Noname manuscript No.
(will be inserted by the editor)

Communication complexity meets cellular automata: Necessary conditions for intrinsic universality

Raimundo Briceño · Ivan Rapaport

Received: date / Accepted: date

Abstract A natural way to interpret a cellular automaton (CA) is as a mechanism that computes, in a distributed way, some function f. In other words, from a computer science point of view, CAs can be seen as distributed systems where the cells of the CAs are nodes of a network linked by communication channels. A classic measure of efficiency in such distributed systems is the number of bits exchanged during the computation process. A typical approach is to look for bottlenecks: channels through which the nature of the function f forces the exchange of a significant number of bits. In practice, a widely used way to understand such congestion phenomena is to partition the system into two subsystems and try to find bounds for the number of bits that *must* pass through the channels that join them. Finding these bounds is the focus of communication complexity theory. Measuring the communication complexity of some problems induced by a CA ϕ turned out to be tremendously useful to give clues regarding the intrinsic universality of ϕ (a CA is said to be intrinsically universal if it is capable of emulating any other). In fact, there exist particular problems P's for which the following key property holds: if ϕ is intrinsically universal, then the communication complexity of P(ϕ) must be maximal. In this tutorial, we intend to explain the connections that were found, through a series of papers, between intrinsic universality and communication complexity in CAs.

Keywords cellular automata · communication complexity · intrinsic universality

R. Briceño
Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Chile
E-mail: raimundo.briceno@mat.uc.cl

I. Rapaport
Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático (UMI 2807 CNRS), Universidad de Chile, Chile
E-mail: rapaport@dim.uchile.cl

1 Introduction

Cellular automata (CAs) are discrete dynamical systems. They were introduced by von Neumann as a formal model of self-reproducing organisms [58]. The simplest description of a CA is as an array of cells whose content evolves in discrete time steps. At each time step each cell is in one of a finite set of possible states. Every cell changes its own state at each clock tick following a local rule which determines its new state as a function of its present state together with the states of its neighbors. For a nice survey on CAs we refer to [34].

Formally, a (one-dimensional) CA is defined by its *local transition rule* $\phi : A^{2r+1} \to A$, where A corresponds to the set of states and r denotes the radius of the local rule. A configuration $c \in A^{\mathbb{Z}}$ is an assignment of states to the lattice $\mathbb{Z}$. To any CA ϕ we associate a *global transition function* $\Phi : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$, which corresponds to the synchronous and uniform application of the local transition rule ϕ. More precisely, for every $i \in \mathbb{Z}$, the updated state at position i is given by $(\Phi(c))_i = \phi(c_{i-r}, \ldots, c_{i+r})$. The t-step iteration of the global function is denoted by $\Phi^t : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$.

This formalism turned out to be tremendously useful for describing all kinds of phenomena, especially of physical [10,17,59,60] and biological [24,27, 51] nature. CAs also proved to be extremely versatile and, indeed, they have been studied from purely theoretical perspectives [33,36,56] and, at the same time, they have been used in very practical applications [5,6,14,39].

From the point of view of computer science, CAs can also be seen as message-passing distributed systems [50]. For that reason, it comes as no surprise the similarity between some message-passing distributed models and some variants of CAs. This convergence between models occurred over time. On the one hand, some researchers of the CA community relaxed certain original restrictions of the model allowing asynchronous and/or random dynamics [25,26,40,52] and less rigid topologies [2,35,44,57]. On the other hand, with the proliferation of huge networks of very weak devices, the distributed computing community became interested in models where processors, instead of being considered as powerful Turing machines, were assumed to be finite automata [8,13,16,23].

Our approach is precisely focused at this point where the two communities —that of distributed computing and that of CAs— meet.

The main idea is to view the cells of the CAs as nodes of an interconnection network linked by communication channels. Our first step consists in applying ideas of communication complexity to CAs and to interpret the dynamic processes as communication protocols [22,28]. After that first phase, a very useful and interesting relationship appears between communication complexity of CAs and *intrinsic universality*. Before explaining this relationship we will refer to the general notion of universality.

Universality has always been present in the theory of computation. It is related to the ability of certain systems to emulate others. This idea has had notable repercussions and gave rise to deep concepts and theories: computability, universal Turing machines, reductions, NP-completeness, among others [3].

Informally, a Turing machine is said to be universal if it can emulate any other Turing machine. A system (for instance a CA) is Turing-universal if it has the same computing power of a universal Turing machine. By interpreting the configurations of one-dimensional CAs as bi-infinite tapes, the emulation of Turing machines, and therefore the existence of Turing-universal CAs, is straightforward. In this framework, Smith constructed in 1971 a Turing-universal CA with 18 states and radius 1 [54]. Later, in 1990, Lindgren and Nordahl constructed a 7-state, radius 1 Turing-universal CA [38]. Finally, published somewhat late (2004), Cook was able to show that Rule 110, a 2-state CA of radius 1, was Turing-universal [15].

It is important to point out that there is actually no consensus on the formal notion of Turing-universality in CAs (see [21] for a discussion about encoding/decoding issues). On the other hand, by observing emulations where the cells of the emulated CA are identified with blocks of cells of the emulating CA, little by little the idea of intrinsic universality was forged. Roughly, a CA is intrinsically universal if it is capable of emulating any other. This notion was first outlined by Banks [4], to be later taken up by Albert and Čulik II [1]. At the end of the 90s, the different types of universalities were clearly and formally distinguished [21,41]. Finally, it is good to note at this point that, although intrinsic universality implies Turing-universality —given any reasonable definition of it— the converse can be shown to be false. For a detailed tour, it is recommended to read Ollinger's survey [48]. As with Turing-universality, small intrinsically universal CAs have been constructed [49].

The notion of Turing-universality is extrinsic to the CA model. On the other hand, emulations between CAs are intrinsic. An advantage of the definition of intrinsic universality is that it accepts well-defined characterizations, without great arbitrariness. These well-defined characterizations allow us to address the question of whether a certain CA is intrinsically universal (in contrast, for the case of Turing-universality, we do not even know what the input/output relationship is like, and therefore, finding negative results is particularly cumbersome).

Now we can precise the relationship between intrinsic universality in CAs and communication complexity [11,29]. This relationship will be developed later, but we can summarize it as follows. Given a CA ϕ, we define a computational problem $\mathrm{P}(\phi)$ parametrized by ϕ. We split the input into two parts: one given to a party called Alice and the other given to a second party called Bob. Then, we view such problem as a communication problem and we prove that, the existence of a CA ψ for which the communication complexity of $\mathrm{P}(\psi)$ is higher than the one of $\mathrm{P}(\phi)$, corresponds to a *certificate* of the fact that ϕ is not intrinsically universal.

The intuition behind the previous result is the following. If a CA ψ can be emulated through another CA ϕ, then, for "any" problem P parametrized by CAs, the "complexity" of $\mathrm{P}(\phi)$ cannot be "lower" than the complexity of

P(ψ). Therefore, if the complexity is indeed lower, then it means that ϕ cannot emulate ψ, and, in particular, ϕ cannot be intrinsically universal.

The organization of the paper is the following. In Section 2, we introduce basic communication complexity ideas. In particular, we define the notion of overlapping, that we apply to the study of CAs in the last section, but which may be of interest by its own. In Section 3, we give the formal definition of emulation between CAs and the corresponding notion of intrinsic universality. In Section 4, we introduce a natural communication problem induced by CAs, known as the *prediction problem*. Through this particular problem, we illustrate the connection between communication complexity and intrinsic universality. More precisely, since the communication complexity of the prediction problem increases monotonically with respect to the order induced by emulations, we can rule out CAs from being intrinsically universal (those for which the prediction problem has low communication complexity). In Section 5, we use the same approach but for other problems: length of cycle, spatial-invasion, temporal-invasion, and controlled-invasion. Finally, in Section 6, through the overlapping relation, we explain how to use simultaneously many problems in order to rule out CAs from being intrinsically universal.

2 Communication complexity

The basic, deterministic two-party communication complexity model was introduced by Yao [63]. For an extensive exposition of the model see [37]. The model is the following. First, there is some function $f : X \times Y \to Z$, where X, Y, Z are finite sets. Typically, $X = Y = \{0,1\}^n$ and $Z = \{0,1\}$. Two parties, Alice and Bob, must cooperate exchanging messages in order to compute $f(x, y)$. The key point is that Alice receives the input $x \in X$ (unknown to Bob) while Bob receives the input $y \in Y$ (unknown to Alice).

We usually refer to a function $f : X \times Y \to Z$ as a *two-party communication problem*. This problem can be solved by a deterministic protocol, which specifies, at each step of the communication between Alice and Bob, whose turn it is to speak and what she/he says (a bit, 0 or 1) as a function of her/his respective input, together with the messages received so far. A protocol specifies when the communication has ended and, in each end state, the corresponding output.

	y_1	y_2	y_3	y_4
x_1	z_1	z_1	z_1	z_2
x_2	z_1	z_1	z_1	z_2
x_3	z_1	z_1	z_1	z_1
x_4	z_1	z_2	z_2	z_2

Table 1 A matrix M_f which represents a function f.

Note that we can represent a function $f : X \times Y \to Z$ as a matrix M_f, where the rows are labeled with the set X of possible inputs for Alice, the columns with the set Y of possible inputs for Bob, and, finally, the entry (x, y) of the matrix is $f(x, y) \in Z$. This matrix is fully known to both Alice and Bob when they agree on a protocol. As an example, see the matrix of Table 1, where $X = \{x_1, x_2, x_3, x_4\}$, $Y = \{y_1, y_2, y_3, y_4\}$, and $Z = \{z_1, z_2\}$.

Formally, a protocol $\mathcal{P}$ over a domain $X \times Y$ with range Z is a binary tree where each internal node v is labeled either with a map $a_v : X \to \{0, 1\}$ or with a map $b_v : Y \to \{0, 1\}$, and each leaf ℓ is labeled with an element $z \in Z$. For instance, in Figure 1, we can see a tree/protocol that solves the function of Table 1.

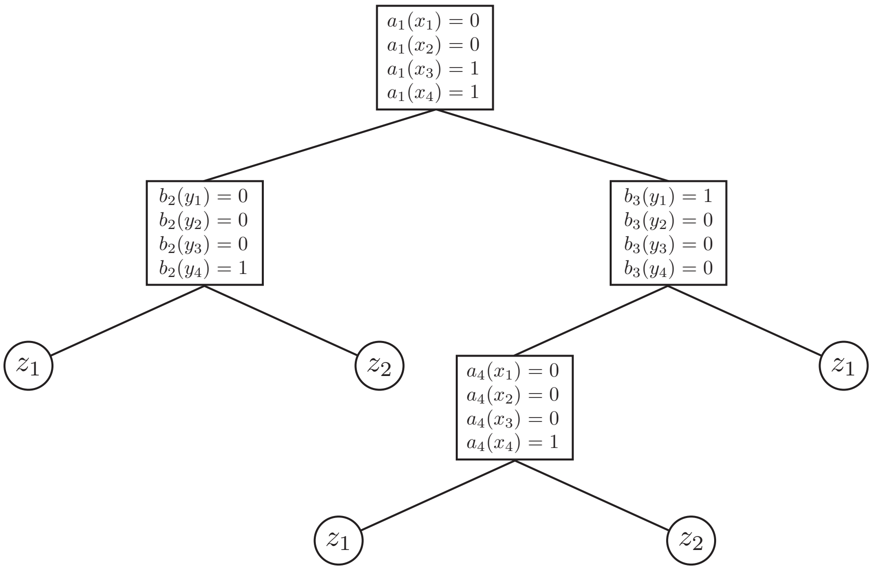

Fig. 1 A communication protocol as a binary tree.

The *value* of protocol $\mathcal{P}$ on input $(x, y) \in X \times Y$ is given by $A_\ell(x)$ (or $B_\ell(y)$), where A_ℓ (or B_ℓ) is the label of the leaf reached by walking on the tree from the root, turning left if $a_v(x) = 0$ (or $b_v(y) = 0$) and right otherwise. We say that a protocol computes a function $f : X \times Y \to Z$ if, for every $(x, y) \in X \times Y$, its value on input (x, y) is precisely $f(x, y)$.

Intuitively, each internal node specifies a bit to be communicated either by Alice or by Bob, whereas at the leaves both parties already know $f(x, y)$.

The cost of a protocol is the maximum number of bits it ever sends, ranging over all inputs, i.e., the depth of the associated tree. For example, the cost of the protocol shown in Figure 1 is 3. The *communication complexity* of a function f is the minimum cost over all protocols that compute it. We denote by $\mathbf{cc}(f)$ the (deterministic) communication complexity of a function $f : X \times Y \to Z$.

A very successful approach for proving lower bounds on the communication complexity of arbitrary functions is based on the so-called *monochromatic rectangles.*

Definition 1 Given a function $f : X \times Y \to Z$, a subset $R = A \times B \subseteq X \times Y$ is an *f-monochromatic rectangle* if f is constant on R.

Lemma 1 ([37]) *If any partition of $X \times Y$ into f-monochromatic rectangles requires at least t rectangles, then* $\mathbf{cc}(f) \geq \lceil \log_2(t) \rceil$.

Remark 1 In order to lighten the notation, in this work we are going to write $\log(t)$ instead of $\lceil \log_2(t) \rceil$.

Let $t(f)$ be the smallest number of f-monochromatic rectangles in a partition of $X \times Y$. There are many approaches for finding lower bounds for $t(f)$. One of these approaches is based on the notion of *fooling set*. Roughly, a fooling set is a (hopefully large) set of input pairs such that no two of them can belong to a single monochromatic rectangle.

Definition 2 Given a function $f : X \times Y \to Z$ and $z \in Z$, a set $\mathcal{F} \subseteq X \times Y$ is a *z-fooling set* for f if

1. for every $(x, y) \in \mathcal{F}, f(x, y) = z$;
2. for every distinct pairs (x_1, y_1) and (x_2, y_2) in $\mathcal{F}$, either $f(x_1, y_2) \neq z$ or $f(x_2, y_1) \neq z$.

The usefulness of fooling sets is given by the following lemma.

Lemma 2 ([37]) *Let $Z = \{z_1, \ldots, z_k\}$. If, for all $1 \leq i \leq k$, $\mathcal{F}_i$ is a z_i-fooling set for f of size t_i, then* $\mathbf{cc}(f) \geq \log(t_1 + \cdots + t_k)$.

Consider now the problem *Equality*, $\mathrm{EQ} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, where

$$\mathrm{EQ}(x, y) = \begin{cases} 1, & x = y, \\ 0, & x \neq y. \end{cases}$$

Note that M_{EQ}, the matrix associated to EQ, is the diagonal matrix of dimension 2^n (see Figure 2). The elements of the diagonal form a natural 1-fooling set that we call $\mathcal{F}_{\mathrm{EQ}}$. More precisely, for all $x, y \in \{0,1\}^n$, $(x, y) \in \mathcal{F}_{\mathrm{EQ}}$ if and only if $x = y$. From Lemma 2, and considering that there exists a 0-fooling set of size at least 1, we obtain that $\mathbf{cc}(\mathrm{EQ}) \geq \log(2^n + 1) > n$. On the other hand, there is a trivial protocol whose cost is $n + 1$: Alice sends the n bits of her input x and then Bob answers with a 1 if $x = y$ and with a 0 otherwise. Hence, $\mathbf{cc}(\mathrm{EQ}) = n + 1$.

A similar situation arises with another well-known communication problem known as *Disjointness*, $\mathrm{DISJ} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, where

$$\mathrm{DISJ}(x, y) = \begin{cases} 1, & \forall i \in \{1, \ldots, n\} : x_i = 0 \vee y_i = 0, \\ 0, & \exists i \in \{1, \ldots, n\} : x_i = y_i = 1. \end{cases}$$

If we interpret x and y as the indicator function of a set of n elements, then $\text{DISJ}(x, y) = 1 \iff x \cap y = \emptyset$. A 1-fooling set of size 2^n is $\mathcal{F}_{\text{DISJ}}$, defined as follows: $(x, y) \in \mathcal{F}_{\text{DISJ}} \iff y = \overline{x}$, where $\overline{x}$ is the complement of x. With a reasoning analogous to the one of problem EQ, we conclude that $\mathbf{cc}(\text{DISJ}) = n + 1$. (The fractal nature of matrix M_{DISJ} is due to the following two facts: On one hand, $\text{DISJ}(x \cup \{n+1\}, y \cup \{n+1\}) = 0$. Second, $\text{DISJ}(x, y \cup \{n + 1\}) = \text{DISJ}(x, y)$.)

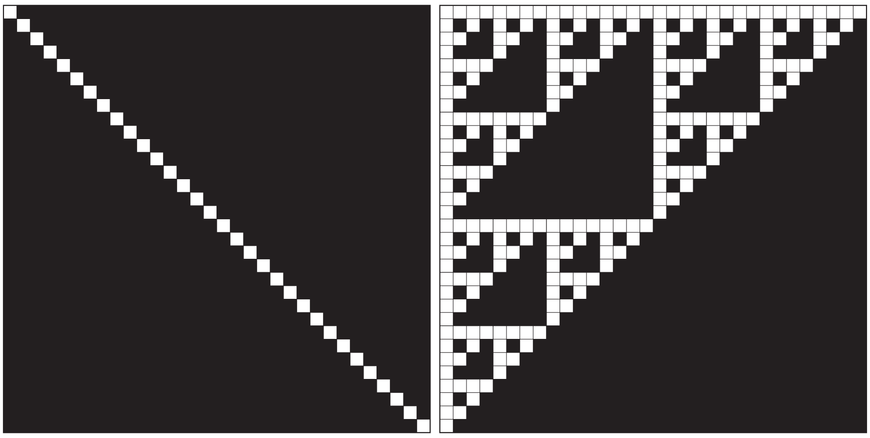

Fig. 2 A matrix representation of EQ and DISJ, where 1s are white and 0s are black).

The deterministic communication complexity notion can be either *restricted* to one-way protocols or *generalized* from functions to relations.

2.1 One-way communication complexity

The protocol we introduced for EQ was extremely simple: Alice sent a message to Bob, and then Bob decided the output and communicated it to Alice. These type of protocols are called AB-one-way protocols. The AB-one-way communication complexity of a function $f : X \times Y \to Z$, denoted by $\mathbf{cc}^{AB}(f)$, is the worst case number of bits Alice needs to send to Bob (in an AB-one-way protocol that computes f). The BA-one-way protocols and the BA-one-way communication complexity are defined analogously, by interchanging the role of Alice and Bob.

By forbidding back-and-forth communication between Alice and Bob, we simplify the analysis. In fact, consider the $|X| \times |Y|$ matrix M_f. Let $x, x' \in X$, and let m_x and $m_{x'}$ be the messages Alice would send to Bob upon receiving the inputs x and x', respectively. It follows that, if the x-row and the x'-row of M_f are different, then it must be that $m_x \neq m_{x'}$.

It is direct from the previous remark, that the AB-one-way communication complexity of f corresponds *exactly* to the logarithm of the number of different rows of M_f, while the BA-one-way communication complexity corresponds exactly to the logarithm of the number of different columns in the matrix M_f. In other words, the optimal AB-one-way protocol is the natural one: Alice sends to Bob the "name" of the row corresponding to her input x (and then Bob computes the output using his input y).

More precisely, $\mathbf{cc}^{AB}(f) = \log(d(M_f))$, where $d(M_f)$ is the number of different rows of the matrix M_f. Obviously, $\mathbf{cc}(f) \leq \mathbf{cc}^{AB}(f) + \log(|Z|)$, and the gap might be very large. In fact, there are problems like *Clique-independent set* [37], denoted CIS, for which $\mathbf{cc}^{AB}(\text{CIS}) = \Theta(n)$ while $\mathbf{cc}(\text{CIS}) = O(\log^2 n)$.

2.2 Relations

A relation $\mathcal{R}$ is a subset $\mathcal{R} \subseteq X \times Y \times Z$. The associated communication problem is the following: Alice receives $x \in X$, Bob receives $y \in Y$, and then they have to find a $z \in Z$ such that $(x, y, z) \in \mathcal{R}$. Note that, for a given relation, there may be more than one z satisfying the above property: Alice and Bob only need to give one such z. However, it might be the case that, for a given $(x, y) \in X \times Y$, there is no $z \in Z$ such that $(x, y, z) \in \mathcal{R}$. In that situation, we say that the input (x, y) is *illegal*. Otherwise, we say that the input is *legal*.

A protocol $\mathcal{P}$ computes a relation $\mathcal{R}$ if, for every legal input $(x, y) \in X \times Y$, the protocol reaches a leaf marked by a value z such that $(x, y, z) \in \mathcal{R}$. We denote by $\mathbf{cc}(\mathcal{R})$ the (deterministic) communication complexity of a relation $\mathcal{R} \subseteq X \times Y \times Z$, which corresponds to the minimal depth of a protocol tree computing $\mathcal{R}$.

Consider the following example related to the function EQ. Let $X, Y \subseteq \{0,1\}^n$ such that $X \cap Y = \emptyset$. Assume that the task of Alice and Bob consists in finding, for any input $(x, y) \in X \times Y$, an index $i \in \{1, \ldots, n\}$ such that $x_i \neq y_i$. Finding non-trivial lower bounds for the communication complexity of this relation turns out to be a major open problem in computational complexity [32,53].

One can also define monochromatic rectangles for relations.

Definition 3 Given a relation $\mathcal{R} \subseteq X \times Y \times Z$, a subset $A \times B \subseteq X \times Y$ is an *$\mathcal{R}$-monochromatic rectangle* if there exists a value $z \in Z$ such that, for every $(x, y) \in A \times B$, either $(x, y, z) \in \mathcal{R}$ or (x, y) is illegal.

Lemma 3 ([37]) *If any partition of $X \times Y$ into $\mathcal{R}$-monochromatic rectangles requieres at least t rectangles, then* $\mathbf{cc}(\mathcal{R}) \geq \log(t)$.

2.3 Overlapping

In Section 6 we are going to use, in the context of CAs, a relation called *overlapping*. This relation may be of interest by its own. The idea is that, given a fixed collection of functions with common domain $X \times Y$, let's say

$\{f_i : X \times Y \to Z_i\}_{i=1}^k$, once Alice and Bob receive their inputs, they may *choose online* the function for which they prefer to give the answer. The formal definition is the following.

Definition 4 Let $\{f_i : X \times Y \to Z_i\}_{i=1}^k$ be a finite family of functions with common domain $X \times Y$. We define the **overlapping** of such family as the relation

$$f_1 \uplus \cdots \uplus f_k \subseteq X \times Y \times (\bigcup_{i=1}^{k} Z_i \times [k]),$$

given by

$$(x, y, (z, i)) \in f_1 \uplus \cdots \uplus f_k \iff f_i(x, y) = z.$$

In other words, $f_1 \uplus \cdots \uplus f_k$ asks about some index i pointing towards a problem f_i together with the answer $z \in Z_i$ to such problem. The communication complexity of $f_1 \uplus \cdots \uplus f_k$ corresponds to the number of bits Alice and Bob need to exchange in order to find a correct answer. Obviously, and this is the key motivation of the overlapping definition, $\mathbf{cc}\,(f_1 \uplus \cdots \uplus f_k) \leq \min_{i=1,\ldots,k} \mathbf{cc}\,(f_i)$.

Let us give an example. Consider the functions EQ and DISJ with common domain $\{0,1\}^n \times \{0,1\}^n$. Recall that $\mathbf{cc}(\mathrm{EQ}), \mathbf{cc}(\mathrm{DISJ}) \in \Omega(n)$. Nevertheless, $\mathbf{cc}(\mathrm{EQ} \uplus \mathrm{DISJ}) \in \Theta(\log n)$. In fact, for the upper bound, consider the following protocol. If $x = 0 \ldots 0$, then Alice sends a 0 to Bob; otherwise, she sends a 1. If Bob receives a 0 or, if $y = 0 \ldots 0$, then he answers $\mathrm{DISJ}(x, y) = 1$; otherwise he sends the position i corresponding to the leftmost 1 in y. Finally, Alice compares x_i with y_i. If $x_i = y_i$, then she answers $\mathrm{DISJ}(x, y) = 0$; otherwise, she answers $\mathrm{EQ}(x, y) = 0$. Therefore, the complexity of the protocol is $O(\log n)$ because of the number of bits needed to encode the index i.

For the lower bound, we will consider the following extension of the definition of a fooling set to this context.

Definition 5 Let $\{f_i : X \times Y \to Z_i\}_{i=1}^k$ be a finite family of functions with common domain $X \times Y$. A subset $\mathcal{F} \subseteq X \times Y$ is called a *fooling set* for $f_1 \uplus \cdots \uplus f_k$ if, for all $1 \leq i \leq k$, there exists a value $z_i \in Z_i$ such that

- for every $(x, y) \in \mathcal{F}$, $f_i(x, y) = z_i$;
- for every two distinct pairs (x_1, y_1) and (x_2, y_2) in $\mathcal{F}$, either $f_i\,(x_1, y_2) \neq z_i$ or $f_i\,(x_2, y_1) \neq z_i$.

The relevance of this definition is that, as in the case of functions, it allows to obtain lower bounds.

Proposition 1 ([11]) *If $\{f_i : X \times Y \to Z_i\}_{i=1}^k$ has a fooling set $\mathcal{F}$ for $f_1 \uplus \cdots \uplus f_k$ of size t, then*

$$\mathrm{cc}\,(f_1 \uplus \cdots \uplus f_k) \geq \log(t).$$

Now, we can prove that the set

$$\mathcal{F} = \{(x, x) \in \{0,1\}^n \times \{0,1\}^n : \sum_{i=1}^{n} x_i = 1\}$$

is a fooling set for EQ ⊎ DISJ (see Figure 3). Let $(x, x), (x', x') \in \mathcal{F}$ such that $x \neq x'$. Then, $\text{DISJ}(x, x) = \text{DISJ}(x', x') = 0$ and $\text{EQ}(x, x) = \text{EQ}(x', x') = 1$. But $\text{DISJ}(x, x') = 1$ and $\text{EQ}(x, x') = 0$. Note that $|\mathcal{F}| = n$ and therefore $\mathbf{cc}(\text{EQ} \uplus \text{DISJ}) \in \Omega(\log n)$.

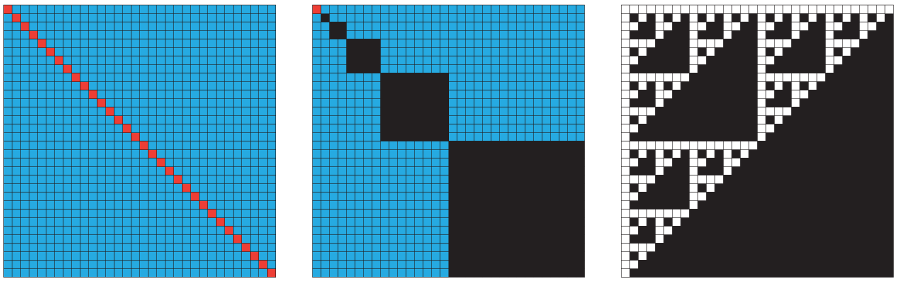

Fig. 3 On the left (in blue and red), the matrix representation of the EQ problem and, on the right (in white and black), the matrix representation of the DISJ problem. At the center, a representation of a partition for EQ ⊎ DISJ into monochromatic rectangles from EQ and DISJ. While EQ and DISJ require exponentially (in n) many monochromatic (blue-red and white-black, respectively) rectangles to be partitioned, EQ⊎DISJ only requires a polynomial number (in n) of monochromatic (blue-red-white-black) rectangles to be partitioned.

3 Intrinsic universality in CAs

The notion of intrinsic universality —which relies on a particular notion of emulation— is very natural: a CA is intrinsically universal if it is able to *emulate* any other [21, 41, 48]. Since intrinsic universality is a very precise concept, it may seem, because of this precision, restrictive. Nevertheless, the intrinsic universality property can be very common in some natural families of CAs. In [7, 55], the authors exhibit natural families where *almost all* the CAs are intrinsically universal.

By completely formalizing the notion of universality, we facilitate the proof of negative results. It is important to point out that, since the notion of intrinsic universality is related to a process by which we change the scale of space-time diagrams, we are answering pure dynamical questions by using computational tools.

The notion of emulation involved in the definition of intrinsic universality induces a partial order in the set of all CAs. Considering this, we can compare CAs according to their power of emulation, being the intrinsically universal CAs the maximum elements of such order, i.e., the most powerful ones.

Remark 2 We would like to point out that if the reader decides to skip the following formal definition of intrinsic universality, it is possible to follow the rest of the tutorial with no problem. This is analogous to the situation in which one studies reductions in computational complexity without the need of understanding the Turing machine model in full detail.

We say that a CA ϕ is a *sub-automaton* of another CA ψ, denoted by $\phi \sqsubseteq \psi$, if we can identify $\Phi : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ in $\Psi : B^{\mathbb{Z}} \to B^{\mathbb{Z}}$ after renaming the states, where Φ and Ψ denote the corresponding global transition functions. Formally, $\phi \sqsubseteq \psi$ if there is an injective map $\iota : A \to B$ such that $\bar{\iota} \circ \Phi = \Psi \circ \bar{\iota}$, where $\bar{\iota} : A^{\mathbb{Z}} \to B^{\mathbb{Z}}$ denotes the uniform extension of ι, i.e.,

$$\bar{\iota}(c) = (\iota(c_i))_{i\in\mathbb{Z}} \quad \text{for } c = (c_i)_{i\in\mathbb{Z}} \in A^{\mathbb{Z}}.$$

We say that a CA ψ *emulates* a CA ϕ if some rescaling of ϕ is a subautomaton of some rescaling of ψ. The ingredients of the rescaling are simple: packing cells into blocks, iterating the rule, and shifting.

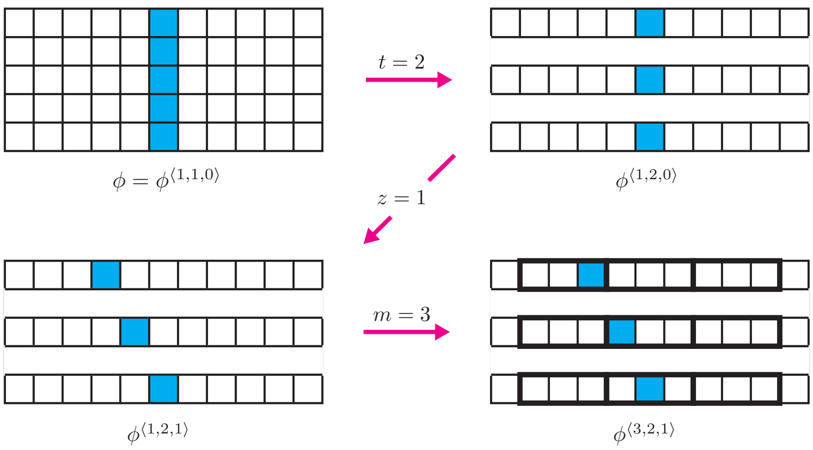

Fig. 4 The rescaling $\langle m, t, z\rangle$ of ϕ with parameters $m = 3$ (packing), $t = 2$ (iterating) and $z = 1$ (shifting).

Formally, given any state set A and any $m \geq 1$, we define the bijective *packing map* $\gamma_m : A^{\mathbb{Z}} \to (A^m)^{\mathbb{Z}}$ by

$$(\gamma_m(c))_i = (c_{mi}, \ldots, c_{mi+m-1}), \qquad \text{for } i \in \mathbb{Z} \text{ and } c = (c_i)_{i\in\mathbb{Z}} \in A^{\mathbb{Z}}.$$

The *shift map* is defined as $\sigma : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$, where $\sigma(c)_i = c_{i+1}$ for $c \in A^{\mathbb{Z}}$.

Then, the rescaling $\langle m, t, z\rangle$ of ϕ by parameters m (packing), $t \geq 1$ (iterating), and $q \in \mathbb{Z}$ (shifting) is the CA $\phi^{\langle m,t,q\rangle}$ (see Figure 4) with set of states A^m and the following global transition function:

$$\gamma_m \circ \sigma^q \circ \Phi^t \circ \gamma_m^{-1}.$$

The fact that the above global transition function is induced by a CA (i.e., a local transition rule) follows from the Curtis-Hedlund-Lyndon Theorem [31], which characterizes CAs as the continuous functions that commute with the shift.

Having this, we say that ψ *emulates* ϕ, denoted by $\phi \preccurlyeq \psi$, if there exist rescaling parameters $m_1, m_2, t_1, t_2 \in \mathbb{N}$ and $q_1, q_2 \in \mathbb{Z}$ such that

$$\phi^{\langle m_1,t_1,q_1\rangle} \sqsubseteq \psi^{\langle m_2,t_2,q_2\rangle}.$$

We say that a CA ψ is *intrinsically universal* if for all CA ϕ it holds that $\phi \preccurlyeq \psi$. It is possible to check that intrinsically universal CAs do exist and, moreover, they are generic in natural classes [7, 55].

4 The Prediction Problem

To simplify the exposition, let us consider here only CAs with radius $r = 1$. There is no much loss of generality if we do this, since, by augmenting the number of states in A, we can emulate an arbitrary CA, of any radius, with some CA of radius 1. In this case, the local transition rule is of the form $\phi : A^3 \to A$. A natural way to interpret a CA is as a mechanism that computes the function $\phi^n : A^{2n+1} \to A$ for any n, where ϕ^n is defined inductively as the top of a light-cone of base $2n+1$ (see Figure 5). For example, if we take $n = 2$,

$$\phi^2(x_{-2}, x_{-1}, x_0, x_1, x_2) = \phi(\phi(x_{-2}, x_{-1}, x_0), \phi(x_{-1}, x_0, x_1), \phi(x_0, x_1, x_2)).$$

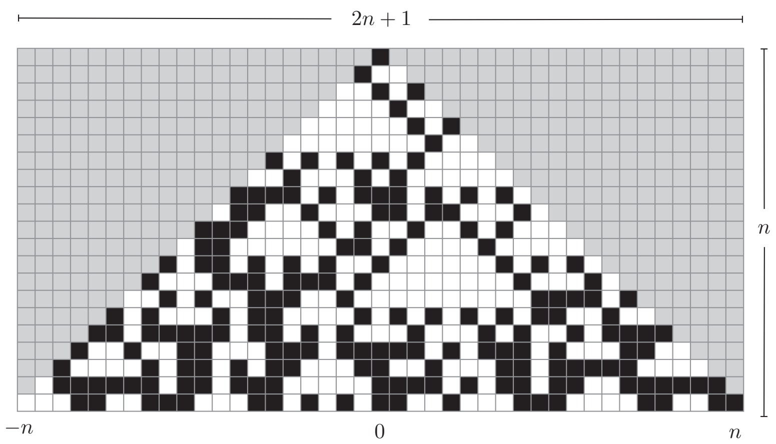

Fig. 5 Light-cone of base $2n + 1$ for the CA $\phi(x_{-1}, x_0, x_1) = x_{-1} + x_1 \mod 1$.

In computer science, functions such as $\phi^n : A^{2n+1} \to A$ can be seen as input/output problems, where the input is an element $x \in A^{2n+1}$ and the corresponding output is $\phi^n(x) \in A$. Moreover, we refer to this problem –which consists in predicting the state at the top of the light-cone– as the *prediction problem*. More precisely, $\textsc{Pred}(\phi) = \phi^n$

We can visualize the cells of a CA as nodes in a one-dimensional arrangement where each node is connected through a channel with its two neighbors (with the exception of the two extreme nodes, which have only one neighbor). A classic measure of efficiency in such distributed systems is the number of bits that cross each channel during the computation process.

For example, when considering the prediction problem $\textsc{Pred}(\phi) : A^{2n+1} \to A$, we can say that, by solving it through the $(2n+1)$-node CA ϕ, 2 states (elements of A) pass through the first and last channel (one in each direction), while in the channels adjacent to the central node, the number of states that pass is $2n$ (n in each direction). In general, if we take $1 \leq i \leq n$, we have that both in the i-th channel and in the $(2n+1-i)$-th channel the number of states that pass is $2i$ (i in each direction).

In distributed computing one typically looks for *bottlenecks*. That is, channels for which the nature of the problem (in this case $\textsc{Pred}(\phi)$) forces the exchange of a significant number of bits. In practice, a widely used way to understand the congestion of a set of channels is to divide the system into two subsystems joined by such channels (that we also understand as cuts) and try to find (lower and upper) bounds for the number of bits that must pass through these channels.

For the upper bounds, we look for protocols. For the lower bounds, we typically use techniques from the theory of communication complexity. Note that, in our context, every channel is a bridge (i.e., by removing it we disconnect the system into two independent components).

Let us introduce formally a key definition that captures previous ideas.

Definition 6 Let ϕ be CA. The two-party communication problem $\textsc{Pred}_i(\phi)$ is defined, for all $1 \leq i \leq 2n$, as follows.

$$\textsc{Pred}_i(\phi) : A^i \times A^{2n+1-i} \to A,$$

where Alice receives $x \in A^i$, Bob receives $y \in A^{2n+1-i}$, and together they must compute $\phi^n(xy)$.

Now, in order to illustrate its usefulness, we are going to study the communication problem $\textsc{Pred}_i(\phi)$ on particular CAs, called *elementary*.

4.1 Elementary CAs

The family of elementary CAs (ECAs) is the family of CAs of radius $r = 1$ and binary alphabet $A = \{0, 1\}$. There are $2^{2^3} = 256$ CAs of this type, each of which corresponds to a local rule $\phi : \{0, 1\}^3 \to \{0, 1\}$ that can be identified with its corresponding Wolfram number defined as $\sum_{x,y,z\in\{0,1\}} 2^{4x+2y+z}\phi(x, y, z)$, which is between 0 and 255.

Thus, for instance, ECA Rule 105 is given by $\phi_{105}(x, y, z) = x + y + z + 1$, where the sum is taken modulo 2. This CA satisfies the following *affine* property [22]: for all $1 \leq i \leq 2n$, $x \in \{0,1\}^i$, $y \in \{0,1\}^{2n+1-i}$,

$$\phi^n_{105}(xy) = \phi^n_{105}(x \underbrace{0\dots0}_{2n+1-i}) + \phi^n_{105}(\underbrace{0\dots0}_{i} y) + n,$$

where $xy \in \{0,1\}^{2n+1}$ denotes the concatenation of x and y.

Note that the communication complexity of $\textsc{Pred}_i(\phi_{105})$ is a lower bound for the number of bits that must pass through the channel that connects the i-th node with the $(i+1)$-th node in order to compute $\textsc{Pred}_i(\phi_{105})$. More precisely, there is one communication channel between Alice and Bob.

The affine property mentioned above allows us to design a very simple protocol, whose communication complexity is 2 bits. The protocol is as follows [22]. Alice computes $b = \phi^n(x0\dots0)$ and sends bit $b \in \{0,1\}$ to Bob. Then Bob computes $b + \phi^n(0\dots0y) + n$ and shares the result with Alice. Hence, $\mathbf{cc}\,(\textsc{Pred}_i(\phi_{105})) = 2$.

In the protocol defined for ECA Rule 105, Alice sends a message to Bob, and then Bob decides the output and communicates it to Alice. As we already mentioned in Section 2, this particular type of protocols are known as *AB*-one-way protocols.

Consider now, for any ECA ϕ, the matrix $M^i_\phi = M_{\textsc{Pred}_i(\phi)}$ with 2^i rows and 2^{2n+1-i} columns, where $M^i_\phi(x, y) = \phi^n(xy)$. As we explained in Section 2, $\mathbf{cc}^{AB}(\textsc{Pred}_i(\phi))$, the *AB*-one-way communication complexity of a CA ϕ corresponds *exactly* to $\log(d(M^i_\phi))$, where $d(M^i_\phi)$ is the number of different rows of the matrix M^i_ϕ.

The fact that the one-way communication complexity corresponds to a parameter of the matrix M^i_ϕ motivated the authors in [22] to measure the one-way communication complexity of the ECAs for different values of n through emulations, by simply counting the number of different rows.

Note that, since $|A| = 2$, $0 \leq \mathbf{cc}^{AB}\,(\textsc{Pred}_i(\phi)) \leq i$. The emulations suggest the existence of well-defined types of behaviors for the ECAs, highly correlated with the four *Wolfram classes* [61]. Indeed, it was observed that the number of different rows of the matrices M^n_ϕ was upper bounded by a constant, grew polynomially, or grew exponentially, depending on the particular ECA ϕ.

Many of these observations could be rigorously proved. In [22,30], the authors exhibited (optimal) one-way protocols for different ECAs. More precisely, for ECAs ϕ for which $\mathbf{cc}^{AB}\,(\textsc{Pred}_n(\phi))$ was constant or equal to $\log(n)$.

In [28], a protocol for ECA Rule 218, where Alice needs to send 2 positions of her input $x \in \{0,1\}^n$ (i.e., $2\log(n)$ bits), was given. More precisely, it was proved that $\mathbf{cc}^{AB}\,(\textsc{Pred}_n(\phi_{218})) \leq 2\log(n)$. By using communication complexity arguments, it was also proved that $\mathbf{cc}^{AB}\,(\textsc{Pred}_n(\phi_{218})) \geq 2\log(n) - 5$. Note that the difference, in terms of the number of different rows in the matrix, between messages of size $\log(n)$ and $2\log(n)$ is n versus n^2.

4.2 The connection with intrinsic universality

As we have already said, given a CA ϕ, the communication complexity of $\textsc{Pred}(\phi)$ attempts to capture the idea of bottleneck, i.e., a splitting that induces the highest traffic. The definition is as follows:

$$\mathrm{CC}(\textsc{Pred}(\phi)) = \max_{1 \le i < 2n+1} \mathbf{cc}\left(\textsc{Pred}_i(\phi)\right).$$

The main connection between bottlenecks and emulation between CAs is that the communication complexity of the prediction problem increases monotonically with respect to the order induced by emulations. More precisely, if a CA emulates another one, its prediction problem has to have the same or higher communication complexity than the emulated one. Combining the previous fact with the existence of just one particular —not necessarily intrinsically universal— CA ϕ such that

$$\mathrm{CC}(\textsc{Pred}(\phi)) \in \Omega(n),$$

one has immediately the following useful result that links intrinsic universality with communication complexity.

Theorem 1 *[29] Let ϕ be a CA such that $\mathrm{CC}(\textsc{Pred}(\phi)) \in o(n)$. Then, ϕ is not intrinsically universal.*

The previous theorem tell us that it is enough that a CA ϕ satisfies that $\mathrm{CC}(\textsc{Pred}(\phi)) \in o(n)$ for ruling it out from being intrinsically universal. Hence, we can conclude that many different CAs such as number-conserving ECAs, monotone ECAs, ECA Rule 105, ECA Rule 218, equicontinuous CAs, etc., are not intrinsically universal [29,30].

The existence of a CA ϕ such that $\mathrm{CC}(\textsc{Pred}(\phi)) \in \Omega(n)$ is interesting by itself. The main technique here is to design CAs that “solve”, by means of their dynamics, well-known two-party communication problems. Then, solving a problem like EQ, DISJ, etc., reduces to predicting the evolution of the CA designed for this purpose.

For example, the prediction problem associated to the CA ϕ depicted in Figure 6 could be used to solve the Equality problem of the strings represented at both sides of the central tower of $\top$'s. The idea is the following: given strings $x = x_1 \cdots x_n$ and $y = y_1 \cdots y_n$, represent both of them as rows of “labeled particles” that move towards the central tower from left to right and from right to left, respectively. In this particular scenario, one can perfectly define a local rule that generates this motion; then, in n steps, it is possible to determine if the two strings are the same or not since any difference between the two strings will eventually reach the central tower (more specifically, if i is the first index such that $x_i \neq y_i$, then the corresponding particles will reach the center in time i).

Therefore, the question of equality between the two strings can be translated into the question of whether any difference will reach the central tower

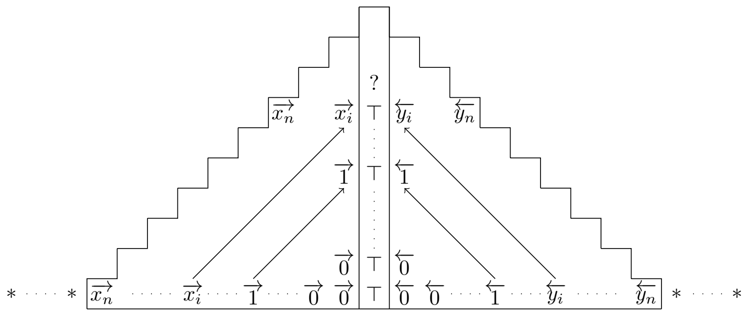

Fig. 6 CA with a prediction problem that solves EQ.

in time n or not, and such information can be registered in the central tower by just considering a special state $\perp$. Then, if the outcome of the prediction problem is $\top$, we know that $x = y$. On the other hand, if the outcome is $\perp$, we know that $x \neq y$. This can be formalized by constructing an explicit fooling set; however, it is not hard to see that one can naturally translate the usual fooling set for EQ into a fooling set for $\textsc{Pred}_n(\phi)$.

5 Communication problems in CAs

Let ϕ be a fixed CA and Φ its corresponding global transition function. For simplicity, we will keep assuming that ϕ has radius 1, although all the definitions introduced in this section can be extended to the general setting. Considering this, we can formulate at least five problems induced by CAs relevant for our purposes. Besides the prediction problem, four additional problems that are also relevant are the following: length of cycle, spatial-invasion, temporal-invasion, and controlled-invasion.

Let $x \in A^n$ be an input.

The **prediction problem** [29], denoted by $\textsc{Pred}(\phi)$, was introduced in Section 4. We consider here a natural generalization related to the definition of the output when the base of the light-cone is of even size. More precisely, $\textsc{Pred}(\phi)$ outputs the value $z \in A$ if n is odd, or $z_1 z_2 \in A^2$ if n is even, obtained after iterating $\lfloor \frac{n-1}{2} \rfloor$ steps the CA ϕ starting from x.

For the following problems, we require some extra notation. Given a finite word u, we denote by p_u the (spatially) periodic configuration $\dots uuu \dots \in A^{\mathbb{Z}}$ constructed by repeating u infinitely many times. Notice that the evolution of Φ starting from p_u becomes (temporally) periodic after a finite number of steps, i.e., there exists $t_0 > 0$ such that $\Phi^{t+t_0}(p_u) = \Phi^t(p_u)$ for al $t \geq 0$, since the spatial period of p_u never increases under the action of Φ.

Given a fixed parameter $k \in \mathbb{N}$, the **length of cycle problem** [29], denoted by $\text{Cycl}^k(\phi)$, considers the evolution of $\Phi^t(p_x)$ and outputs a 1 if the length of the ultimate (temporal) period is less than or equal to k, and 0 otherwise.

Now, we denote by $p_u(x)$ the "perturbed" configuration obtained from p_u by replacing the original content of coordinates from 0 to $n-1$ with x. Then, given a fixed finite word u, we can define the three following problems (see Figure 7):

- the **spatial-invasion problem** [29], denoted by $\text{SInv}^u(\phi)$, compares the time evolution of p_u with $p_u(x)$ and outputs 1 if the distance between the rightmost and leftmost disagreement between $\Phi^t(p_u)$ and $\Phi^t(p_u(x))$ goes to infinity as t goes to infinity, and 0 otherwise;
- the **temporal-invasion problem** [9,11], denoted by $\text{TInv}^u(\phi)$, compares the time evolution of p_u with $p_u(x)$ and outputs 1 if the disagreement between $\Phi^t(p_u)$ and $\Phi^t(p_u(x))$ persists forever, and 0 otherwise;
- the **controlled-invasion problem** [9,11], denoted by $\text{CInv}^u(\phi)$, compares the time evolution of p_u and $p_u(x)$ and outputs 1 if the disagreement between $\Phi^t(p_u)$ and $\Phi^t(p_u(x))$ persists forever but the distance between the rightmost and leftmost disagreement remains bounded as t goes to infinity, and 0 otherwise.

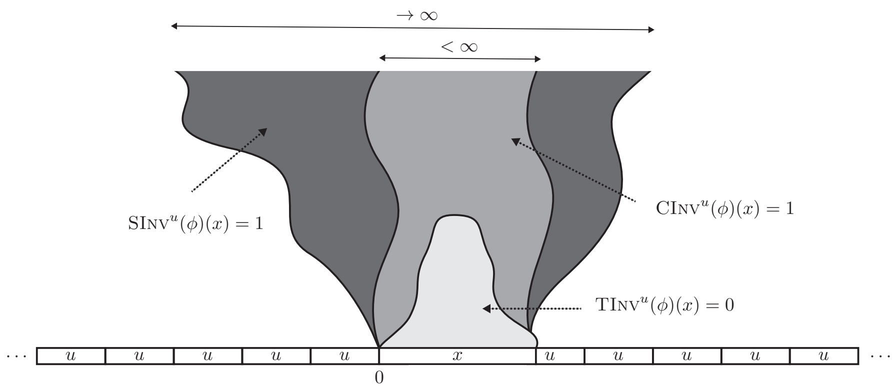

Fig. 7 Possible outcomes for the disagreement between $(\Phi^t(p_u))_{t\geq 0}$ and $(\Phi^t(p_u(x)))_{t\geq 0}$.

Notice that in every problem introduced here, the input is $x \in A^n$. The integer k and the word u are just parameters. Then, for each decision problem, we can define the associated communication problem obtained after splitting the input x.

Now, the key feature is that the communication complexity of the aforementioned problems must be monotone with respect to emulations. Before stating the theorem, let us define the $\prec$ relation. Given $f, g : \mathbb{R}_+ \to \mathbb{R}_+$, we denote $f \prec g$ if

$$\exists a, b, d, x_0 \in \mathbb{R}_+, \exists c \in \mathbb{R} : \forall x \geq x_0, f(ax) \leq dg(b(x+c)).$$

Considering this, we have the following theorem.

Theorem 2 ([9, 11, 29]) *Let ϕ and ψ be two CAs such that $\phi \preccurlyeq \psi$. Then:*

1. $\mathrm{CC}(\textsc{Pred}(\phi)) \prec \mathrm{CC}(\textsc{Pred}(\psi))$.
2. $\forall\ k \in \mathbb{N}$, $\exists\ k_1, k_2 \geq k$ *such that* $\mathrm{CC}(\textsc{Cycl}^{k_1}(\phi)) \prec \mathrm{CC}(\textsc{Cycl}^{k_2}(\psi))$.
3. *For every finite word u, there exists a finite word v such that* $\mathrm{CC}(\mathrm{P}^u(\phi)) \prec \mathrm{CC}(\mathrm{P}^v(\psi))$, *where* $\mathrm{P} \in \{\mathrm{S}\textsc{Inv}, \mathrm{T}\textsc{Inv}, \mathrm{C}\textsc{Inv}\}$.

From Theorem2 it follows that, if a CA ψ is intrinsically universal, then it must be the case that the communication complexity is maximal (in the sense of $\prec$) for every problem. It turns out that for all the problems considered here, there exists some CA ϕ (not necessarily intrinsically universal) such that the communication complexity is of order n. Therefore, if a CA ψ is intrinsically universal, there exist k and u such that

$$\mathrm{CC}(\textsc{Pred}(\psi)), \mathrm{CC}(\textsc{Cycl}^k(\psi)), \mathrm{CC}(\mathrm{P}^u(\psi)) \succ n,$$

where $\mathrm{P} \in \{\mathrm{S}\textsc{Inv}, \mathrm{T}\textsc{Inv}, \mathrm{C}\textsc{Inv}\}$. (A priori, u may depend on P but, with additional work, it is possible to prove that a unique u suffices.)

Another interesting fact is that these problems are in some sense *orthogonal.* In other words, given a pair of problems P_1 and P_2, it is possible to construct a specific CA ϕ such that the communication complexity of P_1 for ϕ is maximal and the communication complexity of P_2 for ϕ is trivial [11, 29]. Therefore, all the problems presented here are independent and none of them can be reduced to another one.

The relevance of having a variety of problems can be appreciated better when studying classes of CAs. Given a CA ϕ and its corresponding global transition function Φ, it is customary to define classes of CAs according to the dynamical behavior of Φ on $A^{\mathbb{Z}}$. Considering the previous problems and developing clever protocols based on the general properties of a given class, it is possible to discard full classes of CAs from being intrinsically universal by proving that the communication complexity of a particular problem is $o(n)$.

Let us consider the following case. A CA ϕ is said to be *nilpotent* if there exists $s \in A$ such that $\Phi^t(c)$ evolves towards the constant configuration $\overline{s} = \ldots sss \ldots$ for every $c \in A^{\mathbb{Z}}$. In such case, given a finite word u and an input $x \in A^n$, then both p_u and its perturbation $p_u(x)$ will evolve towards $\overline{s}$ under Φ. In particular, for large enough t, there is no disagreement between $\Phi^t(p_u)$ and $\Phi^t(p_u(x))$. Therefore, the answer to $\mathrm{S}\textsc{Inv}^u(\phi)$, $\mathrm{T}\textsc{Inv}^u(\phi)$, and $\mathrm{C}\textsc{Inv}^u(\phi)$ is always 0, for any u. In particular, there is no communication needed in order to solve them, and therefore $\mathrm{CC}(\mathrm{S}\textsc{Inv}^u(\phi)), \mathrm{CC}(\mathrm{T}\textsc{Inv}^u(\phi)), \mathrm{CC}(\mathrm{C}\textsc{Inv}^u(\phi)) \in O(1)$.

On the other hand, a CA is *nilpotent restricted to periodic configurations* if there exists $s \in A$ such that $\Phi^t(p)$ evolves towards the constant configuration $\overline{s}$ for every (spatially) periodic configuration $p \in A^{\mathbb{Z}}$ [42]. It is possible to prove that this class of CAs is simple in terms of communication for some problems. Indeed, given any $k \in \mathbb{N}$, we have that $\mathrm{CC}(\textsc{Cycl}(\Phi)^k) \in O(1)$, because every periodic configuration p_x for some input $x \in A^n$ evolves towards $\overline{s}$, which is a fixed point for Φ, so the ultimate (temporal) period has always length 1 and $\textsc{Cycl}(\Phi)^k$ is constant ($\textsc{Cycl}(\Phi)^k = 0$ if $k \geq 2$ and $\textsc{Cycl}(\Phi)^k = 1$ if $k = 1$).

We do not attempt to be exhaustive here with regard to all the definitions involved, but let us mention other meaningful classes. For example, if ϕ is such that Φ is *equicontinuous* or *linear*, it is possible to prove that $\mathrm{CC}(\textsc{Pred}(\phi)) \in O(1)$ (indeed, the protocol for a linear CA is very similar to the one described for ECA Rule 105 in Section 4); if ϕ is such that Φ is *reversible* (or, as we discussed before, nilpotent on periodic configurations), then $\mathrm{CC}\left(\mathrm{CYCL}^k_\Phi\right) = 0$ for all $k \in \mathbb{N}$; as we discussed before, if ϕ is such that Φ is *nilpotent*, then $\mathrm{CC}(\mathrm{S}\textsc{Inv}^u(\phi)) = 0$ for every finite word u; if ϕ is such that Φ is *positively expansive*, then $\mathrm{CC}(\mathrm{C}\textsc{Inv}^u(\phi)) = 0$ for every finite word u; and if ϕ is such that Φ is *surjective* (or more generally, its *limit set* is a *subshift of finite type*), then $\mathrm{CC}(\mathrm{T}\textsc{Inv}^u(\phi)) = 1$ for every finite word u. We refer to [9, 11, 29] for these and other examples.

6 Overlapping of problems in CAs

In the context of CAs and intrinsic universality, one usually looks for communication problems that are monotone with respect to emulations and such that (1) they are easy to solve for a great number of CAs and (2) they are difficult to solve for a particular CA. Then, the more CAs that induce instances of the problem with low communication complexity, the more CAs that will be ruled out from being intrinsically universal. Considering this principle of what makes a problem a good filter for discarding CAs from being intrinsically universal, it is a natural step in that direction to use the overlapping relation defined in Section 2. In fact, solving the overlapping relation is necessarily not harder than solving individually any of the problems involved in its definition.

Considering this, given $k \in \mathbb{N}$ and a finite word u, we define the **overlapping problem** as

$$\textsc{Ovrl}^{k,u}(\phi) := \textsc{Pred}(\phi) \uplus \textsc{Cycl}^k(\phi) \uplus \mathrm{S}\textsc{Inv}^u(\phi) \uplus \mathrm{T}\textsc{Inv}^u(\phi) \uplus \mathrm{C}\textsc{Inv}^u(\phi).$$

Analogously to what happened in previous section, there exist particular CAs which are hard for $\mathrm{CC}(\textsc{Ovrl}^{k,u}(\phi))$. Hence, we have the following result.

Theorem 3 ([11]) *Let ϕ be a CA such that, for all $k \in \mathbb{N}$ and every finite word u, $\mathrm{CC}(\textsc{Ovrl}^{k,u}(\phi)) \in o(n)$. Then, ϕ is not intrinsically universal.*

It is possible to check that if each of the problems involved in the definition of $\textsc{Ovrl}^{k,u}(\phi)$ is compatible with emulations, then the overlapping problem will also be. A natural question is whether the overlapping problem can be strictly easier than all the other problems individually for some particular CA such as in the case of EQ and DISJ. The next proposition answers this question affirmatively.

Proposition 2 ([11]) *There exists a CA ϕ such that*

1. $\mathrm{CC}(\textsc{Pred}(\phi)) \in \Omega(n)$*;*
2. there exists $k \in \mathbb{N}$ *such that* $\mathrm{CC}(\textsc{Cycl}^{k_1}(\phi)) \in \Omega(n)$*;*
3. there exists a finite word u *such that* $\mathrm{CC}(\mathrm{P}^u(\phi)) \in \Omega(n)$*, where* $\mathrm{P} \in \{\mathrm{S}\textsc{Inv}, \mathrm{T}\textsc{Inv}, \mathrm{C}\textsc{Inv}\}$*;*

and, for all $k \in \mathbb{N}$ *and for all finite words* u*,*

$$\mathrm{CC}(\textsc{Ovrl}^{k,u}(\phi)) \in O(1).$$

7 Conclusion

In this tutorial we presented a deep connection between communication complexity and intrinsic universality in (one-dimensional) CAs. The first step was to introduce communication problems parametrized by CAs. Some communication problems turned out to be very useful for ruling out specific CAs from being intrinsically universal. In fact, we were able to find particular problems P's for which the following key property holds: if ϕ is intrinsically universal, then the communication complexity of $\mathrm{P}(\phi)$ must be maximal. We defined five such "canonical" problems. Note that, if the communication complexity of $\mathrm{P}(\phi)$ is low, then we are ruling out ϕ from being intrinsically universal. It is clear that the main goal of our approach is to find a problem P having a small set of CAs ϕ's for which the communication complexity of P_ϕ is high. In such a way, P will be a good filter for ruling out CAs from being intrinsically universal.

Instead of finding new problems like P, we explained in this tutorial how to use all the canonical problems *simultaneously*. The idea is to give much more freedom to Alice and Bob, the two parties of the communication complexity model: depending on the input they receive, they are allowed to *choose* the canonical problem to solve. By definition, this new "overlapping" problem —which we denote Ovrl— will be simpler (in terms of communication complexity) than all the canonical ones. In fact, given an input, in order to solve $\textsc{Ovrl}(\phi)$ it suffices to find *any* canonical problem P for which $\mathrm{P}(\phi)$ is simple.

Therefore, for a non intrinsically universal CA ϕ it is much more likely to obtain a result saying that $\textsc{Ovrl}(\phi)$ has low communication complexity; and this result serves as a certificate to the fact that ϕ is not intrinsically universal. To do this we must be able to prove the existence of at least one CA ϕ for which $\textsc{Ovrl}(\phi)$ is high.

It is known that a necessary condition for a CA ϕ to be intrinsically universal is the P-completeness of the prediction problem $\textsc{Pred}(\phi)$ when viewed as a classical computational problem [47]. In fact, Neary and Woods [45] proved that $\textsc{Pred}(\phi_{110})$ is P-COMPLETE for the ECA Rule 110. However, it is not known yet whether ECA Rule 110 is intrinsically universal. Since it is not difficult to find non intrinsically universal CAs for which Pred is P-COMPLETE [18], we think that our approach is very useful for proving negative results for some particular CAs. The reason is the following: there exist CAs whose prediction problem is P-COMPLETE but for which the communication complexity of (the canonical problem) Pred grows as $o(n)$ [29].

In distributed computing, applying communication complexity is a standard method for proving lower bounds in a variety of network topologies [12]. This can be regarded as an indication that our approach could be extended to CAs of higher dimensions (in particular, dimension 2) and other scenarios.

Acknowledgements The first author was supported by ANID/FONDECYT de Iniciación en Investigación 11200892. The second author was supported by ANID/PIA Apoyo a Centros Científicos y Tecnológicos de Excelencia AFB 170001 and Fondecyt 1170021.

References

1. J. Albert and K. Čulik II. A simple universal cellular automaton and its one-way and totalistic version. *Complex Systems* 1, 1–16, 1987.
2. M. Aldana, S. Coppersmith, and L. P. Kadanoff. Boolean dynamics with random couplings. In *Perspectives and Problems in Nolinear Science* (pp. 23–89), 2003.
3. S. Arora and B. Barak. Computational complexity: a modern approach. *Cambridge University Press*, 2009.
4. E. R. Banks. Universality in cellular automata. *In 11th Annual Symposium on Switching and Automata Theory (SWAT)*, IEEE, 194–215, 1970.
5. M. Batty. Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals. *The MIT press*, 2007.
6. C. Bone, S. Dragicevic, and A. Roberts. A fuzzy-constrained cellular automata model of forest insect infestations. *Ecological Modelling*, 24(2), 247–261, 1997.
7. L. Boyer and G. Theyssier. On local symmetries and universality in cellular automata. *Proceedings of the 26th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 195–206, 2009.
8. D. Brand and P. Zaropulo. On communicating finite-state machines. *J. ACM*, 30:323–342, 1983.
9. R. Briceño and P.-E. Meunier. The structure of communication problems in cellular automata. *DMTCS Proceedings, AUTOMATA 2011*, 59–76, 2011.
10. R. Briceño, P. Moisset de Espanés, A. Osses, and I. Rapaport. Solving the density classification problem with a large diffusion and small amplification cellular automaton. *Physica D: Nonlinear Phenomena*, 261, 70–80, 2013.
11. R. Briceño and I. Rapaport. Letting Alice and Bob choose which problem to solve: Implications to the study of cellular automata. *Theoretical Computer Science*, 468, 1–11, 2013.
12. K. Censor-Hillel, S. Khoury, and A. Paz. Quadratic and near-quadratic lower bounds for the CONGEST model. *In Proceedings of the 31st International Symposium on Distributed Computing (DISC)*,10:1–10:16, 2017.
13. I. Chlamtac and S. Kutten. On Broadcasting in Radio Networks - Problem Analysis and Protocol Design. *IEEE Transactions on Communications* 33(12):1240–1246, 1985.
14. K. C. Clarke, S. Hoppen, and L. Gaydos. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and planning B: Planning and design.*
15. M. Cook. Universality in elementary cellular automata. *Complex Systems*, 15:1–40, 2004.
16. A. Cornejo and F. Kuhn. Deploying wireless networks with beeps. *In Proceedings of the 24th International Conference on Distributed Computing (DISC)*,148-162, 2010.
17. M. Creutz. Deterministic Ising dynamics. *Annals of Physics* 167.1, 62–72, 1986.
18. M. Delorme, J. Mazoyer, N. Ollinger, and G. Theyssier. Bulking II: Classifications of Cellular Automata. *Theoretical Computer Science* 4012(30): 3881–3905, 2011.
19. E. D. Demaine, M. J. Patitz, T. A. Rogers, R.T. Schweller, S. M. Summers, and D. Woods. The two-handed tile assembly model is not intrinsically universal. *Algorithmica*, 74(2), 812-850, 2016.
20. D. Doty, J. H. Lutz, M. J. Patitz, R. T. , Schweller, S. M. Summers, and D. Woods. The tile assembly model is intrinsically universal. *In 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science* (pp. 302–310), IEEE, 2012.

21. B. Durand and Z. Róka. *Cellular automata: a parallel model*, volume 460 of *Mathematics and its Applications*, chapter The game of life: universality revisited, pages 51–74. Kluwer Academic Publishers, 1999.
22. C. Dürr, I. Rapaport, and G. Theyssier. Cellular automata and communication complexity. *Theoretical Computer Science* 322:355–368, 2004.
23. Y. Emek and R. Wattenhofer. Stone age distributed computing. *In Proceedings of the 2013 ACM Symposium on Principles of Distributed Computing*, 137–146, 2013.
24. G. B. Ermentrout and L. Edelstein-Keshet. Cellular automata approaches to biological modeling. *Journal of Theoretical Biology* 160(1), 97–133, 1993.
25. N. Fates, É. Thierry, M. Morvan, and N. Schabanel. Fully asynchronous behavior of double-quiescent elementary cellular automata. *Theoretical Computer Science* 362: 1-3, 1–16, 2006.
26. H. Fukś. Nondeterministic density classification with diffusive probabilistic cellular automata. *Physical Review E*, 66(6), 066106, 2002.
27. Z. P. Gerdtzen, J. C. Salgado, A. Osses, J. A. Asenjo, I. Rapaport, and B. A. Andrews. Modeling heterocyst pattern formation in cyanobacteria. *BMC Bioinformatics.* Vol. 10. No. 6. BioMed Central, 2009.
28. E. Goles, C. Little, and I. Rapaport. Understanding a non-trivial cellular automaton by finding its simplest underlying communication protocol. *Proceedings of the 19th International Symposium on Algorithms and Computation (ISAAC)*, pages 592–604, 2008.
29. E. Goles, P.-E. Meunier, I. Rapaport, and G. Theyssier. Communication complexity and intrinsic universality in cellular automata. *Theoretical Computer Science* 412:2–21, 2011.
30. E. Goles, A. Moreira, and I. Rapaport. Communication complexity in number-conserving and monotone cellular automata. *Theoretical Computer Science* 412:3616–3628, 2011.
31. G.A. Hedlund. Endomorphisms and automorphisms of the shift dynamical systems. *Mathematical System Theory* 3(4):320–375, 1969.
32. M. Karchmer and A. Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *Proceedings of the 27th Annual ACM Symposium on Theory of Computing (STOC)*, pages 539–550. ACM, 1988.
33. J. Kari. Reversibility and surjectivity problems of cellular automata. *Journal of Computer and System Sciences*, 48(1), 149–182, 1994.
34. J. Kari. Theory of cellular automata: A survey. *Theoretical Computer Science* 334(1-3), 3–33, 2005.
35. S. A. Kauffman. The origins of order: Self-organization and selection in evolution. *Oxford University Press*, USA, 1993.
36. P. Kurka. Languages, equicontinuity and attractors in cellular automata. *Ergodic Theory and Dynamical Systems*, 17(2), 417–433, 1997.
37. E. Kushilevitz and N. Nisan. Communication complexity. *Cambridge University Press*, 1997.
38. K. Lindgren and M. G. Nordahl, M. G. Universal computation in simple one-dimensional cellular automata. *Complex Systems*, 4(3), 299–318, 1990.
39. S. Maerivoet and B. De Moor. Cellular automata models of road traffic. *Physics Reports*, 419(1), 1–64, 2005.
40. F. Martinelli, R. Morris, and C. Toninelli. Universality results for kinetically constrained spin models in two dimensions. *Communications in Mathematical Physics* 369.2: 761–809, 2019.
41. J. Mazoyer, and I. Rapaport. Inducing an order on cellular automata by a grouping operation. *Proceedings of the 15th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 116–127, 1998.
42. J. Mazoyer and Rapaport. Global fixed point attractors of circular cellular automata and periodic tilings of the plane: undecidability results. *Discrete Mathematics*, 199(1-3), 103–122, 1999.
43. P. E. Meunier, M. J. Patitz, S. M. Summers, G. Theyssier, A. Winslow, and D. Woods. Intrinsic universality in tile self-assembly requires cooperation. *In Proceedings of the twenty-fifth annual ACM-SIAM Symposium on Discrete Algorithms* , 752–771, Society for Industrial and Applied Mathematics, 2014.

44. P. Moisset de Espanés, A. Osses, and I. Rapaport. Fixed-points in random Boolean networks: The impact of parallelism in the Barabási-Albert scale-free topology case. *BioSystems* 150, 167–176, 2016.
45. T. Neary and D. Woods. P-completeness of cellular automaton Rule 110. *Proceedings of the 25th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 132–143, 1998.
46. N. Ollinger. The quest for small universal cellular automata. *Proceedings of the 29th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 318–329, 2002.
47. N. Ollinger. Intrinsically universal cellular automata. *Proceedings of International Workshop on The Complexity of Simple Programs (CSP)*, pages 318–329, 2008.
48. N. Ollinger. Universalities in cellular automata: a (short) survey. *Proceedings of the First Symposium on Cellular Automata Journées Automates Cellulaires (JAC)*, pages 102–118, 2008.
49. N. Ollinger and G. Richard. Four states are enough! *Theoretical Computer Science* 412:22–32, 2011.
50. D. Peleg. Distributed computing: a locality-sensitive approach. *Society for Industrial and Applied Mathematics*, 2000.
51. P. Prusinkiewicz and A. Lindenmayer. The algorithmic beauty of plants. Springer Science & Business Media, 2012.
52. I. Rapaport, K. Suchan, I. Todinca, and J. Verstraete. On dissemination thresholds in regular and irregular graph classes. *Algorithmica*, 59(1), 16–34, 2011.
53. A. A. Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. *Combinatorica*, 10(1), pages 81–93. 1990.
54. A. R. Smith III. Simple computation-universal cellular spaces. *Journal of the ACM (JACM)*, 18(3), 339–353, 1971.
55. G. Theyssier. How common can be universality for cellular automata? *Proceedings of the 26th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 121–132, 2005.
56. T. Toffoli and N. H. Margolus. Invertible cellular automata: a review. *Physica D: Nonlinear Phenomena*, 45(1-3), 229–253, 1990.
57. M. Tomassini, M. Giacobini, and C. Darabos. Evolution and dynamics of small-world cellular automata. *Complex Systems*, 15(4), 261–284, 2005
58. J. von Neumann. Theory of Self-Reproducing Automata. *University of Illinois Press*, 1966.
59. D. A. Wolf-Gradow, Lattice-gas cellular automata and lattice Boltzmann models: an introduction. Springer, 2004.
60. S. Wolfram. Statistical mechanics of cellular automata. *Reviews of Modern Physics*, 55(3), 601, 1993.
61. S. Wolfram. Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1-2), 1–35, 1984.
62. D. Woods. Intrinsic universality and the computational power of self-assembly. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2046), 20140214, 2015.
63. A. C. C. Yao. Some complexity questions related to distributive computing. *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.