

The Prognosis of Delayed Reactions in Rats Using Markov Chains Method

Sulkhan N. Tsagareli¹, Nino G. Archvadze¹, Otar Tavdishvili², Marika Gvajaia¹

¹Biology Department, Faculty of Exact and Natural Sciences, Iv. Javakishvili Tbilisi State University, Tbilisi, Georgia

²Department of Applied Systems for Pattern Recognition, Institute of Cybernetics, Georgian Technical University, Tbilisi, Georgia

Email: nino.archvadze@tsu.ge, sulkhan.tsagareli@tsu.ge, otartavdi@yahoo.com

Received 5 November 2015; accepted 12 January 2016; published 15 January 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The original modified method of the direct delayed reaction has been used for the evaluation of food-obtaining strategy across spatial learning tasks in T-maze alternation. The optimal behavioral algorithms for each experimental day have been identified so that the animals obtain maximum possible food amount with minimal number of mistakes. Markov chain method has been used for the prognosis of rat's behavioral strategy during the spatial learning task. The learning and decision-making represent the probabilistic transition process where the animal choice at each step (state) depends on the learning experience from previous step (state).

Keywords

Delayed Reactions, Behavioral Algorithm, Markov Chain

1. Introduction

Both humans and animals live in a rich world of constantly changing external signals, thus the detection of changes and adequate reaction to environmental variation is crucial for the survival and successful adaptation. The learning and decision-making abilities are essential for adaptive behavioral strategy.

With no doubt, the behavioral studies remain the most prominent tool for exploring neural mechanisms of learning, memory and decision-making, though the past year experiments in behavioral and cognitive neuroscience [1] [2] have shaped innovate interdisciplinary approaches integrating the findings from different fields of science. Among them are widely-used mathematical-statistical methods [3]-[9]. At the same time, as a result of cross-disciplinary studies, various equations expressing biological relations gain a prominent position within

mathematics [10]. Mathematical approach is believed to explain all phenomena where at least one variable varies towards other variables or affects them [11] [12].

The probabilistic and stochastic processes in cognitive neuroscience are considered as constitute foundations of learning and decision-making [9] [10] [13]-[16]. Application of algebraic expressions has been proposed for description of learning processes [17]-[19]. More explicit mathematical methods are suggested for assessment of cognitive mechanisms involved in the adaptive learning, repeated decision tasks, reinforcement, and strategic changes [20]. The learning ability is studied in a binary choice task, where subjects have to choose from two possibilities (correct vs. incorrect) and the authors assume that the answers follow Bernoulli distribution that depends on a hidden state, reflecting the subject's performance [21]. The probability that animals perform better than by chance is quantified and trial by trial performance is estimated. To develop a dynamic approach for analyzing learning experiments with binary response (correct, incorrect), state-space model of learning is introduced.

The question arises of how we go about understanding the probabilistic and stochastic processes that groundwork learning and memory. Due to a large number of possible parameters involved in learning, memory, and formation of adequate behavioral strategy, sometimes it seems difficult to generalize the results of behavioral studies. From this point, mathematical approach to the problem in general, and quantification of the measured parameters in particular, should be considered as the most reasonable means to identify the behavioral features and to interpret numeric data. Nowadays, such attitude is rather common to behavioral studies. Beside the wide range of traditional statistical methods used for analysis of behavioral parameters, different mathematical approaches and models are proposed for the data analysis in ethology, psychology, neuroscience, etc. [21]-[26].

Markov chains are one of the basic tools in modern statistical computing, providing the basis for numerical simulations conducted in a wide range of disciplines. This mathematical technique undergoes transitions from one state to another through a set of values which process can take. Every state depends only on the current state and not on the sequence of events that precede it. Markov chains have many applications in biological modeling [2] [13] [14] [18] [21] [23] [27] [28]. They are widely applied in neuroscience research, including behavioral studies, neuron assemblies, ion channels and others [7] [14] [22] [29], are the output of the Markov chains which can be different variables, for instance, the decision, or animal's motor output and neural activity [13] [17] [24] [30] [31]. Markov chains are also used in simulations of brain function, such as the activity of mammalian neo-cortex [30].

There are several ways we may go while studying cognitive abilities. We can simply observe, record and analyze neurobiological, behavioral conformities, or/and make an attempt to construct quantitative models in order to understand the computations that underpin cognitive abilities. What different cognitive studies share is the attempt to identify the involvement of various brain states in the behavioral processes. Sometimes the segmenting of observed and measured behavioral processes into consequence elements is needed to explore and quantify transitions between them. Thorough inquiry of animal behavioral conformities across learning process gave huge body of information on behavioral consequences which can be tested with Markov chains analysis. Markov chains have shown to provide better estimation of learning conformities in comparison with other methods used to infer from behavior data treatment. Such modern approaches contribute in studying the cognitive abilities and their behavioral correlates.

From neurobiological point of view, it is interesting to perform extrapolation on the basis of experimental data in order to establish quantitatively the degree of learning and the dynamics of the memory. The results of behavioral experiments can be predicted by means of the mathematical model using one of the main objects of probabilistic-statistical investigation-Markov chains [11]. In this paper the mathematical apparatus describing the direct delayed reactions using the discrete-time Markov chains is considered.

The experiments with using a modified method of direct delayed reactions made it possible to observe the learning process of the animals along with establishing the maximal delay and identifying an optimal algorithm of minimum errors and maximal reward. Here we will discuss the development of optimal algorithms and the dynamics of variability at delays of different duration.

2. Methods

2.1. Subjects

31 albino rats of both sexes (with an average weight of 150 g) have been examined. The animals were indivi-

dually housed in stainless steel cages in conditions of natural light-dark cycle and temperature of $20^{\circ}\text{C} \pm 1^{\circ}\text{C}$. The rats had free access to food and water throughout the whole experiment.

2.2. Apparatus

The rats were tested in a 25 cm-walled wooden T-maze (64 cm in length). The start compartment arm (47 cm in length) joined the two goal arms, each of which was 17 cm in width. Wooden feeders were situated at the far end of each goal arm. The floor under the feeders was electrified. The light and audio signal sources were attached to the top of starting compartment to study rat spatial learning through the different behavioral tasks using different experimental schemes (Figure 1).

2.3. Procedure

The experiments were conducted on white rats using a modified method of direct delayed reactions [16]. Rats were trained in a T-maze-based spatial learning task that required animals to make trial-by-trial choices contingent upon their previous experience. The aim of the experiment was to fixate the complex perception of food in conditions of two feeders. Ten trials were conducted daily with a strictly defined time-spatial program, with strictly defined sequence of signaling feeders, the time interval between trials and the duration of delay. Before the delay, the animal was allowed to move in the experimental cabin without the intervention of the experimenter and obtain food in any feeder defined by the program. The rat could first run up to the feeder with food, and then to an empty one, or *vice versa*, and if it did not return to the starting compartment, the experimenter returned it with force. The animal was allowed to return from the feeder with food without correction to the start compartment so that it did not run to the empty feeder. After the delay the door of start compartment was opened and the animal had the possibility of free behavior. Both direct and indirect reactions were registered. During direct reaction the animal got food, while in case of indirect reaction did not. In the protocol the right move performed by the rat was marked by “1”, and incorrect-by “0”. The distribution of zeroes and ones was recorded in the protocol (Table 1).

Proposed approach gave us the possibility to characterize animal behavior and describe a learning algorithm [16]. For example, the sequence “11001110” means that in pre-delayed behavior (**first 5 digits**), the animal leaves the starting compartment without the interference of the experimenter (1), performs a proper move and runs to the feeder, where it has got food in the previous trial (1). This time it had not received food, since it was not provided by the program (0). So it was necessary to make the correction of the movement to the opposite feeder, as the food was placed there. The animal corrects its movement after the interference of the experimenter (0) and returns to the starting compartment (1). The **last three digits** of the given algorithm describe delayed reactions. The animal itself leaves the starting compartment (1), performs a right move and runs towards the

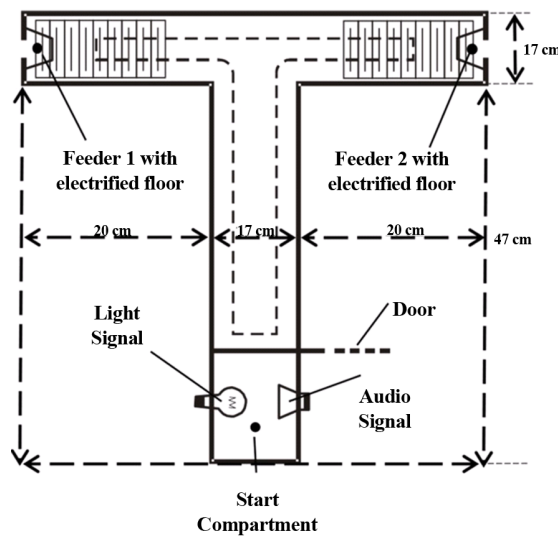


Figure 1. T-maze experimental cabin.

Table 1. The experimental protocol.

Time	Spatial-time Program						Pre-delayed Reactions				Delayed Reactions		
	The sequence of trials	Feeder with food	Intervals between trials in min	Duration of trial in sec	The type of trials: consistent (c) alternate (a)	Exit from the start compartment	The type of move	Food intake	The correction of move	Return to start compartment	Exit from the start compartment	The type of move	Return to start compartment
1	2	3	4	5	6	7	8	9	10	11	12	13	14
	0	1	2	5	-	0	-	1	1	0	1	1	0
	1	1	3	5	c	0	1	1	1	0	1	0	0
	2	2	2	5	a	1	0	1	1	0	0	1	0
	3	1	5	5	a	1	0	1	1	0	1	0	1
11 - 00	4	2	3	5	a	1	0	1	0	1	0	1	0
	5	2	4	5	c	1	1	1	0	1	0	1	0
	6	1	4	5	a	1	1	0	0	1	1	1	0
	7	1	2	5	c	1	1	1	0	1	1	1	1
	8	1	5	5	c	1	1	1	1	0	1	0	1
	9	2	3	5	a	1	1	0	1	1	1	1	1
12 - 24	10	2	2	5	c	1	0	0	1	0	1	1	1

feeder, where before delay it has got reinforcement (1). After this the experimenter returns the animal to the starting compartment (0). **Table 2** demonstrates the dynamics of delayed reactions’ algorithms for pre-delayed behavior and describes the process of optimal behavioral strategy formation. The obtained algorithms are applied for further mathematical analysis.

The schematic diagram of behavioral strategies based on obtained algorithms for one of the experimental days is presented on **Figure 2**. The diagram shows different behavioral strategies identified as bed (1,2,3), medium (4,5,6), good (7,8,9) and the best (10,11,12) learning algorithms.

2.4. Mathematical Description of Behavioral Algorithm

Description of Markov chain is as follows: we have a set of experimental trials (states) $n = \{n_1, n_2, \dots, n_i, n_j\}$. The process starts in one of these states and moves successively from one state to another. Each move is called a step. If the chain is currently in state n_i then it moves to state n_j at the next step with a probability denoted by P_{ij} , and it does not depend upon which states the chain was at before the current state. In Markov chain the probabilities P_{ij} are called transition probabilities. The process can remain in particular state, which occurs with probability P_{ij} . An initial probability distribution, defined on n , specifies the starting state. Usually this is done by specifying a particular state as the starting state of behavioral experiment. We present a behavioral method for estimating these functions.

3. Results and Discussion

The initial state of a system or a phenomenon and a transition from one state to another appear to be principal in the explanation of Markov chain. In our experiments the initial state is 0 or 1; while the transition may occur from 0 to 1, or *vice versa* from 1 to 0, or from 1 to 1.

For studying Markov chains it is necessary to describe the probabilistic character of transitions. It is possible

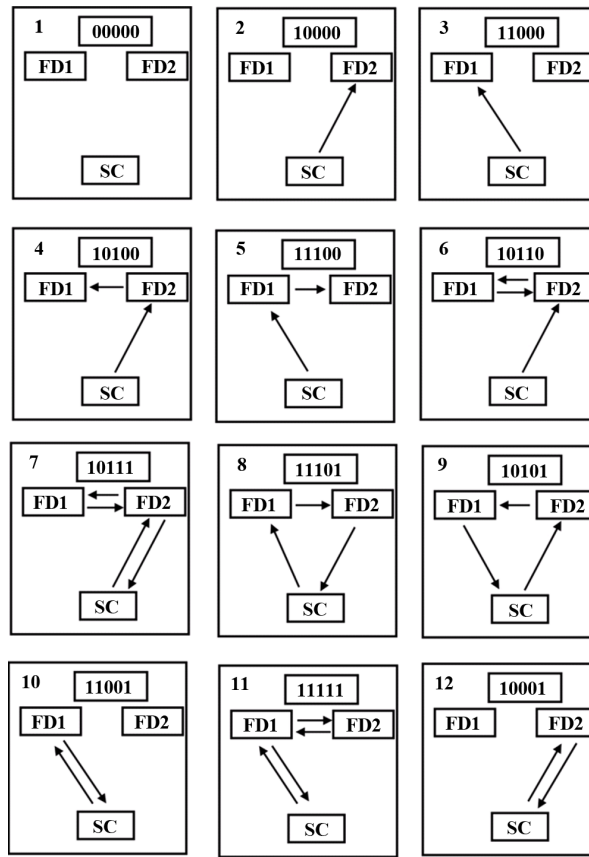


Figure 2. The schematic description of behavioral algorithms. FD1-Feeder 1; FD2-Feeder 2; SC-Start Chamber.

Table 2. The dynamics of delayed reactions' behavioral algorithms.

Trial #	Days									
	1	2	3	4	5	6	7	8	9	10
1	00000	11000	11000	11000	11001	11000	11111	11100	11110	11001
2	01000	10100	11000	10100	11000	11001	11001	11101	11001	10101
3	10010	11110	10100	11111	11001	11101	11001	10101	10111	10101
4	10010	11000	10100	11001	11001	11000	11000	11000	11001	10101
5	11000	11000	11001	10101	11001	11000	11000	11001	11001	11101
6	10100	11110	11110	11001	11001	11000	11000	11001	11111	11101
7	10100	10110	11000	11000	11001	10100	10100	10101	11001	10101
8	11000	11000	11100	11101	11001	10101	10101	11001	11001	10101
9	11000	11000	10100	11001	10101	11001	11001	11001	10101	11001
10	11000	11100	11101	11000	11001	11000	10100	11000	10101	11001

that the time intervals between transitions be permanent.

All states n (in our case $n = 10$) are numbered. If the system is described by Markov chain, then the probability that the system will moves from state i to j ($i = 1, n, j = 1, n$) during the next time interval depends only on the variables i and j and not on the behavior of the system before the transition to state i . In other words, the P_{ij} probability that the system will transit from state i to j does not depend on the type of behavior before state i .

Proceeding from the explanation and features of the probability it is easy to assume that

$$\sum_{j=1}^n P_{ij} = 1; 0 \leq P_{ij} \leq 1.$$

For the modeling the above described experiment, it is convenient to define Markov chain as follows: let's say n number of trials is conducted; if across consecutive trials the conditional probability that any event occurred in trial $n+1$ (we consider the results of previous trials as known) depends only on the result of the last n^{th} trial and does not depend on previous trials, one can say that Markov condition holds. The observed process is called "Markov chain"—random process with discrete (finite) state-space.

Let us assume that the observed behavioral process appears to be random chain with Markov properties, the possible values of which are 0 and 1, and the transition probabilities are determined (estimated) using obtained empirical frequencies. As the initial state is 0 or 1, the transition may occur from 0 to 1, from 0 to 0, or from 1 to 1. We consider the question of determining the probability that, given the chain is in state i , it will be in state j across behavioral treatment. Simply, if we will know the probability that the result of the first trial is 0, then we can define the probability that during the trial n^{th} result will also be 0

Let the conditional probability for transition from initial state 0 for trial n to state 1 in trial $n+1$ be written as:

$$a = P\{1_{n+1}/0_n\}; \text{ analogously,}$$

$$b = P\{1_{n+1}/1_{n+1}\}.$$

It should be noted that the events $A_n = \{\text{the result of trial } n^{\text{th}} \text{ is } 0\}$ and $\bar{A}_n = \{\text{the result of trial } n^{\text{th}} \text{ is } 1\}$ create a complete system of events (*i.e.* a system of such events, one of which necessarily occurs in any trial and any two different events cannot occur simultaneously). Therefore, the event $A_{n+1} = \{(n+1), \text{ the result of the trial is } 0\}$ can be denoted as follows:

$$A_{n+1} = A_{n+1} \cdot A_n + A_{n+1} \cdot \bar{A}_n.$$

It is easy to determine the probability $P(A_{n+1})$ of the event A_{n+1} using total probability formula:

$$P(A_{n+1}) = P(A_{n+1} \cdot A_n + A_{n+1} \cdot \bar{A}_n) = P(A_{n+1} \cdot A_n) + P(A_{n+1} \cdot \bar{A}_n) = P(A_n)P(A_{n+1}/A_n) + P(\bar{A}_n)P(A_{n+1}/\bar{A}_n),$$

where $P(A_{n+1}/A_n)$ and $P(A_{n+1}/\bar{A}_n)$ are conditional probabilities, *i.e.* the transition probabilities of Markov chain.

Now we need to recall formulas:

$$a = P(\bar{A}_{n+1}/A_n), b = P(\bar{A}_{n+1}/\bar{A}_n), P(A_n) = P_n, \text{ i.e. } P(\bar{A}_n) = 1 - P_n,$$

which give recurrence equation for the reliable probability:

$$P_{n+1} = aP_n + b(1 - P_n). \quad (1)$$

In the general case $|a-b| < 1$. Therefore $|a-b|^n \rightarrow 0$ when $n \rightarrow \infty$. As $P(A_n) = P_{n+1}$ when $n \rightarrow \infty$, from the equation (1) we gain formula:

$$P_n = aP_n + b - bP_n.$$

In the case $n \rightarrow \infty$, we obtain:

$$P_n = P_{n+1} = \frac{1}{1-a+b}.$$

Analogously, when $n \rightarrow \infty$ $P(\bar{A}_n) \rightarrow 1 - p = q$. The probabilities p and q are called "final" probabilities. They depend on transition probabilities only and do not depend on the initial state. For a big n we can assume that the process is "balanced"—it's probabilities (independently from n) approximately equal to p or q . This feature of Markov process (chain) is very important in applied science and bears the name of ergodic.

In the general case finding solutions to Equation (1) are very difficult, but in case when P_n does not depend on n , i.e. $P_n = P$, stationary equation can be written:

$$P = aP + b(1 - P). \quad (2)$$

The solution of stationary Equation (2) is simple:

$$P = \frac{1}{1 - a + b}. \quad (3)$$

As solution of stationary Equation (2) is known, we can get the general solution. We only have to suppose that

$$P_n = P + P_n^*. \quad (4)$$

If we substitute (4) into the Equation (1) for P_n , we get:

$$P + P_{n+1}^* = aP + aP_n^* + b(1 - P) - bP_n^*.$$

From the stationary Equation (2) the following recurrence equation is correct

$$P_{n+1}^* = (a - b)P_n^*, \text{ from which}$$

$$P_{n+1}^* = (a - b)^n P_1^*.$$

According to assumption (4) we get $P_n^* = P_n - P$ which gives possibility to move from P_n^* to P_n .

$$P_{n+1} - P = (a - b)^n (P_1 - P), \text{ or } -P + P_{n+1} = P + (P_1 - P)(a - b)^n, \text{ and}$$

$$P_{n+1} = \frac{b}{-a + b - 1} + \left(P - \frac{b}{-a + b + 1} \right) (a - b)^n. \quad (5)$$

The Equation (5) can be used to predict the experimental results: if we know values for probabilities a , b and P_1 for the first n number of trials, we can calculate the probability P_{n+1} of getting 0 the result of trail $n + 1$ will be.

The probabilities a , b and P_1 can be estimated by corresponding empirical frequencies.

To illustrate this, let's discuss in detail an algorithm of calculation of the probability (prognosis) P_{n+1} using the Equation (5) on the basis of the data of 7th column ($n = 10$) of **Table 1**.

1) $a = 0.1$; $b = 0.9$; $P_1 = 0.1$;

$$P_{11}^7 = \frac{0.9}{1 - 0.1 + 0.9} \left(0.1 - \frac{0.9}{1 - 0.1 + 0.9} \right) \times (0.1 - 0.9)^{10}.$$

$$P_{11}^7 = 0.456.$$

Analogically, the following results are obtained for 8th - 14th columns of **Table 1**:

2) $a = 0.1$; $b = 0.6$; $P_1 = 0.4$;

$$P_{11}^8 = 0.400.$$

3) $a = 0.1$; $b = 0.6$; $P_1 = 0.3$;

$$P_{11}^9 = 0.400.$$

4) $a = 0.1$; $b = 0.6$; $P_1 = 0.4$;

$$P_{11}^{10} = 0.400.$$

5) $a = 0.2$; $b = 0.3$; $P_1 = 0.5$;

$$P_{11}^{11} = 0.300.$$

6) $a = 0.2$; $b = 0.4$; $P_1 = 0.3$;

$$P_{11}^{10} = 0.300.$$

7) $a = 0.3$; $b = 0.4$; $P_1 = 0.3$;

$$P_{11}^{13} = 0.370.$$

8) $a = 0.3$; $b = 0.1$; $P_1 = 0.6$;

$$P_{11}^{14} = 0.130.$$

4. Conclusion

In this paper, the original modified method of the direct delayed reaction on the basis of results of the Markov chains theory is developed. The proposed mathematical apparatus allow to calculate the probability (prognosis) P_{n+1} that the results of the following trail $n + 1$ will be 0 if the estimates of the probabilities a , b and P_1 for the first n number of trials are known. It should be noted that the probabilities of possible events of the theoretically calculated reactions coincide with the experimental data. It gives us an opportunity to use widely the above described method in neurophysiological investigations. In addition, it is possible to use them also for delayed reactions carried out by indirect and alternate methods. If we imagine everyday experimental results as one trial, it will be possible to make a long-term prognosis of the animals' behavior.

References

- [1] Hines, W.G.S. (1983) Analysing Qualitative Behavioural Data: A Markov Chain Aid. *Applied Animal Ethology*, **11**, 111-121. [http://dx.doi.org/10.1016/0304-3762\(83\)90120-7](http://dx.doi.org/10.1016/0304-3762(83)90120-7)
- [2] Macdonald, I.L. and Raubenheimer, D. (1995) Hidden Markov Models and Animal Behaviour. *Biometrical Journal*, **37**, 701-712. <http://dx.doi.org/10.1002/bimj.4710370606>
- [3] Granovskiy, B. (2012) Modeling Collective Decision-Making in Animal Groups. Department of Mathematics, Uppsala Dissertations in Mathematics, Uppsala.
- [4] Houillon, A., Lorenz, R.C., Boehmer, W., Rapp, M.A., Heinz, A., Gallinat, J. and Obermayer, K. (2013) The Effect of Novelty on Reinforcement Learning. *Progress in Brain Research*, **202**, 415-439. <http://dx.doi.org/10.1016/B978-0-444-62604-2.00021-6>
- [5] Mathew, B., Bauer, A.M, Koistinen, P., Reetz, T.C., Léon, J. and Sillanpää, M.J. (2012) Bayesian Adaptive Markov Chain Monte Carlo Estimation of Genetic Parameters. *Heredity*, **109**, 235-245. <http://dx.doi.org/10.1038/hdy.2012.35>
- [6] Smith, A.C., Wirth, S., Suzuki, W.A. and Brown, E.N. (2007) Bayesian Analysis of Interleaved Learning and Response Bias in Behavioral Experiments. *Journal of Neurophysiology*, **97**, 2516-2524. <http://dx.doi.org/10.1152/jn.00946.2006>
- [7] Tejada, J., Bosco, G.G., Morato, S. and Roque, A.C. (2010) Characterization of the Rat Exploratory Behavior in the Elevated Plus-Maze with Markov Chains. *Journal of Neuroscience Methods*, **193**, 288-295. <http://dx.doi.org/10.1016/j.jneumeth.2010.09.008>
- [8] Von Sydow, M., Hagmayer, Y. and Meder, B. (2015) Transitive Reasoning Distorts Induction in Causal Chains. *Memory & Cognition*, in press.
- [9] Zin, T.T., Tin, P., Toriu, T. and Hama, H. (2012) A Series of Stochastic Models for Human Behavior Analysis. *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Seoul, 14-17 October 2012, 3251-3256. <http://dx.doi.org/10.1109/icsmc.2012.6378292>
- [10] Solway, A. and Botvinick, M.M. (2012) Goal-Directed Decision Making as Probabilistic Inference: A Computational Framework and Potential Neural Correlates. *Psychological Review*, **119**, 120-154. <http://dx.doi.org/10.1037/a0026435>
- [11] Blaettler, F., Kollmorgen, S., Herbst, J. and Hahnloser, R. (2011) Hidden Markov Models in the Neurosciences. In: Dymarski, P., Ed., *Hidden Markov Models, Theory and Applications*, InTech Publisher, Rijeka, 69-186. <http://dx.doi.org/10.5772/601>
- [12] Hanks, E.M., Hooten, M.B. and Alldredge, M.W. (2015) Continuous-Time Discrete-Space Models for Animal Movement. *The Annals of Applied Statistics*, **9**, 145-165. <http://dx.doi.org/10.1214/14-AOAS803>
- [13] Arantes, R., Tejada, J., Bosco, G.G., Morato, S. and Roque, A.C. (2013) Mathematical Methods to Model Rodent Behavior in the Elevated Plus-Maze. *Journal of Neuroscience Methods*, **220**, 141-148. <http://dx.doi.org/10.1016/j.jneumeth.2013.04.022>
- [14] Khodadadi, A., Fakhari, P. and Busemeyer, J.R. (2014) Learning to Maximize Reward Rate: A Model Based on Semi-Markov Decision Processes. *Frontiers in Neuroscience*, **8**, 101. <http://dx.doi.org/10.3389/fnins.2014.00101>
- [15] Linderman, S.W., Johnson, M.J., Wilson, M.A. and Chen, Z. (2014) A Nonparametric Bayesian Approach to Uncovering Rat Hippocampal Population Codes during Spatial Navigation. CBMM Memo, 027. <http://arxiv.org/abs/1411.7706>
- [16] Stamateli, A. and Tsagareli, S. (2003) Evaluation of Optimum Algorithm by Testing on Modified Direct Delayed Reaction. *Proceedings of the Georgian Academy of Sciences*, **1**, 70-73.
- [17] Ito, M. and Doya, K. (2011) Multiple Representations and Algorithms for Reinforcement Learning in the Cortico-Basal Ganglia Circuit. *Current Opinion in Neurobiology*, **3**, 368-373. <http://dx.doi.org/10.1016/j.conb.2011.04.001>
- [18] Schliehe-Diecks, S., Kappeler, P.M. and Langrock, R. (2012) On the Application of Mixed Hidden Markov Models to

- Multiple Behavioural Time Series. *Interface Focus*, **2**, 180-189. <http://dx.doi.org/10.1098/rsfs.2011.0077>
- [19] Zilli, E.A. and Hasselmo, M.E. (2008) The Influence of Markov Decision Process Structure on the Possible Strategic Use of Working Memory and Episodic Memory. *PLoS ONE*, **3**, e2756. <http://dx.doi.org/10.1371/journal.pone.0002756>
- [20] Lloyd, K., Becker, N., Jones, M.W. and Bogacz, R. (2012) Learning to Use Working Memory: A Reinforcement Learning Gating Model of Rule Acquisition in Rats. *Frontiers in Computational Neuroscience*, **6**, 87. <http://dx.doi.org/10.3389/fncom.2012.00087>
- [21] Huh, N., Jo, S., Kim, H., Jung, H.S. and Min, W.J. (2009) Model-Based Reinforcement Learning under Concurrent Schedules of Reinforcement in Rodents. *Learning and Memory*, **16**, 315-323. <http://dx.doi.org/10.1101/lm.1295509>
- [22] Chen, Z., Kloosterman, F., Brown, E.N. and Wilson, M.A. (2012) Uncovering Spatial Topology Represented by Rat Hippocampal Population Neuronal Codes. *Journal of Computational Neuroscience*, **33**, 227-255. <http://dx.doi.org/10.1007/s10827-012-0384-x>
- [23] Gomes, C.F., Brainerd, C.J., Nakamura, K. and Reyna, V.F. (2014) Markovian Interpretations of Dual Retrieval Processes. *Journal of Mathematical Psychology*, **59**, 50-64. <http://dx.doi.org/10.1016/j.jmp.2013.07.003>
- [24] Huang, Y. and Rao, R.P.N. (2013) Reward Optimization in the Primate Brain: A Probabilistic Model of Decision Making under Uncertainty. *PLoS ONE*, **8**, e53344. <http://dx.doi.org/10.1371/journal.pone.0053344>
- [25] Vevea, J.L. (2006) Recovering Stimuli from Memory: A Statistical Method for Linking Discrimination and Reproduction Responses. *British Journal of Mathematical and Statistical Psychology*, **59**, 321-346. <http://dx.doi.org/10.1348/000711005X73754>
- [26] Whitehead, H. and Jonsen, I.D. (2013) Inferring Animal Densities from Tracking Data Using Markov Chains. *PLoS ONE*, **8**, e60901. <http://dx.doi.org/10.1371/journal.pone.0060901>
- [27] Costa, A.A., Roque, A.C., Morato, S. and Tinós, R. (2012) A Model Based on Genetic Algorithm for Investigation of the Behavior of Rats in the Elevated Plus-Maze. *Intelligent Data Engineering and Automated Learning—IDEAL, the Series Lecture Notes in Computer Science*, **19**, 151-158. http://dx.doi.org/10.1007/978-3-642-32639-4_19
- [28] McKellar, A.E., Roland, L.R., Walters, J.R. and Kesler, D.C. (2014) Using Mixed Hidden Markov Models to Examine Behavioral States in a Cooperatively Breeding Bird. *Behavioral Ecology*, **26**, 148-157. <http://dx.doi.org/10.1093/beheco/aru171>
- [29] Rao, R.P.N. (2010) Decision Making under Uncertainty: A Neural Model Based on Partially Observable Markov Decision Processes. *Frontiers in Computational Neuroscience*, **4**, 146. <http://dx.doi.org/10.3389/fncom.2010.00146>
- [30] Dileep, G. and Hawkins, G.D. (2009) Towards a Mathematical Theory of Cortical Micro-circuits. *PLoS Computational Biology*, **5**, e1000532.
- [31] Zhang, J. (2009) Adaptive Learning via Selectionism and Bayesianism, Part II: The Sequential Case. *Neural Networks*, **22**, 229-236. <http://dx.doi.org/10.1016/j.neunet.2009.03.017>