Models of fish exploratory behavior in mazes

Vladimir G. Red'ko[a,*], Valentin A. Nepomnyashchikh[b], Elena A. Osipova[b]

[a]*Scientific Research Institute for System Analysis, Russian Academy of Sciences, Moscow, Russia*
[b]*I.D. Papanin Institute for Biology of Inland Waters, Russian Academy of Sciences, Borok, Yaroslavl region, Russia*

Abstract. The computer models of fish exploratory behavior in mazes are developed and investigated. These models are inspired by the exploratory behavior of zebrafish, Danio rerio, in mazes. We consider three types of models. Model 1 describes the increase of knowledge acquired by fish about arms of the maze. Model 2 characterizes fish's predictions of future situations in the maze. A fish uses knowledge and predictions to organize its behavior. These two models characterize initial stages of fish exploratory behavior in mazes. Upon completion of these initial stages, the fish obtains some experience of movements in the maze and some knowledge about general features of the maze. Model 3 takes into account this experience. In this hypothetical model, we suppose that after certain exploration of the rather complex maze, the fish is able to form some generalized notions characterizing places in this maze. Using these generalized notions, the fish forms a mental plan of its movement towards the goal situation.

Keywords: modeling animal behavior, exploratory behavior, prediction of future situations, planning behavior.

1. Introduction

The goal of our work is a mathematical modeling and computer simulation of cognitive behavior of zebrafish, Danio rerio, in the course of maze exploration. Our models are based solely on qualitative description of exploratory behavior. Therefore, we describe biological experiments very briefly, focusing only on those qualitative data, which are sufficient for understanding the models.

We design and investigate three types of models. The model 1 describes how the fish acquires knowledge about arms of the maze. The model 2 considers actions of the fish in certain situations in the maze (the situation can be the current arm in which the fish is in the given moment of time). This model characterizes fish's predictions of the next situation for the current situation and action. We believe that the models 1 and 2 characterize initial stages of the fish exploratory behavior in mazes. Upon completion of these initial stages, the fish obtains some experience of movements in the maze and

[*]Corresponding author: Tel.: +7 915 1673584.
E-mail address: vgredko@gmail.com (V.G. Red'ko)

some knowledge about general features of the maze. The model 3 is the hypothetical one; this model is developed for the case of fish movements in a rather complex maze. In this hypothetical model, we suppose that after certain exploration of the maze, the fish is able to form some generalized notions characterizing places in this maze. Using these generalized notions, the fish mentally forms a simple knowledge database; then using this knowledge database, the fish designs a plan of its movement towards the goal situation.

Additionally, we take into account that two opposing trends in the behavior of animals are usually observed (Inglis et al, 2001). One of them is a need for new, unpredictable stimulation, leading to an exploratory behavior, and the other is the desire to predict the results of the behavior. A permanent motivation to obtain information on the environment prevails if basic needs (e.g. the need for food) are satisfied; this corresponds to the hypothesis of the "reduction of uncertainty" (Inglis et al, 1997; Inglis et al, 2001).

These two trends (the aspirations for unpredictable novelty and for predictability) in the behavior of animals are contradictory at first sight. Competition between these trends leads to a balance between novelty and predictability, and results in an efficient exploration of unfamiliar environment. The idea of this balance has been successfully implemented in the control software of robotic dog Aibo (Oudeyer & Kaplan, 2004, 2009).

The paper is organized as follows. Section 2 contains the short description of biological experiments on Danio rerio. Section 3 describes the models 1 and 2 characterizing the knowledge acquisition and formation of predictions by fish during initial stages of maze exploration; the results of computer simulation for these models are also included in this section. Section 4 describes non-trivial hypothetic model 3 that explains how a fish could form a mental plan of movement towards the goal situation in the rather complex maze.

2. Behavioral experiments. Behavior of fish in mazes

Zebrafish behavior was studied in an unfamiliar environment, namely in mazes of two types: the four-arm cross-shaped maze (Fig. 1) and the more complex maze with 11 arms (Fig. 2). A detailed description of experimental procedures and quantitative data will be published elsewhere (Nepomnyashchikh & Osipova, in preparation).

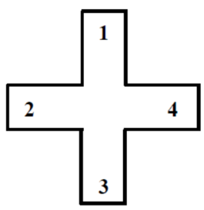

Fig. 1 The cross maze.

2.1. Behavior of zebrafish in the cross maze

The sizes of the maze were as follows. The length of the arm from the entrance into the arm to its end was 65 mm; the width of arm was 33 mm; the center square between the entrances to the arms was 33x33 mm. The height of the maze was 48 mm; the water level in the maze was 38 mm.

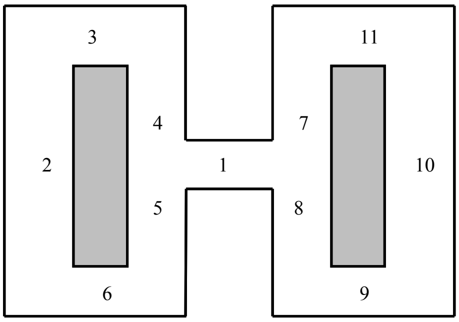

Fig. 2 The maze with 11 arms. Impenetrable barriers (gray) are shown within the west and east halves of the maze.

The main results of the experiment are as follows. 20 male zebrafish was observed (a typical length of fish was 25 mm). Each fish was placed individually into an arm and closed in this arm for 2 minutes by means of a removable hurdle. Then the hurdle was removed, and the fish behavior was observed for 15 minutes.

The basic types of fish movements are described below; the numbering of the arms is shown in Fig. 1. A significant portion of fish movements consisted of stereotyped repetitive sequences, which we will refer to as “motifs”. The most frequent motifs were:

1. Repeated shuttle transitions between adjacent corridors, for example, 12121212 or 141414... This motif was most noticeable.
2. Shuttle-like movements between opposing arms: 131313… and 242424... This motif was less frequent as compared with the first one.
3. Rarely, there was a steady transition from one arm to another adjacent one, clockwise (1234) or in the opposite direction (3214).
4. In addition to these motifs, random movements of the fish were observed rarely, for example, such as 231421.

A typical sequence of visits to arms is shown below:
<u>2121212121232323</u>42<u>23232323</u>1<u>3232323232</u>132312313<u>41234123412342323232323</u>1
32143141214<u>121212</u>34<u>121212121212</u>313<u>2121212341234</u>3<u>24242424131313124242</u>43
121324121412312.

Underlined chains of visits correspond to above mentioned motifs. Random movements are not underlined.

2.2. Behavior of fish in the maze with 11 arms

In the maze with 11 arms (Fig. 2), the length of arms 2 and 10 was 110 mm, arms 3, 6, 9 and 11 had the length 60 mm, arms 4, 5, 7 and 8 had the length 45 mm. The length of the arm 1 was 30 mm. The width of each arm was 20 mm. Twenty male zebrafish (of the same length as in the case of cross maze) were observed.

The following motifs were observed in this maze:

1. Shuttle-like movements in opposite directions along a one of long arms 2, 3, 6, 9, 10, or 11 (but not 1). Reaching the end of an arm, a fish turned back, moved in the opposite direction, reached the opposite end of the arm, again turned back, etc.
2. Shuttle-like movements between adjacent arms, e.g., 232323… or 262626...
3. Shuttle-like movement along arms 4 and 5 or 7 and 8. In these cases, the fish repeatedly moved, for example, from the arm 4 to the arm 5 and back: 454545... During such movements, the entrance into the arm 1 was ignored.
4. Movements 4171417... and 5181518... In this case, the fish entered into the arm 1, whenever it reached this short arm.
5. Fish could repeatedly move along the perimeter of any half of the maze, for example, 2345623456...

Thus, the behavior of fish in both mazes revealed certain regularities. We used these regularities as a starting point for developing computer models of fish behavior.

3. Models of fish movements, accumulation of knowledge, formation of predictions

Our models simulate movements of an agent (modeled zebrafish) in mazes. The time t is discrete, t = 0,1,2... At each step of time, the agent leaves an arm and enters into another one. The models were studied using computer simulation. The quantitative results were usually averaged over 10000 simulations with different starting random seeds.

This section describes two models that characterize initial stages of the fish exploratory behavior in mazes. The model 1 describes acquisition of knowledge about arms of the maze. The model 2 characterizes predictions of the next situation for the current situation and action. For the sake of simplicity, we consider in this section mainly the fish movements in the cross maze.

3.1. Model of knowledge acquisition

The model of knowledge acquisition (i.e. the model 1) describes the agent movements in the maze in accordance with motifs described above, and the increase of knowledge about frequently visited arms. The model assumes that the agent moves according to the two most frequent motifs (see Section 2.1), namely, motifs of the type 1 (the movement of an agent between adjacent arms) and the type 2 (the movement between the opposite arms). We also take into account the possibility of random movements.

In order to take into account transitions between different motifs, we introduce the following model parameters: transition probabilities between the two types of motifs P_{ij}, i,j = 1,2. For example, P_{12} is the probability of transition from the first motif to the second one. We also introduce the probabilities of transitions from a random movement (indexed “0”) to motifs, P_{00}, P_{01}, P_{02}, as well as the probability of transition from motifs to a random movement P_{10} and P_{20}. The following natural restrictions are imposed: $P_{00} + P_{01} + P_{02} = 1$, $P_{10} + P_{11} + P_{12} = 1$, $P_{20} + P_{21} + P_{22} = 1$. It is assumed that at the probabilities P_{11} and P_{22} the motif does not change. Transitions from the certain motif of the first type to another motif of the same type are neglected; we believe that such transition is carried out through a random selection of the movement type. If there are several possible arms for movement in accordance with the described method of motif changing, a next arm is selected at random.

Acquisition of the knowledge about the arms was modeled as follows. It was assumed that the agent has a certain knowledge K_i about each arm, $0 \leq K_i \leq 1$, i = 1,2,3,4. Initial values of knowledge K_i are 0. When the agent visits i-th arm, the value K_i becomes equal to 1. Additionally, all values K_i slightly decrease with time: at any time moment t, values K_i are multiplied by the factor d_K ($0 < d_K < 1$, $1–d_K << 1$).

A special tendency to enter into those arms, which the agent did not visited for a long time, was introduced as follows. The agent at the time step t considers knowledge K_i about all four arms, and at the next step t+1 the agent with a certain probability P_{choice} moves into the arm, which has a minimal value of K_i. Such movement into the arm with a minimal value of K_i occurs independently of the method of motif changing that was described above. The described tendency implies that a fish sometimes checks the arm, which it did not visit for a long time, because some good things (for example, food) could appear in such arm.

It should be noted that the movement into the arm with smallest value of K_i is rather simple in the case of the cross type maze: the agent knows values of K_i for all four arms, selects the arm with minimal K_i and moves into this arm. Such selection and movement into the arm with minimal K_i is essentially more complex for the maze with 11 arms; these processes are described in Section 4.

We analyzed the model by means of computer simulation. The parameters of simulation were as follows: the factor of knowledge decrement was d_K = 0.9, the transition probabilities P_{ij} were: P_{00} = 0.4, P_{01} = 0.4, P_{02} = 0.2; P_{10} = 0.1, P_{11} = 0.8, P_{12} = 0.1; P_{20} = 0.2, P_{21} = 0.6, P_{22} = 0.2, the probability of movement into the arm with minimal knowledge P_{choice} was variable.

The main results of computer simulation for this model are represented in Fig. 3, which shows the dynamics of the sum of knowledge K_{SUM} about all four arms for different values of probability P_{choice}. The maximal possible value of K_i for any arm is 1; therefore, the maximum of the sum K_{SUM} is 4. Owing to the decrease of values K_i (that is due to the factor d_K), the sum K_{SUM} does not reach this maximal possible value.

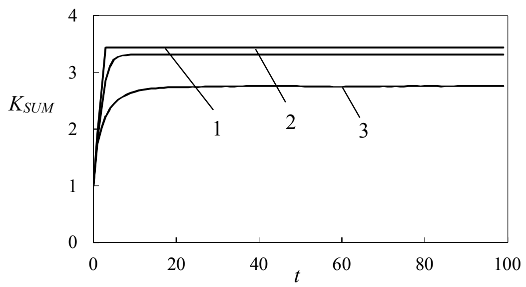

Fig. 3 The dependence of the sum K_{SUM} of knowledge for all arms in the cross maze on the time t for different values of P_{choice} (averaged over 10000 different starting random seeds). 1) P_{choice} = 1. 2) P_{choice} = 0.5. 3) P_{choice} = 0. The factor of decreasing knowledge is d_K = 0.9.

Fig. 3 shows that the summarized knowledge K_{SUM} about the maze grows faster and reaches larger values at large probabilities P_{choice}.

The computer simulation demonstrated that navigation of agents was qualitatively similar to the behavior of real fish. Repetitive shuttling movements between adjacent arms was observed, which corresponds to the motif of the first type. Also, repeated shuttle movements between opposite arms, which correspond to the motif of the second type, were sometimes observed.

A typical example of the arm sequence visited by the agent is as follows ($P_{choice} = 0$ for this example):
414343434343434343434342121131112434341212111114141414141311214342424234343434331313434
411133434343434313112121212121212123232323232343434322141414141423232314432113312121212
143432331414141412121212123332141414141.

As for real fish (see Section 2.1), underlined chains correspond to the motifs described above. Sequences of visited arms for the agent and for the real fish are similar to each other. Therefore, the agent movement is in qualitative agreement with the fish movement.

3.2. Model of predictions of future situations

The computer model that describes the formation of predictions of future situations by the agent has been developed and studied. This model 2 differs from the model 1 described in the Section 3.1: movements of the agent are not in accordance with motifs; these movements are mainly random. Predictions are characterized quantitatively by the values of assurance of these predictions A_S. Namely, for the given initial situation S_t and the action A_t, the assurance of the prediction of the next situation S_{t+1} is characterized by the value A_S ($0 \le A_S \le 1$). Each situation corresponds to a particular arm; the number of different situations equals to 4. When the agent leaves an arm, there are three possible actions A_t: 1) to turn into the right arm, 2) to go into the opposite arm, 3) to turn into the left arm. The agent predicts the next situation S_{t+1}. Thus, the assurances A_S characterize all possible chains $\{S_t, A_t\} \to S_{t+1}$.

To some extent, the considered predictions are similar to the formation of simple acceptor of results of action in the theory of functional systems by P.K. Anokhin (1974).

The behavior of the agent in the model 2 is as follows. Initially, at $t = 0$, all assurances A_S equal 0; then the agent makes some predictions, and the assurances A_S are changing. At any time moment t ($t > 0$), the agent predicts the next situation $S_{pr,t+1}$ for the time step t+1. If the assurance for this prediction is small: $A_S(S_t,A_t,S_{pr,t+1}) < T_A$ (T_A is a certain threshold value), then the agent at time moment t+1 returns to that arm, which it visited at time t-1 (the action A_{t+1} in this case is determined completely). If the assurance for this prediction is rather large ($A_S(S_t,A_t,S_{pr,t+1}) > T_A$), then the agent choose randomly one of the possible actions A_{t+1} and moves randomly into some new arm. Thus, the agent explores the maze.

The described mode of the agent's movement can be considered as an heuristics which corresponds to the opposite trends (mentioned in Introduction): 1) the desire to predict the results of the behavior reliably (in this case the low assurance of the correct prediction increases after repetitions of movements), and 2) the search for a new, unpredictable situation (this corresponds to performing random actions at the high current assurance).

The agent makes predictions in the following manner. For the given current situation S_t and action A_t, the agent checked the assurances $A_S(S_t,A_t,S_{pos,t+1})$ for predictions of all possible next situations $S_{pos,t+1}$. Then the agent determines the situation $S_{max,t+1}$ with the maximal value $A_S(S_t,A_t,S_{pos,t+1})$: $S_{pr,t+1} = \arg\max\{A_S(S_t,A_t,S_{pos,t+1})\}$, and predicts this situation $S_{pr,t+1}$. Only if all values $A_S(S_t,A_t,S_{pos,t+1})$ are too small, the agent predicts the next situation randomly.

The values of assurances A_S are adjusted as follows. At the time step t+1, the agent checks the prediction that it has made at the time step t. If the prediction of the next situation is correct (the predicted situation $S_{pr,t+1}$ coincides with the real situation S_{t+1}: $S_{pr,t+1} = S_{t+1}$), then the assurance of this prediction is increased according to the expression:

$$\Delta A_S(S_t,A_t,S_{pr,t+1}) = d_I[1 - A_S(S_t,A_t,S_{pr,t+1})], \tag{1}$$

where S_t and A_t are the situation and the action in the moment t, $S_{pr,t+1}$ is the predicted situation that the agent expects, d_I is the factor of increasing of assurances ($0 < d_I < 1$).

If the prediction is wrong ($S_{pr,t+1} \neq S_{t+1}$), then the assurance of this prediction decreases:

$$\Delta A_S(S_t,A_t,S_{pr,t+1}) = -\, d_D A_S(S_t,A_t,S_{pr,t+1})], \tag{2}$$

where d_D is the factor of decreasing of assurances ($0 < d_D < 1$). According to expressions (1), (2), the values of assurances are restricted: $0 \le A_S \le 1$.

In addition, it is assumed that all assurances values are multiplied by the factor d_A ($0 < d_A < 1$, $1 - d_A << 1$) at any time step; i.e. all assurances $A_S(S_t,A_t,S_{t+1})$ are slightly reduced.

Thus, the agent forms the assurance of the prediction of the final element of the chain $\{S_t,A_t\} \rightarrow S_{t+1}$. The assurances for all possible chains $\{S_t,A_t\} \rightarrow S_{t+1}$ are stored by the agent.

The model 2 was analyzed by means of the computer simulation. The parameters of simulation were as follows: the threshold for estimations of assurance values was $T_A = 0.9$ (the assurance A_S is large, if $A_S > T_A$), the factors of increasing and decreasing of assurances for adjustment of assurances were $d_I = d_D = 0.3$, the factor of slight reduction of all assurances was $d_A = 0.995$.

The main results of computer simulation for this model are represented in Fig. 4, which shows the dependence of the sum of all assurances for the whole cross maze $A_{S\text{-}SUM}$ on the time t. Initial assurances equal 0; then the summarized assurance $A_{S\text{-}SUM}$ for the maze grows.

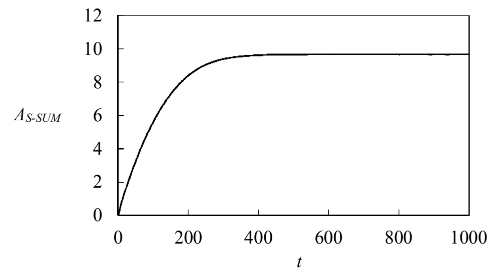

Fig. 4 The dependence of the summarized assurance $A_{S\text{-}SUM}$ on the time t for the whole cross maze (averaged over 10000 different starting random seeds).

In accordance with the number of all situations (that equals 4) and possible actions (that equals 3), there are 12 correct predictions. The maximal summarized assurance $A_{S\text{-}SUM}$ equals this number of correct predictions. This maximal possible value of the summarized assurance (12) is never achieved, because all assurances are slightly reduced at any time step.

Analogously to the cross maze, the computer models for fish movement, accumulation of knowledge, and formation of assurance of predictions for the maze with 11 arms were also developed and analyzed. These models and corresponding results of simulation were very similar to those obtained for the cross maze. However, the 11-arms maze is rather complex and it is difficult to model processes of the agent movement into a selected arm (in particular, into the arm with minimal value of knowledge K_i). Therefore, in order to represent such movements, we developed the new hypothetical model for the maze with 11 arms. This model 3 is described in the next section.

4. Hypothetical model of planning of movement in the maze with 11 arms

The hypothetical computer model assumes that after certain period of maze exploration, the fish is able to form some generalized notions that characterize the essential places (situations) in this maze (see Fig. 5). These 8 notions correspond to 8 situations S_t, in which the agent can be at the time moment t. These notions are: the most west/east arms in the maze (1 and 8 in Fig. 5), the north/south passes in the west/east halves of the maze (2, 3, 6, and 7), the places near the short arm in the west and east halves (4 and 5). The agent can execute the following four actions A_t: 1) to move to north, 2) to move to south, 3) to move to west, 4) to move to east.

It should be noted that such generalized representation of the maze is similar to usual human representation of a territory plan. After some experience, a person considers the territory near her/his home using some simple essential places: the street, the square, the bridge over the river, the shop, etc. Analogously, the agent uses the essential places in the maze.

The model 3 includes: 1) the knowledge acquisition by the agent about eight situations shown in Fig. 5, 2) the formation of assurances of predictions A_S for all possible chains $\{S_t,A_t\} \rightarrow S_{t+1}$, and 3) the planning of movement to that situation, which did not visited for a long time (such situation has minimal value of knowledge K_i).

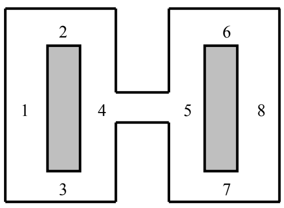

Fig. 5 The maze with 11 corridors. Situations 1-8 correspond to generalized notions.

The knowledge acquisition and the formation of the assurances of predictions A_S are almost the same as in the models described in the Section 3. The only difference is as follows. The model of the knowledge acquisition in the case of the maze with 11 arms does not use the motifs, and there is no returning to the previous arm in case of a small assurance of the current prediction; the exploratory movements in this rather complex maze is manly random.

We analyzed these models of the knowledge acquisition and the formation of the assurances of predictions by means of computer simulation. The parameters of simulation were the following: the factor of decreasing knowledge K_i was d_K = 0.99, the factors of increasing and decreasing of assurances for assurances adjustment according to expressions (1) and (2) were $d_I = d_D = 0.3$, the factor of slight reduction of all assurances was d_A = 0.999.

The results of computer simulation of processes of the knowledge acquisition and the formation of the assurances of predictions A_S are represented in Figs. 6 and 7. Fig. 6 shows the dynamics of the sum of knowledge K_{SUM} about all 8 places of the maze (see Fig. 5). Each place is characterized by its knowledge K_i (i = 1,…,8). Owing to the decrease of values K_i (that is due to the factor d_K), the sum K_{SUM} does not reach the maximum possible value 8.

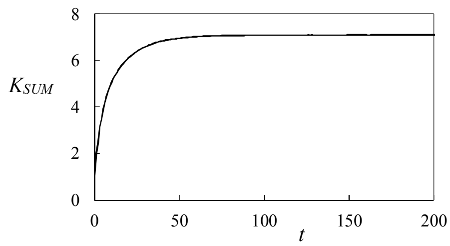

Fig. 6 The dependence of the sum K_{SUM} of knowledge for all places in the maze with 11 arms on the time t (averaged over 10000 different starting random seeds).

Fig. 7 shows the time dependence of the sum $A_{S\text{-}SUM}$ of assurances of predictions A_S for all possible chains $\{S_t,A_t\} \rightarrow S_{t+1}$ for the whole maze with 11 arms. There are 3 possible actions in places 4 and 5, while in other places there are 2 possible actions (Fig. 5). In accordance with the number of all places (that is equal to 8) and numbers of all possible actions, there are 18 correct predictions. The maximal value $A_{S\text{-}SUM}$ is equal to this number of correct predictions. This maximal possible value of $A_{S\text{-}SUM}$ is not achieved, because all assurances are slightly reduced at any time step.

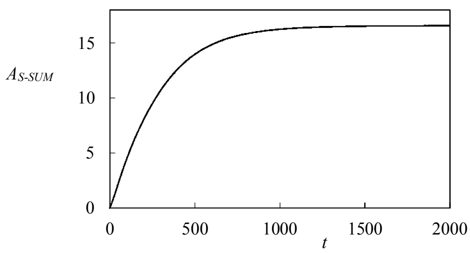

Fig. 7 The dependence of the summarized assurance $A_{S\text{-}SUM}$ on the time t for the whole maze with 11 arms (averaged over 10000 different starting random seeds).

The most non-trivial part of the current model is the scheme of the planning of movement into a selected place; this scheme is described below.

The model assumes that the agent already has formed the knowledge K_i for all situations S_i (i = 1,…,8) and the assurances A_S for all possible chains $\{S_t, A_t\} \rightarrow S_{t+1}$. This corresponds to time moments after t = 2000 in Figs 6, 7. Therefore, the agent is already experienced. Using values of knowledge K_i, the agent intends to move to the situation with the smallest knowledge. The agent also knows already the chains $\{S_t, A_t\} \rightarrow S_{t+1}$, which correspond to large values of assurance. In the described example of the simulation, the assurance values A_S for correct predictions at t =2000 were 0.88-0.98, while the assurances A_S for incorrect predictions were negligible small. Therefore, the agent knows already the correct predictions. For example, the agent knows that performing the action 4 (moving to east) in the situation 2, it will go to the situation 4 (the place near the short arm in the west half of the maze). Using these correct predictions, the agent selects such sequence of situations and actions that result in the movement towards the intended situation.

The following example shows the process of such selection of situations and actions and planning of the movement. We suppose that the starting position of the agent is the most west arm (the situation 1); the goal situation (for which the knowledge K_i is smallest) is the situation 8. Therefore, the agent should create a plan of the movement from the starting situation 1 to the goal situation 8 by itself (using the set of assurances for all possible situations and actions). The agent creates this plan as follows. The agent begins to analyze such situations (6 and 7) and actions that result in the goal situation 8. Then the agent analyzes situations and actions that result in the pre-goal situations 6 and 7, and so on. Thus, the agent begins from the goal situation 8 and analyzes consecutively possible ways to reach this situation. The agent also takes into account the distance from the considered situation S_t to the goal situation; this distance is the number of actions needed to reach the situation 8 from the situation S_t.

The result of agent's analysis is shown in Table 1.

Table 1 Scheme of analysis of movement towards the goal situation.

Step	$S_{previous}$	$A_{previous}$	$S_{current}$	Distance
1	6	4	8	1
2	7	4	8	1
3	5	1	6	2
4	5	2	7	2
5	4	4	5	3
6	2	4	4	4
7	3	4	4	4
8	1	1	2	5
9	1	2	3	5

In this table, $S_{previous}$, $A_{previous}$, and $S_{current}$ are the previous situation, the action in the previous situation, and the final situation respectively in the chain $\{S_{previous}, A_{previous}\} \rightarrow S_{current}$, "Distance" is the distance between the situation $S_{previous}$ and the goal situation 8.

This table demonstrates that initially the agent considers situations from which it can reach the goal situation 8; these are situations 6 and 7; performing the action 4 (moving to east) in any of these situations, the agent goes to the goal situation 8. Then the agent considers how to reach the situations 6 or 7. The distance from situations 6 and 7 (the number of needed actions) to the situation 8 is equal to 1. Then the agent considers possible ways to reach the situation 6 and the situation 7. Both situations are reachable from the situation 5 by means of the actions 1 or 2 (by movements to north or to south); this is demonstrated by steps 3 and 4 in Table 1. The agent continues this analysis of all possible ways from the starting situation 1 towards the goal situation 8. Reaching the starting situation, the agent stops the analysis.

Then the agent creates a simple knowledge database that characterizes possible ways to reach the goal situation 8. This knowledge database is presented in Table 2.

Table 2 Knowledge database.

Initial situation	Number of useful actions	Action	Next situation	Distance for the initial situation	Distance for the next situation
1	2	1	2	5	4
1	2	2	3	5	4
2	1	4	4	4	3
3	1	4	4	4	3
4	1	4	5	3	2
5	2	1	6	2	1
5	2	2	7	2	1
6	1	4	8	1	0
7	1	4	8	1	0

The knowledge database is simple reconstruction of results of analysis presented in Table 1. The second column shows the number of useful actions, which result in a decrease of the distance between the considered situation and the goal situation.

Finally, using this knowledge database, the agent forms a plan of movement. In some situations, there are several possible useful actions; according to Table 2, two possible useful actions exist in situations 1 and 5. The agent randomly chooses one of possible actions in such situations.

The example of the plan of movement towards the goal situation is shown in Table 3. According to this plan, when the agent is in the situation 1, it moves to north and reaches the situation 2, then it moves to east and reaches the situation 4; next, it moves to east to the situation 5, then it moves to north to the situation 6; finally, the agent moves to east and reaches the goal situation 8. In the computer simulation, we observed this plan and other possible 3 plans.

Table 3 Example of the plan of movement towards the goal situation.

Initial situation	Action	Next situation
1	1	2
2	4	4
4	4	5
5	1	6
6	4	8

5. Discussion and conclusion

Thus, the models of fish movement, accumulation of knowledge, formation and use of predictions have been developed and investigated. The most interesting is the hypothetical model of forming the simple database and planning of movement towards the goal situation.

Probably, the described process of forming the plan of movement from the most west place in the maze to the most east place is too complex for fish. Nevertheless, a fish could form more simple analogous plans. For example, the fish could create a plan of movement from the west half of the maze to the east half. Such plan could include only two steps: 1) a movement to the place 4 (the place near the short arm between halves of the maze, see Fig. 5), 2) the movement to east (into the east half of the maze). Additionally, after exploration of a maze and obtaining some experience, a fish could form and use some stereotypes of behavior and move without detailed analysis of possible ways, without a creation of a detailed plan.

We can consider a similar planning of movement by humans. For example, a human could create such plan of her/his movement from a home to a shop: 1) to go from the home to east to the street 1, 2) to go along the street 1 to south to the bridge over the river, 3) to cross the bridge and go to south to the street 2, 4) to go along the street 2 to south to the shop. Of course, initially such analysis with a map could be useful, but after some experience, a stereotype of this movement will be created in the person's mind. E.g., this could be a simple stereotype "the movement to the shop" without details of the movement.

We can consider similar planning for other animals. For example, we can consider a similar model of behavior planning for New Caledonian crows, which really form a mental plan of consecutive actions of the complex behavioral chain, basing on previously obtained knowledge about parts of this chain (Taylor et al., 2010). Actually, the crows predict results of their particular actions in certain situations and use these predictions. Therefore, we can analyze processes of formation of plans of behavior at different evolutionary levels. It is essential that the considered planning is based on predictions of the results of elementary actions.

In this connection, it should be noted that there was the very interesting proposal for using a set of logical predictive rules in close relation with the foundation of the mathematics (Turchin, 1987). Therefore, we can consider the cognitive processes of creation and use of predictions at different evolutionary levels: from fish to humans.

Acknowledgments

This work was partially supported by the Russian Foundation for Basic Research, Grants No 13-01-00399 and No 15-04-06379. The authors thank Tamara I. Sharipova for her help at initial stages of the described research. We are very grateful for the valuable remarks and suggestions of the reviewers, who helped us to reevaluate our work and to correct the paper.

References

Anokhin, P.K. (1974). *Biology and Neurophysiology of the Conditioned Reflex and its Role in Adaptive Behavior*. Oxford etc.: Pergamon Press.

Inglis, I.R., Forkman, B., & Lazarus, J. (1997). Free food or earned food: a review and fuzzy model of contrafreeloading. *Animal Behaviour*, *53*(6), 1171–1191.

Inglis, I.R, Langton, S., Forkman, B., & Lazarus, J. (2001). An information primacy model of exploratory and foraging behaviour. *Animal Behaviour*, *62*(3), 543–557.

Oudeyer, P.-Y., & Kaplan, F. (2004). Intelligent adaptive curiosity: a source of self-development. In L. Berthouze et al. (Eds.). *Proceedings of the 4th International Workshop on Epigenetic Robotics* (Vol. 117, pp. 127–130). Lund University Cognitive Studies.

Oudeyer, P.-Y., & Kaplan, F. (2009). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*. Vol. 1. Article 6. URL: http://journal.frontiersin.org/Journal/10.3389/neuro.12.006.2007/full

Taylor, A.H., Elliffe, D., Hunt, G.R., & Gray, R.D. (2010). Complex cognition and behavioural innovation in New Caledonian crows. *Proceedings of the Royal Society B*, *277* (1694), 2637–2643.

Turchin, V.F. (1987). A constructive interpretation of the full set theory. *Journal of Symbolic Logic*, *52*(1), 172 –201.