

## **МОДЕЛИ АДАПТИВНОГО ПОВЕДЕНИЯ – ЕСТЕСТВЕННОНАУЧНЫЙ ПОДХОД К РАЗВИТИЮ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ\***

В.Г. Редько

Институт оптико-нейронных технологий РАН

E-mail: redko@iont.ru

В статье представлен аналитический обзор направления исследований "Адаптивное поведение", цель которого – исследовать архитектуры и принципы функционирования, позволяющие аниматам (модельным организмам) приспосабливаться к переменной внешней среде. Особое внимание уделяется работам Массачусетского университета (исследования Р. Саттона и Э. Барто по методу обучения с подкреплением) и работам AnimatLab (Париж) по анализу формирования архитектур управления на основе симбиоза 1) индивидуального обучения, 2) онтогенетического развития анимата, и 3) эволюционной оптимизации. Статья также включает изложение оригинальных моделей эволюционного возникновения целенаправленного адаптивного поведения. На основе материала статьи развивается дискуссия о будущем науки, инициированная Г.Г. Малинецким [1].

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? Какие информационные процессы обеспечивают их работу? Какие архитектуры и принципы функционирования биологических систем управления обеспечивают способность животных приспосабливаться, адаптироваться к постоянно меняющимся условиям во внешней среде? До какой степени исследования биологических адаптивных систем могут способствовать развитию информационных технологий?

В настоящей статье речь пойдет об исследованиях, направленных на изучение этих интригующих проблем, а именно о современных исследованиях в области моделирования адаптивного поведения. Это направление исследований сложилось сравнительно недавно и основная его цель – изучение систем управления адаптивным поведением живого или искусственного организма (робота или компьютерной модели). В 1-м разделе мы охарактеризуем в целом направление исследований "Адаптивное поведение", а также тесно связанное с ним направление "Искусственная жизнь". В разделе 2 характеризуются компьютерные методы, используемые в этих исследованиях, в частности, описывается простой и изящный метод обучения с подкреплением, разработанный Р. Саттоном и Э. Барто (Массачусетский университет). В разделе 3 излагаются подходы к исследованиям адаптивного поведения, разрабатываемые в AnimatLab – одной из ведущих лабораторий в этой области. Далее (раздел 4) обсуждается общий концептуальный подход, который естественно использовать в будущих исследованиях, основанный на теории функциональных систем П.К. Анохина. Затем следует описание конкретных моделей целенаправленного адаптивного поведения, разработанных по инициативе автора (раздел 5). И в заключение (раздел 6) на основе материала статьи развивается дискуссия о будущем науки, начатая на страницах этого журнала Г.Г. Малинецким [1].

---

\* Работа выполнена финансовой поддержке программы Президиума РАН "Интеллектуальные компьютерные системы" (проект 2-45), ОИТВС РАН (проект ОИТВС-01, № 1.8) и РФФИ (проект № 02-07-90197).

## 1. Направления исследований "Адаптивное поведение" и "Искусственная жизнь"

В конце 80-х - начале 90-х годов возникли два интересных, тесно связанных между собой направления исследований: "Искусственная жизнь" (английское название Artificial Life или ALife) [2,3] и "Адаптивное поведение" (Adaptive Behavior) [4].

Первая конференция по **Искусственной жизни** состоялась в 1987 году в Лос Аламосе. Как сказал руководитель этой конференции К. Ленгтон, "основное предположение искусственной жизни состоит в том, что «логическая форма» организма может быть отделена от материальной основы его конструкции". Основной мотивацией исследований Искусственной жизни (ИЖ) служит желание понять и промоделировать формальные принципы организации биологической жизни.

Отметим, хотя лозунг "Искусственная жизнь" был провозглашен в конце 80-х, в действительности идейно близкие модели разрабатывались в 50-70-е годы. Приведем два примера из истории отечественной науки.

В 60-х годах блестящий кибернетик и математик М.Л. Цетлин предложил и исследовал модели автоматов, способных адаптивно приспосабливаться к окружающей среде. Работы М.Л. Цетлина инициировали целое научное направление, получившее название "коллективное поведение автоматов" [5,6].

В 70-х годах под руководством талантливого кибернетика М.М. Бонгарда был предложен весьма нетривиальный проект "Животное", характеризующий адаптивное поведение искусственных организмов [7, 8].

Типичные примеры современных моделей "Искусственной жизни" кратко охарактеризованы в [9].

Первую международную конференцию по Адаптивному поведению организовали Жан-Аркадий Мейер и Стюарт Вильсон в 1990 году в Париже.

Основной подход направления "**Адаптивное поведение**" – конструирование и исследование искусственных (в виде компьютерной программы или робота) "организмов", способных приспосабливаться к внешней среде [4, 10-12]. Эти организмы называются "*аниматами*" (от англ. animal + robot = animat). Часто используют близкий термин "агент", подразумевая под этим термином модельный искусственный организм.

Поведение аниматов имитирует поведение животных. Исследователи направления "Адаптивное поведение" (АП) стараются строить такие модели, которые применимы к описанию поведения *как реального животного, так и искусственного анимата*.

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

*Программа-максимум этого направления – попытаться проанализировать эволюцию когнитивных способностей животных и эволюционное происхождение человеческого интеллекта [13].*

Как и для ИЖ, для исследований АП характерен *синтетический подход*: здесь конструируются архитектуры, обеспечивающие "интеллектуальное" поведение аниматов. Причем это конструирование проводится как бы с точки зрения инженера: исследователь сам "изобретает" архитектуры, подразумевая, конечно, что какие-то подобные структуры, обеспечивающие адаптивное поведение, должны быть у реальных животных.

И так же, как для ИЖ, для направления АП были явные провозвестники этого направления до его "официального провозглашения". Яркий пример – хороший обзор ранних работ по адаптивному поведению, представленный в книге М.Г. Гаазе-Рапопорта, Д.А. Поспелова "От амебы до робота: модели поведения" [8].

Ряд современных моделей АП представлен в обзорах В.А. Непомнящих [11,12].

ИЖ и АП имеют много общего: синтетический подход к конструированию жизнеподобных "организмов", попытка промоделировать формальные законы жизни и систем управления, ориентация на компьютерные и математические модели, использование эволюционных концепций и моделей.

Эти направления используют ряд нетривиальных компьютерных методов:

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

Подчеркнем, что АП и ИЖ – активно развивающиеся направления исследований. По этим направлениям регулярно проводятся международные и европейские конференции "Simulation of Adaptive Behavior (From Animal to Animat)", "Artificial Life", "European Conference on Artificial Life". Издаются журналы "Adaptive Behavior" и "Artificial Life".

В целом соотношение между направлениями "Искусственная жизнь" и "Адаптивное поведение", используемыми в них компьютерными методами, их научным значением и их потенциальными применениями иллюстрируется рис. 1.

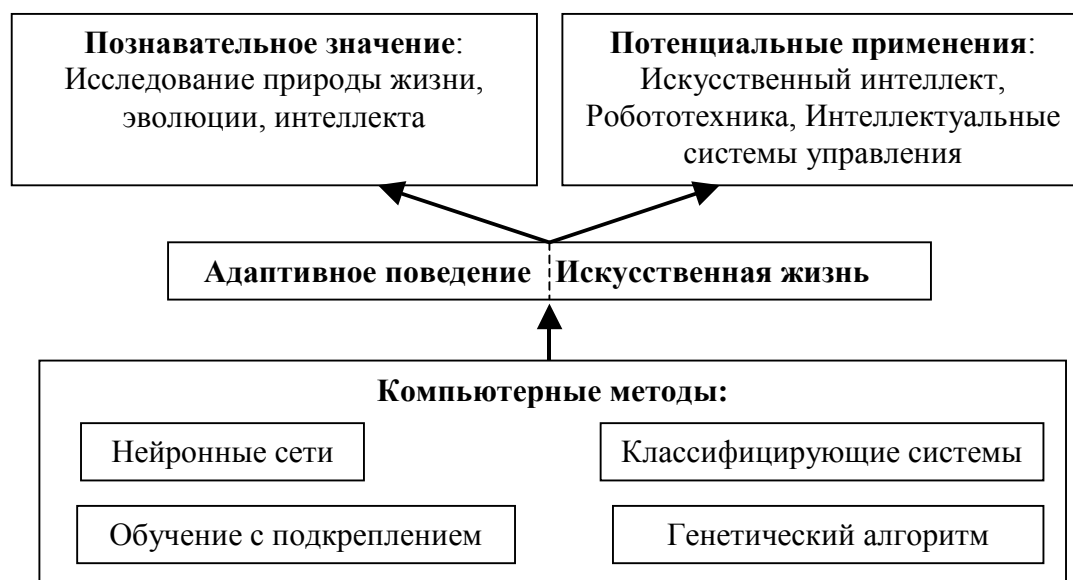


Рис. 1. Направления исследований "Адаптивное поведение" и "Искусственная жизнь".

Отметим, что направление исследований "Адаптивное поведение" выглядит значительно более серьезным, чем "Искусственная жизнь". В работах по ИЖ слишком много "игрушечности", исследователи часто просто играют с теми моделями, которые они придумывают, многих исследователей, возможно, привлекает красивый лозунг "Искусственная жизнь" (Хотя необходимо отметить, что на основе исследований искусственной жизни формируется новое направление прикладных разработок – многоагентное моделирование [18]). Исследования по адаптивному поведению более целенаправлены – они ориентированы на определенную и важную задачу – исследование систем управления, обеспечивающих поведение естественных или искусственных организмов. Более того, схемы управления адаптивным поведением тестируются на модельных (реализованных в компьютерных программах) или реальных роботах, что придает направлению исследований "Адаптивное поведение" определенность и надежность.

Подчеркнем, что в "Адаптивном поведении", как и в "Искусственной жизни", в основном используется *феноменологический подход* к исследованиям систем управления адаптивным поведением. Т.е. 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. По-видимому, и для описания живых организмов может быть аналогичное соотношение феноменологическим (на уровне поведения) и микроскопическим (на уровне нейронов и молекул) подходами. При этом, естественно ожидать, что для исследования систем управления адаптивным поведением феноменологический подход должен быть более эффективен (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.

## 2. Компьютерные методы

Обратим внимание на компьютерные методы, используемые в работах по ИЖ и АП (рис.1). В отличие от нейронных сетей и генетического алгоритма, которые достаточно хорошо известны отечественным ученым, работы по классифицирующим системам и обучению с подкреплением практически не представлены в русскоязычной научной литературе. Кратко охарактеризуем эти работы.

### 2.1. Классифицирующие системы

Классифицирующие системы (Classifier Systems) [15] предназначены для формирования "логических" правил поведения анимата.

Классифицирующие системы работают на основе множества правил (классификаторов) вида:

Если <УСЛОВИЕ(Я)>, то <ДЕЙСТВИЕ>.

Условия представляют собой цепочки бинарных символов:

$$S_k = (1, 0, 1, 0, 0, 1).$$

Каждый классификатор имеет свой вес  $W_k$  (силу классификатора). При этом классификаторы допускают неопределенность – неполное совпадение всех символов в условиях. Т.е., условие может иметь вид:

$$S_k = (1, 0, \#, 0, \#, 1), \quad \# - \text{любой символ из множества } \{0,1\}.$$

Часть действий состоит формирование условий для системы классификаторов, т.е. в процессе работы возможно формирование цепочки последовательных действий.

Популяция классификаторов оптимизируется в результате двух процессов: обучения и эволюции.

При обучении модифицируются веса классификаторов  $W_k$ . Обучение происходит так называемым методом "пожарной бригады": при успехе поощряются не только тот классификатор, который привел к успешному действию, но и его предшественники.

В процессе эволюции с помощью генетического алгоритма формируются новые классификаторы.

Современное состояние исследований классифицирующих систем хорошо отражено в работе [19].

## 2.2. Обучение с подкреплением

Теория обучения с подкреплением была разработана в работах Р. Саттона и Э. Барто (Массачусетский университет).

Идейным вдохновителем этих работ был А.Г. Клопф (Air Force, USA), который в книге "Целеустремленный нейрон" предложил несколько спорную, но достаточно четкую и последовательную методологию исследований памяти, обучения, адаптивного поведения [20].

Общая схема обучения с подкреплением [16] показана на рис.2.

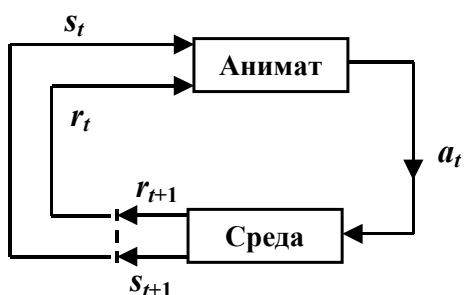


Рис.2. Схема обучения с подкреплением.

Рассматривается анимат, взаимодействующий с внешней средой. Время предполагается дискретным:  $t = 1, 2, \dots$ . В текущей ситуации анимат  $s_t$  выполняет действие  $a_t$ , получает подкрепление  $r_{t+1}$  и попадает в следующую ситуацию  $s_{t+1}$ . Число возможных ситуаций и действий предполагается конечным. Подкрепление может быть положительным (награда) или отрицательным (наказание).

Цель анимата – максимизировать суммарную награду  $\sum_{k=0}^T r_{t+k+1}$ , которую можно получить в будущем в течение длительного периода времени  $T$ ,  $T \gg 1$ . Предполагается, что анимат имеет свою внутреннюю "субъективную" оценку суммарной награды и в процессе обучения постоянно совершенствует эту оценку. Эта оценка определяется с учетом дисконтного фактора:

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad , \quad (1)$$

где  $R_t$  - оценка суммарной награды, ожидаемой после момента времени  $t$ ,  $\gamma$  – дисконтный фактор,  $0 < \gamma < 1$ , дисконтный фактор учитывает, что чем дальше анимат "заглядывает" в будущее, тем меньше у него уверенность в оценке награды – "рубль сегодня стоит больше, чем рубль завтра". Здесь и далее предполагается, что  $T = \infty$ .

В процессе обучения анимат формирует *политику* (стратегию поведения). Политика определяет выбор (детерминированный или вероятностный) действия в зависимости от ситуации. Р. Саттон и Э. Барто [16] исследовали ряд методов формирования политики, основанных на динамическом программировании и методах Мотне-Карло.

Опишем ключевой метод формирования политики – метод временной разности (temporal difference method). Обучение происходит в on-line режиме. В процессе обучения итеративно формируются оценки суммарной величины награды  $Q(s_t, a_t)$ , которую получит анимат, если в ситуации  $s_t$  он выполнит действие  $a_t$ . Так как число действий и число ситуаций конечно, то необходимо найти матрицу  $Q(s_t, a_t)$ , определяющую выбор действий. Величины  $Q(s_t, a_t)$  носят название ценности действий. Математическое ожидание награды равно:

$$Q(s_t, a_t) = E \{ ( r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots ), \} | s=s_t, a=a_t, \quad (2)$$

Из (1), (2) следует  $Q(s_{t-1}, a_{t-1}) = E[r_t + \gamma Q(s_t, a_t)]$ . Ошибку естественно определить так:

$$\delta_t = r_t + \gamma Q(s_t, a_t) - Q(s_{t-1}, a_{t-1}) \quad (3)$$

Здесь  $\delta_t$  – разность между той оценкой суммарной величины награды, которая формируется у анимата для предыдущего момента времени  $t-1$  после совершения действия  $a_t$  в текущей ситуации  $s_t$  в момент времени  $t$ , и предыдущей оценкой этой же величины, которая была у анимата в момент времени  $t-1$ . Предыдущая оценка равна  $Q(s_{t-1}, a_{t-1})$ , новая оценка (после совершения действий  $a_{t-1}$  и  $a_t$  при текущей политике) равна  $r_t + \gamma Q(s_t, a_t)$ , что и отражает формула (3) для величины  $\delta_t$ . В соответствии с этим  $\delta_t$  анимат и обучается (см. ниже, формулу (4)).

Метод временной разности работает следующим образом. Исходные значения величин  $Q(s_t, a_t)$  произвольны. Затем каждый такт времени одновременно происходит как выбор действия, так и обучение анимата.

Выбор действия происходит так:

- в момент  $t$  с вероятностью  $1 - \varepsilon$  выбирается действие с максимальным значением  $Q(s_t, a_t)$ :  $a^* = \arg \max_a \{ Q(s_t, a_t) \}$ ,
- с вероятностью  $\varepsilon$  выбирается произвольное действие,  $0 < \varepsilon \ll 1$ .

Такой выбор действия называют " $\varepsilon$ -жадной" политикой: доля случаев, в которых анимат выбирает действие, соответствующее максимальной оценке ожидаемой суммарной награды, равна  $1 - \varepsilon$ . В малом числе случаев (доля их равна  $\varepsilon$ ) анимат ведет случайный поиск.

Обучение, т.е. переоценка величин  $Q(s_t, a_t)$  происходит в соответствии с оценкой ошибки  $\delta_t$  – к величине  $Q(s_{t-1}, a_{t-1})$  добавляется величина, пропорциональная  $\delta_t$ :

$$\Delta Q(s_{t-1}, a_{t-1}) = \alpha [r_t + \gamma Q(s_t, a_t) - Q(s_{t-1}, a_{t-1})] = \alpha \delta_t, \quad (4)$$

$\alpha$  – параметр скорости обучения.

Метод временной разности идейно связан с методом динамического программирования, и в том и другом случае общая оптимизация многошагового процесса принятия решения происходит путем упорядоченной процедуры одношаговых оптимизирующих итераций, причем оценки эффективности тех или иных решений, соответствующие предыдущим шагам процесса, переоцениваются с учетом знаний о возможных будущих шагах. Например, при решении задачи поиска оптимального маршрута в лабиринте от стартовой точки к определенной целевой точке сначала находится конечный участок маршрута, непосредственно приводящий к цели, а затем ищутся пути, приводящие к конечному участку, и т.д. В результате постепенно прокладывается трасса маршрута от его конца к началу.

Важное достоинство этого метода обучения с подкреплением – его простота. Т.е. анимат получает от учителя или из внешней среды только сигналы подкрепления  $r_t$ . Здесь учитель поступает с обучаемым объектом примитивно: "бьет кнутом" (если действия объекта ему не нравятся,  $r_t < 0$ ), либо "дает пряник" (в противоположном случае,  $r_t > 0$ ), не объясняя обучаемому

объекту, как именно нужно действовать. Это радикально отличает этот метод от таких традиционных в нейронных сетях методах обучения, как метод обратного распространения ошибок, для которого учитель точно определяет, что должно быть на выходе нейронной сети при заданном входе.

Метод обучения с подкреплением был исследован рядом авторов (см. подробную библиографию в [16]) и был использован многочисленных приложениях. В частности, применения этого метода включают в себя:

- оптимизацию игры в триктрак (достигнут уровень мирового чемпиона);
- оптимизацию системы управления работы лифтов;
- формирование динамического распределения каналов для мобильных телефонов;
- оптимизацию расписания работ на производстве.

Подчеркнем, что метод обучения с подкреплением может рассматриваться как развитие автоматной теории адаптивного поведения, разработанной в работах М.Л. Цетлина и его последователей [5,6].

### 3. Исследования AnimatLab

Исследования по адаптивному поведению ведутся в ряде университетов и лабораторий. Кратко охарактеризуем работы одной из ведущих лабораторий – лаборатории AnimatLab, которой руководит один из инициаторов этого направления Жан-Аркадий Мейер.

#### 3.1. Общий подход AnimatLab: Эволюция + обучение + онтогенез

Общий подход этой лаборатории можно охарактеризовать следующим образом [21]. Анимат (рис.3) существует в реальной или модельной среде. У него есть сенсоры, которые воспринимают информацию из внешней и внутренней среды анимата, и эффекторы, посредством которых он взаимодействует со средой, а также система управления, которая координирует восприятие и действия анимата. Поведение анимата считается *адаптивным*, если система управления поддерживает жизненно важные переменные анимата (например,  $V_1$  и  $V_2$  на рис.3) в допустимых пределах. На рис.3 штриховая стрелка показывает возможную траекторию, выходящую за пределы допустимой области (серый фон – недопустимая область переменных). Сплошная стрелка показывает "исправленную" траекторию, откорректированную с помощью системы управления, обеспечивающей поддержание переменных в допустимой (светлой) области.

Если система управления выбирает последовательные цели, которые анимат стремится достичь, то о такой системе можно говорить как о *мотивационной системе* (motivational system). Система управления анимата может формироваться и модифицироваться путем *обучения*, индивидуального *развития* (онтогенеза) и *эволюции*.

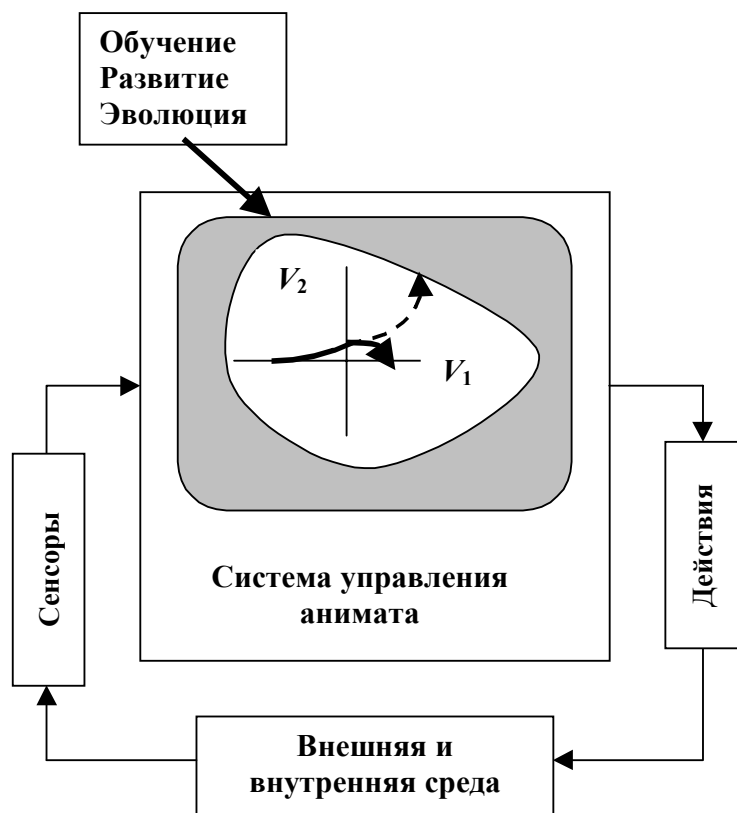


Рис. 3. Общая схема поведения анимата. Подход AnimatLab [21].

### 3.2. Онтогенез нейронной сети анимата – схема формирования структуры "нервной системы" робота

В работах AnimatLab была сделана интересная попытка промоделировать индивидуальное развитие, "онтогенез" нейронной сети анимата [22-24].

Общая схема метода состоит в следующем (рис.4). Нейронная сеть искусственного организма (анимата) формируется с помощью специальной программы, контролирующей процесс конструирования сети. Эта программа имитирует процесс развития нейронной сети в процессе индивидуального взросления организма. Сама программа оптимизируется с помощью эволюционного алгоритма. Нейронная сеть формируется в двумерной ограниченной области. Программа состоит из инструкций (команд), которые определяют процессы возникновения новых нейронов (или исчезновения уже имеющихся нейронов) в этой области, формирование связей между нейронами и задание весов синаптических связей между нейронами. Инструкции программы составляют геном анимата. Работа формирующихся нейронных сетей оценивается по поведению анимата некоторой естественной функцией приспособленности, которая определяет отбор наиболее эффективных программ, кодируемых геномами аниматов. Так как структура сети определяется расположением нейронов в двумерной области, то метод называется геометрическим.





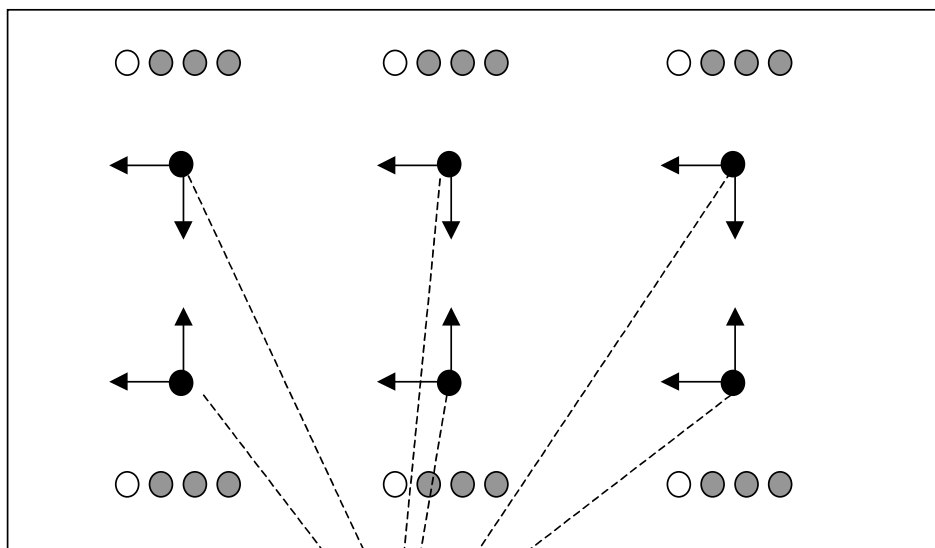
Рис.4. Общая схема геометрического метода формирования структуры нейронной сети.

Эта схема иллюстрирует процесс формирования отдельного модуля нейронной сети. К этому модулю могут добавляться новые модули. Причем сначала формируются модули нижних уровней, определяющие инстинктивное, рефлекторное поведение анимата (например, согласованное движение 6-ти ног при прямолинейном перемещении анимата). А затем формируются модули более высоких уровней (например, модуль управляющий остановкой и возобновлением прямолинейного движения), управляющие работой модулей нижних уровней.

**Программа развития нейронной сети.** Опишем принципы формирования нейронной сети на основе инструкций программы развития. Процесс формирования определяется как экспериментатором, конструирующим нейронную сеть, так и работой программ развития и эволюционной оптимизацией этих программ.

Первоначально экспериментатор выделяет двумерную область (например, прямоугольник). В этой области он задает расположение тех нейронов, которые заведомо необходимы (сенсорные и моторные нейроны). Затем задается множество затравочных нейронов, из которых будет развиваться нейронная сеть.

Пример области с расположенными на ней сенсорными, моторными и затравочными нейронами показан на рис. 5.



Программа развития нейронной сети

- – сенсорные нейроны
- – моторнейроны трех разных типов, управляющие разными "мышцами" анимата
- – затравочные нейроны

Рис. 5. Схема расположения изначально задаваемых нейронов в прямоугольной двумерной области (по работе [22], с изменениями). Симметрия схемы отражает то, что она предназначена для формирования нейронной сети 6-ногого анимата. Каждый из затравочных нейронов имеет свою локальную систему координат, которую он затем использует при формировании дочерних нервных клеток и связей с другими нейронами.

Экспериментатор также задает инструкции, из которых будут формироваться программы развития нейросетей, и некоторые синтаксические ограничения на то, как из этих инструкций можно формировать программы.

В работах [22-24] была использована модель интегрирующего нейрона, согласно которой динамика мембранного потенциала  $i$ -й нервной клетки описывается уравнением:

$$\tau_i \cdot dm_i / dt = -m_i + \sum_j w_{ij} x_j + I_i$$

где  $x_j = \{1 + \exp[-(m_j + B_j)]\}^{-1}$  – частота импульсов нейрона,  $B_j$  – случайный порог нейрона со средним значением  $b_j$ ,  $\tau_i$  – постоянная времени релаксации  $i$ -го нейрона,  $I_i$  – внешний вход  $i$ -го нейрона от определенного сенсора,  $w_{ij}$  – синаптический вес, характеризующий связь  $j$ -го от  $i$ -му нейрону.

Примеры инструкций, определяющих программы нейронов, представлены в таблице 1.

Таблица 1. Инструкции – команды программы развития нейронной сети

DIVIDE ( $\alpha, r$ )	Создать новый (дочерний) нейрон
GROW ( $\alpha, r, w$ )	Создать соединение к другому нейрону
DRAW ( $\alpha, r, w$ )	Создать соединение от другого нейрона
SETBIAS ( $b$ )	Изменить порог нейрона
SETTAU ( $\tau$ )	Изменить постоянную времени релаксации нейрона
DIE	Удалить нейрон

Инструкции характеризуются параметрами, которые определяют, как именно будет развиваться нейронная сеть (пример см. ниже, при описании рис. 6).

Программа развития состоит из подпрограмм. Каждому затравочному нейрону соответствует своя подпрограмма.

Инструкции каждой из подпрограмм скомпонованы в граф, в котором есть корневой узел, команда-инструкция которого применяется к затравочному нейрону. Работа подпрограммы начинается с того, что затравочная нервная клетка исполняет инструкцию корневого узла подпрограммы.

Одна из инструкций (DIVIDE) соответствует делению клетки на материнскую и дочернюю. Формирование дочерней нервной клетки соответствует ветвлению графа подпрограммы; при этом после деления клетки (и соответствующего ветвления графа) инструкция левого узла относится к материнской нервной клетке, а инструкция правого узла – к дочерней клетке. Итак, инструкция DIVIDE соответствует узлу ветвления графа на две ветви. Все остальные инструкции не приводят к ветвлению графа подпрограммы.

Некоторые из команд-инструкций терминальные, после этих инструкций процесс развития той нервной клетки, к которой они применяются, останавливается.

Программы формирования нейронной сети оптимизируются с помощью эволюционного алгоритма. В процессе эволюции программы испытывают мутации и рекомбинации. Рекомбинации представляют собой обмен подграфами подпрограмм – аналогично тому, как это осуществляется в генетическом программировании [25,26]. Пример применения инструкций показан на рис. 6.

- – развивающаяся нервная клетка      ○ – дочерняя клетка
- – другой нейрон

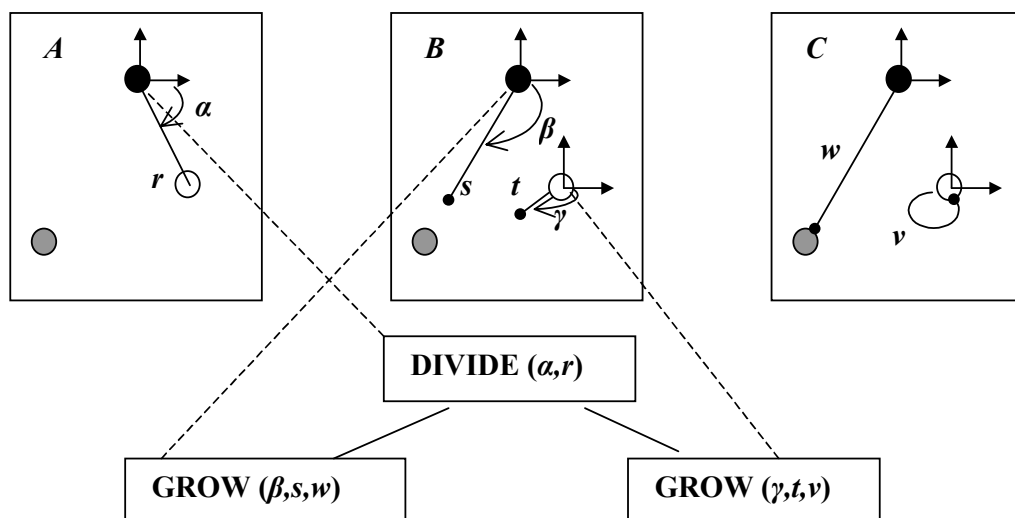


Рис. 6. Пример развития нейронной сети. Развитие происходит под управлением программы, представленной внизу рисунка. Программа содержит команду деления развивающейся нервной клетки  $DIVIDE(a,r)$  и команды формирования связей от материнской  $GROW(\beta,s,w)$  и дочерней  $GROW(\gamma,t,v)$  клеток. Первые два параметра в этих командах определяют углы и расстояния, в соответствии с которыми определяется положение дочерней клетки (для команды  $DIVIDE$ ) и направление и длина синаптической связи (для команды  $GROW$ ). Третий параметр команды  $GROW$  определяет величину синаптической связи.

Как видно из рис. 6А, сначала по команде  $DIVIDE(a,r)$  формируется дочерняя клетка на расстоянии  $r$  от материнской с "азимутом"  $\alpha$ . Дочерняя клетка наследует локальную систему координат материнской клетки. Затем (рис. 6В), как материнская (по команде  $GROW(\beta,s,w)$ ) так и дочерняя (по команде  $GROW(\gamma,t,v)$ ) клетки формируют отростки длиной  $s$  и  $t$  и под углом  $\beta$  и  $\gamma$  соответственно. Далее (рис. 6С) концы этих отростков подсоединяются к ближайшей из клеток и задаются веса ( $w$  и  $v$ ) соответствующих синаптических связей. На этом указанный блок программы выполнен.

Итак, геном анимата представляет собой определенную программу, однозначно определяющую процесс формирования структуры и весов нейронной сети. Программа состоит из инструкций, которые задают правила расстановки нейронов в заданной геометрической области и формирования синаптических связей между нейронами.

**Эволюционный алгоритм.** Эволюционный алгоритм в данном методе состоит в следующем. Поведение анимата оценивается в соответствии с эвристически задаваемой функцией приспособленности, производится отбор в аниматов с большими приспособленностями, и наиболее приспособленные аниматы дают потомков.

При формировании потомков применяются три генетических оператора: 1) совместимая рекомбинация под-графов программ выбранных родителей, 2) формирование новых случайных (но допустимых) под-графов взамен старых, 3) и случайные мутации параметров, входящих в команды-инструкции.

В целом схема эволюции типична для генетического алгоритма, однако, необходимо наложение определенных условий на то, чтобы программы потомков не выходили за рамки ограничений, накладываемых на эти программы.

**Принцип модульности. Примеры использования геометрического метода.** Принцип модульности подразумевает, что модули нейронной сети формируются последовательно – сначала один модуль, затем следующий. Например, в работе [22] сначала был эволюционно сформирован 1-й модуль, определяющий согласованное движение 6-ти ног при прямолинейном одномерном перемещении анимата, а затем 2-й модуль, управляющий остановкой и возобновлением прямолинейного движения.

В работе [23] была сформирована нейронная сеть, управляющая 2-мерным движением 6-ногого анимата. При этом нейронная сеть состояла из трех модулей:

1-й модуль управлял движением анимата,

2-й модуль контролировал работу первого модуля и обеспечивал перемещение анимата к заданной цели,

3-й модуль обеспечивал перемещение в среде с препятствиями, и был предназначен для минимизации столкновений анимата с препятствиями.

Модули формировались последовательно (сначала 1-й, потом 2-й, затем 3-й). При формировании 1-го модуля приспособленность анимата оценивалась по скорости его движения (чем больше скорость перемещения, тем выше приспособленность). При формировании 2-го и 3-го модулей 1-й модуль оставался неизменным, а формировались только связи от нейронов новых модулей к нейронам первого модуля. Приспособленность программ 2-го модуля оценивалась по способности анимата находить источник запаха. Приспособленность программ 3-го модуля оценивалась по способности анимата избегать препятствия, случайно разбросанные в области его движения. Программы формирования каждого из модулей были различными, более того, несколько различался синтаксис инструкций и графов, на основе которых формировались программы.

В полученной нейронной сети 1-й модуль содержал 38 интернейронов (сформированных нейронов, обеспечивающих связи между сенсорными нейронами и мотонейронами) и 100 межнейронных соединений; 2-й модуль содержал 6 интернейронов и 22 межнейронных соединения; 3-й модуль содержал 2 интернейрона и 6 межнейронных соединений.

В работах [22,23] геометрический метод "выращивания" структуры нейронной сети был применен к задаче формирования системы управления аниматами, моделируемыми компьютерной программой. Т.е. аниматы "жили" только в компьютере. В работе [24] этот метод был применен к реальному 6-ногому роботу SECT, который был обучен перемещаться по двумерной плоскости и избегать столкновения с препятствиями. Было продемонстрировано, что имитация поведения робота в компьютерных программах согласуется с поведением реального робота.

Анализ геометрического метода формирования структуры нейронной сети показывает, что этот метод достаточно универсален и может быть применен к широкому классу систем управления адаптивным поведением. Основные принципы этого метода сводятся к следующему.

1. Создается специальный язык команд-инструкций, на основе которого строятся программы развития (онтогенеза) структуры и параметров нейронной сети.
2. Программы онтогенеза нейронных сетей оптимизируются эволюционным путем. Схема эволюционной оптимизации близка к таковой в генетическом программировании.
3. Сложная нейронная сеть строится по модульному принципу: сначала формируются модули нижних уровней управления, а затем модули верхних уровней иерархии управления.

### **3.3. Анимат MonaLysa – пример мотивационной системы**

Интересное направление исследований AnimatLab – конструирование и моделирование мотивационных систем управления аниматами.

Пример мотивационной системы – довольно интеллектуальная архитектура управления аниматом MonaLysa, который, функционируя в сложной среде, способен сам выделять цели и

подцели адаптивного поведения. (MonaLysa – сокращение от MotivatiONAILY autonomouS Animat) [13]. Основная идея данной системы управления состоит в том, что в процессе освоения внешнего мира и накопления опыта анимат стремится разбить задачу достижения глобальной цели на подзадачи, а затем использовать этот опыт при планировании решения новых задач.

В работе [13] исследовалось поведение анимата MonaLysa на примере навигационной задачи. Анимат помещался в центральную нижнюю точку прямоугольника, и нужно было попасть в центральную верхнюю точку (рис. 7), обходя различные препятствия. Анимат мог работать в "планирующем режиме", т.е., как сказано выше, разбивать задачи на подзадачи, и планировать свои действия в соответствии с уже имеющимся опытом. Это поведение сравнивалось с поведением в "реактивном режиме" – без плана, на основе только текущей видимой ситуации.

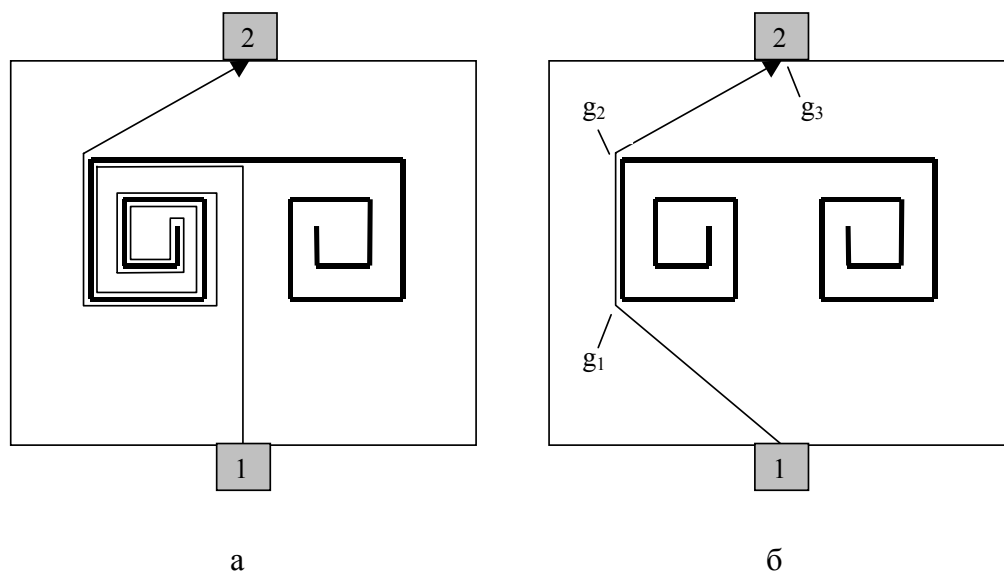


Рис. 7. Траектории движения анимата MonaLysa в реактивном (а) и планирующем (б) режиме работы системы управления. По [13] с изменениями. 1 – исходное положение анимата, 2 – конечная цель движения анимата. Жирной линией показаны препятствия, которые должен обойти анимат, тонкой линией – путь анимата.

В "реактивном режиме" анимат знает, где расположена конечная цель, и если нет препятствия, то движется прямо к этой цели; наткнувшись на препятствие, он обходит его до тех пор, пока не появится снова возможность двигаться к цели (рис.7а). В "планирующем режиме" анимат на основании предшествующего опыта выделяет подцели (точки  $g_1$ ,  $g_2$ ,  $g_3$  на рис.7б), и движется прямо к текущей подцели, причем последняя подцель совпадает с конечной целью движения.

Отметим, что схема анимата MonaLysa была реализована как в компьютерной программе, так и для управления реальным роботом Khepera. Система управления в анимате MonaLysa была основана на простой версии классифицирующей системы.

Здесь мы охарактеризовали работы только одной лаборатории, ведущей исследования в области адаптивного поведения, в близких направлениях ведутся исследования и разработки в ряде университетских центров, таких как:

- Лаборатория искусственного интеллекта в университете Цюриха (руководитель Рольф Пфейфер) [27,28]. Основной подход этой лаборатории – познание природы интеллекта путем создания ("understanding by building"). Он включает в себя 1) построение моделей биологических систем, 2) исследование общих принципов естественного интеллекта животных и человека, 3) использование этих принципов при конструировании роботов и других искусственных интеллектуальных систем.

- Лаборатория искусственной жизни и роботики в Институте когнитивных наук и технологий (Рим, руководитель Стефано Нолфи) [29,30], ведущая исследования в области эволюционной роботики и принципов формирования адаптивного поведения.
- Лаборатория искусственного интеллекта в Массачусетском технологическом институте (руководитель Родни Брукс) [31,32], которая ведет исследования широкого спектра интеллектуальных и адаптивных систем, включая создание интеллектуальных роботов.

Для осмысления многообразия форм адаптивного поведения необходимо не только исследование конкретных моделей, но и разработка общих концепций и схем, позволяющих взглянуть сверху, "с высоты птичьего полета" на эти исследования.

Из одной таких концептуальных теорий может служить теория функциональных систем, предложенная и развитая в 1930-70 годах известным советским нейрофизиологом П.К. Анохиным [33-35].

#### 4. Теория функциональных систем П.К. Анохина как концептуальная основа исследований адаптивного поведения

Функциональная система по П.К. Анохину – схема управления, нацеленного на достижение полезных для организма результатов.

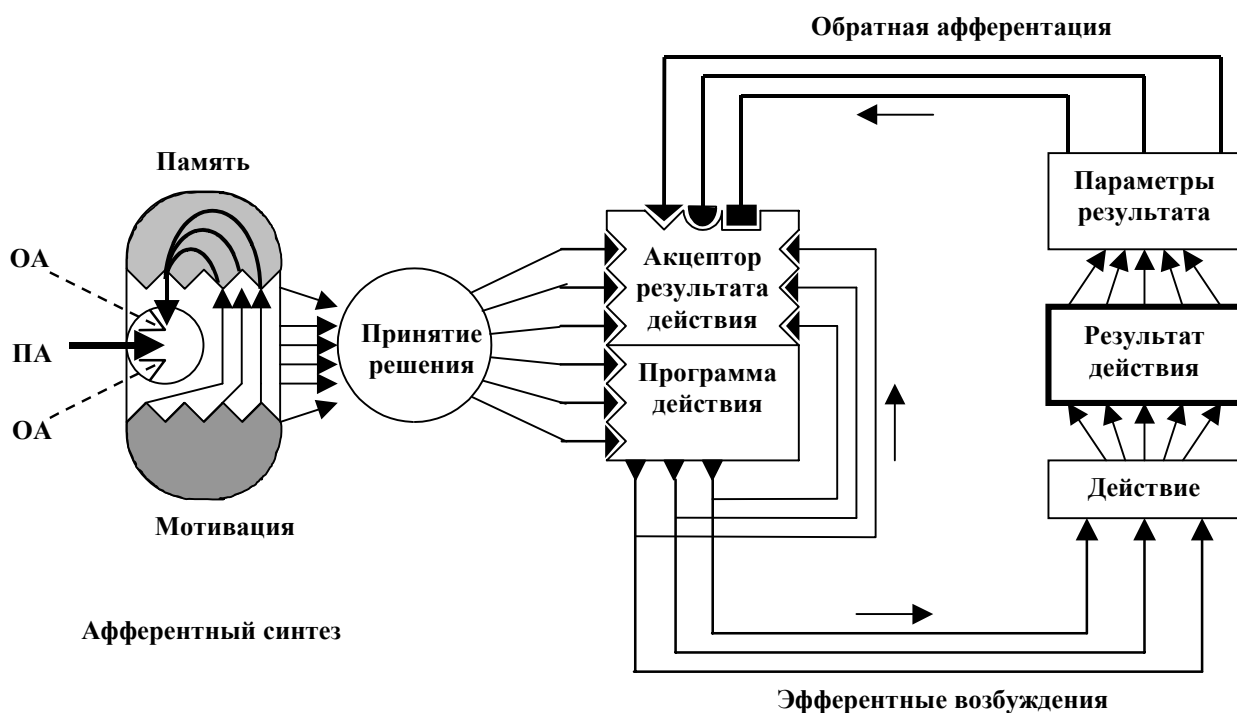


Рис. 8. Структура функциональной системы. ОА – обстановка афферентация, ПА – пусковая афферентация.

Работа функциональной системы (рис. 8) может быть описана следующим образом.

Сначала происходит *афферентный синтез*, который включает в себя нейронные возбуждения, обусловленные 1) доминирующей мотивацией, 2) обстановочной и пусковой афферентацией, 3) врожденной и приобретаемой памятью.

За афферентным синтезом следует *принятие решения*, при котором происходит уменьшение степеней свободы для эфферентного синтеза и выбор конкретного действия в соответствии с доминирующей потребностью животного и с другими составляющими афферентного синтеза.

Затем следует формирование *акцептора результата действия*, т.е. прогноза результата. Прогноз включает в себя оценку параметров ожидаемого результата.

*Эфферентный синтез* – подготовка к выполнению действия. При эфферентном синтезе происходит генерация определенных нейронных возбуждений перед подачей команды на выполнение действия.

Все этапы достижения результата сопровождаются *обратной афферентацией*. Если параметры фактического результата отличаются от параметров акцептора результата действия, то действие прерывается и происходит новый афферентный синтез. В этом случае все операции повторяются, до тех пор, пока не будет достигнут конечный потребный результат.

Таким образом, функциональная система имеет циклическую (с обратными афферентными связями) саморегулирующуюся архитектуру.

Теория П.К. Анохина подразумевает *динамизм функциональных систем*. Для каждого конкретного поведенческого акта может быть сформирована своя функциональная система.

Функциональные системы формируются в процессе *системогенеза*. Теория системогенеза, которая исследует закономерности формирования функциональных систем в *эволюции, индивидуальном развитии и обучении* [36], может рассматриваться как отдельная ветвь теории функциональных систем. Отметим, что указанные составляющие системогенеза соответствуют составляющим формирования систем адаптивного поведения в трактовке AnimatLab (рис. 3).

Каждая функциональная система ориентирована на достижение *конечного потребного результата*.

Необходимо подчеркнуть, что теория функциональных систем была разработана, в первую очередь, для интерпретации нейробиологических данных и зачастую сформулирована в очень интуитивных терминах. Поэтому, хотя она и хорошо известна, она не общепризнана и практически не использовалась при разработке серьезных моделей адаптивного поведения. Можно сказать, что попытки формализации теории функциональных систем только начинаются [37-39]. Тем не менее, эта теория базируется на многочисленных биологических экспериментальных данных и представляет собой хорошую концептуальную основу для исследования широкого спектра проблем адаптивного поведения.

Отталкиваясь от теории П.К. Анохина, можно предложить общую кибернетическую схему управления целенаправленным адаптивным поведением естественного или искусственного организма (рис. 9). Здесь под организмом можно подразумевать как животное, так и робот или социально-экономическую систему: промышленную фирму, государство, человечество.



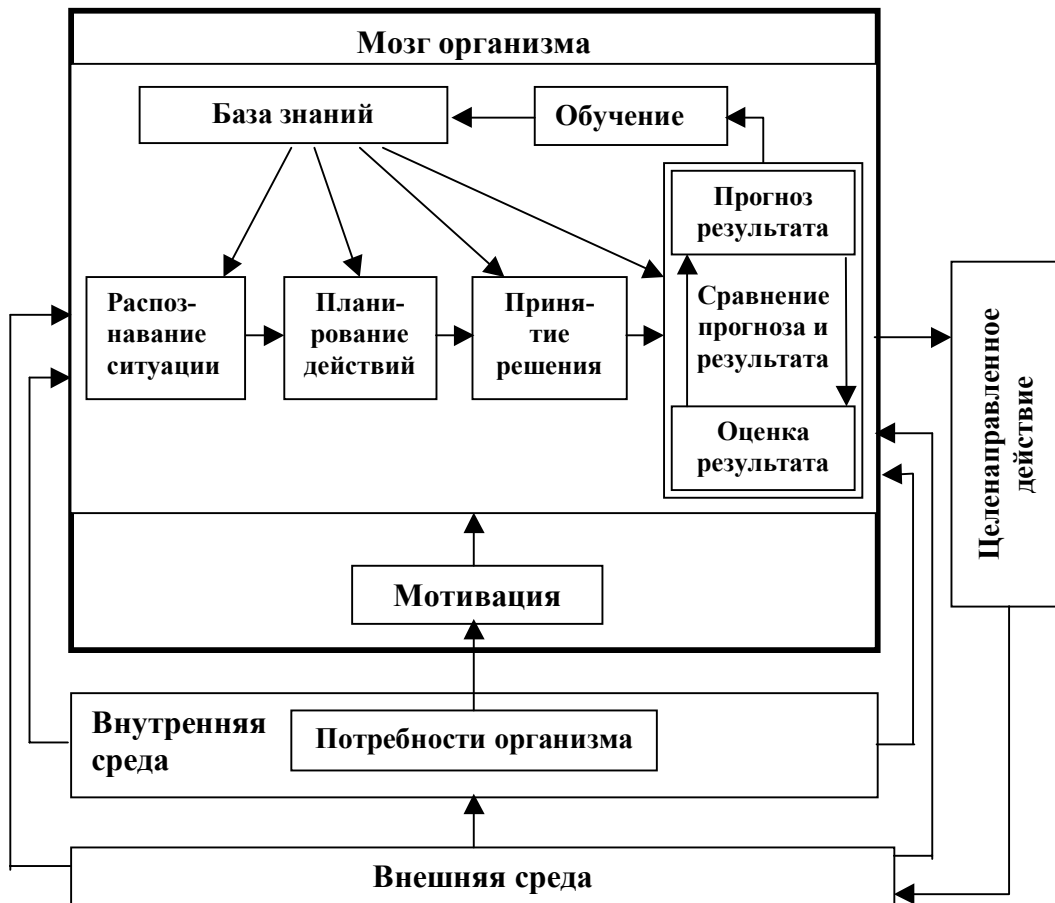


Рис. 9. Общая схема управления целенаправленным адаптивным поведением (в духе П.К. Анохина).

Что же можно делать сразу сейчас? Какие интересные задачи можно исследовать, отталкиваясь от теории функциональных систем?

Важное понятие функциональной системы – *мотивация*. Роль мотивации состоит в формировании цели и поддержке целенаправленных форм поведения. Мотивация может рассматриваться как активная движущая сила, которая стимулирует нахождение такого решения, которое адекватно потребностям организма в рассматриваемой ситуации. И имеет смысл провести моделирование эволюционного возникновения *целенаправленного* адаптивного поведения и анализ роли мотиваций в формировании целенаправленного поведения. Также следует отметить, что целенаправленность могла возникнуть на очень ранних стадиях эволюции, до появления каких-либо форм индивидуально приобретаемой памяти [40], поэтому, следуя пути, пройденному эволюцией, разумно начать с анализа этого свойства. Кроме того, свойство целенаправленности важно само по себе – это существенная особенность поведения *именно живых существ*.

Модели эволюционного возникновения целенаправленного адаптивного поведения были построены и исследованы в работах [41-43]. Основные результаты этого моделирования излагаются в следующем разделе.

## 5. Модели эволюционного возникновения целенаправленного адаптивного поведения

### 5.1. Роль мотиваций в формировании адаптивного поведения [41,42]

В данной модели исследовался возможный механизм эволюционного возникновения целенаправленного поведения, обусловленного мотивациями.

**Основные предположения модели** состоят в следующем:

- Имеется популяция агентов (искусственных организмов), имеющих естественные потребности: 1) *потребность энергии* и 2) *потребность размножения*.
- Популяция эволюционирует в одномерной клеточной среде (рис. 10), в клетках может эпизодически вырастать трава (пища агентов). Каждый агент имеет *внутренний энергетический ресурс  $R$* , который пополняется при съедании травы и уменьшается при выполнении каких-либо действий. Уменьшение ресурса до нуля приводит к смерти агента. Агенты могут скрещиваться, рождая новых агентов.
- Потребности характеризуется количественно *мотивациями*. Если энергетический ресурс  $R$  агента уменьшается, то возрастает мотивация к пополнению энергетического ресурса (соответствующая потребности энергии) и уменьшается мотивация к размножению. При увеличении  $R$  мотивация к пополнению ресурса уменьшается, а мотивация к размножению растет.
- Поведение агента управляется его *нейронной сетью*. Сеть имеет один слой нейронов. На входы нейронов подаются сигналы, характеризующие внешнюю и внутреннюю среду агента, выходы нейронов определяют действия агента. Каждому возможному действию соответствует ровно один нейрон. В каждый такт времени совершается действие, соответствующее максимальному сигналу на выходе нейрона.
- Агенты "близорукие" – агент воспринимает состояние внешней среды только из трех клеток его поля зрения (рис. 10): той клетки, в которой агент находится, и двух соседних клеток.
- Агент может выполнять следующие *действия*: 1) быть в состоянии покоя ("отдыхать"), 2) двигаться, т.е. перемещаться на одну клетку вправо или влево, 3) прыгать через несколько клеток в случайную сторону, 4) есть (питаться), 5) скрещиваться.
- Нейронная сеть имеет специальные входы от мотиваций. Если имеется определенная мотивация, то поведение агента меняется с тем, чтобы удовлетворить соответствующую потребность. Такое поведение можно рассматривать как *целенаправленное* (есть цель удовлетворить определенную потребность).
- Популяция агентов *эволюционирует*. Веса синапсов нейронной сети, управляющей поведением агента, составляют геном агента. Геном потомка формируется на основе геномов родителей при помощи рекомбинаций и мутаций.
- Мотивация к пополнению энергетического ресурса  $M_E$  и мотивация к размножению  $M_R$  определялись как простые функции энергетического ресурса агента  $R$  :

$$M_E = \max \{(R_0 - R) / R_0, 0\} \quad M_R = \min \{R / R_1, 1\},$$

где  $R_0, R_1$  – параметры (обычно полагалось  $R_0 = 2 R_1$ ).

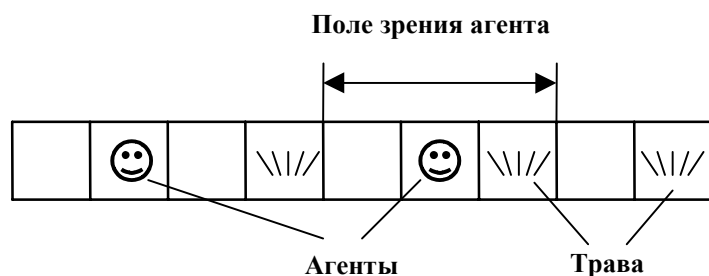


Рис. 10. Агенты в одномерной клеточной среде.

В проведенных компьютерных экспериментах моделировалась эволюция популяции агентов. Нейронная сеть агентов исходной популяции определяла некоторые простые изначальные инстинкты, обеспечивающие питание и размножение агентов. Далее наблюдалось, как в процессе эволюции изменялись нейронная сеть агентов и определяемое ей поведение агентов.

Для того чтобы исследовать влияние мотиваций на поведение агентов, были проведены две серии экспериментов. В первой серии моделировалась эволюция популяции агентов с "выключенными" мотивациями (входы нейронов от мотиваций были "задавлены"), во второй серии мотивации "работали" (так, как это изложено выше).

**Основные результаты** проведенного моделирования таковы:

- Мотивации играют важную роль в исследованных эволюционных процессах. А именно, если сравнить популяцию агентов без мотиваций с популяцией агентов с мотивациями, то, как показывают компьютерные эксперименты, эволюционный процесс приводит к тому, что вторая популяция (с мотивациями) имеет значительные селективные преимущества по сравнению с первой (без мотиваций).
- Результаты моделирования также демонстрируют (рис. 11, 12), что управление поведением агента без мотиваций можно рассматривать как набор простых инстинктов (несколько отличающихся от изначально заданных), а управление агентом с мотивациями – как *иерархическую систему управления*, состоящую из двух уровней: уровня простых инстинктов и метауровня, обусловленного мотивациями. При этом иерархическая система управления обеспечивает более эффективное управление, чем одноуровневая система, в которой поведение определяется одними лишь простыми инстинктами. Переход от схемы управления без мотиваций (рис. 11) к схеме управления с мотивациями (рис. 12) подобен метасистемному переходу от простых рефлексов к сложному рефлексу в теории метасистемных переходов В.Ф. Турчина [44].



Рис. 11. Схема управления агента без мотиваций. Поведение агента состоит из одних только простых безусловных рефлексов, при котором выбор действия напрямую определяется текущим состоянием окружающей среды.



Рис. 12. Схема управления агента, обладающего мотивациями. Мотивации формируют новый уровень иерархии в системе управления агентами.

## 5.2. Возникновение естественной разветвленной иерархии целей

Изложенная модель была развита в работе М.С. Бурцева [43], в которой исследовалось поведение популяции агентов в двумерном мире. При этом дополнительно в модель были введены 1) возможность борьбы между агентами и 2) эволюционное изменение структуры нейронной сети, состоящей из рецепторов, эффекторов и связей между рецепторами и эффекторами.

Как и в предыдущей модели, для агентов исходной популяции задавалась некоторая минимальная система управления, обеспечивающая питание и размножение агентов. Поведение агентов начальной популяции (имеющих минимальный набор рецепторов и эффекторов) схематично представлено на рис. 13.

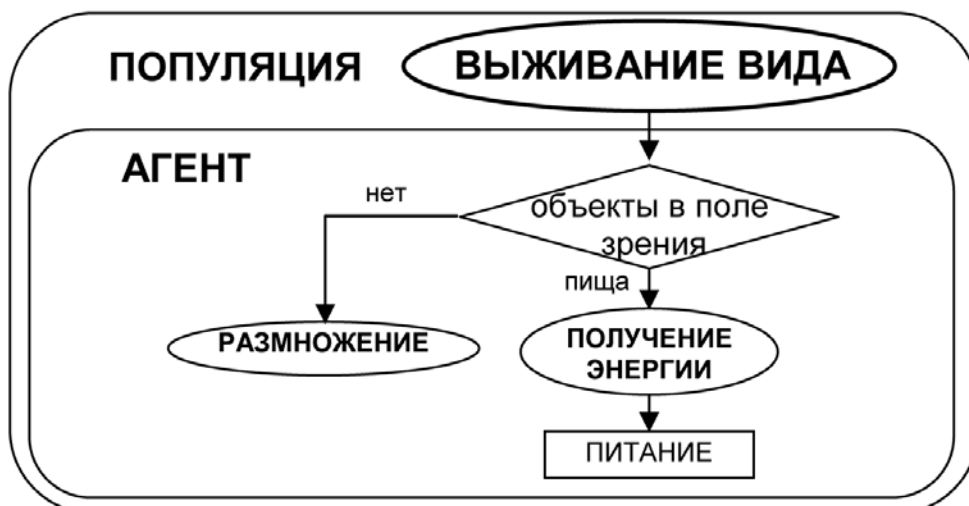


Рис. 13. Дерево условий для управления выбором подцелей агента начальной популяции.

В ходе эволюции поведение агентов структурируется. Стратегия агентов, сформированная в процессе эволюции, может быть представлена в виде схемы, показанной на рис. 14. Видно, что развивается достаточно сложное поведение, которое можно считать целенаправленным. Так первоначальный "инстинкт" агента, направленный на получение энергии, оптимизируется за счет появления еще одного уровня подцелей, направленных, соответственно: на собственно питание, на поиск пищи, на борьбу.

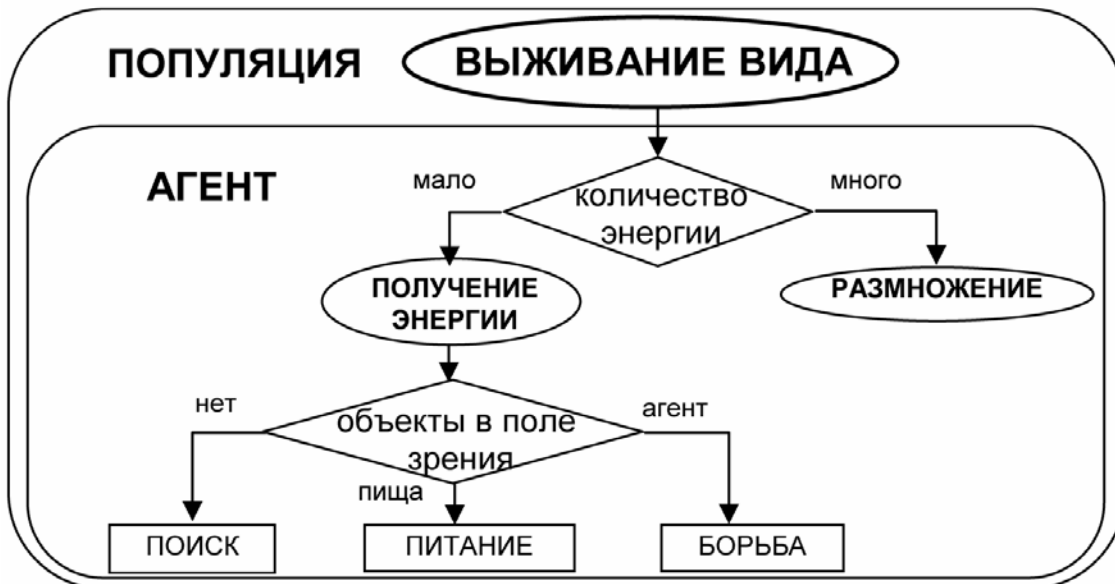


Рис. 14. Дерево условий для управления выбором подцелей, формирующееся в результате эволюции. Количество энергии здесь эквивалентно внутреннему энергетическому ресурсу агента  $R$ .

В целом моделирование, выполненное в работе [43], продемонстрировало, что в процессе исследованных эволюционных процессов возникает естественная иерархическая структура целей и подцелей.

## 6. Сверхзадачи и будущее науки

Изложенное показывает, что работы по моделированию адаптивного поведения, в общем-то, только разворачиваются. Поэтому целесообразно предложить сверхзадачу для стимулирования этих исследований.

### **Сверхзадача 1. Исследование когнитивной эволюции – путь к теории происхождения мышления.**

Существует глубокая гносеологическая проблема: *почему человеческое мышление применимо к познанию природы?* Ведь далеко не очевидно, что те мыслительные процессы, которые мы используем в научном познании, применимы к процессам, происходящим в природе, так как эти два типа процессов различны. Рассмотрим, например, физику, наиболее фундаментальную из естественнонаучных дисциплин. Мощь физики связана с эффективным применением математики. Но математик строит свои теории совсем независимо от внешнего мира, используя свое мышление (в тиши кабинета, лежа на диване, в изолированной камере...). Почему же результаты, получаемые математиком, применимы к реальной природе?

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

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

А теперь перейдем от математика к собаке И.П. Павлова. Пусть у собаки вырабатывают условный рефлекс, в результате в памяти собаки формируется связь "за УС должен последовать БС" (УС - условный стимул, БС - безусловный стимул). И когда после выработки рефлекса собаке предъявляют УС, то она, "помня" о хранящейся в ее памяти "записи": УС  $\rightarrow$  БС, делает элементарный "вывод"  $\{УС, УС \rightarrow БС\} \Rightarrow БС$ . И у собаки, ожидающей БС (скажем, кусок мяса), начинают течь слюнки.

Конечно, применение правила *modus ponens* (чисто дедуктивное) математиком и индуктивный "вывод", который делает собака, явно различаются. Но можем мы ли думать об эволюционных корнях логических правил, используемых в математике? Да, вполне можем – умозаключение математика и "индуктивный вывод" собаки качественно аналогичны.

Мы можем пойти и дальше – можем представить, что в памяти животного есть *семантическая сеть*, сеть, узлами которой являются понятия, образы, а связи характеризуют взаимоотношения между понятиями. Можно далее представить процессы формирования разнообразных семантических сетей в процессе накопления жизненного опыта. Такие семантические сети, формируемые в памяти животных, по-видимому, аналогичны семантическим сетям, исследуемым разработчиками искусственного интеллекта [45].

Итак, мы можем думать над эволюционными корнями логики, мышления, интеллекта. И более того, было бы очень интересно попытаться строить модели эволюционного происхождения мышления. По-видимому, наиболее четкий путь такого исследования – построение математических и компьютерных моделей "интеллектуальных изобретений" биологической эволюции, таких как безусловный рефлекс, привыкание (угасание реакции на биологически нейтральный стимул), классический условный рефлекс, инструментальный условный рефлекс, цепи рефлексов, ..., логика [46] (рис. 15). То есть, целесообразно с помощью моделей представить общую картину эволюции когнитивных способностей животных и эволюционного происхождения интеллекта человека.

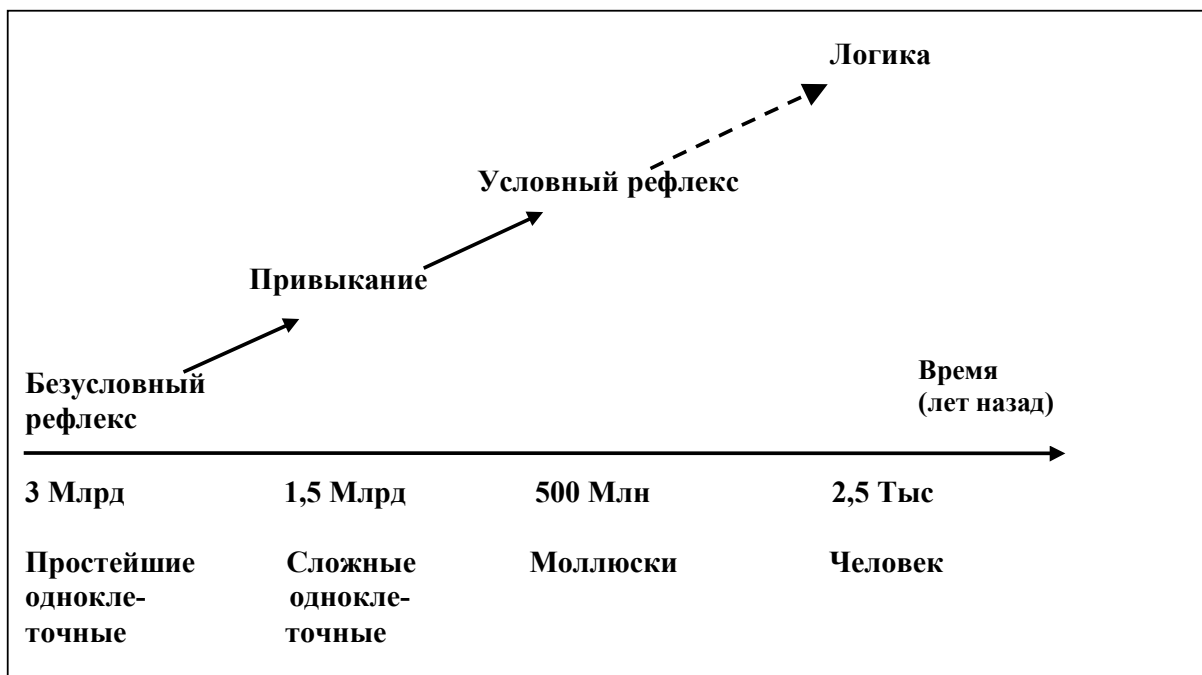


Рис. 15. "Интеллектуальные изобретения" биологической эволюции. "Авторы изобретений" и "даты приоритетов" представлены довольно условно.

Естественно, что такие исследования – это огромный фронт работы, и задачу построения теории происхождения мышления, задачу моделирования когнитивной эволюции можно пока рассматривать как сверхзадачу. Тем не менее, эта задача очень интересна и очень важна с точки зрения развития научного миропонимания. Исследования этой проблемы могли бы обеспечить определенное обоснование применимости нашего мышления в научном познании, то есть, укрепить фундамент всего величественного здания науки.

Подробнее задача моделирования когнитивной эволюции обсуждается в работах [9,47,48].

Отметим, что задача моделирования когнитивной эволюции близка к сформулированной выше (раздел 1) программе-максимум исследований адаптивного поведения, и, естественно, что работы в этой области уже ведутся. Общее состояние моделей адаптивного поведения в контексте исследования когнитивной эволюции примерно таково. Есть множество математических и компьютерных моделей, характеризующих "интеллектуальные" изобретения: модель возникновения безусловного рефлекса на молекулярно-генетическом уровне [49], модели привыкания [50,51], большое количество моделей условных рефлексов [52-56]. Однако эти модели очень фрагментарны, слабо разработаны и пока далеко не формируют общую картину эволюционного происхождения мышления, логики, интеллекта.

Отметим также потенциальное прикладное значение моделирования когнитивной эволюции [57]. Есть направление исследований Искусственный интеллект (ИИ). Это направление испытывает взлеты, падения, периоды энтузиазма и периоды разочарования. Скорее всего, это направление можно рассматривать как прикладное – применение принципов естественного интеллекта в искусственных практически важных для человека компьютерных системах. Судьба прикладных разработок зависит от наличия достаточно серьезного научного фундамента, на котором базируются такие разработки. Например, научной базой развития микроэлектроники во второй половине 20-го века была физика твердого тела. При этом для физиков чисто научные исследования твердого тела были интересны практически независимо от применения их исследований, в результате чего научная основа микроэлектроники интенсивно развивалась. И результаты микроэлектроники, как наукоемкой технологии, впечатляющи.

Моделирование когнитивной эволюции чрезвычайно интересно и важно с точки зрения научного миропонимания. Следовательно, можно ожидать, что такие исследования будут очень интересны для ученых. Но эти исследования могут быть тесно связаны и с разработками ИИ. И, следовательно, могло бы быть взаимное обогащение фундаментальных и прикладных исследований природы интеллекта. И, тем самым, исследования когнитивной эволюции могли бы служить научной основой разработок систем ИИ.

Очерченные в статье работы по моделированию адаптивного поведения могут быть также связаны с широким циклом исследований социально-экономических систем. Эти исследования естественно вести в рамках следующей сверхзадачи.

## **Сверхзадача 2. Исследования эволюции человечества в информационном сообществе.**

В ряде работ последних лет Н.Н. Моисеев высказывал идеи о необходимости создания своеобразного *Совета мудрецов планеты* (вероятно, при ООН или ЮНЕСКО) – научного органа, включающего ученых разных стран и специальностей [58,59]. По мнению Н.Н. Моисеева, такой орган должен стать авторитетным всемирным научным парламентом, обеспечивающим разработки концепций устойчивого развития человечества. К такому органу должны прислушиваться международные организации и руководители отдельных стран.

Понятно, что рекомендации Совета мудрецов должны базироваться на определенных научных разработках. Естественно предположить, что здесь могли бы быть использованы методы исследования адаптивного поведения эволюции коллективов социально-экономических структур.

Чтобы не быть голословным, приведем некоторые результаты [60], отталкиваясь от которых можно предложить важное направление исследований. При исследовании очерченной в разделе 5.2 модели эволюционного формирования целенаправленного поведения в некоторых компьютерных экспериментах наблюдались неожиданные пики в зависимостях численности популяции агентов от времени (рис. 16, нижняя кривая). 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 вывода полная модель (с борьбой между агентами) сравнивалась с моделью, в которой эффекторы борьбы были полностью устранены из системы управления агентов (рис. 16, верхняя кривая). Видно, что в случае такого "глобального пацифизма" установившаяся численность популяции была примерно в 2 раза выше, чем для обычных агентов, имеющих средства для борьбы между собой.

Если пытаться концептуально развить этот подход для социально-экономических систем и 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".



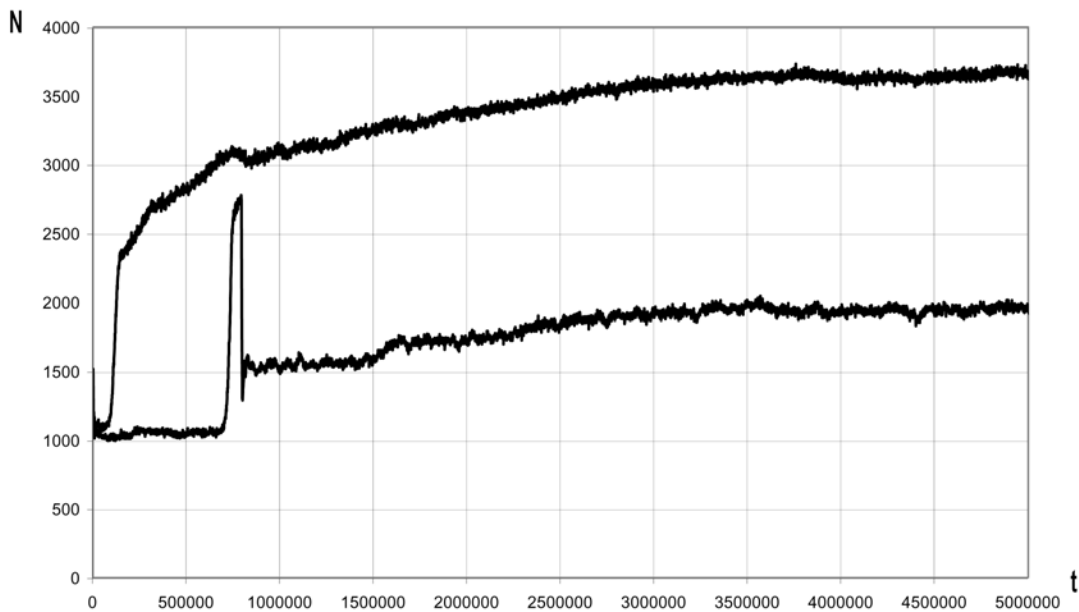


Рис.16. Зависимость численности популяции  $N$  от времени  $t$  в полной модели (нижняя кривая) и в случае "глобального пацифизма" (верхняя кривая).

В последнее время моделирование многоагентных систем активно используется при исследовании социально-экономических структур [61,62]. Достоинства такого моделирования [62]:

- многоагентные системы позволяют исследовать процессы самоорганизации;
- дают возможность естественного описания сложных систем;
- обладают высокой гибкостью.

В качестве примеров использования многоагентных систем для моделирования в социологии можно привести такие исследования, как моделирование информационных потоков в социальных средах [63], коллективного самоопределения [64] или гражданского неповиновения [65]. Ведутся работы по исследованию роли когнитивных процессов в эволюции социально-экономических систем [66,67].

Разработки конкретных моделей сопровождаются серьезным концептуальным анализом. В качестве примера отметим международный Интернет-проект "Principia Cybernetica Project" [68]. Этот проект организовала в начале 1990-х годов небольшая группа исследователей (В.Ф. Турчин, Нью-Йорк, К. Джослин, Лос-Аламос, Ф. Хейлигхен, Брюссель). В проекте делается попытка осмыслить на основе эволюционных концепций вечные вопросы человечества: «Почему мир такой, какой он есть? Откуда мы (люди, человечество) произошли? Кто мы? Куда эволюционирует человечество?». В основу проекта была положена известная книга В.Ф. Турчина "Феномен науки" [44].

Проект Principia Cybernetica [68] представляет собой Интернет-журнал по вопросам эволюции, кибернетики, интеграции человеческого сообщества в единое целое. Проект рассматривается как развивающаяся энциклопедия идей по кибернетической эволюции человечества. Особое внимание уделяется анализу влияния активно развивающихся информационных технологий на эволюцию современного человеческого сообщества.

Одна из наиболее активно обсуждаемых тем в проекте Principia Cybernetica – всемирный мозг [69,70]. Согласно определению Ф. Хейлигхена (Свободный университет Брюсселя) "всемирный мозг" – это метафорическое название возникающей интеллектуальной сети, формируемой человечеством на основе использования компьютеров, баз знаний и связей, которые объединяют все это в единое целое. Эта сеть представляет собой сложную, самоорганизующуюся систему, которая не только обрабатывает информацию, но и постепенно приобретает функции, сходные с

функциями мозга: принятие решений, решение проблем, обучение, формирование новых соединений и открытие новых идей” [71] .

В 2001 году группа специалистов, обсуждающая вопросы всемирного мозга в проекте Principia Cybernetica, организовала конференцию по проблемам всемирного мозга. Программа конференции была хорошо продумана [71], она включала в себя следующие темы:

### **Вопросы теории и технологии**

Поисковые и рекомендующие системы  
Сетевые агенты и сообщества агентов  
Протоколы обмена знаниями  
Представление, обмен и интерпретация распределенных знаний  
Человеко-машинное взаимодействие в коллективных системах  
Анализ сетевой динамики  
Распределенные виртуальные среды  
Распределенная память и распределенные вычисления  
Перспективные сетевые архитектуры  
Поиск знаний и данных в сетях  
Теория сложных адаптивных и эволюционирующих систем  
Глобальное сотрудничество

### **Гуманитарные аспекты**

Образование (дистанционное обучение, электронные университеты)  
Влияние на экономику, социальное развитие и интеграцию  
Влияние на управление государствами (электронная демократия)  
Влияние на права человека и свободу человека  
Коллективное управление знаниями  
Технология сознания  
Глобальное сознание  
Аспекты человеко-сетевого симбиоза  
Сетевая экономика и ее устойчивость  
Интеграция всемирного мозга и глобальной экосистемы  
Человечество как суперорганизм  
Долгосрочные перспективы эволюции человечества

Перечисленные вопросы показывают, что здесь делаются попытки оценить развитие информационно-коммуникационных технологий в контексте эволюции человечества непредвзятым взглядом, с тем, чтобы в дальнейшем открытыми глазами смотреть в будущее развитие этих технологий и в будущее развитие человечества.

### **Краткое обсуждение сверхзадач**

В работе [1] (см. также [72]) обсуждаются общие перспективы развития науки в XXI веке и сформулированы и обсуждаются три сверхзадачи: 1) управление риском и безопасностью сложных систем, 2) нейронаука, 3) теоретическая история.

Области исследований, которые подразумевают эти сверхзадачи, во многом перекрываются с намеченными в настоящей статье. По крупному счету, различие состоит в акцентах и в расстановке приоритетов. Так, конечно же, признавая актуальность исследований по нейронаукам, намеченных в [1], здесь мы смотрим на исследования в области нейронаук под определенным "эволюционным" углом, стараясь выделить наиболее важную для развития научного миропонимания проблему: проблему происхождения мышления, логики, интеллекта человека.

Есть также определенное различие в обсуждаемых в [1] и здесь методах исследования. Обсуждаемые в [1] подходы преимущественно основаны на изучении общих закономерностей на базе методов нелинейной динамики и синергетики. Моделирование адаптивного поведения больше похоже на инженерный подход к анализу естественных "информационных технологий" живых организмов – здесь делается попытка воспроизвести конструкции систем управления, систем обработки информации, обеспечивающих адаптивное поведение животных. И тем самым, качественно воспроизвести с помощью имитационных моделей (моделей, воплощенных в компьютерных программах или в реальных роботах) наиболее характерные и наиболее нетривиальные свойства естественных управляющих систем. То есть, делается попытка вникнуть вглубь конструкций, в природу конструкций биологических информационных систем. Естественно ожидать, что эти два подхода – анализ общих закономерностей методами нелинейной динамики и имитационное моделирование работы систем управления адаптивным поведением – должны быть, в конечном итоге, взаимодополнительны.

Автор благодарен Г.Г. Малинецкому за многочисленные дискуссии, стимулировавшие написание данной статьи.

### Литература:

1. Малинецкий Г. Г. Сценарии, стратегические риски, информационные технологии // Информационные технологии и вычислительные системы. 2002. N 4. С. 83-108.
2. Langton C. G. (Ed.) *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*. – Redwood City CA: Addison-Wesley, 1989.
3. Langton C. G., Taylor C., Farmer J. D., and Rasmussen S. (Eds.) *Artificial Life II: Proceedings of the Second Artificial Life Workshop*. – Redwood City CA: Addison-Wesley, 1992.
4. 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.
5. Цетлин М.Л. Исследования по теории автоматов и моделирование биологических систем. – М.: Наука, 1969. 316 с.
6. Варшавский В.И., Поспелов Д.А. Оркестр играет без дирижера. – М.: Наука, 1984.
7. Бонгард М.М., Лосев И.С., Смирнов М.С. Проект модели организации поведения – Животное // Моделирование обучения и поведения. – М.: Наука, 1975. С.152-171.
8. Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота: модели поведения. – М.: Наука, 1987.
9. Редько В.Г. Эволюционная кибернетика. М.: Наука, 2001, 156 с.
10. Guillot A., Meyer J.-A. From SAB94 to SAB2000: What's New, Animat? // In Meyer et al. (Eds). *From Animals to Animats 6. Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*. The MIT Press. 2000. See also: <http://animatlab.lip6.fr/index.en.html>
11. Непомнящих В.А. Аниматы как модель поведения животных // IV Всероссийская научно-техническая конференция "Нейроинформатика-2002". Материалы дискуссии "Проблемы интеллектуального управления – общесистемные, эволюционные и нейросетевые аспекты". М.: МИФИ, 2003. С. 58-76. См. также <http://www.keldysh.ru/pages/BioCyber/RT/Nepomn.htm>
12. Непомнящих В.А. Поиск общих принципов адаптивного поведения живых организмов и аниматов // Новости искусственного интеллекта. 2002. N. 2. С. 48-53.
13. Donnart J.Y., Meyer J.A. Learning Reactive and Planning Rules in a Motivationally Autonomous Animat // *IEEE Transactions on Systems, Man, and Cybernetics*. Part B: Cybernetics. 1996. V. 26. N.3. See also: <http://animatlab.lip6.fr/index.en.html>
14. 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).
15. Holland J.H., Holyoak K.J., Nisbett R.E., Thagard P. *Induction: Processes of Inference, Learning, and Discovery*. – Cambridge, MA: MIT Press, 1986.

16. Sutton R., Barto A. Reinforcement Learning: An Introduction. – Cambridge: MIT Press, 1998. See also: <http://www-anw.cs.umass.edu/~rich/book/the-book.html>
17. Потапов А.Б., Али М.К. Нелинейная динамика обработки информации в нейронных сетях // Новое в синергетике: Взгляд в третье тысячелетие. (Под ред. Г.Г. Малинецкого и С.П. Курдюмова). М.: Наука, 2002. С. 367-426.
18. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. М.: Эдиториал УРСС, 2002. 352 с.
19. Holland J. H., Booker L.B., Colombetti M., Dorigo M., Forrest S., Goldberg D. G., Riolo R. L., Smith R. E., Lanzi P. L., Stolzmann W., Wilson S. W. What is a Learning Classifier System? // Holland J. H. et. al. (Eds). Learning Classifier Systems. Springer Verlag, 2000. pp. 3-32. See also: [http://www.cs.unm.edu/~forrest/gacs\\_papers.htm](http://www.cs.unm.edu/~forrest/gacs_papers.htm)
20. Klopff A. H. The Hedonistic Neuron: A Theory of Memory, Learning, and Intelligence. Hemisphere, Washington, 1982. 140 p.
21. Сайт AnimatLab: <http://animatlab.lip6.fr/index.en.html>
22. Kodjabachian J., Meyer J.A. Evolution and Development of Modular Control Architectures for 1-D Locomotion in Six-Legged Animats // Connection Science. 1998. V. 10. PP. 211-237. See also: <http://animatlab.lip6.fr/index.en.html>
23. Kodjabachian J., Meyer J.A. Evolution and Development of Neural Controllers for Locomotion, Gradient-Following, and Obstacle-Avoidance in Artificial Insects // IEEE Transactions on Neural Networks 1998. Vol. 9. PP. 796-812. See also: <http://animatlab.lip6.fr/index.en.html>
24. Filliat D., Kodjabachian J., Meyer J.A. Incremental Evolution of Neural Controllers for Navigation in a 6-legged Robot // Sugisaka and Tanaka (Eds). Proceedings of the Fourth International Symposium on Artificial Life and Robotics Oita Univ. Press, 1999. See also: <http://animatlab.lip6.fr/index.en.html>
25. Koza. J. Genetic Programming: On the Programming of Computers by Means of Natural Selection. The MIT Press, 1992.
26. Koza J. Genetic Programming II: Automatic Discovery of Reusable Subprograms. The MIT Press, 1994.
27. Сайт AI Laboratory of Zurich University: <http://www.ifi.unizh.ch/groups/ailab/>
28. Pfeifer R., Scheier C., Understanding Intelligence. MIT Press, 1999.
29. Сайт Laboratory of Artificial Life and Robotics: <http://gral.ip.rm.cnr.it/>
30. Nolfi S., Floreano D. Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines. Cambridge, MA: MIT Press/Bradford Books, 2000. 384 p.
31. Сайт MIT Artificial Intelligence Laboratory: <http://www.ai.mit.edu/>
32. Brooks R.A. Cambrian Intelligence: The Early History of the New AI. MIT Press, 1999.
33. Анохин П.К. Принципиальные вопросы общей теории функциональных систем // Принципы системной организации функций. – М.: Наука, 1973. См. также: <http://www.keldysh.ru/pages/BioCyber/RT/Functional.pdf>
34. Анохин П.К. Системные механизмы высшей нервной деятельности. М.: Наука, 1979. 453 с.
35. Анохин П.К. Очерки по физиологии функциональных систем. – М.: Медицина, 1975.
36. Судаков К.В. (ред.). Теория системогенеза. – М.: Горизонт, 1997.
37. Умрюхин Е.А. Механизмы мозга: информационная модель и оптимизация обучения. М. 1999. 96 с.
38. Моделирование функциональных систем (под ред. Судакова К.В. и Викторова В.А.). – М.: РАМН, РСМАН, 2000. 254 с.
39. 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.
40. Tsitolovsky L.E. A model of motivation with chaotic neuronal dynamics // Journ. of Biological Systems. 1997. V. 5. N.2. PP. 301-323.

41. Бурцев М.С., Гусарев Р.В., Редько В.Г. Модель эволюционного возникновения целенаправленного адаптивного поведения. 1. Случай двух потребностей // Препринт ИПМ РАН. 2000. N. 43. См. также <http://www.keldysh.ru/pages/BioCyber/PrPrint/PrPrint.htm>
42. Бурцев М.С., Гусарев Р.В., Редько В.Г. Исследование механизмов целенаправленного адаптивного управления // Изв. РАН "Теория и системы управления" 2002. №6. С.55-62.
43. Бурцев М.С. Модель эволюционного возникновения целенаправленного адаптивного поведения. 2. Исследование развития иерархии целей // Препринт ИПМ РАН, 2002, № 69.
44. Турчин В.Ф. Феномен науки. Кибернетический подход к эволюции. М.: Наука, 1993. 295с. (1-е изд). М.: ЭТС, 2000. 368 с. (2-е изд). См. также <http://www.refal.ru/turchin/phenomenon/>
45. Искусственный интеллект. – В 3-х кн. Кн.2. Модели и методы. Справочник / Под ред. Д.А. Поспелова. М.: Радио и связь, 1990. 304 с.
46. Воронин Л.Г. Эволюция высшей нервной деятельности. М.: Наука. 1977. 128 с.
47. Red'ko V.G. Evolution of cognition: Towards the theory of origin of human logic // Foundations of Science. 2000, Vol.5. N. 3. PP. 323-338.
48. Редько В.Г. Прологомены к теории происхождения мышления // Статья на сайте Круглого стола IV Всероссийской научно-технической конференции "Нейроинформатика-2002": <http://www.keldysh.ru/pages/BioCyber/RT/Redko/Redko2.htm>
49. Редько В.Г. Адаптивный сайзер // Биофизика. 1990. Т.35. Вып.6. С.1007-1011.
50. Staddon J. E. R. On rate-sensitive habituation // Adaptive Behavior. 1993. Vol. 1. N. 4. PP. 421-436.
51. Guillot A., Meyer J.-A. From SAB90 to SAB94: Four Years of Animat Research. // Cliff et al. (Eds). From animals to animats 3. Proceedings of the Third International Conference on Simulation of Adaptive Behavior. The MIT Press.1994. See also: <http://animatlab.lip6.fr/index.en.html>
52. Ляпунов А.А. О некоторых общих вопросах кибернетики // Проблемы кибернетики. М.: Физматгиз, 1958. Вып. 1. С.5-22.
53. Grossberg S. Classical and instrumental learning by neural networks // Progress in Theoretical Biology. 1974. Vol.3. PP.51-141.
54. Barto A.G., Sutton R.S. Simulation of anticipatory responses in classical conditioning by neuron-like adaptive element // Behav. Brain Res. 1982. Vol.4. P.221-235.
55. Klopff A. H., Morgan J. S., Weaver S. E. A hierarchical network of control systems that learn: modeling nervous system function during classical and instrumental conditioning // Adaptive Behavior. 1993. Vol. 1. N. 3. PP. 263-319.
56. Balkenius C., Moren J. Computational models of classical conditioning: a comparative study // C. Langton and T. Shimohara (Eds.). Proceedings of Artificial Life V. MIT Press, Bradford Books, MA.: 1998. See also: [http://www.lucs.lu.se//Abstracts/LUCS\\_Studies/LUCS62.html](http://www.lucs.lu.se//Abstracts/LUCS_Studies/LUCS62.html)
57. Редько В. Г. Моделирование когнитивной эволюции – естественный путь к искусственному интеллекту // Новости искусственного интеллекта. 2001. N. 2-3. С. 52-56.
58. Моисеев Н.Н. Универсум. Информация. Общество. – М.: Устойчивый мир, 2001. – 200 с.
59. Свободная мысль – XXI. 2002. N. 8. Специальный выпуск, посвященный 85-летию со дня рождения академика Н.Н. Моисеева.
60. Бурцев М.С., Редько В.Г. Влияние агрессии на эволюцию в многоагентной системе // Десятая международная конференция "Проблемы управления безопасностью сложных систем". М., 2002. Тезисы доклада.
61. Epstein J.M., Axtel R.I. Growing Artificial Societies: Social Science from the Bottom Up. Cambridge, MA: MIT Press, 1996.
62. Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems // Proceedings of the National Academy of Sciences. 2002. Vol. 99. Suppl. 3. PP. 7280-7287.
63. Митин Н.А. Новые модели математической психологии и информационные процессы // Новое в синергетике: Взгляд в третье тысячелетие. (Под ред. Г.Г. Малинецкого и С.П. Курдюмова). М.: Наука, 2002. С. 426-460.
64. Lustick I. Agent-based modeling of collective identity: testing constructivist theory // Journal of Artificial Societies and Social Simulations. 2000. Vol. 3. N. 1.

65. Epstein J.M. Modeling civil violence: An agent-based computational approach // Proceedings of the National Academy of Sciences. 2002. Vol. 99. Suppl. 3. PP. 7243-7250.
66. Arthur B.W., Durlauf S., Lane D. (Eds). The Economy as an Evolving Complex System II. Addison-Wesley, Reading, Mass., Series in the Sciences of Complexity, 1997.
67. Arthur B.W. Cognition: The Black Box of Economics // The Complexity Vision and the Teaching of Economics, David Colander (Ed). Edward Elgar Publishing, Northampton, Mass, 2000. Chapter 3.
68. Principia Cybernetica Project: <http://pespmc1.vub.ac.be/> , <http://pcp.lanl.gov/>
69. Heylighen F. Global Brain FAQ // Heylighen F., C. Joslyn and V. Turchin (editors): *Principia Cybernetica Web* (Principia Cybernetica, Brussels), URL: <http://pespmc1.vub.ac.be/GBRAIFAQ.html>
70. Хейлигхен Ф. Мировой Суперорганизм: эволюционно-кибернетическая модель возникновения сетевого сообщества // <http://www.keldysh.ru/pages/BioCyber/Heylig/superorg.htm>
71. F. Heylighen. From Intelligent Networks to the Global Brain: Evolutionary Social Organization through Knowledge Technology. The First Global Brain Workshop (GBrain 0): <http://pespmc1.vub.ac.be/Conf/GB-0.html>
72. Новое в синергетике: Взгляд в третье тысячелетие. (Под ред. Г.Г. Малинецкого и С.П. Курдюмова). М.: Наука, 2002. 478 с.

## **MODELS OF ADAPTIVE BEHAVIOR – NATURAL AND SCIENTIFIC APPROACH TO DEVELOPMENTS OF INFORMATIONAL TECHNOLOGIES**

V.G. Red'ko

Institute of Optical and Neural Technologies, RAS

E-mail: [redko@iont.ru](mailto:redko@iont.ru)

The paper is an analytical review of modern field of investigations "Adaptive Behavior". The goal of these researches is to study of principles and architectures of such control systems of animats (modeled artificial organisms) those should ensure adaptation of animats to permanently variable external environment. The paper pays a special attention to researches of R. Sutton and A. Barto (Massachusetts University) on Reinforcement Learning and to researches of AnomatLab (Paris) on animat control systems based on symbiosis of 1) individual learning, 2) ontogenetic development and 3) evolutionary optimization. The paper includes also the description of the original models of evolutionary origin of purposeful adaptive behavior. On the basis of the review, the discussion on future of science, initiated by G.G. Malinetsky, is developed.