

Моделирование семантики научно-технических текстов для АИС и его теоретические основы

©Лаврёнова О.А.

Российская государственная библиотека
lavr@rsl.ru

Аннотация

Цель доклада - обратить внимание современных разработчиков электронных библиотек на структуру классической системы терминологии информатики и лингвистики в области исследования проблем представления семантики текстов для АИС. Предлагается способ анализа библиотечных классификаций, тезаурусов, и несложных онтологий на базе модели знаний, которая основана на некоторых представлениях о целесообразной деятельности людей.

1 Ситуация

1.1. Несколько десятилетий использования тематического поиска в автоматизированных информационных системах (АИС) не демонстрируют, к сожалению, его существенного развития. Напротив, классический порядок в системе терминологии информатики забывается – и происходит путаница в понятиях. Важное, существенное утрачивается – и во главу угла ставится упрощение всех проблем – и тем самым возвращение к ситуации первых опытов 60х – 70х годов [1]. Разработчиков современных АИС окупают огромные вычислительные возможности современных компьютерных систем и богатство информации в сети Интернет, но при этом в публикациях снова мелькают выводы тех лет, что свободный поиск по КС из полных текстов документов (подумать только!) обуславливает огромный информационный шум и, одновременно, скрытую от пользователя потерю информации. Проблема тематического поиска в электронных библиотеках (ЭБ) отличается от таковой в сфере создания электронных каталогов (ЭК) не столь существенно, как представляется многим, тем более что поиск в ЭБ преимущественно основан на тех же библиографических записях (БЗ) в ЭК или соответствующих им по содержанию метаданных, структурированных иными способами (например, с помощью языка разметки текстов XML).

Существовавшее в информатике понятие “поиск по ключевым словам (КС)” в процессе развития сети Интернет незаметно изменило своё первоначальное значение. Ранее КС приписывались документам вручную или отбирались автоматически из рефератов, аннотаций, заглавий документов. Таким образом, и при автоматическом индексировании в поисковый образ документа, при исключении «стоп-слов», попадали преимущественно слова и словосочетания, отображающие его семантическое содержание. Соответственно, свободные КС или КС из фиксированных перечней считались средством тематического поиска. Современная ситуация уже не выглядит столь очевидной. Индексные файлы (индексы) строятся в ЭК не только на основе тех полей БЗ, которые описывают семантическое содержание документов. Нередко в поисковых системах, властвующих в сети Интернет, поиск по КС означает поиск во всех элементах БЗ или в полных текстах разнородных документов. На этом фоне каждый пользователь интерпретирует “поиск по КС” (и в электронных каталогах, не только в полных текстах) в меру своего опыта: один – как свободный поиск по всем данным, другой – как тематический.

1.2. На этом фоне исследование методов выбора КС, анализ структуры информационно-поисковых тезаурусов, библиотечных классификаций и тщательно проработанных современных онтологий [2-5] как «попытка подтащить смысл данных к самим данным» «при должном уровне формализации» [3] в наше время выглядит весьма актуальным.

При этом обращение к «классике жанра», некоторая «ретроспектива» изучения проблемы представления семантики текстов для АИС, возможно, внесет для целого ряда специалистов некоторую ясность в структуру проблемы.

2 Пресуппозиция текста и абстрактный тезаурус

В связи с решением задачи моделирования семантики текста как носителя информации особое значение имеет вопрос о том, в каком виде преподносится автором информации (знание) для адресатов, где она “заложена”. Дело в том, что, стремясь при создании текста обеспечить адекватную мыслительную деятельность адресата,

автор, как правило, старается обеспечить это минимальными средствами. В данном вопросе ведущую роль играет **понятие пресуппозиции**, заимствованное из философии. Среди лингвистов, похоже, первыми его применили Дж. Катц и П. Постал [6]. Термин “пресуппозиция” трактуется в разных работах самым различным образом.

В общем случае под пресуппозицией в лингвистике понимается “молчаливая предпосылка”, “предварительный договор”, “условия удачи” (happines conditions), которые позволяют речевому высказыванию реализовать свою коммуникативную функцию, но не находят в нем непосредственного выражения, или внеречевые условия, обеспечивающие правильное понимание речевых актов или высказываний [10]. При построении текста часть элементов отображаемого фрагмента системы знаний получает вербальную экспликацию, а большая часть остается в области пресуппозиций. Получивших непосредственное выражение в тексте элементов и знаний должно быть достаточно для понимания текста при восприятии теми адресатами, для которых создается речевое произведение, т.е. и для выявления пресуппозиций текста. При этом речь идет о восприятии не только нового для адресата знания (информации), но и ранее известного.

Систему знаний о мире в плане решения прикладных семантических задач удобно представлять как некий **абстрактный тезаурус**. Наиболее близкой к нашим представлениям является в этом смысле концепция Ю.А.Шрейдера [7 и др. работы]. Для выявления знаний, составляющих пресуппозиции текста, необходимо как минимум их существование в тезаурусе адресата. Здесь мы имеем дело с трактовкой пресуппозиции как **“фонда общих знаний”**.

Для выделения пресуппозиций и в философии, и в лингвистике используют критерий отрицания. Применительно к тексту это может выглядеть как утверждение: “Неверно все, о чем говорится в данном тексте”. При этом отрицание затронет лишь непосредственно выражение в тексте содержание (“план содержания”), а пресуппозиции останутся неизменными.

Для полноты картины важно различать абстрактный тезаурус человеческого общества (универсальная система знаний человечества о мире действительности) и его «подсистемы» [8]:

- тезаурус группы специалистов данной области знаний (система знаний в данной области науки, техники, культуры и т.д.);
- тезаурус индивидуума, являющегося автором текста (система знаний автора);
- фрагмент тезауруса автора, отображаемый в семантике текста (фрагмент системы знаний автора, передаваемый другим индивидуумам с помощью данного текста);
- тезаурус индивидуума - адресата, слушающего, читающего, т.е. воспринимающего данный текст (система знаний адресата);

- места или точки в тезаурусе воспринимающего индивидуума, требующие заполнения, т.е. пустые места (потребности воспринимающего в определенных знаниях или, как принято говорить в информатике, информационные потребности);
- фрагмент тезауруса адресата (возможно, дополненный и перестроенный под воздействием воспринимаемого текста)
- фрагмент системы знаний адресата, который в его представлении соответствует смысловому содержанию текста.

В принципе, можно себе представить передачу информации в виде полных фрагментов тезауруса автора, записанного в некотором смысловом коде. Тогда потеряло бы смысл использование пресуппозиций, позволяющих часть знаний оставлять в “подтексте”. Несмотря на то, что человек, в конечном счете, запоминает именно такие семантические кодовые последовательности [9] для передачи знаний он использует специфические языковые способы, которые обеспечивают экономию используемых средств и ограничение передаваемой каждым сообщением информации.

Моделирование семантики текста для хранения в памяти АИС должно обеспечивать **построение полного семантического представления текста, необходимого для решения определенной мыслительной задачи**, используя и восстановление необходимых пресуппозиций текста.

3 Структура человеческой деятельности и способы семантического представления текстов

3.1. Очевидно, существует множество разнообразных способов фиксации в мозгу человека знаний (различные способы организации абстрактного тезауруса). Каждый элемент знания включается в целый ряд разнообразных систем связей с другими элементами. Вид (структура, состав элементов) таких систем зависит как от содержательной стороны элемента (соотнесенности с тем или иным видом объектов действительности) и уровня сложности знаний, так и от целей, ради которых данное знание фиксируется в памяти. Строение систем знаний, как и вообще строение сознания человека находится в прямой связи со строением его деятельности [11].

Таким образом, **всякое представление знаний должно рассматриваться в структуре соответствующей деятельности человека**, и, таким образом, каждый способ создания машинного тезауруса конкретной АИС, т.е. способ моделирования коллективного тезауруса специалистов в соответствующей области знания, должен базироваться на определенной гипотезе относительно организации знаний в человеческой памяти для решения данной мыслительной задачи в

соответствующей области знания [8, 12, 13]. Важность, существенность сведений, данных об объекте определяется в связи с потребностями человека и может меняться даже в процессе решения одной задачи. Особенности методов моделирования мыслительных процессов существенным образом зависят от типа задач, решаемых человеком [14].

3.2. Результатом **понимания текста** является построение в памяти адресата его глубинной семантической структуры, заключающееся в определении места воспринятых знаний в тезаурусе адресата, т.е. **интерпретация информации в уже организованной картине мира**. Если в качестве адресата текстов документов рассматривать АИС, то понимание как положительный результат восприятия будет заключаться в правильном определении места зафиксированных в тексте элементов и отношений в машинном тезаурусе. Элементы последнего, способные заполнить пустые или уточнить сомнительные места тезауруса адресата, представляют собой информацию для потребителя.

При наличии разнообразных способов организации универсального тезауруса необходимость наиболее точного определения места семантических представлений текстов и запросов в АИС еще раз подтверждает **целесообразность задания определенного способа организации их системы знаний, соответствующей задаче информационного поиска в них**. Кстати, известный Крэнфилдский эксперимент по индексированию, проводившийся в течение нескольких лет Ассоциацией специальных библиотек и информационных бюро Великобритании и не подтвердивший распространенного мнения, что специальные знания важнее навыков индексирования [15] свидетельствует, надо полагать, о том, что выработанный в процессе выбора индексационных элементов взгляд на определенное устройство системы знаний АИС, наиболее удобное для поиска, важнее наличия в мозгу индекса разнообразных структур знаний в данной области. При этом надо отметить, что этот эксперимент позволил сделать вывод: текст монографии при простом полнотекстовом поиске выдается почти на каждый запрос из данной области знания. Таким образом, уже в то время был дан ответ на распространенный ныне вопрос в сфере создания ЭБ: а не запустить ли нам тематический поиск по массиву полных текстов без всяких тезаурусов и классификаций?

4 Принципы моделирования семантики текстов

4.1. Итак, **вид модели знаний определяется не только характером моделируемого объекта, но и спецификой задач, для решения которых она предназначена**. В качестве той задачи, с точки зрения решения которой разрабатывались принципы моделирования семантики текстов,

предложенные нами [8, 12, 13], выступает автоматизация процессов поиска научно-технической информации в АИС, разновидностями которых являются, в частности, электронные каталоги и электронные библиотеки. Способы моделирования семантики текстов для АИС выступают на практике в виде различных видов информационно-поисковых языков (**ИПЯ**). В соответствии с принятой в АИС структурой ИПЯ создается модель знаний некоторого определенного вида. Концепция относительно организации знаний АИС и определяет так называемую "существенность" информации, выбираемой для ввода в систему. Многообразие возможных способов отбора информации по различным критериям существенности обуславливает вариативность представления смыслового содержания текста в целях его индексирования.

4.2. Из сказанного выше естественным образом выводится следующий путь моделирования семантики текстов научно-технического содержания в целях автоматизации информационного поиска: во-первых, необходима обоснованная **гипотеза относительно способа организации человеческих знаний, который используется при решении задачи поиска научно-технической информации в текстах массива документов в соответствии с информационной потребностью, т.е. необходима соответствующая модель системы знаний**; во-вторых, именно **таким способом должны быть организованы знания АИС**, причем как базисные знания (тезаурус), так и заключенные во вводимых текстах фрагменты знаний.

Разработка предлагаемого способа моделирования опиралась на анализ структуры, видов информации в различных областях науки и техники, а также типичных структур информационных потребностей. Данный способ моделирования семантики текстов научно-технического содержания был разработан в свое время для создания ИПЯ АИС, и его следует рассматривать как один из возможных. В основу способа моделирования семантики текстов положена определенным образом организованная модель системы знаний, в качестве фрагментов которой рассматриваются семантические представления текстов - так называемые поисковые образы документов.

Наиболее естественным нам представляется строить модель системы знаний в соответствии с той последовательностью, в которой эта система строится в процессе научного познания человеком окружающей действительности. **В основу построения гипотетической модели представления знаний положены представления о целесообразной деятельности людей**, направленной на удовлетворение исторически сложившихся потребностей человека и общества, и базирующийся на них принцип "предназначения".

Процесс познания стимулируется определенными мотивами, лежащими, в конечном счете, в практической деятельности людей. **Мотив**,

направляющий деятельность, выражает потребность человека в предмете деятельности. **Вся познавательная деятельность мотивируется разнообразными потребностями личности.** Воплощением определенной потребности является **цель деятельности** как идеальный образ предмета потребности. Отличие одной деятельности от другой заключается, в первую очередь, в различии их предметов, которые составляют действительный мотив деятельности и могут быть как вещественными, так и идеальными. Для реализации цели деятельности необходимы определенные **средства**, которые далеко не всегда имеются в готовом виде, а чаще всего должны быть найдены и сами становятся **предметом потребности**, т.е. целью, для достижения которой отыскиваются средства, и т.д.

Ситуации, когда готовые способы, действия, средства неэффективны и необходимо выработать новые способы, действия, средства, называют проблемами или **проблемными ситуациями**, которые характеризуются различием между сложившейся ситуацией и желаемой. В результате познавательной деятельности (в процессе научного исследования, конструирования, проектирования и т.д.) отыскиваются некоторые средства, способы решения научной или технической проблемы. Например, могут быть определены материальные средства или некоторые теоретические методы, разработаны технологические процессы и т.д., позволяющие удовлетворить потребность человека в определенном предмете.

В процессе своей деятельности человек одновременно с созданием предметов создает и совокупность соответствующих представлений об этих предметах и процессах их производства. Познавательный процесс, развиваясь, надстраивает разнообразные разветвленные структуры знаний об окружающей действительности. Таким образом, можно сделать предположение общего характера, что в процессе удовлетворения разнообразных потребностей людей (отдельной личности или общества) **в системе знаний** фиксируются, во-первых, эти **потребности** в качестве некоторого "пустого места", временно занимаемого приблизительным идеальным образом предмета потребности. Во-вторых, в систему знаний включаются **виды деятельности**, их составляющие (действия, операции), характеристики, условия, сопутствующие явления. В-третьих, в результате теоретического или практического получения предмета, удовлетворяющего потребность, "пустое место" заполняется знанием об этом предмете, его свойствах, условиях существования, связанных с ним явлениях. **"Предметы"** выделяются в отдельную категорию как множество "самостоятельных сущностей" в противоположность "признакам", к которым относятся отношения, действия, состояния, характеристики и т.п., т.е. все то, что не относится к числу самостоятельных сущностей, а лишь присуще чему-то другому. Употребляемое в этом смысле понятие "предметы" следует отличать от понятия

"предмет мысли" или "предмет познания". Описания других элементов модели даны в работах [8, 12, 13].

5 Структура модели системы знаний

Принципиальная структура предлагаемой модели универсальной системы знаний основывается на этой общей гипотезе. В качестве "точки отсчета" ("ядра" модели) рассматривается человек (или человеческое общество), т.е. субъект, обладающий биологическими и социальными потребностями, как первая предпосылка всякой деятельности. Допустим, некие процессы, реализующие деятельность по обеспечению существования человека и его свойств, функций, состояний, помещаются в модели универсального тезауруса в некоторый n -й "слой" (например, *кормление, защита, воспитание*). Они имеют свои характеристики, производятся в некоторых условиях, сопровождаются различными явлениями. Необходимость осуществления процессов из n -го «слоя» обуславливает возникновение потребности в некоторых "предметах" (например, *пища, жилище, одежда, воспитатели, книги, приборы* и т.п.). Предметы следующего, $n+1$ -го слоя, обеспечивающие осуществление процессов n -го слоя, также обладают определенными характеристиками, существуют в некоторых условиях, связаны с определенными явлениями, могут находиться в различных состояниях. Для обеспечения существования предметов $n+1$ -го слоя необходимы также различные виды деятельности, которые помещаются в $n+2$ -й слой, представляющий собой слой процессов (например, *строительство, разработка, приготовление, перемещение* и т.д.) с их характеристиками, условиями, сопутствующими явлениями и т.д.

Опыт показал, что **в процессе такого "надстраивания" системы знаний** (в указанном выше смысле) образуются поочередно как бы **«слои» процессов деятельности**, задачей которых является обеспечение получения, существования предметов потребности с определенными свойствами, ликвидация препятствующих факторов или их использование для удовлетворения потребностей, и **«слои» предметов потребности**, т.е. необходимых для осуществления этой деятельности, с их характеристиками (свойствами, параметрами), условиями, сопутствующими явлениями, в том числе отрицательными. Кроме основных слоев - процессов и предметов - необходимо выделить в модели дополнительные слои процессов, осуществляемых по отношению к другим процессам, их условиям, характеристикам, явлениям (например, организация, улучшение, оптимизация, ускорение процесса, создание условий для процесса, повышение показателей процесса и т.д.). Элемент слоя **обеспечивает существование (осуществление), ликвидацию (уменьшение, изоляцию и т.п.) или использование (для отрицательных явлений) хотя бы одного элемента нижележащего «слоя».**

При этом понятие "обеспечение существования" означает обеспечение не только простого существования элемента модели, но и выполнения им его собственного предназначения, сохранения или создания его характеристик. Каждый отдельный слой представляет собой не простой набор предметов или процессов с их характеристиками, условиями, состояниями и т.д., а целую систему элементов, находящихся в определенных отношениях различного рода. Более подробно модель описана в работах [8, 12, 13] с использованием аппарата теории графов.

Рассмотрим простой пример на рис.1. В документе идет речь о предмете "тренажер", предназначенном для осуществления процесса тренировки, который предназначен для обеспечения требуемых характеристик предмета "экипажа" предназначенного для осуществления процесса эксплуатации предмета "самолет". Если бы речь шла о пассажирском самолете, можно было бы добавить ниже процесс "перевозка", а затем – предмет "пассажиры". Слои разделяются горизонтальными линиями. Такое описание смысла текста документа предупреждает пропуски семантических элементов (например, КС). Получается некая цепочка элементов модели.

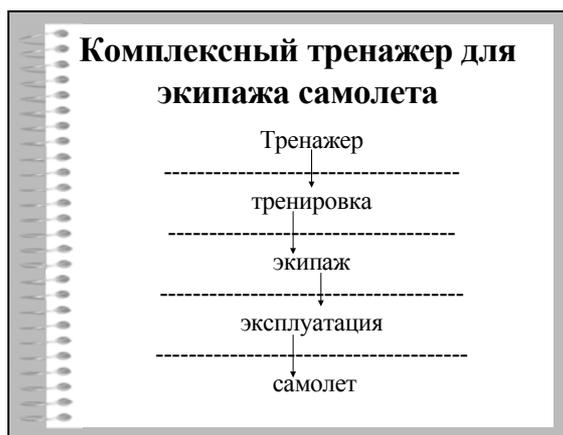


Рис. 1. Пример построения модели смысла текста

На рис. 1 все элементы модели находятся в отношении «предназначения» («для»).

В каждой цепочке, описывающей содержание документа, выделяется основной объект рассмотрения (ООР) как основа этой цепочки. Например, в цепочке на рис. 1 роль ООР может выполнять предмет «тренажер» («Тренажеры для тренировки экипажей самолетов») или процесс «тренировка» («Тренировка экипажей самолетов на тренажерах»).

На рис. 2 показаны основные типы отношений между элементами модели.

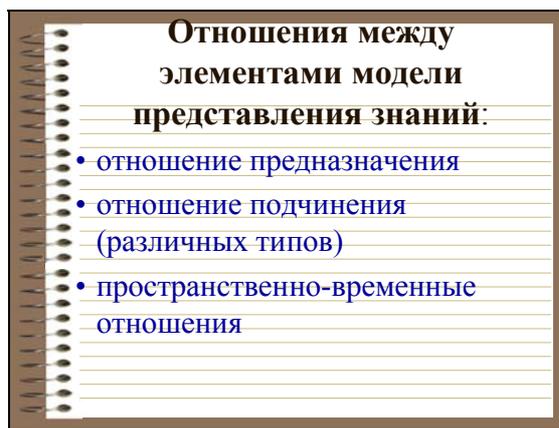


Рис. 2. Типы отношений между элементами модели.

6 Применение способа представления семантики текстов в АИС

Опыт показывает, что структура модели может быть использована для решения следующих задач:

- разработка методики выбора КС для ЭК и ЭБ (реализовано в свое время в методиках индексирования документов для двух отраслевых АИС);
- разработка методики построения информационно-поисковых тезаурусов [16] (использована в различных АИС);
- специальные исследования принципов организации тематического поиска в электронных каталога, т.е. возможностей различных ИПЯ [1] (проведено по отношению к ЭК РГБ и других библиотек в рамках работы, поддержанной РФФИ – проект № 98-07-90037);
- реализация структуры модели при построении онтологий для электронных библиотек на некотором «верхнем» структурном уровне.

В сущности, в целом ряде работ предлагается считать онтологиями широкий спектр структур, представляющих знания о той или иной предметной области [4, 5].

7 Анализ примеров

В качестве примеров проанализируем с помощью предлагаемой модели представления знаний так называемые цепочки словесных формулировок индексов Библиотечно-библиографической классификации (ББК) – нашей национальной библиотечной классификации. Такие цепочки вносятся систематизаторами РГБ в библиографические записи универсального ЭК. Основные фрагменты этих цепочек присутствуют в

машиночитаемых таблицах ББК. При обработке изданий и рукописей, поступающих в фонд Библиотеки, в библиографических записях строятся более сложные индексы с использованием разнообразных структурных элементов ББК, причем взятых из разных иерархических деревьев. Естественно, словесные формулировки этих индексов строятся из элементов таких деревьев. Кроме того, эти же средства ИПЯ используются в **электронной библиотеке диссертаций (ЭБД) РГБ**. Они включены в библиографические записи ЭК, а также в **структуру метаданных открытой ЭБД** на основе ИСИР [1, 18], где представляются в XML-разметке (технология реализуется в РГБ при поддержке РФФИ - проекты №01-07-90310 и 04-07-90154).

Классификационные таблицы формируются из различного рода структурных элементов. Надо отметить, что обычно в классификациях не придается существенного значения семантическим категориям их элементов, как принято в информационно-поисковых тезаурусах. Интересно, что, строя структуру области знаний и оперируя словосочетаниями, обозначающими области деятельности, создатели библиотечных классификаций на некотором уровне начинают оперировать словами и словосочетаниями, которые по своей семантике, строго говоря, представляют собой элементы совершенно других семантических категорий, которые не назовешь наименованиями отраслей знания. Развитие делений может происходить по процессам, свойственным соответствующей области деятельности, по предметам, для обеспечения существования которых процессы предназначены, по предметам, которые являются средствами реализации этих процессов, по характеристикам или явлениям. Понятно, что любое деление можно структурировать и анализировать в соответствии с описанной выше моделью.

Для иллюстрации ниже приведен **ряд примеров словесных формулировок индексов ББК из ЭК диссертаций РГБ** (они же используются в метаданных для диссертаций в ЭБД). Каждая цепочка представляет собой в основном ветвь иерархического дерева ББК, начиная с самого верхнего уровня таблиц и кончая наиболее узким делением. Нередко нижнее деление может быть или не быть последним уровнем иерархии, отображая некоторый аспект темы. Для некоторых элементов цепочки строятся собственные дополнительные ветви отдельного иерархического дерева, представленного в ББК таблицами типовых делений (например, территориальных, специальных и т.д.).

По сути, классификационные таблицы представляют собой также некую модель системы знаний, предназначенную для поиска документов в каталогах библиотек. Для поисковых задач в АИС предназначены и тезаурусы, от которых не отличишь многие фрагменты классификационных таблиц, и языки предметных рубрик, которые при хорошем развитии смысловых отношений

приближаются к тем же тезаурусам, а без них не могут обеспечить полноту информационного поиска. На самом деле, самое важное значение имеет различие между отношениями синтагматическими (текстовыми, создаваемыми при построении текстов) и парадигматическими (зафиксированными в модели знаний заранее и постоянно). Семантическое представление текста строится в координатной системе модели знаний, и для относительно полного представления смысла текста необходимо восстановление его пресуппозиций, релевантных для соответствующих поисковых задач. При этом модель знаний должна иметь средства определения структурных элементов, которых не хватает для задуманного семантического представления текста, как в плане парадигматики, так и в плане синтагматики.

Для краткости **рассмотрим отдельные примеры**, что должно оказаться достаточным для того, чтобы читатель сам смог оценить структурные особенности цепочек в других примерах.

В примере (1) непосредственное содержание текста упрощенно выражено ключевыми словами из заглавия (см. поле 653). Для увеличения полноты поиска эти слова размещаются систематизаторами в семантической модели ББК. Проверим, все ли связи учтены с точки зрения нашей модели знаний:

- *свойства* – «отношение подчинения» - *дорожные покрытия* – «отношения подчинения» - *автозимник* («синоним для» *снежные и ледовые дороги*);
- *срок службы* - «отношение подчинения» - *эксплуатационные параметры* - «отношение подчинения» - *надежность и долговечность* - «отношение подчинения» - *свойства*;
- *фрикционный слой* «отношение подчинения» – *дорожные покрытия* – «отношение подчинения» - *автомобильные дороги*;
- *автомобильные дороги* – «отношение предназначения» - *эксплуатация* - «отношение предназначения» - *автодорожный транспорт* – «отношение подчинения» - *транспорт*.

Получается, что использование КС и ББК позволило построить аккуратное семантическое представление текста, правда, с пропуском элементов «эксплуатационные параметры» и «эксплуатация». В примере (2), кстати, эти элементы присутствуют, но зато нет некоторых КС из заглавия. Кроме того, во второй цепочке пропущен процесс «строительство».

В примере (3) пропущен процесс «лечение» и не использованы элементы из заглавия для развития более низких уровней иерархии понятий, чем предусмотрено в классификации.

Примеры показывают, каким причудливым образом могут формироваться области знания в делениях классификаций. Так, в примере (4) «работают» 3 иерархических дерева парадигматических отношений: *области деятельности, языки, процессы* в данной области деятельности. Строго говоря, языкознание здесь подразделяется на основе дерева языков таким образом: «Изучение и преподавание

индоевропейских языков», «Изучение и преподавание германских языков» и т.д., но структурное решение сделано в пользу одновременного использования разных иерархических деревьев.

(1) *Влияние свойств покрытий автозимников на срок службы фрикционного слоя : Дис. ... канд. техн. наук : 05.23.11*

650 7 |а Транспорт -- Автодорожный транспорт -- Автомобильные дороги -- Снежные и ледовые дороги -- Дорожные покрытия -- Надёжность и долговечность

653 |а автозимники |а фрикционный слой дорожных покрытий |а срок службы

(2) *Определение транспортно-эксплуатационных параметров автомобильных дорог с асфальтобетонными покрытиями на основе шлаковых материалов : Дис. ... канд. техн. наук : 05.23.11*

650 7 |а Транспорт -- Автодорожный транспорт -- Автомобильные дороги -- Асфальтобетонные и дёгтебетонные дороги -- Эксплуатация -- Исследование эксплуатационных параметров

650 7 |а Транспорт -- Автодорожный транспорт -- Автомобильные дороги -- Дорожно-строительные материалы -- Дорожно-строительные материалы на основе органических и неорганических вяжущих -- Асфальтобетон и дёгтебетон -- Исследование

(3) *Особенности функциональной межполушарной асимметрии детей, депривированных по слуху : Дис. ... канд. биол. наук : 03.00.13*

650 7 |а Биологические науки -- Физиология, биофизика и биохимия животных и человека - - Физиология, биофизика и биохимия нервной системы -- Центральная нервная система -- Кора головного мозга -- Динамическая локализация функций в коре -- Возрастная физиология

650 7 |а Здравоохранение. Медицинские науки -- Педиатрия -- Болезни детей -- Оториноларингология детского возраста -- Болезни уха у детей -- Расстройства слуха у детей

(4) *Формирование лингвокультуры студента в образовательном процессе : Дис. ... канд. пед. наук : 13.00.01*

650 7 |а Филологические науки. Художественная литература -- Языкознание -- Индоевропейские языки -- Германские языки -- Западногерманские языки -- Английский язык -- Преподавание и методика изучения языка -- Методика преподавания отдельных аспектов языка

Коды полей MARC21, используемые в примерах:

650_7 - поле для тем; в подполе |а – цепочка словесных формулировок одного индекса ББК; 653 – поле для свободных КС.

В первой строке примера приводится заглавие диссертации.

В структуру метаданных ЭБД индекс и цепочка его словесных формулировок для диссертации или автореферата встраиваются следующим образом (из примера 1):

```
<bib:bbk> <pcv:code> (индекс ББК) </pcv:code>
<pcv:label> Транспорт -- Автодорожный транспорт -
- Автомобильные дороги -- Снежные и ледовые
дороги -- Дорожные покрытия -- Надёжность и
долговечность </pcv:label> </bib:bbk>
```

Литература:

- [1] Лавренова О.А. Тематический поиск в электронных каталогах и электронных библиотеках // Библиотечное дело. – 2004. - №5. - С. 42-50
- [2] Briukhov D.J., Ralinichenko L.A., Zakharov V.N. Diversity of Domane Descriptions in Natural Science: Virtual Observatory as a Case Study // Электронные библиотеки; перспективные методы и технологии, электронные коллекции: Труды седьмой всероссийской научной конференции RCDL'2005 (Ярославль, 4-6 октября 2005 г) – Ярославль, Ярославский гос. ун-т, 2005 - С.23-30
- [3] Марчук А.Г. Semantic Web – на пути к новому поколению информационных систем // Там же – С.122-125
- [4] Добров Б.В., Лукашевич Н.В., Синицын М.Н., Шапкин В.Н. Разработка лингвистической онтологии по естественным наукам для решения задач информационного поиска // Там же. – С.70-79
- [5] Welty C., MCGuinness D., Uschold M., Gruninger M., Lehmann F. Ontologies: Expert Systems all over again. AAAI-1999 Invited Panel Presentation. 1999
- [6] Katz J.J., Postal P.M. An integrated theory of linguistic descriptions. - Cambridge (Mass.): 1965. - 178 pp.
- [7] Шрейдер Ю.А. Информация и метаинформация // Научно-техническая информация б сер.2. – 1974. – С. 3-10
- [8] Лавренова О.А. Моделирование семантической структуры текстов научно-технического содержания в связи с автоматизацией информационных процессов. Дис. ... канд. филол. наук. -. М.: Московский гос. Университет, 1978. – 280 с.
- [9] Кибрик А.Е. Семантическая проблематика гетерологического кодирования // Теоретические проблемы прикладной лингвистики. – М.: 1965. – С. 67-83

- [10] Звегинцев В.А. Предложение и его отношение к языку и речи. – М.: Моск. Ун-т, , 1976. – 308 с.
- [11] Леонтьев А.Н. Проблемы развития психики. – М.: Мысль, 1965. – 572 с.
- [12] Лавренова О.А. Модель системы знаний, используемая в автоматизированных информационных системах // Научно-техническая информация, сер.2. – 1989. - №. - С.5-9
- [13] Лавренова О.А. Семантическое представление текста на основе модели системы знаний. - Научно-техническая информация, сер.2. – 1984. - №4. - С.18-24.
- [14] Напалков А.В. Эвристическое программирование. – Ростов н/Д: 1971. – 127 с.
- [15] Ланкастер Ф., Миллз Дж. Испытание эффективности указателей и вспомогательных средств языка индексирования: Крэнфилдский эксперимент ASLIB // Сборник переводов по вопросам информационной теории и практики, вып. 9. – М.: ВИНТИ, 1967. – С. 33-42
- [16] Лавренова О.А. Методика построения информационно-поискового тезауруса. //Библиотечные компьютерные технологии: Сб. статей. Вып.1-2./ Рос. гос. б-ка. – М., 1997. – С. 21-27.
- [17] Лавренова О.А., Аветисова Т.В. Электронные каталоги Российской государственной библиотеки – реальность // Библиотековедение. – 2000. - №2.
- [18] Лавренова О.А., Вежневцев А.А. Структура и реализация электронной библиотеки диссертаций в РГБ. // Электронные библиотеки; перспективные методы и технологии, электронные коллекции: Труды пятой всероссийской научной конференции RCDL'2003 (Санкт-Петербург, 29-31 октября 2003 г.) – Санкт-Петербург, НИИ Химии СПбГУ, 2003. – С. 308-314