

Редько В.Г. Актуальность моделирования когнитивной эволюции // "Научная сессия НИЯУ МИФИ-2010. Материалы избранных научных трудов по теме "Актуальные вопросы нейробиологии, нейроинформатики и когнитивных исследований". М.: НИЯУ МИФИ, 2010. С. 69-90.

В.Г. РЕДЬКО

Научно-исследовательский институт системных исследований РАН, Москва
vcredko@gmail.com

АКТУАЛЬНОСТЬ МОДЕЛИРОВАНИЯ КОГНИТИВНОЙ ЭВОЛЮЦИИ*

Аннотация. Как моделировать когнитивную эволюцию – эволюцию когнитивных свойств биологических организмов, приведшую к способностям научного познания? Как такое моделирование связано с теорией познания, с основаниями науки, с основаниями математики? Почему такое моделирование актуально? Каков задел исследований когнитивной эволюции? В настоящей статье обсуждаются эти вопросы. Анализируются подходы к математическому и компьютерному моделированию когнитивной эволюции. Очерчен задел по моделям когнитивной эволюции, развитый в направлении исследований «Адаптивное поведение», в котором изучаются модельные «организмы» (аниматы), способные приспособливаться к внешней среде. Характеризуются типичные модели адаптивного поведения; излагается оригинальный проект «Мозг анимата». Предлагаются контуры программы будущих исследований когнитивной эволюции.

1. Ключевая 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

* Работа выполнена при финансовой поддержке РФФИ (проект № 07-01-00180) и Президиума РАН (Программа «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация», проект № 2-15).

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...» «Прологомены¹ ко всякой будущей метафизике, могущей появиться, как наука» [2]. 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 охватить. 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:

¹ Термин «пролегомены» означает предварительные рассуждения, введение в изучение.

«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.

По существу, 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 $\{A, A \rightarrow B\} \Rightarrow B$.

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» {УС, УС --> БС} => БС. И собака ожидает БС. Конечно, применение правила *modus ponens* (чисто дедуктивное) математиком и индуктивный «вывод», который делает собака, явно различаются. Но можем ли мы думать об эволюционных корнях логических правил, используемых в математике? Да, вполне можем – умозаключение математика и индуктивный «вывод» собаки аналогичны.

Следовательно, мы можем анализировать эволюционные корни логического мышления и, более того, 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, подобные *modus ponens*.

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»

Направление «Адаптивное поведение» развивается с начала 1990-х годов [5,6]. Основной подход направления – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы называются «аниматами» (от англ. *animal* и *robot*: *animal* + *robot* = *animat*). 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. 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 адаптироваться к окружающей среде [7]. 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 [8]. Хороший обзор ранних работ по адаптивному поведению представлен в книге М.Г. Гаазе-Рапопорта, Д.А. Поспелова «От амебы до робота: модели поведения» [9].

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 [10],
- генетический алгоритм [11] и другие методы эволюционной оптимизации,
- классифицирующие системы (Classifier Systems) [12],
- обучение с подкреплением (Reinforcement Learning) [13].

Рис. 1 иллюстрирует междисциплинарные связи направления исследований «Адаптивное поведение». Как аргументировано выше, это направление связано с теорией познания, с исследованиями когнитивной эволюции. Моделирование адаптивного поведения связано с исследованиями естественного интеллекта – интеллект необходим для организации поведения [8,14]. Приложения моделей адаптивного поведения – искусственный интеллект, робототехника, модели адаптивного поведения в социально-экономических системах [15-17].

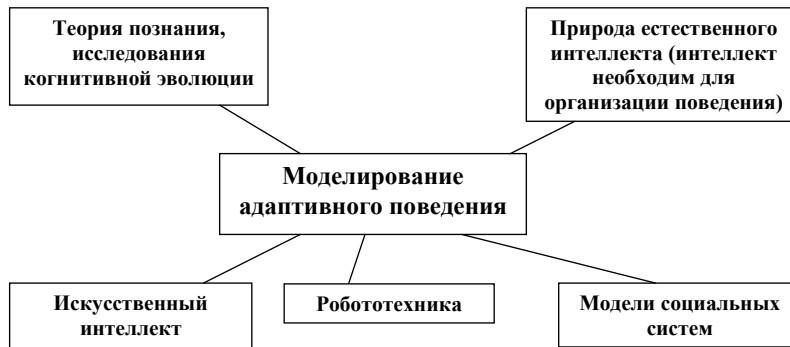


Рис. 1. Междисциплинарные связи направления исследований «Адаптивное поведение».

«Адаптивное поведение» – хорошо сформировавшееся направление исследований. Есть международное общество исследователей, работающих в этом направлении (сайт общества: <http://www.isab.org.uk>), издается журнал Adaptive Behavior, раз в два года проводятся международные конференции по моделированию адаптивного поведения.

Характерные примеры моделей адаптивного поведения кратко излагаются в следующем разделе.

3. Примеры моделей адаптивного поведения

3.1. Модели мозга и поведения в Институте нейронаук Дж. Эдельмана

В Институте нейронаук (сайт института: <http://www.nsi.edu>) уже более 25 лет ведутся разработки поколений моделей работы мозга (Darwin I, Darwin II, ...) и исследования адаптивного поведения искусственного организма NOMAD (Neurally Organized Mobile Adaptive Device), система управления которого построена на базе этих моделей.

Работы по NOMAD'у – исследование поведения адаптивного устройства, использующего модели мозга (авторы называют его также Brain-based device), принципы моделирования которого состоят в следующем:

- 1) устройство помещается в реальную физическую среду,
- 2) имеется некоторая поведенческая задача, которую должно решать устройство,

- 3) поведение устройства контролируется модельной нервной системой, которая отражает архитектуру мозга и динамику процессов в мозге,
- 4) поведение устройства и процессы в модельной нервной системе должны допускать сравнение с экспериментальными биологическими данными.

В одной из последних работ по NOMAD'у [18] промоделировано поведение мыши в лабиринте Морриса.

Исследования поведения мыши или крысы в лабиринте Морриса – один из канонических биологических экспериментов, который состоит в следующем. Имеется бассейн с непрозрачной жидкостью (например, это может быть вода, подкрашенная молоком), на бортах бассейна есть рисунки, которые мышь видит и может использовать для ориентировки. В определенном месте бассейна есть скрытая платформа, которую мышь может найти и тем самым спастись – не утонуть. Мышь бросают в бассейн, она плавает некоторое время и либо находит платформу и спасается, либо начинает тонуть (тогда ее спасает экспериментатор). После ряда экспериментов мышь начинает использовать ориентиры на бортах бассейна и находить платформу за достаточно короткое время.

Поведение NOMAD'а в лабиринте Морриса моделировалась следующим образом [18]. NOMAD представлял собой подвижное устройство на колесах, управляемое нейронной сетью, состоящей из 90000 нейронов, в которой было выделено 50 различных нейронных областей, в частности, были выделены несколько областей гиппокампа. В сети было $1.4 \cdot 10^6$ синаптических контактов (синапсов) между нейронами. Программно нейронная сеть была реализована на основе компьютерного кластера. При моделировании детально исследовались процессы, происходящие в разных нейронных областях.

Сенсорная система NOMAD'а включала зрение, обонятельную систему, позволяющую отслеживать свои собственные следы, систему инфракрасных приемников-излучателей, обеспечивающую избегание столкновений, и специальный детектор скрытой от зрения платформы, позволяющий обнаруживать эту платформу только тогда, когда NOMAD находился непосредственно над ней.

NOMAD помещался в комнату, в которой была скрытая платформа; на стенах комнаты были разноцветные полосы – ориентиры. В начале каждого из компьютерных экспериментов NOMAD помещался в разные участки комнаты, задача NOMAD'а была найти скрытую платформу. Обучение нейронных сетей NOMAD'а осуществлялось по модифицированному правилу Хебба (увеличение или уменьшение веса синаптической связи между

активными нейронами) на основе подкреплений (получаемых при нахождении скрытой платформы) и наказаний (получаемых при приближении к стенам комнаты).

Было продемонстрировано, что

- 1) NOMAD обучается находить платформу достаточно быстро (за 10-20 попыток);
- 2) в модельном гиппокампе формируются *нейроны места*, активные только тогда, когда NOMAD находится в определенных участках комнаты;
- 3) в модельном гиппокампе формируются связи между отдельными нейронными областями, отражающие причинно-следственные зависимости.

Итак, изложенная модель представляет собой интересное эмпирическое компьютерное исследование самообучающегося адаптивного устройства, хорошо продуманное с биологической точки зрения.

3.2. Модель эволюционного возникновения коммуникаций в коллективе роботов

В работе [19] исследовались вопросы: Как могут эволюционно возникнуть коммуникации между модельными организмами? Как в эволюционном процессе может сформироваться сигнальная обработка информации?

Рассматривалась следующая проблема. Есть четыре робота, каждый из которых управляется рекуррентной нейронной сетью (т.е. нейронной сетью с обратными связями), состоящей из 5-ти нейронов. На входы нейронов поступают сигналы от 8-ми инфракрасных датчиков и от 4-х датчиков, воспринимающих звуковые сигналы с разных сторон. Нейронная сеть имела 3 выходных нейрона, два из которых определяли скорость движения двух колес робота, а третий нейрон – интенсивность силы звука, издаваемого в данный момент роботом. В ограниченной области пространства находились две кормушки, и роботам нужно было, используя свои нейронные сети и звуковые сигналы разной интенсивности, как можно быстрее распределиться по кормушкам: по 2 робота на каждую из кормушек.

Нейронные сети роботов оптимизировались эволюционным путем (при этом настраивались веса синапсов нейронов). В результате в течение 2000 поколений у роботов сформировались сигналы 5 различных видов (т.е. разной интенсивности). Используя эти сигналы, роботы достаточно устойчиво находили требуемое распределение по кормушкам.

В целом в работе [19] продемонстрировано, что в эволюционирующей популяции роботов, управляемых рекуррентными нейронными сетями, может формироваться простая сигнальная система, а именно, система коммуникаций, позволяющая решать нетривиальную задачу распределения роботов по кормушкам.

3.3. Бионическая модель поискового адаптивного поведения

Одно из актуальных направлений исследований в рамках моделирования адаптивного поведения – имитация поискового поведения животных. В работе [20] исследовано поисковое поведение на примере личинок ручейников *Chaetopteryx villosa*, обитающих на дне водоемов. Личинки носят на себе «домик» – трубку из песка и других частиц, которые они собирают на дне водоемов. Частицы скрепляются между собой по краям с помощью клейкой белковой нити. Строительство домика требует меньше времени, усилий и белка, если личинки используют относительно крупные и плоские частицы. Однако поиск крупных частиц на дне водоема требует затрат времени и энергии, не известных личинке заранее. Задача осложняется еще и тем, что личинки при поиске частиц не пользуются зрением и могут обнаружить частицу и определить её размер только на ощупь, что требует дополнительных затрат времени.

В [20] построена компьютерная модель поискового поведения личинок ручейников, строящих чехол-домик из частиц разного размера и ведущих поиск скоплений подходящих частиц. Модель использует понятие мотивации, а именно мотивации к прикреплению частиц к домику. Динамика регулирующей поведение мотивации $M(t)$ описывается уравнением:

$$M(t) = k_1 M(t-1) + \xi(t) + I(t), \quad (1)$$

время t дискретно, k_1 – параметр, характеризующий медленную релаксацию мотивации ($0 < k_1 < 1$, $1 - k_1 \ll 1$), $\xi(t)$ – случайные вариации мотивации, величина $I(t)$ характеризует направленное изменение мотивации при тестировании модельной личинкой частиц:

$$I(t) = k_2 (S_{curr} - S_{last}) / S_{last}, \quad (2)$$

где k_2 – положительный параметр, S_{curr} – площадь тестируемой в данный момент частицы, S_{last} – площадь последней протестированной ранее частицы. Если мотивация $M(t)$ достаточно велика, то происходит сбор и при-

крепление частиц к домику, если $M(t)$ мала, то модельная личинка ищет новое место с подходящими размерами частиц.

Построенная модель качественно согласуется с биологическими экспериментальными данными: как в эксперименте, так и в модели к домику преимущественно прикрепляются крупные частицы, есть сильный разброс числа прикрепляемых частиц и момента начала прикрепления, число прикрепляемых частиц в обоих случаях невелико.

Отметим, что согласно (1), (2) динамика мотивации проста и эффективна. Она учитывает инерцию изменения, случайные вариации и направленное изменение $M(t)$. И разработанный подход может быть использован при моделировании регулирования переключений между тактиками поведения в случае нескольких потребностей и целей живого организма или искусственного анимата, например, мобильного робота.

Аналогичная динамика $M(t)$ может быть использована и в других подобных задачах. Например, ее можно использовать при поиске экстремума функции нескольких переменных. Такое моделирование было проведено и показана возможность реализации аналога известного «овражного» метода [21] оптимизации. Суть реализованного нами метода состоит в том, что вводится мотивация к сохранению направления поискового движения, при большой величине $M(t)$ направление движения сохраняется, при малых $M(t)$ направление поиска случайно варьируется. Если минимизируемая функция имеет достаточно глубокий «овраг», в котором она слабо меняется, то сначала в данном методе быстро находится сам овраг, а затем происходит постепенная минимизация функции при движении вдоль оврага.

3.4. Модель эволюции популяции самообучающихся агентов

В данном разделе излагается модель эволюции популяции самообучающихся автономных агентов [22]. Модель отработывалась на примере агента-брокера. Этот пример использовался только для определенности, аналогично можно было бы анализировать функционирование модельного «организма», более подобного биологическим прототипам, например, «организма», помещенного во внешнюю среду, которая определяется зависимостью температуры от времени, аналогичной курсу акций для агента-брокера. В рассматриваемой модели задача агента была достаточно проста: надо было предсказывать курс акций, и при прогнозе роста курса вкладывать весь имеющийся у агента капитал в акции, при падении курса вкладывать капитал в деньги.

При моделировании особое внимание уделялось взаимоотношению между обучением агентов и эволюционной оптимизацией их систем управления.

Описание модели. При моделировании считалось, что время дискретно. Каждый такт времени агент выбирал одно из двух действий: 1) вложить весь свой капитал в акции, либо 2) вложить весь капитал в деньги. Система управления агента представляла собой простой нейросетевой адаптивный критик [23], состоящий из двух нейронных сетей (НС): Модель и Критик. НС Модель предназначена для прогнозирования изменения курса временного ряда. НС Критик предназначена для оценки ожидаемого прироста капитала в течение достаточно длительного времени в будущем для каждого из действий. Самообучение агента происходило методом обучения с подкреплением [13], величина подкрепления в определенный такт времени равна приросту капитала агента в течение этого такта. При этом для формирования оценок ожидаемого прироста капитала нейронной сетью Критика используется стохастический поиск: изредка (с малой вероятностью ϵ) действия агента выбираются случайно.

В процессе обучения веса синапсов НС подстраивались градиентным методом с целью минимизации ошибки предсказания и уточнения оценок ожидаемого прироста капитала.

Кроме обучения осуществлялась эволюционная оптимизация весов синапсов НС. Рассматривалась дарвиновская эволюция популяции n агентов. Продолжительность каждого поколения равна T тактов времени. Эволюция происходит в течение ряда поколений. Начальные веса синапсов обеих НС (Модели и Критика) формируют геном агента \mathbf{G} . Геном \mathbf{G} задается в момент рождения агента (в начале поколения) и не меняется в течение его жизни. В противоположность этому текущие веса синапсов НС \mathbf{W} подстраиваются в течение жизни агента путем обучения.

В конце каждого поколения определяется агент, прирост капитала которого в течение поколения максимален. Этот лучший агент порождает n потомков, которые составляют новое поколение. Геномы потомков \mathbf{G} отличаются от генома родителя небольшими мутациями.

Таким образом, геном \mathbf{G} (начальные веса синапсов, получаемые при рождении агента) изменяется только посредством эволюции, в то время как текущие веса синапсов \mathbf{W} дополнительно к этому подстраиваются посредством обучения. При этом в момент рождения агента $\mathbf{W} = \mathbf{G}$.

Так как передаются по наследству начальные веса синапсов, получаемые при рождении, то эволюционный процесс именно дарвиновский. Ес-

ли бы по наследству передавались веса синапсов, формируемые у лучшего агента в конце поколения, получаемые с помощью обучения, тогда эволюция была бы ламарковской.

Результаты моделирования. Изложенная модель была реализована в виде компьютерной программы. В компьютерных экспериментах использовалось два варианта временного ряда, определяющего курс акций: 1) синусоида, 2) стохастический сглаженный временной ряд. Изучалась роль эволюции и обучения при формировании весов синапсов нейронных сетей агентов. Нейронные сети были достаточно простыми: они содержали один скрытый слой нейронов, число нейронов было порядка 20. Были проанализированы следующие варианты рассматриваемой модели:

- Случай L (чистое обучение); в этом случае рассматривался отдельный агент, который только обучался;
- Случай E (чистая эволюция), т.е. рассматривалась эволюционирующая популяция без обучения;
- Случай LE (обучение + эволюция), т.е. полная модель, изложенная выше.

Было проведено сравнение капитала (ресурса), приобретаемого агентами за 200 временных тактов для этих трех способов адаптации. Для случаев E и LE бралось $T = 200$ (T – продолжительность поколения) и регистрировалось максимальное значение ресурса в популяции $R_{max}(n_g)$ в конце каждого n_g -го поколения. В случае L (чистое обучение) рассматривался только один агент, ресурс которого $R(t)$ для удобства сравнения со случаями E и LE обнулялся каждые $T = 200$ тактов времени. В этом случае индекс n_g увеличивался на единицу после каждых T временных тактов, и полагалось $R_{max}(n_g) = R(T n_g)$.

Графики $R_{max}(n_g)$ для синусоидального изменения курса акций показаны на рис. 2. Чтобы исключить уменьшение значения $R_{max}(n_g)$ из-за случайного выбора действий для случаев LE и L, стохастический выбор действия прекращался после $n_g = 100$ для случая LE и после $n_g = 2000$ для случая L. Результаты усреднены по 1000 экспериментам.

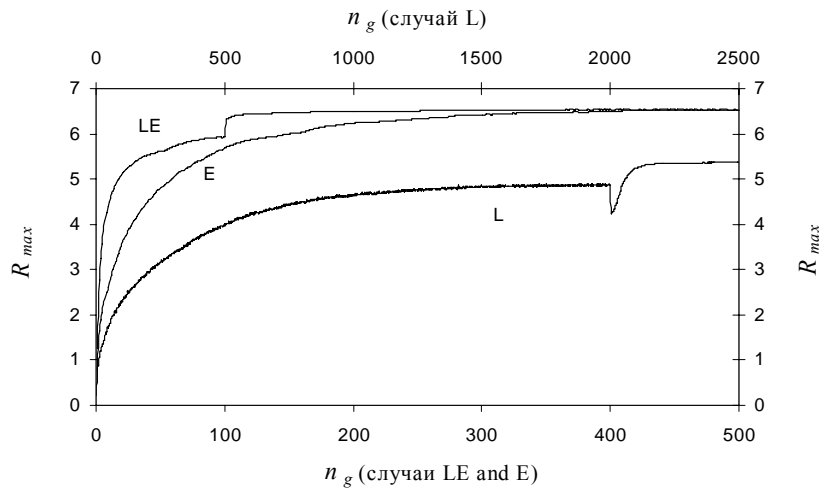


Рис. 2. Зависимости $R_{max}(n_g)$. Кривая LE соответствует случаю обучения, объединенного с эволюцией, кривая E – случаю чистой эволюции, кривая L – случаю чистого обучения. Временная шкала для случаев LE и E (номер поколения n_g) представлена снизу, для случая L (индекс n_g) – сверху. Моделирование проведено для синусоидального изменения курса акций, кривые усреднены по 1000 экспериментам; $n = 10$, $T = 200$.

Рис. 2 показывает, что обучение, объединенное с эволюцией (случай LE), и чистая эволюция (случай E) дают одно и то же значение конечного ресурса $R_{max}(500) = 6.5$. Однако эволюция и обучение вместе обеспечивают нахождение больших значений R_{max} быстрее, чем эволюция отдельно – существует симбиотическое взаимодействие между обучением и эволюцией.

Анализ компьютерных экспериментов, представленных на рис. 2, показал, что в случаях LE (обучение + эволюция) и E (чистая эволюция) оптимальная стратегия агента – вкладывать весь капитал в акции при прогнозе роста курса, вкладывать весь капитал в деньги при прогнозе падения курса – находится. Это соответствует асимптотическому значению ресурса $R_{max}(500) = 6.5$.

В случае L (чистое обучение) асимптотическое значение ресурса ($R_{max}(2500) = 5.4$) существенно меньше. Анализ экспериментов для этого случая показал, что одно обучение обеспечивает нахождение только «субоптимальной» стратегии поведения: агент держит капитал в акциях при

росте и при слабом падении курса и переводит капитал в деньги при сильном падении курса, т.е. агент явно предпочитает вкладывать капитал в акции.

Представленные результаты демонстрируют, что хотя обучение в настоящей модели и несовершенно, оно способствует более быстрому нахождению оптимальной стратегии поведения по сравнению со случаем чистой эволюции.

Тот факт, что эволюционная оптимизация может быть эффективней самообучения, обусловлен тем, что в используемом варианте метода обучения с подкреплением оценка ожидаемого прироста капитала нейронной сетью Критика формируется с использованием стохастического поиска. И это формирование оценки путем самообучения только на основе получаемых подкреплений и наказаний с использованием стохастического поиска довольно непросто.

Если длительность поколения T была достаточно большой (1000 и более тактов времени), то для случая LE часто наблюдалось сильное влияние обучения на эволюционный процесс (рис. 3). В первых поколениях эволюционного процесса существенный рост ресурса агентов наблюдался не с самого начала поколения, а спустя 200-300 тактов, т.е. агенты явно обучались в течение своей жизни находить более или менее приемлемую стратегию поведения, и только после смены ряда поколений рост ресурса начинался с самого начала поколения. Это можно интерпретировать как проявление известного эффекта Болдуина: исходно приобретаемый навык в течение ряда поколений становится наследуемым [24, 25]. Этот пример показывает, что хотя эволюционный процесс и дарвиновский, тем не менее, в рамках эволюции наблюдаются особенности, характерные для ламарковской эволюции, – приобретенные навыки переходят в наследуемые в течение ряда поколений. Причем в наших компьютерных экспериментах это происходило быстро, всего за 5 поколений.

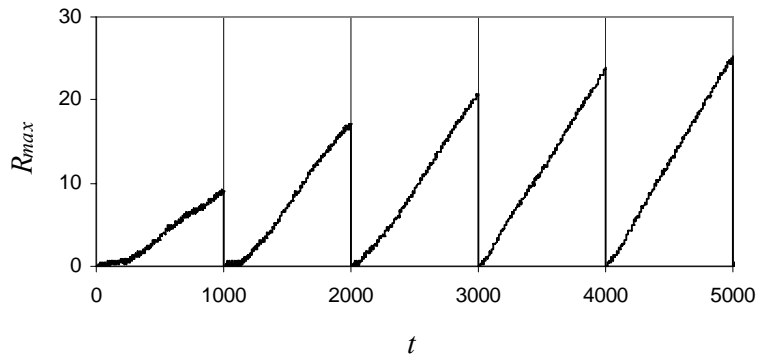


Рис. 3. Зависимость ресурса лучшего в популяции агента R_{max} от времени t для первых пяти поколений. Случай LE (обучение, объединенное с эволюцией); размер популяции $n = 10$, длительность поколения $T = 1000$. Моделирование проведено для синусоидального изменения курса акций. Моменты смены поколений показаны вертикальными линиями. Для первых двух поколений есть явная задержка в 200-300 тактов времени в росте ресурса агента. К пятому поколению лучший агент «знает» хорошую стратегию поведения с самого рождения, т.е. стратегия, изначально приобретаемая посредством обучения, становится наследуемой.

Итак, исследована модель эволюции популяции самообучающихся автономных агентов и проанализирована роль обучения и эволюционной оптимизации на примере простых самообучающихся агентов-брокеров. Показано, что для таких агентов эволюционная оптимизация может быть более эффективна, чем обучение с подкреплением. Продемонстрирован эффект Болдуина: приобретаемые навыки становятся наследуемыми в течение нескольких поколений дарвиновской эволюции.

Возвратимся к общему обсуждению исследований адаптивного поведения. Анализ работ в рамках анимат-подхода показывает, что хотя уже разработано множество моделей, эти модели пока еще во многом фрагментарны. Поэтому в наших работах предложен проект «Мозг анимата», который нацелен на формирование общей схемы построения таких моделей (термин «Мозг анимата» был предложен К.В. Анохиным) [26-28]. Проект основан на теории функциональных систем П.К. Анохина [29].

4. Проект «Мозг анимата»

В проекте предполагается, что система управления аниматом имеет иерархическую архитектуру, базовым элементом которой является отдельная функциональная система (ФС). Верхний уровень соответствует основным потребностям организма: 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» [27] предполагалась простая формализация ФС на основе нейросетевых адаптивных критиков, которые обучаются с помощью метода обучения с подкреплением. Пример модели, соответствующей этой схеме ФС и показывающей, как работают такие ФС, изложен выше (раздел 3.4). Для этого примера были обнаружены интересные и нетривиальные результаты, однако в данной модели, как показывают компьютерные расчеты, есть и слабые свойства: нейронная сеть Модель автономных агентов в определенных условиях может делать неправильные предсказания. А именно, хотя по форме кривой изменения курса акций предсказываются правильно, предсказанные изменения могут сильно отличаться от реальных по величине. И эти, строго говоря, неверные предсказания могут вполне разумно использоваться. Это показывает, что необходима определенная осторожность в выборе базовой модели функциональной системы для проекта «Мозг анимата». Поэтому имеет смысл рассмотреть и другие возможности. С этой целью в [28] начата разработка нейросетевой системы «Мозга анимата» на основе использования понятия хеббовских ансамблей [30], в которой, как и выше, предполагается, что есть эволюционирующая популяция аниматов, оптимизация систем управления которых происходит как в процессе обучения, так и в процессе эволюции популяции. Этот вариант интересен тем, что в нем предусмотрена простая возможность изменения архитектуры системы управления аниматом: связи между функциональными системами и сами системы могут меняться в процессах обучения и эволюции. Это фактически соответствует принципам метаобучения (обучения как в отдельных моделях, блоках, так и обучения путем подбора моделей или блоков), предложенным в вычислительном интеллекте [31]. Тем не менее, необходимо согласование функционирования всех элементов системы управления аниматом и это накладывает довольно серьезные ограничения на ее архитектуру. Следовательно, важны

дальнейшие исследования по данному проекту, в том числе и на основе представленного выше опыта использования нейросетевых адаптивных критиков.

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

5. Модели адаптивного поведения и задача исследования когнитивной эволюции

В рамках направления «Адаптивное поведение» уже ведется определенное моделирование когнитивных способностей животных. Отметим некоторые из таких исследований.

В ряде ранних работ исследовались модели условных рефлексов. Здесь мы приведем только ссылки на некоторые из моделей [9,32,33].

В последнее время активно развиваются исследования антисипаторного поведения, при котором животное предвидит будущие ситуации и использует это предвидение при организации поведения [34].

Интересные работы связаны с методами формализации правил принятия решения. Например, в [35] 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 компьютерное моделирование, подтвердившее эффективность предложенных правил.

Некоторые работы анализируют эволюционные аспекты становления нейронных структур мозга животных, обеспечивающих познавательные процессы. Например, в [36] проанализирована эволюция нейронных структур мозга, играющих важную роль при выборе действий, обеспечивающих адаптивное поведение.

Отметим работы и с другой стороны – со стороны исследования логических процессов на стыке с «логикой» поведения животных. В статье В.Ф. Турчина (1987 г.) [37] предпринята весьма нетривиальная попытка пересмотра оснований математики и рассмотрена возможность построения предиктивных логических процессов в контексте теории множеств и кибернетического подхода к обоснованию математики. В недавних работах Е.Е. Витяева и А.В. Демина [38-40] 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. Отметим, что, развивая подход работ [38-40], 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).

Возвратимся к работам по анализу когнитивных процессов при формировании адаптивного поведения. В качестве иллюстрации этих работ отметим модель [41], которую можно рассматривать как попытку изучения того, как животное в своем поведении делает индуктивный «вывод», формируя и используя определенные логические правила. В [41] моделировалась система управления поведением животного на основе использования одновременно метода классифицирующих систем [12] и метода обучения с подкреплением [13]. А именно, формировалась популяция «логических» правил вида: «если имеет место ситуация $S(t)$, то нужно выполнить действие $A(t)$, после этого действия возникнет следующая ситуация $S(t+1)$ », t – время. Эта популяция правил оптимизировалась 1) путем обучения (посредством изменения вероятности применения того или иного правила и обобщения правил) и 2) путем эволюции (отбора и мутаций правил в популяции). При этом методом компьютерного моделирования анализировался биологический эксперимент [42], в котором крыс обучали следующему. Если в ситуации S_1 крыса нажмет педаль, т.е. выполнит действие A_1 , то она получит пищу одного вида O_1 , если в этой же ситуации S_1 , крыса потянет за цепочку, т.е. выполнит действие A_2 , то она получит пищу другого вида O_2 . В ситуации S_2 вид пищи менялся. Если в этой ситуации S_2 крыса нажмет педаль (действие A_1), то она получит пищу O_2 , если потянет за цепочку (действие A_2), то она получит пищу O_1 . Схематично имеем:

$$\begin{aligned} S_1 : A_1 \rightarrow O_1, A_2 \rightarrow O_2, \\ S_2 : A_1 \rightarrow O_2, A_2 \rightarrow O_1, \end{aligned}$$

Проводили достаточно длительное обучение. После обучения пищу O_1 сопровождали наказанием (добавляли к этой пище «невкусное» вещество). В результате в ситуации S_1 крысы предпочитали действие A_2 , а в ситуации S_2 – действие A_1 , и, тем самым, избегали пищу O_1 . Следовательно,

крысы должны были сделать определенный «логический» вывод, связывающий ранее полученное обучение с ожидаемым результатом, и использовать результат вывода при принятии решения о выполняемом действии. Модель [41] воспроизводила основные особенности биологического эксперимента [42].

Подчеркнем, что компьютерная модель [41] качественно описывает определенную «логику» поведения животных, хотя, конечно, это только одна из попыток моделирования познавательных процессов, используемых животными в поведении.

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.

6. Контуры программы будущих исследований когнитивной эволюции

Таким образом, анализ современных исследований адаптивного поведения показывает, что хотя проделана большая работа и есть много интересных моделей, ученые еще далеки от понимания того, как возникали и развивались системы управления живых организмов, как развитие этих систем способствовало эволюции когнитивных способностей животных, и как процесс когнитивной эволюции привел к возникновению логического мышления. Образно говоря, у нас уже есть некоторые небольшие фрагменты картины, но мы еще не видим всей картины. Предложим контуры программы будущих исследований, нацеленных на моделирование когнитивной эволюции.

А) Моделирование адаптивного поведения аниматов с несколькими естественными потребностями: питания, размножения, безопасности. Фактически, уже начато моделирование в этом направлении [22, 28, 43]; в частности, моделирование на базе проекта «Мозг анимата» предназначено для исследования поведения простых аниматов с естественными потребностями.

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 --> 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 (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. Работы в этих направлениях уже ведутся, но четкой последовательности серьезных, канонических моделей еще нет.

7. Еще раз об актуальности моделирования когнитивной эволюции

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.

8. Заключение

Итак, предложен подход к моделированию когнитивной эволюции. Указана связь этого направления исследований с основаниями математики. Очерчен задел по моделям когнитивной эволюции, развитый в направлении исследований «Адаптивное поведение». Намечены контуры программы будущих исследований когнитивной эволюции.

Список литературы

1. Кант И. Критика чистого разума. Соч. в 6-ти томах. Т.3. М.: Мысль, 1964. С. 69-695.
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.
3. Lorenz K. Kant's doctrine of the a priori in the light of contemporary biology (1941) // In: H. Plotkin (Ed.). Learning, Development and Culture: Essays in Evo-

- lutionary Epistemology. New York: Wiley, 1982. PP. 121-143. См. также перевод: 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.
4. Математическая теория логического вывода (под ред. А.В. Идельсона и Г.Е. Минца). М.: Наука, 1967.
 5. 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.
 6. От моделей поведения к искусственному интеллекту (под ред. В.Г. Редько). М.: Изд-во УРСС (Серия «Науки об искусственном»), 2006.
 7. Цетлин М.Л. Исследования по теории автоматов и моделирование биологических систем. М.: Наука, 1969.
 8. Бонгард М.М., Лосев И.С., Смирнов М.С. Проект модели организации поведения – «Животное» // Моделирование обучения и поведения. М.: Наука, 1975. С. 152-171.
 9. Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота: модели поведения. М.: Наука, 1987. М.: УРСС, 2004.
 10. Хайкин С. Нейронные сети: полный курс, 2-е издание. М.: Издательский дом «Вильямс», 2006.
 11. 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).
 12. Holland J.H., Holyoak K.J., Nisbett R.E., Thagard P. Induction: Processes of Inference, Learning, and Discovery. Cambridge: MIT Press, 1986.
 13. Sutton R., Barto A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. See also: <http://www.cs.ualberta.ca/~sutton/book/the-book.html>
 14. Вайнцвайг М.Н., Полякова М.П. О моделировании мышления // От моделей поведения к искусственному интеллекту. Серия «Науки об искусственном» (под ред. Редько В.Г.). М.: УРСС, 2006. С. 280-286.
 15. Nolfi S., Baldassarre G., Calabretta R., Hallam J., Marocco D., Miglino O., Meyer J.-A., Parisi D. (Eds). From animals to animats 9: Proceedings of the Ninth International Conference on Simulation of Adaptive Behavior. LNAI. Volume 4095. Berlin, Germany: Springer Verlag, 2006.
 16. Wilson S.W. The animat path to AI // In: 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. PP. 15-21.
 17. Pfeifer R., Scheier C. Understanding Intelligence. London: MIT Press, 1999.
 18. 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. PP. 197-221.
 19. Marocco D., Nolfi S. Origins of communication in evolving robots // In: Nolfi S., Baldassarre G., Calabretta R., Hallam J., Marocco D., Miglino O., Meyer J.-A., Parisi D. (Eds.). From Animals to Animats 9: Proceedings of the Ninth International

- Conference on Simulation of Adaptive Behavior. LNAI. Volume 4095. Berlin, Germany: Springer Verlag, 2006. PP. 789-803.
20. Непомнящих В.А., Попов Е.Е., Редько В.Г. Бионическая модель адаптивного поискового поведения // Изв. РАН. Теория и системы управления, 2008. № 1. С. 85-93.
 21. Гельфанд И.М., Цетлин М.Л. Принцип нелокального поиска в задачах автоматической оптимизации // ДАН СССР, 1961. Т. 137. № 2. С. 295-298.
 22. Red'ko V.G., Mosalov O.P., Prokhorov D.V. A model of evolution and learning // Neural Networks, 2005. V. 18. No. 5-6. PP. 738-745.
 23. Редько В.Г., Прохоров Д.В. Нейросетевые адаптивные критики // Научная сессия МИФИ – 2004. VI Всероссийская научно-техническая конференция "Нейроинформатика-2004". Сборник научных трудов. М.: МИФИ, 2004. Часть 2. С. 77-84.
 24. Baldwin J.M. A new factor in evolution // American Naturalist, 1896. V. 30. PP. 441-451.
 25. Turney P., Whitley D., Anderson R. (Eds.). Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect // Special Issue of Evolutionary Computation on the Baldwin Effect. V. 4. No. 3. 1996.
 26. 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.
 27. 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.
 28. 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.
 29. Анохин П.К. Системные механизмы высшей нервной деятельности. М.: Наука, 1979.
 30. Hebb D.O. The Organization of Behavior. A Neuropsychological Theory. New York: Wiley and Sons, 1949.
 31. Duch W. Towards comprehensive foundations of computational intelligence // In: Duch W., Mandziuk J. (Eds.). Challenges for Computational Intelligence, 2008. Berlin, Heidelberg: Springer Verlag. PP. 261-316.
 32. Grossberg S. Classical and instrumental learning by neural networks // Progress in Theoretical Biology. 1974. V.3. PP. 51-141.
 33. 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.

34. 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.
35. Witkowski M. An action-selection calculus // *Adaptive Behavior*, 2007. V. 15. No. 1. PP. 73-97.
36. 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.
37. Turchin V.F. A constructive interpretation of the full set theory // *Journal of Symbolic Logic*, 1987. V. 52. No. 1. PP. 172 -201.
38. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. Новосибирск: НГУ, 2006.
39. 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>
40. 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>
41. Butz M.V., Hoffmann J. Anticipations control behavior: animal behavior in an anticipatory learning classifier system // *Adaptive Behavior*, 2002. V. 10. No. 2. PP. 75-96.
42. Colwill R.M., Rescorla R.A. Evidence for the hierarchical structure of instrumental learning // *Animal Learning and Behavior*, 1990. V. 18. No. 1. PP. 71–82.
43. Бурцев М.С., Гусарев Р.В., Редько В.Г. Исследование механизмов целенаправленного адаптивного управления // *Изв. РАН. Теория и системы управления*. 2002. № 6. С. 55-62.