

Формальная модель семантики художественного текста (на материале русских пословиц)¹

Малкова А. С., Январев В. И.*

Российский государственный гуманитарный университет

125993, Москва, Миусская пл., 6

e-mail: asmalkova@gmail.com

**Московский государственный университет им. М.В. Ломоносова*

119991, Москва, Воробьевы горы

e-mail: janvarev@gmail.com

получена 22 ноября 2006

Аннотация

Предпринята попытка построения формальной модели, отражающей суть моральных предписаний, заключенных в текстах русских пословиц. Модель включает четыре компонента: две оппозиции, связывающее их т.н. “мифологическое” тождество и конфигурацию связи. Модель позволяет применять процедуры интеллектуальной обработки текстов: проводить статистический анализ устойчивых (часто воспроизводящихся) связей *оппозиции + тождество*; осуществлять информационный поиск по семантике текста с возможностью формулирования сложных запросов, например выдать пословицы о “негативных последствиях глупости”. На основе модели была реализована информационно-аналитическая система МЕТАФОРА, находящаяся в открытом доступе в сети Интернет по адресу <http://www.metaphora2.ru/>.

1. Введение

В настоящее время актуальной является задача структурного представления данных, изначально имеющих в неструктурированном виде (в виде текста, написанного на естественном языке). Структурное описание позволяет применять к тексту алгоритмы автоматической обработки содержащейся в нем информации, в том числе осуществлять интеллектуальный поиск, базирующийся на семантическом наполнении текста (а не на его лексическом составе), также применять алгоритмы статистического анализа на больших объемах текстов и проводить сравнительные (компаративные) исследования с применением точных методов.

Подход, ориентированный на семантику, на современном этапе предполагает создание онтологии предметной области – структурного описания, в котором содержательная информация представлена в виде концептуальной схемы: иерархически построенного графа формальных объектов различных классов. Каждому объекту может быть приписан некоторый набор атрибутов, рёбра графа интерпретируются как отношения между объектами. В процессе анализа текста из него вычленяется актуальная информация и на ее основе формируется фрагмент онтологии.

Таким образом, работа по формализации данных предполагает, в первую очередь, проектирование онтологии (описание библиотеки классов и отношений, а также набора правил и ограничений, которые необходимо учитывать при заполнении онтологии реальными фактами). Поскольку содержание текстов, как правило, направлено на описание некоторых явлений, имеющих отношение к вне-текстовой реальности, при проектировании онтологии за основу берется актуальный фрагмент физического мира, а прообразами объектов становятся физические объекты с их реальными отношениями.

Иначе обстоит дело, если предметом текстов являются абстрактные категории, границы которых к тому же четко не определены. В этом случае сами принципы выделения классов и объектов становятся нетривиальной интеллектуальной задачей.

Именно такая ситуация складывается с художественными текстами. Художественный текст несет в себе особого рода информацию, имеющую отношение к духовному миру человека: реализацию морального выбора, постановку и разрешение этических и социальных конфликтов и т.п. Выявление границ и номенклатуры этих специфических “духовных” объектов позволило бы привлекать точные методы для обработки художественных текстов: осуществлять эффективный информационный поиск, оперируя на

¹Работа выполнена при поддержке РГНФ (грант 07-03-00658а).

уровне глубинной семантики текста, подвергать тексты статистическому анализу, выявляя психологические матрицы, лежащие в основании сюжетной канвы.

Целью настоящей работы является попытка разработать формальные принципы структурного представления содержательной информации, заложенной в художественном тексте. Для достижения поставленной цели привлекался опыт исследования поэтики художественного текста, имеющийся в гуманитарных науках: филологии, этнографии, отчасти психологии.

2. История вопроса

История структурных исследований художественного текста насчитывает уже больше трех четвертей века. Но протекала она до настоящего момента почти исключительно в рамках теоретической филологии и имела прикладное значение для внутринаучных филологических задач, таких как классификация материала, частные компаративные изыскания и т.п. Сегодня настал тот момент, когда открытия филологов могут быть востребованы для решения широкого круга прикладных задач (семантического поиска, реферирования и аннотирования художественных текстов, машинного перевода, кросс-культурных исследований и т.п.).

Структурный подход в филологии развивался в рамках фольклористики и этнографии. Фольклор вырастает из мифологии, а мифология, в свою очередь, является одновременно самой базовой и самой примитивной формой художественного текста. Именно в мифе впервые складываются те сюжетные коллизии, та этическая проблематика, которая получает дальнейшее развитие в мировой литературе вплоть до настоящего времени. Мифологические сюжеты универсальны, они не являются достоянием какого-то одного народа, а распространены на всех континентах. Это свидетельствует о том, что в них находят отражение глубинные психологические основы человеческой личности.

Одновременно с этим, мифологическая фабула относительно проста (по сравнению с современной литературой). Мифы постоянно “вращаются” вокруг одних и тех же тем, обыгрывая их на все лады, что дало повод представить мифологический текст как калейдоскоп постоянно рассыпающихся и составляющихся снова в причудливый узор элементарных микросоставляющих: мифологем, микросюжетов. Знаменательно, что из всего литературного фонда именно мифы – это тот материал, стремление формализовать и структурировать который неоднократно зарождалось в среде самих исследователей-гуманитариев (этнографов, фольклористов) – тех, кто по роду деятельности далек от формальных подходов к своему материалу.

Первой подлинно структурной работой была “Морфология волшебной сказки” нашего соотечественника В.Я. Проппа (1928 г.) [23]. Сказка в ней предстает как набор эпизодов, следующих друг за другом в строгой очередности. Каждый из эпизодов обладает уникальной функциональностью (откуда их название – “функции”²) и вносит свой вклад в общее развитие сюжета.

На 60-е годы XX века приходится период расцвета французского структурализма, возглавляемого этнографом К. Леви-Строссом [15, 16]. Исследователь прочно ввел в обиход фольклористики понятие семантической оппозиции. Противопоставления *холодного* и *теплого*, *низкого* и *высокого*, *сырого* и *вареного*, *темного* и *светлого*, *женского* и *мужского*, *мертвого* и *живого* структурируют пространство и время, социальную жизнь “дикаря”. Сюжет мыслился у Леви-Стросса как процесс нейтрализации более сильных оппозиций через более слабые.

Феномену отождествляющей способности мифологического сознания посвящен фундаментальный труд О.М. Фрейденберг “Поэтика сюжета и жанра” (1936 г.) [27]. “Система первобытной образности – это система восприятия мира в форме равенств и повторений”, – пишет Ольга Михайловна. Понятие мифологического тождества является центральным для настоящей модели.

Также при построении модели принимались во внимание разработки последователей Проппа и Леви-Стросса: А. Дандеса [9, 10], супругов Маранда [17], А. Греймаса [7], представителей научного кружка Мелетинского (Е.М. Мелетинский, С.Ю. Неклюдов, Е.С. Новик, Д.М. Сегал [20, 21]), М.Г. Гаазе-Рапопорта [4, 5] и других.

В рамках исследования одновременно разрабатывались два варианта модели: для повествовательных текстов (нарративов) и для пословиц. Модель нарратива в настоящий момент не завершена. Напротив, достаточно детально была проработана, формально описана и реализована в действующей информационной системе модель пословичного текста. Она представляет собой некоторое усечение более распространенной модели нарратива и имеет с ней множество параллелей.

²В.Я. Пропп выделил всего 31 функцию: начальную недостаточность, помощь волшебного помощника, борьбу героя с вредителем и др.

Поскольку далее речь пойдет непосредственно о модели пословицы, следует упомянуть авторов, занимавшихся структурными исследованиями пословиц. Наиболее крупной и авторитетной работой в этом направлении является труд Г.Л.Пермякова “Пословицы и поговорки народов Востока” [25]. Пермяков выделяет в пословице два уровня – уровни “логического содержания” и “образного оформления”. В первом заключен логический смысл пословицы, ко второму относятся выразительные средства, образный язык фольклора. Логический смысл, в свою очередь, складывается из двух компонентов: темы, представимой в терминах семантических оппозиций: *много–мало*, *дружба–вражда*, *труд–лень* и др., и отношения между полюсами темы (например, *качественное*, или *количественное*, или *бытийное* и т.п. *соответствие/несоответствие*).

Модель Пермякова хорошо проработана и подкреплена обширным фактическим материалом³. В ряде других исследований были высказаны интересные идеи, не доведенные, однако, до уровня законченной модели. Ограничимся их кратким обзором. А. Дандес [11] также подчеркивал особое значение бинарных противопоставлений в пословице⁴. Он предлагал при выделении компонентов пословицы ориентироваться на противопоставление *тема – рема*. А. К. Жолковский [12], используя модель Пермякова, пытался заполнить разрыв между логикой и образным оформлением в пословице и выступил с предложением правил последовательного усиления выразительности логических конструкций. Соотнесением образов и понятий в пословице занимался Н. Барли [1]. Образы у него представляли как сужение широких семантических категорий с помощью маркеров. Стоит также отметить работы А. А. Крикманна [13, 14], где внимание было заострено на крайне широких возможностях интерпретации образов в пословице.

3. Основные положения и базовые категории (неформальное описание)

Перечислим основные положения и “опорные точки” модели, выявленные в результате изучения теории мифа и существующих попыток формального описания фольклорного текста.

3.1. В художественном тексте прослеживаются два плана: **глубинный** и **поверхностный**. Поверхностный план соотносится с линейным сюжетом, разворачивающимся в тексте: последовательностью событий и поступков героев. Глубинный план, специфичный именно для художественного текста – это уровень этической оценки перипетий повествования. Этот второй, глубинный, план представляет наибольший интерес для исследователя, однако и наибольшую сложность для извлечения и перевода в формальные категории. Между двумя уровнями существуют сложные неоднозначные отношения. Одни и те же сюжеты допускают различную морально-этическую интерпретацию в исторической перспективе и в разных культурных и социальных контекстах.

3.2. Развитие сюжета в фольклорном тексте строится на двух типах противопоставлений: **мифологических оппозициях** и **мифологических тождествах** (эти категории базируются на когнитивных аспектах мифотворческого сознания, обоснование их выделения см. в работах К. Леви-Стросса и О.М. Фрейденаберг).

И оппозиции, и тождества применимы на каждом из двух уровней (глубинном и поверхностном), несколько варьируясь по функциональной направленности. К поверхностному плану имеют отношение **сюжетные** оппозиции, такие как *обман-разоблачение*, *потеря-находка*, *уход-возвращение*, *голод-сытость*, *поражение-победа*, *заточение-освобождение* и другие. **Этические** оппозиции, такие как *ум-глупость*, *счастье-беда*, *смелость-трусость*, *уродство-красота*, *честь-позор*, *великодушие-коварство* и другие, соотносятся с этическим (глубинным) уровнем текста.

Мифологические тождества являются структурным остовом текста. Их целесообразно подразделить на **категориальные** (те, которые могут быть выражены современными абстрактными категориями): *часть-целое*, *характер-внешность*, *усилия-результат*, *человек-место*, *лидер-коллектив*, *настоящее-будущее*, *характер-судьба* и др. – и **семантические**, в которых связываются понятия, способные в мифологическом сознании символически заменять одно другое: *голод-болезнь-смерть-сон-скитания* или *смерть-бракорождение-разгадка-находка* (списки можно продолжить). Понятия здесь группируются вокруг двух полюсов: положительного и отрицательного. Категориальные тождества тяготеют к глубинному уровню, семантические – к поверхностному.

3.3. Единичный **сюжетный ход** в нарративе (повествовательном тексте) объединяет в себе все четыре типа противопоставлений. С одной стороны, он представляет собой некоторую сюжетную коллизию,

³Несколько десятков тысяч разнонациональных текстов.

⁴“Все пословицы являются утверждениями, содержащими сравнение и/или противопоставление. Сравнение – по своему смыслу – направлено на поиски сходных или одинаковых черт в разных вещах, а противопоставление нацелено на установление и подчеркивание различий между ними” [11, С. 25].

некоторое событие, происходящее с главным действующим лицом (как правило, эта коллизия строится из двух фаз по принципу *действие-противодействие*). С другой стороны, сюжетный ход – это столкновение характеров, поведенческих линий, ценностных систем, что составляет этическую коллизию. Связность текста, плавность переходов от одного звена к другому (развитие сюжета) обеспечивается действием мифологических тождеств.

3.4. В настоящей статье, как уже было сказано, речь пойдет о модели пословицы. Функционально пословица является моральным предписанием, поэтому в ней на первый план выходит этическая составляющая. Для упрощения модели сюжетный план пословицы был опущен. Объединение двух планов – отдельная задача, она решается в рамках построения модели нарратива.

Пословицы в силу жанровой специфики часто явно оперируют на этическом уровне: *труд человека кормит, а лень портит, береженого бог бережет, смелость города берет*. Но нередки также случаи, когда оценка личностных качеств человека производится опосредованно, через регламентацию моделей поведения: *назвался груздем – полезай в кузов, любишь кататься – люби и саночки возить*, либо посредством констатации (с вынесением моральной оценки) некоторого порядка вещей: *яблочко от яблони недалеко падает, два медведя в одной берлоге не уживутся, по Сеньке и шапка, один с сошкой – семеро с ложкой*. В последних случаях моделированию подлежит не сам текст пословицы, а ее интерпретация на языке этических категорий.

4. Формальная модель пословицы

4.1. Модель семантики пословицы базируется на трех предпосылках:

1) Пословица строится на контрастах, ее текст пронизан противопоставлениями. *Лучше хлеб с водой, чем пирог с бедой (голод-сытая жизнь, счастье-горе, духовное-материальное). Собака лает, караван идет (агрессия-спокойствие, маргинал-коллектив, слово-дело). В чужой монастырь со своим уставом не ходят (свой-чужой, правила-территория, маргинал-коллектив)*. Выше уже было сказано, что основными сюжетообразующими средствами любых фольклорных текстов являются бинарные противопоставления: семантические тождества и оппозиции. Задача, таким образом, заключается в том, чтобы унифицировать и классифицировать эти противопоставления, описать их иерархию в пословице, выявить внутренние отношения между ними в тексте.

2) Пословица, в отличие от поговорки, представляет собой суждение⁵. Это означает, что цель пословичного высказывания – утвердить связь между определенными понятиями или явлениями: подчеркнуть параллель, установить причинно-следственное отношение, противопоставить два явления. Минимальная структура в данном случае – два понятия плюс связка между ними.

3) Часто встречаются пословицы-“антиподы”, такие как *для кого труд – радость, для того и жизнь – счастье и скучен день до вечера, коли делать нечего*. При известных различиях в прагматике высказываний, наблюдается их концептуальное сходство. Механизм употребления антонимов в пословице похож на действие логических отрицаний ($x = y \Leftrightarrow \neg x = \neg y$). Для отображения поверхностной структуры, учитывающей “плюс”- или “минус”-реализацию одного и того же понятия, был введен специальный параметр – конфигурация связи.

4.2. Формальная модель пословицы, построенная с учетом этих трех предпосылок, включает:

- два понятия, выраженных каждое в виде оппозиции с “плюс”- и “минус”-полюсами;
- связку между ними, обладающую, в свою очередь, двумя характеристиками: семантической интерпретацией, имеющей характер семантического тождества, и конфигурацией связи.

В формальной записи пословица представляет собой четверку:

$P = \langle T, O1, O2, K \rangle$, где

T – мифологическое тождество (биполярное);

O1 и O2 – оппозиции;

K – конфигурация связи.

Графически структуру пословицы можно представить как семантический квадрат, где по горизонтали располагается ось тождеств, а по вертикали – ось оппозиций.

⁵У В. Даля: “Пословица – ... коротенькая притча, .. суждение, приговор, поучение”. *Даль В. И. Пословицы русского народа*. Т. I-II. – М. 1984. – С. 13.



На приведенной схеме O11 и O21 – положительные полюса оппозиций O1 и O2, а O12 и O22 – соответственно отрицательные. T1 и T2 – две части мифологического тождества.

4.3. Покажем действие модели на примере. Для этого рассмотрим пословицу *дурная голова ногам покоя не дает*. В ней говорится о том, что поспешные, плохо продуманные действия могут повлечь за собой лишние проблемы, хлопоты (по Мокиенко). Два понятия, выраженных в форме оппозиций – *глупый-мудрый* и *покой-заботы*, – объединены причинно-следственной связью, основанной на соотношении душевных качеств человека, его характера с объективными последствиями его поступков – его судьбой (тождество *характер-судьба*). При этом негативные свойства характера влекут негативные последствия (конфигурация $\rightarrow$).



Остановимся подробно на каждом из элементов модели.

4.4. Семантические **оппозиции** представляют собой некоторые общие понятия из разных сфер человеческой деятельности, несущие ярко выраженную морально-этическую оценку и выраженные в двух-частной, полярной форме.

4.4.1. При объявлении номенклатуры оппозиций в первую очередь учитывалась частота их присутствия в пословицах. Оппозиции должны носить достаточно общий характер, чтобы описывать классы текстов, а не уникальные случаи.

4.4.2. Полюса оппозиций имеют ярко выраженную этическую окраску. Это обстоятельство накладывает существенное ограничение на их именование. Дело в том, что в пословичном фонде нередко наблюдается амбивалентное отношение к достаточно фундаментальным понятиям: к труду (*труд человека кормит, а лень портит* и *от работы не будешь богат, а будешь горбат*), к славе и почету (*денег за душой ни гроша, зато слава хороша* и *и честь не в честь, коли нечего есть*) и другим. Необходимость соблюдать полярность приводит к объявлению в подобных случаях двух разных оппозиций: *труд-безделье* и *покой-заботы* в первом случае, *почет-позор* и *слава-безвестность* во втором.

4.4.3. Для удобства использования оппозиций необходимо их классифицировать. В модели принято следующее разделение на классы:

- 1) личностные свойства человека (подклассы: умственные способности, физические способности, моральные, социальные качества, отношение к труду);
- 2) неотчуждаемые, или базовые, ценности (подклассы: самоидентификация (*свой-чужой*), духовное благополучие, физическое благополучие (*гибель-жизнь, болезнь-здоровье*), моральная оценка);
- 3) ценности второго уровня (подклассы: внешний вид, успех, материальное благополучие).

Следует подчеркнуть, что здесь классификация не рассматривалась как специальная научная задача. Приведенное разделение на классы призвано только сделать работу с моделью более удобной. В классификации явно прослеживается ориентация на морально-этическую тематику, что соответствует этическому уровню, на котором работает модель.

4.5. Семантическое **тождество** является глубинной характеристикой отношения между понятиями в пословице. Оно вносит наиболее весомый вклад в общую структуру высказывания.

4.5.1. Тожество задает структурный остов и валентности оппозиций. В процессе апробирования модели на материале русских пословиц были выделены крупные классы текстов, соответствующие традиционному разделению на пословицы о труде и лени, о дружбе и вражде, о характере человека, о своем и чужом и т.п., характеризующихся типичными наборами элементов (тождество + оппозиции). Например, для одного из самых обширных классов пословиц о своем и чужом характерно сочетание одного из тождеств {принадлежность-действие, принадлежность-оценка} с левой оппозицией *свой-чужой*. Для класса пословиц о чертах характера человека типично тождество *характер-судьба*, при этом первой оппозицией идут человеческие качества (*мудрость-глупость, отвага-трусость, терпеливость-нетерпение* и др.), а второй – объективные последствия поступков (*жизнь-гибель, успех-неудача, слава-позор, доход-убыток* и т.п.).

4.5.2. Не только оппозиции, но и конфигурация связи (поверхностное отношение между понятиями), как правило, определяется типом тождества. О типах конфигураций будет сказано ниже. Здесь же важно подчеркнуть взаимосвязь между тождеством и остальными компонентами пословицы. Например, для тождества *сущность-внешний вид* характерны высказывания о зависимости внешнего вида (как частный случай, рода и масштаба деятельности) от внутренних качеств (*птицу видно по полету, большому кораблю – большое плавание*) и, напротив, независимости внутренних качеств от внешнего вида (*в тихом омуте черти водятся, золото и в грязи блестит*). Также частотно сравнение внешнего и внутреннего в пользу последнего (*мал золотник, да дорог*). Подобные типичные типы конфигураций наблюдаются у любого достаточно представительного класса пословиц.

4.6. Наконец, **конфигурация связи** является поверхностной характеристикой отношения между понятиями в пословице. В тексте пословицы в каждой из двух оппозиций актуализируется, как правило, лишь один из полюсов, что порождает различные комбинаторные варианты поверхностных реализаций глубинного смысла. Примеры наиболее частотных конфигураций:

графич. обозначение	конфигурация	пример
	если первое хорошо, то и второе хорошо	<i>смелость города берет</i>
	если первое плохо, то и второе плохо	<i>праздность – мать всех пороков</i>
	лучше какое ни есть хорошее, чем улучшенное плохое	<i>лучше хлеб с водой, чем пирог с бедой</i>
	плохое не совместимо с хорошим	<i>один в поле не воин</i>
	в хорошем иногда бывает плохое	<i>конь о четырех ногах, а спотыкается</i>
...	и др.	...

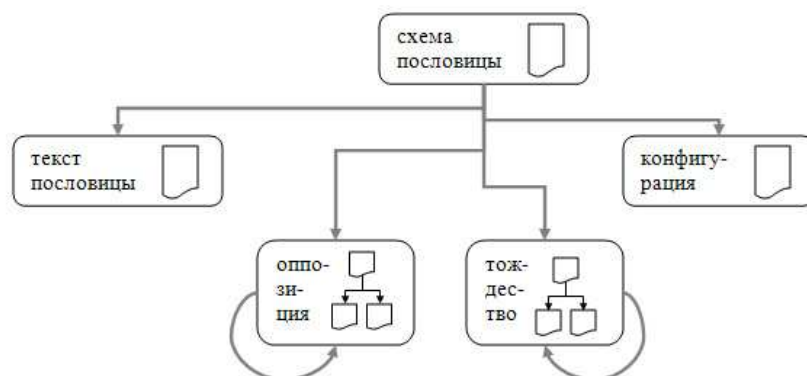
4.6.1. Конфигурации обладают симметрией относительно положительного и отрицательного полюсов, они могут быть сгруппированы в пары: *в хорошем иногда бывает плохое* и *в плохом иногда бывает хорошее*, *плохое портит все хорошее* и *хорошее компенсирует плохое*, *нечто хорошо, а что-то еще лучше* и *нечто плохо, а что-то еще хуже* и т.п.

4.6.2. На множестве конфигураций можно задать отношение сходства. Упомянутая выше пословица *дурная голова ногам покоя не дает* может быть разобрана с участием одной из двух конфигураций: *плохое не совместимо с хорошим* и *если первое плохо, то и второе плохо*. Первый вариант ближе к тексту, второй – к глубинному смыслу. При разборе текстов возможны аналогичные разночтения. Отношение сходства снимает нарушение однозначности соотнесения текста с конфигурацией в подобных случаях⁶.

4.6.3. Наконец, важно подчеркнуть, что словесные интерпретации конфигураций (приведенные в таблице во второй колонке) достаточно условны. В зависимости от типа тождества одна и та же конфигурация может соответствовать различным типам отношений между оппозициями. Ее единственное назначение – отображать взаимное расположение актуализированных полюсов в пословице. Поэтому на практике наиболее удобной была признана графическая запись (приведенная в первой колонке). Именно она используется в информационной системе.

⁶Сходство конфигураций было учтено как специальная возможность в поисковом модуле информационной системы МЕТАФОРА.

4.7. Перечисленные компоненты (тождество, оппозиции, конфигурация) вместе с текстами пословиц и собственно схемой пословицы формируют систему классов онтологии. Диаграмма классов приведена на рисунке.

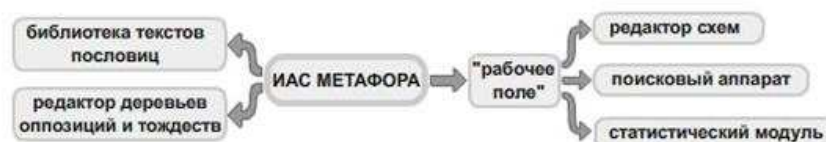


4.8. Модель покрывает до 80% текстов из произвольной выборки (словарь современных русских пословиц под редакцией проф. В.М. Мокиенко [26]).

5. Реализация модели в информационной системе “МЕТАФОРА”

5.1. Модель пословицы была реализована в информационно-аналитической системе МЕТАФОРА. На настоящий момент в системе проанализировано свыше 350 тестовых пословиц, относящихся к разнообразным темам (наиболее представительные выборки из разделов *труд-безделье*, *свой-чужой*, *характер человека*). Составлены иерархически упорядоченные списки тождеств (порядка трех десятков), оппозиций (около шести десятков) и выявлено двадцать два типа конфигураций.

5.2. Система включает три рабочих модуля, не считая вспомогательных (блок авторизации, информационную страницу): 1) библиотеку текстов пословиц, 2) редактор дерева оппозиций и дерева тождеств, 3) основное рабочее поле, объединяющее функциональность редактора схем пословиц, статистического и поискового модуля.



Библиотека текстов, редактор деревьев оппозиций и тождеств и редактор схем пословиц предназначены для заполнения и редактирования базы данных. Их функциональность стандартна.

5.2.1. Статистический модуль отображает распределение значений любого из четырех элементов (тождества, одной из двух оппозиций, конфигурации) при любом наборе фиксированных остальных элементов. Основное назначение данного инструмента заключается в том, чтобы выявлять с его помощью устойчивые сочетания (тождество + оппозиция), (тождество + конфигурация) и т.п., а также отслеживать частоту представления конкретных единиц (тождеств, оппозиций, конфигураций) в выборке.

5.2.2. Поисковый аппарат также позволяет производить поиск при любом наборе фиксированных компонентов схемы. Это дает возможность формулировать сложные запросы, допускающие содержательную интерпретацию на естественном языке. Примеры запросов и результаты поиска приведены в таблице.

неформальное описание, запрос	результаты поиска
амбивалентное отношение к труду: позитивная оценка труда $Q = \{T = ?, O1 = \text{безделье-труд}, O2 = ?, K = \text{хорошее порождает хорошее}\}$	Кто ищет, тот найдет После трудов сладок сон Кончил дело - гуляй смело Трудовой хлеб сладок Что потопаешь, то и полопаешь Для кого труд - радость, для того и жизнь - счастье Хочешь жить - умей вертеться Волка ноги кормят
негативная оценка труда $Q = \{T = ?, O1 = \text{покой-заботы}, O2 = ?, K = ?\}$	Ретивый конь веку не доживает От работы не будешь богат, а будешь горбат Кто везет, того и погоняют Была бы шея, а хомут найдется От трудов праведных не наживешь палат каменных Работа не волк, в лес не убежит
пословицы о том, что за неприглядной внешностью может скрываться добрая натура $Q = \{T = \text{внешность-натура}, O1 = ?, O2 = ?, K = \text{хорошее может компенсировать плохое}\}$	Золото и в грязи блестит Криво дерево, да яблоки сладки Мыло серо, а моет бело

5.3. Система МЕТАФОРА является открытым интернет-приложением, выполненным с использованием передовых технологий разработки RIA (Rich Internet Application) – Adobe Flash CS3 и Adobe Flex 2. Серверная часть написана на PHP на основе CMS Wacko Wiki Quick Start. Система МЕТАФОРА располагается по адресу <http://www.metaphora2.ru/>.

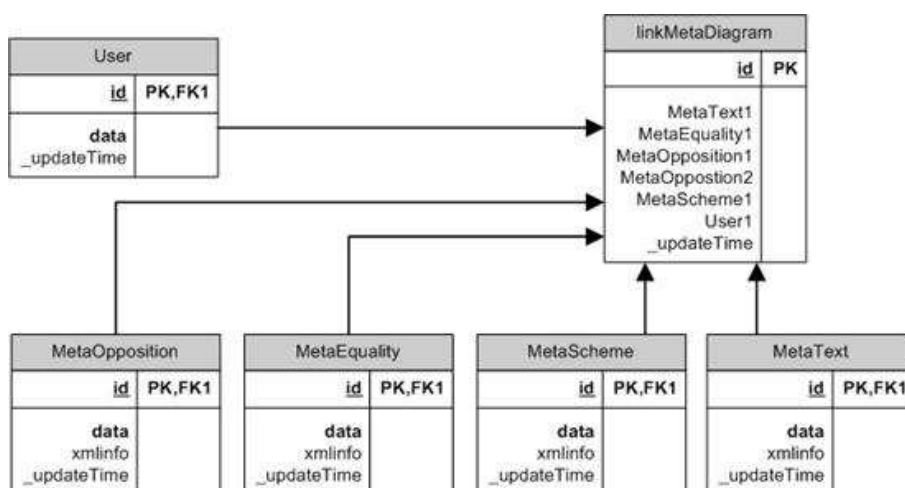
5.4. МЕТАФОРА является вики-системой, т.е. при редактировании схем старые версии сохраняются в базе данных, все изменения авторизованы, предусмотрена возможность оставлять небольшой текстовый комментарий при каждой версии схемы. Данная функциональность востребована, поскольку успешное функционирование системы предполагает дальнейшее уточнение состава и именования оппозиций и тождеств, существенное расширение текстовой базы. Эта работа станет более эффективной, если она будет проводиться в рамках научного сообщества, состоящего из специалистов-предметников – лингвистов и исследователей фольклора.

5.5. Основу технической реализации модели составляет специально написанный сервис онтологий (AdvOntology). Он позволяет определять два типа объектов:

– собственно объекты, характеризующиеся а) текстовой строкой с критической для данного объекта информацией, б) большого текстового поля для хранения XML с произвольным набором второстепенных данных, в) даты создания объекта, необходимой для соответствия вики-формату хранения данных;

– связи, включающие ссылки на произвольное количество любых других объектов и связей и вики-информацию (автора и дату создания связи).

При реализации модели были созданы следующие объекты:



Техническая реализация модуля онтологий осуществлена в соавторстве с В.И. Январевым (ВМиК, МГУ им. Ломоносова).

6. Заключение

Моделирование семантики художественных текстов позволит привлекать для их обработки точные методы: осуществлять эффективный информационный поиск, оперируя на уровне глубинной семантики текста, подвергать тексты статистическому анализу, выявляя психологические матрицы, лежащие в основании сюжетной канвы.

В настоящем исследовании была предпринята попытка построения формальной модели, отражающей суть моральных предписаний, заключенных в текстах русских пословиц. При ее создании учитывался опыт исследования мифологии и фольклора, имеющийся в гуманитарных науках: филологии, этнографии, отчасти психологии. Основные категории, лежащие в основе любого фольклорного текста – это мифологические оппозиции и мифологические тождества.

Модель включает четыре компонента: две оппозиции, связывающее их тождество и конфигурацию связи. Данная модель покрывает до 80% текстов из произвольной выборки (словарь современных русских пословиц под ред. проф. Мокиенко).

Модель позволяет применять процедуры интеллектуальной обработки текстов:

- проводить статистический анализ устойчивых (часто воспроизводящихся) связей *оппозиции + тождество*. Это позволяет составить “национальный портрет” представителей русской культуры, поскольку таким образом отражаются ментальные представления, глубинная связь между понятиями в сознании носителя культуры.

- осуществлять информационный поиск по семантике текста с возможностью формулировать сложные запросы, например, выдать пословицы о негативных последствиях глупости: $Q = \{T = \text{характер-судьба}, O1 = \text{мудрость-глупость}, O2 = ?, K = \text{плохое влечет плохое}\}$.

На основе модели была реализована информационно-аналитическая система МЕТАФОРА, в которой было проанализировано свыше 350 тестовых пословиц, составлены иерархически упорядоченные списки тождеств и оппозиций, выявлено двадцать два типа конфигураций.

Система МЕТАФОРА находится в открытом доступе в сети Интернет и доступна по адресу <http://www.metaphora2.ru>.

Список литературы

1. Барли, Н. Структурный подход к пословице и максиме / Н. Барли // Паремнологические исследования / под ред. Г.Л. Пермякова. — М. : Наука, 1984.
2. Барт, Р. Миф сегодня .Р. Барт // От структурализма к постструктурализму. — М.: Прогресс, 2000.
3. Бремон, К. Логика повествовательных возможностей / К. Бремон // Семиотика и искусствометрия. — М.: Мир, 1972. — С. 108 – 135.
4. Гаазе-Рапопорт, М.Г. Поиск вариантов в сочинении сказок / М.Г. Гаазе-Рапопорт, Р.Х. Зарипов. Машинный поиск вариантов при моделировании творческих процессов. — М. : Наука, 1983.

5. Гаазе-Рапопорт, М. Г. Новые сказки / М. Г. Гаазе-Рапопорт, Д. А. Поспелов, Е. Т. Семенова // Новости искусственного интеллекта. — 1992. — №4.
6. Голосовкер, Я. Э. Логика мифа / Я. Э. Голосовкер. — М.: Главная редакция восточной литературы издательства “Наука”, 1987.
7. Греймас, А. К теории интерпретации мифологического нарратива / А. Греймас // Зарубежные исследования по семиотике фольклора. — М., 1985. — С.109 – 144.
8. Даль, В. Пословицы русского народа. Т. I–II / В. Даль. — М., 1984.
9. Дандес, А. От этических единиц к эмическим в структурном изучении сказок / А. Дандес // Дандес А. Фольклор: семиотика и/или психоанализ. — М.: Восточная литература, 2003.
10. Дандес, А. Структурная типология индейских сказок / А. Дандес // Дандес А. Фольклор: семиотика и/или психоанализ. — М.: Восточная литература, 2003.
11. Дандес, А. О структуре пословицы / А. Дандес // Паремнологический сборник. Пословица. Загадка / под ред. Г.Л.Пермякова. — М. : Наука, 1978.
12. Жолковский, А. К. К описанию выразительной структуры паремий / А. К. Жолковский // Паремнологический сборник. Пословица. Загадка / под ред. Г.Л.Пермякова. — М. : Наука, 1978.
13. Крикманн, А. А. Опыт объяснения некоторых семантических механизмов пословицы / А. А. Крикманн // Паремнологические исследования / под ред. Г.Л.Пермякова. — М. : Наука, 1984.
14. Крикманн, А. А. Некоторые аспекты семантической неопределенности пословицы / А. А. Крикманн // Паремнологический сборник. Пословица. Загадка / под ред. Г.Л.Пермякова. — М. : Наука, 1978.
15. Леви-Стросс, К. Деяния Асдиваля / К. Леви-Стросс // Зарубежные исследования по семиотике фольклора. — М., 1985. — С.35 – 76.
16. Леви-Стросс, К. Первобытное мышление / К. Леви-Стросс. — М.: ТЕРРА — Книжный клуб; Республика, 1999.
17. Маранда, П. Структурные модели в фольклоре. Поиск вариантов в сочинении сказок / П. Маранда, Э. Кенгас-Маранда // Зарубежные исследования по семиотике фольклора. — М., 1985. — С.194 – 260.
18. Мелетинский, Е. М. От мифа к литературе. Курс лекций “Теория мифа и историческая поэтика” / Е. М. Мелетинский — М.: Российск. гос. гуманит. ун-т. 2001.
19. Мелетинский, Е. М. Поэтика мифа / Е. М. Мелетинский. — М. : Восточная литература, Школа “Языки русской культуры”, 1995.
20. Мелетинский, Е. М. Проблемы структурного описания волшебной сказки / Е. М. Мелетинский, С. Ю. Неклюдов, Е. С. Новик, Д. М. Сегал // Структура волшебной сказки. — М. : Российск. гос. гуманит. ун-т. 2001.
21. Мелетинский, Е. М. Статус слова и понятие жанра в фольклоре / Е. М. Мелетинский, С. Ю. Неклюдов, Е. С. Новик // Историческая поэтика. Литературные эпохи и типы художественного сознания. — М. : Наследие, 1994.
22. Мифы народов мира: Энциклопедия. В 2-х т. — М.: Советская энциклопедия, 1980.
23. Пропп, В. Я. Морфология сказки. — 2-е изд. / В. Я. Пропп. — М., 1969.
24. Рафаева, А. В. Еще раз о структурно-семиотическом изучении сказки / А. В. Рафаева, Э. Г. Рахимова, А. С. Архипова // Структура волшебной сказки. — М.: Российск. гос. гуманит. ун-т. 2001. — С. 199 – 225.
25. Пермяков, Г. Л. Пословицы и поговорки народов Востока: Систематизированное собрание изречений двухсот народов / Г. Л. Пермяков. — М. : Лабиринт, 2001.
26. Словарь русских пословиц: ок. 1000 единиц / В. М. Мокиенко, Ю. А. Ермолаева, А. А. Зайнульдинов и др.; под ред. В. М. Мокиенко. — М. : Астрель; АСТ, 2007.

27. Фрейденберг, О. М. Поэтика сюжета и жанра / О. М. Фрейденберг / Ред. и коммент. Н.В. Брагинской.
— М. : Лабиринт, 1997.

Formal model of basic sense in proverb

Malkova A.S., Janvarev V.I.

In the article the formal model of proverb semantics is suggested. The proverb statement is considered as a construction of two semantic oppositions and a so called mythological equation. The opposition lexicon forms a sort of semantic network. This model is to be used as the intellectual module kernel in the information system METAPHORA developed to process proverbs at the semantic aspect.