

Редько В.Г. Подход к моделированию когнитивной эволюции // Сборник трудов Международной научной конференции "Василий Васильевич Налимов - математик и философ", М.: МАКС Пресс, 2011 (372 с., ISBN 978-5-317-03835-9). С.114-122.

ПОДХОД К МОДЕЛИРОВАНИЮ КОГНИТИВНОЙ ЭВОЛЮЦИИ

Редько В.Г.

Научно-исследовательский институт системных исследований РАН
e-mail: vgredko@gmail.com

Аннотация

Обсуждается подход к моделированию когнитивной эволюции – эволюции познавательных способностей биологических организмов. Это моделирование связано с основаниями науки, основаниями математики. Характеризуется направление «Адаптивное поведение», в рамках которого развивается задел исследований когнитивной эволюции. Предлагаются контуры программы будущих исследований когнитивной эволюции. Подчеркивается актуальность этих исследований.

1. Предпосылки исследований когнитивной эволюции

Важность моделирования когнитивной эволюции связана с глубокой гносеологической проблемой: почему формальный логический вывод, сделанный человеком, применим к реальным объектам в природе? Действительно, математики доказывают теоремы. Но математик делает логические выводы, доказывает теоремы независимо от внешнего мира, используя свое мышление. Почему же эти выводы применимы к реальной природе? Или в более общей формулировке: 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...» «Пролегомены 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 просто 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 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]. Согласно К. Лоренцу кантовские априорные категории и другие формы «чистого разума» произошли в результате естественного отбора:

«Не были бы законы разума, необходимые для априорного мышления, совершенно иными, если бы они сформировались абсолютно другим историческим способом и если бы мы, следовательно, были оснащены иным типом нервной системы? И вообще, возможно ли, чтобы законы нашего когнитивного аппарата не были связаны с законами реального внешнего мира?».

«...наши категории и формы восприятия, зафиксированные до индивидуального опыта, адаптированы к внешнему миру в точности по тем же причинам, по которым копыто лошади адаптировано к степному грунту еще до того, как лошадь рождается, а плавник рыбы – к воде до ее появления из икринки».

То есть, 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 поставленный в начале данной статьи вопрос о применимости логического мышления человека к познанию природы.

Как же разобраться в происхождении логических форм мышления? Можно ли промоделировать это происхождение? Какие методы использовать? Как такие исследования связаны с основаниями науки, основаниями математики? Есть ли задел работ по моделированию когнитивной эволюции – эволюции познавательных способностей животных? Далее обсуждаются перечисленные вопросы. Но, прежде всего, кратко остановимся на вопросе: можно ли вообще в принципе вести такие исследования?

Итак, можно ли проследить эволюционные корни логических правил, используемых в дедуктивных доказательствах? По мнению автора настоящей статьи, да, можно. Приведем следующую аналогию.

Одно из элементарных правил, которое использует математик в логических заключениях, – правило *modus ponens*: «если имеет место *A*, и из *A* следует *B*, то имеет место *B*», или

$$\frac{A, A \rightarrow B}{B} .$$

Перейдем от математика к собаке, у которой вырабатывают классический условный рефлекс. При выработке рефлекса в памяти собаки формируется связь «за УС должен последовать БС» (УС – условный стимул, БС – безусловный стимул). Когда после выработки рефлекса собаке предъявляют УС, то она, «помня» о хранящейся в ее памяти «записи» УС → БС, делает элементарный «вывод»

$$\frac{УС, УС \rightarrow БС}{БС} .$$

И собака ожидает БС.

Конечно, чисто дедуктивное применение формального правила *modus ponens* математиком и основанный на обобщении опыта индуктивный «вывод», который делает собака, явно различаются. Тем не менее, и в первом и во втором случаях речь идет о следственной связи между математическими утверждениями либо событиями: из *A* следует *B*, за УС следует БС. Хотя контексты следственной связи, следственной зависимости в этих двух случаях различны: в первом случае *B* есть формальное логическое следствие *A*; во втором случае в процессе наблюдений формируется и запоминается причинно-следственная связь между событиями УС и БС, следующими одно за другим во времени. Итак, применение правила *modus ponens* при дедуктивном выводе аналогично «выводу», происходящему после выработки классического условного рефлекса.

Указанная аналогия позволяет задуматься об эволюционных корнях логических правил, используемых в математике. Не исключено, что

простейшие формы дедуктивных правил возникли как обобщение индуктивных правил.

Таким образом, можно анализировать эволюционные корни логического мышления и строить модели эволюционного происхождения логических правил, используемых в научном познании. Естественно, что при моделировании когнитивной эволюции целесообразно рассматривать использование познавательных способностей животных при их приспособлении к внешней среде. Познавательные свойства были полезны для животных и закреплялись в процессе естественного отбора.

При этом результат эволюции – правила логического вывода, используемые при математических доказательствах, – известны и достаточно хорошо формализованы [4]. В основе этих выводов – элементарные правила, такие как *modus ponens*.

Более того, были попытки пересмотра оснований математики, связанные с логическими правилами. 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».

2. Направление исследований «Адаптивное поведение» – задел исследований 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 приспособливаться 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 нереализованная) – анализ эволюции когнитивных способностей животных и происхождения интеллекта человека [9]. Эта цель близка к задаче моделирования когнитивной эволюции. Отметим модели адаптивного поведения, достаточно близкие к моделированию когнитивной эволюции.

В ряде ранних работ исследовались модели условных рефлексов. Здесь мы приведем только ссылки на некоторые из моделей [10, 11].

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

Интересные работы связаны с методами формализации правил принятия решения. Например, 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. Например, 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 [15-17] 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 известного метода обучения с подкреплением [18]. Отметим, что развивая подход работ [15-17], 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.

3. Контуры программы будущих исследований когнитивной эволюции

Предложим контуры программы будущих исследований, нацеленных на моделирование когнитивной эволюции. 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. По-видимому, 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 (см. выше, раздел 1). И целесообразно разобраться 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. В [19] 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. Модель [19] может рассматриваться как опорная модель, отталкиваясь от которой имеет смысл проводить более близкое к биологическим примерам моделирование. Интересно, что недавно появились идейно близкие работы других исследователей. Например, в работе [20] анализировалось поведение агентов с несколькими потребностями с учетом мотиваций агента.

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 [21, 22]. Тем не менее, пока модели [21, 22] целесообразно рассматривать как довольно упрощенные модели, как определенный начальный этап более полноценных исследований.

Сопоставляя пункты контуров программы с отмеченными работами [10-17, 19-22], 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.

4. 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, развивается 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-ти томах. Т.3. М.: Мысль, 1964. С. 69-695. См. также: <http://www.philosophy.ru/library/kant/01/sod.html>
2. *Кант И.* Прологомены ко всякой будущей метафизике, могущей появиться как наука. Соч. в 6-ти томах. Т.4, часть 1. М.: Мысль, 1965. С. 67-210. См. также: <http://www.philosophy.ru/library/kant/metaf.html>
3. *Лоренц К.* Кантовская концепция а priori в свете современной биологии. Сб. Эволюция. Язык. Познание. (Отв. ред. И.П. Меркулов). М.: Языки русской культуры, 2000. С. 15-41. См. также: http://gfrkjh.narod.ru/lorenz/lorenz_apriori.html
4. Математическая теория логического вывода (под ред. А.В. Идельсона и Г.Е. Минца). М.: Наука, 1967.
5. *Turchin V.F.* A constructive interpretation of the full set theory. *Journal of Symbolic Logic*, 1987. V. 52. No. 1. PP. 172 -201.
6. *Турчин В.Ф.* Феномен науки. Кибернетический подход к эволюции. М.: ЭТС, 2000. См. также: <http://www.refal.ru/turchin/phenomenon/>
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. От моделей поведения к искусственному интеллекту (под ред. В.Г. Редько). М.: Изд-во УРСС (Серия «Науки об искусственном»), 2006.
9. *Donnart J.Y. and 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. No.3. PP.381-395. See also: <http://www.isir.upmc.fr/files/1996ACLI367.pdf>

10. *Grossberg S.* Classical and instrumental learning by neural networks. *Progress in Theoretical Biology*. 1974. V.3. PP. 51-141. See also:
<http://cns-web.bu.edu/~steve/Gro1974ProgressTheorBiol.pdf>
11. *Barto A.G., Sutton R.S.* Simulation of anticipatory responses in classical conditioning by neuron-like adaptive element. *Behav. Brain Res.* 1982. V.4. PP. 221-235. See also:
<http://webdocs.cs.ualberta.ca/~sutton/papers/barto-sutton-82.pdf>
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. No. 1. PP. 73-97. See also:
<http://www.ee.ic.ac.uk/mark/papers/Witkowski%202007%20An%20Action%20Selection%20Calculus%20AB15-1.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. No. 1. PP. 9-31. See also:
http://www.abrg.group.shef.ac.uk/content/attach/0008.prescott_AB07.pdf
15. *Витяев Е.Е.* Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. Новосибирск: НГУ, 2006. См. также:
http://www.math.nsc.ru/AP/ScientificDiscovery/PDF/scientific_discovery.pdf
16. *Витяев Е.Е.* Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова. *Нейроинформатика (электронный рецензируемый журнал)*. 2008. Т. 3. № 1. С. 25-78. <http://www.niisi.ru/iont/ni/Journal/V3/N1/Vityaev.pdf>
17. *Демин А.В., Витяев Е.Е.* Логическая модель адаптивной системы управления. *Нейроинформатика (электронный рецензируемый журнал)*. 2008. Т. 3. № 1. С. 79-108.
<http://www.niisi.ru/iont/ni/Journal/V3/N1/DeminVityaev.pdf>
18. *Sutton R., Barto A.* *Reinforcement Learning: An Introduction*. Cambridge: MIT Press, 1998. See also:
<http://www.cs.ualberta.ca/~sutton/book/the-book.html>
19. *Редько В. Г., Бесхлебнова Г.А.* Моделирование адаптивного поведения автономных агентов. *Нейрокомпьютеры: разработка, применение*. 2010. № 3. С. 33-38. См. также:
<http://www.niisi.ru/iont/ni/Publications/Redko/RedkoBeskhlNc.pdf>
20. *Butz M.V., Shirinov E., Reif K.* Self-organizing sensorimotor maps plus internal motivations yield animal-like behavior. *Adaptive Behavior*. 2010. V. 18. No. 3-4. PP. 315-337. See also:

http://www.coboslab.psychologie.uni-wuerzburg.de/fileadmin/ext00209/user_upload/Publications/2010/ButzShirinovReif2010SelfOrganizingSensorimotorMapsPlusInternalMotivationsYieldAnimalLikeBehavior.pdf

21. *Редько В.Г., Бесхлебнова Г.А.* Модель формирования адаптивного поведения автономных агентов. Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник научных трудов V-й Международной научно-практической конференции. В 2-х томах. Т.1. М.: Физматлит, 2009. С. 70-79. См. также:

<http://www.niisi.ru/iont/ni/Publications/Redko/RedkoBeskhlKolomna.pdf>

22. *Бесхлебнова Г.А., Редько В.Г.* Модель формирования обобщенных понятий автономными агентами. Четвертая международная конференция по когнитивной науке: Тезисы докладов: В 2 т. Томск: ТГУ, 2010. Т. 1. С. 174-175. См. также:

<http://www.niisi.ru/iont/ni/Publications/Redko/BeskhlRedkoTomsk.pdf>