Local Information Transfer – Lizier, Prokopenko and Zomaya

ANRG presentation
3/8/2007

Joseph Lizier
PhD student
The University of Sydney – Prof. Albert Zomaya, Dr Selvakennedy Selvadurai
CSIRO ICT Centre – Dr. Mikhail Prokopenko

Local Information Transfer – Lizier, Prokopenko and Zomaya

Local Information Transfer in Complex Spatiotemporal Systems

Joseph T. Lizier,[1,2,*] Mikhail Prokopenko,[1] and Albert Y. Zomaya[2]

[1] *CSIRO Information and Communications Technology Centre, Locked Bag 17, North Ryde, NSW 1670, Australia*

[2] *School of Information Technologies, The University of Sydney, NSW 2006, Australia*

(Dated: August 2, 2007)

We present a measure of local information transfer, derived from an existing averaged information-theoretical measure, namely transfer entropy. Local transfer entropy is used to produce profiles of the information transfer into each spatiotemporal point in a complex system. These spatiotemporal profiles are useful not only as an analytical tool, but also allow explicit investigation of different parameter settings and forms of the transfer entropy metric itself. As an example, local transfer entropy is applied to cellular automata, where it provides the first quantitative evidence for the long-held conjecture that the emergent travelling coherent structures known as particles, gliders and domain walls are the dominant information transfer agents therein.

- *Aim*: Demonstrate how to profile information transfer locally:
  - i.e. as a function of space-time
- *Application*: Prove particles are the information transfer agents in Cellular Automata (CAs).

CSIRO

Contents

CSIRO

Information-theoretical preliminaries

- Shannon entropy $H(X) = -\sum_{x \in A_X} p(x) \log p(x)$
- Joint entropy $H(X,Y) = -\sum_{xy \in A_X A_Y} p(x,y) \log p(x,y)$
- Conditional entropy $H(X \mid Y) = -\sum_{xy \in A_X A_Y} p(x,y) \log p(x \mid y)$
- Mutual information $I(X;Y) = \sum_{xy \in A_X A_Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$

$$I(X;Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X)$$

- Conditional mutual information

$$I(X;Y \mid Z) = H(X \mid Z) - H(X \mid Y,Z) = H(Y \mid Z) - H(Y \mid X,Z)$$

CSIRO

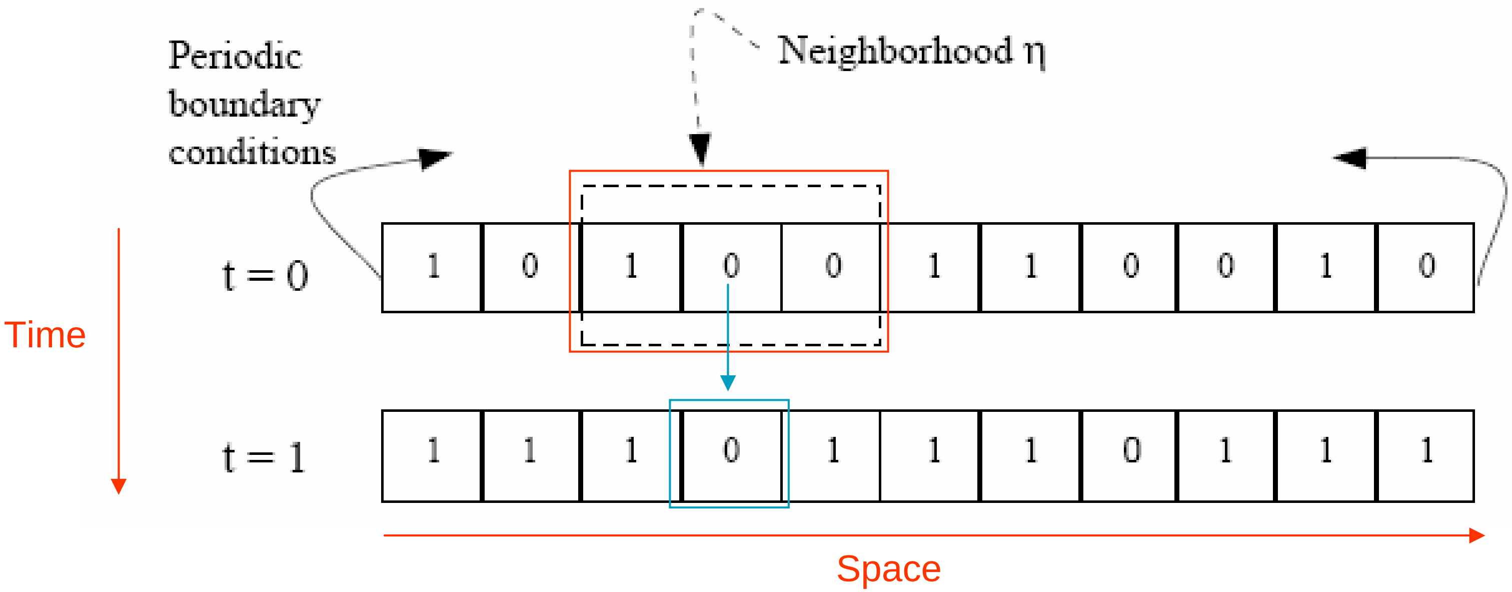
Cellular Automata – micro-level rules

Rule table ϕ:

neighborhood:	000	001	010	011	100	101	110	111
output bit:	0	1	1	1	0	1	1	0

= Rule 0x6e = Rule 110

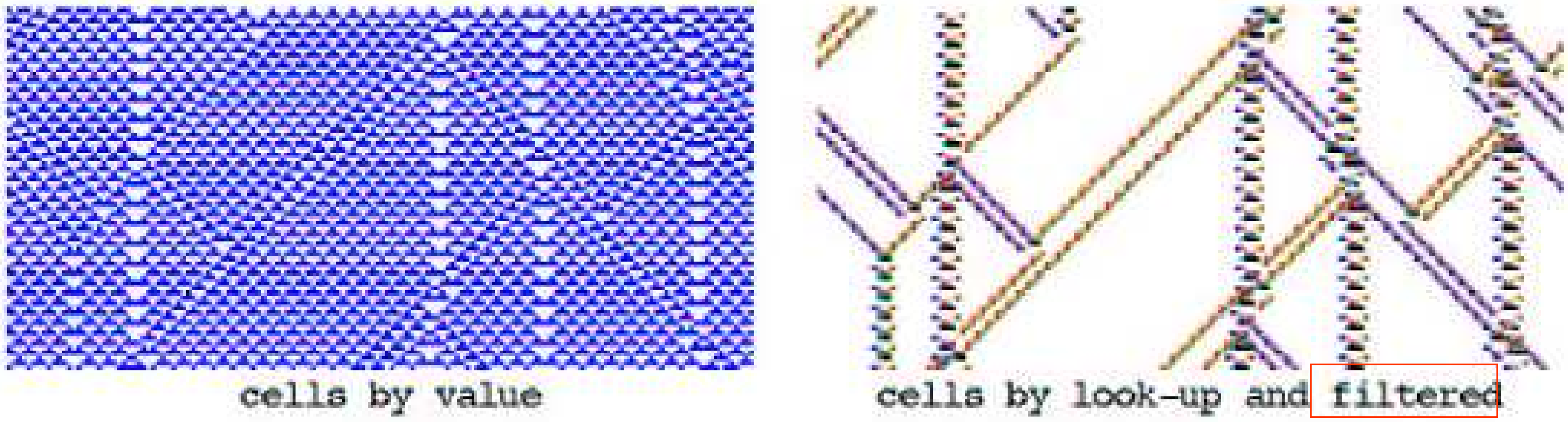
Lattice:



- "Computation in Cellular Automata: A selected review", Mitchell, 1998

Cellular Automata – emergent structure

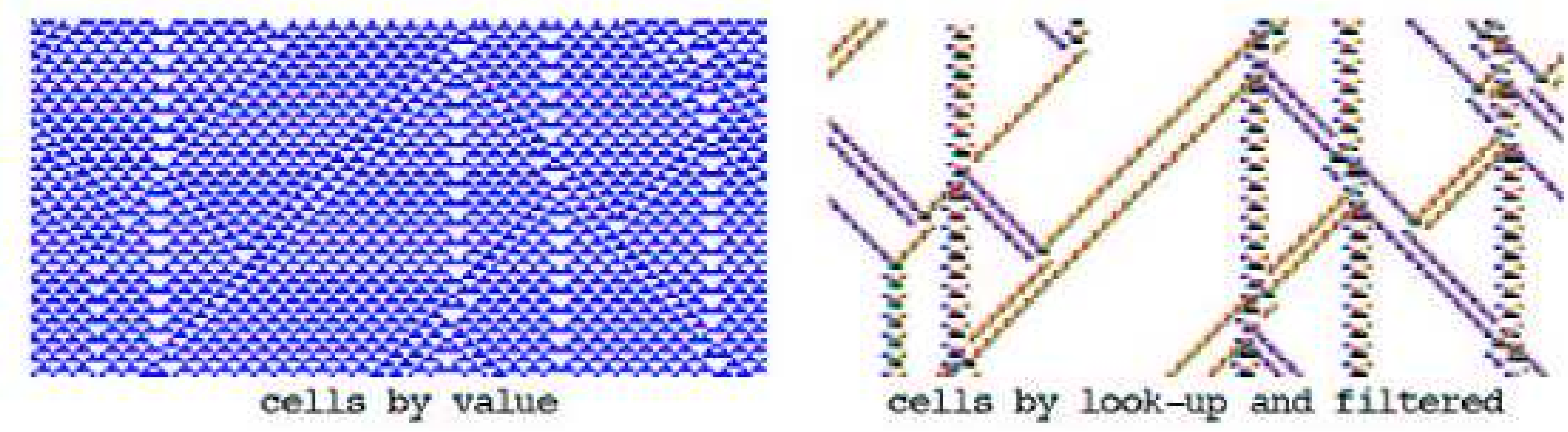
- "Classifying Cellular Automata Automatically …", Wuensche, 1999



- Existing filtering methods to highlight emergent structure:
  - Computational dynamics (finite state transducers) (Crutchfield and Hanson)
  - Frequency of rule execution (Wuesnche)
  - Local statistical complexity and sensitivity (Shalizi et al)
  - Local information (really local spatial entropy rate) (Helvik et al)

Cellular Automata – emergent structure

- “Classifying Cellular Automata Automatically …”, Wuensche, 1999



- Emergent structure:
  - Domain
  - Particles
    - Gliders, Domain wa
  - Collisions

- Conjectured to represent:
  - nformation transfer
    - “
  - Information modification

No quantified evidence !!

CSIRO

Information transfer in other complex systems

- Information transfer in the vicinity of order-chaos phase transitions:
  - Conjectured to be at an intermediate level (Langton, Coffey)
  - Task-based measures of info tx, and mutual information, maximised (Solé and Valverde).
- Concept of empowerment (Klyubin et al), channel capacity of agent's perception-action loop. Maximisation → necessary structure in agent.
- Info tx in an example adaptive system (Sporns and Lungarella): differences between sensors, effect of learning, effect of changes to sensor morphology.

Transfer Entropy

- Mutual information was previously the de facto candidate for measuring information transfer:

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}.$$

- But MI is a symmetric, static measure of shared information.
- Transfer entropy introduced by Schreiber as an asymmetric measure of dynamic info flow:

$$T_{Y \to X} = \sum_{z_n} p(z_n) \log \frac{p(x_{n+1} | x_n^{(k)}, y_n^{(l)})}{p(x_{n+1} | x_n^{(k)})}. \quad z_n = (x_{n+1}, x_n^{(k)}, y_n^{(l)})$$

- Is a mutual information between source and destination, conditioned on the past of the destination

$$T_{Y \to X} = I(Y_n; X_{n+1} | X_n^{(k)}) =$$

$$H(X_{n+1}|X_n^{(k)}) - H(X_{n+1}|Y_n;X_n^{(k)})$$

Deriving Local Transfer Entropy (1/2)

$$T_{Y\rightarrow X} = \sum_{z_n} p(z_n) \log \frac{p(x_{n+1}|x_n^{(k)}, y_n^{(l)})}{p(x_{n+1}|x_n^{(k)})}. \quad z_n = (x_{n+1}, x_n^{(k)}, y_n^{(l)})$$

$$p(z_n) = \frac{c\left(x_{n+1}, x_n^{(k)}, y_n^{(l)}\right)}{N} \qquad p(z_n) = \left(\sum_{m=1}^{c(z_n)} 1\right)/N$$

$$T_{Y\rightarrow X} = \frac{1}{N} \sum_{x_{n+1}, x_n^{(k)}, y_n^{(l)}} \left(\sum_{m=1}^{c\left(x_{n+1}, x_n^{(k)}, y_n^{(l)}\right)} 1 \right) \log \frac{p\left(x_{n+1} \mid x_n^{(k)}, y_n^{(l)}\right)}{p\left(x_{n+1} \mid x_n^{(k)}\right)}$$

Deriving Local Transfer Entropy (2/2)

$$T_{Y \to X} = \frac{1}{N} \sum_{n=1}^{N} \log \frac{p(x_{n+1} | x_n^{(k)}, y_n^{(l)})}{p(x_{n+1} | x_n^{(k)})} = \langle t_{Y \to X}(n+1) \rangle$$

Local transfer entropy

$$t_{Y \to X}(n+1) = \log \frac{p\left(x_{n+1} \mid x_n^{(k)}, y_n^{(l)}\right)}{p\left(x_{n+1} \mid x_n^{(k)}\right)}$$

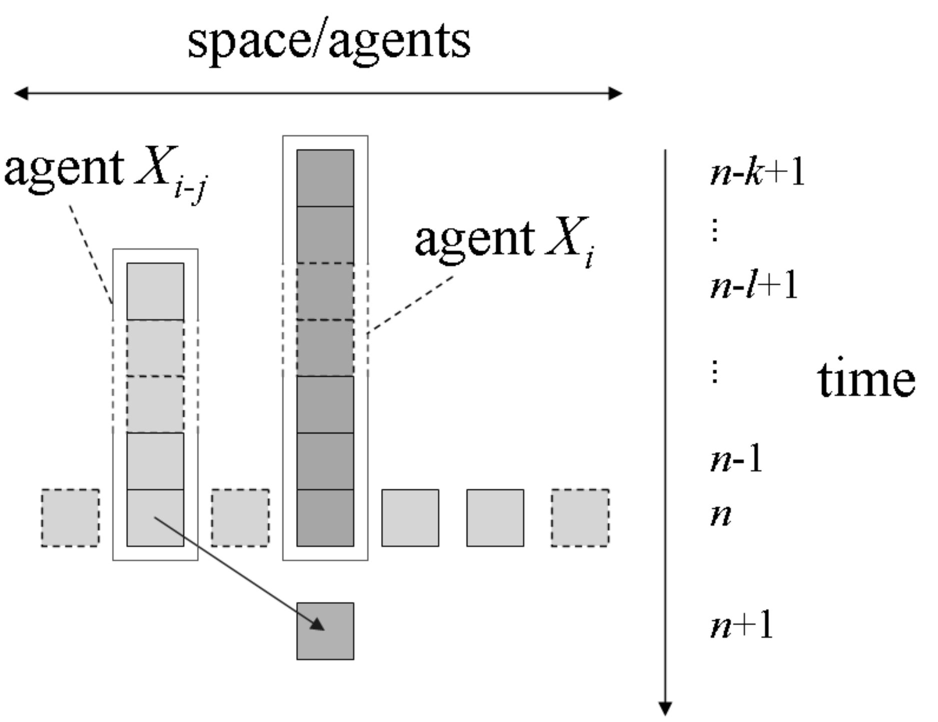
Local transfer entropy with spatially ordered agents *i*

$$t(i, j, n+1) = \log \frac{p(x_{i,n+1} | x_{i,n}^{(k)}, x_{i-j,n}^{(l)})}{p(x_{i,n+1} | x_{i,n}^{(k)})}$$

CSIRO

Local Transfer Entropy in ordered spatiotemporal systems

$$t(i,j,n+1,k,l) = \log \frac{p(x_{i,n+1} | x_{i,n}^{(k)}, x_{i-j,n}^{(l)})}{p(x_{i,n+1} | x_{i,n}^{(k)})}$$

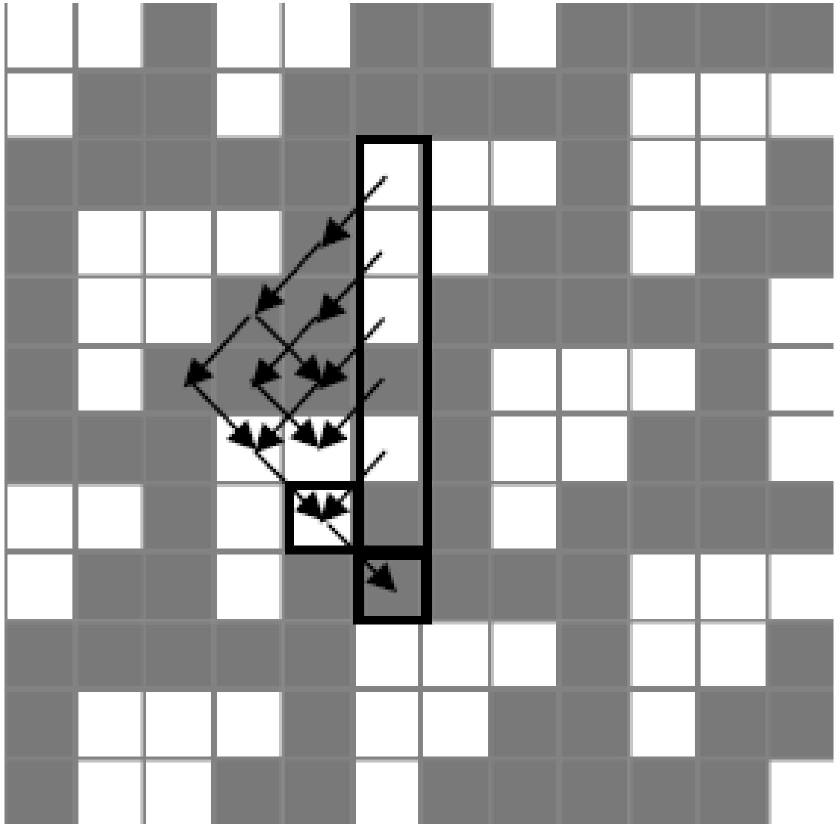


Subtlety 1: Correct only in limit k → ∞

- Destination's own historical values can also influence it's future states.
- This is a non-travelling form of information, like a standing wave, and is eliminated from the measurement by conditioning on the history $x_{i,n}^{(k)}$
- The most correct form is thus in the limit k → ∞:

$$t(i,j,n+1) = \lim_{k\to\infty} \log \frac{p(x_{i,n+1} | x_{i,n}^{(k)}, x_{i-j,n})}{p(x_{i,n+1} | x_{i,n}^{(k)})}$$

- Use *t(i,j,n+1,k)* for finite k estimates.



CSIRO

Subtlety 2: Apparent versus complete transfer entropy

- As stated, local transfer entropy takes no other possible information sources into account.
- As such, we label it *apparent* transfer entropy.
- We introduce *complete* transfer entropy to condition the measure on all other possible information sources.
- For CAs, this means conditioning on all other contributors $d_r(i,j,n)$ in the neighbourhood of the destination:

$$t_c(i,j,n+1) = \log \frac{p\left(x_{i,n+1} \mid x_{i,n}^{(k)}, x_{i-j,n}^{(l)}, d_r(i,j,n)\right)}{p\left(x_{i,n+1} \mid x_{i,n}^{(k)}, d_r(i,j,n)\right)}$$

CSIRO

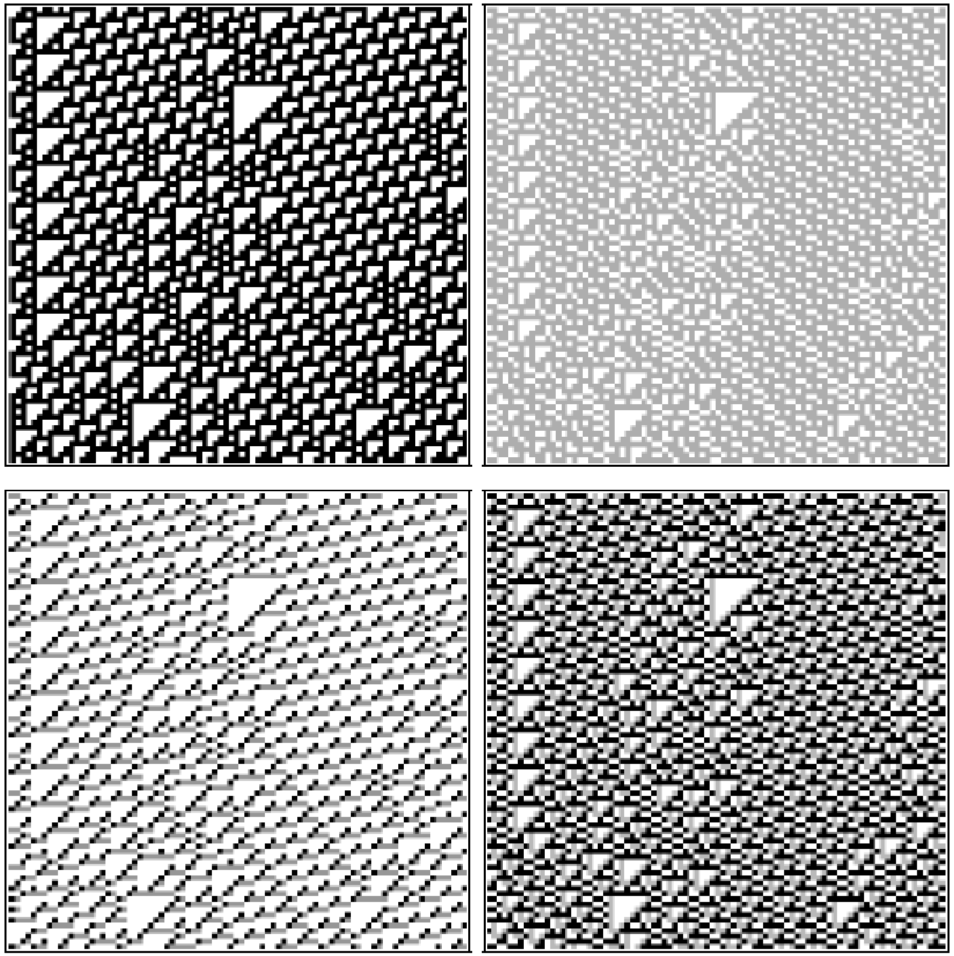
Application to CAs

- Using Elementary CAs (ECAs)
- 10 000 cells, periodic boundary conditions
- First 30 time steps eliminated to allow CA to settle.
- Next 600 time steps kept for estimate of probability distribution functions for $t_c(i,j,n,k)$
- Local complete TE measured at each space-time point
- All results confirmed by several CA runs.

Base cases: k = 0,1 for ECA rule 110

Raw CA

t(i,j=1,n,k=0)



tc(i,j=1,n,k=1)

t (i,j=1,n,k=1)

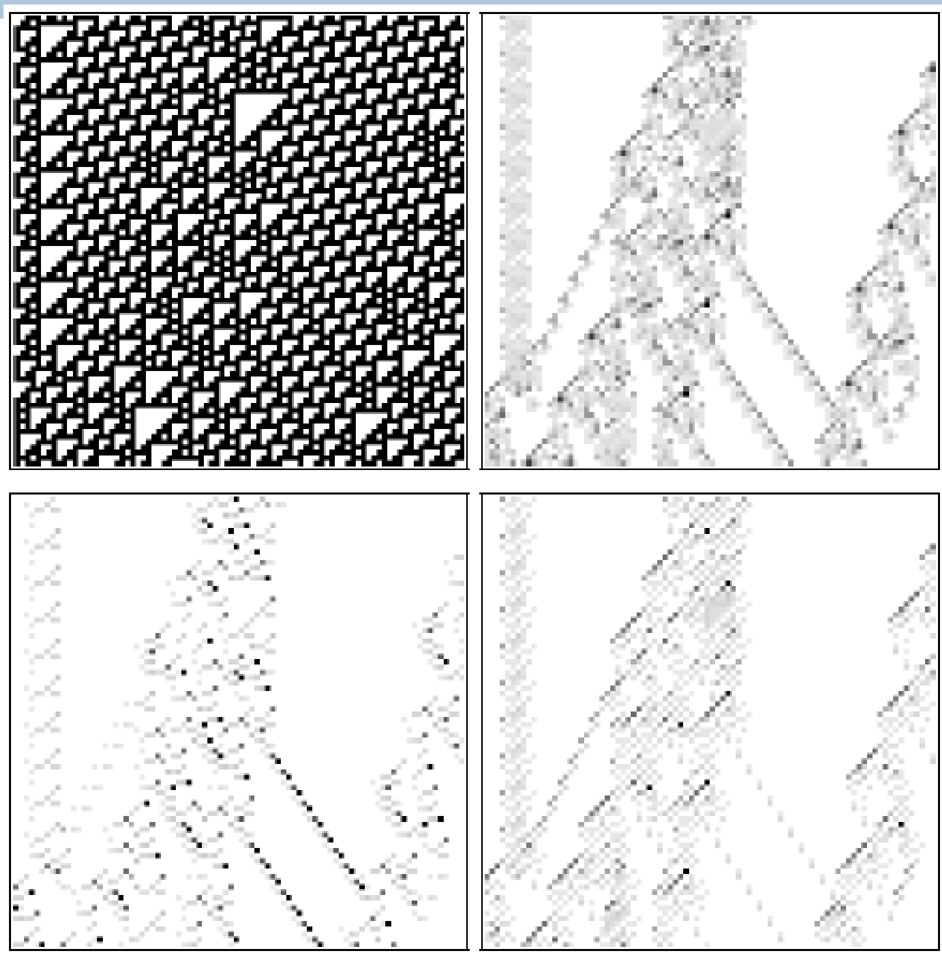
- None of these base cases distinguish the structure with any more clarity than the raw CA

CSIRO

Use of larger k for rule 110

Raw CA

$t_{sc}(i,j=1,n,k=6)$



$t_c(i,j=1,n,k=6)$

$t_c(i,j=-1,n,k=6)$

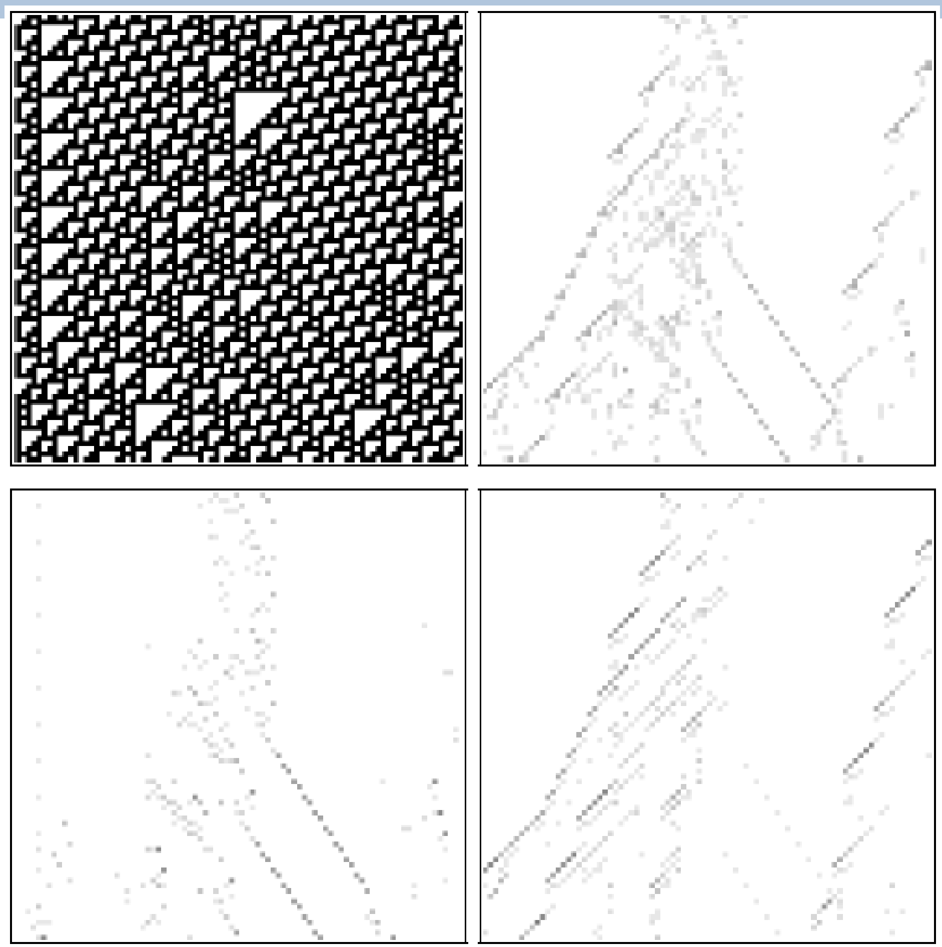
- Gliders distinguished from domain with minimum *k*.
  - Similar highlighting to existing filtering work.
  - Direction of transfer allows more detailed, quantitative view.

CSIRO

Approximate k → ∞ for rule 110

Raw CA

$t_{sc}(i,j=1,n,k=16)$

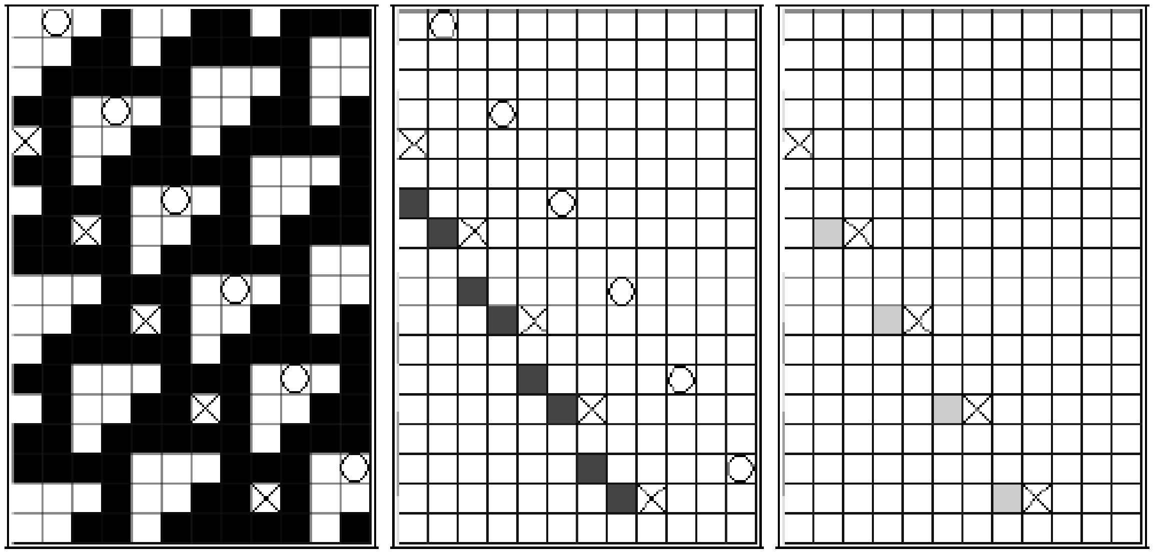


$t_c(i,j=1,n,k=16)$

$t_c(i,j=-1,n,k=16)$

- Gliders established as the information transfer agents in CAs.
- Leading edges highlighted as major travelling info components.

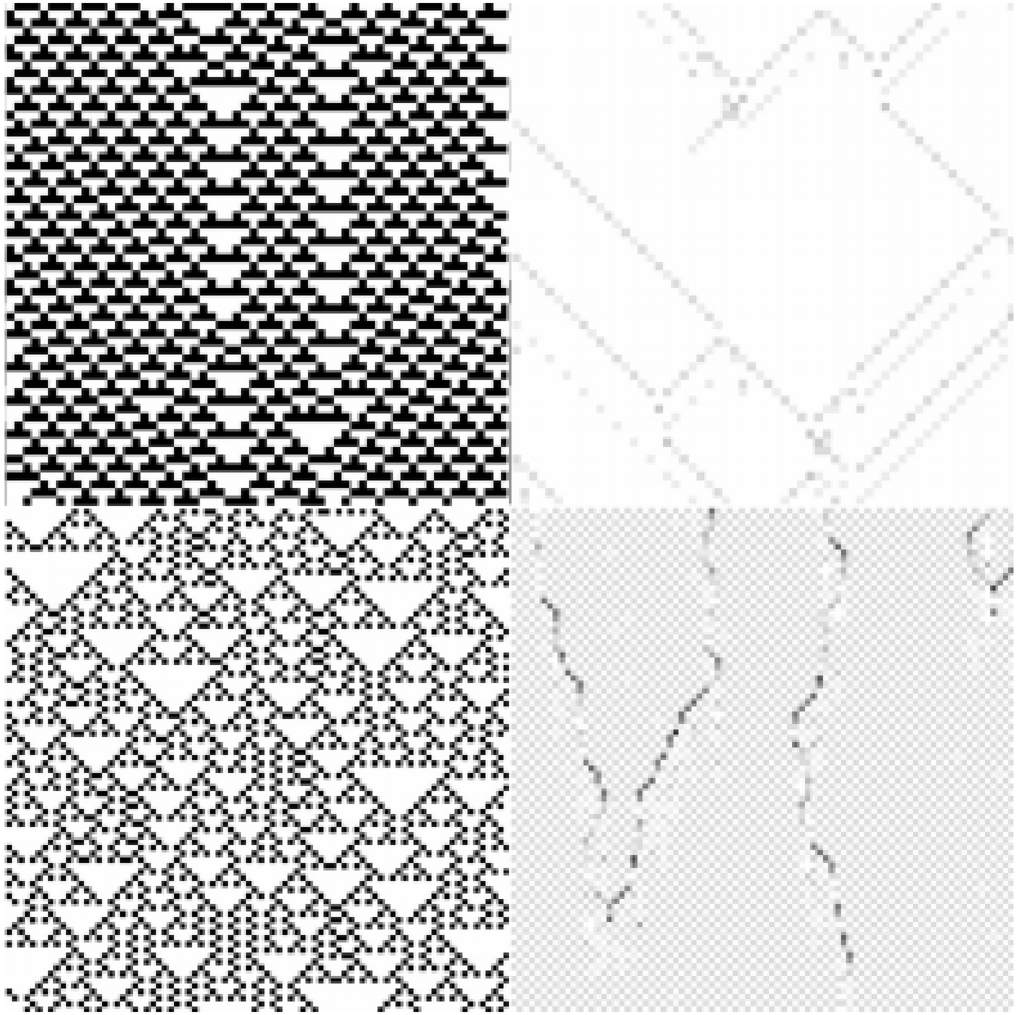
Close-up example



Glider in raw CA $t_c(i,j=1,n,k=6)$ $t(i,j=-1,n,k=6)$ (apparent)

- High transfer to right where cell up and to left provides significant information about the next state.
- Negative apparent transfer entropy (misleading info) found on gliders for measurements in orthogonal direction to glider motion. Source, as part of domain, is misleading about next state.

Information transfer in other rules: 54, 146



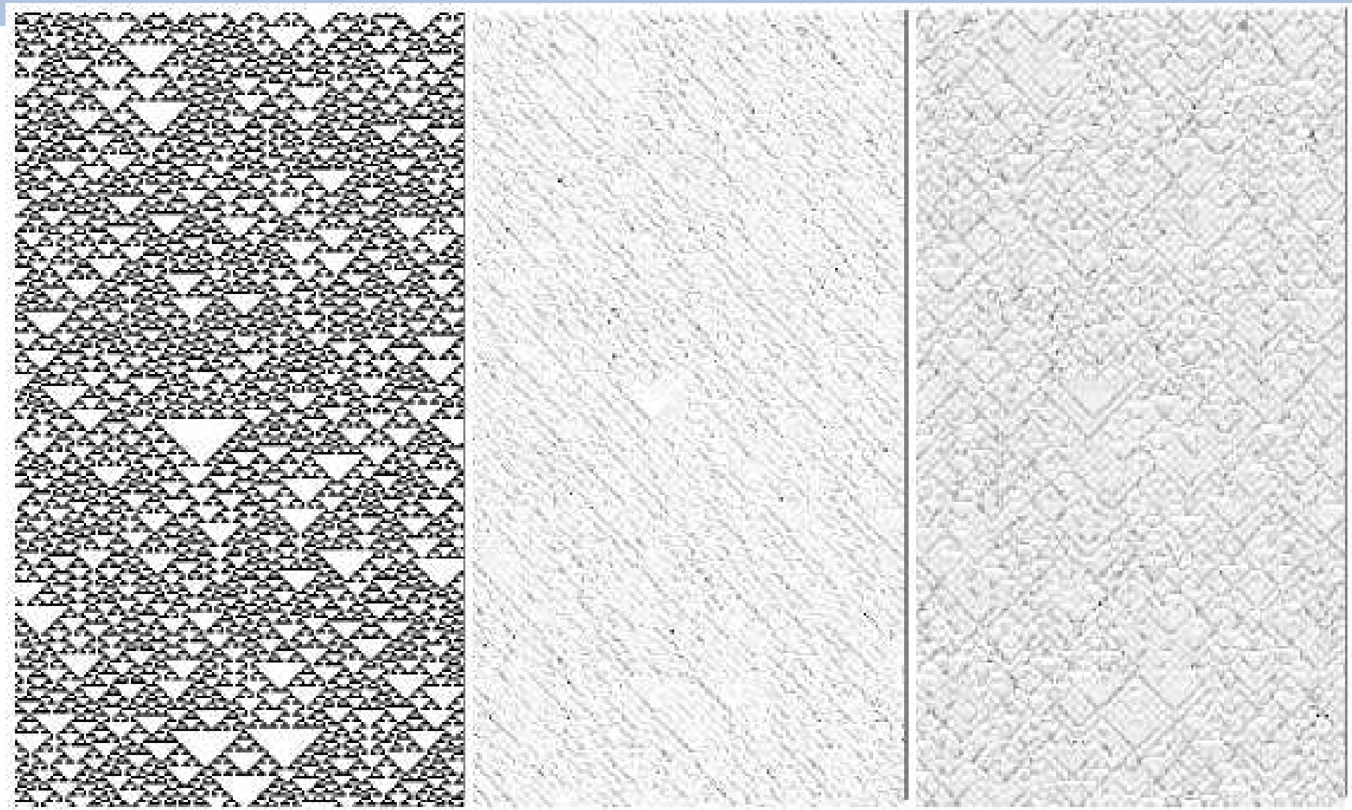
Rule 54 – complex rule

- Further evidence
- Vertical gliders not shown

Rule 146 – domain walls

- Domain walls also dominant info tx agents
- More significant tx in domain.

Information transfer in other rules: 22



Raw CA $t_c(i,j=1,n,k=11)$ t_c summed over j=-1,0,1

- No structure distinguishable.

Local transfer entropy: Conclusions

- Gives insights not obtainable with averaged measure:
  - To parameters of metric: Importance of $k \rightarrow \infty$ demonstrated.
  - To difference between apparent and complete metrics
  - To subtleties within systems.
- New filter for spatiotemporal structure and local information dynamics.
- Long-held conjecture that particles are information transfer agents proven quantitatively.
  - Evidence that transfer entropy is appropriate information transfer metric for complex systems.
- Future work:
  - Comparison to a local version of information flow.
  - Apply to other complex systems and conjectures about signalling therein, e.g. microtubules.

CSIRO