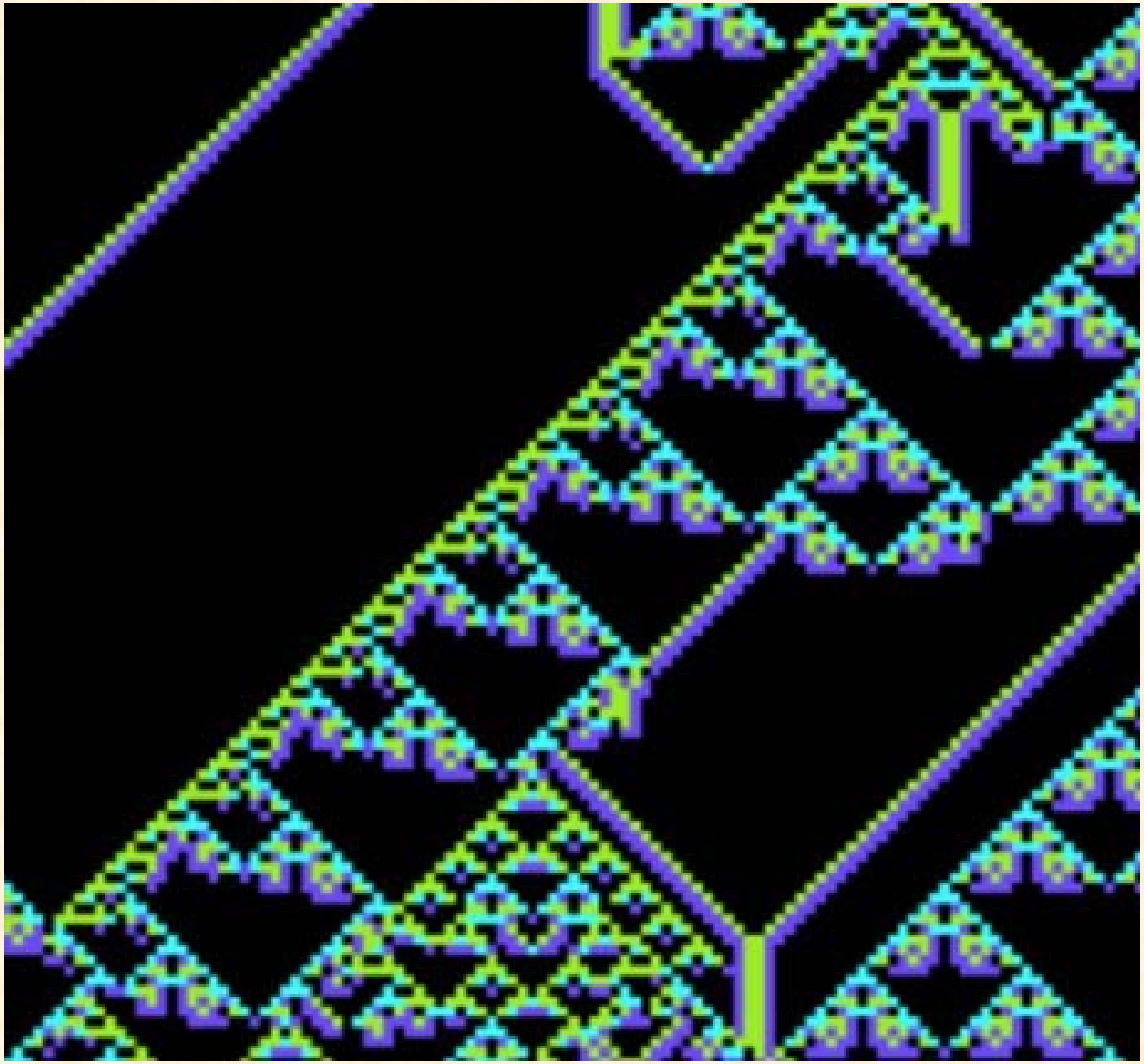


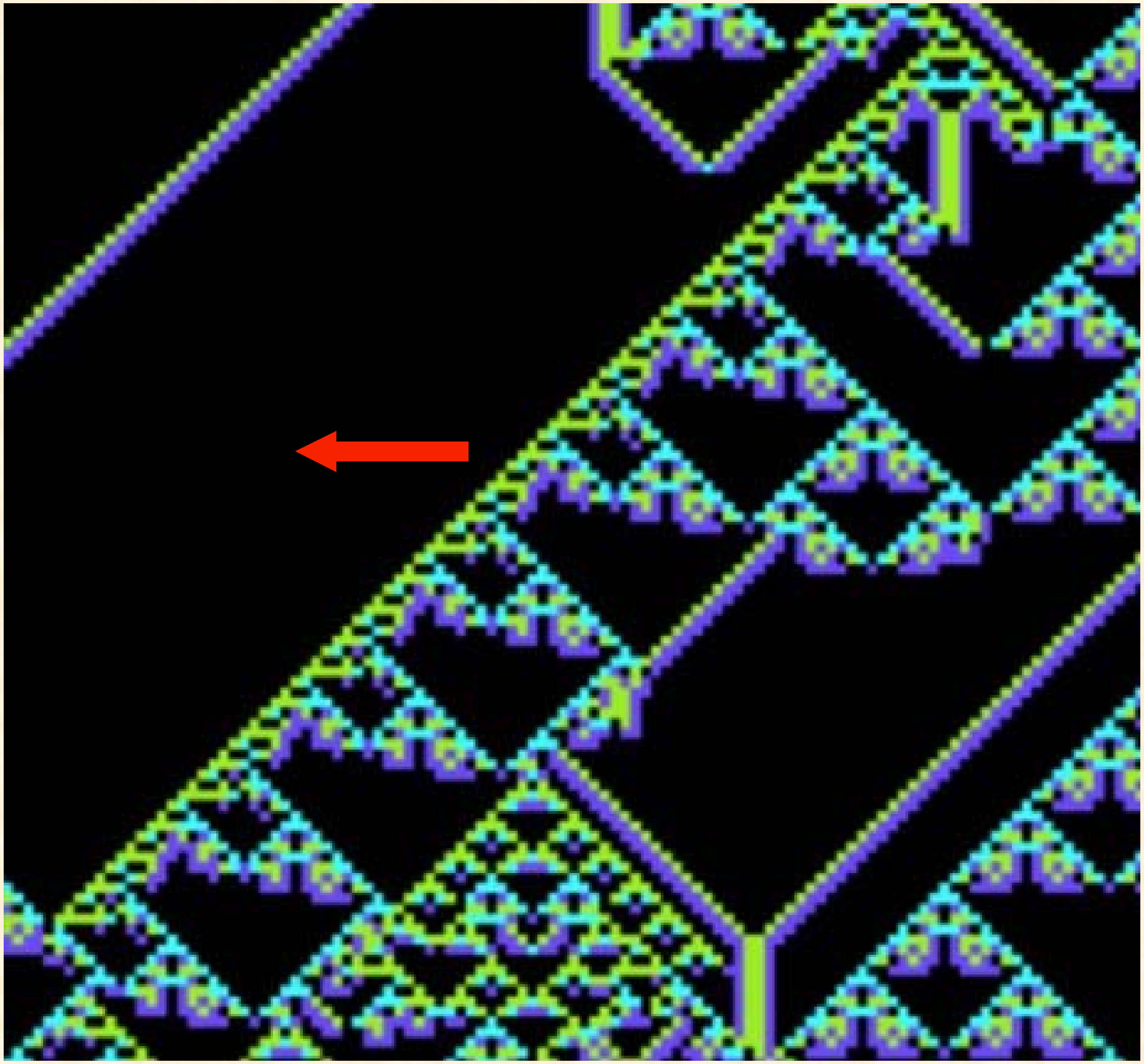
Class IV ($\lambda = 0.34$)

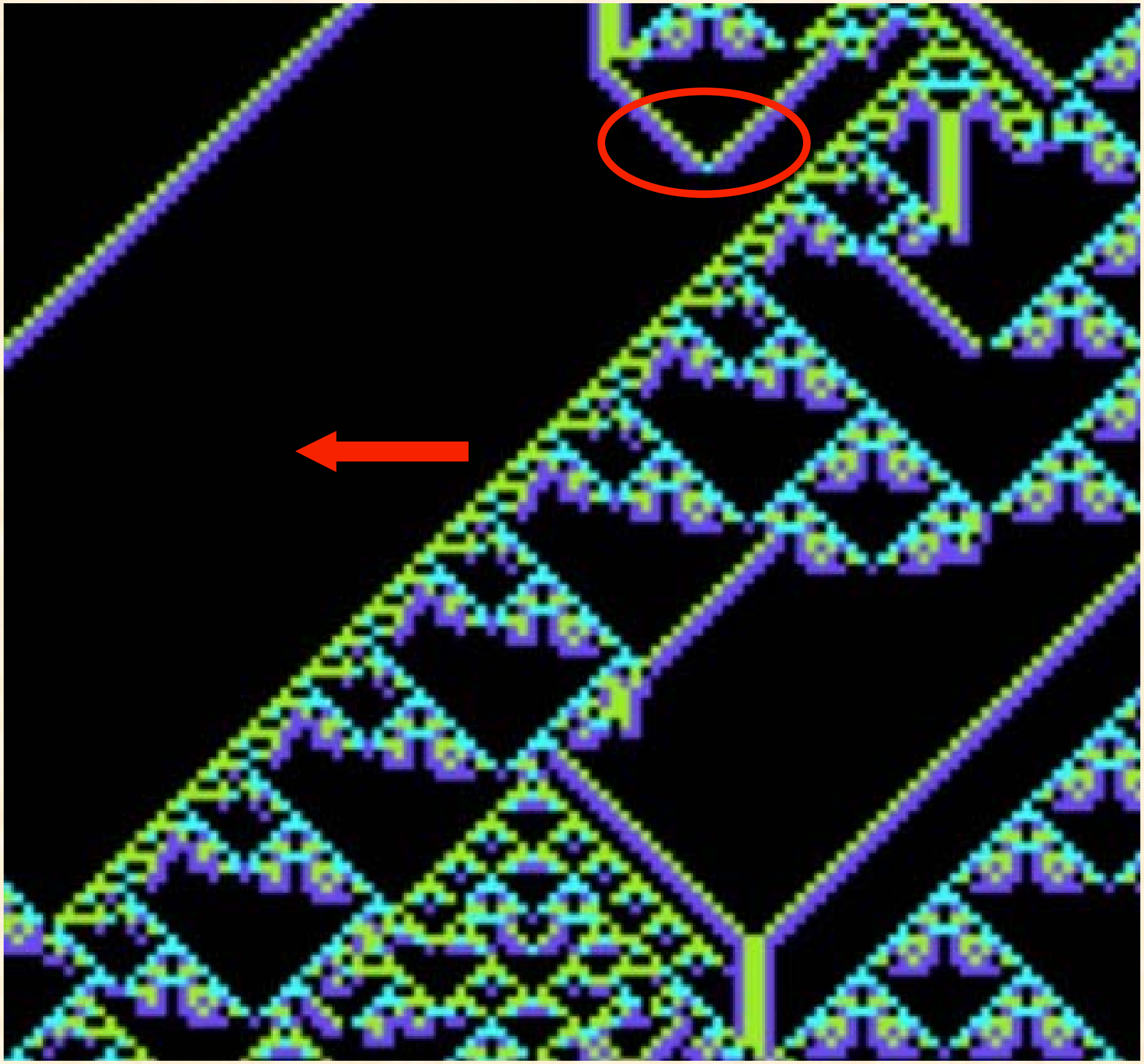


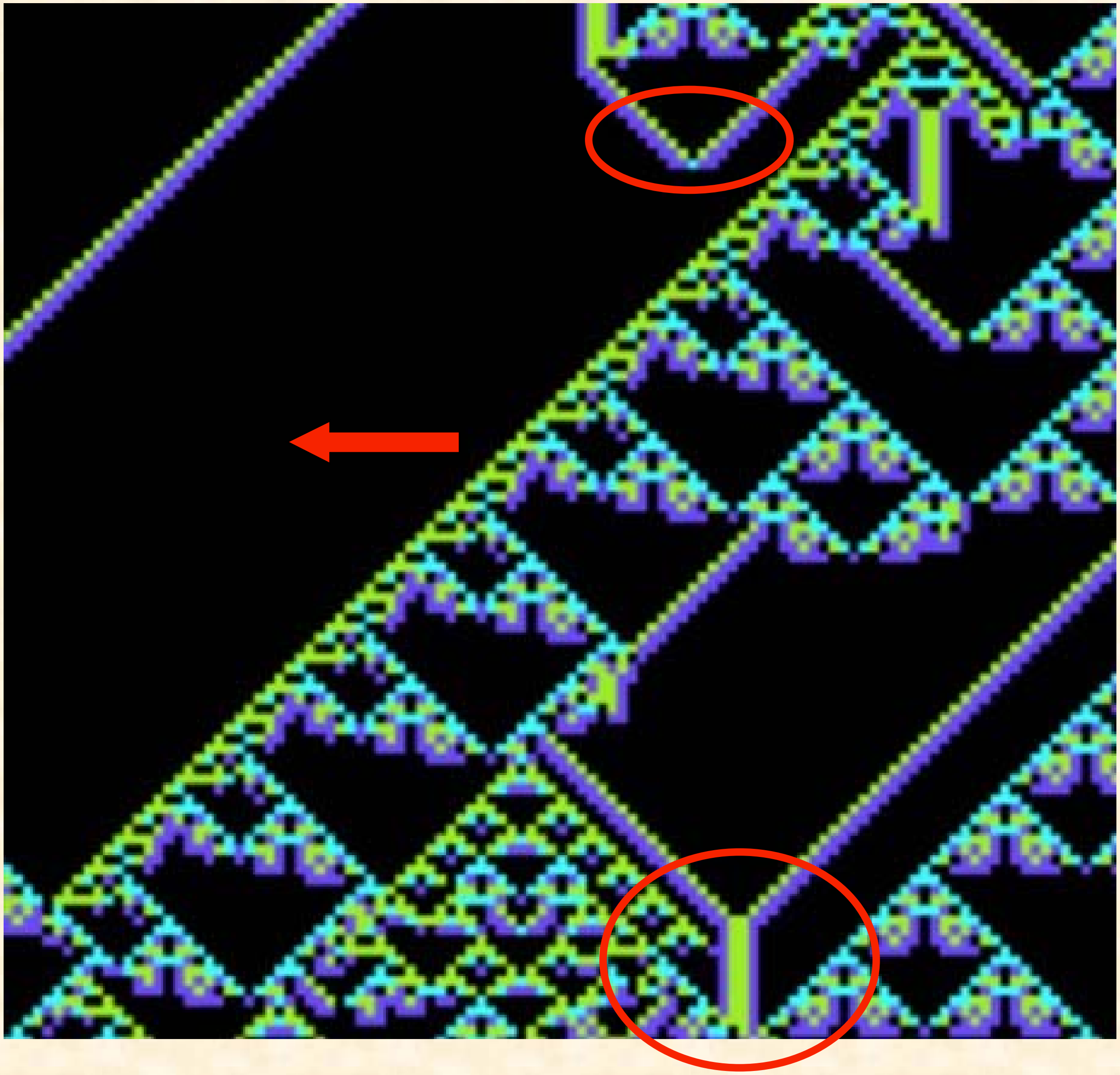


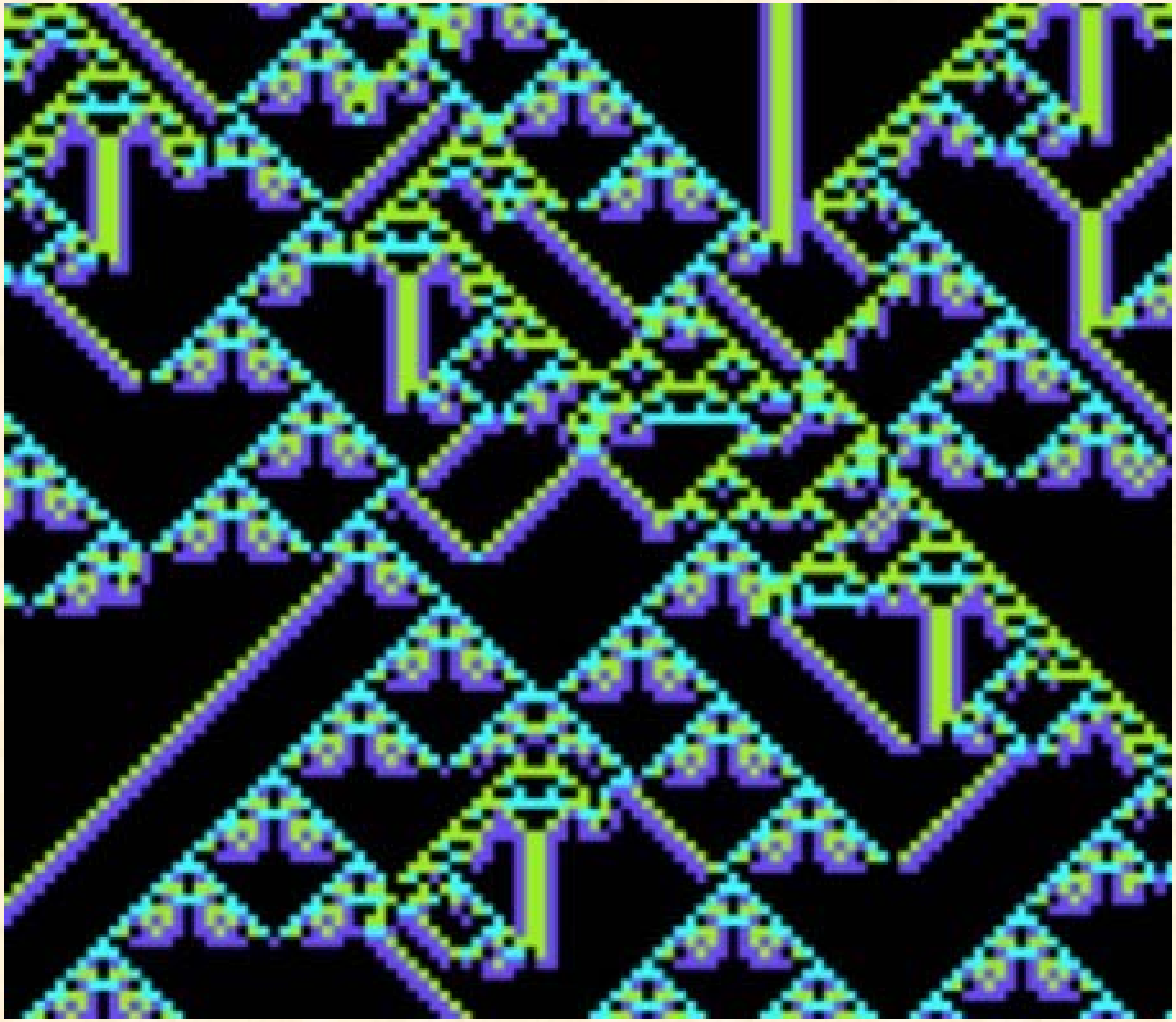


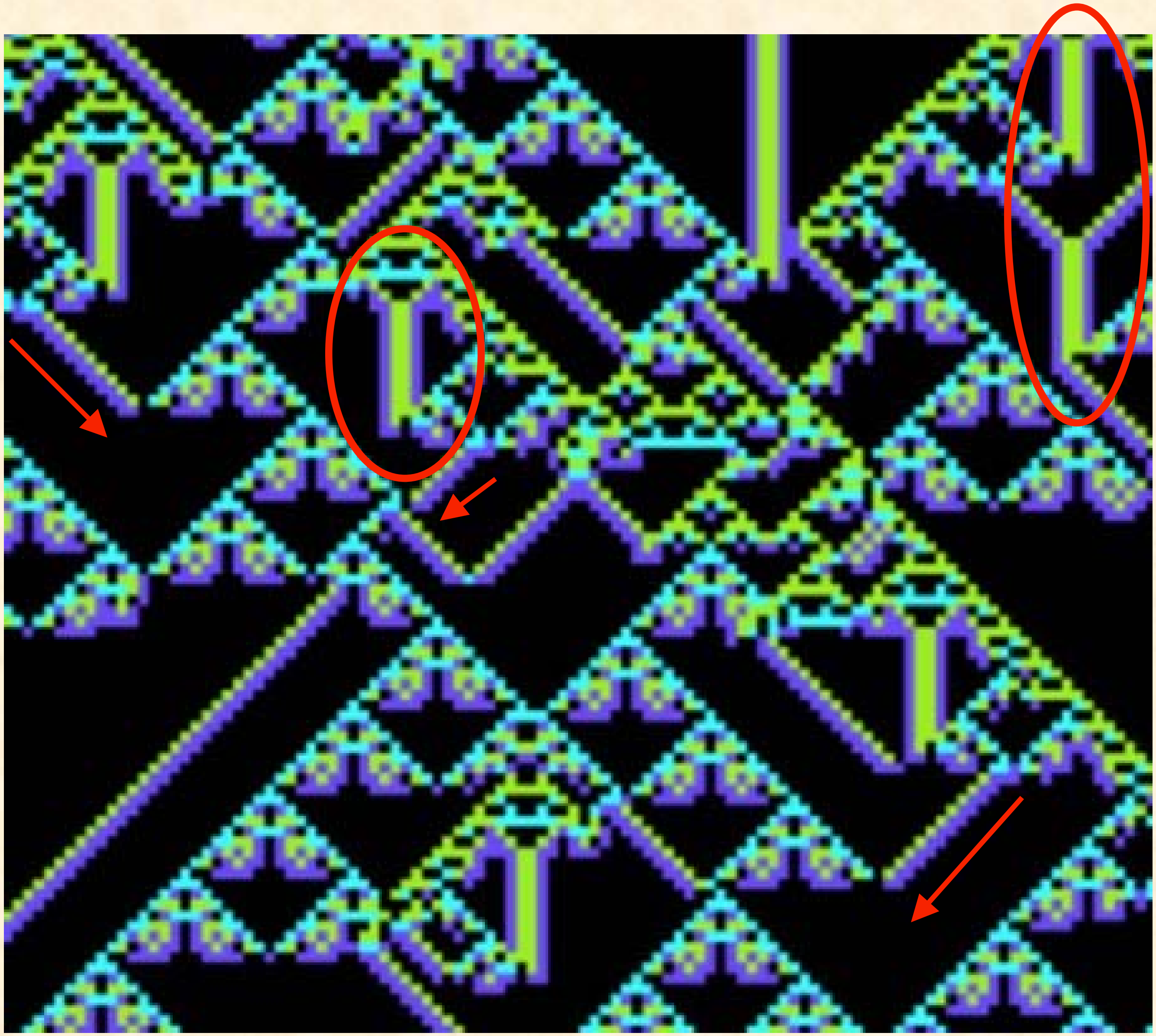












Class IV Shows Some of the Characteristics of Computation

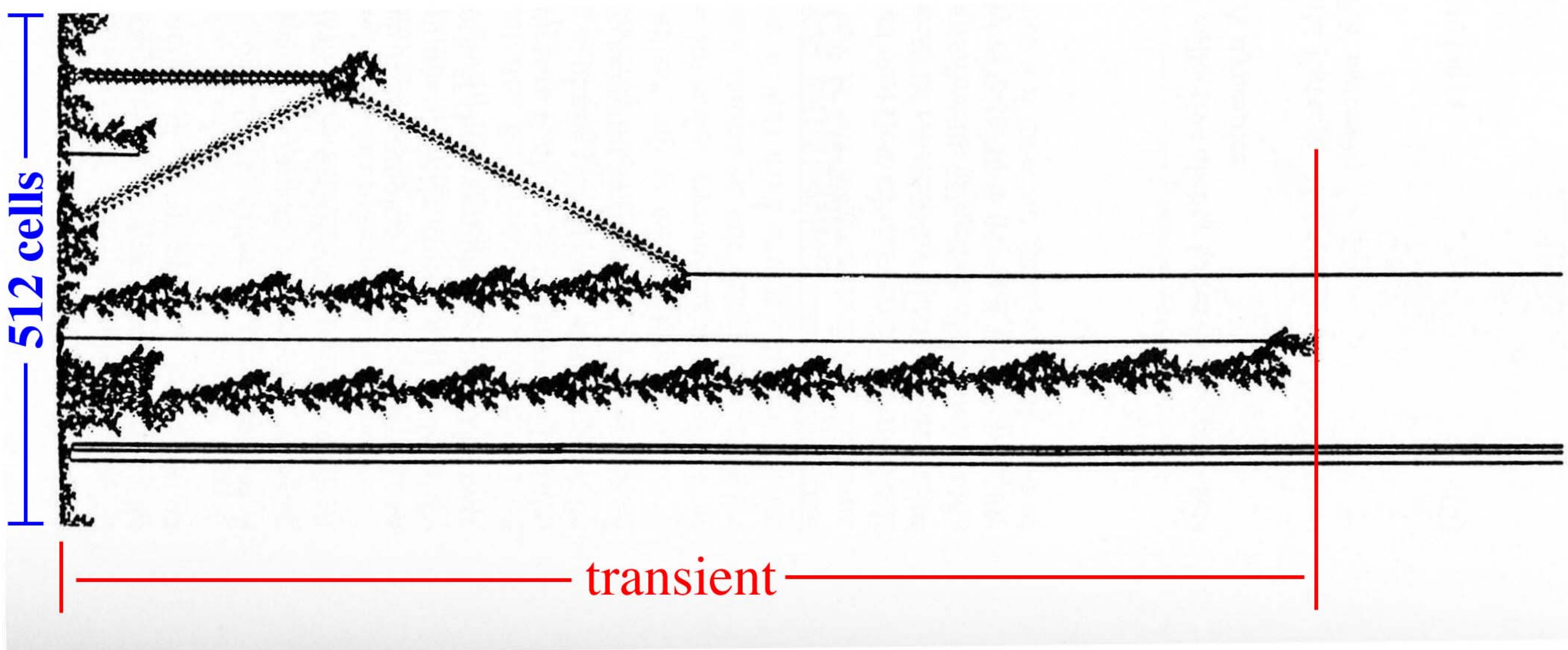
- Persistent, but not perpetual storage
- Terminating cyclic activity
- Global transfer of control/information

λ of Life

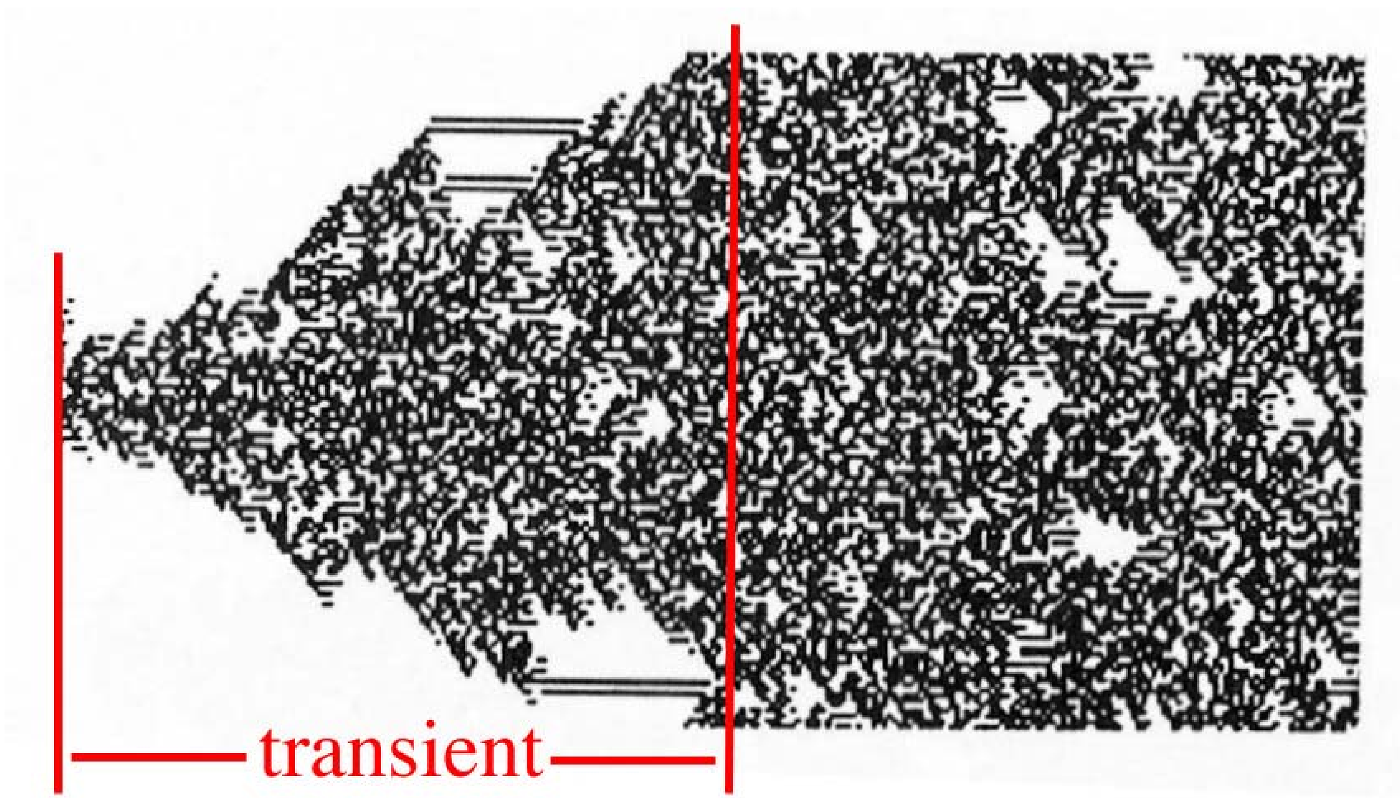
- For Life, $\lambda \approx 0.273$
- which is near the critical region for CAs with:

$$K = 2$$
$$N = 9$$

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

Shannon Information
(very briefly!)

- Information varies directly with surprise
- Information varies inversely with probability
- Information is additive
- $\therefore$ 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

  $H_{\max} = \lg N$

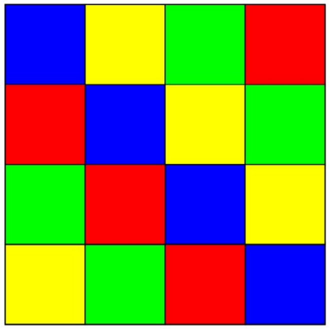
- Minimum entropy occurs when one symbol is certain and the others are impossible

  No uncertainty; no surprise

  $H_{\min} = 0$

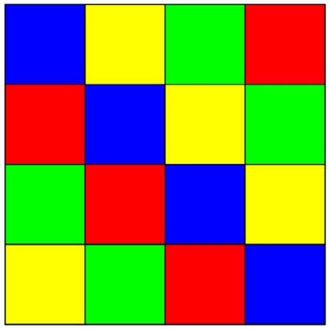
Entropy Examples

Entropy Examples

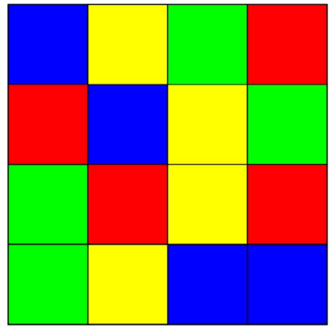


H **= 2.0 bits**

Entropy Examples

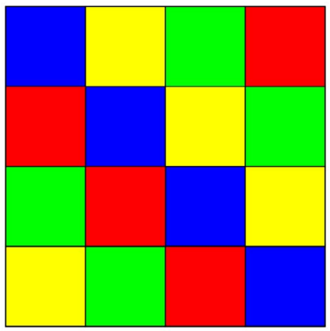


H = 2.0 bits

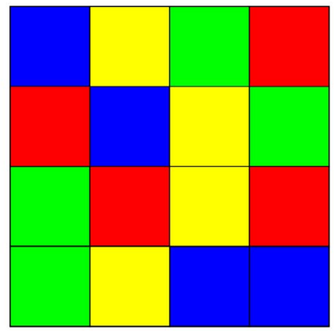


H = 2.0 bits

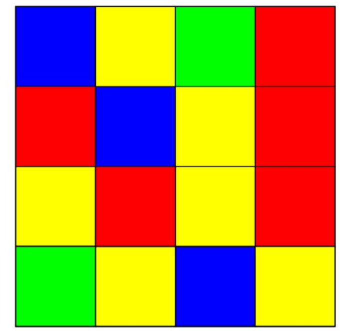
Entropy Examples



H = 2.0 bits

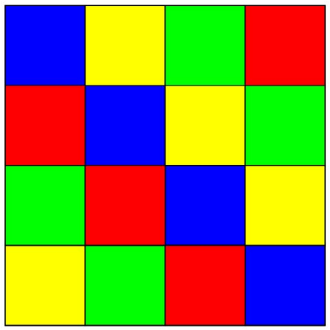


H = 2.0 bits

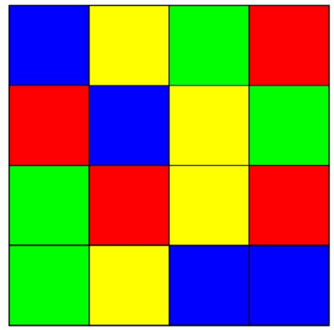


H = 1.9 bits

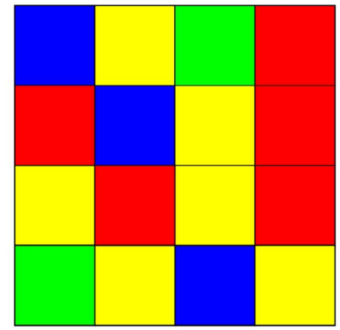
Entropy Examples



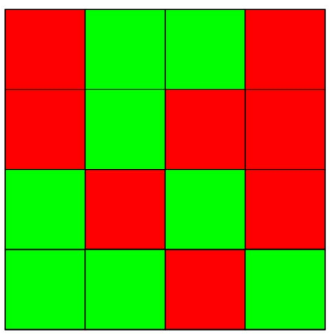
***H* = 2.0 bits**



***H* = 2.0 bits**

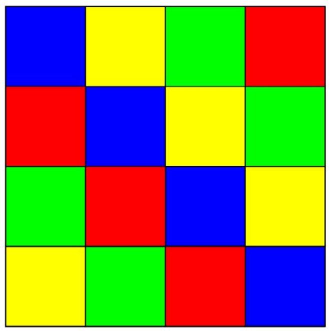


***H* = 1.9 bits**

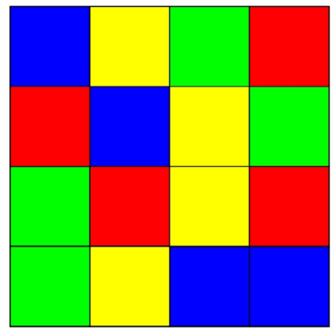


***H* = 1.0 bits**

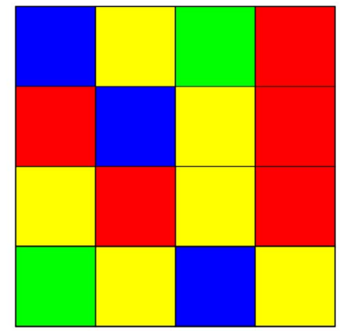
Entropy Examples



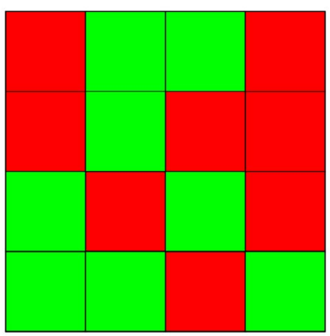
H = 2.0 bits



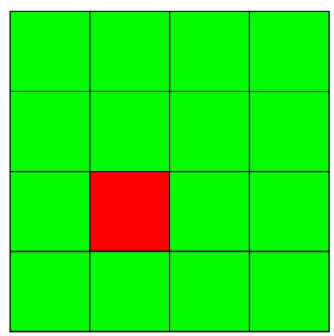
H = 2.0 bits



H = 1.9 bits

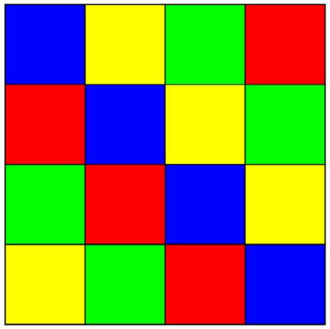


H = 1.0 bits

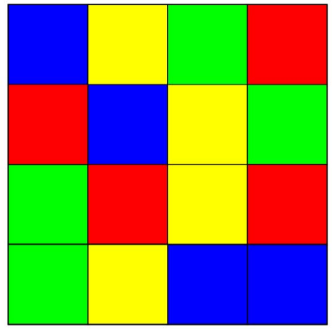


H = 0.3 bits

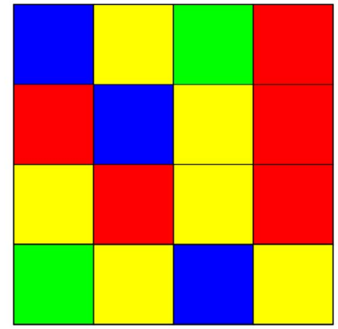
Entropy Examples



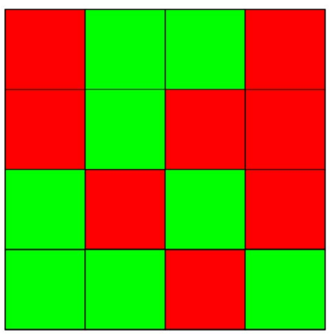
H = 2.0 bits



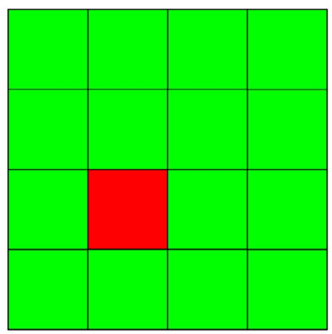
H = 2.0 bits



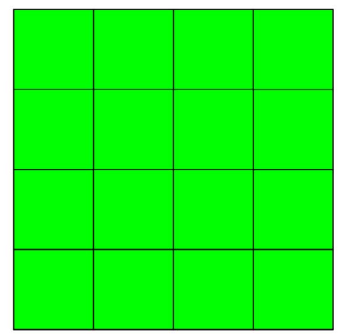
H = 1.9 bits



H = 1.0 bits



H = 0.3 bits



H = 0.0 bits

Entropy of Transition Rules

- Among other things, 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

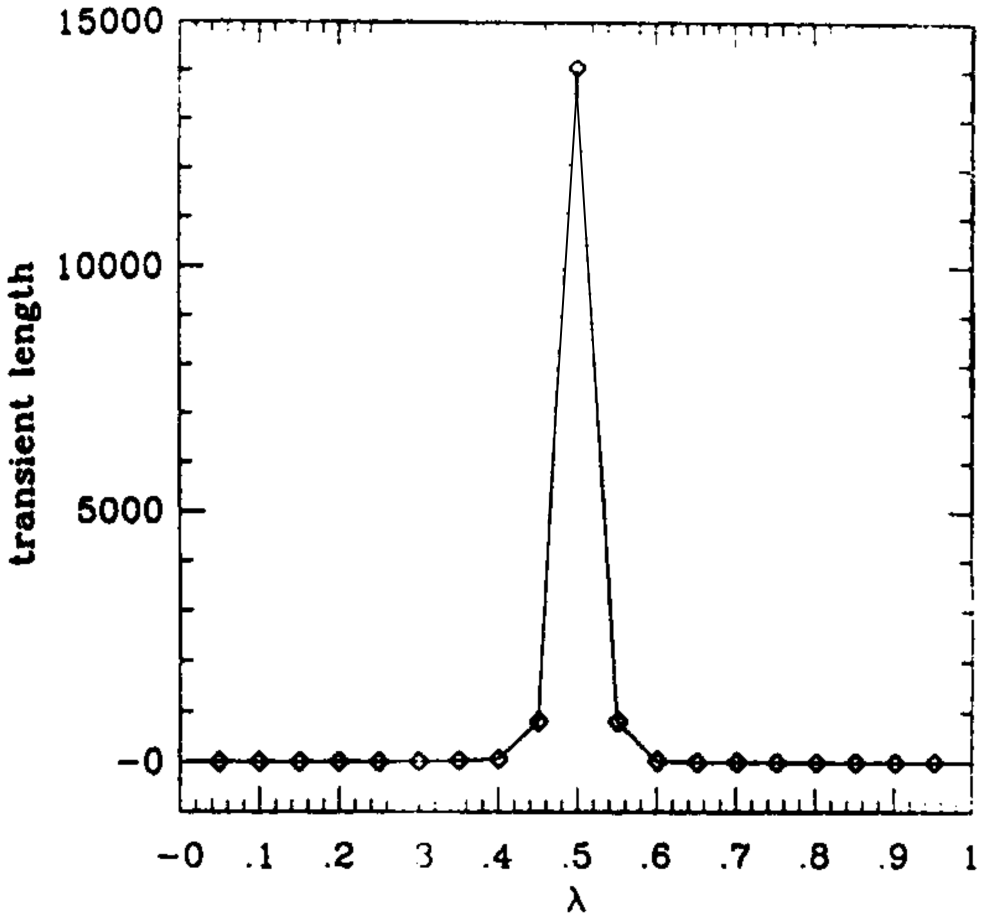
- Maximum entropy ($\lambda = 1 - 1/K$):
  uniform as possible
  all $n_k = T/K$
  $H_{\max} = \lg K$

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

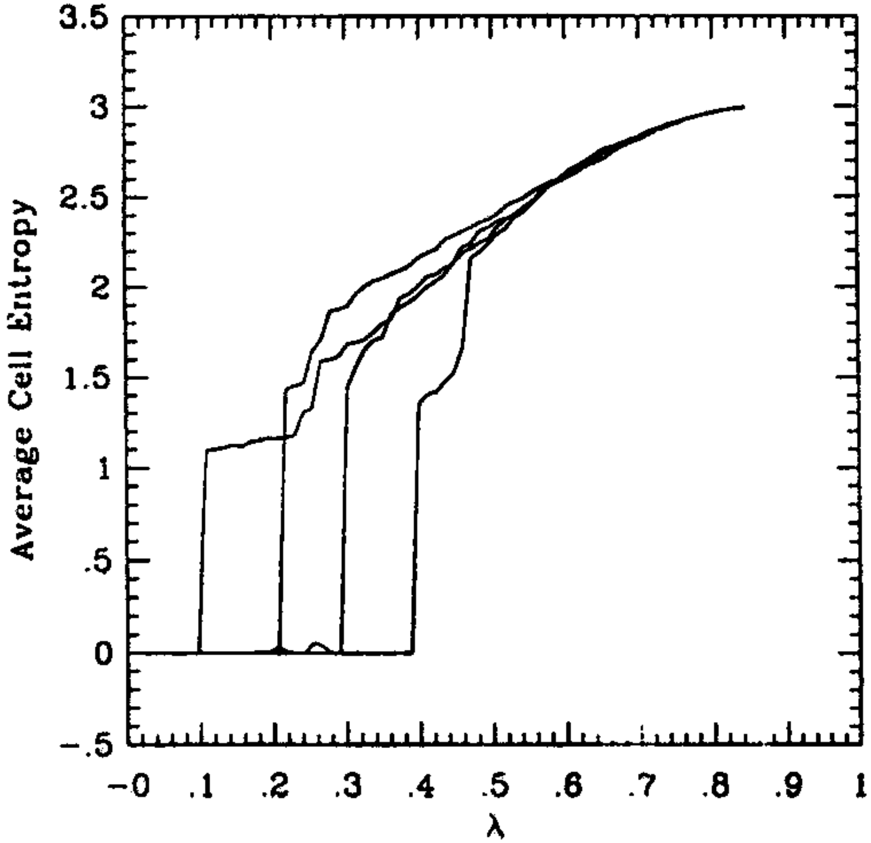
Further Investigations by Langton

- 2-D CAs
- $K = 8$
- $N = 5$
- 64×64 lattice
- periodic boundary conditions

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

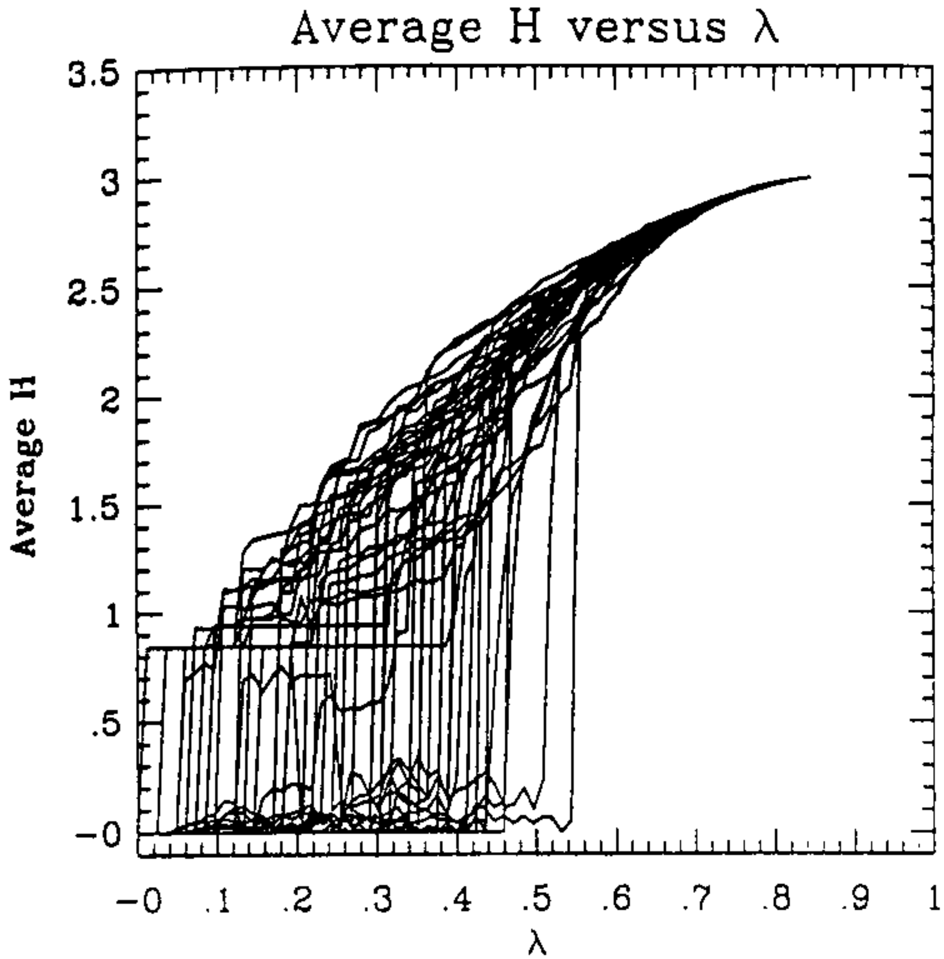


Avg. Cell Entropy vs. λ (*K*=8, *N*=5)

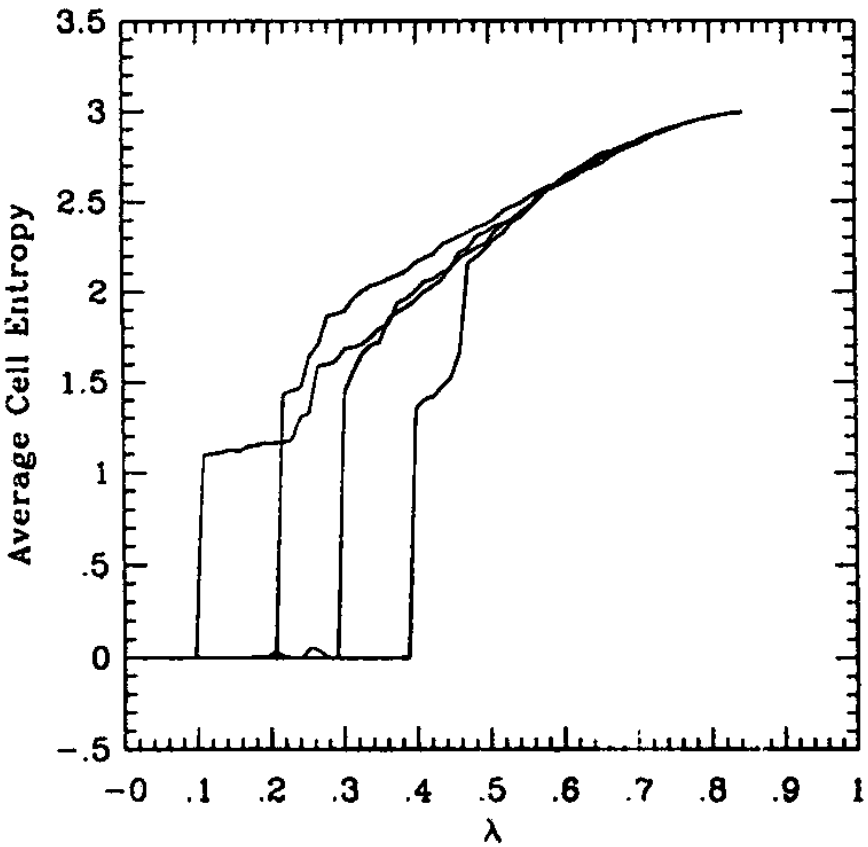


$$H(A) = \sum_{k=1}^{K} p_k \lg p_k$$

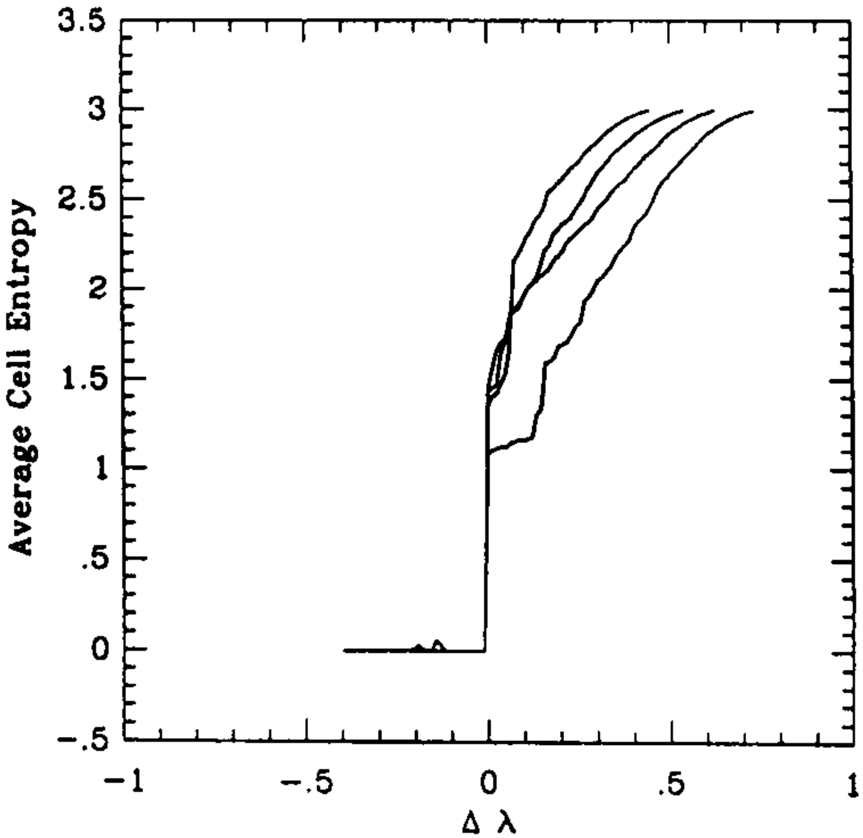
Avg. Cell Entropy vs. λ (K=8, N=5)



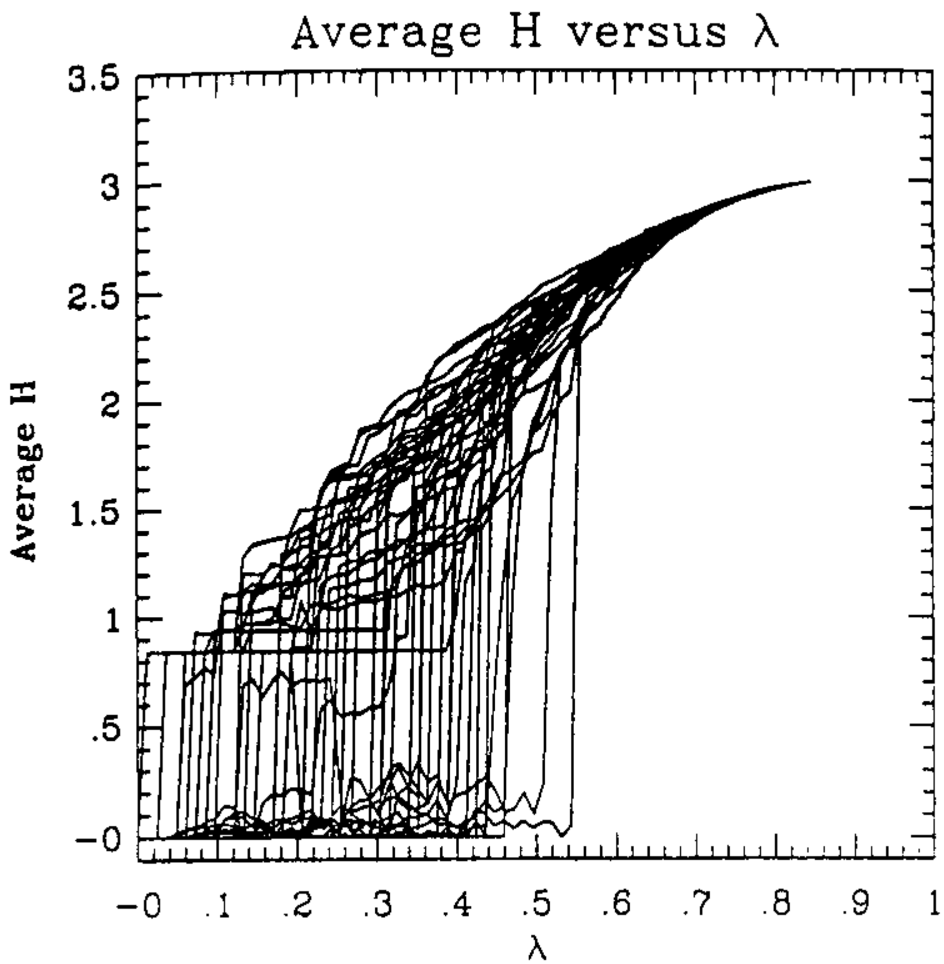
Avg. Cell Entropy vs. λ (K=8, N=5)



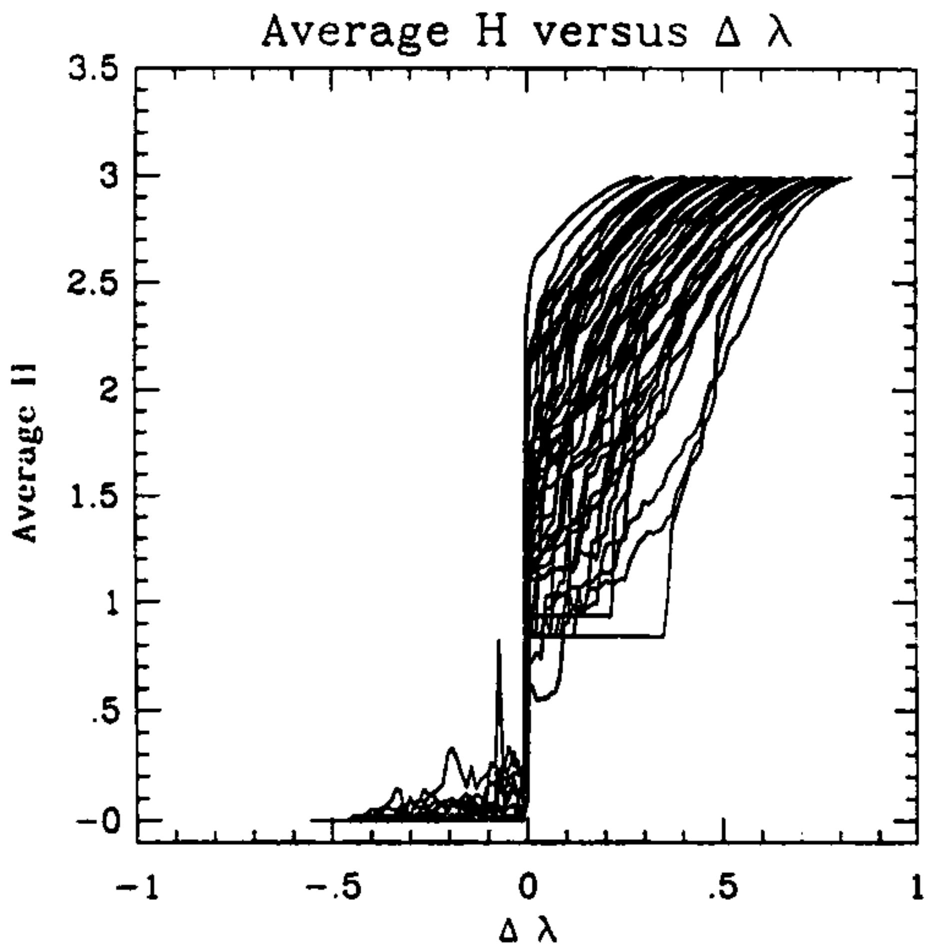
Avg. Cell Entropy vs. Δλ (K=8, N=5)



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 $\Pr\{a_j, b_k\} = \Pr\{a_j\} \Pr\{b_k\} = p_j q_k$

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

Mutual Information

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

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

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

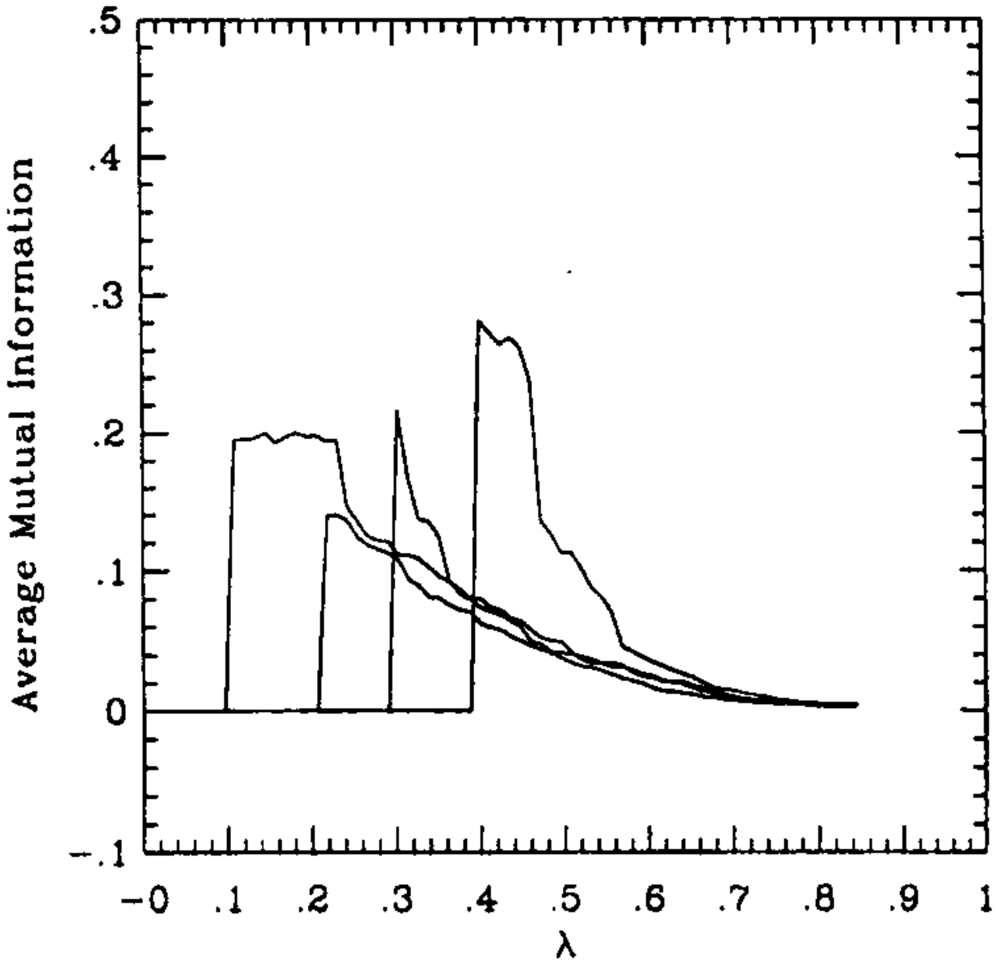
Mutual Information

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

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

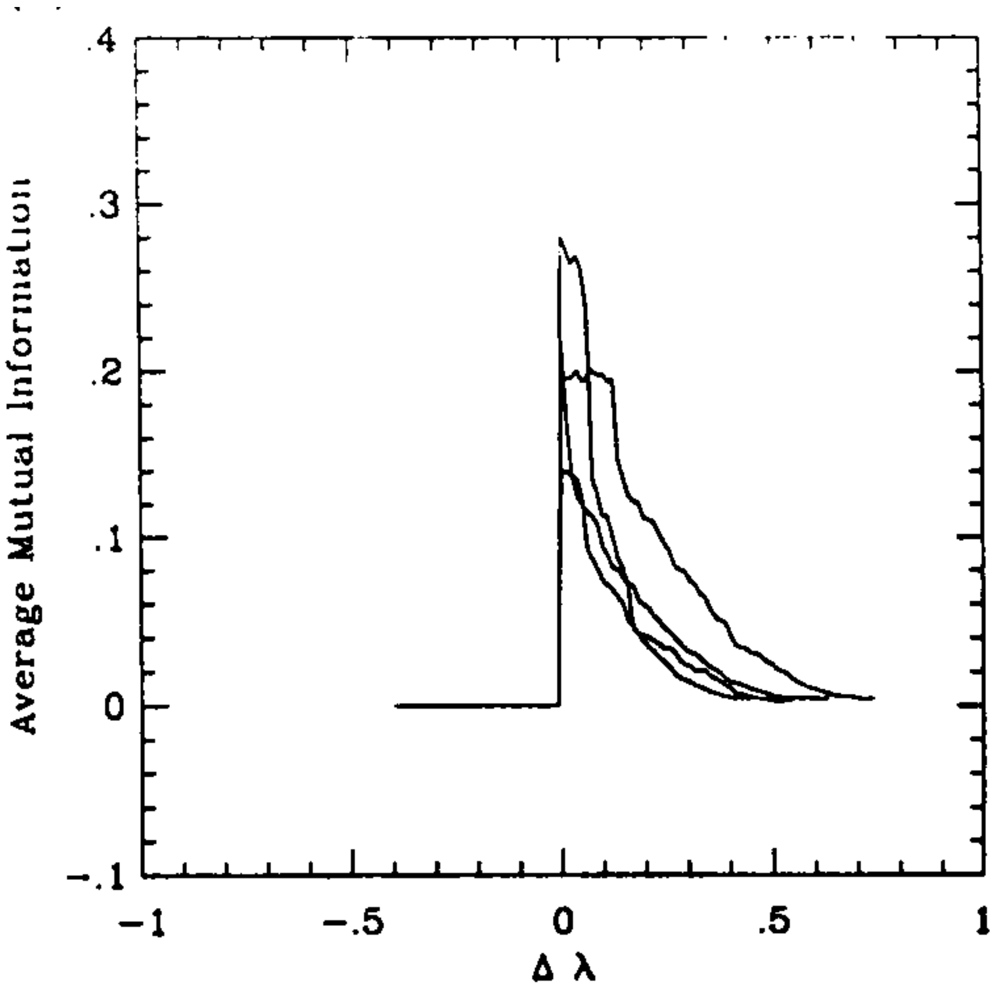
- $I(A,B) = 0$ for completely independent sources
- $I(A,B) = H(A) = H(B)$ for completely correlated sources

Avg. Mutual Info vs. λ
(K=4, N=5)

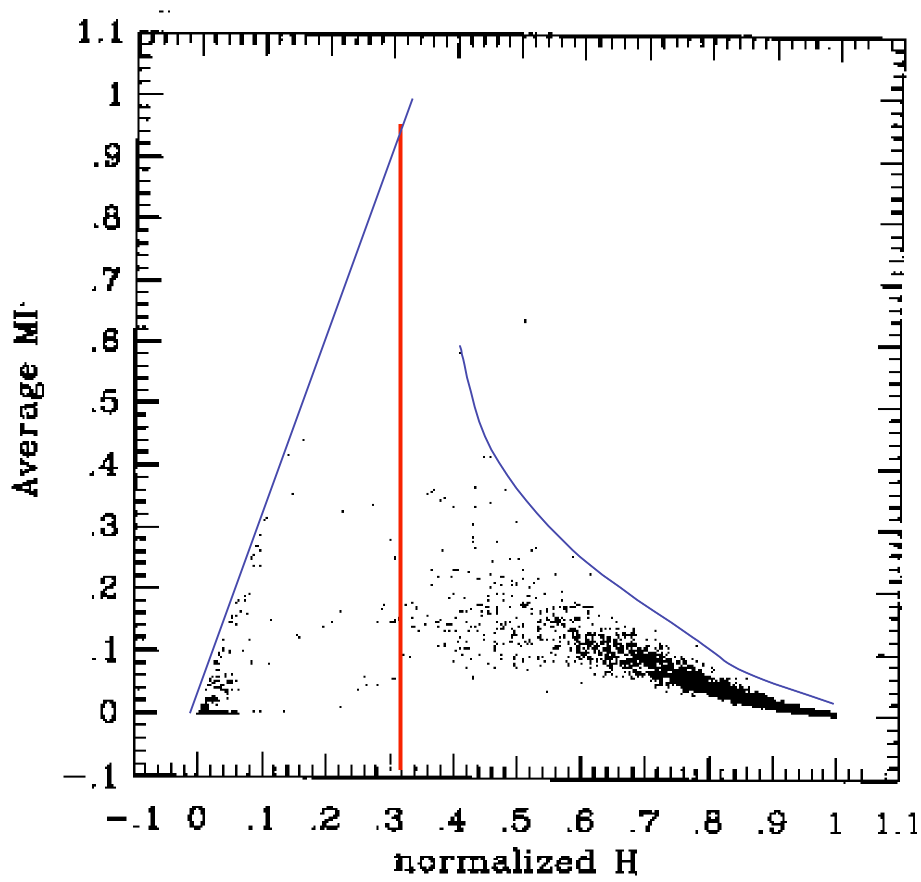


$I(A,B) =$
$H(A) + H(B)$
$- H(A,B)$

Avg. Mutual Info vs. Δλ
(K=4, N=5)



Mutual Information vs. Normalized Cell Entropy



Critical Entropy Range

Critical Entropy Range

- Information *storage* involves *lowering* entropy

Critical Entropy Range

- Information *storage* involves *lowering* entropy
- Information *transmission* involves *raising* entropy

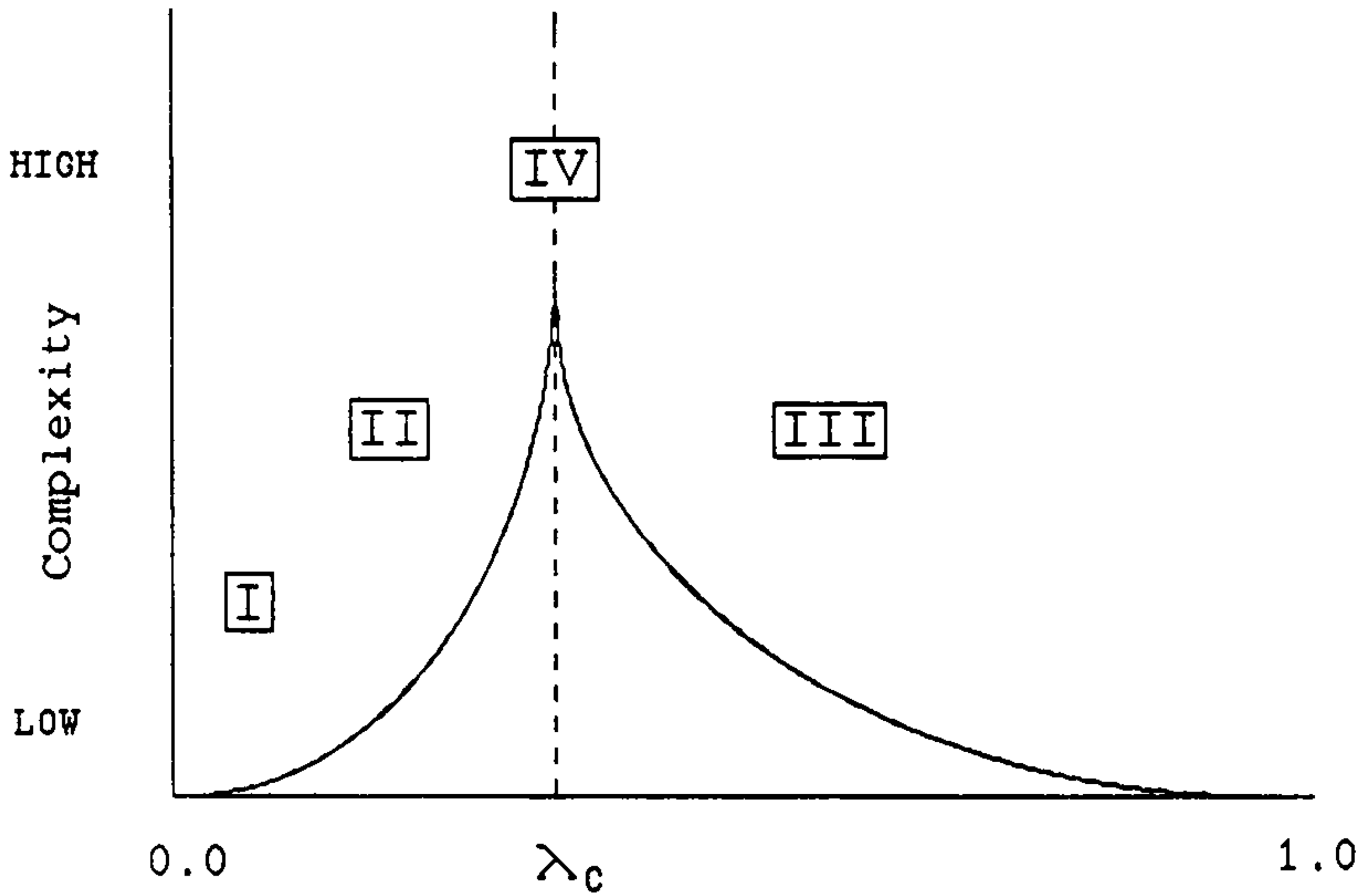
Critical Entropy Range

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

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

Complexity vs. λ



Schematic of CA Rule Space vs. λ

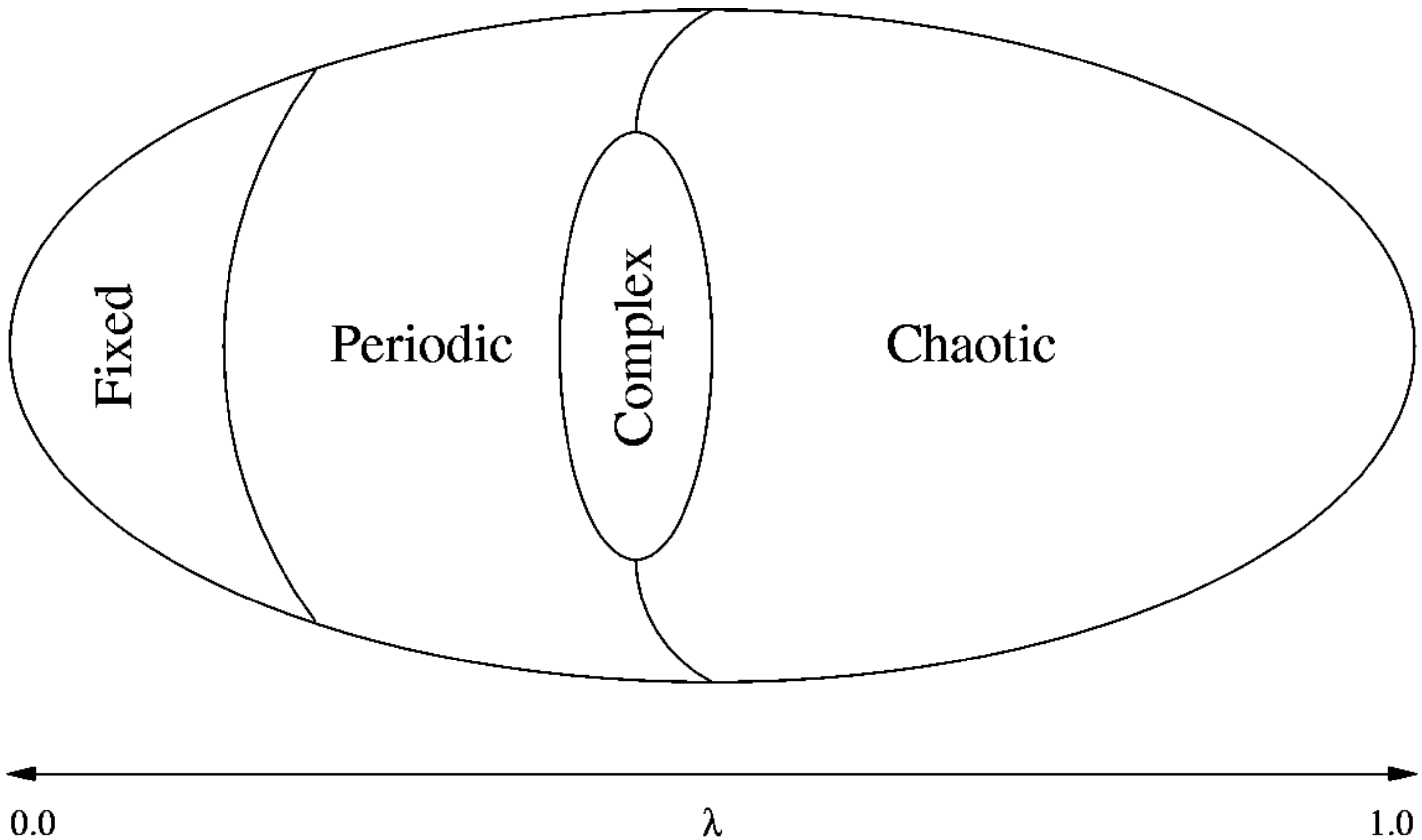


Fig. from Langton, “Life at Edge of Chaos”

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

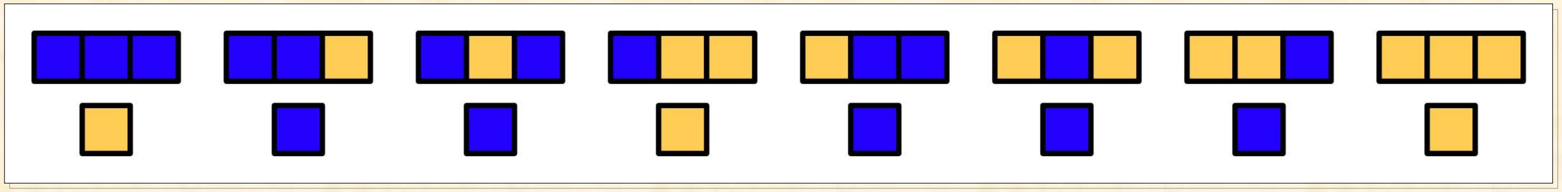
Universality

- 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?

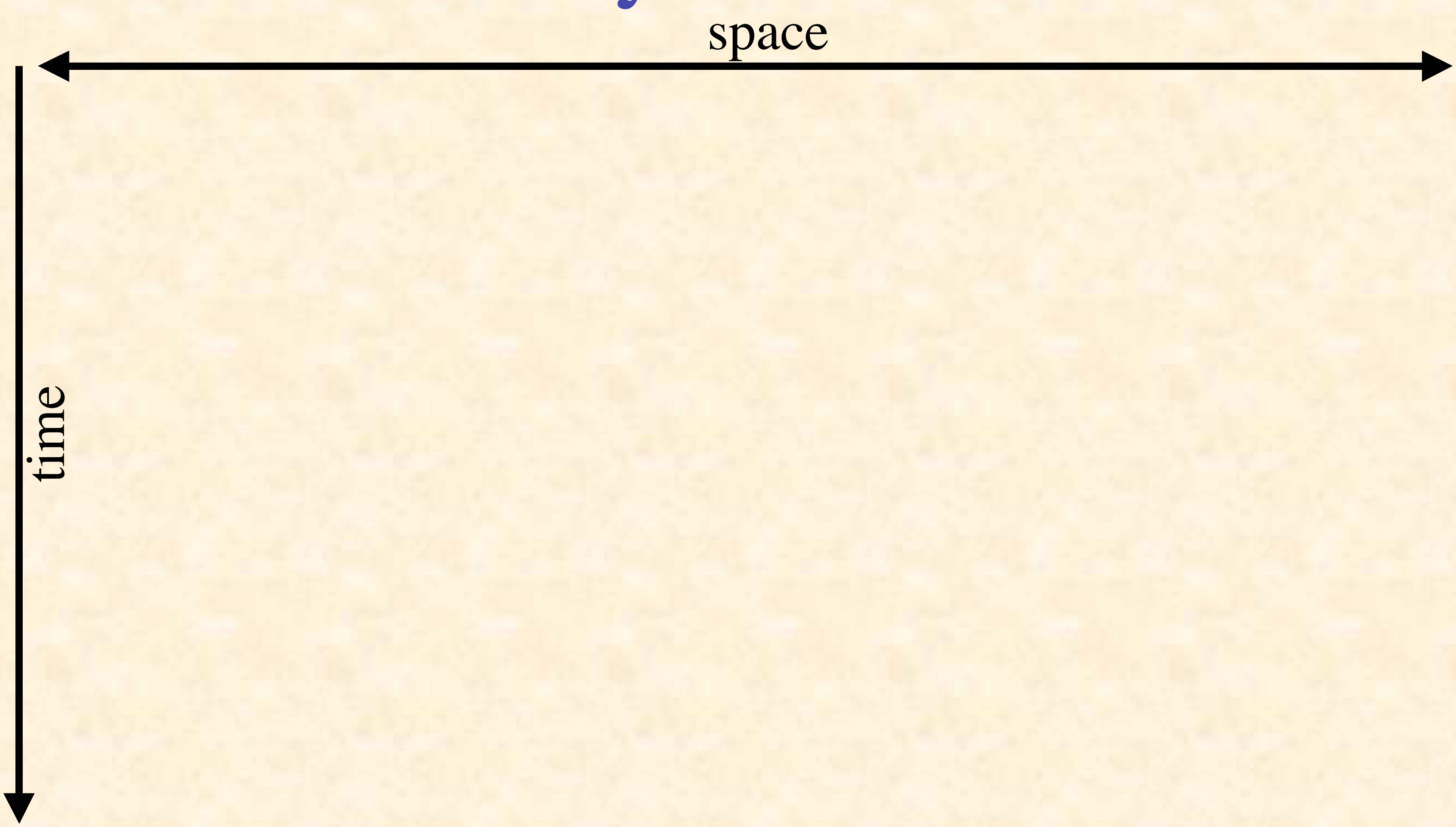
Rule 110: A Universal 1D CA

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

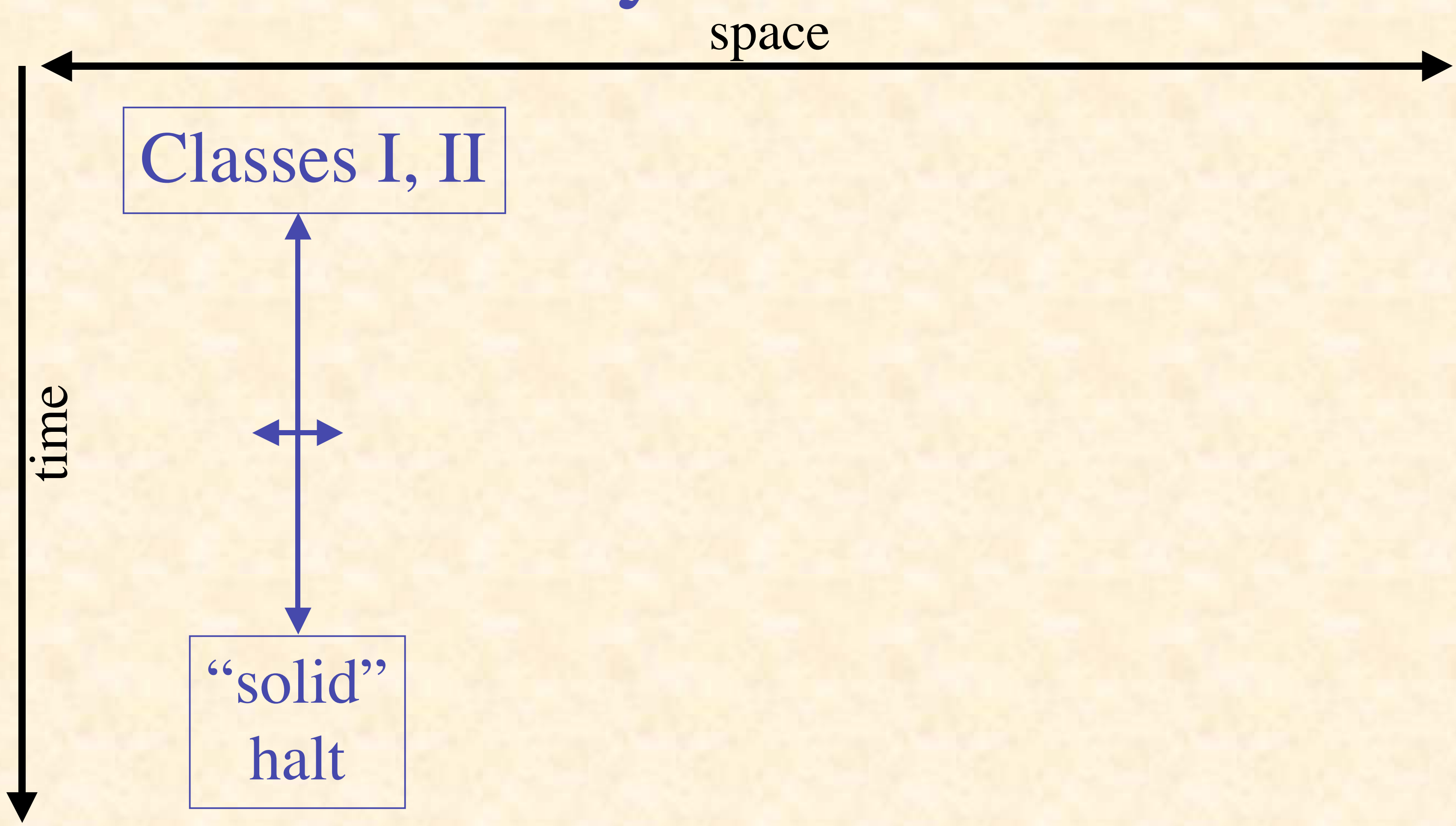
  where p, q, r are neighborhood states
- Proved by Wolfram



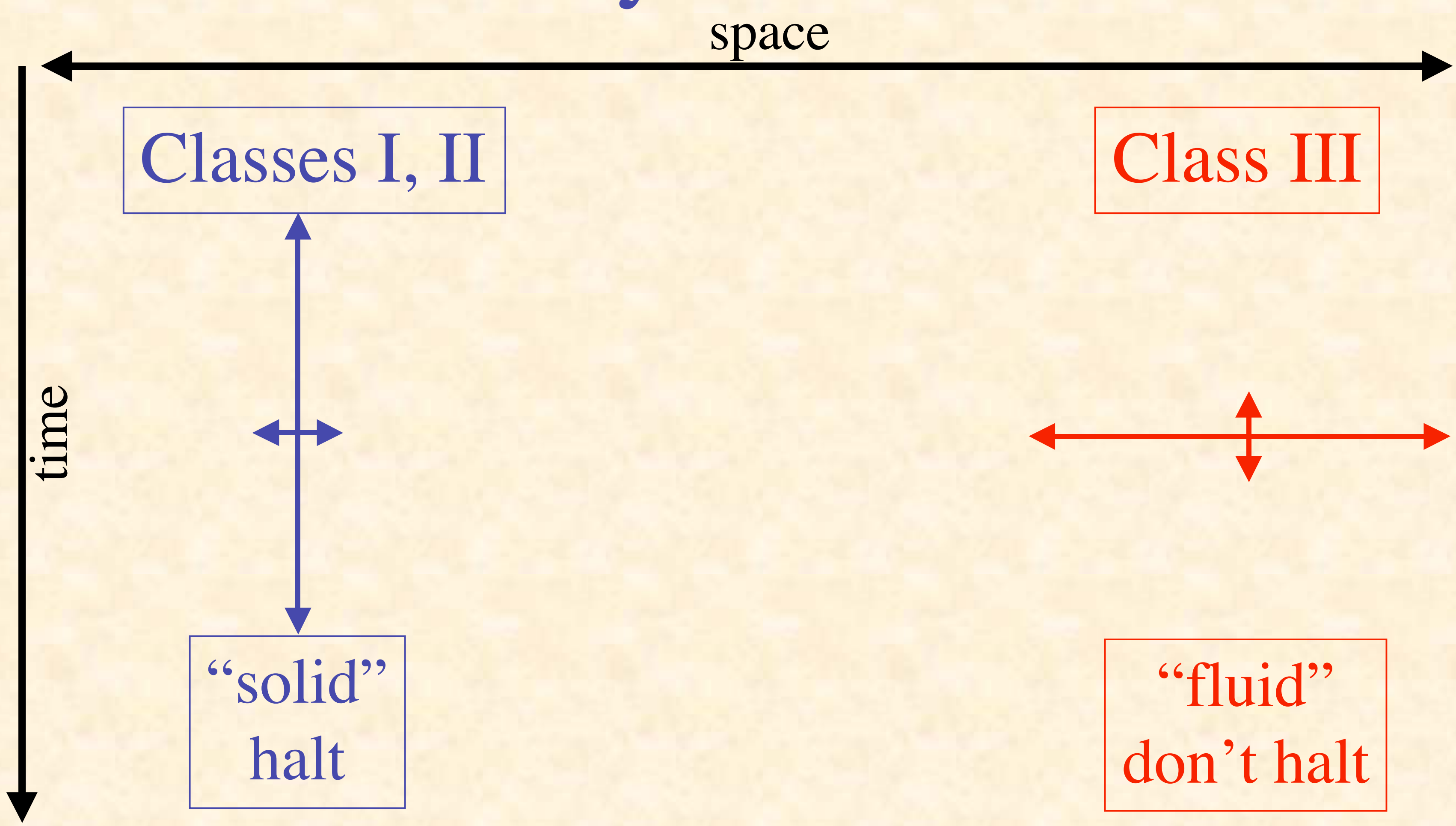
Fundamental Universality Classes of Dynamical Behavior



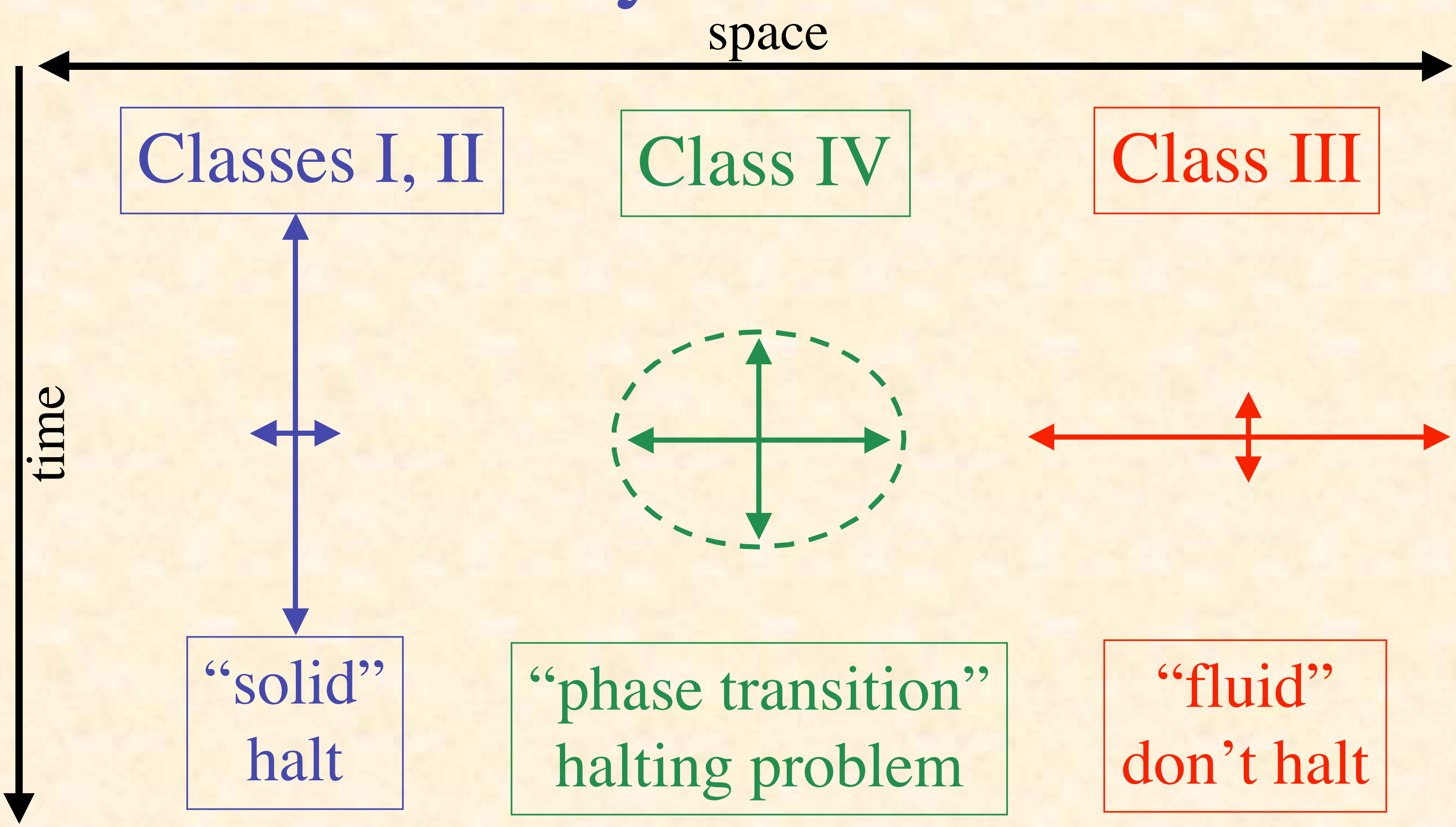
Fundamental Universality Classes of Dynamical Behavior



Fundamental Universality Classes of Dynamical Behavior



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.

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.

Project 1

- Investigation of relation between Wolfram classes, Langton's λ, and entropy in 1D CAs
- **Due Sept. 12**
- Information is on course website (scroll down to "Projects/Assignments")
- Read it over and email questions or ask in class