

## Research Article

# An Information-Based Classification of Elementary Cellular Automata

Enrico Borriello<sup>1,2</sup> and Sara Imari Walker<sup>1,2,3,4</sup>

<sup>1</sup>*Beyond Center for Fundamental Concepts in Science, Arizona State University, Tempe, AZ, USA*

<sup>2</sup>*ASU-SFI Center for Biosocial Complex Systems, Arizona State University, Tempe, AZ, USA*

<sup>3</sup>*School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA*

<sup>4</sup>*Blue Marble Space Institute of Science, Seattle, WA, USA*

Correspondence should be addressed to Enrico Borriello; [enrico.borriello@asu.edu](mailto:enrico.borriello@asu.edu)

Received 7 March 2017; Revised 11 June 2017; Accepted 27 July 2017; Published 7 September 2017

Academic Editor: Sergio Gómez

Copyright © 2017 Enrico Borriello and Sara Imari Walker. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a novel, information-based classification of elementary cellular automata. The classification scheme proposed circumvents the problems associated with isolating whether complexity is in fact intrinsic to a dynamical rule, or if it arises merely as a product of a complex initial state. Transfer entropy variations processed by cellular automata split the 256 elementary rules into three *information classes*, based on sensitivity to initial conditions. These classes form a hierarchy such that coarse-graining transitions observed among elementary rules predominately occur within each information-based class or, much more rarely, down the hierarchy.

## 1. Introduction

Complexity is easily identified, but difficult to quantify. Attempts to classify dynamical systems in terms of their complexity have therefore so far relied primarily on qualitative criteria. A good example is the classification of cellular automaton (CA) rules. CA are discrete dynamical systems of great interest in complexity science because they capture two key features of many physical systems: they evolve according to a local uniform rule and can exhibit rich emergent behavior even from very simple rules [1]. Studying the dynamics of CA therefore can provide insights into the harder problem of how it is that the natural world appears so complex given that the known laws of physics are local and (relatively) simple. However, in the space of CA rules, interesting emergent behavior is the exception rather than the rule. This has generated wide-spread interest in understanding how to segregate those local rules that generate rich, emergent behavior, for example, gliders and particles, computational universality, and so on from those that do not (for the seminal paper on Conway's *Game of Life* cf. [2]). The first outlines of proof of the universality of a two-dimensional Game of Life can be found

in [3, 4]. Proofs of the universality of nonelementary, one-dimensional cellular automata are instead in [5, 6]. For the famous proof of the universality of elementary rule 110 cf. [7]). A complication arises in that the complexity of the output of a CA rule is often highly dependent on that of the input state. This makes it difficult to disentangle emergent behavior that is a product of the initial state from that which is *intrinsic* to the rule. This has resulted in ambiguity in classifying the intrinsic complexity of CA rules as one must inevitably execute a CA rule with a particular initial state in order to express its complexity (or lack thereof).

One of the first attempts to classify CA rules was provided by Wolfram in his classification of elementary cellular automata (ECA) [8]. ECA are some of the simplest CA and are 1-dimensional with nearest-neighbor update rules operating on the two-bit alphabet {■, □}. Despite their simplicity, ECA are known to be capable of complex emergent behavior. Initializing an ECA in a random state leads some rules to converge on fixed point or oscillatory attractors, while others lead to chaotic patterns that are computationally irreducible (such that their dynamics are impossible to predict from the initial state and chosen rule without actually executing the

TABLE 1: Wolfram’s classification of ECA, with only the 88 rule equivalence classes shown (only the lowest number rule of a given equivalence class is listed) [28].

Wolfram’s Class I									
0	8	32	40	128	136	160	168		
Wolfram’s Class II									
1	2	3	4	5	6	7	9	10	11
12	13	14	15	19	23	24	25	26	27
28	29	33	34	35	36	37	38	41	42
43	44	46	50	51	56	57	58	62	72
73	74	76	77	78	94	104	108	130	132
134	138	140	142	152	154	156	162	164	170
172	178	184	200	204	232				
Wolfram’s Class III									
18	22	30	45	60	90	105	122	126	146
150									
Wolfram’s Class IV									
54	106	110							

full simulation). Based on these diverse behaviors, Wolfram identified four distinct complexity classes, shown in Table 1. Class I CA are those whose evolution eventually leads to cells of only one kind. Class II CA lead to scattered stable or oscillating behaviors. Class III CA show an irreducibly chaotic pattern. Class IV CA can exhibit any of the previous behaviors simultaneously and seem to possess the kind of complexity that lies at the interface between mathematical modeling and life studies [9]. Wolfram’s classification stands as a milestone in the understanding of CA properties and still represents the most widely adopted classification scheme. Nonetheless, its qualitative nature represents its main limitation. In fact, as Culik II and Yu showed, formally redefining Wolfram’s classes [10] reveals that determining which class a given ECA belongs to is actually *undecidable*.

Despite this no-go result, efforts to better classify CA have not diminished. Li and Packard [11] introduced a classification scheme refining Wolfram’s originally proposed scheme, with the explicit goal to better distinguish between locally and globally chaotic behaviors. Langton introduced the first attempt at a quantitative classification [12]. His classification implemented a projection of the CA *rule space* over the 1D closed interval  $[0, 1]$ : CA rules with similar qualitative features are then roughly arranged in terms of comparable values of the Langton parameter, with the most complex behaviors found at the boundary separating periodic rules from chaotic ones. Alternative approaches to quantitative classification include generalizations of concepts from continuous dynamical systems applied to CA, such as the maximum Lyapunov Exponent (MLE) [13]. For CA, MLE is defined in terms of the Boolean Jacobian matrix [14] and captures the main properties of its continuous counterpart: encoding the tendency of a CA to reabsorb *damage* in its initial configuration or to let it spread, as a consequence of chaotic dynamics. This approach proved useful in more recent analysis of the stability and the shift in complexity of CA in response to changes in topology [15–18]. A number of other classification schemes have been proposed over the years, based on index complexity [19], density parameter with d-spectrum [20],

communication complexity [21], algorithmic complexity [22], integrated information [23], and so on. A comprehensive review is outside of the scope of this paper. The interested reader is directed to the recent review by Martínez [24] and the literature cited therein.

In this paper, we report on experiments demonstrating new quantitative classification of the intrinsic complexity of ECAs, which differs from earlier attempts by exploiting the main weakness plaguing many approaches to quantitative classification. That is, we explicitly utilize the sensitivity of the expressed complexity of ECA rules to the initial input state. In recent years, there has been increasing interest in using information-theoretic tools to quantify the complexity of dynamical systems, particularly in the context of understanding biological systems [25, 26]. A promising tool in this context is transfer entropy (TE), Schreiber’s measure of the directed exchange of information between two interacting parts of a dynamical system [27]. In what follows, we adopt TE as a candidate quantitative selection criterion to classify the intrinsic complexity of ECAs by exploring the sensitivity of TE to changes in the initial state of a CA.

The paper is structured as follows. In Section 2, we start from the simplest, nontrivial, initial configuration of an ECA and use it to identify the dynamical rules able to produce a complex output (as quantified by TE) by virtue of their intrinsic complexity. In Section 3, we repeat our analysis for more general inputs and identify those outputs whose complexity is instead inherited by the complexity of the input, as opposed to the rule. We then classify ECA rules according to the maximum degree of variability of the output they produce for varying inputs. As we will show, three quantitatively and qualitatively distinct classes naturally emerge from this analysis. In Section 4, we show that this classification induces a partially ordered hierarchy among the rules, such that coarse-graining an ECA of a given class yields an ECA of the same class, or simpler [1]. We conclude by proposing further applications of the classification method presented.

## 2. Intrinsic Complexity

In what follows we identify the complexity of a CA with the *amount of information it processes* during its time evolution. The concept of information processing presented herein is adopted from information dynamics [29–32], a formalism that quantifies computation in dynamical systems using methods from information theory. In information dynamics, Schreiber’s transfer entropy [27] is identified with the information processed between two elements of a dynamical system. In Schreiber’s measure, the transfer of information between a source  $Y$  and a target  $X$  is quantified as the reduction in uncertainty of the future state of  $X$  due to knowledge of the state of  $Y$ . In a CA, the source  $Y$  and target  $X$  are both cells of the CA.

Formally, TE is defined as the mutual information between the source  $Y$  at a certain time  $t$  and the target  $X$  at time  $t + 1$ , conditioned on a given number  $k$  of previous states of  $X$ :

$$T_{Y \rightarrow X} = \sum_{(\mathbf{x}_t^{(k)}, x_{t+1}, y_t) \in \mathcal{S}} p(\mathbf{x}_t^{(k)}, x_{t+1}, y_t) \times \log_2 \frac{p(x_{t+1} | \mathbf{x}_t^{(k)}, y_t)}{p(x_{t+1} | \mathbf{x}_t^{(k)})}, \quad (1)$$

where (see Figure 1)

- (i)  $x_{t+1}$  represents the state of cell  $X$  at time  $t + 1$ ;
- (ii)  $y_t$  is the state of cell  $Y$  at the previous time  $t$ ;
- (iii)  $\mathbf{x}_t^{(k)}$  is the vector  $(x_{t-k+1}, \dots, x_t)$  of the previous  $k$  states of cell  $X$ ;
- (iv)  $\mathcal{S}$  represents the set of all possible patterns of sets of states  $(\mathbf{x}_t^{(k)}, x_{t+1}, y_t)$ .

To calculate TE, one must start with the time series of states of a CA. Given the time series (see Figure 1), the occurrence numbers of patterns of states  $(\mathbf{x}_t^{(k)}, x_{t+1}, y_t)$  are counted for each combination of cells  $X$  and  $Y$  in the array. Once normalized to the total number of (not necessarily distinct) patterns appearing in the time series, these frequencies are identified with the probabilities  $p(\mathbf{x}_t^{(k)}, x_{t+1}, y_t)$  appearing in (1). The conditional probabilities  $p(x_{t+1} | \mathbf{x}_t^{(k)}, y_t)$  and  $p(x_{t+1} | \mathbf{x}_t^{(k)})$  are calculated analogously.

Changing the value of the history length can affect the value of TE calculated for the same time series data. We find very similar values of TE for  $k = 4$  and  $6$  and observe a considerable decrease in TE for  $k < 4$  and  $k > 6$  (cf. Section 4.2 of [32]), relative to the value of TE for  $k = 5$ . We therefore consider  $k = 5$  as the optimal value of  $k$  that properly captures the past history of  $X$  for the results presented here (see Section 3 for a visual explanation of why  $k \sim 5$  represents the optimal history length).

We first generated time series for the simplest initial state for each of the 256 possible ECA rules, which we numerically label following Wolfram’s numbering scheme [33, 34]. Here and in the following, periodic boundary conditions are

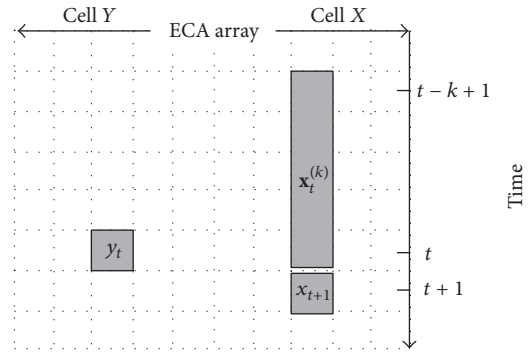


FIGURE 1: Time series of an ECA, with the pattern of cells relevant to the evaluation of the transfer entropy from  $Y$  to  $X$ ,  $T_{Y \rightarrow X}$ , highlighted.

enforced. Therefore, the specific location of the different color cell in the input array is irrelevant. Our initial state is either

$$\square \cdots \square \blacksquare \square \cdots \square \quad (2)$$

50-times    50-times

or the equivalent state obtained through a  $\square \leftrightarrow \blacksquare$  conjugation. For example, rules 18 and 183 are equivalent under a  $\square \leftrightarrow \blacksquare$  conjugation. The input  $\square \cdots \square \blacksquare \square \cdots \square$  is updated using rule 18, and the conjugated input  $\blacksquare \cdots \blacksquare \square \blacksquare \cdots \blacksquare$  using rule 183. A comprehensive study of the symmetry properties of ECA can be found in [11], where the 256 possible ECA rules are grouped into 88 different equivalence classes. Each class contains no more than four rules, which show the same behavior under  $\square \leftrightarrow \blacksquare$  exchange, left-right transformations, or the combination of both transformations. The interested reader can find such classification summarized in Table 1 of the cited work. For Wolfram Classes III and IV (Wolfram’s most recent ECA classification scheme is now implemented in the *Wolfram Alpha* computational engine [28] and reproduced here (Table 1) for the convenience of the reader), the equivalence classes are also shown in the legend of Figure 2.

We calculate the amount of *information processed by a CA* in a space-time patch as the average of  $T_{Y \rightarrow X}$  calculated over that region, designated as  $\langle \text{TE} \rangle$ . To do so, we evolve the CA for 250 time steps and then remove the first 50 time steps from each generated time series. This allows us to evaluate TE over a region of the CA in causal contact with the initial input at each point in space and time. This ensures that the observed dynamics over the relevant space-time patch are driven by both the chosen rule *and* the initial state. The resulting time series is then used to evaluate the  $101^2$  values  $\text{TE}_{Y \rightarrow X}$ . For each rule, the average value  $\langle \text{TE} \rangle$  is shown in Figure 2(a). Equivalent rules, like rules 30, 86, 135, and 149, produce the same values of  $\langle \text{TE} \rangle$  and are represented by a common marker. For clarity, individual Wolfram Classes I and II rules are not shown and are instead replaced by a dashed line corresponding to the highest  $\langle \text{TE} \rangle$  of any individual Class I or II rule. All Class III rules lie above the range of Classes I and II. Interestingly, with only the exception of rule 110 and its equivalent rules, all Class IV rules lie within the range of Classes I and II rules.

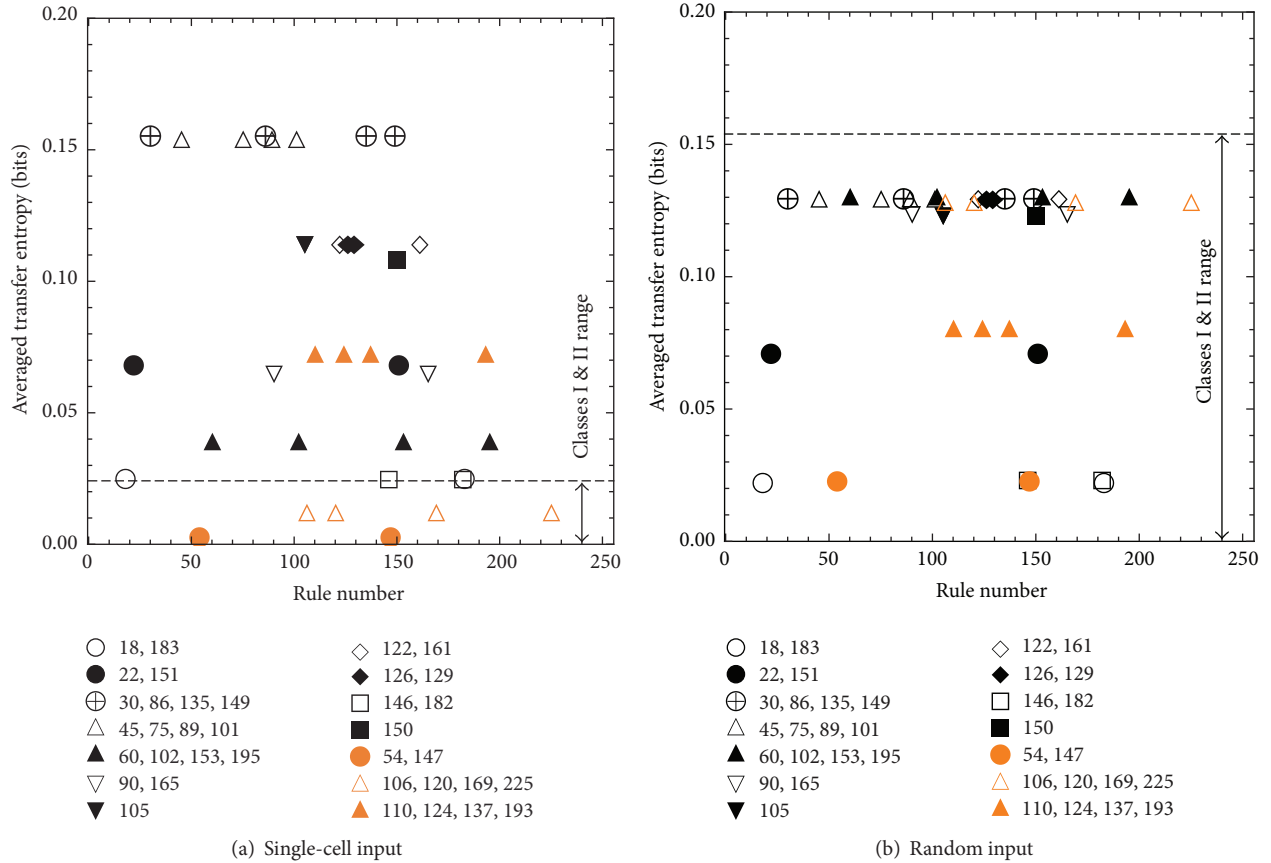


FIGURE 2:  $\langle TE \rangle$  for each rule, shown for two different inputs: single cell (a) (either one black cell or one white cell) and random (b). Equivalent rules are represented by a common marker. The association of equivalent inputs with equivalent rules is chosen heuristically as the one that maximizes the averaged TE. The dashed line corresponds to the highest value of  $\langle TE \rangle$  for any Wolfram Class I or II rule. Individual markers are shown only for Wolfram Classes III (black) and IV (orange) rules.

The single-cell input considered here is extremely rare within the space of all possible initial inputs. The number of black cells in a state randomly extracted among the  $2^{101}$  different possible inputs follows a binomial distribution, meaning that states containing 50 or 51 black and white cells are about  $2 \times 10^{27}$  times more likely than our initially selected input. Our motivation for considering the single-cell input first is that it automatically excludes many trivial rules from our search for the complex ones. Rules that duplicate the input in a trivial way or annihilate it to produce all  $\blacksquare$  or all  $\square$  naturally yield  $\langle TE \rangle \approx 0$ . This is the same approach that has been recently assumed in algorithmic complexity based classification of CAs [22]. It has the advantage of selecting many rules according to their intrinsic complexity, and not the complexity of the input. However, a major shortcoming of choosing the single-cell input is that many Class IV rules look simpler than what they truly are. Class IV rule 106 is a good example. For the simple input of a single black bit, rule 106 functions to shift the black cell one bit to the left at each time-step, generating a trivial trajectory. However, in cases where the input allows two black cells to interact, the full potential of rule 106 can be expressed. This is an explicit example of the sensitivity of the behavior of some ECA rules

to the complexity of their input, as discussed in the introduction. To isolate the complexity of the rule from that inherited by the input, we must therefore consider a random sample of initial states, as we do next.

### 3. Inherited Complexity and Information-Based Classification

We next consider a more generic input, randomly selected among the  $2^{101} \approx 2.5 \times 10^{30}$  different possibilities. As the input is no longer symmetric, we now need to consider reflections in selecting the equivalent input  $\rightarrow$  rule associations. The scenario changes completely in this case, as shown in Figure 2(b), where the highest values of  $\langle TE \rangle$  correspond to Class II rules, including 15, 85, 170, and 240, which generate trivial, wave-like behaviors. These rules behave like rule 106 (especially 170 and 240) when initialized with a single-cell input, but they do not contribute any new, emergent nontrivial features when nearby cells interact. For all purposes they appear *less* intrinsically complex. With only the exception of rule 110 and its equivalent rules, Class IV rules behave like many rules of Class II and exhibit a large increase in complexity, as qualitatively observed and also

as quantitatively captured by  $\langle TE \rangle$ , in response to a more complex input.

It is worth noting that all of the rules whose initial value of  $\langle TE \rangle$  lay above the upper limit for Classes I and II of  $\langle TE \rangle = 0.024$  bits in the simplest input scenario still have calculated  $\langle TE \rangle$  above this value under a change of input. In particular, the rules with the highest values of  $\langle TE \rangle$  are not significantly affected by the change in the input. Let us naively use the upper limit for Classes I and II rules emerging from Figure 2(a) as the border line between what we call *low* and *high* values of  $\langle TE \rangle$ . We can summarize the changes under varying the input as follows. There are rules whose value of  $\langle TE \rangle$  is low for both the inputs. There are rules whose value of  $\langle TE \rangle$  is low in the simplest case, but high in the more complex one. And, finally there are also rules with a high value of  $\langle TE \rangle$  in both cases. The interesting point is that *we find no rule whose value of  $\langle TE \rangle$  is high for the simplest input and low for the random input*. This is equivalent to say that there are no rules generating complexity for simple inputs, but annihilating it for complex ones. The significance of this observation lies in the fact that it enables classification in terms of the shift of  $\langle TE \rangle$  over a space-time region in response to a change in its input by taking advantage of the main limitation that makes quantitative classifications of ECAs so difficult, that is, ECA sensitivity to their initial state, and exploits it in order to achieve such classification.

To confirm this is indeed a viable method for quantitative classification, we must consider more than just two inputs. We therefore first randomly selected twenty different ones. Being interested in how much  $\langle TE \rangle$  can vary as we vary the input, for each rule we selected the maximum absolute value of the percent change of  $\langle TE \rangle$  between the random inputs ( $\langle TE \rangle_r$ ) and the single-cell input ( $\langle TE \rangle_1$ ) considered before:

$$\max \left\{ \frac{|\langle TE \rangle_r - \langle TE \rangle_1|}{\langle TE \rangle_r} \right\}_{\text{random inputs}} \quad (3)$$

The results are shown in Figure 3, where the maximum percent change is plotted as a function of  $\langle TE \rangle_1$ . Equivalent rules are represented by a single marker and share the same value for the maximum relative change of  $\langle TE \rangle$ . In fact, given two rules  $r_1$  and  $r_2$  sharing a symmetry, it is always possible to find two inputs  $i_1$  and  $i_2$  such that rule  $r_1$  initialized to  $i_1$  and rule  $r_2$  initialized to  $i_2$  give rise to the same value of  $\langle TE \rangle$ .

The horizontal dashed line separates rules whose maximum  $\langle TE \rangle$  change is less than one order of magnitude from those that can undergo a change in  $\langle TE \rangle$  of at least a factor 10. The vertical dashed line denotes the highest value of  $\langle TE \rangle_1$  for Wolfram Classes I and II, exactly as in Figure 2(a). The region to the right of the vertical line and above the horizontal line is void of any ECA rules. Points in that region would correspond to values of  $\langle TE \rangle$  that are both high for the simplest input and capable of high variation, for example, CA rules that can annihilate the complexity of the input, which we do not observe. This feature yields a distinctive L-shape in the distribution of rules with the rules in each region sharing well-defined properties.

As a result, a natural, information-based classification of ECAs can be given as follows:

*Class I<sub>1</sub>*:  $\langle TE \rangle$  is very small for the simplest input and stays so for complex inputs. This is the most populated class, including almost all Wolfram Classes I and II rules, rules 18 and 146 and their equivalent Class III rules.

*Class I<sub>2</sub>*:  $\langle TE \rangle$  is small for the simplest input, but it experiences a drastic change (one order of magnitude or more) when the input is complex. This is the case for many Wolfram Class II and some Class IV rules (e.g., 54 and 106 and their equivalent rules).

*Class I<sub>3</sub>*:  $\langle TE \rangle$  has a high value for the simplest input, and this value is approximately unaffected by a change in the input. Most Wolfram Class III rules belong to this class, as well as Class IV rule 110 and its equivalent rules.

Our classification is summarized in Table 2. Despite the arbitrary placement of our boundary lines at the border of the  $I_1$  region, it shows an unambiguous separation between the most majority of rules classified as  $I_2$  and  $I_3$ .

Randomly sampling inputs leads to a bias favoring nearly fifty-fifty distributions of black and white cells, due to their binomial distribution. We therefore also verified this classification using a different distribution of inputs, where the number of black cells is increased in a regular way from 2 to 50, while keeping the specific positions of black cells in the input array random. We considered 20 different inputs, each containing exactly 2, 5, 7, 10, 12, ..., 47, and 50 black cells (higher numbers are not considered due to the  $\square \leftrightarrow \blacksquare$  conjugation). Apart from minor shifts of the data points, applying the same procedure as above yields exactly the same classification as Figure 3, indicating that our classification scheme is robust.

We stress the importance of considering a *large* system (order  $\sim 100$  cells) as opposed to a much smaller one (e.g., order  $\sim 10$ ). While for the latter a scan over the entire space of inputs is computationally feasible (and indeed we performed these experiments), it hides one of the main features enabling information-based classification, the existence of Class  $I_3$  rules, which form the most stable class with respect to our TE based complexity measure. Class  $I_3$  rules produce time series largely independent of the initial state, exemplified by rule 30 in Figure 3, a feature evident only in larger systems (being the result of combinatorial calculus, TE computational time grows exponentially with the number of cells considered. An ECA array of about 100 cells is both the smallest size that we could safely consider much larger than that of the domains of Figure 3 and for which a complete scan over the rule space and a large enough ensemble of inputs could be made. We also considered individual runs with 200 and 400 cells and noticed that our numerical results are stable for CA with 100 or more cells but are fluctuating for CA with 50 or less cells).

The typical, order 10, cell length of the domains appearing in Figure 3 is also the reason why  $k \sim 5$  is the most optimal value for the history length parameter. Smaller and larger values of  $k$  would not be able to resolve this lattice size independent feature of Class  $I_3$  rules [32] (Figure 4).

TABLE 2: Information based classification of ECA, with only the 88 rule equivalence classes shown (represented by the lowest number rule). Wolfram Class III rules are denoted in black, boldface type, and Class IV in italic.

Class I <sub>1</sub>									
0	1	3	4	5	8	9	12	13	<b>18</b>
19	23	25	26	28	29	32	33	35	36
37	40	41	44	50	51	57	58	62	72
73	76	77	78	94	104	108	128	130	132
134	136	140	<b>146</b>	152	154	156	160	162	164
178	200	204	232						
Class I <sub>2</sub>									
2	6	7	10	11	14	15	24	27	34
38	42	43	46	54	56	74	106	138	142
168	170	172	184						
Class I <sub>3</sub>									
<b>22</b>	<b>30</b>	<b>45</b>	<b>60</b>	<b>90</b>	<b>105</b>	<i>110</i>	<b>122</b>	<b>126</b>	<b>150</b>

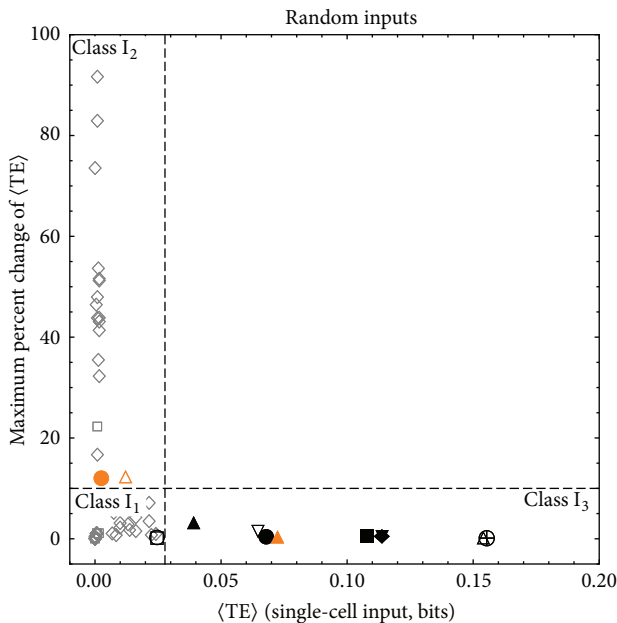


FIGURE 3: Maximum percent change of  $\langle TE \rangle$  as a function of  $\langle TE \rangle_1$  (3). Equivalent rules are represented by a single marker, colored as in Figure 2 but with Wolfram Classes I and II in gray. Rules above the horizontal dashed line can undergo a change of at least one order of magnitude in  $\langle TE \rangle$ . The vertical dashed line is the same as in Figure 2(a).

#### 4. Coarse-Graining and the Information Hierarchy

Perhaps the most interesting feature of our quantitative classification is that Wolfram Class III and Class IV rules are distributed over different information-based classes. This behavior looks less surprising in the light of the coarse-graining transitions among ECA uncovered by Israeli and Goldenfeld [1]. One important aspect of the physical world is that coarse-grained descriptions often suffice to make

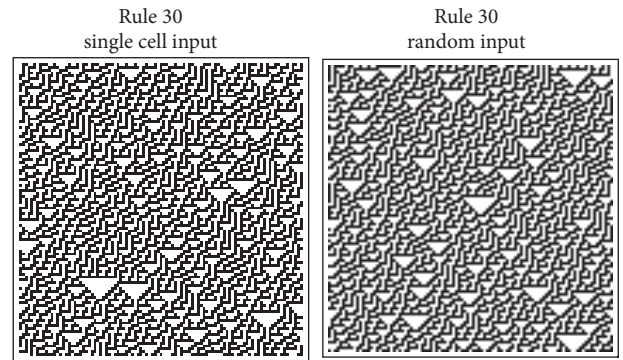


FIGURE 4: Time series for rule 30 generated by the two initial states of Figure 2. The similarity of the two outputs, each yielding color domains with a typical linear size of about  $\sim 10$  blocks, does not depend on the size of the CA and is responsible for the stability of Class I<sub>3</sub> rules to changes in input.

predictions about relevant behavior. Noting this, Israeli and Goldenfeld adopted a systematic procedure to try to coarse grain ECA rules. Their prescription consists in grouping together nearby cells in the CA array into a supercell of a given size  $N$ , according to some specified supercell  $\rightarrow$  single-cell correspondence (for Boolean CAs,  $2^{2^N}$  possible applications of this kind exist), and a given time constant  $T$ . The search for a coarse-graining rule consists in looking for a new CA rule such that running the original CA for  $Tt$  time steps, and then grouping the supercells, produces the same output as grouping the initial array and then adopting the new CA rule for  $t$  time steps. The new CA rule, the time constant  $T$ , and the size of the supercell  $N$  are all unknown; therefore, this search requires a scan over all the possible combinations of these parameters. Israeli and Goldenfeld successfully coarse-grained 240 of the 256 ECA rules, many to other ECA rules, performing a complete scan of all the possibilities compatible with  $N \leq 4$ . Importantly, the rule complexity was never observed to increase under coarse-graining, introducing a partial ordering among CA rules.

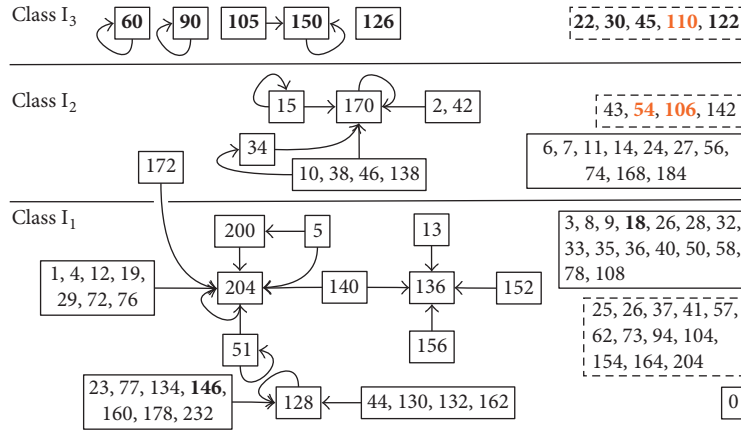


FIGURE 5: Coarse-graining transitions within ECA found in [1], mapped onto a hierarchy of the information-based classes identified here. Only the 88 rule equivalency classes are shown (represented by the lowest number rule). For simplicity, (nontrivial) transitions to rule zero are not shown (isolated boxed with continuous border). Boxes with a dashed border include ECA rules for which a coarse-graining transition to another ECA rule with supercell of size 4 or less does not exist [1].

The same ordering emerges from our information-based classification, as shown in Figure 5, where arrows indicate coarse-graining transitions uncovered in [1]. These transitions introduce a fully ordered hierarchy  $I_3 \rightarrow I_2 \rightarrow I_1$  such that coarse-graining is never observed to move up the hierarchy, and the vast majority of rules may only undergo coarse-graining transitions within the *same information class*. This ordering is sometimes expected, like in the case of many of our Class  $I_2$  rules, where wave-like patterns are coarse-grained to either self-similar, or very similar, patterns. Other times it looks more profound, like in the case of Wolfram Class III rule 146. Rule 146 is in  $I_1$ , and it can be nontrivially coarse-grained to  $I_1$  rule 128 (a Wolfram Class I rule), due to a shared sensitivity of the information processed in a given space-time patch to its input state. As noted in [1], the fact that potentially (a conclusive proof of the computational irreducibility of ECA rule 146 is still missing) computationally irreducible rules [8] like elementary rule 146 can be coarse-grained to predictable dynamics (e.g., rule 128) shows that computational irreducibility might lack the characteristics of a good measure of physical complexity. We are led to conclude that the coarse-graining hierarchy is more likely defined by conserved informational properties, with more complex rules by Wolfram’s classification appearing lower in the hierarchy if they can be coarse-grained to less complex ones with common sensitivity to the input state.

## 5. Conclusions

Physical systems evolve in time according to local, uniform rules but can nonetheless exhibit complex, emergent behavior. Complexity arises either as a result of the initial state or rule, or some combination of the two. The ambiguity of determining whether complexity is generated by rules or states has confounded previous efforts to isolate complexity *intrinsic* to a given rule. In this work, we introduced a quantitative, information-based, classification of ECA. The classification scheme proposed circumvents the difficulties arising due to

the sensitivity of the expressed complexity of ECA rules to their initial state. The (averaged) directed exchange of information (TE) between the individual parts of an ECA is used as a measure of its complexity. The identification of the single-cell input (Section 2) as the nontrivial state with the least complexity is assumed as a working hypothesis and provides a reference point for our analysis of the degree of variability of the complexity of ECA rules for varying inputs. We identified three distinct classes based on our analysis, which vary in their sensitivity to initial conditions. Class  $I_1$  ECA always process little information, Class  $I_3$  always processes high information, and Class  $I_2$  can be low or high depending on the input. It is only for Class  $I_3$  that the expressed complexity is intrinsic and not a product of the complexity carried by the input.

The most complex rules by our analysis are in Class  $I_3$ , which includes the majority of Wolfram’s Class III rules and Class IV rule 110 and its equivalent rules. These rules form a closed group under the coarse-graining transitions found in [1]. The truly complex rules are those that remain complex even at a macrolevel of description, with behavior that is not sensitive to the initial state.

We are currently investigating the possibility of extending our measure to less elementary dynamical systems, as well as to idealized biological systems (e.g., networks of regulatory genes). Inspired by Nurse’s idea [35] that the optimization of information storage and flow might represent the distinctive feature of biological functions, we find interesting that information dynamics identifies the most complex ECA dynamics with the ones that least depend on their initial state. This feature, evocative of the redundancy so ubiquitous in biology, might represent a link between biology and information dynamics [36] that we consider worth of further study.

## Disclosure

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of Templeton World Charity Foundation.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

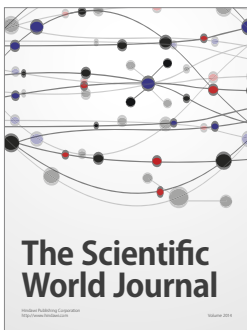
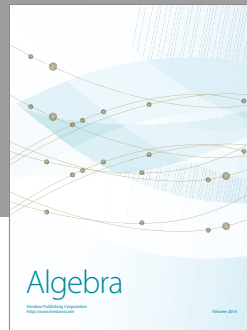
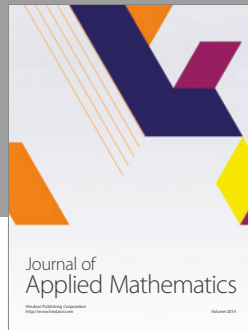
## Acknowledgments

This publication was made possible through support of a grant from Templeton World Charity Foundation. The authors thank Douglas Moore and Hector Zenil for their helpful comments.

## References

- [1] N. Israeli and N. Goldenfeld, "Coarse-graining of cellular automata, emergence, and the predictability of complex systems," *Physical Review E. Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 2, 026203, 17 pages, 2006.
- [2] M. Gardner, "Mathematical games," *Scientific American*, vol. 223, pp. 120–123, 1970.
- [3] E. R. Berlekamp, J. H. Conway, and R. K. Guy, *Winning Ways for Your Mathematical Plays, Vol. 2: Games in Particular*, vol. 2, Academic Press, London, UK, 1982.
- [4] M. Gardner, *Wheels, Life, and other Mathematical Amusements*, Freeman & Company, NY, USA, 1983.
- [5] I. Smith, "Simple computation-universal cellular spaces," *Journal of the Association for Computing Machinery*, vol. 18, pp. 339–353, 1971.
- [6] K. Lindgren and M. G. Nordahl, *Complex Systems*, vol. 4, pp. 299–318, 1990.
- [7] M. Cook, "Universality in elementary cellular automata," *Complex Systems*, vol. 15, no. 1, pp. 1–40, 2004.
- [8] S. Wolfram, "Cellular automata as models of complexity," *Nature*, vol. 311, no. 5985, pp. 419–424, 1984.
- [9] J. von Neumann, *Theory of Self-Reproducing Automata*, A. W. Burks University of Illinois Press, Urbana, ILL, USA, 1966.
- [10] K. Culik and S. Yu, "Undecidability of CA classification schemes," *Complex Systems*, vol. 2, pp. 177–190, 1988.
- [11] W. Li and N. Packard, "Erratum: "The structure of the elementary cellular automata rule space";" *Complex Systems*, vol. 4, no. 1, pp. 281–297, 1990.
- [12] C. G. Langton, "Computation at the edge of chaos: phase transitions and emergent computation," *Physica D: Nonlinear Phenomena*, vol. 42, no. 1–3, pp. 12–37, 1990.
- [13] F. Bagnoli, R. Rechtman, and S. Ruffo, "Damage spreading and Lyapunov exponents in cellular automata," *Physics Letters A*, vol. 172, no. 1–2, pp. 34–38, 1992.
- [14] G. Y. Vichniac, "Boolean derivatives on cellular automata," *Physica D. Nonlinear Phenomena*, vol. 45, no. 1–3, pp. 63–74, 1990.
- [15] J. M. Baetens and B. De Baets, "Phenomenological study of irregular cellular automata based on Lyapunov exponents and Jacobians," *Chaos. An Interdisciplinary Journal of Nonlinear Science*, vol. 20, no. 3, Article ID 033112, 033112, 15 pages, 2010.
- [16] J. M. Baetens, P. Van Der Weeën, and B. De Baets, "Effect of asynchronous updating on the stability of cellular automata," *Chaos, Solitons and Fractals*, vol. 45, no. 4, pp. 383–394, 2012.
- [17] J. M. Baetens and B. De Baets, "Topology-induced phase transitions in totalistic cellular automata," *Physica D. Nonlinear Phenomena*, vol. 249, pp. 16–24, 2013.
- [18] J. M. Baetens and J. Gravner, "Stability of cellular automata trajectories revisited: branching walks and Lyapunov profiles," *Journal of Nonlinear Science*, vol. 26, no. 5, pp. 1329–1367, 2016.
- [19] L. O. Chua, S. Yoon, and R. Dogaru, "A nonlinear dynamics perspective of Wolfram's new kind of science. I. Threshold of complexity," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 12, no. 12, pp. 2655–2766, 2002.
- [20] N. Fatés, "Experimental study of elementary cellular automata dynamics using the density parameter," *Discrete Mathematics and Theoretical Computer Science*, pp. 15–166, 2003.
- [21] C. Dürr, I. Rapaport, and G. Theyssier, *Theoretical Computer Science*, vol. 322, pp. 355–368, 2004.
- [22] H. Zenil and E. Villarreal-Zapata, "Asymptotic behavior and ratios of complexity in cellular automata," *International Journal of Bifurcation and Chaos*, vol. 23, no. 9, Article ID 1350159, 2013.
- [23] L. Albantakis and G. Tononi, "The intrinsic cause-effect power of discrete dynamical systems—from elementary cellular automata to adapting animats," *Entropy*, vol. 17, no. 8, pp. 5472–5502, 2015.
- [24] G. J. Martínez, "A note on elementary cellular automata classification," *Journal of Cellular Automata*, vol. 8, no. 3–4, pp. 233–259, 2013.
- [25] W. Bialek, *Biophysics: searching for principles*, Princeton University Press, 2012.
- [26] S. I. Walker, H. Kim, and P. C. W. Davies, "The informational architecture of the cell," *Philosophical Transactions of the Royal Society A*, vol. 374, Article ID 20150057, 2016.
- [27] T. Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, 2000.
- [28] Wolfram Alpha, <https://www.wolframalpha.com>.
- [29] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, "The information dynamics of phase transitions in random boolean networks," in *Proceedings of the 11th International Conference on the Simulation and Synthesis of Living Systems: Artificial Life XI, ALIFE 2008*, pp. 374–381, 2008.
- [30] J. T. Lizier and M. Prokopenko, *The European Physical Journal B*, vol. 73, 2010.
- [31] X. R. Wang, J. M. Miller, J. T. Lizier, M. Prokopenko, and L. F. Rossi, "Measuring information storage and transfer in swarms," in *Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*, pp. 838–845, 2011.
- [32] T. Bossomaier, L. Barnett, M. Harré, and J. T. Lizier, *An Introduction to Transfer Entropy: Information Flow in Complex Systems*, Springer, 2016.
- [33] S. Wolfram, "Statistical mechanics of cellular automata," *Reviews of Modern Physics*, vol. 55, no. 3, pp. 601–644, 1983.
- [34] S. Wolfram, *A New Kind of Science*, Wolfram Media, Champaign, ILL, USA, 2002.
- [35] P. Nurse, "Life, logic and information," *Nature*, vol. 454, no. 7203, pp. 424–426, 2008.
- [36] P. C. W. Davies and S. I. Walker, "The hidden simplicity of biology," *Reports on Progress in Physics*, vol. 79, no. 10, Article ID 102601, 2016.





# Hindawi

Submit your manuscripts at  
<https://www.hindawi.com>

