

ФИЗИКА НАШИХ ДНЕЙ

Робот и Человек. Где находится предел их сходства?

Г.Р. Иваницкий

Показано, что основная цель робота и человека — обеспечить динамическую стабильность в условиях изменения внешней среды. Человеческий мозг по сравнению с компьютерным "мозгом" робота имеет многоуровневую иерархическую организацию, в которой обработка информации происходит на всех уровнях — от квантового до социального. Человек сам ставит цель, совершенствуя виртуальную модель, которая синтезируется его мозгом. Мозг человека может работать на основе одновременного использования как классической детерминированной логики, так и диалектической вероятностной логики.

Ключевые слова: андроидные роботы, адаптация, креативность, память, иерархия организации, классическая и вероятностная логика

PACS numbers: 87.19.L-, 87.85.St, 89.20.Ff

DOI: <https://doi.org/10.3367/UFNr.2018.03.038302>

Содержание

1. Введение (965).

1.1. Двойственность отношения к робототехнике. 1.2. Виды роботов и прогноз их развития. 1.3. Основные отличия роботов от человека в настоящее время.

2. Проблема креативности (968).

2.1. Пирамида иерархических уровней в организации живых систем. 2.2. Диалоги на разных уровнях иерархической организации. 2.3. Диалог сердца с мозгом. 2.4. Простая модель достижения консенсуса. 2.5. Примеры парадоксов взаимодействия сенсоров. 2.6. Виртуальная модель, синтезируемая нашим мозгом. 2.7. Сенсоры тоже могут ошибаться. Эволюционное предвращение появления ложных образов.

3. Адаптация робота и человека к окружающей среде (978).

3.1. XX век: от кибернетики к синергетике. 3.2. Синергетика — теории кооперативных динамических взаимодействий дисциплинарных систем.

4. XXI век: новая интерпретация старых идей (981).

4.1. Время дорогих проектов. 4.2. Классическая и вероятностная логика. 4.3. Мозг, живущий в вероятностном мире.

5. Заключение (986).

5.1. Математическое сжатие информации. 5.2. Выход за границы обычного пространства в новое информационное пространство. 5.3. Предел сходства Робота и Человека.

6. Выводы (989).

Список литературы (989).

1. Введение

1.1. Двойственность отношения к робототехнике

Робототехника сегодня развивается с экспоненциальным ускорением. Нас будет интересовать в основном *искусственный интеллект (ИИ) креативного андроидного робота¹ (КАР)* в сравнении с *интеллектом человека и предел их сходства*. Разработчики КАР надеются, что именно этот класс машин станет высшим научно-технологическим достижением человечества. В человеческом обществе отношение к этим разработкам двойственное.

С одной стороны, интеллектуальные роботы смогут выполнять любые работы намного дешевле, точнее и быстрее людей и, следовательно, вытеснят последних с рынка труда. При этом, возможно, социальная система человечества изменится так, что нарушится её устойчивость. Нельзя также исключить желание оборонных ведомств и органов правопорядка доверить КАР силовые действия. Всё чаще возникают идеи о гибридных системах человеческий мозг + ИИ. Компания Neuralink разрабатывает интерфейсы, которые объединяют человеческий мозг и компьютер. Эффективно работающая система мозг + компьютер породит касту "суперполицейских" и "супервоинов". При реализации такого проекта главным риском станет сбой программного обеспечения. Хотя современные программы и компьютеры обычно надёжнее людей и при дублировании не ошибаются, тем не менее вероятность сбоя существует. Нельзя исключить атаки хакеров и любые другие случайные или намеренные сбои в работе ИИ. В связи с

Г.Р. Иваницкий. Институт теоретической и экспериментальной биофизики РАН,
ул. Институтская 3, 142290 Пущино, Московская обл.,
Российская Федерация
E-mail: ivanitsky@iteb.ru

Статья поступила 17 октября 2017 г.,
после доработки 14 февраля 2018 г.

¹ Креативность (от англ. *create* — создавать, *creative* — созидательный, творческий) — умение самостоятельно принимать творческие решения для достижения цели. *Андроид* (от греч. слова ἄνθρ — человек и суффикса -oid ("подобие") — человекоподобный) — робот-гуманоид, или синтетический организм, предназначенный для того, чтобы действовать подобно человеку.

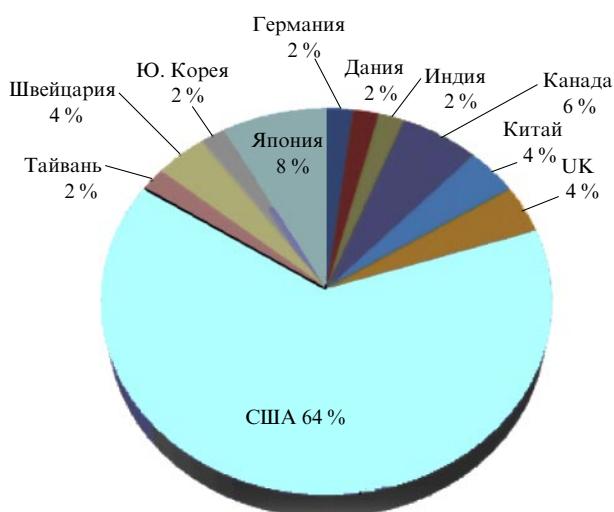


Рис. 1. Диаграмма, составленная по итогам 2015 г. редакцией сайта roboticsbusinessreview.com. Включено 50 компаний из 10 стран, специализирующихся в области робототехники.

быстро действием КАР у оператора в случае их сбоя на ликвидацию последствий времени не будет, что может привести к трагической ситуации. Наборы рисков при создании креативного ИИ часто сравнивают с содержимым ящика Пандоры.

С другой стороны, будущее формируется сегодня. Логично предположить, что программы, заложенные в КАР, будут находиться под контролем людей и взаимодействовать с нами без злого умысла, для пользы всего человечества. Хотя прогнозы составляются, но никто достоверно не знает, каким будет мир, например, в 2040 г. и тем более в 2100 г. Невозможно предсказать отношения между людьми и новыми поколениями суперкомпьютеров, реализующих ИИ и КАР. Невозможно ограничить человеческую любознательность и остановить порождаемый ею научный прогресс. В прошлые годы делались попытки запретить отдельные направления исследований (например, искусственное оплодотворение человеческих яйцеклеток вне организма или манипуляции с генами). Запреты тормозят, но не останавливают развитие науки [1]. Работы по робототехнике проводятся во всех развитых странах (рис. 1).

Развитие робототехники происходит настолько быстро, что сегодня эта диаграмма уже устарела. В последние три-четыре года КНР сделала существенный скачок в этой области. В России редкий технический университет не принимает участия в отечественной программе робототехники. Сюда следует добавить пять академических институтов и семь научно-производственных объединений. Оценка роста рынка робототехники была проведена в 2016 г. консалтинговой компанией Tractica. Рост рынка происходит по экспоненте (рис. 2).

1.2. Виды роботов и прогноз их развития

В 2017 г. в Интернете демонстрировался танец, исполняемый одновременно 1069 роботами Dobi производства компании WL Tech. Это событие происходило в городе Гуанчжоу на юге Китая. Как передаёт Mashable.com, роботами управляли через единую систему. В этом шоу был побит мировой рекорд количества одновременно танцующих роботов. (Предыдущий рекорд — совмест-

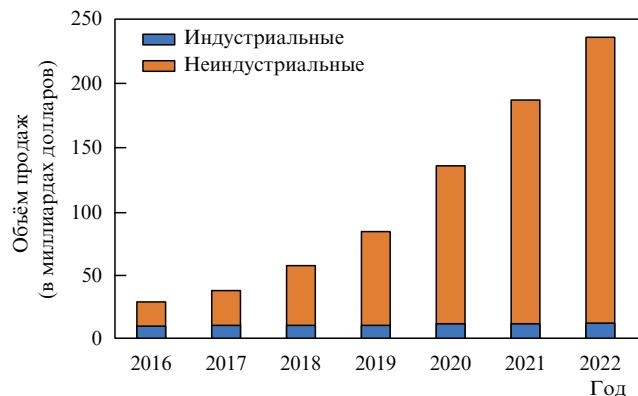


Рис. 2. Предполагаемый рост объема продаж роботов к 2022 г. оценивается в 237,3 млрд долларов. Он вырастет к 2022 г. в пять раз, в основном за счёт покупок роботов неиндустриального назначения (Tractica: Total Industrial and Non-Industrial Robotics Revenue, World Markets (2016–2022 гг.)).

ное исполнение танца 1007 роботами.) Развлекательные роботизированные шоу начинают приобретать массовый характер, подстёгиваемый амбициями разработчиков в борьбе за мировые рекорды. Роботы Dobi способны не только танцевать, но и петь, они также могут изображать приёмы бокса или кунг-фу.

Класс роботов, управляемых оператором, представляет собой не более чем машины, дистанционно управляемые командами с пульта. С таким дистанционным управлением знакомы даже дети, имеющие радиоуправляемые игрушки.

Особое место занимают экзоскелеты, которые соединяются с подвижными частями тела человека и копируют их движение, повышая его мощность и точность. Экзоскелеты активно внедряются в медицину для помощи людям с патологиями пространственного передвижения. Они используются также для дистанционной манипуляции с предметами в агрессивных средах и в работах, требующих высокой точности микродвижений, например в микрохирургии.

При этом для управления можно использовать электрические потенциалы, которые идут из центральной нервной системы к мышцам человека, но при этом исключать трепет в движениях. Потенциалы отводятся с двух нервных окончаний мышц-антагонистов — сгибателей и разгибателей, например, ног, кисти руки или пальцев. В нашей стране ещё в 1960–1970-х годах В.С. Гурфинкелем и А. Кобринским с соавторами была создана управляемая биотоками "искусственная рука" [2, 3]. Е.Б. Бабским был разработан вживляемый в грудную полость пациента стимулятор сердечной деятельности, управляемый ритмом сердечных сокращений с помощью усилителя потенциалов предсердия [4, 5].

Сегодня успешно решаются задачи имитации роботами движений, распознавания лиц в толпе людей или воспроизведения голоса и мимики конкретного человека, а также сравнительно легко осуществляются программируемые диалоги: детерминируемый вопрос ↔ детерминируемый ответ. Роботы имитируют, например, поведение: 1) секретаря, встречающего посетителей; 2) официанта, принимающего заказы; 3) экскурсовода, рассказывающего об экспозиции выставки; 4) сиделок медицинских учреждений или домов престарелых; 5) лек-

тора, отвечающего на вопросы слушателей с заранее запрограммированными ответами на предполагаемые вопросы, и, наконец, (6) продавца товаров [6].

Кроме того, постепенно приобретают популярность шлемы виртуальной реальности (*virtual reality*, VR) и шлемы, смешивающие окружающую реальность с виртуальной. Они используют зрительный и слуховой рецепторы человека, но не в реальном, а в виртуальном, искусственном, мире. Виртуальный мир создаётся программой компьютера. Создание виртуальной реальности уже находит применение при производстве тренажёров для обучения космонавтов, пилотов, диспетчеров, хирургов и др., а также в области дизайна и архитектуры [7].

Особый класс роботов составляют нанороботы и микrorоботы, которые запускаются внутрь организма для медицинской диагностики или микрохирургии, а также для адресной доставки биологически активных соединений. Уже существуют микрокапсулы, нано-платформы с биологически активным содержимым или так называемая умная пыль (англ. *smartdust*), обменивающаяся информацией и объединяющаяся в стаи с целью решения совместных задач [8]. Другой класс наносистем — это наномашины [9], например, весьма опасная саморазмножающаяся "серая слизь" (англ. *grey goo*) [10].

Ограничимся одним примером из нашего опыта. Попытки создать "микро- и наномашинную технологию" делались давно, более 40 лет тому назад [11]. Например, нами были реализованы простые варианты таких технологий при создании газотранспортных плазмозаменителей крови в 1980-х годах. Их идея основана на том, что частицы из комбинации перфторуглеродов размером 30–70 нм переносят кислород из оставшихся после кровопотери эритроцитов в ткань. Эритроцит в артериальной крови, размер которого более чем в 100 раз превосходит размер частиц эмульсии, выполняет функцию "матки-базы", нагруженной кислородом. Частицы из перфторуглеродов периодически циркулируют за счёт пульсирующих гидродинамических потоков плазмы крови между эритроцитами и тканью. Они образуют линейные структуры типа "жемчужных нитей", эстафетно передавая по закону Фика кислород от эритроцита к стенкам кровеносного сосуда, откуда он сам далее диффундирует в ткань [12].

Что касается *креативных андроидных роботов*, то будущая область их использования — это работа в агрессивной среде или при чрезвычайных ситуациях, включая дальние космические полёты. Надеются, что КАР не только будет превосходить человека по быстродействию, точности выполнения различных операций и силовым показателям, но и не уступать ему по когнитивным возможностям. Хотя самостоятельность формирования цели у КАР должна быть ограничена. В противном случае могут возникнуть проблемы при взаимодействии с человеком (см. разделы 5.2 и 5.3).

В этом классе роботов много нерешённых проблем. Создание таких роботов требует развития совместных физико-технических и нейрофизиологических исследований. Все процессы, для которых можно написать алгоритм, могут быть реализованы в системах ИИ. Однако возникают вопросы: есть ли такие процессы, которые не поддаются математическому описанию и алгоритмизации? Имеется ли предел сходства КАР с человеком?

Кроме того, существует и другой научный интерес: мы не сможем эффективно бороться с нейродегенеративными заболеваниями человека, не поняв в полной мере, как работает наш мозг и как мыслит человек. Понять — это значит попытаться создать ИИ и КАР с интеллектом, похожим на человеческий, или сформулировать ограничения в достижении их похожести.

Из медицинской практики известно, что шизофрения — это переход мозга человека в хаотический режим работы. Существует термин *апофения* (введённый в 1958 г. немецким психиатром Клаузом Конрадом) от древнегреческого слова *ἀποφάνω* — высказываю суждение, делаю явным. Апофения — это переживание, заключающееся в способности долго находиться в мире собственных вымыслов. Такое состояние можно классифицировать как начальный этап шизофрении. Мозг человека при поиске решений, возникающих в сложных ситуациях, всегда находится на границе порядка и хаоса или, что то же самое, на границе использования классической и вероятностной логики. Если мозг, войдя в стохастический режим, не может выйти из такого состояния, то процессы мышления нарушаются и появляются аномальные эмоциональные реакции.

В журнале *Успехи физических наук* на темы, связанные с нейрофизиологией, было опубликовано много обзоров [13–22], но настоящий обзор отличается от них тем, что в нём акцент делается на логику процессов, которые происходят в мозге человека и ИИ креативного робота.

Прогнозы — дело неблагодарное, поскольку вследствие нелинейности систем они сбываются не часто. Тем не менее приведём один из их вариантов (табл. 1), подготовленный в 2016 г. российской компанией МИВАР (Многомерная информационная варьирующаяся адаптивная реальность) на основе различных предположений, высказанных в разные годы экспертами (robot-rends.ru).

1.3. Основные отличия роботов от человека в настоящее время

Очевидно, что созданные сегодня роботы — это электромеханические машины. Живые организмы — это физико-химические "машины" на водной основе. Живой мозг и его функции развивались эволюционно *на основе блочно-иерархического отбора* [23, 24] в неразрывной связи с развитием тела организма и изменениями внешней среды на нашей планете. В этом процессе особое место отводится внутренней среде организма, в частности важной роли сердца при взаимодействии с мозгом. Этот факт часто забывают при сравнении ИИ робота с интеллектом человека. Сердце создаёт заметные акустические и электромагнитные поля в нашем организме, топография которых на поверхности тела легко регистрируется с помощью современных методов пространственной кардиологической аускультации (лат. *auscultatio* — выслушивание) и электрокардиографии.

Как известно, акустические волны представляют собой продольные упругие колебания давления в газах, жидкостях и твёрдых телах. В твёрдых и жидких средах каждая частица среды может колебаться около точки равновесия (стоячие волны). Однако следует отметить, что одиночные колебания в нелинейных средах могут создавать одиночную волну, которая ведёт себя подобно частице. В этом случае может возникнуть особый вид

Таблица 1. Перспективы развития робототехники до 2040 года

Год	Предполагаемый результат	Источник с годом прогноза
2018	Робот сдаст экзамен на права вождения автомобилем	Олег Варламов, президент МИВАР, 2016
2019	90 % организаций будут иметь в своём штате такую должность, как директор по управлению данными (Chief Data Officer — CDO)	Cortner, 2016
2020	В США будут эксплуатироваться 30 тыс. беспилотных гражданских летательных аппаратов	Федеральное управление гражданской авиации США, 2012
2022	Роботы научатся понимать поведение человека и реагировать на него	Экспертно-аналитический отчёт Росбизнесконсалтинга (РБК), 2014
2024	Появятся коммерчески доступные автомобили, способные реагировать на изменение ситуации на дороге и передвигаться автономно	IHS Automotive, 2014
2025	Использование роботов в промышленности уменьшит затраты на оплату труда на 16 %	Boston Consulting Group, 2015
2028	Появятся первые автономные медицинские микророботы, способные самостоятельно и направлению перемещаться в теле пациента	"A Roadmap for US Robotics: From Internet to Robotics", 2013
2029	ИИ будет способен к самообучению, пониманию шуток и имитации эмоций	Рэй Курцвейл, технический директор Google, 2014
2030	Появятся коммерчески доступные андроиды, внешний вид и способности которых будут идентичны человеческим	Экспертный семинар "Тренды и перспективы развития отрасли робототехники в России", 2014
2032	Роботы превзойдут людей по интеллектуальным возможностям	Дэйв Эванс, футуролог компании Cisco, 2011
2035	В Японии роботы освоят 49 % специальностей из 600 существующих	Nomura Research Institute, 2015
2040	Роботы будут использоваться для поддержания правопорядка в городах	Профессор Ноэл Шорки, Университет Шеффилда, 2012

волн — *солитоны*, которые проходят сквозь друг друга и распространяются на большие расстояния [25, 26]. Солитоны отличаются от гармонических волн (см. подробнее в разделе 2).

В основе создания роботов в сравнении с человеком пока имеется множество различий как в способах адаптации к внешней среде, так и в интеллектуальной деятельности (табл. 2).

Данные, приведённые в табл. 2, соответствуют сегодняшней ситуации. В неживых системах по отдельности встречаются все признаки, которые характерны для живых систем [23]. Однако живые системы, в отличие от неживых, адаптируются к изменениям внешней среды. Главное требование к адаптации — это динамическая устойчивость системы, находящейся в неравновесном состоянии [27]. *Обработка поступающей извне информации происходит на всех иерархических уровнях организма.*

2. Проблема креативности

2.1. Пирамида иерархических уровней в организации живых систем

В живых системах можно отметить по крайней мере девять иерархических уровней (рис. 3).

Задача построения математической модели иерархической системы состоит в объединении всех уровней в единое целое, но она пока не решена. Казалось бы, при девяти уровнях не так сложно определить результаты взаимодействия между уровнями. При этом количество связей "всех уровней на всех" будет небольшим, соответствуя выражению

$$Q = m(m - 1), \quad (1)$$

где Q — число связей, m — число уровней. Таким образом, при $m = 9$ получим всего $Q = 72$. Однако в действительности всё намного сложнее, поскольку на каждом уровне количество взаимодействующих элементов, объединённых в сети, чрезвычайно велико.

Попытки объединить уровни в единую систему делались многократно. Например, ещё в 1970-х годах Манфред Эйген (Manfred Eigen) выдвинул идею "гиперцикла", сосредоточившись на анализе *макромолекул* [28, 29]. Суть идеи Эйгена заключается в следующем (рис. 3б): в основе размножения элементов живых систем лежит матричный способ репликации ДНК. Обозначим этот цикл буквой I с индексом. Таких микроциклов много, от 1 до n .

Внутри каждого цикла идёт автономная репликация, т.е. перезапись с матрицы на матрицу: негатив — позитив, позитив — негатив и т.д. Для того чтобы эти микроциклы были взаимно регулируемыми, М. Эйген предложил ввести объединяющий их гиперцикл, что, по его убеждению, способствовало бы процессу увеличения сложности организации в далёкой от равновесия химической системе с образованием новых микроциклов с многочисленными петлями обратной связи. Каждый микроцикл, входящий в состав гиперцикла, одновременно с репликацией производит набор факторов связи (ферментов E_i). Эти факторы могут либо селективно повышать скорость и точность репликации, либо уменьшать скорость распада уже синтезированных матриц. Это означает, что любой i -й микроцикл в гиперцикле независимо от того, сколько и чего он кодирует и перезаписывает в режиме репликации, всегда должен зависеть от фактора связи E_{i-1} , а сам должен содержать фактор связи для соседа E_{i+1} . М. Эйген считал, что каждый i -й микроцикл служит координированно регулируемой генетической единицей, а все вместе они работают последовательно по гиперциклу.

Таблица 2. Сравнение основных характеристик робота и человека

№	Основные характеристики	Робот	Человек	Превосходство*
Обобщённые характеристики				
1	Принцип организации	Электромеханическая машина	Физико-химическая "машина" на водной основе	—
2	Силовые усилия	Практически неограниченны	Ограничены	P
3	Утомляемость	Отсутствует	Имеет место	P
4	Коэффициент полезного действия	60–90 %	15–20 %	P
5	Самостоятельное формирование цели	Отсутствует	Имеет место	Ч
6	Социальная фасилитация**	Возможна	Имеет место	Ч
7	Самообъединение в группы	В ряде случаев достигнуто, пример — формирование стай дронов (беспилотников)	Имеет место	Ч
8	Самовоспроизведение	Возможно	Имет место	Ч
9	Количество иерархических уровней, обрабатывающих информацию, поступающую из внешней среды	Пока ограничено и меньше, чем у человека	Ограничено, но больше, чем у современных роботов	Ч
Сравнение характеристик "искусственного мозга" робота и мозга человека				
10	Характеристическое время формирования реакции	Менее 1 мкс	Порядка 0,1 с	P
11	Принцип работы	Дискретный	Дискретно-аналоговый	—
12	Подложка под логическими элементами	Твёрдая, пассивная среда	Жидкая, активная среда	—
13	Пластичность	Низкая	Высокая	Ч
14	Поведение подложки	Стабильное	Пульсирующее	—
15	Термостабилизация	Распределённая, в основном воздушное охлаждение	Охлаждение локализованное, жидкое	Ч
16	Взаимодействие с внутренней средой системы	Сверху вниз	Иерархическое и циклическое с обратными связями сверху вниз и снизу вверх	Ч
17	Влияние шума	Помеха в работе	Шум — созидающий фактор, обеспечивающий переходы между состояниями	Ч
18	Несимметричные ответы на изменение внешней среды	Отсутствуют, не способны к творчеству	Творчество имеет место	Ч
19	Логика работы	Детерминированная (классическая логика)	На границе классической и вероятностной логики	Ч

* Обозначения: Р — робот, Ч — человек.

** *Фасилитация* (от англ. *facilitate* — помогать, способствовать) — управление, основанное на поиске консенсуса при объединении элементов системы, что повышает устойчивость системы как целого. Процесс фасилитации происходит на основе сочетания симбиотического и конкурентного дуализма (обычно в философии такой процесс именуется как "единство и борьба противоположностей").

Другими словами, М. Эйген, описав феноменологию биологического явления на уровне макромолекул, все сложности объяснения его развития и функционирования обошёл, введя своеобразное "управляющее начало" в виде факторов связей E_i , которые "всё умеют", а сами усложняются. Но на вопрос: какова кинетика и какие движущие силы такого самоусложнения? — он не ответил. Таким образом, несмотря на удачную идею циклов, объяснение самого процесса образования гиперцикла осталось неопределенным.

По сути, в основе гиперцикла М. Эйгена лежит проекция расположенных выше уровней пирамиды (рис. 3а) на уровень макромолекул. При такой редукции

специфика кинетики процессов на остальных иерархических уровнях оказалась потерянной.

Другая идея принадлежит М.И. Рабиновичу с соавторами [14, 30]. Как будет ясно из дальнейшего изложения, она близка к идее, рассматриваемой в данном обзоре, а именно представляет собой попытку создать математическую модель сознания на основе теории колебаний в терминах нелинейных уравнений. Однако в ансамбле большого числа степеней свободы при объяснении спонтанно возникающего порядка из хаоса существенную роль играет *удачный случай*, который в подходе М.И. Рабиновича отсутствует. Удачный случай в психологии часто называют *озарением* (см. притчу о смышлён-

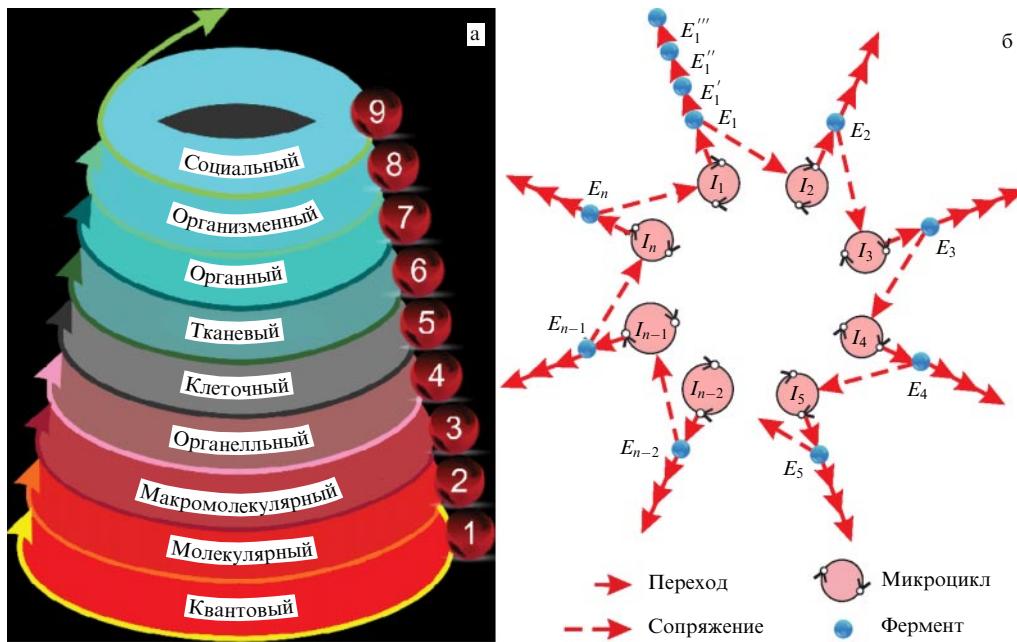


Рис. 3. Иерархическая организация живых систем: (а) уровни организации, (б) гиперцикл М. Эйгена.

ном ребёнке в разделе 5.2). При этом *случайность* является одним из существенных механизмов *возникновения новых межнейронных связей*, а нелинейность включается позднее, предотвращая неограниченный рост. При этом нервные импульсы, возникающие при взаимодействии с внешней средой, имеют своеобразный вид пиков, превышающих уровень шумов [31]. Промежутки между импульсами характеризуются большой протяжённостью. Как известно, общее название такой картины в физике — "перемежаемость" [32]. Отметим, что задача о возникновении перемежаемости в некотором смысле напоминает задачу о возникновении возврата из детерминированного хаоса в упорядоченное движение в динамической системе [33].

Требований к любой предлагаемой модели только два: внутренняя непротиворечивость и возможность предсказания новых режимов, подтверждаемых в результате экспериментальной проверки. Очевидно, что по этой причине для описания сложной системы можно предложить множество разных моделей, отвечающих этим двум требованиям (см. раздел 3).

Р. Пенроуз в книге [34, с.14] писал: «*В формировании нашего сознания с необходимостью есть элементы, которые не могут быть получены из какого бы то ни было набора вычислительных инструкций... "Неалгоритмическим действиям" нельзя найти место в рамках общепринятых сегодня физических теорий. А значит, мы должны искать соответствующее место, где в научной картине существует серьёзный пробел. И я утверждаю, что это "белое пятно" лежит где-то на границе между "субмикроскопическим" миром, в котором правит квантовая механика, и непосредственно воспринимаемым нами макромиром, подчиняющимся законам классической механики».*

Отдавая должное мысли Пенроуза, следует заметить, что есть причины для существования различия как в обработке информации в мозге человека и компьютере робота, так и в виде поступающей энергии для реализа-

ции этих процессов. Энергия, на которой работает робот, не отличается разнообразием — это электроэнергия, а энергия мозга человека (как и его организма в целом) включена в цикл преобразования энергии на нашей планете: Солнце → Земля с её атмосферой ↔ флора ↔ фауна. Человек включен в пищевую цепь биосфера, и он не может существовать автономно.

Создав микроскопы, телескопы и ускорители, быстрые средства передвижения и системы внешней памяти (языки, письменность, различные виды искусства и Интернет), человек существенно расширил размерность наблюдаемого им мира. Но количество естественных иерархических уровней в самой системе его тело + мозг, где происходит обработка поступающей из внешнего мира информации, практически не изменилось. В живых системах можно отметить наличие квантового уровня — это светочувствительные белки — хлорофилл и родопсины, уровня ионов, молекул и макромолекул. К организационно более высоким уровням следует добавить уровни органелл, клеток, тканей (сетей), органов, целостного организма и, наконец, социальный уровень — взаимодействия индивидов между собой на основе акустического языка, языка запахов или языка жестов и мимики, воспринимаемых слухом, обонянием или зрением.

По-видимому, главная причина разнообразия уровней в живых системах состоит в том, что окружающий мир существует в разных обличиях: от абсолютно детерминированного, отвечающего законам Ньютона, до абсолютно хаотического, вероятностного мира (мира большого бильярда) Эйнштейна–Смолуховского [35, 36] или являющегося смесью и того и другого (хаоса или перемежаемости, детерминированности или стохастичности) [32]. Для ориентации в условиях постоянных метаморфоз внешней среды нужна комплементарная ей многоуровневая организация обработки информации, которая быстро адаптируется к внешним условиям, обеспечивая тем самым устойчивость живой системы. Каждое действие, производимое организмом, имеет

свою цену H . Цена действия есть функция

$$H = f(E, \tau), \quad (2)$$

где E — потребляемая энергия, τ — время действия.

В стрессовых ситуациях скачком увеличивается величина затрачиваемой энергии, а время действия τ_i каждого этапа уменьшается и, следовательно, частота Ω_i выполнения всех действий при переходе от уровня к уровню снизу вверх увеличивается. Если имеется N иерархических уровней, каждый из которых может вносить свою лепту в затрату энергии, то суммарные затраты энергии возрастают и достигают максимума. В результате получим выражение

$$E_{\max} = \sum_{i=1}^N H_i \Omega_i, \quad (3)$$

где Ω — частота действия. Это выражение соответствует ситуациям в условиях сильного стресса. Сумма затрат энергии в выражении (3) определяется i -м уровнем с самой большой величиной слагаемого, т.е. нижними уровнями, работающими с высокой частотой (квантовым, молекулярным и макромолекулярным) (рис. 3а).

В нормальных слабострессовых ситуациях, связанных, например, с обучением, минимизируется суммарная цена действия. Все N уровней работают в неполную силу, что соответствует выражению

$$H_{\min} = \sum_{i=1}^N \frac{E_i}{\Omega_i}. \quad (4)$$

При этом в суммарную цену действия H наибольший вклад вносят уровни, работающие с малой частотой: социальный (общественное мнение и поощрение), организменный (комфорт индивида), органный (комфорт внутренней среды при взаимодействии между органами).

Следовательно, устойчивость системы определяется обратной связью переходов от верхних уровней к нижним и обратно. При этом *ускорение изменения внешней среды под действием различных факторов, включая антропогенные, не должно превышать ускорения адаптации живых систем к этим изменениям*. В противном случае живая система не будет успевать адаптироваться к внешней среде, начнёт деградировать и погибнет.

2.2. Диалоги на разных уровнях

иерархической организации

Жизнь человека — это переходный процесс с определённой длительностью и непрерывными разрывками, требующими выбора направления движения во времени и пространстве с оценками рисков выбора. Риск возникает и растёт в условиях неопределенности. Это ограничивает точность нахождения вероятности как позитивного, так и негативного исходов. Например, в системах хищник — жертва неправильная оценка этой вероятности грозит одному из участников гибелью. Восприятие риска живым организмом связано со страхом, генетически заложенным в виде инстинкта, порождающим бессознательный, а с учётом мозга частично или полностью осознаваемый выбор: драться или бежать? Соглашаться или возражать?

В реальном организме, включая мозг, существует дуализм, выражющийся в том, что между различными

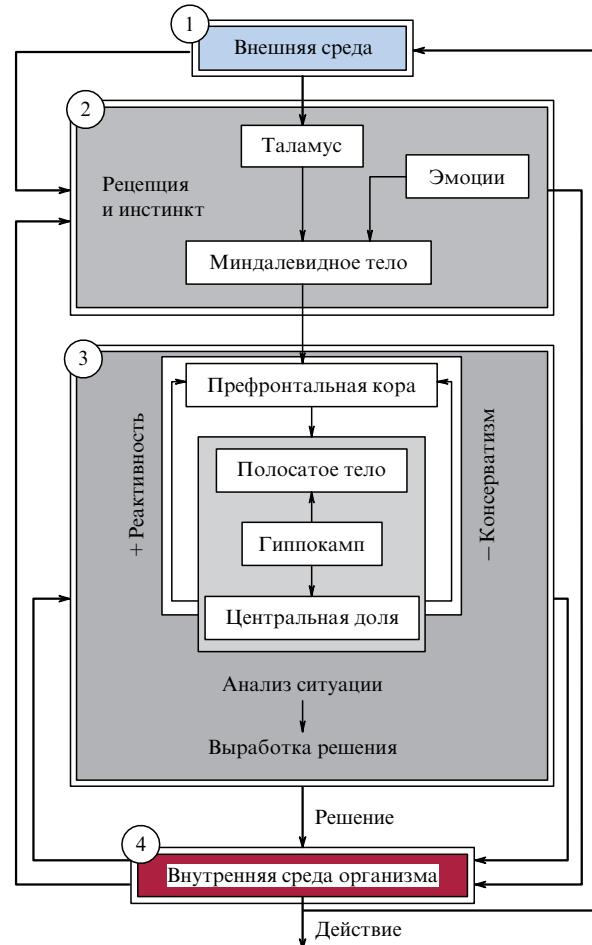


Рис. 4. Пример упрощённой схемы путей на уровне структур мозга с прямыми и обратными связями при принятии решений человеком. На любое решение, принимаемое человеком, влияют как эмоции, так и результат конкуренции между структурами: полосатое тело и центральной зоной мозга. У людей, которые склонны к рискам и быстро принимают решения, более активно полосатое тело. У тех, кто консервативен в принятии решений, активнее работает центральная доля. Компромисс между ними с учётом влияния внешней среды и прошлого опыта (памяти, регулируемой гиппокампом) создаёт устойчивое состояние и вырабатывает сигнал к действию.

его частями (клетками и органами) и внешней средой одновременно имеют место как симбиоз (содружество), так и конкуренция (антагонизм), основанные на положительных и отрицательных обратных связях. Понятие "обратная афферентация" (от лат. слова в род. падеже *afferentis* — приносящий), по-видимому, было введено в научный обиход П.К. Анохиным ещё в начале 1950-х годов [37], а синоним этой связи в кибернетике сохранил старое техническое название — "обратная связь" (feedback). Напомним, что в диалектике это положение со времён Гегеля приобрело форму закона "единства и борьбы противоположностей", а результатом такой борьбы является "снятие противоречия" [38]. Все системы живого организма, включая различные области мозга, пронизаны обратными связями, образующими циклы, что позволяет снимать противоречия, удерживая их конкуренцию в заданных границах. Один из примеров приведён на рис. 4.

Однако подобная конкуренция, спор или диалог (называть этот процесс можно по-разному), на различ-

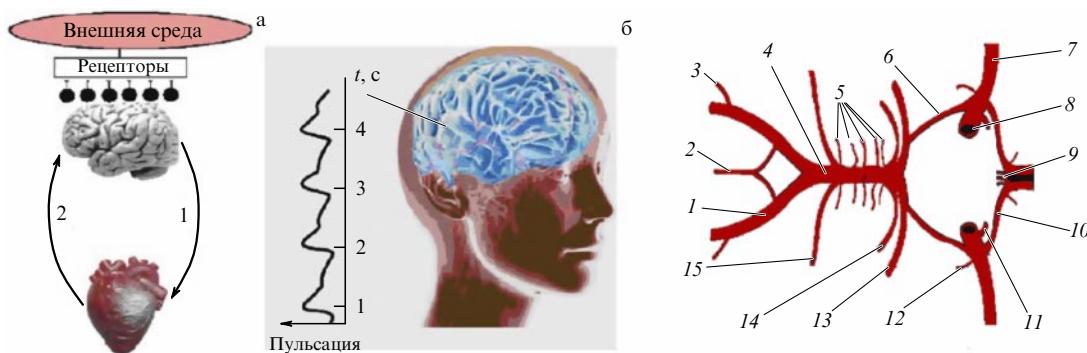


Рис. 5. Взаимодействие сердца с мозгом в объёме спинномозговой жидкости (ликвор). (а) Схема взаимодействия. (б) Пульсация объёма спинномозговой жидкости во времени за счёт сокращения сердца. (в) Система артериального снабжения мозга энергией: 1 — позвоночная артерия, 2 — передняя спинномозговая артерия, 3 — задняя нижняя мозжечковая артерия, 4 — базальная артерия, 5 — мосты артерии, 6 — задняя соединительная артерия, 7 — средняя мозговая артерия, 8 — внутренняя сонная артерия, 9 — передняя соединительная артерия, 10 — передняя мозговая артерия, 11 — главная артерия, 12 — передняя хориоидальная артерия, 13 — задняя мозговая артерия, 14 — верхняя мозжечковая артерия, 15 — нижняя мозжечковая артерия. Вены не указаны.

ных уровнях организации живых систем осуществляется на разных языках: биохимическом, электрическом, акустическом. На биохимическом молекулярном уровне хорошо изучен антагонизм взаимодействия между гормонами, например гормонами страха: адреналином и норадреналином — в соответствии с их воздействием на организм первый иногда называют "гормоном кролика", а второй — "гормоном льва". Цель этих диалогов — достижение согласия.

Во-первых, достижение согласия на каждом уровне — это гарантia устойчивости организма в целом. Во-вторых, при этом происходит сжатие информации, что в случае успеха облегчает её запоминание для использования в будущем. Кроме того, экономится объём памяти при построении виртуальной модели внешней среды внутри мозга. В результате в живых организмах информация, перемещаясь по уровням, претерпевает изменения, например, за счёт образования связей верхних и нижних уровней переходит от головного мозга к спинному, приобретая функцию навыков движения и освобождая тем самым верхние уровни от рутинной работы [39]. Это позволяет также в течение сравнительно короткого интервала времени дополнительно экономить энергию и понижать энтропию. В социальных системах поведение индивида благодаря перемещению информации по уровням системы может изменяться. Например, индивид адаптируется к поведению толпы [40], а животное — к поведению стаи или стада [41–43].

2.3. Диалог сердца с мозгом

В качестве примера на уровне органов приведём диалог сердца с мозгом. Пульсирующее сердце "разговаривает" с мозгом на языке акустики, испуская во внутримозговую жидкость (включая спинномозговую жидкость, содержащую гормоны и энергию — сахара и окислители) инфразвуки со средней частотой ~ 1 Гц. Частота может изменяться в пределах от 0,75 до 2,5 Гц. При этом жидкость перемещается и перемешивается. На дальность распространения звука влияют поглощение неоднородностями среды, рефракция (искривление звуковых лучей в неоднородной среде) и рассеяние звуковых лучей. Кроме того, рефракция выражается тем сильнее, чем больше градиент скорости звука. С понижением час-

тоты звуковых колебаний дальность их распространения увеличивается и могут возникать солитоноподобные движения [25, 26]. Звуковые волны, отражаясь от неоднородностей и слившись вместе, растягивают задний фронт звукового импульса, увеличивая продолжительность его воздействия.

Распространение гормонов в различных частях мозга создаёт "эмоциональную окраску" восприятия окружающей среды. Мозг отвечает сердцу на языке электрических сигналов, распространяющихся по аксонам и дендритам, которые в свою очередь изменяют локальную плотность жидкости на величину до 70 %. Изменение плотности происходит за счёт выброса ионов калия при возбуждении нейрона. Ионы калия имеют большую массу по сравнению с массой ионов натрия, поэтому локальная плотность среды снаружи аксона увеличивается. При этом возрастают и отражение звуковых волн, и скорость прошедших через уплотнение звуковых волн.

Кроме того, мозг регулирует частоту сокращений сердца по цепи обратной связи 1 (рис. 5а), подавая сигналы в синусовый узел ведущего центра сердца по цепи 1. В свою очередь по пути 2 сердце воздействует на мозг (рис. 5б).

В воде звуки могут распространяться с большой скоростью, порядка $1300\text{--}1500\text{ м с}^{-1}$. По сравнению со скоростью распространения звука скорость распространения импульса возбуждения по аксонам нейронов невелика — порядка 25 м с^{-1} .

Даже в состоянии покоя через головной мозг проходит около 15 % объёма крови, при этом мозг потребляет до 20–25 % кислорода, получаемого при дыхании [44]. Кровоснабжение головного мозга осуществляется двумя внутренними сонными и двумя позвоночными артериями. При этом задняя и передняя артерии образуют круг (рис. 5в). Отток крови происходит по двум ярёмым венам.

Казалось бы, роль сердца во взаимодействии с мозгом довольно скромна — поставлять питание и вывозить отходы метаболизма. Однако это не так. Одновременно с этим сердце совместно с глиальной сетью мозга перераспределяет потоки крови как на макроуровне, так и внутри каждого фрагмента нейронной сети. Это перераспределение адаптирует организм к

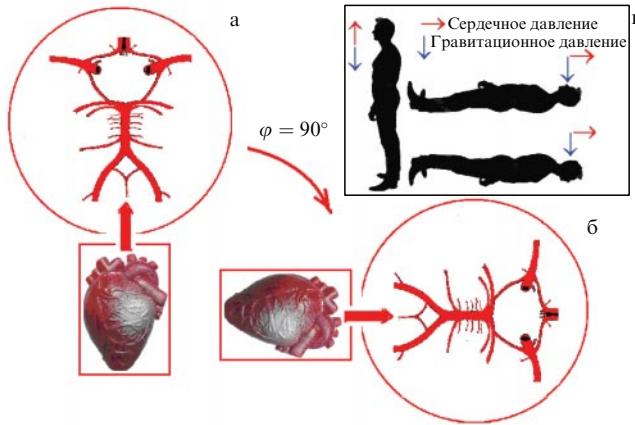


Рис. 6. Изменение позы человека меняет относительное расположение сердца и кровеносной сети мозга (а, б) по отношению к вектору гравитационного давления (в). Мозг может работать лишь при сравнительно постоянной разности между внутричерепными давлениями и температурой и аналогичными параметрами внешней среды.

изменениям среды, влияя в стрессовых ситуациях на скорость принятия решений мозгом посредством доставки гормонов.

На тело человека и его внутренние органы внешняя среда воздействует давлениями двух видов: атмосферным и гравитационным. Сердце под влиянием мозга изменяет частоту пульса, компенсируя это влияние. Поле атмосферного давления колеблется в пределах 6 % и не зависит от позы человека. Гравитационное давление, напротив, в зависимости от позы человека имеет разное расположение по отношению к вектору сердечного давления, направленному в основном вдоль оси тела (рис. 6).

Для описания процессов, происходящих с жидкостью в мозге человека, можно использовать разные варианты законов сохранения [44–46]. Спинномозговая жидкость (или, как её чаще называют, ликвор) наполнена разнообразными соединениями и ионами. В среднем плотность ликвора близка к плотности воды: 1,005–1,007 г мл^{-1} , т.е. отличается от плотности воды всего на 0,5–0,7 %. Следовательно, ликвор можно принять за почти идеальную несжимаемую жидкость и использовать закон Пуазейля для определения вязкости. При вертикальной позе человека имеем

$$\mu_1 = \frac{\pi r_c^4 (p_1 - p_2)}{8 Q_1 l}, \quad (5)$$

где r_c — диаметр капилляров, p_1 — давление сердца, p_2 — давление гравитации, Q_1 — расход жидкости при вертикальной позе, l — длина капилляров.

Для горизонтальной позы человека вследствие векторного сложения давлений вязкость выражается как

$$\mu_2 = \frac{\pi r_c^4 (p_1^2 + p_2^2)^{1/2}}{8 Q_2 l}. \quad (6)$$

При изменении позы получим соотношение

$$\frac{\mu_1}{\mu_2} = \frac{Q_2 (p_1 - p_2)}{Q_1 \sqrt{p_1^2 + p_2^2}}. \quad (7)$$

Если вязкость не изменяется, то для линейных законов на локальном участке

$$\frac{\mu_1}{\mu_2} = 1. \quad (8)$$

Из выражений (5) и (6) получим соотношение

$$Q_1 (p_1^2 + p_2^2)^{1/2} = Q_2 (p_1 - p_2), \quad (9)$$

т.е. при изменении позы возрастание потока жидкости при постоянной вязкости обратно пропорционально давлению. Существуют два варианта регуляции при изменении позы человека: либо мозг по цепи обратной связи, воздействуя на сердце, понижает (повышает) частоту его сокращений и тем самым уменьшает (увеличивает) давление p , либо регулируется масса потока крови посредством сжатия (расширения) сосудов, что обеспечивается системой *сердце → сосуды → глиальные клетки → нейроны*.

Скорость движения потока жидкости, а следовательно, и пульсация мозга в различных участках неодинаковы. Плотность линий потока соответствует величине скорости на данном участке. *Стационарному или установившемуся потоку* соответствует *уравнение неразрывности*, описывающее прохождение одного и того же объёма жидкости через любое выделенное поперечное сечение за одинаковые промежутки времени. Пусть S_1 и S_2 — две площади сечения, а \mathbf{u}_1 и \mathbf{u}_2 — соответственно векторы скорости частиц жидкости в равномерно движущемся потоке, тогда уравнение неразрывности имеет вид $\mathbf{u}_1 S_1 = \mathbf{u}_2 S_2$. Уравнение неразрывности справедливо для всех потоков, т.е. $\mathbf{u}S = \text{const}$. Однако возможен другой динамический вариант поведения жидкости в случае, когда "берега" пульсируют с частотой ω_1 , а объём жидкости пульсирует с частотой ω_2 . Тогда новое инвариантное отношение будет иметь вид

$$\mathbf{u}_1 \omega_1 = \mathbf{u}_2 \omega_2 \text{ или } \mathbf{a}_1 = \mathbf{a}_2, \quad (10)$$

где \mathbf{a}_1 и \mathbf{a}_2 — векторы ускорения двух потоков. Общий объём ликвора у взрослого здорового человека находится в пределах от 140 до 270 мл, что составляет около 20 % от веса мозга. Рассмотрим подробнее динамику подвижности жидкости с плотностью ρ_1 в полусфере.

Если допустить, что полусфера частично заполнена жидкостью (предположим, на 20 %), то любое её положение во внешнем пространстве за счёт переходного процесса за счёт перетекания жидкости будет восстанавливать высоту жидкости $h = f(R)$. Структуры мозга будут замедлять или ускорять перетекание жидкости в зависимости от относительного направления векторов ускорения. Масса перемещаемой жидкости пропорциональна объёму полушария V_s , заполненного жидкостью,

$$V_s = \frac{1}{6} \pi h (h^2 + 3r_1^2) = \pi h^2 \left(R - \frac{1}{3} h \right), \quad (11)$$

и плотности этой жидкости ρ_1 , где h — высота заполнения полусферы жидкостью, r_1 — радиус поверхности сегмента, R — радиус сферы. С учётом выражения (11) получим массу жидкости

$$m = \rho_1 V_s = \pi \rho_1 h^2 \left(R - \frac{1}{3} h \right). \quad (12)$$

Момент импульса \mathbf{L} выражается как

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}. \quad (13)$$

Здесь \mathbf{r} — радиус-вектор, исходящий из центра тяжести массы мозга, \mathbf{p} — вектор импульса силы. Продифференцировав (13), получим

$$\frac{d\mathbf{L}}{dt} = \frac{d\mathbf{r}}{dt} \times \mathbf{p} + \mathbf{r} \times \frac{d\mathbf{p}}{dt} = \mathbf{u} \times \mathbf{p} + \mathbf{r} \times \mathbf{F}_{\text{res}}, \quad (14)$$

где \mathbf{u} — вектор скорости, \mathbf{F}_{res} — вектор результирующей силы. При вертикальной позе произведение $\mathbf{u} \times \mathbf{p}$ равно нулю, поскольку векторы \mathbf{u} и \mathbf{p} параллельны друг другу. Аналогично обращается в нуль и член $\mathbf{r} \times \mathbf{F}_{\text{res}}$, так как центральные силы, создаваемые сердцем, параллельны вектору \mathbf{r} . Таким образом,

$$\frac{d\mathbf{L}}{dt} = 0, \text{ или } \mathbf{L} = \text{const}. \quad (15)$$

При изменении позы первый гравитационный член ($\mathbf{u} \times \mathbf{p}$) в правой части (14) не изменяется и равняется нулю. Второй член представляет собой векторную сумму гравитационной и сердечной сил. Результирующий момент \mathbf{T}_{res} силы выражается как

$$\mathbf{T}_{\text{res}} = \frac{d\mathbf{L}}{dt}. \quad (16)$$

Отсюда следует, что *результирующий момент силы равен скорости изменения момента импульса*. Если системы мозга считать замкнутыми, то $\mathbf{L}_{\text{res}} = \text{const}$, что является прямым следствием законов Ньютона. Отсюда для двух крайних поз: вертикальной и горизонтальной — можно записать ускорение кровотока в трёх вариантах. Ускорение кровотока в вертикальной позе обозначим как \mathbf{a}_\uparrow , а в горизонтальной позе — как \mathbf{a}_\rightarrow . Соответствующие им результирующие силы при вертикальной и горизонтальной позе: $(F_{\text{heart}} - F_{\text{grav}})_\uparrow$ и $(F_{\text{heart}}^2 + F_{\text{grav}}^2)^{1/2}_\rightarrow$. При этом имеем три случая в зависимости от соотношения ускорений:

- 1) $\mathbf{a}_\rightarrow > \mathbf{a}_\uparrow$ при $\frac{F_{\text{heart}} - F_{\text{grav}}}{\sqrt{F_{\text{heart}}^2 + F_{\text{grav}}^2}} > 1,$
- 2) $\mathbf{a}_\rightarrow = \mathbf{a}_\uparrow$ при $\frac{F_{\text{heart}} - F_{\text{grav}}}{\sqrt{F_{\text{heart}}^2 + F_{\text{grav}}^2}} = 1,$
- 3) $\mathbf{a}_\rightarrow < \mathbf{a}_\uparrow$ при $\frac{F_{\text{heart}} - F_{\text{grav}}}{\sqrt{F_{\text{heart}}^2 + F_{\text{grav}}^2}} < 1.$

Из рассмотрения этих трёх случаев следуют пять выводов:

1. Случай 1, когда $F_{\text{heart}} > F_{\text{grav}}$, является нормой.
2. Случай 2, $F_{\text{heart}} \gg F_{\text{grav}}$, соответствует значительному относительному снижению гравитации (невесомости). Это состояние опасно, поскольку оно может привести к инсультам. Для восстановления нормы мозг будет снижать частоту сокращений сердца или увеличивать содержание жидкости. При этом величина отношения $(F_{\text{heart}} - F_{\text{grav}})/(F_{\text{heart}}^2 + F_{\text{grav}}^2)^{1/2}$ стремится к единице, поскольку

$$\lim \left(1 - \frac{F_{\text{grav}}}{F_{\text{heart}}} \right) \rightarrow 1 \text{ при } F_{\text{heart}} \rightarrow \infty.$$

3. Случай 3, $F_{\text{heart}} < F_{\text{grav}}$, соответствует патологии, т.е. сердечной недостаточности, что приводит к невозможности находиться в вертикальном состоянии, так как при оттоке крови от мозга человек теряет сознание.

4. Наконец, особый случай — значительное повышение гравитации, $F_{\text{heart}} \ll F_{\text{grav}}$, когда

$$\lim \left(\frac{F_{\text{heart}}}{F_{\text{grav}}} - 1 \right) \rightarrow -1.$$

Это состояние — опасное, поскольку может приводить к потере сознания. Неслучайно космонавты считают самым сложным при подготовке к полётам испытание на вращающемся кресле или центрифуге.

Главный вывод: *благодаря регуляторным системам организма, включая мозг, жидкости организма всегда находятся в движении, причём могут двигаться с ускорениями. В этом и состоит смысл одного из вариантов адаптации к внешним условиям. Адаптация в данном случае — это способ поддержания устойчивости работы организма при изменении положения тела в трёхмерном пространстве при наличии гравитации*. У робота подобных адаптационных механизмов нет, но они ему и не нужны. Если в некоторых экзотических случаях такие механизмы будут необходимы, то их можно создать исходя из других технических принципов.

2.4. Простая модель достижения консенсуса

Наш мозг — это не только "третейский судья", распределяющий посредством обратных связей внутренние конкурентные ситуации между органами, но и интерфейс, согласующий генетически заложенные инстинкты, лежащие в основе потребностей тела, с их реализацией при изменениях внешней среды (рис. 7а).

Внешняя среда часто предстаёт перед нашими рецепторами в виде случайности (например, детерминированного хаоса [47]). Мозг пытается преодолеть случайность на основе конкуренции между разными видами рецепторов. На примере согласования мнений рецепторов рассмотрим один из вариантов простейшей компьютерной *концептуальной автоматной модели*, в рамках которой нами исследовалась динамика случайного поведения сложной системы, состоящей из N сенсоров с заданными периодически изменяющимися детерминированными правилами взаимодействия под действием внешнего периодического управления [48, 49]. Схема рассматриваемого процесса показана на рис. 7б. Работа этой системы представляет собой периодические переходы хаос \leftrightarrow порядок.

Опишем модель "диалога", в котором в пределе возможны два результата: двухполюсный вариант (диалог не дал результата) и однополюсный (удалось договориться). Допустим, что имеется K сосудов и в каждом из них находится n_i сенсоров, которые могут взаимодействовать между собой (рис. 7б). Полное число сенсоров равно N ,

$$N = \sum_{i=1}^K n_i. \quad (18)$$

Для простоты возьмём $N = \text{const}$, т.е. рассмотрим систему на малом интервале времени, когда размножением можно пренебречь. Каждый из сенсоров имеет свой номер и своё свойство. По оси x на рис. 7б отложены первоначальные свойства сенсоров. Все номера сенсоров

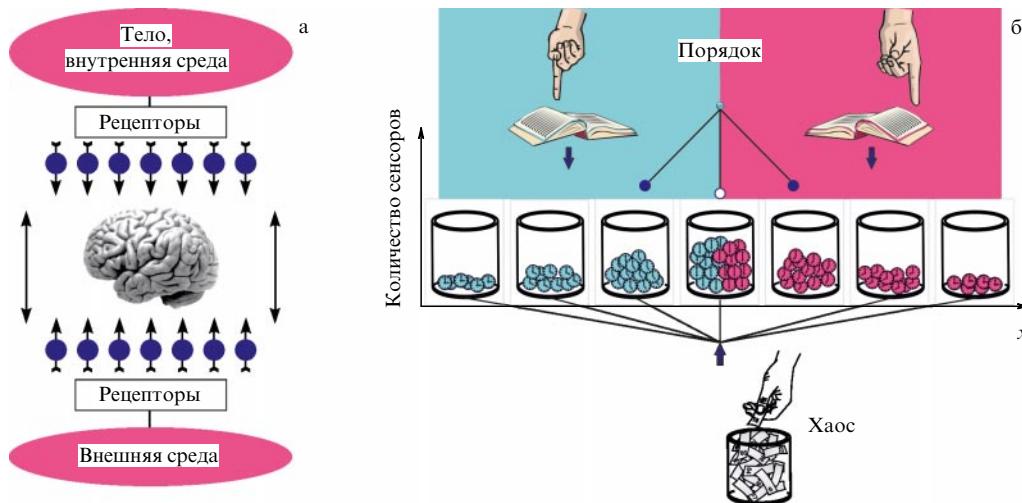


Рис. 7. Поиск равновесия (гомеостаза) между внутренней средой организма и изменением внешней среды: (а) схема мозга как интерфейса, согласующего внешнюю среду с внутренней средой организма; (б) простейшая компьютерная модель преобразования хаоса в порядок на основе двух популяций конкурирующих сенсоров.

перемешаны и находятся в сосуде под названием Хаос. Непрерывно с частотой Ω из сосуда Хаос *случайным образом* вынимаются два номера, что означает встречу и взаимодействие двух сенсоров, первоначально находившихся в своих сосудах на оси x . Выбор их случаен. Они могут находиться как в одном и том же сосуде, так и в разных сосудах. Правила взаимодействия встретившихся сенсоров могут изменяться на противоположные по отношению друг к другу, что определяется системой управления, т.е. положением колеблющегося с частотой ω внешнего маятника, задающего ритм взаимодействия. Это есть биологические часы:

$$\Omega > \omega, \quad (19)$$

где Ω — частота встречи сенсоров. Если маятник находится справа, то встретившимся сенсорам присваивается мнение (свойство) сенсора, находившегося в сосуде справа. Если маятник находится слева, то встретившимся сенсорам присваивается мнение (свойство) сенсора, находившегося в сосуде слева от него. Если до этого оба сенсора находились в одном сосуде, то при встрече их свойства не изменяются. После каждой встречи сенсоры возвращаются в сосуд Хаос и перемешиваются. Таким образом, сенсоры при описанных выше парных взаимодействиях будут изменять во времени кривую распределения суммарных мнений, отвечая на вопрос: как в динамике изменяется виртуальная модель, синтезируемая разными областями мозга, при изменении распределения мнений сенсоров по мере увеличения количества встреч?

По сути, описанная модель демонстрирует диалог двух спорящих коллективов, имеющих разные точки зрения на внешнюю или внутреннюю ситуацию. На рисунке 8 показана динамика изменения мнений во времени как переход от первоначального распределения к конечному через интервал времени, определяемый количеством взаимодействий.

Аналитически описать этот вероятностный процесс трудно. Причина состоит в неопределенности количества встреч сенсоров с одинаковым мнением. Очевидно, что

процесс установления мнений — сходящийся, но при этом имеется неопределенность в величине временного интервала переходного процесса, т.е. интервала времени достижения согласия — устойчивого или квазистабильного (колебательного) состояния.

Процесс в целом может закончиться по двум причинам: 1) время встреч исчерпано, но полная договоренность не достигнута; 2) договоренность достигнута за

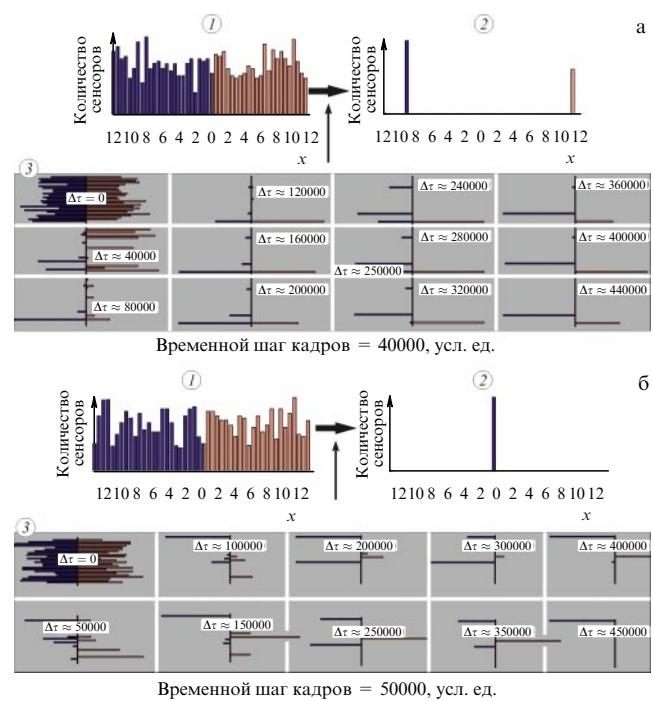


Рис. 8. Два примера фильмограммы динамики изменения мнения двух исходно разных конкурирующих коллективов сенсоров (по 500 сенсоров в каждом) при парных взаимодействиях. В верхней части рис. а и б показаны первоначальное (1) и конечное (2) нормированные распределения мнений. Ниже приведены фильмограммы (3) кинетики переходного процесса с указанным временным шагом демонстрации кадров.

заданный интервал времени, т.е. все разногласия сняты. В последнем случае все сенсоры собираются в одном сосуде.

Важен период движения маятника (смена обстоятельств во внешней или внутренней среде, анализируемых мозгом). Если маятник перемещается очень медленно, то его колебательное управление не будет оказывать особого влияния на процесс. Система в течение одного периода движения маятника, в крайнем случае двух или трёх периодов, успеет прийти к устойчивому состоянию с мнением того коллектива, преимущества которого соответствовали положению маятника в отвечающем ему цикле.

Интересны случаи, в которых на реализацию каждого правила отводятся приблизительно равные интервалы времени, причём много меньше, чем это необходимо для достижения полного согласия. В этом случае система переходит в длительный колебательный режим с неочевидным финальным результатом.

Исследование модели показало, что продолжительность переходного процесса конкуренции между двумя возникшими равносильными кластерами, порождающими противоположные мнения, резко возрастает. Изменение амплитуды распределения происходит с частотой колебаний маятника. Смещение маятника влево приводит к возрастанию по амплитуде мнений левого коллектива с соответствующим уменьшением мнений правого, а при смещении маятника вправо ситуация меняется в обратную сторону.

Динамику переходного процесса можно иллюстрировать также изменением среднеквадратичного отклонения $\sigma(t)$ распределения мнений по ансамблю всех сенсоров во времени (рис. 9).

Исследование модели выявило ряд интересных особенностей. Очевидно, что имеет место зависимость времени установления синхронизации мнений от симметрии исходного состояния и наличия разбросов градаций в исходном распределении (от генетики). Если асимметрия и локальные разбросы градаций в распределении отсутствуют, то вероятность возникновения двух равноценных кластеров и, следовательно, процесса с длительным переходным режимом возрастает.

На "выживание" градаций также оказывает влияние их локальное окружение. Часто в итоге появляется градация распределения, первоначально не соответствующая её максимальному значению в исходном распределении, поскольку сенсоры в процессе смены правил могут переместиться в градацию с исходно меньшей амплитудой. Образно говоря: *случай может уменьшить амбиции первоначально лидирующей группы сенсоров и развенчать их*. Случайное движение сенсоров из градации в градацию следует из случайного описания самого процесса встреч (на рис. 7б — это блок Хаос).

Изменение радиальной плотности во времени в расположении колец на графиках изменения разброса в полярных координатах свидетельствует о неравномерной сходимости к конечным результатам синхронизации в процессе переходного режима (рис. 9б, г).

Время достижения системой состояния синхронизации при введении внешнего периодического воздействия (маятникового управления) увеличивается на несколько порядков (например, изменяется от 10^5 до 10^9 в условных единицах при общем количестве сенсоров $N = 10000$). Периодическая смена правил вызывает

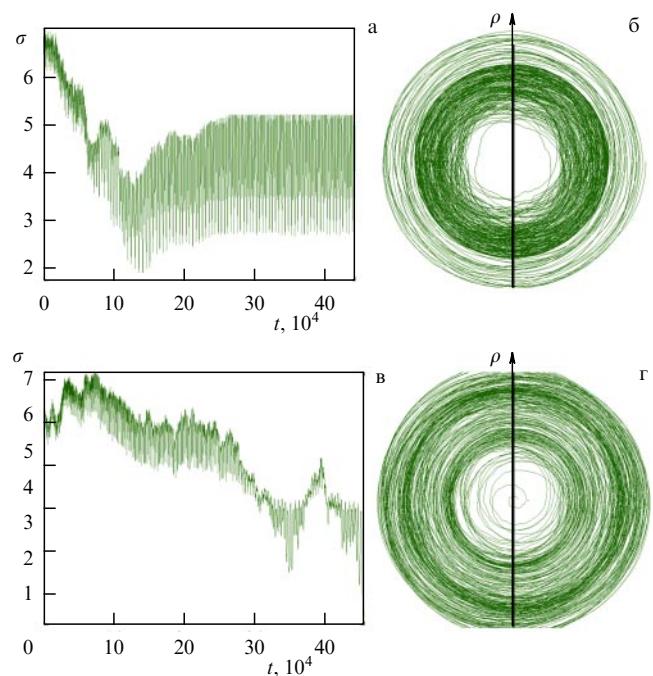


Рис. 9. Изменение во времени среднеквадратичного отклонения σ по ансамблю всех сенсоров. (а) Динамика схождения мнений во времени двух исходно разных конкурирующих коллективов сенсоров, соответствующая фильмограмме на рис. 8а. (б) Тот же процесс в полярных координатах, где $\rho = (\sigma^2 + t^2)^{1/2}$, а угол поворота $\varphi = \arctan(\sigma/t) = \arcsin[\sigma/(\sigma^2 + t^2)^{1/2}]$. (в) Динамика схождения мнений во времени, соответствующая фильмограмме на рис. 8б. (г) Тот же процесс в полярных координатах. Поскольку случайные взаимодействия сенсоров, согласно условиям, заложенным в модель, происходят через равные интервалы, количество их взаимодействий равно прошедшему времени.

переход системы в состояние вынужденного колебательного режима.

Два варианта самоорганизации привели к двум различным итогам. В первом варианте имеет место двухполюсная колебательная ситуация (рис. 8а), во втором — однополюсная статичноустойчивая ситуация (рис. 8б).

Эта модель описывает конкуренцию, которая характерна для парных взаимодействий. В живом организме на каждом его иерархическом уровне такие взаимодействия распределены в пространстве и времени по всему объёму тела, включая мозг. В одних участках тела и мозга консенсус достигается, в других, наоборот, увеличивается рассогласование. Переходы распространяются подобно волнам, перетекая из одного участка в другой и возвращаясь назад. Такие процессы перетекания возбуждений можно наблюдать в мозге с помощью магнитно-резонансной томографии (МРТ).

2.5. Примеры парадоксов взаимодействия сенсоров

Одна из задач, которую решает наш мозг, состоит в сопоставлении сведений, поступающих по разным каналам органов чувств. Одни сведения мозг игнорирует, другие усиливает и, сталкиваясь при этом с парадоксами, приходит в замешательство. Приведём пример с голограммами. В середине 1970-х годов по набору электронно-микроскопических снимков мы научились создавать трёхмерные голографические изображения (рис. 10) [50–52]. О свойствах голограмм тогда многие ещё не знали.

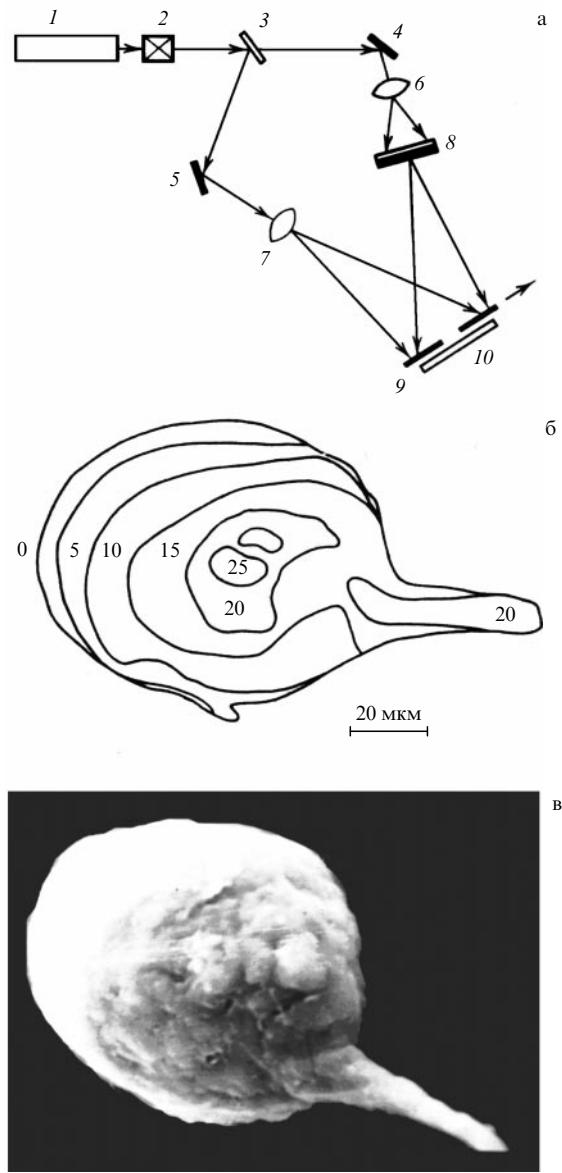


Рис. 10. Трёхмерное голографическое изображение нейрона большого прудовика *Limnaea stagnalis*, появляющееся в задымлённом воздухе как результат рассеяния света, пропускаемого через голографическую пластинку. (а) Схема установки для получения составных голограмм на основе фотографий, снятых в разных ракурсах: 1 — лазер, 2 — затвор, 3 — лучрасщепитель, 4, 5 — зеркала, 6, 7 — линзы, 8 — ракурсные снимки, 9 — перемещаемая щель, 10 — голограмма. (б) Линии равной высоты в объекте (в микрометрах), измеренные с помощью световой метки, вводимой в восстановленное изображение. (в) Фотография голограммы восстановленного изображения.

Человек (или обезьяна), впервые сталкиваясь с голографическим изображением предмета, обнаруживает, что рука проходит сквозь него. Бестелесность трёхмерного объекта может привести индивида либо в восторг, либо в ужас, вызвав или любопытство, или агрессию [50].

Один из вариантов реакции на голограмму был в описан в газетах Нью-Йорка в ноябре 1972 г., когда на Пятой авеню в течение часа движение транспорта остановилось, поскольку образовалась толпа, рассматривавшая женскую руку, которая протягивала из полутенни витрины магазина бриллиантовое ожерелье. Таинственное изображение, которое было голографическим, фор-

мировалось с помощью мощного лазера с голограммы, изготовленной фирмой McDonnel Douglas Electronics Company. Эффект объёмности был настолько поразительным, что кто-то из зрителей в конце концов пришёл в ярость и разбил зонтиком витрину, заявив, что это работа дьявола. Наши наблюдения показали, что противоречие между зрительными и тактильными каналами восприятия быстро устраняется, когда появляется понимание (т.е. коррекция виртуальной модели внешней среды внутри нашего мозга), что голограмма — это другой, особый, класс изображений [51].

Ещё примеры: изображение всех окружающих нас предметов линзой хрусталика проецируется на сетчатку глаза перевёрнутым на 180°. Новорождённый ребёнок довольно долго за счёт хватательного рефлекса пытается набраться опыта, взаимодействуя методом проб и ошибок с подвешенными над кроваткой игрушками. Согласие между воспринимающими информацию зрительными рецепторами и тактильным ощущением устанавливается, когда в мозге сформируется программа переворота изображения внешнего мира на 180°.

Примеров таких временных диссонансов организма при восприятии внешней среды с последующим восстановлением гармонии между рецепторами разной модальности можно привести несколько десятков. Например, при смене часовых поясов даже с трёхчасовой разницей у многих людей ухудшаются сон и работоспособность. У космонавтов в первые дни пребывания на орбитальных станциях возникает необходимость адаптироваться к новым формам перемещения в без gravitационном пространстве. Эксперименты с длительным ношением очков, переворачивающих изображения, требуют для восстановления ориентации в пространстве свыше 10 дней. Переезд по меридиану в другой регион, без изменения временной зоны, также требует адаптации к микрофлоре, микрофауне и качеству воды. Время адаптации сильно различается в человеческой популяции и зависит от возраста, в некоторых случаях адаптация сопровождается заболеваниями (десинхронозом, аллергией, диареей, морской болезнью и т.п.). Все эти адаптивные перестройки проходят через фазу хаотизации активности мозга с последующим восстановлением сенсорного согласия.

2.6. Виртуальная модель, синтезируемая нашим мозгом

О связи виртуальной модели мозга с нашим поведением и об особой роли такой модели говорят многочисленные факты. Например, на основе априорных предположений в сознании ребёнка складывается образ *страшного*. Если ребёнок опасается, что в незнакомой тёмной комнате сейчас находится нечто *страшное*, то мысль об этом вызывает соответствующее ощущение, получаемое мозгом якобы от глаз, ушей и других сенсоров (хотя формально они не дают таких сигналов). *Образ страшного* возникает внутри мозга, включая размер *страшного*, его силу, его необычный вид. Как правило, в отсутствие жизненного опыта все дети боятся темноты (крылатое выражение — "у страха глаза велики").

С одной стороны, ограниченность модели внешней ситуации не играет большой роли, потому что модель — это всегда упрощение действительности. Многие аспекты окружающего мира не учтены в модели, но главное отражено точно — избегать катастрофических ситуаций.

Однако, с другой стороны, у человека (и не только) в онтогенезе имеет место и развивается заложенная эволюцией противоположность подобному состоянию — это *любопытство*. В физике со времён Платона взаимодействие противоположностей называют спором *антиподов* (др.-греч. ἀντίποδες)². Как известно, в лингвистике антиподы именуют *антонимами* (др.-греч. ἀντί — приставка со значением противоположности + ὄνομα — имя), т.е. словами, имеющими противоположное лексическое значение.

Конкуренция и достижение согласия между страхом и любопытством приводит к развитию системы обработки информации, поступающей из внешней среды, что повышает адаптационные возможности как отдельного организма, так и вида в целом. Для мозга между восприятием и действиями существует тесная связь. Наше тело служит нам, чтобы познавать окружающий мир. Мы взаимодействуем с окружающим миром посредством своего тела и тем самым корректируем работу мозга. Виртуальные модели мозга часто дают сбои. Подобные сбои полезны, но иногда и опасны. Причин две.

С одной стороны, сбои позволяют разобраться, как устроен окружающий нас мир. Сбои во всякой ситуации допускают неоднозначную трактовку, например, когда два разных объекта окружающего мира вызывают одно и то же ощущение. Такие проблемы обычно решаются благодаря тому, что одна из возможных трактовок намного вероятнее другой.

С другой стороны, мозг может ошибаться, когда маловероятная трактовка на самом деле является правильной. Многие зрительные иллюзии возникают именно потому, что обманывают наше восприятие подобным образом. Если оранжевую морковь осветить синим светом, то наш мозг будет по-прежнему воспринимать её как оранжевую. При синем свете все объективные способы замера длины световой волны, отражённой от моркови, будут констатировать, что эта отражённая волна относится к диапазону синего цвета, $\lambda = 0,43 - 0,48$ мкм, а не оранжевого, $\lambda = 0,60 - 0,66$. Причина такого парадоксального восприятия связана с наличием виртуальной модели моркови в нашем мозге, составленной на основе прошлого опыта при её наблюдении при разном освещении.

Итак, наше восприятие зависит от априорных убеждений и памяти о прошлом. Наше восприятие начинается со страха, компенсируемого любопытством, что непрерывно уточняет нашу виртуальную модель мира. В коррекции модели априорно ни зрению, ни слуху, ни осязанию, ни другим рецепторам не оказывается предпочтения, поскольку все они могут быть информативными.

Этой способности не хватает компьютерным системам ИИ роботов без сенсоров тела. Если у роботов нет тела, снабжённого рецепторами, то такие роботы не смогут самостоятельно научиться делать предсказания. Они ограничены в возможности корректировать модель своими действиями³. Но это ещё не самое главное.

² Термин "антипод" был введён Платоном в III в. до н.э. в диалоге "Тимей".

³ У людей при повреждении любой доли головного мозга часто возникает нечто подобное. Они перестают корректировать виртуальную модель внешнего мира и, соответственно, цель своих действий. Такие люди не способны реализовать какой-либо план или следовать каким-либо указаниям. Они постоянно отвлекаются на всё, что попадает в их поле зрения.

2.7. Сенсоры тоже могут ошибаться. Эволюционное предотвращение появления ложных образов

Сетчатку нашего глаза, в которой происходит первичная обработка световых сигналов, исходящих от предметов внешней среды, часто называют "фрагментом мозга, вынесенным на периферию". Ричард Фейнман, говоря об устройстве глаза человека, утверждал [53]: "...чувствительные клетки расположены в сетчатке в глубине, так что, прежде чем попасть на рецепторы, свет должен пройти через несколько слоёв других клеток. Сетчатка как бы вывернута наизнанку! В общем, некоторые вещи в устройстве глаза кажутся великолепными, а некоторые — просто глупыми".

Далее в той же лекции Р. Фейнман, описывая глаз осьминога, говорил: "Сетчатка у осьминога, как оказалось, представляет собой тоже часть мозга, и образовалась она при эмбриональном развитии, как у позвоночных животных, но имеется одно очень интересное и поразительное отличие. Чувствительные к свету клетки в ней расположены не позади слоёв других клеток, как у нас, а непосредственно на внутренней поверхности глазного яблока, а клетки, занимающиеся вычислением, — позади них. Теперь мы, по крайней мере, видим, что в расположении клеток в нашем глазе глубокого смысла нет. В другой раз Природе пришлось исправить свою ошибку".

Однако утверждение Р. Фейнмана неправильное. Осьминоги древнее позвоночных животных, и дело не в том, что Природа исправляла в конструкции их глаза свою ошибку. Природа не ошибается, а если ошибается, то организм вымирает. В утопленных в глубь тела глазах человека есть глубокий адаптационный смысл. Для надёжного зрения и распознавания образов зрительные рецепторы (колбочки и палочки) должны находиться в среде с практически постоянной или медленно изменяющейся температурой. В противном случае быстрые колебания температуры во внешней среде, например вследствие ветра, будут создавать тепловые градиенты и вызовут изменение скорости биохимических реакций при преобразовании квантов света в электрические импульсы в сетчатке [54]. Появятся ложные зрительные иллюзии — *призраки*, которые в действительности не существуют. Мозг, борясь с такой ситуацией, начнёт переходить в режим хаотического поведения в поисках ответа на вопрос: что это? Реальность или иллюзии? Для того чтобы устраниТЬ эту проблему, эволюционный отбор погрузил сетчатку у теплокровных организмов, живущих на суше, в глубь глазного яблока, где температура практически постоянна. Осьминоги живут в воде, температура которой изменяется медленно и не оказывает влияния на их зрительное восприятие.

3. Адаптация робота и человека к окружающей среде

3.1. XX век: от кибернетики к синергетике

Более 70 лет назад Нобелевский лауреат Конрад Лоренц (Conrad Lorenz) писал, что в мире животных *обучение практически всегда приводит к адаптации, т.е. к достижению согласованности с внешней средой*. В 1941 г. он опубликовал статью "Кантовская концепция a priori в свете современной биологии" [55]. Вступая в заочный спор с Кантом, К. Лоренц утверждал, что априорные формы мышления и интуицию следует понимать как

адаптацию, т.е. достижение компромисса, приводящего к повышению устойчивости организма.

30 лет назад сфера применения, быстродействие и миниатюризация компьютеров стали стремительно развиваться. Появилось искушение создать ИИ, подобный человеческому [56–62]. 16 лет спустя в книге *Модели человека* Герберт Саймон (Herbert A. Simon), также Нобелевский лауреат, утверждал [63], что через 10 лет компьютеры смогут обыгрывать в шахматы чемпионов мира⁴, писать стихи и музыку. Это предсказание сбылось. Более того, в 2016 г. компьютер обыграл профессионала в китайской игре го, которая имеет большее комбинаторное разнообразие, чем шахматы⁵.

Тем не менее Саймон предполагал, что должны быть и какие-то ограничения. Дефицит времени при решении той или иной задачи заставляет рассматривать не все возможные пути, которых бесконечно много, а только наиболее существенные. При этом Человек, в отличие от Робота, стремится за ограниченное время построить какую-либо стратегию, возможно, далёкую от оптимальной, но обеспечивающую сиюминутный успех в данной конкретной ситуации [63, 64]. Ключевые слова здесь — "за ограниченное время" и "успех". Дефицит времени возникает из-за необходимости преодоления мешающих, а часто угрожающих жизни условий изменения состояния внешней среды. В Институте биофизики АН СССР (ныне Институт теоретической и экспериментальной биофизики РАН) подобную мысль развивал М.М. Бонгард в 1960-х годах [60, 65].

Российский биофизик М.Л. Цетлин в 1950-х годах начал рассмотрение социальных аспектов с точки зрения теории игр. Ранее Джон фон Нейман [66] исследовал ситуации, в которых участники не обладают в полной мере априорной информацией об алгоритме игры. М.Л. Цетлин показал, что если случайное поведение внешней среды изменяется медленно, то роботы даже с линейной тактикой обучаются иправляются с задачей адаптации к условиям игры. Стало очевидным, что фактор времени оказывает существенное влияние на возможности адаптации. Этот результат позволил М.Л. Цетлину в дальнейшем рассмотреть социальные игры с большим числом участников. Итоги этих исследований были объединены в сборнике его работ [62], который вышел в свет в 1969 г., к сожалению, уже после смерти автора.

В 1990 г. австралиец Родни Брукс (Rodney Brooks) написал статью "Слоны не играют в шахматы" [67], в которой обратился к создателям робототехники с призывом — разрабатывать системы ИИ, основываясь на данных нейрофизиологии. Брукс отмечал, что разум — это прежде всего "инструмент" адаптации организма к внешней среде. Положительные эмоции подавляют панику и способствуют принятию творческих решений даже в условиях недостатка информации на основе вы-

⁴ 11 мая 1997 г. компьютер Deep Blue (корпорация IBM) в матче из шести партий обыграл чемпиона мира по шахматам Гарри Каспарова со счётом 2:1 при трёх ничьих.

⁵ С 9 по 15 марта 2016 г. между компьютерной программой AlphaGo, разработанной британской компанией Google DeepMind, и корейским профессионалом игры в го Ли Седолем был проведён матч из пяти партий. Компьютер выиграл со счётом 4:1. В программе AlphaGo использовались 1920 процессоров и 280 графических процессоров, работающих в распределённой сети. Игры транслировались в прямом эфире на YouTube.

хода за пределы "возможного" (крылатое выражение: *кто не рискует, тот не пытёт шампанского*) [37].

Сегодня многие исследователи в области робототехники, например Моше Варди (Moshe Vardi) из Университета Райса в Хьюстоне (США) [68], уверяют, что нет никаких пределов и КАР со временем станут не только быстрее, но и намного умнее человека. Однако реализация этой мечты сложна. Работа нейронной сети Хопфилда (Hopfield network) [69] в своё время очаровывала меня, поскольку сеть после нарушения равновесия самостоятельно находила новое положение равновесия [70]. Однако позднее я понял, что процедура поиска равновесия детерминирована, поскольку она заложена в программу, а алгоритм задаётся программистом. Хотя интервал времени наступления равновесия зависит от многих внешних и внутренних причин, как показано на примере модели, рассмотренной в разделе 2.4. В простейшем случае — это нахождение локального минимума энергии ("потенциальной ямы") на n -мерном кубе.

Все дальнейшие модификации этого алгоритма (например, нейронная сеть Хэмминга для классификации бинарных векторов или персептроны разных видов [71–73]) также приводят к системам с жёстким, детерминированным алгоритмом, созданным программистами. Сегодня процесс "мышления" обучающихся современных роботов представляет собой всего лишь очередной этап развития соединённых в параллельно-последовательные сети детерминированных машин Тьюринга [59]. У человека есть свобода выбора цели. У робота цель формирует человек. Следовательно, у робота нет творчества, если робота такой опцией не снабдил его создатель.

Термин *творчество* можно определить как поиск несимметричного, неожиданного ответа на возникающие препятствия при выборе дальнейшего движения к заданной или возникшей спонтанно цели с учётом рисков, неудач и удач.

3.2. Синергетика — теории кооперативных динамических взаимодействий диссипативных систем

В 1990-х годах в качестве модели коллективного поведения нами была использована популяция бактерий, образующих колонии. В связи с возникающими дискуссиями мы хотели снизу вверх, начиная от простейших организмов, подняться по ступеням формирования процессов *сознания*. В качестве начального этапа самоорганизации была взята операция создания бактериями своих колоний [74, 75]. Однако работу не удалось завершить. Причина — неполнота знаний о переходах от простейших организмов к сложным.

На смену кибернетике 1950–1960-х годов в 1980-х годах пришли синергетика [76] и термодинамика диссипативных структур школы И.Р. Пригожина [77]. Многим казалось (как ранее с кибернетикой), что новые идеи позволяют создать общую теорию и наконец ответить на вопрос: как мыслит человек? Дело в том, что наш мозг то работает, как детерминированная система Лапласа, то вдруг переходит на короткое время к хаотическому поведению, то вновь возвращается к детерминированному. Мысль о хаотизации в приложении к процессам мышления высказывалась давно. В конце 1950-х годов Н. Бор [78, с. 27, 28] отметил: "Основные признаки живых организмов зависят от процессов атомного масштаба, где мы сталкиваемся с существенными ограничениями

для применения понятий классической физики... Из свойственного классической физике детерминизма следует, что всякое возмущение системы, состоящей из огромного числа частей, непременно приводит к хаотическому беспорядку. В квантовой же физике это описание отражает результат взаимодействия между устойчивыми атомными системами, поэтому оно основывается на исходе конкуренции между различными индивидуальными процессами. Эти процессы простым образом определяют состояние новых систем через посредство содержащихся в них атомных частиц, подобно тому как они определяли бы первоначальное состояние системы..."

В 1971 г., опираясь на моделирование нелинейной динамики, Давид Рюэль и Флорис Такенс опубликовали работу "О природе турбулентности" [79]. В ней они подвергли критике сценарий Ландау–Хопфа [80, 81], указав на то, что уже после трёх-четырёх бифуркаций динамика может стать турбулентной, в частности, у системы может возникнуть характерный для случайного процесса сплошной спектр. Рюэль и Такенс связывали это обстоятельство с возникновением "странныго аттрактора" в фазовом пространстве из-за неустойчивости траекторий. Как показали эксперименты, во многих открытых нелинейных системах, далёких от равновесия, происходит самоорганизация [82]. При этом обычно возникают либо пространственно неоднородные стационарные (т.е. медленно изменяющиеся со временем) образования, которые И.Р. Пригожин предложил называть диссипативными структурами [83], либо периодические или квазипериодические пространственно-временные волны, которые стали называть автоволнами [84, 85].

В 1961 г. был создан Международный союз теоретической и прикладной биофизики (International Union for Pure and Applied Biophysics — IUPAB). При организации конгрессов IUPAB была необходима структуризация потоков присланных докладов, поэтому было принято решение разделить биофизику на три части: молекулярная биофизика, биофизика клетки и биофизика сложных систем. Этот рубрикатор положили в основу деления биофизики на указанные три части, в том числе в биофизических журналах. Такая классификация применяется и поныне (например, в российском журнале *Биофизика*), но она небезукоризненна, так как в биологии *трудно найти простые системы*. Что такое сложность, и чем сложные системы отличаются от простых?

Более 30 лет назад в Баварии прошёл симпозиум по синергетике, на котором вопросы сложности разных систем, включая мозг, обсуждались с разных сторон. Однако универсальной теории оценки сложности, так же как и универсальной оценки устойчивости биологических систем, до сих пор не существует [27]. Слишком разнообразны наборы вариантов. Главное их свойство можно описать афоризмом: *изменить одно — изменится всё*. В мозге человека имеют место цепные реакции: перемены в одной деятельности вызывают цепную реакцию и влияют на другую деятельность [86]. Подобные явления свойственны динамическим системам, содержащим существенную хаотическую компоненту [87].

Г.Г. Малинецкий и А.Б. Потапов предположили, что "сложность" основана на дуализме обработки информации [88]. С одной стороны, её можно понимать как сложность устройства, т.е. наличие в некоторой системе большого числа элементов и/или нетривиальных связей между ними. С другой стороны, речь может идти о слож-

ности внешних проявлений системы безотносительно к её внутреннему устройству, т.е. о сложности, выражающейся в нетривиальном поведении. В основе образования диссипативных структур и возникновения автоволновых процессов лежит явление самоорганизации, т.е. наличие переменных параметров двух видов. Первые, которых большинство, соответствуют "мягким быстро перестраивающимся режимам". Вторые, число которых невелико, соответствуют "жёстким режимам". По истечении достаточно продолжительного промежутка времени первые режимы подстраиваются под вторые, т.е. две раздельные системы объединяются. Это объединение приводит к тому, что для описания возникшей объединённой системы требуются новые переменные, которые не обязательно должны совпадать с первоначальным набором переменных, описывающих поведение двух раздельных систем. Другими словами, при самоорганизации описание системы должно меняться (см. простейший пример в разделе 2.4). Такое свойство самоорганизующихся систем называют эмерджентным. Напомним, что термин *эмержентность*, или *эмегентность*, происходит от англ. *emergent* — возникающий, появляющийся неожиданно. В синергетике это означает наличие у какой-либо системы особых свойств, не присущих как составляющим её элементам, так и сумме этих элементов. Эти особые свойства возникают благодаря системообразующим связям между элементами.

В настоящее время на смену эре диссипативных структур и автоволновых процессов в синергетику приходит эра самоорганизованной критичности, поставщиками идей которой становятся такие науки, как нейрофизиология, биофизика, психология, теоретическая история (Big History) и другие области, связанные с анализом влияния *прошлого на настоящее и будущее* [24].

Ю.А. Данилов [89] отмечал, что рождение науки о нелинейных взаимодействиях привело к отказу от ньютонанского представления о траектории как о геометрической линии, т.е. как о "длине без ширины". Физическое описание поведения динамической системы на языке траекторий означало бы, что у нас имеется прибор со столь высокой разрешающей способностью, что он позволяет нам "увидеть" отдельную геометрическую линию. Разумеется, в действительности разрешающая способность любого прибора конечна, и мы можем увидеть не индивидуальную траекторию, а только пучок совмещённых индивидуальных (порою спутанных) траекторий. Все траектории внутри пучка для нас неразличимы, поэтому имеет смысл говорить лишь о некотором *вероятностном распределении*. При этом, по терминологии И.Р. Пригожина и И. Стенгерса [90], от распределения вероятностей невозможно перейти к отдельным траекториям. Внешний наблюдатель их не видит. Несводимые вероятностные распределения коренным образом изменили описание динамических систем и даже понимание физических законов. Ключевые слова здесь: *вероятностное распределение или особая роль наблюдателя*.

И.Р. Пригожин и И. Стенгерс отмечали [90] следующее: «*Традиционно существовали две формулировки физических законов: одна — в терминах траекторий или волновых функций, другая — в терминах статистических ансамблей. Но такая статистическая формулировка не была несводимой. Она была вполне применима к отдельным траекториям или волновым функциям. Иначе говоря, при статистическом подходе не появлялись новые дина-*

мические свойства. В результате необратимое приближение к равновесию традиционно было принято связывать с приближённостью, "крупнозернистостью" описания, а появление стрелы времени — связывать с неполнотой нашего знания. Предложенная нами несводимая формулировка порывает с этой ситуацией. Не обратимость и вероятность становятся объективными свойствами. Они выражают то обстоятельство, что наблюдаемый нами физический мир не может быть сведён к отдельным траекториям или отдельным волновым функциям..."

Из изложенного выше следует, что функцию мозга невозможно полностью понять исходя из анализа на одном или двух иерархических уровнях, например, нейронной сети или набора сетей в виде сэндвича. В таком варианте единственным информационно-значимым событием является процесс обмена электрическими импульсами на основе движения ионов K^+ , Na^+ , Cl^- и Ca^{2+} между отдельными нервными клетками, запускающими через синапсы каскад дальнейших событий. Что касается работы мозга человека, то остаются вопросы: как может быть устроен мозг, чтобы его функция могла рассматриваться одновременно и как детерминированная, и как хаотическая? Как вообще возможно существование физических макросистем на основе дуализма?

Любой человек всегда больше того, что непосредственно извне воспринимается наблюдателем за его поведением. Для человеческого интеллекта, когда мы его рассматриваем как целое, нарушается закон тождества $\{A_{obs}\} \neq \{A_{real}\}$, где $\{A_{obs}\}$ — наблюдаемое множество свойств интеллекта, $\{A_{real}\}$ — реальное множество его свойств. Для человека характерны скрытые параметры, которые проявляются при изменении внешней среды, но априорно в другой конкретной ситуации не обнаруживаются. Таким образом, имеет место неравенство

$$\{A_{real}\} > \{A_{obs}\}. \quad (20)$$

Иными словами, человеческий интеллект и его поведение в своей основе многовариантны. Человек полностью "не раскрывается" перед внешним наблюдателем.

До недавнего времени робот как машина всегда оставался тем, что он есть, т.е. чем-то вполне определённым, но современные самообучающиеся роботы уже стали по поведению похожими на человека. Из компьютера, реализующего самообучающийся ИИ, часто трудно извлечь сведения о том, какие признаки он использовал, решая ту или иную задачу.

Кроме того, робот и человек являются открытыми системами для протока не только энергии, но и информации из внешней среды. Для человека открытость системы приводит к тому, что он не может достоверно осознать правила, которым подчиняются его мышление, сознание и поведение. Эти правила непрерывно меняются под влиянием как внешней среды, так и внутренней среды организма. Предположим, что человек полностью выяснил алгоритм собственного поведения X , который исчерпывающим образом описывает функцию его собственной психики и функцию его мозга $\{A_{real}\}$. Тогда он будет способен точно предсказывать риски при принятии того или иного решения, осуществляя определённые действия и поступки. Допустим, что точно известные величины рисков привели его к заключению, что в ситуации X при осуществлении действия P возникнут не-

приятности. Но тогда он может помешать их неизбежности, сознательно отказавшись от осуществления поступка P , и совершив какой-то другой поступок, вплоть до противоположного по направленности.

Однако принятие того или иного решения всегда происходит в условиях неопределенности поведения конкурента во внешней среде. Следовательно, оценка риска всегда носит вероятностный характер. Другими словами, человек одновременно должен как детерминированная система и не должен как осознавшая себя вероятностная система осуществлять или не осуществлять действие P . Неопределенность приводит к колебательному процессу, т.е. к вечному вопросу Гамлета: быть или не быть? Производить или не производить действие?

4. XXI век: новая интерпретация старых идей

4.1. Время дорогих проектов

В 2009 г. Генри Маркрам (Henry Markram) на основе своих исследований [91–107] представил научной общественности идею создания математической модели мозга на суперкомпьютере [108]. Идея была простой: если в компьютерную программу включить всю сеть из 86 млрд нейронов и 100 трлн синапсов, соответствующих архитектонике мозга человека, то, как предполагал Г. Маркрам, количество само перейдет в новое качество и такая система будет обладать всеми когнитивными качествами мозга человека. Появился проект Human Brain Project, рассчитанный на 10 лет (2013–2023 гг.), с общей стоимостью в 1,63 млрд долларов.

Но уже через год возникли сомнения в целесообразности реализации этого проекта [109]. Более чем 800 нейрофизиологов написали письма, в которых критиковался проект. Критика основывалась на том, что пока отсутствует понимание работы мозга человека в сравнении с работой алгоритмизируемых процессов ИИ на компьютерах. Была создана комиссия, её доклад на 53 страницах резюмировал доводы против проекта Маркрама. Также подверглись критике и способы принятия в Брюсселе проектов затратных исследований без должной научной проработки их основы.

В связи с критикой идеи Г. Маркрама следует упомянуть мегапроект "Коннектом человека" (Human Connectome Project) с более скромной стоимостью (30 млн долларов), который был запущен в 2009 г. в США [110]. Термин *коннект* происходит от англ. *connect* — соединение, например, с сервером для выхода в Интернет. Коннектом — это динамический граф (карта) возникающих, развивающихся и отмирающих связей между самоорганизующимися подсистемами мозга как результат дальнодействующих сил или полей (англ. *guidance* or *pathfinding*). Например, аксональное наведение и поиск пути, т.е. направленный процесс роста сети аксонов и дендритов, стали его основой. В период развития нервной системы (в онтогенезе) аксоны прорастают на большие расстояния, достигая всех органов и клеток, формируя тем самым архитектонику нервных связей как единую систему всего организма. При этом тело и мозг представляют собою единое целое. Изучение сил и взаимодействий, влияющих на выбор пути, важно для понимания того, как строится структура объединения мозга и тела и какие нарушения при их самоорганизации могут приводить к патологиям.

Предполагалось, что в механизме формирования связей между нейронами заключены основные аспекты человеческой индивидуальности (такие как личность, интеллект и креативность), поэтому моделирование коннектома человека могло бы стать существенным шагом к пониманию всего множества мыслительных процессов. Казалось, что для этого потребуется только особый язык программирования [111]. В России эту же мысль развивает К.В. Анохин, используя термин *когнитом* (гиперсеть мозга) [112].

Китайские учёные в августе 2017 г. объявили об открытии научного центра HUST–Suzhou Institute for Brainsmatics (HUST — Huazhong University of Science and Technology), который будет заниматься картированием головного мозга. Новый научный центр получил из бюджета 450 млн юаней (67 млн долларов) на пять лет. Основная его цель — улучшение существующих технологий картирования головного мозга с помощью современных электронных микроскопов методом нарезания ткани на супертонкие нанослои для их последующего изучения.

Атлас мозга с микрометровым разрешением Allen Human Brain Reference Atlas, созданный сотрудниками Алленовского института исследований мозга в США, недавно появился в Интернете в открытом доступе. Его печатная версия опубликована коллективом из 40 авторов и заняла весь выпуск *Journal of Comparative Neurology* [113] от 15 сентября 2016 г.

Однако исходя из своего опыта [114, 115] могу заметить, что исследование анатомии мозга без одновременного изучения динамики его работы на том же объекте существенно не приблизит нас к пониманию механизмов креативности. Мозг каждого человека уникален и индивидуален, поскольку в нём содержится жизненный опыт каждого из нас. Вероятность их полного совпадения даже у одногенетических близнецов стремится к нулю. Кроме того, проблема состоит в том, что любая идея требует экспериментальной проверки. Для того чтобы изучить динамику изменения связей, необходимо иметь метод регистрации формирования как структуры, так и функции во времени и пространстве на одном и том же мозге. Процессы, которые идут в нейронных сетях и их межклеточном пространстве, во времени происходят в миллисекундном диапазоне, а в пространстве — в нанометровом диапазоне. Арсенал методов для наблюдения и совмещения этих процессов, который имеется в распоряжении исследователей сегодня, к сожалению, не обладает таким пространственно-временным разрешением. Позитронно-эмиссионная томография (ПЭТ) даёт невысокое временное разрешение — порядка 10 с, чего явно недостаточно [116]. Электроэнцефалография (ЭЭГ) достигает временного разрешения до 2 мкс [117], но имеет низкое пространственное разрешение. Вживление микроэлектродов повышает пространственное разрешение, но этот метод — инвазивный и использовать его на человеке можно лишь в лечебных целях, да и то в исключительных случаях. Кроме того, для всех инвазивных методов существует проблема неопределенности, поскольку ввод чужеродного тела в мозг вызывает реакцию на его появление. МРТ позволяет регистрировать изменение потребления кислорода, а следовательно, скорость метаболических процессов. Её пространственное разрешение порядка 1 мм, а временное — порядка 2 с [118].

Большие надежды возлагаются на развитие нейрофотоники, например, на использование пространственно-временного анализа работы сети нейронов с помощью гибких оптоволоконных устройств, одновременно возбуждающих и регистрирующих внутреннюю или наведённую флуоресценцию нейронов. Формально временное разрешение этого метода неограничено, а пространственное разрешение определяется длиной волны используемого излучения [119]. Это представляет собой миниатюризацию методов динамической наноэндоскопии в приложении к мозгу. Однако методы нейрофотоники, как и микроэлектродные методы, являются инвазивными с вытекающими отсюда недостатками и неопределенностью результатов.

В книге отца и сына Шерозия [120] также излагался проект создания антропоморфного робота. Стоимость проекта авторы оценили в 10 млрд долларов, а сроки реализации — в 10 лет. Достоинством их проекта является идея учёта в работе мозга обратной связи с телом. По оценке авторов [120] среднее количество степеней свободы при движении тела человека порядка 300, хотя такая оценка у меня вызвала сомнение. Организм человека можно представить как сформировавшийся в процессе эволюции конденсат из одноклеточных организмов, находящихся в едином объёме тела [27]. Эта гигантская колония на всех иерархических уровнях управляет единым генетическим кодом. Число степеней свободы у такой динамической системы будет соответствовать по порядку величины числу клеток в теле взрослого человека, т.е. $\approx 10^{13}$ [27].

4.2. Классическая и вероятностная логика

Попытаемся для описания мышления и сознания использовать другой язык: не от архитектороники к работе сетей мозга, а от логики их работы — к архитектонике. Мысль о таком подходе к анализу динамических процессов не является новой, она рождалась у многих исследователей ещё до XX в., она живет и в XXI в. Можно отметить, например, лекцию В.Г. Редько на IV Всероссийской научно-технической конференции "Нейроинформатика-2002" [121]. При этом, как показано ниже, мы приходим к нетривиальному результату.

Классическая логика, в отличие от вероятностной, обладает чёткими правилами. На её основе развился математический аппарат классической физики. К классической логике относятся все разделы логики, которые удалось представить в символической форме. Существенный вклад в символическую формализацию процессов логического мышления внесли О. де Морган [122], Дж. Буль [123], Дж. Пеано [124], Г. Фреге [125], А. Пуанкаре [126], Д. Гильберт [127], Б. Рассел [128] и др.

Тем не менее всесильность классической логики при получении объективного знания ограничена. Классическая логика утверждает [129], что надо однозначно отождествлять с символами объекты окружающей нас внешней среды (закон тождества). Построение алгоритма должно протекать непротиворечиво по линейной цепи: причина \rightarrow следствие. Логический закон противоречия запрещает в процессе рассуждений и анализа ситуаций противоречить самому себе. Нельзя, например, о положении, которое принимается верным, говорить в то же время как о неверном. Кроме того, согласно закону исключённого третьего недопустимо на вопрос отвечать неопределенno: ни да, ни нет. Наконец, любая

мысль верна лишь тогда, когда она вытекает как следствие из другой мысли, которая служит её основанием (закон достаточного основания). Таким образом, мышление должно быть последовательным.

В рамках этой стройной системы логического мышления человек чувствует себя комфортно, потому что классическая логика замкнута, внутренне непротиворечива и однозначна. Однако практика подсказывает, что в замкнутой системе утверждений находятся высказывания, о которых нельзя сказать, истинны они или ложны. Наиболее яркий вариант — это разные парадоксы, например апории (от др. греч. *απορία* — трудность, безысходность, безвыходное положение) Зенона [130], которые нельзя объяснить на основе классической детерминированной логики.

Ограниченностю классической логики состоит также в том, что она отражает лишь один класс ситуаций во внешней среде, которые соответствуют "порядку" (а не "хаосу" [131]). В её рамках сложно объяснить спонтанное возникновение какой-либо цели у человека или прогнозировать траекторию развития нелинейных процессов, таких как "капризы" погоды, землетрясения, изменения экономических, биологических и социальных систем или появление креативности. Попытки построения будущего на основе прошлого в нелинейных системах или формулировки политических доктрин на основе формальных логических построений часто приводят к догматизму, а следовательно, к крушению замкнутых мировоззренческих концепций.

Начиная с идеей Парменида⁶ бытует утверждение: *лишь умопостигаемое бытие, а не данные органов чувств, обладает действительной реальностью* [132]. Поясним с современных позиций, как можно обосновать гипотезу Парменида. Приведём постулат, например: *законы (модели) природы не должны быть сложнее тех данных, которые они объясняют*. В противном случае понятие сжатия информации отсутствует, поскольку любой необработанный набор данных может существовать сам по себе как "закон". Если память ничем не ограничена, то, казалось бы, можно просто запомнить все ситуации, не производя отбора, и принимать их как частные законы природы. Однако к настоящим законам природы, формулируемым наукой, применяют другой принцип — *принцип простоты, экономии памяти и экспериментальной проверки* [133].

Грегори Чейтин (Gregory Chaitin) в конце XX в. отметил [134], что при этом возникают противоречия. Казалось бы, подавляющее большинство строк символов (например, ряды случайных чисел в иррациональных дробях), описывающих явления внешнего мира, принципиально несжимаемо. Если это так, то мы не можем свести их к более простым, коротким последовательностям. Подобная ситуация возникает каждый раз, когда в наблюдаемой цепи событий не обнаруживается никаких внутренних закономерностей, которые позволили бы её сжать. У нас не остаётся другого выбора, кроме как принять всю последовательность в качестве закона природы. Мы принимаем такую ситуацию без всякого обоснования, просто потому, что не имеем другого выбора.

⁶ Парменид из Элеи — древнегреческий философ, живший во времена около 550 лет до н.э., пытался отделить истину от мнения о ней. Его учеником и последователем был Зенон — автор парадоксов (апорий).

Как отмечалось выше, по-видимому, мозг может самостоятельно порождать виртуальные модели (или строки программ) произвольной сложности. Их сравнение с реальностью позволяет вскрывать локальные (частные) закономерности, существующие во внешней среде. Динамическое сравнение происходит посредством сопоставления синтезированной последовательности с последовательностью, заданной внешней средой, что позволяет при совпадениях фрагментов обнаруживать локальные закономерности, разделённые большими интервалами в пространстве и времени. Например, для текстов это слова или фразы. В 1989 г. мы реализовали похожую идею в виде алгоритма. Спектральный анализ в фурье-пространстве позволял осуществлять поиск повторов, которые отделены друг от друга большими интервалами, например, в нуклеотидных последовательностях ДНК [135].

Нахождение повторов обеспечивало сжатие информации, а случайные несжимаемые части последовательности можно было отнести к разряду "мусора". Тем не менее при этом нельзя объяснить, почему "мусор" не содержит информации. Если изменить исходные фрагменты, то изменится и состав "мусора". Единственным объяснением такой ситуации будет то, что сжатие проведено на основе классической детерминированной логики, не противоречит ей, позволяет решать практические задачи и в данном конкретном случае "мусор" — это шум. Робот действует на основе алгоритмов, построенных на законах классической логики, заданных программистом. Работа мозга Человека основана на преодолении противоречий между антиподами, т.е. между настоящим и прошлым, между воспринимаемым органами чувств внешним миром и памятью о нём из прошлого опыта.

Диалектическую логику часто характеризуют как логику, рождающуюся на основе конкуренции в спорах (крылатое выражение: *в спорах рождается истина*). Б.Б. Кадомцев, с которым мне несколько раз приходилось обсуждать проблему возникновения вероятностного информационного мира, пропагандировал его связь с квантовой механикой [136, 137]. Свои соображения на этот счёт он обобщил в книге *Динамика и информация* [138].

В начале XX в. Д. Гильберт (David Hilbert) обратился к математикам с предложением определить всю конечную совокупность принципов, на основе которых с помощью последовательного использования правил классической математической логики можно найти гармонию во всей математике. К 1922 г. у Гильберта сложился план сделать математику замкнутой путём её полной формализации на основе доказательства отсутствия противоречий внутри неё. Для осуществления этой программы Гильберт, продолжая работы Фрехе [125], разработал логическую теорию доказательств, с помощью которой непротиворечивость математики свелась бы к доказательству непротиворечивости, например, арифметики. При этом Гильберт использовал только общепризнанные логические средства классической, а не вероятностной логики. Как отмечает Г. Чейтин [134, 139], программа Гильберта оказалась невыполнимой, хотя послужила значительным стимулом к развитию логики.

Похожей проблемой в своё время интересовался и Курт Гёдель (Kurt Friedrich Gödel). В 1931 г. Гёдель

показал, что на путях формализации математических теорий появляется проблема противоречивости. На границах любых замкнутых множеств возникают "недоказуемые теоремы" [140]. Смысл утверждения Гёделя состоит в том, что любая теория является неполной и, следовательно, противоречивой. Неполнота означает наличие высказываний, которые на границах множеств нельзя ни доказать, ни опровергнуть исходя из аксиом, находящихся внутри этой теории. Противоречивость — это появление парадоксов, т.е. возможности доказать любое высказывание, приняв истинное за ложное и наоборот.

Как часто бывает в развитии науки, новое — это забытое старое. Ещё в 1686 г. Готфрид Лейбниц (Gottfried Wilhelm von Leibniz) за 250 лет до Гёделя в философской статье "Рассуждение о метафизике" [141] поставил вопрос: как отличить факты, которые можно описать каким-либо законом, от фактов, никакими законами не описываемых? Лейбниц переформулировал мысль Парменида и высказал простой постулат: *теория должна быть проще данных, на основе которых она построена. В противном случае она не нужна.* Что значит "проще"? Выражаясь современным языком: информация должна быть сжата, а её запись должна быть короче, чем исходные данные. Но это сжатие должно содержать правило, которое позволяет без потерь в случае необходимости развернуть сжатую информацию (см. раздел 5).

Наш мозг производит сжатие информации. Вопрос о количестве градаций в иерархических уровнях при создании алгоритмов поведения в живых системах пока остаётся открытым; по-видимому, их число в живой системе очень велико — сопоставимо с числом органелл или клеток, т.е. $10^{13} - 10^{17}$.

Итак, окружающий нас мир и мы сами как его часть устроены иерархически не только по законам классической двоичной логики, но и по законам смешанной троичной логики с дополнением вероятностной логики. В результате мозг работает в троичной системе: "да — нет — неопределённость с разной величиной вероятности". Мозг, живущий лишь в собственном виртуальном мире, оторванном от реального мира, — это мозг сумасшедшего. Синтезированные им модели иногда могут совпадать с реальностью, но это скорее исключение, чем правило.

4.3. Мозг, живущий в вероятностном мире

Будем называть мозг, живущий в вероятностном мире, "байесовским". Простейшая формула с учётом условной вероятности (формула Байеса) известна давно [142, 143]. При учёте условной вероятности $p(A|X)$ происходит увеличение наших знаний о событии А при получении новых сведений о событии X:

$$p(A|X) = \frac{p(X|A)p(A)}{p(X)}. \quad (21)$$

Вероятность $p(A)$ — это основанная на вере вероятность того или иного события А, т.е. фидуциальная (от лат. fides — вера, доверие) вероятность. Вероятность $p(X)$ — это вероятность появления новых сведений, которые могут изменить нашу априорную веру в вероятность реализации события А. Условная вероятность $p(A|X)$ позволяет измерить, в какой степени была правильна наша убеждённость в вероятности $p(A)$ соверше-

ния события А. Если человек абсолютно уверен, то вероятность $p(A) = 1$. Если, напротив, он абсолютно уверен, что этого никогда не случится, то $p(A) \rightarrow 0$. Большинство ситуаций занимает промежуточное положение между нулём и единицей. Приближение к 0 или к 1 изменяется по мере получения новых сведений с вероятностью $p(X)$.

Казалось бы, популярность формулы Байеса обусловлена тем, что она даёт нам возможность точно рассчитать вероятность, которая характеризует достижение цели при получении новых сведений. На этом основана концепция так называемого *идеального байесовского наблюдателя* — воображаемого существа, всегда использующего получаемые сведения наилучшим способом из всех возможных [144]. Однако это воображаемое существо в значительной степени напоминает другое воображаемое существо в термодинамике — демона Максвелла [9]. Напомним, что через *цену действия* (функцию от энергии и времени (см. выражение (2)) эти два воображаемых "существа" (Байеса и Максвелла) связаны между собою. Бесплатно и точно они не могут измерить величину и направление изменения вероятности при получении новых сведений. Но идеальный байесовский наблюдатель, в отличие от демона Максвелла, "*питается*" не только *энергией, но и временем, затрачиваемым на повторение* (крылатое выражение: *повторение — мать учения*).

В формуле Байеса по трём причинам при использовании условных вероятностей $p(X|A)$ надо учитывать следующие обстоятельства.

Во-первых, для достоверного получения вероятности некоторого события А требуется достаточно большое число испытаний. Только в этом случае значение вероятности $p(A)$ стремится к объективной оценке возможности наступления этого события.

Во-вторых, нужно быть уверенным, что вероятности наступления событий А и Х зависимы. Если $p(A|X)$ или $p(X|A)$ не изменяют вероятности наступления события $p(A)$, то события А и Х независимы.

В-третьих, в общем случае $p(A|X) \neq p(X|A)$. Цепочка условных вероятностей может быть асимметричной. При наличии события А событие Х имеет место с вероятностью $p(A|X)$, обратное утверждение, что при событии Х событие А имеет ту же вероятность, может оказаться неверным. *Процесс мышления человека имеет склонность игнорировать общую информацию о частоте событий и сосредоточиваться на специфической информации о конкретном, интересующем лишь его событии, искажая тем самым реальную ситуацию, описываемую вероятностями.* Например, поскольку рак у человека встречается в общей массе населения редко, уверенность человека в том, что он не болен раком, близка к единице, т.е. $p(A) \approx 1$. Если у человека обнаружена растущая опухоль (событие X), то вероятность получить положительный результат теста на рак составит $p(A) \approx \approx 1 - p(A|X)$. Причина — редкость появления злокачественной опухоли в множестве доброкачественных опухолей. Вероятность этого события порядка 15 %, т.е. $p(A|X) \leq 15\%$. Уверенность человека в том, что у него нет рака, также изменится незначительно, т.е. вероятность станет $p(A) \approx 85\%$. Наконец, при получении положительного результата теста на рак с помощью современного метода диагностики, который имеет точность 90 %, вероятность для этого человека $p(A)$ при-

близится к 90 % (крылатое выражение: "надежда умирает последней").

Для внешнего наблюдателя, например врача, $p(A)$ — вероятность встретить больного раком среди всего множества людей X, проходящих у него профилактический осмотр, очень мала, $p(X|A) \leq 1\%$. Другими словами, при приравнивании двух условных вероятностей могут возникать ошибки. Они связаны с игнорированием при постановке задачи исходной частоты встречи того или иного события, так называемого *его базового процента*. Это понятие является столь важным, что поясним его подробнее ещё на одном примере.

Классическим и актуальным примером служит автоматизированный поиск террористов (Ошибка базового процента <https://ru.wikipedia.org>). Допустим, в городе с миллионным населением проживает 100 преступников-террористов и 999900 законопослушных мирных жителей. Пытаясь найти террористов, полиция установила в городе систему тревоги с камерами наблюдения, снабжёнными программами автоматического распознавания лиц на основе архива лиц террористов в памяти распознающих устройств. Программное обеспечение имеет две возможные ошибки: пропустить цель (террориста) с вероятностью 1 % и с той же вероятностью 1 % принять за террориста похожего на него мирного жителя (ложная тревога). Когда камера видит террориста, вероятность его не распознать 1 %, следовательно, правильный сигнал тревоги при появлении террориста прозвучит в 99 % случаев и будет отсутствовать в 1 % случаев. Когда же камера видит жителя города, не принадлежащего к разыскиваемым террористам, но похожего на одного из них, она по ошибке может дать сигнал тревоги в 1 % случаев.

Итак, прозвучал сигнал тревоги. Какова вероятность того, что вызвавший сигнал человек — террорист?

Если не учитывать указанные выше три условия, то можно предположить, что вероятность равна 99 %. Хотя такое предположение кажется правильным, на самом деле оно ложно. Правильный ответ — около 1 %. Заблуждение возникает вследствие смешения природы двух разных процентов ошибки. Количество случаев отсутствия сигнала тревоги на 100 террористов и количество тревожных звонков на 100 мирных жителей являются несвязанными величинами. Ошибку легко понять, если рассмотреть предельный случай, когда подобная система распознавания лиц установлена в городе, где террористов вообще нет. Ошибочно сигнализация тревоги будет срабатывать один раз на каждые 100 мирных жителей города, следовательно, все сигналы будут ложными. Но, поскольку об этом априорно никто не знает, полиция будет вынуждена приезжать по ложному сигналу. Таким образом, в 100 % случаев тревога звучит не в связи с обнаружением террориста, а как ложная. Теперь представим, что все 10^6 жителей города пройдут перед камерой. На ~99 из 100 террористов сработает тревога, но она также сработает примерно на 9999 из 999900 законопослушных жителей, поскольку их значительно больше. В сумме сигнал тревоги прозвучит при прохождении перед камерой около 10098 человек, из которых лишь ~99 будут террористами. Таким образом, вывод о том, что человек, вызвавший сигнал тревоги, является террористом, справедлив в 99 случаях из 10098, что меньше 1 % и намного ниже начального предположения в 99 %.

Редкие явления часто приводят к ошибкам, когда к ним пытаются применить формулу Байеса. Джереми Вулф (Jeremy Wolfe) и его коллеги из Бостона экспериментально проанализировали работу служб безопасности, сканирующих вещи в аэропорту в поисках ножей, взрывчатых веществ и других запрещённых к провозу предметов. Когда искомые предметы встречались часто, испытания показывали результат, соответствующий формуле Байеса. Служащие не заметили лишь около 7 % запрещённых к провозу предметов. Когда такие предметы стали встречаться редко, результат оказался неутешительным. В одном из экспериментов испытуемые не заметили больше 50 % искомых предметов, находившихся лишь в 1 % досмотренных вещей (цитируется по [144]).

Байесовская формула начинает работать лишь после повторений и набора статистики, т.е. при уточнении виртуальной модели внешней среды, имеющей место в нашей голове или в компьютере на основе предыдущего опыта, полученного при взаимодействии с реальным миром. В противном случае мы попадаем в ловушку ложных корреляций [145]. Неистребимое стремление человека найти порядок во внешней среде, даже там, где его нет, может приводить к рождению суеверий [146], а в настоящее время — к формированию гипотез, в которых желаемое выдаётся за действительное [147].

Восприятие реального мира начинается изнутри мозга с априорного убеждения, представляющего собой виртуальную модель мира, в которой объекты и связи между ними занимают "определенное" положение в пространстве и времени. Генетика и прошлый опыт определяют эту модель. Пользуясь ею, наш мозг может предсказать, какие сигналы должны поступать в наши глаза, уши и другие сенсорные системы, но может и ошибаться. Предсказания мозга сравниваются с реальными сигналами, при этом, разумеется, обнаруживаются ошибки. Характер ошибок говорит мозгу, как улучшить собственную модель внешней среды. В итоге цикл повторяется вновь и вновь, до тех пор пока ошибки не станут малы. Для этого обычно достаточно небольшого числа таких циклов (в пределе достаточно двух), на которые мозгу может потребоваться минимально 100 мс (это предельное время простой двигательной реакции). *Наше тело служит нам, чтобы познавать окружающий мир.* Восприятие — это цикл, в котором предсказание постоянно проверяется действиями. Эта способность отсутствовала в ранних системах ИИ и КАР.

Отметим, что после внесения уточнения мозг получает новое представление о мире и может снова повторить ту же процедуру, сделав новое предсказание о характере событий, отслеживаемых органами чувств. С каждым повтором этого цикла ошибка в предсказаниях уменьшается. Когда ошибка оказывается достаточно малой, наш мозг может ориентироваться во внешней среде. Несложно определить, как быстро возрастает величина вероятности $p(A|X)$ при повторных уточнениях.

Вернёмся к приведённому выше примеру диагностики рака: если пациент ещё раз пройдёт диагностический тест, то может кардинально уменьшить неопределённость. Если тест тоже будет положительным, то по теореме Байеса вероятность $p_2(A)$ будет равна результату первого теста — 0,5. Надёжность диагностики осталась без изменения: $p(X|A) = 0,99$. Вероятности получить положительную или отрицательную реакции на тест равны,

т.е. $p_2(X) = 0,5$. Подставив эти значения в выражение (21), получим

$$p(A|X) = \frac{0,99 \times 0,5}{0,5} = 0,99. \quad (22)$$

Следовательно, вероятность наличия у пациента рака будет равна 99 % вместо прежних 50 %. Как показывает этот пример, повторение быстро распутывает клубок из ложных и истинных ситуаций и уточняет ответ.

Относительно повторений можно отметить особую роль учителя. "Учитель" также уменьшает долю ошибок [61]. Очевидно, что необязательно приобретать опыт, общаясь лишь самостоятельно с внешней средой. Казалось бы, программы ИИ лишены этого недостатка, поскольку в них предварительно программистом в качестве учителя заложен его собственный опыт. Однако если алгоритм программы нацелен на выигрыши в какой-либо ситуации, но не оговорены правила потерь, которых надо избежать, то робот будет действовать как таран (как слон в посудной лавке). Он физически уничтожит всё, что мешает достигнуть цели. Понятия этики, гуманизма и нравственности у робота не существует. Следовательно, программист должен предусмотреть в механизме "обучения" не только кратчайший путь к цели, но и цену, которую не следует превысить на пути достижения цели. При достижении границы начинает действовать теорема Гёделя, приводящая к появлению ситуаций, в которых истинная и противоположная ей ложная цели спутаны и неразличимы или, другими словами, нарушена устойчивость решения поставленной программистом задачи.

Программа точно и быстро выполняет именно то и только то, что в ней сформулировано, а не то, что имел в виду или предполагал получить программист в результате её выполнения. К этой же мысли ещё в середине XX в. пришёл Норберт Винер, изложив её в дополнительных главах книги [148]. Наличие неопределённостей чревато для создателя сюрпризами со стороны ИИ.

5. Заключение

5.1. Математическое сжатие информации

Продолжая метафору с развитием логики и математики, высажем гипотезу о существовании пока ещё не открытых супералгоритмов. Будем исходить из следующего постулата: *посредством кодирования можно сжимать информацию по объёму*. Например, для рядов натуральных чисел, арифметической или геометрической прогрессии или чисел Фибоначчи и многих других сходящихся рядов не требуется запоминать весь бесконечный ряд. Необходимо лишь запомнить формулу вычисления любого N -го члена ряда. Эта формула позволяет построить подобные ряды любой протяжённости. Следовательно, такие ряды содержат мало информации.

Однако, возможно, существуют некоторые бесконечные ряды, казалось бы, с неповторяющимися последовательностями чисел, но имеющие некоторые локальные особенности. Приведём вычисление ряда числа $\pi = 3,14159\dots$

Существует множество алгоритмов для его последовательного нахождения: стохастический метод Монте-Карло, интеграл Пуассона или формула Франсуа Виета (François Viète). Остановимся на последней.

Ф. Виет ещё в XVI в. при вычислении числа π вывел формулу для определения бесконечной последовательности членов этого ряда [149]:

$$\frac{2}{\pi} = \sqrt{\frac{1}{2}} \times \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}} \times \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}}} \times \dots \quad (23)$$

Это выражение указывает на то, что вычисление π определяется рекуррентной процедурой. Очевидно, что время вычисления таким способом возрастает с увеличением номера N члена ряда.

В 1997 г. Дэвид Бэйли, Питер Борвайн и Саймон Плафф, используя идею о локальной сходимости фрагментов некоторых рядов, получили другую, более изящную и менее затратную по времени, формулу для вычисления N -го числа в последовательности π [150]. С её помощью можно вычислить любой N -й знак числа π с ограниченным вычислением предыдущих и последующих ближайших членов ряда:

$$\pi = \sum_{k=0}^{\infty} 16^{-k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right). \quad (24)$$

Если требуется вычислить по этой формуле N -й знак числа π , то обе её части надо умножить на 16^N , преобразовав тем самым множитель, стоящий перед скобками в правой части (24), в 16^{N-k} . Затем вычисляется сумма нескольких ближайших членов ряда. Причём необязательно вычислять много членов: из формулы видно, что по мере возрастания k значение 16^{N-k} быстро убывает, так что последующие члены не будут оказывать сильного влияния на значение искомого N -го члена ряда. Формула Бэйли–Борвайна–Плаффа была найдена с помощью алгоритма PSLQ⁷. Таких находок, сжимающих информацию, в истории развития математики немало: например, можно упомянуть формулу Стирлинга (или формулу Муавра–Стирлинга) для вычисления факториала или гамма-функции [151].

Следовательно, можно высказать гипотезу, что над множеством числовых рядов существует следующий иерархический уровень, а именно множество алгоритмов, с помощью которых можно сжимать информацию о рядах. Продолжая эту мысль, можно также допустить, что над множеством алгоритмов находится другой математический мир. Неважно, как его называть, — например множество супералгоритмов или метаалгоритмов. Супералгоритмы этого уровня позволяют анализировать информацию об алгоритмах. В таком *иерархическом варианте* математика не является замкнутой системой. У неё отсутствуют границы, что открывает простор для творчества в области сжатия информации путём движения по ступеням иерархии. Чем выше иерархический подъём, тем меньший интервал времени требуется для решения задачи нижних уровней. Непрерывное расширение границ снимает ограничения, сформулированные Гёделем.

Для биофизики из этой метафоры следовали бы два важных вывода.

⁷ PSLQ — алгоритм анализа на языке дискретных вопросов — один из вариантов поиска целочисленных отношений между набором слагаемых суммы вещественных чисел x_i с коэффициентами a_i . Эта сумма приравнивается к нулю. Программа вычисления либо находит целочисленное соотношение между a_i и x_i , либо определяет, что такого соотношения нет (см. PSLQ Algorithm на веб-сайте Wolfram MathWorld, созданном Эриком Вайсштайном (Eric W. Weisstein)).

Первый вывод: для того чтобы подняться на следующий иерархический уровень, необходимо найти способ выйти за границу предыдущего уровня в другое информационное пространство. Кроме того, подъём по уровням должен быть последовательным — нельзя пропускать уровни-ступеньки.

Второй вывод — это требование устойчивости алгоритма к начальным условиям [27], поскольку прошлое (начальные условия), совмещённое с граничными условиями (т.е. с настоящим), определяет будущее.

Тем не менее фазовые переходы разных типов нарушают правило эволюции, так как они рождаются на основе скачков. Если на нижних уровнях скачки могут быть полезными, поскольку создают *озарения*, порождают творчество и способствуют прогрессу, то на верхних уровнях скачки могут приводить к гибели части или всей системы, а в биосфере — к замещению одного вида живых организмов другим видом. Виртуальный мир нашего мозга тем и хорош, что может ограничивать вероятность скачков (крылатое выражение: *семь раз отмерь и один раз отрежь*).

5.2. Выход за границы обычного пространства в новое информационное пространство

При решении интеллектуальных задач мозг человека пытается найти нестандартные решения. Приведём один простой пример "озарения" — притчу о *смышильном ребёнке*. Отец, которому пришлось сидеть с маленьким ребёнком, попытался заняться своими делами, но чтобы ребёнок ему не мешал, отвлёк его игрой: вырвал из журнала страницу с изображением карты мира и порвал её на куски. Затем отдал все обрывки сыну со словами: "Я свожу тебя в зоопарк, если ты соберёшь из них карту". Ребенок согласился. Отец был уверен, что сын будет заниматься этим не менее полутора часов и оставит его в покое. Однако через десять минут сын собрал карту. Отец удивлённо спросил: "Как тебе это удалось? Ты знаешь географию?" Малыш ответил: "Папа, ты не видел, что изображено на обратной стороне: там — рисунок человека. Я собрал рисунок человека, а с другой стороны получилась карта. А теперь идём в зоопарк!" (Сюжет из книги [152].)

Ребёнок нашёл решение трудной для него задачи посредством выхода в другое понятное ему на основе прошлого опыта информационное пространство и обнаружил связь с поставленной перед ним задачей. Тем самым он произвёл сжатие информации, используя уже накопленный им опыт, и сэкономил время при решении поставленной перед ним новой задачи.

В приведённой в разделе 5.1 задаче определения числа π было продемонстрировано, что временной интервал при сжатии алгоритмической информации зависит от уровня выбранного языка программирования. Например, требуемое время работы алгоритма Δt при вычислении любого N -го значения числа π пропорционально величине порядкового номера искомого знака N , а объём требуемой памяти для программы пропорционален $\ln N$ [150]. Поскольку

$$N \gg \ln N, \quad (25)$$

имеет место очевидный, но важный вывод: память и опыт (знания о прошлом) существенно экономят время решения новых задач.

При создании моделей поведения внешней среды возможность сжатия информации зависит от возможности интегрирования рядов или матриц с различными

свойствами. Обзор теории матричных моделей с точки зрения её связи с интегрируемыми иерархиями приведён в статье [153]. Развитие понятия полной интегрируемости динамической системы, в частности для нужд квантовой механики, дано Л.Д. Фаддеевым [154].

Тем не менее если ряды являются расходящимися, то существуют ситуации, в которых возникают алгоритмические проблемы сжатия информации за конечный интервал времени. А.Н. Колмогоров в середине XX в. предложил взять измеренную, например, в битах длину программы или (что то же самое) алгоритм вычислений (I), который переводит некоторую последовательность $\{Y_j\}$ в последовательность $\{X_i\}$. Такой переход позволяет решить задачу сжатия информации в общем виде [155, 156]. Если сложность задачи невысока (т.е. сжатие возможно), то I существенно меньше длины последовательности $\{X_i\}$. В противном случае процесс не поддаётся алгоритмическому сжатию. Для таких процессов $I \sim N$ -алгоритм перевода $\{X_i\} \rightarrow \{Y_j\}$ сводится к посимвольному запоминанию всей последовательности $\{Y_j\}$.

Пока остаётся открытым вопрос: существуют ли в природе абсолютно случайные несжимаемые процессы или их существование — это математическая абстракция? Возможно, человечество пока находится на недостаточно высокой ступени иерархии алгоритмов, для того чтобы с уверенностью ответить на данный вопрос. Эта проблема волновала в начале XX в. А. Эйнштейна. Всем известно его высказывание в дискуссии с Н. Бором: "Бог не играет в кости" [157]. Н. Бор изящно вышел из этой дискуссии, введя *принцип дополнительности* [158].

5.3. Предел сходства Робота и Человека

Выяснить точно предел сходства Робота и Человека невозможно. Эта область принадлежит не физикам или биофизикам, а писателям-фантастам и футурологам. В *terra incognita* невозможно отделить факты от вымысла. Пытаясь предсказать будущее, мы вступаем на территорию неопределённости. Тем не менее попытаемся сформулировать условия, при которых может появиться *универсальный креативный робот*. Как отмечалось ранее, креативность — это умение самостоятельно формулировать цель поведения и принимать творческие решения для её достижения. Современный робот не может сам себе поставить цель. Её ставит создатель робота, закладывая её в виде программы в ИИ робота.

У самого человека, по-видимому, цель появилась в условиях изменения среды на нашей планете в ходе эволюционного отбора его адаптационных механизмов, направленных на выживание вида. По данным истории развития человечества эпоха возникновения гоминидов началась 4,5 млн лет тому назад и продолжалась в течение $\Delta t = 2,8$ млн лет. Существенным моментом развития стало время, когда в нижнем палеолите в Африке появился наш предок *Homo habilis* — человек умелый. Затем произошёл качественный переход в эволюции человека и начался процесс социального развития. Человек постепенно заселил весь Земной шар. Таким образом, можно весьма условно принять время формирования современного человека в ряду гоминидов в 3,5 млн лет [159–161] (рис. 11).

Какие изменения в мозге отделяют нас от наших предков? Антропологические находки сообщают, что у австралопитеков объём мозга в среднем составлял 459 см³, приблизительно как у некоторых шимпанзе. У человека прямоходящего (*Homo erectus*, 1,6 млн лет назад) мозг увеличился в среднем до 930 см³. У *Homo sapiens* (200 тыс. лет назад) объём мозга достиг в среднем

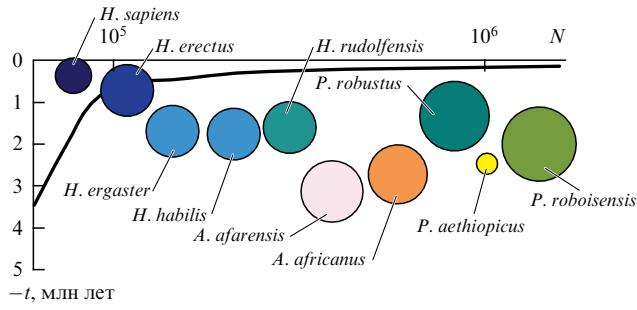


Рис. 11. Эволюционные этапы (t — прошедшее время, N — численность популяции), которые привели к появлению современного человека в результате мутационных изменений гоминидов, направленных на адаптацию к изменяющейся внешней среде. Диаметр окружности соответствует размерам численности популяции, положение окружности — времени существования вида. Наш вид *Homo Sapiens*, появившийся приблизительно 200 тыс. лет назад, обладал наилучшими способностями к адаптации и стал видом, занявшим почти все регионы планеты, достигнув сегодня численности свыше 7 млрд [162, 163].

1330 см³. Но главное заключалось не только и не столько в увеличении всего объёма мозга или изменении его структуры, или в особенностях фронтальной коры (поля Бродмана 10), которая участвует (по данным ЭЭГ) в построении виртуальной модели внешней среды на основе генетически заложенных инстинктов и совершенствуется на основе информации, поступающей от рецепторов (хотя фронтальная зона у *Homo sapiens* увеличилась по сравнению с таковой у других приматов более чем в два раза), а, по-видимому, в том, что расстояние между нейронами возросло на 50 %, предоставив больше пространства для роста дендритов, обеспечивающих память и совершенствующих виртуальную модель внешней среды. Всё это, скорее всего, и заложило возможности для появления и совершенствования языка общения на социальном уровне, существенно расширив динамические возможности при формировании виртуальной модели среды. Открылся новый путь к использованию неформальной логики и творчеству посредством сопоставления и объединения информации, которая поступала не только от разных сенсорных систем в кору мозга, но и от совместных действий коллективов людей и обмена информацией и опытом между ними [164, 165].

Развитие мозга привело к тому, что сформировалась стратегия выживания на основе таких индивидуальных свойств человека, как: страх смерти и стремление избегать чувства боли, способность к оценке возникающих угроз для продолжения рода, защита потомства, добыча пищи и поиск комфортных зон обитания для себя и своих близких. Эта стратегия была дополнена формированием кооперативного поведения людей в соответствии с наблюдением за изменениями внешней среды, запоминанием таких изменений, развитием широкого обмена опытом между людьми.

Можно ли в виде алгоритма заложить виртуальную модель мира при создании самообучающегося робота в его искусственный интеллект? Пока это невозможно.

Во-первых, мы сами не знаем в полной мере такую модель. Процесс биологической эволюции экспериментально воспроизвести не удается. Это некорректная обратная физическая задача. Когда факт возникновения живой материи свершился, обсуждение истории свершения носит гипотетический характер. Особенность некорректно поставленных задач состоит в том, что по "уликам", которые мы наблюдаем в настоящее время,

требуется восстановить картину развёртывания процесса во времени в далёком прошлом. Такие задачи очень чувствительны к начальным условиям, которые мы представляем лишь гипотетически [166–168]. Пока мы имеем лишь один вариант эволюции, и то не в полной мере, — это развитие живой материи на Земле. Если произвести расчёт по формуле Байеса, то, как показано выше, достоверность одиночного наблюдения может быть лишь порядка 50 %. Поэтому не утихают споры вокруг панспермии (т.е. о том, что жизнь на Землю занесена из космоса) [169, 170].

Во-вторых, у робота нет генетического прошлого и причинно-следственных отношений с ним. Тем не менее предположим, что человек создал:

1) робота с "мозгом" на квантовом компьютере, работающим со скоростью света, и наделил его телом, насыщенным разными рецепторами;

2) замкнутый цикл производства (без участия людей) самостоятельно размножающихся роботов (теоретическую возможность такого цикла обсуждал ещё в 1960-х годах Джон фон Нейман [171]).

Что произойдёт далее? Для последующего развития роботов человек уже не нужен. Если имеются источники сырья и энергии, то процесс будет развиваться сам по себе. Сколько лет потребуется для того, чтобы мозг робота прошёл путь эволюции мозга человека (с учётом переданного ему человеком своего опыта) до этапа самостоятельного формирования собственных целей?

Если принять, что для развития от *Человека умелого* (*Homo habilis*) до *Человека мыслящего* (*Homo sapiens*) потребовалось не более 3,5 млн лет, а среда нашей планеты в течение всего этого периода сохраняла сравнительную стабильность, то оценка интервала времени Δt_R , за который робот в результате своей эволюции достигнет уровня способности человека в формулировке целей, будет соответствовать пропорции

$$\frac{v}{c} = \frac{\Delta t_R}{\Delta t_H}, \quad (26)$$

где v — максимальная скорость передачи информации по аксонам и дендритам между нейронами мозга человека, примем значение $v \approx 25 \text{ м с}^{-1}$; c — максимальная скорость передачи информации по связям между микропроцессорами "мозга" КАР (скорость света), $c \approx \approx 3 \times 10^8 \text{ м с}^{-1}$; Δt_H — интервал времени формирования мозга человека в условиях Земли, примем условно, что $\Delta t_H \approx 3,5 \text{ млн лет} \approx 10^{14} \text{ с}$; Δt_R — интервал времени формирования креативного робота, который сможет самостоятельно ставить цели своего существования и развития. Из пропорции (26) можно найти значение Δt_R :

$$\Delta t_R \leq \frac{v}{c} \Delta t_H \approx 9,2 \times 10^6 \text{ с} \approx 3,7 \text{ мес.} \quad (27)$$

Таким образом, чтобы самообучающийся робот мог действовать как человек, самостоятельно формируя свою цель, он должен самосовершенствоваться в течение всего трёх-четырёх месяцев. Появление мира КАР в этом случае следует рассматривать как продолжение биологической эволюции, но на другой материальной основе. В зависимости от того, в какой сценарий эволюции мы верим, такой результат и получим (подобно вычислению вероятности по формуле Байеса без повторения). Однако человечеству ввиду опасности для него не нужны роботы, которые способны самостоятельно формировать цели. Такие роботы будут грустить, любить, радоваться, завидовать, злиться, драться, сомневаться,

мстить и даже сходить с ума. Человек для них в лучшем случае будет подобен домашнему животному.

6. Выводы

На основе изложенного можно сделать пять выводов.

I. Развитие робототехники на ближайший период будет продолжаться по уже существующим траекториям: 1) создания "умного пространства" — роботизированных заводов и территорий; 2) создания "умных и безопасных" городов и регионов; 3) расширения ареала обитания человека в третьем измерении (Sky City, Akva City, CyberVillage); 4) расширения масштабов знания, т.е. освоения астродиапазонов (роботизированные полёты в космос) и нанодиапазонов (роботизированные "полёты" внутри организмов).

II. Мозг человека создаёт *виртуальную модель внешней среды на основе детерминированной логики*. Хотя, по сути, она — вероятностная, так как может быть ошибочной. Эта модель на всех иерархических уровнях непрерывно уточняется на основе генетики, приобретаемого опыта (в результате действия всей совокупности сенсорных систем организма и подвижности тела), включая память о прошлом опыте, и социальных взаимодействий.

III. Хотя в термин "androïd" можно заложить *внешнюю похожесть* робота на человека, однако это не существенный признак. Дизайн роботов определяется спецификой их применения. Внешняя схожесть роботов с человеком является скорее экзотикой, и в большинстве случаев она нецелесообразна.

IV. В искусственном интеллекте роботов не должны формироваться собственные цели. Указание цели всегда должно сохраняться за человеком. Это и есть *условный предел сходства между роботами и человеком*. Программисты в искусственный интеллект роботов должны закладывать лишь алгоритмы, способствующие оптимальному выбору пути к достижению той цели, которую поставил Человек-создатель робота. Постановка цели должна сохраняться за человеком.

V. Логично предположить, что создание КАР — это опасный вариант антропогенного преобразования *человеческим разумом* неживой материи. В случае ошибки он может быть и последним в судьбе человечества. У Робота, самостоятельно формирующего цель, может появиться любая "мысль", вплоть до уничтожения своего создателя.

Благодарности

В заключение я хочу поблагодарить Е.П. Хижняка и особо В.И. Архипова и А.Б. Медвинского за ряд ценных замечаний, а также [Д.С. Чернавского], Г.Н. Борисюка и В.Г. Яхно, обсуждение с которыми на семинарах и симпозиумах идей, близких к рассмотренным в данной статье, привели меня к мысли об актуальности этой темы. Хочу отметить роль М.И. Рабиновича, две работы которого с соавторами: "Нелинейная динамика мозга: эмоции и интеллектуальная деятельность" [14] (2010 г.) и "Математика сознания" [30] (2017 г.) — повлияли на моё желание написать эту статью.

Список литературы

1. Иванецкий Г Р *Найдутся ли ответы?* (Сер. Знак вопроса, № 6) (М.: Знание, 1991)
2. Гурфинкель В С и др. *Биоэлектрическое управление* (М.: Наука, 1972)
3. Кобринский А Е и др. *ДАН СССР* **117** (1) 78 (1957)
4. Бабский Е Б, Ульянинский Л С *Электрическая стимуляция сердца* (М.: Медицина, 1961)
5. Кузин А М (Отв. ред.) *Биологические аспекты кибернетики* (М.: Изд-во АН СССР, 1962) Сб. работ
6. Siegwart R, Nourbakhsh I R, Scaramuzza D *Introduction to Autonomous Mobile Robots* 2nd ed. (Cambridge, Mass.: MIT Press, 2011)
7. Joseph I "Step into a new world — Virtual Reality (VR)", Basic Concepts of Virtual Reality along with Research Challenges Explained in Simple Words (2016)
8. Warneke B et al. *Computer* **34** 44 (2001)
9. Иванецкий Г Р и др. *УФН* **168** 1221 (1998); Ivanetskii G R et al. *Phys. Usp.* **41** 1115 (1998)
10. Drexler K E *Nanosystems: Molecular Machinery, Manufacturing, and Computation* (New York: Wiley, 1992)
11. New Technology (Tokyo: NEDO, 1997)
12. Иванецкий Г Р, Воробьев С И *Биофизика* **41** 178 (1996)
13. Рабинович М И *УФН* **139** 363 (1983); Rabinovich M I *Sov. Phys. Usp.* **26** 186 (1983)
14. Рабинович М И, Мюзинолу М К *УФН* **180** 371 (2010); Rabinovich M I, Muezzinoglu M K *Phys. Usp.* **53** 357 (2010)
15. Борисюк Г Н и др. *УФН* **172** 1189 (2002); Borisyuk et al. *Phys. Usp.* **45** 1073 (2002)
16. Менский М Б *УФН* **175** 413 (2005); Menskii M B *Phys. Usp.* **48** 389 (2005)
17. Клинишов В В, Некоркин В И *УФН* **183** 1323 (2013); Klinshov V V, Nekorkin V I *Phys. Usp.* **56** 1217 (2013)
18. Масленников О В, Некоркин В И *УФН* **187** 745 (2017); Maslenikov O V, Nekorkin V I *Phys. Usp.* **60** 694 (2017)
19. Васильев В А, Романовский Ю М, Яхно В Г *УФН* **128** 625 (1979); Vasil'ev V A, Romanovskii Yu M, Yakhno V G *Sov. Phys. Usp.* **22** 615 (1979)
20. Абарбанель Г Д и др. *УФН* **166** 363 (1996); Abarbanel H D et al. *Phys. Usp.* **39** 337 (1996)
21. Чернавский Д С *УФН* **170** 157 (2000); Chernavskii D S *Phys. Usp.* **43** 151 (2000)
22. Смолянинов В В *УФН* **170** 1063 (2000); Smolyaninov V V *Phys. Usp.* **43** 991 (2000)
23. Иванецкий Г Р *УФН* **180** 337 (2010); Ivanetskii G R *Phys. Usp.* **53** 327 (2010)
24. Иванецкий Г Р *Виражи закономерностей* (М.: Наука, 2011)
25. Цыганов М А и др. *УФН* **177** 275 (2007); Tsyganov M A et al. *Phys. Usp.* **50** 263 (2007)
26. Tsyganov M A, Biktshev V N *Phys. Rev. E* **70** 031901 (2004)
27. Иванецкий Г Р *УФН* **187** 757 (2017); Ivanetskii G R *Phys. Usp.* **60** 705 (2017)
28. Eigen M "Self-organization of matter and the evolution of biological macromolecules" *Naturwissenschaft* **58** 465 (1971); Пер. на русск. яз.: Эйген М *Самоорганизация материи и эволюция биологических макромолекул* (М.: Мир, 1973)
29. Eigen M, Schuster P *The Hypercycle: A Principle of Natural Self-Organization* (Berlin: Springer-Verlag, 1979); Пер. на русск. яз.: Эйген М, Шустер П *Гиперцикл* (М.: Мир, 1982)
30. Рабинович М, Варона П *Изв. вузов. Приклад. нелинейная динамика* **25** (3) 3 (2017)
31. Анищенко В С и др. *УФН* **169** 7 (1999); Anishchenko V S et al. *Phys. Usp.* **42** 7 (1999)
32. Зельдович Я Б и др. *УФН* **152** 3 (1987); Zeldovich Ya B et al. *Sov. Phys. Usp.* **30** 353 (1987)
33. Анищенко В С, Астахов С В *УФН* **183** 1009 (2013); Anishchenko V S, Astakhov S V *Phys. Usp.* **56** 955 (2013)
34. Penrose R *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics* (Oxford: Oxford Univ. Press, 1989); Пер. на русск. яз.: Пенроуз Р *Новый ум короля (О компьютерах, мышлении и законах физики)* (М.: Едиториал УРСС, 2003) с. 14
35. Смолуховский М *УФН* **7** 329 (1927)
36. Смолуховский М *УФН* **93** 724 (1967); Smoluchowski M "Gültigkeitsgrenzen des zweiten Hauptsatzes der Wärmetheorie", in *Vorträge über kinetische Theorie der Materie und der Elektrizität* (Mathematische Vorlesungen an der Universität Göttingen, Vol. 6) (Leipzig: B.G. Teubner, 1914) p. 87
37. Анохин П К *Вопросы психологии* (6) 16 (1955)
38. Соловьев В С "Гегель, Георг Фридрих Вильгельм", в кн. *Энциклопедический словарь Брокгауза и Ефрона* Т. 82 (СПб., 1890–1907)

39. Бернштейн Н А *Очерки о физиологии движений и физиологии активности* (М.: Медицина, 1966)
40. Гумилёв Л Н *Этногенез и биосфера Земли* (СПб.: Кристалл, 2001)
41. Сенина И Н, Тютюнов Ю В *Журн. общей биологии* **63** 494 (2002)
42. Говорухин В Н, Моргулис А Б, Тютюнов Ю В *Докл. РАН* **372** 730 (2000)
43. Ardit R et al. *Theor. Popul. Biol.* **59** 207 (2001)
44. Lightfoot E N *Transport Phenomena and Living Systems. Biomedical Aspects of Momentum and Transport* (New York: Wiley, 1974); Пер. на русск. яз.: Лайтфут Э *Явление переноса в живых системах. Биомедицинские аспекты переноса количества движения и массы* (М.: Мир, 1977)
45. Воларович М П *Изв. АН СССР Сер. физ.* **11** (1) 7 (1947)
46. Бернульи Д *Гидродинамика, или записки о силах и движениях жидкостей* (Сер. Классики науки) (Л.: Изд-во АН СССР, 1950)
47. Schuster H G, Just W *Deterministic Chaos* (Weinheim: Wiley-VCH, 2005); Пер. на русск. яз.: Шустер Х Г *Детерминированный хаос* (М.: Наука, 1988)
48. Иваницкий Г Р, Деев А А *Биофизика* **54** 545 (2009)
49. Иваницкий Г Р, Деев А А, Цыганов М А *Вестник РАН* **81** 1008 (2011); Ivanitskii G R, Deev A A, Tsyanov M A *Herald Russ. Acad. Sci.* **81** 637 (2011)
50. Иваницкий Г Р, Куниский А С *Исследование микроструктуры объектов методами когерентной оптики* (М.: Энергия, 1981)
51. Гарив А М, Куниский А С, Мячин Е Т *Квантовая электроника* **2** 2303 (1975); Gariev A M, Kuniskii A S, Myachin E T *Sov. J. Quantum Electron.* **5** 1254 (1975)
52. Иваницкий Г Р, Куниский А С, в сб. *Биофизика сложных систем и радиационных нарушений* (Под ред. Г М Франка) (М.: Наука, 1977) с. 129
53. Feynman R P, Leighton R B, Sands M *The Feynman Lectures on Physics* Vol. 3 (Reading, Mass.: Addison-Wesley, 1965); Пер. на русск. яз.: Файнман Р, Лайтон Р, Сэндс М *Фейнмановские лекции по физике* Т. 3 (М.: Мир, 1967) с. 179
54. Prosser C L *Comparative Animal Physiology* (Philadelphia: W.B. Saunders Company, 1973); Пер. на русск. яз.: Прессер Л, в кн. *Сравнительная физиология животных* Т. 2 (Под ред. Т М Турпаева) (М.: Мир, 1977) Гл. 9, с. 86
55. Lorenz C *Blätter Deutsche Philosophie* **15** 94 (1941); Пер. на русск. яз.: в сб. *Эволюция. Язык. Познание* (Отв. ред. И П Меркулов) (М.: Языки русской культуры, 2000)
56. Wiener N *Cybernetics or Control and Communication in the Animal and the Machine* (New York: John Wiley and Sons, 1948); Пер. на русск. яз.: Винер Н *Кибернетика, или управление и связь в животном и машине* (М.: Советское радио, 1958)
57. Колмогоров А Н *Успехи математических наук* **5** (4) 175 (1953)
58. Колмогоров А Н "Автоматы и жизнь", в кн. *Возможное и невозможное в кибернетике* (Под ред. А И Берга, Э Я Кольмана) (М.: Наука, 1964)
59. Turing A *Mind* **59** 433 (1950)
60. Бонгард М М *Проблема узнавания* (М.: Физматгиз, 1967); Пер. на англ.: Bongard M M *Pattern Recognition* (New York: Spartan Books, 1970)
61. Chomsky N *Syntactic Structures* (Berlin: Mouton, The Hague, 1957)
62. Цетлин М Л *Исследование по теории автоматов и моделированию биологических систем* (М.: Наука, Физматлит, 1969)
63. Simon H A *Models of Man* (New York: John Wiley, 1957)
64. Simon H A *Proc. Am. Philos. Soc.* **106** 467 (1962)
65. Бонгард М М *Биофизика* **4**(2) 17 (1961)
66. von Neumann J, Morgenstern O *Theory of Games and Economic Behavior* (Princeton: Princeton University Press, 1944); Пер. на русск. яз.: фон Нейман Дж, Моргенштерн О *Теория игр и экономическое поведение* (М.: Наука, 1970)
67. Brooks R A *Robotics Autonomous Syst.* **6** 139 (1990)
68. Vardi M Y *Commun. ACM* **55** 5 (2012)
69. Hopfield J J *Proc. Natl. Acad. Sci.* **79** 2554 (1982)
70. Иваницкий Г Р *Убегающее время* (М.: Наука-пресс, 2001)
71. Rosenblatt F *Principles of Neurodynamic: Perceptrons and the Theory of Brain Mechanisms* (Washington: Spartan Books, 1962); Пер. на русск. яз.: Розенблатт Ф *Принципы нейродинамики: Перцептроны и теория механизмов мозга* (М.: Мир, 1965)
72. Minsky M, Papert S *Perceptrons* (Oxford: MIT Press, 1969); Пер. на русск. яз.: Минский М, Пейперт С *Перцептроны* (М.: Мир, 1971)
73. Осовский С *Нейронные сети для обработки информации* (М.: Финансы и статистика, 2002)
74. Иваницкий Г Р, Медвинский А Б, Цыганов М А *УФН* **161** (4) 13 (1991); Ivanitskii G R, Medvinskii A B, Tsyanov M A *Sov. Phys. Usp.* **34** 289 (1991)
75. Иваницкий Г Р, Медвинский А Б, Цыганов М А *УФН* **164** 1041 (1994); Ivanitskii G R, Medvinskii A B, Tsyanov M A *Phys. Usp.* **37** 961 (1994)
76. Haken H *Synergetik* (Berlin: Springer-Verlag, 1978); Пер. на русск. яз.: Хакен Г *Синергетика* (М.: Мир, 1980)
77. Prigogine I *From Being to Becoming: Time and Complexity in the Physical Sciences* (New York: W.H. Freeman, 1980); Пер. на русск. яз.: Пригожин И *От существующего к возникающему: Время и сложность в физических науках* (М.: Наука, 1985)
78. Bohr N *Atomic Physics and Human Knowledge* (New York: John Wiley, 1958); Пер. на русск. яз.: Бор Н, в кн. *Моделирование в биологии* (Под ред. Н А Бернштейна) (М.: ИЛ, 1963)
79. Ruelle D, Takens F *Comm. Math. Phys.* **20** 167 (1971)
80. Ландau Л Д *ДАН СССР* **44** (8) 339 (1944)
81. Hopf E *Comm. Pure Appl.* **1** 303 (1948)
82. Krinsky V I (Ed.) *Self-Organization. Autowaves and Structures Far from Equilibrium* (Berlin: Springer-Verlag, 1984)
83. Prigogine I, Nicolis G *Self-Organization in Non-Equilibrium Systems* (New York: Wiley, 1977); Пер. на русск. яз.: Николис Г, Пригожин И *Самоорганизация в неравновесных системах* (М.: Мир, 1979) с. 152
84. Иваницкий Г Р, Кринский В И, Сельков Е Е *Математическая биофизика клетки* (М.: Наука, 1978)
85. Зыков В С *Моделирование волновых процессов в возбудимых средах* (М.: Наука, 1984)
86. Арнольд В И *УФН* **141** 569 (1983); Arnold V I *Sov. Phys. Usp.* **26** 1025 (1983)
87. Devaney R L *Introduction to Chaotic Dynamical Systems* (Boulder, CO: Westview Press, 2003)
88. Малинецкий Г Г, Потапов А Б *Современные проблемы нелинейной динамики* (М.: УРСС, 2002)
89. Данилов Ю А *Прекрасный мир науки* (М.: Прогресс–Традиции, 2008)
90. Prigogine I, Stengers I *Order out of Chaos: Man's new dialogue with nature* (New York: Bantam, 1984); Пер. на русск. яз.: Пригожин И, Стенгерс И *Порядок из хаоса: Новый диалог человека с природой* (М.: Прогресс, 1986)
91. Maass W, Legenstein R, Markram H A, in *Biologically Motivated Computer Vision* (Eds C Bulthoff, S-W Wallraven, T Poggio) Vol. 2525 (Berlin: Springer, 2002) p. 1
92. Markram H et al. *Nature Rev. Neurosci.* **5** 793 (2004)
93. Kalisman N, Silberberg G, Markram H *Proc. Natl. Acad. Sci. USA* **102** 880 (2005)
94. Le Be J-V, Markram H *Proc. Natl. Acad. Sci.* **103** 13214 (2006)
95. Wang Y et al. *Nature Neurosci.* **9** 534 (2006)
96. Le Be J-V et al. *Cereb Cortex* **17** (9) 2204 (2007)
97. Markram H *Nature* **445** 160 (2007)
98. Markram H *HFSP J.* **2** (3) 132 (2008)
99. Druckmann S et al. *Biol. Cybern.* **99** (4-5) 371 (2008)
100. King J G et al. *Front Neuroinformatics* **3** 10 (2009)
101. Berger T K et al. *J. Physiol.* **587** (Pt. 22) 5411 (2009)
102. Hay E et al. *PLoS Comput. Biol.* **7** (7) e1002107 (2011)
103. Lasserre S et al. *IEEE Trans. Visualizat. Comput. Graphics* **18** 214 (2011)
104. Druckmann S et al. *PLoS Comput. Biol.* **7** (8) e1002133 (2011)
105. Romand S et al. *Front. Neuroanat.* **5** (2011) <https://doi.org/10.3389/fnana.2011.00005>
106. Anastassiou A C et al. *Nature Neurosci.* **14** 217 (2011)
107. Perin R, Berger T K, Markram H *Proc. Natl. Acad. Sci. USA* **108** 5419 (2011)
108. Markram H *Nature Rev. Neuroscience* **7** 153 (2006); Маркрам Г В мире науки (8) 46 (2012)
109. Тейл С В мире науки (12) 40 (2015)
110. "The Human Connectome Project", NIH Blueprint for Neuroscience Research, National Institutes of Health, retrieved 2013-02-16
111. Братко И *Алгоритмы искусственного интеллекта на языке PROLOG = Prolog Programming For Artificial Intelligence* (М.: Вильямс, 2004)
112. Анохин К В и др. *Матем. биология и биоинформ.* **7** 372 (2012)

113. "Special Issue: The Allen Human Brain Reference Atlas" *J. Comparative Neurology* **524** (16) (2016) Version of Record online: 15 September
114. Блинков С М, Иваницкий Г Р *Биофизика* **10** 817 (1965)
115. Блинков С М, Иваницкий Г Р *Вестн. Акад. мед. наук СССР* (5) 84 (1968)
116. Baile D L et al. (Eds) *Positron Emission Tomography* (New York: Springer, 2005)
117. Rowan A J, Tolunsky E *Primer of EEG: With A Mini-Atlas* (Philadelphia, PA: Butterworth-Heinemann, 2003)
118. Kim S G, Richter W, Ugurbil K *Magn. Reson. Med.* **37** 631 (1997)
119. Доронина-Амитонова Л В и др. *УФН* **185** 371 (2015); Doronina-Amitonova L V et al. *Phys. Usp.* **58** 345 (2015)
120. Шерозия Г А, Шерозия М Г *Человеческий разум, рожденный в сияниях искусственных логических элементов — введение в проект создания нового человека* (Рязань: ПРИЗ, 2013)
121. Редько В Г *Лекция по нейроинформатике* (М.: МИФИ, 2002)
122. Weisstein E W "de Morgan's Laws", <http://mathworld.wolfram.com/deMorgansLaws.html>
123. Владимиров Д А *Булевы алгебры* (М.: Наука, 1969)
124. Peano G *Formulaire de mathematiques* Vol. 2 (Torino: Bocca, 1897)
125. Frege F L G, in *Beiträge zur Philosophie des Deutschen Idealismus III* (1923) p. 36
126. Пуанкаре А *О науке* (М.: Наука, 1983)
127. Гильберт Д *Избранные труды* в 2 т. (Под ред. А Н Паршина) (М.: Факториал, 1998)
128. Russell B *Introduction to Mathematical Philosophy* (London, 1924); Пер. на русск. яз.: Рассел Б *Введение в математическую философию* (М.: Сибирское университетское изд-во, 2009)
129. Arnold van der Nat *Simple Formal Logic: with Common-Sense Symbolic Techniques* (Routledge, 2010)
130. Grünbaum A *Modern Science and Zeno's Paradoxes* (Middletown, Conn.: Wesleyan Univ. Press, 1968)
131. Лоскутов А Ю *УФН* **180** 1305 (2010); Loskutov A *Phys. Usp.* **53** 1257 (2010)
132. Вольф М Н *Философский поиск: Гераклит и Парменид* (СПб.: Изд-во Русской христианской гуманитарной академии, 2012)
133. Конт-Спонвиль А "Бритва Оккама (Rasoir D'Okham)", Философский словарь Litres (2015)
134. Chaitin G J *Limits of Formal Reasoning* (Berlin: Springer-Verlag, 1998)
135. Деев А А, Единцов И М, Иваницкий Г Р, Куниский А С *Биофизика* **34** 564 (1989)
136. Кадомцев Б УФН **164** 449 (1994); Kadomtsev B B *Phys. Usp.* **37** 425 (1994)
137. Кадомцев Б Б УФН **165** 967 (1995); Kadomtsev B B *Phys. Usp.* **38** 923 (1995)
138. Кадомцев Б Б *Динамика и информация* (М.: Изд-во журнала "Успехи физических наук", 1999)
139. Чайгин Г В *мире науки* (6) 38 (2006)
140. Gödel K *Monatshefte für Mathematik und Physik* **38** (1) 173 (1931)
141. Лейбниц Г В *Труды по философии науки* (М.: Либроком, 2010)
142. Гмурман В Е *Теория вероятностей и математическая статистика* (М.: Высшее образование, 2005)
143. Kahneman D et al. *Judgment under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge Univ. Press, 2005)
144. Frit K *Making up the Mind: How the Brain Creates our Mental World* (Malden, MA: Blackwell Publ., 2007); Пер. на русск. яз.: Фрит К *Мозг и душа* (М.: CORPUS, 2010)
145. Канеман Д, Словник П, Тверски А *Принятие решений в неопределенности: Правила и предубеждения* (Харьков: Гуманитарный центр, 2005)
146. Лuria А Р *Об историческом развитии познавательных процессов* (М.: Наука, 1974)
147. Кругляков Э П "Ученые" с большой дороги-2 (М.: Наука, 2005)
148. Wiener N *God and Golem* (Cambridge, Mass.: MIT Press, 1964); Пер. на русск.: Винер Н *Творец и робот* (М.: Прогресс, 1966)
149. Бобылев Д *Энциклопедический словарь Т. VIa* (СПб.: Брокгауз-Ефрон, 1892) с. 616
150. David H et al. *Math. Comput.* **66** 903 (1997)
151. Рыбников К А *История математики* Т. 1, Т. 2 (М.: Изд-во Моск. университета, 1960, 1963)
152. Белл Б, Белл Д *Пособие по английскому языку* (Дубна, 1971)
153. Морозов А Ю УФН **164** 3 (1994); Morozov A Yu *Phys. Usp.* **37** 1 (1994)
154. Фаддеев Л Д УФН **183** 487 (2013); Faddeev L D *Phys. Usp.* **56** 465 (2013)
155. Колмогоров А Н *Проблемы передачи информации* **1** (3) (1965)
156. Колмогоров А Н *Проблемы передачи информации* **5** (3) (1965)
157. Clark R W *Einstein: The Life and Times* (World Publ. Co., 1971)
158. Хютт В П *Концепция дополнительности и проблема объективности физического знания* (Ин-т истории АН ЭССР, Таллин: Валгус, 1977)
159. Jones S, Martin R, Pilbeam D (Eds) *The Cambridge Encyclopedia of Human Evolution* (Cambridge: Cambridge Univ. Press, 1994)
160. Хрисанкова Е Н, Перевозчиков И В *Антропология* (М.: Изд-во МГУ, 1991)
161. Foley R *Another Unique Species* (New York: John Wiley and Sons, 1987); Пер. на русск. яз.: Фоули Р *Еще один неповторимый вид* (М.: Мир, 1990)
162. Wood B *Nature* **355** 783 (1992)
163. Капица С П УФН **166** 63 (1996); Kapitsa S P *Phys. Usp.* **39** 57 (1996)
164. Falk D *Prog. Brain Res.* **195** 255 (2012)
165. Прингл Ч В *мире науки* (5) 54 (2013)
166. Опарин А И *Возникновение жизни на Земле* (М.—Л.: Госиздат биологической и медицинской литературы, 1936)
167. Miller S L *Science* **117** 528 (1953)
168. Miller S L, Urey H C *Science* **130** 245 (1959)
169. Розанов А Ю *Вестн. РАН* **70** 214 (2000); Rozanov A Yu *Herold Rus. Acad. Sci.* **70** 154 (2000)
170. Лопухин А С *Вестн. РАН* **85** 916 (2015); Lopukhin A S *Herold Rus. Acad. Sci.* **85** 453 (2015)
171. Von Neumann J *Theory of Self-Reproducing Automata* (Urbana: Univ. of Illinois Press, 1966); Пер. на русск. яз.: фон Нейман Дж *Теория самовоспроизводящихся автоматов* (М.: Мир, 1971); Von Neumann J, Morgenstern O *Theory of Games and Economic Behavior* (Princeton: Princeton Univ. Press, 1944); Пер. на русск. яз.: фон Нейман Дж, Моргенштерн О *Теория игр и экономическое поведение* (М.: Наука, 1970)

The robot and the human. Where's their similarity limit?

G.R. Ivanitskii

Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences,
ul. Institutskaya 3, 142290 Pushchino, Moscow region, Russian Federation
E-mail: ivanitsky@iteb.ru

It is shown that the primary goal of the robot-human collaboration is to ensure dynamic stability under varying environment conditions. Compared with the robotic computer "brain", its human counterpart has a multi-level hierarchical organization, with information processing occurring at all levels—from the quantum up to the social. A human himself sets a goal and improves the virtual model synthesized by his brain. The human brain can work by simultaneously using both classical deterministic logic and dialectical probabilistic logic.

Keywords: android robot, adaptation, creativity, memory, hierarchy of organization, classical and probabilistic logic

PACS numbers: 87.19.L-, 87.85.St, 89.20.Ff

Bibliography — 171 references
Uspekhi Fizicheskikh Nauk **188** (9) 965–991 (2018)
DOI: <https://doi.org/10.3367/UFNr.2018.03.038302>

Received 17 October 2017, revised 14 February 2018
Physics – Uspekhi **61** (9) (2018)
DOI: <https://doi.org/10.3367/UFNe.2018.03.038302>