

Модели адаптивного поведения – бионический подход к искусственному интеллекту*

Редько В.Г.

НИИ Системных исследований РАН, Москва, Россия

E-mail: vgrecko@gmail.com

Статья представляет собой аналитический обзор современных исследований в области моделирования адаптивного поведения. Основной подход этого направления – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы называются «аниматами» (от англ. animal + robot = animat). Сделан обзор работ последней международной конференции по моделированию адаптивного поведения (Рим, сентябрь, 2006 г.). Представлены исследования по проекту «Мозг анимата».

1. ВВЕДЕНИЕ

На стыке информатики и биологии есть ряд глубоких проблем, имеющих мировоззренческое значение, исследованию которых до сих пор уделяется несоизмеримо малое внимание по сравнению с их значимостью.

Действительно, в процессе биологической эволюции возникли чрезвычайно сложные и вместе с тем удивительно эффективно функционирующие живые организмы. Эффективность, гармоничность и согласованность работы «компонент» живых существ обеспечивается биологическими управляющими системами. Но каковы эти управляющие системы? Как и почему они эволюционно возникли? Какие информационные процессы обеспечивают работу этих управляющих систем? Как животные познают внешний мир и используют это познание для управления своим поведением? Как эволюционное развитие познавательных способностей животных привело к возникновению интеллекта человека? До какой степени исследования причин возникновения естественного интеллекта могут способствовать развитию искусственного интеллекта?

В настоящей статье мы обсудим подходы к изучению этих проблем. Тем не менее, подчеркнем, что это не просто концептуальные подходы, а подходы на основе как концептуальных разработок, так и исследований содержательных математических и компьютерных моделей биологических информационных и когнитивных систем.

Раздел 2 содержит краткий обзор направления исследований «Адаптивное поведение». В разделе 3 будет охарактеризован ряд математических и компьютерных моделей, представляющих возможный сценарий развития информационно-кибернетических систем в процессе происхождения жизни. Раздел 4 описывает проект «Мозг анимата», нацеленный на формирование общей «платформы» для систематического построения моделей адаптивного поведения.

2. НАПРАВЛЕНИЕ ИССЛЕДОВАНИЙ «АДАПТИВНОЕ ПОВЕДЕНИЕ»

2.1. From animals to animats – направление исследований «Адаптивное поведение»

С начала 1990-х годов за рубежом активно развивается направление исследований «Адаптивное поведение» [1-3]. Основной подход этого направления – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы называются «аниматами» (от англ. animal +

* Работа выполнена при финансовой поддержке РФФИ, проект № 07-01-00180.

robot = animat). Данное направление исследований рассматривается как бионический подход к разработке систем искусственного интеллекта.

Подчеркнем, что это активно развивающееся направление исследований. Есть международное общество «International Society for Adaptive Behavior»: <http://www.isab.org>. Регулярно проводятся международные конференции «Simulation of Adaptive Behavior (From Animal to Animat)». Издается журнал «Adaptive Behavior».

В настоящее время исследования адаптивного поведения включают в себя работы по следующим темам [4]:

- сенсорные системы и управление,
- обучение и адаптация,
- выбор действий, навигация и внутренние модели мира,
- антиципаторное адаптивное поведение,
- нейроэволюция (настройка нейронных сетей аниматов эволюционными методами),
- возникновение языка и коммуникаций при адаптивном поведении,
- коллективное и социальное поведение,
- биологически инспирированное адаптивное поведение роботов,
- поведение и мышление как сложные адаптивные системы.

Рис. 1 иллюстрирует междисциплинарные связи этого направления исследований. Моделирование адаптивного поведения связано с изучением естественного интеллекта – интеллект необходим для организации поведения [5]. Приложения моделей адаптивного поведения – искусственный интеллект, робототехника, модели адаптивного поведения в социально-экономических системах.

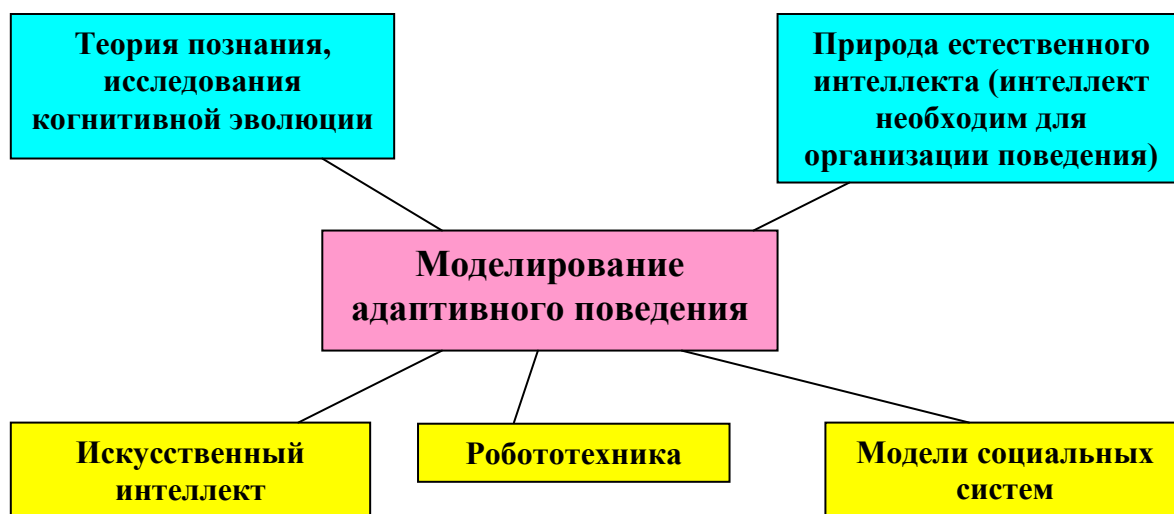


Рис. 1. Схема междисциплинарных связей направления исследований «Адаптивное поведение»

Подчеркнем, что при моделировании адаптивного поведения в основном 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. Такой 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 статистическое описания относительно независимы друг от друга и, вместе с тем,

взаимодополнительны. По-видимому, и для описания живых организмов может быть аналогичное соотношение между феноменологическим (на уровне поведения) и микроскопическим (на уровне нейронов и молекул) подходами. Естественно ожидать, что для исследования систем управления адаптивным поведением феноменологический подход на начальных этапах работ должен быть более эффективен, так как очень трудно сформировать целостную картину поведения на основе анализа всего сложного многообразия функционирования нейронов, синапсов, молекул.

Проиллюстрируем работы по моделям адаптивного поведения типичными примерами.

2.2. Примеры моделей адаптивного поведения

2.2.1. Модели мозга и поведения в Институте нейронаук Дж. Эдельмана. В Институте нейронаук Дж. Эдельмана (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), построенного на базе этих моделей.

Работы по NOMAD'у – исследование поведения адаптивного устройства на основе моделей 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 промоделировано поведение мыши в лабиринте Морриса [6].

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

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 находится 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 платформу. Обучение нейронных сетей NOMAD’а осуществлялось по модифицированному правилу Хебба 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.

Изложенная модель представляет собой эмпирическое компьютерное исследование, хорошо продуманное с биологической точки зрения. Поведение NOMAD’а нетривиально: обучение происходит быстро, формируются «нейроны места» в модельном гиппокампе, исследовано формирование «причинных связей» между областями модельного гиппокампа. Было бы полезно более формализованное исследование, дополнительное к этим эмпирическим работам.

2.2.2. Модель эволюционного возникновения коммуникаций в коллективе роботов (Д. Марокко, С. Нолфи, Институт когнитивных наук и технологий, Рим) [7]. В данной модели 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, 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, а третий нейрон – интенсивность силы звука, издаваемого в данный момент роботом. В ограниченной области пространства имелось 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. 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 [7] продемонстрировано, что в эволюционирующей популяции роботов, управляемых рекуррентными нейронными сетями, может формироваться система коммуникаций, позволяющая решать достаточно нетривиальную задачу распределения роботов по кормушкам.

2.2.3. Бионическая модель поискового адаптивного поведения (В.А. Непомнящих, ИБВВ им. И.Д. Папанина РАН, Е.Е. Попов, МФТИ, В.Г. Редько, НИИСИ РАН). Одно из актуальных направлений исследований в рамках моделирования адаптивного поведения – имитация поискового поведения животных. В [8] исследовано поисковое поведение на примере личинок ручейников *Chaetopteryx villosa* – насекомых, обитающих на дне водоемов. Личинки носят на себе «домик» – трубку из песка и других частиц. Строительство требует меньше времени, усилий и белка, если личинки используют относительно крупные и плоские частицы: общая протяженность швов между составляющими его немногими крупными частицами оказывается меньше, чем в том случае, когда домик сооружается из большого числа мелких частиц. Однако поиск крупных частиц на дне водоема требует затрат времени и энергии, не известных ручейнику заранее. Задача

осложняется еще и тем, что личинки при поиске частиц не пользуются зрением и могут обнаружить частицу и определить её размер только наощупь, что требует дополнительных затрат времени.

В [8] построена компьютерная модель поискового поведения личинок ручейников, строящих чехол-домик из частиц разного размера и ведущих поиск скоплений подходящих частиц. Модель использует понятие мотивации к прикреплению частиц к домику. Продемонстрирована адекватность построенной модели биологическим экспериментальным данным. Модель характеризуется как своей спецификой, обусловленной памятью о размерах последних обработанных частиц, так и общими свойствами инерционного переключения, позволяющими животным выявлять и использовать при адаптивном поведении наиболее общие закономерности взаимодействия с внешней средой.

Представленные модели довольно разнородны. Поэтому в следующем разделе рассмотрим ряд моделей предбиологических молекулярно-генетических информационных систем, для которых уже возможно провести определенное упорядочивание.

3. МОДЕЛИ ПРЕДБИОЛОГИЧЕСКИХ МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ

Как возникли первые молекулярно-генетические информационные системы в процессе происхождения жизни? В 60-80 годы подобные проблемы заинтриговали многих ученых. Целая плеяда Нобелевских лауреатов (М. Эйген, Ф. Крик, Ф. Дайсон, Ф. Андерсон) предприняла попытки представить и промоделировать сценарии возникновения предбиологических информационных систем.

Когда знакомишься с этими работами, то возникает впечатление, что в те годы ученые, получив мировое признание в конкретной науке (физике, биофизике, биохимии) и продолжающие активную научную работу, решали выбрать для себя наиболее интересную область исследований и останавливали свой взгляд на проблеме происхождения жизни.

В результате этих попыток был разработан ряд интересных математических и компьютерных моделей, позволяющих представить возможные схемы эволюции, приведшие к возникновению молекулярно-генетических самовоспроизводящихся систем простейших клеточных организмов.

Принципы функционирования молекулярно-генетических систем живых клеток схематично представлены на рис. 2. В клетке есть двойная спираль ДНК (постоянная память клетки), с которой определенным образом (путем транскрипции) считываются матричные РНК (мРНК). По мРНК (путем трансляции) синтезируются белки. В процессе трансляции участвуют транспортные РНК (тРНК). Синтез белков происходит в сложных макромолекулярных структурах – рибосомах. В процессе синтеза белков полинуклеотидные цепочки мРНК служат матрицами, на основе которых формируется цепочки белковых молекул. Синтезируемые ферменты принимают участие в процессах репликации, транскрипции и трансляции.

Эта схема нетривиальна с молекулярно-генетической точки зрения: для ее функционирования необходим четкий генетический код (соответствие между триплетами ДНК и пептидными основаниями белковых молекул), необходима сложная "механика" работы рибосом, и, самое главное, все в целом должно согласовано работать.

Но каковы могли быть механизмы возникновения такой схемы в процессе происхождения жизни? Попыткам осмыслить этот вопрос и было посвящено моделирование эволюционного возникновения молекулярно-генетических информационных систем. Впечатляющие результаты были получены М. Эйгеном [9-10], который предложил и исследовал модели квазивидов и гиперциклов.

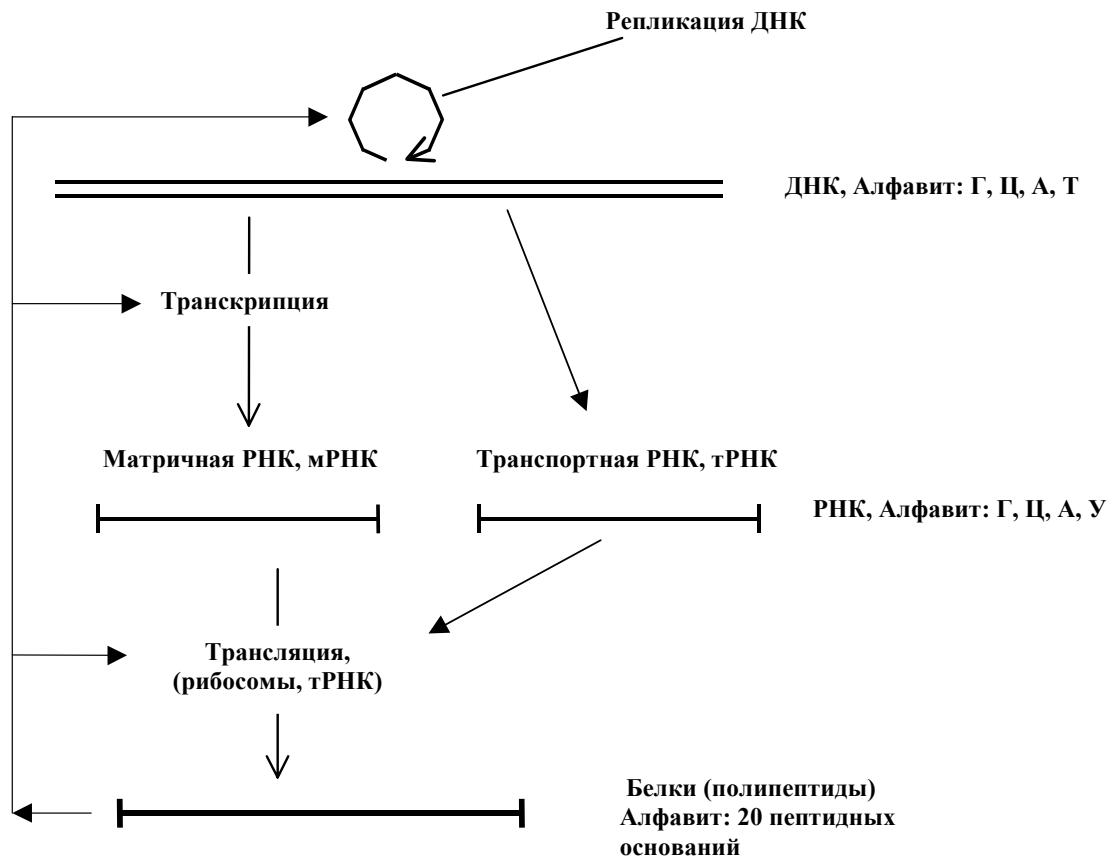


Рис. 2. Схема самовоспроизводящейся молекулярно-генетической системы живой клетки

3.1. Модель квазивидов – простейшая модель эволюции информационных последовательностей

Содержательно модель квазивидов описывает возможный процесс того, как могли бы эволюционно возникнуть самые простейшие макромолекулы, кодирующие наследственную информацию. В модели рассматривается эволюция популяции цепочек РНК. При наличии притока энергии в виде богатых энергией молекул типа АТФ цепочки РНК могут реплицироваться, т.е. размножаться (путем саморепликации или репликации с помощью простейших ферментных систем). При размножении происходит копирование наследуемой информации. Ошибки в процессе копирования приводят к мутациям РНК. Эволюция такой популяции приводит к отбору *квазивида* – такого распределения цепочек РНК, в которое входит как "наилучшая РНК" (размножающаяся с максимальной скоростью), так и близкие к ней цепочки, отличающиеся от этой наилучшей мутационными заменами.

С формальной точки зрения в модели квазивидов рассматривается эволюция информационных последовательностей $S = (S_1, S_2, \dots, S_N)$, символы которых S_i принимают небольшое число ν дискретных значений данного алфавита (для полинуклеотидных последовательностей $\nu = 4$). Рассматривается популяция последовательностей S_1, S_2, \dots, S_n и анализируется дарвиновская эволюция этой популяции. В простейшем случае предполагается, что имеется оптимальная последовательность S_m , а приспособленность произвольной особи S экспоненциально уменьшается с ростом расстояния по Хеммингу $\rho(S, S_m)$ между S и S_m (числа несовпадающих

символов в соответствующих позициях этих последовательностей):

$$f(\mathbf{S}) = \exp[-\beta\rho(\mathbf{S}, \mathbf{S}_m)],$$

где β – параметр интенсивности отбора.

Формальная схема эволюции модели квазивидов представлена в Таблице 1.

Таблица 1. Схема эволюции в модели квазивидов

Шаг 0. <i>Формирование начальной популяции</i> $\{\mathbf{S}_k(0)\}$. Для каждого $k = 1, \dots, n$, и для каждого $i = 1, \dots, N$, выбираем случайно символ S_{ki} , полагая его равным произвольному символу данного алфавита.
Шаг 1. <i>Отбор</i>
Подшаг 1.1. <i>Расчет приспособленностей</i> . Для каждого $k = 1, \dots, n$, вычисляем величину $f(\mathbf{S}_k)$.
Подшаг 1.2. <i>Формирование новой популяции</i> $\{\mathbf{S}_k(t+1)\}$. Отбор n особей в новую популяцию $\{\mathbf{S}_k(t+1)\}$ с вероятностями, пропорциональными $f(\mathbf{S}_k)$.
Шаг 2. <i>Мутации особей в новой популяции</i> . Для каждого $k = 1, \dots, n$, для каждого $i = 1, \dots, N$, заменяем $S_{ki}(t+1)$ на произвольный символ алфавита с вероятностью P . Параметр P характеризует интенсивность мутаций.
<i>Организация последовательности поколений</i> . Повторяем шаги 1, 2 для $t = 0, 1, 2, \dots$

Поясним подшаг 1.2. Формирование новой популяции организуется следующим образом. Представим, что у нас есть рулетка. Для каждого поколения отмечаем на рулетке n секторов, долю k -го сектора (отнесенную ко всей площади круга) полагаем равной $q_k = f_k / [\sum_i f_i]^{-1}$, $f_k = f(\mathbf{S}_k)$. Далее n раз крутим рулетку, каждый раз определяем номер сектора, на котором останавливается стрелка, и соответствующую этому номеру особь выбираем в популяцию следующего поколения. Таким образом, в следующее поколение будут отобраны n особей. При этом для каждого вращения рулетки вероятность k -й особи попасть в следующее поколение пропорциональна ее приспособленности f_k . Некоторые особи будут отобраны в новое поколение несколько раз – это можно интерпретировать как появление в новой популяции нескольких потомков данной особи.

Модель квазивидов исследовали многие авторы. Различают детерминистический и стохастический случаи. В детерминистический случае численность популяции очень велика ($n \gg v^N$), и эволюционная динамика популяции может быть описана системой обыкновенных дифференциальных уравнений [9,10].

В противоположном случае и более интересном случае ($v^N \gg n$) эволюционный процесс существенно стохастический, и для характеристики основных особенностей эволюции целесообразно использовать разумные количественные оценки и методы компьютерного моделирования. Для стохастического случая модель квазивидов исследовалась в работах [11,12]. Были получены оценки скорости сходимости эволюционного процесса и общего количества особей, участвующих в эволюционном процессе. Показано, что достаточно разумных предположениях ($n, N \gg 1$, $n \sim N$, $P \sim N^{-1}$, $\beta \sim 1$, $v = 2$), эволюционный процесс сходится за $T \sim N$ поколений, при этом общее число особей, участвующих в эволюционном процессе, составляет

порядка N^2 . Также было показано, что эволюционный процесс, как алгоритм оптимизации в определенном смысле "субоптимален": он не обеспечивает максимальную скорость поиска, тем не менее, он намного быстрее случайного перебора.

Отметим, что модель квазивидов может рассматриваться как "каноническая" схема эволюции информационных последовательностей.

3.2. Модель гиперциклов или "Цепочки РНК кооперируются"

Модель квазивидов изучает эволюцию информационных последовательностей. Одна из содержательных интерпретаций этой модели – эволюция полинуклеотидных цепочек РНК. Такая эволюция могла иметь место на самых начальных этапах происхождения жизни. Цепочки РНК могли самореплицироваться, однако точность копирования примитивных полинуклеотидов была мала, поэтому длина таких цепочек могла быть только небольшой.

В конце 1970-х годов М. Эйген и П. Шустер предложили модель гиперциклов [9,10]. В гиперцикле к цепочкам РНК добавляются цепочки аминокислот – белки, которые выполняют определенные каталитические функции и вместе с цепочками РНК формируют целостную систему кооперативно взаимодействующих макромолекул. Образно говоря, в гиперцикле цепочки РНК кооперируются, но не сами, а с привлечением примитивных полипептидных ферментов. Модель гиперциклов интерпретирует гипотетическую стадию эволюции, которая могла следовать за квазивидами. Схема гиперцикла представлена на рис. 3.

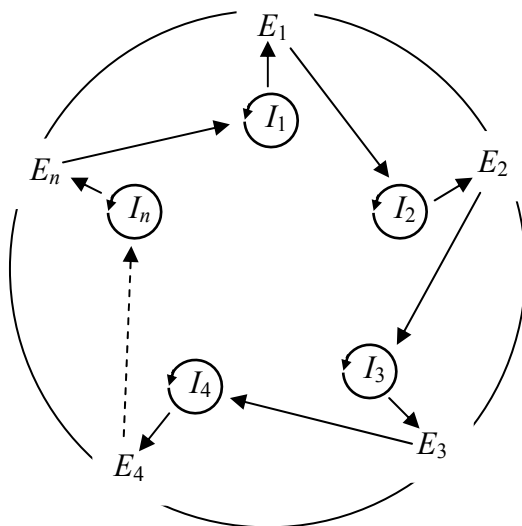


Рис. 3. Структура гиперцикла. I_i – РНК матрицы, E_i – ферменты репликации ($i = 1, 2, \dots, n$)

Гиперцикл – это самовоспроизводящаяся макромолекулярная система, в которой РНК и ферменты кооперируются следующим образом: имеются РНК-матрицы (I_i); i -я РНК кодирует i -й фермент E_i ($i = 1, 2, \dots, n$); ферменты циклически катализируют репликацию РНК, а именно, E_1 способствует репликации I_2 , E_2 способствует репликации I_3 , ..., E_n способствует репликации I_1 . Кроме того, упомянутые макромолекулы кооперативно обеспечивают примитивную трансляцию, так что информация, закодированная РНК-последовательностями, транслируется в структуру ферментов, аналогично обычному механизму трансляции в биологических клетках. Циклическая организация гиперцикла обеспечивает его структурную стабильность.

Схема гиперцикла еще далека от схемы самовоспроизводящейся молекулярно-генетической системы живой клетки (рис. 2), тем не менее, гиперцикл – это определенный шаг к живой клетке по сравнению с квазивидами, включающий кооперацию между полинуклеотидами и белками.

Математическое описание гиперциклов было разработано рядом авторов, при этом динамика концентраций макромолекул описывается системой обыкновенных нелинейных дифференциальных уравнений [10]. Было показано, что для того, чтобы в процессе эволюции отбирались гиперциклы, имеющие лучшие каталитические способности входящих в гиперциклы ферментов, гиперциклы разных видов должны быть размещены по компартментам (таким как коарцеватные капли Опарина).

Итак, модель гиперциклов характеризует систему кооперативно взаимодействующих макромолекул (полинуклеотидных матриц и кодируемых ими ферментов). Такая система может рассматриваться как следующий эволюционный шаг по сравнению с квазивидами.

Еще более близкие «конструкции» к самовоспроизводящейся молекулярно-генетической системе живой клетки схемы – «сайзеры» – были предложены В.А. Ратнером и В.В. Шапиным [13].

3.3. Модель сайзеров

Сайзер включает в себя (рис. 4): полинуклеотидную матрицу I , фермент репликации E_1 , фермент трансляции E_2 и другие возможные белки E_3, \dots, E_n .

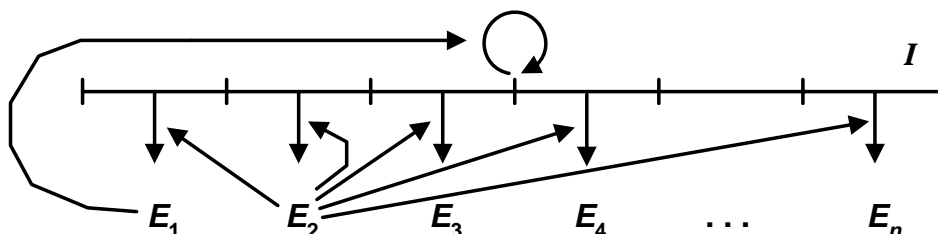


Рис. 4. Общая схема сайзера. I – полинуклеотидная матрица, E_i – ферменты/белки. Круговая стрелка над матрицей иллюстрирует процесс репликации. Стрелки, направленные вертикально вниз, иллюстрируют процессы трансляции. Стрелки от ферментов E_1 и E_2 поясняют, что эти ферменты катализируют процессы репликации и трансляции, соответственно

Полинуклеотидная матрица I кодирует белки, фермент репликации E_1 обеспечивает репликацию матрицы I , фермент трансляции E_2 обеспечивает синтез белков в соответствии с информацией, хранящейся в матрице I .

Термин «сайзер» был предложен В.А. Ратнером и происходит от System of self-reproduction (syser).

Сайзеры имеют определенное сходство с биологическими клетками (рис. 2). Матрица I хранит "генетическую" информацию; ферменты E_1 и E_2 представляют собой простые аналоги систем репликации и трансляции биологических клеток.

Как и для гиперциклов, для сайзеров есть математическое описание на основе системы обыкновенных нелинейных дифференциальных уравнений, которое демонстрирует, что при помещении сайзеров в коарцеваты происходит отбор сайзеров, обладающих наибольшей максимальной суммарной скоростью синтеза макромолекул [13-15]. Выживают сайзеры, обладающие максимальной "производительностью труда".

Модель сайзеров так же, как и модель гиперциклов может рассматриваться как разумная модель возникновения кооперативных макромолекулярных самовоспроизводящихся систем, содержащих как полинуклеотиды, так и ферменты.

Модель сайзеров ближе к простейшим биологическим организмам, чем гиперциклы и позволяет анализировать возможные этапы эволюции от простейших мини-сайзеров до макромолекулярных

систем, обладающих свойствами реальных живых организмов. Например, модель адаптивного сайзера [16] иллюстрирует процесс эволюционного возникновения простейшей системы управления. Подчеркнем, что схема адаптивного сайзера может рассматриваться как модель самой простейшей адаптивной системы, которая могла появиться еще на стадиях предбиологической эволюции.

Были представлены наиболее четкие модели эволюции предбиологических кибернетических систем, предложенные различными авторами в 1970-1990 годах. Рассмотренные модели позволяют представить возможные сценарии возникновения самовоспроизводящейся молекулярно-генетической системы живой клетки (рис. 2). Гипотетический процесс эволюции предбиологических кибернетических систем иллюстрируется рис. 5.

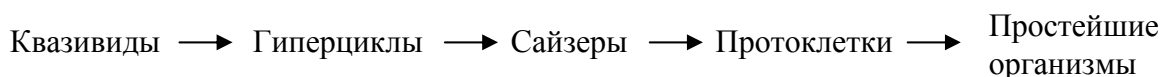


Рис. 5. Схема, иллюстрирующая гипотетический процесс эволюции самовоспроизводящихся макромолекулярных систем

Итак, очерченные модели – это только элементы и только возможного сценария эволюции предбиологических кибернетических систем. Тем не менее, эти модели позволяют четко и на серьезном математическом уровне представить гипотетические этапы предбиологической эволюции. Более подробное описание моделей предбиологических молекулярно-генетических информационных систем содержится в [14,15].

В следующем разделе излагается проект «Мозг анимата» [17,18]. Этот проект нацелен на формирование общей «платформы» для систематического построения моделей адаптивного поведения. Проект основан на теории функциональных систем П.К. Анохина.

4. ПРОЕКТ «МОЗГ АНИМАТА»

Теория функциональных систем была предложена и развита в 1930-70 годах известным советским нейрофизиологом П.К. Анохиным [19]. Функциональная система по П.К. Анохину – схема управления, нацеленная на достижение полезных для организма результатов. Далее излагается основанный на теории функциональных систем проект «Мозг анимата».

4.1. Общая архитектура «Мозга анимата»

Предполагается, что система управления аниматом имеет иерархическую архитектуру. Базовым элементом системы управления является отдельная функциональная система.

Верхний уровень соответствует основным потребностям организма: питания, размножения, безопасности, накопления знаний. Более низкие уровни системы управления соответствуют тактическим целям поведения. Блоки всех этих уровней (включая первый) реализуются с помощью функциональных систем. Управление с верхних уровней может передаваться на нижние уровни (от «суперсистем» к «субсистемам») и возвращаться назад. Предполагается, что система управления аниматом функционирует в дискретном времени. Также предполагается, что каждый такт времени активна только одна функциональная система.

Рассматривается простая формализация функциональной системы на основе нейросетевых адаптивных критиков. Функциональная система моделирует следующие важные особенности ее биологического прототипа: 1) прогноз результата действия, 2) сравнение прогноза и результата, 3) коррекцию прогноза путем обучения в соответствующих нейронных сетях, 4) принятие решения.

Рассматриваемая схема адаптивного критика состоит из двух блоков: Модель и Критик. Предполагается, что Модель и Критик – многослойные персептроны, и что производные по весам нейронных сетей этих блоков могут быть вычислены обычным методом обратного распространения ошибки [20]. Цель адаптивного критика – максимизировать функцию суммарной награды $U(t)$ [21]:

$$U(t_j) = \sum_{k=0}^{\infty} \gamma^k r(t_j + k) \quad (1)$$

где $r(t_j)$ – текущее подкрепление (награда, $r(t_j) > 0$ или наказание, $r(t_j) < 0$), полученное адаптивным критиком в данный момент времени t_j , γ – дисконтный фактор, $0 < \gamma < 1$.

Для развития проекта важно проверить, как функционируют простые схемы адаптивных критиков в конкретных моделях. В следующем разделе излагаются результаты исследования такой модели.

4.2. Модель эволюции популяции самообучающихся агентов на базе нейросетевых адаптивных критиков [22]

В данном разделе исследуется модель эволюции популяции самообучающихся автономных агентов и анализируется взаимодействие между обучением и эволюцией (рис.6). Модель отрабатывается на примере агента-брокера. Этот пример используется только для определенности, аналогично можно рассматривать 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.

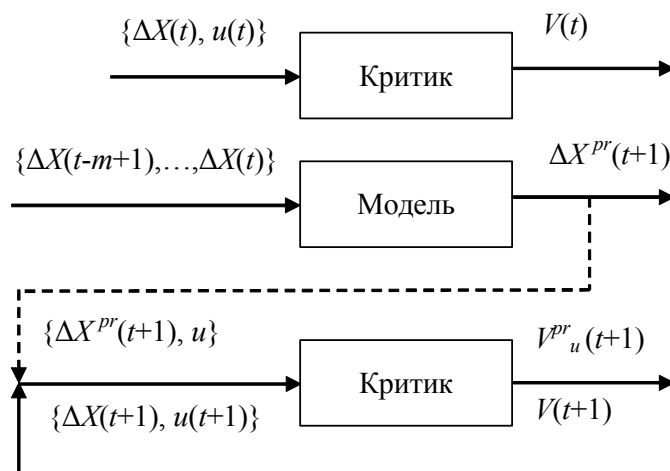


Рис. 6. Схема системы управления агента. НС Критика показана для двух последовательных тактов времени. Модель предназначена для прогнозирования изменения курса временного ряда. Критик предназначен для оценки качества ситуаций $V(S)$ для текущей ситуации $S(t) = \{\Delta X(t), u(t)\}$, для ситуации в следующий такт времени $S(t+1)$ и для предсказываемых ситуаций для обоих возможных действий $S_u^{pr}(t+1) = \{\Delta X^{pr}(t+1), u\}$, $u = 0$ либо $u = 1$

Схема агента. Рассматривается модель агента-брокера, который имеет ресурсы двух типов: деньги и акции; сумма этих ресурсов составляет капитал агента $C(t)$; доля акций в капитале равна $u(t)$. Внешняя среда определяется временным рядом $X(t)$; $t = 0, 1, 2, \dots$; $X(t)$ – курс акций на бирже в момент времени t . Агент стремится увеличить свой капитал $C(t)$, изменяя значение $u(t)$. Динамика капитала определяется выражением [23]:

$$C(t+1) = C(t) [1 + u(t+1) \Delta X(t+1) / X(t)], \quad (2)$$

где $\Delta X(t+1) = X(t+1) - X(t)$ – текущее изменение курса акций. Используется логарифмическая шкала для ресурса агента, $R(t) = \log C(t)$ [24]. Текущее подкрепление агента $r(t) = R(t+1) - R(t)$ равно:

$$r(t) = \log [1 + u(t+1) \Delta X(t+1) / X(t)]. \quad (3)$$

Для простоты предполагается, что переменная u может принимать только два значения $u = 0$ (весь капитал в деньгах) или $u = 1$ (весь капитал в акциях).

Алгоритм обучения. Система управления агента представляет собой простой адаптивный критик, состоящий из двух нейронных сетей (НС): Модель и Критик (рис. 6).

В предположении $\Delta X(t) \ll X(t)$ считаем, что ситуация $\mathbf{S}(t)$, характеризующая состояние агента, зависит только от двух величин, $\Delta X(t)$ и $u(t)$: $\mathbf{S}(t) = \{\Delta X(t), u(t)\}$.

Модель предназначена для прогнозирования изменения курса временного ряда. На вход Модели подается m предыдущих значений изменения курса $\Delta X(t-m+1), \dots, \Delta X(t)$, на выходе формируется прогноз изменения курса в следующий такт времени $\Delta X^{pr}(t+1)$. Модель представляет собой двухслойную НС, работа которой описывается формулами:

$$\mathbf{x}^M = \{\Delta X(t-m+1), \dots, \Delta X(t)\}, \quad y_j^M = \text{th}(\sum_i w_{ij}^M x_i^M), \quad \Delta X^{pr}(t+1) = \sum_j v_j^M y_j^M,$$

где \mathbf{x}^M – входной вектор, \mathbf{y}^M – вектор выходов нейронов скрытого слоя, w_{ij}^M и v_j^M – веса синапсов НС.

Критик предназначен для оценки качества ситуаций $V(\mathbf{S})$ по формуле (1), а именно, оценки функции полезности $U(t)$ для агента, находящегося в рассматриваемой ситуации \mathbf{S} . Критик представляет собой двухслойную НС, работа которой описывается формулами:

$$\mathbf{x}^C = \mathbf{S}(t) = \{\Delta X(t), u(t)\}, \quad y_j^C = \text{th}(\sum_i w_{ij}^C x_i^C), \quad V(t) = V(\mathbf{S}(t)) = \sum_j v_j^C y_j^C,$$

где \mathbf{x}^C – входной вектор, \mathbf{y}^C – вектор выходов нейронов скрытого слоя, w_{ij}^C и v_j^C – веса синапсов НС.

Каждый момент времени t выполняются следующие операции:

- 1) Модель предсказывает следующее изменение временного ряда $\Delta X^{pr}(t+1)$.
- 2) Критик оценивает величину V для текущей ситуации $V(t) = V(\mathbf{S}(t))$ и для предсказываемых ситуаций для обоих возможных действий $V^{pr}_u(t+1) = V(\mathbf{S}^{pr}_u(t+1))$, где $\mathbf{S}^{pr}_u(t+1) = \{\Delta X^{pr}(t+1), u\}$, $u = 0$ либо $u = 1$.
- 3) Применяется ε -жадное правило [21]: действие, соответствующее максимальному значению $V^{pr}_u(t+1)$ выбирается с вероятностью $1 - \varepsilon$, и альтернативное действие выбирается с вероятностью ε ($0 < \varepsilon \ll 1$). Выбор действия есть выбор величины $u(t+1)$: перевести весь капитал в деньги, $u(t+1) = 0$; либо в акции, $u(t+1) = 1$.
- 4) Выбранное действие $u(t+1)$ выполняется. Происходит переход к моменту времени $t+1$. Наблюдаемое значение $\Delta X(t+1)$ сравнивается с предсказанием $\Delta X^{pr}(t+1)$. Веса НС Модели подстраиваются так, чтобы минимизировать ошибку предсказания методом обратного распространения ошибки. Скорость обучения Модели равна $\alpha_M > 0$.
- 5) Критик подсчитывает $V(t+1) = V(\mathbf{S}(t+1))$; $\mathbf{S}(t+1) = \{\Delta X(t+1), u(t+1)\}$. Рассчитывается ошибка временной разности:

$$\delta(t) = r(t) + \gamma V(t+1) - V(t) . \quad (4)$$

6) Веса НС Критика подстраиваются так, чтобы минимизировать величину $\delta(t)$, это обучение осуществляется градиентным методом, аналогично методу обратного распространения ошибки. Скорость обучения Критика равна $\alpha_C > 0$.

Схема эволюции. Рассматривается эволюционирующая популяция, состоящая из n агентов. Каждый агент имеет ресурс $R(t)$, который изменяется в соответствии с подкреплениями агента: $R(t+1) = R(t) + r(t)$, где $r(t)$ определено в (3)

Эволюция происходит в течение ряда поколений, $n_g=1,2,\dots$. Продолжительность каждого поколения n_g равна T тактов времени (T – длительность жизни агента). В начале каждого поколения ресурс каждого агента равен нулю, т.е., $R(T(n_g-1)+1) = 0$.

Начальные веса синапсов обеих НС (Модели и Критика) формируют геном агента $\mathbf{G} = \{\mathbf{W}_{M0}, \mathbf{W}_{C0}\}$. Геном \mathbf{G} задается в момент рождения агента и не меняется в течение его жизни. В противоположность этому текущие веса синапсов НС \mathbf{W}_M и \mathbf{W}_C подстраиваются в течение жизни агента путем обучения.

В конце каждого поколения определяется агент, имеющий максимальный ресурс $R_{max}(n_g)$ (лучший агент поколения n_g). Этот лучший агент порождает n потомков, которые составляют новое (n_g+1) -е поколение. Геномы потомков \mathbf{G} отличаются от генома родителя небольшими мутациями. Более конкретно, предполагается, что в начале каждого нового (n_g+1) -го поколения для каждого агента его геном формируется следующим образом $G_i(n_g+1) = G_{best, i}(n_g) + \text{rand}_i$, $\mathbf{W}_0(n_g+1) = \mathbf{G}(n_g+1)$, где $\mathbf{G}_{best}(n_g)$ – геном лучшего агента предыдущего n_g -го поколения и rand_i – это нормально распределенная случайная величина с нулевым средним и стандартным отклонением P_{mut} (интенсивность мутаций), которая добавляется к каждому весу.

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 $\mathbf{W} = \mathbf{W}_0 = \mathbf{G}$.

4.3. Результаты моделирования

Общие особенности адаптивного поиска. Изложенная модель была реализована в виде компьютерной программы. В компьютерных экспериментах использовалось два варианта временного ряда:

1) синусоида:

$$X(t) = 0.5(1 + \sin(2\pi t/20)) + 1 , \quad (5)$$

2) стохастический временной ряд [23]:

$$X(t) = \exp(p(t)/1200), \quad p(t) = p(t-1) + \beta(t-1) + k_1\lambda(t), \quad \beta(t) = k_2\beta(t-1) + \mu(t), \quad (6)$$

где $\lambda(t)$ и $\mu(t)$ – два нормальных процесса с нулевым средним и единичной дисперсией, $k_1 = 0.3$; $k_2 = 0.9$.

Некоторые параметры модели имели одно и то же значение для всех экспериментов: дисконтный фактор $\gamma = 0.9$; количество входов НС Модели $m = 10$; количество нейронов в скрытых слоях НС Модели и Критика $N_{hM} = N_{hC} = 10$; скорость обучения Модели и Критика $\alpha_M = \alpha_C = 0.01$; параметр ε -жадного правила $\varepsilon = 0.05$; интенсивность мутаций $P_{mut} = 0.1$. Остальные параметры

(продолжительность поколения T и численность популяции n) принимали разные значения в разных экспериментах.

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. рассматривается эволюционирующая популяция без обучения;
- Случай LE (эволюция + обучение), т.е. полная модель, изложенная выше.

Было проведено сравнение 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 поколения. В случае L (чистое обучение) рассматривался только один агент, ресурс которого для удобства сравнения со случаями E и LE обнулялся каждые $T = 200$ тактов времени: $R(T(n_g-1)+1) = 0$. 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 синусоиды (5) показаны на рис. 7. Чтобы исключить уменьшение значения $R_{max}(n_g)$ из-за случайного выбора действий при применении ε -жадного правила для случаев LE и L, полагалось $\varepsilon = 0$ после $n_g = 100$ для случая LE и после $n_g = 2000$ для случая L. Результаты усреднены по 1000 экспериментам; $n = 10$, $T = 200$.

Рис. 7 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.

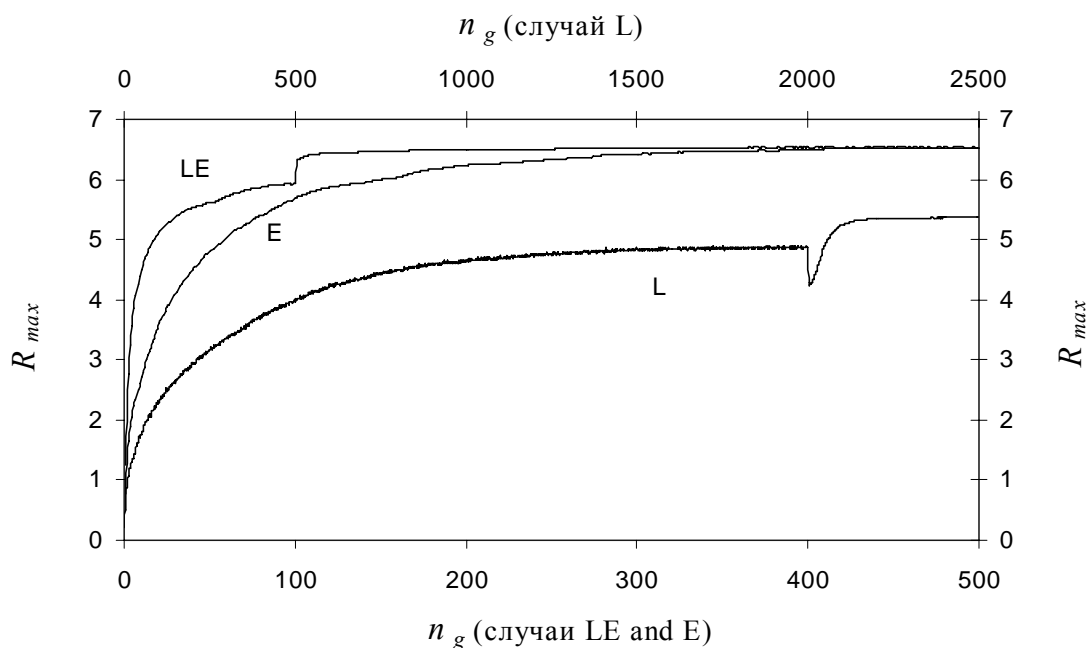


Рис. 7. Зависимости $R_{max}(n_g)$. Кривая LE соответствует случаю эволюции, объединенной с обучением, кривая E – случаю чистой эволюции, кривая L – случаю чистого обучения. Временная шкала для случаев LE и E (номер поколения n_g) представлена снизу, для случая L (индекс n_g) – сверху. Моделирование проведено для синусоиды, кривые усреднены по 1000 экспериментам; $n = 10$, $T = 200$.

Из (3) следует, что существует оптимальная стратегия поведения агента: вкладывать весь капитал в акции ($u(t+1) = 1$) при прогнозе роста курса ($\Delta X(t+1) > 0$), вкладывать весь капитал в деньги ($u(t+1) = 0$) при прогнозе падения курса ($\Delta X(t+1) < 0$). Анализ экспериментов, представленных на рис. 7, показал, что в случаях LE (обучение + эволюция) и E (чистая эволюция) такая оптимальная стратегия 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 следующей «субоптимальной» стратегии поведения: агент держит капитал в акциях при росте и при слабом падении курса и переводит капитал в деньги при сильном падении курса. Та же тенденция к явному предпочтению вкладывать капитал в акции при чистом обучении наблюдалась и для экспериментов на стохастическом ряде (6).

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.

Взаимодействие между обучением и эволюцией. Эффект Болдуина. Если длительность поколения T была достаточно большой (1000 и более тактов времени), то для случая LE часто наблюдалось сильное влияние обучения на эволюционный процесс. 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].

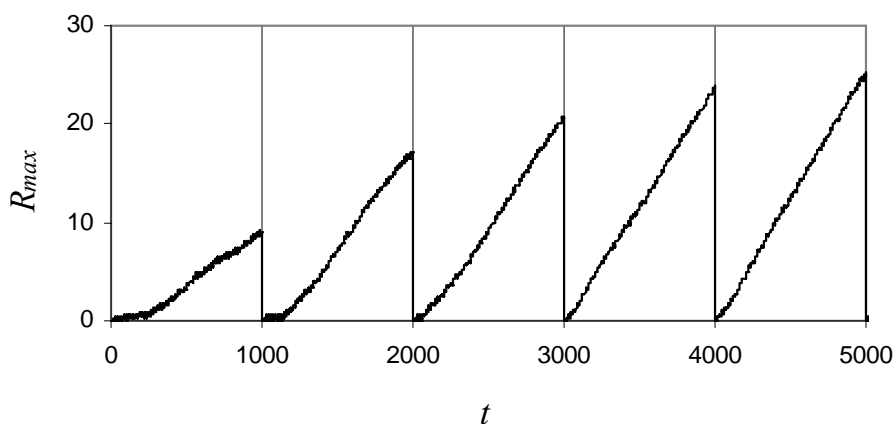


Рис. 8. Зависимость ресурса лучшего в популяции агента R_{max} от времени t для первых пяти поколений. Случай LE (эволюция, объединенная с обучением); размер популяции $n = 10$, длительность поколения $T = 1000$. Моменты смены поколений показаны вертикальными линиями. Для первых двух поколений есть явная задержка в 100-300 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.

Особенности предсказания Модели. Практика не есть критерий истины. Система управления каждого агента включает в себя нейронную сеть Модели, предназначенную для предсказания изменения значения $\Delta X(t+1)$ временного ряда в следующий такт времени $t+1$. Нейронная сеть Модели может давать неверные предсказания, однако агент, тем не менее, может использовать эти предсказания для принятия верных решений – практика не есть критерий истины. Например, рис. 9 показывает предсказываемые изменения $\Delta X^{pr}(t+1)$ и реальные изменения $\Delta X(t+1)$ стохастического временного ряда в случае чистой эволюции (случай E). Предсказания нейронной

сети Модели достаточно хорошо совпадают по форме с кривой $\Delta X(t)$. Однако, предсказанные значения $\Delta X^{pr}(t+1)$ отличаются примерно в 25 раз от значений $\Delta X(t+1)$.

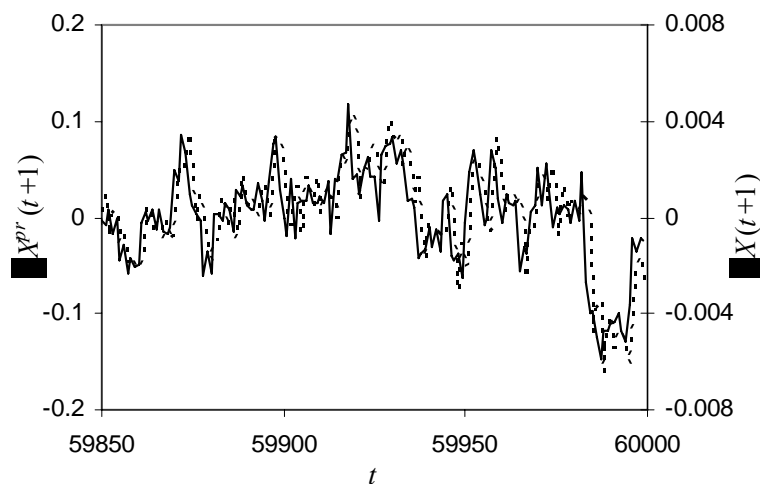


Рис. 9. Предсказываемые $\Delta X^{pr}(t+1)$ (пунктирная линия) и реальные изменения $\Delta X(t+1)$ (сплошная линия) стохастического временного ряда. Случай чистой эволюции. $n = 10$, $T = 200$. Хотя обе кривые имеют сходную форму, по величине $\Delta X^{pr}(t+1)$ и $\Delta X(t+1)$ радикально различаются.

На рис. 10 приведен другой пример особенностей предсказания нейронной сети Модели в случае LE (эволюция, объединенная с обучением). Этот пример показывает, что предсказания нейронной сети Модели могут отличаться от реальных данных не только масштабом, но и знаком.

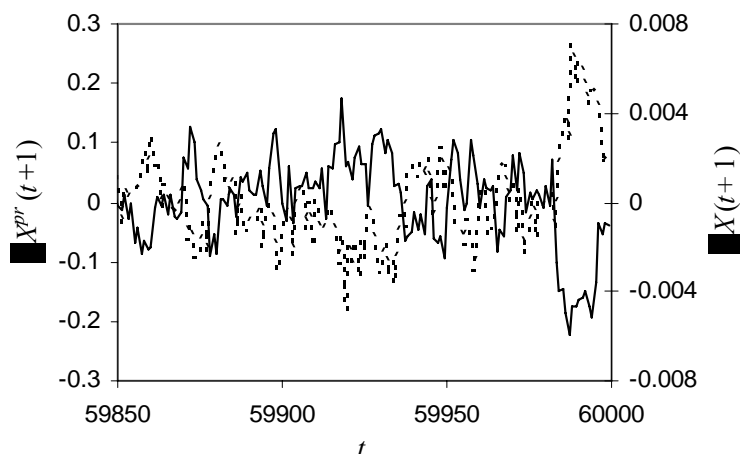


Рис. 10. Предсказываемые $\Delta X^{pr}(t+1)$ (пунктирная линия) и реальные изменения $\Delta X(t+1)$ (сплошная линия) стохастического временного ряда. Случай эволюции, объединенной с обучением. $n = 10$, $T = 200$. Кривые $\Delta X^{pr}(t+1)$ и $\Delta X(t+1)$ различаются как по величине, так и знаком.

По-видимому, наблюдаемое увеличение значений ΔX^{pr} нейронной сетью Модели полезно для работы нейронной сети Критика, так как реальные значения $\Delta X(t+1)$ слишком малы (порядка 0.001). Таким образом, нейронная сеть Модели может не только предсказывать значения $\Delta X^{pr}(t+1)$, но также осуществлять полезные преобразования этих значений.

Эти особенности работы нейронной сети Модели обусловлены доминирующей ролью эволюции над обучением при оптимизации системы управления агентов. На самом деле, из-за малой длительности поколений ($T = 200$) в проведенном моделировании, веса синапсов нейронных сетей изменяются большей частью за счет эволюционных мутаций. Такой процесс делает предпочтительными такие системы управления, которые устойчивы в эволюционном смысле.

Кроме того, важно подчеркнуть, что задача, которую «решает» эволюция в рассматриваемой модели, значительно проще, чем та задача, которую «решает» обучение. Эволюции достаточно обеспечить выбор из двух действий, приводящий к награде. А схема обучения предусматривает довольно сложную процедуру прогноза ситуации S , оценки качества прогнозируемых ситуаций, итеративного формирования оценок качества ситуаций $V(S)$ и выбора действия на основе этих оценок. Эволюция идет к нужному результату более прямым путем, а так как задача агентов проста, то эволюция в определенной степени «задавливает» довольно сложный механизм обучения. Тем не менее, есть определенная синергия во взаимодействии обучения и эволюции: обучение ускоряет процесс поиска оптимальной стратегии поведения.

Эволюция самообучающихся агентов и проект «Мозг анимата». Опыт работы с исследованной моделью показывает важность вопроса о том, какие системы управления автономных агентов являются эволюционно устойчивыми. Под эволюционной устойчивостью подразумеваем свойство фенотипа становиться практически нечувствительным к мутациям соответствующего ему генотипа.

Эволюционная неустойчивость работы рассмотренной схемы адаптивного критика показывает, что необходима определенная осторожность в выборе базовой модели функциональной системы (ФС) для проекта «Мозг анимата». А, именно, хотя схема адаптивного критика, используемая в качестве основы ФС, и моделирует прогноз результата действия и принятие решения на основе этого прогноза, что существенно для теории функциональных систем, тем не менее, имеет смысл рассмотреть и другие возможности для базовой модели ФС. Например, в [27] начата проработка новой версии «Мозга анимата» на основе использования понятия хеббовских ансамблей [28]. Однако проведенный в [27] анализ показывает, что необходимость согласования всех элементов системы управления аниматом накладывает довольно серьезные ограничения на ее архитектуру, и важны дальнейшие исследования по данному перспективному проекту, в том числе и на основе опыта разработки архитектуры на базе нейросетевых адаптивных критиков.

ЗАКЛЮЧЕНИЕ

Итак, рассмотрены подходы к исследованию биологических информационных систем и процессов формирования интеллектуального адаптивного поведения в биологических организмах. Рассмотрение показывает, что имеется широкий фронт актуальной научной работы, включающий в себя 1) разработку моделей адаптивного поведения разного эволюционного уровня, 2) исследование принципов и моделей организации естественного интеллекта, 3) моделирование когнитивной эволюции, 4) сопоставление моделей с биологическими прототипами.

Работы в этой области имеют важное мировоззренческое значение, однако этим исследованиям уделяется несоизмеримо малое внимание по сравнению с их значимостью. Все это указывает на необходимость формирования серьезной академической научной программы «Эволюция, Мозг, Интеллект», в рамках которой велись бы такие работы.

СПИСОК ЛИТЕРАТУРЫ

1. Meyer J.-A., Wilson S. W. (Eds). From animals to animats // Proceedings of the First International Conference on Simulation of Adaptive Behavior. The MIT Press: Cambridge, Massachusetts, London, England. 1990.
2. От моделей поведения к искусственному интеллекту. Серия «Науки об искусственном» (под ред. Редько В.Г.). М.: УРСС, 2006.
3. Непомнящих В.А. Поиск общих принципов адаптивного поведения живых организмов и аниматов // Новости искусственного интеллекта. 2002. N. 2. С. 48-53.
4. 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 Behaviour. LNAI. 2006. Volume 4095. Berlin, Germany: Springer Verlag.

5. Вайнцвайг М.Н., Полякова М.П. О моделировании мышления // В книге [2]. С. 280-286.
6. J.L. Krichmar, A.K. Seth, D.A. Nitz, J.G. Fleischer, G.M. Edelman. Spatial navigation and causal analysis in a brain-based device modeling cortical-hippocampal interactions // *Neuroinformatics*, 2005, Vol.3, N. 3, P. 197-221. See also: <http://www.nsi.edu/nomad/pubs.html>
http://www.nsi.edu/nomad/krichmar_neuroinf_2005.pdf
7. Marocco D., Nolfi S. Origins of communication in evolving robots // В книге [3]. P. 789-803. См. также:
http://laral.istc.cnr.it/nolfi/papers/marocco_sab9.pdf
8. Редько В.Г., Непомнящих В.А., Попов Е.Е. Бионическая модель адаптивного поискового поведения // *Изв. АН. Теория и системы управления* (направлено в печать)
9. Эйген М. Самоорганизация материи и эволюция биологических макромолекул. М.: Мир, 1973. 216 с.
10. Эйген М., Шустер П. Гиперцикл. Принципы самоорганизации макромолекул. М.: Мир, 1982. 270 с.
11. Редько В.Г. Оценка скорости эволюции в моделях Эйдена и Куна // *Биофизика*.1986. Т. 31. N.3. С. 511-516.
12. В.Г.Редько, Ю.Р.Цой. Оценка эффективности эволюционных алгоритмов // *Доклады АН*, 2005. Т. 404. N. 3. С. 312-315.
13. Ратнер В.А., Шамин В.В. Сайзеры: моделирование фундаментальных особенностей молекулярно-биологической организации. Соответствие общих свойств и конструктивных особенностей коллективов макромолекул // *Журн. общ. биологии*. 1983. Т.44. N.1. С. 51-61.
14. Редько В.Г. Эволюционная кибернетика. М.: Наука, 2001, 156 с.
15. Редько В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики. Серия «Синергетика: от прошлого к будущему». М.: УРСС, 2005. 224 с.
16. Редько В.Г. Адаптивный сайзер // *Биофизика*. 1990. Т.35. Вып.6. С.1007-1011.
17. 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.
18. 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, P.1787-1792.
19. Анохин П.К. Системные механизмы высшей нервной деятельности. М.: Наука, 1979. 453 с.
20. Rumelhart D.E., Hinton G.E., Williams R.G. Learning representation by back-propagating error // *Nature*. 1986. Vol.323. N.6088. P. 533-536.
21. Sutton R., Barto A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. See also: <http://www.cs.ualberta.ca/~sutton/book/the-book.html>
22. Red'ko V.G., Mosalov O.P., Prokhorov D.V. A Model of Evolution and Learning // *Neural Networks*, 2005. Vol. 18, N. 5-6, P. 738-745.
23. Prokhorov D., Puskorius G., Feldkamp L. Dynamical neural networks for control // In J. Kolen and S. Kremer (Eds.) *A field guide to dynamical recurrent networks*. NY: IEEE Press, 2001, P. 257-289.
24. Moody J., Wu L., Liao Y., Saffel M. Performance function and reinforcement learning for trading systems and portfolios // *Journal of Forecasting*, 1998, Vol.17, P. 441-470.
25. Baldwin J.M. A new factor in evolution // *American Naturalist*, 1896, Vol. 30, P. 441-451.
26. 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*, Vol.4, N.3, 1996.
27. Red'ko V.G., Anokhin K.V., Burtsev M.S., Manolov A.I., Mosalov O.P., Nepomnyashchikh V.A., Prokhorov D.V. Project "Animat Brain": Designing the Animat Control System on the Basis of the Functional Systems Theory // *The Ninth International Conference on the Simulation of Adaptive Behavior (SAB'06)*, 25 - 29 September 2006, CNR, Roma, Italy, *Third Workshop on Anticipatory Behavior in Adaptive Learning Systems (ABiALS 2006)*, Proceedings.
28. Hebb D.O. *The Organization of Behavior. A Neuropsychological Theory*. N.Y.: Wiley & Sons, 1949. 355 p.