

УДК 32.81

## О МОДЕЛИРОВАНИИ КОГНИТИВНОЙ ЭВОЛЮЦИИ

В.Г. Редько ([vgredko@gmail.com](mailto:vgredko@gmail.com))  
НИИ системных исследований РАН, Москва

Анализируются подходы к моделированию когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, в результате которой произошли логика, мышление, интеллект человека. Очерчены заделы обсуждаемых исследований и предлагаются контуры программы будущего моделирования когнитивной эволюции. Характеризуются начальные шаги намеченного моделирования.

### 1. Задача моделирования когнитивной эволюции

В настоящей работе анализируются подходы к моделированию когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, в результате которой произошло мышление человека, обеспечивающее научное познание природы. Развивается подход, предложенный ранее [Редько, 2005, 2010].

Кратко охарактеризуем постановку исследований когнитивной эволюции. Существует глубокая гносеологическая проблема: почему формальное логическое человеческое мышление, казалось бы, совсем не связанное с реальным физическим миром, применимо к познанию природы? Поясним проблему. Рассмотрим, например, физику, одну из фундаментальных естественнонаучных дисциплин. Мощь физики связана с эффективным применением математики. Но математик делает логические выводы, доказывает теоремы независимо от внешнего мира, используя свое мышление. Почему же эти выводы применимы к реальной природе? В общей формулировке проблема может быть сформулирована так: почему логическое мышление *человека* применимо к познанию *природы*? Для исследования проблемы целесообразно исследовать эволюцию познавательных свойств биологических организмов и с помощью моделирования проанализировать, как в процессе эволюции возникали способности, обеспечивающие познание закономерностей внешнего мира, и как развитие этих способностей привело к возможности научного познания природы.

## **2. Заделы исследований когнитивной эволюции**

В настоящее время в нескольких научных направлениях, таких как «Адаптивное поведение», «Искусственная жизнь», «Когнитивные архитектуры», научные основы искусственного интеллекта, ведутся исследования, близкие к моделированию когнитивной эволюции. В основном это работы по основанным на биологических аналогиях моделям автономных агентов с когнитивными свойствами. Кратко охарактеризуем эти направления.

Направление исследований «Адаптивное поведение» сформировалось в начале 1990-х годов [Meyer et al, 1991]. Основной подход этих исследований – конструирование и изучение искусственных «организмов» (в виде компьютерной программы или робота), способных приспосабливаться к переменной внешней среде. Исследователи адаптивного поведения разрабатывают такие модели, которые применимы к описанию поведения как реального животного, так и искусственного модельного организма. Дальняя цель этих работ – анализ эволюции когнитивных способностей животных в контексте происхождения интеллекта человека – близка к задаче моделирования когнитивной эволюции. Работы отечественных исследователей адаптивного поведения представлены в сборнике [Редько, 2006].

Под когнитивными архитектурами понимаются структуры и принципы функционирования познающих систем, которые можно использовать в искусственном интеллекте. Пример когнитивной архитектуры – система Soar (от англ. State, Operator And Result). Основная цель работ по Soar – создание системы функционирования интеллектуальных агентов, работающих в широкой области: от простейших форм до оперирования в сложных, заранее не предсказуемых условиях. Обзор исследований по когнитивным архитектурам и Soar содержится в работах [Langley et al, 2009], [Lehman et al, 2006]. В последние годы большое внимание уделяется биологически инспирированным когнитивным архитектурам [Samsonovich et al, 2011].

В исследованиях автономных интеллектуальных или когнитивных агентов часто изучаются компьютерные модели основанных на биологических аналогиях агентов. Обзор исследований в этой области содержится в работе [Vernon et al, 2007].

Необходимо подчеркнуть, что автономные агенты вполне могут рассматриваться как объединяющее понятие для указанных направлений: агенты могут моделировать биологические организмы и искусственные аналоги организмов. Вполне естественно изучать познавательные свойства и знания автономных агентов. Также разумно исследовать эволюцию популяций агентов и коммуникации и обмен информацией между агентами.

Подчеркнем также связь исследований когнитивной эволюции с основаниями математики. Выше заострялся вопрос о причинах применения математических доказательств к познанию реальных объектов в природе. Данный вопрос связан с обоснованием методов математического вывода и с возможностью пересмотра оснований математики. Именно в этом контексте в работе [Turchin, 1987] был предложен подход к введению предиктивных логических правил, позволяющих предсказывать будущие ситуации.

### **3. Примеры моделей автономных адаптивных агентов**

Приведем два примера моделей автономных агентов, обладающих когнитивными свойствами.

#### **3.1. Нейробиологическая модель адаптивного поведения, разработанная в Институте нейронаук, руководимом Дж. Эдельманом**

В цикле исследований данного института анализировалось адаптивное поведение искусственного организма – устройства NOMAD (Neurally Organized Mobile Adaptive Device). В работе [Krichmar et al, 2005] поведение NOMAD'a моделировало поведение мыши в лабиринте Морриса, которая в бассейне с непрозрачной жидкостью ищет скрытую от зрения платформу, ориентируясь по рисункам на бортах бассейна. NOMAD представлял собой подвижное роботоподобное устройство на колесах, система управления которого представляла собой большую структурированную нейронную сеть (реализованную в компьютерной программе), построенную по аналогии с реальной нервной системой. NOMAD самостоятельно обучался находить скрытую от зрения платформу, ориентируясь по разноцветным полосам на стенах комнаты. Было продемонстрировано, что NOMAD адекватно имитирует нетривиальное адаптивное поведение мыши.

#### **3.2. Бионическая модель поискового адаптивного поведения**

В работе [Непомнящих и др., 2008] промоделировано поисковое поведение личинок ручейников, обитающих на дне водоемов. Личинки носят на себе защитный чехол-домик – трубку из песчинок и других частиц, собираемых со дна водоемов. Компьютерная модель поискового поведения личинок ручейников, строящих чехол-домик из частиц разного размера и ведущих поиск скоплений подходящих частиц, использовала понятие мотивации  $M$  к прикреплению частиц к домику. Динамика регулирующей мотивации  $M$  учитывала инерцию переключения между тактиками поведения, случайные вариации и направленное изменение мотивации. Если мотивация  $M$  была достаточно велика, то происходил сбор и прикрепление частиц к домику, если  $M$  была мала, то модельная личинка искала новое место с подходящими

размерами частиц. Важно, что в модели использовались знания личинки о размерах последних протестированных и прикрепленных частиц.

Эта модель качественно согласуется с биологическими экспериментальными данными, в частности, как в эксперименте, так и в модели к чехлу-домику преимущественно прикрепляются крупные частицы.

#### **4. Контуры программы будущих исследований когнитивной эволюции**

Отталкиваясь от анализа современных биологически инспирированных моделей автономных «организмов», можно предложить следующие этапы моделирования когнитивной эволюции.

*А) Моделирование адаптивного поведения агентов с несколькими естественными потребностями: питание, размножение, безопасность.* Это могло бы быть моделирование достаточно естественного и полноценного поведения простых модельных организмов. Моделирование в этом направлении уже начато, см. ниже.

*Б) Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов.* Такой переход можно рассматривать, как появление в «сознании» животного свойства «понятие». Обобщенные образы можно представить как мысленные аналоги наших слов, не произносимые животными, но реально используемые ими. Использование понятий приводит к существенному сокращению и требуемой памяти, и времени обработки информации, поэтому оно должно быть эволюционно выгодным.

*В) Исследование процессов формирования причинных связей в памяти животных.* Запоминание причинно-следственных связей между событиями во внешней среде и адекватное использование этих связей в поведении – одно из ключевых свойств активного познания животным закономерностей внешнего мира. Такая связь формируется, например, при выработке условного рефлекса: животное запоминает связь между условным стимулом (УС) и следующим за ним безусловным стимулом (БС), что позволяет ему предвидеть события в окружающем мире и адекватно использовать это предвидение.

Естественный следующий шаг – переход от отдельных причинных связей к логическим выводам на основе уже сформировавшихся знаний.

*Г) Исследование процессов формирования логических выводов в «сознании» животных.* Фактически, уже на базе классического условного рефлекса животные способны делать «логический вывод» вида: {УС, УС → БС} ⇒ БС или «Если имеет место условный стимул, и за условным стимулом следует безусловный, то нужно ожидать появления безусловного стимула». В определенной степени такие выводы подобны

выводам математика, доказывающего теоремы. И целесообразно разобратся в системах подобных выводов, понять, насколько адаптивна логика поведения животных и насколько она подобна нашей, человеческой логике.

Д) *Исследование коммуникаций, возникновения языка.* Наше мышление тесно связано с языком, с языковым общением между людьми. Поэтому целесообразно проанализировать: как в процессе биологической эволюции возникал язык общения животных, как развитие коммуникаций привело к современному языку человека, как развитие коммуникаций и языка способствовало развитию логики, мышления, интеллекта человека.

Перечисленные пункты очерчивают круг исследований от моделирования простейших форм поведения к логическим правилам, используемым в математике.

Опираясь на эти пункты, мы начали соответствующее моделирование.

## 5. Начальные шаги моделирования

### 5.1. Модель автономных агентов с естественными потребностями

В работе [Коваль и др., 2012] построена и исследована компьютерная модель автономных агентов, обладающих естественными для живых организмов потребностями: питание, безопасность и размножение. Потребностям агента соответствовали три фактора: питания ( $F_F$ ), безопасности ( $F_S$ ) и размножения ( $F_R$ ). Для каждого фактора имелся порог ( $T_F, T_S, T_R$ ), при превышении фактором которого удовлетворялась соответствующая потребность.

Между потребностями вводилась следующая иерархия: 1) пищевая потребность (наиболее приоритетна), 2) потребность безопасности, 3) потребность размножения (наименее приоритетна). Потребность считалась ведущей, если она была наиболее приоритетной из всех потребностей, для которых фактор потребности меньше соответствующего ему порога.

Рядом с агентом имелся хищник, активность которого периодически менялась; защищаясь от хищника, агент мог обеспечивать свою безопасность.

Система управления агента была основана на наборе правил вида:  $S_k \rightarrow A_k$  (в ситуации  $S_k$  нужно выполнить действие  $A_k$ ),  $k$  – номер правила. Веса правил  $W_k$  настраивались методом обучения с подкреплением [Саттон и др., 2011]. Ситуация  $S$  определялась активностью хищника рядом с агентом, предыдущим действием агента и ведущей потребностью. Каждый такт времени агент мог выполнять одно из следующих действий  $A$ : 1) поиск пищи, 2) питание, 3) подготовка к размножению, 4) размножение, 5) оборона, 6) покой.

Агент имел ресурс  $R$ , который немного уменьшался при выполнении действий и существенно увеличивался при питании. Активный хищник также мог отнимать у агента значительный ресурс.

После выполнения цепочки действий «подготовка к размножению», «размножение» агент рождал потомка, которому отдавал часть своего ресурса  $R$ . Выполняя действие «оборона», агент защищался от активного хищника.

Преимущественно выполнялись действия, соответствующие максимальным весам правил  $W_k$ . Факторы  $F_F, F_S, F_R$  увеличивались при выполнении действий, направленных на удовлетворение ведущих потребностей. Веса правил  $S_k \rightarrow A_k$  менялись следующим образом [Саттон и др., 2011]:

$$\Delta W_{t-1} = \alpha (F_{L,t} - F_{L,t-1} + \gamma W_t - W_{t-1}),$$

где  $F_{L,t}, F_{L,t-1}$  – факторы потребности, ведущей в такты времени  $t$  и  $t-1$ ,  $W_t$  и  $W_{t-1}$  – веса правил, примененных в такты  $t$  и  $t-1$ ,  $\alpha$  – параметр скорости обучения,  $\gamma$  – дисконтный фактор.

Результаты компьютерного моделирования иллюстрируются рис. 1, на котором показана типичная динамика факторов  $F_F, F_S$  и  $F_R$ . Видно, что при достаточно большом времени все три фактора достигают порога, т.е. все потребности удовлетворяются. Формируется цикличность поведения агента, характерный цикл показан участком «а». В цикле агент сначала накапливает внутренний ресурс  $R$ , затем его действия направлены на максимизацию безопасности, а когда обе потребности (пищевая и потребность безопасности) удовлетворены, агент размножается.

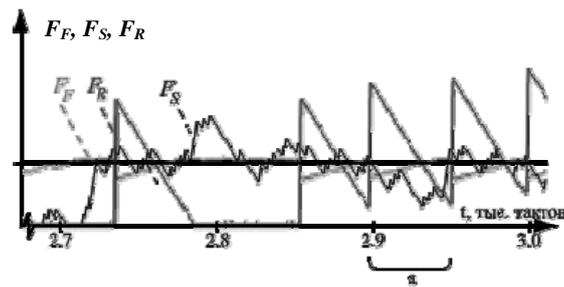


Рис. 1. Зависимость факторов  $F_F, F_S$  и  $F_R$ , соответствующих потребностям агента, от времени  $t$ ; горизонтальная линия показывает порог, соответствующий всем трем факторам

Таким образом, построена и исследована модель автономных агентов, обладающих несколькими потребностями (пункт *A* контуров программы). Показано, что происходит формирование циклов поведения, в которых последовательно удовлетворяются потребности питания, безопасности и размножения.

### **5.2. Модель формирования обобщенных эвристик и простых обобщенных образов**

Формирование обобщенных эвристик и простых обобщенных образов (пункт *B*) в процессе обучения агентов при поиске агентами пищи в двумерной клеточной среде было продемонстрировано в модели [Бесхлебнова и др., 2010], хотя пока эту модель целесообразно рассматривать как определенный начальный этап более полноценных исследований.

## **6. Заключение**

Сопоставляя пункты контуров программы будущих исследований когнитивной эволюции с упомянутыми моделями и с другими близкими по тематике работами, можно заключить, что уже имеются отдельные элементы, соответствующие каждому из пунктов. Образно говоря, у нас уже есть некоторые небольшие фрагменты картины, но мы еще не видим всей картины. Четкой последовательности серьезных, канонических моделей, которые показывали бы общую картину происхождения логического мышления, пока еще нет.

Тем не менее, в целом ряде направлений исследований развивается задел моделирования когнитивной эволюции. При этом объединяющим понятием для этих исследований может служить понятие *автономный агент*. С помощью автономных агентов можно моделировать когнитивные и «интеллектуальные» свойства как биологических организмов, так и искусственных аналогов живых организмов.

Также необходимо подчеркнуть междисциплинарные связи моделирования когнитивной эволюции: эти исследования связаны с основаниями науки, с основаниями математики, с теорией познания, с широким кругом когнитивных работ.

**Благодарности.** Работа выполнена при финансовой поддержке РФФИ (проект № 10-01-00129) и программы президиума РАН «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация».

## Список литературы

- [Бесхлебнова и др., 2010] Бесхлебнова Г.А., Редько В.Г. Модель формирования обобщенных понятий автономными агентами // Четвертая международная конференция по когнитивной науке: Тезисы докладов: В 2 томах. – Томск: ТГУ, 2010. Т. 1.
- [Коваль и др., 2012] Коваль А.Г. Редько В.Г. Поведение модельных организмов, обладающих естественными потребностями и мотивациями // Математическая биология и биоинформатика (электронный журнал). 2012. Т. 7. № 1. URL: [http://www.matbio.org/2012/Koval2012\(7\\_266\).pdf](http://www.matbio.org/2012/Koval2012(7_266).pdf)
- [Непомнящих и др., 2008] Непомнящих В.А., Попов Е.Е., Редько В.Г. Бионическая модель адаптивного поискового поведения // Известия РАН. Теория и системы управления. 2008. № 1.
- [Редько, 2005] Редько В.Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. – М.: УРСС, 2005.
- [Редько, 2006] Редько В.Г. (ред.). От моделей поведения к искусственному интеллекту. – М.: УРСС, 2006.
- [Редько, 2010] Редько В.Г. Моделирование когнитивной эволюции – перспективное направление исследований на стыке биологии и математики // Математическая биология и биоинформатика (электронный журнал). 2010. Т. 5. № 2. URL: [http://www.matbio.org/downloads/Redko2010\(5\\_215\).pdf](http://www.matbio.org/downloads/Redko2010(5_215).pdf)
- [Саттон и др., 2011] Саттон Р.С., Барто Э.Г. Обучение с подкреплением. – М.: Бинном, 2011.
- [Krichmar et al, 2005] Krichmar J.L., Seth A.K., Nitz D.A., Fleischer J.G., Edelman G.M. Spatial navigation and causal analysis in a brain-based device modeling cortical-hippocampal interactions // Neuroinformatics. 2005. V. 3. No. 3
- [Langley et al, 2009] Langley P., Laird J.E., Rogers S. Cognitive architectures: Research issues and challenges // Cognitive Systems Research. 2009. V.10. No. 2.
- [Lehman et al, 2006] Lehman J.F., Laird L., Rosenbloom P. A gentle introduction to Soar: Architecture for human cognition: 2006 update (электронный ресурс). URL: <http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf>
- [Meyer et al, 1991] Meyer J.-A., Wilson S.W. (Eds.). From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior. – Cambridge: MIT Press, 1991.
- [Samsonovich et al, 2011] Samsonovich A.V., Johansdottir K.R. (Eds.). Biologically Inspired Cognitive Architectures 2011. Proceedings of Second Annual Meeting of the BICA Society. – Amsterdam et al.: IOS Press, 2011.
- [Turchin, 1987] Turchin V.F. A constructive interpretation of the full set theory // Journal of Symbolic Logic. 1987. V. 52. No. 1.
- [Vernon et al, 2007] Vernon D., Metta G., Sandini G. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents // IEEE Transactions on Evolutionary Computation (special issue on Autonomous Mental Development). 2007. V. 11. No. 2.