Surjectivity and reversibility in cellular automata: A review

Silvio Capobianco

Institute of Cybernetics at TUT
silvio@cs.ioc.ee

Kääriku, 31 January 2008

Revised: February 4, 2009

Overview

- Cellular automata (CA) are synchronous distributed systems where the next state of each device only depends on the current state of its neighbors.
- Their implementation on a computer is straightforward, making them very good tools for simulation and qualitative analysis.
- It is instead very difficult to recover the properties of the global dynamics by only looking at the local description.

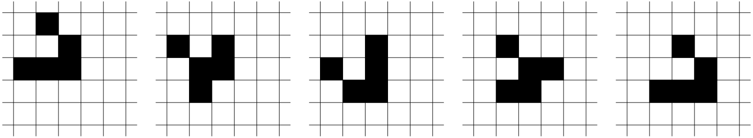
Life is a Game

Ideated by John Horton Conway (1960s) popularized by Martin Gardner.
The checkboard is an infinite square grid.
Each case of the checkboard is “surrounded” by those within a chess' king's move, and can be “living” or “dead”.

1. A “dead” case surrounded by exactly three living cases, becomes living.
2. A living case surrounded by two or three living cases, survives.
3. A living case surrounded by one or no living cases, dies of isolation.
4. A living case surrounded by four or more living cases, dies of overpopulation.

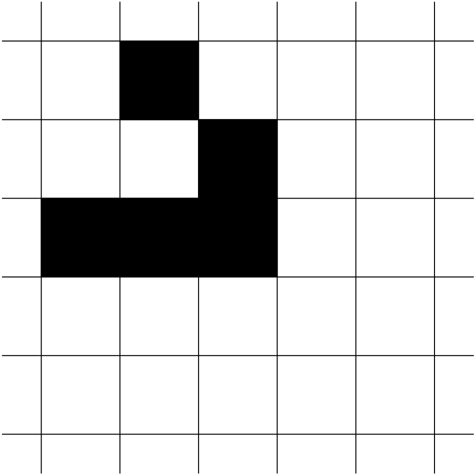
Simple rule, complex behavior

The structures of the Game of Life can exhibit a wide range of behaviors.
This is a glider, which repeats itself every four iterations, after having moved:

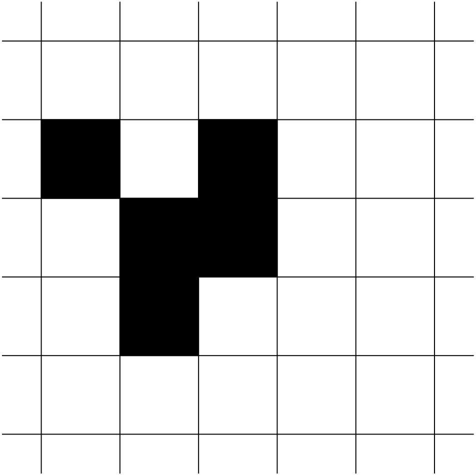


Gliders can transmit information between regions of the checkboard.
Actually, using gliders and other complex structures, any planar circuit can be simulated inside the Game of Life.

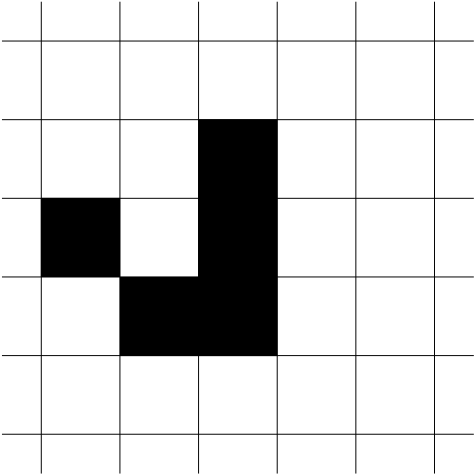
Glider in motion, $t = 0$



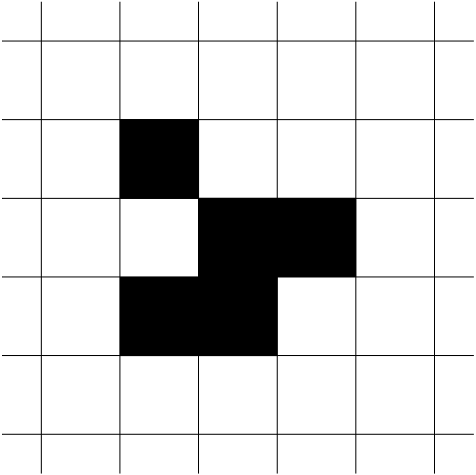
Glider in motion, $t = 1$



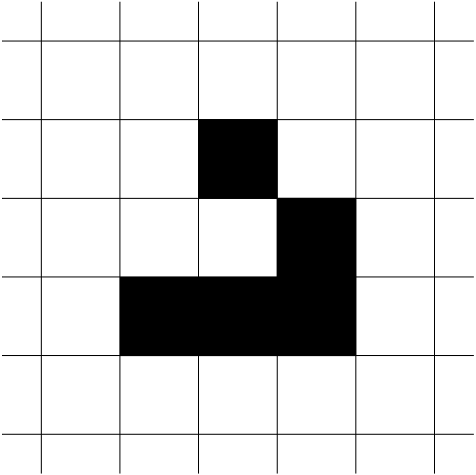
Glider in motion, $t = 2$



Glider in motion, $t = 3$



Glider in motion, $t = 4$



The ingredients of a recipe

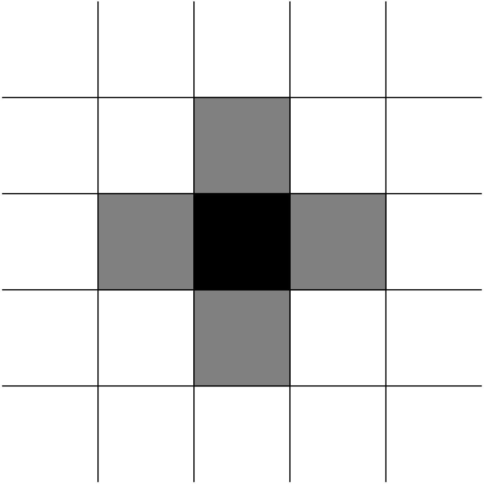
A cellular automaton (CA) is a quadruple $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$ where

- $d > 0$ is an integer—dimension
- $Q = \{q_1, \ldots, q_n\}$ is finite nonempty—set of states
- $\mathcal{N} = \{n_1, \ldots, n_k\}$ is a finite subset of $\mathbb{Z}^d$—neighborhood
- $f : Q^{\mathcal{N}} \to Q$ is a function—local map

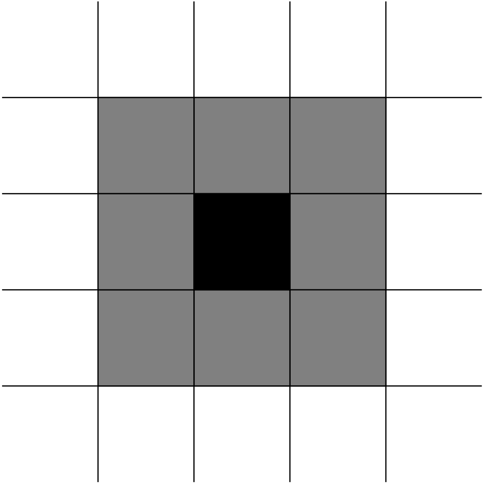
Special neighborhoods are:

- the von Neumann neighborhood of radius r
  $\mathrm{vN}(r) = \{x \in \mathbb{Z}^d \mid \sum_{i=1}^{d} |x_i| \leq r\}$
- the Moore neighborhood of radius r
  $\mathrm{M}(r) = \{x \in \mathbb{Z}^d \mid \max_{1 \leq i \leq d} |x_i| \leq r\}$

For $d = 2$, this is von Neumann's neighborhood $\mathrm{vN}(1)$...



and this is Moore's neighborhood $\mathrm{M}(1)$.



Configurations

A d-dimensional configuration is a map $c : \mathbb{Z}^d \to Q$.
We consider the following distance on configurations:

$$\text{if } c_1 \text{ and } c_2 \text{ differ on } \mathrm{M}(n) \text{ but coincide on } \mathrm{M}(n-1)$$
$$\text{then } d_{\mathrm{M}}(c_1, c_2) = 2^{-n}$$

Two configurations are "near" according to d_{M} iff they are "equal on a large zone around the origin".
d_{M} induces the product topology—which makes $Q^{\mathbb{Z}^d}$ compact.
We also consider translations given by

$$c^x(y) = c(x+y) \text{ for all } y \in \mathbb{Z}^d$$

From local to global

Let $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$ be a CA.
The global map of $\mathcal{A}$ is $F_{\mathcal{A}} : Q^{\mathbb{Z}^d} \to Q^{\mathbb{Z}^d}$ defined by

$$(F_{\mathcal{A}}(c))(x) = f(c(x + n_1), \ldots, c(x + n_k))$$

We say that $\mathcal{A}$ is injective, surjective, etc. if $F_{\mathcal{A}}$ is.

Hedlund's theorem (1969)

Let $F : Q^{\mathbb{Z}^d} \to Q^{\mathbb{Z}^d}$.
The following are equivalent:

1. F is a CA global map
2. F is continuous and commutes with the translations

Reason why:

- translation invariance $\Rightarrow$ only need to determine $F(c)(\mathbf{0})$
- $Q^{\mathbb{Z}^d}$ compact $\Rightarrow$ F uniformly continuous $\Rightarrow$ $F(c)(\mathbf{0})$ only depends on $c|_{\mathrm{M}(n)}$ for n large enough

Consequence: a composition of CA yields a CA.
(This can also be seen from the local rules.)

Hedlund's theorem (1969)

Let $F : Q^{\mathbb{Z}^d} \to Q^{\mathbb{Z}^d}$.
The following are equivalent:

1. F is a CA global map
2. F is continuous and commutes with the translations

Reason why:

- translation invariance $\Rightarrow$ only need to determine $F(c)(\mathbf{0})$
- $Q^{\mathbb{Z}^d}$ compact $\Rightarrow$ F uniformly continuous $\Rightarrow$ $F(c)(\mathbf{0})$ only depends on $c|_{\mathrm{M}(n)}$ for n large enough

Consequence: a composition of CA yields a CA.
(This can also be seen from the local rules.)

Special configurations and states

- Periodic configurations cover the d-dimensional space with a repeated regular pattern.
- q-finite configurations only have finitely many points in states other than q.
- Quiescent states satisfy $f(q, \ldots, q) = q$.
  In this case, we call $\mathcal{A}_{(q)}$ the restriction of $\mathcal{A}$ to q-finite configurations.

The state 0 in Conway's Game of Life is quiescent.

Garden-of-Eden configuration and orphan patterns

A Garden-of-Eden (GoE) for a CA $\mathcal{A}$ is a configuration c that has no predecessor according to the global law of $\mathcal{A}$—that is,

$$F_{\mathcal{A}}(c') \neq c \quad \forall c' \in Q^{\mathbb{Z}^d}$$

A pattern p is orphan if every configuration where it occurs is a GoE.

An orphan pattern for a simple CA

Consider the AND CA on two neighbors

- $d = 1$
- $Q = \{0, 1\}$
- $\mathcal{N} = \{0, 1\}$
- $f(a, b) = a \text{ AND } b$

The pattern 101 is orphan:

···	1	0	1	·	···
···	1	1	1	1	···
···	·	↑	·	·	···

An orphan pattern for a simple CA

Consider the AND CA on two neighbors

- $d = 1$
- $Q = \{0, 1\}$
- $\mathcal{N} = \{0, 1\}$
- $f(a, b) = a \text{ AND } b$

The pattern 101 is orphan:

$$\begin{array}{cccccc} \cdots & 1 & 0 & 1 & \cdot & \cdots \\ \cdots & 1 & 1 & 1 & 1 & \cdots \\ \cdots & \cdot & \uparrow & \cdot & \cdot & \cdots \end{array}$$

An orphan pattern for a simple CA

Consider the AND CA on two neighbors

- $d = 1$
- $Q = \{0, 1\}$
- $\mathcal{N} = \{0, 1\}$
- $f(a, b) = a \text{ AND } b$

The pattern 101 is orphan:

$$\begin{array}{cccccc} \cdots & 1 & 0 & 1 & \cdot & \cdots \\ \cdots & 1 & 1 & 1 & 1 & \cdots \\ \cdots & \cdot & \uparrow & \cdot & \cdot & \cdots \end{array}$$

An orphan pattern for a simple CA

Consider the AND CA on two neighbors

- $d = 1$
- $Q = \{0, 1\}$
- $\mathcal{N} = \{0, 1\}$
- $f(a, b) = a \text{ AND } b$

The pattern 101 is orphan:

$$\begin{array}{cccccc} \cdots & 1 & 0 & 1 & \cdot & \cdots \\ \cdots & 1 & 1 & 1 & 1 & \cdots \\ \cdots & \cdot & \uparrow & \cdot & \cdot & \cdots \end{array}$$

An orphan pattern for a simple CA

Consider the AND CA on two neighbors

- $d = 1$
- $Q = \{0, 1\}$
- $\mathcal{N} = \{0, 1\}$
- $f(a, b) = a \text{ AND } b$

The pattern 101 is orphan:

$$
\begin{array}{cccccc}
\cdots & 1 & 0 & 1 & \cdot & \cdots \\
\cdots & 1 & 1 & 1 & 1 & \cdots \\
\cdots & \cdot & \uparrow & \cdot & \cdot & \cdots
\end{array}
$$

The Garden-of-Eden lemma

Let $\mathcal{A}$ be a CA. The following are equivalent:

1. $\mathcal{A}$ has a GoE configuration
2. $\mathcal{A}$ has an orphan pattern

Proof:

- For each n, consider the restriction p_n of c to $\mathrm{M}(n)$.
- $\mathcal{A}$ has no orphan pattern $\Rightarrow$ each p_n has a predecessor $\Rightarrow$ extend that to a configuration c_n'.
- $Q^{\mathbb{Z}^d}$ compact $\Rightarrow$ the sequence $\{c_n'\}$ has a limit point c'.
- Then $F_{\mathcal{A}}(c') = c$ by continuity.

Corollary: CA surjectivity is co-r.e.
Reason why: try all patterns until one has no predecessors.

The Garden-of-Eden lemma

Let $\mathcal{A}$ be a CA. The following are equivalent:

1. $\mathcal{A}$ has a GOE configuration
2. $\mathcal{A}$ has an orphan pattern

Proof:

- For each n, consider the restriction p_n of c to $\mathrm{M}(n)$.
- $\mathcal{A}$ has no orphan pattern $\Rightarrow$ each p_n has a predecessor $\Rightarrow$ extend that to a configuration c_n'.
- $Q^{\mathbb{Z}^d}$ compact $\Rightarrow$ the sequence $\{c_n'\}$ has a limit point c'.
- Then $F_{\mathcal{A}}(c') = c$ by continuity.

Corollary: CA surjectivity is co-r.e.

Reason why: try all patterns until one has no predecessors.

"Not injectivity, but almost"

Cellular automata are "not finitar, but almost".
It seems reasonable that surjectivity for CA may be equivalent to "not injectivity, but almost".
Say that two distinct patterns $p_1, p_2 : E \to Q$ are mutually erasable (m.e.) for $\mathcal{A}$ if

- $(c_i)|_E = p_i$ and
- $(c_1)|_{\mathbb{Z}^d \setminus E} = (c_2)|_{\mathbb{Z}^d \setminus E}$

imply $F_{\mathcal{A}}(c_1) = F_{\mathcal{A}}(c_2)$.
Call pre-injective a CA that has no two m.e. patterns.

"Not injectivity, but almost"

Cellular automata are "not finitar, but almost".
It seems reasonable that surjectivity for CA may be equivalent to "not injectivity, but almost".
Say that two distinct patterns $p_1, p_2 : E \to Q$ are mutually erasable (m.e.) for $\mathcal{A}$ if

- $(c_i)|_E = p_i$ and
- $(c_1)|_{\mathbb{Z}^d \setminus E} = (c_2)|_{\mathbb{Z}^d \setminus E}$

imply $F_{\mathcal{A}}(c_1) = F_{\mathcal{A}}(c_2)$.
Call pre-injective a CA that has no two m.e. patterns.

Moore-Myhill's theorem (1962)

The following are equivalent:

1. $\mathcal{A}$ is surjective
2. $\mathcal{A}$ is pre-injective

Reason why:

- Call boundary of E (w.r.t. $\mathcal{N}$) the sets of neighbors of points of E that are not in E
- Then the size of the boundary of a dD hypercube is bounded by a polynomial of degree $d-1$
- The thesis follows by a counting argument

Corollary: (Richardson, 1972)

1. injective CA are surjective
2. if $\mathcal{A}$ has a quiescent state q, then

$$\mathcal{A} \text{ surjective} \Leftrightarrow \mathcal{A}_{(q)} \text{ injective}$$

Moore-Myhill's theorem (1962)

The following are equivalent:

1. $\mathcal{A}$ is surjective
2. $\mathcal{A}$ is pre-injective

Reason why:

- Call boundary of E (w.r.t. $\mathcal{N}$) the sets of neighbors of points of E that are not in E
- Then the size of the boundary of a dD hypercube is bounded by a polynomial of degree $d-1$
- The thesis follows by a counting argument

Corollary: (Richardson, 1972)

1. injective CA are surjective
2. if $\mathcal{A}$ has a quiescent state q, then

$$\mathcal{A} \text{ surjective} \Leftrightarrow \mathcal{A}_{(q)} \text{ injective}$$

Moore-Myhill's theorem (1962)

The following are equivalent:

1. $\mathcal{A}$ is surjective
2. $\mathcal{A}$ is pre-injective

Reason why:

- Call boundary of E (w.r.t. $\mathcal{N}$) the sets of neighbors of points of E that are not in E
- Then the size of the boundary of a dD hypercube is bounded by a polynomial of degree $d-1$
- The thesis follows by a counting argument

Corollary: (Richardson, 1972)

1. injective CA are surjective
2. if $\mathcal{A}$ has a quiescent state q, then

$$\mathcal{A} \text{ surjective} \Leftrightarrow \mathcal{A}_{(q)} \text{ injective}$$

A surjective, non-injective CA

Let $d = 1$, $Q = \{0, 1\}$, $\mathcal{N} = \{-1, 0, 1\}$, $f(a, b, c) = a \oplus c$,

- Non-injectivity: put

$$c_0(x) = 0 \ \ \forall x \in \mathbb{Z} \ \ ; \ \ c_1(x) = 1 \ \ \forall x \in \mathbb{Z}$$

  then $F_{\mathcal{A}}(c_0) = F_{\mathcal{A}}(c_1) = c_0$.
- Surjectivity:
  1. for every a and k, the equation $a \oplus x = k$ has a unique solution
  2. for every b and k, the equation $x \oplus b = k$ has a unique solution

  Thus every configuration has exactly four predecessors.

Balancement

Let $\mathcal{A} = \langle d, Q, \mathrm{M}(r), f \rangle$ be a CA.
(Observe that all CA may be written as such.)
For each n, define $F_n : Q^{\{1,\dots,n+2r\}^d} \to Q^{\{1,\dots,n\}^d}$ as

$$(F_n(p))(x) = f(p(x+n_1), \dots, p(x+n_{|\mathrm{M}(r)|}))$$

We say that $\mathcal{A}$ is *n*-balanced if

$$|(F_n(p))^{-1}| = Q^{(n+2r)^d - n^d} \quad \forall p \in Q^{\{1,\dots,n+2r\}^d} ,$$

i.e., if every pattern on a d-hypercube of side n has the same number of pre-images.

Balancement

Let $\mathcal{A} = \langle d, Q, \mathrm{M}(r), f \rangle$ be a CA.
(Observe that all CA may be written as such.)
For each n, define $F_n : Q^{\{1,\ldots,n+2r\}^d} \to Q^{\{1,\ldots,n\}^d}$ as

$$(F_n(p))(x) = f(p(x+n_1), \ldots, p(x+n_{|\mathrm{M}(r)|}))$$

We say that $\mathcal{A}$ is *n*-balanced if

$$|(F_n(p))^{-1}| = Q^{(n+2r)^d - n^d} \quad \forall p \in Q^{\{1,\ldots,n+2r\}^d} ,$$

i.e., if every pattern on a d-hypercube of side n has the same number of pre-images.

A 1-balanced, nonsurjective CA

The majority CA is defined by the local function

$$f(a,b,c) = \begin{cases} 0 & \text{if } a+b+c \leq 1 \\ 1 & \text{if } a+b+c \geq 2 \end{cases}$$

Then the string 01001 is a GoE:

$$\begin{array}{ccccccc} \cdots & 0 & 1 & 0 & 0 & 1 & \cdots \\ \cdots & 1 & 0 & 1 & 0 & 0 & \cdots \\ \cdots & & & & & \uparrow & \cdots \\ \cdots & 0 & 1 & 0 & & \cdots & \\ \cdots & & \uparrow & & & & \cdots \end{array}$$

A 1-balanced, nonsurjective CA

The majority CA is defined by the local function

$$f(a,b,c) = \begin{cases} 0 & \text{if } a+b+c \leq 1 \\ 1 & \text{if } a+b+c \geq 2 \end{cases}$$

Then the string 01001 is a GoE:

$$\begin{array}{ccccccc} \cdots & 0 & 1 & 0 & 0 & 1 & \cdots \\ \cdots & 1 & 0 & 1 & 0 & 0 & \cdots \\ \cdots & & & & & \uparrow & \cdots \\ \cdots & 0 & 1 & 0 & & \cdots & \\ \cdots & & \uparrow & & & & \cdots \end{array}$$

A 1-balanced, nonsurjective CA

The majority CA is defined by the local function

$$f(a,b,c) = \begin{cases} 0 & \text{if} \;\; a+b+c \leq 1 \\ 1 & \text{if} \;\; a+b+c \geq 2 \end{cases}$$

Then the string 01001 is a GoE:

$$\begin{array}{ccccccc}
\cdots & 0 & 1 & 0 & 0 & 1 & \cdots \\
\cdots & 1 & 0 & 1 & 0 & 0 & \cdots \\
\cdots & & & & & \uparrow & \cdots \\
\cdots & 0 & 1 & 0 & & \cdots & \\
\cdots & & \uparrow & & & & \cdots
\end{array}$$

The balancement theorem (Maruoka and Kimura, 1976)

Let $\mathcal{A} = \langle d, Q, \mathrm{M}(r), f \rangle$, $U \subseteq \mathbb{Z}^d$. The following are equivalent:

1. $\mathcal{A}$ is surjective
2. $\mathcal{A}$ is n-balanced for all n

Reason why:

- Boundary grows slower than support
- If n-balanced for all n then no pattern is orphan
- If not n-balanced for some n, employ “rarest” patterns to find (larger) orphan pattern

The balancement theorem (Maruoka and Kimura, 1976)

Let $\mathcal{A} = \langle d, Q, \mathrm{M}(r), f \rangle$, $U \subseteq \mathbb{Z}^d$. The following are equivalent:

1. $\mathcal{A}$ is surjective
2. $\mathcal{A}$ is n-balanced for all n

Reason why:

- Boundary grows slower than support
- If n-balanced for all n then no pattern is orphan
- If not n-balanced for some n, employ "rarest" patterns to find (larger) orphan pattern

Loss of output state in CA

- The Garden-of-Eden theorem says that non-surjective CA are precisely those that lose output state within finite range
- How does one measure the amount of lost state?

Given $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$, let $\mathrm{Out}_f(n)$ be the number of non-orphan patterns with support a d-hypercube of side n.
Consider then the loss of state at side n

$$\Lambda_{\mathcal{A}}(n) = n^d - \log_{|Q|} \mathrm{Out}_f(n)$$

measured in qits ($q = |Q|$; 1 qit$=\log_2 q$ bits)
Then $\mathcal{A}$ is surjective iff $\Lambda_{\mathcal{A}}$ is identically zero.

Loss of output state in CA

- The Garden-of-Eden theorem says that non-surjective CA are precisely those that lose output state within finite range
- How does one measure the amount of lost state?

Given $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$, let $\mathrm{Out}_f(n)$ be the number of non-orphan patterns with support a d-hypercube of side n.
Consider then the loss of state at side n

$$\Lambda_{\mathcal{A}}(n) = n^d - \log_{|Q|} \mathrm{Out}_f(n)$$

measured in qits ($q = |Q|$; 1 $q$it$=\log_2 q$ bits)
Then $\mathcal{A}$ is surjective iff $\Lambda_{\mathcal{A}}$ is identically zero.

How much information do non-surjective CA lose?

Theorem (Capobianco, 2008)

If $\mathcal{A}$ is nonsurjective then $\lim_{n\to\infty} \Lambda_{\mathcal{A}}(n) = +\infty$

Proof: (for $d = 1$)

- "Large" non-orphan is juxtaposition of "small" non-orphans $\Rightarrow \mathrm{Out}_f(m+n) \leq \mathrm{Out}_f(m) \cdot \mathrm{Out}_f(n)$ for all m and n
- By Fekete's lemma, there exists $\delta < 1$ such that

$$\frac{\log_{|Q|} \mathrm{Out}_f(n)}{n} < \delta$$

  for all n large enough
- For those values of n, $\Lambda_{\mathcal{A}}(n) > n \cdot (1-\delta)$

How much information do non-surjective CA lose?

Theorem (Capobianco, 2008)
If $\mathcal{A}$ is nonsurjective then $\lim_{n \to \infty} \Lambda_{\mathcal{A}}(n) = +\infty$

Proof: (for $d = 1$)

- "Large" non-orphan is juxtaposition of "small" non-orphans $\Rightarrow \mathrm{Out}_f(m+n) \leq \mathrm{Out}_f(m) \cdot \mathrm{Out}_f(n)$ for all m and n
- By Fekete's lemma, there exists $\delta < 1$ such that

$$\frac{\log_{|Q|} \mathrm{Out}_f(n)}{n} < \delta$$

  for all n large enough
- For those values of n, $\Lambda_{\mathcal{A}}(n) > n \cdot (1 - \delta)$

Reversibility

A CA $\mathcal{A}$ is reversible if

1. $\mathcal{A}$ is invertible, and
2. $F_{\mathcal{A}}^{-1}$ is the global evolution function of some CA.

Thus, a CA is reversible iff the reverse CA exists.

This seems more than just existence of inverse global evolution.

Reversible CA are important in physical modelization because Physics, at microscopical scale, is reversible.

Fact CA reversibility is r.e.

Reason why: try composing $\mathcal{A}$ with other CA in all possible ways until a combination yields the identity CA.

Reversibility

A CA $\mathcal{A}$ is reversible if

1. $\mathcal{A}$ is invertible, and
2. $F_{\mathcal{A}}^{-1}$ is the global evolution function of some CA.

Thus, a CA is reversible iff the reverse CA exists.
This seems more than just existence of inverse global evolution.
Reversible CA are important in physical modelization because Physics, at microscopical scale, is reversible.
Fact CA reversibility is r.e.
Reason why: try composing $\mathcal{A}$ with other CA in all possible ways until a combination yields the identity CA.

Reversibility

A CA $\mathcal{A}$ is reversible if

1. $\mathcal{A}$ is invertible, and
2. $F_{\mathcal{A}}^{-1}$ is the global evolution function of some CA.

Thus, a CA is reversible iff the reverse CA exists.

This seems more than just existence of inverse global evolution.

Reversible CA are important in physical modelization because Physics, at microscopical scale, is reversible.

Fact CA reversibility is r.e.

Reason why: try composing $\mathcal{A}$ with other CA in all possible ways until a combination yields the identity CA.

Richardson's reversibility principle (1972)

The following are equivalent:

1. $\mathcal{A}$ is reversible
2. $\mathcal{A}$ is bijective

Thus, existence of inverse CA comes at no cost from existence of inverse global evolution.

Reason why:

- $Q^{\mathbb{Z}^d}$ compact metrizable $\Rightarrow F_{\mathcal{A}}^{-1}$ continuous
- $F_{\mathcal{A}}$ commutes with shift $\Rightarrow F_{\mathcal{A}}^{-1}$ does
- apply Hedlund's theorem

Richardson's reversibility principle (1972)

The following are equivalent:

1. $\mathcal{A}$ is reversible
2. $\mathcal{A}$ is bijective

Thus, existence of inverse CA comes at no cost from existence of inverse global evolution.

Reason why:

- $Q^{\mathbb{Z}^d}$ compact metrizable $\Rightarrow F_{\mathcal{A}}^{-1}$ continuous
- $F_{\mathcal{A}}$ commutes with shift $\Rightarrow F_{\mathcal{A}}^{-1}$ does
- apply Hedlund's theorem

A scheme of the current situation

Reversible (r.e.)	Properly Surjective	Non−Surjective (r.e.)

Can reversibility be checked?

Let $\mathcal{C}$ be a class of cellular automata.
The invertibility problem for $\mathcal{C}$ states:

given an element $\mathcal{A}$ of $\mathcal{C}$,
determine whether $F_{\mathcal{A}}$ is invertible.

Meaning: invertibility of the global dynamics of any CA in $\mathcal{C}$ can be inferred algorithmically by looking at its local description.

One is too little...

Theorem (Amoroso and Patt, 1972)
The invertibility problem for 1D CA is decidable.

Reason why:

- surjectivity of 1D CA can be determined via a suitable graph
- injectivity of 1D surjective CA can be checked "within finite range"

Additional results:

1. surjectivity of 1D CA is decidable.
2. there are computable bounds for inverse neighborhood radius of 1D CA—though none is known that is polynomial

One is too little...

Theorem (Amoroso and Patt, 1972)

The invertibility problem for 1D CA is decidable.

Reason why:

- surjectivity of 1D CA can be determined via a suitable graph
- injectivity of 1D surjective CA can be checked "within finite range"

Additional results:

1. surjectivity of 1D CA is decidable.
2. there are computable bounds for inverse neighborhood radius of 1D CA—though none is known that is polynomial

... but two is too much

"Although the techniques we employ are in principle adaptable to arrays of higher dimension, it turns out that they are difficult to manage beyond dimension one."

Theorem (Kari, 1990)

The invertibility problem for 2D CA—and consequently for dD CA with $d > 2$—is **un**decidable.

Reason why: undecidability of Hao Wang's Tiling Problem:

given a set of square tiles with colored sides,
determine if there is a tiling of the plane
where pairs of adjacent sides always have same color

… but two is too much

"Although the techniques we employ are in principle adaptable to arrays of higher dimension, it turns out that they are difficult to manage beyond dimension one."

Theorem (Kari, 1990)

The invertibility problem for 2D CA—and consequently for dD CA with $d > 2$—is **un**decidable.

Reason why: undecidability of Hao Wang's Tiling Problem:

given a set of square tiles with colored sides,
determine if there is a tiling of the plane
where pairs of adjacent sides always have same color

… but two is too much

"Although the techniques we employ are in principle adaptable to arrays of higher dimension, it turns out that they are difficult to manage beyond dimension one."

Theorem (Kari, 1990)

The invertibility problem for 2D CA—and consequently for dD CA with $d > 2$—is **un**decidable.

Reason why: undecidability of Hao Wang's Tiling Problem:

given a set of square tiles with colored sides,
determine if there is a tiling of the plane
where pairs of adjacent sides always have same color

Kari's method

- Consider a set of tiles T.
- Consider a special set of tiles S, whose tiles also have arrows. (This set has a special, "plane filling" property.)
- Construct a CA with $Q = T \times S \times \{0, 1\}$ and whose rule says:
  - if both tilings are correct then XOR with pointed neighbor
  - otherwise do nothing
- Then there is a valid tiling with T iff the CA is non-reversible.

Corollary: for $d \geq 2$ there is no computable bound for inverse neighborhood radius.

Kari's method

- Consider a set of tiles T.
- Consider a special set of tiles S, whose tiles also have arrows. (This set has a special, "plane filling" property.)
- Construct a CA with $Q = T \times S \times \{0, 1\}$ and whose rule says:
  - if both tilings are correct then XOR with pointed neighbor
  - otherwise do nothing
- Then there is a valid tiling with T iff the CA is non-reversible.

Corollary: for $d \geq 2$ there is no computable bound for inverse neighborhood radius.

Kari's method

- Consider a set of tiles T.
- Consider a special set of tiles S, whose tiles also have arrows. (This set has a special, "plane filling" property.)
- Construct a CA with $Q = T \times S \times \{0, 1\}$ and whose rule says:
  - if both tilings are correct then XOR with pointed neighbor
  - otherwise do nothing
- Then there is a valid tiling with T iff the CA is non-reversible.

Corollary: for $d \geq 2$ there is no computable bound for inverse neighborhood radius.

From infinite to finite

- For now, we have only considered CA on infinite grids.
- We now consider laws induced the same way, on toroidal supports—equivalently, on periodic configurations.
- If $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$ and a hypercube of side n contains $\mathcal{N}$, call $\mathcal{A}_n$ the transformation induced by $\mathcal{A}$ on $Q^{(\mathbb{Z}/n\mathbb{Z})^d}$.
- We call locally non-reversible those CA local rules that induce non-surjective transformations for some values of the size n.

From infinite to finite

- For now, we have only considered CA on infinite grids.
- We now consider laws induced the same way, on toroidal supports—equivalently, on periodic configurations.
- If $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$ and a hypercube of side n contains $\mathcal{N}$, call $\mathcal{A}_n$ the transformation induced by $\mathcal{A}$ on $Q^{(\mathbb{Z}/n\mathbb{Z})^d}$.
- We call locally non-reversible those CA local rules that induce non-surjective transformations for some values of the size n.

Local (non-)reversibility

Local reversibility is co-r.e.

- Reason why: Try all periodic configurations until a GoE is found.

Reversible CA are locally reversible.

- Reason why: $\mathcal{A}$ reversible $\Rightarrow$ (pre)image of a periodic configuration is also periodic—with same period(s)

Non-surjective CA are locally non-reversible.

- Reason why: Extend an orphan pattern to a periodic configuration.

Local (non-)reversibility

Local reversibility is co-r.e.

- Reason why: Try all periodic configurations until a GoE is found.

Reversible CA are locally reversible.

- Reason why: $\mathcal{A}$ reversible $\Rightarrow$ (pre)image of a periodic configuration is also periodic—with same period(s)

Non-surjective CA are locally non-reversible.

- Reason why: Extend an orphan pattern to a periodic configuration.

Local (non-)reversibility

Local reversibility is co-r.e.

- Reason why: Try all periodic configurations until a GoE is found.

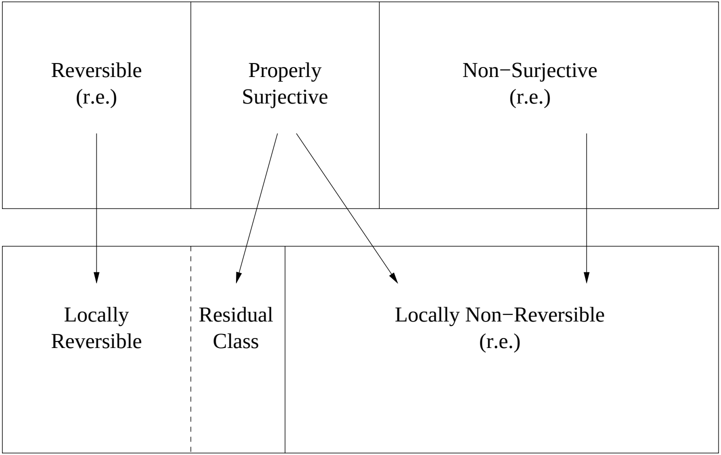
Reversible CA are locally reversible.

- Reason why: $\mathcal{A}$ reversible $\Rightarrow$ (pre)image of a periodic configuration is also periodic—with same period(s)

Non-surjective CA are locally non-reversible.

- Reason why: Extend an orphan pattern to a periodic configuration.

A scheme of the updated situation



The residual class (Toffoli and Margolus, 1990)

It is made of local rules that

- always determine reversible CA on hypercubes
- always determine properly surjective CA on the whole space

It is non-r.e. (in particular, non-empty)

Reason why:

- Suppose otherwise
- Then global non-reversibility is union of r.e. properties
- But global reversibility is r.e. $\Rightarrow$ violation of Kari's theorem

The residual class (Toffoli and Margolus, 1990)

It is made of local rules that

- always determine reversible CA on hypercubes
- always determine properly surjective CA on the whole space

It is non-r.e. (in particular, non-empty)

Reason why:

- Suppose otherwise
- Then global non-reversibility is union of r.e. properties
- But global reversibility is r.e. $\Rightarrow$ violation of Kari's theorem

The complexity issue

We can consider the finitary version of the invertibility problem:

given $\mathcal{A}$ and n,
determine if $\mathcal{A}_n$ is reversible

Theorem (Clementi, 1994)
The invertibility problem for hypercubic CA is co-NP-complete.
Reason why: a polynomial reduction such that

- a Turing machine stops within given time from empty tape
- iff a toroidal 2D CA is non-injective

The complexity issue

We can consider the finitary version of the invertibility problem:

given $\mathcal{A}$ and n,
determine if $\mathcal{A}_n$ is reversible

Theorem (Clementi, 1994)
The invertibility problem for hypercubic CA is co-NP-complete.
Reason why: a polynomial reduction such that

- a Turing machine stops within given time from empty tape
- iff a toroidal 2D CA is non-injective

Suggested readings

- T. Toffoli, N. Margolus. (1990) Invertible cellular automata: A review. *Physica D* **45**, pp. 229–253. http://pm1.bu.edu/~tt/publ/ica.ps
- J. Kari. (2005) Theory of cellular automata: a survey. *Theor. Comp Sci.* **334**, pp. 3–33. doi:10.1016/j.tcs.2004.11.021
- T. Toffoli, S. Capobianco, P. Mentrasti. (2008) When—and how—can a cellular automaton be rewritten as a lattice gas? *Theor. Comp Sci.* **403**, pp. 71–88. doi:10.1016/j.tcs.2008.04.047

Thank you for attention!

Any questions?