CDMTCS
Research
Report
Series

Embedding Cellular Automata into Reversible Ones

Peter Hertling
Department of Computer Science
University of Auckland

CDMTCS-065
November 1997

Centre for Discrete Mathematics and
Theoretical Computer Science

Embedding Cellular Automata into Reversible Ones

Peter Hertling*
Department of Computer Science, University of Auckland
Private Bag 92019, Auckland, New Zealand
email: `hertling@cs.auckland.ac.nz`

Abstract

Toffoli showed that every cellular automaton of an arbitrary dimension d can be embedded into a reversible cellular automaton of dimension $d+1$. He asked "whether an arbitrary cellular automaton can be embedded in a reversible one having the same number of dimensions" and conjectured that this is not possible. We show that his conjecture is true. Even if one imposes only a weak, natural condition on embeddings, no cellular automaton which possesses a Garden of Eden configuration can be embedded into a reversible cellular automaton of the same dimension.

1 Introduction

Cellular automata are a mathematical model for the simulation of complex processes. They were introduced by Ulam and von Neumann [16] in order to study the evolution in time of complex, self-reproducing biological, physical or mathematical systems which exhibit uniform behaviour over a certain region of space. Cellular automata show this behaviour with the additional simplifying assumptions of a local discrete state space with a regular structure and discrete time. The space on which a cellular automaton operates is a homogeneous lattice of cells each of which is in one of a finite number of states. In each time step each cell is updated according to the same rule as all other cells. The next state of a cell depends only on the states of finitely many cells. A precise definition of a cellular automaton will be given in the following section.

The issue of reversibility of mathematical, physical or biological processes has a long history, see Toffoli [14] and the references cited therein. In modelling these, also the reversibility of cellular automata has received great attention. A cellular automaton is called reversible if it is injective, that means, if any configuration can have at most one predecessor. By a result of Richardson [12], in this case every configuration has exactly one predecessor and the inverse of an injective cellular automaton is a cellular automaton itself. For an overview over reversible cellular automata the reader is referred to Toffoli and Margolus [15]. Cellular automata are also computing devices and a strong motivation for studying reversible computing devices comes from the wish to reduce heat dissipation in computing machinery, compare Landauer [7] and Bennett [1, 2]. It has turned out that reversible cellular automata are capable of universal computation: they can simulate arbitrary and hence, also universal Turing machines, see Toffoli [14]. The question arose whether reversible cellular automata can also simulate arbitrary cellular automata. This was answered positively by Toffoli [14] who showed that any cellular automaton of an arbitrary dimension d can be embedded into a reversible cellular automaton of dimension $d+1$. He posed the question whether perhaps the same dimension suffices, that is, whether any cellular automaton can be embedded into a reversible cellular automaton of the same dimension. He conjectured that the answer

*The author was supported by the DFG Research Grant No. HE 2489/2-1.

is negative. This question was restated by Culik, Hurd, and Yu [4] in the context of computational aspects of cellular automata.

We show that this is in general impossible, thus proving Toffoli's conjecture. Only when one does not impose any effectivity condition on an embedding, one can embed an arbitrary cellular automaton into a reversible cellular automaton of the same dimension, and in fact, then even into a reversible cellular automaton of dimension one. But as soon as one demands that the embedding respects the uniform structure of cellular automata in at least a weak sense — which leads to the notion of a *weak embedding* — one obtains a negative answer.

We recall that a configuration c in the configuration space of a cellular automaton with global map F is called a *Garden of Eden configuration* if there is no configuration from which c can be reached via F. Hence, a cellular automaton has a Garden of Eden configuration if and only if it is not surjective. Starting with Moore's paper [10], in which the notion of a Garden of Eden configuration was introduced, a series of papers by Myhill [11], Richardson [12], Maruoka and Kimura [9], and others have led to interesting other characterizations of surjective cellular automata and the relations between injectivity and surjectivity for cellular automata. The following two results will be derived in the last section from the technical main result, which is stated in terms of spatially periodic configurations.

> *A cellular automaton which possesses a Garden of Eden configuration cannot be weakly embedded into a reversible cellular automaton of the same dimension.*

For one-dimensional cellular automata we obtain a slightly stronger result.

> *An irreversible cellular automaton of dimension one cannot be weakly embedded into a reversible cellular automaton of dimension one.*

In the following section we provide precise definitions of a cellular automaton and of the embedding notions which we consider. Then, in the section about positive results we restate Toffoli's [14] result and give a complete and simplified proof. Furthermore we show that every cellular automaton can be embedded in an ineffective, purely set-theoretic way into a very simple, reversible, one-dimensional cellular automaton. Since this embedding does not make much sense, in Section 4 we consider weak embeddings, which respect the basic structure of cellular automata in a weak sense. First, we show that no cellular automaton can be weakly embedded into any cellular automaton of smaller dimension. Then the main result is proved.

> *A cellular automaton which is not injective on all spatially periodic configurations cannot be weakly embedded into a cellular automaton of the same dimension which is injective on all spatially periodic configurations.*

From this result the two statements above are deduced. We conclude the paper with some remarks about reversible computation by cellular automata.

2 Cellular Automata and Embeddings

This section contains precise definitions of cellular automata and the considered embedding notions. First we introduce the full shift spaces $S^{\mathbb{Z}^d}$ and the shift mappings operating on them. By $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ we denote the set of integers.

Let $d \in \{1, 2, \ldots\}$ be a positive integer and S be a finite set containing at least two elements. The *full shift space*

$$S^{\mathbb{Z}^d} = \{c : \mathbb{Z}^d \to S\}$$

is the infinite product space over S using the lattice $\mathbb{Z}^d$ as an index set (notice that we demand $|S| \geq 2$). Endowed with the product topology of the discrete topology on S the space $S^{\mathbb{Z}^d}$ is a compact topological space by Tychonoff's theorem. We call its elements *configurations*. On this space we have natural shift mappings. Each integer vector $a = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}^d$ induces a bijection $\sigma_a^{(d)} : S^{\mathbb{Z}^d} \to S^{\mathbb{Z}^d}$ defined by $\sigma_a^{(d)}(c)_b := c_{b+a}$, for any $b \in \mathbb{Z}^d$. It is called the *shift map associated with* a. It is clear that

$$\sigma_a^{(d)} \circ \sigma_b^{(d)} = \sigma_{a+b}^{(d)} \tag{1}$$

for arbitrary vectors $a, b \in \mathbb{Z}^d$. One concludes

$$(\sigma_a^{(d)})^n = \sigma_{na}^{(d)} \tag{2}$$

for arbitrary $a \in \mathbb{Z}^d$ and $n \in \mathbb{Z}$. The shift map $\sigma_{e_i^{(d)}}^{(d)}$ associated with the unit vector $e_i^{(d)} = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{Z}^d$ having a 1 at position i and zeros at all other positions is also written $\sigma^{(d),i}$. The shift mapping $\sigma^{(1),1}$ is the usual left shift in the one-dimensional case.

Cellular automata are functions which operate on one full shift space.

Definition 1 A *cellular automaton* (short: *CA*) is a triple (S, d, F) consisting of a finite set S containing at least two elements, called the *set of states*, a positive integer d, called the *dimension*, and a continuous function $F : S^{\mathbb{Z}^d} \to S^{\mathbb{Z}^d}$ which commutes (this means: $F \circ \sigma^{(d),i} = \sigma^{(d),i} \circ F$) with the shift mappings $\sigma^{(d),i}$ for $i = 1, \ldots, d$. The function F is called the *global map* of the CA.

This definition does not reflect the usual characterization via a so-called local function. Since the space $S^{\mathbb{Z}^d}$ is a compact metric space (use for example the metric d defined by $d(c, c') := 2^{-m(c,c')}$ where $m(c, c') := \min\{r \in \{0, 1, 2, \ldots\} \mid \text{there exists an } a \in \{-r, \ldots, 0, \ldots, r\}^d \text{ with } c_a \neq c'_a\}$, for $c, c' \in S^{\mathbb{Z}^d}$ and where $\min \emptyset = \infty$) any continuous function $F : S^{\mathbb{Z}^d} \to S^{\mathbb{Z}^d}$ is uniformly continuous. Hence, if F is continuous and commutes with the shift mappings, then there exist a finite set $A \subseteq \mathbb{Z}^d$ and a function $f : S^A \to S$ such that $F(c)_b = f(c_{b+A})$, for all $c \in S^{\mathbb{Z}^d}$ and $b \in \mathbb{Z}^d$, where $c_{b+A} \in S^A$ is defined in the obvious way: $(c_{b+A})_a := c_{b+a}$ for all $a \in A$. The function f is called a *local* function for F and we say that F *is induced* by f. Obviously, one could choose A to be the d-dimensional cube $\{-r, -r+1, \ldots, 0, \ldots, r-1, r\}^d$ for some sufficiently large r. On the other hand it is clear that any function F induced by a local function f is the global map of a cellular automaton.

Special classes of cellular automata are of great interest. A cellular automaton (S, d, F) is called *reversible* if there is a cellular automaton (S, d, G) such that $F \circ G$ and $G \circ F$ are the identity on $S^{\mathbb{Z}^d}$. In other words, a CA (S, d, F) is reversible if its global map F is bijective and the inverse map F^{-1} defines a cellular automaton (S, d, F^{-1}) as well. By a result of Richardson [12], see also Hedlund [5] for dimension one, a cellular automaton is reversible if and only if its global map F is injective (that means: $F(c) = F(c') \Rightarrow c = c'$, for $c, c' \in S^{\mathbb{Z}^d}$). Another important class is the set of all surjective (that means: for all $c' \in S^{\mathbb{Z}^d}$ there is a $c \in S^{\mathbb{Z}^d}$ with $F(c) = c'$) cellular automata. If a cellular automaton (S, d, F) is not surjective, then a configuration $c \in S^{\mathbb{Z}^d}$ which cannot be reached via F from any configuration in $S^{\mathbb{Z}^d}$ is called a *Garden of Eden configuration*. The classes of surjective and reversible cellular automata have received great attention in the past and were analyzed by Moore [10], Myhill [11], Richardson [12], Maruoka and Kimura [9], and others. For more details on cellular automata the reader is referred to Culik, Hurd, and Yu [4] and the literature cited therein.

We are interested in embedding one cellular automaton into another one, or, in other words, in simulating a cellular automaton by another one. Therefore we first introduce two natural classes of mappings between full shift spaces. The first condition expresses that the mapping commutes at least in a weak sense with the shift mappings: a shift by a vector in the domain must correspond to a shift by some — not necessarily the same — vector in the range.

Definition 2 Let S, T be finite sets with at least two elements and d, e be positive integers.

1. A function $\mu : S^{\mathbb{Z}^d} \to T^{\mathbb{Z}^e}$ is called a *weak morphism* if for each $i \in \{1, \ldots, d\}$ there exists a vector $a^{(i)} \in \mathbb{Z}^e$ such that $\mu \circ \sigma^{(d),i} = \sigma_{a^{(i)}}^{(e)} \circ \mu$.

2. A function $\mu : S^{\mathbb{Z}^d} \to T^{\mathbb{Z}^e}$ is called a *morphism* if it is a weak morphism and continuous.

The continuity condition is another formulation of the condition that the value of a cell of $\mu(c)$ depends only on the values of finitely many cells of c. Note that in this terminology a cellular automaton is given by a morphism with $S = T$, $d = e$, and $a^{(i)} = e_i^{(d)}$ for all $i \in \{1, \dots, d\}$.

Which conditions should an embedding of one cellular automaton into another cellular automaton satisfy? The first condition is certainly that one can read off the behaviour of the first automaton in the proceeding time from the second automaton. This is reflected by the first, set-theoretical notion in the following definition, which is essentially copied from Toffoli's paper [14]. We do not demand that one time step in the first CA corresponds to one time step in the second CA but instead allow a slow-down by a constant factor. A further natural condition is that the embedding should reflect the lattice structure of the full shift spaces and the fact that a CA behaves uniformly in each cell. Furthermore, one can demand that an embedding mapping μ is continuous, hence that the value in a cell of an image $\mu(c)$ of a configuration c depends only on finitely many cells of c.

Definition 3 Let (S, d, F) and (T, e, G) be two cellular automata.

1. A *set-theoretic embedding of* (S, d, F) *into* (T, e, G) is a triple (μ, ν, k) consisting of a mapping $\mu : S^{\mathbb{Z}^d} \to T^{\mathbb{Z}^e}$, a mapping $\nu : T^{\mathbb{Z}^e} \to S^{\mathbb{Z}^d}$, and a positive integer k satisfying

$$F^t = \nu \circ G^{kt} \circ \mu \qquad \text{for all} \quad t \geq 0\,. \tag{3}$$

   The number k is called the *delay factor*. In the case $k = 1$ we speak of a set-theoretic embedding *without delay*.

2. A set-theoretic embedding (μ, ν, k) is called a *weak embedding* if additionally μ is a weak morphism.

3. A set-theoretic embedding (μ, ν, k) is called a *strong embedding* if μ is a morphism.

For $t = 0$, Equation (3) says $c = F^0(c) = \nu \circ G^0 \circ \mu(c) = \nu \circ \mu(c)$ for all configurations $c \in S^{\mathbb{Z}^d}$. Therefore, the mapping μ must be injective. Stronger notions of embeddings can be obtained by imposing conditions also on the mapping ν.

3 Positive Results on Embeddings into Reversible Cellular Automata

The following theorem is the fundamental result of Toffoli [14] on the embedding of a cellular automaton into a reversible cellular automaton stated in the language introduced in the last section.

Theorem 4 (Toffoli [14]) *Every cellular automaton of arbitrary dimension d can be strongly embedded without delay into a reversible cellular automaton of dimension $d + 1$.*

Toffoli gives a complete construction of an embedding only for the case of a cellular automaton which possesses a quiescent configuration (that is, an automaton (S, d, F) such that there is a state $s \in S$ with $F(c) = c$ if $c \in S^{\mathbb{Z}^d}$ is the constant configuration with $c_a = s$ for all $a \in \mathbb{Z}^d$). Furthermore, his construction can be simplified. Therefore we give a complete, simplified proof, which is still based on his ideas.

Proof. Let (S, d, F) be a cellular automaton. We wish to construct a reversible cellular automaton $(T, d + 1, G)$ and a strong embedding $(\mu, \nu, 1)$ of (S, d, F) into $(T, d + 1, G)$.

We can assume without loss of generality $S = \{1, 2, \dots, q\}$. The construction can be viewed as consisting of two steps. First, we add a new state 0 to S in order to obtain a quiescent configuration. Then the new CA is embedded into a CA $(T, d + 1, G)$ whose set of states T consists of pairs over the state set $S \cup \{0\}$. In each time step the left half of a state in T contains the essential information and is changed in the appropriate way while the right half contains a copy of the left half at the previous time step. This is done in order to ensure that the CA $(T, d+1, G)$ is reversible. The idea behind the embedding of (S, d, F) into $(T, d+1, G)$ is to map the configurations in $S^{\mathbb{Z}^d}$ into a hyperplane in $T^{\mathbb{Z}^{d+1}}$ and to use the additional dimension in order to preserve (transformed) copies of previous configurations in parallel hyperplanes.

We will define the CA $(T, d+1, G)$ directly without separating the first step. Besides the cellular automaton $(T, d+1, G)$ we also construct its inverse $(T, d+1, H)$. The set of states T is the set of all pairs over $\{0\} \cup S$:

$$T := \{0, 1, 2, \ldots, q\}^2 = (\{0\} \cup S)^2 = \{(s^{(l)}, s^{(r)}) \mid s^{(l)}, s^{(r)} \in \{0, 1, 2, \ldots, q\}\}$$

The functions G and H will be defined via local functions. Since F is continuous there is a finite set $A \subseteq \mathbb{Z}^d$ and a local function $f : S^A \to S$ such that F is induced by f in the sense explained after Definition 1. We can assume that $(0, \ldots, 0) \in A$.

We define a new local function $\tilde{f} : \{0, 1, \ldots, q\}^A \to \{0, 1, \ldots, q\}$ by

$$\tilde{f}(c) := \begin{cases} f(c) & \text{if } c \in S^A \\ 0 & \text{otherwise} \end{cases}$$

for all $c \in \{0, 1, \ldots, q\}^A$. We will define G and H by local functions $g : T^B \to T$ and $h : T^C \to T$. The finite set $B \subseteq \mathbb{Z}^{d+1}$ is defined by

$$\begin{array}{rcl} B & := & A \times \{0\} \cup \{(0, \ldots, 0, 1)\} \\ & = & \{(\alpha_1, \ldots, \alpha_d, \alpha_{d+1}) \in \mathbb{Z}^{d+1} \mid \begin{array}{l} \text{either } (\alpha_1, \alpha_2, \ldots, \alpha_d) \in A \text{ and } \alpha_{d+1} = 0 \\ \text{or } \alpha_1 = \ldots = \alpha_d = 0 \text{ and } \alpha_{d+1} = 1\}. \end{array} \end{array}$$

We write a state $s \in T$ as $(s^{(l)}, s^{(r)})$ with $s^{(l)}, s^{(r)} \in \{0, 1, \ldots, q\}$ and an element $c \in T^B$ as $c = (c^{(l)}, c^{(r)})$ with $c^{(l)}, c^{(r)} \in \{0, 1, \ldots, q\}^B$. The local function $g : T^B \to T$ is defined by

$$\begin{array}{rcl} g(c)^{(l)} & := & c^{(r)}_{(0,\ldots,0,1)} \oplus \tilde{f}((c^{(l)}_{(\alpha_1,\ldots,\alpha_d,0)})_{(\alpha_1,\ldots,\alpha_d) \in A}), \\ g(c)^{(r)} & := & c^{(l)}_{(0,\ldots,0,0)} \end{array}$$

for $c \in T^B$. Here $\oplus$ denotes the addition on $\{0, 1, \ldots, q\}$ modulo $q+1$. Remember that we assume $(0, \ldots, 0) \in A$ and hence $(0, \ldots, 0, 0) \in B$. The function $G : T^{\mathbb{Z}^{d+1}} \to T^{\mathbb{Z}^{d+1}}$ induced by g copies each left component of a cell c_b of a configuration $c \in T^{\mathbb{Z}^{d+1}}$ to its right component and assigns the modulo $q+1$ sum of the right component of the cell above the considered cell c_b and of the $\tilde{f}$-value of the left components of the cells in a neighborhood (corresponding to A) of c_b to its left component. We define the mapping $H : T^{\mathbb{Z}^{d+1}} \to T^{\mathbb{Z}^{d+1}}$ as the mapping induced by the following local function $h : T^C \to T$ with

$$\begin{array}{rcl} C & := & A \times \{-1\} \cup \{(0, \ldots, 0, 0)\}, \\ h(c)^{(l)} & := & c^{(r)}_{(0,\ldots,0,0)}, \\ h(c)^{(r)} & := & c^{(l)}_{(0,\ldots,0,-1)} \ominus \tilde{f}((c^{(r)}_{(\alpha_1,\ldots,\alpha_d,-1)})_{(\alpha_1,\ldots,\alpha_d) \in A}) \end{array}$$

for $c \in T^C$. Here $\ominus$ denotes subtraction modulo $q+1$ on $\{0, 1, \ldots, q\}$. It is straightforward to check that $H \circ G(c) = c$ for all $c \in T^{\mathbb{Z}^{d+1}}$. Hence, G is injective, and thus, by Richardson's [12] result, the cellular automaton $(T, d+1, G)$ is reversible and the cellular automaton $(T, d+1, H)$ is its inverse.

Finally we have to show that the CA (S, d, F) can be embedded strongly and without delay into $(T, d+1, G)$. Therefore we map each cell in $S^{\mathbb{Z}^d}$ to the left component of the corresponding cell with coordinate in the hyperplane $\mathbb{Z}^d \times \{0\}$, i.e. $\mu : S^{\mathbb{Z}^d} \to T^{\mathbb{Z}^{d+1}}$ is defined by

$$\mu(c)_{(\alpha_1,\ldots,\alpha_d,\alpha_{d+1})} := \begin{cases} (0,0) & \text{if } \alpha_{d+1} \neq 0 \\ (c_{(\alpha_1,\ldots,\alpha_d)}, 0) & \text{if } \alpha_{d+1} = 0\,. \end{cases}$$

The inverse mapping $\nu : T^{\mathbb{Z}^{d+1}} \to S^{\mathbb{Z}^d}$ is defined by

$$\nu(c)_{(\alpha_1,\ldots,\alpha_d)} := \begin{cases} c^{(l)}_{(\alpha_1,\ldots,\alpha_d,0)} & \text{if } c^{(l)}_{(\alpha_1,\ldots,\alpha_d,0)} \in S \\ 1 & \text{otherwise.} \end{cases}$$

The value 1 in the second case is irrelevant. It is only assigned in order to make ν a totally defined function. It is straightforward to check that $(\mu, \nu, 1)$ is a strong embedding of (S, d, F) into $(T, d+1, G)$. Therefore note especially that

$$g(c)^{(l)} = f((c^{(l)}_{(\alpha_1,\ldots,\alpha_d,0)})_{(\alpha_1,\ldots,\alpha_d)\in A})$$

for $c \in T^B$ if $c^{(r)}_{(0,\ldots,0,1)} = 0$ and if $(c^{(l)}_{(\alpha_1,\ldots,\alpha_d,0)})_{(\alpha_1,\ldots,\alpha_d)\in A}$ lies in S^A. This ends the proof. □

It should be noted that the embedding mappings μ and ν are very simple mappings: they are given by projections mapping one cell to another cell.

The result poses the question whether in order to obtain an embedding of a CA into a reversible CA one necessarily has to increase the dimension. After stating and proving his result Toffoli [14, page 277] asked "whether an arbitrarily cellular automaton can be embedded in a reversible one having the same number of dimensions" and conjectured that this is not possible.

In his definition of an embedding he leaves open which conditions an embedding should satisfy besides being (in our language) a set-theoretic embedding without delay. It turns out that in fact every cellular automaton can be embedded in the purely set-theoretical sense even into a very simple reversible one-dimensional cellular automaton.

Theorem 5 *Every cellular automaton of any dimension can be embedded set-theoretically without delay into the reversible one-dimensional cellular automaton* $(\{0,1\}, 1, \sigma^{(1),1})$.

Proof. A configuration $c \in \{0,1\}^{\mathbb{Z}}$ is called *periodic* if there is a positive integer n with $(\sigma^{(1),1})^n(c) = c$. The set of periodic configurations in $\{0,1\}^{\mathbb{Z}}$ is countable. And the forward and backward orbit $\{(\sigma^{(1),1})^n(c) \mid n \in \mathbb{Z}\}$ of an arbitrary configuration $c \in \{0,1\}^{\mathbb{Z}}$ is countable. Hence, the set of orbits of non-periodic configurations has the cardinality of the continuum.

Let (S, d, F) be an arbitrary cellular automaton. We have just seen that the set $S^{\mathbb{Z}^d}$ and the set of orbits of non-periodic configurations in $\{0,1\}^{\mathbb{Z}}$ have the same cardinality. By the axiom of choice we can define a function $\mu : S^{\mathbb{Z}^d} \to \{0,1\}^{\mathbb{Z}}$ which maps each configuration in $S^{\mathbb{Z}^d}$ to a non-periodic configuration in $\{0,1\}^{\mathbb{Z}}$ such that for two different $c, c' \in S^{\mathbb{Z}^d}$ the orbits $\{(\sigma^{(1),1})^n(\mu(c)) \mid n \in \mathbb{Z}\}$ and $\{(\sigma^{(1),1})^n(\mu(c')) \mid n \in \mathbb{Z}\}$ are disjoint. We can clearly define a mapping $\nu : \{0,1\}^{\mathbb{Z}} \to S^{\mathbb{Z}^d}$ such that (3) with $G = \sigma^{(1),1}$ is satisfied for $k = 1$. Thus, the triple $(\mu, \nu, 1)$ is a set-theoretic embedding of (S, d, F) into $(\{0,1\}, 1, \sigma^{(1),1})$. □

But such an ineffective mapping is certainly not what one would understand under an embedding of cellular automata. An embedding of cellular automata should reflect the lattice structure of the configuration spaces and the fact that cellular automata behave uniformly at every cell. That is, an embedding should at least in a weak sense commute with the shift mappings. This is modeled by our weak embeddings. A further natural condition would be that the embedding mapping μ is continuous and hence, that the value in the cell of an image $\mu(c)$ depends only on finitely many cells of the configuration c. This is formulated in our strong embeddings. In the next section we shall see that already the weaker condition prohibits the embedding of a cellular automaton which is not surjective into a reversible cellular automaton of the same dimension. This shows that Toffoli's conjecture is true.

4 Negative Results on Embeddings into Reversible Cellular Automata

After the two positive results of the last section — strong embedding of a CA into a reversible CA with dimension increased by 1 (Toffoli), and ineffective, purely set-theoretical embedding into a reversible CA of dimension 1 — we shall see in this section that already the weak morphism condition is an obstacle to embedding an arbitrary CA into a reversible CA of the same dimension or smaller dimension.

We start with an observation about injective weak morphisms.

Lemma 6 *Let S, T be finite sets containing at least two elements, and let d, e be positive integers. Let $\mu : S^{\mathbb{Z}^d} \to T^{\mathbb{Z}^e}$ be a weak morphism and for each $i \in \{1, \ldots, d\}$ let $a^{(i)} \in \mathbb{Z}^e$ be an integer vector such that $\mu \circ \sigma^{(d),i} = \sigma^{(e)}_{a^{(i)}} \circ \mu$. If μ is injective, then the d vectors $a^{(i)}$ are linearly independent.*

Proof. Let us assume that the d vectors $a^{(i)}$ for $i \in \{1, \ldots, d\}$ are not linearly independent. Then there is a non-trivial rational linear combination which is equal to the zero vector in $\mathbb{Z}^e$. By multiplying with the least common multiple of the denominators of the coefficients we can assume that we have a non-zero integer vector $b = (\beta_1, \ldots, \beta_d) \in \mathbb{Z}^d \setminus \{(0, \ldots, 0)\}$ with

$$\sum_{i=1}^{d} \beta_i \cdot a^{(i)} = (0, \ldots, 0)\,.$$

Since μ is a weak morphism we obtain by (1) and (2) for an arbitrary configuration $c \in S^{\mathbb{Z}^d}$:

$$\mu \circ \sigma^{(d)}_b(c) = \sigma^{(e)}_{\sum \beta_i \cdot a^{(i)}} \circ \mu(c) = \sigma^{(e)}_{(0,\ldots,0)} \circ \mu(c) = \mu(c)\,.$$

Because of $|S| \geq 2$ there are configurations $c \in S^{\mathbb{Z}^d}$ with $\sigma^{(d)}_b(c) \neq c$. We conclude that μ is not injective. This proves the assertion. □

The following result is no surprise. But it is interesting that one does not need to impose a continuity condition on embeddings in order to arrive at the conclusion that no CA can be embedded into a CA of smaller dimension.

Theorem 7 *A cellular automaton cannot be weakly embedded into a cellular automaton of smaller dimension.*

Proof. Assume that a CA (S, d, F) can be weakly embedded via (μ, ν, k) into a CA (T, e, G). The weak morphism μ is injective. By Lemma 6 there are d linearly independent vectors in $\mathbb{Z}^e$. Hence, $d \leq e$. That is the assertion. □

Our main result is based on and formulated with spatially periodic configurations.

Definition 8 Let d be a positive integer and S a finite set with at least two elements. A configuration $c \in S^{\mathbb{Z}^d}$ is called *spatially periodic* iff there exists a positive integer n such that $(\sigma^{(d),i})^n(c) = c$ for each $i \in \{1, \ldots, d\}$.

Let $\mathrm{Spatial}(S, d)$ denote the set of all spatially periodic configurations in $S^{\mathbb{Z}^d}$. We say that a cellular automaton (S, d, F) is *injective on all spatially periodic configurations* if the restriction $F|_{\mathrm{Spatial}(S,d)}$ of the global map to the set of spatially periodic configurations is injective. Cellular automata with this or similar properties and with a quiescent configuration have been analyzed by Sato and Honda [13]. The following theorem is the technical main result of the paper. We will use it later in order to derive the first two results stated in the introduction.

Theorem 9 *A cellular automaton which is not injective on all spatially periodic configurations cannot be weakly embedded into a cellular automaton of the same dimension which is injective on all spatially periodic configurations.*

The proof is based on a series of lemmata. We start with an alternative description of spatially periodic configurations.

Lemma 10 *Let d be a positive integer and S a finite set with at least two elements. A configuration $c \in S^{\mathbb{Z}^d}$ is spatially periodic iff there exist d linearly independent vectors $a^{(i)} \in \mathbb{Z}^d$ such that $\sigma^{(d)}_{a^{(i)}}(c) = c$ for each $i \in \{1, \ldots, d\}$.*

Proof. The only if part is clear by taking the vectors $a^{(i)} := ne_i^{(d)}$ where n is the number n of Definition 8. We prove the if part. Let us assume that there exist d linearly independent vectors $a^{(i)} \in \mathbb{Z}^d$ such that $\sigma^{(d)}_{a^{(i)}}(c) = c$ for each $i \in \{1, \ldots, d\}$. We fix a coordinate $j \in \{1, \ldots, d\}$. Since the vectors $a^{(i)}$ for $i = 1, \ldots, d$ are linearly independent we can write the unit vector $e_j^{(d)}$ as a rational linear combination of the $a^{(i)}$. By multiplying with the least common multiple of the denominators of the coefficients we obtain a positive integer n_j and an integral vector $b^{(j)} = (\beta_1^{(j)}, \ldots, \beta_d^{(j)}) \in \mathbb{Z}^d$ with

$$\sum_{i=1}^{d} \beta_i^{(j)} \cdot a^{(i)} = n_j \cdot e_j^{(d)} = (0, \ldots, 0, n_j, 0, \ldots, 0)$$

where the number n_j stands at position j. Let n be the smallest common multiple of the numbers n_j for $j = 1, \ldots, d$. By using (1) and (2) we obtain for each $j \in \{1, \ldots, d\}$:

$$(\sigma^{(d),j})^n(c) = \sigma^{(d)}_{ne_j^{(d)}}(c) = \sigma^{(d)}_{(n/n_j)\sum \beta_i^{(j)} a^{(i)}}(c) = c$$

where for the last equality we used $\sigma^{(d)}_{a^{(i)}}(c) = c$ for each $i \in \{1, \ldots, d\}$. This shows that c is spatially periodic according to Definition 8. □

Lemma 11 *Every injective, weak morphism between full shift spaces of the same dimension maps spatially periodic configurations to spatially periodic configurations.*

Proof. Let S, T be finite sets containing at least two elements, d be a positive integers, and $\mu : S^{\mathbb{Z}^d} \to T^{\mathbb{Z}^d}$ be an injective weak morphism. Let $c \in S^{\mathbb{Z}^d}$ be a spatially periodic configuration. We have to show that also $\mu(c)$ is a spatially periodic configuration.

For each $i \in \{1, \ldots, d\}$ let $a^{(i)} \in \mathbb{Z}^d$ be a vector such that $\mu \circ \sigma^{(d),i} = \sigma^{(e)}_{a^{(i)}} \circ \mu$. By Lemma 6 the d vectors $a^{(i)}$ for $i \in \{1, \ldots, d\}$ are linearly independent. Let n be a positive integer with $(\sigma^{(d),i})^n(c) = c$ for all $i \in \{1, \ldots, d\}$. We obtain

$$\sigma^{(d)}_{na^{(i)}} \circ \mu(c) = \mu \circ \sigma^{(d)}_{ne_i^{(d)}}(c) = \mu \circ (\sigma^{(d),i})^n(c) = \mu(c)$$

for $i = 1, \ldots, d$. Since the vectors $a^{(i)}$ are linearly independent, also the vectors $na^{(i)}$ are linearly independent. The assertion follows now from Lemma 10. □

Let (S, d, F) be a cellular automaton. A configuration $c \in S^{\mathbb{Z}^d}$ is called *F-periodic* if there is a positive integer such that $F^n(c) = c$. The following lemma was shown by Sato and Honda [13] for cellular automata with quiescent configuration.

Lemma 12 *A cellular automaton (S, d, F) is injective on all spatially periodic configurations if and only if every spatially periodic configuration is F-periodic.*

Proof. For each positive integer m let

$$\mathrm{Spatial}(S,d,m) := \{c \in S^{\mathbb{Z}^d} \mid (\sigma^{(d),i})^m(c) = c \text{ for each } i \in \{1,\dots,d\}\}.$$

Then $\mathrm{Spatial}(S,d) = \bigcup_{m=1}^{\infty} \mathrm{Spatial}(S,d,m)$. Each set $\mathrm{Spatial}(S,d,m)$ is finite. The global map F commutes with the shift mappings $\sigma^{(d),i}$ for $i \in \{1,\dots,d\}$. Hence, it maps the set $\mathrm{Spatial}(S,d,m)$ into itself.

Assume that the CA (S,d,F) is injective on all spatially periodic configurations. Then, for any $m \geq 1$, the mapping F maps the finite set $\mathrm{Spatial}(S,d,m)$ injectively into itself and therefore bijectively onto itself. We conclude that every element of $\mathrm{Spatial}(S,d,m)$ is F-periodic (more generally: if M is a finite set and $H : M \to M$ a bijection, then every element of M is H-periodic). Hence, every spatially periodic configuration is F-periodic.

For the inverse implication we assume that every spatially periodic configuration is F-periodic. If c is in $\mathrm{Spatial}(S,d,m)$ for some $m \geq 1$, and $k \geq 1$ is a number with $F^k(c) = c$, then $c' := F^{k-1}(c)$ is an element of $\mathrm{Spatial}(S,d,m)$ with $F(c') = c$. Therefore, F maps the set $\mathrm{Spatial}(S,d,m)$ surjectively onto itself. Since this set is finite, F maps it bijectively onto itself. Finally, if c and $\tilde{c}$ are two spatially periodic configurations, then there is an $m \geq 1$ such that $\mathrm{Spatial}(S,d,m)$ contains c as well as $\tilde{c}$. We conclude that $F(c)$ and $F(\tilde{c})$ are different if c and $\tilde{c}$ are different. Thus, the CA F is injective on all spatially periodic configurations. $\square$

We remark that for an arbitrary cellular automaton (S,d,F) every spatially periodic configuration is ultimately F-periodic, i.e. for any $c \in S^{\mathbb{Z}^d}$ there is a $k \geq 0$ such that $F^k(c)$ is F-periodic.

Proof of Theorem 9. Let (S,d,F) be a cellular automaton and let (T,d,G) be a cellular automaton of the same dimension which is injective on all spatially periodic configurations. We assume that there is a weak embedding (μ,ν,k) of (S,d,F) into (T,d,G) and show that this assumption implies that also (S,d,F) must be injective on all spatially periodic configurations.

In view of the if part of Lemma 12 it is sufficient to show the following claim:

$$\text{if } c \in S^{\mathbb{Z}^d} \text{ is spatially periodic, then it is } F\text{-periodic.} \tag{4}$$

Let $c \in S^{\mathbb{Z}^d}$ be spatially periodic. By Lemma 11 also $\mu(c) \in T^{\mathbb{Z}^d}$ is spatially periodic. By the only if part of Lemma 12 $\mu(c)$ is G-periodic, i.e. there is a positive integer l with $G^l(\mu(c)) = \mu(c)$. This implies $G^{kl}(\mu(c)) = \mu(c)$, of course. Using the fact that (μ,ν,k) is a weak embedding we obtain

$$F^l(c) = \nu G^{kl} \mu(c) = \nu\mu(c) = c\,.$$

Hence, c is F-periodic. We have proved (4) and thus also Theorem 9. $\square$

By using well-known facts we can deduce from Theorem 9 the first two results stated in the introduction.

Theorem 13 *A cellular automaton which possesses a Garden of Eden configuration cannot be weakly embedded into a reversible cellular automaton of the same dimension.*

First, we formulate separately a simple observation, compare Sato and Honda [13]. For completeness sake we give the proof.

Lemma 14 *If a cellular automaton is injective on all spatially periodic configurations, then it is surjective.*

Proof. Let (S,d,F) be a cellular automaton which is injective on all spatially periodic configurations. Let $c \in S^{\mathbb{Z}^d}$ be an arbitrary configuration. We have to show that there is a configuration $\tilde{c} \in S^{\mathbb{Z}^d}$ with $F(\tilde{c}) = c$.

For each $k \geq 0$ there exists a spatially periodic configuration $c^{(k)}$ with $c_a^{(k)} = c_a$ for all $a \in \{-k, \ldots, k\}^d$. Since by Lemma 12 $c^{(k)}$ is F-periodic, there is a configuration $\tilde{c}^{(k)}$ with $F(\tilde{c}^{(k)}) = c^{(k)}$. The sequence $(\tilde{c}^{(k)})_{k \geq 0}$ has an accumulation point $\tilde{c}$ in the compact space $S^{\mathbb{Z}^d}$. By continuity of F we conclude $F(\tilde{c}) = c$. □

Proof of Theorem 13. The assertion follows from Lemma 14, the fact that every reversible cellular automaton is especially injective on all spatially periodic configurations, and from Theorem 9. □

Using a result stated by Culik, Hurd, and Yu [4] we obtain the following version for dimension one.

Theorem 15 *An irreversible cellular automaton of dimension one cannot be weakly embedded into a reversible cellular automaton of dimension one.*

Proof. It is known that a cellular automaton of dimension one is reversible if and only if it is injective on all spatially periodic configurations, see Culik, Hurd, and Yu [4, Theorem 38]. The assertion follows from this fact and from Theorem 9. □

In the last proof we used the fact that for dimension one the class of reversible cellular automata coincides with the class of cellular automata which are injective on all spatially periodic configurations. For dimension greater than one this is not true, due to a result by Kari [6]. But the difference between the two classes of cellular automata seems to be subtle. It is interesting that according to Theorem 9 the answer to the embedding question is for dimension greater than one given by a line drawn not between reversible and irreversible CA's but drawn between CA's which are injective on all spatially periodic configurations and between CA's which are not.

5 Conclusion

We conclude with some remarks on reversible computation. The interest in reversible computation stems from the observation by Landauer [7] that not necessarily all elementary operations in a physical computing device must produce thermal entropy but only the irreversible ones. Hence, reversible computing devices are likely to dissipate less energy than irreversible ones, compare Bennett [1, 2]. In this respect, reversible cellular automata appear to be good candidates for a computation model which takes physical constraints into account. There are reversible cellular automata which are capable of universal computation, see Toffoli [14], for example the billiard ball model cellular automaton of Margolus [8]. Our results show a limitation of such computing devices: for example a three-dimensional reversible CA cannot simulate a non-surjective three-dimensional CA if under simulation one understands that the second CA should at least be weakly embedded into the first one according to Definition 3. This can informally be described by saying that a three-dimensional reversible CA cannot in real time and cell for cell simulate a three-dimensional non-surjective CA or any process which is to be modeled by such a CA. It remains to be seen whether this limitation of the computing power of reversible cellular automata is of practical importance.

References

[1] C. H . Bennett. Logical reversibility of computation. *IBM Journal of Research and Development*, 17(6):525–532, 1973.

[2] C. H . Bennett. The thermodynamics of computation — a review. *International Journal of Theoretical Physics*, 21:905–940, 1982.

[3] A. W. Burks. *Essays on Cellular Automata.* University of Illinois Press, Urbana, 1970.

[4] K. Culik II, L. P. Hurd, and S. Yu. Computation theoretic aspects of cellular automata. *Physica D*, 45:357–378, 1990.

[5] G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Mathematical Systems Theory*, 3:320–375, 1969.

[6] J. Kari. Reversibility and surjectivity problems of cellular automata. *Journal of Computer and System Sciences*, 48:149–182, 1994.

[7] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5:183–191, 1961.

[8] N. Margolus. Physics-like models of computation. *Physica*, D 10:81–95, 1984.

[9] A. Maruoka and M. Kimura. Injectivity and surjectivity of parallel maps for cellular automata. *Journal of Computer and System Sciences*, 18:47–64, 1979.

[10] E. F. Moore. Machine models of self reproduction. *American Mathematical Society, Proceedings of Symposia in Applied Mathematics*, 14:17–33, 1962. Reprinted in [3], 187–203.

[11] J. Myhill. The converse to Moore's Garden-of-Eden theorem. *Proc. of the AMS*, pages 685–686, 1963. Revised version in [3], 204–205.

[12] D. Richardson. Tesselations with local transformations. *Journal of Computer and System Sciences*, 6:373–388, 1972.

[13] T. Sato and N. Honda. Certain relations between properties of maps of tesselation automata. *Journal of Computer and System Sciences*, 15:121–145, 1977.

[14] T. Toffoli. Computation and construction universality of reversible cellular automata. *Journal of Computer and System Sciences*, 15:213–231, 1977.

[15] T. Toffoli and N. Margolus. Invertible cellular automata: a review. *Physica*, D 45:229–253, 1990.

[16] J. von Neumann. *Theory of Self-Reproducing Automata.* University of Illinois Press, Champaign, Illinois, 1966. Edited and completed by A. Burks.