

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
МИНИСТЕРСТВО ПРОМЫШЛЕННОСТИ, НАУКИ И ТЕХНОЛОГИЙ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
РОССИЙСКАЯ АССОЦИАЦИЯ НЕЙРОИНФОРМАТИКИ  
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ  
(ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ)

---

**НАУЧНАЯ СЕССИЯ МИФИ–2002**

**НЕЙРОИНФОРМАТИКА–2002**

**IV ВСЕРОССИЙСКАЯ  
НАУЧНО-ТЕХНИЧЕСКАЯ  
КОНФЕРЕНЦИЯ**

**ЛЕКЦИИ  
ПО НЕЙРОИНФОРМАТИКЕ**

**Часть 1**

По материалам Школы-семинара  
«Современные проблемы нейроинформатики»

Москва 2002

УДК 004.032.26 (06)

ББК 32.818я5

М82

**НАУЧНАЯ СЕССИЯ МИФИ–2002. IV ВСЕРОССИЙСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ «НЕЙРОИНФОРМАТИКА–2002»: ЛЕКЦИИ ПО НЕЙРОИНФОРМАТИКЕ. Часть 1.** – М.: МИФИ, 2002. – 164 с.

В книге публикуются тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 23–25 января 2002 года в МИФИ в рамках IV Всероссийской конференции «Нейроинформатика–2002».

Материалы лекций связаны с рядом проблем, актуальных для современного этапа развития нейроинформатики, включая ее взаимодействие с другими научно-техническими областями.

Ответственный редактор

*Ю. В. Тюменцев*, кандидат технических наук

ISBN 5–7262–0400–X

© *Московский государственный инженерно-физический институт (технический университет), 2002*

## Содержание

<b><i>В. Г. Редько. Эволюционная кибернетика</i></b>	<b>29</b>
Введение . . . . .	29
Гносеологическая проблема и подходы к ее решению . . . . .	31
Философия глазами физика — лирическое отступление . . . . .	31
На пути к теории происхождения логики — естественно-научный подход к теории познания . . . . .	33
«Интеллектуальные изобретения» биологической эволюции — путь к логике . . . . .	34
Направления исследований «Искусственная жизнь» и «Адаптивное поведение» . . . . .	40
Искусственная жизнь — жизнь, какой она могла бы в принципе быть . . . . .	40
Типичные модели искусственной жизни . . . . .	41
From Animal to Animat — модели адаптивного поведения животного и робота . . . . .	45
Общая характеристика направлений «Искусственная жизнь» и «Адаптивное поведение» . . . . .	45
Функциональная система по П. К. Анохину — общая схема адаптивного поведения . . . . .	48
«Кузнечик» — модель эволюционного возникновения целенаправленного адаптивного поведения . . . . .	52
Модель искусственной жизни в Интернете . . . . .	54
Описание модели . . . . .	54
Результаты моделирования . . . . .	58
О перспективах эволюционной кибернетики . . . . .	59
Благодарность . . . . .	61
Литература . . . . .	61

**В. Г. РЕДЬКО**

Институт прикладной математики им. М. В. Келдыша РАН, Москва

**E-mail: redko@keldysh.ru**

## **ЭВОЛЮЦИОННАЯ КИБЕРНЕТИКА**

### **Аннотация**

Обсуждается проблема моделирования эволюционного происхождения интеллекта. Сделан обзор направлений исследований «Искусственная жизнь» и «Адаптивное поведение». Излагаются две конкретные модели: модель эволюционного возникновения целенаправленного адаптивного поведения и модель искусственной жизни в Интернете.

**V. G. RED'KO**

M. V. Keldysh Institute for Applied Mathematics, RAS, Moscow

**E-mail: redko@keldysh.ru**

## **EVOLUTIONARY CYBERNETICS**

### **Abstract**

A modelling problem for evolutionary intelligence emergence is discussed. Investigations in the “Artificial Life” and “Adaptive Behavior” fields are reviewed. Two particular models are described: a model of evolutionary emergence for purposeful adaptive behavior and a model of artificial life in Internet.

## **Введение**

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 биоконструктивных систем и познавательных способностей животных привело к возникновению интеллекта человека? Какие уроки из знаний о естественных «биоконструктивах» можно извлечь для разработки искусственных компьютеров и программных продуктов? До какой степени исследования причин возникновения естественного интеллекта могут способствовать развитию искусственного интеллекта?

Может ли какая-либо научная дисциплина ответить на все эти интригующие вопросы? По мнению автора, такая научная дисциплина только формируется. Эту дисциплину можно назвать «эволюционная кибернетика». Основным предметом ее исследования мог бы стать теоретический анализ (на основе математических и компьютерных моделей) эволюции биологических систем обработки информации и кибернетических свойств живых организмов.

Общее рассмотрение математических и компьютерных моделей эволюционной кибернетики, характеризующих, в частности, 1) эволюционное возникновение биологических кибернетических систем в ранней биосфере Земли, 2) молекулярно-генетическую эволюцию в целом, 3) прикладные эволюционные алгоритмы, проведено в работе [1]. В данной лекции мы сконцентрируем внимание на наиболее интересной из возникающих здесь проблем: нельзя ли промоделировать эволюцию познавательных способностей животных и подойти к моделированию эволюционного возникновения интеллекта? Для конкретности будут также рассмотрены две новые модели из области «Искусственной жизни».

Задача моделирования эволюционного возникновения интеллекта связана с глубокой гносеологической проблемой: почему мышление, логика, интеллект человека применимы к познанию природы? Поскольку сейчас придется говорить о философии для явно нефилософской аудитории, то в свое оправдание сделаю небольшое «лирическое отступление».

## Гносеологическая проблема и подходы к ее решению

### Философия глазами физика — лирическое отступление

Когда я учился на втором курсе МФТИ, мне попала в руки хорошо написанная биография Альберта Эйнштейна. Из этой книги я неожиданно для себя узнал, что Эйнштейн не только занимался физикой и математикой, но и серьезно интересовался философией. Более того, по словам А. Эйнштейна, изучение философии способствовало его научной работе. Особенно ему помогли работы Д. Юма и Э. Маха. Этот интерес Эйнштейна к философии заинтриговал меня и я решил почитать трактаты тех философов, которых изучал Эйнштейн.

Это было во второй половине 1960-х годов, в самый «расцвет застоя». К философии у нас, студентов-физиков, было скептическое отношение, как «неестественной» науке, которую нас заставляли изучать в принудительном порядке. Высшее образование подразумевало обязательное изучение довольно ограниченного курса марксистско-ленинской философии. Тем не менее, в те застойные времена в книжных магазинах наряду с серьезной и обширной физико-математической литературой можно было купить собрания сочинений Платона, Аристотеля, Б. Спинозы, Д. Юма, И. Канта и многих других — философская классика тогда издавалась достаточно хорошо. Так что свою любознательность — что же полезного мог вынести А. Эйнштейн из трудов философов — я вполне мог попытаться удовлетворить.

Начал я с трудов нидерландского философа XVII века Бенедикта Спинозы — этого философа тоже читал Эйнштейн, и я, подражая ему, решил познакомиться с трактатами Спинозы. Сначала меня удивило, что серьезная работа мысли возможна не только в естественных науках, но и в философии — при изучении принудительного курса общественных наук как-то этого не чувствовалось, а при прикосновении к классике философии это сразу бросилось в глаза. Но самое интересное произошло дальше — однажды при чтении «Краткого трактата о боге, человеке и его счастье» Б. Спинозы я задумался над вопросом: «А почему, собственно говоря, математика применима к физике?»

Действительно, физик, изучая природные явления, использует доказанные математиками результаты (теоремы, методы решения уравнений и т. п.). Математики же обосновывают эти результаты чисто логическим

путем, совершенно независимым от реальной природы. Так почему же результаты математиков применимы к природным явлениям?

Поясним этот вопрос простым примером.

Допустим, физик, изучая динамику некоторого объекта, сумел в определенном приближении свести описание объекта к дифференциальному уравнению (рис. 1). Далее он интегрирует полученное уравнение согласно известным из математики правилам и получает характеристики движения объекта (в рамках используемого приближения). Переход от дифференциального уравнения к характеристикам движения происходит в сознании физика, а используемые при этом правила интегрирования получены логическим путем математиками. Переход этот чисто дедуктивный, но, если быть предельно строгим, применимость данного перехода к физическому объекту надо обосновывать: поведение физического объекта совершенно необязательно должно соответствовать правилам человеческой логики.

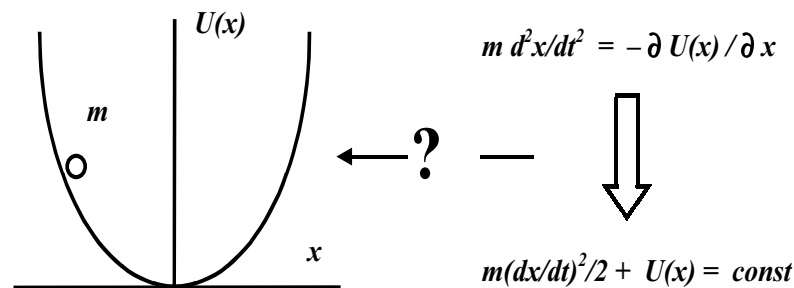


Рис. 1. Почему *человеческая* логика применима к познанию *природы*?

Отметим, что проблема применимости математического знания в естественных науках в нескольких близких контекстах обсуждалась рядом ученых. Например, М. Клайн в книге «Математика. Поиск истины», посвященной исследованию природы математического знания, задает вопрос: «Почему теоремы, доказанные человеческим разумом в тиши кабинетов, должны быть применимы к реальному миру. . .?» [2]. Проблема природы математического знания и «непостижимой эффективности ма-

тематики в естественных науках» обсуждалась такими известными учеными как А. Пуанкаре и Ю. Вигнер [3, 4].

В более общей формулировке рассматриваемую гносеологическую проблему следует поставить так: почему *человеческая* логика применима к познанию *природы*? Действительно, с одной стороны, логические процессы вывода происходят в нашем, человеческом мышлении, с другой стороны, процессы, которые мы познаем посредством логики, относятся к изучаемой нами природе. Эти два типа процессов различны. Поэтому далеко не очевидно, что мы можем использовать процессы первого типа для познания процессов второго типа.

Рассматриваемая гносеологическая проблема — фундаментальная проблема. Она касается принципиальных возможностей человеческого познания, и, в особенности, возможностей научного познания природы. И, следовательно, эта проблема заслуживает тщательного исследования.

Как же подойти к решению этого вопроса? По моему мнению, наиболее естественный путь — проанализировать процесс возникновения познавательных способностей животных в биологической эволюции и постараться понять: как и почему возникали «интеллектуальные» способности, обеспечивающие познание природы.

#### **На пути к теории происхождения логики — естественнонаучный подход к теории познания**

Действуя максимально упрощенно, мы можем рассуждать следующим образом. В процессе биологической эволюции животные приобретали способности познавать внешний мир. Эти познавательные способности помогали им адаптироваться, приспосабливаться к окружающей среде. Приспособленности тех животных, которые приобретали «хорошие» познавательные способности, увеличивались. Животные с повышенными приспособленностями распространялись в популяциях, следовательно, «хорошие» познавательные способности фиксировались эволюцией. Шаг за шагом, эволюционно возникали все более сложные и более интеллектуальные способности, что, в конечном итоге, привело к появлению человеческого мышления, человеческой логики.

Но этих простых рассуждений явно недостаточно. Проблема принципиальной способности познавать природу — фундаментальная философ-



ская проблема, касающаяся *обоснования всего научного познания*, и она должна быть проанализирована настолько глубоко, насколько это возможно.

Логика (в общем смысле: дедуктивная и индуктивная) — наиболее четкая часть, составляющая мышления. Более того, для логики есть определенное формальное описание, есть формулы логики, и мы можем думать над тем, как такие формулы (или их предшественники, прототипы) постепенно появлялись в процессе биологической эволюции в «мышлении» животных. Так что наиболее интересная задача, которая должна быть исследована, может быть поставлена в следующей форме: как и почему в процессе биологической эволюции возникли логические системы, обеспечивающие научное познание природы?

Для понимания процесса возникновения логики и осмысления того, как и почему в этом процессе появились логические формы, обеспечивающие познание природы, имеет смысл построить модельную теорию эволюционного происхождения человеческой логики. Такая теория могла бы включать математические модели наиболее важных «интеллектуальных изобретений» биологической эволюции, посредством которых животные «познают» закономерности во внешнем мире, а также модели эволюционных переходов между «интеллектуальными изобретениями» разных уровней.

Надежнее всего начать с «самого начала» — с происхождения жизни — и проследить весь путь биологической эволюции от простейших до человека, выделяя на этом пути наиболее важные эволюционные «изобретения», ведущие к логике. Чтобы представить круг вопросов, которые могут составить предмет модельной теории происхождения логики, отметим некоторые уровни «интеллектуальных изобретений», укажем соответствующие им примеры, а также модели, характеризующие эти уровни.

#### **«Интеллектуальные изобретения» биологической эволюции — путь к логике**

**Уровень первый** — организм различает состояние среды, память об этих состояниях записана в геноме и передается по наследству, организм адекватно использует различие сред, меняя свое поведение с изменением среды.

Пример этого уровня — свойство регулирования синтеза белков живой клеткой в ответ на изменение питательных веществ во внешней среде по схеме Ф. Жакоба и Ж. Моно [5]. Например, бактерия кишечной палочки обычно питается глюкозой, но если нет глюкозы, а есть лактоза, то в бактерии включается синтез специальных ферментов, перерабатывающих лактозу в глюкозу, что и обеспечивает жизнь бактерии в лактозной среде.

Простая модель такого свойства — адаптивный сайзер<sup>1</sup> — разработана в [6]. Согласно модели адаптивные сайзеры имеют селективные преимущества перед подобными им неадаптивными, однако не всегда, а только в тех случаях, когда чередование сред происходит достаточно часто. Если же среда постоянна и богата питанием, то неадаптивные сайзеры имеют селективные преимущества перед адаптивными, так как последние вынуждены постоянно расходовать свои жизненные силы на синтез веществ, отслеживающих состояние внешней среды.

Описанное свойство — это фактически *безусловный рефлекс* на молекулярно-генетическом уровне.

**Второй уровень** — временное запоминание организмом состояния среды и адекватное (также временное) приспособление к среде.

Пример этого уровня — *привыкание*, а именно, постепенное угасание реакции раздражения на биологически нейтральный стимул. Этот простейший приобретаемый навык появляется на уровне сложных одноклеточных организмов. Охарактеризуем кратко это свойство, следуя опытам В. Кинастовского на инфузориях [7]. В этих опытах на инфузорию воздействовали биологически нейтральным стимулом (падающая капля воды) и наблюдали реакцию инфузории. Сначала в ответ на действие стимула у инфузории возникает реакция раздражения. Но, если этот стимул повторяется многократно, то реакция раздражения постепенно угасает. Отметим, что привыкание отличается от утомления: если на инфузорию, привыкшую к падающей капле воды, подействовать другим нейтральным стимулом, то реакция раздражения восстанавливается. Согласно опытам В. Кинастовского, время выработки привыкания у инфузорий составляет 10–30 минут, сохраняется привыкание в течение 1–3 ч. Память о раздра-

---

<sup>1</sup>Сайзер (от SYstem of SElf-Reproduction) — Самовоспроизводящаяся система ферментов и кодирующих их полинуклеотидов. Эта модель была предложена новосибирскими учеными В. А. Ратнером и В. В. Шаминам в 1980 году (см. также [1]). — Прим. ред.

жителях, выработанная в процессе привыкания у одноклеточных осуществляется, по-видимому, за счет реорганизации химических взаимодействий в цитоплазме клеток [8]. Интересно, что свойством привыкания обладают отдельные нервные клетки [9].

Подчеркнем, что привыкание — простейшее свойство *индивидуального приспособления*. Память о состояниях внешней среды, формируемая при привыкании, кратковременная.

Модели автоматов, способных временно запоминать состояния внешней среды и использовать приобретаемый при этом опыт, уровень «интеллектуальности» которых примерно соответствует уровню привыкания, были разработаны и разносторонне исследованы М. Л. Цетлиным [10] и его последователями. В последнее время модели привыкания строятся исследователями направления «Адаптивное поведение» (см. обзор [11]).

**Третий уровень** — запоминание устойчивых связей между событиями в окружающей организм природе.

Пример — классический *условный рефлекс* [12], в котором происходит долговременное запоминание связи между условным и безусловным стимулами. Скажем, собака запоминает связь между звуковым сигналом (условный стимул) и пищей (безусловный стимул). Формирование условного рефлекса (УР) происходит в три стадии [13] (рис. 2). Первая стадия — прегенерализация, во время которой еще нет реакции на условный стимул, однако повышается электрическая активность разных областей мозга. За ней следует стадия генерализации, при которой реакция возникает как на условный стимул, так и на различные подобные ему (дифференцировочные) раздражители. Затем происходит специализация, при которой реакция на дифференцировочные стимулы постепенно ослабевает, и сохраняется только реакция на условный стимул.

Память о связи между условным и безусловным стимулами долговременная: УР сохраняется в течение многих недель у низших позвоночных и до нескольких лет, а может быть, и всю жизнь у высших животных [14]. Характерное свойство классического УР — самовосстановление, наступающее через несколько десятков минут или часов после угасания.

Биологическое значение условного рефлекса — *предвидение событий* во внешней среде и адекватное использование этого предвидения. Появляется классический условный рефлекс примерно на уровне моллюсков.

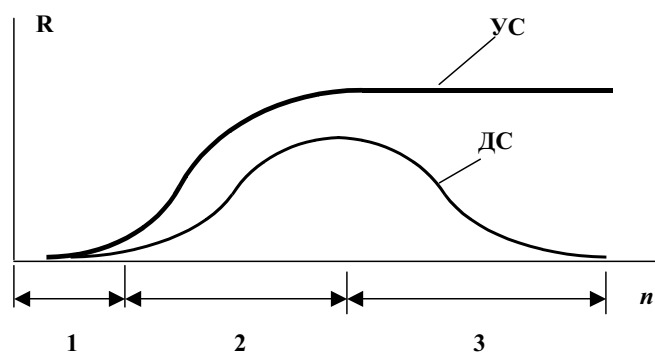


Рис. 2. Зависимость реакции на условный (УС) и дифференцировочный (ДС) стимулы от числа сочетаний  $n$  между условным и безусловным стимулами при выработке классического условного рефлекса. Очень схематично. 1 — прегенерализация, 2 — генерализация, 3 — специализация.

Математические и кибернетические модели образования связи между условным и безусловным стимулами при выработке УР предлагались рядом авторов. Одна из наиболее простых моделей, описывающая выработку и угасание УР, принадлежит А. А. Ляпунову [15]. Модели условного рефлекса разрабатывали такие известные кибернетики, как С. Гроссберг [16], А. Барто и Р. Саттон [17]. Популярный обзор математических и кибернетических моделей УР сделан М. Г. Гаазе-Рапопортом и Д. А. Поспеловым [18]. В последнее время достаточно серьезный анализ моделей УР сделали Х. Балкениус и Дж. Морен [19]. Однако, насколько автору известно, несмотря на изобилие моделей, до сих пор нет математической модели, единым образом отражающей основные свойства УР — самовосстановление, генерализацию, реакцию на дифференцировочные стимулы — и соответствующей биологическому смыслу УР (предвидение будущих событий и использование этого предвидения). Отметим, что аналогичная точка зрения высказана и авторами работы [19].

Между классическим условным рефлексом и логикой лежит целый ряд *промежуточных уровней*. Здесь мы только упомянем некоторые из них.

**Инструментальный условный рефлекс** отличается от классического тем, что здесь для получения поощрения животному необходимо совершить заранее неизвестное ему действие.

**Цепь условных рефлексов** – система реакций, формирующаяся на основе ранее хранившихся в памяти животного условных связей.

Начиная с некоторого уровня, у животных возможно формирование **моделей внешнего мира** [20, 21].

Последний из рассматриваемых уровней – **логика**. Пример этого уровня – системы логического вывода [22]. Примеры правил, входящих в эти системы:

$$\frac{A \Rightarrow B, B \Rightarrow C}{A \Rightarrow C}$$

(если из А следует В и из В следует С, то из А следует С),

$$\frac{A \Rightarrow B, \neg B}{\neg A}$$

(если из А следует В и В ложно, то А ложно).

Математические модели логики широко разработаны: есть исчисление высказываний, исчисление предикатов, математические теории логического вывода [22–24]; активно ведутся работы по математической формализации индуктивного вывода [25–27].

Можно отметить, что математическая логика дает ответы на вопросы: «Каковы правила человеческой логики?» и «Как использовать правила логики?» Рассматриваемая здесь пока лишь чисто умозрительно теория происхождения логики могла бы дать ответы на более глубокие вопросы: «Почему правила человеческой логики таковы, каковы они есть?» и «Почему правила логики могут корректно использоваться?»

Итак, можно выделить несколько ключевых «изобретений» и расположить их в последовательный ряд эволюционных достижений Природы (рис. 3). В этом ряду происходит постепенное усовершенствование «познания» закономерностей в окружающей среде. Что же сделано в области теоретического моделирования эволюции «интеллектуальных изобретений» и осмысления того, как в процессе биологической эволюции возникали и развивались познавательные свойства биологических организмов? Общая ситуация примерно такова. Есть множество математических

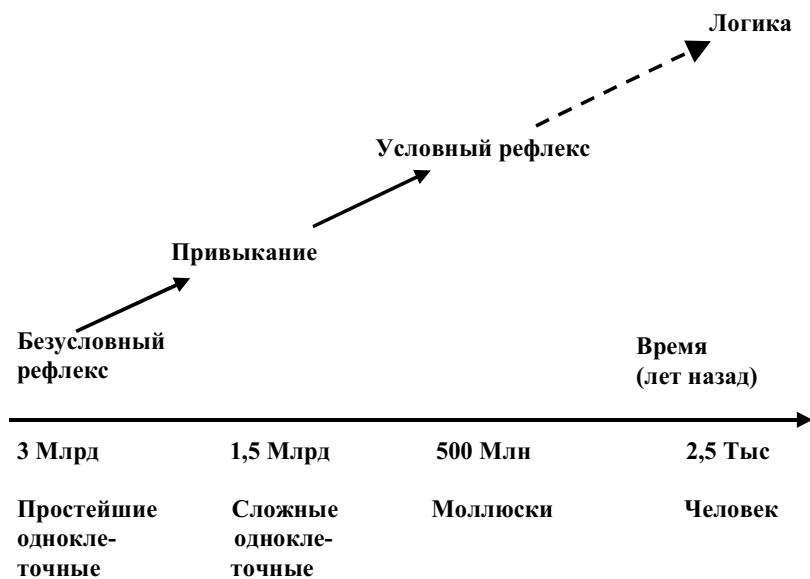


Рис. 3. «Интеллектуальные изобретения» биологической эволюции. «Авторы изобретений» и «даты приоритетов» представлены довольно условно.

и компьютерных моделей, характеризующих «интеллектуальные изобретения»: модель возникновения безусловного рефлекса на молекулярно-генетическом уровне [6], модели привыкания [11], большое количество моделей условных рефлексов [11, 15–19]. Однако эти модели очень фрагментарны, слабо разработаны и не формируют общую картину эволюционного происхождения логики.

Пока можно говорить только об определенном заделе в направлении теоретических исследований «интеллектуальных изобретений» биологической эволюции. Но этот задел активно развивается. Сравнительно недавно (конец 1980 – начало 1990 годов) появились два интересных направления исследований «Искусственная жизнь» и «Адаптивное поведение», которые тесно связаны с моделированием «интеллектуальных

изобретений» биологической эволюции. Дадим краткую характеристику этим направлениям.

### **Направления исследований «Искусственная жизнь» и «Адаптивное поведение»**

#### **Искусственная жизнь — жизнь, какой она могла бы в принципе быть**

В конце 80-х – начале 90-х годов возникли два интересных, тесно связанных между собой направления кибернетических исследований: «Искусственная жизнь» (английское название Artificial Life или ALife) [28, 29] и «Адаптивное поведение» (Adaptive Behavior) [30].

Основной мотивацией исследований искусственной жизни служит желание понять и промоделировать формальные принципы организации биологической жизни. Как сказал руководитель Первой международной конференции по искусственной жизни К. Лангтон, «основное предположение искусственной жизни состоит в том, что “логическая форма” организма может быть отделена от материальной основы его конструкции».

Сторонники направления «Искусственная жизнь» часто считают, что они исследуют более общие формы жизни, чем те, которые существуют на Земле. То есть изучается жизнь, какой она могла бы в принципе быть (“life-as-it-could-be”), а не обязательно та жизнь, какой мы ее знаем (“life-as-we-know-it”).

Искусственная жизнь — это синтетическая биология, которая по аналогии с синтетической химией пытается воспроизвести биологическое поведение в различных средах. Это жизнь, созданная человеком, а не природой (“life made by Man rather than by Nature”). Исследования искусственной жизни направлены не только на теоретические исследования свойств жизни, но и (подобно синтетической химии) на практические приложения, такие как подвижные роботы, медицина, нанотехнология, исследования «жизни» социальных и экономических систем и т. п.

Большую роль в исследованиях искусственной жизни играет математическое и компьютерное моделирование. Очень часто «организмы» в искусственной жизни — это придуманные людьми объекты, «живущие» в мире компьютерных программ.

Отметим, что хотя лозунг «Искусственная жизнь» был провозглашен в конце 80-х годов, в действительности идейно близкие модели разрабатывались уже 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» [10, 31].

В 60–70-х годах под руководством талантливого кибернетика М. М. Бонгарда была построена весьма нетривиальная модель «Животное», характеризующая адаптивное поведение искусственных организмов, живущих на плоскости, разбитой на клетки и обладающих рядом конкурирующих между собой потребностей [18, 32].

### Типичные модели искусственной жизни

Приведем некоторые примеры характерных исследований искусственной жизни:

- Исследование динамики жизнеспособных структур в клеточных автоматах (К. Лангтон) [33].
- ПолиМир (PolyWorld) Л. Ягера — компьютерная модель искусственных организмов, которые имеют структурированную нейронную сеть, обладают цветовым зрением, могут двигаться, питаться (и увеличивать тем самым свою энергию), могут скрещиваться и бороться друг с другом [34]. При моделировании эволюции в ПолиМире возникал ряд нетривиальных стратегий поведения организмов.
- Тьерра (Tierra) Т. Рэя — модель эволюции самовоспроизводящихся компьютерных программ [35]. «Организмы» Тьерры содержат геномы, которые определяют инструкции исполнительных программ. Взаимодействия между организмами приводят к эволюционному возникновению сложного «биоразнообразия» самовоспроизводящихся программ.



- Авида (Avida) К. Адами с сотрудниками [36]; эта модель — развитие модели Тьерра. По сравнению с Тьеррой Авида проще и обладает большей общностью. Модель исследовалась аналитическими методами. Были получены характеристики распределения особей в эволюционирующих популяциях. Исследования на модели Авида количественно поддерживают ту точку зрения, что эволюция движется скачками, а не непрерывно.
- Анализ взаимодействия между обучением и эволюцией, выполненный Д. Экли и М. Литтманом [37]. Эта работа продемонстрировала, что обучение и эволюция вместе более успешны в формировании адаптивной популяции, чем обучение либо эволюция по отдельности.
- Эхо (Echo) Дж. Холланда [38]. Эта модель описывает эволюцию простых агентов, которые взаимодействуют между собой путем скрещивания, борьбы и торговли. Взаимодействие между агентами приводит к формированию различных экологических систем: «войны миров», симбиозов и т. п.
- Модель эволюции двух конкурирующих популяций, одна из которых есть популяция программ, решающих определенную прикладную проблему (задачу сортировки), а вторая — популяция задач, эволюционирующих в направлении усложнения проблемы (Д. Хиллис) [39]. Первая из популяций может рассматриваться как популяция особей-хозяев, а вторая — как популяция паразитов. Моделирование показало, что коэволюция в системе «паразит-хозяин» приводит к нахождению значительно лучших решений проблемы по сравнению с тем решением, которое можно найти в результате эволюции одной только первой популяции (популяции особей-хозяев).
- Модели эволюции клеточных автоматов, например модели М. Митчелл с сотрудниками, описывающие эволюционный поиск клеточных автоматов, которые могут выполнять простые вычисления [40].
- «Муравьиная ферма» (AntFarm) Р. Коллинза и Д. Джефферсона. Эта модель разработана на базе «Коннекшен-машины» (Connection Machine). Модель имитирует поведение при поиске пищи в огромных эволюционирующих популяциях искусственных муравьев [41].

- Классифицирующие системы Дж. Холланда с сотрудниками [42]. Это модель эволюции когнитивного процесса. Классифицирующая система есть система индуктивного вывода, которая основана на использовании набора логических правил. Каждое правило имеет следующую форму: «если *условие*, то *действие*». Система правил оптимизируется как посредством обучения, так и эволюционным методом. В процессе обучения меняются приоритеты использования правил (т.е. меняются коэффициенты, характеризующие «силу» правил). При обучении используется так называемый алгоритм «пожарной бригады»: при успехе поощряются не только те правила, которые непосредственно привели к успешному действию, но и те, которые были предшественниками успеха. Поиск новых правил осуществляется эволюционным методом.

Модели искусственной жизни — активно развивающаяся область исследований. Большинство моделей — остроумные компьютерные эксперименты. Серьезное математическое описание здесь только начинается. Хороший пример серьезного математического исследования — работы К. Адами с сотрудниками по анализу распределения особей в эволюционирующих популяциях [43]. Этот анализ основан на теории самоорганизованной критичности [44] и разумно интерпретирует как компьютерные эксперименты на моделях Тьерра и Авида, так и реальные биологические данные.

Исследования искусственной жизни тесно связаны с другими интересными направлениями: моделями происхождения жизни, автоматами С. А. Кауффмана [45], работами по прикладному эволюционному моделированию, по теории нейронных сетей. Эволюция популяций искусственных организмов — одно из ведущих направлений исследований искусственной жизни. Модели эволюции здесь часто основаны на генетическом алгоритме [46]. Правда, в моделях искусственной жизни часто не вводится явно функция приспособленности, как это обычно делается в генетическом алгоритме. Приспособленность проявляется естественным путем: особи рождаются, когда их родители готовы дать потомков, и погибают, когда не хватает пищи или когда их убивает и съедает хищник. В этом случае — при отсутствии явной функции приспособленности — говорят, что приспособленность эндогенна.

Управление поведением искусственных организмов часто моделируется с помощью нейронных сетей.

Модели искусственной жизни проливают новый свет на эволюционные явления. Отличный пример, иллюстрирующий этот тезис — исследование *эффекта Болдуина*. Более 100 лет назад, в 1896 году Дж. Болдуин [47] предложил механизм, согласно которому первоначально приобретенные навыки организмов могут в дальнейшем стать наследуемыми. Эффект Болдуина работает в два этапа. На первом этапе эволюционирующие организмы (благодаря соответствующим мутациям) приобретают свойство обучиться некоторому полезному навыку. Приспособленность таких организмов увеличивается, следовательно, они распространяются по популяции. Но обучение имеет свои недостатки, так как оно требует энергии и времени. Поэтому возможен второй этап (который называют *генетической ассимиляцией*): приобретенный полезный навык может быть «повторно изобретен» генетической эволюцией, в результате чего он записывается непосредственно в геном и становится наследуемым. Второй этап длится множество поколений; устойчивая окружающая среда и высокая корреляция между генотипом и фенотипом облегчают этот этап. Таким образом, полезный навык, который первоначально был приобретенным, может стать наследуемым, хотя эволюция имеет дарвиновский характер. Ряд исследователей (Г. Хинтон и С. Новлан, Д. Экли и М. Литтман, Г. Мейлей и многие другие) анализировали эффект Болдуина. Они показали, что этот эффект может играть существенную роль в процессе эволюции искусственных особей. Подробнее см. сборники статей [48, 49].

В Институте прикладной математики им. М. В. Келдыша РАН мы начали работу над моделью эволюционного возникновения целенаправленного адаптивного поведения [50] в духе «Искусственной жизни». Модель может рассматриваться как развитие работ Л. Ягера [34], а также Д. Экли и М. Литтмана [37]. Мы кратко опишем нашу модель ниже (см. с. 52–53).

**From Animal to Animat — модели адаптивного поведения животного и робота**

С начала 90-х годов активно развивается направление «Адаптивное поведение» [11, 30]. Основной подход этого направления — конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы называются «аниматами». Термин «анимат» происходит от слов *animal* (животное) и *robot*: ANIMAL + ROBOT = ANIMAT.

Поведение аниматов имитирует поведение животных. Исследователи направления «Адаптивное поведение» 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 [51].

Как и для «Искусственной жизни», для исследований «Адаптивного поведения» 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.

**Общая характеристика направлений «Искусственная жизнь» и «Адаптивное поведение»**

Направления «Искусственная жизнь» и «Адаптивное поведение» 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.

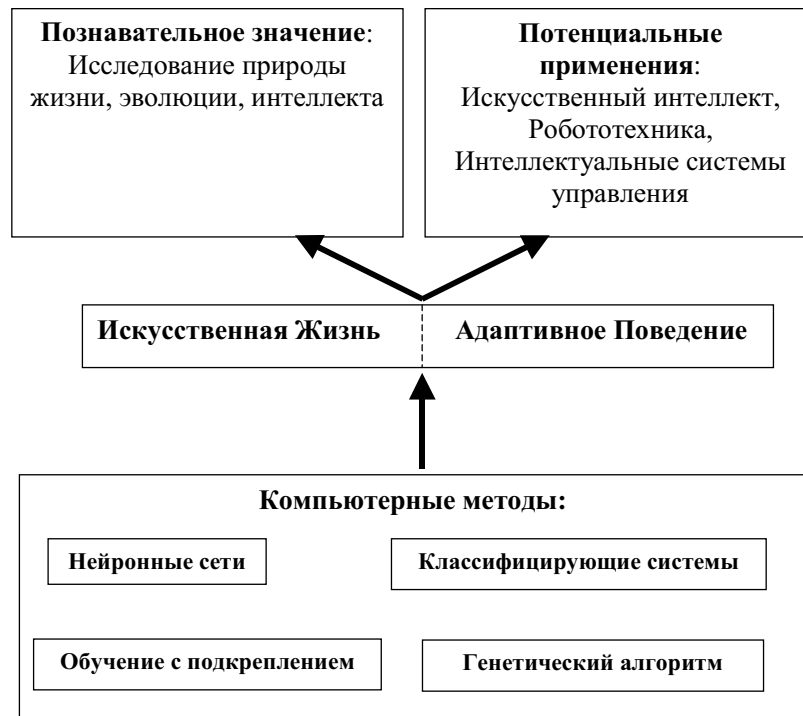


Рис. 4. Схема междисциплинарных связей направлений «Искусственная жизнь» и «Адаптивное поведение».

Эти направления используют ряд нетривиальных компьютерных методов: нейронные сети, классифицирующие системы (Classifier Systems) [42], обучение на основе подкрепления (Reinforcement Learning) [52], генетический алгоритм [46] и другие методы эволюционной оптимизации.

Отметим, что обучение на основе подкрепления (Reinforcement Learning) [52] — это самостоятельное серьезное направление кибернетических исследований. Достоинство этого метода — его сравнительная простота: наблюдают действия обучаемого объекта и в зависимости от ре-

зультата поощряют, либо наказывают данный объект. То есть учитель поступает с обучаемым объектом примитивно: «бьет кнутом» (если действия объекта ему не нравятся), либо «дает пряник» (в противоположном случае), не объясняя обучаемому объекту, как именно нужно действовать. Подчеркнем, что роль учителя может играть внешняя среда. В этом методе, как и в классифицирующих системах, большое внимание уделяется поощрению/наказанию не только текущих действий, которые непосредственно привели к положительному/отрицательному результату, но и тех действий, которые предшествовали текущим.

При моделировании поведения искусственных организмов часто используются эволюционные методы, это позволяет находить конкретные структуры управления «организмов» естественным, самоорганизующимся путем, без навязывания воли автора модели процессу конструирования.

Отметим, что исследования «Искусственной жизни» и «Адаптивного поведения» относятся не к «микроскопическому» уровню (уровню детальной структуры нейросети, составленной из отдельных нейронов), а к поведенческому, феноменологическому уровню.

Подчеркнем, что исследование поведения на феноменологическом уровне вполне имеет право на существование. Приведем аналогию с физикой. Пример феноменологического рассмотрения в физике — термодинамика (изучение свойств физических тел в терминах макроскопических понятий: температуры, энергии, энтропии, объема, и т. п.). Микроскопический аналог термодинамики — статистическая механика (изучение тех же свойств на базе статистического анализа большого числа частиц: молекул, атомов, электронов).

И соотношение между феноменологическим и микроскопическим уровнями исследования в нейрокибернетике может быть аналогично соотношению между термодинамикой и статистической механикой. Естественно, что микроскопический и феноменологический подходы в нейрокибернетике могли бы быть взаимодополнительными так же, как дополняют друг друга термодинамика и статистическая механика.

В целом соотношение между направлениями «Адаптивное поведение» и «Искусственная жизнь», используемыми в них компьютерными методами, их научным значением и потенциальными применениями можно представить в виде схемы, показанной на рис. 4. Подчеркнем, что

это активно развивающиеся направления исследований. По ним регулярно проводятся международные и европейские конференции “Artificial Life” (international), “European Conference on Artificial Life”, “Simulation of Adaptive Behavior (From Animal to Animat)”. Издаются журналы “Artificial Life” и “Adaptive Behavior”.

### **Функциональная система по П. К. Анохину — общая схема адаптивного поведения**

Вернемся к рассмотрению путей построения теории происхождения логики. По мнению автора, есть достаточно универсальная схема адаптивного поведения организмов, которая может быть положена в основу исследований эволюции «интеллектуальных изобретений». Эта схема — *функциональная система*, разработанная советским нейрофизиологом П. К. Анохиным в 1930–1970-х годах [53].

Функциональная система по П. К. Анохину — кибернетическая схема управления организмом, нацеленная на достижение полезных для организма результатов. Опишем наиболее общие особенности функциональной системы в виде схемы управления поведением животного, характеризуя основные ее свойства, и переводя — там, где это естественно — биологические представления на кибернетический язык. Функциональная система характеризует следующие свойства схемы управления поведением животного (рис. 5):

- целенаправленность, связанную с необходимостью удовлетворения потребностей животного;
- мотивацию, задающую предпосылки (например, обусловленные потребностями) для формирования цели;
- доминанту по А. А. Ухтомскому [54], обеспечивающую мобилизацию ресурсов животного на достижение приоритетной цели, в том числе мобилизацию интеллектуальных ресурсов (концентрацию внимания);
- распознавание ситуации;
- «планирование» действий;

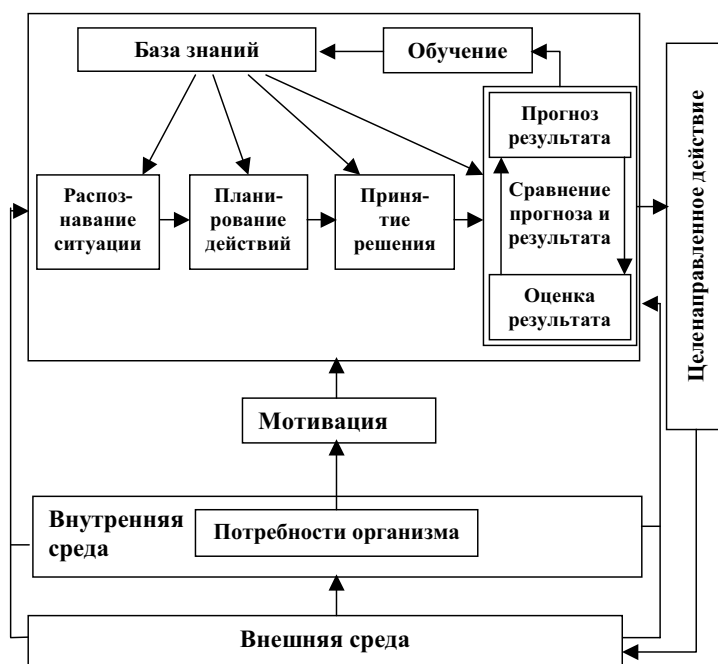


Рис. 5. Кибернетическая схема функциональной системы (в духе теории П. К. Анохина).

- принятие решения;
- прогноз результата действия;
- выполнение самого целенаправленного действия;
- оценку результата действия;
- сопоставление прогноза и результата;
- поиск нужного решения и корректировку базы знаний (в случае рассогласования прогноза и результата) — обучение.



Распознавание, планирование, принятие решения основываются на использовании базы знаний, которая пополняется при обучении.

Важное понятие функциональной системы — *мотивация*. Роль мотивации — формирование цели и поддержка целенаправленных форм поведения. Мотивация может рассматриваться как активная движущая сила, которая стимулирует нахождение такого решения, которое адекватно потребностям животного в рассматриваемой ситуации. Мотивация тесно связана с понятием *доминанты*, которое ввел А. А. Ухтомский. Доминанта мобилизует ресурсы животного на достижение заданной цели. В частности, мобилизуются нервные ресурсы, так что внимание животного концентрируется на приоритетной цели. Отметим, что модели доминанты разрабатывал В. И. Крюков [55]<sup>2</sup>.

Перечисленные свойства характеризуют общую схему системы управления поведением животного. И в целом эта схема почти тривиальна. Однако следует подчеркнуть, что есть большой потенциал для ее глубокого развития. Наметим вопросы, задающие направления развития:

1. Какова могла бы быть структура памяти в базе знаний? Естественно предположить существование постоянной (генетической, передаваемой по наследству) памяти (памяти инстинктов), долговременной памяти, в которую записывались бы надежно выработанные приобретенные навыки, и кратковременной памяти, в которую записывались бы промежуточные результаты, формируемые в процессе выработки приобретаемых навыков. Каковы приоритеты использования того или иного вида памяти? В каком виде могут записываться данные в память? Как производится сжатие информации при записи в память?
2. Какова функциональная роль памяти в базе знаний, как в нее включаются категории, понятия и отдельные образы? Можно ли ее рассматривать как семантическую сеть, включающую понятия, объединенные в сеть смысловыми связями (аналогичную семантическим сетям в разработках искусственного интеллекта)? Можно ли ввести «синтаксис», характеризующий структуру базы знаний?
3. Каковы программы поведения, основанные на базе знаний? Как они формируются в процессе обучения? Какова степень параллелизма

<sup>2</sup>См. также лекцию игумена Феофана (Крюкова) в настоящем сборнике. — Прим. ред.

обработки информационных данных в процессе функционирования программ?

4. Как видоизменяется целенаправленное поведение в процессе взросления животного? Какова роль любопытства, игр в процессе пополнения базы знаний? Какова роль родителей животного, передающих потомкам накопленный опыт поколений?
5. И самый нетривиальный, самый интересный вопрос. Какова «логика умозаключений», используемая животными при планировании, прогнозе, построении их собственных «моделей» ситуаций, коррекции и пополнении базы знаний? Каковы особенности этой «логики умозаключений» (степень нечеткости, параллелизма, степень эмоциональной окраски)? До какой степени «логику умозаключений» животных можно сопоставить с человеческой логикой (повседневной и научной)? Каковы эволюционные корни «логики умозаключений» животных и человеческой логики?

Как сказано выше, функциональная система по П. К. Анохину — общая база для построения моделей адаптивного поведения, и, опираясь на нее, целесообразно строить модели теории происхождения логики.

Что же можно начать делать сейчас? Какие наиболее интересные задачи можно наметить на ближайшее время?

По-видимому, имеет смысл начать с моделирования эволюционного возникновения целенаправленного адаптивного поведения. Приведем некоторые аргументы в поддержку этого мнения:

1. Целенаправленность могла возникнуть на очень ранних стадиях эволюции, до появления каких-либо форм индивидуально приобретаемой памяти [56], поэтому, следуя пути, пройденному эволюцией, разумно начать с анализа этого свойства.
2. Свойство целенаправленности важно само по себе — это существенная особенность поведения именно живых существ.

Разработка модели эволюционного возникновения целенаправленного адаптивного поведения уже начата [50], и полученные результаты кратко представлены в следующем разделе.

**«Кузнечик» — модель эволюционного возникновения  
целенаправленного адаптивного поведения**

В данной модели явно используется понятие мотивации — одной из главных составляющих общей схемы управления функциональной системы П. К. Анохина (рис. 5).

Основные предположения модели состоят в следующем:

- Имеется популяция агентов (искусственных организмов), имеющих две естественные потребности: 1) потребность энергии и 2) потребность размножения.
- Популяция эволюционирует в простой клеточной среде, в клетках может эпизодически вырастать трава (пища агентов). Каждый агент имеет внутренний энергетический ресурс, 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.
- Каждая потребность характеризуется количественно мотивацией. Например, если энергетический ресурс агента мал, то появляется мотивация найти пищу и пополнить энергетический ресурс.
- Поведение агента управляется его нейронной сетью, которая имеет специальные входы от мотиваций. Если имеется определенная мотивация, то поведение агента меняется с тем, чтобы удовлетворить соответствующую потребность. Такое поведение будем называть целенаправленным (есть цель удовлетворить определенную потребность).

Для простоты мы предполагаем, что среда, в которой живут агенты, представляет собой одномерный массив клеток. Агенты могут перемещаться в соседние клетки и перескакивать через несколько клеток. В силу способности агентов прыгать, мы называем их «кузнечиками». В каждой клетке может находиться только один агент.

Предполагаем, что время дискретно, и в каждый такт времени каждый агент популяции может выполнить строго одно действие. Эти действия таковы:

- быть в состоянии покоя («отдыхать»);
- двигаться, то есть перемещаться на одну клетку вправо или влево;
- прыгать через несколько клеток в случайную сторону;
- есть (питаться);
- скрещиваться.

Управление агентом (выбор действий) осуществляется его нейронной сетью. Параметры нейросети (веса синапсов) модифицируются в процессе эволюции популяции агентов. Веса синапсов нейронной сети составляют геном агента.

Описанная выше модель была реализована в виде программы на языке C++, затем было проведено моделирование с целью исследовать влияние мотиваций на поведение отдельного агента и популяции в целом.

Основные результаты проведенного моделирования таковы:

1. Целенаправленное поведение действительно возникает в ходе проанализированного нами эволюционного процесса. А именно, если сравнить популяцию агентов, имеющих мотивационные входы нейронной сети, с популяцией агентов, в которых эти входы искусственно «подавлены», то эволюционный процесс приводит к тому, что первая популяция (с мотивациями) имеет значительные эволюционные преимущества по сравнению со второй (без мотиваций). Мотивации обеспечивают эволюционное возникновение целенаправленности.
2. Переход от схемы управления агентом без мотиваций к схеме с мотивациями можно интерпретировать как формирование нового уровня иерархии управления, а именно, как метасистемный переход [20] от простых рефлексов к сложному рефлексу.

Подробнее модель эволюционного возникновения целенаправленного адаптивного поведения изложена в работах [50].

### Модель искусственной жизни в Интернете

В данном разделе излагаются предварительные результаты исследования модели искусственной жизни в Интернете [57].

Модель основана на исследованиях фирмы Webmind, Inc. [58] и на модели эволюционного происхождения целенаправленного адаптивного поведения, разработанной в Институте прикладной математики им. М. В. Келдыша РАН [50]. Работа над моделью выполнена при финансовой поддержке фирмы Webmind, Inc.

Целью модели было проанализировать процессы эволюции и самоорганизации популяции искусственных «организмов» (агентов), распределенной в Интернете.

#### Описание модели

Основные предположения модели состоят в следующем:

- Есть мир — определенное множество узлов Интернета (компьютеры, локальные сети и т. п.), в котором живет популяция агентов. Популяция агентов состоит из подпопуляций, каждая из которых находится в своем узле (рис. 6).
- Каждый агент имеет внутренний энергетический ресурс.
- Агенты могут выполнять определенные действия. Совершая действия, агенты расходуют свой ресурс.
- Уменьшение ресурса до нуля приводит к смерти агента.
- Агенты могут решать определенные задачи, в зависимости от качества решения задачи агенты получают поощрение или наказание.
- При поощрении агенту позволено съесть некоторое количество пищи, что приводит к увеличению энергетического ресурса агента. При наказании ресурс агента уменьшается.
- В каждом узле есть случайный процесс, определяющий приток пищи.

- Агенты могут переговариваться между собой, передавая друг другу свой «жизненный опыт» и накапливая знания о мире.
- В каждом из узлов агенты могут скрещиваться, рождая новых агентов. При рождении нового агента, часть энергии родителей передается потомкам.
- Агенты могут перемещаться по миру, перелетая с одного узла на другой.
- Агент имеет две нейронные сети, которые управляет его поведением. Первая (основная) нейронная сеть определяет выбор действия, которое совершает агент в текущий момент времени. Вторая нейронная сеть используется при решении заданий.
- Геном агента состоит из двух хромосом. Первая хромосома кодирует веса основной нейронной сети, вторая хромосома — начальные веса (получаемые агентом при рождении) второй нейронной сети.
- При рождении потомка его геном формируется путем кроссинговера хромосом родителей (для каждой из хромосом) и слабых мутаций генов. Имеется процедура обучения второй нейронной сети, основанная на методе обратного распространения ошибки. То есть первая нейронная сеть оптимизируется только эволюционным путем, вторая — как посредством эволюции, так и путем индивидуального обучения.
- Агенты имеют две потребности: энергия и знания. Каждая из потребностей характеризуется количественным параметром — мотивацией к удовлетворению этой потребности.

Агенты могут выполнять следующие действия:

- 1 – отдыхать;
- 2 – послать ширококвещательное сообщение на весь мир;
- 3 – установить дружеский контакт с каким-либо агентом;
- 4 – произвести детальный обмен информацией с агентом-другом;

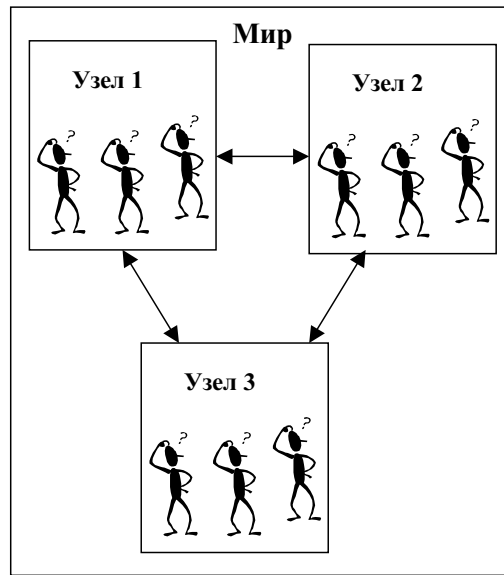


Рис. 6. Схема популяции агентов, распределенной в Интернете.

- 5 – перелететь в произвольный узел (случайно);
- 6 – перелететь в выбранный узел;
- 7 – решить задание;
- 8 – скреститься.

Расходы внутреннего энергетического ресурса  $E$  агента на выполнение данных действий ранжированы следующим образом:

$$\Delta E_1 < \Delta E_2 \approx \Delta E_3 < \Delta E_4 < \Delta E_5 \approx \Delta E_6 < \Delta E_7 \approx \Delta E_8,$$

в частности, расходы на отдых минимальны, расходы на скрещивание — максимальны.

Поясним кратко действия агентов.

Действия 2–4 — коммуникационные. С помощью действий 2 и 3 агент может находить друзей. Действие 4 — детальный «обмен опытом» между друзьями, который позволяет агентам увеличивать свои знания о ситуациях в различных узлах.

Выполняя действие 5, агент может проводить случайный поиск «хорошего места» в мире. Используя свои знания о мире, агент может выбрать наиболее предпочтительный с его точки зрения узел и перелететь в него (действие 6).

Действие 7 (решение задания) требует небольшого пояснения. При решении задания агенты получают поощрения либо наказания, которые определяются 1) текущей ситуацией  $\mathbf{S}_k$  и 2) поддействием  $a_j$ , которое совершает агент. При этом ситуации  $\mathbf{S}_k$  и поддействия  $a_j$  рассматриваются абстрактно: каждая ситуация характеризуется своим случайным вектором  $\mathbf{S}_k = (S_{k1}, S_{k2}, \dots, S_{kn})$ , а выполнение поддействия  $a_j$  означает только выбор одной из возможных альтернатив. Получаемое подкрепление  $R_{kj}$  зависит от ситуации  $\mathbf{S}_k$  и выбранного поддействия  $a_j$ .

Предполагается, что подкрепление  $R_{kj}$  зависит также и от узла, в котором находится агент, а именно, подкрепление имеет как общую часть  $Q_{kj}$ , характеризующую мир в целом, так и специфическую составляющую  $r_{kj}$ , характеризующую конкретный узел  $N$ :

$$R_{kj}(N) = Q_{kj} + r_{kj}(N),$$

где  $Q_{kj}$  и  $r_{kj}(N)$  — случайные матрицы, элементы которых выбираются в соответствии с определенной статистикой.

Если подкрепление  $R_{kj}$  положительное, то агент съедает количество пищи, равное  $R_{kj}$  (если таковое есть в узле, если же количество пищи меньше  $R_{kj}$ , то агент съедает всю наличную пищу). При этом энергетический ресурс агента увеличивается на величину съеденной пищи.

При отрицательном  $R_{kj}$  энергетический ресурс агента уменьшается на величину  $|R_{kj}|$ .

Первая нейронная сеть осуществляет выбор действия агента. Эта сеть представляет собой один слой нейронов с логистической активационной функцией. На входы нейронов подаются сигналы внешней среды, все входы связаны со всеми нейронами. Каждый нейрон соответствует какому-либо одному действию. На входы нейронов подаются сигналы из внешней (количество пищи в данном узле, количество агентов в узле и



т. п.) и внутренней (мотивации, характеризующие стремление к повышению энергетического ресурса и знаний) среды. При этом агент выбирает то действие, которое соответствует нейрону с максимальным выходным сигналом.

Вторая нейронная сеть аналогична первой, она также состоит из одного слоя нейронов. Каждый нейрон соответствует одному поддействию. На входы нейронов подаются компоненты векторов  $S_k$ , характеризующие ситуационные сигналы. Поддействие выбирается по максимальному выходному сигналу нейронов.

Схема эволюции представляет собой достаточно естественный процесс рождения и гибели агентов: когда в узле находятся два агента, готовых к скрещиванию, они производят потомка, если энергетический ресурс какого-либо агента уменьшается до нуля, то этот агент погибает.

### Результаты моделирования

Работа над этой моделью начата сравнительно недавно и результаты пока скромные. Разработана компьютерная программа, моделирующая поведение агентов, проведены первые эксперименты. Эти эксперименты показали, что хотя в модель была заложена «тонкая структура интеллекта» агентов (три типа коммуникаций; накопление знаний как за счет обучения, так и за счет «обмена опытом»; влияние мотиваций), эволюционный поиск пока не выявил эффективного проявления этой «тонкой структуры интеллекта». А именно, при достаточно большом количестве пищи наблюдалась простейшая стратегия выживания популяции агентов: 80 % времени агенты скрещивались, давая потомков, а 20 % времени агенты решали задачи простейшим способом, зарабатывая при этом энергию, которая требуется для размножения. Если приток пищи уменьшался так, что популяция агентов находилась на грани выживания, то агенты сочетали упомянутую простейшую стратегию со случайным поиском — жизнь агентов происходила как бы «на грани хаоса и порядка» [45].

Отметим возможный путь развития модели, при котором можно ожидать эффективной работы «тонкой структуры интеллекта» агентов. По-видимому, было бы целесообразно поручить агентам выполнять не описанные выше абстрактные задания, а более полезные и более богатые содержанием задания, специфические для сети Интернет: поиск нужной

информации на сайтах, прогноз экономических показателей рынка используя данные Web-страниц, поиск надежных бизнес-партнеров на базе информации, представленной в Интернете и т.п.

### **О перспективах эволюционной кибернетики**

Здесь мы охарактеризовали только некоторые аспекты исследований эволюции кибернетических систем. Более детальное рассмотрение показывает [1], что сфера исследований эволюционной кибернетики обширна: от моделей возникновения жизни до концепций будущего развития человечества. В наиболее интересных направлениях: моделирование эволюции высшей нервной деятельности, анализ проблемы происхождения интеллекта, анализ глубокого философского вопроса — почему человеческое мышление применимо к познанию природы — эти исследования только начинаются.

Прежде чем подвести окончательный итог, сделаем еще одно существенное замечание о прикладных аспектах эволюционной кибернетики. Перспективы ее развития связаны не только с научным интересом к исследованию эволюции познавательных свойств биологических организмов, но с развитием современных информационных технологий.

Например, в последние годы появилась концепция «Всемирного мозга», формирующегося на базе сети Интернет. Предполагается, что активной частью этого «мозга» будут эволюционирующие популяции программ-агентов, созданных на основе методов «Искусственной жизни». Модель, изложенная на с. 54–59, дает некоторое представление о работах в этом направлении.

Появились также прикладные работы по моделированию экономических структур методами «Искусственной жизни». В частности, появилось новое направление экономической науки — моделирование экономических систем на базе взаимодействующих агентов (Agent-Based Computational Economics). Для этого направления характерен типичный для исследований «Искусственной жизни» синтетический подход — сначала в модель вводятся характеристики отдельных экономических объектов (агентов), а также принципы взаимодействия агентов между собой и с окружающей средой, а затем в результате моделирования определяются закономерности эволюции формирующейся экономической системы. На-

пример, есть модели популяции брокеров, играющих на бирже, брокеры обладают разными приспособленностями. Брокеры с высокими приспособленностями вытесняют остальных из популяции. Хорошо играющие брокеры могут продавать свои знания плохо играющим — в моделях это соответствует продаже «мозгов» (нейронной сети) одного брокера другому.

Модели «Адаптивного поведения» активно используются в робототехнике. Хотя, как правило, поведение аниматов в этих разработках довольно простое — на уровне системы безусловных рефлексов, — наиболее «интеллектуальные» аниматы имеют свои собственные цели и мотивации — схемы их управления приближаются (хотя и очень медленно) к очерченной выше системе целенаправленного адаптивного управления в духе П. К. Анохина (рис. 5).

Разумеется, есть и перспектива использования принципов целенаправленного адаптивного управления в многочисленных промышленных, экономических и социальных системах управления. Можно ожидать, что такого рода прикладные исследования будут служить дополнительным стимулом развития эволюционной кибернетики.

Недавно, весной 2000 года, авторы и редакторы (В. Ф. Турчин, Ф. Хейлигхен, К. Джослин) международного Интернет-проекта Principia Cybernetica Project провозгласили лозунг «Эволюционная кибернетика» как общую концепцию широкого спектра кибернетических исследований возникновения и развития целенаправленных структур в Природе, Мышлении, Обществе и Технологиях<sup>3</sup>. По мнению этих авторов, эволюционная кибернетика должна служить основой для исследований кибернетической эволюции всего человечества, включая разработку будущих иерархических систем управления человеческим сообществом.

<sup>3</sup>Сайт проекта Principia Cybernetica:  
URL: <http://pespmc1.vub.ac.be/EVOLCYB.html>

Подводя общий итог, можно сказать, что исследования эволюционной кибернетики:

- ценны с научной точки зрения, так как весьма интересно проанализировать, как развивались в процессе эволюции кибернетические свойства биологических организмов;
- важны с точки зрения развития теории познания, так как могли бы прояснить причины возникновения познавательных способностей человека и, в определенной степени, обосновать применимость нашего мышления в научном познании;
- могут служить естественно-научной основой многочисленных прикладных работ, от простых эволюционных методов оптимизации в инженерных задачах до разработки иерархических систем управления человеческим сообществом.

### Благодарность

Работа выполнена при финансовой поддержке Российского гуманитарного научного фонда. Код проекта 00-03-00093.

### Литература

1. *Редько В. Г.* Эволюционная кибернетика. – М.: Наука, 2001 (в печати); см. также: *Редько В. Г.* Лекции по эволюционной кибернетике. – 1999. URL: <http://www.keldysh.ru/BioCyber/Lectures.html>
2. *Клайн М.* Математика. Поиск истины. – М.: Мир, 1988. – 295 с.
3. *Пуанкаре А.* О науке. – М.: Наука, 1990. – 736 с.
4. *Вигнер Ю.* Непостижимая эффективность математики в естественных науках // В кн.: *Вигнер Ю.* Этюды о симметрии. – М.: Мир, 1971. – с. 182–198.
5. *Жакоб Ф., Моно Ж.* Регуляция активности генов // В сб.: *Регуляторные системы клетки.* – М.: Мир, 1964. – с. 278–304.
6. *Редько В. Г.* Адаптивный сайзер // *Биофизика.* – 1990. – т. 35, № 6. – с. 1007–1011.

7. *Kinastowski W.* Der Einfluss der mechanischen Reise auf die Kontraktilität von *Spirostomum ambiguum* Ehrbg. // *Acta Protozool.* – 1963. – v. 1, No. 23. – pp. 201–222.
8. *Туималова Н. А.* Функциональные механизмы приобретенного поведения у низших беспозвоночных. – М.: Изд-во МГУ, 1986. – 109 с.
9. *Соколов Е. Н.* Исследование механизма памяти на уровне отдельного нейрона // *Журн. высш. нерв. деят.* – 1967. – т. 17, № 5. – с. 909–924.
10. *Цетлин М. Л.* Исследования по теории автоматов и моделирование биологических систем. – М.: Наука, 1969. – 316 с.
11. *Meyer J.-A., Guillot A.* From SAB90 to SAB94: Four years of Animat research // In.: *Proc. of the Third International Conference on Simulation of Adaptive Behavior.* – Cambridge: The MIT Press, 1994.  
URL: <http://www-poleia.lip6.fr/ANIMATLAB/>
12. *Павлов И. П.* Двадцатилетний опыт объективного изучения высшей нервной деятельности (поведения) животных. Условные рефлексы. – М.–Л.: Госиздат, 1928. – 388 с.
13. *Котляр Б. И., Шульговский В. В.* Физиология центральной нервной системы. – М.: Изд-во МГУ, 1979. – 342 с.
14. *Воронин Л. Г.* Эволюция высшей нервной деятельности. – М.: Наука, 1977. – 128 с.
15. *Ляпунов А. А.* О некоторых общих вопросах кибернетики // В сб.: *Проблемы кибернетики.* – М.: Физматгиз, 1958, № 1. – с. 5–22.
16. *Grossberg S.* Classical and instrumental learning by neural networks // In.: *Progress in Theoretical Biology.* – 1974. Vol. 3. – pp. 51–141.
17. *Barto A. G., Sutton R. S.* Simulation of anticipatory responses in classical conditioning by neuron-like adaptive element. // In.: *Behav. Brain Res.* – 1982. – Vol. 4.
18. *Гаазе-Рапопорт М. Г., Поспелов Д. А.* От амёбы до робота: модели поведения. – М.: Наука, 1987. – 288 с.
19. *Balkenius C., Moren J.* Computational models of classical conditioning: A comparative study // In.: *Proc. of Artificial Life V.* – MIT Press, Bradford Books, MA., 1998.
20. *Турчин В. Ф.* Феномен науки. Кибернетический подход к эволюции. – М.: Наука, 1993. – 295 с. (1-е изд); М.: ЭТС, 2000. – 368 с. (2-е изд).
21. *Jantsch E.* The self-organising universe. – Oxford a.o.: Pergamon Press, 1980. – 340 pp.

22. Генцен Г. Исследования логических выводов // В сб.: *Математическая теория логического вывода*. – М.: Наука, 1967. – с. 9–76.
23. Клини С. Математическая логика. – М.: Мир, 1973. – 480 с.
24. Клини С. Введение в метаматематику. – М.: ИЛ, 1957. – 526 с.
25. Финн В. К. О машинно-ориентированной формализации правдоподобных рассуждений в стиле Ф. Бэкона – Д. С. Милля // В сб.: *Семиотика и информатика*. – М.: ВИНТИ, 1983. – Вып. 20. – с. 35–101.
26. Angluin D., Smith C. H. Inductive inference: theory and methods // *Comp. Surveys*. – 1983. – Vol. 15, No. 3. – pp. 237–269.
27. Zadeh L. A., Klir G. J., Yuan B. (Eds.) Fuzzy sets, fuzzy logic, and fuzzy systems: Selected papers by Lotfi A. Zadeh (Advances in Fuzzy Systems – Applications and Theory, Vol. 6). – Singapore: World Scientific, 1996.
28. 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. – 655 p.
29. Langton C. G., Taylor C., Farmer J. D., Rasmussen S. (Eds.) Artificial Life II: Proceedings of the Second Artificial Life Workshop. – Redwood City CA: Addison-Wesley, – 1992. – 854 p.
30. Meyer J.-A., Wilson S. W. (Eds.) From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior. – Cambridge, Massachusetts, London, England: The MIT Press, 1990.
31. Вариавский В. И., Поспелов Д. А. Оркестр играет без дирижера. – М.: Наука, 1984. – 208 с.
32. Моделирование обучения и поведения. – М.: Наука, 1975.
33. Langton C. G. Life at the edge of chaos // In [29]. – pp. 41–91.
34. Yaeger L. Computational genetics, physiology, metabolism, neural systems, learning, vision, and behavior or Polyworld: Life in a new context // *Artificial Life III*. – Redwood City CA: Addison-Wesley, 1994. – pp. 263–298.  
URL: <http://www.beanblossom.in.us/larryy/PolyWorld.html>
35. Ray T. S. An approach to the synthesis of life // In [29]. – pp. 371–408.  
URL: <http://www.hip.atr.co.jp/~ray/tierra/tierra.html>
36. Adami C., Brown C. T. Evolutionary learning in the 2D Artificial Life system “Avida” // In: *Artificial Life IV*. – Cambridge, MA: MIT Press, 1996.  
URL: <http://www.krl.caltech.edu/avida/>

37. *Ackley D., Littman M.* Interactions between learning and evolution // In: [29]. – pp.487–509.
38. *Holland J. H.* Echoing emergence: Objectives, rough definitions, and speculations for Echo-class models // In.: *Complexity: Metaphors, Models and Reality, volume XIX of Santa Fe Institute Studies in the Sciences of Complexity.* – Reading, MA: Addison-Wesley, 1994. – pp. 309–342.  
URL: <http://www.santafe.edu/projects/echo/>
39. *Hillis W. D.* Co-evolving parasites improve simulated evolution as an optimization procedure // In: [29]. – pp. 313–322.
40. *Mitchell M., Crutchfield J. P., Das R.* Evolving cellular automata with genetic algorithms: A review of recent work // In.: *Proc. of the First International Conference on Evolutionary Computation and Its Applications (EvCA'96).* – Moscow, Russia: Russian Academy of Sciences, 1996.  
URL: <http://www.santafe.edu/projects/evca/evabstracts.html>
41. *Collings R. J., Jefferson D. R.* AntFarm: Towards simulated evolution // In: [29]. – pp. 579–601.
42. *Holland J. H., Holyoak K. J., Nisbett R. E., Thagard P.* Induction: Processes of Inference, Learning, and Discovery. – Cambridge, MA: MIT Press, 1986.
43. *Adami C., Seki R., Yirdaw R.* Critical exponent of species-size distribution in evolution // In.: *Artificial Life VI.* – MIT Press, 1998. – pp. 221–227.
44. *Bak P.* How Nature works: The science of self-organized criticality. – Berlin: Springer, 1996.
45. *Kauffman S. A.* Origins of order: self-organization and selection in evolution. – New York: Oxford University Press, 1993.
46. *Holland J. H.* Adaptation in Natural and Artificial Systems. – Ann Arbor, MI: The University of Michigan Press, 1975 (2nd ed. – Boston, MA: MIT Press, 1992).
47. *Baldwin J. M.* A new factor in evolution // *American Naturalist.* – 1896. – v. 30. – pp. 441–451:  
URL: <http://www.santafe.edu/sfi/publications/Bookinforev/baldwin.html>
48. *Belew R. K., Mitchell M.* (Eds.) Adaptive individuals in evolving populations: Models and algorithms. – Massachusetts: Addison-Wesley, 1996:  
URL: <http://www.santafe.edu/sfi/publications/Bookinforev/ipep.html>
49. *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.* – 1996. – v. 4, No. 3:  
URL: <http://ai.iit.nrc.ca/baldwin/toc.html>

50. *Бурцев М. С., Гусарев Р. В., Редько В. Г.* Модель эволюционного возникновения целенаправленного адаптивного поведения. 1. Случай двух потребностей // Препринт ИПМ РАН. – 2000, № 43.  
*Бурцев М. С., Гусарев Р. В., Редько В. Г.* Нейросетевая модель эволюционного возникновения целенаправленного адаптивного поведения // III Всероссийская конференция «Нейроинформатика-2001». – Сб. науч. трудов. – Часть 1. – М.: МИФИ, 2001. – с. 153–160.
51. *Donnart J. Y., Meyer J. A.* Learning reactive and planning rules in a motivationally autonomous animat // *IEEE Transactions on Systems, Man, and Cybernetics*. – 1996. – Part B: Cybernetics. – v. 26, No. 3. – pp. 381–395:  
URL: <http://www-poleia.lip6.fr/ANIMATLAB/>
52. *Sutton R., Barto A.* Reinforcement learning: An introduction. – Cambridge: MIT Press, 1998. – 432 pp.:  
URL: <http://www-anw.cs.umass.edu/rich/book/the-book.html>
53. *Анохин П. К.* Системные механизмы высшей нервной деятельности. – М.: Наука, 1979. – 453 с.
54. *Ухтомский А. А.* Доминанта. – М.-Л.: Наука, 1966. – 273 с.
55. *Kryukov V. I.* An attention model based on the principle of dominanta // In.: *Proc. in Nonlinear Science. Neurocomputers and Attention I: Neurobiology, Synchronization and Chaos*. – 1989, Ed. by *A. Y. Holden* and *V. I. Kryukov*. – pp. 319–351.
56. *Tsitlovsky L. E.* A model of motivation with chaotic neuronal dynamics // *J. of Biological Systems*. – 1997. – v. 5, № 2. – pp. 301–323.
57. *Goertzel B., Macklakov Yu. V., Red'ko V. G.* Model of evolution of Web agents // Report at First International Conference on Global Brain (Brussels, July, 2001):  
URL: <http://www.keldysh.ru/BioCyber/webagents/webagents.htm>
58. *Goertzel B., Bugaj S. V.* WebWorld. A conceptual and software framework for Internet Alife // Report at VII International Conference on Artificial Life, 2000.

**Владимир Георгиевич Редько**, доктор физико-математических наук, старший научный сотрудник Института прикладной математики им. М. В. Келдыша РАН. Область научных интересов — эволюционная кибернетика.