

DOI: 10.12737/21052

СОХРАНЕНИЕ, ПЕРЕДАЧА И ИСПОЛЬЗОВАНИЕ ЗНАНИЙ НА ОСНОВЕ ПРИМЕНЕНИЯ ТЕХНОЛОГИИ «ПРЯМОГО НАЛОЖЕНИЯ ЗНАНИЙ»

Г.Б. БРОНФЕЛЬД

Федеральное государственное бюджетное образовательное учреждение высшего образования «Нижегородский государственный технический университет им. Р.Е. Алексеева», ул. Минина, 24, Нижний Новгород, Нижегородская обл., 603155, Россия

Аннотация. В статье рассматривается проблема сохранения, передачи и использования знаний на основе технологии «прямого наложения знаний». Описываются новый способ моделирования знаний, новая технология прямого наложения знаний, новые виды интеллектуальных систем – элинги, АСУЗ, БАНЗ. Показывается, что технология прямого наложения знаний и элинги/АСУЗ/БАНЗ могут обеспечить «практически вечное» сохранение, как минимум, знаний, необходимых людям для практической деятельности.

Ключевые слова: знание, модель знаний, текст, книга, молинга, элинга, база знаний, машинный носитель, сохранение знаний.

THE PRESERVATION, TRANSMISSION AND USE OF KNOWLEDGE-BASED THE TECHNOLOGY OF DIRECT KNOWLEDGE OVERLAY"

G.B. BRONFELD

Federal state budget-ing the institution of higher education-about "Nizhny Novgorod State Technical University. RE Alekseev", Str. Minin, 24, Nizhny Novgorod, Nizhny Novgorod region., 603155, Russia

Abstract. The problem of preservation, transmission and uses of knowledge-based technologies, "direct blending of knowledge." It describes the new ACT-sob knowledge modeling, the new technology of direct blending of knowledge, new in-intellectual systems - Eling, BMS, Banse. It is shown that the technology of direct blending of knowledge and Eling / BMS / Banse mo-gut to provide "almost eternal" preserves-tion, as a minimum, knowledge, people need to practice.

Key words: knowledge, knowledge-model tions, text, book, molinga, Eling, knowledge base-tion, machine support, preservation of knowledge.

Введение. В последние десятилетия резко активизировались работы в области создания *интеллектуальных систем* (ИС), которые помогут решить проблему «сатурации» («пере-насыщения информацией») [8]. Одной из таких является *технология прямого наложения знаний* (ТПНЗ) [2-6]. Первым ТПНЗ в 90-х годах использовал Д. Грей [13] при создании «виртуальной астрономической обсерватории» для обработки численных данных наблюдений за звездным небом. Однако, видимо Д. Грей не сообразил, что этот эффект применим в широком аспекте в иных предметных областях для разных видов знаний при ис-

пользовании разных ИС.

ТПНЗ основана на подходе, использующем интеграцию знаний. Вообще знания содержатся в разной форме. В частности, и известный философ К. Поппер, и академик Д.А. Поспелов, говорили, что знания об окружающем мире отражены и в книгах, и документах, созданных человеком. Это – тексты в соответствии с принятым подходом в лингвистике. Специалисты при решении текущих проблем интересуют тексты, множество которых определенного предметного направления называется сверхтекстом.

Используемая модель знаний. По-

пробуем промоделировать знания новым методом. Известно из лингвистики, что «простое предложение ... было и остается основной единицей синтаксиса текста...», «наиболее существенной чертой предложения является его способность формировать и выражать мысль» [9]. Используем это.

Метод описания знаний моделью, названной молингой [2-6], заключается в следующем. Каждое простое предложение текста (сложные делятся на простые) представляется семантической сетью с составлением словарей терминов (включая синонимы), отношений, качественных признаков и т.д. Из предложений убирается эмоциональная окраска, они упрощаются с сохранением основных риторических отношений.

Молинги имеют формальный вид:

$$D ; P; Z; K; O; N.$$

здесь D – множество идентификаторов, с помощью них молинга выделяется из всего множества молинг. При выдаче ответа может восстанавливаться близкий к исходному текст. Элемент P -условие применимости ядра молинги. Ядро молинги Z представляет собой простое предложение текста. В K указаны номера словарей, фиксирующих положение в ядре молинги – терминов, отношений, качественных признаков и т.д. В O зафиксированы уровни достоверности молинг, например, в виде факторов уверенности, впервые примененных в *экспертной системе (ЭС) MYCIN*. Элемент N описывает постуловия молинги, которые включают в себя визуальные и графические образы, запрограммированные формулы, расчетные модели и подключаемые внешние пакеты программных средств.

В результате база знаний (БЗ) ИС представляются огромным набором коротких семантических сетей специального вида, внешне читаемых, как вполне ясные и достаточно короткие предложения, подробнее в [2-6]

Технология «прямого наложения знаний». При вводе знаний в виде молинг в БЗ, указываются идентификаторы из какого текста-источника и места в исходном тексте. Повторение знаний, уже содержащихся в БЗ

и полученных ранее из других источников, приводит к повышению достоверности имеющихся знаний. Чего нет, вводится и увеличивает объем БЗ, т.е. с каждым новым введенным текстом БЗ подобной ИС представляет собой новую версию. Введение противоположного утверждения уменьшает достоверность уже введенного. Расчетные модели, визуальные и графические образы запоминаются и подключаются через постуловия. Подобный подход позволяет сохранять знания в актуальном состоянии столетиями и тысячелетиями без всяких специальных дополнительных затрат [2-6].

Такой процесс создания БЗ ИС для интеграции знаний и будем называть ТПНЗ (рис.1).

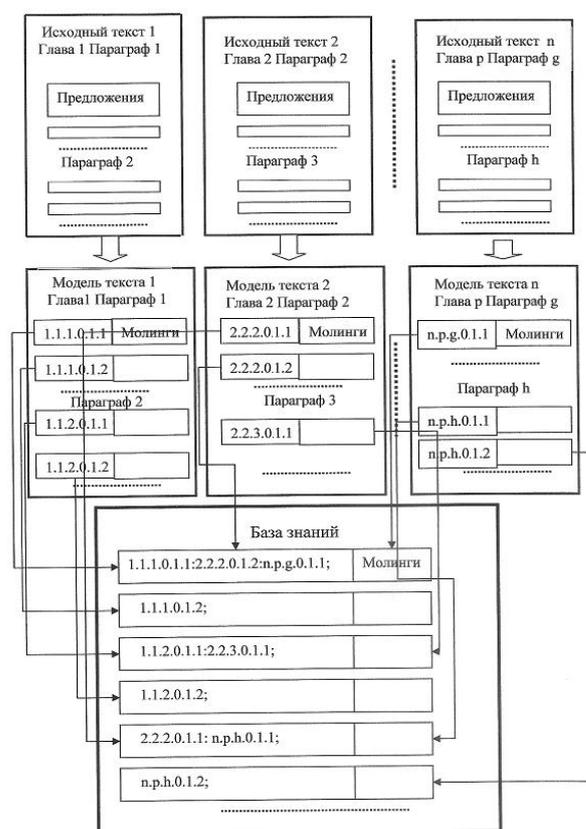


Рис.1. Технология прямого наложения знаний

Ключевая особенность – значительно меньший объем БЗ сверткестов по сравнению с другими способами получения БЗ с иными моделями знаний. А это приводит к ускорению поиска ответа даже с одинаковыми методами логического вывода от десятков и сотен раз быстрее до практической бесконечности при полном переборе.

Разновидности новых ИС. ИС с БЗ на основе ТПНЗ могут иметь следующие три базовых варианта, – с массовым производством для индивидуального потребителя в варианте *интеллектуальная электронная книга* (ИЭК), названная автором – элингой, в индивидуальном варианте для отдельного предприятия (организации) в варианте *аналитическая система управления знаниями* (АСУЗ) [2-5] и в виде *библиотеки аналитического накопления знаний* (БАНЗ) [6] для традиционного массового потребителя на почти бесплатной основе пользования.

Идею работы над общим полем знаний в понятном для него варианте высказал в 1945 г. бывший советник по науке президента Рузвельта В.Буш, что даже подтолкнуло развитие гипертекстовой технологии.

Именно в частичном разрешении этой основной проблемы и заключается основной эффект проекта ИЭК – элинга.

Элинги позволяют:

- заносить в них постепенно большое количество знаний;

- получать ответ почти со скоростью расчетных программных средств без изучения технической литературы, задавая вопросы элинге;

- приобретать их по низкой стоимости (о чем мечтал В.Буш);

- работать в широкой предметной области (в отличие от ЭС, которые работают только в узкой области и для ограниченного набора задач);

- пользователю постепенно получать со временем результаты на уровне лучших специалистов мира и выше.

В состав элинги входят программные средства, реализующие машину вывода, интеллектуальный интерфейс, БЗ и сервисные и вспомогательные программы. БЗ включает в себя словари терминов, отношений и др. (и их синонимов), собственно знания, содержащиеся в текстах в виде молинг, расчетные модели (реализованные в виде отдельных подпрограмм, ускоряющих работу элинги, и повышающих ее ценность), визуальные и графические образы.

Физически элинги представляют собой съемное устройство хранения инфор-

мации со специальной машиной вывода и БЗ. Одновременно могут создаваться элинги по многим направлениям знаний. Каждые 1-2 квартала будет выходить новая версия элинги определенного направления с новой расширенной версией БЗ.

Могут существовать две модификации элинг – просто в виде комплекса программ и БЗ, так и виде специально разработанного и массово производимого *машинного носителя знаний* (МНЗ) на основе изобретения [3]. В МНЗ те же алгоритмы обработки, как и в первом варианте, но реализуются с помощью встроенных микропроцессорных средств, обеспечивающих более высокую скорость работы и усложняющих возможность копирования и воспроизводства.

Практически «вечное» сохранение знаний. Производитель вначале выбирает печатные источники для ввода информации (знаний) в объединенную БЗ. Затем эксперт-редакторы моделируют тексты молингами и вводят молинги в БЗ. Затем происходит массовый выпуск элинг на конкретном специальном (стандартном) машинном носителе на тот период. Эксплуатация элинг-потребителями вряд ли будет более 1 года (мало кто из людей узнает новости из вчерашних газет и журналов). Постепенно формируется запрос от потребителей на новый вариант элинги с более полной БЗ. И далее по «бесконечному» циклу, пока существует человечество или не придумают нечто гораздо более лучшее, чем элинга.

Элинга выступает в данном случае для потребителя, как «носитель знаний», или иначе – локальная библиотека знаний в конкретной предметной области или областях. С АСУЗ и БАНЗ будет происходить близкая ситуация с периодической заменой используемых технических средств и программного обеспечения.

Попутно решается еще очень важная проблема, кроме чисто технических (постепенный беспроблемный переход на новые технические носители с полным сохранением всех накопленной ранее информации и знаний за годы, десятилетия и сотни лет). С течением времени меняется и сам язык, появляются новые понятия, а старые меняются [4,10]. Поэтому так плохо восприни-

маются многие знания и факты, изложенные в древних индийских, буддийских текстах, Библии и т.п. А многие письменные тексты вообще не в состоянии расшифровать. Элинга/АСУЗ/БАНЗ позволят постепенно адаптироваться к языковым изменениям и сохранять семантическое содержание устаревших понятий и технологий и состыкуют с новыми.

Диалого-ассоциативный поиск. Логический вывод в элинге проводится на основе модифицированного *modus ponens* (в простейшем виде «если *A*, то *B*»), где в качестве *A* и *B* используются термины из словарей терминов, которые помечены в каждой молинге [5,6]. Причем не имеет значения, в каком порядке находятся термины в молинге. При этом выводятся последовательно варианты возможных сочетаний молинг, упорядоченные по уровню фактора уверенности. Но выводятся блоками, не перегружая информацией пользователя.

Логический вывод в ИС ведет нередко к неограниченному перебору вариантов, даже при относительно простых запросах, что создает проблемы в общении с пользователем. Однако элинги создают здесь новые возможности.

Для чего пользователь обращается к системе?

Чтобы получить знания, как ему, например, обрабатывать конкретную деталь. Есть, например, отработанные автоматизированные системы технологической подготовки. Ну а если в детали есть некие особенности конструкции или материалов, непредусмотренные технологической системой?

Тогда ему приходится искать решение самостоятельно – вот тут требуется элинга. Пользователь задает вопрос системе и пошла цепочка ответов в виде блоков взаимосвязанного текста на основе разных принципов, в том числе при логическом выводе – через связь понятий или использования имманентных свойств исходных текстов – когезии (локальной связности) и когерентности (связности в пределах текста). И хотя получаемый текст – совершенно искусственен (совместное творчество компьютера с пользователем), он обладает свойством интертекстуальности, является линейным (хотя получен нелиней-

ным образом) в основных блоках, и обладает свойством семантической связности [5,6,10].

В элинге возможна выдача значительной части БЗ даже по одному запросу. Но это не нужно, поскольку пользователь не в состоянии воспринять выдаваемые элингой огромные объемы знаний. А критерии решения часто и самому пользователю не ясны [1,7], и в задаче нет всех необходимых условий и ограничений.

В теории решения изобретательских задач (ТРИЗ) [1] и методе «мозгового штурма» – используется путь создания определенной среды и условий для эффективной работы по поиску решений для пользователя.

В элинге также создается своя специальная среда для ускоренного и эффективного нахождения рациональных решений.

Такой подход к восприятию решения задач высказал и Пойа [11] – «найти решение задачи – это значит установить связь между заранее дифференцированными объектами или идеями (...предпосылкой и заключением)...».

Назовем применяемый в элинге (а также в АСУЗ/БАНЗ) подход – диалого-ассоциативным поиском [2-6]. Каждый раз при подаче запроса сложно сказать, найдет ли пользователь необходимый ответ и каков он будет, но зато можно уверенно утверждать, что пользователь разберется в интересующем его вопросе намного лучше. И, одна из главных особенностей, – быстро, поскольку, не надо обращаться вначале – ни в другие книги, ни в другие библиотеки, ни к другим специалистам, ни в другие организации (конечно, при достаточной развитости элинги).

Описание одного из зарубежных конкурентов. В 2013 г. был опубликован патент *US 8583422 System and method for automatic semantic labeling of natural language texts* (Система и способ для автоматического семантического маркирования текстов естественного языка) [14], авторы Д. Тодхантер, И. Совпель, Д. Пастанохау из компании *Invention Machine Corporation* (ИМС, США, основные разработчики находятся в Минске, Беларусь). Основателем

компания *ИМС* является Цуриков В.М., разработчик программного комплекса «Изобретающая машина» (1988 г.), который поддерживал технологию ТРИЗ, созданную Г. Альтшуллером [1]. Теперь уже с помощью компьютерной системы ТРИЗ позволяла находить решения на уровне изобретений. Затем уже в США, основав фирму *ИМС*, Цуриков В.М. развил и преобразовал эту систему, получив постепенно десятки патентов в разных странах. В нулевые годы его основная система стала называться *Invention Machine Goldfire*. Одними из важных дополнительных возможностей стали лингвистический и семантический анализ текстов в Интернете и реферирование интересных результатов с выдачей пользователю [12,15]. Этот подход подходил для реферирования текстов, но конечно совершенно не мог конкурировать с возможностями подхода [2-6], в первую очередь, в связи с ограниченными возможностями используемых моделей знаний – *SAO*-структур. Однако эта группа патентов *ИМС* послужила хорошим опорным прототипом для получения патента [3]. В дальнейшем логика решаемых проблем подвела *ИМС* к необходимости расширения возможностей используемых моделей знаний, а соответственно возможностей логического вывода. И вот в новых патентах, в частности [15],

они сделали решительный скачок и перешли к расширенной модели знаний *eSAO*, которая по возможностям и размеру соответствует молинге. Кроме того, они впервые ввели в описание патента возможность участия эксперта в работе системы в диалоге с компьютером, что сразу заложено в элинге, в т.ч. зафиксировано в патенте [3]. При этом автоматический логический вывод, зафиксированный в патентах для почти одинаковых моделей знаний, тоже оказывается близким, даже прямо судя по описанию. Однако присутствует разная терминология и разное целевое назначение и все-таки разница в некоторых возможностях.

Надо отметить, что в 2012 г. компания *ИМС* была куплена компанией *IHS inc.* и теперь основное программное средство называется *IHS Goldfire*.

Заключение. В целом, можно констатировать то общее, что проявилось, – появление нового направления в развитии ИС – интеллектуальные системы создания нового смысла, что прямо обеспечивается в работе участием человека в процессе диалога элинги с пользователем, а у компании *IHS* уже поддерживается алгоритмической основой и фактически уже заложено в описании патента. Это все создает новые возможности для сохранения, передачи и использования знаний для массового пользователя.

Литература

References

1. Альтшуллер Г.С. Найти идею. Введение в теорию решения изобретательских задач. Новосибирск: Наука, 1986. 209 с.
2. Бронфельд Г.Б. Об одном подходе к построению интеллектуальных электронных книг и о молинге // Труды Восьмого международного симпозиума «Интеллектуальные системы». М.: РУСАКИ, 2008. С. 230–234.
3. Бронфельд Г.Б. Система для работы с интеллектуальной электронной книгой – элингой. Пат. № 2440610 РФ, опубл. 20.01.12.
4. Бронфельд Г.Б. Прямое наложение знаний и «практически вечное» сохранение знаний и информации // Труды Нижегородского государственного технического университета им. Р.Е. Алексева. 2013. №4. С.173–182.
5. Бронфельд Г.Б. Основы искусственного интеллекта: учеб. Пособие. Н. Новгород: Нижегород. гос. техн. ун-т. им. Р.Е.Алексева,
- Al'tshuller GS. Najti ideju. Vvedenie v teoriju reshenija izobretatel'skih zadach. Novosibirsk: Nauka; 1986. Russian.
- Bronfel'd GB. Ob odnom podhode k postroeniju intellektual'nyh jelektronnyh knig i o molinge. Trudy Vos'mogo mezhdunarodnogo simpoziuma «Intellektual'nye sistemy». Moscow: RUSAKI; 2008. Russian.
- Bronfel'd GB, inventors. Sistema dlja raboty s intellektual'noj jelektronnoj knigoj – jelingoj. Russian Federation patent RU 2440610. 2012. Russian.
- Bronfel'd GB. Prjamoe nalozhenie znaniy i «prakticheski vechnoe» sohranenie znaniy i informacii. Trudy Nizhegorodskogo gosudarstvennogo tehničeskogo universiteta im. R.E. Alekseeva. 2013;4:173-82. Russian.
- Bronfel'd GB. Osnovy iskusstvennogo intellekta: ucheb. Posobie. N. Novgorod: Nizhegorod. gos. tehn. un-t. im. R.E.Alekseeva; 2014. Russian.

2014. 253 с.

6. Бронфельд Г.Б. Прямое наложение знаний и его возможности. Анализ, методология, новая модель знаний, алгоритмы, возможности «невозможности». Saarbrucken, Deutschland: LAP LAMBERT Academic PUBLISHING, 2014. 236 с.
Bronfel'd GB. Prjamoje nalozhenie znanij i ego vozmozhnosti. Analiz, metodologija, novaja model' znanij, algoritmy, vozmozhnosti «nevozhnosti». Saarbrucken, Deutschland: LAP LAMBERT Academic PUBLISHING; 2014. Russian.
7. Брукс Ф.П., мл. Проектирование процесса проектирования: записки компьютерного эксперта. М.: ООО «И.Д. Вильямс», 2013. 464 с.
Bruks FP, ml. Proektirovanie processa proektirovanija: zapiski komp'juternogo jeksperta. Moscow: ООО «I.D. Vil'jams»; 2013. Russian.
8. Вихнин А.Г., Сакипов Н.З. Штурм четвертого мегапроекта: кто будет новым Биллом Гейтсом? Системный анализ и выбор стратегии. М.: Изд-во «Диалог-МИФИ», 2008. 288 с.
Vihnin AG, Sakipov NZ. Shturm chetvertogo megaproekta: kto budet novym Billom Gejtsom? Sistemnyj analiz i vybor strategii. Moscow: Izd-vo «Dialog-MIFI»; 2008. Russian.
9. Современный русский язык. Теория. Анализ языковых единиц: в 2 ч./ Диброва Е.И. [и др.]. М.: Изд. центр «Академия», 2008. Ч.1. 480 с.; Ч.2. 624 с.
Dibrova EI, et al. Sovremennyy russkij jazyk. Teorija. Analiz jazykovyh edinic: v 2 ch. Moscow: Izd. centr «Akademija»; 2008. Russian.
10. Лукин В.А. Художественный текст: Основы лингвистической теории. Аналитический минимум. М.: Изд-во «Ось-89», 2009. 560 с.
Lukin VA. Hudozhestvennyj tekst: Osnovy lingvisticheskoj teorii. Analiticheskij minimum. Moscow: Izd-vo «Os'-89»; 2009. Russian.
11. Пойа Дж. Математическое открытие (Решение задач: основные понятия, изучение и преподавание). М.: Издательство «Наука», 1976. 448 с.
Poja Dzh. Matematicheskoe otkrytie (Reshenie zadach: osnovnye ponjatija, izuchenie i prepodavanie). M.: Izdatel'stvo «Nauka», 1976. 448 s.
12. Совпель И.В. Система автоматического извлечения знаний из текста и ее приложения // Искусственный интеллект. 2004. № 3. С. 668–677.
Sovpel' IВ. Sistema avtomaticheskogo izvlechenija znanij iz teksta i ee prilozhenija. Iskusstvennyj intellekt. 2004;3:668-77. Russian.
13. Gray J., Czalay A. The World – Wide Telescope, an Archetype for Online Scienc, 2002. 6 p.
Gray J, Czalay A. The World – Wide Telescope, an Archetype for Online Scienc; 2002.
14. Todhunter J., Sovpel I., Pastanohau D. US Patent No 8,583,422 System and method for automatic semantic labeling of natural language texts (12.11.2013)
Todhunter J, Sovpel I, Pastanohau D; inventors. US Patent No 8,583,422 System and method for automatic semantic labeling of natural language texts (12.11.2013)
15. Tsourikov V., Batchilo L., Sovpel I. US Patent No 6,167,370 Document semantic analysis/selection with knowledge creativity capability utilizing subject-action-object (SAO) structures (26.12.2000)
Tsourikov V, Batchilo L, Sovpel I; inventors. US Patent No 6,167,370 Document semantic analysis/selection with knowledge creativity capability utilizing subject-action-object (SAO) structures (26.12.2000)