

В. И. Городецкий, д-р техн. наук, проф., vladim.gorodetsky@gmail.com,
Санкт-Петербургский государственный электротехнический университет "ЛЭТИ"
имени В. И. Ульянова (Ленина),

Р. М. Юсупов, член-корр. РАН, проф., yusupov@iias.spb.su,
Санкт-Петербургский институт информатики и автоматизации РАН

Искусственный интеллект: метафора, наука и информационная технология

В настоящее время научное направление в информатике, метафорически называемое искусственным интеллектом (ИИ), а также информационные технологии, построенные на его основе, приобретают стратегическую важность как теоретический и технологический базис цифровой экономики XXI века. Ведущие экономики мира инвестируют огромные деньги в исследования и разработки по ИИ. В ряде стран, например, в США и Европе, начаты работы по стандартизации технологий ИИ. В Российской Федерации Президентом подписана "Национальная стратегия развития технологий в области ИИ".

Однако нужно признать, что до настоящего времени среди исследователей и разработчиков, включая специалистов индустрии ИИ, пока что нет единого мнения об объекте исследований ИИ, о его целях, научном содержании, методологии, методах и алгоритмах ИИ, о существовании его технологий и об актуальных направлениях его развития.

В данной работе ИИ рассматривается как научная дисциплина, которая является разделом информатики. В ней объектом исследований являются знания, а научное содержание определяется методологиями, методами и алгоритмами получения, формального представления, а также использования знаний в прикладных системах принятия решений. Информационные технологии, построенные на теоретической базе ИИ, называются интеллектуальными технологиями, или технологиями ИИ.

Для обоснования этой точки зрения в работе кратко описываются основные этапы развития ИИ, анализируются различные аспекты исследований ИИ и интеллектуальных технологий, а также различные точки зрения на них в смежных областях и в междисциплинарном контексте. Проводится анализ содержания базовых научных и прикладных проблем получения, представления и использования знаний, а также конкретных прикладных задач ИИ, которые, по мнению авторов работы, должны составить основное содержание дорожной карты исследований и прикладных разработок в рамках упомянутой Национальной стратегии развития технологий ИИ.

Ключевые слова: искусственный интеллект, информатика, знания, получение знаний, представление знаний, использование знаний, принятие решений

Введение

В настоящее время научное направление в информатике, называемое метафорически "искусственный интеллект" (ИИ), а также информационные технологии (ИТ), построенные на его основе, приобретают стратегическое значение как теоретический и технологический базис цифровой экономики XXI века. Ведущие экономики мира вкладывают огромные средства в исследования и разработки в области ИИ и интеллектуальных технологий. В ряде стран, например, в США начаты работы по стандартизации технологий ИИ. В России Президентом подписана "Национальная стратегия развития технологий в области искусственного интеллекта" (далее для краткости — НС).

В НС ИИ рассматривается как "комплекс технологических решений, позволяющих имитировать когнитивные функции человека, получая результаты, сопоставимые как минимум с результатами интеллектуальной деятельно-

сти человека" [1]. Такой взгляд на ИИ, по меньшей мере, настораживает, поскольку он явно провозглашает задачу "имитации когнитивных функций человека" как самоцель стратегии, для реализации которой и должен быть создан "комплекс технологических решений".

Однако следует признать, что в настоящее время среди ученых и специалистов в области информатики и ИТ, включая индустриальных разработчиков новейших приложений технологий ИИ, нет единого мнения об объекте исследований ИИ, о целях и существовании этих исследований и разработок, о научном содержании и специфической методологической базе ИИ, что ведет к неоднозначному мнению и о существовании технологий ИИ.

Среди ученых и специалистов в этой области нередко встречаются мнения об ИИ как об альтернативе естественному интеллекту, реализованной в вычислительной среде. Заметим, что точка зрения НС по существу провозглашает практически ту же точку зрения на

уровне официального мнения. Другие представители научного сообщества настаивают на "ведущей роли компьютерных наук в развитии искусственного интеллекта" и утверждают, что "Основная проблема разработки "комплекса технологических решений", о котором идет речь в НС, прежде всего, в архитектуре и "программном обеспечении" самих "интеллектуальных" вычислителей" [2]. Существуют и другие трактовки существа ИИ, которые в большинстве своем рассматривают содержание и цели ИИ либо однобоко, либо вне контекста ИТ. К сожалению, в дискуссиях на эту тему нередко основные усилия расходятся на философскую трактовку ИИ, и значительно реже затрагиваются детали его научного содержания. Кроме того, этот плюрализм мнений сильно усложняется традиционным невежеством журналистов, создающих информационный шум, заглушающий мнения специалистов и активно продвигаемый в органы исполнительной власти. В результате такого разнообразия в понимании сути ИИ как науки, эмоциональных и часто искаженных оценок появляются весьма однобокие трактовки целей и задач НС. Следствием этого может быть провал всех передовых *технологических* намерений НС.

Представляется, что ключевое содержание НС представлено термином "комплекс технологических решений", и совершенно очевидно, что речь идет о таких решениях в области интеллектуальных ИТ. В этом состоит прагматическая цель НС, которая в своей преамбуле явно ассоциируется с такими программами, как "Стратегия развития информационного общества РФ на 2017–2030 годы", национальная программа "Цифровая экономика РФ", дорожные карты "Национальной технологической инициативы", и с другими программами, "эффективность реализации которых может быть повышена за счет использования технологий ИИ" [1, стр. 2].

В данной работе научное направление с названием "искусственный интеллект" трактуется с прагматических позиций как один из разделов информатики, в котором принятие решений основывается на *знаниях*, а его научное содержание определяется методологиями, методами и алгоритмами *получения*, компьютерного *представления* и *использования знаний* в прикладных системах принятия решений. В свою очередь, интеллектуальные ИТ рассматриваются как "комплекс технологических решений", о котором идет речь в НС.

ИИ как раздел информатики не ставит цели воспроизведения естественного интеллекта с какой-то мерой аналогии и не претендует на создание альтернативы естественному интеллекту. НС тоже не претендует на создание "сильного ИИ", полностью реализующего возможности естественного интеллекта. В ближайшее время, которое охватывает временной горизонт реализации НС, это будет абсолютно невозможно, и в этом авторы данной работы солидарны с мнением авторов работы [2]. В контексте современного понимания ИИ его связь с естественным интеллектом видится в том, что ИИ включает в свой методологический базис методы принятия решений на основе знаний, которые в чем-то аналогичны или схожи с методами, применяемыми человеком, как они понимаются сейчас.

Однако основу методов принятия решений в системах ИИ, включая и методы, рассматриваемые как аналоги процессов принятия решений человеком, составляют все-таки знания, а не конкретные и тем более обязательно человеческие механизмы. В рамках ИТ термин "искусственный интеллект" используется в качестве метафоры, удобной для позиционирования методов, алгоритмов, инструментов и технологий ИИ в контексте более общей науки, называемой "Информатика" и технологий, называемых "Информационные технологии". Такой прагматический взгляд на ИИ, очищенный от философских аспектов, отделенный от проблем физиологии мозга и защищенный от информационного шума "специалистов", необходим для того, чтобы рассматривать НС с прагматических позиций, диктующих адекватную дорожную карту исследований и разработок в интересах создания целевого "комплекса технологических решений" на основе ИИ и их последующего использования для эффективного решения задач цифровой экономики.

Для достижения поставленных целей в дальнейшем материале работы сначала кратко характеризуются основные этапы развития ИИ и интеллектуальных ИТ, анализируются различные проблемы ИИ и разные точки зрения на их реализацию в виде интеллектуальных технологий. Основное содержание работы касается характеристики ИИ и интеллектуальных ИТ как неотъемлемых компонентов информатики. В этом контексте приводится анализ базовых проблем ИИ, определяющих современные технологические тренды, решение

которых, по мнению авторов, должно определять существо *дорожной карты* исследований и прикладных разработок, и которые позволяют создать целевой "комплекс технологических решений" как результат выполнения НС.

Немного истории

Строго очертить какое бы то ни было научное направление невозможно в принципе, и доказательством тому служат многочисленные безуспешные попытки представить иерархическую классификацию науки, встроить в нее новые междисциплинарные проблемы и новые научные направления, которые появляются регулярно. Еще сложнее было бы описать как-то строго научное направление "Искусственный интеллект", в название которого входит такой совсем уже неопределенный кросс-доменный термин, как "интеллект", используемый в ряде наук и в очень разнообразном смысле.

Важно отметить, что в период зарождения ИИ как раздела кибернетики его описание не было связано напрямую с интеллектом человека. Согласно Дж. Маккарти ИИ-исследователи вольны "использовать методы, которые не наблюдаются у людей, если это необходимо для решения конкретных проблем" [3]. В этой цитате самыми важными являются слова — "для решения конкретных проблем". Это подчеркивает тот факт, что сам термин "искусственный интеллект" для основателей ИИ как научного направления не являлся важным, важнее было содержание, которое они в него вкладывали.

В процессе более чем 60-летних исследований и разработок в области ИИ изменялось содержание его задач и существо его методов, инструменты для исследований и приоритетность разделов ИИ. Соответственно изменялось содержание и множилось число трактовок термина "искусственный интеллект". Эта динамика имеет место сейчас, и она сохранится в будущем.

Однако большинство "противоречий" в трактовке понятия "искусственный интеллект" обусловлено различием базового смысла, который вкладывается авторами трактовок в этот термин. Если говорить коротко, то эти "противоречия" обусловлены или традиционной омонимией — множеством различных смыслов термина, и тогда противоречия естественны, или они обусловлены кросс-доменной сущностью термина "искусственный интеллект".

Автор термина "искусственный интеллект" Дж. Маккарти описал его как интеллект "вообще" — это "вычислительная" составляющая того, что помогает субъекту достигать заданных целей. Это значит, что для создания или развития интеллектуальных информационных технологий и систем ИИ человеку, с точки зрения Дж. Маккарти, и не надо разбираться в том, как именно работает его мозг.

Описанное выше понимание было дано "отцами" ИИ, которые были специалистами в области компьютерных наук и ввели этот термин как метафорическое название определенного класса методологий, методов и алгоритмов вычислений, использование которых позволяло им решать новые, более сложные классы задач и программировать новые классы приложений. Эта точка зрения, по существу, отвечает взгляду на ИИ и его технологии как на раздел информатики. В настоящее время эта точка зрения разделяется многими специалистами в области ИИ, которые занимаются исследованиями и разработками интеллектуальных технологий и приложений [4].

Важно отметить, что исследователи в области нейробиологии, психологии и философии сознания использовали термин "интеллект" задолго до появления вычислительного взгляда на этот термин в контексте информатики и ИИ. Естественно, что этот термин и сейчас используется ими в "своем" смысле, а именно как "человеческий интеллект", т. е. как название для того, что определяет мотивацию, поведение, принятие решений, эмоции человека и многое другое, свойственное ему. При этом они рассматривают мозг как физический носитель структур и процессов, ответственных за проявление интеллекта. Поэтому появление научного направления ИИ, эксплуатирующего слово "интеллект", специалистами названных наук стало ассоциироваться с механизмами работы мозга человека. И на начальном этапе становления и развития ИИ это не вызвало каких-либо возражений со стороны сторонников вычислительного взгляда на ИИ, поскольку они увидели в этих ассоциациях потенциальный источник дополнительных моделей, методов и алгоритмов для расширения практических возможностей уже провозглашенной ими вычислительной парадигмы ИИ. По сути, это "молчаливое согласие" породило кросс-доменное научное направление на стыке нейробиологии и ИТ.

Это казалось естественным, поскольку практически в это же время была предложена формальная модель персептрона, которая ввела в обиход концепцию нейросети. И она оказалась привлекательной для специалистов в области ИИ не только как некоторая новая модель обучения распознаванию и принятию решений, но и как шаг в направлении обогащения ИИ формальными (алгоритмическими) моделями, имитирующими, хоть и совсем отдаленно, механизмы обучения и принятия решений человеком и, как тогда казалось, другие функции его головного мозга. В последующем эта концепция развивалась специалистами ИИ в рамках его направления, названного *коннекционизмом* (англ. *connectionism*), в котором формальная модель под названием "нейросеть", заимствованная у нейробиологов, стала основной.

Успехи этих наук, а также новые возможности электроники, ИТ и коннекционистских (нейросетевых) моделей ИИ привели в начале 2000-х годов к появлению на стыке нейробиологии и ИИ кросс-доменного научного направления под названием "Информатика мозга" (англ. *Brain informatics*) [5]. С тех пор в этом направлении плодотворно и активно сотрудничают специалисты обоих направлений. Заметным практическим результатом этих работ к настоящему времени является создание нейроинтерфейсов, которые позволяют использовать сигналы, структуры и механизмы управления головного мозга для управления физическими устройствами, главным образом протезами.

Новое звучание проблемы вычислительного ИИ получили с середины 2000-х годов в связи с достижениями в области семантических вычислений и технологий, которые, можно сказать, "очеловечили" вычислительный ИИ [6, 7]. Заметим, что вычисления называют *семантическими*, если они выполняются с контентом, семантика которого четко определена в терминах естественного языка (ЕЯ), и результат также получается с четко заданной семантикой. В частности, эти вычисления решают задачи вычисления смысла контента, семантической интеграции гетерогенных знаний и данных. Они реализуют семантические интерфейсы, пользовательские и машиночитаемые, предоставляют ЕЯ-средства для взаимодействия роботов и человека и т. п.

Появление модели *глубокого обучения* как нового этапа развития нейросетевых технологий ИИ примерно к 2010 году привело к новой волне достижений на основе нейросетевых

моделей ИИ [8]. В частности, это коснулось широкого сегмента трудных задач обработки изображений, обработки речи и текстов на ЕЯ. К сожалению, это вызвало не вполне адекватную эйфорию среди апологетов нейросетевых технологий. Некоторые из них уверены в том, что все проблемы ИИ решаемы с помощью нейросетевых моделей глубокого обучения, и они активно продвигают это мнение "в массы".

Обобщая приведенный краткий и фрагментарный исторический обзор результатов и тенденций в области ИИ, его компоненты можно проиллюстрировать графически (рис. 1, см. вторую сторону обложки). Современные модели и технологии ИИ можно разделить на три части. Одна из них, которая выше по тексту была названа коннекционизмом, представляет ту часть моделей и технологий ИИ, которую чаще называют нейросетями. В настоящее время ее основными предметами исследований являются модели структур и методы обучения глубоких нейросетей. Этот новый вид нейросетей отличается от своих предшественников не только числом слоев, но и более сложной технологической архитектурой, которая использует многослойную систему нелинейных фильтров для предобработки данных обучения, для поиска признаков в разных слоях и для их преобразования в иерархическую структуру с признаками различного уровня абстракции. Модель глубокого обучения нейросети способна использовать одновременно оба механизма обучения — с учителем и без учителя. Заметим, что успеху этой модели ИИ во многом способствовали успехи в создании высокоскоростных компьютеров.

К этим исследованиям примыкают (справа на рис. 1, см. вторую сторону обложки) исследования в области информатики мозга, нейробиологии и другие, которые иногда называют когнитивными науками [9]. К ним часто относят также и нейросетевые модели, однако с этим можно согласиться с большой натяжкой, поскольку кроме общего слова "нейросеть", используемого в них в разных смыслах, их мало что объединяет. По крайней мере, модель сети глубокого обучения появилась не в результате каких-то достижений в области изучения мозга человека (появление этой модели относят к 1980 г.), а в результате совершенствования нейросетевых механизмов обработки данных в области ИИ. Реально нейросети глубокого обучения имеют мало общего с реальными нейросетями головного мозга. Однако поскольку

концепция нейросети мотивировалась изначально как попытка моделировать эти механизмы, то отнесение модели глубокого обучения к когнитивным наукам не вызывает особого возражения ни в ИИ, ни в среде специалистов когнитивных наук.

К полярной в некотором смысле области ИИ как научного направления в информатике следует отнести вычислительный интеллект, который очень далек от когнитивных моделей и который акцентирует внимание на эффективной реализации вычислительных моделей ИИ с помощью программно-аппаратных реализаций и суперкомпьютеров. Действительно, использование высокоскоростных компьютеров и программно-аппаратных реализаций ИИ предоставляет новые технологические возможности по сравнению с вариантами программной реализации моделей ИИ на обычных компьютерах. Хорошо известно, что победы, одержанные шахматными программами и программами игры в Го над профессионалами, были бы невозможны без использования суперкомпьютеров, и эти достижения принято относить к достижениям ИИ. Следует также отметить перспективу суперкомпьютерных технологий в обработке больших объемов сырых данных, поскольку их последующая обработка для извлечения знаний и машинного обучения предъявляет уже значительно более слабые требования к мощности компьютера.

Однако основная область задач, моделей методов, алгоритмов, инструментов и технологий современного ИИ находится между двумя названными выше компонентами ИИ (рис. 1, см. вторую сторону обложки). Системы ИИ, которые относятся к этой области и которые охватывают большую часть того, что принято называть методами ИИ, часто называют также системами, основанными на знаниях. В среде специалистов уже относительно давно оба эти названия отождествляют, а сам термин "знания" чаще всего используется специалистами ИИ с момента его зарождения и по настоящее время. По-видимому, именно знания являются ключевым понятием ИИ и нуждается в отдельном обсуждении.

Знания, семантика знаний, семантические технологии

Знания как объект исследований, разработок и использования всегда присутствуют в при-

кладных системах и технологиях ИИ, в отличие от систем и технологий других типов. При этом *знания* составляют ядро любой интеллектуальной системы, и именно знания (их объем, корректность, полнота и др.) и используемые методы работы с ними определяют потенциальные возможности конкретной интеллектуальной системы. Вместе с тем термин "знания" является ключевым словом и при описании интеллекта человека. В работе [10] один из ее авторов, Д. А. Поспелов, наиболее авторитетный специалист ИИ в СССР и России, в разделе "Становление информатики в России" рассматривает ИИ как раздел информатики и характеризует его как "представление знаний, вывод на знаниях, обучение, экспертные системы и т. п." (стр. 8), и ИИ — это "технологии решения задач, опирающиеся на идею использования знаний о предметной области" (стр. 37), по сути, однозначно выделяя предмет исследований ИИ и его базовые задачи.

Это же, хотя и косвенно, но однозначно утверждается и в работе [11]: "...при определении искусственного интеллекта следует отталкиваться от определения естественного интеллекта, то есть "способности мозга человека решать интеллектуальные задачи путем приобретения, запоминания и целенаправленного преобразования знаний и использования этих знаний для управления средой". Заметим, что в этих высказываниях не говорится о том, что ИИ должен в чем-то повторять естественный интеллект. В них, как и в приведенном высказывании Дж. Маккарти, акцентируется внимание лишь на *сходстве задач*, решаемых человеком и программой ИИ, но не утверждается обязательное сходство методов их решения. И именно в этом смысле термин "искусственный интеллект" является метафорой, которая показалась создателям этого научного направления весьма удобной в качестве названия, но не более того, и эта метафора сразу была принята в сообществе специалистов.

Ключевая роль знаний в информационных технологиях ИИ состоит в том, что именно благодаря их использованию стало возможным эффективное решение многих вычислительных задач, которые являются задачами экспоненциальной сложности и для которых любое кратное увеличение производительности вычислительных машин не в состоянии гарантировать получение оптимального решения за разумное время. Заметим, что иногда

можно встретить утверждения о том, что большая часть достижений в области ИИ напрямую связана с ростом производительности компьютеров. Эту связь отрицать невозможно, но далеко не все проблемы ИИ могут быть успешно и эффективно решены только за счет этого фактора, а использование знаний совместно с ростом производительности компьютеров позволяет этого достичь. Хорошо известно, что человек успешно решает многие вычислительно сложные задачи с помощью когнитивных эвристик. Механизм принятия решений на основе правил является одним из самых распространенных механизмов решения задач ИИ. И этот механизм ИИ является аналогом когнитивных эвристик человека. Многие вычислительно сложные переборные задачи успешно решаются с использованием знаний и принципа ограниченной рациональности, который ориентирует исследователя на поиск "хороших" решений вместо оптимальных, но за разумное или заданное время.

Часто возникает вопрос о том, а почему по отношению к одному типу информации мы используем термин "данные", а по отношению к другому — "знания", и почему именно знания, а не данные, позволяют решать сложные задачи, которые метафорически называют интеллектуальными. Конечно, однозначную границу между этими типами информации провести вряд ли можно, однако семантические их модели, которые твердо вошли в обиход исследователей и разработчиков ИИ в последние годы, позволяют эту границу наметить. Появление семантических моделей знаний и данных — это настолько важный шаг в развитии ИИ, что их практическое использование следует трактовать как новый этап его развития, и об этом следует сказать особо.

Семантические вычисления — это область ИИ, которая имеет целью исследования и разработки компьютерных технологий для решения задач поиска, создания, формального представления, манипулирования и использования контента и доступных компьютерных ресурсов, а именно различных данных, тестовых документов, изображений, видео- и аудиоинформации, инструментальных средств, людей, устройств, агентов и т.д. в терминах ЕЯ [7, 12—16]. Основные задачи семантических вычислений в ИИ включают в себя:

— *семантический анализ* — вычисление смысла контента, представленного текстом, изо-

бражением, аудио- и видеоинформацией, анализ семантики контента социальных сетей, работу с контентом в семантическом стиле и др.;

— *семантическая интеграция знаний* и данных, например, для создания единого семантического информационного пространства на множестве распределенных баз данных, знаний и онтологий;

— *семантические интерфейсы* (пользовательские и машиночитаемые, средства взаимодействия роботов и человека);

— *семантические приложения* (поисковые машины, семантические веб-сервисы, базы знаний текстов, изображений, видео-, аудиоинформации и др.).

Работа со знаниями в терминах ЕЯ является фокусом семантических технологий. Основные решения в области семантических вычислений и семантических технологий строятся с помощью онтологических моделей знаний и данных. Теоретические основы онтологий, языковые средства их формального описания, а также семантические ресурсы и инструменты для их практического использования развиты в настоящее время достаточно хорошо [14—16], и они продолжают активно развиваться.

Как видно из материалов данного раздела, знания — это ключевое понятие ИИ, и работа со знаниями является отличительной чертой системы ИИ по отношению к компьютерным системам других классов. Это позволяет утверждать, что именно *знания* являются основным *объектом исследований* и разработок в области ИИ. Более того, анализ научной литературы в области ИИ показывает, что основные исследования и разработки в ИИ концентрируются вокруг трех проблем, и в фокусе каждой из них находятся различные аспекты технологии работы со знаниями, а именно:

- получение знаний,
- представление и преобразование знаний и
- использование знаний.

Рассмотрим кратко научное содержание этих проблем и покажем, что они охватывают практически все задачи, которые принято относить к ИИ.

Получение знаний

Для решения задач *получения знаний* привлекаются самые разнообразные источники данных и знаний. Для этих целей разработа-

ны разнообразные методы, алгоритмы и программные инструменты, известные из математики и математической статистики. Исследователями в области ИИ созданы свои специфические методы, алгоритмы и инструменты для решения задач извлечения знаний из различных источников. Это направление в ИИ обычно называют "инженерией знаний". К ним, например, относятся специфические методы и инструменты получения знаний напрямую от экспертов, а также из данных с помощью методов машинного обучения (англ. *Machine Learning*) или методов обнаружения зависимостей в данных (англ. *Data Mining*). Рисунок 2 (см. вторую сторону обложки) иллюстрирует основные источники данных и знаний, источники знаний о семантике понятий и текстов, а также типы инструментов, которые к настоящему времени доступны специалистам для решения задач инженерии знаний.

Важно отметить, что совместное использование методов, известных из математики, и методов инженерии знаний обычно настолько естественно и привычно, что многие исследователи, а особенно разработчики программного обеспечения автоматически причисляют их к методам машинного обучения ИИ. Яркими примерами такого отождествления "по привычке" являются многочисленные методы приближенной оптимизации на основе случайного поиска, например, генетические алгоритмы, алгоритмы случайного поиска на основе мета-эвристик, заимствованных у природы (муравьиные алгоритмы, пчелиные и другие алгоритмы из этой серии) и др. Строго говоря, сюда же следовало бы отнести и методы машинного обучения с подкреплением, а также методы оптимизации на основе случайного поиска, которые активно развивались вне задач ИИ и иногда задолго до инженерии знаний. Об этом часто забывают и студентам читают их в курсах инженерии знаний ИИ.

Однако некоторые варианты случайного поиска были созданы именно для машинного обучения по примерам, и их следует относить к ИИ. Заметим, что методы обучения нейросетей, например, методы глубокого обучения, являются методами случайного поиска с подкреплением, но именно для решения задач ИИ, и потому их правомерно относить к инженерии знаний.

Результатами задач инженерии знаний являются паттерны знаний либо с семантиче-

ской интерпретацией (например, в виде правил, предметные переменные которых являются понятиями онтологии), либо в чисто синтаксической форме, в которой такая интерпретация явно отсутствует. Последний вариант характерен для результатов обучения нейросетей. Заметим, что онтология сама по себе — это одна из форм знаний с явным представлением семантики, поэтому построение онтологий относится к получению знаний.

Отметим, что выше шла речь об описании знаний на концептуальном уровне. Их называют также концептуальными моделями знаний. Для их описания обычно используют различные средства естественных и/или формальных языков, куда относят также формулы, таблицы, графы, тексты и прочее. На этом уровне описания знания еще не готовы к использованию компьютерными программами, поскольку машинный язык сильно отличается от языков описания концептуальных моделей знаний. Исключение составляют разве что модели знаний, представленные в форме нейросети. Они не являются, строго говоря, моделями знаний концептуального уровня, поскольку уже представляются виде вычислительных структур в битовой форме.

Преобразование знаний и данных к формальной модели, ее трансляция на машинный язык — это задача следующего этапа работы со знаниями.

Представление и преобразование знаний и данных

Как уже отмечалось, концептуальная модель знаний не может быть использована компьютером напрямую, поскольку машинный язык — это язык бинарных последовательностей. Представление знаний концептуального уровня в виде структур, которые могут быть далее формально преобразованы к машиночитаемой форме, — это одна из задач второй группы типовых проблем ИИ. Выше она была названа проблемой *представления знаний*, причем под представлением знаний понимается их спецификация в терминах конкретных структур, описываемых на некотором формальном языке. Построенные структуры знаний и данных могут далее использоваться в качестве входа для программ, выполняющих их трансляцию в форму, понятную компьютеру. Несколькими упрощая, можно сказать, что задача

данного этапа — это преобразование знаний из формального представления, понятного человеку, в форму, понятную компьютеру. На этом этапе знания и данные, описанные, например, текстами на ЕЯ или представленные в другой форме, понятной человеку, преобразуются в машиночитаемую форму.

Ранее отмечалось, что в настоящее время в качестве модели описания знаний концептуального уровня обычно рассматривается модель онтологии. В ней понятиям *онтологии* соответствуют реальные или виртуальные объекты (классы) таксономии предметной области разных *уровней обобщения*. Каждый класс понятий (ему отвечает множество классов объектов реального или виртуального миров) описывается конкретной *структурой*, множеством *атрибутов*, а также непустым *множеством примеров* (в приложениях).

Такое представление знаний и данных принято называть семантическим представлением. Хорошо известно, что семантика любого термина ЕЯ и в том числе семантика понятия онтологии не может быть описана точно. Для нее могут быть определены только некоторые размытые границы. В *онтологии приложения* семантика понятия определяется множеством его *примеров*. При этом чем больше примеров понятия имеется в онтологии, тем точнее определена его семантика, так как каждый новый пример понятия уточняет (расширяет) его смысл, снижая тем самым его смысловую неопределенность.

Представление знаний в форме онтологии принято еще называть формализацией концептуализации. Действительно, онтология концептуального уровня в модели представления сначала описывается на некотором формальном языке, например, на *RDF*, *OWL* или на каком-то другом языке, но это пока форма, понятная человеку. После трансляции такое представление преобразуется в структуры данных и знаний, читаемые компьютером.

Человек предпочитает вводить знания и данные в компьютер в форме, понятной ему. В той же форме он предпочитает получать промежуточные и конечные результаты решения задачи. Но компьютер "понимает" только машиночитаемую форму данных и знаний. Поэтому на этапе представления знаний рассматриваются обе формы знаний и данных и их преобразование в обе стороны. Рисунок 3 (см. третью сторону обложки) иллюстрирует один из случаев использования, когда выполняются

преобразования знаний и данных из формы, понятной человеку, в форму, понятную компьютеру, и наоборот. Во всех этих преобразованиях используется модель представления знаний в онтологической форме.

Использование знаний

Обратимся к задаче использования знаний, т. е. к *технологиям* работы со знаниями, принятым в различных прикладных интеллектуальных системах. *Использование знаний* в ИИ — это применение конкретных технологий ИИ и инструментальных средств их программной поддержки. Технологии ИИ, например, многоагентные технологии, технологии глубокого обучения и т. д., и инструменты их программной поддержки сильно зависят от класса приложений. Именно на этом этапе дорожная карта НС должна различать классы приложений, такие как компьютерное зрение, многомодальные интерфейсы, автономные роботы и групповое управление в различных случаях, сценариях и миссиях использования, распределенные системы и сети класса интернета вещей и/или киберфизических систем, интеллектуальные сенсорные сети и "умные" облачные сервисы в различных приложениях, системы поддержки принятия решений (в производстве, на транспорте, для обеспечения безопасности в киберпространстве, финансовой сфере, биологии, фармацевтике и т. д.), рекомендательные системы, например, персональные рекомендательные системы в медицине, в управлении активами и т. д.

В качестве одного из классов задач использования знаний следует также рассматривать задачи обмена информацией между компонентами (распределенного) приложения, реализуемые коммуникационными технологиями, в которых содержание сообщений описано на языках представления знаний.

Обсуждение: важные задачи дорожной карты

Представленная факторизация базовых задач ИИ на задачи получения, представления и использования знаний позволяет структурировать работы по созданию "комплекса технологических решений", запланированных в НС.

В настоящему времени теория и технология ИИ достигли высокого уровня зрелости, и они по достоинству оценены индустрией. За

шестьдесят с лишним лет развития наука об управлении на основе знаний уже преодолела этап начальной эйфории, когда специалисты всерьез говорили о возможности быстрого создания ИИ как альтернативы или даже как расширения естественного интеллекта. Сейчас это направление ИИ остается областью профессиональной деятельности нейробиологов, специалистов в области философии сознания и информатики мозга (см. рис. 1 на второй стороне обложки). Он остался также областью профессиональной деятельности специалистов в области когнитивных наук, которые изучают когнитивные процессы головного мозга человека в интересах познания, моделирования и использования получаемых знаний.

ИИ как направление в информатике болезненно и с большой потерей времени преодолел этап всеобщей уверенности во всемогущести формальных логических моделей и логического вывода как универсального средства решения проблем ИИ. Сейчас уже мало тех, кто использует Пролог и механизмы принятия решений, построенные на его основе. В истории развития ИИ определенные достижения связаны с периодом активного развития идей и технологий экспертных систем как замкнутых систем принятия решений, построенных на экспертных знаниях. Однако все это уже отошло в прошлое.

В настоящее время палитра актуальных научных исследований и прикладных разработок в области ИИ достаточно обширна и разнообразна. Однако ее наиболее важные, наиболее продуктивные и наиболее проблемные направления, как и на начальных этапах развития ИИ и интеллектуальных технологий, относятся, прежде всего, к *получению знаний*. Именно качество и полнота знаний, используемых той или иной прикладной интеллектуальной системой, практически однозначно определяют ее возможности. Тематика получения знаний остается актуальной уже более 40 лет. В разные периоды она концентрировалась вокруг разных источников, методологий и алгоритмов извлечения знаний и программных средств их инструментальной поддержки. Сейчас в этой по-прежнему ведущей и наиболее трудной области ИИ преобладают два направления. Одно — это семантическая обработка больших данных для *извлечения знаний* и построения механизмов принятия решений, а второе — это глубокое обучение нейросетевых структур принятия решений на основе тех же данных. Эти два во

многом альтернативных подходы к решению одной и той же проблемы, тем не менее, дополняют друг друга.

Извлечение знаний из больших данных имеет целью построение моделей данных и знаний, в которых семантика представлена на ЕЯ. Методология получения знаний в такой форме использует строгие математические постановки задач, методы и алгоритмы их решения. Эти методы принято называть символическими, хотя понятие "символ" в современных языках представления и использования знаний значительно отличается от аналогичного понятия ИИ, скажем, 30 лет тому назад. Отличается оно, прежде всего, переходом от чисто синтаксической формы символических языков к семантической. Если 30 лет тому назад основные механизмы формального представления знаний концентрировались вокруг языков формальных логических исчислений, то сейчас эти модели ушли в прошлое ввиду слабой выразительности и вычислительной неэффективности. Настоящее время — это время семантики, семантических вычислений, семантических технологий и семантических приложений [7], в которых каждая строка символов (переменная или константа) имеет четко заданную и формально "вычисляемую" семантику в терминах ЕЯ. Семантические вычисления в ИИ — это сейчас один из самых заметных и долговременных трендов в области ИТ, а семантические технологии должны стать составляющей общего "комплекса технологических решений" НС.

В отличие от этого нейросетевые модели и знания, представленные в нейросетях, пока остаются чисто синтаксическими. Глубокое обучение есть управляемый механизм случайного поиска с подкреплением, и не стоит строить иллюзий относительно его всемогущести. Интуитивно этот подход является попыткой воспроизведения механизмов обучения принятия решений, которые реализуются в когнитивной модели мозга человека. Но пока достоверных знаний об этих механизмах нет, а потому не вызывают доверия утверждения о том, что эти механизмы аналогичны механизмам естественного интеллекта. Тем не менее, успехи современных нейросетевых моделей принятия решений очевидны и значимы для практики. Они позволили решить ряд вычислительно трудных задач обучения структур принятия решений на основе размеченных данных.

Но эти успехи, как и возможности современных механизмов глубокого обучения, силь-

но преувеличены теми, кто религиозно верит в них. Главный недостаток технологии глубокого обучения — это катастрофические требования к объему обучающих данных, что свойственно всем механизмам случайного поиска. И хотя эти требования могут быть частично ослаблены за счет механизмов предобработки обучающих данных (например, за счет механизмов нелинейной фильтрации), но они пока все равно остаются высокими.

Высокие требования к объему обучающих данных для глубокого обучения в последнее время вызвали определенный кризис в этой области, хотя большинство специалистов закрывают на это глаза. Эти требования стимулировали появление средств инжиниринга обучающих данных с помощью математического моделирования. Но это всего лишь полумера, ведущая только к временному ослаблению остроты проблемы. По мере усложнения объектов и систем, для работы с которыми используются механизмы глубокого обучения, потребности в объемах размеченных данных будут расти, и проблема обязательно обострится снова. Очевидно, что нейросети головного мозга человека опираются в значительной мере на семантику и на контекст принятия решений, а потому пути для преодоления проблем дефицита размеченных обучающих данных следует искать также в их обогащении семантикой ЕЯ.

В области моделей представления знаний основной проблемой является автоматизация построения онтологии, в частности, онтологии данных. Это направление уже достаточно долго остается одним из ключевых в области ИИ, и в этом направлении уже имеется значительный прогресс по созданию семантических ресурсов и типовых онтологий, которые предоставляет семантический веб и семантические ресурсы проекта *Linked Data Web* [15, 17], о чем ранее уже шла речь. Однако существующие многочисленные инструменты проекта *Linked Data Web* и используемые в них семантические ресурсы пока еще не обладают нужной точностью по семантическому анализу данных при выявлении понятий и их структуризации.

Другая группа проблем представления знаний и семантических вычислений, которая частично может быть отнесена и к множеству проблем использования знаний, связана с решением задач трансформации знаний и данных из человекопонятной формы в форму,

понятную компьютеру, и наоборот (см. пример на рис. 3, см. третью сторону обложки). В настоящее время, к сожалению, отсутствуют эффективные механизмы, пригодные для выполнения таких преобразований знаний и данных с большими объемами баз знаний. Однако имеются новые разработки в области графовых баз данных, реализованных новыми системами управления базами данных, например, *Neo4j* и *OrientDB* с большим потенциалом для задач семантического представления и использования знаний.

Что касается проблем использования знаний в прикладных системах ИИ, то здесь можно констатировать наибольшее число острых и весьма сложных проблем, требующих совместных усилий научного сообщества и бизнеса. В этой области тенденции определяются новыми приложениями, в частности, приложениями класса интернета вещей и киберфизических систем. Будущее этих систем принято связывать с их реализацией в виде экосистем умных сервисов. Центральным компонентом таких реализаций является (распределенная) цифровая платформа, которая играет роль программно-коммуникационной инфраструктуры, трансформирующей множество ее гетерогенных компонентов (распределенных приложений, сенсоров, облачных ресурсов и сервисов, разделяемых ресурсов и сервисов, которые реализуются цифровой платформой и т. д.) в единую "систему систем" [18–20].

Кратко задачи экосистемы умных сервисов, интеллектуального ядра процессов координации и управления, должны состоять в том, чтобы (рис. 4, см. третью сторону обложки):

- "погрузить" все компоненты в единое коммуникационное, информационное и семантическое пространство знаний и данных и обеспечить интерфейсы доступа к ним со стороны программ и пользователей;
- обеспечить совместимость компонентов экосистемы по форматам данных и знаний при решении задач кооперации и координации поведения;
- поддерживать решение задач распределения, планирования и оперативного управления общими ресурсами экосистемы в реальном времени;
- обеспечить узлы сети компонентов экосистемы онлайн-доступом к облачным ресурсам и сервисам и к внешним источникам информации и

- обеспечить накопление и хранение данных, генерируемых приложениями, для машинного обучения в целях повышения качества их работы.

Одна из особенностей современных приложений класса интернета вещей и киберфизических систем состоит в том, что в них обычно объектом управления является групповое поведение множества автономных объектов, и это важный новый класс объектов управления. Управление групповым поведением в этом классе объектов называют *групповым управлением*. В ближайшее время следует ожидать вирусного распространения приложений, в которых требуется решать задачи группового управления. Эффективные решения в этой области могут быть построены на основе сетей агентов и агентских технологий, важного компонента в "комплексе технологических решений" НС.

Заключение

Ясное понимание сути проблем ИИ и интеллектуальных технологий должно способствовать ускоренному внедрению радикально новых ИТ, позволяющих решать научные и технические задачи современной цифровой эры. В работе показана ключевая роль знаний в ИИ и в интеллектуальных ИТ. Показано, что ИИ следует рассматривать как научную дисциплину, объектом изучения в которой являются знания, а методологическую основу составляют модели, алгоритмы и программные инструменты извлечения, представления и использования знаний в приложениях.

В работе проанализировано существо базовых задач ИИ и сделан вывод о том, что дорожная карта развития ИИ и интеллектуальных технологий для создания "комплекса технологических решений" ИИ может быть структурирована в соответствии с этими тремя базовыми проблемами ИИ. В контексте этих задач перечислены проблемы и основные тренды их эволюции, существенные для долговременной стратегии развития ИИ и интеллектуальных технологий в различных приложениях, в том числе в цифровой экономике.

Список литературы

1. Указ Президента РФ "О развитии искусственного интеллекта в Российской Федерации" от 10 октября 2019 г.

№ 490. URL: <http://www.kremlin.ru/acts/news/61785>. (дата обращения 24.01.2020)

2. Каляев И., Заборовский В. Искусственный интеллект. От метафоры к техническим решениям // Control Engineering. Россия. Октябрь 2019. С. 78—83.

3. Искусственный интеллект — Википедия. URL: https://ru.wikipedia.org/wiki/Искусственный_интеллект (дата обращения 24.01.2020).

4. Юсупов Р. М. К семидесятилетию отечественной информатики // Труды СПИИРАН. 2019. Т. 18, № 1. С. 5—29.

5. Zhong N., Jeffrey M., Bradshaw J. M., Liu J., Taylor J. G. Brain Informatics // IEEE Intelligent Systems, September/October. P. 16—20.

6. Sheu P., Yu H., Ramamoorthy C. V., Joshi A., Zadeh L. A. (eds.) Semantic Computing // IEEE. Wiley, 2010. 552 p.

7. Hoppe T., Humm B., Reibold A. (eds.) Semantic Applications. Berlin, Heidelberg, Springer Vieweg, 2018.

8. Goodfellow I., Bengio Y., Courville A. Deep Learning. An MIT Press book. URL: <http://www.deeplearningbook.org> (дата обращения 24.01.2020).

9. Станкевич Л. А. Когнитивные системы и роботы. СПб: Политех-Пресс, 2019. 630 с.

10. Очерки истории информатики в России / Ред.-сост. Д. А. Поспелов, Я. И. Фет. Научно-издательский центр ОИГМ СО РАН, 1998.

11. Каляев И. А. Гонка за цифровым призраком. Есть ли у России шанс поучаствовать в борьбе за первенство в разработке ИИ // Огонек. № 24 (5569). URL: <https://www.kommersant.ru/doc/4003879> (дата обращения 24.01.2020).

12. Bartussek W., Bense H., Hoppe T. et al. Introduction to Semantic Applications // Semantic Applications. Methodology, Technology, Corporate Use. Springer Nature, 2018. P. 1—13.

13. Bizer C., Lehmann J., Kobilarov G. et al. DBpedia — A crystallization point for the Web of Data // Web Semantics: Science, Services and Agents on WWW. 2009. Vol. 7. P. 154—165.

14. Flisar J., Podgorelec V. Improving Short Text Classification using Information from DBpedia Ontology // Fundamenta Informaticae 2018. Vol. XXI. IOS Press. P. 1—35.

15. Heath T., Bizer C. Linked Data: Evolving the Web Into a Global Data Space. Morgan & Claypool Publishers, 2011.

16. Городецкий В. И., Тушканова О. Н. Семантические технологии для семантических вычислений. 1. Базовые компоненты // Искусственный интеллект и принятие решений. 2018. № 4. С. 61—71.

17. Linked Data. URL: <http://linkeddata.org> (дата обращения 24.01.2020).

18. Городецкий В. И., Ларюхин В. Б., Скобелев П. О. Концептуальная модель цифровой платформы для киберфизического управления предприятием. Часть 1. Цифровая платформа и цифровая экосистема // Мехатроника, автоматизация, управление. 2019. Т. 20, № 6. С. 323—332.

19. Городецкий В. И., Ларюхин В. Б., Скобелев П. О. Концептуальная модель цифровой платформы для киберфизического управления современным предприятием. Часть 2. Цифровые сервисы // Мехатроника, автоматизация, управление. 2019. Т. 20, № 7. С. 387—397.

20. Gorodetsky V., Larukhin V., Skobelev P. Conceptual Model of Digital Platform for Enterprises of Industry 5.0 // In Proceedings of the International Conference on Intelligent Distributed Computing. Lecture Notes in Computer Science. St. Petersburg, October 6—7, 2019. Springer, 2019.

Artificial Intelligence: Metaphor, Science and Information Technology

V. I. Gorodetsky, vladim.gorodetsky@gmail.com,

St. Petersburg State Electro-technical University (LETI) St. Petersburg, 197376, Russian Federation,

R. M. Yusupov, yusupov@iias.spb.su,

St. Petersburg Institute for Informatics and Automation of RAS, St. Petersburg, 199178, Russian Federation

Corresponding author: **Gorodetsky Vladimir I.**, Ph.D. St. Petersburg State Electro-technical University (LETI), 197376, St. Petersburg, Russian Federation, e-mail: vladim.gorodetsky@gmail.com

Accepted on January 27, 2020

Abstract

To the present days, the scientific direction of computer science metaphorically called "Artificial Intelligence" (AI) as well as corresponding information technologies is becoming of strategic value as the theoretical and the technological basis of digital economy of the 21-th century. The leading economies of the world invest huge money in R&D of AI area. In many countries, e.g. in the USA and Europe the standardization effort in AI scope are initiated. In Russia, the President issued a decree "On development of artificial Intelligence in the Russian Federation" dated 10 October 2019. One must acknowledge, however, that up to the current days there is no common opinion, among AI scientists and engineers, on what is the research subject of AI, which are its R&D goals, scientific contents, methodology and technologies. In the paper, Artificial Intelligence is considered as a scientific discipline that is a section of Computer Science. In it, the knowledge acts as the main research subject and its scientific contents is composed of methodologies, methods and algorithms of the three AI subsections those are knowledge acquisition, knowledge representation and knowledge usage for decision making. This point of view is justified in the paper by analysis of AI historical issues, composition of its scientific topics and technologies. The aforementioned analysis is underway with regard to the basic R&D problems that should be in the focus of the AI roadmap of the Russian scientific and industry communities in context of the Presidential decree on the strategy of the development of Artificial Intelligence in Russian Federation.

Keywords: Artificial Intelligence, Computer Science, knowledge, knowledge acquisition, knowledge representation, knowledge usage, decision making

For citation:

Gorodetsky V. I., Yusupov R. M. Artificial Intelligence: Metaphor, Science and Information Technology, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2020, vol. 21, no. 5, pp. 282–293.

DOI: 10.17587/mau.21.282-293

References

1. **Presidential** Decree of the Russian Federation No 490 dated 10 October 2019 "On development of artificial Intelligence in the Russian Federation", available at: <http://www.kremlin.ru/acts/news/61785> (accessed on 24.01.2020) (in Russian).
2. **Kaliyev I., Zaborovsky V.** Artificial Intelligence: From metaphor to technical solutions, *Control Engineering*, Russia, October 2019, pp. 78–83 (in Russian).
3. **Artificial Intelligence** — Wikipedia, available at: https://ru.wikipedia.org/wiki/Artificial_Intelligence (accessed on 24.01.2020) (in Russian).
4. **Yusupov R. M.** Towards 70-th anniversary of computer science, *Transactions of SPIIRAS*, 2019, vol. 18, no. 1, pp. 5–29 (in Russian).
5. **Zhong N., Jeffrey M., Bradshaw J. M., Liu J., Taylor J. G.** Brain Informatics, *IEEE Intelligent Systems*, September/October, pp. 16–20.
6. **Sheu P., Yu H., Ramamoorthy C. V., Joshi A., Zadeh L. A.** (eds.) Semantic Computing, *IEEE/Wiley*, 2010, 552 p.
7. **Hoppe T., Humm B., Reibold A.** (eds) Semantic Applications. Springer Vieweg, Berlin, Heidelberg, 2018.
8. **Goodfellow I., Bengio Y., Courville A.** Deep Learning, MIT Press, available at: <http://www.deeplearningbook.org> (accessed on 24.01.2020).
9. **Stankevich L. A.** Cognitive systems and robots, St. Petersburg, Polytechnic Press, 2019, 630 p. (in Russian).
10. **Pospelov D. A., Fet Ya. I.** Essays of informatics history in Russia, Scientific Publishing Center OIGGV SO RAN, Novosibirsk, 1998 (in Russian).

11. **Kaliyev I. A.** Race for digital ghost. Does Russia have a chance to participate in competition to win in Artificial Intelligence?, *Ogoniok*, no. 24 (5569), 24.06.2019, available at: <https://www.kommersant.ru/doc/4003879> (accessed on 24.01.2020) (in Russian).

12. **Bartussek W., Bense H., Hoppe T. et al.** Introduction to Semantic Applications, In T. Hoppe, B. Humm, A. Reibold (Eds.). Semantic Applications. Methodology, Technology, Corporate Use, Springer Nature, 2018, pp. 1–13.

13. **Bizer C., Lehmann J., Kobilarov G., Auer S., Becker C., Cyganiak R., Hellmann S.** DBpedia — A crystallization point for the Web of Data, *Web Semantics: Science, Services and Agents on WWW*, 2009, 7, Elsevier, pp. 154–165.

14. **Flisar J., Podgorelec V.** Improving Short Text Classification using Information from DBpedia Ontology, *Fundamenta Informaticae* 2018, XXI, IOS Press, pp. 1–35.

15. **Heath T., Bizer C.** Linked Data: Evolving the Web Into a Global Data Space, Morgan & Claypool Publishers, 2011.

16. **Gorodetsky V. I., Tushkanova O. N.** Semantic technologies for semantic computing. Part 1. The basic components of semantic technologies, *Artificial Intelligence and Decision making*, 2018, no. 4, pp. 61–71 (in Russian).

17. **Linked Data**, available at: <http://linkeddata.org> (дата обращения 24.01.2020).

18. **Gorodetsky V. I., Lariukhin V. B., Skobelev P. O.** Conceptual model of a digital platform for cyber-physical management of a modern enterprises/ Part 1. Digital platform and digital ecosystem, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2019, vol. 20, no. 6, pp. 323–332 (in Russian).

19. **Gorodetsky V. I., Lariukhin V. B., Skobelev P. O.** Conceptual model of a digital platform for cyber-physical management of enterprise. Part 2. Digital services, *Mekhatronika, Avtomatizatsiya, Upravlenie*, vol. 20, 2019, no. 7, pp. 387–397 (in Russian).

20. **Gorodetsky V., Larukhin V., Skobelev P.** Conceptual Model of Digital Platform for Enterprises of Industry 5.0, *Proceedings of the International Conference on Intelligent Distributed Computing, Lecture Notes in Computer Science*, Springer, 2019, St. Petersburg, October 6–7, 2019.