

Проблема происхождения интеллекта*

В.Г. Редько

Научно-исследовательский институт системных исследований РАН, Москва,
Россия

E-mail: vgreddko@gmail.com

Аннотация. Обсуждаются подходы к исследованию проблемы происхождения интеллекта человека путем математического и компьютерного моделирования когнитивной эволюции. Характеризуются философские предпосылки этих работ. Обсуждается направление исследований «Адаптивное поведение», которое может рассматриваться, как задел моделирования эволюции познавательных способностей биологических организмов. Предлагается эскизный план будущих исследований когнитивной эволюции.

Ключевые слова: проблема происхождения интеллекта, когнитивная эволюция, адаптивное поведение, моделирование

Цель настоящей работы – привлечь внимание к исследованию проблемы эволюционного происхождения интеллекта человека. Рассмотрены подходы к исследованию этой проблемы на основе математического и компьютерного моделирования когнитивной эволюции.

* Работа выполнена при финансовой поддержке РАН (Программа «Фундаментальные проблемы информатики и информационных технологий», проект 2-45) и РФФИ (проект 07-01-00180).

1. Философские предпосылки исследований когнитивной эволюции

1.1. Гносеологическая проблема и подходы к ее решению

Существует глубокая гносеологическая проблема: *почему человеческое мышление применимо к познанию природы?* Ведь далеко не очевидно, что те мыслительные процессы, которые мы используем в научном познании, применимы к процессам, происходящим в природе, так как эти два типа процессов различны. Охарактеризуем проблему. Рассмотрим, например, физику – одну из фундаментальных естественнонаучных дисциплин. Мощь физики связана с эффективным применением математики. Но математик делает логические умозаключения, доказывает теоремы независимо от внешнего мира. Почему же результаты, получаемые математиком, применимы к реальной природе? Или в общей формулировке: почему наше логическое мышление применимо к познанию природы?

Можно ли конструктивно подойти к решению этих вопросов? Скорее всего, да. Чтобы продемонстрировать такую возможность, будем рассуждать следующим образом.

Рассмотрим одно из элементарных правил, которое использует математик в логических заключениях, правило *modus ponens*: «если имеет место утверждение A , и из A следует B , то имеет место утверждение B », или $\{A, A \rightarrow B\} \Rightarrow B$.

А теперь перейдем от математика к собаке И.П. Павлова. Пусть у собаки вырабатывают условный рефлекс, в результате 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 \rightarrow БС, делает элементарный «вывод»: $\{УС, УС \rightarrow БС\} \Rightarrow БС$. И собака ожидает появление

БС. Конечно, применение правила *modus ponens* (чисто дедуктивное) математиком и индуктивный «вывод», который делает собака, явно различаются. Тем не менее, умозаключение математика и «вывод» собаки качественно аналогичны, поэтому можно исследовать эволюционные корни логических правил, используемых в математике. И эти произошедшие в результате эволюции правила логического вывода известны и достаточно хорошо формализованы [1]. В основе этих выводов – элементарные правила, подобные *modus ponens*.

Итак, мы можем поставить задачу исследования эволюционного происхождения логики, мышления, интеллекта. По-видимому, это исследование должно в себя включать построение математических и компьютерных моделей «интеллектуальных изобретений» биологической эволюции, таких как безусловный рефлекс, привыкание, классический условный рефлекс, инструментальный условный рефлекс, цепи рефлексов, ..., логика [2]. Целесообразно с помощью моделей представить общую картину эволюции когнитивных способностей животных и эволюционного происхождения интеллекта человека. Естественно, что такие исследования – огромный фронт работы, и задачу моделирования когнитивной эволюции можно пока рассматривать как сверхзадачу. Исследования в этом направлении могли бы обеспечить определенное обоснование применимости нашего мышления в научном познании, т.е., укрепить фундамент всего величественного здания науки.

Но прежде чем обсуждать аспекты моделирования когнитивной эволюции, давайте посмотрим, кто еще думал над близкими вопросами. Проследим цепочку: Д. Юм -> И. Кант -> К. Лоренц.

1.2. Д. Юм -> И. Кант -> К. Лоренц

В «Исследовании о человеческом познании» (1748) Давид Юм подверг сомнению понятие причинной связи [3]. А именно, он задался вопросом: почему когда мы видим, что за одним явлением A постоянно следует другое B , мы приходим к выводу, что A является причиной B ? Например, когда мы многократно наблюдаем, что Солнце освещает камень, и камень нагревается, мы говорим, что солнечный свет есть причина нагревания камня.

Фактически Юм задался вопросом: что заставляет нас делать выводы о происходящих в природе явлениях? Что лежит в основе этих выводов? Юм попытался понять, откуда мы берем основание заключать, что A есть причина B . Он посмотрел на этот вопрос, как он пишет, со всех сторон и не нашел никакого другого основания, кроме некоторого внутреннего чувства привычки. Есть какое-то наше внутреннее свойство, которое заставляет нас утверждать, что если за A постоянно следует B , то A есть причина B . И это внутреннее чувство заставляет нас после того, когда мы сделали такое умозаключение, и снова видим событие A , ожидать, что за A вновь последует и событие B .

Юм взглянул на наш познавательный процесс со стороны, извне. Он как бы вышел на некий мета-уровень рассмотрения наших собственных познавательных процессов и задался вопросом о том, откуда взялись эти познавательные процессы и почему они работают. Острота сомнений Юма состояла в том, что он задался вопросом о принципиальной способности человека познавать мир.

Остроту сомнений Юма очень хорошо почувствовал Иммануил Кант. Но Кант также видел мощь и силу современной ему науки. Во времена Канта уже была глубокая, серьезная и развитая математика, и была мощная

ньютоновская физика, дающая *картину мира*, позволяющая объяснить множество явлений на основе немногих четких предположений, использующая многозвенные и сильные математические дедуктивные выводы. И как же нужно было поступать Канту? Подвергнуть сомнению все эти познавательные процессы? И развивая сомнения Юма дальше, отвергнуть всю науку?

Конечно же, Кант, как научно образованный человек, не стал отвергать современную ему науку, а постарался разобраться, как же работают познавательные процессы. В результате появились знаменитая «Критика чистого разума» [4] и ее популярная интерпретация – «Пролегомены ко всякой будущей метафизике, могущей появиться, как наука» [5]. Кант провел исследование познавательных процессов в определенном приближении – приближении фиксированного мышления взрослого человека. Он не анализировал происхождение познавательных способностей, он констатировал факт, что они существуют, и исследовал, как они работают. В результате этого анализа Кант пришел к выводу, что существует система категорий, концепций, логических правил и методов вывода (таких как заключения относительно причинных связей между событиями), 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: Кант утверждает – и вполне логично – что так как «чистый разум» априорен, то наш рассудок в познавательном процессе предписывает свои законы природе ([5], стр. 140):

«... хотя в начале 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 Канта было разумно ограничиться приближением фиксированного мышления взрослого человека – всё сразу не охватишь. Кроме того, не было еще теории Чарльза Дарвина. Скорее всего, если бы Кант знал теорию происхождения видов, то он явно задумался бы об эволюционном происхождении «чистого разума». Тем более что эволюционные подходы были не чужды Канту – вспомним его знаменитую гипотезу происхождения Солнечной системы. Но использованное им приближение накладывает свои ограничения – оно не позволяет ответить на вопросы – откуда же взялись познавательные способности, познаем ли мы истинные законы природы или наш рассудок «предписывает их ей».

Естественно, что после появления теории происхождения видов Ч. Дарвина должна была произойти ревизия концепции априорного «чистого разума». И она произошла. Очень четко ее выразил Конрад Лоренц в знаменитой статье «Кантовская доктрина априорного в свете современной биологии» [6]. Согласно К. Лоренцу, кантовские априорные категории и другие формы «чистого разума» произошли в результате естественного отбора ([6], стр. 19):

«...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 этом контексте «чистый разум» совсем *не априорен*, а имеет явные эволюционные *эмпирические* корни.

Но в работе К. Лоренца была сделана только общая критика позиции Канта, которая еще не показывала, как подойти к исследованию проблемы

происхождения интеллекта. Отметим, что К. Лоренц был одним из основателей эволюционной эпистемологии, которая была развита в работах К. Поппера, Д. Кэмпбелла, Г. Фоллмера (см., например, [7]). Рассмотрим далее, насколько эволюционная эпистемология близка к изучению этой проблемы.

1.3. Возможен ли симбиоз теории познания с естественнонаучными исследованиями?

Согласно известной работе К. Поппера, два основных тезиса эволюционной эпистемологии состоят в следующем ([8], стр. 57):

«Первый тезис. Специфически человеческая способность познавать, как и способность производить научное знание, являются результатами естественного отбора. Они тесно связаны с эволюцией специфически человеческого языка».

«Второй тезис. Эволюция научного знания представляет собой в основном эволюцию в направлении построения все лучших и лучших теорий. Это – дарвинистский процесс. Теории становятся лучше приспособленными благодаря естественному отбору. Они дают нам все лучшую и лучшую информацию о действительности. (Они все больше и больше приближаются к истине.) Все организмы – решатели проблем: проблемы рождаются вместе с возникновением жизни».

При этом первый тезис считается почти тривиальным, а второй разворачивается и всесторонне исследуется. Согласно этим тезисам основное внимание в эволюционной эпистемологии уделяется изучению того, каковы познавательные процессы и насколько их можно сопоставить с процессами накопления информации в процессе биологической эволюции, но

эволюционная эпистемология практически не занимается исследованием эволюционного происхождения познавательных способностей человека.

Более близки к проблеме происхождения интеллекта концептуальные исследования когнитивной эволюции, развитые в монографии [9], в которой, в частности, особое внимание уделяется анализу процесса формирования логического мышления на этапах перехода от примитивного мышления первобытных племен к формальному логическому (от племен охотников до Аристотеля).

Но, насколько известно автору настоящей статьи, в современных философских работах как-то не ощущается понимания остроты проблемы о принципиальной способности человека познавать мир. Проблемы: почему с помощью нашего *человеческого* мышления, нашей логики, нашего интеллекта, нашего «чистого разума» мы в принципе способны познавать *природу*. Как правило, много говорится о том, каковы методы познания, формализуются эти методы, говорится о том, что трудно формализовать все их особенности, но нигде не ставится задача – разобраться в том, почему они применимы в принципе. Тем не менее, проблема достаточно острая. Надо либо подвергнуть сомнению все научные знания, которые получены с помощью человеческого мышления, в применимости которого к научному познанию можно сомневаться, либо заняться обоснованием этой применимости. Естественный подход к решению проблемы – исследовать биологические корни наших познавательных способностей и постараться разобраться, почему эти способности возникли, как и почему в процессе их эволюционного развития появилась возможность познания природы. Это исследование целесообразно вести на основе математических и компьютерных моделей эволюции познавательных свойств биологических организмов. Причем здесь, как это ни удивительно, возможен симбиоз

философской науки, теории познания с естественнонаучными исследованиями.

Как же конкретно вести моделирование когнитивной эволюции? Существует ли задел исследований в этой области? Оказывается, что да, существует. Такой задел развивается в направлении исследований «Адаптивное поведение», дальняя цель которого близка к задаче исследования когнитивной эволюции. В следующем разделе кратко характеризуется это направление, указываются методы исследования адаптивного поведения, перечислены ключевые лаборатории и группы, ведущие эти работы. Также представлен начальный этап работ по перспективному проекту «Мозг анимата», нацеленному на формирование общей платформы для изучения систем управления адаптивного поведения модельных «организмов».

2. Модели адаптивного поведения – задел исследований когнитивной эволюции

2.1. From Animal to Animat – исследование общих принципов поведения животного и робота

С начала 1990-х годов активно развивается направление «Адаптивное поведение» [10]. Основной подход этого направления – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы называются «аниматами» (от англ. animal + robot = animat). Поведение аниматов имитирует поведение животных. Исследователи адаптивного поведения стараются разрабатывать именно такие модели,

которые применимы к описанию поведения как реального животного, так и искусственного анимата [11].

Программа-минимум направления «Адаптивное поведение» – исследовать архитектуры и принципы функционирования, которые позволяют животным или роботам жить и действовать в переменной внешней среде.

Программа-максимум этого направления – попытаться проанализировать эволюцию когнитивных способностей животных в контексте происхождения интеллекта человека [12]. Программа-максимум соответствует очерченной выше задаче моделирования когнитивной эволюции.

Для исследований «Адаптивного поведения» характерен синтетический подход: здесь конструируются архитектуры, обеспечивающие «интеллектуальное» поведение аниматов. Причем это конструирование аналогично тому, как это делает инженер: исследователь сам изобретает архитектуры, подразумевая, конечно, что какие-то подобные структуры, обеспечивающие адаптивное поведение, должны быть у реальных животных. При этом данное направление исследований рассматривается как бионический подход к разработке систем искусственного интеллекта.

Отметим, что хотя «официально» направление «Адаптивное поведение» было провозглашено в 1990 году, были явные провозвестники этого направления. Приведем примеры из истории отечественной науки. В 1960-х годах блестящий кибернетик и математик М.Л. Цетлин предложил и исследовал модели автоматов, способных адаптивно приспосабливаться к окружающей среде. Работы М.Л. Цетлина инициировали целое научное направление, получившее название «коллективное поведение автоматов» [13]. В 1960-70-х годах под руководством талантливого кибернетика М.М. Бонгарда был предложен интересный проект «Животное», направленный на моделирование адаптивного поведения искусственных организмов с

иерархией целей и подцелей [14,15]. Хороший обзор ранних работ по адаптивному поведению представлен в книге М.Г. Гаазе-Рапопорта, Д.А. Поспелова «От амебы до робота: модели поведения» [15].

Подчеркнем, что в современных исследованиях адаптивного поведения используется ряд нетривиальных компьютерных методов (для краткости приводим только ссылки на ключевые монографии по этим методам):

- нейронные сети [16],
- генетический алгоритм [17] и другие методы эволюционной оптимизации,
- классифицирующие системы (Classifier Systems) [18],
- обучение с подкреплением (Reinforcement Learning) [19].

2.2. Исследователи адаптивного поведения

Перечислим университеты и лаборатории, в которых проводятся активные исследования по адаптивному поведению:

- AnimatLab (Париж, руководитель лаборатории – один из инициаторов анимат-подхода Жан-Аркадий Мейер) [20]. Исследования AnimatLab предполагают, что система управления анимата может формироваться и модифицироваться посредством 1) *обучения*, 2) *индивидуального развития* (онтогенеза) и 3) *эволюции*.
- Лаборатория искусственного интеллекта в университете Цюриха (руководитель Рольф Пфейфер) [21]. Основной подход этой лаборатории – познание природы интеллекта путем создания («understanding by building»). Он включает в себя 1) построение моделей биологических систем, 2) исследование общих принципов естественного интеллекта животных и человека, 3) использование этих принципов при конструировании роботов и других искусственных интеллектуальных систем.

- Лаборатория искусственной жизни и роботики в Институте когнитивных наук и технологий (Рим, руководитель Стефано Нолфи) [22], ведущая исследования в области эволюционной роботики и принципов формирования адаптивного поведения.
- Институт нейронаук Дж. Эдельмана (Сан-Диего, США), где ведутся разработки поколений моделей работы мозга (Darwin I, Darwin II, ...) и исследования поведения искусственного организма NOMAD (Neurally Organized Mobile Adaptive Device), построенного на базе этих моделей [23].

В России исследования адаптивного поведения пока ведутся скромными усилиями ученых-энтузиастов, среди работ которых следует отметить [24]:

- модели поискового адаптивного поведения на основе спонтанной активности, приводящей к переключениям между разными тактиками поиска, например, тактикой движения по градиенту источника запаха (при поиске пищи) и тактикой случайного поиска (В.А. Непомнящих, Институт биологии внутренних вод им. И.Д. Папанина РАН);
- концепции и модели автономного адаптивного управления на основе аппарата эмоций (А.А. Жданов, Институт точной механики и вычислительной техники им. С.А. Лебедева РАН);
- разработку принципов построения систем управления антропоморфных и гуманоидных роботов (Л.А. Станкевич, Санкт-Петербургский государственный политехнический университет);
- разработку нейросетевых моделей поведения роботов и робототехнических устройств (А.И. Самарин, НИИ нейрокибернетики им. А.Б. Когана ЮФУ);
- модели адаптивного поведения на основе эволюционных и нейросетевых методов (В.Г. Редько, М.С. Бурцев, О.П. Мосалов, НИИ системных

исследований РАН, Институт прикладной математики им. М.В. Келдыша РАН).

2.3. Проект «Мозг анимата»

Анализ исследований в рамках анимат-подхода показывает, что разработанные модели пока еще очень фрагментарны и иллюстрируют только отдельные стороны адаптивного поведения. Поэтому в наших работах предложен проект «Мозг анимата»^{*)}, который нацелен на формирование общей схемы построения таких моделей [25-27]. Проект основан на теории функциональных систем П.К. Анохина [28].

Предполагается, что система управления аниматом имеет иерархическую архитектуру, базовым элементом которой является отдельная функциональная система (ФС). Верхний уровень соответствует основным потребностям организма: потребностям питания, размножения, безопасности, накопления знаний. Более низкие уровни соответствуют тактическим целям поведения. Блоки всех этих уровней реализуются с помощью ФС. Управление с верхних уровней может передаваться на нижние уровни и возвращаться назад.

В одном из первых вариантов архитектуры «Мозга анимата» предполагалась простая формализация ФС на основе нейронных сетей, которые обучаются методом обучения с подкреплением (без учителя, путем непосредственного взаимодействия с внешней средой, на основе только поощрений и наказаний, получаемых от внешней среды) [26]. Каждая ФС состоит из двух нейронных сетей: Модели и Критика. Нейронная сеть Модели предназначена для прогнозирования будущей ситуации при выполнении данного действия в данной текущей ситуации. Роль Критика

^{*)} Термин «Мозг анимата» был предложен К.В. Анохиным

состоит в оценке качества ситуаций. Качество ситуации определяется как количество ожидаемой суммарной награды, которую можно получить, исходя из данной ситуации. На основе прогнозов и оценок выбираются те действия анимата, которые приводят к ситуациям с максимальными значениями качества, сами оценки качества ситуаций постепенно уточняются в процессе обучения. Таким образом, в нейронных сетях Моделей и Критиков запоминаются следующие знания о внешнем мире: 1) знания об ожидаемых ситуациях и 2) знания о качестве ситуаций. Эти знания распределены в иерархической системе управления анимата. Система управления работает в соответствии с потребностями модельного организма.

В настоящее время проект «Мозг анимата» активно развивается. Предложены новые нейросетевые архитектуры, реализующие иерархию ФС, начато компьютерное моделирование аниматов, имеющих естественные потребности (питания, размножения, безопасности).

3. Контуры программы будущих исследований

Анализ современных исследований адаптивного поведения показывает, что хотя проделана большая работа, ученые еще очень далеки от понимания того, как возникали и развивались когнитивные способности животных, как процесс когнитивной эволюции привел к возникновению интеллекта человека [24,29]. Предложим эскизный план будущих исследований когнитивной эволюции, направленных на изучение проблемы происхождения интеллекта.

А) Разработка схем и моделей адаптивного поведения на базе проекта «Мозг анимата». Воплощение в конкретные модели конструкций систем управления аниматов и их изучение разумно начать с анализа целостного

адаптивного поведения простых аниматов, имеющих потребности питания, размножения, безопасности. Компьютерное моделирование в этом направлении уже начато.

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 произносимых животными, но реально используемых ими. Например, у собаки явно имеются понятия «хозяин», «свой», «чужой», «пища». Важно осмыслить, как такой весьма нетривиальный переход к сокращенному описанию взаимодействий с внешним миром на основе понятий мог произойти в процессе эволюции.

В) Исследование процессов запоминания причинных связей. 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: $\{УС, УС \rightarrow BC\} \Rightarrow BC$ или «Если имеет место условный стимул, и за условным стимулом следует безусловный, то нужно ожидать появления безусловного стимула». Можно даже говорить, что такие выводы подобны выводам математика, доказывающего теоремы (см. выше). 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. Работы в этих направлениях уже ведутся, но четкой последовательности серьезных, канонических моделей еще нет.

ЛИТЕРАТУРА

1. Математическая теория логического вывода (под ред. А.В. Идельсона и Г.Е. Минца). М.: Наука, 1967.
2. Воронин Л.Г. Эволюция высшей нервной деятельности. М.: Наука, 1977.
3. Юм Д. Исследование о человеческом познании. Соч. в 2-х томах. Т.2. М.: Мысль, 1966. С. 5-169.

4. Кант И. Критика чистого разума. Соч. в 6-ти томах. Т.3. М.: Мысль, 1964. С. 69-695.
5. 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.
6. 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.
7. Кузнецов Н.А., Баксанский О.Е., Гречишкина Н.А. Происхождение знания: истоки и основания // Информационные процессы, 2007. Т. 7. N. 1. С. 72-92. <http://www.jip.ru/2007/72-92-2007.pdf>
8. Поппер К. Эволюционная эпистемология // Сб. Эволюционная эпистемология и логика социальных наук: Карл Поппер и его критики. Составление Д.Г. Лахути, В.Н. Садовского, В.К. Финна. М.: Изд-во УРСС, 2000. С. 57-74.
9. Меркулов И.П. Когнитивная эволюция. М.: Наука, 1999.
10. Meyer J.-A., Wilson S. W. (Eds) From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior. The MIT Press: Cambridge, Massachusetts, London, England. 1991.
11. Непомнящих В.А. Поиск общих принципов адаптивного поведения живых организмов и аниматов // Новости искусственного интеллекта. 2002. N. 2. С. 48-53.
12. 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. Vol. 26. N. 3. PP.381-395.
13. Цетлин М.Л. Исследования по теории автоматов и моделирование биологических систем. М.: Наука, 1969.

14. Бонгард М.М., Лосев И.С., Смирнов М.С. Проект модели организации поведения – «Животное» // Моделирование обучения и поведения. М.: Наука, 1975. С.152-171.
15. Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота: модели поведения. М.: Наука, 1987. М.: Изд-во УРСС, 2004.
16. Хайкин С. Нейронные сети: полный курс, 2-е издание. М.: Издательский дом «Вильямс», 2006.
17. Holland J.H. Adaptation in Natural and Artificial Systems. Ann Arbor, MI: The University of Michigan Press, 1975 (1st edn). Boston, MA: MIT Press., 1992 (2nd edn).
18. Holland J.H., Holyoak K.J., Nisbett R.E., Thagard P. Induction: Processes of Inference, Learning, and Discovery. Cambridge, MA: MIT Press, 1986.
19. Sutton R., Barto A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. See also: <http://www.cs.ualberta.ca/~sutton/book/the-book.html>
20. Сайт AnimatLab: <http://animatlab.lip6.fr/index.en.html>
21. Сайт Artificial Intelligence Laboratory, University of Zurich: <http://www.ifi.unizh.ch/groups/ailab/>
22. Сайт Laboratory of Artificial Life and Robotics: <http://gral.istc.cnr.it/>
23. Сайт Neuroscience Institute: <http://www.nsi.edu/>
24. От моделей поведения к искусственному интеллекту (под ред. В.Г. Редько). М.: Изд-во УРСС (Серия «Науки об искусственном»), 2006.
25. Анохин К.В., Бурцев М.С., Зарайская И.Ю., Лукашев А.О., Редько В.Г. Проект «Мозг анимата»: разработка модели адаптивного поведения на основе теории функциональных систем // Восьмая национальная конференция по искусственному интеллекту с международным участием. Труды конференции. М.: Физматлит, 2002. Т. 2. С. 781-789.

- 26.Red'ko V.G., Prokhorov D.V., Burtsev M.S. Theory of functional systems, adaptive critics and neural networks // International Joint Conference on Neural Networks, Budapest, 2004. PP. 1787-1792.
- 27.Red'ko V.G., Anokhin K.V., Burtsev M.S., Manolov A.I., Mosalov O.P., Nepomnyashchikh V.A., Prokhorov D.V. Project “Animat Brain”: designing the animat control system on the basis of the functional systems theory // In: 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. PP. 94-107.
- 28.Анохин П.К. Системные механизмы высшей нервной деятельности. М.: Наука, 1979.
- 29.Редько В.Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. М.: Изд-во УРСС (Серия «Синергетика: от прошлого к будущему»), 2005.