

Подходы к моделированию происхождения мышления^{*}

Редько В.Г.

Научно-исследовательский институт системных исследований, Российская академия наук, Москва, 117218, Россия
vgredko@gmail.com

Аннотация. Анализируются подходы к моделированию происхождения мышления. Обсуждается новое направление исследований – моделирование когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, в результате которой произошли логика, мышление, интеллект человека. Характеризуются заделы исследований когнитивной эволюции и предлагаются контуры программы будущих исследований когнитивной эволюции. Аргументируется, что при изучении происхождения мышления важную роль должны играть модели автономных агентов, обладающих когнитивными свойствами. Анализируются попытки приблизиться к созданию интеллекта человеческого уровня. Обсуждаются методы, которые могут быть использованы при рассматриваемом моделировании. Рассматриваются биологические аспекты, которые должны быть важны при моделировании происхождения мышления, в частности, аспекты, связанные со способностями, которые можно характеризовать понятием «элементарное мышление животных». Также кратко характеризуются прикладные аспекты, связанные с рассматриваемым моделированием.

Ключевые слова: на пути к теории происхождения мышления, когнитивная эволюция, автономные агенты, методы моделирования, познавательные способности биологических организмов.

ВВЕДЕНИЕ

В статье анализируются подходы к моделированию происхождения мышления. В первой части статьи обсуждается новое интересное направление исследований – моделирование когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов. Аргументируется, что актуальность исследований когнитивной эволюции связана с глубокой гносеологической проблемой: почему формальное логическое человеческое мышление, казалось бы, совсем не связанное с реальным физическим миром, применимо к познанию природы? Показано, что имеется задел моделирования когнитивной эволюции, развиваемый в направлении исследований «Адаптивное поведение». Характеризуются начальные шаги моделирования, непосредственно направленные на анализ когнитивной эволюции. Обсуждаются перспективы моделирования когнитивной эволюции.

Вторая часть статьи посвящена анализу широкого круга вопросов, связанных с исследованиями, направленными на изучение происхождения мышления. Аргументируется, что при изучении происхождения мышления важную роль должны играть модели автономных агентов, обладающих когнитивными, интеллектуальными свойствами. Эти модели, как правило, основаны на биологических аналогиях. Обсуждается связь рассматриваемых исследований с попытками приблизиться (путем моделирования, а возможно, и создания) к интеллекту человеческого уровня. Кроме этого обсуждаются экспериментальные биологические работы по «элементарному

* Работа выполнена при поддержке РФФИ, проект № 10-01-00129.

мышлению животных». В конце второй части рассматриваются прикладные аспекты, связанные с анализируемыми подходами.

Статья построена следующим образом. В части 1 обсуждается постановка исследований по моделированию когнитивной эволюции. В разделе 1.1 обсуждается указанная выше гносеологическая проблема. В разделе 1.2 характеризуется направление исследований «Адаптивное поведение». Типичные примеры моделей адаптивного поведения кратко излагаются в разделе 1.3. В разделе 1.4 предлагаются контуры программы будущих исследований когнитивной эволюции. Начальные шаги моделирования когнитивной эволюции, соответствующие этапам предложенной программы, характеризуются в разделе 1.5. В разделе 1.6 обсуждаются перспективы моделирования когнитивной эволюции. В части 2 анализируются различные аспекты исследований, направленных на изучение происхождения мышления. В разделе 2.1 анализируются направления вычислительного интеллекта, связанные с моделированием происхождения мышления. В разделе 2.2 рассмотрены попытки продвинуться к интеллекту человеческого уровня. Биологические эксперименты по «элементарному мышлению животных» и связанное с ними моделирование кратко характеризуются в разделе 2.3. В разделе 2.4 обсуждаются некоторые прикладные направления, которые могут возникнуть на базе рассматриваемых исследований. В разделе 2.5 кратко говорится о рассматриваемых исследованиях в контексте развития научного миропонимания. Итоги статьи подводятся в заключении.

1. О МОДЕЛИРОВАНИИ КОГНИТИВНОЙ ЭВОЛЮЦИИ

1.1. Гносеологическая проблема

По-видимому, наиболее серьезные и глубокие когнитивные процессы – это процессы научного познания. Но насколько способен человек познавать внешний мир? Почему формальный логический вывод, сделанный *человеком*, применим к реальным объектам в *природе*? Поясним эти вопросы. Рассмотрим, например, физику, одну из фундаментальных естественнонаучных дисциплин. Мощь физики связана с эффективным применением математики. Но математик делает логические выводы, доказывает теоремы независимо от внешнего мира, используя свое мышление. Почему же эти выводы применимы к реальной природе? В более общей формулировке проблема может быть сформулирована так: почему логика человеческого мышления применима к познанию природы?

Близкие вопросы давно интересовали ученых. В 1781 году появилась знаменитая «Критика чистого разума» И. Канта [1], а два года спустя вышло популярное изложение «Критики...» «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 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.

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

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

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

По существу, И. Кант и К. Лоренц показали, что если не рассматривать эволюционное происхождение методов познания, то нет ответа на 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 эволюции познавательных способностей животных? Далее обсуждаются перечисленные вопросы. Но, прежде всего, кратко остановимся на вопросе: можно ли вообще в принципе провести 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

$$\frac{A, A \rightarrow B}{B} \quad (1)$$

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»

$$\frac{УС, УС \rightarrow БС}{БС} \quad (2)$$

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» на основе 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*.

Важно подчеркнуть, что были и попытки пересмотра оснований математики в близком к исследованиям когнитивной эволюции контексте. В статье В.Ф. Турчина (1987 г.) [5] предпринята весьма нетривиальная попытка пересмотра оснований математики и рассмотрена возможность построения предиктивных логических правил в контексте теории множеств и кибернетического подхода к обоснованию математики. В.Ф. Турчиным также была предложена концептуальная теория метасистемных переходов [6], которая может быть использована при моделировании когнитивной эволюции.

Как же конкретно вести моделирование когнитивной эволюции? Есть ли задел таких исследований? Оказывается, что да, такой задел в той или иной форме существует в нескольких современных направлениях вычислительного интеллекта. В частности, модели познавательных свойств живых организмов в близких к изучению когнитивной эволюции аспектах развиваются в рамках направления исследований «Адаптивное поведение».

1.2. Направление исследований «Адаптивное поведение»

Направление «Адаптивное поведение» развивается с начала 1990-х годов [7,8]. Основной подход направления – конструирование и исследование искусственных «организмов» (в виде компьютерной программы или робота), способных приспособиваться 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 [9]. Дальняя цель этих работ (пока еще нереализованная) – анализ эволюции когнитивных способностей животных и происхождения интеллекта человека [10]. Эта цель близка к задаче моделирования когнитивной эволюции.

Отметим, что хотя «официально» направление «Адаптивное поведение» было провозглашено в 1990 году, были явные провозвестники этого направления. Приведем примеры из отечественной науки. В 1960-х годах М.Л. Цетлин предложил и исследовал модели автоматов, способных адаптивно приспособиваться к окружающей среде [11]. В 1960-70-х годах под руководством М.М. Бонгарда был предложен интересный проект «Животное», характеризующий целенаправленное адаптивное поведение искусственных организмов [12]. Хороший обзор ранних работ по адаптивному поведению представлен в книге М.Г. Гаазе-Рапопорта, Д.А. Поспелова «От амебы до робота: модели поведения» [13].

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 [14],
- генетический алгоритм [15] и другие методы эволюционной оптимизации,
- классифицирующие системы (Classifier Systems) [16],
- обучение с подкреплением (Reinforcement Learning) [17].

Направление исследований «Адаптивное поведение» часто рассматривается как бионический подход к разработкам искусственного интеллекта. Приложениями исследований адаптивного поведения являются также робототехника, модели адаптивного поведения в социально-экономических системах.

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

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

Характерные примеры моделей адаптивного поведения кратко излагаются в следующем разделе.

1.3. Примеры моделей адаптивного поведения

1.3.1. Модели мозга и поведения в Институте нейронаук

В Институте нейронаук, руководимом лауреатом Нобелевской премии Дж. Эдельманом (сайт института: <http://www.nsi.edu>), более 25 лет ведутся разработки поколений моделей работы мозга (Darwin I, Darwin II,...). В последние годы идут 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 роботоподобного 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 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 поощрений (получаемых при нахождении скрытой платформы) и наказаний (получаемых при приближении к стенам комнаты).

Было продемонстрировано, что

- 1) NOMAD самостоятельно 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.

1.3.2. Бионическая модель поискового адаптивного поведения

Одно из актуальных направлений исследований в рамках моделирования адаптивного поведения – имитация поискового поведения животных. В работе [19] исследовано поисковое поведение на примере личинок ручейников *Chaetopteryx villosa*,

обитающих на дне водоемов. Личинки носят на себе «домик» – трубку из песчинок и других частиц, собираемых со дна 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 [19] 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:

$$M(t) = k_1 M(t-1) + \xi(t) + I(t), \quad (3)$$

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:

$$I(t) = k_2 (S_{curr} - S_{last}) / S_{last}, \quad (4)$$

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

Важно, что в модели использовались знания личинки о размерах последних протестированных и прикрепленных 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 (3), (4) 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 [20] оптимизации. Суть реализованного нами метода [21] 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.

Далее отметим некоторые работы по адаптивному поведению, непосредственно связанные с моделированием когнитивной эволюции.

1.3.3. Модели адаптивного поведения агентов с когнитивными свойствами

В ряде ранних работ исследовались модели условных рефлексов. Здесь мы приведем только ссылки на некоторые из моделей [22, 23].

В последнее время активно развиваются исследования антисипаторного поведения, при котором животные предвидят будущие ситуации и используют это предвидение при организации адаптивного поведения, см. например [24].

Интересные работы связаны с методами формализации правил принятия решения. Например, в [25] 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 [26] проанализирована эволюция нейронных структур мозга, играющих важную роль при выборе действий, обеспечивающих адаптивное поведение.

В работах Е.Е. Витяева и А.В. Демина [27-29] 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 функционировать быстрее обычного метода обучения с подкреплением [17]. Отметим, что, развивая подход работ [27-29], 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 разных эволюционных уровней 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.

1.4. Контуры программы будущих исследований когнитивной эволюции

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

А. Моделирование адаптивного поведения аниматов с несколькими естественными потребностями: питания, размножения, безопасности. Это могло бы быть моделирование достаточно естественного и полноценного поведения простых модельных организмов. Моделирование в этом направлении уже начато, см. ниже.

Б. Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов. Такой переход можно рассматривать, как появление в «сознании» животного свойства «понятие». Обобщенные образы можно представить как мысленные аналоги наших слов, не произносимые животными, но реально используемые ими. Например, у собаки явно есть понятия «хозяин», «свой», «чужой», «пища». 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 1.1). 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 возник 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. При этом пункты *Б*, *В*, *Г* характеризуют следующие друг за другом и связанные между собой этапы эволюции познавательных способностей биологических организмов. И можно ожидать, предлагаемые по данной программе исследования помогут понять причины возникновения существенных познавательных свойств организмов в процессе эволюции.

Опираясь на эти пункты, мы начали соответствующее моделирование, которое характеризуется в следующем разделе.

1.5. Начальные шаги моделирования когнитивной эволюции

Компьютерная модель автономных агентов, имеющих естественные потребности. В [30, 31] была построена и исследована компьютерная модель автономных адаптивных агентов, обладающих естественными для живых организмов потребностями питания, безопасности и размножения (пункт *А* контуров программы). Предполагалось, что каждый агент имеет определенный ресурс R . Ресурс агента R уменьшался при выполнении действий и увеличивался при питании. Кроме того, ресурс агента существенно уменьшался, когда на агента нападал активный хищник. Предполагалось, что рядом с агентом имелся хищник, активность которого периодически менялась: активный хищник отнимал ресурс у агента, неактивный хищник был нейтральным.

Потребностям агента соответствовали три фактора: фактор питания F_F , фактор безопасности F_S , фактор размножения F_R . Для каждого фактора имелся порог T_F , T_S и T_R , при превышении фактором которого удовлетворялась соответствующая потребность.

Между потребностями вводилась следующая иерархия: 1) пищевая потребность (наиболее приоритетна), 2) потребность безопасности, 3) потребность размножения

(наименее приоритетна). Потребность считалась ведущей, если она была наиболее приоритетной из всех потребностей, для которых фактор меньше своего порога, т.е. $F_N < T_N$.

Указанная иерархия потребностей была подтверждена с помощью эволюционного моделирования, при котором для каждой потребности вводился коэффициент значимости, и иерархия потребностей выстраивалась в соответствии с убыванием значимости. Проводилось компьютерное моделирование согласно основной модели, только дополнительно при размножении значимости потребностей передавались потомку со случайными мутациями. После достаточно большого времени моделирования в популяции агентов преимущественно отбирались агенты с описанной выше иерархией.

Система управления агента была основана на наборе правил вида: $S_k \rightarrow A_k$, где S_k – ситуация, A_k – действие. Веса правил W_k настраивались методом обучения с подкреплением [17]. Ситуация S_k определялась активностью хищника рядом с агентом, предыдущим действием агента и ведущей потребностью. Агент мог выполнять одно из следующих действий A_k : 1) покой, 2) поиск пищи, 3) питание, 4) подготовка к размножению, 5) размножение, 6) оборона.

Время t предполагалось дискретным, каждый такт времени каждый агент выполнял действие, определяемое его системой управления.

При выборе действия использовалось ϵ -жадное правило [17]. Каждый такт времени с вероятностью $1-\epsilon$ выполнялось то действие, для которого вес W_k соответствующего ему правила для текущей ситуации максимален, с вероятностью ϵ выполнялось случайное действие.

Ресурс агента R уменьшался при выполнении действий (а также под воздействием активного хищника) и увеличивался после выполнения последовательной цепочки действий «поиск пищи», «питание». После выполнения цепочки действий «подготовка к размножению», «размножение», агент рождал потомка. Агент-родитель передавал агенту-потомку часть своего ресурса R . При этом агент-потомок наследовал веса правил W_k родителя. Выполняя действие «оборона», агент защищался от активного хищника и соответствующей потери ресурса. Когда ресурс агента становился меньше 0, этот агент погибал.

Факторы F_F , F_S , F_R увеличивались при выполнении действий, направленных на удовлетворение ведущих потребностей, а именно, это осуществлялось следующим образом. Фактор питания F_F увеличивался при питании агента и соответствующем росте его ресурса R . Фактор безопасности увеличивался при выполнении агентом действия «оборона» и немного уменьшался при выполнении любого другого действия. Фактор размножения значительно возрастал при успешном рождении потомка, не менялся при выполнении действия «подготовка к размножению» и немного уменьшался при выполнении других действий.

В модели использовалось обучение с подкреплением [17]. Веса правил $S_k \rightarrow A_k$ менялись следующим образом:

$$\Delta W_{t-1} = \alpha (F_{L,t} - F_{L,t-1} + \gamma W_t - W_{t-1}), \quad (5)$$

где $F_{L,t}$, $F_{L,t-1}$ – факторы потребности, ведущей в такты времени t и $t-1$, W_t и W_{t-1} – веса правил, примененных в такты t и $t-1$, α – параметр скорости обучения, γ – дисконтный фактор ($\alpha = 0.1$, $\gamma = 0.8$).

Результаты компьютерного моделирования иллюстрируются рис. 1, на котором показана типичная динамика факторов F_F , F_S и F_R . Видно, что при достаточно большом времени все три фактора достигают порога, т.е. все потребности удовлетворяются. Формируется цикличность поведения агента, характерный цикл показан участком «а». В цикле агент сначала накапливает внутренний ресурс, затем его

действия направлены на максимизацию безопасности, а когда обе потребности (пищевая и потребность безопасности) удовлетворены, агент размножается.

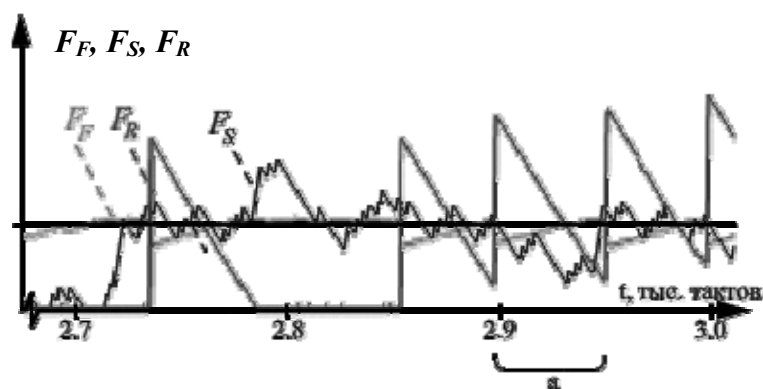


Рис. 1. Зависимость факторов F_F , F_S и F_R , соответствующих потребностям агента, от времени t .

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

Завершая обсуждение изложенной модели, отметим, что в настоящее время проводятся идейно близкие работы по моделированию поведения автономных агентов с несколькими потребностями [32]. Хотя в [32] не наблюдалось формирования цепочек действий и циклов поведения, как это представлено выше, в этой работе введена потребность поискового поведения, что позволяет моделировать отложенное обучение (первоначальное исследование лабиринта, а затем использование накопленного знания при поиске пищи), известное для живых организмов [33]. Отметим, что достаточно формальный подход к моделированию отложенного обучения был предложен и продемонстрирован на простом примере также в работе [25]. Таким образом, в дальнейшем возможен более полный анализ поведения агентов с несколькими потребностями, основанный на методах работ [25, 30-32].

Модель формирования обобщенных эвристик и простых обобщенных образов. Формирование обобщенных эвристик и простых обобщенных образов (пункт *B*) в процессе обучения агентов при поиске агентами пищи в двумерной клеточной среде было продемонстрировано в модели [34, 35]. Модель предполагала, что в части клеток были порции пищи. При съедании пищи ресурс агента пополнялся. Система управления агента также была основана на правилах вида $S \rightarrow A$, веса правил оптимизировались методом обучения с подкреплением. В результате обучения агент формировал обобщающие эвристики. Эти эвристики таковы: если имеется пища в той же клетке, в которой находится агент, то нужно съесть пищу; если в клетке агента нет пищи, но есть пища в клетке впереди или справа/слева от агента, то нужно выполнить действие «перемещение вперед» или «поворот направо/налево». Использование эвристик приводило к формированию последовательных цепочек действий, приводящих к пополнению ресурса агента. Дополнительно в компьютерную программу вводилась процедура усреднения: вычислялось среднее число применений данного действия для той или иной ситуации [35]. В результате усреднения агент формировал внутренние понятия «имеется пища в моей клетке», «имеется пища в клетке впереди меня», «имеется пища в клетке справа/слева от меня».

Таким образом, построена модель формирования обобщенных эвристик и простых обобщенных образов, хотя пока эту модель целесообразно рассматривать только как начальный этап более полноценных исследований.

1.6. Перспективы моделирования когнитивной эволюции

Обсуждая перспективы моделирования когнитивной эволюции, сначала отметим ряд скептических сомнений и наметим подходы к разрешению этих сомнений.

Начнем с яркой работы Ю. Вигнера [36], который заострил внимание на проблеме «непостижимой эффективности математики в естественных науках». Однако в отличие от Ю. Вигнера, который подчеркивал многочисленные трудности понимания «непостижимой эффективности математики», здесь мы намечаем подходы к конструктивному анализу применимости математики в естественных науках путем исследования эволюции познавательных процессов, приведших к логическим правилам, используемых в математике. Естественно, что мышление человека не сводится только к правилам доказательства теорем, но эти правила достаточно четко формализованы [4] и лежат в основе математического знания и на них имеет смысл сконцентрироваться на начальном этапе исследований.

Также отметим интересную работу Н.Б. Новикова «Происхождение человеческой логики» [37], в которой излагаются результаты экспериментальных исследований когнитивных, интеллектуальных способностей высших животных. Сформировано определенное понимание эволюционного происхождения интеллекта животных и человека. Кроме того, хорошо подчеркивается роль индукции и аналогии в процессах творческого мышления высших животных и человека, в том числе в процессах научного познания. Однако на основании приведенных данных делается несколько странный вывод о том, что проблема эволюционного происхождения человеческой логики решена. В статье [37] на ряде биологических примеров хорошо охарактеризованы отдельные интеллектуальные способности высших животных. Но почему эти способности возникали? Что они обеспечивали? Почему они эффективны? Каковы были причины возникновения когнитивных способностей? Как продемонстрировать с помощью компьютерных моделей эволюционное возникновение и эффективность когнитивных способностей? Эти вопросы пока открыты.

Более того, есть нетривиальная гносеологическая проблема (раздел 1.1): почему формальный логический вывод, сделанный человеком, применим к реальным объектам в природе? Почему результаты, полученные математиками формальным логическим путем, применимы к реальным физическим объектам? Проблема, действительно, нетривиальна. Причем проблема, несомненно, важна: ее исследование – это определенное обоснование всего научного познания.

Также важно, что в ряде направлений вычислительного интеллекта уже имеются заделы по компьютерному моделированию познавательных способностей человека и биологических организмов. Выше был кратко очерчен задел, развиваемый в направлении исследований «Адаптивное поведение». Близкие работы ведутся и в ряде других направлений. В частности, в настоящее время имеется огромное количество работ по автономным агентам [38]. Таким образом, в настоящее время как раз формируются предпосылки для моделирования когнитивной эволюции. И сопоставляя пункты контуров программы моделирования когнитивной эволюции с уже отмеченными работами [22-32, 34, 35, 38], можно заключить, что имеются отдельные небольшие 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.

В заключение раздела укажем потенциальные междисциплинарные связи будущих исследований когнитивной эволюции и подчеркнем аргументы в пользу этих исследований.

- 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, а также в ряде близких направлений 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.

2. НА ПУТИ К ТЕОРИИ ПРОИСХОЖДЕНИЯ МЫШЛЕНИЯ

В этой части статьи рассматривается широкий спектр подходов к исследованию происхождения мышления.

2.1. Направления и методы, которые могут использоваться

В вычислительном интеллекте сложилось несколько направлений исследований, которые связаны с моделированием происхождения мышления и которые могут быть использованы при таком моделировании.

Адаптивное поведение. Выше уже характеризовалось направление «Адаптивное поведение», модели которого применимы к описанию поведения как реального животного, так и искусственного модельного организма и дальняя цель которого – анализ эволюции когнитивных способностей животных в контексте происхождения интеллекта.

Заделом исследований происхождения мышления также являются работы по искусственной жизни, когнитивным архитектурам, моделям автономных агентов, близкие работы по научным основам искусственного интеллекта.

Искусственная жизнь. Близким направлением к «Адаптивному поведению» является интересное направление исследований «Искусственная жизнь». Основной мотивацией исследований искусственной жизни служит желание понять и промоделировать формальные принципы организации биологической жизни. Как сказал руководитель первой международной конференции по искусственной жизни К. Лангтон, «основное предположение искусственной жизни состоит в том, что "логическая форма" организма может быть отделена от материальной основы его конструкции».

Сторонники направления «Искусственная жизнь» часто считают, что они исследуют более общие формы жизни, чем те, которые существуют на Земле. Т.е. изучается жизнь, какой она могла бы в принципе быть (“life-as-it-could-be”), а не обязательно та жизнь, какой мы ее знаем (“life-as-we-know-it”). Искусственная жизнь – это синтетическая биология, которая по аналогии с синтетической химией пытается воспроизвести биологическое поведение в различных средах. Это жизнь, созданная человеком, а не природой (“life made by Man rather than by Nature”). Исследования искусственной жизни направлены не только на теоретические исследования свойств

жизни, но и (аналогично синтетической химии) на практические приложения, такие как подвижные роботы, медицина, нанотехнология, исследования «жизни» социальных и экономических систем.

Обзор ранних работ по искусственной жизни содержится в [39].

Когнитивные архитектуры. Под когнитивными архитектурами понимаются структура и принципы функционирования познающих систем, которые можно использовать в искусственном интеллекте. Пример когнитивной архитектуры – система Soar (от англ. State, Operator And Result). Soar – это основанная на символических представлениях достаточно общая когнитивная архитектура развивающихся систем, которая обладает свойствами интеллектуального поведения. Основная цель работ по Soar – создание системы функционирования интеллектуальных агентов, работающих в широкой области: от простейших форм до оперирования в сложных, заранее не предсказуемых условиях. Систему Soar предложили специалисты в области искусственного интеллекта еще в 1980 годах, тогда ее инициировали, как попытку построить унифицированную теорию познания. Обзор исследований по когнитивным архитектурам и Soar содержится в работах [40, 41]. В последние годы большое внимание уделяется биологически инспирированным когнитивным архитектурам (Biologically Inspired Cognitive Architectures, BICA) [42].

Интеллектуальные автономные агенты. Это близкое к когнитивным архитектурам направление исследований, в котором большое внимание уделяется биологически обоснованным автономным агентам и компьютерным моделям агентов, обладающих когнитивными и некоторыми интеллектуальными свойствами. Обзор исследований по автономным агентам содержится в работе [38]. Необходимо подчеркнуть, что автономные агенты вполне могут рассматриваться как объединяющее понятие для указанных направлений: агенты вполне естественно могут моделировать биологические организмы и искусственные аналоги организмов. Вполне естественно изучать познавательные свойства и знания автономных агентов. Также разумно исследовать эволюцию популяций агентов и коммуникации и обмен информацией между агентами.

Близкие работы по научным основам искусственного интеллекта. Такие работы включают исследования когнитивных систем роботов [43, 44], изучение нестандартных логик [45, 46], исследование познавательных методов, используемых человеком [47, 48].

Теперь кратко остановимся на методах, которые целесообразно использовать при моделировании проблемы происхождения мышления.

О методах моделирования. Помимо хорошо известных методов (нейронные сети [14], эволюционная оптимизация [15], обучение с подкреплением [17]) имеется ряд интересных и нетривиальных подходов к исследованию когнитивных, интеллектуальных свойств реальных и модельных организмов. Отметим некоторые из таких подходов.

Эмоции и стремление к накоплению знаний. В работах А.А. Жданова, Е.Е. Витяева и Л.И. Перловского [28, 29, 49, 50] подчеркивается важная роль эмоций и стремления автономных агентов к накоплению новых знаний.

Семантический вероятностный вывод. Автономный агент может использовать также нетривиальный метод семантического вероятностного вывода [29], в котором используется символическая информация, и который может быть более эффективен, чем обычный метод обучения с подкреплением.

Повторный вход, re-entrance и поле внимания. В ряде нейробиологических работ (см. например, работы А.М. Иваницкого и Дж. Эдельмана [51, 52]) исследуются нейронные механизмы возвратного возбуждения, повторного входа, при которых сенсорная информация повторно поступает на вход нервной системы. Использование повторного входа при моделировании предлагается В.Г. Яхно [53]. Возможно, что такой повторный вход существенен для поступления информации в поле внимания и осознания наиболее важной информации об объектах и событиях во внешней среде. Роль поля внимания и осознания той информации, которая помещается в поле внимания, отмечались в работах [12, 51, 54].

Разработка нейросетевых архитектур систем управления автономных агентов на основе теории функциональных систем П.К. Анохина была начата в работах [28, 29, 55, 56].

Еще один важный вопрос: как перейти от имеющейся системы управления поведением автономного агента к системе управления более высокого уровня? Концептуальная теория метасистемных переходов, характеризующая такие переходы с уровня на уровень, была предложена В.Ф. Турчиным [6].

Кратко суть теории метасистемных переходов сводится к следующему: переход от нижних уровней системной иерархии к верхним происходит путем метасистемных переходов. Каждый метасистемный переход можно рассматривать как объединение ряда подсистем S_i нижнего уровня и появление дополнительного механизма управления C объединенными подсистемами. В результате метасистемного перехода формируется система S' нового уровня ($S' = C + \sum_i S_i$), которая может быть включена как подсистема в следующий метасистемный переход.

В.Ф. Турчин рассматривает метасистемный переход как некий кибернетический аналог физического фазового перехода. Он уделяет особое внимание количественному накоплению «потенциала развития» в подсистемах S_i перед метасистемным переходом на качественно новый уровень иерархии, а также процессу размножения и развития подсистем предпоследнего уровня иерархии после метасистемного перехода.

Завершая раздел, отметим, что поскольку поставленная проблема моделирования происхождения мышления чрезвычайно сложна, то целесообразно использовать метод последовательных приближений. Например, в качестве первого приближения можно использовать очерченный выше (раздел 1.2) феноменологический подход.

2.2. Попытки найти путь к интеллекту человеческого уровня

На нескольких международных конгрессах по вычислительному интеллекту (World Congress on Computational Intelligence, Ванкувер, 2006; Гонконг, 2008) проводились представительные дискуссии по подходам к моделированию и возможному созданию интеллектуальных систем человеческого уровня.

На конференции по биологически инспирированным когнитивным архитектурам [42] был предложен такой подход: создать интеллектуального робота, поведение которого было бы настолько «разумно», что люди могли бы выбрать этого робота президентом страны [57]. Однако, по мнению автора настоящей статьи, можно подойти к вопросу и другим образом.

В работе М.Н. Вайнцвайга [58] анализировались модели автономных агентов, которые познают закономерности механики, наблюдая за столкновениями тел, например, шаров. По мнению М.Н. Вайнцвайга, исследование таких автономных агентов, в конечном итоге, нацелено на то, чтобы агенты могли бы прийти к самостоятельному открытию трех законов Ньютона. Ну, чем это не интеллект человеческого уровня (конечно, пока только в замыслах) – открытие законов природы?!

И можно представить, как мог бы развиваться «интеллект» автономного агента,

чтобы агент смог прийти к уровню ученого. У агента должна быть база знаний, должно быть стремление к получению новых знаний и к обобщению знаний, должна быть любознательность, направляющая агента к постановке вопросов о внешнем мире и решению этих вопросов путем постановки экспериментов. Агенты должны учитывать многочисленные связи между уже имеющимися знаниями. Должен быть коллектив агентов, исследующий внешний мир, и должны быть коммуникации между агентами. Должно быть самосознание агентов, эмоциональная самооценка результатов своей деятельности и стремление агента достигнуть наиболее высоких результатов в коллективе агентов. Агенты должны иметь стремление к получению наиболее ясных, четких и компактных знаний, таких как законы Ньютона или аксиомы Евклида. И конечно, агенты должны освоить возможности логических выводов, позволяющих получить многочисленные следствия законов и аксиом.

Понятно, что путь к интеллекту такого уровня весьма не прост, но, все же, он просматривается.

Отметим еще важный метасистемный переход (В.Ф. Турчин [6]), который характеризует конечные этапы этого пути. Это переход от первобытного мышления к критическому (от мышления первобытных племен к возникновению формального логического мышления). Критическое мышление отличается от первобытного тем, что возникает оценка мыслительного процесса самим мыслящим субъектом. «Критическое мышление рассматривает каждое объяснение (языковую модель действительности) наряду с другими, конкурирующими объяснениями (моделями), и оно не удовлетворится, пока не будет показано, чем данное объяснение лучше, чем конкурирующее» [6].

Отметим еще одно важное понятие, которое целесообразно использовать при анализе путей к интеллекту человеческого уровня, понятие «модель». Под моделями здесь подразумеваются «внутренние модели» внешнего мира, те модели, которые формируются в «базе знаний» автономного агента или животного.

Человек, естественно, тоже имеет модели, характеризующие его представления о внешнем мире. Более того, общая научная картина мира – создаваемая всем международным научным сообществом – также может рассматриваться как совокупность моделей.

2.3. Биологические эксперименты по «элементарному мышлению животных»

В последние годы проведены интересные биологические исследования, показывающие, что элементарные формы мышления присущи не только высшим животным, но и достаточно простым, например, насекомым (пчелам, муравьям), а также врановым птицам, которые способны самостоятельно создавать и целесообразно использовать «орудия труда». В данном разделе кратко характеризуются биологические экспериментальные данные и некоторое связанное с ними моделирование.

Пчелы сообщают другим пчелам, где искать медоносные цветы. Еще в 20-е годы XX века Карл фон Фриш начал исследования способа передачи пчелами информации друг другу. К. фон Фриш обратил внимание на то, что возвращающиеся в улей пчелы-сборщицы совершают танец, привлекающий внимание других пчел. Угол, составленный осью танца и вертикалью, соответствует углу между направлением на пищу и направлением на солнце. Причем этот угол для танцующей пчелы меняется в соответствии с движением солнца. Скорость танца соответствует расстоянию между пищей и ульем. Танцующая пчела также пахнет, ее запах характеризует медоносный цветок. Другие пчелы наблюдают этот танец и затем находят те цветы, с которых прилетела танцующая пчела (на расстоянии до нескольких километров). Долгое время

результаты К. фон Фриша вызывали споры, которые продолжались и после того, как в 1973 г. К. фон Фриш получил за свое открытие Нобелевскую премию.

Интересно, что в 1990-е годы Б. Андерсен и А. Михельсен создали искусственную пчелу-робота, которая также танцевала и с помощью танца передавала информацию живым пчелам. Живые пчелы летели из улья на поляну, руководствуясь указаниями пчелы-робота. Подробнее об этой передаче символической информации см. обзор Ж.И. Резниковой [59].

Муравьи могут передавать информацию со скоростью около 1 бита в минуту и запоминать небольшие целые числа. В семье рыжих лесных муравьев можно выделить рабочие группы, состоящие из одного разведчика и 3-8 фуражиров. Каждый разведчик, найдя пищу, вступает в контакт со своей группой и передает информацию, где нужно искать пищу. В экспериментах и расчетах Ж.И. Резниковой и Б.Я. Рябко исследовалось поведение муравьев с помощью специально изготовленных ветвящихся лабиринтов небольшой глубины: каждый путь разветвлялся на два, было до 5-6-ти развилок в лабиринте. Было установлено, что муравьи-разведчики могли запомнить путь к кормушке, а затем путем контактов с фуражирами передавать им информацию о том, в каком месте лабиринта находится пища [59, 60]. Зная глубину лабиринта и время контакта, Ж.И. Резникова и Б.Я. Рябко определяли количество информации, передаваемой разведчиками фуражирам и скорость передачи этой информации. В результате оказалось, что скорость передачи информации составляла около 1 бита в минуту.

Причем, если путь был достаточно простой, например, в каждой развилке лабиринта надо было поворачиваться в левую ветку, то разведчики могли использовать «сжатую» информацию и передавать ее разведчикам с большей скоростью. Образно говоря, муравьи «освоили» как информацию по Шеннону, так и информацию по Колмогорову (подробнее см. [59-61]).

Кроме этого, исследовалась передача информации между муравьями-разведчиками и фуражирами для более простых лабиринтов-гребенок, состоящих из одного основного канала, от которого ответвляются в одну сторону одинаковые каналы (до 40 веток) [60]. Пища находилась в одном из ответвлений. Муравьи оказались способными запоминать и передавать друг другу сведения о номере ветки, т.е. они умели определенным образом считать.

Новокаледонские вороны могут самостоятельно изготавливать орудия труда. Обычно в природе новокаледонские вороны могут обкусывать веточки так, что получаются простые орудия (заостренные палочки или крючки). На заостренную палочку ворона может насаживать личинки насекомых, а крючком вытаскивать личинок из-под коры.

В университете Оксфорда проводили исследования с воронами, находившимися долгое время в неволе [62]. Двум воронам (молодой самке и самцу постарше) предлагали добывать ведро с пищей со дна прозрачного вертикального цилиндра. Рядом с цилиндром были прямая проволока и проволока, согнутая крючком. Ведро можно было вытащить крючком, но не прямой проволокой. Раньше с проволокой вороны дела не имели. Тем не менее, они сразу поняли, что ведро можно вытащить с помощью крючка.

Неожиданность произошла, когда самец утащил крючок. Тогда самка сначала немного попыталась подцепить ручку ведра прямым куском проволоки (что было безуспешно), а затем быстро научилась делать из прямой проволоки крючок, зажимая один конец проволоки в одной из щелей экспериментальной установки и загибая проволоку. А затем с помощью изготовленного ей крючка доставала ведро с пищей. В дальнейшем, если крючок убирали, а оставляли прямую проволоку, ворона сразу

делала крючок и доставала пищу. Отметим, что самец не перенял опыт самки, он наблюдал за ней и иногда отнимал у нее пищу.

Таким образом, ворона сама, без какого-либо обучения, без каких-либо инструкций, изобрела способ изготовления орудия труда. Подробнее см. [62] и сайт исследователей новокаледонских ворон университета Оксфорда [<http://users.ox.ac.uk/~kgroup/index.html>, <http://users.ox.ac.uk/~kgroup/tools/introduction.shtml>].

Новокаледонские вороны могут мысленно составлять планы цепочек целенаправленных действий. Еще одно интересное исследование с новокаледонскими воронами провели исследователи из Новой Зеландии [63]. Задание для ворон состояло из 3 следующих частей. 1) Сначала надо было подтянуть к себе шнуром и освободить от шнура маленькую палочку, которая висела на шнуре. 2) Затем с помощью маленькой палочки достать из одного зарешеченного контейнера длинную палочку. 3) Наконец, с помощью длинной палочки надо было достать пищу из второго контейнера. Причем, без короткой палочки нельзя было достать длинную, а короткой палочкой нельзя было дотянуться до пищи во втором контейнере. Т.е. весь процесс добывания пищи должен был состоять из трех последовательных шагов, на которых надо было использовать три «инструмента»: шнур, маленькую палочку, длинную палочку.

Предварительно вороны тренировались в более простых условиях: они могли использовать часть или все эти инструменты по отдельности. Когда им нужно было выполнить полное задание из трех шагов, то те вороны, которые имели опыт использования всех трех инструментов по отдельности, выполняли задание с первого раза, а те, которые предварительно освоили инструменты частично, выполняли задание не всегда с первого раза, но, тем не менее, тоже быстро научились с ним справляться.

Таким образом, вороны научились продумывать план решения новой задачи, мысленно связывая в плане ранее освоенный опыт.

«Пространственный интеллект», запоминание и использование когнитивных карт. Многие животные (птицы, грызуны, насекомые), ориентируясь в пространстве или запасая корм, формируют в своей памяти когнитивные карты. Например, колумбийские сойки, обитающие на юго-западе США, в конце лета начинают собирать семена сосны. Наполнив подъязычный мешок, сойка улетает на расстояние до нескольких километров, чтобы припрятать семена на южных склонах холмов, где зимой бывает мало снега. Птица может запастись до 33 000 семян, по 4–5 штук в одном месте, т. е. получается несколько тысяч тайников. Зимой и весной птица наведывается в те места, в которых она прятала пищу в конце лета, и выкапывает пищу из своих складов. Подробнее о различных формах пространственного интеллекта животных см. [59].

В заключение раздела подчеркнем, экспериментальные биологические работы и моделирование когнитивных свойств биологических организмов по смыслу близки друг к другу, и возможно налаживание серьезных междисциплинарных связей между этими исследованиями. Возможно даже сочетание между экспериментальными (биологические эксперименты) и теоретическими (моделирование когнитивных способностей) исследованиями, подобное сочетанию между экспериментальной и теоретической физикой.

2.4. Попутные прикладные направления

До сих пор рассматривались исследования, направленные на изучение фундаментальной проблемы – моделирование происхождения мышления. Но при

серьезных исследованиях неизбежно попутно возникают и связанные с ними прикладные направления. В этом разделе отметим два направления возможных прикладных работ: 1) устранения причин агрессивной конкуренции на основе моделирования эволюции конкурирующих агентов, которые могут вести борьбу между собой, и 2) возникновение и использование коллективных знаний в децентрализованной многоагентной системе, использующей агентов-посыльных для обмена информацией между агентами.

Вымирание генов агрессивности в эволюционирующей популяции конкурирующих агентов. В работах [64, 65] была построена и исследована компьютерная модель адаптивного поведения агентов в эволюционирующей популяции. В модели рассматривался искусственный двухмерный клеточный мир, в клетках могли находиться агенты и их пища. В каждый такт времени агент выполнял одно из следующих действий: 1) питаться, 2) двигаться на одну клетку вперед, 3) повернуть налево, 4) повернуть направо, 5) отдыхать, 6) делиться (размножиться), 7) ударить агента в клетке спереди (напасть), 8) защищаться.

Каждый агент обладал внутренним энергетическим ресурсом R , который пополнялся, когда агент выполнял действие «питаться» и расходовался при выполнении других действий. При нападении одного агента на второго (выполнении действия «ударить») нападающий агент отнимал ресурс у второго агента, если второй агент не выполнял действие «защищаться»; но если второй агент защищался, то первый агент просто расходовал значительный ресурс.

Система управления агента представляла собой однослойную нейронную сеть, на вход сети подавалась сенсорная информация о ближайшем окружении агента и о внутреннем ресурсе агента, выходы сети определяли действия агента. Входы нейронной сети представляли собой сенсоры агента, выходы сети – эффекторы агента. Структура нейронной сети и веса синаптических связей между входом и выходом сети определяли геном агента. Геном агента (т.е. его нейронная сеть) в течение жизни агента не менялся. При делении потомок агента наследовал измененную мутациями нейронную сеть агента-родителя. При мутациях модифицировались как веса синаптических связей, так и структура нейронной сети, в частности, могли исчезать и появляться отдельные сенсоры или эффекторы. Популяция агентов эволюционировала во времени за счет вариаций структуры и весов синапсов нейронных сетей и отбора тех агентов, которые достаточно быстро набирают ресурс и размножаются.

В процессе компьютерных экспериментов была показана возможность возникновения иерархии целей у агентов в процессе искусственной эволюции. Кроме этого, при исследовании модели можно было наблюдать эффекты, связанные с агрессивной конкуренцией между агентами. В некоторых компьютерных экспериментах наблюдались неожиданные пики в зависимостях численности популяции агентов от времени (кратковременное резкое возрастание численности). Анализ системы управления агентов продемонстрировал, что пики соответствовали тем моментам времени, когда в нейронной сети агентов отсутствовали эффекторы, ответственные за борьбу между агентами. Для подтверждения этого полная компьютерная модель (с борьбой между агентами) сравнивалась с моделью, в которой эффекторы борьбы были полностью устранены из системы управления агентов. В случае такого «глобального пацифизма» установившаяся численность популяции была примерно в 2 раза выше, чем для обычных агентов, имеющих средства для борьбы между собой. Этот эффект связан с тем, что наличие агрессивной конкуренции приводило к тому, что агенты тратили свой ресурс R именно на борьбу друг с другом. Этим агентам не хватало ресурса на «созидательную» жизнь, для активного накопления ресурса и расходования его на процессы размножения, приводящие к возрастанию численности популяции. Таким образом, устранение борьбы между агентами,

исчезновение «генов агрессивности» приводило к увеличению численности популяции агентов.

Необходимо подчеркнуть, что наблюдавшийся при этом эффект исчезновения «генов агрессивности» носил кратковременный характер: наблюдавшиеся пики в численности популяции были довольно узкими. Кратковременность была связана с тем, что мутационное появление новых агрессивных агентов (с эффекторами борьбы) быстро приводило к тому, что агрессивные агенты вели активную борьбу с «мирными» агентами (не имевшими эффекторов борьбы) и выживали их из популяции. Агрессивная конкуренция приводила к уничтожению мирных агентов. Выживали только агенты, имеющие средства для агрессивной конкуренции.

На основе изложенной модели возможен анализ процессов исчезновения агрессивной конкуренции для различных социально-экономических приложений: для популяций конкурирующих фирм, регионов, государств. Например, если рассматривать эволюцию агентов как эволюцию государств, то можно думать о научном обосновании концепции глобального пацифизма, можно даже предложить идею проекта на Нобелевскую премию мира «Разработка научных основ всемирного разоружения».

Модель формирования и использования коллективных знаний в многоагентной системе. В работе [66] была построена многоагентная модель, предназначенная для решения задачи управления дорожным движением города с целью минимизации количества заторов и повышения скорости передвижения автомобилей. Модель основана на многоагентных системах, использующих «легких» агентов-посыльных, которые исследовали окружающую среду автомобилей, то есть определяли транспортную ситуацию и делали прогнозы заторов, эти прогнозы передавались более «тяжелым» агентам (агентам автомобилей), располагавшихся на автомобилях. Благодаря этим прогнозам, автомобили могли менять свой маршрут. Другие «легкие» агенты-посыльные по выбранным маршрутам автомобилей передавали от агентов автомобилей информацию о намеченных маршрутах агентам инфраструктуры, располагавшихся в транспортных узлах города (таких как перекрестки). Агенты инфраструктуры коллективно формировали карту транспортной ситуации во всем городе. Эта карта определяла окружающую среду автомобилей.

Целью предложенного механизма координации движения автомобилей являлось сокращение длительности и дальности поездки. Конечный выбор маршрута оставался на усмотрение водителя автомобиля, но он тесно взаимодействовал с агентом автомобиля. Этот механизм децентрализован, то есть, с одной стороны, агенты автомобилей самостоятельно делали прогнозы характеристик движения по тому или иному маршруту и на основе этих прогнозов выбирали оптимальный текущий маршрут, а с другой стороны, агенты автомобилей, зная выбранные маршруты, формировали общую карту транспортной ситуации в городе. Этот механизм самосогласованный: агенты автомобилей одновременно 1) формировали и 2) использовали общую карту транспортной ситуации, и динамичный: как выбираемые автомобилями маршруты, так и карта ситуации динамически менялись.

Работоспособность и эффективность этой схемы многоагентной системы, включающей в себя агентов-посыльных, была продемонстрирована путем компьютерного моделирования. Компьютерные эксперименты были выполнены при моделировании в реальных условиях дорожного движения на дорожной сети города Лёвен (Leuven), Бельгия. Результаты моделирования показали, что предложенная схема децентрализованной маршрутизации на основе прогнозов и нетривиального использования агентов-посыльных может обеспечить достижение автомобилями их пунктов назначения на 35% быстрее, чем для ранее разработанных аналогичных схем.

Важно то, что в данной модели все агенты автомобилей конкурируют между собой, но, благодаря обмену информацией и коллективному формированию общих знаний об

инфраструктуре всей транспортной системы в городе и о динамике этой инфраструктуры, они способны находить приемлемое для всех решение.

Заключая раздел, отметим, что очерченные модели важны как для понимания когнитивных свойств автономных агентов и процессов формирования распределенных знаний в многоагентной популяции, так и с прикладной точки зрения.

2.5. О развитии научного миропонимания

Исследование проблемы эволюционного происхождения мышления могло бы внести серьезный вклад в развитие научного миропонимания. Как в процессе эволюции возникли логика, мышление, интеллект человека – это и интересно любознательным людям, и важно с научной точки зрения.

Более того, исследование проблемы происхождения мышления могло бы способствовать более гармоничному развитию человечества. Ведь во многих регионах мира возникает напряженность в международных отношениях, идут локальные войны, связанные с различием традиций и религий разных народов.

Различие мировоззрений, сложившееся в разных религиях, связано с различными точками зрения на происхождение человека и его мышления. И для налаживания взаимопонимания между людьми, придерживающихся разных мировоззрений, был бы полезен широкий международный диалог по проблемам происхождения мышления человека, основанный на наиболее серьезном и глубоком миропонимании – научном миропонимании. Например, представим себе, что молодые люди в Палестине и Израиле совместно занимаются научным исследованием проблемы происхождения мышления человека – так уж ли будут важны для них их религиозные разногласия, и нужно ли им будет воевать друг с другом? Серьезные совместные научные исследования, направленные на развитие научного миропонимания, могут вытеснить религиозную вражду.

ЗАКЛЮЧЕНИЕ

Итак, сделан краткий обзор подходов к моделированию происхождения мышления. Видно, что происходит формирование широких междисциплинарных исследований в рассматриваемой области. Имеется ряд заделов моделирования происхождения мышления в таких направлениях, как «Адаптивное поведение», «Искусственная жизнь», «Когнитивные архитектуры». В этих направлениях большое внимание уделяется биологически обоснованным автономным агентам и компьютерным моделям агентов, обладающих когнитивными и интеллектуальными свойствами. С помощью автономных агентов естественно исследовать поведение биологических и модельных организмов, а также эволюцию познавательных свойств агентов. Более того, можно представить, как мог бы развиваться «интеллект» автономного агента до такого уровня, чтобы агент смог бы прийти к уровню ученого, например, чтобы агент мог бы прийти к самостоятельному открытию трех законов Ньютона (раздел 2.2).

Еще одно близкое направление работ – биологические исследования, направленные на изучение «элементарного мышления животных». Хотя работ на стыке биологии и вычислительного интеллекта, направленных на моделирование «интеллектуального» поведения животных, пока еще явно не достаточно.

Среди подходов к моделированию происхождения мышления наиболее четко очерчен подход к исследованию эволюционного происхождения логического мышления, используемого в математических доказательствах (см. часть 1 и [67]). Существенное достоинство этого подхода – ясен конечный результат эволюции:

правила логического вывода, используемые при математических доказательствах, известны и достаточно хорошо формализованы.

Итак, моделирование эволюционного происхождения мышления могло бы представить собой чрезвычайно интересную область исследований, направленную на существенное развитие научного миропонимания.

Литература

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: Learning, Development and Culture: Essays in Evolutionary Epistemology. Ed. Plotkin H. New York: Wiley, 1982. P. 121-143. См. также перевод: Лоренц К. Кантовская концепция а priori в свете современной биологии. // В: Эволюция. Язык. Познание. Под ред. Меркулова И.П. М.: Языки русской культуры, 2000. С. 15-41.
4. Математическая теория логического вывода. Под ред. Идельсона А.В. и Минца Г.Е. 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.: Наука, 1993 (1-е изд.). М.: ЭТС, 2000 (2-е изд.). См. также: <http://www.refal.ru/turchin/phenomenon/>
7. From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior. Eds. Meyer J.-A., Wilson S.W. Cambridge: MIT Press, 1991.
8. От моделей поведения к искусственному интеллекту. Под ред. Редько В.Г. М.: УРСС, 2006.
9. Непомнящих В.А. Поиск общих принципов адаптивного поведения живых организмов и аниматов // Новости искусственного интеллекта. 2002. № 2. С. 48-53.
10. 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.
11. Цетлин М.Л. Исследования по теории автоматов и моделирование биологических систем. М.: Наука, 1969.
12. Бонгард М.М., Лосев И.С., Смирнов М.С. Проект модели организации поведения – «Животное» // Моделирование обучения и поведения. М.: Наука, 1975. С. 152-171. См. также [8], С. 61-81.
13. Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота: модели поведения. М.: Наука, 1987. М.: УРСС, 2004, 2011.
14. Хайкин С. Нейронные сети: полный курс (2-е изд.). М.: Издательский дом «Вильямс», 2006.
15. 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).
16. Holland J.H., Holyoak K.J., Nisbett R.E., Thagard P. Induction: Processes of Inference, Learning, and Discovery. Cambridge: MIT Press, 1986.
17. Sutton R., Barto A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. Саттон Р.С., Барто Э.Г. Обучение с подкреплением. М.: Бином, 2011.
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. Непомнящих В.А., Попов Е.Е., Редько В.Г. Бионическая модель адаптивного поискового поведения // Известия РАН. Теория и системы управления. 2008. № 1. С. 85-93.
20. Гельфанд И.М., Цетлин М.Л. Принцип нелокального поиска в задачах автоматической оптимизации // ДАН СССР. 1961. Т. 137. № 2. С. 295-298.
21. Непомнящих В.А., Редько В.Г. Биологически обоснованный метод формирования поискового поведения // Двенадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2010. Труды конференции. Т.4. М.: Физматлит, 2010. С. 122-127.
22. Grossberg S. Classical and instrumental learning by neural networks // Progress in Theoretical Biology. 1974. V. 3. PP. 51-141.
23. 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.
24. Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior. Eds. Butz M.V., Sigaud O., Pezzulo G., Baldassarre G. LNAI 4520, Berlin, Heidelberg: Springer Verlag, 2007.
25. Witkowski M. An action-selection calculus // Adaptive Behavior. 2007. V. 15. No. 1. PP. 73-97.
26. 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.
27. 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
28. 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>
29. 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>
30. Коваль А.Г. Введение в систему управления аниматов естественных потребностей и мотиваций // XIV Всероссийская научно-техническая конференция «Нейроинформатика-2012». Сборник научных трудов. В 3-х частях. Часть 2. М.: МИФИ, 2012. С. 125-133.
31. Коваль А.Г. Редько В.Г. Поведение модельных организмов, обладающих естественными потребностями и мотивациями // Математическая биология и биоинформатика (электронный журнал). 2012 (в печати).
32. 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.
33. Tolman E.C. Purposive behavior in animals and men. New York: Appleton, 1932.
34. 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.
35. Бесхлебнова Г.А., Редько В.Г. Модель формирования обобщенных понятий автономными агентами // Четвертая международная конференция по когнитивной науке: Тезисы докладов: В 2-х томах. Томск: ТГУ, 2010. Т. 1. С. 174-175.
36. Вигнер Е. Непостижимая эффективность математики в естественных науках // Успехи физических наук. 1968. Т. 94, вып. 3. С. 535-546.

37. Новиков Н.Б. Происхождение человеческой логики // *Нейроинформатика* (электронный рецензируемый журнал), 2010. Т. 4. № 1. С. 1-30.
<http://www.niisi.ru/iont/ni/Journal/V4/N1/Novikov.pdf>
38. Vernon D., Metta G., Sandini G. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents // *IEEE Transactions on Evolutionary Computation* (special issue on Autonomous Mental Development). 2007. V. 11. No. 2. PP. 151-180.
39. Редько В.Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. М.: УРСС, 2005.
40. Langley P., Laird J.E., Rogers S. Cognitive architectures: Research issues and challenges // *Cognitive Systems Research*. 2009. V.10. No. 2. PP. 141-160.
41. Lehman J.F., Laird L., Rosenbloom P. 2006. A gentle introduction to Soar: Architecture for human cognition: 2006 update (электронный ресурс). URL: <http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf>
42. *Biologically Inspired Cognitive Architectures 2011. Proceedings of Second Annual Meeting of the BICA Society*. Eds. Samsonovich A.V., Johannsdottir K.R. Amsterdam et al.: IOS Press, 2011.
43. Станкевич Л.А. Когнитивный подход к управлению гуманоидными роботами // В книге [8]. С. 386-443.
44. Станкевич Л.А. Искусственные когнитивные системы // Научная сессия МИФИ-2010. XII Всероссийская научно-техническая конференция «Нейроинформатика-2010»: Лекции по нейроинформатике. М.: НИЯУ МИФИ, 2011. С. 106-160.
45. Финн В.К. О машинно-ориентированной формализации правдоподобных рассуждений в стиле Ф.Бэкона - Д.С. Милля // *Семиотика и информатика*. 1983. М.: ВИНТИ. Вып. 20. С. 35-101.
46. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. М.: Физматлит, 2008.
47. Осипов Г.С. Лекции по искусственному интеллекту. М.: УРСС, 2009.
48. Anshakov O., Gergely T. *Cognitive Reasoning: A Formal Approach*. Heidelberg et al.: Springer, 2010.
49. Жданов А.А. Автономный искусственный интеллект. М.: Бином, 2008.
50. Перловский Л.И. К физической теории мышления: теория нейронных моделирующих полей // *Нейроинформатика* (электронный рецензируемый журнал), 2006. Т. 1, № 2. С. 175-196.
URL: <http://www.niisi.ru/iont/ni/Journal/N2/Perlovsky.pdf>
51. Иваницкий А.М. Проблема «Сознание и мозг» и искусственный интеллект // Научная сессия МИФИ-2006. VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006»: Лекции по нейроинформатике. М.: МИФИ, 2006. С. 74-87.
52. Edelman G.M. Group selection and phasic reentrant signaling: A theory of higher brain function // In: *The Mindful Brain*. Cambridge: MIT Press, 1978. PP. 51-100.
53. Яхно В.Г. Динамика нейроноподобных моделей и процессы «сознания» // Научная сессия МИФИ-2006. VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006»: Лекции по нейроинформатике. М.: МИФИ, 2006. С. 88-111.
54. Вайнцвайг М.Н., Полякова М.П. О моделировании мышления // В книге [8]. С. 61-81.
55. 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.

56. 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: Eds. Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior. LNAI 4520, Berlin, Heidelberg: Springer Verlag. 2007. PP. 94-107.
57. Chella A., Lebiere C., Noelle D.C., Samsonovich A.V. On a Roadmap to Biologically Inspired Cognitive Agents // In [42], PP. 453-460.
58. Вайнцвайг М.Н. Обучение управлению поведением в мире объектов пространства-времени // Научная сессия МИФИ-2011. XIII Всероссийская научно-техническая конференция «Нейроинформатика-2011»: Лекции по нейроинформатике. М.: НИЯУ МИФИ, 2011. С. 111-129.
59. Резникова Ж.И. Современные подходы к изучению языкового поведения животных // Сборник «Разумное поведение и язык. Коммуникативные системы животных и язык человека», 2008. М.: «Языки славянских культур». С. 293 - 337. См. также: <http://www.reznikova.net/Lang09.pdf>
60. Резникова Ж. И., Рябко Б. Я. Теоретико-информационный анализ «языка» муравьев // Журнал общей биологии. 1990. Т. 51. № 5. С. 601-609.
61. Резникова Ж.И., Рябко Б.Я. Передача информации о количественных характеристиках объекта у муравьев // Журнал высшей нервной деятельности, 1995. Т. 45. С. 500 - 509.
62. Weir A.A.S., Chappell J., Kacelnik A. Shaping of hooks in New Caledonian Crows // Science. 2002. V. 297. No. 5583. PP. 981-983.
63. Taylor A.H., Elliffe D., Hunt G.R., Gray R.D. Complex cognition and behavioural innovation in New Caledonian crows // Proc. R. Soc. B. 2010. V. 277. PP. 2637–2643.
64. Бурцев М.С. Модель эволюционного возникновения целенаправленного адаптивного поведения. 2. Исследование развития иерархии целей // Препринт ИПМ РАН, 2002, № 69.
65. Бурцев М.С., Редько В.Г. Влияние агрессии на эволюцию в многоагентной системе // Сб. трудов 9-ой Международной конференции «Проблемы управления безопасностью сложных систем». М., ИПУ. 2002.
66. Claes R., Holvoet T., Weyns D. A decentralized approach for anticipatory vehicle routing using delegate multiagent systems // IEEE Transactions on Intelligent Transportation Systems, 2011. V. 12, issue 99. PP. 364-373.
67. Редько В.Г. Моделирование когнитивной эволюции – перспективное направление исследований на стыке биологии и математики // Математическая биология и биоинформатика (электронный журнал). 2010. Т. 5. № 2. С. 215-229. URL: [http://www.matbio.org/downloads/Redko2010\(5_215\).pdf](http://www.matbio.org/downloads/Redko2010(5_215).pdf)