

Аналитические системы и новые архитектуры машин для обеспечения стратегического анализа

А. О. Поляков

Санкт-Петербургский институт информатики и автоматизации РАН
199178, Санкт-Петербург, 14-я линия В.О., д.39

www.polyakov.com

УДК 681.3

А. О. Поляков. **Аналитические системы и новые архитектуры машин для обеспечения стратегического анализа** // Труды СПИИРАН. Вып. 1, т. 1 — СПб: СПИИРАН, 2002.

Аннотация. Предлагаются новые программно-технические решения, учитывающие реальные успехи и возможности отечественной науки в области аналитических систем. Обсуждаются возможности создания эффективных стратегических аналитических систем «партнерского» класса, не имеющих мировых аналогов. — Библ. 7 назв.

UDC 681.3

А. О. Polyakov. **Analytical systems and new architectures of computers for the strategic analysis support** // SPIIRAS Proceedings. Issue 1, v. 1. — SPb: SPIIRAS, 2002.

Abstract. The new program-technical solutions using real successes and possibilities of domestic science in the field of analytical systems are offered. The possibilities of creation of effective strategic analytical systems of the "partner" class, having no analogies in the world, are discussed. — Bibl. 7 items.

1. Аналитические системы и машины — второй этап противостояния

По ряду причин скорее психологического и экономического, чем технического характера, западные разработки программного обеспечения стали в нашей стране де-факто стандартом практически во всех областях его применения.

Почти все имеющиеся человеческие программистские ресурсы за малым исключением оказались у нас направлены на освоение и эксплуатацию этих программных продуктов, хотя, как известно, это «путь к отставанию навсегда». Однако, если такое положение с программированием естественно в сфере бизнеса и частной экономики, то в приложениях стратегического значения это абсолютно недопустимо.

Проиграв гонку в области техники и технологии системотехнических решений, что, вообще говоря, не является, вопреки сложившемуся мнению, крупной потерей¹, мы, однако, пока еще не проиграли в области «программистских

¹ Считается, что страна проиграла аппаратную гонку. Но мы не одиноки в проигрыше попыток нахождения лучших реализаций фон Неймановской архитектуры. Фактический провал аналогичных попыток Франции, Японии, Англии, Швеции означает, что это не крах технической политики, но просто «смысл предмета борьбы» подошел к границе своей целесообразности. Фон Неймановская архитектура исчерпала все возможности своего развития, и гонку выиграл «технологически сильный».

Далее аппаратная конкуренция будет продолжаться на поле принципиально других технических решений и других архитектур. Речь идет не о переводе фон Неймановских машин на квантовую, оптическую и другую экзотическую базу, не о многопроцессорных машинах, но о принципиально новых конструкциях и «невывчислительных» технологиях. Рассматриваемые в настоящей статье проблемы, кроме всего прочего, суть стартовая площадка создания новых архитектур машин, имеющих, прежде всего, стратегическое значение. Мы уже располагаем не-

мозгов». Это произошло потому, что наш специалист (из числа тех, кто не ушел в бизнес «программирования на подхвате» типа банковских систем и в науку, «грантированную» западом) не обременен гонкой за программистской технологией очередных модных решений известных фирм, с которыми надо согласовывать языки и интерфейсы, не зависим от «модных» направлений исследований. Мы свободны в наших тенденциях развития, и нас не держит ожидание очередных стандартов от фирм, заинтересованных в движении вперед, не опережающем коммерческую прибыль, не держит ничто, кроме собственных амбиций наших программистов и специалистов, рыночная стоимость которых зависит сегодня только от знания ими зарубежных продуктов и способности выполнять исследования, соответствующие общепринятым тенденциям.

Прекрасно понимая это, западные эксперты направленно сдерживают наши последние стратегические возможности в области программирования, подсовывая нам легально и нелегально решения (в частности, в области баз данных и хранилищ), не лезущие ни в какие разумные рамки. Иначе просто невозможно объяснить наш путь по СУБД от dBase'ов и ADABAS'ов к Oracle и иже с ней, аналогичный пути по микромодульным и прочим заимствованиям негодных технологических решений для технических систем, включая весь ряд печально известных машин класса ЕС.

Если мы можем понять, что товарных отечественных технических решений «по фон Нейману» нам не дождаться, то, по крайней мере, мы не должны попадать в зависимость от иноземных программных решений на прикладном уровне в стратегических аналитических разработках, не должны решать неразрешимые проблемы сертификации импортных баз, операционных систем и прочих аналогичных «пакетов программ».

Отметим, что сертификация — это проблема не техническая, а организационно-политическая. Пора смириться с тем, что если «дырки» в вашем программном продукте затыкают чужие программисты, то никакая сертификация не поможет. Это тем более важно, что обеспечение стратегических решений заключается не в том, чтобы делать машины «лучше всех» (конечно, это было бы не плохо), а в том, чтобы уметь использовать имеющиеся и будущие машины лучше потенциального соперника.

Аналитические системы-партнеры, с автоматическим принятием решения в контакте с профессиональным аналитиком — важнейшая составляющая современного стратегического оружия, без которых его эффективность падает многократно. Очень хорошо, что мы не должны тратить силы и время на создание для них фон Неймановских машин и операционных систем, а можем заняться сразу системами такого рода.

Пусть заимствованные операционные системы и машины будут сколь угодно не сертифицированными — это не поможет несанкционированному доступу, если прикладные программы стратегического значения делать самим и так, чтобы они были не только платформи-независимыми², но и общаться могли только с человеком, непрерывно идентифицируемым ими не

которыми проработками такой направленности [1]. *Сегодня важно принять правильные организационно-стратегические решения о выборе дальнейших путей развития, отвечающих интересам России, ибо первый проигрыш был технический, т.е. опасный не более чем тактически, а во втором случае можно проиграть противостояние информационное, то есть проиграть все.*

² Известны очень неплохие системы уже более 30 лет существующие и функционирующие вне зависимости от всех технических революций, произошедших за это время.

по кодам, ключам, зрачкам и пальцам, а по менталитету и контексту восприятия известных только ему далеких событий, передать которые кому-либо невозможно даже теоретически.

Как будет ясно из дальнейшего, преодоление такого барьера доступа путем встраивания «закладок» и обхода модуля защиты исключено. Защиту данных не сможет преодолеть даже ведущий программист разработки. По этой же причине полностью похищенный вместе с машиной объем данных без подготовленного вместе с ними человека-системного оператора не читаем или, при необходимости, может быть сделан читаемым «с целенаправленным искажением смысла записей».

Более того, аналитическая система, выполненная без предварительной структуризации базы, вообще имеет смысл только в «живом», непрерывно актуализируемом состоянии, а оторванная от постоянной связи со своими привычными источниками сообщений и запросов формально превращается в реляционную базу без априорно заданной структуры, т.е. в нечто «равномерно размазанное по памяти».

Конечно, это накладывает некоторые дополнительные требования к обеспечению работы таких систем, и делает их не самыми дешевыми, но, как будет показано, и возможности удовлетворения всех мыслимых требований к аналитическим способностям машины, в условиях безопасного хранения данных мы получаем неограниченные.

При рассматриваемом подходе обеспечение безопасности системы может стать не грузом, а облегчением процедуры ее создания. Это следует из того, что и аналитическое решение машины о стратегической обстановке и целесообразном поведении в контакте с источниками сообщений и системным аналитиком, и решение об идентификации личности в контакте со сведениями, которые эта личность знает о себе, и которые анализируются машиной в контексте предыдущего общения — аналитические задачи одного порядка, *неразрешимые без предыстории общения.*

2. К проблеме «аналитического» противостояния

Защита и получение дополнительных данных традиционно трактуются как сохранение их конфиденциальности и обеспечение выгоды от разного объема знания сравнительно с противостоящей стороной. Все это было верно в период противостояния технико-экономического, но постепенно такая постановка теряет свою стратегическую значимость, в связи с противостоянием аналитическим переходит в раздел тактических средств и задач.

Сегодня к стратегическому вооружению, прежде всего, необходимо отнести не столько энергетические и обеспечивающие их «информационно-фактографические», сколько аналитические системы, *автоматически, как равноправный партнер, вырабатывающие в контакте с аналитиком стратегию поведения вооруженных сил и самого государства в соответствии с глобальной оценкой суммы частных оперативных обстановок, ее экономических, географических, политических и прочих составляющих.*

Выигрывает тот, кто быстрее и правильнее обработает потоки сообщений, приходящих в самом разном виде из множества источников. Быстрее всегда будет работать машина. Правильнее — она же, но только через определенный период партнерской работы. Никаким увеличением численности аналитиков

или мыслимыми затратами на получение дополнительных сведений решить задачу «совмещения быстроты и правильности» невозможно.

Техническое решение этих проблем следует искать в новых концепциях хранения данных, и, как будет рассмотрено далее, в практически полезном, естественном, а не модельном подходе к понятию выработки аналитической информации, и в создании принципиально новых, не фон Неймановских архитектур технических средств анализа сообщений.

Мы не утверждаем, что эту задачу можно решить быстро и просто. Мы утверждаем только, что задачу *решить можно* уже сегодня, и указываем один из путей ее решения. Решать такую задачу, по уровню важности равную проблемам создания нового стратегического оружия, нам придется все равно, несмотря на значительные затраты, но потом все это обойдется многократно дороже, как это было уже указано в примеч. 1, ибо «заимствовать» англоязычные системы для стратегических задач рассматриваемого класса невозможно.

Для того чтобы на текущем этапе выделить основное направление таких работ укажем, что «аналитическое оружие» можно представлять состоящим из следующих основных составляющих:

- системы стратегической аналитики;
- системы поддержки тактических решений;
- системы извлечения, уничтожения и искажения информации в технических средствах и сетях (здесь речь не о вирусах и механических воздействиях).

Глобальный смысл каждого пункта целесообразно рассмотреть в отдельных работах, а здесь только укажем, что все эти направления связаны единым подходом — проблемой создания информодинамической машины, пригодной, разумеется, и для множества других задач. О структуре машины можно получить некоторые сведения в [1], стоимость ее разработки не может быть ниже затрат на создание нового класса технических средств высоких категорий технологической сложности. Однако можно уверенно утверждать, что страны, обладающие машинами такого рода, получают существенное преимущество во всех областях, и, прежде всего, в военной.

Начальным этапом работ может являться создание систем стратегической аналитики «на пределе возможностей фон Неймановской архитектуры». Указанный этап рассматривается в настоящей статье. Он является базовым для разворачивания работ по конструированию новой обрабатывающей «неисчислительной» машины, и без него новая информационная техника реализована быть не может.

3. Концепция структурно-свободного подхода

Начнем с того, что и реляционная, и объектная концепции хранения и разработки хранилищ данных являются концепциями «насильно привнесенной структуры». Достаточно ясно, что при создании хранилищ для систем стратегической аналитики необходимо отказаться от идеологии проектируемых баз и найти возможность обеспечения быстрой и эффективной структуризации хранимого в зависимости от изменяющейся внешней обстановки.

Прежде всего, разберемся с фактически навязанными нам (за непризнанием нами же собственных приоритетных разработок) подходами Кодда и прочими «классическими решениями» в области концепций баз данных. Фактически, если это не «модели от лукавого», то уж точно «модели от системного про-

граммиста», модели от индивидуального восприятия, но не от собственной структуры организации данных в некоторой текущей ситуации. Отметим, что в сколько-нибудь серьезных базах интересной является не столько статистика записей, сколько динамика связей данных.

Достаточно ясно, что про это разработчикам целесообразно молчать по причине их неспособности обеспечить сколько-нибудь быстрое изменение внутренних связей при любом из известных подходов к проектированию структуры баз. Если же мы прекратим заимствование подходов, то увидим, что задача разрешима. Монополия проектируемых баз и объектно-ориентированного подхода уже давно могла быть нарушена существованием баз самоструктурирующихся, структурно-свободным подходом, активным построением (от потока входных сообщений) внутренних взаимосвязей записей в поле машинной памяти.

Как известно, объектно-ориентированный подход «от Simul'ы», идущий на смену построениям Кодда, суть та же структурная мертвечина, что и просто реляционный. Если в одном случае мы имеем обобщенное описание, а в другом описание по частям, названными объектами, пусть даже и снабженными некоторыми «действиями», то это удобство чисто проектное, хотя и структурное, но не динамическое. Поэтому рассмотрим совершенно другой подход, первой особенностью которого является его прозрачность и простота восприятия.

Откажемся от существующих невразумительных определений понятия «знание» и примем простое и понятное решение. *Данные* суть записи, факты, выраженные на каком угодно языке. *Знание* — текущее контекстное именованное отношение между данными. В каждый следующий момент отношение между данными может измениться или стать воспринимаемым в другом контексте. Соответственно должна быть обеспечена эффективная актуализация знания (при необходимости и данных) в реальном времени существования системы.

Исходя из сказанного, можно создать на основе «естественного» языка проблемно-ориентированный язык (ПОЯ) некоторой профессиональной области интересов, построенный из предикатов³ стандартного вида ARB , где A и B — термины, R — реляция, отражающая взаимоотношения явлений, объектов, процессов и другого, заданного указанными терминами⁴.

Теперь мы имеем все, чтобы сконструировать искомую нами базу с динамической структурой

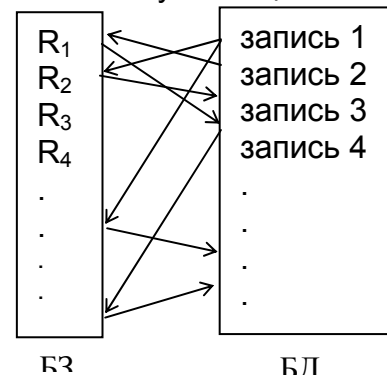


Рис.1. Архитектура структурно-свободной базы

³ Мы не утверждаем, что предикатный вариант преобразования языка лучший из возможных или то, что он должен быть именно таким. Мы только показываем путь разделения данных и знания, ведущий к построению динамической архитектуры.

⁴ Двухместные (бинарные) отношения, являющиеся эквивалентом простых высказываний к которым всегда может быть сведено высказывание сложное (высказывание на «естественном» языке) **без существенной потери того, что называют его проблемной семантикой**, представляются в качестве синтагматических отношений (отношений целостных смысловых единиц) ПОЯ. Для отношений устанавливаются формально-логические и эмпирические свойства, определяются операции, лежащие в основе тождественных преобразований элементарных высказываний, содержащих те или иные отношения, изучаются правильно составленные простые синтагмы. На первом этапе это может быть даже язык предикатов первого порядка, но, если не гнаться за формально-логическими построениями, а обратить внимание на непрерывно изменяющийся набор «эвристик», которыми пользуется «естественный» язык, то задача такого рода преобразования языков представляется весьма интересной.

рой. Рассмотрим рис. 1. Любая структурно-свободная (не проектируемая, самоструктурирующаяся) база состоит как минимум из двух частей – записей фактов (данных) и отношений между ними (знания, по определению знания). Отношение (реляция R) всегда представляет собой «знание». Совокупность возможных в данной области отношений образует базу знания (БЗ). Соответственно левая и правая часть предложения относительно R – «данные», соединенные конкретным именованным отношением. Их совокупность образует базу данных (БД). На каждом последующем шаге общения могут изменяться не только данные, но и отношение между ними. Все имеющиеся тексты, записи и прочие сведения, составляющие предмет хранения, могут быть занесены в такую базу без существенной потери их семантического содержания (о соотношении семантики и информации см. в следующем разделе) в виде ARB или в соответствии с рис. 1:

«Запись $A \rightarrow R_n \rightarrow$ Запись B ».

Ясно, что любые новые сообщения (в том числе и запросы) на входе структурно-свободной базы также можно рассматривать на основе их предикатного варианта построения. Тогда каждое новое входное сообщение фактически несет в себе изменение данных и/или этого отношения, что и является командой актуализации.

Некоторые дополнительные подробности будут рассмотрены в последующих разделах, а сейчас нам необходимо для перехода от концепции структурно-свободной базы к хранилищу рассмотреть понятие так называемого «информационного обмена», без точного уяснения сущности которого невозможно говорить о структурировании базы под потоком входных сообщений.

4. Место и смысл информации в аналитической системе

К сожалению, в науке имеются существенные разночтения не только понятий «знание» и «интеллект», но даже и определения понятия «информация» и «информационный обмен». Не будем забывать, что основатели того, что называют теперь «теорией информации» информацией как таковой не занимались. Напомним, что изначально были созданы «статистическая теория количества информации» и «математическая теория связи», не предусматривавшие собственного определения информации, но только ее количества, и впоследствии названные просто «теорией информации», имеющей косвенное отношение к повсеместно употребляемому термину «информация».

Это повлекло множество неприятных последствий (включая утверждения ряда ученых, что информация является понятием неопределяемым), из которых здесь нас напрямую затрагивает понятие «информационного обмена». Сегодня непредубежденному специалисту достаточно ясно, что никакого информационного обмена между источником и получателем сообщения или сигнала нет и быть не может. Есть сигнал или сообщение (как структурированный сигнал), знаковая посылка, на основе которой в приемном техническом аппарате по заложенному в нем алгоритму, а в человеке — по результатам контекстного анализа, вырабатывается нечто новое, не содержавшееся в посылке ни в явном, ни в закодированном виде⁵.

⁵ Не надо путать получение на вход системы данных (даже с отношениями между ними) с получением информации. Интерпретация входного сигнала приемником-человеком, тем и отличается от интерпретации сигнала приемником-машиной, что первый сначала рассматривает

Именно это новое порождение, которое существует только в контексте множества возможных интерпретаций входного сообщения имеет смысл называть информацией. Конечно, для аналитики главное не как это называть, а принципиальная необходимость работать именно с этим феноменом.

В [3] показано, что входной сигнал является не более чем программой запуска обрабатывающей «машины выработки информации» — контекстно-зависимой «неисчислительной» интерпретации сигнала в мозгу человека. Никакой «объективно содержащейся в сигнале семантики» без интерпретации сигнала в приемнике не существует. Соответственно в компьютере с заложенным алгоритмом интерпретации сообщения никакой информации, ничего нового не вырабатывается — существует только процесс актуализации априори предусмотренной алгоритмом трактовки полученного и не более того.

При постановке аналитической задачи мы просто по определению аналитики, ставим задачу получения нового, информации в ее нормальном, указанном выше понимании. Однако, по факту применения традиционной, спроектированной структуры базы и фон Неймановской машины, можем получить только алгоритмическую интерпретацию входного сообщения. *Аналитика и алго-*

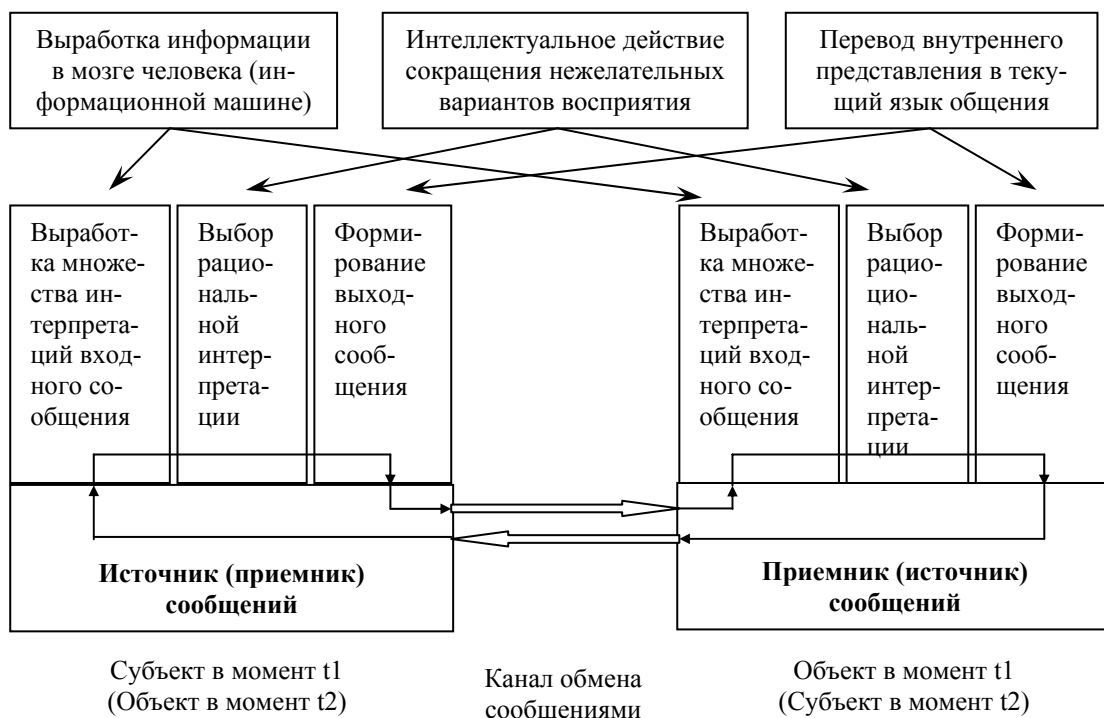


Рис.2. К проблеме определения места и смысла информации как явления

ритмы интерпретации несовместимы.

В любой стандартной базе для поиска того, что на самом деле является не информацией, а фактографическими сведениями, нам требуется все то же самое — формирование поискового образа как эквивалента входного сообщения для заложенного алгоритма его интерпретации. Шаг влево или вправо от точного формирования поискового образа ведет к непредсказуемым последствиям, а

полученное в том или ином контексте (*вырабатывая их в заранее не ограниченном количестве*), после чего может принять или не принять присланное в структуру своей памяти, а машина обязана сделать это в соответствии с заранее заложенным алгоритмом ее поведения — трактовкой сигнала в одном из *заранее предписанных* контекстов.

ведь чаще всего мы не можем точно сформулировать, что же мы ищем. Аналитик ищет непредусмотренное, случайное, неожиданное⁶. Предусмотреть все остальное — дело соответствующих служб. Как известно, постановка аналитического вопроса — уже более половины ответа на него.

Все существующие решения по хранению данных в виде, пригодном для их последующего нахождения и использования, для аналитических задач абсолютно непригодны и пригодны не будут, ибо так мы не решим задачу *творения информации* в машине по входным сообщениям. И конечно, нельзя забывать при этом обратную задачу — выдачи вразумительного для человека сообщения по результатам этого внутримашинного процесса (рис. 2).

Кстати, процедуру выбора приемлемого варианта интерпретации сообщения из всего выработанного «информационного массива» надо называть интеллектом, по определению интеллекта в его общечеловеческом понимании рационального мышления, т.е. безотносительно его технической или биологической реализации и сентенций о том, что это, мол, образная метафора. Понятие интеллекта обычно искажают до крайности и используют в основном для повышения рыночной стоимости разработки. Для нас указанный выше факт важен именно потому, что аналитическая машина может быть только интеллектуальной или никакой — не по причине желания назвать ее красиво, но по сути поставленной перед ней задачи.

Все сказанное укладывается в иллюстративную схему, приведенную на рис. 2, дополнительный комментарий к которой уже не требуется, разве что комментарий к известным понятиям субъекта и объекта мы дадим чуть ниже. Именно так мы формируем постановку задачи организации информационного обмена через получение и посылки сообщений, межбазового взаимодействия и выработки информации по множеству контекстов интерпретации сообщений, без чего говорить об аналитических работах просто невозможно.

Теперь вполне естественно утверждать, что хранилище *информации*, отнюдь не то, что объявил Бил Инмон, и наши программисты с удовольствием повторяют, а, если такой термин вообще приемлем, то суть *множество структурно-свободных баз, способных порождать информацию* под воздействием входного сообщения в контексте той или иной задачи его интерпретации.

Вот на такой основе действительно можно непротиворечиво двигаться вперед и в область технических решений, и в область организации аналитических работ с использованием технических средств сегодняшней (чисто вычислительной) и завтрашней (информационно-динамической) архитектуры.

5. Организация межбазового взаимодействия в хранилищах данных

Можно для простоты временно согласиться, что все сказанное выше относительно представления знания, данных и предикатного ПОЯ является правилами создания и непрерывной реструктуризации соответствующих реляционных баз, реализующих в процессе обмена сообщения общезвестные понятия «субъекта» и «объекта», без которых не может существовать ни информация, ни интеллект как атрибут системы [4]. Рассмотрим эту процедуру (рис. 3).

⁶ Это «самое краткое задание» на работу аналитической системы. Либо наша система может научиться это делать, либо она не аналитическая.

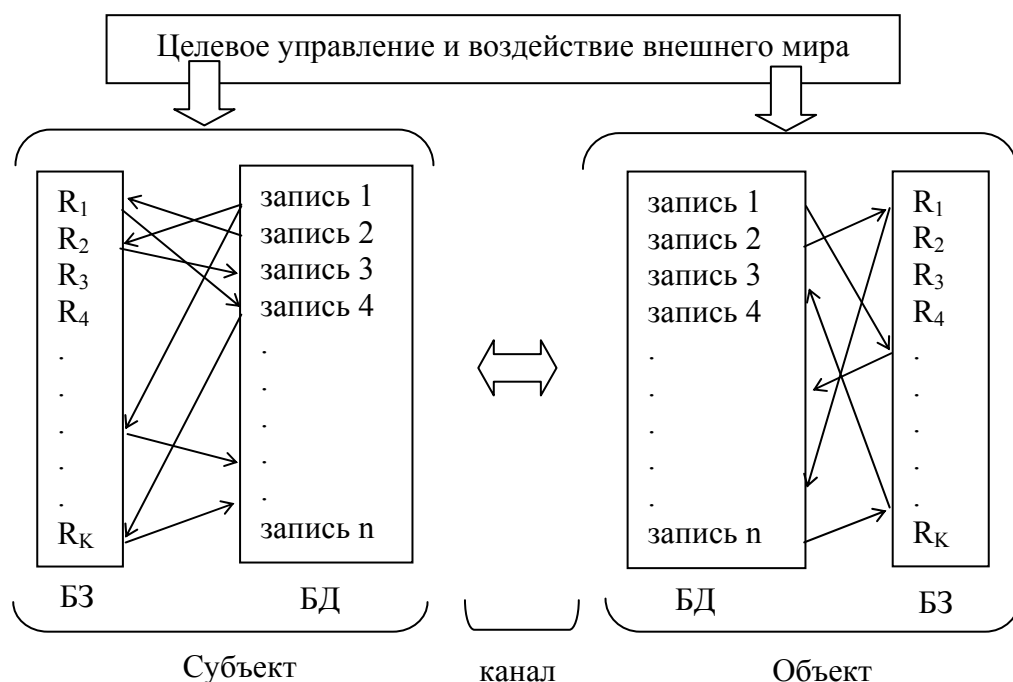


Рис. 3. Организация субъектно-объектного взаимодействия

Предположим, что в левой части рис. 3 для некоторого конкретного момента диалога существует *субъект* в составе характеризующих его внутренних базы данных (совокупность всех сформированных данных «запись 1 — запись n») и базы знания (совокупность разрешенных отношений « R_1 – R_k »), а в правой части — *объект* аналогичной структуры.

Напомним, что такая схема является схемой интеллектуального управления, ибо субъектно-объектное взаимодействие характерно для представления процессов обмена сообщениями и в биологических системах. Смысл управления заключается в передаче сообщения (управляющего сигнала) от активного источника (субъекта) к пассивному (объекту), выработке в последнем реакции на управление как ответного сообщения, после чего, в момент отправки выработанного ответного сообщения субъект и объект меняются ролями⁷.

Достаточно понятно, что эта схема является даже не моделью, а прямой инструкцией нормального общения и, в нашем случае, полностью соответствует понятию организации аналитического исследования в режиме не столько вопросно-ответном, сколько «совместных размышлений и совместной выработки согласованной точки зрения», а это уже настоящая аналитика.

Конечно, если некоторое сообщение в точности совпадает с уже имеющимися в другой части записями, то, в соответствии с «классическим» пониманием информации, оно не несет разнообразия и никакой реструктуризации не производится. Но это только внешне похоже на подход Эшби, ибо первичным контекстом здесь является сообщение о том, что реально известно «собеседнику» или источнику сообщений. Зачастую именно это и является основной целью общения.

⁷ Такого рода «информационный мультивибратор» образуется в любом сочетании технических и биологических составляющих процесса. Отсюда же вытекает и понятие «информационного резонанса» (!), использование которого, как и любого резонансного явления, может являться основой многих новых построений.

Если же отношение и терминология в чем-то не совпадают с известными собеседнику записями (синтагмами, предикатами), то возникает необходимость в коррекции структуры его базы для изменения или введения нового отношения. Коррекция может быть выполнена безусловно, то есть без оценки степени достоверности полученного сообщения или после проведения анализа достоверности (приемлемости с точки зрения получателя) сообщения.

Здесь все происходит и может происходить только так же, как и в общении людей — просто нет других возможностей оценки сообщения в процессе беседы, кроме опоры на некоторые уже имеющиеся сведения о предмете сообщения и априорную оценку надежности источника сообщения. Фактически это «реакция через накопленный образ» — управление на основе информационной⁸ [1, 4], а не сигнальной обратной связи.

Уже здесь становится ясно, что защита информации реализуется настройкой аналитической системы на каждого конкретного ее оператора, на его понимание контекстов сообщений и предыстории *своего собственного* существования. Подробнее об этом будем говорить далее, а пока продолжим тему субъектно-объектных отношений.

Так как общение субъекта в разное время происходит с разными объектами, то становится ясно, что речь идет об особенностях организации субъектно-объектного взаимодействия, о ситуации, когда внутренняя структура каждого субъекта (объекта), какова бы она ни была в начальный момент, всегда отличается от структуры любого другого субъекта (объекта). Более того, в каждом отдельном общении вся структура недоступна, она проявляется для взаимодействия только в тех объемах и контекстах, которые сложились в предыдущих актах общения.

Мы получаем некоторое построение, обладающее индивидуальностью, которую оно не теряет ни в каком случае, в точности (на уровне возможностей представления естественного языка с помощью ПОЯ) повторяющее организационное построение взаимодействия естественных (биологических) систем.

Отметим, что исследование правил преобразования бинарных отношений данных в набор синтагм является ключевым моментом разработки ПОЯ системы. В литературе все обычно сводится к расширению имеющегося в реляционной базе набора синтагм по заранее известным основным (формально-логическим) и (в самом «продвинутом» случае) выделенным дополнительным (эвристическим) правилам логического вывода.

Конечно, если ограничиться сказанным, то все это не более чем алгоритмическое решение и инженерные ухищрения вокруг реляционных баз, как мы и согласились в начале этого раздела. Но в процессе исследований выяснилось, что надежды на достаточность разового выделения «чисто математических» свойств отношений и эмпирик логически непродуктивны. Живые системы аналитического класса «живут в изменяющейся логике», логике, правила применения и изменения которой существенно зависят от взаимодействия имеющейся структуры базы и структуры входного сообщения. Поэтому здесь необходимо говорить о разработке «логики с изменяющимися в процессе общения отношениями» и «ситуационном изменении формально-логических и эмпирических правил построения (расширения имеющегося набора) синтагм».

⁸ Понятие информационной обратной связи совершенно невозможно ввести при традиционном количественном понимании информации, но естественно, инженерно понятно и совершенно необходимо существует для работ в рассматриваемой постановке.

Возвращаясь к высказанному в начале раздела предположению, отметим, что это уже не реляционный подход и не алгоритмический по определению алгоритма. Это прямое указание на практическую значимость ряда работ связанных с понятием невычислимости и в чисто математическом смысле, и в лингвистическом, и в биологическом [5]. Подробности мы оставим пока для структурных лингвистов. Важно, что наше построение для будущей информационной машины не замкнулось аппаратом математической логики, не стало формальным аппаратом для порождения синтагм, но, как это и положено открытой⁹ системе, осталось под воздействием входного потока сообщений.

6. Проблема доверия источнику сообщения

В принципе, проблема доверия источника и приемника информации друг другу существует всегда. Изначально она трактовалась как проблема помехоустойчивости или помехозащищенности. В нашем подходе этот этап дополнен с одной стороны необходимостью контекстного анализа сообщений — «в каком смысле сказано», какая программа интерпретации подразумевалась отправителем для конкретной знаковой посылки. С другой стороны необходима проверка сообщения на факт добросовестного заблуждения отправителя или даже на факт преднамеренного введения приемника в заблуждение.

Такого рода проблемы не являются надуманными дополнениями или проблемами только защиты, это нормальные проблемы общения высокоразвитых систем, как технического, так и биологического уровня, повседневно и многократно решаемые человеком. Аналитическая система сколько-нибудь высокого уровня начинается именно с такого понимания проблем двухстороннего субъектно-объектного общения.

В открытой системе с внешним потребителем сообщений (человеком, например) критерием правильности логических выводов и необходимости их текущих изменений является, естественно, согласие или несогласие потребителя информации с утверждениями, полученными в системе путем логического вывода.

В полностью автоматической системе такого рода критерием может являться только базовый набор правил проблемной области, нарушение которых по какой-либо причине недопустимо, а все остальные выводы эквивалентны формированию «собственной точки зрения» интеллектуальной системы.

Последнее подтверждает, что в некотором, пусть еще и достаточно примитивном смысле, мы начинаем иметь дело с системой, проявляющей свою индивидуальность, имеющей свой взгляд на представление проблемной области и реализующей упомянутое выше понятие *невычислимости* в цепи, указанной на рис. 2.

Обратите внимание — мы не изобретали новый контекстно-зависимый язык, использование которого строго необходимо для аналитической системы, но только преобразовали представление существующего естественного языка и

⁹ К сожалению, стандарты IEEE, на которые опираются многие наши специалисты, дают чисто инженерное определение открытой системы, как системы, обладающей возможностями, облегчающими реализацию приложений из различных платформ от многих поставщиков и обеспечивают взаимодействие с другими системными приложениями и с пользователем. Безусловно, это определение интегрируемых систем, а не открытых. Понятие открытости было резервировано общей теорией систем по фон Берталанди для систем реального мира в отличие от систем замкнутых, модельных, от внешнего мира изолированных.

нормального диалога с его использованием. За счет этого мы *впервые получили обоснованное терминологическое право говорить об интеллектуальной информационной поисковой (аналитической) системе*, с пониманием информации как процедуры выработки неограниченного множества контекстов восприятия входного сообщения и интеллекта — как процесса рационального выбора в этом множестве.

Естественно, что право на ошибку всегда останется за системой, а наше право не принимать новые выводы системы останется за нами. Не объемы данных, а устанавливающиеся под разными влияниями отношения между ними предопределяют выводы, правильности которых мы можем доверять.

«Абсолютной правильности» вывода не бывает, ибо в системах с самостоятельным поведением правильность определяется только достигнутыми соглашениями в ходе диалогового обмена сообщениями. Абсолютная правильность является атрибутом модели, в соглашениях которой можно терминологически и на уровне формального аппарата договориться о понимании правильного. Но что тогда будет считаться правильным? Только заранее предписанная истина. Вряд ли это может являться результатом аналитического поиска решений.

Но только ли «словесным диалогом» с аналитиком ограничивается предлагаемое построение? Конечно, нет. При известном уровне надежности собственных решений это самостоятельное автоматическое управляющее устройство, как этого и следовало ожидать от действительно интеллектуальной системы.

Укажем, что существование интеллектуальных (т.е. принимающих аналитические, а значит, управляющие решения) систем должно обеспечиваться не менее чем тремя структурами, каждая из которых аналогична конструкции, представленной на рис. 3. В сложных системах существуют как минимум:

- управление для поддержания собственной стабильности системы как некоторый процесс «внутреннего переосмысления», внутренней реструктуризации, изменения целевых установок в процессе анализа изменения ее внутренней структуры под влиянием входных сообщений;
- управление для компенсации возмущений от внешнего мира;
- внешнее управление, как целенаправленное воздействие некоторой другой системы, желающей в данный момент выступить как субъект.

Все эти управления реализуются однотипно на основе рассмотренной процедуры согласования связей соответствующих им данных. В реальности нельзя представить себе систему, получающую непротиворечивые сведения от внешнего мира и составляющих ее субъектов и объектов. Но мы обязаны помнить, что ничего, кроме этих сведений система получать не может. Значит, будем учитывать при конструировании, что *все указания об изменении «правил мышления» содержатся в потоке входных сообщений*, в предикатном ПОЯ, раз уж этот язык выбран нами, исходя из его «фон Неймановской реализуемости».

Приемлемость для сложной системы использования противоречивых фактов до некоторой степени дает возможность обеспечить начальное функционирование рассмотренной структуры и инженерное решение построения необходимой нам «изменяющейся формальной логики», отвечающее процедуре переосмысления, характерной для действительно интеллектуальных систем.

Некоторые подробности таких работ можно посмотреть в [2, 4], в работах других авторов по проблемам представления ПОЯ. В целом они сводятся к весьма долгому и достаточно плохо автоматизируемому процессу «обучения

системы требуемой логике выводов» (см., например, [7]). Здесь речь идет о повторении на уровне машины процесса обучения реального специалиста «всему нужному» и «всему не нужному», причем последнее, как выяснилось, имеет важнейшее значение при интерпретации сообщений, имеющих отклонения в своей формулировке и тезаурусе от чистого ПОЯ. Нельзя достигнуть свободы общения с машиной, не понимающей некоторого набора общепринятых «околоспециальных» комментариев.

7. На пути к созданию контекстных аналитических машин

Все известные на сегодняшний день стратегии поиска и обработки информации, несмотря на внешнее разнообразие программных и технических решений, имеют в себе один главный и неустранимый недостаток — все они предполагают, что проблемная область задается некоторой информационной базой, на создание и обслуживание которой тратится основное время обслуживающего персонала. Неявно используется допущение, что семантика и вообще все, что можно вытащить из базы, заложено в нее накоплением данных (записей), и главное — «уметь ловить в этом омуте», искать семантику в морфологии. Самые смелые предложения заключаются в дополнительном анализе системы запросов, с той же целью «выявления семантики», но уже семантики вопроса к системе. Такие подходы суть все то же восприятие информации как сигнала, как знаковой системы, как семиотики, вообразившей себя наукой о семантике.

Мы уже договорились выше о пересмотре самой сущности информации как предмета исследования для практических целей. Рассмотрим некоторые возникающие при этом проблемы организации семантического поиска. Нас будет интересовать обработка связных текстов и сообщений на ПОЯ с извлечением из них всего, что может быть полезно не вообще, а *в каждом конкретном случае* — в пределе: «от чтения между строк, до анализа качества печати отдельных символов».

В приведенных выше словах «*в каждом конкретном случае*» заключается основной тезис предложенной парадигмы информации, нового восприятия смысла ее обработки: *контекст, система взглядов, точка зрения, накопленный образ ситуации и прочая, и прочая первичны относительно любого входящего сообщения, любой принятой знаковой системы*. Необходимо ставить задачу создания смыслового окружения входного сообщения, задачу обеспечения вариантов контекстного восприятия, как задачу первичную.

Строго говоря, ничего нового относительно Природы и человека здесь не выдвигается: чтобы обработать сообщения, лучший в мире приемник сообщений — человек долгие годы специально подготавливается, накапливая горы сведений, которые в частности никогда и не понадобятся, но в целом дадут возможность понять и проанализировать почти все из любых связных текстов. Ну а сколько раз человек поймет при этом «все наоборот» и говорить не придется.

Во всех сколько-нибудь серьезных поисковых системах, претендующих на работу с семантикой, в том или ином виде ставится задача формирования внешнего мира как задача накопления так называемого «знания» — сведений обобщающего характера типа словаря родовидовых отношений, синонимов и эвристик. Это позволяет искать «информацию» путем написания алгоритмов, внешне проявляющих свое «знание», типа того, что если запрос идет о постав-

щиках мебели, то поставщики стульев и столов также представляют интерес, а поставщики стенок — только в том случае, если есть дополнительные сведения, что это за стенки.

Достаточно ясно почему такие системы не получают значительного пространства и быстро «умирают» уже на уровне полумодельного существования, хотя и содержат родовидовые отношения, требуемые и нами для ПОЯ. Стандартным вариантом работы таких «продвинутых» поисковых систем является переключивание формирования всего прочего «избирающего материала» на формулировку запроса.

От нищеты иных возможностей или просто начального интереса пользователь обычно включается в эту игру, но не надолго. По каким-то своим интуитивным понятиям он быстро формирует круг своих интересов и далее обходится прямой адресацией или просто косвенной информацией из других источников. Для поддержания его интереса требуется непрерывная подстройка системы под все новые варианты запросов, что ведет к превалирующему росту затрат по сопровождению над реальной пользой достижимых результатов.

Как бы ни называли такие системы, все они не более чем фактографические, примитивно подтверждающие факт существования или не существования некоторых записей, упорядочиваемых поисковой системой по весовому критерию целесообразности их просмотра. Конечно, от системы кибернетического уровня, всегда работающей на контекстно-свободном языке, ничего другого ждать и не приходится. Язык же другого уровня компьютеру без специальных ухищрений не доступен, и ждать чего-либо другого от систем, ныне объявляемых «интеллектуальными», не приходится.

Сформируем совокупность задач, без совместного решения которых какое-либо движение в сторону создания семантических систем невозможно, если под семантикой понимать именно то, что она представляет собой в задаче семантического поиска. Речь идет о том, что в реалии на «практическую семантику (мы имеем здесь ввиду не прагматику, но инженерию решений)», интуицию, априорное знание собеседника и предмета разговора человек переключает основную задачу организации работы с информацией — увязывание полученных сведений с собственным текущим пониманием ситуации и выяснение правильности понимания своих вопросов собеседником.

Итак, мы констатируем достаточно очевидные утверждения, без учета которых ни о какой семантике общающейся с человеком информационной системы не может быть и речи:

- человек не в состоянии четко (без «подразумеваемого») сформулировать все свои текущие интересы;
- человек не в состоянии сформулировать все ограничения на свои интересы;
- человек выражает свои мысли на том подмножестве естественного языка (проблемно-ориентированного языка — ПОЯ), которое ему привычно и которым он лично владеет;
- разные люди называют одинаковые вещи по-разному и разные вещи одинаково, строя при этом собственные деревья родовидовых отношений;
- требуют учета и другие ситуации, связанные, например, с неформализуемыми оценками событий, что связано в основном не с вероятностными оценками и информированностью, но с принципом действия информодинамической машины или мозга [1].

Все эти особенности диалоговой обработки информации и должны возлагаться на «практическую семантику» — обеспечение возможности самостоятельного «доконструирования» системой полученного сообщения.

8. О «не формализме» языка системы семантического уровня

Необходимо прекратить попытки создания информационных систем на базе контекстно-независимых языков. К этой области мы должны сразу отнести и все работы, связанные с формированием формальных (по определению формальности) логик, ибо первое, что надо сделать, чтобы убить семантику: создать «некроинтеллект» — построить универсальные правила производства выводов.

Семантическая составляющая восприятия текста не спрятана в сигнале-сообщении, а присутствует в потребителе сообщения, причем отправитель, неважно осознанно или нет, всегда ориентируется на эту составляющую, считая ее хотя бы частично совпадающей со своей. С нашей точки зрения новым здесь является не само это утверждение, но то, что все это уже пора учитывать и уже практически можно учитывать при создании аналитических систем.

Механизм создания логик производства выводов, изменяющихся под потоком входных запросов, конечно гораздо сложнее выработки очередного варианта исчисления, но потери при априорной формализации в точности равны тому, что мы перечислили, когда говорили о семантике, возлагаемой человеком «по умолчанию» на поисковую систему.

Изменение логики вывода под потоком входных запросов рассматривалось нами еще в [2], применительно к фактографическому поиску, но для семантических систем оно имеет дополнительные особенности. Логика изменяющихся формально-эмпирических отношений (например, предикатного типа) суть логика, порожденная процессом познания, существующая в зависимости от познанных нами особенностей предметной области, т.е. зависящая больше от глубины познания предмета описания, чем от объемов терминов конкретного ПОЯ.

Исследования наводят на мысль, что механизм такого рода должен быть одинаков для всех языков и всех отраслей науки. В подтверждение этой гипотезы можно привести тот факт, что, при любой терминологии ПОЯ, логические выводы выше третьего порядка (третьей «итерации») какой-либо достоверности практически не имеют.

9. О множественности вариантов контекстного восприятия

Необходимо перейти от попыток формирования проблемной области из фактографических данных и запросов к ним к широкому использованию тематики области как таковой, в первую очередь, используя для ее формирования аналитические обзоры, отчеты и публикации оценочного и методологического характера. Именно круг упомянутых документов является «задатчиком» реальных отношений между терминами и лексическими единицами проблемной области. Именно такие документы выражают логику мышления, которую пользователь хотел бы иметь «внутри» семантической контекстной поисковой машины.

Поясним основные моменты, подлежащие учету при постановке работ.

Во-первых, нельзя забывать, что «абсолютная» истина, являющаяся результатом формальной логики, в реальной жизни места практически не имеет. Это означает, что наше требование перехода к логике изменяющихся отношений требует постоянного обучения поисковой машины, указанию, какие из ее выводов сегодня считаются правильными, и соответственно, какие из правил вывода приходится считать в текущий момент правильными. Обучение тут надо понимать правильно — не традиционное «обучение на тестовых примерах», а постоянный индивидуальный контакт с каждым ее пользователем, индивидуальное участие во всех их аналитических работах, настройка под конкретный «образ мышления». Иначе, это будет не «аналитическая машина для меня», а некто другой, чьи выводы придется учитывать критически, сообщениям которой мы будем доверять не более чем всем другим сообщениям.

Во-вторых, как известно, логика изменяющихся отношений состоит из некоторого подмножества разрешенных формально-логических отношений, выполняющихся на данном отрезке времени в данной проблемной области и дополняющих ее эмпирически установленных отношений, фактически являющихся «индивидуальным образом мышления конкретного потребителя информации».

Проблема выявления такого рода «добавок» к разрешенному подмножеству формальных свойств является одной из сложнейших при создании семантических систем. При всем желании установить здесь единственный набор разрешенных отношений этого сделать не удастся. Система просто погибнет через некоторое время, отстав от изменяющихся «требований к процессу ее мышления».

Еще более нереальной явится обычно желаемая постановка задачи, связанная с поиском «оптимальной» системы логических выводов. Систему можно «вырастить» по образцу лучших аналитиков, заставить работать быстрее них, но лучше — никак¹⁰, вполне достаточно, что быстрее, без усталости и грубых логических ошибок. Иначе это будет уже совсем другая задача, задача не поисково-аналитическая, но задача перспективного планирования, с созданием очередной модели под задачу оптимизации.

Следовательно, семантический поиск, может быть обеспечен только при индивидуальной настройке системы на текущую логику пользователя в непрерывном контакте с ним. Кроме того, это означает, что с разными пользователями система будет общаться по-разному. Все как в реальной жизни, когда с собеседниками с разным образом мыслей нам приходится говорить об одном и том же разными словами и в разной логике (или останавливаться перед фактом невозможности достижения взаимопонимания).

Вот мы и подошли к понятию исключительности доступа к такого рода аналитическим системам. Достаточно ясно, что оператор такой системы войдет с ней в лучший контакт только тогда, когда будет «обучаться вместе с ней» буквально с поступления в некоторое учебное заведение. Их общение, понимае-

¹⁰ Это в шахматах можно найти способ играть лучше всех — правила не изменяются, к интеллектуальным способностям задача прямого отношения не имеет. Она только иллюстрирует достаточно простой факт, что есть ситуации, когда неисчислительный механизм работает быстрее вычислительного. Однако конечный проигрыш интеллектуального человека неинтеллектуальной машине обеспечен всегда, ибо задача алгоритмическая. В аналитической работе правил нет — есть только открытия как новые аналогии и нетрадиционные выводы, побеждает не лучший алгоритмист, а лучший «творец рациональных выборов из информационных контекстов».

мое как индивидуальное совпадение тезаурусов и трактовки свойств отношений (производства логических выводов), должно быть постоянно и логика мышления подготавливаемого аналитика будет понятна системе как задание на именно такие требования для анализа сообщений, которые *подразумевает* аналитик.

Это будет «последний вариант» подчинения аналитической машины человеку. При всех других ситуациях логика принятия решения будет заимствованной или выработанной самостоятельно как «смесь способов мышления» разных операторов. Хорошо это или плохо — проблема выходит за рамки статьи. Остаются только два замечания.

Во-первых, без дополнительных комментариев ясно, что мы без затрат на обеспечения безопасности хранимого набора записей получаем *индивидуализацию доступа для определенного класса специально подготовленных людей*. Подробности и степень доступности информации можно варьировать, но когда защитным механизмом становится образ мышления и предыстория общения с машиной «механистическая» защита доступа становится абсолютной.

Во-вторых, фон Неймановская архитектура машин для такого рода задач в обозримом будущем заменится архитектурой другого порядка — безадресной, обеспечивающей самостоятельность аналитического мышления искусственного устройства [1]. Надо ли стремиться к такой аналитике или бежать от нее?

10. О контекстных машинах

Перечислим три предложенных нами подхода к организации работы с контекстно-зависимой информацией, описанные в [1].

Во-первых, это выведение фон Неймановского автомата в переменный режим «структурирование информации — принятие решений». Назовем такой режим псевдоинтеллектуальным, режимом, внешне аналогичным «вспышкам сознания» по Р. Пенроузу [5]. Фактически он описан выше и является просто трудоемким, но необходимым вариантом перехода к существенному повышению того, что сейчас называют «вычислительными мощностями», хотя сам переход будет связан именно с созданием *невыхисляющих аналитических механизмов*.

Во-вторых, упомянем об альтернативном решении, предложенном нами в отличие от термина «машина баз данных» под названием «машина баз знания» [1]. Ее сущностью является параллельная обработка отношений, заданных предикатами, в поле однородных ячеек, программой для которых являются те самые «правила текущей логики».

Для этого нам надо иметь достаточно гибкую структуру, на которой можно эффективно и безболезненно вести реструктуризацию, обеспечивающую реализацию поведения (принятия решения), не сводимого к конечной модели.

Исходя из сказанного, рассмотрим следующую архитектуру МБЗ (рис. 4), в которой машина состоит из необходимого числа n отдельных модулей, на вход каждого из которых после преобразования входной текстовой посылки в язык предикатов поступает тройка вида $(ARB)_i$, $i \in 1 \dots n$ для ее последующего непосредственного раскрытия по тем или иным законам логики.

В связи с тем, что реальное число законов преобразования невелико, здесь эффективно использование динамически образующейся матрицы вида, приведенного на рис. 4. Строки этой матрицы соответствуют не только порядку

преобразования входных предикатов, но и составляют таблицу всех формально-логических отношений, использованных в текущей информационной посылке. К каждой следующей операции приема входной информации машина переходит в состояние готовности принять входные сигналы на отдельные входы, для чего в процессе поступления сигнала формируется очередная матрица.

ARB ₁	Пра- вила пре- обра- зова- ния	ARB ₁₁	ARB ₁₂	ARB ₁₃	ARB ₁₄	ARB ₁₅	ARB ₁₆	ARB ₁₇
ARB ₂		ARB ₂₁	ARB ₂₂	ARB ₂₃	ARB ₂₄	ARB ₂₅	ARB ₂₆	
ARB ₃		ARB ₃₁	ARB ₃₂	ARB ₃₃	ARB ₃₄	ARB ₃₅	ARB ₃₆	ARB ₃₇
ARB ₄		ARB ₄₁	ARB ₄₂	ARB ₄₃	ARB ₄₄			
ARB ₅		ARB ₅₁	ARB ₅₂	ARB ₅₃	ARB ₅₄	ARB ₅₅	ARB ₅₆	ARB ₅₇
...		...						
ARB _n		ARB _{n1}	ARB _{n2}	ARB _{n3}	ARB _{n4}	ARB _{n5}	...	

Рис. 4. Архитектура машин баз знания
(реализация на основе языка исчисления предикатов)

Выходным сообщением каждого модуля является набор расширяющих входной текст предикатов вида $ARB_{i,l}$, $i \in 1 \dots n$, $l \in 1 \dots k$, содержащий в себе все возможные логические преобразования исходного предиката. В связи с тем, что этот процесс существенно параллелен, возможна организация каждого отдельного модуля как параллельной машины со своим выходом.

Несколько входов полезно предусматривать и для получения выводов «второго порядка» — выводов следующих из совокупной оценки нескольких различных входных предикатов. Ясно, что тем самым модули получают полную универсальность и могут по мере окончания своей работы на первом уровне преобразования использоваться на втором, а если надо – то и последующих уровнях.

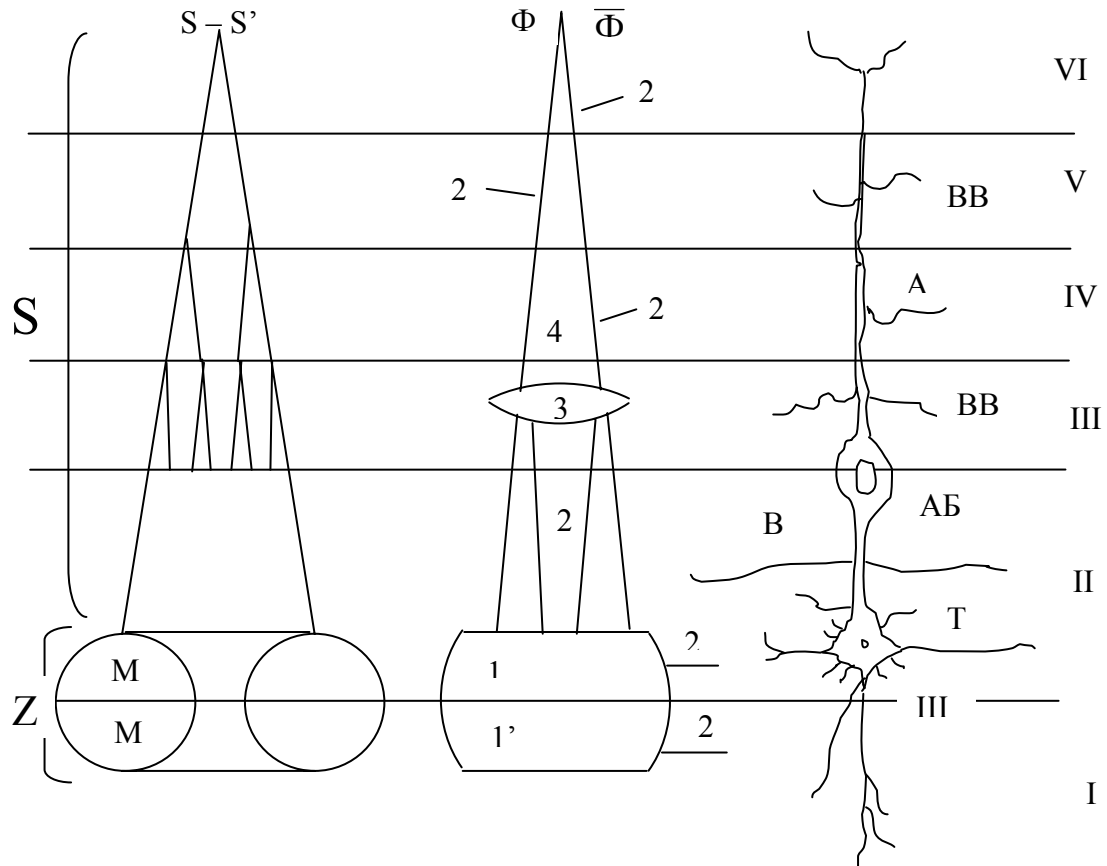
Сравним возможности машины фон Неймановской архитектуры и МБЗ применительно к предикатному варианту механизма обработки контекстно-зависимого языка.

В первой мы имеем классический вариант архитектуры составных частей, ориентированный на программно-алгоритмическую организацию процесса работы с данными. Это существенно последовательная машина с известными проблемами отладки и конечной вычислимости процесса. В МБЗ входная информация является и управлением и, одновременно, командой для организации действий по ее обработке. Это достаточно естественно, ибо речь идет о входном языке контекстно-зависимого уровня. При желании говорить в терминологии алгоритмов здесь можно сказать, что алгоритм работы задан в обрабатываемой контекстно-зависимой записи.

Отметим, что МБЗ внешне представляющаяся конечным автоматом, в каждый конкретный момент времени изменяет свою структуру под воздействием входного потока информации, реализуя тем самым случай моделирования в частотном смысле нерекурсивных объектов при ситуации «несводимости к конечной модели, но конечной моделируемости в частотном смысле».

Этот достаточно простой по звучанию тезис заключает в себе на самом деле целый спектр фундаментальных положений. По существу это неявно высказанное положение о существовании «инженерного» или «механистического» интеллекта, некоторого «псевдоразума», способного не только обучаться, но и

«обучаться обучению». На самом деле пока именно это мы и хотим, ставя задачу создания МБЗ, а остальные свойства интеллекта считаем «излишними» для такой системы.



Действительно, как-то неуютно оказаться в зависимости от некоторой машины, которая сильно «облегчает нам жизнь», но в обмен на это мы будем вынуждены считаться с тем, что она имеет «свою собственную трактовку» полученных ей данных, «свое понимание предмета». Иначе говоря, *создавая МБЗ мы желаем создать нечто, способное к пониманию, но хотим априори избавиться себя от обязанности взаимопонимания.*

а) абстрактная машина б) оптическая машина в) нейрон

а) абстрактная машина: Z — зона «зеркала», трехмерная память для двух трехслойных структур данных; S — структура (метаструктура), S, S' — «выходы».

б) оптическая машина: 1, 1' — трехслойная оптическая (голографическая) память для «запомненной» и «воспринимаемой» структур данных; 2 — пучки «малых» оптических волокон; 3 — «сумматор» волновых фронтов; 4 — «большое волокно» — формирователь структуры — решения; $\Phi, \bar{\Phi}$ — «синфазный» и «парафазный» выходы.

в) нейрон: Т — тело; Ш — шипиковый механизм; В — волокна «обучающих» и «рабочих» входов; АБ — аксонный бугор; ВВ — волокна «верхних связей»; А — аксонное волокно.

Рис. 5. Сравнение вариантов «схемы» информационной машины

В-третьих, предложена концепция и практический подход к созданию новой, не фон Неймановской безадресной структуры, названной нами «Вертикальной машиной», в отличие от предыдущего случая «самостоятельно избавляющей нас от возможностей полного понимания ее действий» (рис. 5). По нашему мнению, на сегодняшний день именно ее структура сравнительно со

всеми прочими наиболее близка к контекстно-зависимому механизму обработки информации в «Природной аналитической машине» — мозге человека.

Кратко описать конструкцию такой машины здесь невозможно, поэтому приведем только иллюстративную схему (рис. 5) предполагаемой оптической реализации такой машины в сравнении с ее абстрактным (логическим) представлением и структурой реального «высокого нейрона», дав к ней самый минимальный комментарий.

Предполагается, что хранилищем данных является некоторая центральная память, построенная по типу голографической. В каждый момент времени над этой памятью «поднимаются» две структуры по типу бинарных деревьев, одна из которых соответствует всей хранимой совокупности данных, а другая — пришедшему запросу, новому сообщению и т.п. Деревья образуют из себя конусы, вложенные и вращающиеся друг в друге специально организованным образом. Каждый из возможных моментов хотя бы частичного совпадения листьев и узлов деревьев представляет собой некоторое «контекстное» осмысление ситуации. Далее адресуем читателя к рис. 2 настоящей публикации.

Конечно, мы отдаем себе отчет в том, что сказанное здесь является не более чем попыткой заинтересовать читателя для ознакомления с подробностями обратиться к материалам работы [1].

В связи с аналогией с высоким нейроном отметим еще и еще раз — с существующими нейронными сетями и нейрочипами здесь нет ничего общего. Однако дальнейшие подробности, функциональное назначение остальных типов нейронов и других компонентов мозга могут быть существенно полезны для практического создания вертикальной машины путем сопоставления логики работы систем в целом и топологии связей элементов.

Отметим, что при абсолютно жестком, одинаковом принципе действия «высокого нейрона» и унитарной общей схеме их соединения, количество входов каждого типа меняется от экземпляра к экземпляру более чем на порядок, то же можно сказать и про соотношение количества входов разного типа (уровня). То есть при абсолютной жесткости принципа устройства наблюдается большая свобода реализации, что говорит о превалировании принципа работы подобных систем над их частными конструкционными параметрами.

12. Заключение

Все изложенное выше является достаточным материалом для начала опытно-конструкторских работ по созданию вполне реализуемой системы контекстного, т.е. семантического анализа потока сообщений из некоторой связной области. Последнее утверждение отражает желательность более скромной начальной постановки «претензий к разуму системы», но отнюдь не сводит задачу к существенно меньшему объему работ.

Постановочно дело здесь совсем не в требовании сотен лет сбора в некоторую базу новой конструкции того, что ныне называют информацией — дело, прежде всего, в том, чтобы не попасть в постановке в болото кибернетического подхода, болото формальной логики и математического «моделирования информации» и ее хранилищ. Не склад информации надо разрабатывать, но процесс, *технология существования* совокупности «информационных» сообщений во всех смыслах и контекстах их возможного восприятия.

Кратко о том, что должно быть в реальной технической постановке работ. Приводим несколько очень непростых направлений практических работ, реализуемых только на уровне опытно-конструкторской работы:

- разработка родовидовых словарей весьма больших объемов;
- создание информационной базы не менее чем из двух частей — собственно сообщения входные и «описатели смысла интерпретации» — интерпретаторы «точки зрения вопрошающего» (с балансировкой на уровне «не потерять семантику — вложиться в возможности реального компьютера»);
- разработка языка уточнений запросов (на уровне контекстно-зависимого языка);
- дальнейшая разработка «вариационной» логики — логики, изменяющейся от потока входных сообщений (некоторое обобщающее исследование по «нетривиальным логикам»);
- выбор (или создание) операционной системы и надстроек над ней для обеспечения «самоструктурирования» информационных баз к логике информационных взаимосвязей с полным отказом от «тяжелых баз», из-за проектного подхода к структуре которых все информационные системы во всех странах неизбежно заканчивают свое существование. Надо ли объяснять, почему эта «смертность» происходит так неотвратимо?

О политическом, экономическом, военном и прочих смыслах создания, более того — необходимости проведения таких работ уже говорилось в начале статьи. Здесь можно только предложить вернуться к ее началу и, на основе полученных сведений («информации») еще раз оценить военно-экономическую или другую, аналогичную необходимость проектирования аналитических систем стратегического уровня.

Достаточно ясно, что все это явится необходимым условием перехода к созданию самой вертикальной машины. Еще раз подчеркнем ее «потребительские» особенности. Как известно, все новые решения появляются в связи с их грядущей необходимостью. *Необходимость вертикальной машины обусловлена неизбежным переходом от математического моделирования задач, принципиально не имеющих возможности интегро-дифференциального описания для представления их в сколько-нибудь открытом, связанном с изменяющейся обстановкой виде, к их логико-лингвистическому представлению, требующему, конечно, использования той логики, которая адекватна и ситуации в ее восприятии сознанием обработчика, и цели обработки получаемых сообщений.*

Поэтому дополнительно укажем: *предложение «предварительно сделать маленькую демонстрационную систему, иллюстрирующую возможности большой», вполне равноценно требованию построить «интеллектуальную» систему на персональном компьютере, или создать нейрон на «нейрончиках» — все это будет игрой в слова и профанацией проекта.* Ведь по какой-то причине никто не ставил задачу предварительного создания атомной бомбы с тротильным эквивалентом в двести грамм. А задачи по масштабу и степени возможного воздействия сравнимы и, скорее всего, реальные последствия успехов разработок типа «вертикальной машины» мы пока еще оценить не можем.

Все остальные проблемы такого порядка — дело новой науки, поэтому перейдем еще к одному стратегическому аспекту значения разработки такого рода не вообще, а в некоторой конкретной стране.

Достаточно ясно, что каждый разработчик таких систем будет делать их на своем языке. Даже если предположить доведение таких систем где-то до «коробочного варианта», то и тогда их заимствование будет упираться в два неприятных момента.

Во-первых, логика работы будет предусматривать особенности логических выводов на некотором определенном чужом языке. Только «физическая» структура машины и формально-логические построения не зависят от языка. Все говорит о том, что внутренняя структура неформального логического вывода существенно зависит от конкретной реализации контекстно-зависимого языка (этим ни в коей мере не утверждается, что результаты будут разные, разными будут пути вывода).

Во-вторых, системы такого рода требуют внутренние словари, прямой перевод которых на другой язык невозможен. В соответствии с классическими понятиями государственности, это прямой путь к получению мирового приоритета (для стран языка разработчика) и его потери (в остальных странах).

Являются ли такие тенденции интеграционными или разрушительными пусть думают те, кому об этом ведать надлежит. Но вот недавно полученное сообщение: в США разработана послойная голографическая память, хранящая уже 10 млн. Гбайт в объеме 1 см³. Плюс сообщение о сверхпроводящих свойствах омигайда (специальной конструкции оптического волокна). Интересно, зачем это все делают «в сумме»?

Литература

- [1] *Лачинов В. М., Поляков А. О.* Информодинамика или Путь к Миру открытых систем. — СПб.: Изд. СПбГТУ, 1999. — 432 с.
- [2] *Булкин Г. А., Григорьев В. В., Поляков А. О.* Организация геологической автоматизированной информационно-логической системы на основе языка предикатов // Алгоритмические модели автоматизации исследований. М.: Наука, 1980. — с. 221–235
- [3] *Поляков А. О.* От количественной информации к информодинамической машине. — СПб.: Изд. СПбГТУ, 2001. — 80 с.
- [4] *Ерофеев А. А., Поляков А. О.* Интеллектуальные системы управления. — СПб.: Изд. СПбГТУ, 1999. — 264 с.
- [5] *Penrose R.* Shadows of the Mind. — Vintage, 1995.
- [6] *Поляков А. О., Лачинов В. М.* Некоторые вопросы методологии разработки автоматизированных систем управления. Ленинград: Изд. НИИЭФА, 1983. — 32 с.
- [7] *Однобоков В. В.* Компьютерное моделирование процессов обучения и контроля. Псков: Изд. ППИ, 2002. — В печати.