УДК 111, 168
DOI:
10.15372/PS20200308

Е.А. Безлепкин

СООБЩЕСТВО РАЗУМА: ИДЕИ М. МИНСКОГО ОБ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТАХ

В статье рассмотрены основные подходы к моделированию слабого искусственного интеллекта. Концепция интеллектуальных агентов – один из перспективных подходов, предложенный М. Минским. Цель статьи – осуществить реконструкцию его идей относительно принципов работы сознания и принципов построения искусственного интеллекта. Рассмотрены базовые понятия и концепции представлений об устройстве сознания и устройстве памяти. Проведена классификация концепции Минского с философских позиций. Сделано заключение, что в онтологическом плане с точки зрения структурного анализа концепция представляет собой совокупность подходов коннекционизма и элементаризма, с точки зрения системного анализа – совокупность сетевой и иерархической моделей. В методологическом плане базой концепции явились подходы редукционизма и компьютерной метафоры.

Ключевые слова: слабый искусственный интеллект; М. Минский; интеллектуальный агент; коннекционизм; сознание; память

E.A. Bezlepkin

THE SOCIETY OF MIND: M. MINSKY'S IDEAS ABOUT INTELLECTUAL AGENTS

The article discusses the basic approaches to modeling weak artificial intelligence. The concept of intelligent agents is one of the promising approaches proposed by M. Minsky. The purpose of the article is to reconstruct his ideas regarding principles of consciousness and principles of creating artificial intelligence. The article considers the basic notions of the concept, the structure of consciousness and the structure of memory. Minsky's concept is classified from a philosophical perspective. The conclusion is made that in the ontological terms, with regard to the structural analysis, the concept is a set of approaches of connectionism and elementarism, while with regard to the system analysis, it is a set of a network model and a hierarchical one. Methodologically, the concept is based on the approach of reductionism and the approach using a computer metaphor.

Keywords: weak artificial intelligence; M. Minsky; intelligent agent; connectionism; consciousness; memory

Введение: Подходы к моделированию искусственного интеллекта

В настоящее время можно четко выделить несколько основных подходов к моделированию слабого искусственного интеллекта (о понятиях слабого и сильного искусственного интеллекта см. [2]). Подходы классифицируются достаточно интуитивно в зависимости от положенных в их основу концепций (похожую классификацию также можно найти в статье С. Брингсйерда и Н.С. Говиндараюлу «Искусственный интеллект» [8]).

Наиболее успешным как в коммерческом плане, так и в плане теоретической разработанности является подход, основанной на концепции нелогического искусственного интеллекта (ИИ). С. Брингсйерд и Н.С. Говиндараюл пишут: «…Нелогический ИИ – это ИИ, осуществляемый на основе определенных формализмов, отличных от логических систем. Конечно, эти формализмы не включают в себя знания в обычном смысле» [8].

Основное направление в рамках этого подхода – моделирование искусственных нейронных сетей, опирающееся на методологию коннекционизма. Суть коннекционизма заключается в возможности моделирования строения мозга через моделирование биологических нейронных сетей. Основной принцип состоит в следующем: мышление как процесс может быть описано с помощью простых вычислительных элементов, которые связаны в сети. Если для биологической сети элементами являются нейроны, а связями – синапсы, то для искусственной сети элементами являются, например, физические элементы памяти, а связями – провода (или все это может быть смоделировано в компьютерной программе). Если «идеи представлены нейронами, совокупности (ассоциации) идей – связями между нейронами, а воспоминание – клеточным ансамблем или синоптической цепочкой» [7, с. 104], то в теории знание точно так же может быть формализовано как нейронная сеть, интегрированная с правилами вывода.

Второй подход, логический ИИ, предполагает, что осмысленные действия субъекта могут моделироваться как символьные манипуляции. В рамках этого подхода считается, что без символьных вычислений осмысленные действия невозможны и, наоборот, наличия символьных вычислений достаточно для того, чтобы действие можно было считать осмысленным. Эта идея известна как гипотеза Ньюэлла–Саймона: «физи-

ческая символьная система имеет необходимые и достаточные средства для произведения основных интеллектуальных операций» [12].

Главный посыл состоит в том, что интеллектуальный агент трансформирует восприятия из внешнего мира в символы, выраженные в логической системе, и на основе своей базы знаний выводит, какие действия должны выполняться для достижения целей агента. Наиболее успешная на данный момент реализация этого подхода – экспертные системы. Они являются программами, которые моделируют знания эксперта в определенной области (например, в медицине или юриспруденции).

Существуют и смешанные направления, например направление, опирающееся на концепцию когнитивных архитектур. Когнитивная архитектура – это «гипотеза о фиксированных структурах, которые являются условием появления разума, будь то в естественных или искусственных системах, и о том, как они работают вместе – в сочетании со знаниями и навыками, воплощенными в архитектуре, для обеспечения интеллектуального поведения в разнообразных сложных средах» [9].

Например, в теории когнитивной архитектуры Soar [10] выдвинуты гипотезы о том, что такое познание и как построить его вычислительную реализацию. Soar предлгает несколько гипотез о вычислительных структурах, лежащих в основе интеллекта. Центральная гипотеза заключается в том, что целенаправленное поведение может быть смоделировано как вектор в математическом пространстве возможных состояний. Каждому элементарному действию соответствует оператор, который можно применить к текущему состоянию субъекта, что приводит к ментальным изменениям. В свою очередь, измененное внутреннее состояние порождает целенаправленные действия во внешнем мире.

Последний подход, о котором и пойдет речь в статье, основан на идее интеллектуальных агентов. Интеллектуальный агент – это сущность, которая получает восприятия из окружающей среды и выполняет действия. Ниже мы будем рассматривать концепцию интеллектуальных агентов, предложенную М. Минским в книге «Сообщество разума» («The Society of Mind»), вышедшей в свет в 1986 г. На основе идей Минского С. Рассел и И. Норвиг опубликовали ставшую классической монографию «Искусственный интеллект», в которой определяют понятие интеллектуального агента следующим образом: «Агентом является все, что может рассматриваться как воспринимающее свою среду с помощью датчиков и воздействующее на эту среду с помощью исполнительных механизмов. …Поведение некоторого агента может быть описано с помощью функции агента, которая ото-

бражает любую конкретную последовательность актов восприятия на некоторое действие» [6, с. 75–76].

Основные понятия концепции интеллектуальных агентов

Два базовых понятия концепции – это «агенты» и «операторы». Высокая степень абстрактности этих понятий позволяет дать им множество интерпретаций.

Во-первых, с биологической точки зрения агентом можно считать как отдельный нейрон, так и совокупность нейронов, выполняющих какую-то одну функцию. Под оператором в самом общем виде можно понимать соединения нейронов, т.е. дендриты и синапсы.

Во-вторых, с философской точки зрения агенты представляют собой сущности. С их помощью описывается статическое состояние системы. Операторы – это функции над сущностями. С их помощью описывается динамическое состояние системы.

В-третьих, с логической точки зрения через понятие агента можно представить любой требуемый для исследований уровень абстракции. Необходимое условие заключается в том, что элементы этого уровня должны быть способными действовать как единое целое. Например, если двигаться сверху вниз по ступеням абстрагирования, то в качестве агента можно представить сознание в целом, функцию внимания, нейроны зоны Брока, «нейрон бабушки». Понятие оператора с этой точки зрения отражает изменения в слоях и взаимодействия между слоями как одного, так и разных уровней абстракции.

Теперь приведем определения этих понятий, данные Минским. «Сознание, – пишет он, – представляется состоящим из множества мелких процессов. Указанные процессы мы будем называть агентами. Каждый ментальный агент по отдельности выполняет некое простое действие, для чего не требуется ни разум, ни мышление. Тем не менее, когда мы объединяем указанных агентов в сообщества – это ведет к возникновению подлинного интеллекта» [4, с. 2].

Агенты в обществе разума организованы по степени абстрактности их функции. Например, существует агент «Играть», который управляет агентом «Играть с кубиками», который управляет агентом «Строитель» или «Крушитель», которые управляют агентами «Начать», «Добавить», «Закончить». Как можно видеть, мы выстроили иерархию агентов, обладающих разной степенью функциональной аб-

стракции. При этом агенты, обладающие одной степенью функциональной абстракции, находятся на одном уровне и образуют сетевые структуры. Минский пишет, что структура агентов, составляющих общество разума, напоминает иерархическое общество, в котором «агент на каждой крупной "ветви" полностью ответственен за агентов на малых "ветках", отходящих от нее» [4, с. 18]. Важной идеей в концепции является способность верхних уровней абстрагировать результаты работы нижних уровней. Например, агенты зрения занимаются обработкой зрительных сигналов, в то время как агенты зрительной коры концептуально обрабатывают эти данные, в то время как агенты сознания занимаются семантической обработкой полученных с нижележащего уровня данных и т.д.

Вернемся к понятию оператора. Оператор – это «любой элемент мыслительного процесса, который достаточно просто вычленить и понять, хотя взаимодействие между группами операторов способно порождать более сложные явления» [4, с. 285].

Существует много видов операторов, среди которых наиболее важными являются следующие два. Во-первых, это нома – «оператор, выходной сигнал которого воздействует на агента каким-либо "предопределенным" способом… деятельность этого оператора больше зависит от генетической архитектуры, чем от обучения на собственном опыте» [4, с. 285]. Во-вторых, это нема – «оператор, выходной сигнал которого репрезентирует фрагмент идеи или ментального состояния» [2, с. 285]. Например, к неме можно отнести оператор, который активирует фрагмент воспоминания.

Концепция устройства памяти по М. Минскому

Свою концепцию памяти Минский основывает на модели Д. Нормана, который выделяет две базовые структуры: первичную память, хранящую временную информацию, и вторичную память, хранящую информацию длительное время.

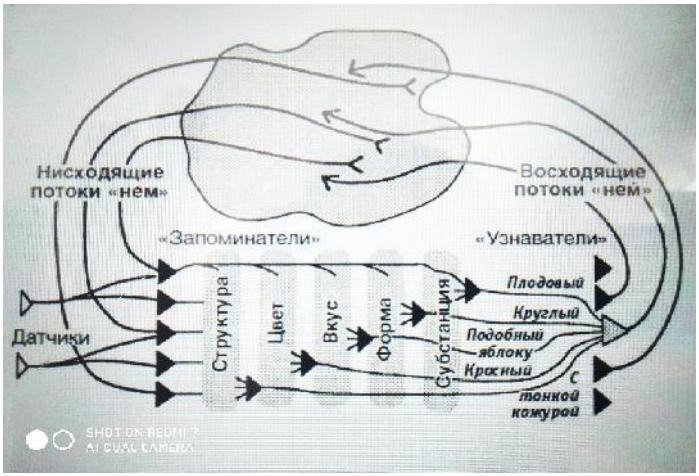
Минский пишет: «Когда вам "приходит отличная идея", когда вы справляетесь с задачей или приобретаете существенный опыт, происходит активация строки З, призванной "выразить" это воспоминание. Строка З есть подобная проволоке структура, которая "прикрепляется" к активным ментальным агентам, когда мы решаем задачу или обдумываем отличную идею. Если активировать строку З, впоследствии

связанные с нею агенты возбуждаются, вводя нас в “психическое состояние”, схожее с тем, которое сложилось, когда мы решили или породили идею… Другими словами, мы “запоминаем” то, о чем думаем, составляя список агентов, участвующих в этой деятельности» [4, с. 59]. Эта идея перекликается с моделью синапсов Д. Хебба, который писал, что если, грубо говоря, один нейрон постоянно или многократно активирует второй нейрон, то связь между ними становится более сильной.

Далее Минский отмечает, что в разуме отсутствует общая, или единая, система памяти. «Каждый новый слой возникает как совокупность строк З, которые порождаются усвоением любых навыков, полученных предыдущим слоем. Всякий раз, когда слой приобретает некоторое полезное и существенное умение, он склонен прерывать обучение и останавливать перемены, но затем другой слой начинает изучать возможности предыдущего. Каждый новый слой обучается, осваивает новые способы использования умений старых слоев. Далее он замедляет скорость обучения – и становится одновременно объектом изучения и “наставником” для слоев, которые формируются впоследствии» [4, с. 71].

В модели устройства языка и памяти Минского необходимо выделить еще один важный элемент – фреймы. «Фрейм есть своего рода костяк – или нечто вроде бланка заявки со множеством пустых граф, ожидающих заполнения. Мы будем называть эти “бланки” терминалами; они используются как точки подключения, к которым можно присоединить информацию. Например, фрейм, репрезентирующий “стул”, может иметь терминалы для сиденья, спинки, ножек, а фрейм, репрезентирующий “человека”, будет иметь терминалы для тела, головы, рук и ног. Чтобы представить стул или конкретного человека, мы просто заполняем терминалы соответствующего фрейма конструкциями, которые более подробно описывают особенности спинки, сиденья и ножек данного стула (или облика данного человека)» [4, с. 208].

Теперь мы можем описать процесс распознавания образов, возникновение воспоминаний и устройство кратковременной памяти по Минскому [4, с. 175]. Этот процесс изображен на рисунке.



Цикл активации воспоминания по М. Минскому [4, с. 174]

Давайте представим, что мы увидели яблоко. Перед нами нечто красноватое по форме напоминающее яблоко с характерным ароматом. Как только мы начинаем думать над этим, активируются агенты распознавания образа. Далее они активируют агентов, связанных с определением предмета по его свойствам, и передают им несколько определенных свойств: краснота, размер, запах. Агенты определения предметов подключены к терминалам фреймов. Переданные свойства сравниваются этими агентами с терминалами фреймов до тех пор, пока не будет найдено максимальное соответствие. Если найдено несколько подходящих фреймов, то перед нами возникает ситуация неопределенности, в которой мы не знаем, что точно находится перед нами: маленький мячик или яблоко. Для устранения ситуации неопределенности необходимо дальнейшее взаимодействие с предметом и вычленение других его свойств.

После узнавания предмета запускается процесс нахождения ближайших кратковременных воспоминаний о нем. Как только мы начинаем думать над словом «яблоко», активируется наш агент распознавания слова. Далее он, в свою очередь, активирует строки З, которые ведут к агентам, отвечающим за главные, по нашему мнению, свойства яблока. Это действие пробуждает частичное воспоминание о яблоках. После

этого события запускается цикл: частичное воспоминание активирует связанных с ним агентов, отвечающих за остальные концептуальные свойства яблока, а те, в свою очередь, активируют строки З агентов, связанных с воспоминаниями.

Минский выражает подлинное восхищение тем, что описанный выше процесс можно закодировать в компьютерной программе, имеющей устройства ввода и вывода: «Простой цикл способен реконструировать целое из сведений о некоторых частях объекта!» [4, с. 175]. Отметим, что современные нейросети в достаточной степени могут справляться с задачей классификации изображений и предметов.

Концепция устройства сознания по М. Минскому

Решение проблемы устройства сознания требует ответа на следующие стандартные вопросы [3].

1. Каким образом создаваемая мозгом модель реальности оказывается не простым набором звуковых, визуальных, обонятельных и прочего рода данных, а миром, воспринимаемым как целостный и единственный?
2. Почему мы воспринимаем созданную мозгом субъективную реальность не как модель, а как реальный внешний мир?
3. Что такое субъективное переживание, существует ли оно вообще и можно ли передать его от человека к человеку?
4. Что такое эго (самосознание)? Каков его онтологический статус? Необходимо ли эго для получения субъективного опыта?

Минский не пытается ответить на все эти вопросы, но показывает, что найти ответ принципиально возможно. При этом его ответы либо вытекают из концепции интеллектуальных агентов, либо согласуются с ней.

Прежде всего, Минский в «Сообществе разума» не дает четкого определения термина «сознание», однако описывает ключевую, по его мнению, функцию сознания: «сознание проявляет себя в основном в ситуациях, когда другие наши системы начинают испытывать проблемы» [4, с. 47]. Под ментальным состоянием он понимает «активное состояние группы агентов в конкретный момент времени» [4, с. 284].

Реконструируем ответ на первый вопрос. По мнению Минского, единство сознания связано со «слоистой» организацией памяти, при которой слои могут быть связаны горизонтально и вертикально. При этом слои связаны «управляющими» отношениями. «Наша “цельность” приобретается посредством последовательности “этапов развития”. Каждый новый этап сначала опирается на результаты предыдущих, происходит усвоение знаний, ценностей и целей. Затем роли меняются, и уже этот этап становится “учителем” предыдущих» [4, с. 146].

Ответ на второй вопрос звучит так: «Идея единого, центрального “я” ничего не объясняет. Нечто без составных частей не имеет ничего такого, что возможно было бы использовать как фрагмент объяснения! Тогда почему мы столь часто озвучиваем ту странную мысль, что наши поступки и дела совершает Другой, то есть наше “я”? Да потому что значительная часть нашего разума надежно скрыта от тех элементов личности, которые вовлекаются в осознанное вербальное общение» [4, с. 31]. Эта идея Минского созвучна с частью модели Т. Метцингера [3], в которой тот описывает свойство транспарентности сознания. Транспарентная модель мира – это такая модель, построение которой нашим мозгом мы никак не воспринимаем, поэтому она ощущается нами как единственная и реальная.

По поводу субъективности Минский в целом не говорит ничего определенного. Однако у него можно найти множество высказываний об эмоциях и даже целую книгу, посвященную эмоциям [11]. Как известно, эмоции осуществляют оценочную функцию сознания: они не дают информации об объектах, но характеризуют отношение человека к объектам, значимость определенных объектов для человека. Например, В.М. Аллахвердов писал, что «эмоции нужны, чтобы извещать сознание о происходящем в не осознаваемой сфере» [1, с. 156]. Примерно той же идеи придерживается и Минский [11]: поскольку сознание является высокоуровневым элементом психики, ему необходимы сведения о функционировании низкоуровневых элементов, которые и предоставляются эмоциями.

Что касается последнего вопроса, то, как нам кажется, Минский сводит самосознание к представлению об обратной связи в компьютерных системах. Например, показательна его попытка объяснить такой важный аспект сознания, как понимание. «Откуда вообще берется понимание? Думаю, почти всегда оно проистекает из того или иного вида аналогии, то есть мы сопоставляем всякий новый объект с теми, которые нам уже известны. Когда принципы устройства и работы нового объекта

представляются слишком чуждыми, странными или слишком сложными, мы описываем те части устройства, которые нам доступны, в терминах более привычных знаков. Тем самым мы вынуждаем каждую новинку походить на что-то обыденное, что-то знакомое» [4, с. 36].

Идеи об искусственном интеллекте

На страницах книги «Сообщество разума» можно найти несколько интересных идей о моделировании искусственного интеллекта, которые мы обсудим ниже.

Прежде всего отметим, что Минский принимает принцип Тьюринга, утверждающий, что принципиально возможно построить универсальный компьютер. Под универсальным компьютером подразумевается машина, которую можно запрограммировать для выполнения любого вычисления, выполнимого любым другим физическим объектом. Вариант формулировки Минского звучит так: «возможно запрограммировать компьютер для решения любых задач методом проб и ошибок без готовых вариантов решений при условии, что мы располагаем способом установить факт решения задачи» [4, с. 50]. Затем Минский он углубляет понимание этого принципа в методологическом ключе, добавляя, что метод проб и ошибок (т.е. процесс поиска решения) должен опираться на понятие «прогресс» и на умение зафиксировать достижение частичного прогресса в процессе поиска решения.

Далее хотелось бы привести ключевые высказывания Минского о двух главных направлениях моделирования слабого искусственного интеллекта: нейронных сетях и экспертных системах.

Интересно, что в свое время Минский невысоко оценил нейронные сети. Он опубликовал математическое доказательство ограниченности перцептрона [5] (т.е. его неспособности решать некоторые простейшие задачи), вследствие чего интерес к нейронным сетям резко упал. Однако в связи с появлением многослойных сетей и с открытием и развитием метода обратного распространения ошибки удалось показать ошибочность выводов Минского.

По теме универсальных экспертных систем Минский высказал интересное наблюдение о том, что проблема их моделирования упирается в проблему формализации здравого смысла: «Чтобы считаться “экспертом”, нужно обладать изрядными познаниями относительно сравнительно немногочисленных фактов. По контрасту, “здравый смысл” обычного

человека подразумевает гораздо большее разнообразие понятий, которые требуют более сложной системы управления» [4, с. 50].

Заключение: Философская классификация концепции М. Минского

Под философской классификацией мы будем понимать отнесение концепции к соответствующим онтологическим, эпистемологическим и методологическим основаниям.

Под онтологическими основаниями можно в первом приближении понимать те, которые описывают структурные и системные свойства исследуемого объекта. Отсюда вытекает возможность двойственной онтологической классификации концепции Минского. Эта концепция рассматривает мозг и сознание, во-первых, через призму взаимодействия агентов одной степени абстракции (т.е. структурно), а во-вторых, через призму взаимодействия агентов разной степени абстракции (т.е. системно).

Со структурной точки зрения основой концепции является философия коннекционизма. Идеи коннекционизма в первом приближении связаны с тем, что мыслительные процессы, поведение и в конечном счете устройство мозга можно вывести из совокупности сетей связанных между собой простых элементов-нейронов.

Другая вытекающая отсюда философская идея – это элементаризм. Под элементаризмом мы будем понимать следующее: любой исследуемый уровень является структурой, которую можно разбить на две составляющие, а именно элементы (в нашем случае – агенты) и их взаимосвязи (в нашем случае – операторы). Вторая составляющая элементаризма – идея приоритета элементов перед тем целым, которое они образуют. Целое (в нашем случае – мозг) предстает перед нами как интегральная характеристика из составляющих элементов.

Концепции агента примечательна тем, что мы и сам мозг можем рассматривать как агента, если абстрагируемся от его внутреннего устройства. И тогда разговоры о первичности и вторичности элементов и целого теряют смысл, потому что нельзя абсолютно точно сказать, что является элементом, а что – целым: любой объект (от нейрона до мозга) мы можем рассмотреть как агента со специфическими функциями.

С системной точки зрения основой концепции является идея иерархической организации агентов. С одной стороны, агенты одного уровня

абстракции находятся во взаимосвязи друг с другом (сетевая организация). С другой стороны, агенты разного уровня абстракции находятся в отношении управления и подчинения. Все вместе можно рассматривать как пирамиду, где на вершине находится агент-сознание, а у основания – агенты-нейроны.

Рассмотрим методологические основания. Первое – это идея редукционизма. С развитием науки в Новое время утвердилась программа редукционизма, которая ставила целью объяснить сложные явления с помощью комбинации простых явлений. Наиболее ярким продуктом этой идеи является концепция атомизма, которая стремится объяснить явления природы с помощью сочетаний и свойств простых элементов-атомов. В каком-то смысле редукционизм отрицает появление эмерджентных свойств при количественных скачках усложнения системы, потому что это приводит к разрыву преемственности объяснения на пути от простого уровня к более сложному. Например, Минский пишет следующее: «Мы желаем объяснить интеллект как совокупность простых явлений и процессов» [4, с. 8].

Второе методологическое основание – компьютерная метафора. Редукционизм в методологическом плане стремится найти универсальную метафору для исследования любого процесса. В Новое время часы и автоматы были метафорой для понимания устройства мира и разума. Принципиально важно здесь то, что эти механизмы строились в соответствии с рассмотренными нами онтологическими принципами: сложность любого механизма можно было свести к связи простых элементов. И это являлось ключевой идеей.

В XX веке механической модели в области психологии соответствовал бихевиоризм с его отказом от рассмотрения сознания. Если говорить кратко, то идея бихевиоризма сводится к тому, что сознания не существует. Представление о сознании возникает вследствие наличия языка и долговременной памяти. Позже с появлением и развитием компьютерной техники возникает идея объяснить устройство мозга и сознания на основе принципов устройства компьютера, т.е. через взаимосвязи простых элементов-транзисторов, выполняющих заложенную в них программу. Так рождается идея, что компьютерные программы действуют точно так же, как сознание человека. Таким образом, компьютерные программы становятся моделью для понимания процессов обработки информации в сознании.

Концепция интеллектуальных агентов Минского в настоящее время является весьма привлекательным направлением в изучении и модели-

ровании слабого искусственного интеллекта. Ее привлекательность во многом обусловлена тем, что она базируется на классическом концептуальном каркасе. Интеллект, построенный с помощью этой концепции, будет работать на понятных основаниях и проводить вычисления по понятной логике, во многом схожей с человеческой. Это отличает концепцию интеллектуальных агентов от, например, концепции нейронных сетей, в которых процесс вычисления не детерминирован. Однако, как нам кажется, для настоящего прорыва в области моделирования слабого искусственного интеллекта необходимо возникновение гибридных концепций.

Литература

1. *Аллахвердов В.М.* Сознание как парадокс (Экспериментальная психологика, т. 1). – СПб.: Издательство ДНК, 2000. – 528 с.

2. *Безлепкин Е.А.* Искусственный интеллект: философские основания и перспективы // Философия науки. – 2019. – № 4 (83). – С. 134–146.

3. *Метцингер Т.* Наука о мозге и миф о своем Я: Тоннель эго. – М.: АСТ, 2017. – 430 с.

4. *Минский М.* Сообщество разума. – М.: АСТ, 2018. – 592 с.

5. *Минский М., Пейперт С.* Персептроны. – М.: Мир, 1971. – 264 с.

6. *Рассел С., Норвиг П.* Искусственный интеллект: Современный подход. – 2-е изд. – М.: ИД «Вильямс», 2006. – 1408 с.

7. *Сеунг С.* Коннектом: Как мозг делает нас тем, что мы есть. – М.: БИНОМ. Лаборатория знаний, 2014. – 443 с.

8. *Bringsjord S., Govindarajulu N.S.* Artificial intelligence // The Stanford Encyclopedia of Philosophy. – URL: https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence (дата обращения: 02.05.2020).

9. *Cognitive* Architecture. University of Southern California Institute for Creative Technologies. – URL: https://cogarch.ict.usc.edu (дата обращения: 02.05.2020).

10. *Laird J.E.* The Soar Cognitive Architecture. – MIT Press, 2012. – 390 p.

11. *Minsky M.* The Emotional Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind. – N.Y.: Simon & Schuster, 2006. – 400 p.

12. *Newell A., Simon H.* Computer science as empirical inquiry: symbols and search // Communications of the ACM. – 1976. – Vol. 19, No. 3. – P. 113–126.

References

1. *Allakhverdov, V.M.* (2000). Soznanie kak paradoks (Eksperimentalnaya psikhologika, t. 1) [Consciousness As a Paradox (Experimental Psychologics, Vol. 1)]. St. Petersburg, DNK Publ., 528.

2. *Bezlepkin, E.A.* (2019). Iskusstvennyy intellekt: filosofskie osnovaniya i perspektivy [Artificial intelligence: philosophical foundations and prospects]. Filosofiya nauki [Philosophy of Science], 4 (83), 134–146.

3. *Metzinger, T.* (2017). Nauka o mozge i mif o svoem Ya: Tonnel ego [The Ego Tunnel: The Science of the Mind and the Myth of the Self]. Moscow, AST Publ., 430. (In Russ.).

4. *Minsky, M.* (2018). Soobshchestvo razuma [The Society of Mind]. Moscow, AST Publ., 592. (In Russ.).

5. *Minsky, M. & S. Papert.* (1971). Perseptrony [Perceptrons]. Moscow, Mir Publ., 264. (In Russ.).

6. *Russel, S. & P. Norvig.* (2007). Iskusstvennyy intellekt: Sovremennyy podkhod [Artificial Intelligence: A Modern Approach], 2nd ed. Moscow, Viliams Publishing House, 1408. (In Russ.).

7. *Seung, S.* (2014). Konnektom: Kak mozg delaet nas tem, chto my est [Connectome: How the Brain's Wiring Makes Us Who We Are]. Moscow, BINOM, Laboratoriya Znaniy Publ., 443. (In Russ.).

8. *Bringsjord, S. & N.S. Govindarajulu.* (2020). Artificial intelligence. In: The Stanford Encyclopedia of Philosophy. Available at: https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence (date of access: 02.05.2020).

9. *Cognitive* Architecture. University of Southern California Institute for Creative Technologies. (2020). Available at: https://cogarch.ict.usc.edu (date of access: 02.05.2020).

10. *Laird, J.E.* (2012). The Soar Cognitive Architecture. MIT Press, 390.

11. *Minsky, M.* (2006). The Emotional Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind. New York, Simon & Schuster, 400.

12. *Newell, A. & H. Simon.* (1976). Computer science as empirical inquiry: symbols and search. Communications of the ACM, Vol. 19, No. 3, 113–126.

Информация об авторе

Безлепкин Евгений Алексеевич – кандидат философских наук, научный сотрудник Института философии и права СО РАН (630090, г. Новосибирск, ул. Николаева, 8)

evgeny-bezlepkin@mail.ru

Information about the author

Bezlepkin Evgeniy Alekseevich – Candidate of Sciences (Philosophy), the scientific worker, Institute of Philosophy and Law, Siberian Branch of the Russian Academy of Sciences (8 Nikolaeva str., Novosibirsk, 630090, Russia)

evgeny-bezlepkin@mail.ru

Дата поступления 22.07.2020