Logical probabilistic biologically inspired cognitive architecture[1]

E.E. Vityaev[1], A.V. Demin[1], Y.A. Kolonin[2]

[1]Sobolev institute of mathematics SD RAS, Novosibirsk, Russia
[2]Novosibirsk State University, Mathematics and Mechanics Department, Russia
vityaev@math.nsc.ru

Abstract. We consider a task-oriented approach to AGI, when any cognitive problem, perhaps superior a human ability, have sense given a criterion of its solution. In the frame of this approach, we consider the task of purposefulness behavior in a complex probabilistic environment, where behavior is organized through self-learning. For that purpose we consider the cognitive architecture based on the Functional Systems Theory (TFS). The architecture is based on the main notions of this theory: goal, result, anticipation of the result. The logical structure of this theory was analyzed and used for the control system of the purposeful behavior development. This control system contains the hierarchy of functional systems that organize the purposeful behavior. The control system was used for the agents modeling that solve the foraging task.

Keywords: Architecture, Functional Systems Theory, adaptive control system, purposefulness behavior, goal-directness behavior

1 Introduction

At the moment there is no unitary approach to AGI development. Maybe the most popular approach for now is neural networks of different kind. While this approach is widespread and practically useful for some “intellectual” tasks, it still has its well-known flaws: huge amount of data needed for the network to become effective, high computational cost and the infamous "black box problem", preventing us from understanding how the result of calculations was obtained. All these problems compel us to look for other approaches.

We consider a task-oriented approach to AGI, when any cognitive problem, perhaps superior a human ability, have sense given a criterion of its solution. In the frame of this approach, we consider the task of purposefulness behavior in a complex probabilistic environment, where behavior is organized through self-learning.

Purposefulness behavior was deeply studied in the USSR and Russia in the framework of the Theory of Functional Systems (TFS), which describes the organization of purposefulness behavior aimed at satisfying a certain need [1]. In this theory

[1] The first and second authors are financially supported by the Russian Science Foundation grant #17-11-01176 in part concerning the formal model and the first author also by the Russian Foundation for Basic Research # 19-01-00331-a in other parts.

the elaboration of an action plan to achieve the goal is carried out on the basis of existing experience by predicting the achievement of the goal and all its subgoals, organized hierarchically. This prediction, even before any action begins, is accompanied by the formation of a mechanism controlling the achievement of the goal and subgoals by the corresponding groups of receptors responsible for recording the achievement of the subgoals and the goal. These groups of receptors form a certain complex receptor for achieving the subgoals and goals, which is called the acceptor of action results. Thus, TFS is quite consistent the with task-oriented approach to AGI and, in addition, it was worked out in detail and experimentally confirmed.

In this paper, we present a formalization of TFS based on logical-probabilistic learning driven by detecting the most specific rules of behavior. Prediction of achieving goals and sub-goals is carried out by an inductive-statistical inference of predictions based on these most specific rules. Such rules have a number of important properties. Firstly, they can be detected by special logical-probabilistic neurons that satisfy the Hebb rule [2]. Secondly, their predictions in accordance with the inductive-statistical conclusion are consistent [3]. The preference of a particular actions plan is carried out taking into account the probability of predicting the goal achievement. This model may be implemented in the frame of probabilistic programming [4].

In our approach we can see some parallels to Jeff Hawkins's Hierarchical Temporal Memory (HTM), as it is also based on prediction and biologically inspired. But with regard to the organization of purposeful behavior, TFS has been worked out in much more detail. Our system is more structurally simple due to the difference in mathematical foundations and actions plan, based purely on a prediction with a highest probability of the goal achievement.

Another relatively close approach is SOAR [5], a classic architecture that solves multiple tasks including the purposefulness behavior. Its inference is also based on "if-then" rules, but not on probabilistic predictions.

Our architecture is not only plausible from a biological point of view, but also quite effective: it learns to explore the environment and achieve goals in it much faster than reinforcement learning and neural networks. Also it can achieve more complex, two-stage goals in the same environment, when the classic approaches cannot do anything with. The results of experiments confirming this are presented in section 5.

2 The Theory of Functional Systems of Brain Function

The theory of functional systems (TFS) developed by P.K. Anokhin and many other distinguished scientists of his school is, at the moment, one of the few known theories in which the concepts of goal, purpose, result, and goal-directed activity are principal ones and which exposes the physiological mechanisms that implement these concepts.

Desire is not passive. It makes no sense to desire if there is no possibility to get closer to satisfying the desire by some actions or activity. Desire is active, but meaningless without purposefulness – it causes the organism to be active and display some behavior in order to satisfy it. Thus the concept of goal emerges. Activity and actions are always goal-directed. If there is no goal for an action, it is unclear when it should be terminated. Let us define the goal as such an activity/behavior that is aimed at satisfying certain criteria. A goal cannot be attained without having a criterion of its attainment; otherwise we can always assume that the goal has already been attained. Such a definition of goal allows us to define the result of attaining the goal as that, what we obtain by meeting the criterion and attaining the goal (fulfilling the desire). Between the concepts of goal and result, the following relationship holds: the result is obtained when the goal is attained and the criterion of its availability is "triggered". But when the goal is being set, we have the goal but not the result.

The definition of goal is paradoxical since the activity/behavior of satisfying some criteria does not essentially presuppose knowledge of how to attain the goal; you can set a goal without defining either how it can be attained, or by what means, or when. This paradoxicalness of the goal concept we call the goal paradox. For the paradox solution we need an experience. As will be seen later on, in the framework of the theory of functional systems, brain activity during goal-directed behavior is seen as being constantly occupied by solving the goal paradox, and determining by what means, when, and how to attain goals.

Let us proceed to outline the theory of functional systems, in which the concepts of goal, result, and goal-directed activity are principal ones, and which analyzes the physiological mechanisms of these concepts.

The theory of functional systems (TFS) is a theory of systems, whose function is to attain goals (satisfy needs) by solving the goal paradox. Therefore, we will outline the theory of functional systems as a theory of solving goal paradoxes, and describe how the brain determines by what means, when, and how goals can be attained.

As achieving results consists in satisfying some criteria, this achievement should be registered in some way. In the physiological sense, what constitutes a criterion for registering the attainment of a result? According to P.K. Anokhin, this is physiologically realized by a “special receptor apparatus” [7]. The signaling of this receptor apparatus about obtaining a result (i.e., on the lack of deviation from the optimal level of metabolism) and attaining the goal is called reverse afferentation.

Now we can explain, within the framework of TFS, how goals are being physiologically set by the organism. An organism needs to constitute a goal in TFS. The goal (and its attainment criterion) firstly signals by means of reverse afferentation that there is a lack of some need; secondly, it sets a goal to wait for a signal, indicating that the results have been attained; and thirdly, it provides energy and actually forces the organism to attain the goal. Thus, the physiological mechanism of goal-setting in fact consists of the emergence of a need.

The interaction of different goals and results is organized in several ways according to TFS: by the “principle of the dominant”, by the “hierarchy of results” and by “result models”.

3 Central mechanisms of functional systems

“According to P.K. Anokhin, the central mechanisms of functional systems that support goal-directed behavioral acts have a similar structure” [7]. Let us examine in detail the architecture of goal-directed activity, as well as the physiological mechanisms of solving the goal paradox.

Afferent synthesis. The afferent synthesis, which includes the synthesis of motivational excitation, memory, contextual and triggering afferentation, constitutes the initial stage of a behavioral act of any complexity. ***Motivational Excitation***. As we know, the goal is set by an emerging need. But in case of goal-directed behavior, it transforms into a motivational excitation. But a motivational stimulus does not consist in the excitation of receptors which stand 'on guard' for some physiological constant – it is rather an excitation of 'central brain structures' initiated by the arising need. It is the motivational stimulus that constitutes the goal set in the organism in case of goal-directed behavior. As in the case of needs, the motivational stimulus not only sets a goal but also energetically supports its attainment. ***Memory***. The whole sequence of stimuli that has led to goal attainment is recorded during reinforcement, starting with the motivational stimulus. Motivational stimulus extract from memory all previous sequences of actions which have led to attaining the result. ***Situational Afferentation***. While recording a memory trace, the situation in which the result is attained is also being recorded. This situation is registered, along with the motivation, as a necessary precondition for attaining the result. Thus, the motivational stimulus in this situation "extracts from memory" only those ways of attaining the goal that are possible in the given situation. ***Triggering Afferentation***. The fourth component of afferent synthesis is the triggering afferentation. It is essentially the same as the situational afferentation with the difference that it involves the time and place of attaining the result.

Consequently, the goal paradox is solved for the most part during afferent synthesis, as it's here that the "what", "how", and "when" of goal attainment are determined. Therefore, taking experience and environment into account, the motivational excitation as a goal automatically solves the goal paradox and determines by what means, how, and when can the goal be attained.

Decision-making. At the stage of afferent synthesis, motivational excitation can extract from memory several ways of attaining the goal. At the stage of decision making, only one of them is selected – thus forming the "program of actions".

Acceptor of Action Results. Suppose a program of actions is chosen. At that point, there is no guarantee yet that the final result will necessarily be attained, nor even intermediate ones. The goal can only be attained if each of the intermediate results of the current program of actions will be attained. Motivational excitation "extracts from memory" the entire sequence and the hierarchy of results that should be attained during the program of actions. This sequence and hierarchy of results are defined in TFS as the ***acceptor of action results***. Therefore, while being transformed into a particular goal, the motivational excitation extracts from the memory a particular criteria of this goal attainment. This consists of the whole sequence and the hierarchy of criteria of results which must be attained in the process of attaining the goal and performing the program of actions, i.e. the acceptor of action results. Thus, the acceptor of action results anticipates the particular criteria of attaining the goal.

Transforming motivational excitation as a goal into a particular goal transforms the original paradoxical goal – for which it is not determined by what means, how, and when it can be attained – into a non-paradoxical particular goal, for which the final goal (and result) is divided into sub-goals (and sub-results), so that for each sub-goal it is already known by what means, how, and when it can be attained.

4 Formal model of TFS

Now let us assume that our model constitutes the control system of some animat that operates in discrete time $t = 0, 1, ...$ as it was done in [6]. Suppose the animat has a set of sensors $S_1, ..., S_n$ which characterize both the state of the animat itself and of external environment. Each sensor S_i has a set of possible indications VS_i. The animat also has a set of available actions in the environment $A = \{a_1, ..., a_m\}$. Any action that animat performs at a moment t_i may result at a moment $t_i + 1$ in some changes in the environment, and, consequently, in its sensors indications. Since the animat «perceives» the world only through its sensors, then from its point of view the system's state at any given point in time can be written as a vector of all sensors indications $V(t) = (v_1, ..., v_n)$, where $v_i \in VS_i$ is the indications of the the *i*-th sensor at the moment *t*, and the states with same sensor indications are indistinguishable for it. The set of all possible states of the system is denoted by $S = (VS_1 \times VS_2 \times ... \times VS_n)$.
On a set of states of the system $S = (VS_1 \cup VS_2 \cup ... \cup VS_n)$ we define a set of predicates $PS = \{P_1, ..., P_k\}$ each of which is calculated on the basis of sensors indications. Each state of the system can thus be written as a vector $s = (p_1, ..., p_k),\ p_i \in \{0,1\}$ of predicates values from PS where 1 means validity of a predicate and 0 – its falsity. The state may be described by a subset of predicates $s = (p_{i_1}^e, ..., p_{i_e}^e),\ p_{i_1}^e, ..., p_{i_e}^e \subseteq p_1, ..., p_k$. The animat's task is to attain a certain goal. Let us define a goal *Goal* as a state of the system $s_{Goal} = (p_{i_1}^{goal}, ..., p_{i_{goal}}^{goal})$ which it is required to attain. A notation $(p_{i_1}^{goal}, ..., p_{i_{goal}}^{goal})$ means that predicates $p_{i_1}^{goal}, ..., p_{i_{goal}}^{goal}$ should be true when the goal is attained.
Let us clarify concepts of event and history. By an event $e = (s_0, s_e, a)$ we will understand a singular fact of transferring the system from the state $s_0 = (p_1^0, ..., p_k^0)$ into a state $s_e = (p_1^e, ..., p_k^e)$ as a result of an action a and by a history of events – a set of pairs (e_t, t) where $e_t = (s_t, s_{t+1}, a)$ is an event and *t* is a moment in time when this event has occurred.

Let us define a rule *R* that predicts a change of a state(s) after the execution of an action *a* as a transformation $R = (s_0 \xrightarrow[p]{a} s_e)$, where: s_0 – is the initial state of the system $(p_{i_1}^0, ..., p_{i_0}^0)$; s_e – is the final state of the system $(p_{i_1}^e, ..., p_{i_e}^e)$; a – is an action that transforms the initial state into the final one; p – is the probability with which the action transforms the initial state into the final one.

Let us first define a functional system $FSC = (s_{Goal}, R_1, ..., R_n, p_{FSC})$ that realizes one action. Functional system FSC performs transformations $s_0 \xrightarrow[p_{FSC}]{R_1, ..., R_n} s_{Goal}$, where $s_{Goal} = (p_{i_1}^{goal}, ..., p_{i_{goal}}^{goal})$ – is the target state of the functional system, $R_1, ..., R_n$ – are rules of the form $s_0 \xrightarrow[p]{a} s_{Goal}$, using which from various initial states s_0 and

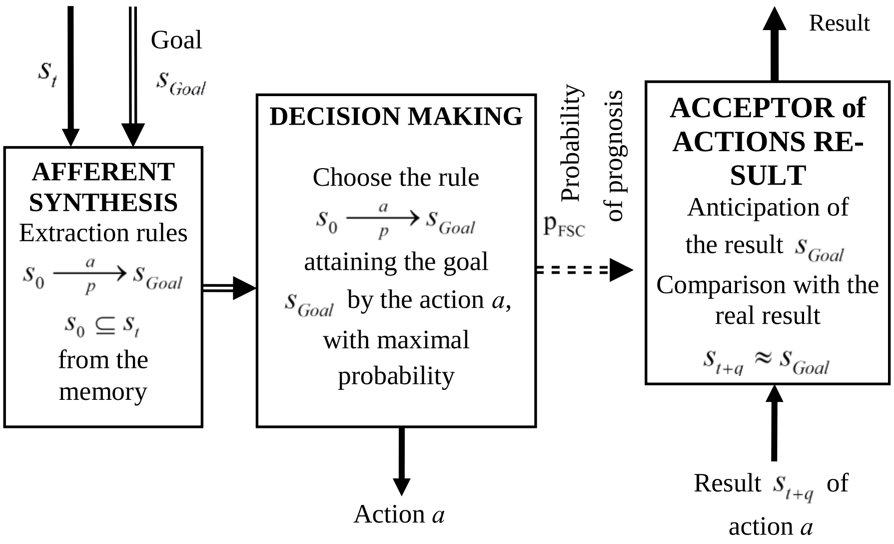

Fig. 1. Functional system that implements sensory corrections.

some action a the system can get to the target state s_{Goal} (fig. 1). An estimate of the probability of attaining a goal by a functional system can be calculated based on the statistics of attaining goals: if n is the number of cases in which a request to attain a goal s_{Goal} was received and m is the number of cases in which the selected rules and sequences/hierarchies of actions led to attaining the goal s_{Goal}, then $p_{FSC} = m/n$.

In general case functional systems are sequences and the hierarchies of the functional systems FSC. A functional system $FS = (s_{Goal}, \langle FSC_1, ..., FSC_n \rangle, p_{FS})$ that combines a sequence of functional systems of the form FSC is defined as:

$$FS = s_0 \xrightarrow[\to s_1 \to s_2 \to ... \to s_{goal} \quad p_{FS} = p_{FSC_1} \bullet ... \bullet p_{FSC_n}]{FSC_1, ..., FSC_n} s_{goal} \text{, where}$$

$$FSC_1 = (s_0 \xrightarrow[p_{FSC_1}]{R_1^1, ..., R_{v_1}^1} s_1), FSC_2 = (s_1 \xrightarrow[p_{FSC_2}]{R_1^2, ..., R_{v_2}^2} s_2) \ ... \ FSC_n = (s_{n-1} \xrightarrow[p_{FSC_n}]{R_1^n, ..., R_{v_n}^n} s_{goal})$$

are functional systems of the type FSC. The goal of the functional system FS is to successively attain goals $s_1 \to s_2 \to ... \to s_{goal}$ using functional systems $FSC_1, ..., FSC_n$ with a resulting probability $p_{FS} = p_{FSC_1} \bullet ... \bullet p_{FSC_n}$.

5 Experiments

For investigation of the control system behavior two experiments were set. We explored the foraging task. In this task some agent explores the area and gathers pabulary objects. There are no subgoals in this task, so in the second phase we have complicated this task by introducing a «tablet» that is needed to eat the pabulary object. In this case the subgoal is eating the tablet before eating the pabulary object.

The virtual world was modeled in which the agent can gather the pabulary objects. This world includes the rectangular field with 25×25 cells. Each cell may be empty or include the «pabulary object» or «barrier». Barriers are placed only on the perimeter of the field. Agent is placed on one of the cells and may be oriented in one of the four directions. The possible actions $\{a_1, a_2, a_3\}$ of the agent are: step on the one cell ahead, turn left, turn right.

In the first experiment some pabulary objects are placed randomly on the field. To eat the pabulary object agent needs to take a step on the cell where the pabulary object is located. In that case the pabulary object disappears from the cell and randomly appears on some other cell.

Agent has sensors $S_1, ..., S_9$, in which $S_1, ..., S_8$ stand for the area around the agent and inform the agent about the objects placed on these cells, and S_9 informs the agent about the object placed on the cell that agent occupies.

The second experiment is more complicated than the first. In the cells of the field a «tablet» objects may be placed, which are randomly distributed. To eat the pabulary object agent needs to have a «tablet» object, which he needs to gather on the field. When the agent eats a pabulary object the gathered «tablet» object disappears and for eating a new pabulary object the agent needs to gather a new «tablet» object. The agent gathers a «tablet» object if it occupies the cell with this object. The agent may gather only one «tablet» object. When agent gathers a «tablet» object the cell becomes empty and a new «tablet» object randomly appears on the field.

In the second experiment agent has ten sensors $S_1, ..., S_9, S_{pill}$, where first nine are the same as in the first experiment and sensor S_{pill} informs the agent about availability of the «tablet» object.

For the estimation of the effectiveness of the control system we compared it with control systems based on the reinforcement learning, described in the work [8]. For comparison we used two control systems based on the Q-Learning. These algorithms consist in consecutive refinement of the estimation of the reward $Q(s_t, a_t)$ summary, if in the state s_t the system acts as a_t:

$$Q^{(i+1)}(s_t, a_t) = Q^{(i)}(s_t, a_t) + \alpha(r_t + \gamma \max_A Q^{(i)}(s_{t+1}, a) - Q^{(i)}(s_t, a_t)).$$

The first system (Q-Lookup Table) uses table, which includes Q-values of all possible states and acts. Initially, the table is fulfilled randomly. Then the system in each act specifies the Q-value.

The second system (Q-Neural Net) uses approximation of the function $Q(s_t, a_t)$ using neural networks. In that case for each act a_i a special neural network is used. In each time period the system chooses an action and neural network produces a greater value of the estimation Q-value. Then the action accomplishes, and weights of the neural nets are changed.

For the estimation of the systems the period of agent functioning was divided on stages for 1000 steps. The estimation consists of the volume of the pabulary objects

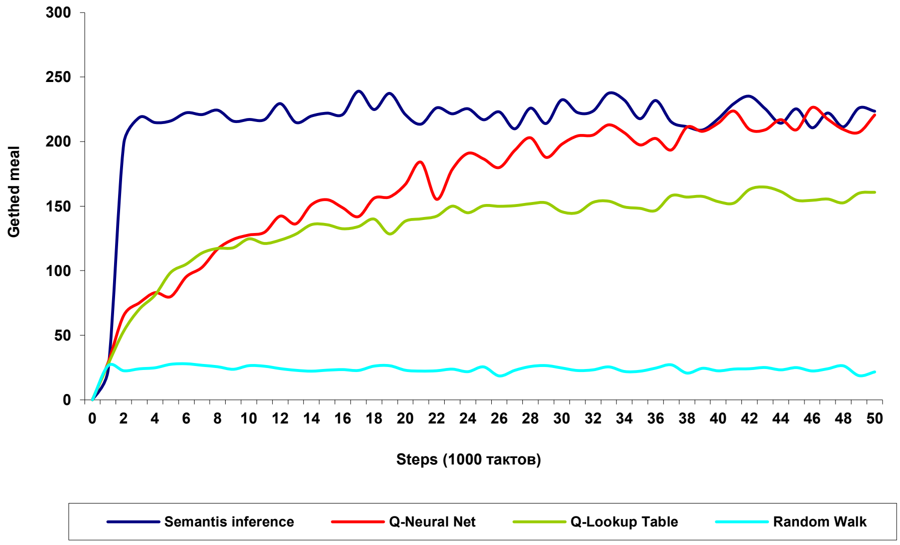

Fig 2. Amount of the food gathered by the agent with different control systems.

gathered for a step of the work. After learning every system reaches some optimal value. During the experiment we can estimate the learning speed and corresponding optimal value.

6 Results of the first experiment

In the first experiment there were 24 predicates for sensors – three predicates $(S_i = empty)$, $(S_i = block)$, $(S_i = food)$ for each sensor S_i, $i = 1,...,8$. At the beginning the control system contained only one functional system with purpose S_{Goal} = $(S_9 = food)$, when sensor S_9 informs about pabulary objects in the central cell.

This experiment had no subgoals. The main task of this experiment is the estimation of the effectiveness of the functional system and its learning.

On the fig. 2 there are results of comparison for different control systems. For each control system the mean values for 20 experiments are presented. The duration of each experiment is 50,000 steps of the agent. The number of pabulary objects on the field is 100.

It is seen from the figure that the control system based on the semantic probabilistic inference is fully learned during the 1000 steps. Control systems based on the neural nets (Q-Neural Net) learn more slowly and become fully learned after nearly 10,000 steps. The slow learning of the Q-Lookup Table follows from the huge number 2496 of states with three possible actions.

The results of this experiment demonstrate that the control system based on the semantic probabilistic inference works rather effectively and more effectively than systems based on the Reinforcement Learning.

7 Results of the second experiment.

The following experiment is principally different as the task may be divided in two parts: at first – to find a «tablet» object and then to find pabulary objects. The purpose of this experiment is to demonstrate the ability of automatic subgoals formation.

The agent now has 32 predicates – four predicates for each sensor S_i, $i = 1,...,8$: $(S_i = empty)$, ($S_i = block$), ($S_i = food$) и ($S_i = pill$) and one predicate ($S_{pill} = yes$) for the state when the agent has a «tablet» object and one predicate $(S_9 = food)$ for the state when the pabulary object is in the central cell under the agent.

At the beginning the control system of the agent has only one functional system with the purpose $S_{Goal} = (S_{pill} = yes) \& (S_9 = food)$, when the agent has a «tablet» object and finds a pabulary object.

During the experiment the control system of the agent had always found the sub-goal $S^2_{Goal} = (S_{pill} = yes)$ and formed a corresponding functional system. When the agent had no tablet in possession, the control system passed the control to the subsystem for the search of a tablet, and, after finding the tablet and achieving the goal S^2_{Goal} = $(S_{pill} = yes)$, the upper level control system started searching the pabulary objects.

The results of the experiment are presented on the figure 3. On the figure the mean values for 20 experiments are presented for each control system. In each experiment the agent had 100,000 steps. The number of pabulary objects and «tablet» objects on the field was 100 for each.

As seen from the figure, the control system based on the semantic probabilistic inference was working more effectively than systems based on the reinforcement learn-

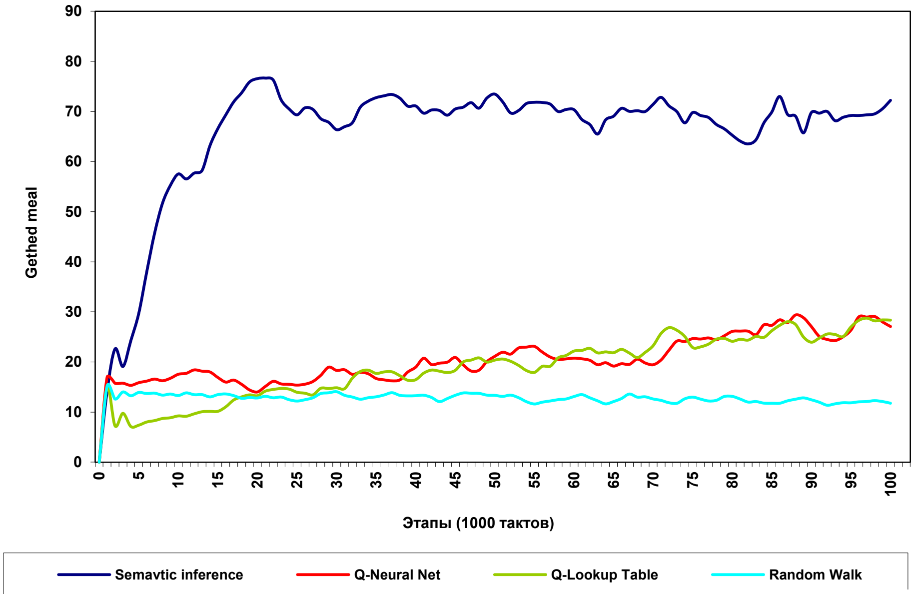

Fig 3. Amount of food gathered by the agent with different control systems in the presence of the «tablet» objects.

ing. Control systems based on the reinforcement learning showed almost no learning ability and worked unstable. They cannot learn the need of the «tablet» objects for the goal achievement during the reasonable time and have passed by «tablet» objects after 100,000 steps of learning. Additional experiments demonstrated that control system (Q-Neural Net) can sometimes learn during the 300,000 – 500,000 steps.

References

1. Anokhin, P.K.: Biology and neurophysiology of the conditioned reflex and its role in adaptive behaviour, Oxford etc.: Pergamon press, pp. 574. (1974).
2. Vityaev E.E. A formal model of neuron that provides consistent predictions. Biologically Inspired Cognitive Architectures 2012. Proceedings of the Third Annual Meeting of the BICA Society (A. Chella, R.Pirrone, R. Sorbello, K.R. Johannsdottir, Eds). In Advances in Intelligent Systems and Computing, v.196, Springer: Heidelberg, New York, Dordrecht, London. 2013, pp. 339-344.
3. Vityaev, E., Odintsov, S. How to predict consistently? // Trends in Mathematics and Computational Intelligence In: Studies in Computational Intelligence, 796, María Eugenia Cornejo (ed), 2019, 35-41.
4. Avi Pfeffer. Practical Probabilistic Programming. Manning Publications, 2016, pp.456.
5. Laird, John E. The Soar Cognitive Architecture. MIT Press, 2012.
6. Muhortov V.V., Khlebnikov S.V., Vityaev E.E.: Improved algorithm of semantic probabilistic inference in task of 2-dimention animat // Neuroinformatics, v. 6(1), pp. 50-62. (in Russian). 2012.
7. Sudakov K.V.: The general theory of functional systems. Moscow: Medicine, p.222. (in Russian). 1984.
8. Sutton R., Barto A. Reinforcement Learning: An Introduction. – Cambridge: MIT Press, 1998. – p. 342.