2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

Gardens of Eden: where Nonlinear Dynamics and Formal Languages meet

Giovanni E. Pazienza† and Marion Oswald‡

†Cellular Sensory and Wave Computing Laboratory, MTA - SZTAKI, Budapest, Hungary
and Pazmany University, Budapest, Hungary
‡Cellular Sensory and Wave Computing Laboratory, MTA - SZTAKI, Budapest, Hungary
and Vienna University of Technology, Vienna, Austria
Email: gpazienza@sztaki.hu, marion@sztaki.hu

Abstract—Many classifications of the Cellular Automata rules have been proposed in the last decades. Some of them are based on empirical observations, and hence their rigor is doubtful, while others are based on precise, though often complicated, mathematical methods. In this paper, we present two new classifications for Cellular Automata rules, which come from two different disciplines; namely, Nonlinear Dynamics and Formal Language Theory. These classifications are straightforward to obtain and give remarkably similar results.

1. Introduction

Cellular Automata (CA) have been thoroughly studied in the last decades, but they are far from being completely characterized. In the last few years, CA have been seen "through the eyes" of Nonlinear Dynamics [1]; this new perspective led, among other results, to a classification of CA rules into six groups, depending on their dynamics. In this paper, we propose a variant of such classification. We were inspired by a classical work [2] in Nonlinear Dynamics that proofs that in a continuous function defined over an interval the presence of a period-3 orbit implies the presence of orbits with any other period. Obviously, this result does not apply directly to Cellular Automata because, in general, they describe a discontinuous function. Nevertheless, we noticed that all rules in the last two groups of Chua's classification have at least one periodic orbit with $\tau = 3$ (the notion of 'periodic orbit' and the meaning of the parameter τ will be explained in Sec. 2.2) and, at the same time, all rules with $\tau = 3$ included in the remaining four groups of Chua's classification tend to have a more complex behavior than the others [3]. For this reason, in our classification we distinguish only two groups: i) rules having exclusively periodic orbits with $\tau = 1$ and/or $\tau = 2$; ii) rules having periodic orbits with $\tau \geq 3$.
Also, we propose yet another classification of the CA local rules, based on the formal languages formed by the so-called *orphan patterns*. In this paper, we will show that there is an excellent correspondence between these two classifications and discuss its implications.
The paper is structured as follows: in Sec. 2, we discuss the basic features of CA and our classification of the CA rules based on the properties of their periodic orbits; in Sec. 3, we introduce a few fundamental notions of Formal Language Theory; in Sec. 4, we define the concepts of *Garden of Eden* and *orphan pattern*, and show how they can be used to classify CA rules; in Sec. 5, we present the comparison between our two classifications; in Sec. 6, we draw the conclusions and outline our future work.

2. Brief notes on Cellular Automata

2.1. Generalities

Cellular Automata [4] consist of regular uniform lattice of cells assuming a finite number of states; here, we consider one-dimensional CA in which cells are arranged in an array of length $L = I + 1$ and can take only two states: 0 and 1. For instance, a bit string $\mathbf{x}$ at the generic time step n is

$$\mathbf{x}^n = (x_0^n x_1^n \dots x_{I-1}^n x_I^n) \tag{1}$$

where the subscript indicates the position of the cell in the array. Hereafter, letters in bold indicate bits strings, and letters in italics are used for the single bits.
Cells are updated synchronously and the time evolution of a bit string can be effectively summarized by the notation

$$\mathbf{x}^{n+1} = f(\mathbf{x}^n) \tag{2}$$

in which the superscript indicates the iteration. The state of each cell at iteration $n+1$ depends on the states of its neighbors (here we consider only the nearest neighbors) at iteration n

$$x_i^{n+1} = f(x_{i-1}^n x_i^n x_{i+1}^n). \tag{3}$$

In the following, we use periodic boundary conditions, which means that

$$x_0^{n+1} = f(x_I^n x_0^n x_1^n) \quad \text{and} \quad x_I^{n+1} = f(x_{I-1}^n x_I^n x_0^n) \tag{4}$$

Under the restrictions detailed above, there are only 256 possible functions f, called rules, which we can be denoted by f_0 up to f_{255}. For instance, the notation:

$$\mathbf{x}^{n+2} = f_{110}(f_{110}(\mathbf{x}^n)) \tag{5}$$

indicates the application of rule 110 to the bit string $\mathbf{x}^n$ two times to obtain the bit string $\mathbf{x}^{n+2}$. However, the 256 local rules show only 88 distinct dynamics, since each CA rule can have up to four *globally-independent* rules, obtained by complementing and/or mirroring the original one, as described in [1]. Therefore, from now on we will often refer to the 88 globally-independent Cellular Automata rules, making it clear that the choice of the representative specimen for each of the 88 equivalence classes is arbitrary.

2.2. Complexity index and Periodic orbits

We can associate an *index of complexity* κ to all CA local rules thanks to the procedure defined in [1]. There are 104 rules (38 of which are globally-independent) with $\kappa = 1$, 126 rules (41 of which are globally-independent) with $\kappa = 2$, and 26 rules (9 of which are globally-independent) with $\kappa = 3$. Observe that the universal rule 110 (and its three globally-equivalent rules 124, 137, and 193) has $\kappa = 2$. Therefore, the dynamics of rules with $\kappa = 3$ is not necessarily more complex than those with $\kappa = 2$.
If the functions f_i are deterministic and the length L of the bit string is finite, then the evolution of an arbitrary initial state under an arbitrary rule f_N will end up in a periodic cycle, in the sense that there exist p and T such that

$$\mathbf{x}^p = \mathbf{x}^{p+T} \tag{6}$$

Obviously, $\mathbf{x}^m = \mathbf{x}^{m+T}$, for all $m > p$. The bit strings from $\mathbf{x}^0$ to $\mathbf{x}^{p-1}$ are said to belong to the *transient*, which has length p, while the bit strings from $\mathbf{x}^p$ on are said to belong to the *periodic orbit*, which has length T. Furthermore, it may happen that the there exists a τ for which:

$$\mathbf{x}^p = \mathcal{S}^{\sigma}(\mathbf{x}^{p+\tau}) \tag{7}$$

where $\mathcal{S}^{\sigma}$ indicates a shift, left or right, by σ positions. Therefore, a periodic orbit can be characterized by its parameters $\tau > 0$ and $\sigma \geq 0$.
Among the 88 globally-independent 1D binary Cellular Automata rules, some are particularly relevant, such as rule 110 which has been proved to be universal, while others do not perform any interesting task, such as rule 0 because all possible initial bit strings evolve into $(0 \ldots 00)$ in one iteration. In order to group the rules according to the properties of their periodic orbits, both Wolfram and Chua proposed a classification according to the dynamics obtained by using a long (L greater than 100) random bit string as initial state of the Cellular Automaton.
In this paper, we use a variant of Chua's classification that, instead of using a *long* random bit string as initial state, uses all bit strings with length L, where $L \leq 15$, as initial states checking whether there is any periodic orbit with $\tau \geq 3$. In fact, as already pointed out by [5] and [6], there are some rules for which 'interesting' periodic orbits exist, but they cannot be found through the method used by Wolfram and Chua because the probability of obtaining them from random initial bit string is extremely low.

3. Formal Languages

In this section, we introduce a few basic concepts on formal languages, which will be used in the following; more details can be found in the classical literature, e.g., [7].
An *alphabet* Σ is a non-empty set of symbols. A finite sequence of symbols from Σ is called a *word* and its *length* is the number of symbols it contains. The word with length 0 is called the *empty word* and denoted by λ. The set of all words over an alphabet Σ (including the empty word) is denoted by Σ^*, while the set of nonempty words is denoted by Σ^+, i.e., $\Sigma^* = \Sigma^+ \cup \{\lambda\}$. Any set of words $\mathcal{L} \subseteq \Sigma^*$ is called a (formal) *language*. A formal language is called *finite* if it contains only a finite number of words. An infinite language in which the number of words with a given length l is bounded by a constant independent of l, for all $l \in \mathbb{N}$, is called *slender*. For a word $w \in \Sigma^*$, where $w = xuy$ for $x, u, y \in \Sigma^*$, we call x *prefix*, u *subword* and y *suffix* of w. For a language $\mathcal{L}$, the set $\overline{\mathcal{L}} = \Sigma^* - \mathcal{L}$ is called the *complement* of $\mathcal{L}$ containing all the words over an alphabet Σ that are not contained in $\mathcal{L}$.
In the following, we will be mainly concerned with *factorial* languages: a language $\mathcal{L}$ is called factorial if for every word $w \in \mathcal{L}$ all the subwords of w also belong to $\mathcal{L}$. Then any word $w \in \Sigma^+$ is called *distinct excluded block* (DEB, for short) or *forbidden word*, if $w \notin \mathcal{L}$ but every proper substring of w is in $\mathcal{L}$. Let $\mathcal{L}'$ be the set of distinct excluded blocks of a language $\mathcal{L}$, then we can write $\overline{\mathcal{L}} = \Sigma^* \mathcal{L}' \Sigma^*$ and $\mathcal{L} = \Sigma^* - \Sigma^* \mathcal{L}' \Sigma^*$. Hence a factorial language can also be defined by $\mathcal{L}'$, sometimes called *the antidictionary*. It is worth noting that $\mathcal{L}'$ eliminates far more words from $\mathcal{L}$ then only the ones contained in $\mathcal{L}'$ (namely, all those having a DEB as proper substring).

4. Gardens of Eden and orphans patterns

For most of the rules, there are some strings, called *Gardens of Eden*, that cannot be generated through the evolution of a given rule. In other words, the bit string $\mathbf{x}$ is said to be a Garden of Eden of rule N if there is no $\mathbf{x}'$ such that $\mathbf{x} = f_N(\mathbf{x}')$.
In this paper, we consider the set of the Gardens of Eden of length L that do not contain any Garden of Eden of length $L-1$ as proper substring, which are also called *orphan patterns*. Such set corresponds with the set $\mathcal{L}'$ of DEB of the factorial language $\mathcal{L}$, where $\mathcal{L}$ is the language consisting of all bit strings of arbitrary sizes generated in one iteration by the rule N. We then analyzed the language $\mathcal{L}'$ of DEB or orphans (see Sec. 3) for all 88 globally-independent rules, and grouped them according to their formal language theoretic properties:

1. $\mathcal{L}'$ is *empty*, i.e., there are no orphans patterns;
2. $\mathcal{L}'$ is *finite*, i.e., there is only a finite number of orphan patterns;

3. $\mathcal{L}'$ is *slender*, i.e., there is an infinite number of orphan patterns but their number for each given string length L is bounded by a constant independent of L;

4. $\mathcal{L}'$ is *non-slender*, i.e., there is an infinite number of orphan patterns and their number for each given string length L is growing with L.

5. Results

First of all, we classified the 88 globally-independent rules into the two groups, as defined in Sec. 1. The first group includes all rules having only $\tau = 1$ and/or $\tau = 2$, while the second group includes all rules having also $\tau \geq 3$. Second, we classified the same 88 globally-independent rules into the four classes defined in Sec. 4. The correspondence between these two classifications is very good, as it can be appreciated in Tables 1, 2, and 3.
In particular, from Table 1 we can observe that the orphan patterns of all rules having only orbits with $\tau = 1$ and/or $\tau = 2$ constitute either finite or slender languages. Furthermore, the first case corresponds to all rules with $\kappa = 1$. This is consistent with Chua's definition of complexity index: the threshold of complexity beyond which rules can exhibit 'interesting' dynamics, and whose orphan patterns can form non-finite languages, is 2.

From Table 2 we notice that all CA rules having orbits with $\tau \geq 3$ constitute (infinite) non-slender languages. This implies that the number of words with a given length L is growing with L.
The third and last group is formed by all rules having no orphans. Such rules are *surjective* when L is infinite, in the sense that given a rule N, for any bit string $\mathbf{x}$ there will be another bit string $\mathbf{x}'$ such that $\mathbf{x} = f_N(\mathbf{x}')$. Nevertheless, this is not true when L is finite. In this last case, some rules – namely, 15, 51, 170, and 204 – are still surjective, while other rules – namely, 30, 45, 60, 90, 105, 106, 150, 154 – have Gardens of Eden. From Table 2 we can see that also this characteristic links well the two classifications.
Unfortunately, there are six exceptions to such a well-built formal construction. On the one hand, the languages defined by the orphan patterns of rules 104, 134, and 152 seem to be non-slender but none of these rules has any orbit with $\tau \geq 3$; on the other hand, the languages defined by the orphan patterns of rules 9, 54, and 126 seem to be finite, but these rules have orbits with $\tau \geq 3$. We still do not know the reason for such a discrepancy, but possibly a more accurate study of the two classifications may explain it.

6. Conclusions and future work

In this paper, we have proposed two different classifications of Cellular Automata local rules. The first classification is a variant of the one proposed by Chua, and

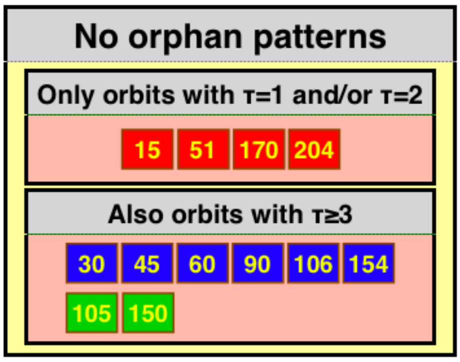
Table 1: The orphan patterns of all rules having only orbits with $\tau = 1$ and/or $\tau = 2$ constitute either finite or slender languages. Observe that all rules with $\kappa = 1$ (red background) correspond to finite languages; as a result, all slender languages correspond to rules with $\kappa = 2$ (blue background) or $\kappa = 3$ (green background).

Only orbits with τ=1 and/or τ=2						
Finite						
0	1	2	3	4	5	7
8	10	11	12	13	14	19
23	32	34	35	42	43	50
76	77	128	136	138	140	142
160	162	168	178	200	232	
6	24	28	33	36	40	56
57	108	130	132	156		
29	184					
Slender						
38	44	72				
27	46	58	78	172		

Table 2: All CA rules having orbits with $\tau \geq 3$ constitute (infinite) non-slender languages.

Also orbits with τ≥3						
Non-slender						
18	22	25	26	37	41	62
73	74	94	110	122	146	164

Table 3: The rules with no orphans can be divided into two groups: those surjective for L finite (upper half of the table) and those non-surjective for L finite (lower half of the table). Look how the classification of rules according to their complexity index and their periodic orbits matches perfectly the classification made by using the Formal Language theory.



it is based on the characteristics of the periodic orbits obtained by using all bit strings with $L \leq 15$, while the second classification is based on the languages defined by the orphan patterns which are found in the Gardens of Eden. We cannot overemphasize that these two classifications refer to two different characteristics of CA rules. For this reason, it is remarkable that the correspondence explained above exists. In the near future, we plan to give a rigorous explanation for such a phenomenon and find the reasons behind the six exceptions aforementioned.

References

[1] L. O. Chua, *A Nonlinear Dynamics Perspective of Wolframs New Kind of Science. Volumes I, II, III.* World Scientific, 2009.

[2] T. Li and J. Yorke, "Period three implies chaos," *American mathematical monthly*, vol. 82, no. 10, pp. 985–992, 1975.

[3] G. E. Pazienza, "Robust and Non-Robust ω-limit Orbits in 1D Cellular Automata," in *Proceedings of the NOLTA 2010*, 2010.

[4] S. Wolfram, "A new kind of science. Wolfram Media," *Inc., Champaign, IL*, 2002.

[5] D. Eppstein. (2008) Gliders and Wolfram's classification. [Online]. Available: http://www.ics.uci.edu/ eppstein/ca/wolfram.html

[6] F. Ohi, "Chaotic properties of elementary CA rule 168," in *Automata 2008: EPSRC Workshop Cellular Automata Theory and Applications*, Bristol, UK, Jun.12–14 2008.

[7] G. Rozenberg and A. Salomaa, *Handbook of Formal Languages, Vol. 1–3*. Springer, Berlin, 1997.