

УДК 32.81

## БИОЛОГИЧЕСКИ ИНСПИРИРОВАННЫЕ МОДЕЛИ КОГНИТИВНЫХ АВТОНОМНЫХ АГЕНТОВ<sup>1</sup>

В.Г. Редько (*vgredko@gmail.com*)

Научно-исследовательский институт системных  
исследований РАН, Москва

В.А. Непомнящих (*nepom@ibiw.yaroslavl.ru*)

Институт биологии внутренних вод им. И.Д. Папанина  
РАН, Борок, Ярославская область

М.С. Бурцев (*burtsev.m@gmail.com*)

НИЦ «Курчатовский институт», Москва

Обобщены и развиты биологически инспирированные модели познавательных способностей автономных агентов. Модели характеризуют процессы 1) накопления знаний, 2) формирования предсказаний, 3) планирования цепочек целенаправленных действий на основе предсказаний. Модели основаны на биологических экспериментах с рыбами, которые познают лабиринты, и с новокаледонскими воронами, формирующими планы достаточно сложного поведения на основе ранее освоенных элементов этого поведения.

**Ключевые слова:** биологически инспирированные модели, познавательные способности, автономные агенты

### Введение

Одно из перспективных направлений когнитивной науки – моделирование когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, той эволюции, в результате которой произошло мышление человека [Редько, 2015; Редько, 2016]. Основной метод исследования когнитивной эволюции – построение и изучение моделей познавательных способностей животных разного эволюционного уровня. Причем модели целесообразно строить,

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проекты №№ 15-01-00223, 15-04-06379).

отталкиваясь от конкретных биологических экспериментов. Сочетание биологических экспериментов и теоретических моделей может весьма эффективным, аналогично тому, как эффективно сочетание эксперимента и теории в физике.

В настоящей работе исследованы непосредственно связанные с биологическими экспериментами модели когнитивных способностей достаточно простых животных. Обобщены и развиты модели познавательных способностей 1) рыб, изучающих лабиринты, 2) новокаледонских ворон, формирующих планы довольно сложного поведения. Модели развивают наши предыдущие работы [Непомнящих и др., 2015; Red'ko et al., 2015; Red'ko & Nepomnyashchikh, 2015].

### 1. Модели познавательных способностей рыб, изучающих лабиринты

Модели основаны на биологическом эксперименте [Непомнящих и др., 2015], в котором изучалось поведение рыб в незнакомой им среде: в простом крестообразном лабиринте с 4-мя коридорами (рис. 1) и в более сложном лабиринте с 11 коридорами (рис. 2).

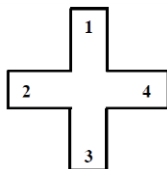


Рис. 1. Крестообразный лабиринт

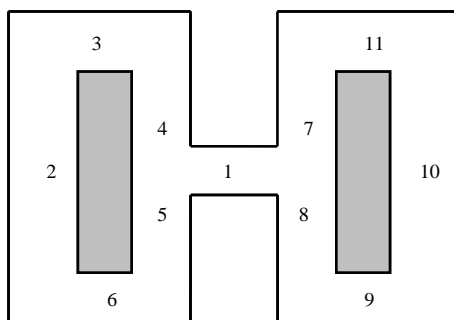


Рис. 2. Лабиринт с 11 коридорами. Серым показаны непрозрачные барьеры внутри левого и правого отсеков

**Модель накопления знаний.** Для крестообразного лабиринта была построена модель накопления рыбами знаний о коридорах. Считалось, что агент (модельная рыба) имеет определенное знание  $K_i$  о каждом из коридоров,  $0 \leq K_i \leq 1$ ,  $i = 1, 2, 3, 4$ . Предполагалось, что в момент посещения агентом  $i$ -го коридора величина  $K_i$  становится равной 1, а для остальных коридоров величины  $K_i$  постепенно уменьшаются.

**Модель формирования уверенностей в предсказаниях.** Для крестообразного лабиринта была также построена компьютерная модель, в которой рассчитывалась уверенность агентов в предсказании будущих ситуаций. Уверенность характеризовалась величиной  $A_s$ , меняющейся от 0 до 1. Модель предполагала, что имеются исходные ситуации  $S_t$  (каждая ситуация соответствует коридору, в котором агент находится);  $t$  ( $t = 0, 1, 2, \dots$ ) – время. В каждой ситуации агент может выполнять три действия  $A_t$ : при выходе из коридора, агент может 1) повернуть в правый коридор, 2) пройти в противоположный коридор, 3) повернуть в левый коридор. При выполнении действия агент предсказывает, в какой ситуации  $S_{t+1}$  он дальше окажется, если выполнит действие  $A_t$ .

Далее агент определяет, сбылось ли его предсказание. Если предсказание сбылось, то уверенность в данном предсказании увеличивается, если предсказание не сбылось, то такая уверенность уменьшается. Таким образом, формируются уверенности в предсказании конечного элемента цепочки  $\{S_t, A_t\} \rightarrow S_{t+1}$ .

Множество уверенностей для всех возможных цепочек  $\{S_t, A_t\} \rightarrow S_{t+1}$  запоминалось агентом. Кроме того, считалось, что уверенности в предсказании ожидаемого результата для всех цепочек  $\{S_t, A_t\} \rightarrow S_{t+1}$  все время немного уменьшаются.

Компьютерное моделирование показало, что как знания о коридорах, так и уверенности в предсказаниях результатов действий возрастают.

**Гипотетическая модель планирования в лабиринте с 11 коридорами.** Для лабиринта с 11 коридорами была построена гипотетическая модель, предполагавшая, что после определенного освоения лабиринта рыба способна сформировать некоторые обобщенные понятия, характеризующие места в этом лабиринте, такие как самый западный или самый восточный коридор лабиринта, северный или южный проход в западном/восточном секторе лабиринта (см. рис. 2). Эти понятия соответствуют ситуациям  $S_i$ , в которых может находиться агент. Агент может выполнять четыре действия  $A_i$ : двигаться на север/юг/запад/восток.

В этой модели предполагается, что модельная рыба уже хорошо освоила лабиринт, т.е. агент уже имеет сформированные знания  $K_i$  для всех обобщенных ситуаций  $S_i$  и уверенности в предсказаниях для всех возможных цепочек  $\{S_i, A_i\} \rightarrow S_{i+1}$ . Используя эти знания  $K_i$  и

уверенности, агент может создать план движения к месту, для которого величина  $K_i$  минимальна, т.е. к тому месту, которое он давно не посетил.

Агент формирует план, сначала мысленно анализируя пути движения от целевой ситуации к исходной. В результате у агента создается простая *база знаний*, которая характеризует ситуации, действия, прогнозы результатов действий, оценку расстояний между текущими ситуациями и целевой ситуацией (в наших моделях это расстояние равно числу действий, необходимых для перехода из текущей ситуации в целевую). Используя эту базу знаний, агент формирует план движения от исходной ситуации к целевой, например, план движения из самого западного коридора в самый восточный коридор лабиринта.

## **2. Модель планирования новокаледонскими воронами**

Модель основана на биологическом эксперименте с новокаледонскими воронами [Taylor et al., 2010]. В этом эксперименте вороны предварительно обучались выполнению отдельных элементов довольно сложного поведения. После обучения воронам предлагалось решить следующую трехзвенную проблему: (1) подтянуть привязанную на шнуре короткую палочку и освободить ее от шнура, (2) с помощью короткой палочки достать длинную палочку из зарешеченного контейнера, (3) длинной палочкой извлечь пищу из второго глубокого прозрачного контейнера. При этом невозможно было извлечь пищу из глубокого контейнера с помощью короткой палочки или клюва и извлечь длинную палочку из зарешеченного контейнера клювом. Поэтому, чтобы добыть пищу, вороне надо было выполнить четко определенную цепочку указанных последовательных действий:  $1 \rightarrow 2 \rightarrow 3$ . Были более полно обученные вороны, которые обучались оперированию со всеми тремя инструментами (шнур, короткая палочка и длинная палочка) по отдельности, и менее обученные вороны, которые обучались оперированию только с двумя из этих инструментов (шнур и длинная палочка). В результате более обученные вороны решали трехзвенную проблему с первой попытки; менее обученные вороны тоже решали эту проблему, но медленнее и не всегда с первой попытки.

Итак, вороны обучались отдельным элементам поведения, направленного на решение трехзвенной проблемы. А затем самостоятельно мысленно планировали полностью все это поведение, мысленно связывая отдельные элементы в единую цепочку.

Построим модель мысленного формирования плана поведения, нацеленного на решение трехзвенной проблемы. Рассматриваем модельных ворон (автономных агентов). Считаем, что имеются следующие существенные ситуации  $S_i$  и действия  $A_i$ :

$S_1$ : короткая палочка висит привязанная на шнуре; длинная палочка лежит в зарешеченном контейнере; пища находится во втором глубоком прозрачном контейнере

$S_2$ : короткая палочка свободна; длинная палочка лежит в зарешеченном контейнере; пища находится во втором глубоком прозрачном контейнере

$S_3$ : длинная палочка свободна; пища находится во втором глубоком прозрачном контейнере

$S_4$ : пища свободна

$A_1$ : подтянуть привязанную на шнуре короткую палочку и освободить ее от шнура

$A_2$ : с помощью короткой палочки достать длинную палочку из зарешеченного контейнера

$A_3$ : длинной палочкой извлечь пищу из второго глубокого прозрачного контейнера.

При этом  $S_1$  – исходная ситуация,  $S_4$  – целевая ситуация.

В результате предварительной тренировки все вороны обучались делать предсказание  $\{S_3, A_3\} \rightarrow S_4$  (длинной палочкой можно достать пищу из глубокого прозрачного контейнера). Более обученные вороны также могли делать предсказание  $\{S_2, A_2\} \rightarrow S_3$  (короткой палочкой можно достать длинную). Все вороны не могли делать четкого нужного предсказания  $\{S_1, A_1\} \rightarrow S_2$  (хотя при предварительной тренировке можно было подтягивать шнур с привязанным на нем куском пищи [Taylor et al., 2010], т.е. до этого предсказания можно было догадаться).

В данной работе полагаем, что механизм формирования плана решения трезвенной проблемы включает в себя следующие этапы (рис. 3):

1. прямое мысленное рассмотрение от  $S_1$  к  $S_4$  (до тех пока не будет найдена правильная цепочка действий, ведущих к цели)
2. проверка найденного решения путем тестирующего обратного и прямого мысленного рассмотрения
3. генерация стереотипа целенаправленного поведения после тестирования.

При прямом мысленном рассмотрении модельные вороны анализируют возможности достижения цели  $S_4$ , исходя из ситуации  $S_1$ . При этом вороны могут тестировать мысленно или реально пробовать отдельные действия и догадываться с определенными вероятностями до нужных предсказаний  $\{S_1, A_1\} \rightarrow S_2$  и  $\{S_2, A_2\} \rightarrow S_3$ . Считаем, что более обученные вороны угадывают предсказание  $\{S_1, A_1\} \rightarrow S_2$  с вероятностью  $P_1$ . Менее обученные вороны догадываются о предсказаниях  $\{S_1, A_1\} \rightarrow S_2$  и  $\{S_2, A_2\} \rightarrow S_3$  с вероятностями  $P_2$  и  $P_3$ , соответственно. При компьютерных расчетах полагалось  $P_1 = 0.9$ ,  $P_2 = 0.7$ ,  $P_3 = 0.5$ . Также (как и в модели для рыб) считаем, что формируются уверенности в предсказаниях, которые увеличиваются/уменьшаются при

правильном/неправильном предсказании. Если сразу решение не найдено, то происходит повтор прямого мысленного рассмотрения.

После первого нахождения результата происходит его проверка, при этом при тестирующем обратном мысленном рассмотрении (рис. 3) оцениваются расстояния между текущими ситуациями и целевой и при успешной проверке формируется база знаний (табл. 1), характеризующая ситуации, действия, прогнозы результатов действий, оценки расстояний между текущими ситуациями и целевой ситуацией.

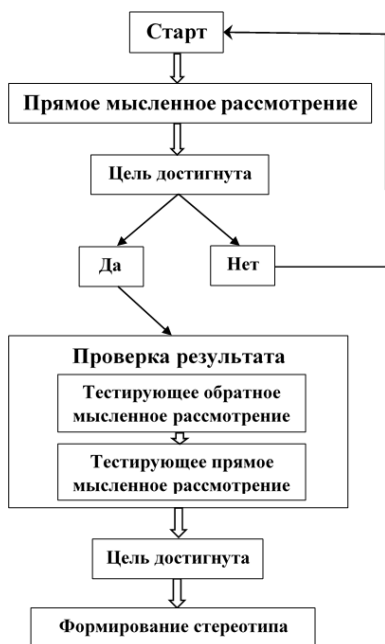


Рис. 3. Схема формирования плана решения трехзвенной проблемы. При прямом мысленном рассмотрении происходит анализ от исходной ситуации  $S_1$  к целевой  $S_4$ ; при обратном рассмотрении – от  $S_4$  к  $S_1$ . При тестирующем обратном рассмотрении оцениваются расстояния между текущими ситуациями и целевой и формируется база знаний. При тестирующем прямом рассмотрении на основе базы знаний выбираются действия, приближающие к цели. После успешной проверки формируется план решения трехзвенной проблемы и стереотип целенаправленного поведения

Табл. 1. База знаний модельных ворон

Текущая ситуация, $S_{current}$	Текущее действие, $A_{current}$	Следующая ситуация, $S_{next}$	$\rho(S_{current}, S_4)$	$\rho(S_{next}, S_4)$
$S_1$	$A_1$	$S_2$	3	2
$S_2$	$A_2$	$S_3$	2	1
$S_3$	$A_3$	$S_4$	1	0

Расстояния  $\rho(S_{current}, S_4)$  и  $\rho(S_{next}, S_4)$  – это число действий, необходимое для достижения цели  $S_4$  из ситуаций  $S_{current}$  и  $S_{next}$ , соответственно.

Для надежности проводится и тестирующее прямое мысленное рассмотрение (рис. 3), которое уже использует базу знаний. На основе базы знаний строится план решения трехзвенной проблемы.

Отметим, что в модели была введена проверка полученного результата путем обратного и прямого тестирующего мысленного рассмотрения. Это тестирование подобно быстрому обратному (а, возможно, и прямому) возбуждению клеток места в гиппокампе животных после достижения результата путем перемещения в пространстве [Foster et.al., 2006; Diba et.al., 2007]. Возможно, что такая проверка играет важную роль в надежном запоминании полученного решения.

После проверки формируется стереотип целенаправленного поведения. При выполнении действий в соответствии со стереотипом уже не нужны мысленные усилия для формирования плана поведения, следовательно, стереотипное поведение должно происходить быстро.

Было проведено компьютерное моделирование по приведенной схеме формирования планов воронами. На рис. 4 представлена зависимость средней частоты успешного составления планов в зависимости от времени. Один такт времени соответствовал одной попытке нахождения решения трехзвенной проблемы,  $\langle F_{success} \rangle = 1$  соответствует достижению надежного решения трехзвенной проблемы.

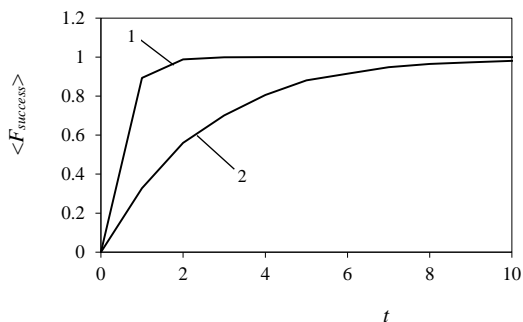


Рис. 4. Зависимость средней частоты успешного планирования  $\langle F_{success} \rangle$  от времени  $t$  для более (1) и менее (2) обученных модельных воронов.  $P_1 = 0.9$ ,  $P_2 = 0.7$ ,  $P_3 = 0.5$ . Усреднено по 1000 различным расчетам

Согласно рис. 4 более обученные вороны находят успешное решение трехзвенной проблемы практически с первой попытки, менее обученным воронам для решения проблемы достаточно нескольких первых попыток. Это качественно согласуется с биологическим экспериментом [Taylor et al., 2010].

## Заключение

Таким образом, построены и исследованы биологически инспирированные модели когнитивных автономных агентов, познающих закономерности взаимодействия с внешним миром. Обобщая свойства модельных рыб и модельных ворон, можно выделить следующие характерные черты рассмотренных когнитивных агентов:

- 1) Оценка *знаний* о каждой из возможных ситуаций.
- 2) Движение к ситуации, знания о которой минимальны. Тем самым *агент может самостоятельно формировать новые текущие цели своего поведения*. Например, рыба в лабиринте может решить двигаться в тот коридор, для которого знания минимальны, т.е. в тот коридор, который она давно не посещала.
- 3) Формирование *предсказаний* о результатах действий. Оценка *уверенности в предсказаниях*, стремление к надежным предсказаниям, уверенность в которых высока.
- 4) Формирование простой *базы знаний*, характеризующей ситуации, действия, прогнозы результатов действий, оценки расстояний между текущими ситуациями и целевой ситуацией.
- 5) Использование базы знаний при планировании целенаправленного поведения.
- 6) *Прямые и обратные мыслительные процессы* (от исходной ситуации к целевой и обратно).
- 7) Формирование *стереотипа* целенаправленного поведения.

**Благодарности.** Авторы благодарны анонимным рецензентам за замечания, способствовавшие улучшению статьи.

## Список литературы

- [Непомнящих и др., 2015] Непомнящих В.А., Редько В.Г., Бесхлебнова Г.А., Осипова Е.А., Шарипова Т.И. Моделирование процессов накопления знаний и формирования предсказаний рыбами, осваивающими лабиринты // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сб. науч. трудов VIII-й международной научно-технической конференции. В 2-х томах. Т.2. – М.: Физматлит, 2015.
- [Редько, 2015] Редько В.Г. Моделирование когнитивной эволюции: На пути к теории эволюционного происхождения мышления. – М: ЛЕНАНД/URSS, 2015.
- [Редько, 2016] Моделирование когнитивной эволюции – формирующаяся научная дисциплина // Седьмая международная конференция по когнитивной науке. Тезисы докладов. – М: Изд-во «Институт психологии РАН», 2016.
- [Diba et.al., 2007] Diba K., Buzsaki G. Forward and reverse hippocampal place-cell sequences during ripples. Nature Neuroscience. 2007. V. 10. No. 10.
- [Foster et.al., 2006] Foster D.J., Matthew A. Wilson M.A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature. 2006. V. 440. No. 7084.



[Red'ko & Nepomnyashchikh, 2015] Red'ko, V.G., Nepomnyashchikh, V.A. Model of plan formation by New Caledonian crows // Procedia Computer Science. 2015. V. 71.

See also: <http://www.sciencedirect.com/science/article/pii/S1877050915036820>

[Red'ko et al., 2015] Red'ko V.G., Nepomnyashchikh V.A., Osipova E.A. Models of fish exploratory behavior in mazes // Biologically Inspired Cognitive Architectures. 2015. V. 13. See also: <https://www.niisi.ru/iont/ni/rygpubl/Redkoetal2015.pdf>

[Taylor et al., 2010] Taylor A.H., Elliffe D., Hunt G.R., Gray R.D. Complex cognition and behavioural innovation in New Caledonian crows // Proceedings of the Royal Society B. 2010. V. 277. No. 1694.

## **BIOLOGICALLY INSPIRED MODELS OF COGNITIVE AUTONOMOUS AGENTS**

Vladimir G. Red'ko (vgredko@gmail.com)  
Scientific Research Institute for System Analysis, Russian  
Academy of Sciences, Moscow

Valentin A. Nepomnyashchikh (nepom@ibiw.yaroslavl.ru)  
I.D. Papanin Institute for Biology of Inland Waters, Russian  
Academy of Sciences, Borok, Yaroslavl region

Mikhail S. Burtsev (burtsev.m@gmail.com)  
National Research Centre "Kurchatov Institute", Moscow

Biologically inspired models of the cognitive abilities of autonomous agents have been developed. These models characterize the following processes 1) knowledge acquisition, 2) generation of predictions of results of actions, 3) planning of chains of actions. Planning uses the predictions of results of actions. Models are based on biological experiments on fish that explore mazes and on New Caledonian crows that are planning the rather complex behavior.

**Keywords:** biologically inspired models, cognitive abilities, autonomous agents