

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ИЛЛЮЗИИ, ПРОБЛЕМЫ, ПЕРСПЕКТИВЫ

Р.А. Дурнев, зам. нач. упр. ФГБУ «Российская академия ракетных и артиллерийских наук», докт. техн. наук, доцент, rdurnev@rambler.ru

К.Ю. Крюков, рук. секр-та ФГБУ «Российская академия ракетных и артиллерийских наук», канд. психол. наук, raran@bk.ru

И.В. Жданенко, ст. научн. сотр. ФГБУ «Всероссийский научно-исследовательский институт по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России» (Федеральный центр науки и высоких технологий), izhdanenko@yandex.ru

Проведен анализ проблем развития искусственного интеллекта. Показан возможный путь создания «сильного» искусственного интеллекта. Рассмотрены преимущества «умных» машин, созданных на его базе.

Ключевые слова: слабый и сильный искусственный интеллект, алгоритм, автоматизация, мышление, неокортекс, нейроны.

ARTIFICIAL INTELLIGENCE: ILLUSIONS, PROBLEMS, PROSPECTS

R.A. Durnev, Deputy Head of Department, Federal State Budgetary Institution «Russian Academy of Rocket and Artillery Sciences», Ph. D., rdurnev@rambler.ru

K.Yu. Kryukov, Head of Secretariat, Federal State Budgetary Institution «Russian Academy of Rocket and Artillery Sciences», Doctor of Psychology, raran@bk.ru

I.V. Zhdanenko, Senior Researcher, All-Russian research Institute on Problems of Civil Defense and Emergency Situations, izhdanenko@yandex.ru

The analysis of the problems of the development of artificial intelligence is carried out. A possible way of creating a «strong» artificial intelligence is shown. The advantages of «smart» machines created on its basis are considered.

Keywords: slack and strong artificial intelligence, algorithm, automation, thinking, neocortex, neurons.

Введение

В статье [1] проведен анализ возможных направлений применения искусственного интеллекта (ИИ) в военном деле. Отмечено, что вопреки ожиданиям 50–60-летней давности современные технологии ИИ не позволяют говорить о качественном моделировании тех видов человеческой деятельности, которые традиционно считаются интеллектуальными, о возможности интеллектуальных систем выполнять творческие функции, считающиеся прерогативой человека. Практически не изменился состав методов ИИ, на протяжении полувека мы говорим о тех же нейросетях, эвристическом программировании, теории нечетких множеств (ТНМ), многозначной логике, генетических алгоритмах, экспертных системах. И если появляется информация о каких-то новых способах моделирования мыслительных процессов, то в большинстве случаев речь идет о комбинации упомянутых. Так, гибридные модели [2] – это комбинация нейросетей и ТНМ, большой пласт методов обучения (с учителем и без, с подкреплением, глубинный и др.), основанный на нейросетях, возможно-вероятностные методы базируются на теории вероятности и ТНМ, экспертные системы включают в себя импликации ТНМ.

Повышение возможностей современных технологий ИИ связано в большей степени с прогрессом микроэлектроники (например, развитием видеокарт для применения в сверточ-

ных и других глубоких сетях), наличием доступных больших баз данных (статистической информации, видеоряда и т. п.) для обучения нейросетей, в меньшей – с успехами в развитии математического аппарата. Так, практически невозможно найти информацию по новым видам нейронных сетей после 80 – середины 90 гг. XX века. В основе почти всех моделей – все те же полносвязные, сверточные, рекуррентные сети или их комбинации и модификации – порождающие, рекурсивные, состязательные, «с вниманием», «с подкреплением» и др. Основным методом их обучения по-прежнему является открытый в середине 70 гг. XX века метод обратного распространения ошибки, а оптимизации – градиентный спуск [3, с. 103].

Определенная иллюзия прогресса ИИ связана с увеличением компактных устройств, наделенных элементами «разумного поведения» (гаджеты, «умная» бытовая техника, системы навигации и управления современными автомобилями). Причиной этого является увеличение их мощности до величин, доступных ранее только большим ЭВМ, на которых обрабатывались алгоритмы ИИ. Да и сам этот прогресс имеет не научно-технический, а инновационный характер, связанный с коммерциализацией тех знаний по ИИ, которые были получены несколько десятилетий назад.

Поэтому, хотя практические приложения ИИ становятся более удобными, доступными, массовыми в применении, по всей видимости, их очередной всплеск развития приближается к плато очередной логистической кривой, в связи с чем вложение значительных средств уже не приведет в ближайшее время к значимому повышению качества систем машинного перевода текстов, технологий распознавания речи в условиях шумов и искажений голоса, технического зрения, позволяющего ориентироваться в незнакомой сложной обстановке, с одновременным распознаванием множества размытых, нечетких, смещенных, повернутых, деформированных образов различной природы.

Одним из характерных примеров несовершенства систем, имитирующих интеллектуальную деятельность человека, являются спамботы – программы, самостоятельно заполняющие электронные поля на различных сервисах в Интернете, как правило, с противозаконными целями. Для противодействия им был разработан CAPTCHA (Completely Automatic Public Turin Test to tell Computers and Humans apart [4]) – полностью автоматизированный публичный тест Тьюринга для различения компьютеров и людей. В этом тесте пользователь должен ввести несколько знаков (цифр, букв), изображения которых искажены (перечеркнуты, деформированы и т. п.), или выбрать элементы матрицы, являющиеся нечеткими фотографиями витрин, дорожных знаков, автомобилей и т. п. Судя по длительности существования таких тестов, спамботы, не способные самостоятельно их пройти, обладают низкой степенью интеллектуализации и не могут справиться с задачей, не представляющей труда для человека.

Основная часть

Следует отметить, что непрерывное увеличение памяти, быстродействия компьютеров не вносит решающий вклад в развитие ИИ. Шахматный суперкомпьютер Deep Blue компании IBM, выигравший у Г. Каспарова в 1997 г., осуществлял вычисления в пространстве 10^{120} позиций, и для реализации метода «поиска с альфа-бета-отсечением» осуществлял перебор в среднем 126 млн позиций в секунду, анализируя 8 тыс. признаков каждой позиции. В 2011 году ЭВМ Watson компании IBM, победившая двух соперников в телеигре Jeopardy!, обрабатывала данные со скоростью 500 Гб/с (что эквивалентно миллиону книг в секунду) и располагала 16 Тб оперативной памяти. Для распознавания кошачьих мордочек Cat Detector от Google использовал 1000 серверов с 16 000 ядер. А в 2016 году компьютер победил в игре го одного из лучших игроков в мире Ли Седоля. Размерность этой игры уже 10^{400} , что на много порядков превышает число элементарных частиц в видимой части Вселенной. Программа игры в го AlphaGo работала со 1920 г. CPU (центральный процессор) и 280 GPU (графический процессор) и использовала метод «поиска по дереву Монте-Карло», в соответствии с которым для оценки текущей позиции запускались случайные симуляции, начи-

нающиеся с нее, затем оценивались ветки дерева ходов, в которых черные или белые выиграли больше законченных партий и в конце рекурсивно повторялся этот поиск в самых перспективных узлах получающегося дерева¹. Но даже такое стремительное увеличение вычислительных мощностей компьютеров, объединение их в различные сети (например, SETY@home для обработки астрономических данных в целях поиска сигналов внеземных цивилизаций) не приводит, к сожалению, к значимому прогрессу в области искусственного осуществления творческих функций человека, а позволяет пока только выигрывать в так называемые дискретные игры с полной информацией.

Недостаточное развитие ИИ зачастую объясняются тем, что компьютеры – это синтаксические объекты, оперирующие правилами построения и функционального взаимодействия различных символов (например, языковых единиц в предложениях, словосочетаниях и др.) без учета их смысла.

С точки зрения системного подхода система не может произвести другую систему одного с ней уровня сложности. Это связано с тем, что система – совокупность взаимодействующих элементов, обладающих определенными свойствами. При этом свойства самой системы не определяются алгоритмическими операциями над свойствами ее элементов. Возникает эмерджентность, то есть то, что отличает систему от простой кучи элементов. Не существует приемлемого объяснения механизма получения новых неизвестных свойств системы из известных свойств ее элементов. Поэтому система не может породить эквивалентную систему, а интеллект человека навряд ли способен создать эквивалентный ему интеллект. Это под силу только сверхсистеме. Но человек может произвести систему более низкого уровня, например интеллект животного.

Некоторые исследователи [5] говорят о проблемах функционального рассогласования искусственных интеллектуальных систем и естественного мышления человека. Отмечается, что для разработчиков систем искусственного интеллекта интерес представляет не весь мыслительный процесс, а только часть, связанная с сознанием, разумом. Функционально за это отвечает кора больших полушарий мозга (неокортекс), которая возникла на самых поздних этапах эволюции и продолжает эволюционировать. При этом совершенно не рассматривается связь процесса мышления с работой других отделов головного мозга – лимбической системы, отвечающей за чувства и эмоции, «рептильного мозга» (самого раннего в эволюционном смысле), на который завязана сфера инстинктов. В этой связи разработчики ИИ копируют, пытаются усовершенствовать и развивают недоработанные эволюцией высшие психические функции (в том числе и мыслительные процессы) и еще в большей степени усугубляют функциональный разрыв между «древними» отделами мозга и эволюционно наиболее молодыми структурами.

По мнению других [6, с. 3], необходим радикальный поворот к принципам «сильного ИИ» (который должен мыслить как человек в отличие от «слабого ИИ» для решения частных задач) путем перенесения личности, сознания человека на небиологические носители. С этой целью в будущем могут использоваться нанороботы, которые, внедряясь в человеческий мозг, оцифруют каждый синапс, нейронную связь, а затем передадут полученную информацию на компьютер.

Другой способ переноса сознания на «винт» может быть связан с нейрохирургическим вмешательством [7, с. 1261], в процессе чего будут последовательно извлекаться из мозга

¹ Следует отметить, что существенна не только размерность задачи для шахмат и го, но и марковский (без последствия, без памяти) характер игр – важна позиция сама по себе, а не предыстория того, какими предыдущими ходами она была достигнута. Этим свойством не отличается карточная игра покер, в которой важно не только значение текущего размера ставок, но и история ставок в текущей задаче, ретроспектива взаимодействия игроков и т.п. Наверное, поэтому (да и еще из-за возможности «блефа») ИИ пока еще не победил человека в покер.

небольшие количества нейронов и дублироваться транзисторами на компьютере, который будет связан с оставшимся живым мозгом. В итоге все нейроны будут заменены транзисторами, и в «голове» робота будет размещаться сознание.

Созданный с помощью нанороботов или нейрохирургов так называемый «коннектом» будет сопряжен с роботизированным телом (аватаром), которому доступны визуальные, слуховые, тактильные и другие ощущения, и перенесенное на суперкомпьютер сознание будет проявлять себя через совершенное, мобильное, суррогатное тело.

Существует также мнение о том, что настоящий ИИ разовьется на основе «брайн-нета» («мозговой сети»), в котором будут использоваться возможности телепатии (на основе нанозондов в мозгу и сканирующих устройств), телекинеза, загрузки воспоминаний и стимулирования мозга [8]. Это позволит человеку мысленно связываться с другими людьми или компьютерными устройствами. Такая суперсеть будет средой для зарождения и развития сильного ИИ.

Еще одна из возможных причин незначительного прогресса в области создания ИИ, подобного человеческому, – это невычислимые процессы сознания, понимания, разума [9]. Здесь можно говорить о хорошо структурированных математических или физических задачах, не имеющих общего аналитического решения. Такой задачей является нахождение решений произвольной системы диофантовых уравнений (полиномиальные уравнения с коэффициентами и решениями в виде целых чисел), алгоритма замощения плоскости многоугольниками без зазоров и перекрытий и др. Применительно к первой задаче Ю. Матиясевичем доказана невозможность создания алгоритма, способного систематически определять, имеет ли система диофантовых уравнений решение. Для второй задачи Р. Бергером было доказано, что задача замощения вычислительными (компьютерными) средствами неразрешима, при этом другими исследователями найдены апериодические процедуры замощения, при которых никакой участок законченного узора замощения нигде не повторяется. Для физики нет общего аналитического способа решения задачи гравитационного взаимодействия трех тел, но найден ряд частных решений при расположении тел на одной прямой, в вершинах равнобедренного треугольника и т.д. Есть также проблемы, которые человек способен поставить в формальном виде, но для которых пока не только не найдено формального решения, но даже и не понятно, конечен ли соответствующий вычислительный алгоритм (например, найти четное число больше двух, не являющихся суммой двух простых чисел). Обо всех этих неразрешимых задачах человек в отличие от компьютера может сделать вполне определенные выводы.

Причины такой невычислимости сознания заключаются в неопределенности квантово-механических эффектов, происходящих в микротрубочках цитоскелета нервных клеток головного мозга [9, с. 547]. Установлено, что, помимо самих нейронов, за мыслительную деятельность отвечает их цитоскелет (который есть во всех эукариотических клетках). Так, например, у инфузории (парамеции) отсутствуют нервные клетки и присутствуют микротрубочки цитоскелета. Благодаря им она «разумно» перемещается в воде, преследует бактерии, огибает препятствия, скрывается от опасности и даже способна к простейшему обучению. Кроме того, сознание человека отключается при воздействии на него препаратов общего наркоза (закиси азота, эфира, хлороформа, галотана, изофлурана, ксеона и др.). Это связано с тем, что анестезирующее вещество, просачиваясь в нервную клетку, благодаря электрическим дипольным свойствам останавливает работу микротрубочек, сознание отключается, и человек перестает быть мыслящим существом. При этом воздействие электрических дипольных свойств на микротрубочки обладает выраженным квантово-механическим эффектом, не являющимся вычислимым. В этом процессе нейроны, по утверждению Р. Пенроуза [9, с. 574], выполняют функции, скорее всего, увеличительных стекол, посредством которых микроскопические цитоскелетные процессы поднимаются на уровень, на котором возможно воздействие на другие структуры, например на мышцы. Поэтому нейронный уровень описания является не более чем тенью цитоскелетных про-

цессов более глубокого уровня, именно там и находится, возможно, физический фундамент разума.

Определенные надежды в создании ИИ возлагают на замену кремниевых транзисторов более быстродействующими и экономными с точки зрения потребляемой энергии нанотрубочками или квантовым компьютером, который в отличие от обычного, находящегося одновременно только в одном из двух состояний — 0 и 1, может быть как в этих состояниях, так и в их суперпозиции, то есть сразу в нескольких. Это позволяет реализовать квантовый параллелизм, когда изменение состояния отдельного квантового бита (кубита) вызывает изменение состояния всей системы кубитов, находящихся в запутанных квантовых состояниях. Эволюция каждого такого состояния соответствует вычислениям над собственным набором данных, которые выполняются параллельно. Предполагается, что такие машины смогут решать невычислимые (неалгоритмизированные) задачи в рамках квантовых процессов в искусственных нейронах и создадут основу для появления квантового искусственного интеллекта. Однако создание таких компьютеров, несмотря на то что они были предложены Р. Фейманом более 35 лет назад, в ближайшее время не предвидится.

Возможно, причина застоя в разработке «сильного ИИ» и в том, что он не может быть наделен основным признаком, отличающим человека от животных. Это рефлексия, то есть приобретенная сознанием способность сосредотачиваться на самом себе и овладевать самим собой как предметом, обладающим своей специфической устойчивостью и своим специфическим значением, — способностью уже не просто познавать (что свойственно и животным), а познавать самого себя; не просто знать (что им также доступно), а знать, что знаешь [10]. Еще более сильным отличием является способность ощущать, что мы знаем, в сравнении с тем, что знают другие люди, способность моделировать их мысли, в том числе и о нас, по-новому ощущать, кто мы есть. И когда, по мнению некоторых специалистов, машина станет учитывать, что знает она и что знает пользователь, она станет мыслящей.

Множество вариантов (причин, поводов и т. п.) возникновения сильного ИИ рассматривается и в художественной литературе. Так, например, в [11] беспилотные разведывательно-ударные летательные аппараты «Freedom Liberator», управляемые нейронными сетями «Free D.O.M.», скачком переходили от слабого искусственного интеллекта к сильному в момент гибели, при отрыве от выполнения боевой задачи и замыкании в «эмоциональной вольтовой дуге».

Все эти и другие причины вызывают определенный интерес с точки зрения познания соотношений естественного и искусственного разума, определения загоризонтных тенденций развития «умных» машин. Однако более значимой и реальной в разрешении представляется проблема, связанная со стремлением разработчиков воссоздать человеческое мышление, поведение на основе собственных мнений, взглядов на предметные области, без учета природы и механизмов функционирования биологического разума. Так, например, при разработке системы технического зрения используется терминология пространственных форм, трехмерных изображений, характеризующая не само зрительное восприятие человека, а упрощенное объяснение его механизмов. Аналогично при создании «говорящего» компьютера применяются алгоритмы, в основе которых положены правила синтаксиса, грамматики и семантики, то есть опять теоретические основы, объясняющие структуру и функции речи, языка. Построение экспертных систем с нечеткой логикой осуществляется с помощью формальных функций принадлежности и базы правил, отражающей мнения (объяснения) экспертов в определенных областях. При разработке роботов с эмоциями и темпераментом [12, с. 169] в основу закладываются информационные теории эмоций, двухпараметрические модели типа «возбуждение—торможение» и т. п. Даже самые современные (в основном с точки зрения «рекламы инноваций») нейросети с глубоким обучением для распознавания речи вначале используют преобразование сигнала в признаки специального вида, далее эти признаки превращаются в гипотезы для выбора вариантов фонем, а сами фонемы объединяются в предположения относительно произнесенных слов, для которых существуют специальные языковые модели [13, с. 285]. Для обработки естественных языков

используются нейросетевые модели представления слов и посимвольные модели, в основу которых положено отображение слова (морфемы, буквы) из словаря в некоторый вектор в евклидовом пространстве и семантические отношения между словами выражаются в виде геометрических отношений в самом евклидовом пространстве [13, с. 312].

То есть во всех этих случаях основу методов ИИ составляют модели предметных областей, а не сами механизмы зрительного, слухового, тактильного и иного восприятия, мышления, воображения и т. п.

Да и в целом в основном только в начале создания искусственных нейросетей была использована аналогия с биологическими нейронами, небольшое более позднее исключение составляют сверточные сети, в элементы которых положены определенные механизмы зрительного восприятия. Поэтому в последнее время превалирующим стал вывод о том, что искусственные нейросети – это не аналог реальной структуры мозга, а абстрактная модель, созданная для решения оптимизационных задач, иначе – «формализм машинного обучения, а не попытка что-то подсмотреть у природы» [13, с. 23]. Так, живые нейроны передают друг другу не вещественные числа (например, значения градиентов поверхности ошибки), а бинарные сигналы. Кроме того, у них нет соединений, работающих сразу в обе стороны, – вход идет через дендриты, выход через аксоны, а никак не наоборот. И, наконец, живые нейроны распознают только нервные импульсы (сигналы одного типа), а искусственные «умеют» работать как с функцией сигналов, так и с вышеуказанными градиентами.

Составляющие искусственных нейронных сетей в том или ином виде давно известны в областях, развивавшихся независимо от ИИ. Так, в исследовании операций и теории решений давно известна процедура взвешенных сумм (сумма произведений значений показателей на их веса), подобие метода обратного распространения ошибки всегда существовал в теории автоматического управления, движение по градиенту поверхности в пространстве функции ошибки и весов сигналов аналогично градиентному подъему или спуску по факторной гиперповерхности в теории планирования эксперимента, в основе обучения с подкреплением заложено основное уравнение динамического программирования Беллмана. При этом сами сети – это один из направленных графов с вершинами в виде математических операций (так называемый граф вычислений) и точками входа и выхода для данных. Поэтому в «Википедии» нейросети относят к частному случаю методов распознавания образов, дискриминантного анализа, кластеризации, нелинейной оптимизации, адаптивного управления и т. п., то есть по-разному в зависимости от классификационных признаков в различных отраслях науки.

Применительно к тем же нейросетям в работе [14, с. 35] приводится мнение о том, что формальное поверхностное повторение структуры расположения и взаимосвязи групп нейронов совсем не свидетельствует о совпадении механизмов обработки, хранения, передачи информации в нейросетях и мозге человека в целом. Можно провести аналогию гипотетического исследования структуры и функций компьютера человеком, слабо знакомым с областями науки и техники, необходимыми для его создания. Человек может выяснить, что компьютер состоит из миллионов транзисторов, связанных между собой в точную и сложную схему. При этом принципы построения и работы такой схемы остаются для него непонятными. Даже случайное соединение нескольких транзисторов в определенном порядке и получение усилителя сигналов не объясняет работу компьютера в целом, кроме того что он и усилитель состоят из транзисторов. Также и изучение нейронной сети не позволяет объяснить принципы работы мозга человека, состоящего из нейронов. Поэтому и не теряют популярности статьи с названиями «Может ли биолог починить радиоприемник» [15] или «Может ли нейробиолог понять микропроцессор» [16].

Простое подражание каким-то отдельным свойствам биологических нейронов (типа порогового сумматора или функции активации, правила Хебба) не позволяет достичь уровня самообучения, сопоставимого с естественными нейронами. Поэтому искусственными нейрон-

ными сетями невозможно реализовать более высокую логику, чем логику высказываний (для нейронов МакКаллока–Питса), распознавать зрительные образы, изменяющиеся при существенном вращении, сдвиге и масштабировании (для перцептронов, когнитронов, неокогнитронов, нейроглиальных сетей) и т. п. Все они эффективно решают задачи там, где исходные данные носят одномерный характер, как, например, для сенсорных звуковых, визуальных и других сигналов. Творческие задачи им не под силу, и навряд ли когда-нибудь нейросети смогут доказать, все ли четные числа являются суммой двух простых, определить, какие из пар чисел (2384 и 2023, 3891 и 1389, 5261 и 1093) принадлежат к множеству пар (3751 и 1357, 4382 и 2348, 7289 и 2798), просто понять, что «в этих солвах бли допущены опечатки» или о какой косе идет речь в выражении «по деревне шла девушка с косой» [13, с. 318]. В целом вполне обоснованным представляется вывод о том, что искусственные нейросети могут спрогнозировать только то, что было в обучающем примере, то есть сделать сугубо локальное обобщение. Человек же способен мыслить гипотетически, расширяя свою модель мира за пределы того, что непосредственно испытал напрямую, приспособляясь к абсолютно новым ситуациям или планируя будущее, а именно: совершать глобальное обобщение.

Следует также отметить, что на определенных этапах мышления, например математического доказательства, человек зачастую прибегает к интуиции, которая не может быть формализована. Это подтверждается теоремой Геделя о неполноте, упрощенную формулировку которой можно представить следующим образом – в рамках заданной системы аксиом (то есть любой формальной системы, а сейчас методы ИИ являются именно такой системой) найдутся положения, которые невозможно ни доказать, ни опровергнуть. Мышление же человека зачастую неформально, поэтому в сложных ситуациях он принимает решения, которые ни в момент их принятия, ни после этого логически трудно или вовсе нельзя объяснить.

В этой связи представляется, что основной проблемой создания систем ИИ являются попытки их построения на основе моделей (упрощенных объяснений) различных явлений и процессов, представление их в виде алгоритмов, что приводит все к той же автоматизации. Под алгоритмами, положенными в основу автоматизации деятельности, понимается совокупность шагов или правил, приводящих к определенному результату, решению задачи, набор инструкций, описывающих порядок действий исполнителя для достижения некоторого результата (ru.m.wikipedia.org). Количество шагов в алгоритме, приводящих к решению задачи, должно быть конечным, и сами шаги необходимо однозначно определить. Если решение задачи может быть выражено в виде алгоритма, она является разрешимой (вычислимой). Именно для таких задач и была разработана концепция универсальной машины Тьюринга, которая наряду с принципами архитектуры Дж. Фон Неймана привела к бурному развитию электронно-вычислительных машин. При этом современные нейронные сети, реализуемые на компьютере, не эквивалентны универсальной машине Тьюринга, так как для этого они должны быть бесконечны или хотя бы не ограничены [17, с. 286].

Считается, что компьютеры – это физическое воплощение концепции алгоритма, которое было введено для формализации мышления в процессе решения математических задач [17, с. 28]. Электронные машины автоматически решают вычислительные задачи, выраженные алгоритмами, и, к сожалению, не позволяют работать с невычислимыми задачами, с которыми должен справляться ИИ, основанный на работе естественного мозга.

Подтверждением этому является парадокс, возникающий в мысленном эксперименте Дж. Серла [17, с. 210]. В нем представляется некоторая комната, в которой расположены таблички с китайскими иероглифами, а также книга правил (алгоритм) для людей, незнакомых с этим языком. В окошко в комнате подается вопрос на китайском языке, и человек, используя алгоритм, перекладывает иероглифы так, чтобы получить правильный ответ. Книга правил достаточно подробная, время не ограничено, поэтому ответы человека, незнакомого с языком, его носителям кажутся разумными и даже с примесью «дзен-буддизма». То есть,

абсолютно не понимая вопросов и ответов на них, человек может «разумно общаться» с представителями другого этноса. При временных ограничениях возможно представить деятельность миллиардов людей, переключивших таблички с иероглифами и передающих их друг другу. Это сродни нейронам, каждый из которых не понимает букв, слов, выражений, однако в рамках целого мозга как системы они действуют разумно. Такое же объяснение дается и для парадокса «китайской комнаты». В соответствии с ним пониманием текста обладает не человек, выполняющий манипуляции с символами, а комната вместе с человеком, книгой правил, табличками с иероглифами и корзинками. Сам Серл полагал, что понимание возникает из физических особенностей устройства мозга, и даже не столько из его эмерджентности, сколько из свойств самих нейронов. Поэтому «китайская комната» или компьютер не могут понимать без воспроизведения определенных физических и биологических механизмов мозга.

С учетом сказанного можно сделать вывод о том, что необходимо не программировать на компьютере решение задач, поддающихся алгоритмизации, а всемерно учитывать природу, механизмы функционирования биологического разума, сосредоточиться на поиске возможностей воссоздания «в кремнии» способности клеток, зон мозга к мышлению. Для этого, как представляется, необходимо сосредоточить усилия по двум основным направлениям. Первое из них связано с построением структурной модели мозга, так называемого коннектома, отражающего связи всех 10^{12} нейронов с их 10^3 связей в каждом. И сейчас такое «обратное проектирование мозга», Брайн-проекты, активно развиваются во многих лабораториях мира.

Второе направление предполагает построение функциональной модели мозга, прежде всего его важнейшей части, отвечающей за разумность, – коры больших полушарий, неокортекса. И здесь прежде всего необходимо принимать во внимание принципиальные отличия работы отдельных нейронов. Так, живые нейроны передают друг другу не вещественные числа (например, значения градиентов поверхности ошибки, как в нейросетях), а бинарные сигналы. Кроме того, у них нет соединений, работающих сразу в обе стороны: вход идет через дендриты, выход через аксоны, а никак не наоборот. И, наконец, живые нейроны распознают только нервные импульсы (сигналы одного типа), а искусственные «умеют» работать как с функцией сигналов, так и с вышеуказанными градиентами.

Кроме того, необходимо учитывать и особенности работы коры головного мозга в целом, связанные с алгоритмом восприятия сенсорных сигналов, способами запоминания окружающего мира и формирования прогнозов мышления и действий. Принципы использования в указанных целях механизмов деятельности неокортекса рассматриваются в оригинальной работе [14]. Так, в соответствии с ней у неокортекса существует универсальный алгоритм, выполняемый одинаково всеми зонами коры головного мозга, отвечающими за распознавание лиц, восприятие музыки, доказательство теорем и т.п. Несмотря на то что звуковой образ воздействует благодаря чувствительности к колебанию воздушных волн, зрительный – к свету, осязательный – к давлению на кожу, температуре, неокортекс использует одни и те же принципы для обработки слуховых, зрительных, двигательных и других сигналов. Поступая в головной мозг от различных органов чувств, данные сигналы превращаются в нервные импульсы и сразу становятся равноценными. Мозг находится в абсолютной изоляции (темноте, тишине и т.п.) и только благодаря поступающим сенсорным сигналам узнает о существовании этого мира, его структуре, протяженности, длительности. В этой связи понятно, почему мозг людей с врожденной глухотой перерабатывает зрительную информацию в зонах, ответственных в обычных случаях за слуховое восприятие, слепые от рождения используют большую часть зоны зрительного восприятия при работе с тактильным шрифтом Брайля. Известен опыт сенсорного замещения, когда потерявшему зрение человеку на лоб помещали видеокамеру, размещали во рту «языковой дисплей» (устройство с десятками электродов, создающих слабое покалывание на языке соответственно яркости пикселей на картинке видеокамеры), и он начинал видеть через язык.

В компьютерах же за прием и обработку сигналов отвечают принципиально разные системы, отличающиеся как аппаратной частью (видеокарты, звуковые карты и т.п.), так и соответствующим программным обеспечением, то есть алгоритмами. Это существенное отличие компьютерных и мозговых механизмов восприятия сенсорных сигналов необходимо учитывать при создании ИИ.

Следующее отличие заключается в том, что компьютер вычисляет, а мозг извлекает из памяти готовое, наиболее близкое, решение. При этом он запоминает последовательность элементов, а не отдельные элементы окружающего мира. Примерами таких последовательностей могут являться, например, воспоминания своего дома, расположения помещений, мебели, предметов обихода, которые вызываются из памяти не сразу в целом, а при мысленном его прохождении, то есть последовательно. Аналогично и знакомая мелодия вспоминается не сразу целиком, а тоже последовательно. Воспоминания хранятся в ограниченном числе синоптических связей между нейронами, и в каждое отдельное мгновение из памяти извлекается только крошечная часть информации.

Данные последовательности запоминаются ассоциативно, с возможностью их полного воспроизведения на основе недостаточной или искаженной входной пространственной и временной информации. Например, заметив за занавеской кроссовки спрятавшегося ребенка, мгновенно создается его полный образ.

Рассматриваемые последовательности формируются инвариантно, что позволяет компенсировать бесконечно изменяющиеся формы и быстротечность событий окружающего мира. Мозг не запоминает все с абсолютной точностью, не привязывается к отдельным элементам действительности, а учитывает только ее важнейшие взаимосвязи. Например, лицо друга узнается под разными углами зрения, ракурсами, при различном освещении, изменяющейся мимике, положении головы и т.п., на основе относительных измерений расстояний, цветов и пропорций, а не фиксированного последнего воспоминания.

Мир чрезвычайно изменчив, и единственный способ его познания – поиск инвариантной структуры (инвариантных репрезентаций, своего рода гештальтов) для переменного потока информации. В отличие от этого память компьютера сохраняет информацию в максимально неизменном виде, и даже незначительная ошибка в ней может вызвать сбой.

Внешний мир имеет иерархический характер (ноты – группы нот – мелодические фразы – песня – альбом, буквы – слоги – слова – предложения – фразы), и для его познания функциональная структура коры головного мозга также носит иерархический характер. Так, при рассмотрении лица сигналы от сетчатки, поступающие в первичную зону зрительного восприятия, сочетаются в линейные сегменты, линейные сегменты в последующих зонах объединяются в сложные формы, те – в образы (например, нос человека). Нос с глазами и ртом формируются в человеческое лицо. При этом компьютеру для сравнения огромного количества реальных точек, линий, углов и т.п. с хранящимися в базе данных приходится делать миллионы операций, мозг же выполняет эту процедуру за несколько шагов, зависящим в основном от количества задействованных зон.

В каждой из них образуются инвариантные репрезентации и передаются из низших высшим (при восприятии) и наоборот (при движении). В низших зонах они специфичные, быстроменяющиеся и отражают только характеристики объектов восприятия. Применительно к зрению – это точки, линии, углы, которые формируют активность нейронов в ответ на зрительные сигналы из крошечной части сетчатки. В высших зонах инвариантные репрезентации пространственно-неспецифические, медленно меняющиеся, включающие в себя целые объекты. Их нейроны возбуждаются и остаются активными, когда в поле зрения попадают лица, автомобили, деревья и т.п.

Вся информация от различных органов чувств и соответствующих зон неокортекса собирается в ассоциативные зоны. Благодаря им, получающим информацию от различных органов чувств, понимается связь между видом мухи, ползущей по руке, и щекочущим чувством на коже.

Репрезентации объектов из нижних иерархических зон (например, буквы, слога, слова) могут повторно использоваться для формирования разных последовательностей в высших зонах (предложения, абзацы, рассказы). В этой связи нет необходимости (в отличие от компьютера) запоминать эти буквы, слоги слова один раз для написания предложения, другой раз – для произнесения его вслух.

Помимо существования общего алгоритма обработки сенсорных сигналов, а также особенностей запоминания информации (последовательностями, ассоциативно, иерархически, с помощью инвариантных репрезентаций), неокортекс человека имеет еще одну отличительную черту: он непрерывно составляет прогнозы, которые являются сущностью понимания, так как понимать – это составлять об этом предположение.

Каждый миг мозг формирует низкоуровневые сенсорные прогнозы относительно того, что он ожидает ощутить (увидеть, услышать и т.п.). Эти прогностические процессы протекают параллельно – в зрительных (формы, цвета, расположения объектов), слуховых (источник, высота, громкость звуков), соматосенсорных (поверхность, контуры, температура, давление) и других зонах. Когда человек осматривается в комнате, его мозг формирует прогноз ожидаемого видения, восприятия. При этом такое формирование происходит до реальной оценки комнаты. Сам прогноз направлен не на отдельные объекты, находящиеся в комнате, а на структуру мира, то есть комнаты в целом. Механизм мгновенного нахождения чего-то нового в знакомой комнате основывается на том, что сохраненные воспоминания используются для осуществления прогноза относительно того, что человек воспринимает сенсорно.

В отличие от этого разработчики ИИ, например, для роботов составляют списки свойств всех предметов в комнате. В такую сформированную базу вводятся исходные данные роботизированной оценки комнаты, и элементы этой оценки сравниваются с каждым предметом обстановки в комнате. При этом количество предметов и их свойств, а также самих комнат (большинство из которых незнакомы) может быть огромным.

Движение, поведение также являются частью прогнозирования. Для прогноза того, что мы почувствуем дальше, необходимо знать, что делаем сейчас. Моторное и сенсорное восприятие тесно связаны, и большинство зон коры головного мозга принимают участие в движении. При этом моторное поведение также должно быть представлено в иерархии инвариантных репрезентаций. Человек генерирует движения, необходимые для осуществления определенного действия, не размышляя о нем. Например, для перемещения из гостиной на кухню мозгу необходимо переключиться с инвариантной репрезентации гостиной на инвариантную репрезентацию кухни. Данное переключение вызывает сложное разворачивание последовательностей моторных команд, заставляющих перейти из гостиной на кухню и перемещать взгляд в движении. Во время этого информационные потоки одновременно проходят по иерархии зон неокортекса в восходящем и нисходящем направлениях, причем последние имеют большую информационную насыщенность. Так, при чтении вслух высшие зоны коры головного мозга посылают к первичным зонам зрительного восприятия намного больше информации, чем получают ее от взгляда, бегущего по строкам книги. То, что является обратной связью для сенсорных областей, одновременно является выходной информацией для моторных, и наоборот. Кора головного мозга не проводит различия между ощущениями и поведением и воспринимает их как сигналы.

Само прогнозирование действует не только на сенсорном, но и на более высоких уровнях. При прохождении теста IQ человек решает прогностические задачи: оценивает следующее число, проекцию, слово и т.п. в последовательности. В науке идет формирование прогностических гипотез, в искусстве – заблаговременное представление объекта искусства в целом, в разработке дизайна – предвидение запросов потребителей.

Человек может совсем не совершать никаких движений, никак внешне не проявлять себя и в то же время мыслить, то есть составлять прогнозы и сравнивать их с образами из своей памяти. Поэтому именно прогнозирование, а не поведение по Тьюрингу составляет основу различения разума.

Кроме рассмотренных в настоящей статье самых общих принципов работы неокортекса, в работе [14] объяснены механизмы восприятия сенсорных сигналов, запоминания окружающего мира, формирования прогнозов мышления и действий, приведены примеры особенностей мышления человека с учетом этих механизмов, достаточно детально отражены принципы взаимодействия различных иерархий зон коры головного мозга, их сенсорной и моторной частей, слоев каждой зоны, состоящих из различных нейронов, обмена информацией между колонками нейронов и многие другие подробности, приемлемые для начала практической реализации этой запоминающе-прогностической модели в интересах создания «сильного ИИ».

С учетом сказанного возможно объяснить, чем принципиально отличаются искусственные и естественные системы распознавания речи, зрительного восприятия и другие «продвинутое» области ИИ. Так, программы распознавания речи соотносят звуковые сигналы с шаблонами слов, внесенными в память без учета их значения. Для некоторого улучшения качества распознавания принимается во внимание вероятность произнесения слов, их перехода, сочетаемости с другими языковыми конструктами, правила синтаксиса и грамматики. Все это оставляет такие системы (к примеру, «GoogleNow», «AppleSiri» «Алиса Яндекс» и т.п.) успешно работающими лишь в небольшом числе ситуаций, когда количество слов, которое произносится в каждое отдельное мгновение, ограничено, произношение осуществляется без аграмматизмов, запинок, перебивания говорящего и многоголосия.

А людям распознавание даже искаженной речи дается без труда. Это связано с тем, что неокортекс не только воспринимает отдельные звуки, но и прогнозирует целые слова, фразы, предложения, даже идеи, а также рамки общего контекста. Например, ничего не значащая для одного из авторов фраза «демократический централизм» у других вызывает из памяти целые смысловые блоки из регулятивных документов времен ВЛКСМ и КПСС.

Аналогичные рассуждения можно привести и по поводу моделирования зрительного восприятия, на современном этапе которого, как и почти полвека назад [18], опять требуется база данных по линиям, геометрическим фигурам, формам, теням, соотношению этих элементов с поступающими сигналами и т.п. То же верно и для «интеллектуальных» транспортных средств, в памяти которых должно содержаться огромное количество данных, начиная от типов дорог и автомобилей и заканчивая вероятностями появления ребенка перед автомобилем после выкатывания на дорогу мяча.

Использование принципов работы биологического разума [14, с. 202] позволит создавать действительно умные машины. Они, прежде всего, должны обладать системой сенсоров (радио, инфракрасных, ультрафиолетовых, ультразвуковых, радиационных и других), соответствующих механизмов обработки сигналов и превращения их в одинаковые по сути искусственные нервные импульсы, а также иерархической системой памяти и алгоритмом прогнозирования.

Такие машины необходимо обучать до тех пор, пока у них не будет создана модель собственного мира, такого, каким они его видят через призму собственных ощущений. При этом не потребуется их программирования с использованием баз данных, фактов, знаний или других концепций высшего порядка, несвойственных искусственному интеллекту. Разумные машины будут учиться на собственном опыте, иногда получая входные сигналы от инструктора-человека или другой машины, обучаться в полноразмерной среде. При этом само обучение в среде необходимо усложнять по этапам: «разработчик – машина», «разработчик – машина – другие машины», «машина – люди – другие машины», «машина – люди – другие машины – открытая природно-техногенная среда». Ну а для того, чтобы машина получилась мыслящей, нужно создать для нее среду с чем-то наподобие естественного (скорее искусственного) отбора или какой-то мотивацией, приводящей к необходимым изменениям.

И, как только разумная машина создаст модель своего собственного мира, она сможет распознавать аналогии, основанные на прошлом опыте, для прогнозирования событий в будущем, предлагать решение новых задач и делиться своим опытом с людьми и другими машинами.

Технические трудности при создании таких машин будут, во всей видимости, связаны с формированием иерархической системы памяти, обладающей значительной емкостью и связанностью [14, с. 204]. И если емкость мозга достижима на современных компьютерах, то проблема связности более сложна в решении. В человеческом мозге под тонким покрытием коры имеется белое вещество, состоящее из миллионов аксонов и связывающее области иерархии коры головного мозга между собой. Так, отдельная клетка неокортекса может быть связана с 5–10 тыс. других клеток. Такой тип масштабного соединения пока невозможно внедрить на основе традиционных технологий производства кремниевых чипов, предусматривающих поочередное нанесение ограниченного числа слоев металла («проводов») и изоляторов. Но в то же время в отличие от аксона, соединенного лишь с единственным нейроном, провод в чипе может обеспечить несколько связей.

Другие, менее значимые, проблемы – размеры, затраты на энергопотребление, стоимость производства – будут решаться традиционными методами.

Преодолев указанные технические трудности, системы ИИ, построенные на основе принципов работы неокортекса, будут обладать рядом преимуществ перед человеком во многих областях деятельности, но особенно в военном деле и прежде всего из-за скорости. Быстродействие нейронов связано с природой передаваемых нервными клетками электрохимических сигналов, измеряется в миллисекундах и в процессе эволюции было обусловлено скоростью изменения самого внешнего мира. Для эволюционных механизмов не было смысла ускорять мышление во много раз при изначальной инертности, статистической устойчивости окружающей среды. В случае крайне хаотичного, неустойчивого, сингулярного мира возможность появления разумных существ вообще бы отсутствовала, так как невозможно спрогнозировать хаос, «дурную неопределенность» и т. п. Применение искусственного интеллекта с реакцией в кремниевых чипах в наносекунды (в миллион раз больше человеческого) может сыграть решающую роль при ведении боевых действий, когда обстановка меняется каждый миг.

Вторым преимуществом ИИ является емкость памяти. Размер, а следовательно, и емкость мозга *homo sapiens* ограничивается рядом факторов. Важнейшими из них являются соотношение размера черепа ребенка и диаметра таза матери, а также большие метаболические затраты на поддержание мозга. Вес последнего составляет всего около 2% веса тела при потреблении 20% вдыхаемого кислорода. Увеличение емкости искусственной памяти позволит разумной боевой машине запоминать больше подробностей (разведывательных и иных), проявлять более высокую точность восприятия (при обнаружении цели, непрерывном управлении средствами поражения), создавать более детальные, качественные модели мира (траектории движения), распознавать сигналы более высокого порядка (радиолокационные и другие).

В-третьих, это реплицируемость. Каждый новый биологический мозг растет и обучается на протяжении до нескольких десятков лет. Человек в течение своей жизни проходит множество кругов обучения в семье, социальной и профессиональной среде. Несмотря на то что этот путь несметное количество раз повторяется другими людьми, модель мира в коре головного мозга каждого создается в индивидуальном порядке. Разумным машинам нет необходимости проходить подобную спиралевидную кривую обучения, поскольку хранящуюся информацию можно копировать неограниченное количество раз. Сам процесс разработки, настройки, испытания «мозгов» боевого робота может длиться долго. Но, как только получен финишный продукт, его можно запускать в производство и изготавливать армии таких роботов. При этом для некоторых из них, наверное, потребуется ограничение их возможностей, с тем чтобы они работали известными и испытанными способами. Как только умная машина усвоит все, что ей положено знать, будут приняты меры к тому, чтобы у нее не образовались вредные привычки и не закреплялись самостоятельно ложные аналогии. Так возможно будет избежать большинства рисков искусственного интеллекта, вплоть до бунта машин [1].

Но для других приложений может оказаться целесообразной возможность постоянного обучения мозгоподобных систем памяти, например, для интеллектуальных станций обнаружения целей. Для них необходимо каждый раз применять старые решения для новых проблем – идентификации своих и чужих летательных аппаратов, птиц и дронов, мушек и мини-роботов, органической пыли и наномеханизмов.

Четвертое преимущество может быть связано с новым решением в различных областях, например в глобальных сенсорно-интеллектуальных системах метеорологического обеспечения полетов. Для этого необходимо разместить органы чувств такой системы (сенсоры) на всем земном и околоземном пространстве. Привязав такую сенсорную «паутину» к памяти, работающей по принципу головного мозга, возможно научить систему предсказывать погоду с огромной точностью на длительные периоды времени. Это позволило бы более достоверно прогнозировать баллистические траектории, увеличить точность попадания в цель, определения мест падения отделяемых блоков ракет и т. п.

Кроме того, при наличии соответствующих ощущений и небольшом структурном изменении памяти коры головного мозга разумные машины будущего могли бы существовать в виртуальном мире математики, физики, химии, биологии и т. п. И возможно, что они окажут бы неоценимую услугу разработчикам новых видов вооружений, для которых могли бы использоваться теории не четырехмерного линейного, а многомерного искривленного пространства-времени и т. п.

Прогнозируя временные рамки создания таких машин, часто говорят о середине XXI столетия. По мнению Рэя Курцвайля, приведенному в [8, с. 380], к этому моменту наступит сингулярность и ИИ превзойдет человеческий. Такой интеллект начнет самостоятельно производить все более и более умные машины. При этом в ближайшем будущем станет уже невозможным увеличивать мощность компьютеров путем уменьшения размеров транзисторов. Сегодня самые тонкие слои кремния на микросхемах Pentium составляют примерно 20 атомов в толщину. При уменьшении этих размеров до 5 атомов вступит в силу соотношение неопределенностей Гейзенберга – не получится точно определить местонахождение электронов. Они могут уйти с проводящего слоя и привести к короткому замыканию в микросхеме. Поэтому рост мощности будет, скорее всего, связан с увеличением как размеров, так и энергопотребления компьютеров. Поэтому более умным машинам потребуются полезные ископаемые в огромных количествах, не ограниченных рамками ни планеты Земля, ни Солнечной системы. Единственным фактором, который позволит сдерживать экспоненциальный рост интеллектуальности машин, будет являться скорость света.

Заключение

Возвращаясь к более реальным проблемам использования «сильного» ИИ, следует отметить, что наибольшая востребованность технологий небиологического мышления существует в военном деле. При этом и возможности создания такого ИИ для нашей страны сосредоточены именно в оборонно-промышленном комплексе. Сейчас большинство технологических разработок в мире не транслируется из военного сектора в гражданский, как было раньше в атомной отрасли, сфере управляемых полетов ракетных комплексов и др., а, наоборот, военными структурами отбирается, селекционируется из бизнеса, промышленности, науки все передовое, что можно использовать в вооруженной борьбе. Но применительно к ИИ такая тенденция неприменима из-за масштабности и глубины проблемы, необходимости сосредоточения усилий больших коллективов разработчиков. Кроме того, в частном бизнесе, основной целью которого является извлечение прибыли, ставятся частные задачи совершенствования своей продукции в узкой области применения, и это приводит, как правило, к развитию систем автоматизации и реже – систем «слабого ИИ».

В заключение следует отметить, что вряд ли тот, кто создаст сильный искусственный интеллект, безоговорочно победит в глобальном военно-политическом, экономическом и другом противостоянии. Скорее всего, создание разумных машин в военной сфере значи-

тельно увеличит военную мощь армии, позволит перенести сферу вооруженной борьбы в более гуманное киберпространство (без людских потерь в процессе компьютерной борьбы за управление экономическими мощностями, территориями, сознанием людей), будет являться наряду с традиционным ядерным оружием важнейшим фактором сдерживания любой агрессии любым противником и, самое главное, позволит сдвинуть решение проблем «сильного ИИ» с мертвой точки во всех сферах жизнедеятельности.

Список литературы

1. Буренок В.М., Дурнев Р.А., Крюков К.Ю. Разумное вооружение: будущее искусственного интеллекта в военном деле. М.: Журнал «Вооружение и экономика», вып. 1(43), 2018.
2. Пегат А. Нечеткое моделирование и управление. М.: Издательство «Бином», 2011. 798 с.
3. Тарик Р. Создаем нейронную сеть. Пер. с англ. СПб.: ООО «Альфа-книга», 2018. 272 с.
4. Наука. Величайшие теории: выпуск 15: Размышления о думающих машинах. Тьюринг. Компьютерное исчисление / Пер. с исп. М.: Де Агостини, 2015. 152 с.
5. Зверева С.В. Человеческий фактор; проблемы функционального рассогласования искусственных интеллектуальных систем и естественного мышления человека. Доклад на XXI Всероссийской научно-практической конференции «Актуальные проблемы защиты и безопасности». СПб., ВУНЦ ВМА им. Н.Г. Кузнецова, 2018.
6. Глобальное будущее 2045: антропологический кризис. Конвергентные технологии. Трансгуманистические проекты: Материалы Первой всероссийской конференции, Белгород, 11–12 апреля 2013 г. / Под ред. Д.И. Дубровского, С.М. Климовой. М.: «Канон+» РООИ «Реабилитация», 2014. 272 с.
7. Рассел С., Норвиг П. Искусственный интеллект: современный подход. Пер. с англ. М.: Издательский дом «Вильямс», 2006. 1409 с.
8. Каку Митио. Будущее разума. Пер. с англ. М.: Альпина нон-фикшн, 2018. 502 с.
9. Пенроуз Р. Тени разума: в поисках науки о сознании. Москва–Ижевск: Институт компьютерных исследований, 2005. 688 с.
10. Тейяр де Шарден П. Феномен человека. М.: АСТ: Астрель: Политграфиздат, 2012. 384 с.
11. Пелевин В.О. Ананасная вода для прекрасной дамы. М.: Эксмо, 2011. 352 с.
12. Карпов В.Э. Эмоции и темперамент роботов. Поведенческие аспекты. Известия РАН. Теория и системы управления, вып. 5. 2014. 232 с.
13. Николенко С., Кадуринов А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. СПб: Питер, 2018. 480 с.
14. Хоккинс Дж. Об интеллекте. Пер. с англ. М.: ООО «И.Д. Вильямс», 2007. 240 с.
15. Lazebnik Y. Can a Biologist Fix a Radio? - or, What I Learned while Studying Apoptosis // Biochemistry (Moscow), 2004. Vol. 69, no. 12.
16. Jonas E., Kording K. Could a Neuroscientist Understand a Microprocessor? // bioRxiv, 2016.
17. Потапов А.С. Искусственный интеллект и универсальное мышление. СПб.: Политехника, 2012. 711 с.
18. Саймон Г.А. Науки об искусственном. М.: Мир, 1972. С. 146.

References

1. Burenok V.M., Durnev R.A., Kryukov K.Yu. (2018) *Razumnoe vooruzhenie: budushchee iskusstvennogo intellekta v voennom dele* [Reasonable weapons: the future of artificial Intelligence in military affairs] *Zhurnal «Vooruzhenie i ekonomika»* [Magazine «Armament and Economics»]. Issue 1(43). Moscow.
2. Pegat A. (2011) *Nechetkoe modelirovanie i upravlenie* [Fuzzy modeling and control] *Izdatel'stvo «Binom»* [Publishing house «Binom»]. Moscow. P. 798.
3. Tarik R. (2018) *Sozdaem neyronnuyu set'. Per. s angl.* [Creating a neural network. Translation from English] *Al'fa-kniga* [Alfa-Kniga]. St. Petersburg. P. 272.

4. (2015) *Nauka. Velichayshie teorii Razmyshleniya o dumayushchikh mashinakh. T'yuring. Komp'yuternoe ischislenie* [Reflections on Thinking Machines. Turing Computer calculus] *Nauka. Velichayshie teorii* [Science. The Greatest Theories]. Moscow. Issue 15. P. 152.
5. Zvereva S.V. (2018) *Chelovecheskiy faktor; problemy funktsional'nogo rassoglasovaniya iskusstvennykh intellektual'nykh sistem i estestvennogo myshleniya cheloveka. Doklad na XXI Vserossiyskoy nauchno-prakticheskoy konferentsii «Aktual'nye problemy zashchity i bezopasnosti»* [Human factor; problems of functional misalignment of artificial intellectual systems and human natural thinking. Report at the XXI All-Russian Scientific and Practical Conference «Actual problems of protection and safety»] *VUNTs VMA im. N.G. Kuznetsova* [Military training and research center of the Navy «Naval Academy named after Admiral of the Fleet of the Soviet Union N.G. Kuznetsov»]. St. Petersburg.
6. (2014) *Global'noe budushchee 2045: antropologicheskiiy krizis. Konvergentnye tekhnologii. Transgumani-sticheskie proekty: Materialy Pervoy vs Rossiyskoy konferentsii, Belgorod, 11–12 aprelya 2013 g. Pod red. D.I. Dubrovskogo, S.M. Klimovoy* [Global Future 2045: anthropological crisis. Convergent technology. Transhumanist projects Materials of the First All-Russian Conference. Belgorod. April 11–12, 2013. Ed. D.I. Dubrovsky, S.M. Klimova] «Kanon+» ROOI «Reabilitatsiya» [Publishing House «Canon +» ROOI «Rehabilitation»]. P. 272.
7. Russell S., Norvig P. (2006) *Iskusstvennyy intellekt: sovremennyy podkhod. Per. s angl.* [Artificial Intelligence: A Modern Approach. Translation from English] *Izdatel'skiy dom «Vil'yams»* [Williams Publishing House]. Moscow. P. 1409.
8. *Kaku Michio (2018) Budushchee razuma. Per. s angl.* [The future of reason. Translation from English]. *Al'pina non-fikshn* [Alpina non-fiction]. Moscow. P. 502.
9. Penrose R. (2005) *Teni razuma: v poiskakh nauki o soznanii* [Shadows of the mind: in search of a science of consciousness] *Moskva–Izhevsk. Institut komp'yuternykh issledovaniy* [Moscow–Izhevsk: Institute for Computer Studies]. P. 688.
10. Teilhard de Chardin P. (2012) *Fenomen cheloveka* [The phenomenon of man] *AST. Astrel'. Politgrafizdat* [AST. Astrel. Politgrafizdat]. Moscow. P. 384.
11. Pelevin V.O. (2011) *Ananasnaya voda dlya prekrasnoy damy* [Pineapple water for a beautiful lady] *Eksmo* [Eksmo]. Moscow. P. 352.
12. Karpov V.E. (2014) *Emotsii i temperament robotov. Povedencheskie aspekty* [Emotions and temperament of robots. Behavioral aspects] *Izvestiya RAN. Teoriya i sistemy upravleniya* [Proceedings of the RAS. Theory and control systems]. Vol. 5. P. 232.
13. Nikolenko S., Kadurin A., Arkhangelskaya E. (2018) *Glubokoe obuchenie. Pogruzhenie v mir neyronnykh setey* [Deep Learning. Immersion in the world of neural networks] *Piter* [Peter]. St. Petersburg. P. 480.
14. Hokkins J. (2007) *Ob intellekte. Per. s angl.* [On the intellect. Translation from English] *I.D. Vil'yams* [Williams Publisher]. Moscow. P. 240.
15. Lazebnik Y. (2004) Can a Biologist Fix a Radio? - or, What I Learned while Studying Apoptosis. *Biochemistry*. Moscow. Vol. 69. No. 12.
16. Jonas E., Kording K. (2016) Could a Neuroscientist Understand a Microprocessor? «BioRxiv».
17. Potapov A.S. (2012) *Iskusstvennyy intellekt i universal'noe myshlenie* [Artificial intelligence and universal thinking] *Politekhnik* [Polytechnic]. St. Petersburg. P. 711.
18. Simon G.A. (1972) *Nauki ob iskusstvennom* [Science of artificial] *Mir* [Mir]. Moscow. P. 146.