

## Зачем и как моделировать когнитивную эволюцию

В.Г. Редько

Научно-исследовательский институт системных исследований РАН, Москва

В докладе обсуждается новое направление исследований – моделирование когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов. Актуальность исследований когнитивной эволюции связана с глубокой гносеологической проблемой: почему формальное логическое человеческое мышление, казалось бы, совсем не связанное с реальным физическим миром, применимо к познанию природы? Исследование когнитивной эволюции естественно вести путем математического, компьютерного моделирования. Существенно, что имеется определенный задел моделирования когнитивной эволюции, развиваемый в направлении исследований «Адаптивное поведение». В докладе характеризуются начальные шаги рассматриваемого моделирования и предлагаются контуры программы будущих исследований когнитивной эволюции. Подчеркивается актуальность моделирования когнитивной эволюции.

### Ключевая гносеологическая проблема

По-видимому, наиболее серьезные и глубокие когнитивные процессы – это процессы научного познания. Но насколько способен человек познавать внешний мир? Почему формальный логический вывод, сделанный человеком, применим к реальным объектам в природе? Поясним эти вопросы. 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 более общей формулировке ключевая проблема может быть сформулирована так: почему логика человеческого мышления применима к познанию природы?

Близкие вопросы давно интересовали ученых. В 1781 году появилась знаменитая «Критика чистого разума» И. Канта [1], а два года спустя вышло популярное изложение «Критики...» «Прологомены ко всякой будущей метафизике, могущей появиться, как наука» [2] (термин «пролегомены» означает предварительные рассуждения, введение в изучение). И. Кант провел исследование познавательных процессов в определенном приближении – приближении фиксированного мышления взрослого человека. Он не задавался вопросом, откуда берутся познавательные способности, он

констатировал факт, что они существуют, и исследовал, как они работают. В результате этого анализа И. Кант пришел к выводу, что существует система категорий, концепций, логических правил и методов вывода, которые используются в познании природы. Эта система «чистого разума» имеет априорный характер – она существует в нашем сознании прежде всякого опыта – и является основой научного познания природы.

Естественно, что приближение фиксированного мышления человека наложило свой отпечаток: И. Кант утверждает – и в рамках данного приближения вполне логично – что так как «чистый разум» априорен, то наш рассудок в познавательном процессе предписывает свои законы природе [2]:

«...хотя вначале это звучит странно, но тем не менее верно, если я скажу: *рассудок не черпает свои законы (a priori) из природы, а предписывает их ей*».

Наверно, во времена И. Канта было разумно ограничиться приближением фиксированного мышления взрослого человека – все сразу не охватишь. Кроме того, не было еще теории происхождения видов Ч. Дарвина. Естественно, что после появления этой теории должна была произойти ревизия концепции априорного «чистого разума». И она произошла. Очень четко ее выразил один из основателей этологии (науки о поведении животных), лауреат Нобелевской премии К. Лоренц в статье «Кантовская концепция a priori в свете современной биологии» [3] (статья была написана в 1941 г.). Согласно К. Лоренцу кантовские априорные категории и другие формы «чистого разума» произошли в результате естественного отбора. Составляющие «чистого разума» возникали постепенно в процессе эволюции, в результате многочисленных взаимодействий с внешним миром. В эволюционном контексте «чистый разум» совсем не априорен, а имеет явные эволюционные эмпирические корни.

Фактически, И. Кант и К. Лоренц показали, что если не рассматривать эволюционное происхождение методов познания, то нет ответа на ключевой вопрос о применимости логического мышления человека к познанию природы.

Таким образом, важно исследовать когнитивную эволюцию, эволюцию познавательных способностей биологических организмов и с помощью моделирования разобраться, как в процессе эволюции возникали способности, обеспечивающие познание закономерностей внешнего мира, и

как развитие этих способностей привело к научному познанию природы.

При этом интересующий нас результат эволюции – правила логического вывода, используемые при математических доказательствах, – известны и достаточно хорошо формализованы [4]. Важно подчеркнуть, что были и попытки пересмотра оснований математики в близком к исследованиям когнитивной эволюции контексте. В статье В.Ф. Турчина (1987 г.) [5] предпринята весьма нетривиальная попытка пересмотра оснований математики и рассмотрена возможность построения предиктивных логических правил в контексте теории множеств и кибернетического подхода к обоснованию математики. В.Ф. Турчиным также была предложена концептуальная теория метасистемных переходов [6], которая может быть использована при моделировании когнитивной эволюции.

Как же конкретно вести моделирование когнитивной эволюции? Есть ли задел таких исследований? Оказывается, что да, есть. Модели познавательных свойств живых организмов развиваются в рамках направления исследований «Адаптивное поведение».

#### **Направление исследований «Адаптивное поведение» – задел моделирования когнитивной эволюции**

Направление «Адаптивное поведение» развивается с начала 1990-х годов [7, 8]. Основной подход направления – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы часто называются «аниматами» (от англ. animal и robot: animal + robot = animat), агентами, автономными агентами. Исследователи адаптивного поведения разрабатывают такие модели, которые применимы к описанию поведения как реального животного, так и искусственного анимата. Дальняя цель этих работ (пока еще не реализованная) – анализ эволюции когнитивных способностей животных и происхождения интеллекта человека [9]. Эта цель близка к задаче моделирования когнитивной эволюции.

Приложения моделей адаптивного поведения – искусственный интеллект, робототехника, модели адаптивного поведения в социально-экономических системах.

Подчеркнем, что в современных исследованиях адаптивного поведения используется ряд нетривиальных компьютерных методов:

- нейронные сети,
- генетический алгоритм и другие методы эволюционной оптимизации,
- классифицирующие системы (Classifier Systems) [10],
- обучение с подкреплением (Reinforcement Learning) [11].

В моделях адаптивного поведения часто используется феноменологический подход. То есть

предполагается, что существуют формальные правила адаптивного поведения, и эти правила не обязательно связаны с конкретными микроскопическими нейронными или молекулярными структурами, которые есть у живых организмов. Скорее всего, такой феноменологический подход для исследований адаптивного поведения вполне имеет право на существование. В пользу этого тезиса приведем аналогию из физики. Есть термодинамика, и есть статистическая физика. Термодинамика описывает явления на феноменологическом уровне, статистическая физика характеризует те же явления на микроскопическом уровне. В физике термодинамическое и стат-физическое описания относительно независимы друг от друга, и вместе с тем взаимодополнительны. Повидимому, и для описания живых организмов может быть аналогичное соотношение между феноменологическим (на уровне поведения) и микроскопическим (на уровне нейронов и молекул) подходами. При этом естественно ожидать, что для исследования систем управления адаптивным поведением и моделей когнитивной эволюции феноменологический подход должен быть более эффективен, чем микроскопический (по крайней мере, на начальных этапах работ), так как очень трудно сформировать целостную картину поведения на основе анализа всего сложного многообразия функционирования нейронов, синапсов, молекул.

В следующем разделе отмечены работы по адаптивному поведению, непосредственно связанные с моделированием когнитивной эволюции.

#### **Модели адаптивного поведения агентов с когнитивными свойствами**

В последнее время активно развиваются исследования антиципаторного поведения, при котором животные предвидят будущие ситуации и используют это предвидение при организации адаптивного поведения [12].

Интересные работы связаны с методами формализации правил принятия решения. В [13] предложены обобщенные, соответствующие разным уровням биологической эволюции, правила принятия решений. Правила учитывают ассоциативные модели, модели классического и инструментального условного рефлексов, модели предвидения результатов действия. Проработаны схемы обучения и принятия решения на основе этих правил, проведено соответствующее компьютерное моделирование, подтвердившее эффективность предложенных правил.

Некоторые работы анализируют эволюционные аспекты становления нейронных структур мозга животных, обеспечивающих познавательные процессы. В [14] проанализирована эволюция нейронных структур мозга, играющих важную роль при выборе действий, обеспечивающих адаптивное поведение.

В недавних работах Е.Е. Витяева и А.В. Демина [15,16] начато интересное исследование

«индуктивной логики» аниматов на основе биологических теорий П.К. Анохина и П.В. Симонова. В частности, разработана логическая модель адаптивной системы управления анимата и на примере задачи фуражирования показано, что логическая модель может быть эффективней метода обучения с подкреплением [11]. Отметим, что, развивая подход работ [15,16], можно попытаться проанализировать переход от индуктивной логики (точнее, той «логики», которую используют животные при организации своего поведения) к дедуктивной (той, которую используют математики при доказательстве теорем).

Итак, работы по когнитивным моделям адаптивного поведения ведутся. Хотя пока это скорее накопление материала, а не последовательное изучение когнитивной эволюции.

Каковы же эволюционные уровни, на которых стоит остановиться? Как от простых форм адаптивного поведения идти к логическим формам, используемым в научном познании? Попытка выделения ключевых эволюционных уровней сделана в следующем разделе.

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

Предложим контуры программы будущих исследований, нацеленных на моделирование когнитивной эволюции. При этом постараемся выделить наиболее существенные этапы, ведущие к логическому мышлению.

**А) Моделирование адаптивного поведения аниматов с несколькими естественными потребностями: питания, размножения, безопасности.** 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. Например, у собаки явно есть понятия «хозяин», «свой», «чужой», «пища». И важно осмыслить, как такой весьма нетривиальный переход мог произойти в процессе эволюции. 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.** Повидимому, 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:  $\{УС, УС \rightarrow БС\} \Rightarrow БС$  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 [17]. И целесообразно разобраться в системах подобных выводов, понять, насколько адаптивна логика поведения животных и насколько она подобна нашей, человеческой логике.

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

Опираясь на эти пункты, мы начали соответствующее моделирование. В [18] была разработана формальная модель упрощенных агентов, которые обладают потребностями питания, размножения, безопасности (пункт А). Модель показала естественное поведение агентов, также была продемонстрирована важная роль размножения при эволюционной оптимизации систем управления агентов. Пример достаточно нетривиального формирования обобщенных эвристик и простых обобщенных образов (пункт Б) в процессе обучения агентов при поиске агентами пищи в двумерной клеточной среде был продемонстрирован в [19, 20]. Тем не менее, пока модели [18–20] целесообразно рассматривать как довольно упрощенные модели, как определенный начальный этап более полноценных исследований.

Сопоставляя пункты контуров программы с отмеченными работами [12–20], 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.

#### **Еще раз об актуальности моделирования когнитивной эволюции**

В заключение просуммируем аргументы в пользу актуальности исследований когнитивной эволюции.

- Эти исследования связаны с основаниями науки, с основаниями математики, с серьезной проблемой: почему логические выводы, математические доказательства применимы к реальной природе.
- Данные исследования интересны с философской, эпистемологической точки зрения – они нацелены на прояснение причин применимости человеческого мышления в познании природы.
- В направлении исследований «Адаптивное поведение», дальняя цель которого близка к задаче моделирования эволюции познавательных способностей живых организмов, развита задел по математическому и компьютерному моделированию когнитивной эволюции.
- Эти исследования интересны с точки зрения развития когнитивных наук, так как они связаны с важными когнитивными процессами – процессами научного познания.

### Литература

1. Кант И. Критика чистого разума. Соч. в 6 т. М.: Мысль, 1964. Т. 3. С. 69–695.  
URL: <http://www.philosophy.ru/library/kant/01/sod.html>
2. Кант И. Прологомены ко всякой будущей метафизике, могущей появиться как наука. Соч. в 6 т. М.: Мысль, 1965. Т. 4. Ч. 1. С. 67–210.  
URL: <http://www.philosophy.ru/library/kant/metaf.html>
3. Лоренц К. Кантовская концепция а priori в свете современной биологии // Эволюция. Язык. Познание. М.: Языки русской культуры, 2000. С. 15–41.  
URL: [http://gfrkjh.narod.ru/lorenz/lorenz\\_apriori.html](http://gfrkjh.narod.ru/lorenz/lorenz_apriori.html)
4. Йодельсон А.В., Минц Г.Е. (Ред.) Математическая теория логического вывода. М.: Наука, 1967.
5. Turchin V.F. A constructive interpretation of the full set theory // Journal of Symbolic Logic. 1987. V. 52. № 1. P. 172–201.
6. Турчин В.Ф. Феномен науки. Кибернетический подход к эволюции. М.: Наука, 1993.
7. 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.
8. Редько В.Г. (Ред.) От моделей поведения к искусственному интеллекту. М.: УРСС, 2006.
9. Donnart J.Y., Meyer J.-A. Learning reactive and planning rules in a motivationally autonomous animat // IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics. 1996. V. 26. № 3. P. 381–395.  
URL: <http://www.isir.upmc.fr/files/1996ACLI367.pdf>
10. Holland J.H., Holyoak K.J., Nisbett R.E., Thagard P. Induction: Processes of Inference, Learning, and Discovery. Cambridge: MIT Press, 1986.
11. Sutton R., Barto A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. URL: <http://www.cs.ualberta.ca/~sutton/book/the-book.html>
12. Butz M.V., Sigaud O., Pezzulo G., Baldassarre G. (Eds.). Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior. LNAI 4520, Berlin, Heidelberg: Springer Verlag, 2007.
13. Witkowski M. An action-selection calculus // Adaptive Behavior. 2007. V. 15. № 1. P. 73–97.  
URL: [http://www.ee.ic.ac.uk/mark/papers/Witkowski\\_2007\\_An\\_Action\\_Selection\\_Calculus\\_ABI5-1.pdf](http://www.ee.ic.ac.uk/mark/papers/Witkowski_2007_An_Action_Selection_Calculus_ABI5-1.pdf)
14. Prescott T.J. Forced moves or good tricks in design space? Landmarks in the evolution of neural mechanisms for action selection // Adaptive Behavior. 2007. V. 15. № 1. P. 9–31. URL: [http://www.abrg.group.shef.ac.uk/content/attach/0008.prescott\\_AB07.pdf](http://www.abrg.group.shef.ac.uk/content/attach/0008.prescott_AB07.pdf)
15. Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова // Нейроинформатика (электронный журнал). 2008. Т. 3. № 1. С. 25–78. URL: <http://www.niisi.ru/iont/ni/Journal/V3/N1/Vityaev.pdf>
16. Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления // Нейроинформатика (электронный журнал). 2008. Т. 3. № 1. С. 79–108. URL: <http://www.niisi.ru/iont/ni/Journal/V3/N1/DeminVityaev.pdf>
17. Редько В.Г. Моделирование когнитивной эволюции – перспективное направление исследований на стыке биологии и математики // Математическая биология и биоинформатика (электронный журнал). 2010. Т. 5. № 2. С. 215–229. URL: [http://www.matbio.org/downloads/Redko2010\(5\\_215\).pdf](http://www.matbio.org/downloads/Redko2010(5_215).pdf)
18. Редько В.Г., Бесхлебнова Г.А. Моделирование адаптивного поведения автономных агентов // Нейрокомпьютеры: разработка, применение. 2010. № 3. С. 33–38. URL: <http://www.niisi.ru/iont/ni/Publications/Redko/RedkoBeskhleNc.pdf>
19. Редько В.Г., Бесхлебнова Г.А. Модель формирования адаптивного поведения автономных агентов // Интегрированные модели и мягкие вычисления в искусственном интеллекте: V Международная научно-практическая конференция. М.: Физматлит, 2009. Т. 1. С. 70–79. URL: <http://www.niisi.ru/iont/ni/Publications/Redko/RedkoBeskhleKolomna.pdf>
20. Бесхлебнова Г.А., Редько В.Г. Модель формирования обобщенных понятий автономными агентами // Четвертая международная конференция по когнитивной науке. Томск: ТГУ, 2010. Т. 1. С. 174–175. URL: <http://www.niisi.ru/iont/ni/Publications/Redko/BeskhleRedkoTomsk.pdf>