



**ДЗЕГЕЛЕНОК Игорь Игоревич** -  
доктор технических наук, профессор  
Московского энергетического института  
Адрес: 111250, г. Москва, Красноказарменная ул., 14,  
МЭИ, кафедра ВМСиС  
e-mail: dzegeli@gmail.com

## НА ПУТИ К ПРЕОБРАЗОВАНИЮ ИНФОРМАЦИОННЫХ РЕСУРСОВ ИНТЕРНЕТА В АКТУАЛЬНОЕ ЗНАНИЕ

### Введение

Неудержимый рост информационных ресурсов Всемирной паутины не означает того, что они сохнут востребованы. Возможно, они важны, нужны и полезны. Но зачастую они не делают погоды в силу их разрозненности и отсутствия явной зависимости от определяемых целей существования (и тем более развития) рассматриваемых объектов данной предметной области. В этом отношении значительно больший интерес представляет, по возможности, своевременно выверенное на практике причинно-следственное знание, которое таит в себе Всемирная паутина в виде структурированного информационного ресурса [1-5]. Собственно в этом и состоит проблема актуализации знаний.

Получаемое знание будет иметь ценность тогда и только тогда, когда оно станет отражать реальное состояние дел на настоящее время, то есть когда оно будет основано на актуальной информации [14]. При определенных далее ограничениях данная проблема может быть решена применением методологии «открытий», где «открытие» рассматривается с позиций субъекта в виде ранее неизвестных ему областей существования эффективных и конкурентоспособных решений. Конструктивность предлагаемой методологии позволяет положить ее в основу создания и дальнейшего развития сетевых образовательных технологий нового поколения. Раскрытию этих возможностей как раз и посвящена данная работа. Начнем ее изложение с определения, казалось бы, непреодолимых трудностей в постижении Истины [20].

### 1. Факторы «НЕ», препятствующие актуализации знаний

Искомое знание будем представлять в виде некоторой закономерности «профессионального успеха»  $g(x_1, \dots, x_n)$  относительно варьируемых переменных  $x_i$ ,  $i=1, \dots, n$ , отображающих свойства предметной области (в нашем случае - вида предметной деятельности). Конкретные проявления каждого свойства условимся отображать именными (качественными) значениями  $a_{ij}$ , что позволяет абстрагироваться от задания в явном виде значений физических величин при построении искомой модели. Определяемым таким образом предметным переменным, а точнее их именованным значениям, можно придать причинную интерпретацию, определяющую их взаимосвязь и степень влияния на скалярную величину  $g(x)$  обобщенного показателя качества. В таком понимании искомую модель можно так же трактовать, как некоторое каузальное или причинно-следственное знание. Понятно и то, что сохнут построить такую модель не представляется возможным. И вот почему.

Начнем с того, что в теории искусственного интеллекта сложилась традиция качественной оценки (на уровне общих представлений) сложности плохо формализуемых задач посредством так называемых факторов «НЕ». Построению модели препятствует, по меньшей мере, семь НЕ - факторов, а именно:

- НЕопределенность цели;
- НЕаналитичность предъявляемых требований-ограничений;
- НЕдостаточность системного видения рассматриваемых свойств в виде определяемых предметных переменных;
- НЕмонотонность в сохранении качественных оценок при введении новых предметных переменных;
- НЕреализуемость ряда комбинаций при вариации значений предметных переменных;
- НЕполнота фактов, представляющих искомую модель;
- НЕлинейность выявляемого знания в целом.

Необходимость преодоления упомянутых факторов указывает на то, что если и существует возможность формализованного построения и искомой модели, то она должна быть многоэтапной в осуществлении целенаправленного диалога человека с машиной.

Можно говорить лишь о частичной автоматизации этого процесса, который во многом будет носить познавательный характер.

Переходя к более пристальному рассмотрению процесса формализованного синтеза искомого знания, заметим следующее. Обучающее множество фактов никогда не бывает представительным, от чего страдает достоверность искомой модели. Парадокс в том, что для получения представительного обучающего множества нужна достоверная модель. А получение достоверной модели знания может быть осуществлено на представительной обучающей последовательности. Этот заколдованный круг разрывается определением рассматриваемого далее общего подхода, основанного на решении серии открытых задач.

**2. Понятие открытой задачи познавательного проектирования**

Вначале дадим такое предельно простое определение [11]. Открытая задача - это конструкция, включающая:

- 1) жестко не фиксированный набор предметных переменных  $x_i$ , определяющих свойства изучаемого мира объектов  $x$ , как правило, в виде разрозненных фактов;
- 2) имеющиеся «под рукой» примеры удачных/неудачных, возможно, разрозненных фактов, индуктивное обобщение которых приводит к получению искомого причинно-следственного знания  $g(x)$ ;
- 3) свободно задаваемое правило качественной оценки  $Q(x)$  эмпирических гипотез  $x_0$ , выдвижение и оценка которых работает на повышение достоверности искомого знания.

Последующая формализация этого понятия позволит нам двинуться дальше в конструктивном воплощении базовой информационной технологии в виде Решателя открытых задач [12].

Пусть зафиксирован набор внутренних параметров  $p_i, i=\overline{1, n}$ , определяющий возможные альтернативы строения объектов, предметов, процессов или явлений, рассматриваемых в поисковом пространстве  $D^n$  мерности  $n$ . Каждому параметру  $p_i$  сопоставляется предметная (качественная) переменная  $x_i \subset M_i$  на базовом множестве  $M_i$  имен  $a_j, j=\overline{1, k_i}$  сопоставляемых или оговоренным значениям, или интервалам значений параметра  $p_i$ , или любым его возможным проявлениям (например, вид, тип аппаратных или программных средств).

Полагая, что порядок задаваемых имен зафиксирован по индексу  $j$ , каждую переменную  $x_i$  можно рассматривать как именную переменную, принимающую соответствующие именные значения  $j=\overline{1, k_i}$ , где  $i=\overline{1, n}$ . Тогда каждой альтернативе сопоставляется упорядоченный набор значений именных переменных в виде вектора  $x=(x_1, \dots, x_n)$ .

Не исключено, что свободно определяемое поисковое пространство  $D^n$  - противоречиво, то есть существуют такие  $I$ -запрещенные комбинации значений порядковых переменных  $x_i$ , которые определяют множество  $\Phi_{запр}$  запрещенных фигур (не имеющих физического смысла альтернатив)  $x \subset \Phi_{запр} \subseteq V$ , где  $V=M_1 \times \dots \times M_n$ , и тогда  $\Omega=V \setminus \Phi_{запр}$  - множество допустимых решений.

Пусть далее существует следующий способ свободно задания цели. Предлагается ограничиться лишь вынесением качественных оценок об альтернативах  $x$  с помо-

щью некоторого правила-предиката  $Q(x)$ , имеющего лишь две градации:  $Q(x)=1$ , если  $x$  признается как удачное решение, и  $Q(x)=0$  в противном случае. Здесь и далее через  $T$  обозначим множество удачных, а через  $F$  - неудачных решений. Если же принято соглашение об оценке  $x$  с помощью набора внешних показателей качества  $f_i(x) \rightarrow \min$ , то для определения области истинности предиката  $Q(x)$  можно ограничиться заданием набора граничных значений  $f_i^0$ , где  $I=\overline{1, L}, T=\{x | f_i(x) \leq f_i^0, x \in \Omega\}$ .

И, наконец, принципиальный момент. Текущие сведения об «удачах» и «неудачах» ограничены, что можно выразить заданием соответствующих множеств:  $T^{(u)} \subset T, F^{(u)} \subset F$ , где  $u$  - порядковый номер сеанса приобретения эмпирических знаний. Таким образом, список  $\langle T^{(u)}, F^{(u)} \rangle$  определяет текущую базу фактов  $Ex^{(u)}$  (от англ. - example). Индуктивное обобщение последних составляет основу для получения приближений  $g^{(u)}(x)$  к искомой закономерности «удачи»  $g(x)$ . Однако на практике приходится считать с тем, что обычно:  $|T^{(1)}| \ll |T|, |F^{(1)}| \ll |F|$ , и потому 1-е приближение  $g^{(1)}(x)$  может оказаться слишком далеким от Истины. Следовательно, нужен некоторый механизм целенаправленного пополнения фактов. Представим его в виде оператора  $P \rightarrow \{x_0\}$ , где  $x_0$  - эмпирические гипотезы о достоверности  $g^{(u)}(x)$ , требующие экспериментальной проверки.

**Определение.** Открытая задача познавательного проектирования - это конструкция

$$\langle D^n, Ex^{(u)}, Q(x), \{g(x)\}, P \rightarrow \{x_0\} \rangle,$$

действие которой направлено на восстановление искомой модели-закономерности «удачи»  $g(x)$  на всей области вариации значений порядковых переменных в поисковом пространстве  $D^n$  (с учетом запрещенных комбинаций  $I$ ) в результате многократного ( $u=1, \dots, U$ ) обобщения активно пополняемой базы фактов  $Ex^{(u)}$ , образуемой путем целенаправленного выдвижения эмпирических гипотез  $x_0$  и их качественной оценки по правилу  $Q(x_0)$ .

Данное определение - первый шаг к конструктивному воплощению предлагаемого подхода. Он позволяет в свободной форме, «на лету» методически единообразно ставить и решать открытые задачи по выявлению недостающего эмпирического знания, не прибегая к построению и применению трудноуловимых и во многом неадекватных аналитических моделей [11-12].

Нужно лишь подчеркнуть следующее. Разрешимость открытых задач постулируется следующим образом.

**Свойство представительности.** В поисковом пространстве  $D^n$  существуют такие наиболее информативные множества  $T_0, F_0$  опорных решений  $x_0 \in T, F$ , которые достаточно полно представляют множества  $T, F$  и границу их разбиения  $g(x) = 0$  такую, что  $g(x) \geq 0$ , если  $x \in T$ ;  $g(x) < 0$ , если  $x \in F$ .

Данное свойство, определяемое и воспринимаемое как постулат, хорошо согласуется с представлением о существовании внутреннего порядка в виде «островов» и «архипелагов», а точнее - многомерных многообразий удачных решений в поисковом пространстве  $D^n$ . Только тогда можно говорить о существовании именно закономерности «удачи» в классе  $\{g(x)\}$ , обладающей панорамным видением всех решений. Только тогда обнаруженный порядок может оказать влияние на степень сокращения числа эмпирических гипотез  $x_0$  в практически неограниченной области  $\Omega \setminus Ex^{(u)}$ .

Постановка открытой задачи является первым и наиболее важным этапом решения, от которого зависит конечный успех выявления скрытого знания. Именно на этом этапе в наибольшей степени проявляется творческий характер решаемых задач и талант проектировщика [11].

### 3. Стратегия активного формирования искомого знания

Вместе с тем, подавляющее большинство объектов удобно описывать булевыми

$$g(x) = \sum_{i=1}^n \sum_{j=1}^{k_i} c_{ij} x_{ij} - c_0 - Z$$

переменными  $x_{ij} \in \{0,1\}$  с двойной индексацией, где  $i$  - номер параметра,  $j$  - номер его именного значения. Тогда интересующее нас знание или закономерность «удачи»  $g(x)$  можно представить в виде обобщенной скалярной функции выбора с весовыми коэффициентами  $c_{ij}$  при каждой переменной  $x_{ij}$ ; нелинейной компонентой  $Z$  сколь угодно высокого порядка и порогом  $c_0$ .

Преимущество представления искомого знания в виде обобщенной скалярной функции выбора  $g(x)$  заключается в возможности достижения эффекта приоритетного различения альтернатив-решений, а именно:

$x_a$  «лучше»  $x_b$ , если только  $g(x_a) > g(x_b)$ .

С точки зрения выполнимости некоторой обобщенной цели, заложенной в определяемое правило  $Q(x)$ , такая возможность возникает не просто так. Она является результатом выявления «островов удачи», существование которых постулируется свойством представительности. Продолжением этого рассуждения является нахождение наилучшего решения

$x^* = \text{Arg max}_x g(x)$  среди  $x \in T$ ,

которое можно трактовать как «открытие», если окажется, что  $x^* \in T^{(1)}$ .

Для синтеза искомого знания  $g(x)$  нужна информация об опорных множествах  $T_0, F_0$ . Но их нахождение возможно лишь при наличии уже синтезированной функции  $g(x)$  [4]. Для выхода из, казалось бы, заколдованного круга предлагается следующая стратегия.

В общем случае  $u$ -й уровень знаний индуктивно обобщается в виде приближения  $g^{(u)}(x)$  к искомой функции выбора. Для этого решается открытая к пополнению система неравенств:  $g(x) \geq 0$  для всех  $x \in T^{(u)}$  и  $g(x) \leq -1$  для всех  $x \in F^{(u)}$ .

Не исключено, что текущее приближение к искомому знанию - далеко от совершенства. Для проверки этого факта относительно  $g^{(u)}(x)$  порождаются **эмпирические гипотезы**, определяемые в виде опорных решений двух сортов: положительные гипотезы  $x_0^+ \in S_0^{(u)}$ , для которых  $u(x_0^+) \geq 0$ , и отрицательные гипотезы  $x_0^- \in E_0^{(u)}$ , когда  $u(x_0^-) < 0$ . Механизм выдвижения гипотез строго формализован. В основе их выдвижения лежит критерий топологической полноты обследуемых подпространств. В первом приближении - это точки  $x_0$ , находящиеся вблизи гиперповерхности  $g^{(u)}(x)=0$ .

Полученные таким образом гипотезы подвергаются качественной оценке с применением правила  $Q(x)$ . Если окажется, что все  $x_0^+ \in T$  и все  $x_0^- \in F$ , то есть все гипотезы полностью подтвердятся по знаку, то  $g^{(u)}(x)$  - искомое знание. Однако на начальных стадиях  $u=1,2,\dots$  индуктивного обобщения фактов обнаруживаются **противоречия**, т.е. часть решений  $x_0^+ \in F$ , а часть решений и потому искомое знание - недостоверны.

Тогда выявленные противоречия автоматически превращаются в факты в виде соответствующих множеств-пополнений:  $\Delta F^{(u)} = \{ x_0^+ \in F \}$ ,  $\Delta T^{(u)} = \{ x_0^- \in T \}$ , что позволяет перейти на следующий ( $u+1$ ) уровень знаний. Описанный процесс порождения и качественной оценки эмпирических гипотез повторяется снова и снова. Критерием остановки итеративного процесса служит состояние отсутствия новых гипотез. Сами же гипотезы выдвигаются по критерию полноты тестирования искомого знания или закономерности «удачи». Геометрически эта закономерность представляет границу разделения всего мира объектов (а не его части в виде начальной базы факторов!) на «удачи» и «неудачи».

Стратегия выдвижения гипотез основана на выявлении, прощупывании только этой границы, а не всего поискового пространства. В результате резко сокращается число выдвигаемых гипотез или вопросов к эксперту при одновременном повышении достоверности искомого знания.

Практическая польза от применения предлагаемой стратегии заключается в следующем [8]. Во-первых, она направлена на выявление как можно большего числа противоречий для функции выбора, что означает, что  $g^{(u+1)}(x)$  не может быть хуже  $g^{(u)}(x)$ . Во-вторых, при благоприятном стечении обстоятельств (многократное подтверждение гипотез) обеспечивается резкое сокращение числа рассматриваемых альтернатив по сравнению с полным перебором.

Важно и то, что рассмотренные принципы и технологические особенности активного формирования эмпирического знания обладают большой степенью инвариантности к предметным областям (техника, архитектура и строительство, биология, медицина, психология, педагогика, военное дело и т.д.).

### 4. Решатель открытых задач как инструмент познания

В пределах данной технологии человек (учащийся, слушатель, исследователь-аналитик) выполняет следующие функции: 1) фиксирует поисковое пространство; 2) задает начальную базу фактов; 3) оценивает гипотезы по правилу качественной оценки. Тем самым человек ставит открытую задачу.

Наиболее трудоемкие в вычислительном и логическом отношении функции: синтез текущего и искомого знания, главное, выдвижение гипотез берет на себя такой мощный инструмент, реализованный программно, как Решатель открытых задач.

Основные блоки Решателя - блок индуктивных обобщений и генератор эмпирических гипотез.

Успех применения Решателя определяется пользовательским интерфейсом, в состав которого входят редактор поискового пространства, редактор запрещенных фигур, а также средства наглядного отображения добываемых знаний. Одна из наиболее продвинутых версий Решателя - «Аргонавт v.3.0» - подробно представлена в сборнике «100 компьютерных программ для бизнеса» [12, 16].

Что же дает сформированное знание?

*Во-первых*, каждое значение определенной именной переменной автоматически приобретает некоторый вес. По величине этого веса можно судить о степени влияния данного знания, а в целом именной переменной, характеризующей свойства изучаемого мира объектов, на обобщенный показатель качества в виде правила качественной оценки.

*Во-вторых*, по величине значений закономерности «удачи» (которая является скалярной функцией) можно оценить ранее «незнакомый» объект, если только его свойства определяются поисковым пространством.

*В-третьих*, можно проранжировать все интересующие объекты в смысле - «лучше-хуже» по оговоренному обобщенному показателю качества.

*В-четвертых*, нетрудно найти лучшее решение, которое занимает первое место теперь уже во взвешенном мире изученных объектов. Если это решение не принадлежит конечной базе фактов, накопленной в Решателе, то оно приобретает статус «открытия».

*В-пятых*, следует отметить и такую замечательную возможность, как «открытие» при условии выполнения вполне определенных свойств познаваемого мира объектов. Эти свойства могут, в частности, отражать некоторые характеристики внешней среды, оказывающей влияние на существование возможных объектов. Получаемый

результат напоминает решение оптимизационной задачи на условный экстремум, когда фиксируются конкретные значения ряда переменных. Отмеченная возможность, особенно когда получаемые «открытия» ставятся в зависимости от конкретных свойств внешней среды, достигается за счет появления нелинейности в искомом знании.

Исторически необходимость решения открытых задач возникла задолго до появления учений об открытых архитектурах в области Computer science. Первые задачи были решены при выполнении ряда НИР и НИОКР, связанных с поиском вариантов алгоритмической и структурной организации проблемно-ориентированных мульти-компьютерных систем.

Стало ясно, технология применения Решателя во многом адекватна движению инженерной мысли, предусматривающей добывание недостающих знаний для поиска и аргументации проектных решений. Начальный опыт внедрения технологии решения открытых задач в учебный процесс также оказался успешным. Наибольший педагогический эффект был достигнут в рамках читаемой автором дисциплины «Поисковое проектирование вычислительных систем». В дальнейшем накопленный опыт был перенесен и в другие вузы. Так, в рамках дисциплины «Интеллектуальные информационные технологии и стратегии» ставились и решались учебные задачи в таких далеких от вычислительной техники областях, как ранжирование банковских технологий, прогнозирование финансовых рисков, оценка недвижимости, выбор бытовой техники, качественная оценка военной техники и объектов гражданского строительства.

Без преувеличения можно сказать, что именно учебный процесс стимулировал развитие рассматриваемой технологии.

## 5. Сетевая технология коллективного творчества

Следующий шаг в развитии предлагаемой технологии - организация согласованного сетевого взаимодействия коллектива локальных экспертов, направленная на построение объемных и емких по содержанию профессионально-ориентированных моделей с применением композиции Решателей открытых задач [14]. В данном случае возможно удаленное сетевое взаимодействие N локальных экспертов (ЛЭ) и главного эксперта.

Каждый эксперт работает со своим Решателем, ставя перед ним частную открытую подзадачу в собственном поисковом подпространстве. В свою очередь результаты, получаемые локальными Решателями (наилучшие решения и генерируемые эмпирические гипотезы), передаются «наверх» Решателю главного эксперта по компьютерной сети Internet/Intranet. Возможна также передача подобной информации в обратном направлении. Все зависит от вида реализуемых сценариев. В общих чертах сценарий таков.

Сложный объект профессиональной деятельности (например, образовательный стандарт, педагогическая технология, оценочная процедура результатов инновационной деятельности, электроэнергетическая система, крупная электростанция, обслуживающий вычислительный комплекс, летательный аппарат, реализуемый бизнес-процесс) представляется композицией отдельных частей.

За каждую смысловую часть отвечает локальный эксперт (ЛЭ). Деятельность ЛЭ координируется главным экспертом (ГЭ). В результате формируется обобщенная причинно-следственная модель интегрированного знания, отвечающая требованию системного видения сложного объекта.

Реализуемые сценарии представляют собой композиции «треугольников познания»: 1-е ребро треугольника символизирует индуктивное обобщение фактов, 2-е

ребро - выдвижение эмпирических гипотез, 3-е ребро - качественную оценку гипотез, что обеспечивает направленное пополнение базы фактов Решателя.

Первый базовый сценарий таков: «Композиция лучших решений локальных экспертов - расщепление гипотез главного эксперта». Так, при реализации первого сценария активность исходит от главного эксперта, который не дает усложняться на достигнутом локальным экспертам.

Второй сценарий обеспечивает формирование более представительной модели за счет большей комбинаторики. Здесь уже активность исходит от локальных экспертов, которые заставляют главного эксперта проверять и перепроверять композиции их гипотез.

## 6. Банк открытых задач в сетевой среде

Как показывает опыт учебной деятельности, постановка студентами своих выстраданных открытых задач носит явно выраженный творческий характер. Каждая задача во многом уникальна и представляет интерес для дальнейшего углубления познавательного процесса.

За время эксплуатации Решателя было поставлено и решено значительное количество открытых задач, что привело к возникновению вопроса их систематизации и упорядоченного хранения. Решатель не содержит механизмов, позволяющих решить данную проблему, и решаемые задачи хранятся в виде группы разрозненных файлов на диске. Такое хранение существенно затрудняет поиск и использование необходимых знаний, полученных в результате решения открытых задач.

Очевидно, что решенные задачи могут быть использованы проектировщиками для постановки собственных задач, а также для анализа проектных решений на основе знаний, сформированных предыдущими разработчиками. Однако при имеющейся технологии хранения эффективное использование накопленных знаний представляется невозможным.

Таким образом, необходим программный продукт, который обеспечивал бы, как минимум, систематизацию и надежное хранение решенных открытых задач. Кроме того, данный программный продукт должен обеспечивать наглядное отображение имеющихся знаний и обеспечивать быстрый доступ к ним.

Так мы приходим к необходимости создания Автоматизированного банка, обеспечивающего архивирование, систематизацию, хранение и сопоставление открытых задач с возможностью оценки новизны получаемых решений при подведении результатов конкурсного проектирования [14].

## 7. Технологическая схема актуализации знаний в интернете

Возможность актуализации знаний в интернете достигается путем встраивания Решателя открытых задач в единую технологическую цепочку согласованного информационного поиска и преобразования набора неструктурированных HTML-страниц в структурированную базу данных реляционного типа так, как это показано на **рис. 1**.

Стартовые позиции технологии определяются **тезаурусом**, который представляет собой систематизированный набор терминов и относительно ясных, выверенных на практике понятий, дающих по возможности полное представление о данной предметной области. В частности, рассматривается класс древовидных тезаурусов. Узлам дерева сопоставляются понятия и термины предметной области, которые на последующих ярусах все более уточняются. Для формирования информационных запросов по каждому понятию в узлах дерева размещаются **слоты** для занесения наборов **ключевых слов**.

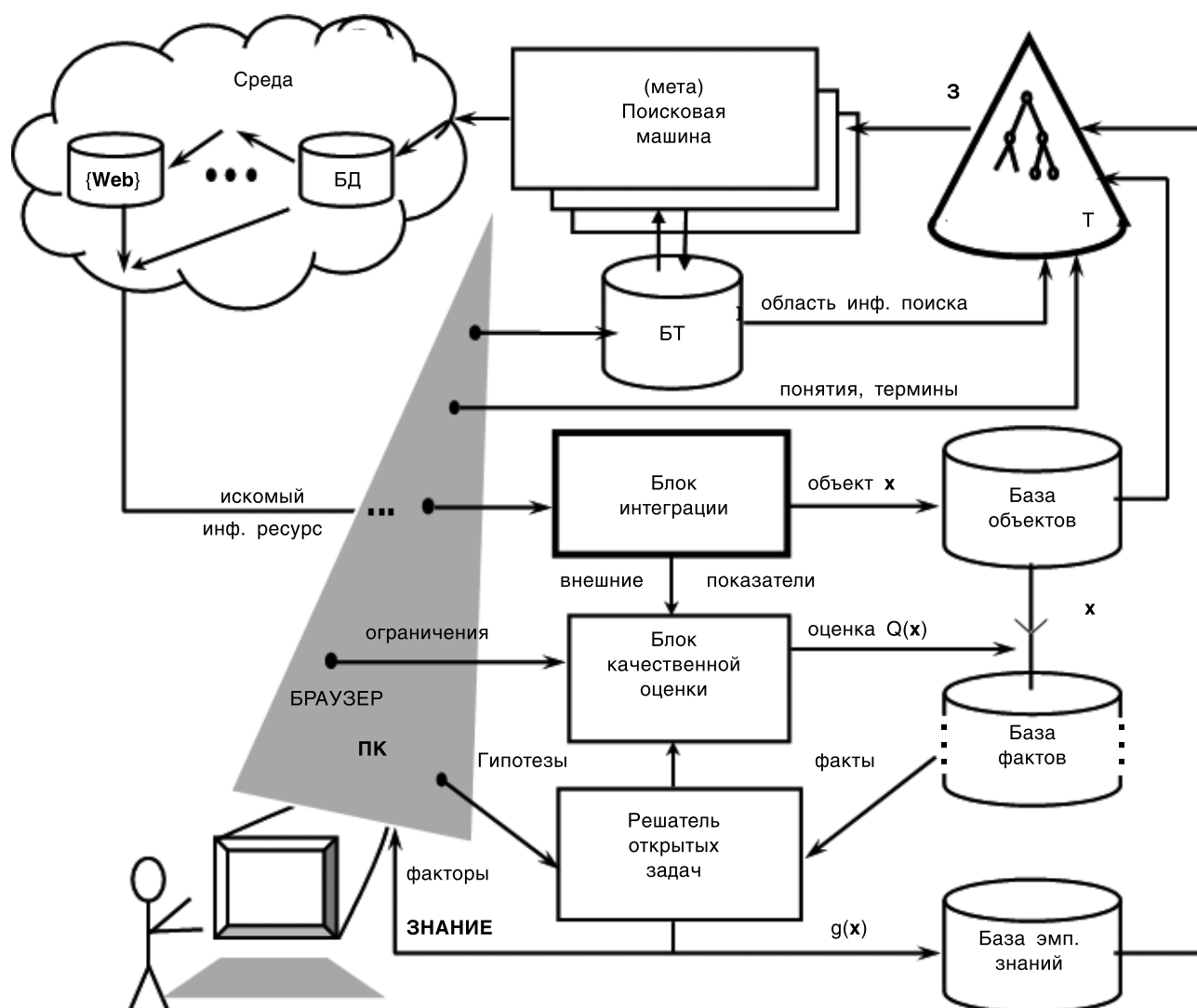


Рис. 1. Укрупненная блок-схема технологии актуализации знаний в среде WWW интернета

Блок интеграции служит для преобразования поступающих из интернета HTML-страниц в целевую базу объектов, характеризующих изучаемую слушателем предметную область. Это наиболее сложный блок [18]. Его работу можно представить в виде двух основных этапов.

На первом этапе из интернета на основании тезауруса извлекается набор наиболее «подходящих» HTML-страниц. Как показано в [15], такая возможность достигается на основе программной реализации адаптивного агента. Свойство адаптивности заключается в осуществлении обратной связи со стороны пользователя, который вмешивается на начальном шаге оценки степени пригодности полученных HTML-страниц.

На втором этапе по виду скорректированного тезауруса производится формирование структуры базы объектов и далее получение записей в базу объектов по результатам извлечения табличной информации, скрытой в HTML-страницах. Для извлечения необходимых записей из найденных HTML-страниц используется язык шаблонов в виде регулярных выражений. Такая возможность подтверждена результатами работы экспериментальной версии программы, созданной в МЭИ.

Посредством блока качественной оценки осуществляется преобразование базы объектов в базу фактов. Последняя служит питательной средой для окончательно-

го превращения добытой в интернете информации в актуальное знание.

Суть реализуемого подхода состоит в организации не отрывочного, спонтанного информационного поиска, а многосвязного, ассоциативного поиска, основанного на предварительном построении уточняемого тезауруса запросов в виде растущего дерева ключевых слов. Запросы, сгенерированные из различных узлов дерева, взаимно дополняют друг друга. В результате удается сократить общее время построения предметного дерева. По сравнению со стандартными средствами информационного поиска в интернете вводится посреднический слой в виде базы знаний и **брокера запросов**.

Опытная версия программного продукта получила название (с учетом перспектив развития) «**Web-Knowledge-Miner**». По сути, данная программа представляет собой инструмент, позволяющий добывать информационный «капитал», структурированный по узлам дерева. При этом пользователь освобождается от многих рутинных действий во время сеансов работы в интернете и потому может сосредоточиться на творческой работе по содержательному формированию древовидного тезауруса.

Рассмотренные технологии и инструментальные средства отработывались и проверялись во многих образовательных учреждениях. Среди них: кафедра вычислитель-

ных машин, систем и сетей МЭИ (ТУ), Военная Краснознаменная академия противовоздушной обороны имени маршала Г.К. Жукова, Институт безопасности бизнеса, Исследовательский центр проблем качества подготовки специалистов при МИСиС, кафедра систем автоматического управления и электротехники Калужского филиала МГТУ им. Н.Э. Баумана, кафедра САПР Московского государственного горного университета. Накопленный опыт проведения с использованием рассмотренных инструментальных средств и технологий подтвердил их высокую результативность в плане активизации творческой деятельности слушателей, аспирантов, студентов и магистров.

## Выводы

1. Основная особенность предлагаемой технологической платформы - отказ от заманчивой идеи извлечения из интернета готового знания, если даже оно вписывается в интеллектуальную прослойку в виде уже существующих онтологий или иных средств формализации знаний (как, например, в [17]).

2. При определении предлагаемого подхода к извлечению знаний делается ставка на его синтез, и, поскольку оно не является готовым, искомый результат во многом зависит от успеха в поиске подходящего информационного ресурса.

3. Для преодоления многочисленных препятствий в виде оговоренных выше (раздел 1) факторов НЕ процесс синтеза искомого знания представлен в рамках понятийного аппарата открытых интеллектуальных систем, предусматривающих, как это показано в разделе 2, постановку и решение открытых задач познавательно-исследовательского проектирования.

4. Принципиально важное отличие определяемого процесса машинного обучения при решении откры-

тых задач - обеспечение представительности обучающей последовательности, что достигается выдвиганием строго определяемых гипотез. Их двувальтернативная оценка приводит к преодолению неопределенности результата машинного обучения и, в конечном счете, - к повышению достоверности искомого знания.

5. Технологическая платформа, воплощающая предлагаемый подход, может послужить основой ее практической реализации в виде ряда оригинальных сетевых информационных технологий, таких как сетевая версия Решателя открытых задач, сетевая организация взаимодействия многих Решателей, сетевой банк открытых задач (точнее синтезированных знаний). При этом ведущей сетевой технологией должна явиться интегрированная инструментальная среда инженерии добываемого из интернета знания относительно предварительной построенной априорной модели предметной области.

6. Накопленный автором опыт постановки и решения многочисленных и разнообразных открытых задач подтверждает работоспособность и высокую эффективность имеющегося инструментария решения открытых задач, связанных с выполнением ряда НИР (подробнее [13]) в части поиска конкурентоспособных решений.

7. Что же касается профессионального образования, то здесь до сих пор существует определенный разрыв в совместном применении высших достижений педагогики (активное, игровое, контекстное и другие виды обучения) и сетевых компьютерных телекоммуникаций. Главное состоит в обеспечении оперативного доступа к мировым информационным ресурсам с обеспечением возможности их превращения в актуальное знание любой отрасли. Речь идет о развитии навыков инженерного творчества посредством обучения через «открытие». Этот тезис мы и пытались отстоять в данной работе.

## Литература:

1. Chakrebarti S. Mining the Web: Discovering Knowledge from Hypertext... Publisher: Morgan Kaufmann, 2003. - 345 p.
2. Cooley R., Mobasher B., and Srivastava J. Web mining: Information and pattern discovery on the World Wide Web // Proc. 9th IEEE Int. Conf. Tools with Artificial Intelligence, pp. 558-567, Nov. 1997.
3. Kosla R. and Blocheel H. Web mining research a survey, SIG KDD Explorations, vol. 2, pp. 1-15, July 2000.
4. Ланде Л.В. Поиск знаний в Internet. Профессиональная работа. - М.: Диалектика. - 272 с.
5. Бизнес-аналитика: от данных к знаниям / Н.Б. Паклин, В.И. Орешков. - СПб: Питер, 2010. - 704 с.
6. Анализ данных и процессов / А.А. Барсегян, М.С. Куприанов, М.Д. Тэсс, С.И. Елизаров. - СПб: БХВ-Петербург, 2009. - 512 с.
7. Chakrebarti S. Mining the Web: Discovering Knowledge from Hypertext... Publisher: Morgan Kaufmann, 2003. - 345 p.
8. Техническое творчество: теория, методология, практика. Энциклопедический словарь-справочник / Под редакцией А.И. Половинкина и В.В. Попова. - М.: Информ-система, 1995. - 408 с.
9. Добряков А.А. Психолого-педагогические основы подготовки элитных специалистов как творческих личностей (содержательные элементы субъект-объектной педагогической технологии). - М.: Логос, 2001.
10. Дзегеленок И.И. Логика поиска проектных решений - М.: Моск. энерг. ин-т, 1984. - 64 с.
11. Дзегеленок И.И. Открытые задачи поискового проектирования - М.: Моск. энерг. ин-т, 1991. - 66 с.
12. АРГОНАВТ - система приобретения эмпирических знаний / В кн. Справочно-практическое издание

«СОФТЕЛЬ - Сто компьютерных программ для бизнеса». Вып. 1997-1998. - М.: ХАМТЕК ПАБЛИШЕР, 1997. - С. 117-125.

13. Дзегеленок И.И. Методология поискового проектирования вычислительных систем // Информационная математика. - 2004. - № 1(4). - С.110-119.

14. Дзегеленок И.И. Сетевые образовательные технологии актуализации знаний // Информационные технологии в проектировании и производстве. - 2003. - № 3. - С.10-15.

15. Дзегеленок И.И., Филиппов И.В. Подход к созданию интеллектуального агента согласованного информационного поиска в интернете / 9-я Национальная конференция по искусственному интеллекту. Труды конференции. Том 2. - М.: Физматлит, 2004. - С. 715-722.

16. Дзегеленок И.И., Аляева Ю.В., Кузнецов А.Ю. Инструментальные средства поискового проектирования / Лабораторные работы №1-3: методическое пособие. - М.: Издательский дом МЭИ, 2007. - 40 с.

17. Токмаков Д.И. Использование средств языка RDF в аннотировании Интернет-ресурсов // Информационные ресурсы России. - 2007. - № 5. - С. 30-33.

18. Рыков В.В. Семиотическое моделирование процессов понимания // Информационные ресурсы России. - 2010. - № 6. - С. 31-35.

19. Становление и развитие научной школы искусственного интеллекта в Московском энергетическом институте / Вагин В.Н., Еремеев А.П., Колосов О.С., Дзегеленок И.И., Фролов А.Б. // Программные продукты и системы. - 2010. - № 3. С. 3-16.

20. Еремеев А.П. Разработка и применение перспективных информационных ресурсов на основе методов искусственного интеллекта // Информационные ресурсы России. - 2009. - № 6. - С. 17-20.