Theoretical Computer Science 641 (2016) 2–10

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Unraveling simplicity in elementary cellular automata

Pierre-Étienne Meunier

Aix Marseille Université, CNRS, LIF UMR 7279, 13288, Marseille, France

A R T I C L E I N F O

Article history:
Received 3 March 2015
Received in revised form 23 November 2015
Accepted 6 January 2016
Available online 8 March 2016
Communicated by J. Kari

Keywords:
Elementary cellular automata
Communication complexity
Intrinsic universality

A B S T R A C T

We provide a mathematical proof that a large number of elementary cellular automata are computationally simple. This work is the first systematic classification of elementary cellular automata based on a formal notion of computational complexity.

It contrasts with previous approaches in the simplicity of the method – most proofs are just a few lines long and require no heavy computational explorations. More importantly, this type of short proof not only provides evidence for the presence of simple patterns, it also provides *reasons* for this simplicity.

Moreover, thanks to the generality of communication complexity, we hope that this work finds new applications to other natural systems such as neural networks and gene regulatory networks, in particular for real data.

1. Introduction

Many computational processes can be seen as a sequence of information exchanges between parts of the space. Moreover, a large number of natural systems, either physical or biological, bear close similarity to algorithmic processes, in the sense that they are dynamical processes, in which local information exchanges play an important role, as noticed for instance by Maxwell [19] in his "daemon" thought experiment, or more recently by biologists working with DNA, or by quantum information theorists [1].

In this work, we present a method to analyze such systems, and apply it systematically to a popular class of cellular automata, showing that most of them are computationally simple. This task was previously tackled, using a different approach, by Israeli and Goldenfeld [14,15], who showed that many elementary cellular automata can be "coarse-grained", meaning that they were able to build a simulation hierarchy. In their work, the property that A simulates B means that B is equal to some projection of some rescaling of A. However, that notion of simulation is not well understood, as even the existence of a universal cellular automaton for it is a long-standing open problem of the field [3]. The relation between their work and ours may shed new light on this problem.

That approach is thus physically and computationally relevant, but relies on the exhaustive exploration, using computer programs, of a large number of possible projections. In this work, we bring this search for simplicity to a new level, by giving a simple proof method showing that our objects of study are simple. It must be stressed that although algorithmic explorations are good at determining *whether* something is true, mathematical proofs also explain *why* it is true. Therefore, our approach, as well as the discrete nature of cellular automata, allows for a "hybrid proof method" in which both approaches are used. We hope that it can be applied to other systems, and especially to real world data, to find simple patterns.

E-mail address: pierre-etienne.meunier@lif.univ-mrs.fr.

http://dx.doi.org/10.1016/j.tcs.2016.01.004

1.1. Main results

Although a number of terms have not yet been defined, we state the main result of our paper. The proof of this theorem is presented in Section 3.

Theorem 1.1. *56 of the 88 non-isomorphic elementary cellular automata are not intrinsically universal.*

Our result makes extensive use of the tools developed in [4]. Its main contribution is to exhibit, for a large number of cellular automata, a *communication protocol* of sublinear complexity. Similarities between the protocols and problems used form more relevant complexity classes than previous "experimental classifications" such as [28].

Moreover, it extends the results of [10,4,11] by giving simple protocols for several elementary cellular automata that [10] could not prove non-intrinsically universal.

1.2. Key technical ideas and methods

Our main tool in this task is called communication complexity, which is a computational model introduced by Yao [29] to prove lower bounds in VLSI design (see [18] for a full introduction).

Cellular automata are a model of computation primarily consisting of a large number of simple local interactions. This kind of dynamics is ubiquitous in many physical or biological processes; developing powerful tools to analyze these objects therefore seems an important step in the study of these systems.

Originally introduced by von Neumann [27] to study self-reproduction of computationally meaningful "organisms", cellular automata have given rise to a rich theory: in particular, their computational capabilities have been extensively studied [17,12,16,21]. The idea of universality *modulo rescaling*, also called *intrinsic universality*, has also originated in this model [23,26,2,9], before inspiring research in other fields, most notably in tile self-assembly [24,8,7,6], where it helped understand long-standing open problems of that domain [20].

This notion of simulation is stronger than Turing universality, and allows reasonings on geometric or topological properties, that are most of the time preserved by this operation. However, it is not *too* strong to be meaningful: indeed, *intrinsically universal* cellular automata have been known for some time [23].

The simplest examples of cellular automata, the one-dimensional ones with two states and two neighbors, are called *elementary*. Since the simulation works of Wolfram [28], they have received a lot of attention, culminating in a result by Cook showing that one of these rules (called "rule 110"), is capable of arbitrary Turing computation [5]. His construction was later improved by Neary and Woods [22].

However, being able to *program* a system, i.e. to use it to perform nontrivial algorithmic operations, does not mean that we understand its computational capabilities. On the contrary, complexity *lower bounds* require such an understanding: for example, showing that a class of logic circuits is not able to compute some function requires a full understanding of its capabilities, to make sure that no "hidden trick" allows it to compute that function.

1.3. With experimental data

A key point of our approach, is that it is directly adaptable to experimental data: indeed, the communication complexity of a function can also be computed experimentally. For instance, in Section 3.1, we provide experimental measurements of the one-round variant of communication complexity for all the elementary cellular automata that we have not been able to understand.

The method for these calculations is explained in [18]. Briefly, the algorithm consists in cutting the initial configuration into two parts of equal length, for all configurations of size $2n+1$ for some integer n. One part is given to Alice, the other one is given to Bob. Then, we build a matrix M where each row represents one possible input a for Alice, and each column represents one possible input b for Bob. Then, $M_{a,b}$ is the value of the function we wish to calculate; in the case of Fig. 2, the evolution of the central cell after n steps. If Alice's configurations are of size $n+1$, and Bob's are of size n, this means that this matrix is of size $2^{n+1} \times 2^n$.

Using experimental or simulation data can be done by studying the *one-round communication complexity*, which is equal to the log of the number of lines in M [18]. The one-round communication complexity is the communication complexity restricted to simpler protocols where only one of the players is allowed to speak.

2. Definition and preliminaries

Let Q be a finite set called the set of *states*.

A *cellular automaton* is a map of $Q^{\mathbb{Z}}$ to itself, defined by a *local rule* $f : Q^{2r+1} \to Q$ for some integer r. The cellular automaton defined by a local rule f is the map F such that for all $x \in Q^{\mathbb{Z}}$, all $i \in \mathbb{Z}$, $(F(x))_i = f(x_{i-r}, x_{i-r+1}, \dots, x_{i+r})$. In this case, F is also called the *global rule*. An important particular case of cellular automata is the *shift operator* σ, which is the map defined for all $x \in Q^{\mathbb{Z}}$ by $(\sigma(x))_i = x_{i+1}$ for all $i \in \mathbb{Z}$. Remark for example that any cellular automaton commutes