

Probabilistic 2D Cellular Automata Rules for Binary Classification

Mirosław Szaban

Institute of Computer Science,
 Siedlce University of Natural Sciences and Humanities, Poland
 Email: mszaban@uph.edu.pl

Abstract—In this paper are presented classification methods with use of two-dimensional three-state cellular automata. This methods are probabilistic forms of cellular automata rule modified from wide known almost deterministic rule designed by Fawcett. Fawcett's rule is modified into two proposed forms partially and fully probabilistic. The effectiveness of classifications of these three methods is analysed and compared. The classification methods are used as the rules in the two-dimensional three-state cellular automaton with the von Neumann and Moore neighbourhood. Preliminary experiments show that probabilistic modification of Fawcett's method can give better results in the process of reconstruction (classification) than the original algorithm.

I. INTRODUCTION

In a classification problem, we wish to determine to which class new observations belong, based on the training set of data containing observations whose class is known. The binary classification deals with only two classes, whereas in a multiclass classification observations belong to one of the several classes. The well-known classifiers are neural networks, support vector machines, k -NN algorithm, decision trees, and others. The idea of using cellular automata (CA) in the classification problem was described by Maji et al. [4], Povalej et al. [7] and recently by Fawcett [1]. Fawcett designed the heuristic rule based on the von Neumann neighborhood (so-called voting rule) moreover, tested its performance on different data sets. Recently, GA was considered as a tool to select CA rules with von Neumann neighbourhood for binary classification problem by Piwonska et al. [6].

Despite the fact that CAs have the potential to efficiently perform complex computations [9]; the main problem is a difficulty of designing CAs which would behave in the desired way. One must not only select a neighborhood type and size, but most importantly the appropriate rule (or rules). In some applications of CAs one can design an appropriate rule by hand (e.g. the GKL rule designed in 1978 by Gacs, Kurdyumov and Levin for density classification task [2]) or can use partial differential equations describing a given phenomenon [5]. Since the number of possible rules is usually huge, this is the extremely hard task, and it is not always possible to select them by hand. Therefore, in the 90-ties of the last century Mitchell et al. proposed to use GAs to find CAs rules able to perform one-dimensional density classification task [3].

In this paper are used rules based on methods of classification like the Fawcett's [1] method, and new proposed

modifications into a probabilistic form of such method. Above rules are applied to the rectangular grid with both neighbourhood types i.e. von Neumann and Moore neighbourhood, and the effectiveness of the modified rules is compared with the effectiveness of the original Fawcett's rule.

This paper is organized as follows. Section 2 describes two-dimensional CAs and binary classification problem. In section 3 are proposed new CA-based classifier as a modification of the Fawcett's rule. Experimental results are presented in Section 4. Last section contains conclusions and future works.

II. TWO-DIMENSIONAL CELLULAR AUTOMATA AND BINARY CLASSIFICATION PROBLEM

A two-dimensional CA considered in this paper is a rectangular grid of $N \times M$ cells, each of which can take on k possible states. After determining initial states of all cells (i.e. the initial configuration of a CA), each cell changes its state according to a rule - transition function TF which depends on states of cells in a neighborhood around it. In this paper is considered finite CA with the periodic boundary conditions. This is usually done synchronously, although asynchronous mode is used too. Two types of the neighborhood are commonly used: the von Neumann neighborhood (the four cells orthogonally surrounding the central cell) and can be described as: $a_{i,j}^{(t+1)} = TF[a_{i,j-1}^{(t)}, a_{i-1,j}^{(t)}, a_{i,j}^{(t)}, a_{i+1,j}^{(t)}, a_{i,j+1}^{(t)}]$, where $a_{i,j}^{(t)}$ denotes the state of a cell at position i, j in the two-dimensional cellular grid, at time step t , and also the Moore neighbourhood (the eight cells around the central cell) and can be described as: $a_{i,j}^{(t+1)} = TF[a_{i-1,j-1}^{(t)}, a_{i,j-1}^{(t)}, a_{i+1,j-1}^{(t)}, a_{i-1,j}^{(t)}, a_{i,j}^{(t)}, a_{i+1,j}^{(t)}, a_{i-1,j+1}^{(t)}, a_{i,j+1}^{(t)}, a_{i+1,j+1}^{(t)}]$.

The evolution of a CA is usually presented using so-called 'space-time diagrams' displaying grid of cells at subsequent time steps, with each state marked with different color.

The square state of the data space in classification problem should be i.e. $[0, 1] \times [0, 1]$. Suppose that $N \times M$ data-points $p_{(i,j)} = (x_i, y_j)$, where $i=1, 2, \dots, N$ and $j=1, 2, \dots, M$ are given as a training set from two classes: class 1 and class 2. When each of $p_{(i,j)}$ data-points is known as one of two classes then we have the classification. On the other hand, when even one of the data-points is not one of two known classes we have the classification problem. Moreover, to answer the question, what kind of class 1 or class 2 are unclassified data points it should be applied the classification method. In CA the data space of such problem should be mapped from $[0, 1] \times [0, 1]$ into the

grid of $N \times M$ cells (in this paper $N \times N$ for the simplicity). Each of cells can take one of 3 states, classified the state 1 (class 1) and state 2 (class 2) and also unclassified state (class 0). Classifier - the rule of CA will analyze the unclassified cells and changes its states into one of two known.

III. PROPOSED CA-BASED CLASSIFIER

The classification problem described in [1] is the base of this paper. The rule of CA known as $n4_V1_nonstable$ and presented there was the starting point to create a better classifier. The classification with use this rule is defined as:

- classify as *class 0*, if *class 1 neighbors + class 2 neighbors = 0*,
- classify as *class 1*, if *class 1 neighbors > class 2 neighbors*,
- classify as *class 2*, if *class 1 neighbors < class 2 neighbors*,
- classify as $rand(\{class 1, class 2\})$, if *class 1 neighbors = class 2 neighbors*.

The Fawcett's rule is productive enough for classification problems in little CA grids. This rule was presented in [1] as better than other, but it was tested and compared for only 81×81 wide CA grid. Therefore, in this paper are proposed modifications of Fawcett's rule into two patterns: partially and fully probabilistic. A proposed modification should strengthen an original and more accurate classify binary data, specially for large CA grid. The partially probabilistic modification ($n4_V1_nonstable_PP$) is proposed as follows:

- classify as *class 0*, if *class 1 neighbors + class 2 neighbors = 0*,
- classify as *class 1*, with probability $p(1)$,
- classify as *class 2*, with probability $p(2)$,

where probability are calculated form classified neighbours in the neighbourhood, i.e.: $p(i) = \frac{class\ i\ neighbors}{class\ 1\ neighbors + class\ 2\ neighbors}$, where $i = \{1, 2\}$. It means that unclassified cell will change in to one of classified state with probability calculated from known states in the neighbourhood and suitable state.

The full probabilistic modification ($n4_V1_nonstable_FP$) is proposed as follows:

- classify as *class 0*, with probability $p(0)$,
- classify as *class 1*, with probability $p(1)$,
- classify as *class 2*, with probability $p(2)$,

where probabilities are calculated form each class of neighbours in the neighbourhood, i.e.: $p(i) = \frac{class\ i\ neighbors}{\sum_{j=0}^2 class\ j\ neighbors}$, where $i = \{0, 1, 2\}$. It means that unclassified cell could change in to one of classified state or stay unclassified with probability calculated from states of cells in neighbourhood.

The $n4_V1_nonstable$ rule presented in [1] and newly proposed modifications will be examine on the sinusoidal testing sets (CA grids), shown in the Fig. 1.

In the Fig. 2(b, c and d) one can observe a classification process of the linear goal Fig. 2(a). In the classification was used $n4_V1_nonstable$ rule in CA size: 100×100 . In the first step (see, Fig. 2(b)), 1% cells of known state (classified as class

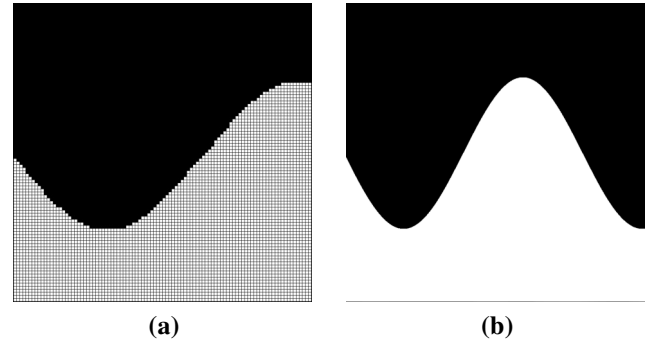


Fig. 1. Two-dimensional classifications of sinusoidal goals (examples): 100×100 (b) and 800×800 (c).

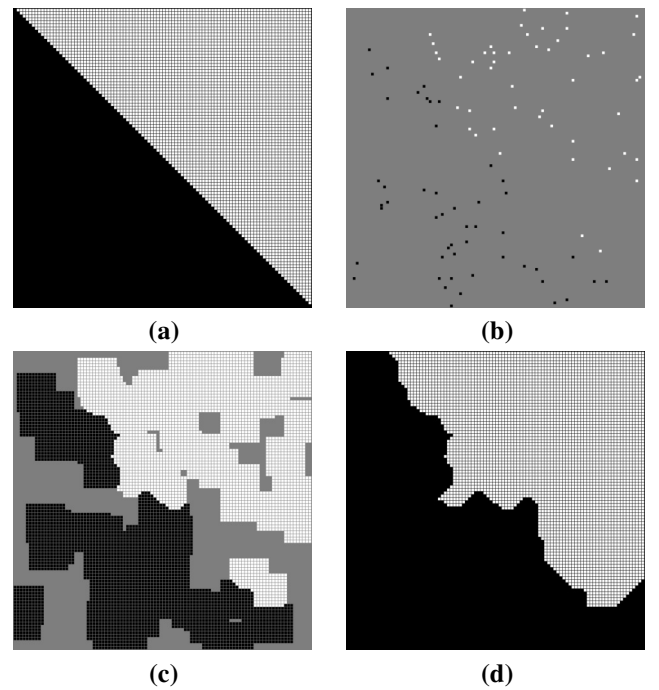


Fig. 2. An example of classification process with use of $n4_V1_nonstable$ rule in 2D CA with size 100×100 for linear goal (a), initial configuration of CA - 1% of classified cells (b), temporary CA state - classification in progress (c), final CA state - after classification (d).

1 - cells in black state or class 2 - cells in the white state) was randomly chosen from the linear goal. This initial configuration of CA was subject to the classification and the unclassified (cells in the gray state) with the use of $n4_V1_nonstable$ rule (temporary state of CA under classification process, see Fig. 2(c)). After the classification, the obtained result shows Fig. 2(d). The effectiveness of classification in presented example achieved level 93, 38% it means that 662 CA cells from 10000 CA cells have an incorrect classification.

Suppose, the $G_{N \times M}$ is the binary matrix of the classification goal. Also, the $F_{N \times M}$ is the binary matrix of the final configuration. The $E_{N \times M} = |G_{N \times M} - F_{N \times M}|$ is the absolute difference between two matrixes. Therefore, the Effectiveness (in %) is calculated by the formula: $Effectiveness = \frac{N * M - \sum_{i=1}^N \sum_{j=1}^M e_{(i,j)}}{N * M} * 100$, where $e_{(i,j)} \in E_{N \times M}$.

IV. EXPERIMENTAL RESULTS

Proposed modifications of Fawcett's CA rule was tested and compared with original one. The results of analysis for the efficiency of classification are described above. During each test were used $N \times N$ CA, where $N \in \{100, 200, \dots, 700, 800\}$, according to each of CA rules ($n4_V1_nonstable$, and newly proposed: $n4_V1_nonstable_PP$, $n4_V1_nonstable_FP$) with the von Neumann and Moore neighbourhood. In the goal, a 99% random states of CA cells was changed into the unclassified. So, only 1% CA cells stayed as classified. Such form of the goal was the initial configuration of CA to classification. After classification (reconstruction of goal) we obtained the finite CA state, and it was compared with the goal CA state; then the differences between both states (incorrect classified CA cells) and next were calculated *Effectiveness*.

A. An analysis of the incorrect classifications for CA rules.

Classification in CA with the use of the Fawcett's rule, and new proposed probabilistic modification with von Neumann and Moore neighbourhood was realized in [8] for a linear goal during 1000 tests with above-presented conditions and assumptions. In [8] we can see that the effectiveness grows with the CA size, but it only depends on the fast growing area with the same class of data where the border with the different classes is increasing very slowly. The results showed that in general, the proposed modifications give better results in a classification for a linear goal than the original rule. For von Neumann neighbourhood, it could be observed for higher CA sizes (from 300 to 800) with both modifications in particular partially probabilistic one $n4_V1_nonstable_PP$, where we can see that the effectiveness of new proposed modifications is better than for the original Fawcett's rule. Some kind anomaly one can see for CA size equal to 700×700 (see, Table 1 in [8]), where can be observed the best result of classification for the fully probabilistic classifier ($n4_V1_nonstable_FP$), better than Fawcett's rule and much better than partially probabilistic one. For Moore neighbourhood both modification gives better results for each CA sizes except the size equal to 700×700 where Fawcett's rule classified effective (see, Table 2 in [8]).

In this paper is presented analysis mentioned above rules for sinusoidal goal (see, Fig. 1) with both von Neumann and Moore neighbourhood.

In the Tab. I one can see the results of classification for a sinusoidal goal in CA with the use of mentioned above three CA rules with Moore neighbourhood for 1000 tests. We can observe that the effectiveness of new proposed modifications is better than for the original Fawcett's rule for the sinusoidal goal. It could be observed for both modifications (marked in bold), in particular for fully probabilistic one ($n4_V1_nonstable_FP$), except the tests for CA size equal to 400×400 and 600×600 (the underlined numbers mean the best results in general, from all analysed rules). We can also observe that effectiveness of classification with use of CA with Moore neighbourhood for a sinusoidal goal in much higher than for linear goal, it is easy to see in comparison the incorrect classification from Tab. I and Tab. 2 from [8].

TABLE I

CLASSIFICATION RESULTS OF GOAL SINUSOIDAL SET FOR CA RULES: $n4_V1_nonstable$, $n4_V1_nonstable_PP$, $n4_V1_nonstable_FP$ WITH MOORE NEIGHBOURHOOD. THE WORST EFFECTIVENESS AND (HIGHEST NUMBER OF INCORRECT CLASSIFIED CA CELLS) FROM 1000 TESTS.

CA size $N \times N$	CA rule Fawcett's	CA rule partially probab.	CA rule fully probab.
100×100	92,7% (730)	93,69% (631)	93,69% (631)
200×200	96,81% (1275)	97,1% (1160)	97,35% (1062)
300×300	98,21% (1613)	98,34% (1493)	98,4% (1436)
400×400	98,78% (1959)	98,82% (1882)	99,12% (1977)
500×500	99,04% (2388)	98,99% (2513)	99,12% (2203)
600×600	99,28% (2606)	99,22% (2817)	99,26% (2661)
700×700	99,37% (3076)	99,37% (3100)	99,4% (2937)
800×800	99,45% (3528)	99,47% (3413)	99,46% (3465)

TABLE II

CLASSIFICATION RESULTS OF GOAL SINUSOIDAL SET FOR CA RULES: $n4_V1_nonstable$ (FAWCETT'S RULE), $n4_V1_nonstable_PP$, $n4_V1_nonstable_FP$ WITH VON NEUMANN NEIGHBOURHOOD. THE WORST EFFECTIVENESS AND (HIGHEST NUMBER OF INCORRECT CLASSIFIED CA CELLS) FROM 1000 TESTS.

CA size $N \times N$	CA rule Fawcett's	CA rule partially probab.	CA rule fully probab.
100×100	92,75% (725)	94,31% (569)	93,53% (647)
200×200	97,08% (1169)	97,43% (1028)	97,35% (1061)
300×300	98,44% (1403)	98,35% (1482)	98,44% (1402)
400×400	98,9% (1757)	98,92% (1732)	98,78% (1952)
500×500	99,13% (2184)	99,16% (2102)	99,08% (2310)
600×600	99,3% (2507)	99,28% (2588)	99,29% (2561)
700×700	99,41% (2875)	99,39% (2974)	99,41% (2913)
800×800	99,49% (3259)	99,49% (3246)	99,48% (3342)

For the rule $n4_V1_nonstable_FP$ the difference between number incorrect classification in linear and sinusoidal goal is equal to $\{210, 432, 436, 526, 726, 1008, 1171, 923\}$ for CA sizes $\{100, \dots, 800\}$, and also for $n4_V1_nonstable_PP$ differences are equal to $\{185, 167, 402, 564, 471, 617, 921, 1086\}$ respectively.

Similar results but not so high, are obtained for classification the sinusoidal goal with this three methods for CA with von Neumann neighbourhood. In the Tab. II one can see the results of classification for a sinusoidal goal in CA with the use of mentioned above three CA rules with von Neumann neighbourhood for 1000 tests. We can observe that the effectiveness of partially probabilistic modification ($n4_V1_nonstable_PP$) is better than other rules for the sinusoidal goal (marked in bold and underlined). It can be observed for tests realised for most of analysed CA sizes (compare, Tab. 1 from [8] and Tab. II).

The observed highest effectiveness for modification of Fawcett's rule maybe is not so spectacular, but we should interpret it from the another point of view; it means the number of incorrectly classified CA cells should be analyzed. One can see that the modifications have in general the lowest number of incorrectly classified CA cells than the original rule (see, numbers in bold in Tab. II, Tab. I also Tab. 1 and Tab. 2 from [8]). For example, for the linear goal in the Tab. I for CA size 200×200 , the fully probabilistic modi-

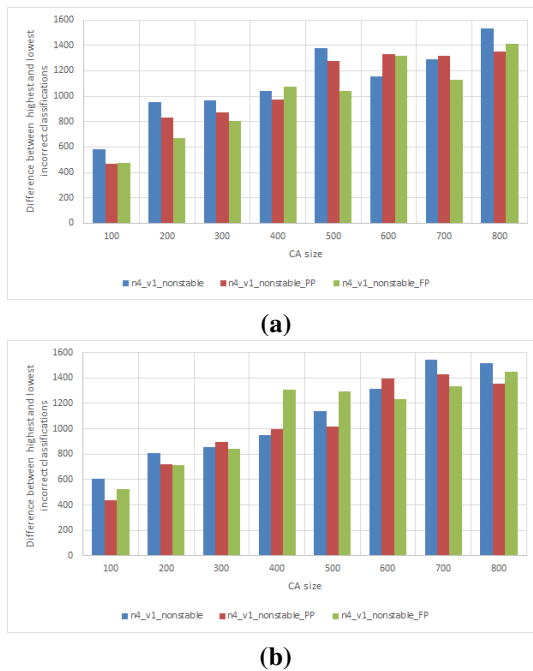


Fig. 3. Ranges of scattering for incorrect classifications for 1000 tests with a sinusoidal goal: (a) Moore neighbourhood, (b) von Neumann neighbourhood.

fication ($n4_V1_nonstable_FP$) has 213 CA cells less than the original rule, and also partially probabilistic modification ($n4_V1_nonstable_FP$) has 115 CA cells less.

So, the analysis of the worst effectiveness and highest number of incorrect classification for data presented above in Tables leads to the conclusion; if the worst results for classification with use of proposed modifications are better than for the original rule, then the new classifiers are more efficient because their the worst classifications are much better.

B. An analysis of the scattering ranges for incorrect classifications for CA rules.

The scattering range is the difference between the highest number of incorrect classifications and the lowest number of incorrect classification. The scattering ranges were calculated for each analysed CA rules for linear and sinusoidal goals with the use of both neighbourhoods von Neumann and Moore. The highest and lowest number of incorrect classifications were selected from 1000 tests for random initial configuration of CA.

We can observe in the Fig. 3 the scattering ranges for sinusoidal goal with Moore neighbourhood Fig. 3(a) and von-Neumann neighbourhood Fig. 3(b). One can see in Fig. 3(a) the scattering ranges for newly proposed modified CA rules are shorter in general than for original rule, in particular for fully probabilistic modification ($n4_V1_nonstable_FP$) with the use of Moore neighbourhood. Moreover, remembering such rule has, in general, the lowest incorrect classifications we can conclude that fully probabilistic classifier shows up the more accurate and consistent with Moore neighbourhood for the sinusoidal goal.

Similarly, in Fig. 3(b) can be observed in general the shortest scattering ranges for proposed modification in particular for partially probabilistic modification ($n4_V1_nonstable_FP$) of CA rule with the use of von Neumann neighbourhood.

For the linear goal, also we can observe the shortest scattering ranges in most cases of CA size, for the proposed new classifiers (see, [8] in particular the Moore neighbourhood).

V. CONCLUSIONS AND FUTURE WORKS

In the paper was presented problem of classification with use of three-state two-dimensional cellular automaton. Among classifiers were analysed the wide known Fawcett's rule and two proposed probabilistic modifications of such rule. The conducted experiments show the better effectiveness for classification applying a newly proposed classifiers (modifications) to reconstruction goals from state consisted of only 1% classified states. Moreover, the proposed modifications result in a much lower number of incorrectly classified CA cells.

Also, the analysis the data from Tab. I, Tab. II; Table 1 and Table 2 from [8] leads to the conclusion that seems to be a weak relationship between kind of modification and type of CA neighbourhood. For the von Neumann neighbourhood better classifications have obtained for partially probabilistic modification, and for Moore neighbourhood fully probabilistic modification are better.

Moreover, from Fig. 3 and Fig. 2 from [8] ensue in general the shortest scattering range for proposed modifications of Fawcett's rule than the original one. From that and above conclusions, we can understand that proposed probabilistic classifiers characterized by more accurate and consistent classification.

In the future works the newly proposed rules will also be examined with use of new testing goals like e.g.: parabolic, closed area (circular, square, concave boundary, ...), disjunctive and other. Also, will be analysed the effectiveness of classification depending on the number of unclassified states of CA cells initial configuration.

REFERENCES

- [1] T. Fawcett, "Data mining with cellular automata." *ACM SIGKDD Explorations Newsletter*, 10(1), pp. 32–39, 2008.
- [2] P. Gacs, G. Kurdyumov, and L. Levin, "One dimensional uniform arrays that wash out finite islands." *Problemy Peredachi Informatsii*, 12, pp. 92–98, 1978.
- [3] M. Mitchell, P. Hraber and J. Crutchfield, "Revisiting the edge of chaos: Evolving cellular automata to perform computations." *Complex Systems*, 7, pp. 89–130 1993.
- [4] P. Maji, B. Sikdar, and P. Chaudhuri, "Cellular automata evolution for pattern classification." *Lecture Notes in Computer Science 3305*, pp. 660–669. Springer Verlag 2004.
- [5] S. Omohundro, "Modelling cellular automata with partial differential equations." *Physica 10D*, 10(1-2), pp. 128–134, 1984.
- [6] A. Pivowska, F. Serebinski, M. Szaban, "Learning Cellular Automata Rules for Binary Classification Problem." *The Journal of Supercomputing*, Volume 63, Issue 3, pp. 800–815, 2013.
- [7] P. Povalej, M. Lenic, and P. Kokol, "Improving ensembles with classificational cellular automata." *Lecture Notes in Computer Science 3305*, pp. 242–249. Springer Verlag 2004.
- [8] M. Szaban, "Probabilistic Binary Classification with Use of 2D Cellular Automata." *12th International conference on Cellular Automata for Research and Industry - ACRI 2016*, 2016 (in print).
- [9] S. Wolfram: *A New Kind of Science*. Wolfram Media, 2002.