

ФЕНОМЕН ИДЕНТИФИКАЦИИ

АЛЕКСАНДРОВ В.В., КУЛЕШОВ С.В., ШАННАГ Б.

УДК 004.6

Александров В.В., Кулешов С.В., Шаннаг Б. Феномен идентификации.

Аннотация. Рассматривается феномен идентификации в его развитии. Рассмотрены различные математические модели и подходы к идентификации. Показано, что понятие программируемой технологии А.Н. Колмогорова адекватна возможностям цифровой коммуникационной среды.

Ключевые слова: развитие, цифровая цивилизация, идентификация, инфокоммуникация.

Alexandrov V.V., Kuleshov S.V., Shannaq B. Phenomenon of identification.

Abstract. This paper considers the phenomenon of identification in its development. Some mathematical models and approaches to identification are shown. The adequacy of Kolmogorov's programmed technology and digital communication environment is shown.

Keywords: grow, digital civilization, identification, infocommunication.

1. Введение. У Платона в конце диалога «Федр» есть такой эпизод. Главный землемер фараона Тевт демонстрирует свое изобретение — алфавит и письменность — фараону Тамусу: «Эта наука, царь, сделает египтян более мудрыми и понятливыми, так как найдено средство для памяти и мудрости». Однако фараон был не в восторге: «В души научившихся алфавиту и письменности они вселят забывчивость, так как будет лишена упражнений память: припоминать станут извне, доверяясь письму, по посторонним знакам, а не изнутри, сами собой. Стало быть, ты нашел средство не для памяти, а для припоминания. Ты даешь мнимую (виртуальную — прим. автора), а не истинную мудрость». В этом диалоге, по сути, содержатся концепция цивилизационного развития инфокоммуникаций, проблема появления специфического многообразия языковых форм и профессионально ориентированных языков (языков математики, физики), а также предвидение негативных последствий социального расслоения общества (власть знания становится конкурентной средой элиты).

Так называемое единое информационное пространство Интернет-сообщества — в виртуальном мире аналог библейского мифа о вавилонской башне [1] — порождает «языковый» хаос (скопление различных текстов, неупорядоченных символов, понятий и терминологий). Семантико-смысловое содержание, информация и знание по-прежнему остаются уделом профессионалов.

2. Десемантизация смысла. Профессионально ориентированные языки расцветают и угасают по мере потребности обслуживать тот или

иной технологический процесс, в том числе и онтологический процесс представления знаний.

Все сенсоры человека, в первую очередь, зрение и слух, призваны трансформировать сигналы в нашем мозгу, формируя специфически функционирующую память, способную идентифицировать через *имя—слова* свойства воспринимаемого объекта. Это один из элементов реализации сложнейшего для анализа акта коммуникации. Например, анализ и обработка потока аудио- и видеоданных бессмысленны без адаптации к целям реализации коммуникативного акта.

Нашу биологическую природную ограниченность сенсоров и психофизиологических свойств зрения, слуха, а также до сих пор не понятую эффективность организации памяти и разума человека мы постоянно пытаемся распространить на окружающие нас процессы, зачастую путая причины и следствия, несоразмерность временных интервалов существования космических, геологических и прочих макро- и микропроцессов, сопоставляя и воспринимая их через призму интервала существования персонализирующего производителя и носителя знаний. Знания сами по себе не содержатся в книгах, журналах, газетах, компьютерных дисках, ресурсах Интернета и прочих «папирусах» — информационных носителях. Это лишь форма хранения сведений, фактов, данных, которые инициализируются каждым индивидуумом в зависимости от его персональной подготовки, полученного образования, социально-культурных аспектов окружающего его общества. Отсюда следует естественность вопроса: «Почему история нас ничему не учит?». Естественные и искусственные катаклизмы в природе и обществе повторяются с завидной регулярностью с цикличностью от 60 до 100 лет — временем биологической памяти двух поколений.

Интеграция свойств воспринимаемого объекта транслируется через *слово—имя* в *слово—смысл*, а совокупность этих процессов и составляет такое понятие, как разум. Персонализируемый интеллект всегда опирается на свое внутреннее воображение (иллюзии сна — прототип индивидуального мира воображений) [2].

В настоящий момент пространство инфокоммуникации (Интернет) на 80 % засорены спамом, блогосфера превратилась в «помойку» массмедиа. Это параллельный мир, в котором царят дезинформация и популизм, в котором мечты и фантазии их создателей преподносятся как «эсклюзивная информация». Кто несет ответственность за достоверность распространяемых данных? Возникшая проблема не в технологии «свободы слова», а в фильтрации семантико-смыслового содержания и ее целевого использования.

Материалы Интернета обезличены, что приводит к десемантизации — приписыванию слову смысла, которого у него не было, и далее использование этого слова в новом смысле, при одновременном попирании смысла старого.

Интернет-сообщество поддерживает краткосрочную тактику и неспособно следить за стратегией. Известен пример игры в шахматы гроссмейстера с телевизионной аудиторией, где каждый телезритель на ход гроссмейстера предлагает свой ответ. Ответный ход «усредняется» и предъявляется гроссмейстеру. Стратегия и тактика гроссмейстера одина. Усредненная тактика и стратегия телеигроков проводят к бессмыслице.

Деперсонализация запутывает реализацию любой концепции, но в то же время, соблюдая единую стратегию и тактику построения текстов (лингвистическое зомбирование), позволяя управлять и контролировать смысловое восприятие, разрушая традиции и культуру (примером могут служить любые экспресс-опросы).

Цифровая цивилизация — технологическое обеспечение глобализации. В ней проблема обучаемости и распространения знаний в текстовой форме перетекает в проблему самообучения (autopedia) идентификации смысла.

Отличительной чертой человека, его необходимым и достаточным свойством является идентификация себя (персонализация) через сопоставление с другими (логика «Другого»), которая порождает конкурентную среду.

В контингентной системе любой объект значим в той мере, в какой он *этот*, а не *другой*. Смысл проявляется, если хотя бы один из элементов системы имеется в оппозиции. Если же такого элемента нет, то и отсутствие *другого* никто не заметит.

3. Постулат Платона. Вся история развития цивилизаций отражается в эволюции языка как эффективном инструменте идентификации и инфокоммуникации. Инвариантность форм и специфических свойств представления знаний в знаковой форме и тексте вытекает из постулата Платона: *имя предмета, его речевое и образное обозначение (запись) должно отражать его свойства и быть уникальным идентификатором.*

... имена у вещей от природы и что не всякий мастер имен, а только тот, кто обращает внимание на присущее каждой вещи по природе имя и может воплотить этот образ в буквах и слогах.

...законодатель, о котором мы говорили, тоже должен уметь воплощать в звуках и слогах имя, причем — то самое, какое в каждом

случае назначено от природы. Создавая и устанавливая всякие имена, он должен также обращать внимание на то, что представляет собою имя как таковое, коль скоро он собирается стать полноправным учредителем имен.

...Имя «человек» означает, что, тогда как остальные животные не наблюдают того, что видят, не производят сравнений, ничего не сопоставляют, человек, как только увидит что-то, а можно также сказать «уловит очами», тотчас начинает приглядываться и размышлять над тем, что уловил. Поэтому-то он один из всех животных правильно называется «человеком», ведь он как бы «очеловец» того, что видит...[3]

Эксклюзивное свойство природного феномена человека — «видеть семантические универсалии» и выражать их через онтологию, как взаимно-однозначное соответствие между понятием, именем и свойством объекта (например, «солнце всходит и заходит»). При этом критерием адекватного синонимического соответствия, в отличие от вычислений некоторых мер (близости, похожести, релевантности и др.), является поиск семантических универсалий. Обратим внимание, что эмпирически разработчики компьютерного пользовательского интерфейса используют иконографическое представление как субязыковые семантические универсалии, иначе — идентификаторы.

4. Семантические универсалии. На рисунке приведена эволюция языка. Заметим, что сначала был семантический образ идентифицируемого предмета, затем стилизация для не умеющих рисовать, следующая уровень — алфавит. От простоты к сложности и обратно.

Семантические универсалии (Шумеров)	Поздняя клинопись	Древнеавилонские	Ассирийские	Онтологическое понятие
				Птица
				Рыба

Движение от семантических универсалий к абстрактным понятиям.

Простота восприятия конкурирует со сложностью записи исходного семантико-смыслового представления. Пиктограммы эволюционируют к простоте формы: знака, символа, алфавита и языка. Удивительно, но факт, что при отсутствии на тот момент каких либо средств связи древние центры цивилизации Египет, Китай, Мексика приходи-

ли к аналогичным системам счета и языковым формам записи. Исторически складывалось так, что цивилизации, достигнув разнообразия приблизительно в 700 пиктограмм, вступали в кризисную фазу цивилизационного развития.

Можно утверждать, что образ или событие любой природы имеют некоторую собственную семантику. Эта собственная семантика универсальна и контекстно-интерпретируема в некотором конечном наборе смыслов или событий. Это утверждение основано на возможности осмысления, хотя возможно и не однозначного, практически любого образа или события внешнего мира, даже лишённого собственного контекста.

Однако при этом возникает проблема смысловой идентификации, которая требует сложного и длительного обучения лингвистическим премудростям для отображения смысла на язык текстового описания.

5. Математические модели идентификации. Для тождественной идентификации смысла текстовых дипломатических договоренностей кардинал Ришелье заказал Р. Декарту метод поиска консенсуса (изложенный в труде «Рассуждения о методе»), в котором Декарт предложил математическую модель идентификации (иными словами, мощное информационное оружие): *... цель этого изобретения Декарта — не что иное, как наделение точек X плоскости именами, с помощью которых мы могли бы их различать и опознавать. Наделение именами должно производиться каким-то систематическим способом; так как точек бесконечно много, систематичность тем более необходима, что все точки, в отличие от людей, совершенно одинаковы и, следовательно, мы в состоянии различать их, лишь прикрепив к ним какие-нибудь ярлыки. Имена, которые мы используем, оказываются парами чисел (x, y) ...*[4]

В иной форме для подобных целей информационного анализа Г. Лейбниц предложил принцип идентификации неразличимости на основе формулировки понятия конгруэнтности. Конгруэнтность, или «универсальная характеристика» по определению Лейбница, открывала возможность *...человеку рисовать, не имея дара художника...* [5]

Решение проблемы смысловой неоднозначности потребовало развития математических методов поиска взаимно однозначного соответствия тождественного, подобного и похожего: идентификации данных физического процесса и математической модели (формализации записи).

Фильтр Р. Калмана, признанный в теории управления, почти через 50 лет привел автора к пониманию «революционности открытия» И. Ньютона [6]:

Кеплеровские законы \Leftrightarrow ньютоновские законы гравитации.

Здесь символ \Leftrightarrow означает изоморфизм или математически тождественное суждение. Слева дано идеализированное строгое описание кеплеровского движения планеты. Справа дана формулировка, претендующая быть точным количественным описанием невидимой силы, прямое измерение которой мыслимо только в специальных условиях. Эта теорема утверждает, что эти два совершенно различных утверждения абстрактны и тождественны.

Другими словами, утверждается следующее: феномен взаимной идентификации находится в однозначной связи между данными, относящимися к действительности, физической реальностью, и простотой математической модели. То есть, если имеются хорошие данные (наблюдаемые или экспериментальные), то существует только одна модель, способная объяснить их. Чтобы выяснить, какая однозначная модель соответствует данным, следует заниматься не физикой, а математикой [6].

Таким образом, Р. Калман проясняет и уточняет область работоспособности его фильтра. Идентификация понятия данных в его математической модели требует строгого соответствия понятия числа и разрешенных операций над ним, которые отнюдь не адекватны структурам данных в компьютере.

Другое понятие идентификации данных в теории управления было предложено Р.М. Юсуповым как установление параметров модели объекта:

«Под идентификацией будем понимать процедуру построения оптимальной (наилучшей) в определенном смысле математической модели объекта по реализациям его наблюдаемых входных и выходных сигналов» [7].

Формальная математическая модель идентифицируемости — это отображение программой некоторой исходной записи X в уникально идентифицируемую кодированную последовательность Y , которая, будучи переданной, формирует битовый транспортный поток. В этом случае транспортный поток не имеет ничего общего с исходным контентом и к нему не применимо понятие информация.

С развитием современной вычислительной техники будет понято, что в очень многих случаях изучение реальных явлений разумно

вести, избегая промежуточного этапа их стилизации в духе представлений математики бесконечного и непрерывного [8].

Понятие алгоритма по А.Н. Колмогорову заключается в существовании программы отображения «записи» в номер и обратно. При этом поиск оптимальной программы опирается на критерий сложности как длину минимальной программы. Понятие программы одновременно выполняет роль идентификатора и содержания «записи». Отсюда следуют различные понятия информации:

— по Шеннону — энтропийная (ансамблевая) идентификация процесса;

— по Колмогорову — объектная идентификация (конкретный элемент из множества).

б. Цифровые программируемые технологии. Цифровые технологические реализации семейства коммуникационных протоколов (в частности TCP/IP) оказались адекватной базой программируемой технологии А.Н. Колмогорова.

Цифровая программируемая технология в своей основе опирается на принцип идентификации из определения записи как битовой последовательности, обобщая любые формы «записи» информационного содержания, кардинально уменьшая энергетические затраты при их хранении и транспортировке [9].

В такой постановке задача передачи битовой последовательности становится эквивалентна задаче шифрования: незнание программы отображения номера в запись (транспортного потока в контент) не позволяет восстановить (расшифровать) исходное сообщение.

Использование программируемой технологии позволяет производить многократное преобразование промежуточных представлений (преобразование форматов, компрессию, упаковку в транспортные фреймы, маршрутизацию и др.) Все эти действия являются гомоморфными относительно исходного объекта, т. е. способны менять промежуточное представление объекта для достижения максимальной плотности пропускания или минимального энергопотребления, не влияя на идентификацию смыслового содержания объекта.

В качестве примера можно привести технологию SDR (soft defined radio), которая иллюстрирует принципиальную возможность не различать реализацию транспортных потоков телевидения, радиовещания и других средств передачи данных, демонстрируя перспективы единой коммуникационной среды.

Уже сейчас использование единой коммуникационной среды (в качестве которой выступают IP-сети вообще и Интернет в частности)

делают возможным все перечисленные выше принципы передачи данных — объектную идентификацию (сохранение характерных свойств передаваемого информационного объекта вне зависимости от формы представления его в транспортном потоке), адресную идентификацию (доставку информационного объекта получателю по уникальному адресу).

Литература

1. *Александров В.В.* Развивающиеся системы. В науке, технике, обществе и культуре: Ч. 1. Теория систем и системное моделирование. СПб.: Изд-во СПбГТУ, 2000. 244 с.
2. *Александров В.В., Сарычев В.А.* Инфотварь — homo infocreature // Материалы Междунар. конф. и Российской науч. школы «Системные проблемы надежности, качества, информационных и электронных технологий в инновационных проектах (Инноватика-2006)». Ч. 3. Т. 1. М., 2006. С. 3–11.
3. Античные теории языка и стиля. М.—Л.: ОГИЗ, 1936. 343 с.
4. *Вейль Г.* Симметрия. М.: Наука, 1968. 18 с.
5. *Лейбниц Г.* Новые опыты о человеческом разумении автора системы предустановленной гармонии. В 4-х т. Т. 2. М.: Мысль, 1983. 686 с.
6. *Калман Р.* Открытие или изобретение: Ньютонианская революция в технологии систем // Авиакосмическое приборостроение. 2004. № 6. 12 с.
7. *Дмитриев А.К., Юсупов Р.М.* Идентификация и техническая диагностика. М.: Изд. МО СССР, 1987. 521 с.
8. *Колмогоров А.Н., Успенский В.А.* К определению алгоритма // Успехи математических наук. 1958. Т. 13, вып. 4. С. 3
9. *Александров В.В., Кулешов С.В., Цветков О.В.* Цифровая технология инфокоммуникации. Передача, хранение и семантический анализ текста, звука, видео. СПб.: Наука, 2008. 244 с.

Александров Виктор Васильевич — д-р техн. наук, профессор; заведующий лабораторией автоматизации научных исследований Учреждения Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: рекурсивная обработка информации текст и аудио-видео потоки, развивающиеся системы, семантический и инфологический анализ. alexandr@iias.spb.su, www.sial.iias.spb.su; СПИИРАН, 14-я линия В.О., д.39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

Alexandrov Viktor Vasilyevich — Dr.Sc., professor; head of Laboratory of Research Automation, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: artificial intelligence for modeling, control and automation, computer simulation of developing structures in nature, society and culture, semiological approach to data processing. alexandr@iias.spb.su, www.sial.iias.spb.su; SPIIRAS, 39, 14th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812) 323-5139, fax +7(812)328-4450.

Кулешов Сергей Викторович — канд. техн. наук, старший научный сотрудник лаборатории автоматизации научных исследований Учреждения Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: исследования в области обработки изображений и видеосигнала, виртуали-

зация каналов передачи данных и информационных носителей, исследования и построение семиологических информационных систем. kuleshov@iias.spb.su, www.sial.iias.spb.su; СПИИРАН, 14-я линия В.О., д.39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

Kuleshov Sergey Viktorovich — Ph.D. in Tech.; researcher, Laboratory of Research Automation, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: semiologic information systems, video data streams processing. kuleshov@iias.spb.su, www.sial.iias.spb.su; SPIIRAS, 39, 14th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)323-5139, fax +7(812)328-4450.

Боумедин Шаннаг — аспирант лаборатории автоматизации научных исследований Учреждения Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: инфология, ассоциативный поиск, исследование и построение семиологических информационных систем. alexandr@iias.spb.su, www.sial.iias.spb.su; СПИИРАН, 14-я линия В.О., д.39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

Boumedyen Shannaq

Post Gradient Student Laboratory of Research Automation, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: Infology, associative search, research and development of the semiology information system. alexandr@iias.spb.su, www.sial.iias.spb.su; SPIIRAS, 39, 14th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812) 323-5139, fax +7(812)328-4450.

Рекомендовано лабораторией автоматизации научных исследований, зав. лаб. В.В. Александров, д-р техн. наук, проф.
Статья поступила в редакцию 06.07.2009.

РЕФЕРАТ

Александров В.В., Кулешов С.В., Шаннаг Б. **Феномен идентификации.**

Единое информационное пространство Интернет-сообщества порождает «языковой» хаос (скопление различных текстов, неупорядоченных символов, понятий и терминологий). Профессионально ориентированные языки появляются и исчезают по мере потребности обслуживать тот или иной технологический процесс, в том числе и онтологический процесс представления знаний.

Знания сами по себе не содержатся в книгах, журналах, газетах, компьютерных дисках, ресурсах Интернета и прочих информационных носителях. Это лишь формы хранения сведений, фактов, данных, которые инициализируются каждым индивидуумом в зависимости от его персональной подготовки, полученного образования, социально-культурных аспектов окружающего его общества.

Вся история развития цивилизаций отражается в эволюции языка как эффективном инструменте идентификации и инфокоммуникации. Инвариантность форм и специфических свойств представления знаний в знаковой форме и тексте вытекает из постулата Платона: имя предмета, его речевое и образное обозначение (запись) должно отражать его свойства и быть уникальным идентификатором.

Решение проблемы смысловой неоднозначности потребовало развития математических методов поиска взаимно однозначного соответствия тождественного, идентификации данных физического процесса и математической модели (формализации записи). Другими словами, утверждается следующее: феномен взаимной идентификации находится в однозначной связи между данными, относящимися к действительности, физической реальности, и простотой математической модели. Если имеются реальные данные (наблюдаемые или экспериментальные), то существует только одна модель, способная объяснить эти данные.

SUMMARY

Alexandrov V.V., Kuleshov S.V., Shannaq B. **Phenomenon of identification.**

The uniform information space of Internet community results in chaos of “languages” (different texts, disordered symbols, terms and terminology). Professional oriented languages appear and disappear in order to serve various technology processes including ontology process of knowledge representation.

Knowledge itself is not contain in books, magazines, newspapers, computer disks, Internet and other information carriers. This is only form of storage for data, facts which is being initialized by each individual depending on personal training level, given education, social and cultural aspects of surrounding community.

The whole civilization history reflexes in evolution of language – effective instrument for identification and infocommunication. The invariance of forms and specific features of knowledge representation in symbolic form and text emerges from Plato’s postulate: subjects name and it’s figurative and speech designation must reflect it’s properties and to be it’s unique identifier.

The solving of semantic ambiguity had demanded to develop mathematic search methods for one-to-one correspondence of identical, physical process and mathematical model data identification (record formalization). In another words: phenomena of mutual identification is in unequivocal relation with the physical reality data and simple mathematical model. For reliable (observable and experimental) data only one model to explain them exists.