

Random walks in cellular automata

Kari Eloranta†

Institute of Mathematics, Helsinki University of Technology, 02150 Espoo, Finland

Received 19 October 1992

Recommended by P Grassberger

Abstract. Topological defects or phase boundaries discerned in a number of one-dimensional cellular automata appear to perform random walks as well as simpler motions. We analyse their properties rigorously using probabilistic methods. This results in a complete classification in the partially permutative case. The paper complements the companion paper in this issue where the general framework of tilings and subpermutivity was introduced and non-probabilistic properties were analysed.

AMS classification scheme numbers: 60J15, 82C41

Introduction

In this paper we give a comprehensive analysis of random walks in one dimensional cellular automata. The results are rigorous and based on probabilistic methods with which we analyse closely related random walks on directed graphs. The paper is mathematically self-contained and can be read as it is. However an applications-oriented reader's motivation would likely be enhanced by at least a superficial acquaintance with the companion paper [1]. There we present the theory of partial permutative cellular automata including the physical motivation to signal/random walk models, the general algebraic structures, the spectrum of interaction types in the case of multiple walks as well as the subtle walks that still remain unexplained in some cellular automata.

The structure of the presentation is as follows. We first briefly list the basic definitions and results concerning partially permutative cellular automata. Some of these can be found in more elaborated form in [1] but we believe that frequent cross-references there would make the reading too cumbersome. After this we proceed with identifying the boundary motion and unveiling its graph theoretic formulation. The graph naturally carries a Markovian random walk which in turn uniquely determines the possibly non-Markovian motion of the boundary walk. The analysis then branches into two cases (sections 2.1 and 2.2) depending on essentially the strength of the subalphabet interaction. In the first case both the degenerate case of a rectilinearly moving boundary motion and a proper random walk are characterized via the graph representation. The second case only supports random walks but their structure is more involved. The mixed case is elaborated in section 2.3. Along the way we present results that identify the types as well as parameters of the boundary motions explicitly. Some examples as well as results relevant to e.g. the well documented elementary cellular automata are also included. However the main bulk of the interface to empirical studies is in [1] where we have collected some pointers to the earlier approaches to identify topological defects in lattice models.

† Research partially supported by the Academy of Finland and The Finnish Cultural Foundation.

1. Basic definitions

Let $S = \{0, 1, \dots, |S| - 1\}$ be a finite *alphabet* i.e. the set of *symbols* and $X = S^{\mathbb{Z}}$ and $X^{(1/2)} = S^{\mathbb{Z}+1/2}$ be the sets of *configurations*. Equipped with the product topology they become compact metric spaces homeomorphic with a Cantor set. Let σ be the *left shift* by one on a sequence in $X \cup X^{(1/2)}$: $\sigma(x)_j = x_{j+1} \forall j$. A *two-block map* $f(x_j, x_{j+1})$ defines the new value of the coordinate $x_{j+1/2}$.

Definition 1.1. One-dimensional cellular automaton (c.a.) is a dynamical system on $X \cup X^{(1/2)}$ defined by a two-block map which commutes with the left shift.

The global map from X to $X^{(1/2)}$ or from $X^{(1/2)}$ to X defined by requiring $F(x)_{j+1/2} = f(x_j, x_{j+1})$ for all $j \in \mathbb{Z}$ or $\mathbb{Z} + 1/2$ is continuous and conversely any such continuous map that commutes with the shift is induced by a block map (argued as in [5]). The block map f is also called the *rule* of the automaton.

Our definition is superficially different from the usual definition of a c.a. However as shown in [1] (definition 1.3) there is a simple way via substitutions or tilings to generate from an arbitrary n -block map a two-block map on a larger alphabet. Since this is particularly useful in analysing permutivity we present the definition most natural to the subsequent analysis. This by no means restricts the applicability of the results—almost all random walks arising in one-dimensional cellular automata (with a rule of any block length) are still covered. The binary operation (multiplication) represented by the two-block map is conveniently expressed in the form of a Cayley table (see e.g. figure 1 below).

The following sets are of paramount importance to this paper.

Definition 1.2. A set $S_r \subset S$ is a right-invariant subalphabet if $f(r, S_r) = S_r$, $\forall r \in S_r$ i.e. $f(r, \cdot)$ is right permutive on S_r for each $r \in S_r$. Left-invariant subalphabets are defined in the obvious symmetric way.

One usually wants to consider maximal such subalphabets i.e. ones that cannot be augmented by any element from the complement without loss of the permutivity property. If this set is the full alphabet the c.a. is (left/right)-permutive ([5]). If a non-trivial maximal invariant subalphabet exists we call the c.a. *partially permutive*. These c.a. are much more abundant than the permutive ones.

For example in [5] and in [7] permutive rules are considered in detail and a number of important results are established. Two key ones for our analysis are distilled in the following theorem. Recall that the Bernoulli measure $B(1/|S|, \dots, 1/|S|)$ is the product measure with uniform weights $1/|S|$ on symbols.

Theorem 1.1. Permutive cellular automata are onto and preserve the uniform Bernoulli measure $B(1/|S|, \dots, 1/|S|)$ (on X and $X^{(1/2)}$ appropriately).

A set of configurations is generated by a subalphabet S' if all its elements have their coordinates in this set.

Corollary 1.1. The action of a partially permutive c.a. on the set of configurations generated by a permutive subalphabet S' is permutive and preserves the $B(1/|S'|, \dots, 1/|S'|)$ -measure.

2. Characterization of the boundary motion

From here we only consider the strictly subpermutive case. Since the full alphabet is rarely referred to we use the symbol S for a subalphabet. When two different permutive subalphabets exist for a c.a. we have two different permutive actions on configurations generated by them separately. Let the subalphabets be S and T . The natural question to ask then is what happens under the c.a. iteration if these two 'phases' are mixed, i.e. $x \in X$ consists of blocks from these subalphabets. In this paper we confine the analysis to the basic case of two semi-infinite blocks generated from the subalphabets. A compact notation for the set of all such configurations is $\underline{ST}$. In order to preserve this set-up under the iteration of the rule we furthermore require that the interaction between the subalphabets is closed i.e. that for all $s \in S$ and $t \in T$ $f(s, t)$ belongs to one of the two subalphabets.

Definition 2.1. Given two subalphabets S and T let $A = S \cap T$ be the set of ambiguous symbols. If it is nonempty it is by itself an invariant subalphabet. Ambiguous symbols are receding i.e. $f(s, a) \in S \setminus A$ for all $s \in S \setminus A$, $a \in A$ and T identically. The configurations in the set $\underline{ST}(j) = \{ \{s_k\} | s_k \in S \forall k \leq j, s_k \in T \forall k > j \text{ and } s_j, s_{j+1} \notin A \}$ are said to have a boundary point at $j + 1/2$.

Note that if A is empty i.e. the subalphabets are disjoint then every configuration from $\underline{ST}$ is unambiguously in some $\underline{ST}(j)$. If $A \neq \emptyset$ then any configuration of the form $\underline{SAT}$ where $\underline{A}$ is a finite block of symbols from A is eventually reduced to the form $\underline{ST}$. Therefore the definition above applies again and we define the location of the boundary point in between these instances by interpolating.

2.1. The unambiguous interaction

We now proceed to characterize the underlying graphs that determine the motion of the boundary point. In this section we restrict to the case where $f(s, t)$ is unambiguous whenever at least one of s and t is unambiguous. The general case is analysed in sections 2.2 and 2.3.

Let $S = \{s_1, s_2, \dots, s_n\}$ and $T = \{t_1, t_2, \dots, t_m\}$ be two permutive subalphabets that generate the set of configurations $\underline{ST}$ for a two-block map f . Consider the set N of all boundary pairs (i.e. pairs that can be seen around the boundary point) (s, t) , $s \in S$ and $t \in T$. If $|A| = k$ then N has exactly $nm - k^2$ elements. Since now $f(s, t)$ belongs to either $S \setminus A$ or $T \setminus A$ the c.a. action induces a directed graph G on the node set N . These nodes are called *type I*. The fan-out i.e. the number of possible successors of a given node is n or m depending on whether $f(s, t) \in T$ (t dominates) or $f(s, t) \in S$ (s dominates). The fan-in is unrestricted i.e. between zero and $nm - k^2$. The graph can have self-loops but no parallels and in general it is just weakly connected.

In figure 1 we have the Cayley table of a simple c.a. together with the graph. The permutive subalphabets are $S = \{1, 2\}$ and $T = \{3, 4\}$. For simplicity we have chosen the rule to be symmetric but given the ordering $\underline{ST}$ we are really only interested in the shaded elements in the framed square. At the nodes of the graph the pair (s, t) is on top of $f(s, t)$.

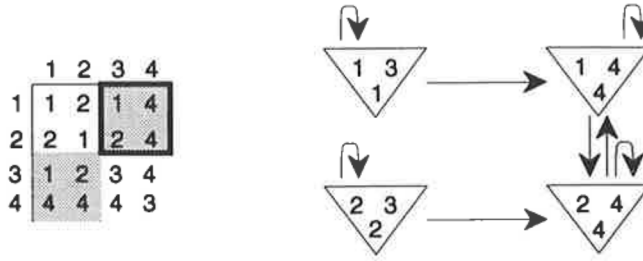


Figure 1.

Note that the ambiguous symbols can, up to bookkeeping, be treated exactly as the unambiguous symbols. The location of a boundary point is defined by interpolation for all configurations in a given evolution starting from any element in ST . Hence whenever the boundary pair is of the form (s, a) , $s \in S \setminus A$, $a \in A$ we know that the symbol a should be counted to belong to T ((a, t) analogously; a is in S).

The node set naturally splits in two subsets. We call N_{tr} the set of *transient* nodes if for a node $n \in N_{tr}$ either (i) there exists a transition from it such that after that it is impossible to re-enter n or (ii) a node of type (i) can be reached in a finite number of steps from n . The complement of N_{tr} is the set of *recurrent* nodes N_{rec} . In the forthcoming analysis all transition probabilities on the edges will be positive so to obtain the equilibrium characteristics of the boundary motion it will suffice to restrict ourselves to the set N_{rec} . In a moment we will investigate under which conditions this set is strongly connected.

Transitions on G result in a walk $\{X_j\}_{j \geq 0}$ on the graph which in turn uniquely determines the motion of the boundary point. Depending on the edge chosen the boundary point either jumps to the right or to the left by $1/2$. Let this increment function be $\Delta(n)$. By keeping track on the partial sum $S_i = \sum_{j=1}^i \Delta(X_j)$ we will be able to locate the boundary at the i th period. The walk on the graph is Markovian and the S_i -process is stationary but in general non-Markovian.

We will now present a lemma that explains how the successor node is selected.

Lemma 2.1.1. *Suppose that at the i th iterate of the c.a. starting from a configuration in ST the boundary point is at $1/2$. The past of the boundary motion is then determined by the block from $-i/2$ to $i/2 + 1$ endpoints included. Let the boundary pair be (s, t) and $f(s, t) = s'$. Then the successor node for the graph walk is (s', t') where t' is uniquely determined by the past and present of the boundary motion and the entry at $i/2 + 2$ in the initial configuration. The left jumps are determined analogously.*

Proof. From the spacetime evolution this result becomes obvious. In figure 2. the past (backward cone of (s, t)) generated by $\{-i/2, \dots, i/2 + 1\}$ is the large triangle around the $1/2$ -line. By the permutivity of the 2-block map given this past the entry at the initial configuration at $i/2 + 2$ determines the entry after one iterate at $(i+3)/2$. But this argument can obviously be iterated $i+1$ times (the entries under arrow in the figure) and therefore given the past and present of (s, t) the entry at $i/2 + 2$ uniquely determines t' . $\square$

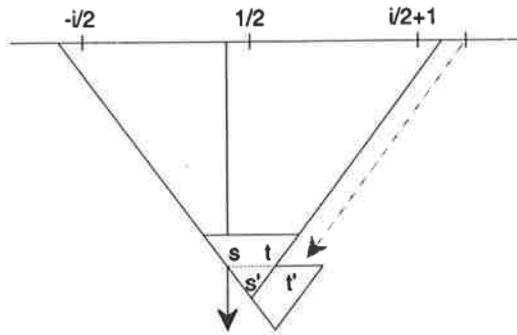


Figure 2.

For a generic initial condition on $\underline{ST}$ the transition probabilities on the graph can be easily determined. By corollary 1.1 we know that the appropriate Bernoulli product measures are the invariant distributions on configurations generated by S and T . The genericity is in the sense of these measures.

Let $\mathbb{Z}_-$ be the set of non-positive integers.

Proposition 2.1.1. *Suppose that on $\underline{ST}(0)$ we have the product measure which is uniform i.e. $B(1/m, \dots, 1/m)$ on $T^{\mathbb{Z}_+}$ and $B(1/n, \dots, 1/n)$ on $S^{\mathbb{Z}_-}$. Then at each node of G the transition probabilities are uniform.*

Proof. Suppose that (s, t) is the boundary pair and $f(s, t) = s'$. By the Lemma 2.1 we know that given the past of the boundary pair (s, t) the follower node (s', t') is determined permutively by a single entry in the positive part of the initial configuration because t' is. But these symbols are $B(1/n, \dots, 1/n)$ -distributed. $\square$

From here we restrict ourselves to the generic case i.e. assume the initial condition to be distributed as indicated in the proposition above. We call this measure the *natural measure*.

Note that in view of the proof of lemma 2.1.1 our initial assumption on the permutivity of S and T can be weakened. When ordered $\underline{ST}$ we only need S to be left-permutive and T to be right-permutive.

We now establish the dichotomy of the boundary motions.

Definition 2.1.1. *A signal is a boundary motion that eventually moves monotonically to either right or left with maximum speed i.e. $1/2$ at each iterate.*

This motion is obviously statistically degenerate i.e. has drift equal to $\pm 1/2$ and vanishing dispersion (variance).

Definition 2.1.2. *Suppose X is a Markovian random walk with uniform transition probabilities on the graph G . If the tail of the jump sequence $\{\pm 1/2\}$ that it generates is not deterministic the boundary point is said to perform a random walk.*

Remark. In many cases the boundary process is also Markovian and then this definition agrees with the usual definition of a random walk. Due to the generating mechanism we however feel that it is appropriate to call all boundary motions of the second type random walks.

Theorem 2.1.1. *The motion of a boundary point starting from an initial configuration distributed according to the natural measure is either a signal or a random walk.*

Proof. We will show that if the boundary motion travels at a speed strictly less than $1/2$ then any tail of the $\pm 1/2$ -sequence describing its motion must be non-deterministic.

Let $s_{i_1} \in S$ appearing in a boundary pair be dominating i.e. $f(s_{i_1}, T) \in S \setminus A$. This now generates a set of chains:

$$s_{i_2} = f(s_{i_1}, T_1), s_{i_3} = f(s_{i_2}, T_2), \dots, s_{i_k} = f(s_{i_{k-1}}, T_{k-1}), f(s_{i_k}, T) \in T$$

$$s'_{i_2} = f(s_{i_1}, T'_1), s'_{i_3} = f(s'_{i_2}, T'_2), \dots, s'_{i_k} = f(s'_{i_{k-1}}, T'_{k-1}), f(s'_{i_k}, T) \in T$$

etc. where $T_i \subset T$, $T_1 \cup T'_1 \cup T''_1 \cup \dots = T$, etc. Here we just list all the sequences generated by different choices of t 's in the boundary pair upto a dominating t . All these chains have to be finite to obtain a speed less than $1/2$. Moreover the sequences $\{s_{i_1}, \dots, s_{i_k}\}$ need to be of equal length since otherwise we would have at some iterate ambiguity whether the motion turns (depending on which $\{s_{i_1}, \dots, s_{i_k}\}$ -block we follow). So let us consider the first block. Suppose s_{i_k} is located at (j, i) . Let s_l be its left descendant at $(j - 1/2, i + 1)$. Now we have $f(s_l, t) \in S$ for example when $l = i_1$ for all $t \in T$. On the other hand $f(s_l, t) \in T$ for $s_l = i_k$ for all $t \in T$. But from lemma 2.1 we know that the entry s_l is determined permutively from the initial condition. Hence if the speed is less than $1/2$ we are bound to have a random choice between a left and a right jump. $\square$

The proof of the theorem immediately implies the following result which is useful in finding the type of dynamics directly from the Cayley table. It generalizes the earlier notion of a dominant symbol.

Corollary 2.1.1. *The dominant chain condition, $\exists S' \subset S$ such that $\forall s' \in S', f(s', T) \in S'$, (and its symmetric counterpart for T' -dominance) is a necessary and sufficient condition for a boundary motion not to be a random walk.*

Remark. The c.a. in figure 1 has a dominant symbol $t = 4$ i.e. a dominant chain $T' = \{4\}$. After a finite transient the c.a. exhibits a left propagating signal.

The motion type that the random walk on a given graph can generate need not be unique. The uniqueness is related to the graph topology in the following fashion.

Theorem 2.1.2. *If the random walk X_i restricted to the recurrent part of the graph generates a boundary random walk the recurrent part must be strongly connected.*

Since the random walk on a strongly connected graph uniquely determines the statistical properties of the boundary motion Theorem 2.1.2 immediately implies a co-existence result.

Corollary 2.1.2. *A signal and a boundary random walk or two different boundary random walks cannot be generated from the same strongly connected graph.*

The existence of a signal is a consequence of the existence of a closed (no transitions out) subgraph in G the nodes of which generates only left or only right jumps. Since several such subgraphs may exist in a weakly connected graph multiple signals may exist and in particular propagating to either direction. If however the subgraph is all of G i.e. one of the subalphabets is dominant only one signal exists.

Proof of theorem 2.1.2. Pick two recurrent nodes (s, t) and (s', t') . Form the follower sets F and F' of both i.e. sets of nodes that can be reached from them in any number of steps. Their elements are recurrent nodes since a follower of a recurrent node is recurrent. If the nodeset N is thought as an $n \times m$ array minus a $k \times k$ corner its subsets F and F' both contain rows and columns of full length. This is because the existence of a boundary random walk generated from the set of recurrent nodes guarantees the absence of dominant chains and hence both left and right jumps are bound to happen starting from either one of (s, t) or (s', t') . But by the geometry of F and F' they intersect and from any element in the intersection the starting points can be reached by recurrence. Hence (s, t) and (s', t') communicate. $\square$

Since the transition probabilities given by proposition 2.1.1 are positive for all edges the strong connectedness implies that the random walk on the recurrent part is irreducible and the nodes are positively recurrent. If the transition probability matrix is denoted by P then the equilibrium distribution, π , on the nodes is the solution of $\pi P = \pi$. From this we get the characterization of the parameters of the random walk by a simple application of the Ergodic Theorem. Let $N^+ \subset N_{\text{rec}}$ be the subset of nodes with right jumps i.e. $\Delta(N^+) = +1/2$.

Theorem 2.1.3. Suppose that we have a c.a. on ST with an unambiguous boundary action. Let the initial distribution be according to the natural measure. If the resulting boundary motion S_i is a random walk then its expected spatial shift in unit time i.e. the drift is

$$d \triangleq \lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I \Delta(X_i) = \sum_{n \in N_{\text{rec}}} \Delta(n) \pi(n) = \pi(N^+) - \frac{1}{2}$$

and the unit squared variation equals to

$$\sigma^2 \triangleq \lim_{I \rightarrow \infty} \frac{1}{I} \sum_{i=1}^I (\Delta(X_i) - d)^2 = \sum_{n \in N_{\text{rec}}} (\Delta(n) - d)^2 \pi(n) = \frac{1}{4} - d^2.$$

Note that is of course compatible with our earlier result on the drift and variance of a signal in the case of a strongly connected graph on N_{rec} .

Example 2.1.1. Suppose that we have the subalphabets $S = \{1, 2\}$ and $T = \{3, 4\}$ and a rule on $\{1, 2, 3, 4\}$ represented by a Cayley table in figure 3. Note that it is only slightly different from that of figure 1. But the corresponding graph is now strongly connected, each transition has probability $1/2$ and the equilibrium distribution π is uniform. The boundary walk generated is Markovian since $\{\Delta(X_i)\}$ is an independent sequence. By the theorem the walk has zero drift and variance $1/4$.

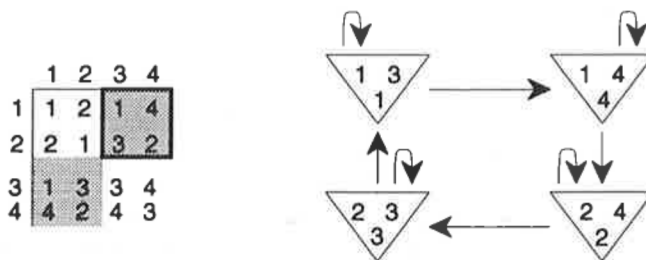


Figure 3.

The fan-in of a type II node is unrestricted and the fan-out equals to $(n - k)(m - k)$. By the lemma they in fact map onto M .

An important special case is an automaton for which all interactions between S and T result in an ambiguous symbol. So let us suppose that all interaction on M is of type II. Then the graph restricted to M is strongly connected (in one step i.e. the directed graph is complete). Moreover the complement $N \setminus M$ is clearly transient so $N_{\text{rec}} = M$. Apart from a possible finite initial transient of jumps on $N \setminus M$ the following result pins down the resulting motion.

Theorem 2.2.1. *Let the subalphabets have the cardinalities $|S| = n$, $|T| = m$ and $|A| = k$, $1 \leq k < \min\{n, m\}$. Suppose that $f(s, t) \in A$ whenever neither s nor t is ambiguous. Let the random variables Δ and τ be as in the lemma. Then the boundary motion performs a random walk with independent and identically distributed holding times of length τ and i.i.d. increments Δ . The drift of the motion is*

$$d = \frac{E(\Delta)}{E(\tau)} = \frac{k(n - m)}{2nm - (n + m)k}$$

and the unit variance equals to

$$\sigma^2 = \frac{k}{4(2nm - k(n + m))} \left(\frac{m(n - k)}{m - k} + \frac{n(m - k)}{n - k} \right).$$

In particular in the symmetric case $n = m$ the drift vanishes and the dispersion reduces to $k/(4n - 4k)$.

Proof. We first observe that in computing Δ and τ we can restrict to the case where $s' = s$ and $t' = t$ in lemma 2.2. But this means that all the nodes of the the generating graph have same (Δ, τ) -distributions associated with them. Moreover since the pairs (Δ, τ) at different nodes are independent we only need to consider the case of a graph with a single node and a loop. This generates a boundary random walk with independent increments Δ delayed by τ iterates. Now the delays $\{\tau_i\}_{i \geq 0}$ form a recurrent renewal process hence an application of the Renewal Theorem to the cumulative jumps $\{S_i\}_{i \geq 0}$ yields the drift

$$d = \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{j=1}^{I(t)} \Delta(X_j) = \frac{E(\Delta)}{E(\tau)}$$

where $I(t)$ is the usual counting function i.e. the number of delay periods up to time t . But by the independence of L and R the expectations are easily calculated to be

$$E(\Delta) = \frac{k(n - m)}{2(n - k)(m - k)} \quad E(\tau) = \frac{2nm - nk - mk}{(n - k)(m - k)}.$$

Since $\sigma^2(\Delta) = (\sigma^2(L) + \sigma^2(R))/4$ the unit variance formula follows via a similar argument. $\square$

Remarks. 1. By choosing $n = m = 2$ and $k = 1$ this result covers a number of random walks arising e.g. in the context of elementary c.a. In particular the walks in the rules 18 (treated in [2]) and rule 22 (see [1]) are just special cases of the theorem above.

2. It is possible to reduce the ambiguous case to the unambiguous one via an extension. If each of the ambiguous symbols is duplicated and assigned to one of the subalphabets the new subalphabets are disjoint. On the enlarged alphabet one can then define a new c.a. dynamically identical to the original one (i.e. possessing statistically identical boundary motion).

2.3. The mixed case

The previous results make the general case of unrestricted $f(s, t)$ now accessible. We still consider a directed graph G with $nm - k^2$ nodes of the two types described and the random walk X_i on it. Let again $M = (S \setminus A) \times (T \setminus A)$. Since off- M entries do not form a closed subgraph (symbols in A are receding) the novel case to be treated arises when M contains nodes of both type.

We recall that the graph restricted to type II nodes is strongly connected and from any node in this set we map onto M . So in terms of recurrence the critical question is how do the follower sets of the nodes of type I lie in M . If their union is in the complement of type II nodes then N_{rec} is of type I and we reduce to the case treated in section 2.1 (this phenomenon is analogous to the dominant chain case in section 2.1). Note also that the size of such type I invariant set is bounded from below by the cardinality of A (just because (s, a) has to be in the set if s is for all $a \in A$). Hence the larger A is the harder it is to type I to confine the action on itself.

Let N_I and N_{II} be a partition of N_{rec} into type I and II nodes and let $N_I^+ \subset N_I$ be the subset of nodes with right jumps. Assume that the graph on N_{rec} is strongly connected and denote again by P and π is the transition probability matrix and equilibrium distribution on N_{rec} . Define the expected displacement and visit time at a node by

$$\bar{\Delta} \triangleq \sum_{n \in N_{\text{rec}}} E(\Delta(n)) \pi(n) = \pi(N_I^+) - \frac{1}{2} + \left(\frac{1}{2} + E(\Delta|N_{II}) \right) \pi(N_{II})$$

and

$$\bar{V} \triangleq \sum_{n \in N_{\text{rec}}} E(\tau(n)) \pi(n) = \pi(N_I) + E(\tau|N_{II}) \pi(N_{II}).$$

These can be readily evaluated using theorem 2.2.1.

Theorem 2.3.1. *Let a c.a. act on ST with the natural measure. Suppose that the graph on N_{rec} is strongly connected and $N_{II} \neq \emptyset$. Then the boundary motion generated is a stationary random walk with drift $d = \bar{\Delta}/\bar{V}$ and variance*

$$\sigma^2 = \frac{1}{\bar{V}} \left\{ \left(\frac{1}{2} + \bar{\Delta} \right)^2 - 2\bar{\Delta}\pi(N_I^+) + \left(E((\Delta - \bar{\Delta})^2|N_{II}) - \left(\frac{1}{2} + \bar{\Delta} \right)^2 \right) \pi(N_{II}) \right\}.$$

Remark. For $N_{II} = \emptyset$ and $N_{II} = M$ we have the unambiguous and ambiguous case respectively so the novelty here appears when $\emptyset \neq N_{II} \neq M$. $N_{II} = \emptyset$ is excluded from the theorem since then (and only then) degeneracy can occur.

Proof. By the assumptions on the existence and communication of the type I and II nodes we know that a unique random walk prevails and is stationary. For the drift we write

$$\frac{1}{t} \sum_{j=1}^{I(t)} \Delta(X_j) = \frac{I(t)}{t} \left\{ \frac{1}{I(t)} \sum_{j=1}^{I(t)} \Delta(X_j) \right\}$$

where $I(t)$ counts the number of nodes visited by time t . The graph walk X_j is positive recurrent on N_{rec} so by the Renewal Theorem $I(t)/t \rightarrow 1/\bar{V}$. Moreover the random variable Δ is integrable (on N_I it is bounded and on N_{II} it is the difference between geometric random variables) so by the Ergodic Theorem the limit $I \rightarrow \infty$ of the remaining Cesaro average equals to $\bar{\Delta}$.

The asymptotic squared variation is obtained via an analogous argument and some manipulation starting from the expression $1/t \sum_{j=1}^{I(t)} (\Delta(X_j) - \bar{\Delta})^2$. $\square$

All the expressions in the theorem can be explicitly computed. Perhaps an application is in order to illustrate this.

Example 2.3.1. Consider the simplest case where nodes of both type can co-exist in the recurrent component. Let $S = \{0, 1\}$, $T = \{0, 2, 3\}$ so $A = \{0\}$ and N consists of five nodes. We have a choice in terms of which node in M is of which type and which ones in its complement are transient. Of these six c.a. two have a dominant symbol i.e. they generate a signal. The remaining four yield akin random walks and are all analysed in a similar fashion. In figure 5 we have illustrated one of these (if the shaded element in the Cayley table were equal to 3 a signal-case would result). The two rightmost elements of the graph are transient. Clearly $N_{\text{rec}} = \{(1), (2), (3)\}$, $N_I^+ = N_I = \{(1), (3)\}$ and $N_{II} = \{(2)\}$.

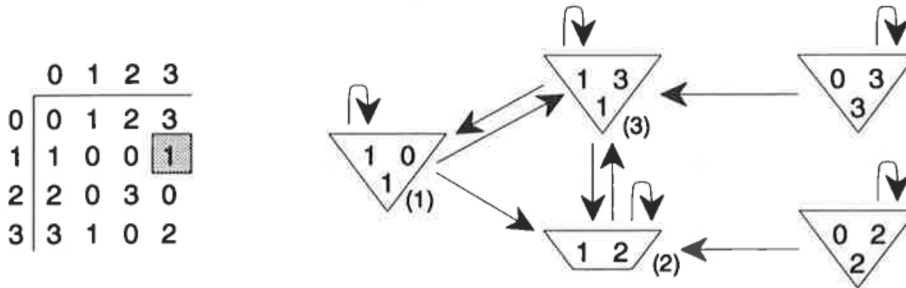


Figure 5.

The transition matrix and equilibrium visit probabilities are

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad \text{and} \quad \pi = \left(\frac{1}{5}, \frac{2}{5}, \frac{2}{5}\right).$$

The (Δ, τ) -pair associated with the nodes (1) and (3) is $(1/2, 1)$. At (2) it has the distribution indicated in lemma 2.2.1. The expectation $E(\Delta, \tau)$ equals to $(-1/4, 7/2)$ by the formulae in the proof of theorem 2.2.1. Hence by theorem 2.3.1 a boundary random walk is generated and with some computation one finds that the unit drift equals to $1/10$ and variance approximately to 0.205. The reader is to judge whether it is obvious that the walk drifts so slowly to the right although the only sure transitions are to this direction (at type I nodes).

When the type II is recurrent no co-existence result of the type discussed in section 2.1 is possible. However even though the random walk that theorem 2.3.1. specifies is unique it can be rather bizarre and actually look like two distinctly different boundary motions intertwined. Our final example shows how to 'design' c.a. like this.

Example 2.3.2. Suppose that $n = m \gg k = 1$ and that there are unambiguous special symbols $\tilde{s}$ and $\tilde{t}$ such that $f(\tilde{s}, t) = \tilde{s}$ except for $t = \tilde{t}$. Also let $f(s, \tilde{t}) = \tilde{t}$ for all s except for $s = s_1$. Let all the rest of the nodes in $M = (S \setminus A) \times (T \setminus A)$ be of type II. Now the theorem applies but by design the sets N_I and N_{II} communicate only with difficulty i.e. jumps between them are rare. They also generate very different motions. The type I motion

consists of long monotone sequences of jumps of the same sign and size $(1/2)$ so that the motion is a "piecewise signal". The right propagating boundary point eventually switches into a left propagating one whereas this ultimately (upon visit to $(s_1, \tilde{t})$) gives birth to a type II symmetric random walk (by theorem 2.2.1 since $n = m$). This in turn is longlived since type II covers nearly all of M but it is not immortal because type II nodes are onto this set.

Clearly this design can be widely varied by changing S or T , by introducing bias to either type, by extending A (and hence N_I) etc. and still generate pairs of very distinct and persistent motions. The key design principle is to partition the graph on the recurrent nodes into two almost disconnected parts supported by N_I and N_{II} each generating a distinctly different motion.

The physical phenomenon this example seems to suggest is that of *metastability*. Considering the amount of freedom in our example this behaviour can indeed be widespread in all but the simplest one-dimensional cellular automata.

References

- [1] Eloranta K 1993 Partially permutative cellular automata *Nonlinearity* **6** 1009–23
- [2] Eloranta K and Nummelin E 1992 The kink of the elementary cellular automaton Rule 18 performs a random walk *J. Stat. Phys.* **69** 5/6
- [3] Feller W 1966 *An Introduction to Probability Theory and Its Applications* vol II (New York: Wiley)
- [4] Harary F 1969 *Graph Theory* (Reading, MA: Addison-Wesley)
- [5] Hedlund G A 1969 Endomorphisms and automorphisms of the shift dynamical system *Math. Sys. Th.* **3** 320–75
- [6] Petersen K E 1983 *Ergodic Theory* (Cambridge: Cambridge University Press)
- [7] Shivani M and Rogers T D 1991 On ergodic one-dimensional cellular automata *Commun. Math. Phys.* **136** 599–605