

PAPER • OPEN ACCESS

Neural networks playing 'matching pennies' with each other: reproducibility of game dynamics

To cite this article: T Dolgova and S Bartsev 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **537** 042002

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices
to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of
every title for free.

Neural networks playing ‘matching pennies’ with each other: reproducibility of game dynamics

T Dolgova¹ and S Bartsev^{1,2}

¹ Siberian Federal University, 79 Svobodny pr., Krasnoyarsk, 660041, Russian Federation

² Institute of Biophysics SB RAS, Federal Research Center, Krasnoyarsk Scientific Center SB RAS, 50, Akademgorodok, Krasnoyarsk, 660036, Russian Federation

E-mail: bartsev@yandex.ru

Abstract. Reflection is an essential feature of consciousness and possibly the single most important one. This fact allows us to simplify the objective of the concept of ‘neural correlates of consciousness’ and to focus investigations on reflection itself. Reflexive games are the concentrated and pure embodiment of reflection manifestation without the addition of other higher cognitive functions. In this paper, we use the game ‘matching pennies’ (“Odd-Even”) in order to trace the strategies and possible patterns of recurrent neural network operation. Experimental results show the splitting of all considered game patterns into two groups. A significant difference was observed in these groups of patterns, indicating a qualitative difference in game dynamics apparently due to the qualitatively different dynamic patterns of neuron excitations of the networks. A similar splitting of all players into two groups was found by other authors for human players, which differ in terms of the reflection availability. By this, we can assume that one of the causes of the splitting is that the presence of reflection in a particular group of recurrent neural networks dramatically changes the game meta-strategy.

1. Introduction

The question regarding the nature of consciousness is one of the fundamental problems of science and philosophy. Currently, cognitive neuroscience is actively developing with the aim of detecting the neural correlates underlying cognitive phenomena. The long-term goal of such studies is to establish what kind of neurological events correlate with those or other states and the contents of consciousness. Ultimately, such studies should determine the minimum sufficient number of neurons or the level of complexity of a nervous system, which necessarily accompanies a given conscious experience. This is the standard definition of the concept of ‘neural correlates of consciousness’ (NCC) [1], [2].

The orientation of the NCC concept on the identification of minimal neural structures accompanying some conscious experience justifies moving from studying the complex functions of the human brain [3], [4], [5], [6], [7], [8] to investigating the separate phenomena of consciousness implemented by simple systems.

The theoretical argument in favour of the possibility of an extremely simple formal description of consciousness phenomena is made by V. Lefevre [9]. He argued that subjectivity (consciousness, mind) is associated with the presence of an internal presentation of the external world including the image of the subject itself. Lefevre called this type of construction of the internal representation



‘reflection’. Using a simple mathematical model of reflection built on the basis of simple axioms, Lefevre explained the results of a number of psychological experiments that had not been satisfactorily explained previously.

This experience of successfully modelling the empirically complex property of consciousness known as reflection by means of a simple mathematical model and the orientation of the NCC concept at identifying minimal neural structures leads to the suggestion that many properties of consciousness, including reflection, can be explored with simple natural and artificial neural networks.

Indeed, experiments with bees and bumblebees having no more than 1 million neurons [10], [11], [12] demonstrated that these organisms are able to operate by concepts and abstractions, and also to manipulate the positions of perception, since the ability to repeat the previously unknown behaviour of another bumblebee is impossible without transitions between the 1st and 3rd positions of perception, which is the manifestation of reflection.

To reveal the essence of reflexive processes, it makes sense to investigate such behaviour that contains the minimal contribution of other cognitive functions, for example logical reasoning, pattern recognition and memory. Reflexive games almost perfectly correspond to these requirements [13], [14], representing a special kind of Nash equilibrium game with mixed strategies [14]. Among the simplest reflexive games are ‘matching pennies’ (‘Odd-Even’), ‘rock-paper-scissors’ and the ‘Believe It or Not?’ card game. These games have no fixed winning strategies and each player, in order to win, must constantly change his strategy depending on the style of play of the opponent.

Representing the simplest form of player, recurrent neural networks (NN) playing with each other can be used to study the internal mechanism of reflection freed from the emotions, greed and fatigue of human players. We can expect that in the course of a reflexive game, new game strategies are constantly generated depending on the opponent's previous moves. Since the weight coefficients of NN are float digits and the initial untrained state (weight matrix) of NN is provided by the generator of random digits, we can expect that the number of different game score dynamics (representing the interaction of two strategies) is very big. Thus, the experiment with small NN can answer the principal question of whether the experiments with playing NN are reproducible, and whether some typical game patterns exist or not. Moreover, the presence of player reflection is expected to change game patterns dramatically, which can be detected in the simulation experiments.

2. Methods and materials

In order to get an answer to these questions, recurrent full-connected NN consisting of a small (10) number of neurons playing ‘matching pennies’ were used. It is a game for two. Each player chooses ‘0’ or ‘1’ in secret and their choice is then presented and compared. An even sum means that one player receives one point, whereas an odd sum means that the other player receives one point. ‘Matching pennies’ is an example of a reflexive game which has no fixed winning strategy and the strategy for a draw is random choice. If the players are not able to choose randomly or try to avoid a drawn result, then the game becomes non-trivial: whoever predicts the moves of their opponent better is the winner.

The functioning of a neural network is determined by a matrix of coupling coefficients, which also describes its structure [15]:

$$\alpha_i^{n+1} = \frac{\rho_i^n}{a + |\rho_i^n|}, \quad \rho_i^n = \sum_j w_{ij} \alpha_j^n + A_i^n \quad (1)$$

where α_i^n is the output signal of the i -th neuron at the n -th instant of time, w_{ij} is the connection weight matrix of a full-connected recurrent network, and A_i^n is the input signal of i -th neuron at n -th time moment.

The information about the partner's choice is fed via input neurons of the NN. Two output neurons determine the NN's choice for the next move: the sign of the output value difference gives the NN's choice ‘0’ or ‘1’.

The target functions for the NNs are different because one NN wins when its output (game move) coincides with the move of the second NN, but the second NN wins when it makes a move different from the one of the first NN:

$$H_1(\alpha_i^n, n) = \frac{1}{2} \left[(\alpha_3^n - \text{move}2)^2 + (\alpha_4^n - (1 - \text{move}2))^2 \right] \quad (2)$$

$$H_2(\alpha_i^n, n) = \frac{1}{2} \left[(\alpha_3^n - (1 - \text{move}1))^2 + (\alpha_4^n - \text{move}1)^2 \right]. \quad (3)$$

After each n -th move, the errors of NN $\left(\frac{\partial H(\alpha_i^n, n)}{\partial \alpha_i^n} \right)$ are calculated and distributed among all neurons in accordance with the next formulae describing the backward time propagation of errors [16], [17]:

$$\Delta_i^{n-m-1} = \frac{1}{a} \left(\sum_j w_{ji} \Delta_j^{n-m;n} + \delta_{m0} \frac{\partial H(\alpha_i^n, n)}{\partial \alpha_i^n} \right) (1 - |\alpha_i^{n-m}|)^2 \quad (4)$$

where $\Delta_i^{n-m;n}$ is the normalized increment of H (distributed among neurons error); δ_{m0} is Kronecker's symbol, $\delta_{m0} = \begin{cases} 1, m=0 \\ 0, m \neq 0 \end{cases}$. It is necessary to calculate $\Delta_i^{n-m;n}$ from $m=1$ to $m=k$, where k is the depth of the error propagating back in time.

Then the summarized gradient of H with respect to the weight coefficients at the n -th move is:

$$\frac{dH}{dw_{ij}} = \sum \Delta_i^{n-m;n} \alpha_j^{n-m} \quad (5)$$

Then it is possible to organize an adaptation procedure based on the gradient descent:

$$w_{ij}^{n+1} = w_{ij}^n - s \frac{dH(\alpha_i^n, n)}{dw_{ij}}, \quad s > 0 \quad (6)$$

For processing the results of the game with dynamic outputs, several approaches were used:

1. Autocorrelation calculation. The calculation of the autocorrelation function of the game score time series was used to detect possible similarities among patterns in the game dynamics. At this preliminary stage of the study, the expectable non-stationarity of the time series is ignored. For processing, the autocorrelation function of the form is used:

$$p(k) = \frac{1}{N-k} \sum_{i=1}^{N-k} \frac{(\bar{y} - y_i)(\bar{y} - y_{i+k})}{\sigma_y^2} \quad (7)$$

where N is the size of time series, \bar{y} is the average value of a dynamic variable, and σ_y^2 – dispersion.

2. Fourier transform. When using this MS Excel function, the initial time series of dynamics are decomposed into elementary components, namely the harmonic oscillations of different frequencies and phases. It is assumed that changes in phases represent changes in game strategy in some way.

3. Cluster analysis was used in order to identify whether there are similar patterns in functions and Fourier transforms obtained before autocorrelation.

For the simulation open soft, 'Lazarus' (<https://www.lazarus-ide.org/>) was used. For statistic analyses of experimental data, MS Excel (autocorrelation function, Fourier transform) and open soft 'Tanagra' (<http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html>) (cluster analyses) were used.

3. Results

There were conducted 50 ‘network-vs-network’ games. As was expected, there were no identical realizations of game dynamics (score time series). Just for illustration, four realizations of game dynamics are presented in figure 1. On the background of visible mismatching in the realization of dynamics, the question concerning the existence of general or common game patterns in the game arises.

To identify possible “hidden” patterns, the obtained score time series of the game were analysed using several approaches.

3.1 Autocorrelation function analyses

The calculation of the autocorrelation function of the series allows the presence of periodically recurring patterns in the game score to be estimated. Figure 2 shows the autocorrelation functions for some games. Visually, it is difficult to assess the similarity of various curves and the number of similar patterns.

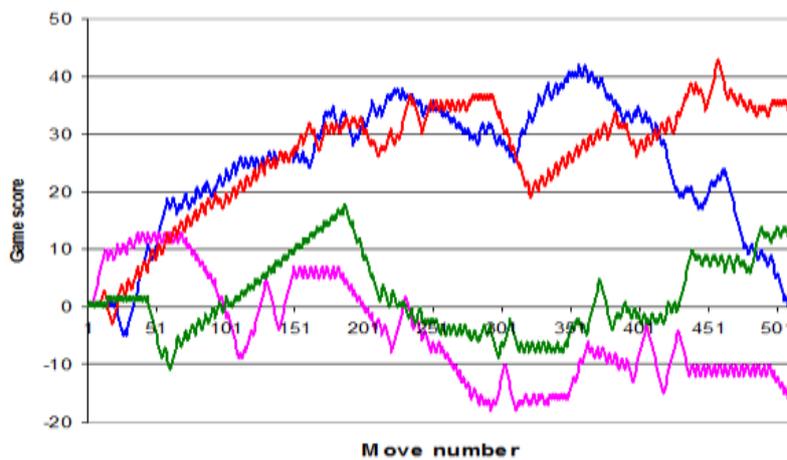


Figure 1. Examples of ‘network-vs-network’ game dynamics.

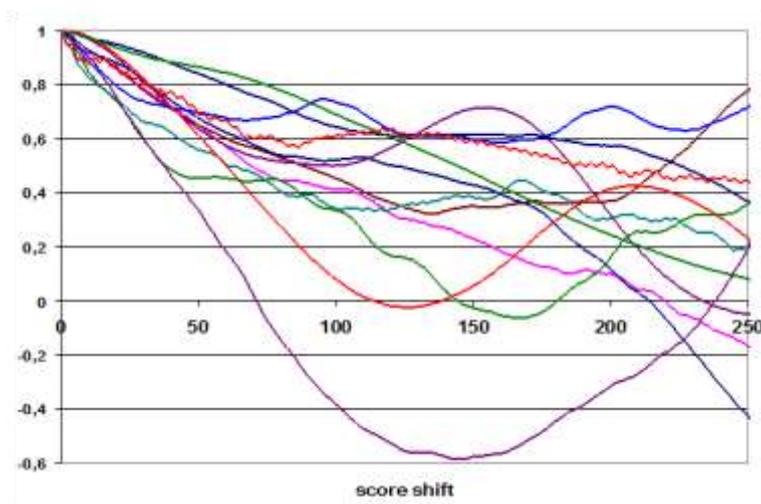


Figure 2. Examples of autocorrelation functions of some game score dynamics.

3.2 Fourier transform analyses

Fourier transform allows the presence of selected frequencies (similar game score change periods) to be evaluated, and in addition, allows a phase change to be evaluated that can be interpreted as a sudden switching of strategies. The analysis revealed that the amplitude frequency spectra of the

games do not show significant differences, while the graphs of the phase shift show extremely different patterns (figure 3). However, we received the same situation as for the autocorrelation function, namely that it was not possible to detect the similarity of the phase curves or select any similar patterns. Therefore, cluster analysis was applied to the obtained autocorrelation patterns and the results of the Fourier transform.

3.3 Cluster analyses

Open source ‘Tanagra’ software providing a wide range of statistical data processing was used for conducting cluster analyses (Hierarchical Clustering, Ward's criterion). The cluster analysis applied to the set of autocorrelation functions and also for the amplitudes and phases of Fourier transforms surprisingly showed the presence of only two very different patterns of game score dynamics (figure 4). It looks unusual since this significant difference arose on the background of the continuous generation of new strategies, depending on the partner's moves, which, in turn, depend on the prehistory of the game, in an expectedly complex way.

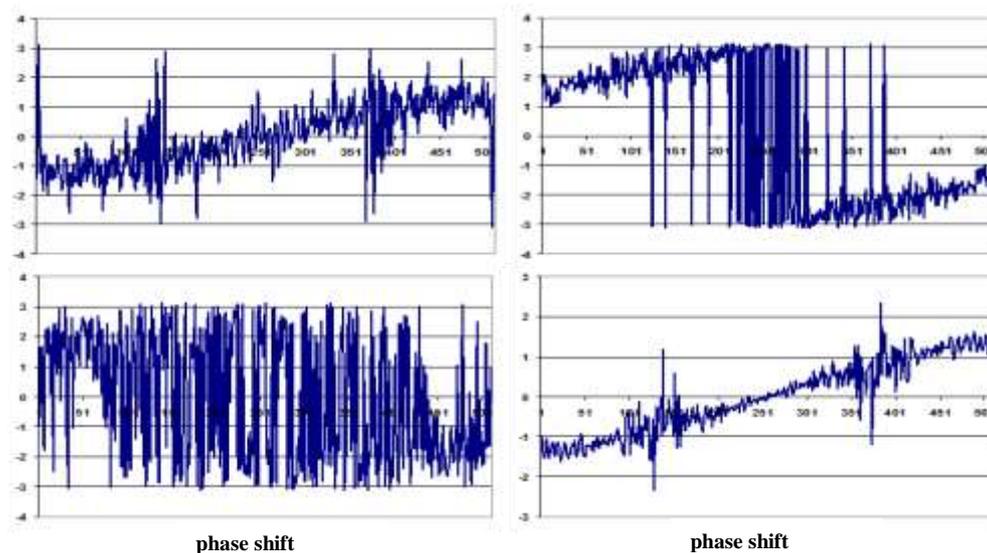


Figure 3. Examples of phase changes obtained by Fourier transform.

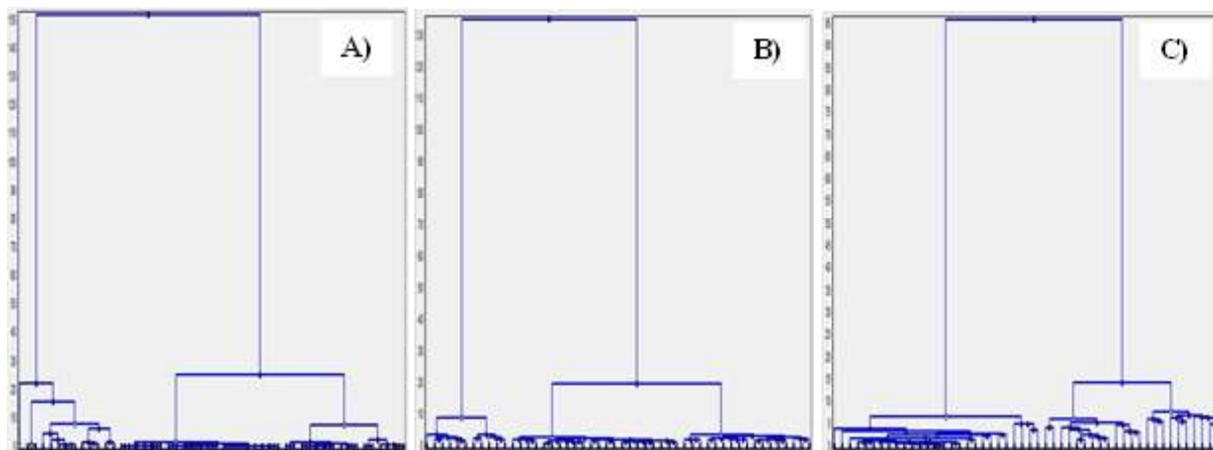


Figure 4. Cluster trees for autocorrelation functions (A), amplitudes (B) and phases (C) of Fourier transform of game score time series.

Analyses of the distributions of NN dynamics between these two big clusters demonstrates a very interesting result. The distributions of NN with respect to autocorrelation functions, amplitudes and phases of Fourier transforms do not coincide. However, it can be seen visually from Figure 4 if the number of NN dynamics in each cluster is compared.

This difference may indicate a qualitative difference in game patterns as a manifestation of qualitative difference in the state of neural networks. The cluster analyses revealed that there are at least three parameters or structural characteristics which equally as clearly distinguish the game dynamics.

4. Discussion

It is interesting that a similar splitting into two qualitatively distinct groups was observed during mass research of decision making in non-cooperative strategic interactions [18]. The subjects played the ‘rock-paper-scissors’ game. The study revealed the presence of two types of strategies in human subjects: 1) ‘won-leave, lose-change’, characteristic of the majority of subjects; 2) a strategy based on knowledge of the above-described ‘basic human’ strategy.

The second strategy involves transferring yourself to the position of your opponent, which is clearly a reflexive action. At the same time, it obviously allows the player to prevail over the first strategy.

It should be noted that according to the theory of games, in any finite non-cooperative game there is an “invincible” strategy, known as the Nash equilibrium. In the case of reflexive games, the Nash equilibrium is a random choice of one of the options, but typical neural networks after training are deterministic systems, and their response depends only on the input data and previous experience, and theoretically can be predicted. However, the emergence of reflection is expected to make the behaviour of a neural network more sophisticated because each one of its moves depends not only on the current move of its opponent and previous experience, but also on the internal reflexive representation of the opponent’s strategy.

5. Conclusions

It was shown that game patterns are predictably very variable. However, their binary splitting into two clusters with respect to different features is rather surprising. The observed significant difference in these groups of patterns indicates a qualitative difference in game dynamics apparently due to the qualitatively different dynamic patterns of neuron excitations of the networks.

It can be hypothesized that the cause of such a large difference in game patterns is the presence of a reflexive representation of an opponent in one or both neural network players. However, since the distributions of game dynamics with respect to the autocorrelation function, amplitude and phase of Fourier transforms do not coincide, we can suggest there are additional features of structure and functioning which result in their binary splitting. This very interesting question requires further investigation.

References

- [1] Crick F and Koch C 1990 Towards a neurobiological theory of consciousness. *Seminars in Neuroscience* **2** 263–75
- [2] Crick F and Koch C 2003 Framework for consciousness *Nature Neuroscience* **6**(2) 119–26
- [3] Dehaene S 2009 Conscious and Nonconscious Processes: Distinct Forms of Evidence Accumulation? *Seminaire Poincare XII* 89-114
- [4] Dehaene S and Naccache L 2001 Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* **79** 1–37
- [5] Dehaene S and Changeux J-P 2011 Experimental and Theoretical Approaches to Conscious Processing. *Neuron* **70** 200–27
- [6] Kiefer M and Pulvermüller F 2012 Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions *Cortex* **48** 805–25

- [7] Mehta N and Mashour G 2013 General and specific consciousness: a first-order representationalist approach. *Frontiers in Psychology: Consciousness Research* **4** 407–10
- [8] Tononi G, Boly M, Massimini M and Koch C 2016 Integrated information theory: from consciousness to its physical substrate *Nature Reviews: Neuroscience* **17** 450–61
- [9] Lefebvre V 2003 *Reflexion* 496
- [10] Alem S, Perry C and Zhu X 2016 Associative Mechanisms Allow for Social Learning and Cultural Transmission of String Pulling in an Insect *PloS Biol.* **14(10)** e1002564
- [11] Avargués Weber A and Giurfa M 2013 Conceptual learning by miniature brains *Proc R Soc B* 280
- [12] Giurfa M 2007 Behavioral and neural analysis of associative learning in the honeybee: a taste from the magic well *J Comp Physiol A* 24
- [13] Camerer C, Ho T H and Chong J K 2004 A Cognitive Hierarchy Model of Games. *The Quarterly Journal of Economics* **119(3)** 861-98
- [14] Novikov D and Chkhartishvili A 2013 *Reflexion and control: mathematical models* (Moscow: Physical-chemical publishing) p 412
- [15] Bartsev S 2010 *Heuristic Neural Network Models in Biophysics: Application to the Problem of Structural and Functional Conformity* (Krasnoyarsk: Sib. Fed. Univ.) p 115
- [16] Khaikin S 2006 *Neural Networks: The Full Course* (Moscow: Williams) p 1104
- [17] Rumelhart D and Hinton G 1988 Learning representations by error back-propagating errors *Nature* **323** 533-6
- [18] Wang Z, Xu B and Zhou H-J 2014 Social cycling and conditional responses in the Rock-Paper-Scissors game *Scientific reports* **4** 5830