

ОТКРЫТАЯ КАФЕДРА

Об интеллектуальных методах обработки информации

Ю. И. ЖУРАВЛЕВ

(ОТДЕЛЕНИЕ МАТЕМАТИЧЕСКИХ НАУК РАН, МОСКОВСКИЙ ГУМАНИТАРНЫЙ УНИВЕРСИТЕТ)

В марте 2014 г. в Институте фундаментальных и прикладных исследований МосГУ прошел методологический семинар, на котором доклад на тему «Об интеллектуальных методах обработки информации» сделал академик РАН, директор Центра математического и компьютерного моделирования ИФПИ МосГУ Ю. И. Журавлев. Среди участников семинара были сотрудники ИФПИ, студенты университета. Выдающийся российский математик рассказал об истоках идеи применения специальных методов интеллектуальной обработки данных в ситуациях решения слабоформализуемых задач, трудностях и перспективах создания искусственного интеллекта. Он привел многочисленные примеры, в том числе из собственной научной работы, когда удавалось построить алгоритм решения таких задач. Доклад был выстроен с учетом интересов слушателей-гуманитариев, которым академик рассказывал о математических проблемах очень доступно, подчеркнув, что каждый ученый должен уметь рассказывать о своей работе так, чтобы его могли понимать все.

Ключевые слова: информация, интеллектуальная обработка данных, искусственный интеллект.

К интеллектуальной обработке данных можно подойти с разных сторон. Можно делать уклон в одну сторону, можно в другую. Я постарался выбрать золотую середину, которая подойдет и технарям, и гуманитариям. Конечно, в большей степени ориентировался на гуманитариев.

«Интеллектуальная обработка данных» — это один из тех терминов, который широко употребляется, но никто до конца не знает, что это такое. Это один из тех терминов, который вошел в мир, так что каждый более или менее имеет представление о том, что он значит. Но вы нигде не найдете определения термина «интеллектуальная обработка данных». Будет описание. Как правило, описание будет связано с тем, чем человек занимается, он исходит из своей области, которая ему наиболее близка.

Придумали термин в свое время американцы. Это у них распространенная вещь: когда появляется какое-то направление в том, что делали раньше, изобретается новый термин. Потом все подгоняется под этот термин. Так в свое время появилось «программирование».

Попробую дать определение, вернее будет сказать — описание того, как мы это понимаем в своей сфере. Начну с отдаленной исторической параллели (некоторые из параллелей вам могут показаться не имеющими никакого прямого отношения к интеллектуальной обработке данных, но, поверьте, это не так очевидно).

В 1920-е годы некий ученый, который часто ездил за границу, однажды попал в Швейцарии на завод, где делают швейцарские часы, и обратил внимание на то, что по залу, огромному цеху, ходит человек, у которого нет никаких обязанностей. Посмотрит в этом месте, пощупает в другом. Наш ученый задает вопрос руководителю предприятия:

— Этот человек — чем он занят?

— Это математик.

— А какие у него обязанности?

— А никаких. Он просто ходит по залу и время от времени говорит, что нужно поправить.

Такой «смотрящий» не имел никаких прав. Не имел права отдавать приказы. Он мог только без доклада ходить к начальнику штаба, его заму и говорить: «Вот мне кажется, что здесь лучше бы сделать по-другому».

Приведу другой яркий пример.

Как известно, Великобритания год воевала в одиночку против фашистской Германии. Это был период, когда Англия и Франция объявили войну после вторжения Германии в Польшу. Мы тогда оставались в стороне. Америка в войну не вступила, Великобритания осталась один на один с агрессором, поскольку Франция, по существу, сопротивления не оказала (там единственная часть, которая воевала, — это танковая дивизия генерала де Голля). Великобритания была в блокаде, нуждалась в продовольствии и боеприпасах. Все необходимое для того, чтобы успешно воевать, подвозили США, в основном на кораблях. Эти корабли немцы пытались уничтожить и делали это с помощью подводных лодок и авиации. Но подводных лодок при всем желании было не так много. Американцы запустили серию транспортных кораблей «Либерти»: корабль собирался за месяц, потом воевал пару месяцев, а потом его топили. Пошел массовый поток перевозок, и потери были огромные. Основную роль сыграла авиация, подводные лодки не справлялись. Тогда адмиралтейство решило, что пора поставить на эти корабли зенитные орудия, хотя бы одну пушку, чтобы она стреляла и стреляла... Поставили. Стреляла. Потом немцы усилили авиационный нажим на Англию. Страшные бомбежки, некоторые города сровняли с землей. На заседании Совета адмиралтейства посмотрели, сколько самолетов сбили пушки, поставленные на кораблях, и показалось, что практически ничего не сбили. Тогда решили вернуть пушки на защиту городов. И вот тут «смотрящий» сказал: мы неправильно формулировали критерии; пушки стоят не для того, чтобы сбивать самолеты, а для того, чтобы охранять корабль. Давайте посчитаем, сколько кораблей потоплено, у которых пушки были, и сколько — где пушек не было. Оказалось, что разница колоссальная. Действительно, самолет такими пушками не сбивался, но точность бомбометания становилась совсем другой. Казалось бы, совсем незначительное замечание привело к сохранению огромных ресурсов.

Я не буду приводить другие примеры, были удачные, были неудачные. Считается, что тут как раз и начинается интеллектуальная обработка данных. Действительно, люди, которые были приглашены, были совсем не военными людьми, но вот в деле интеллектуальной обработки данных у них, применявших способности своего мышления, получалось очень неплохо.

Потом начался век вычислительных машин. Сначала машина на 512 500 операций в секунду. Затем появились машины «Стрела», БЭСМ. Это были машины, которые выполняли простую счетную работу. Но решение уравнений — это не интеллектуальный труд. Дальше появились персоналки, появился Интернет. Первым, по существу, Ин-

тернет предложил В. М. Глушков. Но его идеи не были поняты до конца, поэтому мы такую сеть не создали.

Наконец, появились современные машины, мы знаем, как много и быстро можно на них делать. Но имеется огромное количество задач, которые чисто формальным путем не могут решаться. Например, в большом городе торгуют молоком, 500 магазинов обслуживают 30–40 грузовиков, нужно создать оптимальный график работы. Современная сверхмашина будет работать пару миллионов лет, обсчитывая такое число операций. И математиками доказано, что быстрее нельзя. Формальным путем такая задача не решается.

Итак, первое — это «смотрящие», второе — когда возникло понимание, что есть задачи, которые на уровне чистой формализации не разрешимы. Между тем есть масса ситуаций, когда надо принять решение. Пример. Есть фирма, в которую вы хотите вложиться. Есть огромное количество параметров, числа, высказывания людей, графики, психология людей, карты. Если все это учесть, может быть, вы и примите правильное решение. Нужно придумать какую-то методологию, как решать такую задачу. Решение такого рода задач, когда ситуация описывается и словами, и числами, и мнениями, когда информация разнородная, не сводящаяся к единому общему, — вот это и есть основа того, что стало называться интеллектуальной обработкой данных. В этой области работают коллективы исследователей и очень мощная вычислительная техника в тесном контакте, симбиозе.

Пример. В Калифорнии столкнулись с задачами, которые ни одна машина не решала, не могла построить оптимальный план. Они подумали: а люди в своей практике с такими задачами сталкивались? Нашли людей, которые не имели ни малейшего представления о математике, но такие задачи решали. Это были банковские служащие. Они предлагали клиенту банка разместить деньги. Если клиент получал прибыль, то консультант получал долю в 15–20%. Если клиент нес убытки, то они быстро возмещались, сумма не уменьшалась. Если консультант оказывался неразумным человеком, то он быстро разорялся. Как правило, люди давали правильные прогнозы. Они в уме решали такие математические задачи, которые машины не решали, с огромным количеством переменных. Они сами не знали, как они это делали. Это была интуиция, интеллект. Это и была интеллектуальная обработка данных.

Группа математиков, инженеров, экономистов создала кружок, пригласив успешных людей (они были тогда на пенсии). На протяжении нескольких месяцев их допрашивали, как шел у них процесс мышления. Каждый высказывался, нередко противореча другому, накапливалась определенная информация, которая потом чистилась, обрабатывалась, просматривалась многими людьми. Вырисовывался некий алгоритм, некоторая система правил, по которым люди создавали прогнозы. Как говорят математики, противоречащий пример строится элементарно. Но почему-то в реальной жизни противоречивые примеры не встречаются, а то, что встречается в реальной жизни, имеет совсем другой характер. В реальных задачах, которые возникали, то, что успешные люди предлагали, с точки зрения математиков было совершенно некорректно, но — работало. И вот с этого начинается интеллектуальная обработка данных. Чтобы реализовать все, что вытащили из консультантов, понадобились машины, и счет, и разные математические фокусы.

В Советском Союзе в начале 1960-х годов возник микрокризис — резко истощались золотые месторождения. У нас золота было много, но оно было не южноафриканского типа (есть золотоносные месторождения южноафриканского типа, где бе-

рется огромное количество золота, но они образовывались в специфических условиях). Южной Африке повезло: золото у них залегало на небольшой глубине, там шло все быстро. У нас таких мощных месторождений не было; основная добыча — промывка и самородки. Всего шесть месторождений. Было проведена мощнейшая разведка по всей стране, но оказалось, что ничего не нашли. Вот тогда и вызвали нас: географ, геофизик, математик — и сказали: вот географические карты, любые справочники, укажите, где искать золото. Был создан некий метод расчета, алгоритм придумал я, потом его еще называли шаманским алгоритмом. Он дал четыре точки на территории Советского Союза, где вероятно, может быть то, что мы искали. В двух из них было начато пробное бурение — одно в Узбекистане, другое в Сибири. Там и там золото южноафриканского типа было найдено (другие точки не бурили).

С чего начинается интеллектуальная обработка данных? Вы имеете информацию, подсчитать которую даже на самой мощной машине не получается. Информация, как уже было сказано, разнородная: и числовая, и высказывания, графики, мнения. Всю информацию надо превратить в более-менее точный алгоритм, или последовательность алгоритмов, или систему алгоритмов, которые подстраиваются, меняются, самоорганизуются. Такая система сегодня составляет одну из основ обработки данных. Одна большая ветвь реализуется во всем мире (в США, России, Европе). Думаю, что по сложности решения задач мы впереди.

У меня есть ученик Александр Яковлев — доктор наук, ведущий математик. Развивающиеся страны объявляют конкурс на лучшее предсказание: будет ли данная страна развиваться либо впадет в стагнацию. А. Яковлев выиграл первое место, и газета «Нью-Йорк таймс» написала, что лучший прогнозист живет в России.

Одна из последних вещей, которая произошла в области интеллектуальной обработки данных: фирма «Перекресток» объявила международный конкурс. Задача была такая: определить, какие продукты необходимо завозить на неделю вперед. Давалась информация по предыдущим годам: о том, сколько было завозов и сколько было не реализовано. И вот по нескольким годам необходимо было составить прогноз. Наша команда выиграла с огромным перевесом. Решить такие задачи без техники и сложной математики не могли, но надо было привлекать и собственные решения, и это был искусственный интеллект.

Другая ситуация связана с робототехникой. Связана с одним из колоссальнейших просчетов крупнейшего математика начала XX столетия немца Д. Гильберта. Гильберт был человек своеобразный. Его аспиранты не имели права жениться, женившийся аспирант немедленно изгонялся, считалось, что это несовместимо, сначала защитись — потом женись. Гильберт поставил задачу полной формализации математики, и это была единственная вещь, когда математик с треском провалился. В конце 1930-х годов другой немецкий математик Гёдель доказал, что даже арифметика не формализуется. А если взять более сложную область, ее тем более формализовать нельзя.

Роботу вы задаете набор определенных правил, например: «Нельзя наносить вред человеку». Это для него аксиома, он действует в рамках этого правила. Но вот он столкнулся с ситуацией, которая правилами не охвачена. Если он не может на анализе этой ситуации, отходя от тех правил, которые у него есть, создать новое правило, то он никогда не заменит человека. Человек такие вещи совершает неосознанно. Значит, нужно уметь на основе ранее построенного формализма и вновь поступившей информации сформулировать новые правила действия. Это и есть искусственный интеллект. Искусственный — потому что делает робот, а интеллект — потому что это неформаль-

ная вещь. Формализм в него был уже вложен, а эта задача на сегодняшний день безумно далека от решения. Кое-что сделать удастся. И вот это еще одно мощное направление, которое принято называть интеллектуальной обработкой данных: в работа вкладываются не только система аксиом, но система правил, с помощью которых он может модернизировать систему аксиом, систему правил, наработать новые правила, оценивать их. Вот это все делает математик, программист, инженер. Кстати, гуманитарий тоже широко работает в этой области, особенно на Западе. Мышление гуманитариев гораздо более продуктивно, чем решение формального математика. Гуманитарий часто дает неправильное объяснение своего решения, а решение принимает правильное.

Дата поступления: 20.04.2014.

ON INTELLECTUAL INFORMATION PROCESSING METHODS

Yu. I. Zhuravlev

*(The Department of Mathematical Sciences of the Russian Academy of Sciences,
Moscow University for the Humanities)*

In March 2014 a methodological seminar took place in the Institute of Fundamental and Applied Studies at Moscow University for the Humanities (IFAS MosUH). Yuri I. Zhuravlev, full member of the Russian Academy of Sciences, the director of the Mathematical and Computer Modelling Centre of the IFAS MosUH presented a report on intellectual information processing methods. Among the participants of the seminar there were researchers of the IFAS, students of the university. The outstanding Russian mathematician told about the beginnings of the idea of the application of special methods for intellectual data-processing operation in situations of underformalizable problems solving, difficulties and prospects of the development of artificial intelligence. He made numerous examples, including ones from his own research work when they succeeded in building of an algorithm for the solution of such objectives. The report was set forth taking into consideration the interests of humanities scholars. The academician told them about the mathematical issues very lucidly and emphasized that every scientist must be able to talk about his work in such a way as everyone could understand him/her.

Keywords: information, intellectual information processing, artificial intelligence.

Submission date: 20.04.2014.

Журавлев Юрий Иванович — академик РАН, председатель секции «Прикладная математика и информатика» Отделения математических наук РАН, заместитель директора Вычислительного центра им. А. А. Дородницына РАН по научной работе, директор Центра компьютерного и математического моделирования Института фундаментальных и прикладных исследований Московского гуманитарного университета; иностранный член Испанской королевской академии, Национальной академии наук Украины, Европейской академии наук, доктор физико-математических наук, профессор. Адрес: 119333, Россия, г. Москва, ул. Вавилова, д. 40. Тел.: + 7 (499) 135-61-39. Эл. адрес: wcan@ccas.ru

Zhuravlev Yuri Ivanovich, full member of the Russian Academy of Sciences, the Chairman of the Section “Applied Mathematics and Informatics” of the Department of Mathematical Sciences of the RAS, deputy director for research work of the Computing Centre of the RAS, the director of the Mathematical and Computer Modelling Centre of the Institute of Fundamental and Applied Studies, Moscow University for the Humanities; foreign member of the Spanish Royal Academy of Sciences, the National Academy of Sciences of Ukraine, the European Academy of Sciences, Doctor of Science (physics and mathematics), professor. Postal address: 40, Vavilova str., Moscow, 119333, Russia. Tel.: + 7 (499) 135-61-39. E-mail: wcan@ccas.ru