# Attractor learning with predictor networks

Mikhail Burtsev
NRC “Kurchatov Institute”,
P.K. Anokhin Institute of Normal Physiology of RAMS,
M.V. Keldysh Institute of Applied Mathematics of RAS
Moscow, Russia

Adaptive behavior requires a robust execution of specific action sequences as well as a fast switching between them. This type of sequential dynamics can be implemented by the system with a collection of attracting metastable sets in the phase space joined by heteroclinic orbits [1], [2] which is called stable heteroclinic sequence (SHS). Design of such SHS for a particular problem consists of defining metastable sets and their basins of attraction. This is non-trivial task for the majority of real world applications even with the full knowledge of necessary sequences of actions. The situation is much worse when only partial information about solution space is available or some parameters change with time. In this case the system has to learn some new sequences to preserve functionality.

The problem of learning for the dynamical system can be broken on following parts [3]. The first part is an initiation of learning. The system should “know” a target trajectory in a phase space and constantly estimate deviation from it. The second part is a search of new action sequence that ends up at the target attractor. And the last part is saving of acquired adaptive trajectory as a supplement to the already existing trajectory. As a result of these steps, with every learning episode the system should develop more and more complex behavior underlined by progressively refined structure of metastable sets and their basins of attraction.

Principles of attractor learning outlined above were used for the development of new predictor network architecture called functional systems network. Every node of the network is functional system that activates to participate in goal-directed control and simultaneously predicts outcome of its’ own activation. Learning is implemented as an addition of new nodes to the network (fig. 1(c)). The benchmarking of the architecture demonstrated adaptive learning in single and multi-goal hypercube environments (fig. 1 (a)-(c)) [4]. The performance was comparable to Q-learning in the rooms problem (fig. 1(d)) and superior in non-stationary two-arms maze (fig. 1(e)-(f)).

**References**

[1] V. S. Afraimovich, V. P. Zhigulin, и M. I. Rabinovich, «On the origin of reproducible sequential activity in neural circuits», *Chaos Interdiscip. J. Nonlinear Sci.*, т. 14, с. 1123, 2004.

[2] V. S. Afraimovich, M. I. Rabinovich, и P. Varona, «HETEROCLINIC CONTOURS IN NEURAL ENSEMBLES AND THE WINNERLESS COMPETITION PRINCIPLE», *Int. J. Bifurc. Chaos*, т. 14, № 04, с. 1195–1208, апр. 2004.

[3] M. Burtsev, «Basic principles of adaptive learning through variation and selection», в *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, 2008, с. 88–93.

[4] M. A. Komarov, G. V. Osipov, и M. S. Burtsev, «Adaptive functional systems: Learning with chaos», *Chaos Interdiscip. J. Nonlinear Sci.*, т. 20, № 4, с. 045119, 2010.

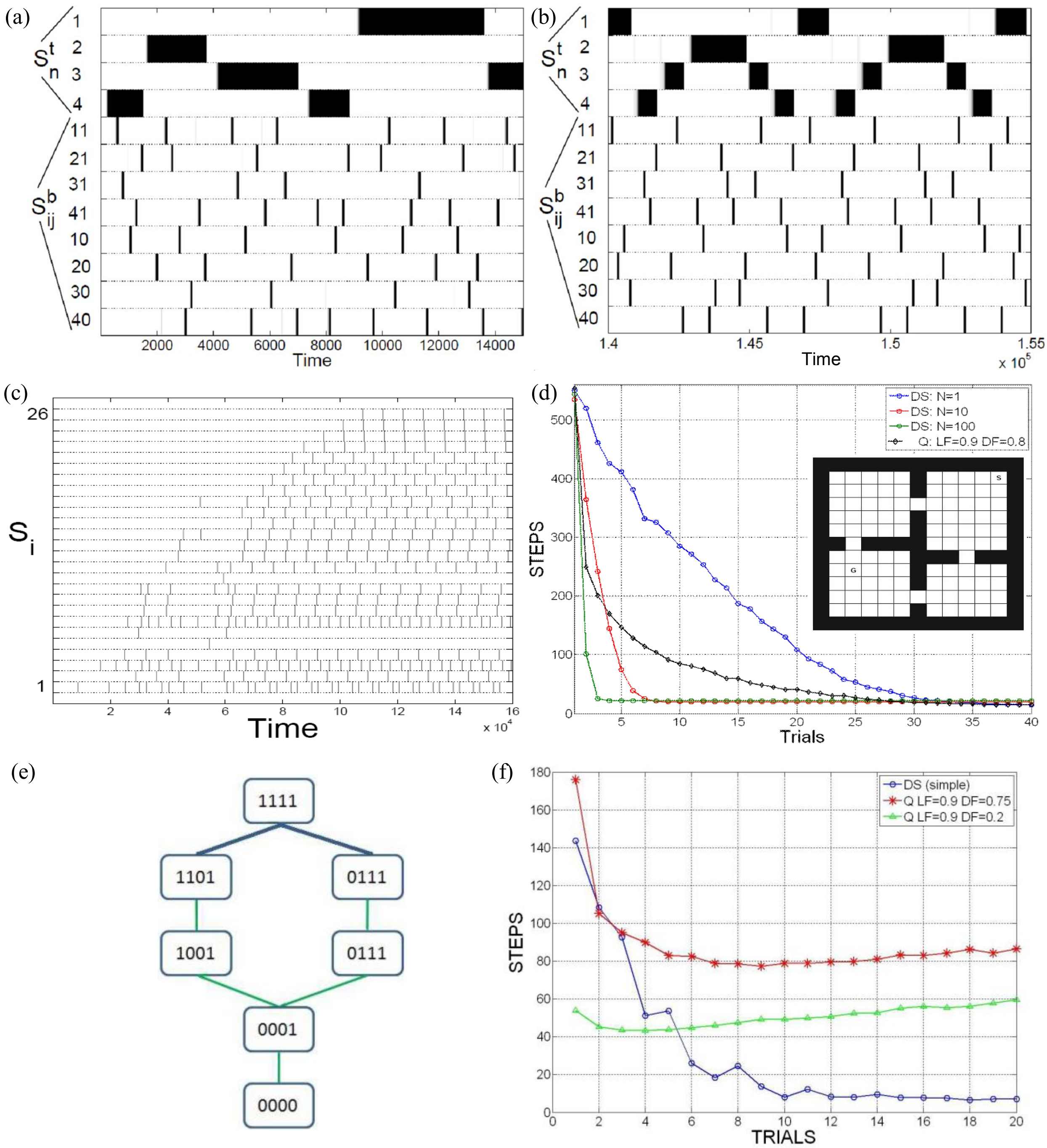


**Fig. 1.** (a) The naive agent in the hypercube environment with multiple goals. Chaotic activations of the action functional systems (FSs) prevail. (b) Behavior of the agent after learning [the same simulation run as (a)]. During the period of adaptation, the secondary systems start to contribute by controlling the selection of goal-directed actions. After learning, the agent reaches the goal states and satisfies motivations (decrease in activity of motivational functional systems $S_n^t$) periodically and much faster compared to (a). (c) A number of active secondary functional systems grow when the agent is learning and then become stable [the same simulation run as (a) and (b)]. (d) The rooms problem (inset). S: start; G: goal. Learning curves for the rooms problem. DS is FS-learning where N is the upper bound on the number of functional systems added to the predictor network during every learning episode. Q is Q-learning, LF is learning rate α and DF is discount factor γ. (e) Non-stationary two-arms maze. Transitions marked with blue lines are concurrently available with equal probabilities. (f) Learning curves for the non-stationary two-arms maze (DS: N=1).