

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 нейронинформатики»

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

Предисловие	5
<i>В. Г. Редько. Эволюционная кибернетика</i>	29
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 эволю- ции — путь к логике	34
Направления исследований «Искусственная жизнь» и «Адап- тивное поведение»	40
Искусственная жизнь — жизнь, какой она могла бы в прин- ципе 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 адап- тивного поведения	48
«Кузнечик» — модель эволюционного возникновения целенаправ- ленного адаптивного поведения	52
Модель искусственной жизни в Интернете	54
Описание модели	54
Результаты моделирования	58
О перспективах эволюционной кибернетики	59
Благодарность	61
Литература	61
УДК 004.032.26 (06) Нейронные сети	3

Игумен Феофан (Крюков). Модель внимания и памяти, основанная на принципе доминанты	66
Введение. Основные проблемы моделирования внимания	67
Принцип доминанты и основные проблемы внимания	70
Фазовые переходы в мозге и принцип доминанты	78
Доминантная архитектура обработки информации в мозге	88
«Нейролокатор» — Центральный Управитель	96
Заключительные замечания	104
Литература	108
Ю. И. Нечаев. Нейросетевые технологии в бортовых интеллектуальных системах реального времени	114
Введение	115
Особенности динамики сложных нелинейных систем	116
Концепция использования ИНС в интеллектуальных системах реального времени	121
Управление и принятие решений	124
Концепция «мягких вычислений» в бортовых интеллектуальных системах управления	125
Когнитивный подход при проектировании нейросетевых контроллеров	135
Многослойная нейронная сеть на нечетких нейронах	139
Идентификация экстремальной ситуации	145
Оценка динамических характеристик объекта и внешней среды	149
Интеллектуальные датчики на основе ИНС	154
Заключение	159
Литература	160

ПРЕДИСЛОВИЕ

В этой книге (она выходит в двух частях) содержатся тексты лекций, прочитанных на Школе-семинаре «Современные проблемы нейроинформатики», проходившей 23–25 января 2002 года в МИФИ в рамках IV Всероссийской научно-технической конференции «Нейроинформатика–2002».

Как и для первой Школы [1], основной целью было дать представление слушателям о современном состоянии и перспективах развития важнейших направлений, связанных с теорией и практикой нейроинформатики, ее применениями, а также с некоторыми смежными вопросами. При подготовке программы Школы особенно приветствовались лекции, лежащие по охватываемой тематике «на стыке наук», рассказывающие о проблемах не только собственно нейроинформатики (т. е. проблемах, связанных с нейронными сетями, как естественными, так и искусственными), но и о взаимодействиях нейроинформатики с другими областями мягких вычислений (нечеткие системы, генетические и другие эволюционные алгоритмы и т. п.), с системами, основанными на знаниях, с традиционными разделами математики, инженерной теории и практики. При этом изложение материала должно было строиться с таким расчетом, чтобы содержание лекции не только было бы интересным для членов нейросетевого сообщества, но и доступно более широкой аудитории, особенно студентам-старшекурсникам и аспирантам (в определенной степени моделью такого рода изложения могут служить брошюры знаменитой серии «Математика, кибернетика», выпускавшейся в течение 30 лет издательством «Знание»).

Предлагаемая подборка текстов лекций — это не учебник, охватывающий всю нейроинформатику или хотя бы значительную ее часть. Целью лекторов, приглашенных из числа ведущих специалистов в области нейроинформатики и ее приложений, было дать живую картину работы «на переднем крае» нейроинформатики, рассказать о ее взаимодействии с другими научно-техническими областями, причем сделать это, по-возможности, на примерах проблем, наиболее актуальных и активно изучаемых на данный момент.

Как и положено работам «с переднего края», каждая из них содержит, хотя и в разной степени, элементы дискуссионности. Не со всеми положениями, выдвигаемыми авторами, можно безоговорочно согласиться,

но это только повышает ценность предлагаемых материалов — они стимулируют возникновение дискуссии, поиск альтернативных ответов на поставленные вопросы, альтернативных решений сформулированных задач.

В программу Школы-семинара «Современные проблемы нейроинформатики» на конференции «Нейроинформатика–2002» вошли лекции В. Г. Редько, игумена Феофана (Крюкова), Ю. И. Нечаева, С. А. Шумского, С. А. Терехова и Н. Г. Макаренко¹.

Открывался данный цикл лекцией **В. Г. Редько** «Эволюционная кибернетика». И это было не случайно.

Наука, техника, многие другие области человеческой деятельности немалымы без создания и исследования моделей, в том числе и такого важнейшего их класса, как модели символьные, базирующиеся на одной из знаковых систем — это и всевозможные математические и другие формальные модели, и различного рода компьютерные программы, и тексты на естественных языках, и разнообразные комбинации этих элементов.

Уже сама возможность применения символьных (в частности, математических) моделей в естественных науках, в технике, представляет собой факт достаточно нетривиальный. Вопрос можно поставить и шире, как это делается в лекции В. Г. Редько: «Почему *человеческая* логика применима к познанию *природы*?»

Эти проблемы — взаимоотношений математики и естествознания, причин применимости человеческой логики к познанию природы, и вообще — «непостижимой эффективности математики в естественных науках» (по известному выражению Юджина Вигнера) обсуждали и продолжают обсуждать многие видные ученые. Наряду с работами Ю. Вигнера, М. Клайна и А. Пуанкаре, упоминаемыми в лекции В. Г. Редько, по этим вопросам можно также рекомендовать обратиться к книгам [2–9].

Создание теоретических моделей для достаточно сложных объектов и процессов — в высшей степени непростая задача. Традиционный путь решения такой задачи состоит в получении требуемой модели сразу на заданном уровне сложности. То обстоятельство, что вначале, чаще всего, решается серия так называемых «модельных задач», сути дела не меняет, поскольку эти модельные задачи представляют собой просто усечен-

¹Первые три из перечисленных лекций публикуются в части 1, а оставшиеся три — в части 2 сборника «Лекции по нейроинформатике.»

ные различным образом варианты основной задачи, но концептуально ее «дух» всегда остается неизменным. Базой для подобного рода процесса решения служит изучение строения требуемого объекта (процесса) и его составных частей, взаимодействия этих частей между собой, а также объекта в целом с окружающей средой (см., например, [10–12]).

Можно не углубляться в изучение внутреннего строения объекта, его «природы», а рассматривать его как «черный ящик», про который известно лишь, как он реагирует на некие представляющие интерес воздействия, возмущающие и/или управляющие. И мы получаем таким способом еще одну разновидность упомянутого выше подхода, поскольку суть дела опять же не изменилась – по-прежнему мы пытаемся получить модель объекта сразу на требуемом уровне сложности.

В значительной степени наука, а вместе с ней и инженерная теория, в течение всей своей истории развивались именно так в попытках познания мира и создания искусственных объектов.

Но есть и в последнее время довольно активно начинает развиваться другой подход, в своих концептуальных установках диаметрально противоположный первому. Он состоит в том, чтобы в качестве исходных взять некоторые очень простые модели и добавить к ним механизмы развития, позаимствованные у Природы. Тогда задача получения модели сложной системы (а в ряде случаев и самой этой системы!) сводится к «выведению», «выращиванию» такой модели эволюционным путем из модели более простой системы (или совокупности моделей простых систем).

Основное содержание лекции В. Г. Редько как раз и посвящено изложению ряда основных концепций этого (второго) направления и основной вопрос, которым задается здесь автор – «... нельзя ли промоделировать эволюцию познавательных способностей животных и подойти к моделированию эволюционного возникновения интеллекта?»

Эволюционное направление как в создании моделей систем, так и самих систем представляется весьма перспективным и многообещающим. Оно открывает возможность заменить процесс создания модели сразу как целого процессом подготовки некоторой «затравки», на которую «напускаются» механизмы эволюционного развития. Такой путь может оказаться перспективным с точки зрения преодоления пресловутого «порога сложности», возникающего при создании систем.

Нельзя сказать, что данному направлению раньше совсем не уделялось внимания. Напротив, предьстория его довольно богата.

Известно, что первые вычислительные машины появились в связи с потребностями выполнения больших объемов вычислений, например, в баллистике, авиационной и ракетной технике, атомной технике и др.

Но уже с самого начала, примерно с середины 50-х годов, ЭВМ пытались использовать не только для проведения расчетов, но и для моделирования интеллектуальных систем. Уже тогда сформировались основные направления работ в этой области, существующие и в настоящее время.

Сразу же сформировалось два конкурирующих направления исследований, получивших наименование нисходящего и восходящего подходов.

Сторонники *нисходящего подхода* пытались воспроизводить (моделировать) достаточно сложные интеллектуальные операции и виды деятельности (игры — шашки, шахматы; доказательство теорем; поиск решений и т. п.). Работы в этом направлении привели, в частности, к появлению экспертных систем и, шире, систем, основанных на знаниях (см., например, [14–16]).

Исследователи, работавшие в рамках *восходящего подхода*, пытались идти от простых аналогов нервной системы примитивных существ с очень малым числом нейронов к сложнейшей нервной системе человека. Это направление привело, в частности, к появлению обширного класса моделей, именуемых искусственными нейронными сетями (см., например, [15, 17–20]; см. также «тему номера» в журнале «Компьютерра» [21]).

Но тогда же, практически одновременно с упомянутыми двумя, возник еще и третий подход к созданию интеллектуальных систем, называемый *эволюционным программированием*. Целью его было, как отмечал А. Г. Ивахненко в предисловии к русскому переводу книги [22] (оригинал ее был издан в 1966 году), «заменить процесс моделирования человека моделированием процесса его эволюции».

Ранняя история данного направления связана с работами Л. Фогеля и его сотрудников [22] по сообществам эволюционирующих конечных автоматов (в определенной степени развитием работ данного направления стали книги [23–25]), работами 60-х годов М. Л. Цетлина по моделям автоматов, адаптивно приспособляющихся к окружающей среде, а также работы 60–70-х годов М. М. Бонгарда по адаптивному поведе-

нию искусственных организмов на плоскости, разбитой на клетки². Наряду с этими работами следует также упомянуть активное обсуждение проблемы «Автоматы и жизнь», проходившее в 60-е годы с участием таких видных отечественных и зарубежных ученых, как Н. М. Амосов, И. И. Артоболевский, Н. Винер, В. М. Глушков, А. А. Дородницын, А. Г. Ивахненко, А. Е. Кобринский, А. Н. Колмогоров, У. Р. Эшби и др. Спектр мнений по данной проблеме был самый широкий — от безудержного оптимизма («Только автомат? Нет, мыслящее существо!») до полнейшего пессимизма («Машина не может жить, плесень не способна мыслить!»)³. Некоторые материалы дискуссии «Автоматы и жизнь» (статьи и доклады разных лет) содержатся в сборнике [30].

В тот же период времени начались исследования по такой сложнейшей проблеме, как *самовоспроизводящиеся искусственные системы*; одними из первых здесь были работы Дж. фон Неймана по самовоспроизводящимся автоматам [31].

Идейно близки к перечисленным работам и быстро развивающиеся сейчас направления — генетические алгоритмы, генетическое программирование, эволюционные вычисления [26–29].

Идеи и методы эволюционного моделирования активно использовались в возникшем в конце 80-х годов интереснейшем направлении, именуемом «Искусственная жизнь» (Artificial Life, или просто ALife), основные элементы которого также рассматриваются в лекции В. Г. Редько.

Обсуждение ряда элементов ALife есть в тематическом разделе («тема номера») журнала «Компьютерра» [32]. В одной из статей этого номера рассказывается об эволюционном процессе, реализованном аппаратно — на уровне электронных микросхем. Здесь же содержится целый ряд ссылок по теме ALife на ресурсы Интернет.

В лекции В. Г. Редько приводится целый ряд примеров модельной реализации идей ALife на программном или аппаратном уровне. Список этот, разумеется, не может претендовать на исчерпывающую полноту.

Хотелось бы обратить внимание читателей на один достаточно показательный пример, не вошедший в этот список.

²Ссылки на работы М. Л. Цетлина и М. М. Бонгарда можно найти в лекции В. Г. Редько и списке литературы к ней.

³Заголовки разделов в сборнике [30].

Речь идет о работах Марка Тилдена (Mark W. Tilden) из Лос-Аламосской национальной лаборатории США (Los Alamos National Laboratory) по направлению, которое он называет «Живые машины». Русский перевод (в сокращении) одной из статей М. Тилдена (совместно с Б. Хасслахером) был опубликован в журнале «Природа» [33].

М. Тилден с сотрудниками построили около сотни действующих образцов «биоморфных машин» («биоморфов», или «жизнеподобных»), главная задача которых — преодолевать незнакомые сложные ландшафты в поисках «пищи». Управляющее ядро этих машин представляет собой аналоговую нейросеть осцилляторного типа с очень небольшим числом нейронов в ней (как правило, менее десятка). Эти машины продемонстрировали очень высокую приспособляемость к меняющемуся рельефу местности.

Кроме статьи [33], информацию о работах М. Тилдена можно найти по адресам Интернет, перечисленным под номером [34] в списке литературы в конце предисловия. Среди этих ресурсов можно найти патент М. Тилдена на нейросеть, используемую им в биоморфных машинах.

Пересказывать содержание этой многоплановой и интересной лекции здесь нет никакой необходимости, укажем лишь ряд дополнительных источников, с помощью которых можно более глубоко проработать затронутые в лекции вопросы.

Различные аспекты зарождения и развития жизни на Земле, общие законы функционирования живого освещаются в книгах [35–43]. Принципы биологической эволюции, ее механизмы и модели рассматриваются в книгах [44–69]. Об эволюционном возникновении интеллекта можно прочитать в книгах [70, 71], об организации психики человека, происхождении, формировании и развитии высших потребностей познания — в книгах [72, 73]. Попытка мысленно представить эволюционное возникновение иерархии биологических систем управления сделана в прекрасной книге В. Ф. Турчина [13].

Общая схема адаптивного поведения, рассматриваемая В. Г. Редько, основывается на *функциональной системе*, разработанной советским нейрофизиологом П. К. Анохиным [74]. Функциональная система характеризует такие свойства схемы управления поведением, как целенаправленность, мотивацию для формирования цели, доминанту по А. А. Ухтомскому для мобилизации ресурсов животного на достижение приоритетной

цели (в том числе и мобилизацию интеллектуальных ресурсов — концентрацию внимания), а также ряд других.

Как показано в лекции **игумена Феофана (Крюкова)** «Модель внимания и памяти, основанная на принципе доминанты», важнейшая роль в этом перечне свойств принадлежит доминанте.

В лекции описаны *шесть основных проблем внимания*: проблема селективности стимулов (почему из нескольких одновременно предъявленных стимулов одни привлекают внимание и получают таким образом доступ к высшей сенсорной обработке, а другие не получают?); проблема долговременной памяти (каков механизм взаимодействия внимания и долговременной памяти?); проблема интеграции (как и где происходит реконструкция интегрального образа для стимулов, обрабатывавшихся параллельно?); проблема инерции (какова основа сохранения длительного внимания в случаях, когда стимулы предъявляются кратковременно?); проблема торможения и подавления помех (что происходит со стимулами, которым не оказывается внимания?); проблема Центрального Управителя (существует ли отдельная структура для координации процессов внимания и памяти или же здесь работают процессы самоорганизации новой коры?).

В лекции показано, что на основе принципа *доминанты А. А. Ухтомского* удастся найти ответы на все шесть перечисленных выше вопросов. Показано, что в основе учения о доминанте лежит физическое явление фазовых переходов, а также трактовка нейронной сети как системы связанных нелинейных осцилляторов. Приводятся доказательства того, что неравновесные фазовые переходы действительно происходят в мозге.

Нейрофизиологический материал, необходимый для понимания материала лекции игумена Феофана (Крюкова), можно почерпнуть, например, в общем курсе биологии [39], а также в книгах [75, 76]. Об исследованиях мозга говорится в книгах [77, 78]. О связях высшей нервной деятельности с психологией рассказывается в книге [79], здесь рассматривается и роль доминанты А. А. Ухтомского для понимания процессов высшей нервной деятельности.

На важность и перспективность использования в обработке информации *колебательных моделей*, включая и колебательные (осцилляторные) нейронные сети, автор данной лекции обращал внимание нейросетевого сообщества в ходе «Дискуссии о нейрокомпьютерах», состоявшейся

в рамках конференции «Нейроинформатика–99» (см. [80], с. 29–33, выступление В. И. Крюкова). Им утверждалось, в частности, что «... материальным носителем биологической памяти, если таковой существует, является не синаптическая система, а скорее целостная нервная ткань, как это предсказывается, исходя из принципа доминанты».

Того же мнения о значимости колебательных нейронных сетей придерживается и Р. М. Борисюк, который на той же самой дискуссии в ответе на вопросы о наиболее значительных достижениях в теории нейронных сетей и в понимании работы мозга, полученных в течение 90-х годов (см. [80], с. 13–16) отметил: «Одним из основных достижений можно считать создание теории осцилляторных нейронных сетей и демонстрацию того, что принцип синхронизации нейронной активности является важным принципом обработки информации в структурах мозга. Детальная разработка этой теории, имеющей глубокие корни в работах выдающегося физиолога А. А. Ухтомского, была начата в нашей стране В. И. Крюковым, а на Западе в работах К. фон-дер Мальсбурга (Christoph von der Malsburg). Дальнейшее развитие теории показало, что на основе принципа синхронизации можно решать задачи распознавания образов, запоминания информации, интеграции признаков объекта в цельный образ, формирования и управления фокусом внимания и др.».

Вопросам, связанным с осцилляторными нейронными сетями, постоянно уделялось внимание и на конференциях «Нейроинформатика» (см. [81–87]).

Здесь уместно будет отметить, что работы М. Гилдена по «живым машинам», упоминавшиеся выше, также основываются на использовании осцилляторных нейронных сетей.

В лекции игумена Феофана (Крюкова) в противовес традиционной коннекционистской архитектуре нейросетевых систем предлагается доминантная архитектура обработки информации в мозге. Кроме того, в ней ставится вопрос о неудовлетворительности существующей концептуальной базы (парадигмы⁴) нейроинформатики и делается вывод о необходимости смены этой парадигмы: «Почти все теоретики мозга ищут

⁴Концепция *парадигмы* в науке была сформулирована Томасом Куном в начале 60-х годов: «... Под парадигмами я подразумеваю признанные всеми научные достижения, которые в течение определенного времени дают научному сообществу модель постановки проблем и их решений (см. [88], с. 11)». Смена одной парадигмы на другую трактуется Т. Куном как *научная революция*.

не истину, а подтверждения хеббовской программы, приняв гипотезу за незабываемый факт. А истина лежит совсем в другом месте — в учении А. А. Ухтомского о доминанте».

В лекции В. Г. Редько отмечается, что удивительная эффективность функционирования живых организмов, гармоничность и согласованность работы органов («компонент») живых существ обеспечивается биологическими управляющими системами. Относительно этих систем возникает целый ряд вопросов, в том числе и такой важнейший, как пути возникновения интеллекта.

Другой аспект этой же проблемы рассматривался в лекции игумена Феофана (Крюкова), где показано, как на основе принципа доминанты А. А. Ухтомского можно адекватно моделировать такие, не менее важные, свойства живых существ, как память и внимание.

Но ведь управляющие системы встречаются не только в живых системах, но и в системах, создаваемых человеком, они являются важнейшим элементом, определяющим в значительной мере уровень возможностей той или иной системы.

Лекция **Ю. И. Нечаева** «Нейросетевые технологии в бортовых интеллектуальных системах реального времени» посвящена вопросам создания управляющих систем именно такого рода, а также систем анализа и интерпретации измерительной информации о поведении динамического объекта.

Эта лекция представляет собой один из примеров того междисциплинарного подхода, что упоминался выше как весьма желательный для Школы-семинара.

Предметом рассмотрения в лекции Ю. И. Нечаева являются *бортовые интеллектуальные системы*, обеспечивающие управление динамическим объектом, идентификацию экстремальных ситуаций, оценку параметров динамического объекта и внешней среды.

Эти задачи решаются с привлечением целого ряда новых подходов, в число которых входят: геометрическая интерпретация динамических моделей на основе теории хаотических систем и принципов самоорганизации; нейросетевые технологии; методы построения систем, основанных на знаниях; методы нечеткой (размытой) логики и нечетких систем; методы теории возможностей; эволюционное моделирование (генетические алгоритмы и т. п.); различные комбинированные технологии (нейро-

нечеткие, нейро-генетические и т. д.).

Целесообразность применения этой совокупности методов и средств, взаимодействие их между собой, последовательно демонстрируется на конкретных примерах задач для динамических объектов, таких как управление движением подводного аппарата, идентификация экстремальных ситуаций для плавучих динамических объектов, оценка динамических характеристик объекта и внешней среды, создание интеллектуальных нейросетевых датчиков.

В лекции Ю. И. Нечаева показано, что сложности, присущие традиционным подходам к созданию бортовых измерительных и управляющих систем, могут быть в значительной мере преодолены, если воспользоваться технологиями мягких вычислений (включая нейросети, нечеткие системы, генетические алгоритмы и т.п.). Рациональное использование этих технологий позволяет обеспечить измерительным и управляющим системам гибкость и способность адаптироваться к изменяющимся условиям внешней и внутренней среды динамического объекта.

Дополнительные сведения по затронутым в лекции Ю. И. Нечаева вопросам можно получить в следующих книгах: по нелинейной динамике, хаотическим системам, самоорганизации — в [90–103] (см. также журнал «Компьютерра» [89] с темой номера «Хаос»); по системам, основанным на знаниях — в [14–16]; по нечеткой логике, нечетким системам — в [104–113] (см. также журнал «Компьютерра» [114] с темой номера «Нечеткая логика»); теория возможностей — в [115–117]; по нейросетевым технологиям — в [15, 17–21]; по генетическим алгоритмам, эволюционному моделированию — в [22–28]; по смешанным технологиям мягких вычислений — в [28, 29]; по информационной обработке и управлению на основе технологий мягких вычислений — в [118–126]. Значительное число программ и публикаций по таким темам, как искусственные нейронные сети, нечеткие системы, генетические алгоритмы, а также их применения можно найти через портал научных вычислений, адрес которого содержится в позиции [127] списка литературы к предисловию.

В начале данного предисловия было сказано о двух диаметрально противоположных подходах к построению моделей систем — традиционном и эволюционном. Эти два подхода вовсе не исключают, а скорее дополняют друг друга.

Примерами, основанными на традиционном подходе являются и лекция игумена Феофана (Крюкова), и лекция Ю. И. Нечаева. А именно, в лекции игумена Феофана (Крюкова) используется подход, типичный для науки: изучение объекта, его особенностей и т. п. В лекции Ю. И. Нечаева наряду с данным вариантом широко применяется и подход типа «черный ящик», реализующийся в искусственных нейросетях (но здесь широко используются и обычные математические модели движения динамических объектов, записанные в виде систем дифференциальных уравнений).

Еще дальше идет **С. А. Шумский** в своей лекции «Байесова регуляризация обучения». В ней речь идет о системе типа «черный ящик», для которой есть только некие описывающие ее эмпирические данные.

Рассматривается задача *машинного обучения*, цель решения которой — выявление закономерностей в эмпирических данных.

Как отмечает С. А. Шумский: «В противоположность математическому моделированию, изучающему следствия из известных законов, машинное обучение стремится воссоздать причины, наблюдая порожденные ими следствия — эмпирические данные».

Отсюда следует, что рассматриваемая задача относится к классу обратных задач, которые в общем случае являются плохо определенными или некорректными. Вследствие повышенной чувствительности некоторых из решений таких задач к данным, для нахождения устойчивых решений приходится применять процедуру так называемой *регуляризации*, которая приводит к ограничению класса допустимых решений.

При этом надо, с одной стороны, не потерять чувствительность к данным, чтобы оставалась возможность объяснения всех имеющихся фактов, а с другой — не переусложнить модель так, что она станет реагировать не только на требуемую закономерность, но и на случайные события в обучающей выборке. Или, как замечает С. А. Шумский, «пройти между Сциллой переупрощения и Харибдой переусложнения».

В лекции С. А. Шумского подробно рассматривается один из наиболее эффективных способов решения этой проблемы — *байесова регуляризация*, основанная не на оценке ожидаемой ошибки, как это принято в традиционных методах математической статистики, а на выборе наиболее правдоподобной (с учетом имеющихся данных) модели.

Иллюстрируется данный подход на задачах оценки параметров, интерполяции функций и кластеризации; одна из практически интересных задач здесь — определение рационального числа элементов в скрытом слое искусственной нейросети.

По теме лекции С. А. Шумского можно рекомендовать следующую дополнительную литературу: некорректные задачи и регуляризация — [128, 129]; традиционная математическая статистика — [130–133]; байесовский подход [134] (здесь управление трактуется как процесс обучения, подробно рассматривается теорема Байеса и ее применение).

Есть задачи, они особенно часто встречаются в ряде областей численного анализа и оптимизации, для решения которых есть, казалось бы, все необходимое — теоретическая база, алгоритмы, даже компьютерные программы. Но тем не менее, решение почти каждой такой задачи представляет собой «штучную работу», в значительной степени опирающуюся на ранее полученный опыт решения аналогичных задач.

Пример решения именно такого рода задачи демонстрируется в лекции **С. А. Терехова** «Нейросетевые аппроксимации плотности распределения вероятности в задачах информационного моделирования». Здесь, как и в лекции С. А. Шумского, изучается проблема построения эмпирических моделей на основе числовых данных. При этом рассматривается обучение без учителя на примерах, в условиях неопределенности в характере модели.

Эта задача аппроксимации плотности распределения вероятности, описывающего множество многомерных экспериментальных данных.

К такой постановке сводятся многие важные прикладные задачи: задача распознавания образов, проблема заполнения пропусков в таблицах данных, вероятностный прогноз и т. п.

В лекции С. А. Терехова дается сопоставление нескольких подходов к аппроксимации плотности распределения, в числе которых параметрические методы аппроксимации и методы непараметрической статистики. Рассматриваются также *байесовы сети*, представляющие собой одно из наиболее важных достижений последнего десятилетия в области искусственного интеллекта.

В качестве еще одного подхода предлагается заменить задачу аппроксимации эквивалентной ей задачей классификации. Здесь опять возникает проблема регуляризации, о которой, хотя и в несколько ином плане

говорилось в лекции С. А. Шумского.

Дополнительную информацию по затронутым в лекции С. А. Терехова вопросам можно получить из книг [130–133] (математическая статистика), а также [15, 17–20] (искусственные нейросети и их применений). Популярное изложение материала о байесовых сетях, а также пакет расширения (Bayes Net Toolbox) для Matlab содержится по адресам, указанным в позиции [135] списка литературы к предисловию.

Наряду с лекцией Ю. И. Нечаева, лекция **Н. Г. Макаренко** «Фракталы, аттракторы, нейронные сети и все такое» представляет собой яркий образец междисциплинарного подхода. Ценность его — в демонстрации глубоких взаимосвязей между различными областями науки, в том числе и такими, что возникли и развивались вначале совершенно независимо друг от друга.

Изложение в лекции Н. Г. Макаренко начинается с изложения концепции дробной размерности и фрактала. Затем вводятся системы итеративных функций в пространстве компактов.

Изучение предельной динамики систем итеративных функций ведет к теории дискретных динамических систем. Далее показано, что процесс аппроксимации аттрактора системы итеративных функций эквивалентен работе бинарной нейронной сети.

Как замечает Н. Г. Макаренко: «Таким образом, термины “фрактал” в геометрии и “странный аттрактор” в динамике оказываются синонимами, а систему итеративных функций (СИФ) можно рассматривать как рекуррентную асимметричную нейросеть. С другой стороны, Фернандо Ниньо в 2000 году установил, что случайная итеративная нейронная сеть (гипернейрон) топологически эквивалентна динамической системе с заданным аттрактором. Круг замкнулся, образовав Единый Контекст, объединяющий *фракталы, СИФ, аттракторы и нейронные сети*. Цель лекции — показать взаимную связь этих предметов, потому что *единое лучше, чем всё вместе, но по-отдельности*».

Дополнительные сведения по фракталам можно найти в книгах [136, 137], по динамическим системам — в книгах [138–141].

* * *

Как это уже было в [1], помимо традиционного списка литературы каждая из лекций сопровождается списком интернетовских адресов, где можно найти информацию по затронутому в лекции кругу вопросов, включая и

дополнительные ссылки, позволяющие расширить, при необходимости, зону поиска.

Вызвано это тем, что ссылки в списке литературы на традиционные «письменные» источники обычно трудно «разрешимы», материалы, на которые они указывают, в современной ситуации мало доступны, особенно вне столиц. В то же время, в Интернете можно найти сейчас информацию практически по любой тематике, часто — те же статьи, которые включены в список литературы — надо только знать, где их искать. Включение в лекции ссылок на интернетовские ресурсы дает подобного рода сведения тем, кто заинтересуется соответствующей тематикой и захочет более подробно разобраться в ней. Учитывая все расширяющиеся возможности доступа к Интернету, это обеспечивает доступ к разнообразным данным практически всем желающим.

Перечень проблем нейроинформатики и смежных с ней областей, требующих привлечения внимания специалистов из нейросетевого и родственных с ним сообществ, далеко не исчерпывается, конечно, вопросами, рассмотренными в предлагаемом сборнике.

В дальнейшем предполагается расширение данного списка за счет рассмотрения насущных проблем собственно нейроинформатики, проблем «пограничного» характера, особенно относящихся к взаимодействию нейросетевой парадигмы с другими парадигмами, развиваемыми в рамках концепции мягких вычислений, проблем использования методов и средств нейроинформатики для решения различных классов прикладных задач. Не будут забыты и взаимодействия нейроинформатики с такими важнейшими ее «соседями», как нейробиология, нелинейная динамика (синергетика — в первую очередь), численный анализ (вейвлет-анализ и др.) и т.п.

Замечания, пожелания и предложения по содержанию и форме лекций, перечню рассматриваемых тем и т.п. просьба направлять электронной почтой по адресу tium@mai.ru Тюменцеву Юрию Владимировичу.

Литература

1. *Лекции по нейроинформатике*: По материалам Школы-семинара «Современные проблемы нейроинформатики» // III Всероссийская научно-техническая конференция «Нейроинформатика-2001», 23–26 января 2001 г. / Отв. ред. Ю. В. Тюменцев. – М.: Изд-во МИФИ, 2001. – 212 с.

2. *Адамар Ж.* Исследование психологии процесса изобретения в области математики: Пер. с франц. – М.: Сов. радио, 1970. – 152 с.
3. *Блехман И. И., Мышкис А. Д., Пановко Я. Г.* Механика и прикладная математика: Логика и особенности приложений математики. 2-е изд., испр. и доп. – М.: Наука, 1990. – 360 с.
4. *Вейль Г.* Математическое мышление: Сб. статей: Пер. с англ. и нем. – М.: Наука, 1989. – 400 с.
5. *Кац М., Улам С.* Математика и логика: Ретроспектива и перспективы: Пер. с англ. – М.: Мир, 1971. – 251 с. (Серия «Современная математика: Популярная серия»)
6. *Клайн М.* Математика: Утрата определенности: Пер. с англ. – М.: Мир, 1984. – 434 с.
7. *Курант Р., Роббинс Г.* Что такое математика? Элементарный очерк идей и методов: Пер. с англ., 3-е изд. – Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001. – 592 с.
8. *Поля Д.* Математическое открытие: Решение задач – основные понятия, изучение и преподавание: Пер. с англ. – М.: Наука, 1970. – 452 с.
9. *Поля Д.* Математика и правдоподобные рассуждения: Пер. с англ. 2-е изд., испр. – М.: Наука, 1975. – 464 с.
10. *Калашиников В. В.* Сложные системы и методы их анализа. – М.: Знание, 1980. – 64 с. (Новое в жизни, науке, технике. Серия «Математика, кибернетика», вып. 9, 1980)
11. *Калашиников В. В., Немчинов Б. В., Симонов В. М.* Нить Ариадны в лабиринте моделирования. – М.: Наука, 1993. – 192 с. (Серия «Кибернетика: неограниченные возможности и возможные ограничения»)
12. *Шрейдер Ю. А., Шаров А. А.* Системы и модели. – М.: Радио и связь, 1982. – 152 с. (Серия «Кибернетика»)
13. *Турчин В. Ф.* Феномен науки: Кибернетический подход к эволюции. 2-е изд. – М.: ЭТС, 2000. – 368 с.
14. *Нильсон Н.* Принципы искусственного интеллекта: Пер. с англ. – М.: Радио и связь, 1985. – 376 с.
15. Компьютер обретает разум: Пер. с англ. Под ред. *В. Л. Стефанюка.* – М.: Мир, 1990. – 240 с.
16. Будущее искусственного интеллекта / Ред.-сост. *К. Е. Левитин* и *Д. А. Поспелов.* – М.: Наука, 1991. – 302 с.

17. Горбань А. Н., Россиев Д. А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996. – 276 с.
18. *Нейрокомпьютер как основа мыслящих ЭВМ*: Сб. науч. статей / Отв. ред. А. А. Фролов и Г. И. Шульгина. – М.: Наука, 1993. – 239 с.
19. Уоссерман Ф. Нейрокомпьютерная техника: Теория и практика: Пер. с англ. – М.: Мир, 1992. – 240 с.
20. Ежов А. А., Шумский С. А. Нейрокомпьютинг и его приложения в экономике и бизнесе. – М.: МИФИ, 1998. – 222 с.
21. *Нейросети* (тема номера, 4 статьи) // Компьютерра. – № 4 (333), 8 февраля 2000 г. – с. 19–31.
URL: <http://www.computerra.ru/offline/2000/333/>
22. Фогель Л., Оуэнс А., Уолли М. Искусственный интеллект и эволюционное моделирование: Пер. с англ. – М.: Наука, 1969. – 231 с.
23. Букатова И. Л. Эволюционное моделирование и его приложения. – М.: Наука, 1979. – 231 с.
24. Букатова И. Л. Эволюционное моделирование: Идеи, основы теории, приложения. – М.: Знание, 1981. – 64 с. (Новое в жизни, науке, технике. Серия «Математика, кибернетика», вып. 10, 1981)
25. Букатова И. Л., Михасев Ю. И., Шаров А. М. Эвоинформатика: Теория и практика эволюционного моделирования. – М.: Наука, 1991. – 206 с.
26. Special Issue “*Evolutionary Computations*” / Ed.: David B. Fogel and Lawrence J. Fogel // IEEE Transactions on Neural Networks. – January 1994. – v. 5, No. 1. – pp. 1–147.
27. Special Issue “*Genetic Algorithms*” / Eds.: Anup Kumar and Yash P. Gupta // Computers and Operations Research. – January 1995. – v. 22, No. 1. – pp. 3–157.
28. Special Issue “*Artificial Intelligence, Evolutionary Programming and Operations Research*” / Eds.: James P. Ignizio and Laura I. Burke // Computers and Operations Research. – June 1996. – v. 23, No. 6. – pp. 515–622.
29. Special Issue “*Neuro-Fuzzy Techniques and Applications*” Eds.: George Page and Barry Gomm // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Apr. 8, 1996. – v. 79, No. 1. – pp. 1–140.
30. Кибернетика: Итоги развития / Ред.-сост.: В. Д. Пекелис. – М.: Наука, 1979. – 200 с. (Серия «Кибернетика: неограниченные возможности и возможные ограничения»)
31. фон Нейман Дж. Теория самовоспроизводящихся автоматов: Пер. с англ. – М.: Мир, 1971. – 382 с.

32. *Искусственная жизнь* (тема номера, 5 статей) // Компьютерра. – № 11 (289), 16 марта 1999 г. – с. 17–31.
URL: <http://www.computerra.ru/offline/1999/289/>
33. *Хасслахер Б., Тилден М.* Живые машины // Природа. – 1995. – № 4. – с. 32–46. Это сокращенный русский вариант статьи: *B. Haszlacher and M. W. Tilden.* Living machines // Robotics and Autonomous Systems. – 1995. – v. 15. – pp. 143–169.
34. Ресурсы Интернет, касающиеся работ М. Тилдена:
- информация о действующих образцах биоморфных машин:
URL: <http://biosat.lanl.gov/>
URL: <http://cism.jpl.nasa.gov/biocomputing/workshop>
URL: <http://citeseer.nj.nec.com/6446.html>
URL: http://www.beam-online.com/Robots/Galleria_other/tilden.html
URL: <http://www.geocities.com/SouthBeach/6897/beam2.html>
 - патент на нейросеть, применяемую в биоморфных машинах:
URL: <http://microcore.solarbotics.net/patent.html>
 - популярное объяснение ее устройства:
URL: http://bftgu.solarbotics.net/starting_nvnet.html
 - нейроконтроллера на ее основе:
URL: http://biosat.lanl.gov/pubs/SPIE/ABSTRACT_SPIE_19981.html
 - а также пример применения в шагающем роботе-жуке:
URL: <http://tnewton.solarbotics.net/robot2.html>
URL: <http://www.iguana-robotics.com/RobotUniverse/BiomorphicRobots.htm>
 - Здесь – большое интервью с М. Тилденом:
URL: <http://fargo.itp.tsoa.nyu.edu/~kevin/tilden/>
35. *Ичас М.* О природе живого: Механизмы и смысл: Пер. с англ. – М.: Мир, 1994. – 496 с.
36. *Медников Б. М.* Аксиомы биологии: *Biologia axiomatica.* – М.: Знание, 1982. – 136 с. (Серия «Наука и прогресс»)
37. *Рьюз М.* Философия биологии: Пер. с англ. – М.: Прогресс, 1977. – 319 с.
38. *Чернов Г. Н.* Законы теоретической биологии. – М.: Знание, 1990. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 1, 1990)
39. *Вилли К., Детье В.* Биология: Биологические процессы и законы: Пер. с англ. – М.: Мир, 1975. – 822 с.
40. *Кемп П., Армс К.* Введение в биологию: Пер. с англ. – М.: Мир, 1988. – 671 с.
41. *Сингер М., Берг П.* Гены и геномы. В двух томах. Том 1: Пер. с англ. – М.: Мир, 1998. – 373 с.

42. *Сингер М., Берг П.* Гены и геномы. В двух томах. Том 2: Пер. с англ. – М.: Мир, 1998. – 391 с.
43. *Франк-Каменецкий М. Д.* Самая главная молекула. – М.: Наука, 1983. – 160 с. (Библиотечка «Квант». Вып. 25)
44. *Антонов А. С.* Генетические основы эволюционного процесса. – М.: Знание, 1983. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 4, 1983)
45. *Кайданов Л. З.* Генетика популяций. – М.: Высшая школа, 1996. – 320 с.
46. *Кейлоу П.* Принципы эволюции: Пер. с англ. – М.: Мир, 1986. – 128 с.
47. *Арена биологической эволюции: Сборник.* – М.: Знание, 1986. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 6, 1986)
48. *Бердников В. А.* Эволюция и прогресс. – М.: Наука, 1991. – 192 с. (Серия «Человек и окружающая среда»)
49. *Борзенков В. Г.* Философские основания теории эволюции. – М.: Знание, 1987. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 1, 1987)
50. *Георгиевский А. Б., Попов Е. Б.* «Белые пятна» эволюции. – М.: Просвещение, 1987. – 96 с. (Серия «Мир знаний»)
51. *Голубев В. С.* Эволюция: От геохимических систем до ноосферы. – М.: Наука, 1992. – 110 с. (Серия «Человек и окружающая среда»)
52. *Горбань А. Н., Хлебопрос Р. Г.* Демон Дарвина: Идея оптимальности и естественный отбор. – М.: Наука, 1988. – 208 с. (Серия «Проблемы науки и технического прогресса»)
53. *Грант В.* Эволюция организмов: Пер. с англ. – М.: Мир, 1980. – 407 с.
54. *Грант В.* Эволюционный процесс: Критический обзор эволюционной теории: Пер. с англ. – М.: Мир, 1991. – 488 с.
55. *Докинз Р.* Эгоистичный ген: Пер. с англ. – М.: Мир, 1993. – 318 с.
56. *Камишилов М. М.* Эволюция биосферы. 2-е изд., доп. – М.: Наука, 1979. – 256 с. (Серия «Человек и окружающая среда»)
57. *Лима-де-Фариа А.* Эволюция без отбора: Автоэволюция формы и функции: Пер. с англ. – М.: Мир, 1991. – 455 с.
58. *Моран П.* Статистические процессы эволюционной теории: Пер. с англ. – М.: Наука, 1973. – 288 с.
59. *Назаров В. И.* Финализм в современном эволюционном учении. – М.: Наука, 1984. – 284 с.

60. *Нейфах А. А., Лозовская Е. Р.* Гены и развитие организма. – М.: Наука, 1984. – 188 с. (Серия «От молекул до организма»)
61. *Пианка Э.* Эволюционная экология: Пер. с англ. – М.: Мир, 1981. – 400 с.
62. Проблемы теории молекулярной эволюции / *В. А. Ратнер, А. А. Жарких, Н. А. Колчанов, С. Н. Родин, В. В. Соловьев, В. В. Шамин.* Отв. ред. *Р. И. Салганик.* – Новосибирск: Наука, 1985. – 263 с.
63. *Северцов А. С.* Основы теории эволюции. – М.: Изд-во МГУ, 1987. – 320 с.
64. *Скворцов А. К.* Микроэволюция и пути видообразования. – М.: Знание, 1982. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 9, 1982)
65. *Солбриг О., Солбриг Д.* Популяционная биология и эволюция: Пер. с англ. – М.: Мир, 1982. – 488 с.
66. *Татаринов Л. П.* Палеонтология и эволюционное учение. – М.: Знание, 1985. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 9, 1985)
67. *Татаринов Л. П.* Эволюция и креационизм. – М.: Знание, 1988. – 64 с. (Новое в жизни, науке, технике. Серия «Биология», вып. 8, 1988)
68. Эволюция: Сборник: Пер. с англ. под ред. *М. В. Миной.* – М.: Мир, 1981. – 265 с.
69. *Яблоков А. В., Юсуфов А. Г.* Эволюционное учение: Дарвинизм. 4-е изд., стер. – М.: Высшая школа, 1998. – 336 с.
70. *Кликс Ф.* Пробуждающееся мышление: У истоков человеческого интеллекта. Пер. с нем. – М.: Прогресс, 1983. – 302 с.
71. *Сергеев Б. Ф.* Ступени эволюции интеллекта. – Л.: Наука, 1986. – 192 с. (Серия «От молекулы до организма»)
72. *Веккер Л. М.* Психика и реальность: Единая теория психических процессов. – М.: Смысл, 2000. – 685 с.
73. *Симонов П. В., Еришов П. М., Вяземский Ю. П.* Происхождение духовности – М.: Наука, 1989. – 352 с. (Серия «Общество и личность»)
74. *Анохин П. К.* Системные механизмы высшей нервной деятельности. – М.: Наука, 1979. – 453 с.
75. *Алейникова Т. В., Думбай В. Н., Кураев Г. А., Фельдман Г. Л.* Физиология центральной нервной системы. 2-е изд., доп. и испр. – Ростов н/Д.: Феникс, 2000. – 384 с.
76. *Данилова Н. Н., Крылова А. Л.* Физиология высшей нервной деятельности. – Ростов н/Д.: Феникс, 1999. – 400 с.

77. Блум Ф., Лейзерсон А., Хофстедтер Л. Мозг, разум и поведение: Пер. с англ. – М.: Мир, 1988. – 248 с.
78. Мозг: Сборник: Пер. с англ. под ред. и с предисл. П. В. Симонова. – М.: Мир, 1982. – 280 с.
79. Симонов П.В. Мотивированный мозг: Высшая нервная деятельность и естественнонаучные основы общей психологии. – М.: Наука, 1987. – 269 с.
80. Дискуссия о нейрокомпьютерах // Всероссийская научно-техническая конференция «Нейроинформатика-99», 19–21 января 1999 г. / Отв. ред. А. А. Фролов и А. А. Ежов. – М.: Изд-во МИФИ, 2000. – 224 с.
81. Борисяк Р.М., Виноградова О.С., Денэм М., Казанович Я.Б., Хоппенштедт Ф. Модель детекции новизны на основе частотного кодирования информации // 2-я Всероссийская научно-техн. конференция «Нейроинформатика-2000», 19–21 января 2000 г. – М.: Изд-во МИФИ, 2000. – с. 145–156.
82. Борисяк Р.М., Виноградова О.С., Денэм М., Казанович Я.Б., Хоппенштедт Ф. Модель детекции новизны на основе осцилляторной нейронной сети с разреженной памятью // III Всероссийская научно-техн. конференция «Нейроинформатика-2001», 24–26 января 2001 г. – М.: Изд-во МИФИ, 2001. – с. 183–190.
83. Кузьмина М.Г., Манькин Э.А., Сурина И.И. Оценка памяти в замкнутых однородных цепочках осцилляторов // 2-я Всероссийская научно-техн. конференция «Нейроинформатика-2000», 19–21 января 2000 г. – М.: Изд-во МИФИ, 2000. – с. 94–99.
84. Кузьмина М.Г., Манькин Э.А., Сурина И.И. Модель осцилляторной сети, имитирующая основанное на синхронизации функционирование зрительной коры // III Всероссийская научно-техн. конференция «Нейроинформатика-2001», 24–26 января 2001 г. – М.: Изд-во МИФИ, 2001. – с. 191–200.
85. Лагутина Н.С. Модель импульсного нейрона. Колебания в простейшей сети из трех нейронов. Самоорганизация полносвязной сети импульсных нейронов // III Всероссийская научно-техн. конференция «Нейроинформатика-2001», 24–26 января 2001 г. – М.: Изд-во МИФИ, 2001. – с. 200–205.
86. Мирошников С.А. Интеграция импульсных и осцилляторных сетей в нейropsychологической системе // III Всероссийская научно-техн. конференция «Нейроинформатика-2001», 24–26 января 2001 г. – М.: Изд-во МИФИ, 2001. – с. 205–213.
87. Сухов А.Г., Бездудная Т.Г., Медведев Д.С. Ритмическая активность как фактор самоорганизации и пластичности нейронной сети // III Всероссийская научно-техн. конференция «Нейроинформатика-2001», 24–26 января 2001 г. – М.: Изд-во МИФИ, 2001. – с. 213–220.

88. *Кун Т.* Структура научных революций. 2-е изд.: Пер. с англ. – М.: Прогресс, 1977. – 300 с. (Серия «Логика и методология науки»)
89. *Хаос* (тема номера, 3 статьи) // Компьютерра. – № 47 (275), 1 декабря 1998 г. – с. 20–35.
URL: <http://www.computerra.ru/offline/1998/275/>
90. *Баблянец А.* Молекулы, динамика и жизнь: Введение в самоорганизацию материи: Пер. с англ. – М.: Мир, 1990. – 375 с.
91. *Заславский Г. М., Сагдеев Р. З.* Введение в нелинейную физику: От маятника до турбулентности и хаоса. – М.: Наука, 1988. – 368 с.
92. *Лоскутов А. Ю., Михайлов А. С.* Введение в синергетику. – М.: Наука, 1990. – 272 с.
93. *Малинецкий Г. Г.* Хаос. Структуры. Вычислительный эксперимент: Введение в нелинейную динамику. – М.: Эдиториал УРСС, 2000. – 256 с.
94. *Малинецкий Г. Г., Потапов А. Б.* Современные проблемы нелинейной динамики. – М.: Эдиториал УРСС, 2000. – 336 с.
95. *Николис Дж., Пригожин И.* Познание сложного. Введение: Пер. с англ. – М.: Мир, 1990. – 344 с.
96. *Табор М.* Хаос и интегрируемость в нелинейной динамике: Пер. с англ. – М.: Эдиториал УРСС, 2001. – 320 с.
97. *Хакен Г.* Синергетика: Пер. с англ. – М.: Мир, 1980. – 404 с.
98. *Хакен Г.* Синергетика. Иерархия неустойчивостей в самоорганизующихся системах и устройствах: Пер. с англ. – М.: Мир, 1985. – 423 с.
99. *Хакен Г.* Информация и самоорганизация. Макроскопический подход к сложным системам: Пер. с англ. – М.: Мир, 1991. – 240 с.
100. *Шустер Г.* Детерминированный хаос. Введение: Пер. с англ. – М.: Мир, 1988. – 240 с.
101. *Эбелинг В., Энгель А., Файстель Р.* Физика процессов эволюции. Синергетический подход: Пер. с нем. – М.: Эдиториал УРСС, 2001. – 328 с.
102. *Эткинс П.* Порядок и беспорядок в природе: Пер. с англ. – М.: Мир, 1987. – 224 с.
103. *Эфрос А. Л.* Физика и геометрия беспорядка. – М.: Наука, 1982. – 176 с. (Библиотечка «Квант», вып. 19)
104. *Борисов А. Н., Алексеев А. В., Меркурьева Г. В., Слядзь Н. Н., Глушков В. И.* Обработка нечеткой информации в системах принятия решений. – М.: Радио и связь, 1989. – 304 с.

105. *Заде Л.* Понятие лингвистической переменной и его применение к принятию приближенных решений: Пер. с англ. – М.: Мир, 1976. – 165 с. (Серия «Новое в зарубежной науке: Математика», вып.3 / Ред. серии *А. Н. Колмогоров* и *С. П. Новиков*)
106. Классификация и кластер / Под ред. *Дж. Вэн Райзина*: Пер. с англ. – М.: Мир, 1980. – 389 с.
107. *Кофман А.* Введение в теорию нечетких множеств: Пер. с франц. – М.: Радио и связь, 1982. – 432 с.
108. *Кузьмин В. Б.* Построение групповых решений в пространствах четких и нечетких бинарных отношений. – М.: Наука, 1982. – 168 с. (Серия «Теория и методы системного анализа»)
109. *Малышев Н. Г., Бернштейн Л. С., Боженьюк А. В.* Нечеткие модели для экспертных систем в САПР. – М.: Энергоатомиздат, 1991. – 136 с.
110. *Мелихов А. Н., Бернштейн Л. С., Коровин С. Я.* Ситуационные советующие системы с нечеткой логикой. – М.: Наука, 1990. – 272 с.
111. *Орлов А. И.* Задачи оптимизации и нечеткие переменные. – М.: Знание, 1980. – 64 с. (Новое в жизни, науке, технике. Серия «Математика, кибернетика». Вып.8, 1980)
112. *Орловский С. А.* Проблемы принятия решений при нечеткой исходной информации. – М.: Наука, 1981. – 208 с. (Серия «Оптимизация и исследование операций»)
113. Прикладные нечеткие системы / Под ред. *Т. Тэрано, К. Асаи* и *М. Сугэно*: Пер. с япон. – М.: Мир, 1993. – 368 с.
114. *Нечеткая логика* (тема номера, 4 статьи) // Компьютерра. – № 38 (415), 9 октября 2001 г. – с. 18–31.
URL: <http://www.computerra.ru/offline/2001/415/>
115. *Дюбуа Д., Прад А.* Теория возможностей. Приложения к представлению знаний в информатике: Пер. с франц. – М.: Радио и связь, 1990. – 288 с.
116. Нечеткие множества и теория возможностей: Последние достижения / Под ред. *Р. Р. Ягера*: Пер. с англ. – М.: Радио и связь, 1986. – 408 с.
117. *Пытьев Ю. П.* Возможность: Элементы теории и применения. – М.: Эдиториал УРСС, 2000. – 192 с.
118. Special Issue “Fuzzy Information Processing” / Ed.: *Dan Ralescu* // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Feb. 10, 1995. – v. 69, No. 3. – pp. 239–354.

119. Special Issue "Fuzzy Signal Processing" / Eds.: Anca L. Ralescu and James G. Shanahan // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Jan. 15, 1996. – v. 77, No. 1. – pp. 1–116.
120. Special Issue "Fuzzy Multiple Criteria Decision Making" / Eds.: C. Carlsson and R. Fullér // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – March 11, 1996. – v. 78, No. 2. – pp. 139–241.
121. Special Issue "Fuzzy Modelling" / Ed.: J. M. Barone // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – May 27, 1996. – v. 80, No. 1. – pp. 1–120.
122. Special Issue "Fuzzy Optimization" / Ed.: J.-L. Verdegay // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – July 8, 1996. – v. 81, No. 1. – pp. 1–183.
123. Special Issue "Fuzzy Methodology in System Failure Engineering" / Ed.: Kai-Yuan Cai // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Oct. 8, 1996. – v. 83, No. 2. – pp. 111–290.
124. Special Issue "Analytical and Structural Considerations in Fuzzy Modelling" / Ed.: A. Grauel // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Jan. 16, 1999. – v. 101, No. 2. – pp. 205–313.
125. Special Issue "Soft Computing for Pattern Recognition" / Ed.: Nikhil R. Pal // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Apr. 16, 1999. – v. 103, No. 2. – pp. 197–367.
126. Special Issue "Fuzzy Modeling and Dynamics" / Eds.: Horia-Nicolai Teodorescu, Abraham Kandel, Moti Schneider // Fuzzy Sets and Systems: Intern. J. of Soft Computing and Intelligence. – Aug. 16, 1999. – v. 106, No. 1. – pp. 1–97.
127. Портал научных вычислений (Matlab, Fortran, C++ и т.п.)
URL: <http://www.mathtools.net/>
128. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. 3-е изд., испр. – М.: Наука, 1986. – 288 с.
129. Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягода А.Г. Численные методы решения некорректных задач. – М.: Наука, 1990. – 232 с.
130. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное издание. – М.: Финансы и статистика, 1983. – 471 с.
131. Айвазян С. А., Бухитабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Классификация и снижение размерности. Справочное издание. – М.: Финансы и статистика, 1989. – 607 с.

132. *Бендат Дж., Пирсол А.* Прикладной анализ случайных данных: Пер. с англ. – М.: Мир, 1989. – 540 с.
133. *Боровков А. А.* Математическая статистика: Оценка параметров, проверка гипотез. – М.: Наука, 1984. – 472 с.
134. *Моррис У. Т.* Наука об управлении: Байесовский подход. Пер. с англ. – М.: Мир, 1971. – 304 с.
135. Bayes net toolbox for Matlab:
URL: <http://www.cs.berkeley.edu/~murphyk/Bayes/bnt.html> A
Brief Introduction to Graphical Models and Bayesian Networks:
URL: <http://www.cs.berkeley.edu/~murphyk/Bayes/bayes.html>
136. *Пайтген Х.-О., Рихтер П. Х.* Красота фракталов. Образы комплексных динамических систем: Пер. с англ. – М.: Мир, 1993. – 176 с.
137. *Шредер М.* Фракталы, хаос, степенные законы. Миниатюры из бесконечного рая: Пер. с англ. – М.: Мир, 1993. – 176 с.
138. *Боуэн Р.* Методы символической динамики. Сб. статей: Пер. с англ. под ред. *В.М.Алексеева*. – М.: Мир, 1979. – 245 с. (Серия «Новое в зарубежной науке: Математика», вып. 13 / Ред. серии *А.Н.Колмогоров* и *С.П.Новиков*)
139. *Каток А. Б., Хассельблат Б.* Введение в современную теорию динамических систем: Пер. с англ. – М.: Факториал, 1999. – 768 с.
140. *Палис Ж., Ду Мелу В.* Геометрическая теория динамических систем. Введение: Пер. с англ. – М.: Мир, 1986. – 301 с. (Серия «Современная математика: Вводные курсы»)
141. Теория систем: Математические методы и моделирование. Сб. статей: Пер. с англ. – М.: Мир, 1989. – 384 с. (Серия «Новое в зарубежной науке: Математика», вып. 44 / Ред. серии *А. Н. Колмогоров* и *С. П. Новиков*)

Редактор материалов выпуска,
кандидат технических наук *Ю. В. Тюменцев*

E-mail: tium@mai.ru

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.

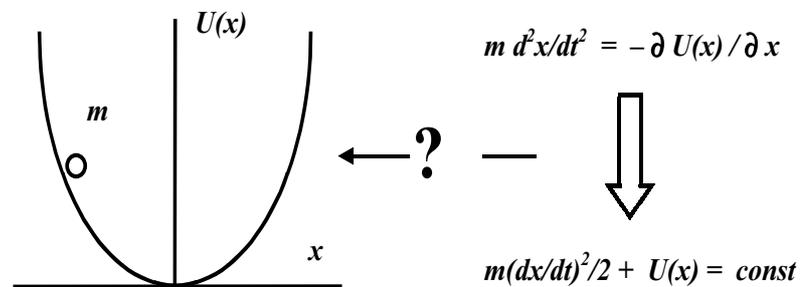


Рис. 1. Почему *человеческая* логика применима к познанию *природы*?

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 сайзер¹ — 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-

¹Сайзер (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.

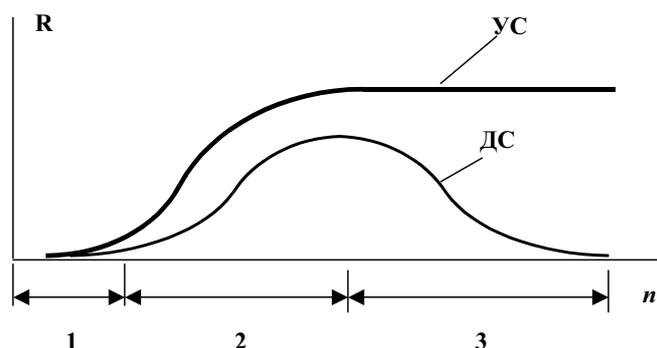


Рис. 2. Зависимость реакции на условный (УС) и дифференцировочный (ДС) стимулы от числа сочетаний n между условным и безусловным стимулами при выработке классического условного рефлекса. Очень схематично. 1 — прегенерализация, 2 — генерализация, 3 — специализация.

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

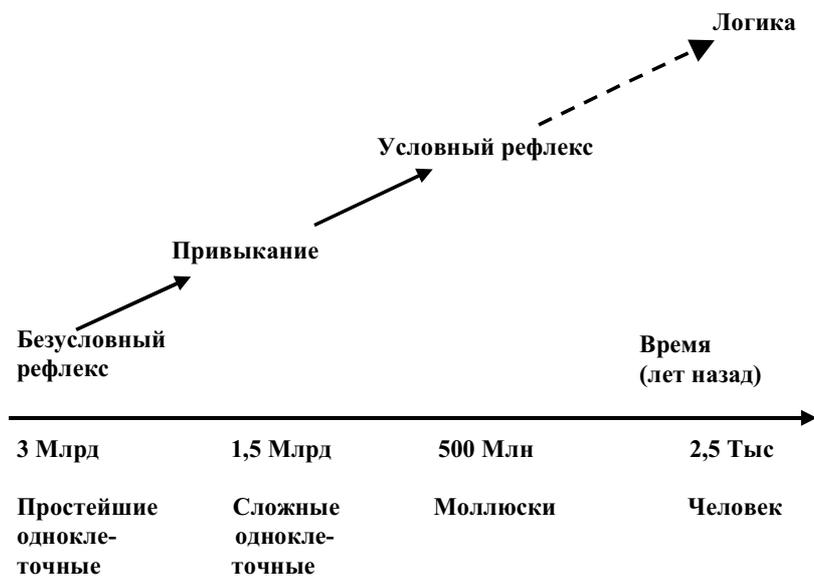


Рис. 3. «Интеллектуальные изобретения» биологической эволюции. «Авторы изобретений» и «даты приоритетов» представлены довольно условно.

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 Тьерой 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. Адама 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.

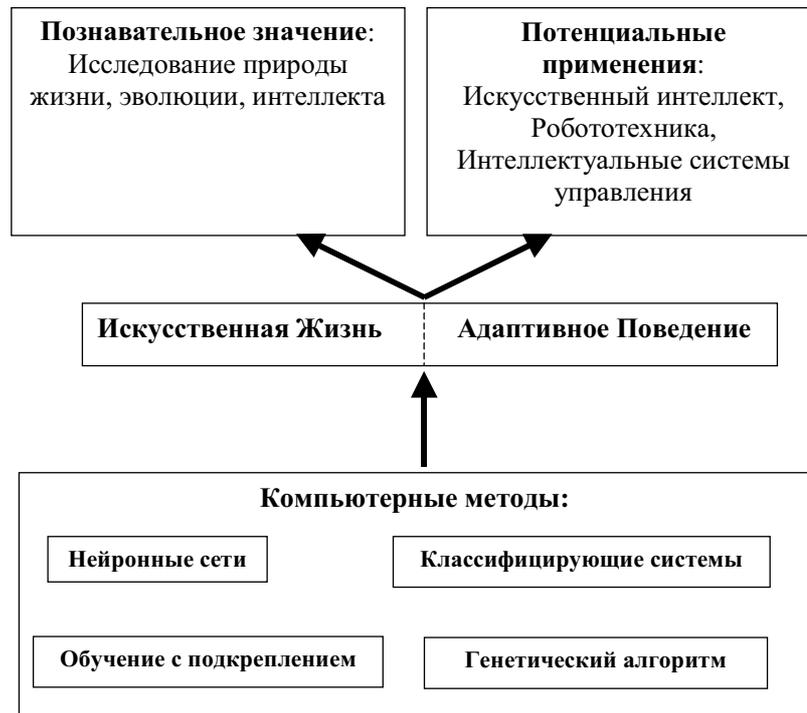


Рис. 4. Схема междисциплинарных связей направлений «Искусственная жизнь» и «Адаптивное поведение».

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;

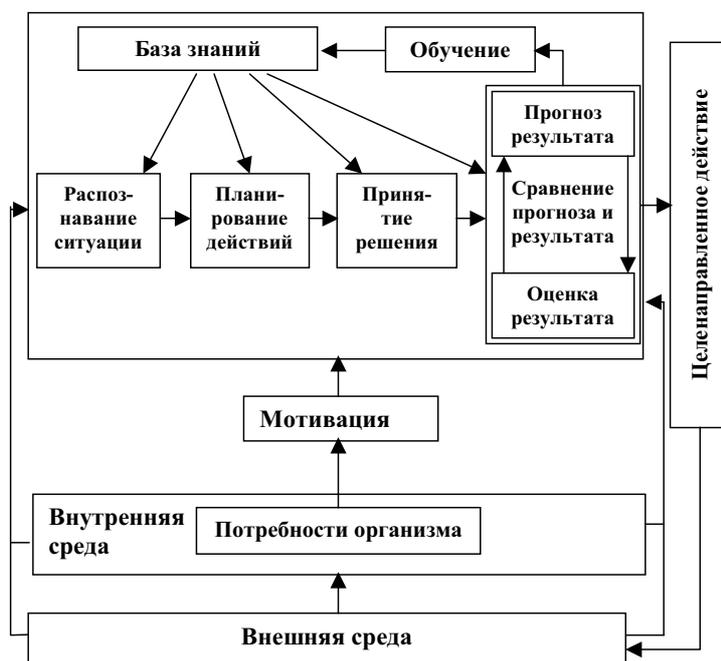


Рис. 5. Кибернетическая схема функциональной системы (в духе теории П. К. Анохина).

- 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]².

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

²См. 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;

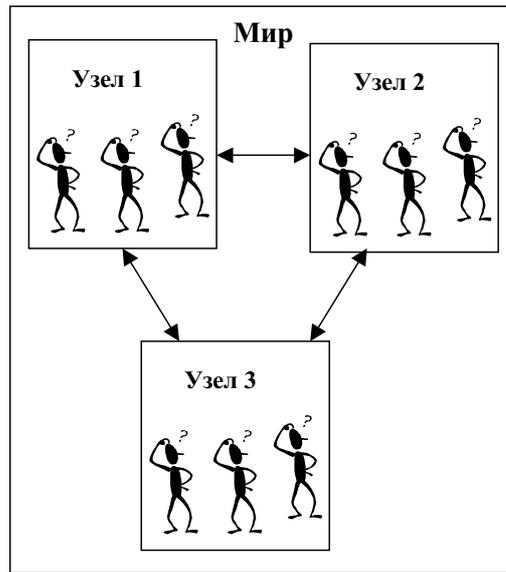


Рис. 6. Схема популяции агентов, распределенной в Интернете.

- 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 Технологиях³. По мнению этих авторов, эволюционная кибернетика должна служить основой для исследований кибернетической эволюции всего человечества, включая разработку будущих иерархических систем управления человеческим сообществом.

³Сайт 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.

ИГУМЕН ФЕОФАН (КРЮКОВ)

Свято-Данилов монастырь, Москва

E-mail: kryukov@mail.cnt.ru

**МОДЕЛЬ ВНИМАНИЯ И ПАМЯТИ,
ОСНОВАННАЯ НА ПРИНЦИПЕ ДОМИНАНТЫ**

Аннотация

Описаны шесть основных проблем внимания в терминах доминанты А. А. Ухтомского. Последняя, в свою очередь, модельно описывается как системное проявление фазовых переходов в мозге. Приводится обзор теоретических и экспериментальных оснований для существования в мозге метастабильных состояний со временем жизни порядка одной секунды и более. Это позволяет решить все основные проблемы внимания одновременно, в том числе проблемы интеграции и Центрального Управителя. В этом отношении приведены новые данные, подтверждающие ранее сделанные предсказания. Работа может быть полезной при разработке доминантного нейрокомпьютера, свободного от недостатков коннекционистских архитектур, а также для модельной интерпретации данных о функционировании мозга, полученных с помощью современной экспериментальной техники функционального картирования.

HEGUMEN THEOPHAN (KRYUKOV)

St. Daniel Monastery, Moscow

E-mail: kryukov@mail.cnt.ru

**AN ATTENTION AND MEMORY MODEL
BASED ON THE PRINCIPLE OF DOMINANTA**

Abstract

The six basic problems of attention are described in terms of the dominant focus by Ukhtsky. The latter in its turn is modeled as a system expression of the phase transition in the brain. The theoretical and experimental grounds for the existence of the metastable states in the brain with life time of about one second and more are given. This solves simultaneously all the above problems of interaction between attention and memory, binding and central executive problems are included. The previous predictions are in this respect confirmed. We hope that the work will be of use for development of neurocomputers of the new type free from the limitations of connectionist architectures and for model interpretation of the new data on brain functional imaging.

Светлой памяти
Ольги Сергеевны Виноградовой
посвящается

Введение. Основные проблемы моделирования внимания

Более столетия назад Джеймс, один из основателей когнитивной психологии, писал: «Всякий знает, что такое внимание. Это обладание умом в ясной и отчетливой форме одним из нескольких одновременно возможных объектов или последовательностей мыслей. Его сущность — фокусирование, концентрация сознания. Оно подразумевает отвлечение от некоторых предметов, чтобы более эффективно иметь дело с остальными».

Не будем искать более современного, более точного определения внимания. Приведем лишь перечень функций, которые обычно ему приписывают и кратко опишем ряд проблем внимания, с ними связанных. Под *вниманием* будем понимать психофизиологический процесс, тесно связанный с процессами памяти и выполняющий следующие функции:

- игнорирование некоторых объектов для того, чтобы более эффективно анализировать другие объекты;
- формирование кратковременной и долговременной памяти;
- реализация межмодальной интеграции;
- контроль общего возбуждения (будем именовать его далее аразузлом);
- торможение ненужной или несущественной текущей активности для получения более легкого или более эффективного отклика на новые или жизненно важные стимулы.

Главная проблема моделирования внимания состоит в следующем: если на вход модели поступает смесь сигналов от различных источников, она должна быть способной узнать, опознать любую компоненту смеси. Другие функции внимания, которые были перечислены выше, также

трудны для моделирования, но они сильно зависят от этой основной проблемы.

Проиллюстрируем важность проблемы на примере. Изучение человеческого зрения свидетельствует о том, что мозг по-разному обрабатывает зрительную информацию в зависимости от поставленной задачи. Эта обработка может быть параллельной (*предвнимание*) или последовательной (*фокусированное внимание*). В первом случае обработка быстрая, не зависящая от числа объектов в визуальном поле, и глобальные свойства изображения эквивалентны локальным свойствам. Во втором случае анализ идет сравнительно медленнее, но целое (опознанный сигнал) «больше», чем сумма его частей, потому что отброшен мало значимый фон и акцент поставлен на анализе «фигуры», наиболее важной части изображения. Предполагается, что именно оперативная и автоматическая комбинация этих двух типов анализа и объясняет весь секрет превосходства человека в распознавании образов над самыми мощными компьютерами. Этот пример также показывает, что внимание должно быть глобальным свойством системы мозга, неразрывно связанным с его функционированием как целого, а не с одной, скажем, визуальной системой.

Как же на самом деле внимание связано с восприятием, памятью и другими когнитивными функциями? Эта проблема всегда была центральной как для психологии, так и для нейронаук. Но прогресс был медленным из-за плохого понимания процессов центрального управления вниманием и представления памяти в мозге.

В последнее время роль внимания изучалась с помощью двух типов моделей. Первый тип модели, который можно назвать *коннекционистским*, основан на теоретических предположениях относительно нейронов и их связей и широко использует компьютерную имитацию для того, чтобы определить, как параллельная обработка свойств может быть в принципе соединена с последовательным вниманием, для того, чтобы увеличить общую эффективность обработки системы. Здесь известные физиологические данные относительно нейронов и их связей обычно значительно упрощаются для того, чтобы можно было применить простой математический аппарат или компьютерное моделирование. Этот тип модели не столько проясняет реальные функции внимания, сколько предлагает некоторые полезные вычислительные архитектуры для решения технических проблем идентификации сигнала. Хорошим примером тако-

го рода модели может служить модель, описанная в работе [64] (Strong & Whitehead, 1989). Однако в последнее время коннекционистские модели пытаются объяснить взаимодействие внимания и памяти [22] (Braver & Cohen, 2001), [31] (Dehaene et al., 1998). Другой тип модели – тип *психологический* или *психофизиологический*. Он предназначен для того, чтобы с наименьшим количеством теоретических допущений и концепций объяснить детальное влияние внимания на восприятие, поведение, кратковременную и долговременную память. В этих моделях архитектура моделей определяется фактическими данными, а не возможностью компьютерной имитации или применимостью конкретного математического аппарата для анализа модели. Наиболее внутренне согласованная модель этого типа дается в работах [26] (Cowan, 1988) и [27] (Cowan, 1995). В настоящей работе приводится обзор результатов моделирования третьего, промежуточного типа. В нем, подобно второму типу, ставится задача лучшего понимания психологических и нейрофизиологических данных по вниманию, но применяется, подобно первому типу модели, некоторое упрощение реальности, которое позволяет исследовать модель аналитическими методами или с помощью компьютерного имитирования. В частности, мы предлагаем модель для обсуждения следующих проблем внимания, сформулированных в слегка отличающейся форме нашими предшественниками [12] (Ухтомский, 1966), [19] (Baddeley & Della Sala, 1996), [26] (Cowan, 1988), [46] (Koch & Ullman, 1985), [61] (Sejnowsky, 1986), [67] (Treisman & Gelade, 1980).

- I. **Проблема селективности стимулов.** Каков механизм селективности стимулов: почему некоторые одновременно предъявленные стимулы привлекают внимание и таким образом получают доступ к высшей сенсорной обработке, в то время как другие не получают?
- II. **Проблема долговременной памяти.** Каков механизм взаимодействия внимания и долговременной памяти?
- III. **Проблема интеграции.** Поскольку параллельная обработка стимулов происходит путем выделения свойств этих стимулов на ранних стадиях обработки, то возникает вопрос, как и где происходит реконструкция интегрального образа.
- IV. **Проблема инерции.** Существует большой разрыв между длительностью эффектов внимания, таких, например, как привыкание, и

длительностью нейронных событий, скажем, таких, как средний межспайковый интервал. Этот разрыв составляет иногда несколько порядков. Каков физиологический механизм замедления времени при переходе от нейронов к поведению? Какова нейронная основа сохранения длительного внимания к некоторым стимулам даже и в тех случаях, когда стимул предъявляется на сравнительно короткое время?

V. **Проблема торможения и подавления помех.** Какого рода обработке подвергаются стимулы, которым не оказывается внимания: тормозятся ли они активно или просто не допускаются к дальнейшей обработке?

VI. **Проблема Центрального Управителя.** Существует ли отдельная от памяти нейронная модально-неспецифическая структура, координирующая функционирование всех звеньев процесса внимания и памяти, или же внимание есть следствие самоорганизации структур новой коры?

Принцип доминанты и основные проблемы внимания

Доминанта в физиологии, согласно Большой Советской Энциклопедии, есть «фокус возбуждения, который временно определяет характер реакции организма на внешнюю и внутреннюю стимуляцию». Сам А. А. Ухтомский в 1923 году писал: «Под именем доминанты моими сотрудниками понимается более или менее устойчивый очаг повышенной возбудимости центров, чем бы он ни был вызван, причем, вновь приходящие в центр возбуждения служат усилению (подтверждению) возбуждения в очаге, тогда как в прочей центральной нервной системе широко разлиты явления торможения» [12] (Ухтомский, 1966, с. 7). Десятью годами позже он писал: «Доминанта есть не теория и даже не гипотеза, но преподносимый из опыта принцип очень широкого применения, эмпирический закон, вроде закона тяготения, который может быть сам по себе и не интересен, но который достаточно назойлив, чтобы было возможно с ним не считаться. Я считаю ее за «принцип» работы центров не потому, что она кажется мне как-нибудь очень рациональной, но потому что она

представляется очень постоянной чертой деятельности центров. В действительности доминанта может становиться и совсем не рациональной чертой работы центров, а только очень устойчивой чертой их работы. Во всяком случае, доминанта — один из скрытых факторов нашей нервной деятельности, притом не невинный, как может показаться сначала» [12] (Ухтомский, 1966, с. 125).

Из множества классических данных, на которых основывается доминанта, мы здесь выделим лишь пять наиболее тесно связанных с соответствующими пятью проблемами, описанными во Введении.

1. В нейромышечном препарате существует оптимальная частота стимуляции, соответствующая максимальной *лабильности* (см. ниже) нервной ткани и максимальной интенсивности мышечного *тетануса* (слитного сокращения). При дальнейшем увеличении частоты стимуляции можно пройти через критическое значение, после которого мышечная активность начинает уменьшаться до тех пор, пока не будет достигнуто состояния пессимума или состояние абсолютной невозбудимости. Подобное явление оптимума и пессимума можно наблюдать и по отношению к интенсивности стимула [2] (Голиков, 1970).
2. Наряду с распространяющимся импульсным возбуждением типа «все или ничего» в нервной системе существуют градуальные *локальные стационарные возбуждения*. Они могут быть созданы как нормальной стимуляцией, так и местным охлаждением, местным электрическим током или некоторыми химическими веществами. Состояние таким образом измененных локусов может быть определено функциональным параметром *физиологической лабильности*. Он соответствует максимальной частоте стимуляции, выше которой в нервной ткани происходит трансформация ритма [2] (Голиков, 1970).
3. Кроме оптимальной частоты силы раздражения и интенсивности стимуляции существует *оптимальная лабильность* и *оптимальная поляризация*, при которых стационарное возбуждение наиболее легко трансформируется в локальную ритмическую активность, которая может распространяться за пределы начального локуса [2] (Голиков, 1970).

4. При оптимальном уровне лабильности и поляризации наблюдается феномен *усвоения ритма*, состоящий в том, что возникающая синхронно с ритмом приходящих раздражений ритмическая активность сохраняется в течение некоторого времени после прекращения подачи раздражения [2] (Голиков, 1970).
5. Переходы от локального стационарного возбуждения к возбуждению распространяющемуся наблюдаются во время выработки или угашения условных рефлексов. Передача возбуждения в коре из центра сигнала в центр эффектора сопровождается синхронностью биопотенциалов в этих центрах, причем распространение возбуждения тем легче, чем большая поверхность вовлечена в синхронную когерентную деятельность и чем выше коэффициент взаимной корреляции «замыкаемых» центров. Для некоторых животных, например кроликов, указанная синхронность и когерентность осуществляется на частоте тета-ритма. Угасательное торможение («размыкание» центров) сопровождается расхождением частот сочетаемых центров, падением синхронности и синфазности медленных биопотенциалов, снижением лабильности в этих центрах [5] (Ливанов, 1975).

Теперь, имея в виду приведенные выше данные, легче понять системный характер принципа доминанты (табл. 1 на с. 73).

С другой стороны доминанта характеризуется следующими пятью локальными свойствами:

1. **Повышенная возбудимость:** для того, чтобы группа стимулов могла войти в доминанту, порог возбудимости доминанты должен быть ниже, чем сила приходящего возбуждения.
2. **Стойкость возбуждения:** для того, чтобы возбуждение могло произвести заметный поведенческий эффект оно не должно чрезмерно быстро изменяться во времени.
3. **Способность суммировать возбуждение:** способность накапливать возбуждение не только от специфической, но и от неспецифической стимуляции.
4. **Инерция**, т. е. способность сохранять состояние возбуждения, если первоначальный стимул уже миновал.

5. **Сопряженное торможение**, т. е. способность исключать из доминанты те центры, чья активность функционально не совместима с активностью доминантной конstellляции.

Уже в одной из самых ранних публикаций о доминанте А. А. Ухтомский писал: «В высших этажах и в коре полушарий принцип доминанты является физиологической основой акта внимания и предметного мышления» [12] (Ухтомский, 1966, с. 11). Теперь, зная основные глобальные и локальные свойства доминанты, обсудим возможность решения проблем внимания, описанных во Введении.

I. **Проблема селективности** связана с повышенной возбудимостью доминанты лишь в отношении к определенным внешним стимулам, а также и с конкретным ритмическим колебательным механизмом отбора этих стимулов. Учитывая, что доминанта — это конstellляция нейронных групп и ансамблей с определенными свойствами, указанными в табл. 1, становится почти очевидным следующее простое, но немаловажное утверждение:

ТАБЛИЦА 1

Принцип доминанты А. А. Ухтомского

В каждый момент времени в нервной системе существует лишь одна активная доминирующая конstellляция (созвездие, очаг) возбужденных локальных нейронных групп или центров, характеризующаяся единым ритмом и единым действием (поведением). Одни и те же отдельные центры или группы нейронов могут входить в состав различных доминирующих конstellляций, причем вхождение в данную конstellляцию или выключение из нее определяется способностью этих групп или центров усваивать единый темп и ритм активности. Следы прежних доминант длительно сохраняются в высших отделах нервной системы и при полном или частичном восстановлении первоначальных условий могут вспыхнуть вновь полностью или частично.

«Хотя селективное внимание хорошо обоснованно поведенческими экспериментами, оставалось неясным, что в биологических терминах селективируется и как. Теперь нами¹ предложено, что селективное внимание включает выбор одного из целого ряда различных глобальных клеточных ансамблей» [53] (Miller, 1989). Чтобы оценить значение этого нам хорошо известного, но на Западе нового результата и тем самым оценить силу принципа доминанты приведем часто повторяющийся в различных вариантах один и тот же непростой вопрос, неразрешимый, по-видимому, без принципа доминанты.

«Доказано, что эффект внимания воздействует на нейроны в первичных зонах, так что сигналы управления вниманием в большинстве исследований идут сверху вниз, от высших этажей к нижним. Но как эти сигналы возвращаются на раннюю стадию обработки, чтобы селективировать некоторые синаптические входы из множества других входов — остается глубокой тайной». [23] (Britten, 1996).

«Известно, что внимание влияет на ранние стадии зрительного восприятия. Проблема в том, как префронтальные клетки из системы управления вниманием узнают, какие клетки зрительной системы надо активировать» [70] (Ungerleider, 1995).

Ответ в терминах доминанты в обоих случаях простой: сигналам высшего уровня не нужно специально отыскивать и селективировать клетки нижнего уровня первичного анализа, так как согласно принципу доминанты они вместе и одновременно войдут в одну и ту же доминантную констелляцию, если усвоят единый темп и ритм активности. Как это реализуется в модели, будет описано позже.

II. Проблема долговременной памяти — это в терминах доминанты проблема долгоживущих состояний возбуждения, т. е. проблема сохранения следов прошлых доминант. При таком отождествлении принцип доминанты, как мы увидим далее, открывает новые широкие возможности в изучении проблемы памяти и, в частности, проблемы связи долговременной и кратковременной памяти. Опять, без доминанты эту проблему, по-видимому, решить не удастся, потому что она многосложная, как свидетельствуют специалисты: «Необходимо дальнейшее исследование вопроса о связи восприятия, памяти и отклика. Некоторые модели памяти используют различные синхронности, чтобы удерживать в кратковремен-

¹Автор не знаком с доминантой. — И. Ф.

ной памяти сразу несколько представлений. Но долговременная память для объектов и событий должна использовать более долговременные следы» [68] (Treisman, 1996).

III. Проблема интеграции, или, как ее еще называют, binding-проблема с точки зрения доминанты тоже оказывается тесно связанной с проблемой памяти, а не только с синхронизацией элементарных свойств объекта, как это сейчас обычно понимается в большинстве случаев. Дело в том, что каждая переживаемая нами доминанта восстанавливается не только по сигнальным компонентам или признакам объекта, но и по всем кортикальным компонентам этой же, но ранее пережитой доминанты. При этом доминанта может быть пережита как мимолетное воспоминание с ничтожной инерцией. «И тогда она без изменения, как постоянный и однозначный интегральный образ, скрывается опять в складках памяти. Но она может быть восстановлена и пережита вновь с почти прежней полнотою, с оживлением работы во всей соматической констелляции. Тогда она вновь надолго занимает своей инерцией работу центров, подбирает вновь биологически интересные для нее раздражения из новой среды и обогащает мозг новыми данными. После такого же оживленного переживания доминанты соответствующий образ скрывается вновь переработанным и уходит в складки памяти более или менее переинтегрированным» [12] (Ухтомский, 1966, с. 36)

По этой причине мы утверждаем, что проблема интеграции не может быть решена отдельно от всех основных проблем внимания и памяти. К этому же, по существу, выводу приходит фон дер Мальсбург [51] (von der Malsburg, 1995), предлагая шесть критериев решения проблемы интеграции. Аналогичные выводы делают и другие исследователи этой проблемы. Вот два примера.

«Современные модели не предлагают возможных решений для координации памяти эпизодов, внимания и рабочей памяти. Это представляет пример того, что повсеместно считается центральной загадкой исследования сознания: проблема интеграции» [55] (Newman & Grace, 1999)

«Раскрытие процесса, управляющего мультимодальной интеграцией, продолжает представлять громадную проблему». [52] (Mesulam, 1998).

IV. Проблема инерции — это центральная проблема доминанты, проблема понимания механизмов «возбуждения больших интервалов», т. е. нейронной активности большой длительности, значительно превосхо-

дядей, скажем, постоянную времени мембранного потенциала нервной клетки: «В основе доминанты лежит своего рода принцип инерции: наклонность отвечать на раздражения все в одну сторону, в сторону текущей “рефлекторной установки”. Маховое колесо. Апериодичность реакции ЦНС на среду» [12] (Ухтомский, 1966, с. 247).

Приведем некоторые новые данные, подтверждающие актуальность проблемы. Кортикальные нейроны обезьян, выполняющих задание на кратковременную память, сохраняют активность в течение нескольких секунд после устранения стимула [50] (Lu et al., 1992), [76] (Zipser et al., 1993), [71] (Uusitalo et al., 1996), [36] (Fuster, 2001).

Нейроны префронтальной коры человека и приматов развили необычную способность обрабатывать информацию «на линии» — способность, которая, как многие думают, лежит в основе понимания, внимания и мышления. Эти нейроны обладают способностью транзитивно удерживать информацию во внимании в течение до 20 секунд. Обычно этот факт объясняется некоторым явно встроенным медленным химическим процессом [37] (Goldman-Rakic, 1996). Однако такое объяснение не всегда пригодно, поскольку величина временной задержки часто оказывается зависящей от функциональной нагрузки теста и от внешних условий, как например, в следующем опыте.

Магнитоэнцефалограмма человека во время задания на кратковременное удержание информации в памяти имеет необычную динамику колебаний активности в диапазоне 4–12 Гц. Эта активность начинается одновременно с подачей стимула (так называемый *reset*), и может продолжаться до 600 мсек. Длительность этой активности увеличивается линейно с увеличением нагрузки на память человека, т. е. с увеличением от одного до пяти объектов во внимании [66] (Tesche & Karhu, 2000).

Объяснение этих данных в терминах доминанты на первый взгляд не представляет затруднений. Проблема инерции внутренним образом связана с ключевой концепцией доминанты — физиологической лабильностью. Существенный пункт учения о доминанте состоит в том, что лабильность может изменяться и изменяется по ходу реакции, т. е. под действием стимулов и нагрузки. Соответственно меняется длительность реакции. Лабильность можно рассматривать как меру инерции активного состояния. Большая инерция означает малую способность прекращать развившееся возбуждение. Однако в «микроскопической» теории доми-

нанты возникает та же проблема: откуда возьмется большая инерция, если нейронная сеть, по предположению, это агрегат малоинерционных элементов. Эта проблема подробнее обсуждается в следующем разделе (с. 78–87).

V. Проблема Центрального Управителя — одна из наиболее острых, нерешенных, проблем не только психологии, но и самой доминанты. С одной стороны: «Нервная координация или соподчинение физиологических приборов в порядке их нервной увязки, есть процесс вынужденный и не предполагает для себя вмешательства какого-либо дополнительного «координирующего» центра. Координация в смысле соподчинения нервных актов получается по необходимости из невозможности осуществить одновременно требующиеся механизмы в одном и том же, относительно узко очерченном исполнительном аппарате» [12] (Ухтомский, 1966, с. 197). С другой стороны: «Почему центр господствует? Наиболее сильный. Наиболее биологически важный. Плеоназмы (многословие) вместо ответов. Ответа на «почему» все-таки нет!» [12] (Ухтомский, 1966, с. 250). Мы покажем далее, что ответ на «почему» все-таки есть, и что «координирующий» центр, он же Центральный Управитель, возникает, точнее, начинает функционировать как новое системное свойство вблизи точки фазового перехода.

VI. Проблема торможения и подавления помех прямо соотносится с важнейшим свойством доминанты — «сопряженным» торможением. Эту связь можно уяснить на следующем примере. Schillen и Konig (см. [60], 1994), моделируя интеграцию признаков в единый объект с помощью синхронизации осцилляторов, ввели десинхронизирующие связи между некоторыми осцилляторами с целью разрушить ошибочные интеграции разных, лишь временно перекрывающихся в пространстве объектов. Возникает вопрос: можно ли эту идею обобщить на случай, когда общими у разных объектов, кроме местоположения, будут еще какие-то другие признаки. Кажется очевидным, что при этом десинхронизирующее торможение должно быть динамическим, функциональным, а не преформированным, чтобы учесть различные варианты одинаковых признаков. Другими словами с каждой доминантой желательно связать только ей одной присущее торможение других мешающих ей доминант. Это и будет «сопряженное» торможение. Оно в принципе возможно, когда «возбуждение и торможение оказываются не противоположными процессами, но

родовым образом одним и тем же процессом с противоположным конечным эффектом, в зависимости от условий его осуществления. Основное и определяющее условие для эффекта дается степенью лабильности действующего эффектора в данный момент времени» [12] (Ухтомский, 1966, с. 601). Как такое возможно? Можно ли это понять с помощью простой физической аналогии?

Фазовые переходы в мозге и принцип доминанты

Идея возникновения торможения из возбуждения, наподобие физических фазовых переходов в противоположность торможению за счет структурно-фиксированной тормозной субстанции, пронизывает все аспекты учения Введенского–Ухтомского. Это учение до сих пор сталкивается с трудностями и непониманием вследствие того, что довольно сложно, малопонятно и нередко парадоксально лежащее в его основе физическое явление *фазовых переходов*. А. А. Ухтомский полагал, что дальнейшее развитие физики магнетизма и математической теории нелинейных колебаний прояснит многие из этих парадоксов. В этом направлении он предложил в 1940 году совместно с П. Гуляевым первую модель нелинейных колебаний, объясняющую *усвоение ритма, оптимум и пессимум частоты и силы стимуляции, эффекты принудительной синхронизации, инерционного отставания, деления частоты* и другие эффекты связанных нелинейных осцилляторов, частично упомянутые в начале предыдущего раздела (с. 71–72). В этой же работе он ставит важный вопрос: «Нельзя ли приписать общие черты поведения физиологических систем переменной лабильности и нелинейных колебательных систем тому, что и там и здесь дело идет, во-первых, об инерционных системах и, во-вторых, о различных выражениях преодоления инерции» [13] (Ухтомский, 1978, с. 194), и сам же отвечает на этот вопрос в другой работе: «Уж если роль следов во времени заставляет учитывать себя в *магнитах*, в коллоидном *гистерезисе*, в отрезке нерва, то надо думать, что в клетке, да еще в нервной, да еще кортикальной, передача следов от момента к моменту должна иметь первенствующую роль» [13] (Ухтомский, 1978, с. 602).

Другими словами, инерционность как центральное свойство доминанты, а с ним и все остальные свойства доминанты аналогичны коллективным эффектам в ферромагнетиках, в которых, как известно, существу-

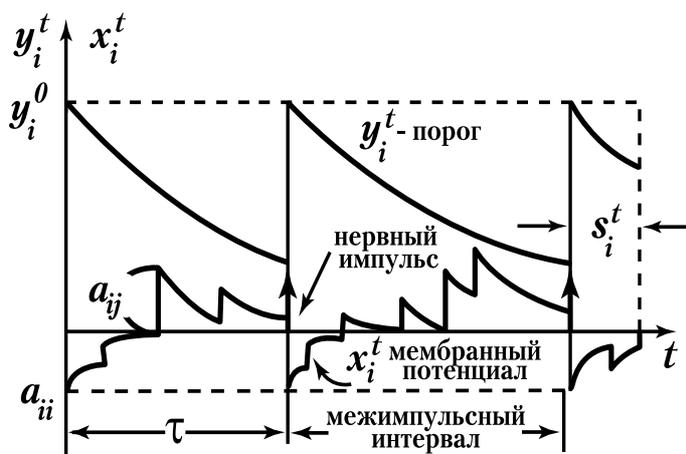


Рис. 1. Схематическое изображение допущений базовой модели нейронной сети

Генерация нервного импульса происходит в момент пересечения мембранного потенциала x_i^t (скачкообразный процесс) и затухающего порога y_i^t , где a_{ij} — синаптический вес связи с i -го на j -й нейрон — удовлетворяет условиям изинговского ($a_{ij} = 0, |i - j| > 1$) ферромагнетика ($a_{ij} > 0, |i - j| = 0, 1$). Главное отличие базовой нейронной модели от стохастической модели Изинга, со спинами в непрерывном пространстве состояний, в том, что в каждом узле плоской решетки развиваются локально-взаимодействующие процессы генерации нервного импульса, изображенные выше.

ет так называемое критическое замедление кинетики вблизи точки фазового перехода. Таким образом, А. А. Ухтомский предсказывает существование фазовых переходов и метастабильности в коре больших полушарий мозга: «На доминанту можно смотреть как на состояние *неустойчивого равновесия* молекул, способное к каталитическому (детонационному) разложению, в то время как обычные реакции — равновесие *обратимых реакций*» [14] (Ухтомский, 1996, с. 241).

Но если теория фазовых переходов и метастабильности в мозге способна единообразно объяснить все указанные свойства доминанты, тогда,

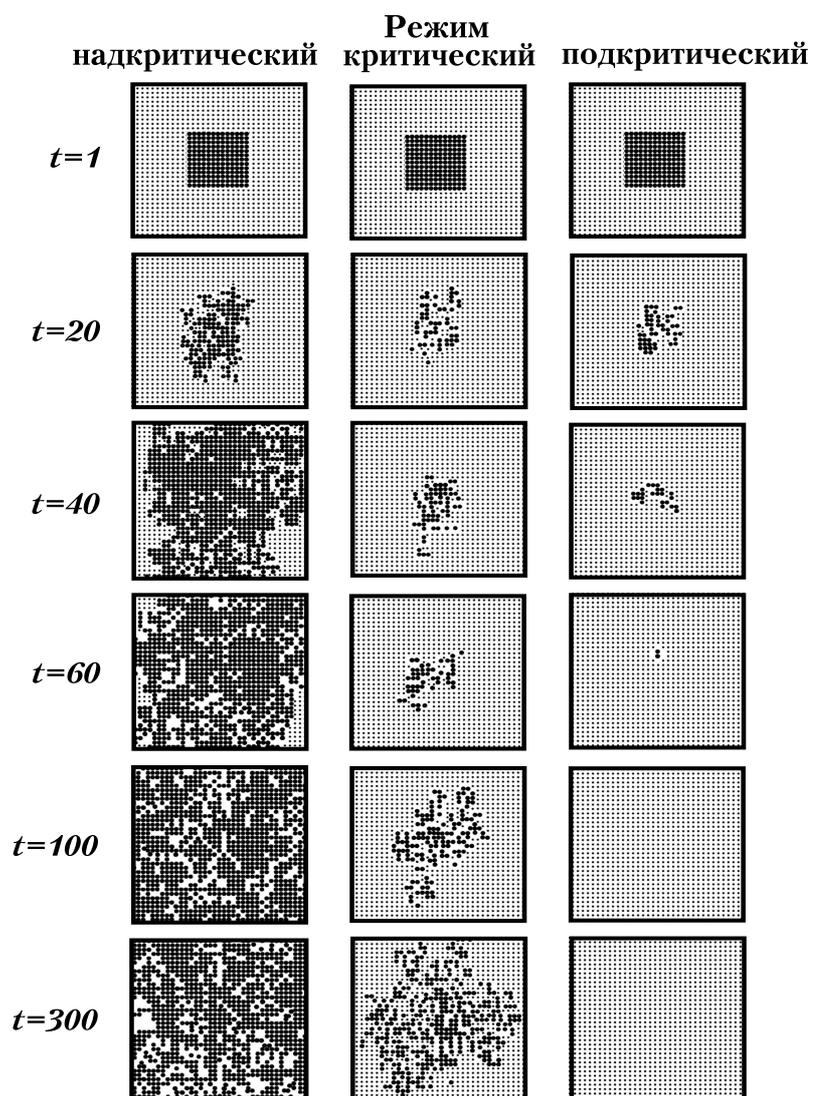


Рис. 2. Эффект сохранения пятна

учитывая их тесную связь с описанными выше проблемами внимания, мы получаем возможность решить все эти проблемы одновременно, если действительно окажется, что в коре больших полушарий происходят неравновесные фазовые переходы. Но как доказать, что фазовые переходы в мозге действительно происходят?

В последние годы концепцию неравновесных фазовых переходов в мозге развивает известный немецкий ученый Г. Хакен [39] (Haken, 1983), [40] (Haken, 1996). Он считает, что эта концепция будет полезна для теории мозга. В частности, он почти дословно повторяет А. А. Ухтомского, когда пишет, что «мозг работает вблизи точек неустойчивости, если сравнивать мозг с некоторой динамической системой» [39] (Haken, 1983, с. 603). Однако доказать существование фазовых переходов в мозге он не может в силу того, что его синергетическая модель мозга, по существу, — феноменологическая, а не микроскопическая, т. е. его модель заранее содержит утвердительный ответ на поставленный вопрос в предлагаемой им специальной форме дифференциального уравнения. Правильный подход здесь был бы в том, чтобы основным элементом модели взять не уравнение кинетики, а модельный пороговый нейрон с затуханием, учитывающий основные экспериментальные данные о связях и внутреннем устройстве реального нейрона. Тогда уже дело строгой математической теории ответить на вопрос о возможности или невозможности фазовых переходов в сети, собранной из таких «физиологических» нейронов. Именно этот подход был реализован в наших исследованиях. Краткая история вопроса такова.

Первой математической моделью нейронной сети с локальным взаимодействием, в которой удалось строго доказать существование кооперативного эффекта, соответствующего физическому фазовому переходу, была бесконечная сеть из формальных нейронов в дискретном времени, предложенная Ставской и Пятецким–Шапиро (1968, [11]). Условия отсутствия фазовых переходов для аналогичной модели в непрерывном времени привел Добрушин (1971, [3], [4]). Он же построил пример, основанный на кинетической модели Изинга, в котором фазовый переход действительно существует, и высказал гипотезу: для конечной сети локальное взаимодействие компонент приведет к очень длительному переходу в стационарное состояние, если параметры и начальные условия выбраны должным образом. Работа содержала по существу первую строгую по-

становку задачи об описании долгоживущих состояний в довольно широком классе марковских процессов. Существенным ограничением здесь, однако, оказалось *условие обратимости* процесса во времени, а также применимость лишь к сетям из формальных нейронов.

Литл [49] (Little, 1974) высказал гипотезу о существовании в мозге долгоживущих состояний по аналогии с магнитными доменами, возникающими в модели Изинга при температуре ниже критической. Нейроны у него снова формальные, время дискретное и марковская цепь *обратимая*, что оставляет гипотезу недоказанной и очень маловероятной.

В работе [6] (Крюков с соавт., 1986) предложена сравнительно несложная, но достаточно близкая к физиологии модель сети из неформальных нейронов, названная *базовой нейронной моделью*, и строго доказано существование в ней фазового перехода при некоторых жестких формальных ограничениях. Модель учитывает импульсную природу нейронной активности, экспоненциальное затухание мембранного потенциала и рефрактерность (рис. 1 на с. 79). Очевидно, что без учета этих факторов говорить о существовании фазовых переходов в нервной системе невозможно. Наоборот, доказав существование фазовых переходов в базовой модели, мы получили возможность сопоставить следствия, вытекающие из этого факта, с реальными экспериментами и, в конце концов, пришли к заключению, что фазовые переходы уже не только в модели, но и в реальном мозге действительно происходят. Изложим кратко последовательность шагов на этом пути.

Базовая модель представляет из себя марковскую систему, полученную из большого числа N первоначально независимых компонент, объединенных в систему с помощью локальных связей типа физического парного потенциала. Из микротeorии систем такого типа следует, что существование фазового перехода, когда N стремится к бесконечности, влечет при некоторых дополнительных условиях существование метастабильных состояний для конечных N . Численные эксперименты на ЭВМ подтвердили это положение. При этом был обнаружен новый «эффект сохранения пятна» (рис. 2 на с. 80), состоящий в том, что начальная констелляция активности нейронов в форме компактного пятна определенного размера вблизи критического режима имеет время жизни на несколько порядков превышающее характерные времена релаксации мембранного потенциала. Тем самым существование фазового перехода в базовой

модели открывало путь для разработки адекватной динамической теории кратковременной нейронной памяти как критического явления, существующего вблизи точки фазового перехода.

Довольно неожиданным оказалось, что небольшая модификация теории среднего поля, изложенная в работе [6] (Крюков с соавт., 1986) качественно объясняет большое количество очень трудно интерпретируемых экспериментов из самых различных отделов ЦНС, в том числе все основные свойства доминанты А. А. Ухтомского (табл. 2).

ТАБЛИЦА 2

Основные результаты исследования базовой модели

- Доказано существование «далекого порядка» при некоторых ограничениях. Предсказано существование метастабильного состояния.
- На имитационной модели получено подтверждение предсказанного эффекта метастабильности.
- Предложен и исследован нейронный осциллятор с необычными свойствами, основанный на эффекте метастабильности.
- Указанные результаты применены для интерпретации трудных для понимания данных из различных мозговых структур.
- Сделан вывод о том, что функциональной основой работы мозга являются метастабильные и неустойчивые состояния, как нейронный субстрат доминанты А. А. Ухтомского.

Экспериментальные подтверждения математической теории о существовании в нервной сети равновесных и неравновесных фазовых переходов и временно устойчивых состояний были собраны в той же работе из опубликованных до 1985 года данных по четырем нейронным структурам: гиппокампу, септуму, мозжечку и новой коре. То обстоятельство,

что одна и та же модель и один и тот же динамический эффект метастабильности оказались пригодными для объяснения большого набора трудных для понимания данных весьма непохожих нейронных структур свидетельствует о том, что мозг работает по единому принципу, а все различия между его структурами имеют, грубо говоря, количественный, а не качественный характер.

Мы приводим таблицу соответствия между фактами ферромагнетизма и данными нейрофизиологии (табл. 3 на с. 85), которая свидетельствует о неслучайности магнитной аналогии для мозга. Обращает на себя внимание прежде всего сходство, иногда тождественность понятий и терминов в столь различных областях (например, спонтанная активность, восприимчивость, гистерезис). Затем можно отметить значительное разнообразие весьма трудных для объяснения эффектов, которые кажутся самостоятельными, независимыми друг от друга, а по существу являются различными аспектами одного и того же явления фазового перехода.

Наконец, в последней строке таблицы записана гипотеза, которую высказывали прежде очень многие нейрофизиологи и специалисты по моделированию нервной системы. Теперь, на наш взгляд, она является неизбежным следствием всех приведенных нами фактов теоретических и экспериментальных.

Перечислим кратко, какие задачи и проблемы удается решить с помощью теории фазовых переходов в мозге. Существование фазовых переходов впервые позволило:

- понять механизм инерционности доминанты как критического замедления кинетики микроочагов;
- устранить основные дефекты реверберационной гипотезы кратковременной памяти, такие, как малый период рециркуляции и низкая его стабильность;
- представить непротиворечивую интерпретацию пяти локальных свойств доминантного очага;
- объяснить несколько парадоксов, связанных с динамикой электрической активности мозга при выработке и угашении условных рефлексов;

ТАБЛИЦА 3. Фазовые переходы и метастабильность в мозге (Крюков с соавт., 1986; см. [6])

Ферромагнетизм	Нейрофизиология
Далекий порядок	Корреляция активности удаленных нейронов
Спонтанная намагниченность	Спонтанная активность
Критическое замедление	Инерционность (Ухтомский, 1978; см. [13])
Критическое поле	Критический уровень деполяризации (Голиков, 1970; см. [2])
Критическое уширение автокорреляций	Уширение автокорреляций (Ebner, Bloedel, 1981; см. [32])
Критическая восприимчивость	Повышенная восприимчивость (Ebner, Bloedel, 1981; см. [32]). Повышенная реактивность (Голиков, 1970; см. [2])
Равновесный фазовый переход первого рода	Скачкообразные, случайные во времени переходы между возбужденными и тормозными состояниями (Colin et al., 1980; см. [25])
Неравновесный фазовый переход. Гистерезис	Выработка и угашение условного рефлекса (Мнухина, 1977; см. [8]). Гистерезис (Colin et al., 1980; см. [25])
Кластеры, домены	Микроочаги возбуждения

- предложить в качестве нейронной основы долговременной памяти микроочаги возбуждения в колонках новой коры;
- решить основную проблему мозжечка о функциональном назначении двух его основных входов и на этой основе предложить новую модель организации движения, как системы фазовой автоподстройки частоты; при этом единство восприятия и действия, требуемое принципом доминанты, реализуется взаимодействием двух идентичных систем фазо-частотного регулирования.

Существование долгоживущих микроочагов возбуждения позволяет выдвинуть следующее новое предсказание о конкретной форме связи долговременной и кратковременной памяти (рис. 3). Если кратковременную память отождествить с активностью ревербераторов, не обязательно синхронных, чаще асинхронных, а долговременную память — со стационарным локальным возбуждением в микроочагах в корковых слоях, то переход из одной формы в другую должен быть тождественным с хорошо изученным в теории доминанты переходом от локального стационарного возбуждения к распространяющемуся [2] (Голиков, 1970). При этом доминанта А. А. Ухтомского, как констелляция совозбужденных на одной частоте множества колонок — это способ «оживить» неактивные долговременные следы, легко и быстро отобрать из них лишь те, которые были «записаны» когда-то в аналогичной ситуации (т. е. при той же частоте и фазе колебаний), скомбинировать различные модальности в единый интегральный образ и снова перевести результат в неактивную форму стационарного возбуждения, хорошо защищенную от внешних и внутренних помех. Важно, что это предсказание относится, как к сенсорной, так и к двигательной системе. Недавно это предсказание до некоторой степени подтверждено в опытах по обучению новым движениям [21] (Brashers-Krug et al., 1996).

Однако проблема устойчивости доминанты, поставленная еще А. А. Ухтомским (1966, [12], с. 264), пока остается нерешенной. Теперь с учетом долгоживущих микроочагов возбуждения ее можно было бы сформулировать следующим образом: как совместить высокую надежность и устойчивость нейронной памяти (следы прошлых доминант) с принципиально неустойчивым режимом метастабильных состояний.

Предлагаемое ниже, в разделе «Нейролокатор — Центральный Управитель» (с. 96–104) решение проблемы устойчивости доминанты состоит

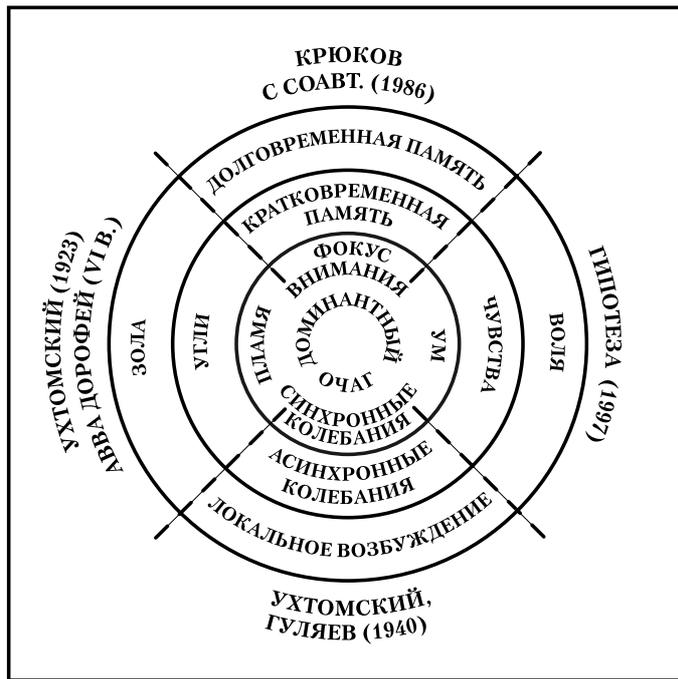


Рис. 3. Четыре представления доминанты А. А. Ухтомского

в том, что в мозге, кроме долгоживущих микроочагов возбуждения, должна существовать фазо-частотная следящая система, синхронизирующая эти микроочаги и вызывающая в них критический режим лишь на периоды времени, необходимые для фиксации или считывания информации из долговременной памяти. Здесь решение проблемы устойчивости можно уподобить равновесию движущегося велосипедиста: исходно неустойчивое состояние при определенных условиях оборачивается источником гибкого и экономного управления ЦНС с помощью *Центрального Управителя*. Эта новая концепция психологии, предложенная в работе [17] (Baddeley & Hitch., 1974) радикально меняет всю существующую архитектуру обработки информации в мозге.

Доминантная архитектура обработки информации в мозге

Мозг человека — многофункциональный орган, обеспечивающий, главным образом, целостность восприятия мира, фиксацию опыта, планирование, реализацию действий душевной и духовной жизни. Обработка сенсорной информации — не самая главная из его функций. Тем не менее, мы рассмотрим архитектуру такой обработки, отвлекаясь от высших психических функций, чтобы уяснить тесную связь внимания и памяти человека. Затем в следующем разделе кратко опишем входящие в нее компоненты.

Чтобы понять отличие доминантной архитектуры от более традиционной коннекционистской, напомним основные постулаты последней.

1. Функциональной единицей обработки информации является здесь нейроподобный элемент, формальный нейрон. Каждый элемент выполняет простую операцию над сигналами своих входов.
2. Вся информация накапливается в весах связи между элементами, которые модифицируются в процессе обучения на основе локальной информации по так называемому хеббовскому правилу.
3. Типичная коннекционистская архитектура представляется параллельной трехслойной сетью из таких элементов, в которых кроме входного и выходного слоя присутствует внутренний, «скрытый» слой. В вариантах архитектуры более развитых и лучше учитывающих данные нейробиологии число слоев и число входов может быть увеличено до четырех и более (рис. 4 на с. 89).
4. Наряду с локальными слоями иногда вводится так называемое глобальное рабочее пространство. Последнее выполняет роль коммутатора локальных пространств или локальных процессоров.
5. Внимание соответствует активации одного из локальных процессоров, функция которого в том, чтобы усилить, высветить, как прожектором, некоторую часть глобального рабочего пространства, усиливая ему соответствующие связи и ослабляя связи от мешающих слоев.

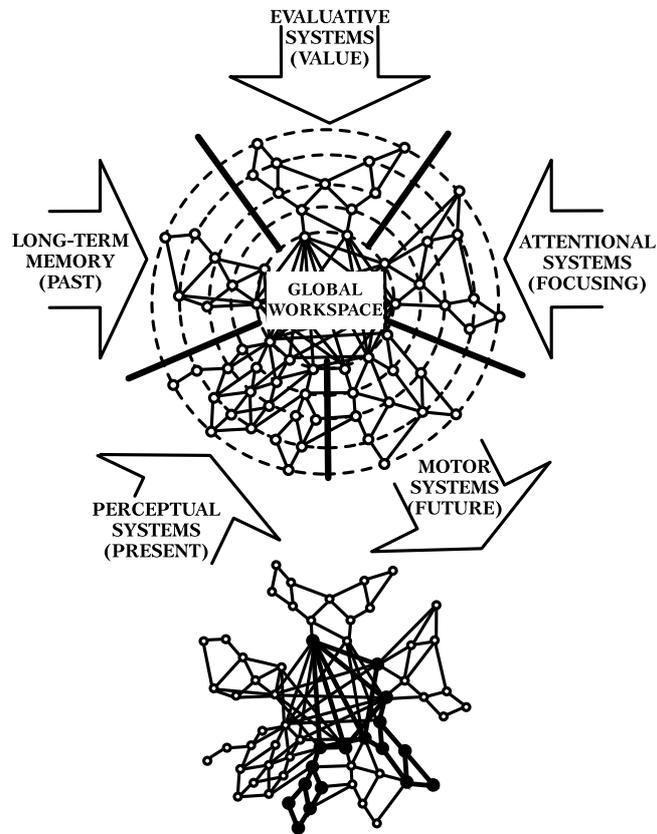


Рис. 4. Схема коннекционистской архитектуры (Dehaene et al, 1998; см. [31])

Вверху: Схематическое представление пяти главных процессов, взаимодействующих через глобальное рабочее пространство (global workspace): evaluative systems (value) — системы оценивания (значение); attentional systems (focusing) — системы внимания (фокусирование); motor systems (future) — моторные системы (будущее); perceptual systems (present) — системы восприятия (настоящее); long-term memory (past) — долговременная память (прошлое). *Внизу:* Пример активации при решении одной из задач распознавания: два процессора установили связь через глобальное рабочее пространство.

Доминантная архитектура отличается от коннекционистской в каждом из пяти указанных признаков.

1. Функциональной единицей является не одиночный нейрон, а сравнительно небольшая сеть из «физиологических» нейронов (пороговых импульсных элементов с затуханием), входящих в микроколонку коры и соединенных между собой таким образом, чтобы возник элементарный кортикальный осциллятор.
2. Информация накапливается в локальных очагах возбуждения в форме *графов изолабильности*. Элементы этого графа способны как к взаимной синхронизации на общей частоте, так и к затухающей активности, сохраняя при этом память о своей оптимальной лабильности.
3. Упрощенная доминантная архитектура, изображенная на рис. 5 (см. с. 89), представляется в виде системы из группы независимых периферических осцилляторов и одного центрального осциллятора, имеющего двусторонние связи с каждым из периферических.
4. Центральный осциллятор выполняет функции Центрального Управителя, «дирижера», способного изменять свою собственную частоту и вовлекать некоторые периферические осцилляторы в различные конstellации, условно показанные в нижней части рис. 5.
5. Появление внимания соответствует возникновению синхронных и до некоторой степени синфазных колебаний центрального и одного или нескольких периферических осцилляторов.

В целом, сравнивая эту схему с коннекционистской схемой (рис. 4), замечаем некоторое формальное их сходство. Однако, различие между ними принципиальное, качественное, и оно примерно соответствует различию между телеграфом и радиосвязью. В доминантной архитектуре все связи между осцилляторами, кроме одной — с центральным осциллятором, — функциональные, а не фиксированные. Поэтому здесь отсутствуют все известные проблемы коннекционизма, такие, как комбинаторный взрыв, катастрофа переполнения, катастрофа объединения и др. Более детально доминантная архитектура показана на рис. 6 (с. 92), где

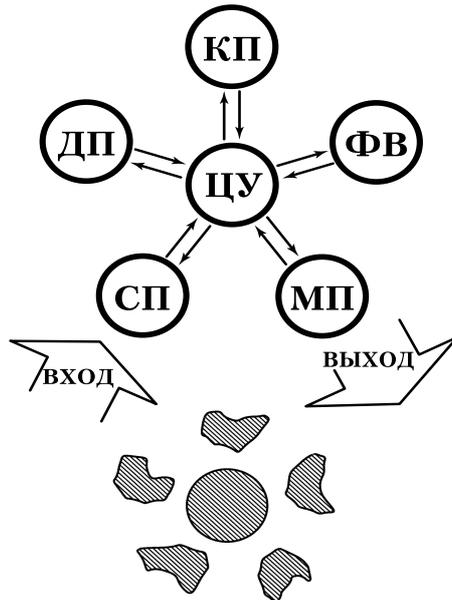


Рис. 5. Упрощенная схема доминантной архитектуры
Вверху: Схематическое представление пяти главных групп осцилляторов, взаимодействующих через центральный осциллятор. *Внизу:* Пример конstellации осцилляторов доминанты: все заштрихованные осцилляторы установили взаимную связь через центральный осциллятор. Все обозначения как на рис. 6.

учтены как свойства доминанты, так и данные психологии, нейропсихологии и нейрофизиологии [26] (Cowan, 1988), [53] (Miller, 1989), [74] (Vinogradova et al., 1991). Здесь, во-первых, учтено иерархическое строение доминанты: фокус внимания (ФВ) входит в кратковременную память (КП), а последняя, в свою очередь, входит в долговременную память (ДП), как это показано на рис. 3. Во-вторых, здесь явно указано возможное участие ЦУ во всех основных переходах и его отделенность от памяти [18] (Baddeley, 1996). В-третьих, в явном виде указаны все

возможные переходы из одной памяти в другую под действием ЦУ или внешних условий. В-четвертых, явно указана совместимость и до некоторой степени взаимозаменяемость сенсорной и двигательной систем, о чем мы писали в предыдущем разделе.



Рис. 6. Доминантная архитектура обработки сенсорной информации в мозге (Кругиков, 1991; см. [48])

СП — сенсорная память; КП — кратковременная память; МП — моторная память; ДП — долговременная память; ФВ — фокус внимания; ЦУ — Центральный Управитель.

Следует учесть, что на рис. 6 не показана явно важная черта доминантной архитектуры, связанная с регулированием общего неспецифического возбуждения или аразула, о чем будет речь в дальнейшем, а также не указаны реально существующие локальные связи между кортикальными осцилляторами, позволяющие сохранить некоторую работоспособ-

ность системы при отсутствии ЦУ. Однако в отсутствие ЦУ многие важные функции системы утрачиваются, что сравнимо по последствиям с удалением у человека такой важной структуры как гиппокамп. Утрачивается, например, способность переводить новую информацию из КП в ДП, сильно ухудшается способность к воспроизведению эпизодических событий [54] (Nadel & Moscovitch, 1998).

Итак, доминантная архитектура позволяет реализовать *последовательную* обработку поступающей информации от формирования ФВ до записи в ДП. Промежуточным, буферным состоянием в этом процессе является формирование КП. В ФВ находится та группа кортикальных осцилляторов, которая в данный момент синхронизирована по частоте и по фазе с ЦУ. Активированные внешним воздействием кортикальные осцилляторы, не входящие в ФВ или уже вышедшие из ФВ, образуют КП. Переключение ФВ не приводит к немедленной десинхронизации уже сформированной группы синхронных осцилляторов. Эти группы какое-то время сохраняют синхронность без контроля со стороны ЦУ. Благодаря этому осуществляется переход из КП в ДП. Для этого используется феномен *усвоения ритма*, описанный в разделе «Принцип доминанты и основные проблемы внимания» (см. с. 72) и в данном случае означающий, что осцилляторы, проработавшие достаточно большое время в синхронном режиме, постепенно меняют свою собственную частоту на ту, на которой они были синхронизованы. Эта частота, точнее, ей соответствующая *лабильность*, сохраняется и после угашения колебаний, она воспроизводится после того, как эти осцилляторы под внешним воздействием снова станут активными. Таким образом, запомненный в ДП паттерн активности может быть возвращен в КП или в ФВ. Попадание осцилляторов в ФВ может быть вызвано тремя причинами: появлением «нового»² стимула, заменой имеющегося стимула на «новый», спонтанной или целенаправленной активацией с помощью ЦУ. Последнее обеспечивает возможность ассоциативной связи между последовательными доминантными состояниями сети, генерируя, таким образом, связанную последовательность представлений (предметная «мысль» или «живое» движение).

²Под «новым» понимается не только отсутствующий в КП и ДП стимул, но и вообще любой другой стимул, достаточно сильный или биологически значимый, способный захватить ФВ.

Теперь приведем новые данные, существенно подтверждающие предложенную архитектуру.

1. КП и ДП имеют один и тот же нейронный субстрат в новой коре. Этот субстрат состоит из системы рассеянных, рассредоточенных по коре и перекрывающихся иерархически организованных сетей кортикальных нейронов [34], [35] (Fuster, 1998).
2. Одни и те же клетки входят в различные памятные конфигурации, включающие разные входы и разные модальности. В частности, такое совмещение имеет место в перцептивной и моторной коре [35] (Fuster, 1998b).
3. ЦУ по данным когнитивной психологии [18] (Baddeley, 1996) способен выполнять следующие функции:
 - координировать выполнение одновременно двух отдельных заданий;
 - генерировать последовательность представлений, извлекая их из долговременной памяти;
 - тормозить другие мешающие воздействия;
 - создавать ДП и ею манипулировать.
4. Гиппокамп играет критическую роль в образовании ДП [38] (Grasby et al., 1993), критическую роль в байндинге многих входов [24] (Cohen et al., 1999). Он нужен не только для записи информации в ДП, но и при считывании ее из ДП [54] (Nadel & Moscovitch, 1998). Эти данные подтверждают доминантную архитектуру, в которой ЦУ отделен от памяти, как гиппокамп отделен от новой коры.
5. Подтверждена гипотеза, что консолидация памятных следов осуществляется через образование кортикального клеточного ансамбля, который сообщает когерентность множеству рассеянных по неокортексу памятных следов, причем прерывание этой консолидации приводит к ухудшению памяти, наблюдаемой у амнептиков [56] (Paller, 1997).

6. Прерывание процесса консолидации заученных движений путем обучения второй двигательной задаче сразу же после обучения первой задаче приводит к разрушению моторной памяти первой задачи. Но такое разрушение отсутствует, если прошло четыре часа после обучения первой задаче. Сравнивая это наблюдение с консолидацией следов после сенсорного обучения, делается вывод, что две различные нейронные системы моторной и сенсорной памяти имеют сходные характеристики кодирования и запоминания новой информации [21] (Brashers-Krug et al., 1996). Заметим, что это есть лишь одно из следствий идентичности механизма этих систем, предсказанное в работе [48] (Kryukov, 1991).

Последний факт не потребует изменения предложенной архитектуры, изображенной на рис. 6, если мы предположим, что ЦУ содержит две параллельно работающие и, по-видимому, тесно взаимодействующие фазо-частотные управляющие системы. Причем, одна из них базируется на септо-гиппокампальной системе, а другая на оливо-мозжечковой, о чем была речь в одном из предыдущих разделов.

В заключение выскажем гипотезу о возможном взаимодействии *декларативной* и *процедурной* памяти, которые обычно противопоставляются друг другу: первая требует внимания, и ее содержание можно сообщить другим, в то время как вторая — автоматическая и поэтому уже не требует внимания, но и не сообщается другим. Гипотеза состоит в том, что декларативная память, реализуемая с помощью гиппокампальной системы [33] (Eichenbaum et al., 1996), с точки зрения теории автоматического управления принципиально не будет отличаться от *процедурной* (бессознательной, автоматической) памяти, если учесть, что последняя, подобно первой, реализуется системой фазо-частотной синхронизации на основе оливо-мозжечковой системы. По этой гипотезе обе системы могут работать на одной и той же частоте тета-ритма и могут до некоторой степени заменять одна другую. Не в этом ли причина удивительных данных о возможности обучения без сознания, об успешности бессознательной двигательной активности, о фактах слепого зрения (blindsight) и других загадочных явлениях, до сих пор не имеющих удовлетворительного объяснения [30] (Danckert & Goodale, 2000), [43] (Kentridge et al., 1999).

«Нейролокатор» — Центральный Управитель

Мозг во многих отношениях подобен радарной системе, причем, как на поведенческом, так и на нейронном уровне. Действительно: «Всякий раз, как имеется симптомокомплекс доминанты, имеется предопределенный ею вектор поведения» [12] (Ухтомский, 1966, с. 604)]. Подобную зависимость П. В. Симонов назвал «принципом радара», понимая под ним избирательную готовность мозга к ответу на определенный стимул при его поведении в среде, активный поиск этого стимула [10] (Симонов, 1981, с. 37). Еще точнее и конкретнее эта аналогия прослеживается в психологической теории человека-оператора [28] (Craig, 1948), где широко и успешно применяется теория автоматических следящих систем импульсного регулирования. Не удивительно, что похожие идеи неоднократно в прошлом выдвигались и на нейронном уровне. Так, в теории нейронной памяти Джон свою радарную аналогию организует вокруг центрального узла радарной системы — *компаратора*, осуществляющего сравнение времен прихода импульса запроса и импульса ответа из долговременной памяти (гипотеза детектора совпадения) [41] (John, 1967). Здесь Джон, *опираясь на доминанту А. А. Ухтомского*, высказал одну из важных идей теории нейронной памяти о том, что возможен почти мгновенный поиск конкретного памятного следа без всякого сканирования по всей памяти. Им будет наиболее вероятный пространственно-временной паттерн, лучше всего соответствующий заданным условиям внешней и внутренней среды животного, и будет именно тот, который наиболее легко отзывается на импульс запроса при заданном симптомокомплексе.

Поясним, что здесь под памятью, в соответствии с учением о доминанте понимается фиксация топологии коактивности нейронов в форме синхронизации по частоте и по фазе различных кортикальных и гиппокампальных областей на частоте тета-ритма [5] (Ливанов, 1975). Распознавание и считывание памятных следов соответствует возрождению прежде бывшей конфигурации коактивных нейронов при прежних или похожих условиях.

В работах [1] (Виноградова, 1975), [73] (Vinogradova, 1978), [75] (Vinogradova, 2001) подробно исследовано применение гипотезы компаратора к гиппокампу и новой коре, и эти результаты легли в основу «Нейролокатора». Итак, в соответствии с компараторной гипотезой Джона и Виноградовой, мозг, подобно радару, производит в подкорке запрос-

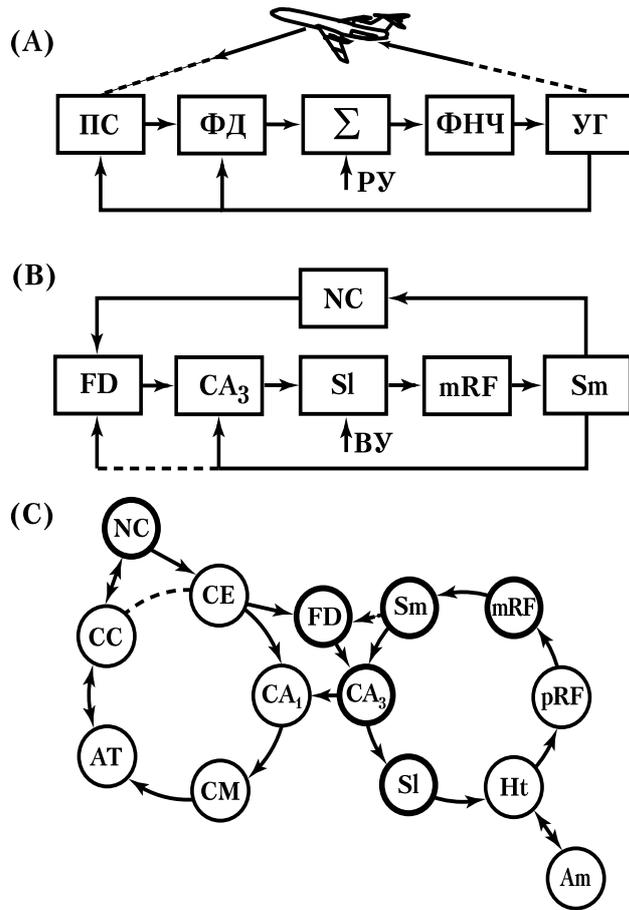
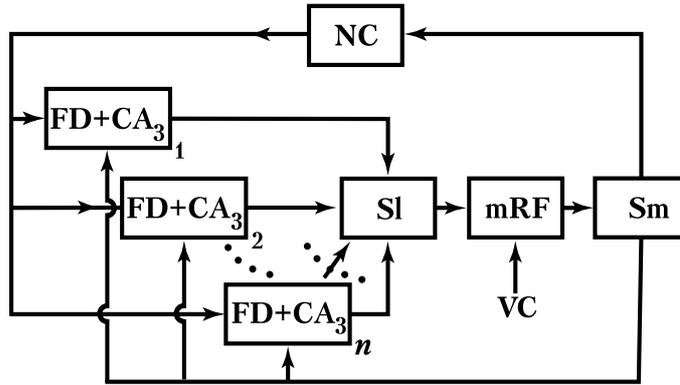


Рис. 7. Блок-схема радарной следящей системы (А), «нейро-локатора» (В) и лимбической системы мозга (С)
 ПС — приемник-смеситель; ФД — фазовый детектор или компаратор; Σ — сумматор; ФНЧ — фильтр низких частот; УГ — управляемый генератор; ПУ, ВУ — ручное и волевое управление, соответственно; NC — неокортекс; FD — зубчатая фасция; CA₁, CA₃ — поля гиппокампа; Sm, Sl — медиальное и латеральное ядро септума; mRF и pRF — мезенцефалическая и пентобулбарная часть ретикулярной формации (Виноградова, 1975; см. [1]).

ный импульс, посылает его во все корковые структуры мозга и после некоторой задержки получает оттуда ответ на свой запрос, наподобие радарного импульса, отраженного от цели. Ответный импульс, с помощью упомянутого компаратора, сравнивается с эндогенно-генерируемым опорным импульсом, и при совпадении их по частоте и фазе вырабатывается сигнал ошибки, уменьшающий возможное рассогласование по фазе между опорным и ответным импульсом. После нескольких таких успешных циклов совпадения на частоте тета-ритма в коре формируется целая *последовательность конфигураций* данного эпизода.

Упрощенная блок-схема обычной радарной следящей системы изображена на рис. 7А (с. 97), а ее нейронная реализация — на рис. 7В. При этом мы опираемся на хорошо установленные функции различных частей лимбической системы, описанные в книге [1] (Виноградова, 1975). Гиппокампальная зубчатая фасция (FD) — это смеситель и интегратор специфических входов; медиальный септум (Sm) — это синхронизатор не только для поля CA₃ гиппокамп, но и для многих других структур мозга; латеральный септум (Sl) — это выходной смеситель полей CA₃ отдельных ламелл, т. е. параллельных структурно-функциональных независимых секций гиппокампальной формации [15] (Amaral & Witter, 1989), соответствующих различным независимым группам кортикальных осцилляторов. Все эти структуры соединены между собой, образуя замкнутый контур, как показано на рис. 7С. Фактически весь комплекс лимбической структуры, частями которого являются выделенные выше структуры, организован в виде двух замкнутых контуров. Первый контур — информационный — включает в себя гиппокампальные поля CA₃ и CA₁, передний таламус, неокортекс, зубчатую фасцию и другие структуры, сохраняющие, по крайней мере частично, сигнальную специфичность. Второй контур — не специфическая регуляторная петля, включающая в себя ретикулярную формацию, гипоталамус, амигдалу и другие структуры, обеспечивающие неспецифическую активацию мозга или аразул. Заметим, что поскольку гиппокамп тормозит активирующую ретикулярную формацию (это хорошо установленный факт), вторая петля является фактически петлей отрицательной обратной связи как для регулирования аразула в мозге, так и для регулирования частоты септального осциллятора. Этот факт позднее будет использован для объяснения «сопряженного» торможения всех сигналов, не попавших в фокус внимания. Обе



$$\frac{d\varphi_i}{dt} = \Lambda_{oi} - \left[\sum_{j=1}^n A_{oj} g(\varphi_j) + N_j(t) \right] F(p), \quad (i=1, \dots, n)$$

$$\frac{d\varphi}{dt} = \Lambda_0 - [A g(\varphi) + N(t)] F(p), \quad (n=1)$$

Рис. 8. Блок-схема и уравнение «Нейролокатора» с учетом ламеллярности гиппокампа

NC — неокортекс; FD — зубчатая фасция; CA₃ — поле гиппокампа; Sm, SI — медиальное и латеральное ядро септума; mRF — мезенцефалическая часть ретикулярной формации; φ_i — разность фаз колебаний септального и i -го кортикального осцилляторов, Λ_{oi} — их частотная расстройка, A_{oi} — сила их связи; $g(\varphi)$ — нелинейная функция фазового дискриминатора; $N_i(t)$ — «белый» шум i -й ламеллы; $F(p)$ — передаточная функция блока усиления mRF.

петли имеют решающее значение для нашей модели, и они обе упрощенно изображены на рис. 7В. Детальное обсуждение всех остальных структур лимбической системы нам пока не потребуется, и мы представим только минимальный вариант системной модели, чтобы раскрыть главную функцию кортико-септо-гиппокампальной системы, которая, как мы полагаем, содержит основной секрет живого мозга.

ТАБЛИЦА 4

Основные свойства «Нейролокатора»

1. Внимание невозможно как при низком араузале ($AK < \Lambda_0$), так и при слишком высоком араузале ($AK \gg \Lambda_0$).
2. Внимание возникает скачком при возрастании араузала ($AK > \Lambda_0$) и проявляется как «выбор» одного из множества внешних сигналов того сигнала, который в данный момент и «захватывает» управление «Нейролокатора».
3. Внимание плавно контролируется изменением расстройки (Λ_0) как автоматически, так и волевым усилием.
4. Внимание унитарно, но при сравнительно невысоком араузале AK оно делимо по правилу Миллера (7 ± 2) или по правилу Коуэна (4 ± 0).
5. Срыв внимания может быть внезапным (в результате действия шумов или синхронизации от другого кортикального осциллятора) или градуальным. В последнем случае наблюдается интересный нелинейный эффект «проскальзывания цикла».

1. Системная модель «Нейролокатор» включает в себя шесть главных подсистем (см. рис. 7В): септальный осциллятор (S_m), M независимых кортикальных осцилляторов (NC), фазовый детектор или компаратор (CA_3), низкочастотный фильтр (mRF), входной смеситель (FD) и выходной смеситель (SI).
2. Каждая подсистема является небольшой модификацией базовой нейронной модели, описанной в разделе «Фазовые переходы в мозге и принцип доминанты» (см. с. 78–87), у которой некоторая часть возбуждательных нейронов получает входы от предшествующей подсистемы и у которой тормозной нейрон действует как выход подсистемы. Пороги тормозных нейронов предполагаются

**Основные психофизические данные, объясняемые
«Нейролокатором» (Крюков, 1991; см. [48])**

- высокая стабильность системной доминанты, несмотря на метастабильность и неустойчивость элементов системы;
- фокусированное, разделенное и регулируемое внимание;
- закон Еркеса–Додсона об оптимальном аразале;
- интеграция или мультимодальный синтез интегрального образа;
- струп-эффект или эффект внутри и межмодальной интерференции;
- одномоментное обучение;
- полифункциональность гиппокампа;
- привыкание и функция долговременной потенциации в гиппокампе;
- генерация временных последовательностей конфигураций.

фиксированными, однако пороги возбудительных нейронов изменяются у всех нейронов одновременно на величину, пропорциональную текущей неспецифической активации (аразал).

3. Все кортикальные осцилляторы имеют сенсорные входы определенной модальности, являясь, таким образом, анализатором свойств стимула (высоты звука, ориентации, яркости и т. д.). Их фиксированные пороги распределены таким образом, что средние частоты осцилляторов распределены равномерно, в альфа-тета частотном диапазоне.

ТАБЛИЦА 6

Новые данные, существенно подтверждающие системную модель внимания «Нейролокатор»

- гиппокамп критически необходим для интеграции множества входных сигналов (Cohen et al., 1999 – см. [24]; Verfaellie & Keane, 1997 – см. [72]);
- подтверждено существование «сопряженного» торможения доминанты: внимание к одному месту приводит к подавлению активности во всех остальных местах (Smith et al., 2000 – см. [63]);
- тэта-ритм, трудно наблюдаемый у человека в обычных условиях, появляется при выполнении когнитивных тестов (Kahana et al., 1999 – см. [42]; Sarnthein et al., 1998 – см. [59]);
- память находится не в гиппокампе, а в новой коре; гиппокамп играет ключевую роль в организации памяти человека (Parkin, 1996 – см. [57]; Tulving & Markowitsch, 1997 – см. [69]; Ungerleider, 1995 – см. [70]; Mesulam, 1998 – см. [52]);
- гиппокамп всегда необходим как для записи и консолидации, так и для воспроизведения разных видов памяти: пространственной, декларативной и эпизодической (Riedel et al., 1999 – см. [58]; Nadel & Moscovitch, 1998 – см. [54]);
- градиент ретроградной амнезии (Parkin, 1996 – см. [57]) свидетельствует об участии гиппокампа не только при регистрации, но и при считывании сенсорной информации.

4. Фазовый детектор поля CA_3 имеет два главных входа — септальный и неокортикальный. Септальный вход, по предположению, имеет фиксированную амплитуду, т. е. меру выраженности пачек импульсов; неокортикальный вход имеет переменную амплитуду $A = (\sum A_i)/M$, где A_i — мера пачечности в i -ом ламеллярном входе, которая зависит от араузала.
5. Низкочастотный фильтр mRF характеризуется передаточной функцией $F(p) = K/(1 + Tp)$. (Так как K действует на систему наподобие A , то мы их объединяем, считая AK одним параметром, измеряющим араузал).
6. Главное предположение модели состоит в том, что внимание может быть описано операционно как глобальное свойство системы, блок-схема которой представлена на рис. 7В, и это свойство соответствует фазовой синхронизации многих первоначально независимых кортикальных осцилляторов на доминантной частоте.
7. Когда число синхронизованных кортикальных осцилляторов превысит критическое значение, оптимальные лабильности осцилляторов уравниваются, т. е. осцилляторы усваивают общий ритм и соответственно меняют собственную частоту.

Как было показано [44] (Kirillov et al., 1989), все указанные выше функции могут быть реализованы с помощью одной и той же базовой нейронной модели, но при различных значениях параметров. В работе [47] (Kryukov et al., 1990) показано, что система описывается интегродифференциальным уравнением, представленным на рис. 8.

Эти уравнения почти аналогичны системе уравнений стандартной фазовой следящей системы, обобщенной на случай нескольких синхронизованных осцилляторов [7] (Линдсей, 1978, с. 131). На этом основании мы в приводим основные свойства «Нейролокатора» как стандартной фазовой следящей системы 1-го порядка, т. е. для случая, когда $F(p) = K$ — см. табл. 4 на с. 100.

В работе [48] (Kryukov, 1991) показано, что «Нейролокатор» способен объяснить большой набор классических психологических эффектов, перечисленных в табл. 5 на с. 101.

ТАБЛИЦА 7

Основное предсказание «Нейролокатора» (1991)

В ЦНС существует фазо-частотная следящая система по меньшей мере второго порядка, управляющая вниманием и памятью человека и функционирующая в диапазоне частот альфа и тета-ритмов. В этой системе септо-гиппокампальная формация играет ведущую роль задающего осциллятора и фазового дискриминатора. Эта система, взаимодействуя с другими корковыми и подкорковыми структурами (такими, например, как mRF), выполняет функции Центрального Управителя.

За последнее время появилось огромное количество новых данных, которые можно рассматривать как подтверждение модели «Нейролокатора». Приведем лишь некоторые из них (табл. 6 на с. 102). Более полный перечень новых подтверждений а также новых данных, необъяснимых с помощью существующих моделей внимания, но подходящих для «Нейролокатора», мог бы стать предметом отдельной работы.

Все указанные выше новые данные свидетельствуют о справедливости основного предсказания модели «Нейролокатор» (табл. 7), сформулированного в работе (Круиков, 1991; см. [48]).

Заключительные замечания

Мы утверждаем, что существует совместное одновременное решение всех шести проблем моделирования внимания, описанных во Введении. И это решение дается моделью доминанты А. А. Ухтомского. Все, что нам осталось сделать, это описать физический механизм, лежащий в основе доминанты и выразить его в терминах текущих психологических и нейробиологических исследовательских работ. Сделав это по отношению к каждой проблеме, упомянутой выше, мы резюмируем наши результаты следующим образом.

I. Проблема селективности стимулов. «Нейролокатор» выбирает группы кортикальных осцилляторов на основе способности этих осцилляторов усваивать единый ритм и единую систему управления с помощью частотно фазовой синхронизации на частоте тета-ритма. Сам процесс выбора происходит при специальных условиях, подобных условиям физического фазового перехода. В частности, необходима малая взаимная частотная расстройка между субкортикальными и кортикальными осцилляторами и достаточно сильные двусторонние связи между ними.

II. Проблема долговременной памяти. Функциональным элементом нашей модели внимания является не одиночный нейрон, а небольшой ансамбль, локальная группа нейронов (субмодуль), который способен произвести кооперативные действия, кооперативный эффект локально-стационарной активности («эффект сохранения пятна»). Под действием внешней стимуляции и неспецифической активации от ретикулярной формации эта локальная активность переходит в колебательную, распространяющуюся по сети активность. Знания импульсной активности и схемы синаптических соединений нейронов недостаточно в нашей модели, чтобы объяснить этот переход, потому что градуальные и «шумовые» компоненты мембранного потенциала существенны для понимания происходящих процессов. В статье предполагается, что этот переход является механизмом фактического перехода кратковременной памяти в долговременную и обратно. Долговременная память — это свойство нервной ткани поддерживать почти неизменной свою оптимальную лабильность, установленную предшествующей доминантой.

III. Проблема интеграции локальной активности в единый цельный образ решается в нашей модели путем фазочастотной синхронизации большого числа удаленных друг от друга модулей с помощью или локальных связей, или с помощью Центрального Управителя, или с помощью обоих этих механизмов. Возникающее при этом частотно-фазовое кодирование в системе инвариантно к действию большой группы преобразований внешних сигналов, обеспечивая, таким образом, инвариантное распознавание объектов. Поскольку при таком синтезе интегрального образа нет необходимости устанавливать постоянные связи между коактивированными модулями, то и исчезают коннекционистские проблемы, такие, как комбинаторный взрыв, экспоненциальный рост числа связей при увеличении числа входов и другие проблемы. В частности, динамическое

соединение элементов эквивалентно в коннекционистских терминах введению в архитектуру обработки «скрытого слоя» произвольно высокого порядка. Эта же метафора может помочь сторонникам коннекционизма уяснить себе необычные ассоциативные функции гиппокампальной формации [65] (Sutherland, & Rudy, 1989). Однако следует подчеркнуть, что метафора «скрытого слоя» дает только частичное объяснение очень сложной функции гиппокампа, подробно описанной в работе [47] (Kryukov, et al., 1990).

IV. Проблема инерции решается в модели автоматическим возникновением критического режима во время активного восприятия внешних стимулов. Отдельные модули и вся система внимания в целом работают вблизи точки неустойчивого равновесия, которое является ближайшим аналогом физического эффекта метастабильности. Он сопровождается так называемым критическим замедлением системной динамики, что и объясняет временной зазор между психологическим и нейробиологическим масштабами времени. Долгоживущие кортико-гиппокампальные реверберации обеспечивают, в частности, состояния, необходимые для решения многих жизненных проблем «на линии». Достаточно продолжительное критическое состояние является необходимым для фиксации информации в процессе консолидации путем отключения соответствующих осцилляторов от входных воздействий и стимулов. Однако энграммы запоминаемой информации нигде не накапливаются, а распределены по всей системе в форме локальных возбуждений. Они могут быть восстановлены в целостный образ, путем преобразования их в распространяющиеся колебания, после того, как принимается стимул, напоминающий первоначальный. Такая память практически безгранична по емкости, поскольку она хранится в неактивном состоянии локального возбуждения, а каждый нейрон может быть активно вовлечен во множество различных доминантных конфигураций.

V. Проблема торможения и подавления помех. Кроме торможения за счет десинхронизации, в модели «Нейролокатор» автоматически действует «сопряженное» торможение и торможение по типу «торможения возвращения» [45] (Klein, 2000).

VI. Проблема Центрального Управителя. Функции, которые способен выполнять Центральный Управитель, очень обширны: наряду с перечнем, приведенным в работе [18] (Baddeley, 1996) сюда относятся та-

кие функции: интеграция, выделение «новизны», селекция объектов, различного вида память (декларативная, эпизодическая, семантическая, рабочая, кратковременная, долговременная, одномоментная), кодирование, считывание, консолидация, генерация последовательностей. Некоторые из этих функций недавно были промоделированы пушинской группой математиков [20] (Borisuyk et al., 2001). Нейролокатор решает проблему моделирования этих функций *все одновременно и совместно*.

В заключение приведем замечание о современном состоянии теории мозга. В 1979 году Ф. Крик [29] (Crick, 1979) предсказывал: «... если прорыв в изучении мозга действительно произойдет, он, вероятно, будет на уровне всеобщего управления системой. Если бы система была бы такой хаотической, какой она иногда представляется, она была бы неспособной выполнять удовлетворительно даже простейшие задания. Возможным хотя и маловероятным примером будет открытие, что мозговая обработка продвигается регулированием фазы некоторого рода периодическим устройством, наподобие компьютера, и оно, вероятно, составило бы главный прорыв».

В 1995 году фон дер Малсбург уже более определенно предсказывает [51] (Malsburg, 1995): «Мы в середине научной революции, результатом которой будет установление интеграции как фундаментального аспекта нейронного кода наряду с классическим частотным кодом. В случае удачи, эта революция приведет к плодотворным парадигмам, превосходящим текущую сосредоточенность на одиночную клетку к более внутренне согласованным концепциям функций мозга и восприятия и к решению многих остающихся проблем в терминах явных моделей».

В этом предсказании уже яснее проявляются черты будущего прорыва как общей закономерности науки, по замеченной еще Томасом Куном³ неизбежности научных революций для преодоления застойных явлений и отживающих парадигм.

Примерно в это же время другой крупный исследователь мозга С. Роуз в своей книге [9] (Роуз, 1995) пишет: «Согласно концепции Имре Лакатоша, поступательное движение науки осуществляется путем «исследовательских программ». Программа может быть прогрессивной, и в этом случае она реализуется в постановке задач, которые могут быть решены в

³См. Кун Т. Структура научных революций. 2-е изд.: Пер. с англ. – М.: Прогресс, 1977. – 300 с. (Серия «Логика и методология науки». – Прим. ред.)

ее рамках, оставаясь плодотворной в отношении познания мира, или она может быть дегенеративной, тогда она будет порождать все более выраженные аномалии. В этом смысле ассоциационистская программа Хебба прогрессивна, поскольку может включить в кэнделовский «клеточный алфавит памяти» и *долговременную потенцию* в качестве модельного механизма, тогда как эксперименты Гарсия и мои эксперименты с повреждением мозга — это такие аномалии, которые могут указывать на дегенеративный характер хеббовской программы; иными словами, нам нужна революция, если использовать термин Куна». Заметим, что упомянутая здесь *долговременная потенция*, согласно обзора большого числа экспериментальных работ [62] (Shors & Madzel, 1997), признана непригодной в качестве модельного механизма обучения.

Кто же из них прав? В середине ли мы научной революции, или она еще не началась? Судя по тому, что теоретики мозга, в отличие от экспериментаторов, почти нигде не пишут о дегенеративном характере хеббовской программы, революция еще не начиналась. Почти все теоретики мозга ищут не истину, а подтверждения хеббовской программы, приняв гипотезу за незыблемый факт. А истина лежит совсем в другом месте — в учении А. А. Ухтомского о доминанте.

Литература

1. *Виноградова О. С.* Гиппокамп и память. М.: Наука, 1975.
2. *Голиков Н. В.* Проблема местного и распространяющегося возбуждения в современной нейрофизиологии // В сб.: Механизмы местной реакции и распространяющегося возбуждения. — Л.: Наука, 1970.
3. *Добрушин Р. Л.* Марковские процессы с большим числом локально взаимодействующих компонент — существование предельного процесса и его эргодичность // Пробл. передачи информ. — 1971. — т. 7, вып. 2. — с. 70–87.
4. *Добрушин Р. Л.* Марковские процессы с большим числом локально взаимодействующих компонент — обратимый случай и некоторые обобщения // Пробл. передачи информ. — 1971. — т. 7, вып. 3. — с. 57–66.
5. *Ливанов М. Н.* Нейронные механизмы памяти // Успехи физиол. наук. — 1975. — т. 6. — с. 66–89.
6. *Крюков В. И., Борисюк Г. Н., Борисюк Р. М., Кириллов А. Б., Коваленко Е. И.* Метастабильные и неустойчивые состояния в мозге. — Пущино, 1986.
7. *Линдсей В.* Системы синхронизации в связи и управлении: Пер. с англ. — М., 1978.

8. *Мухина П. С.* О механизме замыкания связи в свете теории Введенского-Ухтомского // В сб.: Механизмы нервной деятельности. – Л.: Изд-во ЛГУ, 1977. – с. 112–126.
9. *Роуз С.* Устройство памяти. От молекул к сознанию: Пер. с англ. – М.: Мир, 1995.
10. *Симонов П. В.* Эмоциональный мозг. – М.: Наука, 1981.
11. *Ставская О. Н., Пятецкий-Шапиро И. И.* Об однородных сетях из спонтанно-активных элементов // Пробл. кибернетики. – Вып. 20, 1968. – с. 91–106.
12. *Ухтомский А. А.* Доминанта. – М.: Наука, 1966.
13. *Ухтомский А. А.* Избранные труды. М.: Наука, 1978.
14. *Ухтомский А. А.* Интуиция совести. – СПб.: 1996.
15. *Amaral D. G., Witter M. P.* The three-dimensional organization of the hippocampal formation: A review of anatomical data // *Neuroscience*. – 1989. – 31(3). – pp. 571–591.
16. *Alvarez P., Squire L. R.* Memory consolidation and the medial temporal lobe: A simple network model // *Proc. Nat. Acad. Sci. USA*. – 1994. – v. 91. – pp. 7041–7045.
URL: <http://www.pnas.org/cgi/reprint/91/15/7041>
17. *Baddeley A. D., Hitch G. J.* Working memory // In *G. Bower*, Ed.: Recent advances in learning and motivation, Vol. 8. – New York: Academic Press, 1974. – pp. 47–90.
18. *Baddeley A. D.* Exploring the central executive // *The Quarterly J. of Experimental Psychology*. – 1996. – v. 49A. – pp. 5–28.
19. *Baddeley A., Della Sala S.* Working memory and executive control // *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* – 1996. 29;351(1346). – pp. 1397–1403.
20. *Borisyuk R. M. et al.*, 2001.
URL: <http://www.tech.plym.ac.uk/soc/staff/roman/person.htm>
21. *Brashers-Krug T., Shadmehr R., Bizzi E.* Consolidation in human motor memory // *Nature*. – 1996. – 382(6588). – pp. 252–255.
22. *Braver T. S., Cohen J. D.* Working memory, cognitive control, and the prefrontal cortex: Computational and empirical studies // *Cognitive Processing*. – 2001. – v. 2. – pp. 25–55.
23. *Britten K. H.* Cortical neurophysiology. Attention is everywhere // *Nature*. – 1996. – 382. – pp. 497–498.

24. *Cohen N.J., Ryan J., Hunt C., Romine L., Wszalek T., Nash C.* Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies // *Hippocampus*. – 1999. – 9(1). – pp. 83–98.
25. *Colin F., Manil J., Desclin J.C.* The olivocerebellar system. I. Delayed and slow inhibitory effects: an overlooked salient feature of cerebellar climbing fibers // *Brain Res.* – 1980. – v. 187. – pp. 3–27.
26. *Cowan N.* Evolving conceptions of memory, storage, selective attention and their mutual constraints within human information processing systems // *Psych. Rev.* – 1988. – v. 104, pp. 163–191.
27. *Cowan N.* Attention and memory: An integrated framework // *Oxford Psychology Series, No. 26.* New York: Oxford University Press. – 1995. (Paperback edition: 1997)
28. *Craik K.J.W.* Theory of the human operator in control. I. The operator as an engineering system. II. Man as an element in control system // *Brit. J. Psychol., Gen. Sect.* – 1948. – v. 38. – pp. 56–61, 142–148.
29. *Crick F.* Thinking about the brain // *Sci. Am.* – 1979. 241. – pp. 181–189.
30. *Danckert J., Goodale M.A.* Blindsight: A conscious route to unconscious vision // *Current Biology*. – 2000. v. 10. – pp. R64–R67.
URL: <http://download.bmn.com/pdf/journals/list/>
31. *Dehaene S., Kerszberg M., Changeux J.-P.* A neuronal model of a global workspace in effortful cognitive tasks // *Proc. Nat. Acad. Sci. USA.* – 1998. – v. 95, – pp. 14529–14534.
URL: <http://www.pnas.org/cgi/reprint/95/24/14529>
32. *Ebner T.J., Bloedel J.R.* Temporal patterning in simple spike discharge of Purkinje cells and its relationship to climbing fiber activity // *J. Neurophysiol.* – 1981. – v. 45. – pp. 933–947.
33. *Eichenbaum H., Schoenbaum G., Young B., Bunsey M.* Functional organization of the hippocampal memory system // *Proc. Nat. Acad. Sci. USA.* – 1996. – v. 93(24). – pp. 13500–13507.
URL: <http://www.pnas.org/cgi/reprint/93/24/13500>
34. *Fuster J.M.* Cellular dynamics of network memory // *Z. Naturforsch [C]*. 1998. – v. 53. – pp. 670–676.
35. *Fuster J.M.* Distributed memory for both short and long term // *Neurobiol. Learn. Mem.* – 1998. – v. 70(1/2). – pp. 268–274.
36. *Fuster J.M.* The prefrontal cortex — an update: time is of the essence // *Neuron*. – 2001. – v. 30. – pp. 319–333.

37. *Goldman-Rakic P.S.* Regional and cellular fractionation of working memory // *Proc. Nat. Acad. Sci. USA.* – 1996. – v. 93, pp. 13473–13480.
URL: <http://www.pnas.org/cgi/reprint/93/24/13473>
38. *Grasby P.M., Frith C.D., Friston K., Frackowiak R.S., Dolan R.J.* Activation of the human hippocampal formation during auditory-verbal long-term memory function // *Neurosci. Lett.* – 1993. – v. 163(2). – pp. 1851–1888.
39. *Haken H.* Synopsis and introduction // In: *Synergetics of the Brain / (E. Basar, H. Flohr, H. Haken, A. J. Mandell, Eds.)*. – N.Y.: Springer, 1983. – pp. 3–25.
40. *Haken H.* Principles of brain functioning: A synergetic approach to brain activity, behavior and cognition. Berlin; New York: Springer, 1996.
41. *John E.R.* Mechanisms of memory. – New York: Academic Press, 1967.
42. *Kahana M.J.* Human theta oscillations exhibit task dependence during virtual maze navigation // *Nature.* – Jun. 1999. – 24; 399(6738). – pp. 781–784.
43. *Kentridge R.W., Heywood C.A., Weiskrantz L.* Attention without awareness in blindsight // *Proc. R. Soc. Lond. B. Biol. Sci.* – Sep. 1999. 7; 266(1430). – pp. 1805–1811.
44. *Kirillov A.B., Borisyuk G.N., Borisyuk R.M., Kovalenko Ye.I., Makarenko V.I., Chulaevsky V.A., Kryukov V.I.* A model of neural oscillator for a unified submodule // In: *Advance in Neural Information Processing Systems I / D.S.Touretzky (Ed.)*, San Mateo, CA: Morgan Kauffman, 1989. – pp. 560–567.
45. *Klein R.M.* Inhibition of return // *Trends in Cognitive Science.* – 2000. – v. 4. – pp. 138–147.
46. *Koch C., Ullman S.* Shifts in selective visual attention, 1985, 4, pp. 219–227.
47. *Kryukov V.I., Borisyuk G.N., Borisyuk R.M., Kirillov A.B., Kovalenko Ye.I.* Metastable and unstable states in the brain // In: *Stochastic Cellular Systems: Ergodicity, Memory, Morphogenesis / Eds.: R.L. Dobrushin, V.I. Kryukov and A.L. Toom*, Manchester, UK and New York, NY: Manchester Univ. Press. – 1990. – pp. 226–357.
48. *Kryukov V.I.* An attention model based on the principle of dominant // In: *Neurocomputers and Attention. I: Neurobiology, Synchronization and Chaos / Eds.: A.V. Holden, V.I. Kryukov.* – Manchester University Press, 1991. – pp. 319–351.
49. *Little W.A.* The existence of persistent states in the brain // *Math. Biosci.* – 1974. – v. 19. – pp. 101–120.
50. *Lu Z.-L., Williamson S.J., Kaufman L.* Physiological measures predict behavioral lifetime of human auditory sensory memory // *Society for Neuroscience Abstracts* 18, 1992. – pp. 1212.

51. *von der Malsburg C.* Binding in models of perception and brain function // *Current Opinion in Neurobiology.* – 1995. – v. 5. – pp. 520–526.
52. *Mesulam M.-M.* From sensation to cognition // *Brain.* – 1998. – v. 121. – pp. 1013–1052.
53. *Miller R.* Cortico-hippocampal interplay: Selforganizing phase-locked loops for indexing memory // *Psychobiol.* – 1989. – v.17. – pp. 115–128.
54. *Nadel L., Moscovitch M.* Hippocampal contributions to cortical plasticity // *Neuropharmacology.* – 1998. – v. 37(4–5). – pp. 431–439.
55. *Newman J., Grace A.A.* Binding across time: The selective gating of frontal and hippocampal systems modulating working memory and attentional states // *Consciousness and Cognition.* – 1999. – v. 8. – pp. 196–212.
56. *Paller K.A.* Consolidating dispersed neocortical memories: The missing link in amnesia // *Memory.* – 1997. – v. 5(1–2). – pp. 73–88.
57. *Parkin A.J.* Human memory: The hippocampus is the key // *Curr. Biol.* Dec. 1996. v. 1; 6(12). – pp. 1583–1585.
URL: <http://download.bmn.com/pdf/journals/list/>
58. *Riedel G., Micheau J., Lam A.G., Roloff E., Martin S.J., Bridge H., Hoz L., Poeschel B., McCulloch J., Morris R.G.* Reversible neural inactivation reveals hippocampal participation in several memory processes // *Nat. Neurosci.* Oct. 1999. v. 2(10). – pp. 898–905.
59. *Sarnthein J.* Synchronization between prefrontal and posterior association cortex during human working memory // *Proc. Nat. Acad. Sci. USA.* Jun. 1998. – v. 9; 95(12). – pp. 7092–7096.
URL: <http://www.pnas.org/cgi/reprint/95/12/7092>
60. *Schillen T.B., Konig P.* Binding by temporal structure in multiple feature domains of an oscillatory neuronal network // *Biol. Cybern.* – 1994. – v. 70(5). – pp. 397–405.
61. *Sejnowsky T.J.* Open questions about computation in cerebral cortex // In: *Parallel Distributed Processing. Exploration in Microstructure of Cognition, Vol. 2: Psychological and Biological Models / McClelland J. L., Rumelhart D. E. & PDP Research Group.* – MIT Press. – 1986. – pp. 372–389.
62. *Shors T.J., Matzel L.D.* Long-term potentiation: What's learning got to do with it? // *Behavioral and Brain Sciences.* – 1997. – v. 20(4). – pp. 597–655.
URL: <http://www.cogsci.soton.ac.uk/bbs/Archive/bbs.shors.html>
63. *Smith A. T., Singh K. D., Greenlee M. W.* Attentional suppression of activity in the human visual cortex // *Neuroreport.* – Feb. 2000. – v. 7; 11(2). – pp. 271–277.

64. *Strong G. W., Whitehead B. A.* A solution to the tag-assignment problem for neural networks // Behavioral and Brain Sciences. – 1989. – v. 12. – pp. 381–433.
65. *Sutherland R. J., Rudy J. W.* Configural association theory: The role of the hippocampal formation in learning, memory and amnesia // Psychobiol. – 1989. v. 17. – pp. 129–144.
66. *Tesche C. D., Karhu J.* Theta oscillations index human hippocampal activation during a working memory task // Proc. Nat. Acad. Sci. USA. – 2000. v. 97. – pp. 919–924. URL: <http://www.pnas.org/cgi/reprint/97/2/919>
67. *Treisman A., Gelade G.* A feature integration theory of attention // Cognitive Psychology. – 1980. – v. 12. – pp. 97–136.
68. *Treisman A.* The binding problem // Current Opinion in Neurobiology. – 1996. – v. 6. – pp. 171–178.
69. *Tulving E., Markowitsch H. J.* Memory beyond the hippocampus // Current Opinion in Neurobiology. – 1997. – v. 7. – pp. 209–216.
70. *Ungerleider L. G.* Functional brain imaging studies of cortical mechanisms for memory // Science. – 1995. Nov. 3. – No. 270(5237). – pp. 769–775.
71. *Uusitalo M. A., Williamson S. J., Seppa M. T.* Dynamical organisation of the human visual system revealed by lifetimes of activation traces // Neurosci Lett. – 1996, Aug. 9. – v. 213(3). – pp. 149–152.
72. *Verfaellie M., Keane M. M.* The neural basis of aware and unaware forms of memory // Semin. Neurol. – Jun. 1997. v. 17(2). – pp. 153–161.
73. *Vinogradova O. S.* Discussion // In: Function of Septo-Hippocampal System. – Amsterdam: Elsevier, 1978. – p. 402.
74. *Vinogradova O. S., Brazhnik E. S., Stafekina V. S.* Septo-hippocampal system. Rhythmic oscillations and information selection // In: Neurocomputers and Attention. I: Neurobiology, Synchronization and Chaos / Eds.: A. V. Holden, V. I. Kryukov. – Manchester University Press, 1991.
75. *Vinogradova O. S.* Hippocampus as comparator: Role of the two input and two output systems of the hippocampus in selection and registration of information // Hippocampus. – 2001. v. 11(5). – pp. 578–598.
76. *Zipser D., Kehoe B., Littlewort G., Fuster J.* A spiking network model of short-term active memory // J. Neurosci. – Aug. 1993. – v. 13(8). – pp. 3406–3420.

Игумен Феофан (Крюков), кандидат физико-математических наук, игумен Свято-Данилова монастыря. Область научных интересов — теория функций мозга, взаимодействие науки и религии.

Ю. И. НЕЧАЕВ

Государственный морской технический университет,
Институт высокопроизводительных вычислений и баз данных,
Санкт-Петербург

E-mail: int@fn.csa.ru

**НЕЙРОСЕТЕВЫЕ ТЕХНОЛОГИИ В БОРТОВЫХ
ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ РЕАЛЬНОГО ВРЕМЕНИ**

Аннотация

Обсуждается проблема практического применения искусственных нейронных сетей (ИНС) при создании бортовых интеллектуальных систем (ИС) реального времени. Основное внимание уделено использованию ИНС в трудноформализуемых задачах анализа и интерпретации измерительной информации о поведении динамического объекта (ДО) в различных условиях эксплуатации, в том числе и в экстремальных ситуациях. Материалы статьи отражают опыт автора и его научного коллектива при разработке и испытании бортовых ИС в области судостроения и морской техники.

Yu. NECHAEV

Saint-Petersburg State Marine Technical University,
Institute for High-Performance Computing and Information Systems,
Saint-Petersburg

E-mail: int@fn.csa.ru

**NEURAL NETWORK TECHNOLOGY IN REAL-TIME ON-BOARD
INTELLIGENCE SYSTEMS**

Abstract

A practical usage problem is discussed for artificial neural networks (ANN) to develop on-board real-time intelligent systems (IS). Attention is devoted mainly to ill-formalized problems associated with analysis and interpretation of measurement data for dynamic object induced with various maintenance conditions, including extreme ones. This article reflects author's and his research team experience in the field of development and testing for shipboard and marine engineering intelligent systems.

Введение

Для современного этапа развития науки и техники характерна бурная интеграция идей и методов исследования. На базе научно-технических достижений создается принципиально новая техника, отличающаяся чрезвычайной сложностью, расширением круга решаемых задач и диапазоном эксплуатационного использования. В этих приложениях все чаще находят применение компьютерные технологии, основанные на использовании методов искусственного интеллекта (ИИ) [6]. Одной из центральных задач становится интеграция результатов различных подходов и направлений, построение моделей, которые более гибко и адекватно описывают феномен интеллектуального поведения в неопределенной среде. Эта парадигма ИИ опирается на широкую реализацию концепции параллельных вычислений, использование современных суперкомпьютерных платформ, нейрокомпьютеров и процессоров нечеткой логики. Именно здесь просматривается одно из ведущих направлений развития бортовых интеллектуальных систем (ИС), сочетающих фактически данные динамических измерений с широким использованием современного математического аппарата и высокопроизводительных средств вычислений.

Возросшая сложность бортовых ИС связана с необходимостью разработки нового математического обеспечения. Программная реализация используемых алгоритмов должна обеспечивать стратегии принятия решений при анализе и прогнозе ситуаций и гарантировать уровень надежности выдаваемых практических рекомендаций в различных условиях эксплуатации. Перспективой развития бортовых ИС является совершенствование теоретического, методологического и технического обеспечения процесса функционирования плавучего динамического объекта (ДО) в сложных условиях (тяжелый шторм, интенсивное обледенение, плохая видимость, аварийная ситуация).

Внимание специалистов при разработке бортовых ИС все чаще привлекает возможность использования искусственных нейронных сетей (ИНС) [2, 3, 6, 16, 17, 32]. Использование нейросетевых вычислительных структур тесно связано с созданием компьютеров нового поколения на основе теории ИНС. Это открывает возможности решения широкого круга задач, связанных с применением новых подходов обработки информации в режиме реального времени. Среди задач, эффективно решаемых с помощью ИНС в бортовых ИС, следует выделить реализацию нейросе-

тевых принципов управления, идентификацию экстремальных ситуаций, контроль и диагностику измерительных систем, оценку параметров ДО и внешней среды и т. д. Ниже дается характеристика ряда наиболее важных практических задач, исследованием которых в последние годы занимается представляемый мною научный коллектив.

Особенности динамики сложных нелинейных систем

Одной из важных задач при разработке бортовых ИС со сложной динамикой является анализ поведения исследуемого ДО в различных условиях эксплуатации, в том числе и в экстремальных ситуациях. Математические модели, описывающие поведение ДО, обладают специфическими свойствами — единой структурой и общим признаком нелинейности. Слабая нелинейность — достаточно хорошо изученная область математики. Здесь существуют общие методы решения и разработаны эффективные алгоритмы для практических приложений. Методы решения задач с более сильными проявлениями нелинейности основаны на сложных математических построениях и использовании современных вычислительных средств. Анализ нелинейностей позволяет вскрыть особенности задачи, найти, где заложены возможности нетривиальных решений, построить общую теорию, описывающую поведение нелинейных систем.

Углубленное исследование особенностей динамики сложных нелинейных систем требует осмысления физических закономерностей, эффектов и явлений. Эти особенности не всегда укладываются в устоявшиеся понятия и концепции и требуют развития новых подходов. Среди них следует выделить геометрическую интерпретацию динамических моделей на основе теории хаотических систем и принципов самоорганизации. В рамках такой интерпретации удастся получить качественную и количественную информацию о динамических явлениях, связанных с поведением системы при различном характере внешних возмущений. Использование указанных методов предоставляет новый эффективный аппарат для исследования физических систем со сложной структурой с целью организации эффективного управления или создания оптимальных динамических систем.

Применительно к плавучим объектам, их динамика может быть описана системой нелинейных дифференциальных уравнений:

$$F_i(\ddot{x}_i, \dot{x}_i, x_i, X_{i1}, \dots, X_{im}, Y_{i1}, \dots, Y_{in}, t) = 0, \quad (1)$$

где $F_i(\bullet)$ — нелинейные функции; x_i — линейные и угловые перемещения; X_{i1}, \dots, X_{im} — параметры, характеризующие судно как динамическую систему (инерционные, демпфирующие и восстанавливающие компоненты); Y_{i1}, \dots, Y_{in} — возмущающие силы и моменты; t — время; $i = 1, 2, \dots, 6$.

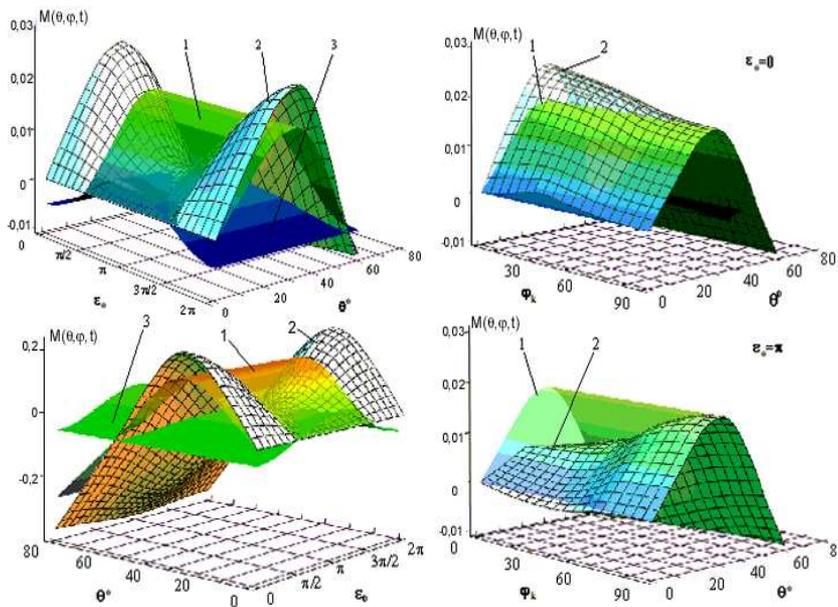


Рис. 1. Нелинейная пространственная функция, характеризующая восстанавливающую компоненту

Наиболее сложной функцией в системе (1) является восстанавливающая компонента, входящая в дифференциальное уравнение вращения ДО относительно продольной центральной оси. Восстанавливающая компонента характеризуется существенной нелинейностью, сложностью и

многозначностью. Непрерывно изменяясь во времени и пространстве, эта функция в значительной степени определяет физические картины взаимодействия ДО с внешней средой. Для общего представления о виде рассматриваемой нелинейной функции в задаче о динамическом наклоне плавучего ДО под воздействием внешних возмущений, на рис. 1 приведена реализация поверхности, характеризующей восстанавливающую компоненту при различных значениях определяющих параметров. Математическое описание этой поверхности представляется в виде полинома, зависящего от параметров морского волнения, элементов ДО, скорости его движения и ориентации относительно направления бега волн:

$$M(\theta, \varphi, t) = DB \cdot \Omega(\lambda/L, \theta, \varphi)(\Phi(h_w/\lambda, \theta, \varphi) + \sum_m A_m f_m(\theta, \varphi) + \sum_n B_n F_n(\theta, \varphi) + \sum_p C_p E_p(\theta, \varphi)) \quad (2)$$

Здесь $\Phi(\bullet)$ и $\Omega(\bullet)$ — нелинейные функции, учитывающие влияние параметров волнения (относительной длины λ/L и высоты h_w/λ волны); $f(\bullet)$, $F(\bullet)$, $E(\bullet)$ — нелинейные функции, характеризующие влияние элементов формы корпуса ДО и относительной скорости (числа Фруда) при различных значениях угла крена θ , курсового угла φ и фазы ε , фиксирующей положение судна относительно волны; D, B — водоизмещение и ширина ДО.

Наличие качки большой амплитуды плавучего ДО при интенсивных динамических нагрузках включает в действие новую физическую связь, «дремавшую» в случае слабых возмущений. В результате происходит резкое изменение картины взаимодействия ДО с внешней средой. Качественные исследования нелинейных динамических систем при различном уровне внешних возмущений свидетельствуют о сложной эволюции фазовых траекторий, нарушающих симметрию фазового пространства и приводящих к образованию аттракторных множеств. Особенно сложными оказываются случаи, характеризующие динамику бистабильной и трехстабильной систем (рис. 2). На разных стадиях своего развития исследуемые фазовые портреты напоминают поведение самоорганизующихся систем. Процесс самоорганизации заключается в переходе системы из одного стабильного состояния в другое. При этом фазовые траектории как бы «втягиваются» в тот или иной аттрактор. При одних со-

стояниях системы возникают аттракторы типа «предельный цикл», а при других состояниях равновесия оказываются зависимыми от времени. Учет этих особенностей при формализации знаний с использованием встроженных процедур, основанных на стандартных математических методах, представляет собой достаточно сложную задачу анализа и интерпретации результатов математического моделирования.

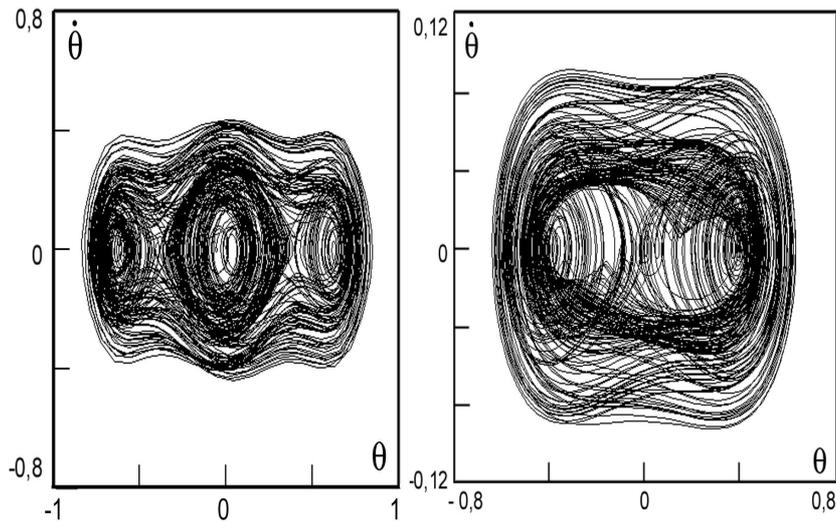


Рис. 2. Фазовые портреты трехстабильной и бистабильной систем

Результаты исследования позволяют выделить различные сценарии образования аттракторных множеств. При определенном уровне внешних возмущений и параметров системы может происходить слияние двух и более состояний равновесия, часть из которых устойчива и часть неустойчива. Этот феномен связан с качественными изменениями в поведении системы, когда система скачком переходит из одного режима движения в другой. Наличие скачка (катастрофы) может иметь и более сложный характер в зависимости от параметров системы. В отдельных случаях возможно исчезновение «предельного цикла» и переход системы либо в устойчивый «предельный цикл», либо в более сложный аттрактор. Вы-

числительные эксперименты с математической моделью плавучих ДО показывают, что наличие сильной периодической составляющей восстанавливающей компоненты приводит к более сложной топологии фазового пространства. При этом частота исходного периодического движения оказывается близкой к доминирующей частоте в спектре аттрактора. Анализ результатов моделирования также указывает на существование общих принципов и закономерностей преобразования информации в системах со сложной динамикой. На этой основе могут быть построены различные математические структуры (одномерные и многомерные отображения), реализующие процессы обработки информации. Важным результатом исследования нелинейной динамики плавучего ДО, является практическая невозможность использования методов статистической линеаризации, широко применяемых при анализе нелинейных систем. Сопоставление топологии фазового пространства для нелинейной и линеаризованной математической модели свидетельствует о практической невозможности сохранения фазовых потоков, порождающих странные аттракторы и детерминированный хаос, в линеаризованных системах.

Использование алгоритмов преобразования информации при оценке характеристик ДО в условиях нормального функционирования связано с решением проблемы вычислительной технологии. Эта проблема связана с анализом некорректно поставленных задач¹. Некорректность приводит к неустойчивым решениям из-за плохой обусловленности информационной матрицы при определении исследуемых характеристик. Для получения устойчивого решения используют методы регуляризации. Улучшение традиционной схемы рекуррентного оценивания параметров достигается за счет взвешенного псевдорешения Мура-Пенроуза. Основное противоречие проблемы разрешается эмиссией информации с высокоинформативных участков идентификации на часто встречающиеся при нормальном функционировании ДО низкоинформативные участки [6].

Приведенная краткая характеристика особенностей поведения нелинейных динамических систем имеет важное значение при разработке моделей формализованной системы знаний бортовых ИС. Структура этих моделей при анализе сложных (особенно экстремальных ситуаций) в antecedentной части логических правил содержит различные встроенные

¹См., например, *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. 3-е изд., испр. – М.: Наука, 1986. – 288 с. — Прим. ред.

процедуры, связанные с выполнением достаточно громоздких вычислений. Это подчеркивает важность использования в бортовых ИС различных методов и подходов, позволяющих обеспечить надежную оценку ситуации и прогноз ее развития в зависимости от характера внешних возмущений.

Как показывают результаты тестирования и натурных испытаний, реализация нейросетевых технологий в формализованной системе знаний открывает новые возможности, обеспечивающие повышение эффективности функционирования ИС. Это позволяет совершенно с других позиций подойти к решению сложных задач, определяющих динамику взаимодействия плавучего ДО с внешней средой. Рассмотренные выше нелинейные эффекты и физические закономерности могут быть учтены при адаптации нейронной сети с целью улучшения качества ее работы и способности самостоятельно менять принципы своей организации.

Концепция использования ИНС в интеллектуальных системах реального времени

Управление бортовой ИС, содержащей различные функциональные подсистемы, представляет собой нетривиальную задачу коллективного интеллекта. В процессе функционирования ИС каждая из подсистем работает в соответствии с принятой коллективной стратегией, определяющей поведение ДО в разрешенном пространстве состояний, сформулированном с учетом текущих состояний других подсистем. Качество работы ИС при реализации коллективной стратегии определяется функцией рассогласования, характеризующей расхождение технических состояний ДО для принятого решения (прогноз) и фактического состояния в рассматриваемый момент времени. Значение функции рассогласования зависит от надежности алгоритмов, заложенных в формализованную систему знаний. Для эффективного функционирования ИС в различных условиях эксплуатации необходимо постоянно «отслеживать» динамические свойства объекта и внешней среды, а также реализовывать полученные данные в алгоритме принятия решений. Выработка управляющих воздействий при организации базы знаний ИС осуществляется на основе результатов динамических измерений параметров состояния ДО и внешних возмущений после выполнения определенных вычислительных и логи-

ческих операций. Среди них важное значение имеют процедуры идентификации, обеспечивающие необходимой информацией адаптивную компоненту базы знаний.

Опыт разработки бортовых ИС позволяет сформулировать концепцию, определяющую основные принципы использования информации при функционировании комплекса. На данной стадии исследований динамики сложных объектов происходит формирование теоретического базиса проблемы создания информационных технологий бортовых ИС реального времени. Сложности, возникающие на пути ее решения, связаны с неполнотой исходной информации и неопределенностью внешней среды, в которой протекает конкретная ситуация. Сформулированный подход определяет концепцию создания бортовых ИС, как систем, основанных на знаниях, непрерывное изменение и совершенствование которых происходит в процессе эксплуатации. Реализация специальных требований к информационной, вычислительной и измерительной технологии, а также условий организации и использования знаний осуществляется с учетом принципов компактности, непрерывности и предсказуемости. Структура знаний обеспечивает возможное их расширение и корректировку в соответствии с новыми теоретико-экспериментальными данными, результатами практических наблюдений и опыта эксплуатации.

Одной из центральных проблем использования методов ИИ в бортовых ИС является формирование инвариантного ядра системы, включающего в себя предметную область, базу знаний и базу данных. Для создания этой совокупности сформулирована концептуальная модель ИС, удовлетворяющая комплексу противоречивых требований. Настраиваясь на предметную область и конкретный объект моделирования, такая модель имеет универсальную схему и формализованное описание [6]. Концепция предусматривает повышение эффективности функционирования комплекса за счет улучшения механизмов формализации сложных задач принятия решений по управлению ДО в условиях неопределенности и неполноты исходной информации. При этом используются следующие преимущества нейросетевых технологий [2, 3, 17, 32]:

- способность к самообучению, замена сложного математического аппарата необходимым объемом информации, исходной функцией ее преобразования и структурой ИНС;

- возможность реализации нелинейных отображений с существенными нелинейностями в скрытых нейронах многослойных ИНС;
- высокая степень параллелизма, возможная для ИНС, обеспечение работы системы в режиме реального времени.

Важной особенностью использования ИНС является возможность сочетания нейросетевых методов с моделями поиска решений, оперирующими образными представлениями с механизмами символично-логического уровня [18].

Теоретический базис разработанной концепции основан на исследованиях, результаты которых обобщены в монографиях, научных статьях и докладах на национальных и международных конференциях [6, 7–13, 25–27]. Оригинальные научно-технические решения защищены патентами Российской Федерации 1994–2000 гг. Основные результаты этих исследований позволяют сформулировать базовые принципы организации алгоритмического обеспечения бортовых ИС реального времени. Эти принципы определяют интеграцию знаний, параллельную обработку информации и визуализацию результатов функционирования ИС.

Преодоление основного недостатка многослойных ИНС, связанного с медленной сходимостью алгоритмов оперативного обучения, достигается за счет использования различных подходов к ускорению реализации этой процедуры. Среди них следует выделить [2]:

- предварительное обучение, эффективные процедуры инициализации и алгоритмы обратного распространения;
- встраивание знаний об особенностях поведения ДО в структуру многослойной ИНС, использование методов имитационного моделирования;
- применение гибридных технологий, в которых ИНС связывается со структурами, разработанными на основе других эффективных методов преобразования информации.

Оценка эффективности общей стратегии использования нейросетевых алгоритмов в бортовых ИС реального времени и получения лучших практических результатов работы достигается в процессе доводки системы на основе данных натурального эксперимента. При этом реализация

ИНС-технологий может быть обеспечена путем последовательного все более полного использования преимуществ различных типов нейросетевых алгоритмов (наращивание структурных элементов системы, создание многоуровневых иерархических архитектур и т. д.).

Таким образом, концепция создания бортовой ИС как системы параллельного действия предусматривает использование ИНС не только как основного метода анализа и интерпретации данных динамических измерений в системах управления ДО, но и в качестве конкурирующей технологии обработки информации при анализе сложных ситуаций. В результате анализа альтернатив предпочтение отдается той технологии, которая обеспечивает более надежную оценку исследуемой ситуации, исходя из принятого дискриминирующего правила. Именно эта технология получает дальнейшее развитие при решении текущей задачи анализа и интерпретации данных динамических измерений.

Стратегия выбора альтернатив реализуется в такой последовательности:

- генерация альтернатив с помощью нейросетевого логического базиса и традиционных алгоритмов контроля характеристик ДО и внешней среды;
- анализ альтернатив и выбор наилучшей модели;
- развитие предпочтительной альтернативы и оценка результата.

Сформулированная концепция имеет важное значение как на стадиях разработки и отладки ИС, так и в процессе ее функционирования в условиях эксплуатации. Сравнение различных алгоритмов определения динамических характеристик объекта управления и внешней среды с результатами моделирования на основе ИНС позволяет оценивать достоинства и недостатки используемых методов, выбирать рациональные пути формирования информационного базиса в порождающих правилах логического вывода.

Управление и принятие решений

С позиций современной теории автоматического управления применение ИНС, как регуляторов динамических объектов, адекватно задачам

аналитического синтеза системы, когда управление становится весьма трудоемкой проблемой из-за сложности и недостоверности строгого математического описания [16]. Применение алгоритмического обеспечения системы управления на основе ИНС и соответствующей аппаратно-программной среды дает возможность:

- уменьшить влияние неточности и неопределенности информации за счет прогнозирующих свойств нейросетевых процедур;
- повысить скорость вычислений путем использования «мелкозернистой» параллельной структуры алгоритма.

Эти достоинства обеспечивают применение нейросетевых алгоритмов в системах управления сложными динамическими объектами.

Концепция «мягких вычислений» в бортовых интеллектуальных системах управления

Концептуальные основы создания бортовых ИС управления ДО базируются на теоретических принципах, определяющих архитектуру системы и уровни ее организации. Эти принципы предполагают учет особенности функционирования ИС реального времени и широкое использование достижений ИИ. Теоретическая база создания ИС формируется на основе эффективного сочетания накопленной системы знаний с новыми подходами и парадигмами ИИ. Наряду с традиционными методами «инженерии знаний» здесь находят применение концепция «мягких вычислений» (soft computing) [6, 9, 25, 28, 33], объединяющая нечеткий логический базис с теорией ИНС и генетических алгоритмов (ГА), а также когнитивная парадигма [5, 6], средства мультимедиа и виртуальной реальности, позволяющие повысить эффективность взаимодействия оператора с ИС. При описании информационных и процедурных элементов «мягких» вычислительных процессов положения нечеткой логики позволяют обеспечить толерантность результата по отношению к неточности и неопределенности, а нейронные сети — формализовать сложные алгоритмы преобразования информации и организовать обучение в процессе выполнения вычислений и их высокопроизводительную реализацию. Развитие математического фундамента ГА обеспечивает оптимизацию базы знаний и

параметров нечетких логических регуляторов, а также выбор стратегических решений в задачах нечеткого моделирования.

В качестве примера применения ИНС в рамках концепции «мягких вычислений» рассмотрим особенности проектирования алгоритмического обеспечения системы автоматического управления (САУ) движением подводного аппарата (ПА) в маршевом режиме [25]. ИС обеспечивает измерение текущих параметров пространственной ориентации ПА, глубины движения, близости дна акватории и скорости хода. На базе данных измерений в режиме связанного управления формируются выходные сигналы на движительно-рулевой комплекс для регулирования или стабилизации заданных параметров движения. Интеллектуальность САУ обеспечивается реализацией в законе управления методологии нечеткой динамической ИС [6, 9, 25, 28].

Структура САУ представлена на рис. 3. Значения измеренных датчиками параметров движения поступают в бортовую вычислительную систему, которая ориентирована на работу со знаниями. По этим данным в блоке анализа текущей ситуации выбирается модель поведения ПА, характеризующая требуемый маршевый режим движения. В соответствии с выбранной моделью формируется база знаний, использующая в своей работе информацию из базы данных. В блоке нечеткой оценки каждое значение измеренного параметра преобразуется в форму, учитывающую меру его неопределенности. В блоке вывода по этой информации и на основе текущего представления базы знаний определяются рекомендации по управлению движением ПА. База знаний вместе с механизмом вывода образует динамическую ИС. В блоке точной интерпретации результаты полученных выводов преобразуются в управление. Сигналы управления поступают в движительно-рулевой комплекс, вырабатывающий управляющее воздействие на ПА. По информации об измеренных значениях параметров движения, сформированном управлении и текущем состоянии базы данных в блоке самонастройки осуществляется ее коррекция.

Обобщенное описание динамики ПА в терминах пространства состояний представляется векторно-матричным дифференциальным уравнением (рис. 4):

$$\begin{aligned} \partial x / \partial t = f(x, s, u, t), \quad x(t) \in R^n, \quad u(t) \in R^m, \\ x(t_0) \in V(t_0), \quad s \in R^k, \quad z(t) = \pi[x(t), t], \quad z(t) \in R^r, \end{aligned} \quad (3)$$



Рис. 3. Структурная схема интеллектуальной системы автоматического управления

где $x(t)$ – вектор состояния, $x(t) \in X \subset R^n$; $u(t)$ – вектор управления, $u(t) \in U \subset R^m$; $z(t)$ – вектор измерения, $z(t) \in Z \subset R^r$; s – вектор неопределенных параметров, $s \in S \subset R^k$; $V(t_0)$ – множество начальных состояний, $V(t_0) \subset X$; $f(\dots)$ и $\pi(\dots)$ – известные вектор-функции; t – время.

Множества X, U, Z, S – заданные компактные выпуклые подмножества соответствующих пространств.

Закон движения ПА задается вектор-функцией (3), которая для рассматриваемого случая имеет вид:

$$f(x, s, u, t) = \left[\frac{\Omega^{-1}(F - B\Omega q)}{\omega} \right], \quad (4)$$

где $q = [v_x, v_z, \omega]^T$ – вектор обобщенных скоростей;

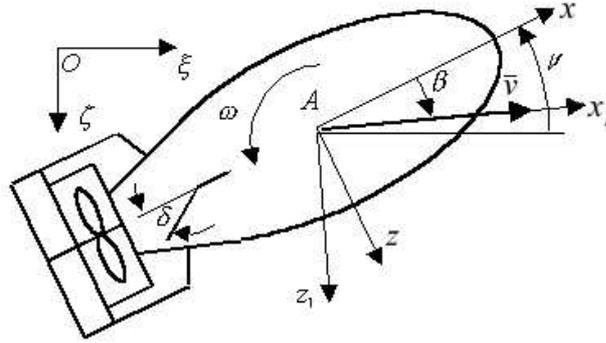


Рис. 4. Схема подводного аппарата и система координат

$$\Omega = \begin{bmatrix} \lambda_{11} + m & 0 & 0 \\ 0 & \lambda_{33} + m & \lambda_{35} - mx_c \\ 0 & \lambda_{35} - mx_c & \lambda_{55} + J \end{bmatrix}$$

есть матрица инерции ПА; $\lambda_{11}, \lambda_{13}, \dots$ — присоединенные массы; m — масса ПА; x_c — координата центра масс;

$$B = \begin{bmatrix} 0 & \omega & 0 \\ -\omega & 0 & 0 \\ v_z & -v_x & 0 \end{bmatrix}$$

представляет собой матрицу обобщенных скоростей;

$$F = \begin{bmatrix} T - 0.5c_x q V^{2/3} \\ 0.5c_z^\beta \beta q V^{2/3} + 0.5c_z^\delta \delta q V^{2/3} + 0.5c_z^\omega \omega q V \\ 0.5m^\beta \beta q V + 0.5m^\delta \delta q V + 0.5m^\omega \omega q V^{4/3} + v_x v_z (\lambda_{11} - \lambda_{33}) \end{bmatrix}$$

задает вектор внешних сил и моментов; $q = \rho v^2$; T — сила тяги движителя ПА; V — объемное водоизмещение ПА; ρ — массовая плотность среды; $c_x, c_z^\beta, c_z^\delta, c_z^\omega, m^\beta, m^\delta, m^\omega$ — гидродинамические коэффициенты; $\beta = \arctan(v_z/v_x)$.

Определим множество ситуаций $C = c_e, e = \overline{1, E}$. Элементами этого множества являются лингвистические значения возможных режимов

управляемого движения ПА: «СТАБИЛИЗАЦИЯ ГЛУБИНЫ», «СТАБИЛИЗАЦИЯ КУРСА», «ЗАДАННАЯ СКОРОСТЬ» и т. д. Преобразование $\theta : Z \rightarrow C$ считается известным.

Для каждой ситуации $c_e \in C$ сформирован показатель качества управления:

$$J^e = \int_{t_0}^T f_0[x(t), x^e(t), u(t), t] dt, \quad (5)$$

где $x(t)$ — вектор состояний, $x^e(t)$ — эталонная траектория движения на интервале $[t_0, T]$; $u(t)$ — вектор управления.

База знаний ИС создана методами концептуального моделирования с использованием продукционных моделей с конструкцией типа **if-then-else**.

Учет неопределенности априорных знаний об элементах вектора $s \in S$ при синтезе базы знаний осуществлялся на основе аппарата теории нечетких множеств [33]. С каждым элементом векторов измерений $z(t)$ и управления $u(t)$ связывались лингвистические переменные, которые описываются кортежем

$$\langle L^i, G^i(L^i), Q^i \rangle, \quad (6)$$

где L^i — имя лингвистической переменной; $G^i(L^i)$ — терм-множество переменной L^i , $G^i(L^i) = \overline{L^i_{a_i}}$; $a_i = \overline{1, A_i}$; $Q^i = q^i$ — универсум с базовой переменной q^i . Функции принадлежности определялись на основе имитационного моделирования поведения ПА с учетом параметрической неопределенности. Моделирование производилось для каждого вектора $x(t_0) \in V(t_0)$. В соответствии с выбранной ситуацией $c_e \in C$ для каждого $s \in S$ на основе (5) решалась задача поиска минимума функционала при заданных ограничениях. По значению $z(t_0) = \pi[x(t_0), t_0]$ и заданным $\mu_{L^i_{a_i}}$, $i = \overline{1, r}$ определялся номер продукции и статистически оценивались значения функции принадлежности $\mu_{L^i_{a_{r+1}}}, \dots, \mu_{L^i_{a_{r+m}}}$ для соответствующих базовых переменных $q^i \in Q^i$, $i = \overline{r+1, r+m}$.

Нечеткий вывод по информации о $\overline{L^1} \dots \overline{L^r}$ осуществляется с использованием композиционного правила, в соответствии с которым

$$\mu_{\overline{L^i}}(q^i) = \bigvee_{k \in K} \left(\bigwedge_j \left(\bigvee_{q^j \in Q^j} (\mu_{L^j_H}(q^j) \wedge \mu_{\overline{L^j}}(q^j)) \right) \wedge \mu_{L^i_H}(q^i) \right),$$

$$i = \overline{1, r}, j = \overline{r+1, r+m}.$$

где \wedge – операция \min ; $V_{k \in K}, V_{q^j \in Q^j}$ – определение наименьшей верхней границы соответствующего множества элементов.

С целью сокращения вычислительных затрат при поиске минимального значения критерия (5) по векторному аргументу использовался ГА. Рассматриваемая задача определялась в терминах генерации набора хромосом, их тестирования и отбора наилучших по требуемым свойствам.

Каждое значение элемента u_l вектора управления u аппроксимируется соотношением:

$$u_l = \frac{\sum_{k=1}^{K_l} \sigma_k^l u_l^k}{\sum_{k=1}^{K_l} \sigma_k^l}, \quad u_l \in [u_{\min}^l, u_{\max}^l], \sigma_k^l \in \{0, 1\}, \quad (7)$$

где u_{\min}^l, u_{\max}^l – граничные значения допустимого диапазона изменения u_l ; K_l – количество точек разбиения диапазона $[u_{\min}^l, u_{\max}^l]$.

Бинарный вектор $\sigma = [\sigma_1^T, \sigma_2^T, \dots, \sigma_m^T]^T$, соответствует хромосоме длиной $K = \sum_{l=1}^m K_l$.

Двоичный элемент σ_j^i определяет значение j -го гена i -ой хромосомы в популяции, состоящей из I хромосом. Начальные значения для генов хромосом выбираются случайным образом. Каждой хромосоме ставится в соответствие оценка критерия качества процесса управления J_e^i . Для хромосомы σ^i вычисляется значение меры пригодности p^i как линейной формы

$$p^i = aJ_r^i(x, x^e, u^i) + b, \quad (8)$$

в которой u^i вычисляется согласно приведенной выше аппроксимации, а коэффициенты a, b являются решением следующей системы

$$\begin{cases} (\sum_{i=1}^I J_e^i)/I = a[(\sum_{i=1}^I J_e^i)/I] + b, \\ \gamma J_e^{\max} = aJ_e^{\max} + b, \\ J_e^{\max} = \max_i J_e^i, \end{cases}, \quad (9)$$

где γ – коэффициент, характеризующий фактор усиления пригодности.

На каждом шаге выполнения ГА с помощью генетических операторов рекомбинации, кроссинговера и мутации решалась задача формирования новой популяции, усиливающей при наследовании требуемые свойства, что соответствует поиску максимума показателя качества (5). Механизм анализа и изменения популяции продолжался до выполнения условия,

когда фактор пригодности элитной хромосомы сохранялся в течение заданного числа итераций. В результате поиска на основе ГА определяются значения элементов u_i вектора управления u для каждого $z(t_0)$.

По данным определения принадлежности каждого $z(t_0)$ к antecedентной части нечеткого продукционного правила и решения оптимизационной задачи по показателю (5) рассчитывались функции принадлежности

$$\mu_{L_{a_{r+1}}^{r+1}}, \dots, \mu_{L_{a_{r+m}}^{r+m}} \quad (10)$$

элементов терм-множеств лингвистических переменных, входящих в консеквентную часть правил.

Входной информацией для формирования вектора u в соответствии с интеллектуальным нечетким законом управления являются текущие измерения вектора z . Эти данные используются для оценки пространственного положения ПА в блоке анализа текущей ситуации. Этот блок, реализующий преобразование θ , представляет собой базу данных реляционного типа, в которой по значениям z выполняются следующие действия: классифицируется ситуация $c_e \in C$, определяется эталонная траектория $x^e(t)$ и вырабатываются рекомендации по выбору соответствующей базы знаний с нечеткими лингвистическими переменными.

Нечеткий вывод по информации об $\bar{L}^1, \dots, \bar{L}^r$ реализуется с помощью композиционного правила :

$$\mu_{\bar{L}^i}(q^i) = V_{k \in K} (\wedge_{j \in Q^j} (\mu_{L_H^j}(q^j) \wedge \mu_{\bar{L}^j}(q^j))) \wedge \mu_{L_H^i}(q^i),$$

$$i = \overline{1, r}, j = \overline{r+1, r+m}.$$

где \wedge — операция \min ; $V_{k \in K}$, $V_{q^j \in Q^j}$ — определение наименьшей верхней границы соответствующего множества элементов.

Преобразование входных нечетких лингвистических переменных $\bar{L}^{r+1}, \dots, \bar{L}^{r+m}$ в четкую оценку управления u осуществляется на основе процедуры точной интерпретации с использованием метода «центра тяжести». Эта процедура реализуется в блоке точной интерпретации и является результатом работы интеллектуального нечеткого закона управления движением ПА. Поддержка открытости ИС управления обеспечивается наличием блока самонастройки. Вычислительная процедура формируется на основе алгоритма обучения нечеткого автомата. Нечеткое состояние автомата используется в алгоритме композиционного правила вывода.

Алгоритмическая сложность интеллектуального закона управления и необходимость функционирования в режиме реального времени требует использования вычислительных процедур, ориентированных на параллельное выполнение в мультипроцессорной вычислительной среде. Реализация этих требований достигнута с использованием многослойной ИНС. Высокий уровень внутреннего параллелизма нейросетевых алгоритмов обеспечивает формирование на выходе ИНС требуемых сигналов по соответствующей совокупности информации от датчиков измерительной системы. Процесс обучения ИНС осуществлялся на заранее выбранном тестовом множестве сочетаний входных и выходных сигналов. Настраиваемыми параметрами сети являлись веса синаптических связей.

Обучающее множество для ИНС интеллектуальной системы управления ПА представляет собой последовательность $\Xi = (z(t), u(t)), t = t_0, t_1, \dots, T$, где элементы вектора измерений $z(t)$ являются входными сигналами сети, а элементы вектора управления $u(t)$ — выходными. Последовательность формируется по результатам имитационного моделирования динамики ПА с использованием интеллектуального закона управления.

Исследование свойств модели динамики ПА позволяет оценить границы интервалов изменения параметров обучающей последовательности $z(t)$ и $u(t)$. Тогда можно воспользоваться следующим стандартным преобразованием, которое для u_l имеет вид

$$\hat{u}_l = \frac{2u_l - u_{\max}^l - u_{\min}^l}{|u_{\max}^l - u_{\min}^l|}, \quad l = \overline{1, m}, u \in \Xi, \quad (11)$$

где \hat{u}_l — выход нейрона последнего слоя ИНС.

Аналогичное преобразование выполняется и для всех элементов вектора $z \in \Xi$.

Процесс обучения ИНС с постоянной топологией заключается в настройке значений синаптических весов таким образом, чтобы при поступлении на ее вход сигналов \hat{z} на выходе сети формировался сигнал, близкий в смысле выбранной оценки точности к нормированному значению \hat{u} .

За критерий качества работы ИНС выбран квадратичный функционал ошибки вида

$$E = 0.5 \sum_{l=1}^m (v_l^D - \hat{u}_l)^2, \quad (12)$$

где D — номер последнего слоя нейронов в ИНС; v_l^D — выход нейрона последнего слоя, характеризующий реакцию ИНС на входные сигналы, соответствующие нормированным значениям элементов вектора \hat{z} ; \hat{u}_l — нормированное значение желаемого выходного сигнала ИНС.

Выбор в качестве функции активации гиперболического тангенса требует нормировки и центрирования элементов обучающей последовательности.

В качестве алгоритма обучения принят стандартный алгоритм обратного распространения ошибки. Согласно этому методу коррекция связей осуществляется на основе процесса градиентного спуска в пространстве весов. Рекуррентные формулы для этой процедуры приведены в [2, 3, 16, 17, 32]. В результате обучения устанавливаются веса синаптических связей. При этом каждой ситуации $c_e \in C$ соответствует своя матрица весов, а коррекция этой матрицы выполняется по реакциям на все элементы обучающей последовательности Ξ . Обобщенная структура многослойной ИНС для рассматриваемой задачи представлена на рис. 5.

Рассмотренная методология проектирования реализована при синтезе интеллектуального закона стабилизации заданной глубины движения ПА с поворотным движительно-рулевым комплексом. Основные характеристики ПА: длина 5,25 м, диаметр 0,19 м, масса 1120 кг, скорость 2,5 м/с. Ограничение максимальной скорости перемещения в вертикальной плоскости равняется 0,1 м/с. База знаний ИС сформирована по результатам имитационного моделирования на основе ГА. Параметры алгоритма: размер популяции 40, длина хромосомы 20, фактор усиления 2, вероятность кроссинговера 0,8, вероятность мутации 0,001. Пример базы знаний в виде таблицы лингвистических правил для ситуации «СТАБИЛИЗАЦИЯ ГЛУБИНЫ» (при нулевом дифференте) дается на рис. 6, где L^1 — лингвистическая переменная «ГЛУБИНА», L^2 — «ДИФФЕРЕНТ», L^3 — «УГОЛ ПОВОРОТА ДВИЖИТЕЛЬНО-РУЛЕВОГО КОМПЛЕКСА», L_j^i — элементы терм-множеств, NB — большое отрицательное, NM — среднее отрицательное, ZE — нулевое, PM — среднее положительное, PB — большое положительное. Для каждого выходного элемента $L_{a_3}^3$ определена функция принадлежности $\mu_{L_{a_3}^3}(q^3)$.

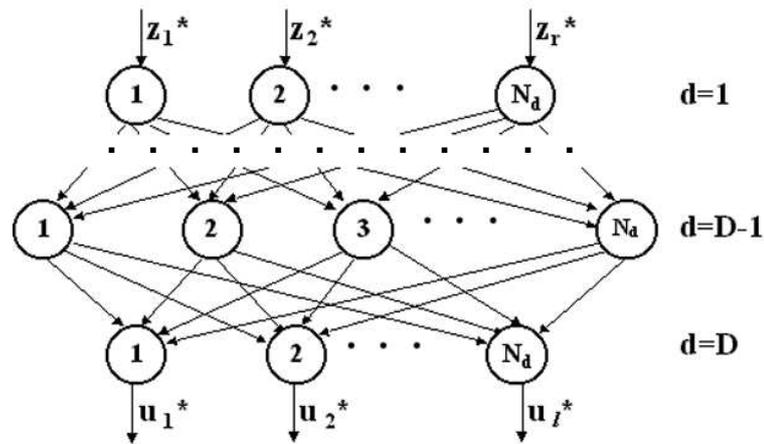


Рис. 5. Структура многослойной нейронной сети

Разработанную нечеткую динамическую ИС аппроксимирует трехслойная ИНС с двумя входами, одним выходом и «скрытым» слоем из восьми нейронов. Процесс стабилизации глубины при начальном отклонении 0,9 м иллюстрирует рис. 7. При этом скорость изменения глубины находится в заданном ограничительном диапазоне.

В результате проведенного исследования реализована методика синтеза интеллектуальных законов управления параметрически неопределенного ПА, функционирующего в условиях пространственно-временной изменчивости внешней среды. Анализ полученных данных позволяет выделить два основных направления развития бортовых ИС ПА. Первое направление объединяет вопросы создания принципиально нового для подводной техники научного и методического обеспечения процесса проектирования различных систем ПА на основе методов ИИ. Второе направление связано с развитием принципов построения бортовых вычислительных систем высокой производительности, разработки их аппаратного и программного обеспечения, ориентированных на обработку знаний как на этапе создания ИС, так и в процессе всего периода эксплуатации.

Stabilisation of depth		L_1^2	L_2^2	L_3^2	L_4^2	L_5^2
		NM	NS	ZE	PS	PM
L_1^1	NB	NB	NM	NM	ZE	PM
L_2^1	NM	NB	NM	NM	ZE	PM
L_3^1	ZE	NM	ZE	ZE	PM	PM
L_4^1	PM	NM	NM	ZE	ZE	PB
L_5^1	PB	NB	NM	ZE	PM	PB

Рис. 6. Фрагмент базы знаний в виде таблицы лингвистических правил

Когнитивный подход при проектировании нейросетевых контроллеров

Перспективным подходом к синтезу нейросетевых алгоритмов является использование когнитивных методов представления и обработки входной информации. В основу этих методов заложен естественный для живых организмов принцип образного представления данных [4]. Идеи когнитивной парадигмы достаточно широко используются в задачах распознавания образов. Применение этого аппарата в системах управления открывает большие возможности реализации программного комплекса в реальном времени и связано с решением ряда научно-технических проблем. Формирование управления сложными динамическими объектами на основе современных достижений в области интеллектуального нейросетевого управления и когнитивного подхода определяет концепцию представления знаний в трудноформализуемых средах.

На рис. 8 представлена структура системы управления, сформированная на основе когнитивного подхода. Блок датчиков осуществляет измерение параметров движения объекта управления и формирует входную информацию для контроллера в виде двоичных векторов. Работа блока формирования когнитивных образов заключается в графической интерпретации входной информации.

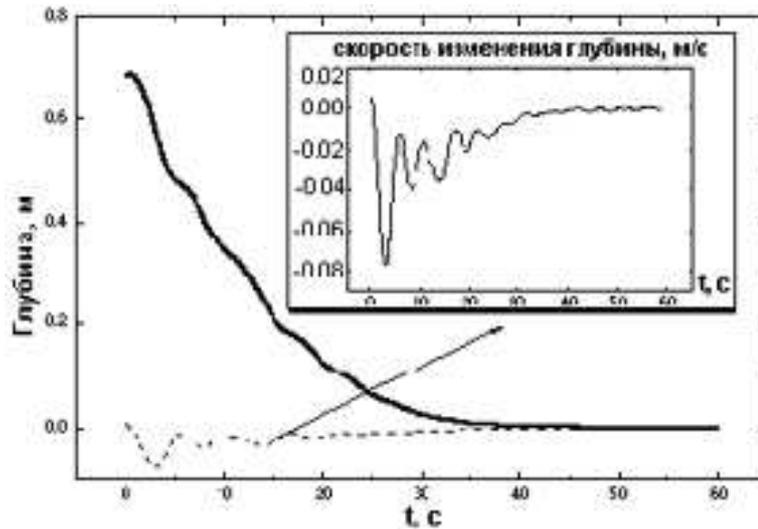


Рис. 7. Процесс стабилизации глубины

Одним из путей реализации когнитивного подхода является использование двумерных самоорганизующихся карт Кохонена [20]. Модификацией соревновательного обучения таких ИНС можно добиться ситуации, когда положение нейрона в выходном слое будет коррелировать с положением прототипов в многомерном пространстве входов сети. При этом близким выходным нейронам будут соответствовать близкие значения входов. В результате появляется возможность строить топографические карты, обеспечивающие визуализацию многомерной информации. Это позволяет сочетать квантование данных с отображением, понижающим размерность. Для повышения эффективности решения используют соревновательные слои в виде двумерных сеток. В результате такого обучения реализуется не только квантование входов, но и упорядочение входной информации в виде одно- или двумерной карт. Каждый многомерный вектор на такой сетке имеет свою координату. Чем ближе координаты двух векторов на карте, тем ближе они и в исходном пространстве.

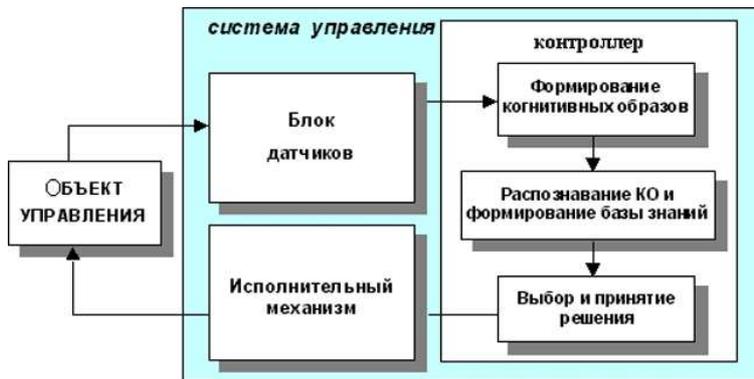


Рис. 8. Система управления

Таким образом, топографическая карта дает наглядное представление о структуре данных в многомерном пространстве. Геометрию такого пространства трудно представить с помощью иных средств анализа и интерпретации информации. Визуализация многомерной информации является одним из основных преимуществ самоорганизующихся карт Кохонена. Этот алгоритм позволяет получить когнитивный образ в виде, представленном на рис. 9.

Блок распознавания когнитивного образа и формирования базы знаний описанной системы (рис. 8) выполняет функции «расшифровки» поступающих образов и их аккумуляции в базе данных. Накопленная информация используется для повышения эффективности управления ДО. В базе знаний когнитивной системы фиксируются закономерные связи между сформированными образами и входными воздействиями.

Блок выбора и принятия решений генерирует команды управления на основе анализа текущей ситуации, формализованной системы знаний и поставленных задач. Эти команды передаются на исполнительные механизмы для формирования управляющих воздействий на ДО.

При проектировании алгоритмического обеспечения контроллера, представленного на рис. 8, необходимо выполнение следующих операций:

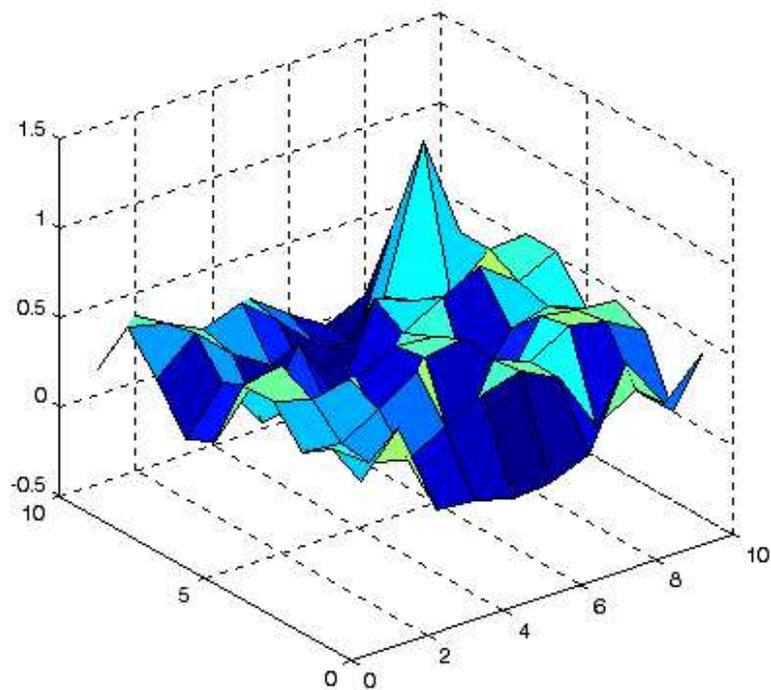


Рис. 9. Когнитивный образ

- математическое моделирование динамики объекта управления с учетом неточности и неопределенности информации о ДО и условиях его функционирования;
- выбор рационального вида нейросетевого алгоритма, формирование библиотеки когнитивных образов и команд управления;
- обучение ИНС на тестовых примерах, представляющих собой предварительно составленную библиотеку когнитивных образов и команд управления.

Анализ качества динамических процессов ведется на основе имитационного моделирования поведения ДО с учетом организации управления

с помощью нейросетевого алгоритма.

Использование когнитивного подхода к представлению информации, поступающей от датчиков измерительной системы, позволяет повысить качество управления. Это достигается за счет феномена когнитивного образа, обладающего более высокой (по сравнению с цифровой информацией) динамичностью и информативностью. Как показывают проведенные исследования, применение когнитивной парадигмы эффективно при проектировании контроллеров для сложных ДО, функционирование которых происходит в условиях неопределенности и неполноты исходной информации. В результате открываются перспективы использования нейрокомпьютеров в системах управления ДО для решения широкого спектра задач обработки информации и принятия решений в различных (особенно в нестандартных и экстремальных) ситуациях.

Многослойная нейронная сеть на нечетких нейронах

Под многослойной нейронной сетью будем понимать набор элементарных преобразователей информации (формальных нейронов), связанных между собой каналами обмена информацией (синаптическими связями) для совместной работы в параллельном режиме. Введем следующие обозначения: x_1, \dots, x_n — входы нейрона; v_1, \dots, v_n — синаптические веса; $\nu_i(v_i, x_i)$ — функция синаптического преобразования; z_1, \dots, z_n — выходы синаптических связей; $\gamma(z_1, \dots, z_n)$ — функция агрегирования выходов синаптических связей; $f_a(b)$ — функция активации; y — выход нейрона.

В зависимости от форм представления функций $\nu_i(v_i, x_i)$ и $\gamma(z_1, \dots, z_n)$ будем различать стандартный и нечеткий нейроны. Для стандартного нейрона функции, характеризующие преобразования в синаптических связях и агрегирование их результатов, имеют вид:

$$\nu_i(v_i, x_i) = v_i \cdot x_i, \quad \gamma(z_1, \dots, z_n) = \sum_{i=1}^n \nu_i(v_i, x_i). \quad (13)$$

В качестве функции активации $f_a(b)$ выбираются монотонные (обычно неубывающие) функции, удовлетворяющие условию $|f_a(b)| < 1$.

Рассмотрим представление нечетких продукционных систем в терминах многослойной ИНС с нечеткими нейронами. Нечетким будем на-

зывать нейрон, структура которого соответствует формальному нейрону с четкими или нечеткими входными сигналами и/или синаптическими весами. Однако в этом нейроне для вычисления функций $\nu_i(v_i, x_i)$, $\gamma(z_1, \dots, z_n)$ могут использоваться t -нормы, t -конормы или другие непрерывные операторы, а функция $f_a(b)$ представляет собой любую функцию от входа [20].

В работе [18] показано, что объединение таких элементов в параллельно-распределенные структуры является универсальным аппроксиматором, то есть может интерпретировать любую непрерывную нечеткую функцию. Построение решений в условиях неопределенности и неполноты исходной информации для прогнозирования состояния ДО основано на формировании логической структуры, описываемой нечетким логическим оператором. Нейроаппроксимацию системы нечетких продукционных правил представим в виде:

$$c : \langle \text{if } \tilde{X}_1 \text{ is } \tilde{X}_1^c \text{ and } \dots \text{ and } \tilde{X}_n \text{ is } \tilde{X}_n^c, \\ \text{then } \tilde{Y}_1 \text{ is } \tilde{Y}_1^c \text{ and } \dots \text{ and } \tilde{Y}_m \text{ is } \tilde{Y}_m^c \rangle; \quad (14) \\ (c = \overline{1, C}),$$

где $\tilde{X} = [\tilde{X}_1, \dots, \tilde{X}_n]^T$ – вектор входных (в общем случае нечетких) данных системы; $\tilde{Y} = [\tilde{Y}_1, \dots, \tilde{Y}_m]^T$ – вектор выходных данных; $\tilde{X}_1^c, \dots, \tilde{X}_n^c, \tilde{Y}_1^c, \dots, \tilde{Y}_m^c$ – нечеткие множества, заданные в виде соответствующих функций принадлежности; c, C – номер правила и их общее количество.

Запишем (14) как композицию следующих продукционных систем:

$$c : \langle \text{if } \tilde{X}_1 \text{ is } \tilde{X}_1^c \text{ and } \dots \text{ and } \tilde{X}_n \text{ is } \tilde{X}_n^c, \\ \text{then } \tilde{Y} \text{ is } \tilde{Y}_c \rangle; \quad (c = \overline{1, C}), \quad (15)$$

где \tilde{Y} – скалярная величина.

Нечеткую нейроаппроксимацию системы (15) выполним на основе работы [29] с использованием нечетких нейронов AND и OR. При их определении в качестве t -норм и t -конорм используются операции \max и \min .

Структура нечеткого нейрона AND терминах обобщенного формального нейрона представлена на рис. 10А. Здесь $\tilde{X}'_1, \dots, \tilde{X}'_n$ – нечеткие (или четкие) данные, поступающие на вход AND нейрона. Оператор

Pos_i ($i = \overline{1, n}$) обозначает процедуру вычисления меры возможности для i -го входа по формуле:

$$\text{Pos}_i = \sup_x \min(\mu_{\tilde{X}'}(x), \mu_{\tilde{X}}(x)). \quad (16)$$

AND–нейрон представляет композицию min–max операций, которая формирует выход

$$y = \min_i \max(w_i, \text{Pos}_i), \quad \forall i \in \overline{1, n}. \quad (17)$$

Структура нейрона OR терминах обобщенного формального нейрона представлена на рис. 10В. Выход нечеткого нейрона OR определяют с помощью max–min композиции

$$y = \max_i \min(w_i, x_i), \quad i = \overline{1, n}. \quad (18)$$

Следует отметить, что входы и выход OR–нейрона являются скалярными величинами.

Рассмотрим принципы построения многослойной ИНС на нечетких нейронах c -го правила системы (15). Пусть заданное правило имеет P различных antecedентов для консеквентной части вида $\langle \tilde{Y} \text{ is } \tilde{Y}^c \rangle$. Каждый p -й antecedент определяет меру возможности $\text{Pos}_{\tilde{Y}^c}^p$. Объединение в правило “с” выполняется на основе операции max

$$\text{Pos}_{\tilde{Y}^c} = \max_p (\text{Pos}_{\tilde{Y}^c}^p), \quad p = \overline{1, P}. \quad (19)$$

Формула (19) может быть реализована с помощью OR–нейрона. В этом случае наличие P элементов учитывается введением синаптических весов w_p . Тогда имеем:

$$\text{Pos}_{\tilde{Y}^c} = \max_p \min(w_p, \text{Pos}_{\tilde{Y}^c}^p). \quad (20)$$

Для вычисления $\text{Pos}_{\tilde{Y}^c}^p$ используется AND–нейрон с n входами.

Система (14) представляет собой композицию по выходам систем типа (15). При этом количество выходных нейронов OR определяется количеством m конъюнктивных составляющих консеквентной части логического правила. Для P нейронов AND вычисление значений Pos_i , ($i =$

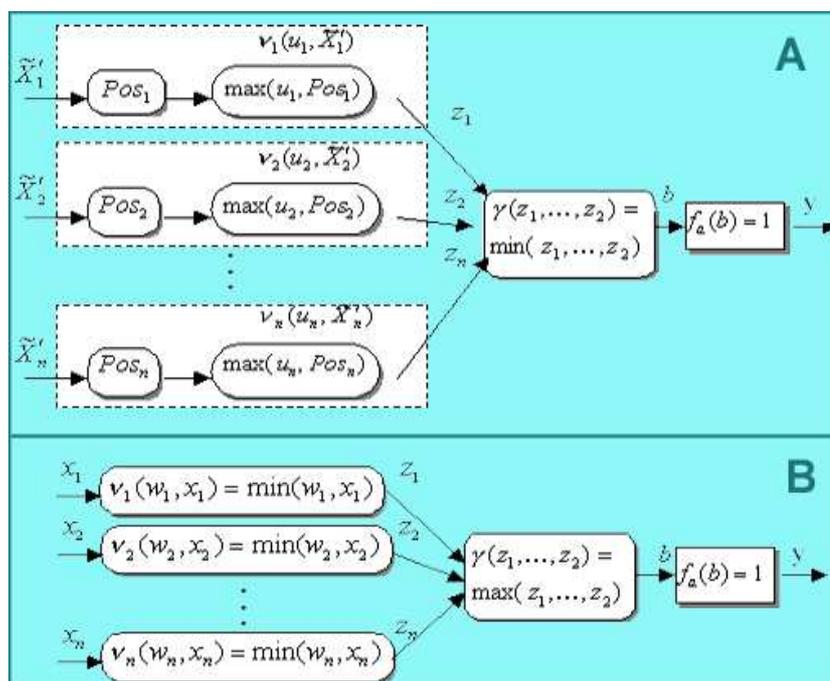


Рис. 10. Структура нечеткого нейрона AND (A) и OR (B)

$\bar{1}, n$) выполняется по одним и тем же формулам. Поэтому они выделяются в отдельный распределительный слой ИНС. Структура сети для c -го правила представлена на рис. 11.

Выходы y_1, \dots, y_m сети трактуются как $Pos_{\tilde{y}_c}$. Их значения используются в вычислительной процедуре нечеткого логического вывода. Синаптические коэффициенты u_{pi} и w_{jp} определяются в процессе параметрического синтеза многослойной ИНС по результатам ее обучения. Для системы S нечетких продукционных правил аппроксимирующая ее сеть содержит такое же количество рассмотренных параллельных нечетких ИНС. Каждая сетевая структура позволяет вычислять меру возможности консеквентной части соответствующего правила.

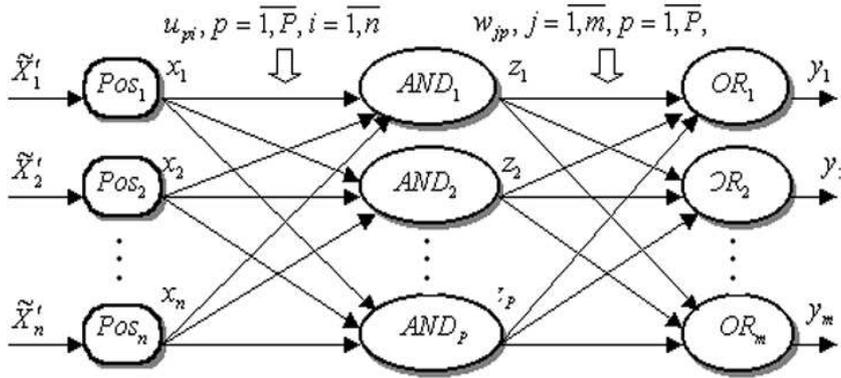


Рис. 11. Структура нечеткой многослойной НС

Рассмотренная многослойная ИНС на нечетких нейронах позволяет адекватно реализовывать в бортовой вычислительной среде систем реального времени нечеткие алгоритмы обработки информации в терминах логико-лингвистических моделей. Реализация нечеткой ИНС осуществлена при создании бортовой ИС обеспечения безопасности мореплавания [6, 9]. Разработка программного комплекса основана на следующих парадигмах ИИ. Основным элементом информационной модели рассматриваемой предметной области является динамическая нечеткая база знаний:

$$\langle B, F, S, I \rangle, \tag{21}$$

где B – база нечетких продукционных правил; F – комплекс алгоритмов формальных процедур обработки измерительной информации; S – набор статистических и нейросетевых процедур анализа и прогноза исследуемых ситуаций; I – система имитационных моделей, описывающая поведение ДО в различных ситуациях.

Нечеткая продукционная модель принятия решений сформулирована на основе (21). При интерпретации ядра продукции в antecedente импликации используются различные встроенные процедуры. Эти процедуры обеспечивают получение необходимой информации о состоянии судна и прогнозирования его изменения в рассматриваемой экстремальной ситуации, а также оценку возможности практической реализации

вырабатываемых рекомендаций. По заданной цепи импликаций формируется нечеткое преобразование, ставящее в соответствие нечеткому дискретному множеству X некоторое нечеткое множество Y . На основании этого преобразования конструируется логическая система, осуществляющая поиск решения с участием оператора. В процессе разрешения нечеткой логической системы определяется взаимосвязь между входными и выходными нечеткими множествами, устанавливаются функции и логические конструкции модели (функции принадлежности, композиционное правило вывода, матрица нечетких отношений). Разрабатываемые стратегии обеспечивают учет характеристик реальных внешних возмущений и особенностей динамики исследуемого объекта в конкретной ситуации. Исходная информация, используемая при функционировании логической системы, формируется на основе обработки данных динамических измерений и математического моделирования на основе (1).

Важная роль при интерпретации информации принадлежит нейросетевой модели базы знаний. На ее основе организация вычислений антецедентной части каждого правила выполняется в нечеткой многослойной ИНС в соответствии с описанным выше алгоритмом. Для вычисления элементов вектора управления U на основе нечеткой продукционной системы используется структура, приведенная на рис.11. Вычисления по композиционному правилу логического вывода выполняются в блоках определения максимума "max". Четкие значения вектора U формируются в соответствии с методом дефаззификации (defuzzification) в блоках "df".

Нейросетевой подход к организации вычислений по интеллектуальному алгоритму управления движением плавучего ДО позволил обеспечить функционирование ИС в реальном масштабе времени и получить важные для обеспечения безопасности эксплуатации практические результаты.

Использование ИНС в задачах управления и принятия решений не ограничивается рассмотренными примерами. При реализации нечетких систем управления возникает необходимость автоматического построения функций принадлежности на основе измерительной информации. Решение этой задачи ведется на основе стратегии конкурентного отбора с использованием ИНС и стандартных алгоритмов преобразования информации.

Большой интерес при функционировании бортовых интегрированных ИС реального времени представляет задача анализа альтернатив. В условиях неопределенности и неполноты исходной информации в многокритериальных системах анализ альтернатив связан со значительными вычислительными сложностями. Результаты исследований [6, 10–12] показывают, что использование нейросетевых технологий в этом случае оказывается более предпочтительным, особенно при их реализации в рамках концепции «мягких вычислений».

Идентификация экстремальной ситуации

Одной из сложных проблем при функционировании бортовой ИС обеспечения безопасности мореплавания является идентификация экстремальной ситуации, связанная с оценкой случая затопления отсеков при получении плавучим ДО пробоины. Оценка и прогноз развития аварийной ситуации в условиях непрерывного изменения динамики объекта и внешней среды связаны с решением чрезвычайно сложной проблемы. Подход к ее решению на основе традиционных алгоритмов сформулирован в работе [6]. Не останавливаясь на особенностях используемого алгоритма преобразования информации, отметим, что при реализации механизма логического вывода в этом случае приходится использовать различные встроенные процедуры. Эти процедуры направлены на выяснение законов распределения исследуемых случайных величин, изучение особенностей поведения фазовых траекторий вблизи точек устойчивого равновесия и математическое моделирование процесса взаимодействия ДО с внешней средой. Практическая реализация указанных процедур даже при использовании высокопроизводительных вычислительных средств представляет определенные трудности в связи с жесткими временными ограничениями при функционировании бортовой ИС реального времени. Поиск эффективных алгоритмов анализа экстремальных ситуаций привел к задаче распознавания на базе ИНС. В качестве исходной информации используются когнитивные структуры, построенные по данным измерений нестационарных случайных процессов, характеризующих динамику аварийного объекта на волнении.

Алгоритм преобразования информации при реализации механизма логического вывода в задаче идентификации представлен на основе про-

дукционной модели. При интерпретации ядра продукции в antecedенте импликации используются различные встроенные процедуры. Проверка достоверности распознавания экстремальной ситуации произведена методами имитационного моделирования. Решение задачи распознавания с использованием когнитивной парадигмы [5, 17] проводилось путем преобразования измерительной информации в рамках формально-логического описания. Такой подход позволяет устанавливать возникновение и развитие колебательных режимов аварийного ДО на волнении в зависимости от характера затопления и уровня внешних возмущений. Поиск конкретной модели отображения осуществлялся с помощью когнитивной спирали, позволяющей «сжать» исходную информацию о нелинейной качке со случайным изменением периодов. Построение когнитивной спирали на случайной функции приводит к «выравниванию» периодов и преобразованию их к одному значению путем аффинного сжатия или растяжения интервалов исследуемой функции до заданного размера.

Когнитивная спираль как альтернатива фазовому портрету несет в себе больше той информации, которую можно выделить визуально, ориентируясь на структуру когнитивного образа (рис. 12). Используя такую форму представления измерительной информации, можно построить корреляционную функцию (или спектральную плотность) процесса для определенных сечений когнитивной спирали, соответствующих участкам квазистационарности. В результате устанавливается число входных нейронов и определяется структура ИНС, а главное — появляется свобода выбора длины реализации, исходя из условий надежности функционирования ИС при оценке экстремальных ситуаций. Для повышения точности распознавания и сокращения времени обучения ИНС из используемых реализаций удалялись скользящее среднее и производился расчет корреляционной функции для «очищенных» данных.

Разработанная ИНС представляет собой последовательное соединение ряда слоев формальных нейронов в виде конусообразной конфигурации (рис. 13). Входной слой имеет число нейронов, соответствующее числу точек в подаваемом образе. Выходной слой имеет три нейрона — по одному нейрону на каждую классифицируемую ситуацию. Функции активации выбираются следующим образом: либо для всех трех слоев используется сигмоидная функция, либо первые два скрытых слоя

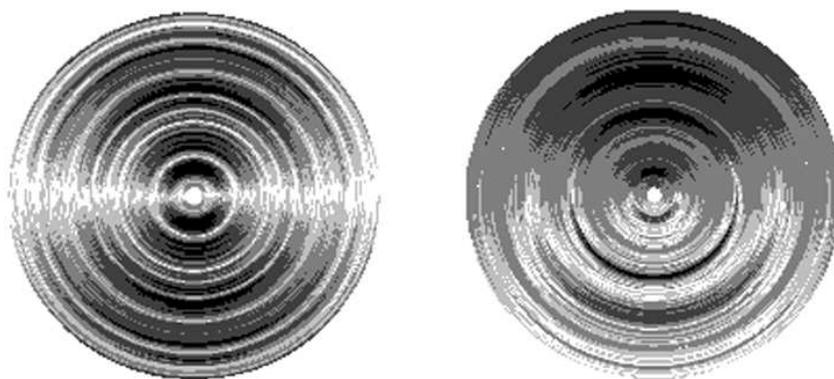


Рис. 12. Когнитивная структура

имеют сигмоидную функцию, а третий, выходной, слой использует модифицированную линейную функцию (в этом случае сокращается время обучения сети). Полученная структура многослойной ИНС обеспечивает сложное нелинейное преобразование входного вектора с учетом входа инициализации в зависимости от векторов синаптических связей и нелинейных функций активации. Связь между входом и выходом ИНС устанавливается нелинейным рекуррентным уравнением [15].

В процессе обучения ИНС производилась настройка весов синаптических связей с использованием итерационной процедуры обратного распространения ошибки. Обучение осуществлялось на корреляционных функциях разрезов когнитивных спиралей. Когнитивные образы строились по 100 периодам реализации бортовой качки ДО на волнении. Затем выделялись разрезы образа-спирали в определенном сечении, из которых удалялись средние значения. При обучении использовалось 45 примеров, что обеспечивало необходимую точность распознавания около 100%. На вход ИНС подавались первые 16 значений корреляционной функции. Практическая реализация полученной многослойной ИНС в задачах распознавания экстремальных ситуаций показала удовлетворительные результаты для различных предъявляемых комбинаций затопленных отсеков. Однако в отдельных случаях при классификации динамически изменяющихся ситуаций возникали затруднения. Топология

фазового пространства в таких ситуациях сохраняется стабильной только на ограниченных интервалах времени, что приводит к неустойчивой работе ИНС. Для решения этой проблемы в качестве конкурирующей использовалась модель ИНС в виде самоорганизующейся карты Кохонена [20].

Эффективность описанного подхода проверялась на распознавании наиболее сложных из пяти классических случаев затопления. При правильно заданной начальной весовой матрице входные данные классифицировались достаточно корректно и ошибка обучения была заметно меньше, чем при использовании чистой конкуренции нейронов.

Для оценки эффективности работы ИНС проводилось исследование возможности использования полученных результатов для решения практических задач классификации при функционировании бортовой ИС. Исследование включало генерацию реального морского волнения и моделирование процесса взаимодействия аварийного ДО с внешней средой с помощью специально разработанного инструментального средства тестирования и отладки ИС. Это инструментальное средство дополнено подпрограммами построения и обработки когнитивных структур и обучения ИНС. В ходе исследования выполнялось качественное сравнение

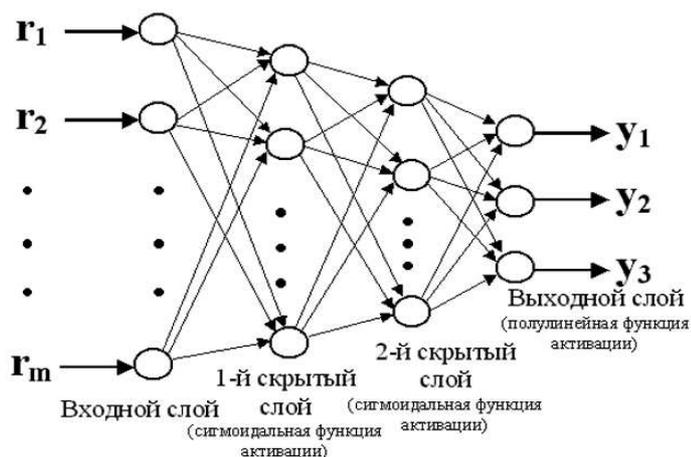


Рис. 13. Многослойная ИНС

результатов классификации для рассматриваемых образов в виде корреляционной функции и спектральной плотности, а также для «коротких» и «длинных» реализаций нелинейных случайных процессов, характеризующих бортовую качку аварийного ДО на нерегулярном волнении.

Конфигурация и основные характеристики сравнительного анализа:

- классификационный образ в виде корреляционной функции: число нейронов входного слоя 16, первого «скрытого» слоя 12, второго «скрытого» слоя 8, выходного слоя 3;
- классификационный образ в виде спектральной плотности: число нейронов входного слоя 40, первого «скрытого» слоя 30, второго «скрытого» слоя 20, выходного слоя 3.

Из этих данных следует, что использование корреляционной функции обеспечивает более эффективное решение задачи классификации, что особенно важно при функционировании ИС реального времени. Интересные результаты получены при исследовании длины реализации и вида функции активации на работу ИНС. Использование этих результатов в практических приложениях позволило сократить время обучения при адаптации сети в случае, когда развитие экстремальной ситуации происходит в течение короткого интервала времени и необходимо быстро оценить обстановку и принять меры по сохранению живучести ДО. Реализация самоорганизующихся карт Кохонена для сложных ситуаций при тех же начальных условиях приводила к определенным преимуществам перед многослойной ИНС. Это позволило оценивать эффективность используемых нейросетевых алгоритмов в зависимости от информации о поведении ДО в экстремальной ситуации.

Оценка динамических характеристик объекта и внешней среды

Использование ИНС при обработке измерительной информации позволяет извлекать скрытые закономерности из потока данных. При наличии достаточного числа примеров и необходимого объема входной информации ИНС способна настраиваться на решение задачи даже при отсутствии алгоритма или принципа решения, зашумленности и частичной

противоречивости данных. Если между входными и выходными данными существует определенная связь, трудно обнаруживаемая традиционными методами, то ИНС способна автоматически настроиться на нее. Это свойство ИНС можно использовать в практических задачах оценки характеристик ДО не зная правил и располагая только набором необходимых примеров. Следует также отметить еще очень важное свойство ИНС в задаче оценки динамических характеристик объекта и окружающей среды — возможность работы с малой обучающей выборкой. Результаты экспериментальных исследований показывают, что эффективным средством устранения ошибок, возникающих при наличии малой обучающей выборки, является подход, основанный на сочетании различных нейросетевых архитектур.

Для реализации на основе ИНС были рассмотрены задачи, решение которых в рамках стандартных алгоритмических процедур вызывает определенные затруднения. Среди этих задач следует выделить такие, как:

- идентификация и прогноз параметров морского волнения по результатам динамических измерений;
- оценка и прогноз равновесных параметров аварийной ватерлинии при качке поврежденного ДО на волнении;
- оценка и прогноз инерционных характеристик, поперечной метацентрической высоты и периода собственных колебаний ДО на волнении.

Другие задачи применения нейросетевых технологий в бортовых ИС связаны с контролем возникающих напряжений, параметров шума и вибраций с целью предотвращения превышения допустимых предельных значений, определяемых действующими нормативами.

Указанные задачи составляют алгоритмический базис оценки и прогноза развития ситуаций в широком спектре внешних условий, в том числе и в резко изменяющейся обстановке. В соответствии с принятой концепцией, на основании результатов измерений в реальном масштабе времени обеспечивается самообучение и адаптация системы в динамически изменяющейся внешней среде. Концепция использования нейросетевых технологий предусматривает повышение эффективности функ-

ционирования ИС за счет высокой производительности вычислений и особенностей решения задач, адекватных по возможностям задачам ИИ.

При использовании ИНС в качестве альтернативного метода моделирования нейронная сеть рассматривается как идентификатор (эмулятор) системы. Здесь в качестве исходной модели исследуемой характеристики используется простая математическая модель (при этом точное описание может быть достаточно сложным и содержать различные виды нелинейностей). В случае несоответствия между действительным выходным сигналом и выходом модели нейронная сеть осуществляет самонастройку системы с целью минимизации ошибки расхождения. Таким образом, эмулятор здесь рассматривается как «сумма» условной модели и ИНС, однако настраивается только нейронная сеть. Аналогичная стратегия находит применение при разработке «интеллектуальных» датчиков. Только в этой задаче выход нейронной сети, имитирующей работу датчика при настройке ИНС, сравнивается с действительным сигналом парируемого датчика.

Рассмотренные стратегии использования ИНС в задачах оценки и прогноза динамических характеристик плавучего объекта демонстрируются на примере контроля важнейшего мореходного качества плавучего ДО — остойчивости, то есть способности противостоять воздействию динамических кренящих нагрузок. Исследование проводилось на основе измерительной информации, полученной во время натурных испытаний контейнеровоза в Средиземном море и в Атлантическом океане. В качестве исходной информации использовались записи бортовой качки на морском волнении различной интенсивности. В процессе исследования отработывался нейросетевой алгоритм оценки инерционных характеристик и поперечной метацентрической высоты. С помощью этих характеристик можно установить положение центра масс и провести математическое моделирование поведения ДО в рассматриваемой ситуации. Результаты анализа сравнивались с данными, полученными во время натурных испытаний. Реализация алгоритма осуществлена на основе многослойной ИНС прямого распространения. В качестве исходной информации при анализе использованы записи бортовой качки, произведенные на коротких интервалах времени в процессе развивающегося шторма. Каждая из записей содержит около 2000 экспериментальных точек с интервалом дискретизации 3 Гц. Предварительная обработка информации выполнена

с помощью фильтра Калмана. Использование корреляционной функции позволило упростить топологию ИНС и сократить время ее обучения. В процессе нелинейного преобразования, задаваемого сигмоидной функцией, область значений входных данных отображается в более узкую область на выходе внутреннего слоя нейронов. Это сжимающее преобразование делает ИНС гибким инструментом моделирования, и в то же время создает определенные трудности при обучении. Для повышения сходимости поисковых алгоритмов используются подходы, основанные на задании начальных условий для элементов весовых матриц и учитывающие специфические особенности сигмоидных функций.

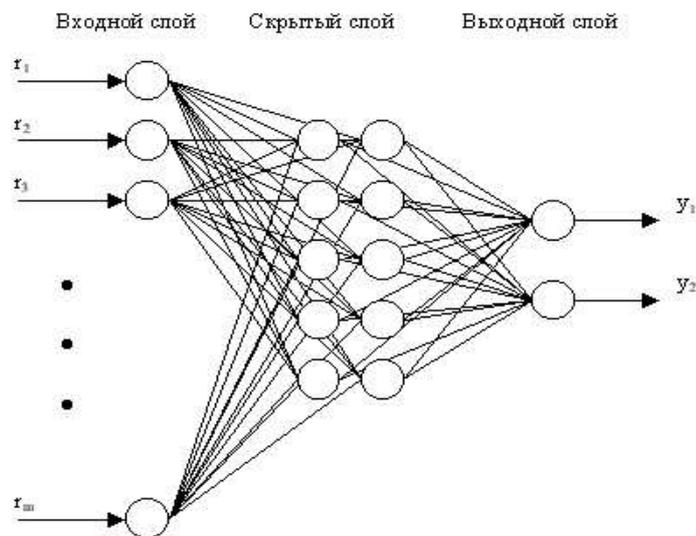


Рис. 14. Топология ИНС

Топология ИНС представлена на рис. 14. Сеть имеет 20 входных нейронов, 10 нейронов в «скрытом» слое и 2 выходных нейрона. Обучение сети производилось на основании алгоритма обратного распространения. В качестве меры расхождения принята функция квадрата ошибки. Для обновления весов применялся метод топологического упорядочения. В процессе обучения программа выбирает решения, принадлежащие допу-

стимулу множеству, что позволяет получить более устойчивую модель. Графики показывают процесс обучения сети при различных начальных инициализациях. Сплошные кривые характеризуют ошибку сети на обучающем множестве, а пунктирные — на проверочном множестве, не используемом при обучении. Неустойчивость возникает из-за недостаточного количества векторов в обучающей выборке. Она обусловлена различными решениями, удовлетворяющими обучаемому множеству. Часть из этих решений не удовлетворяет проверяющим векторам. В процессе обучения сеть может стремиться к любому из этих решений, приближаясь или удаляясь от решения для проверяющего множества.

Сравнительные данные оценки метацентрической высоты и инерционного коэффициента, свидетельствуют о том, что величина расхождения данных при реализации нейросетевой модели не выходит за пределы обычной погрешности гидродинамического эксперимента при исследовании характеристик ДО на волнении.

Другая важная задача оценки динамических характеристик связана с определением текущих значений равновесных параметров аварийной ватерлинии ДО на волнении. Непрерывный контроль этих параметров в процессе затопления отсеков позволяет устанавливать положение аварийной ватерлинии и осуществлять прогноз интервала времени до наступления критического состояния ДО, характеризуемого резким ухудшением мореходных качеств. Исходной информацией для решения этой задачи служат данные динамических измерений параметров бортовой, килевой и вертикальной качки ДО на волнении.

Таким образом, разработанный подход к использованию нейросетевых моделей для оценки динамических характеристик может быть использован в задачах обеспечения безопасности судна на волнении. Важную роль при организации обработки измерительной информации с использованием нейросетевых технологий имеет адаптивная компонента. Одной из ее функций является сравнение реально полученных параметров с предсказанными значениями для генерации и корректировки управления и принятия решений. Афферентный синтез, осуществляемый этой компонентой, лежит в основе целенаправленной деятельности. Адаптивная компонента использует нелинейные модификации алгоритмов калмановской фильтрации и самоорганизации. При получении информации от фильтра Калмана алгоритм нелинейной самоорганиза-

ции с использованием разработанной стратегии конкурирующего отбора строит модели оценки и прогноза исследуемых характеристик ДО.

Моделирование и интерпретация моделей — основное средство и инструмент анализа сложных процессов и явлений в ИС реального времени. Отображение динамики ситуации с учетом данных измерений и моделирования в графическом режиме осуществляется с помощью совокупности интерактивных программных средств, реализующих диалоговое взаимодействие оператора с ЭВМ с подключением механизмов представления и обработки знаний. Динамическая модель работает совместно с графическим интерфейсом, обеспечивающим наглядное отображение процесса развития ситуации, индикацию внешних воздействий и оперативного изменения управляющих сигналов.

Интеллектуальные датчики на основе ИНС

Важным результатом исследований, связанных с повышением качества функционирования измерительного комплекса бортовой ИС, является создание специальных интеллектуальных датчиков (ИД). Традиционные направления совершенствования таких датчиков реализуются как путем улучшения конструкции и элементов аппаратной части, так и за счет разработки методов формального описания и анализа измерительных процедур. Развитию аналитического подхода при синтезе интеллектуальных датчиков способствует создание алгоритмов оптимальной обработки данных на основе все более сложных статистических моделей входных воздействий на датчик. В качестве конкурирующей стратегии рассматривается способ реализации «интеллектуального» датчика на основе ИНС.

Архитектура отказоустойчивой измерительной системы бортового комплекса с «интеллектуальным» датчиком представлена на рис. 15. Система содержит два логических блока. Первый блок устанавливает неисправность и настраивает моделируемую систему обеспечения функционирования нужного датчика. В частности, для интеллектуального датчика на основе ИНС осуществляется настройка весов синаптических связей. Второй логический блок обеспечивает достоверность информации интеллектуальных датчиков IS_1, \dots, IS_N . Интеллектуальный контроллер ИК устанавливает факт возникновения отказа датчиков X_1, \dots, X_n .

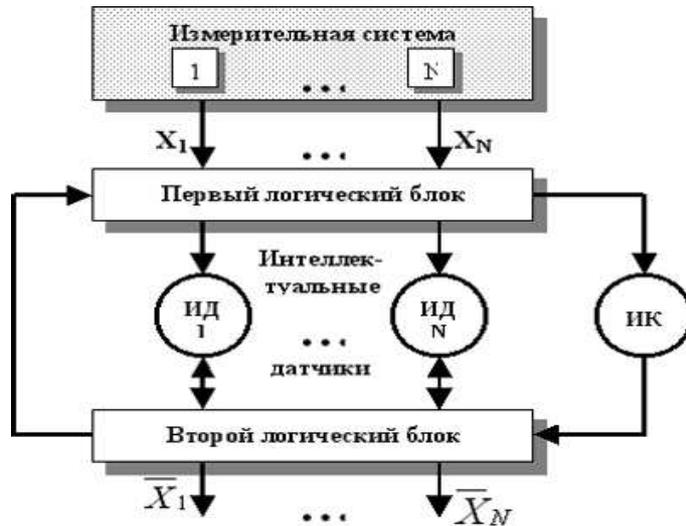


Рис. 15. Схема отказоустойчивой измерительной системы

Алгоритмы функционирования интеллектуальных датчиков разрабатываются с учетом реальных данных, характеризующих поведение исследуемого ДО в заданных условиях эксплуатации. В частности, для обеспечения информации о колебательном движении ДО на волнении разработана система обеспечения режимов работы интеллектуальных датчиков бортовой, килевой и вертикальной качки на реальном морском волнении различной интенсивности. Система способна настраиваться на условия развивающегося шторма, тяжелого обледенения, затопления отсеков и других экстремальных ситуаций в практике эксплуатации судов и плавучих технических средств освоения океана.

При разработке алгоритмов предусматривается определенная последовательность операций:

- составляется спецификация режимов работы бортовой ИС в различных условиях, в том числе и в экстремальных ситуациях;

- производится обучение интеллектуального датчика в зависимости от принципа его организации (на основе математического моделирования или теории ИНС);
- формируется логика функционирования диагностической системы при отказе конкретного датчика и замене его выходного сигнала соответствующим выходом ИС.

При применении самоорганизующейся ИНС каждому входному параметру $x_i (i = \overline{1, n})$ сопоставляется измерение многомерного пространства. Размерность этого пространства равна числу параметров n , возможные значения которых задаются специальной шкалой. Задача состоит в определении свойств точек $x_j = (x_{j1}, \dots, x_{jn})$ n -мерного пространства. Величина x_{ji} характеризует значение входного параметра i точки j в примерах, используемых при обучении ИНС. Точки многомерного пространства образуют класс связанных областей точек, обладающих определенными свойствами. При обучении ИНС этот класс запоминается и используется при функционировании ИС для решения задачи контроля и управления.

Таким образом, создается искусственная интеллектуальная среда, включающая в себя множество функциональных подсистем в виде совокупности интеллектуальных датчиков. Обучение ИНС осуществляется в составе бортовой ИС (рис. 16). Входными значениями ИНС X являются характеристики сигналов с датчиков. Выходы обученной ИНС Y совпадают со значениями сигналов с датчиков. Если разность сигналов E превышает допустимый уровень, блок диагностики выдает сигнал об отказе датчика. После этого в качестве выходного сигнала датчиков принимается его значение на выходе нейронной сети.

Система принятия решений по использованию интеллектуального датчика реализуется на основе нечеткого логического базиса. Нечеткий вывод осуществляется с помощью специальных таблиц принятия решений [1]. Разработка таких таблиц ведется для нормальных условий эксплуатации и при отказе датчиков (нештатные режимы).

Сравнительный анализ моделей сигналов, реализуемых интеллектуальным датчиком на основе ИНС, а также фактических результатов, полученных в процессе экспериментальных исследований ИС, проведен на

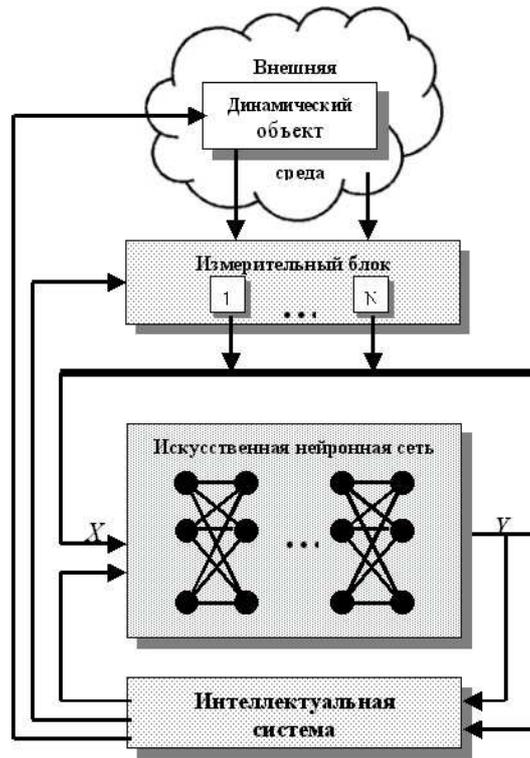


Рис. 16. ИС управления с интеллектуальным датчиком

основе статистического критерия:

$$K_s = \frac{1}{\bar{y}} \sqrt{\sum_{i=1}^n (\bar{y}_i - y_i)^2 / n}. \quad (22)$$

Этот критерий выражает в стандартизированном масштабе степень устойчивости исследуемых динамических процессов: отношение среднеквадратического отклонения к среднему уровню.

Эксперименты, проведенные с использованием реальных данных о функционировании ИС обеспечения безопасности мореплавания с помощью датчиков колебательного движения плавучего ДО при бортовой и килевой качке, свидетельствуют о практической реализуемости разработанного подхода. Значения коэффициентов аппроксимации рассматриваемых регрессионных моделей оказывается выше, чем у моделей ИНС. Замена датчика бортовой качки интеллектуальным датчиком, генерирующим случайный существенно нелинейный процесс, позволила на основе специально разработанного алгоритма:

- восстанавливать статистически характеристики морского волнения и параметров качки;
- обеспечивать работу механизма логического вывода при идентификации ситуаций, связанных с возникновением резонансных режимов качки при движении плавучего ДО с различными курсовыми углами на волнении в условиях развивающегося шторма.

Преимущества нейросетевых моделей в подобных задачах состоит также в том, что их модификация и настройка ведется по мере получения новых данных наблюдений. Важная особенность применения моделей ИНС в бортовых ИС заключается в возможности функционирования системы на основе временных последовательностей с малым интервалом наблюдений. Для надежного функционирования интеллектуальных датчиков важное значение имеет информация, накапливаемая в процессе эксплуатации бортовой ИС. На базе этой информации производится окончательная корректировка алгоритмов, заложенных в формализованную систему знаний при построении интеллектуальных датчиков. При этом для обеспечения надежности работы ИС в экстремальных условиях целесообразно использовать данные динамических измерений, близкие к предельным значениям характеристики в непрерывно изменяющихся внешних условиях.

Практическая реализация разработанного подхода связана с проблемой старения информации, что требует непрерывного поддержания работоспособности ИС на базе фактической информации о поведении плавучего ДО в различных условиях эксплуатации.

Заключение

Опыт использования нейросетевых технологий в бортовых ИС реального времени позволяет рассматривать проблематику их разработки, эффективности функционирования и надежности в области систематического анализа и практических приложений. Отмечается большое разнообразие типов и структур ИНС, находящих применение в бортовых ИС. Их отличает сложность, размеры решаемых задач и функциональные характеристики. О том, насколько адекватно описана исследуемая предметная область и реализованы процедуры логического вывода с использованием информации, полученной с помощью ИНС и других методов анализа и преобразования данных, можно судить об эффективности функционирования ИС в рассматриваемой ситуации. Синергетический эффект интеграции нечетких систем, теории ИНС, эволюционного моделирования и когнитивной парадигмы составляет научную основу для глубинной интеграции сложных интеллектуальных комплексов, объединяющих бортовые ИС различного назначения.

Анализ особенностей разработки бортовых ИС показал, что сложности, присущие традиционному подходу к синтезу таких систем в значительной степени могут быть преодолены за счет применения нейросетевых технологий. Рациональное использование ИНС позволяет обеспечить гибкость и способность адаптироваться к изменяющимся внешним условиям. При этом сохраняется устойчиво высокое качество работы и реализуется недостижимый ранее уровень вычислительной мощности.

Полученные результаты свидетельствуют о больших возможностях использования нейросетевых технологий в бортовых ИС анализа и прогноза поведения плавучих ДО в сложной гидрометеорологической обстановке. Эти возможности базируются на принципиально новых алгоритмах решения сложных задач динамики в рамках нейросетевого логического базиса. Характерные черты ИНС — параллельность, распределенность, адаптивность (способность к обучению) позволяют автономно «изучать» свойства исследуемого ДО на основе результатов измерений. Накапливая подобного рода информацию, можно организовать процесс обучения таким образом, чтобы принять лучшее решение в условиях неполноты и неопределенности исходной информации. Проведенные эксперименты свидетельствуют о работоспособности и эффективности использования нейросетевых технологий в бортовых ИС. Параллелизм и

возможность нелинейного преобразования информации открывают перспективы практической реализации ИНС в сложных задачах анализа и интерпретации данных динамических измерений.

Ресурсы Интернет по теме лекции

URL: <http://www.basegroup.ru/>
URL: <http://neuropower.de/>
URL: <http://www.orc.ru/~stasson/>
URL: <http://www.neuroproject.ru/>
URL: <http://www.neuralbench.ru/>
URL: <http://neurnews.iu4.bmstu.ru/>

Литература

1. *Валеев С.С., Васильев В.И., Ильясов Б.Г., Сун Жан-Гуо.* Отказоустойчивые системы управления сложными динамическими объектами с использованием ИНС // *Нейрокомпьютеры: разработка и применение.* – 2000. – №1. – с. 32–35.
2. *Галушкин А.И.* Теория искусственных нейронных сетей. – М.: ИПРЖР, 2000.
3. *Горбань А.Н., Россиев Д.А.* Нейронные сети на персональных компьютерах. – Новосибирск: Наука, 1996.
4. *Жданов А.А., Одинцов К.В.* Возможности представления измерительной информации с помощью формальных нейронов // *Сб.научн.тр. «Искусственный интеллект в технических системах».* – 1999. – Вып. 20. Гос.ИФТМ, 1999, с.111-146.
5. *Зенкин А.А.* Когнитивная компьютерная графика. – М.: Наука, 1991.
6. *Интеллектуальные системы в морских исследованиях и технологиях.* – Санкт-Петербург: ГМТУ, 2001.
7. *Лукьяница А.А., Олейниченко Л.Г., Торшин А.Д., Тюменцев Ю.В.* Динамические интеллектуальные системы и нейроинформатика // *«Динамические интеллектуальные системы в управлении и моделировании».* Тр. семинара. М.: УРДЗ, 1996. – с. 84–91.
8. *Нечаев Ю.И., Сиек Ю.Л., Васюнин Д.А.* Нейросетевые технологии в интеллектуальных системах морской техники // *Тр. 6-й Национальной конференции по искусственному интеллекту с международным участием КИИ-98.* – Пушино, 1998. – т. 2. – с. 361–368.

9. *Нечаев Ю.И., Сиек Ю.Л.* Концепция мягких вычислений в бортовых интеллектуальных системах реального времени // Тр. Международной конференции по мягким вычислениям и измерениям SCM-99. – Санкт-Петербург, 1999. – т. 1 – с. 64–68.
10. *Нечаев Ю.И., Сиек Ю.Л., Бондарев В.А.* Нейросетевой логический базис в интеллектуальных системах управления морскими динамическими объектами // Тр. 2-й Всероссийской научно-технической конференции «Нейроинформатика-2000». М.: МИФИ, 2000ю – ч. 1. – с 254–262.
11. *Нечаев Ю.И., Дегтярев А.Б., Кирюхин И.А.* Интеллектуальные системы поддержки принятия решений с использованием нейросетевых технологий // Тр. Международного симпозиума «Интеллектуальные системы», INTELS-2000. – М., 2000. – с. 171–173.
12. *Нечаев Ю.И., Дегтярев А.Б., Сиек Ю.Л.* Принятие решений в интеллектуальных системах реального времени с использованием концепции мягких вычислений // Искусственный интеллект. – 2000. – №3. – с. 525–533.
13. *Нечаев Ю.И., Дегтярев А.Б., Кирюхин И.А.* Идентификация нечетких ситуаций с использованием искусственных нейронных сетей и когнитивных структур // Тр. 7-й Национальной конференции по искусственному интеллекту с международным участием КИИ-2000. – Переславль-Залесский, 2000. – т. 2. – с. 492–499.
14. *Сигеру Омату и др.* Нейроуправление и его приложения. – М.: ИПРЖР, 2000.
15. *Сиек Ю.Л., Рудинский А.В., Ермоленко А.С.* Нейросетевой алгоритм обнаружения гидроакустических сигналов // Тр. 2-й научно-технической конференции «Нейроинформатика-2000». – М.: МИФИ, 2000. – ч. 1. – с. 268–275.
16. *Сиек Ю.Л., Воронин Р.И.* Когнитивный подход при проектировании нейросетевых контроллеров // Тр. Международной конференции по мягким вычислениям и измерениям SCM-2001. – Санкт-Петербург, 2001. – т. 1. – с. 243–247.
17. *Терехов В.А., Ефимов Д.В., Тюкин И.Ю., Антонов В.Н.* Нейросетевые системы управления. – СПб.: Изд-во СПбГУ, 1999.
18. *Уоссермен Ф.* Нейрокомпьютерная техника. – М.: Мир, 1992.
19. *Фоминых И.Б.* Интеграция нейронных и символично-логических моделей в интеллектуальных технологиях // Тр. 7-й Национальной конференции по искусственному интеллекту с международным участием КИИ-2000. – Переславль-Залесский, 2000. – т. 2, – с. 588–596.
20. *Buckley J., Hayashi Y.* Fuzzy neural networks // *L.A.Zadeh and R.R.Yager*, Eds. Fuzzy Sets, Neural Networks and Soft Computing. – N.Y.: Van Nostrand Reinhold, 1994. – pp. 233–249.

21. *Fuller R.* Neural fuzzy systems. – Abo, 1995.
22. *Kohonen T.* Self-organizing maps. – Berlin: Springer-Verlag, 1995.
23. *Kosko B.* Fuzzy cognitive maps // International Journal of Man-Machine Studies. – 1986. – v. 24.
24. *Le Blanc L.A., Hashemi R.R., Rucks C.T.* Pattern development for vessel accidents a of statistical and neural computing techniques // Expert Systems with Applications. – 2001. – No. 20. – pp. 163–171.
25. *Nechaev Yu.I., Degtiarev A.B., Boukhanovsky A.V.* Analysis of Extremal Situations and Ship Dynamics on Seaway in Intelligent System of Ship Safety Monitoring // Proceedings of 6th International Conference on Stability of Ships and Ocean Vehicles. – “STAB’97”, Varna. 22–27 Sep. 1997. – Vol. 1. – pp. 351–359.
26. *Nechaev Yu.I., Siek Yu.L.* Design of ship-board control system based on the soft computing conception // Proceedings of 11 th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA-98-AIE. Benicassim, Castellon, Spain, June 1998. Springer. – Vol. 11. – pp. 192–199.
27. *Nechaev Yu.I., Siek Yu.L.* Artificial intelligence and soft computing conception in ship-board adaptive control of underwater vehicle // Proc. 13th International Conference on Hydrodynamics in Ship Design. 2nd International Symposium on Ship Maneuvering. Gdansk–Ostroda, Poland, 1999. – pp. 350–361.
28. *Nechaev Yu.I., Siek Yu.L., Vasunin D.A.* Soft computing conception in a problem of safety conditions insuring in a stormy sea // Proc. 13th International Conference on Hydrodynamics in Ship Design. 2nd International Symposium on Ship Maneuvering. – Gdansk–Ostroda, Poland, 1999. – pp. 236–342.
29. *Pedrycz W.* Fuzzy neural networks with reference neurons as classifiers // IEEE Trans. Neural Networks. – 1992. – Vol. 3. – pp. 770–775.
30. *Pedrycz W., Rocha A.F.* Fuzzy-set based model of net and knowledge-based networks // IEEE Trans. Fuzzy Syst. – 1993. – Vol. 1. – pp. 254–266.
31. *Ripley B.D.* Pattern recognition and neural networks. Cambridge: Cambridge University Press, 1996.
32. *Wasserman P. D.* Combined backpropagation Cauchy machine. Proc. of Network Society. New York: Pergamon Press, 1988.
33. *Zadeh L.* Fuzzy logic, neural networks and soft computing // Commutation on the ACM. – 1994. – v. 37, No. 3, – pp. 77–84.

Нечаев Юрий Иванович, доктор технических наук, профессор кафедры вычислительной техники и информационных технологий Государственного морского технического университета, заведующий лабораторией систем поддержки принятия решений Института высокопроизводительных вычислений и баз данных. Область научных интересов — разработка интеллектуальных систем обеспечения безопасности мореплавания и посадки летательных аппаратов морского базирования. Научная школа — подготовил 7 докторов технических и физико-математических наук, 15 кандидатов наук. Автор более 370 научных работ и изобретений (монографии, учебники, справочники, научные статьи), в том числе 3 монографии и 132 научные публикации изданы за границей.

НАУЧНАЯ СЕССИЯ МИФИ–2002

НЕЙРОИНФОРМАТИКА–2002

**IV ВСЕРОССИЙСКАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ
КОНФЕРЕНЦИЯ**

**ЛЕКЦИИ
ПО НЕЙРОИНФОРМАТИКЕ
Часть 1**

Оригинал-макет подготовлен Ю. В. Тюменцевым

ЛР №020676 от 09.12.97 г.

Подписано в печать 10.12.2001 г. Формат 60 × 84 1/16

Печ. л. 10,25. Тираж 250 экз. Заказ №

*Московский государственный инженерно-физический институт
(технический университет)*

Типография МИФИ

115409, Москва, Каширское шоссе, 31