

Г.В. КАНЬГИН

## ПРЕОБРАЗОВАНИЕ МЕТАФОРИЧЕСКОГО КОНЦЕПТА В КОЛИЧЕСТВЕННЫЙ ОБЪЕКТ

### ПОСТАНОВКА ПРОБЛЕМЫ

Разработчиками программных средств для гуманитарных приложений активно развивается вычислительная поддержка качественных методов (ВПКМ). Наиболее известными программными средствами являются *Atlasti*, *BEST*, *HyperResearch*, *NUD\*IST*, *NVivo*, *WinMAX* и др., представленные на сайте издательской корпорации *Scolari* [1].

Проблематика ВПКМ обсуждается преимущественно в рамках дискуссии о соотношении качественных и количественных (КК) методов в социологическом исследовании. При конструктивной реализации этого соотношения разработчики ВПКМ исходят из представления о типах данных, с которыми имеет дело социолог [2, с. 55]. Программы, создаваемые для ВПКМ, позволяют работать с нетрадиционными форматами (графикой, гипертекстом, аудио- и видеозаписями и т. д.). В статье рассматриваются принципы создания базовых программных структур ВПКМ, которые реализуют качественно-количественную модель социологического исследования. В основе нашего подхода лежит разделение данных на метафорическую и количественную компоненты, что позволяет структурно моделировать неоднозначность социологических концептов. Такое моделирование представляется более продуктивным при анализе качественно-количественных различий социологических понятий, чем «измерение в широком смысле», «онтологический и гносеологический факты» и т. п. [3, с. 102].

Каждый из программных продуктов основан на представлении об определенной предметной области. Не являются исключением и системы поддержки качественных методов. Разработчики *NVivo* отмечают, что анализ качественных данных это собирательное название разнообразных методов работы с плохо структурированной информацией, которая не поддается адекватному переводу в числовую форму. Исследователь, обращающийся к таким данным, как правило, сталкивается с задачей понять изучаемое явление. При этом он либо исходит из речевого поведения самих респондентов, либо анализирует собственные полевые наблюдения, либо черпает факты из документов [4]. Аналогичным образом очерчиваются задачи системы *ATLAS.ti*, позволяющей анализировать «неструктурированные» данные, т. е. исследовательские материалы, которые не могут быть удовлетворительно изучены с помощью формальных или статистических методов» [5]. Фактически, каче-

---

**Каньгин Геннадий Викторович** — кандидат экономических наук, руководитель группы Социологического института РАН. Адрес: 198005 Санкт-Петербург, Измайловский проспект, д. 14. Телефон: (812) 316-21-62; Факс: (812) 316-29-29. Электронная почта: [genaka@hotmail.com](mailto:genaka@hotmail.com)  
Исследование выполнено при поддержке Российского фонда фундаментальных исследований (проект №00-06-80184) и Российского гуманитарного научного фонда (проект № 00-03-00279а).

ственный анализ представляет собой набор методов, которые позволяют обрабатывать материалы «неколичественной» природы. Однако отличить этот вид материалов от данных, которые поддаются обработке (например, с помощью традиционных статистических техник), непросто. Понятие качественного анализа обычно переопределяется через термины, заимствованные из информатики. Типично сведение качественных данных к их информационным прообразам: мемо, форматированному мемо, ссылке, изображению и т. п. Такое переопределение страдает существенным недостатком. Хотя качественный анализ, реализованный в компьютерных представлениях, становится достаточно конструктивным, специфика «качественных материалов» выражается в информационных терминах. Это принципиально неверно. Прикладное программное обеспечение должно развиваться исходя из требований своей предметной области, расширяя возможности информационных технологий, а не подстраиваясь под них.

#### ПРОБЛЕМАТИКА КАЧЕСТВЕННОГО АНАЛИЗА

Основной научной проблемой, решаемой с помощью ВПКМ, является преобразование авторских определений, возникающих в различных предметных областях, из многозначно интерпретируемой формы в однозначно понимаемые термины и наоборот. Конструктивное решение этой проблемы ищется в процессе уточнения текстовых формулировок исходных концептов, привлечении аудио- и визуальных средств, графического представления связей между концептами, построении систем гиперссылок и т. п.

Концепты, разрабатываемые социологом для описания своей предметной области, являются метафорами, которые содержат набор ссылок на другие аналогичные объекты (посредством ссылок метафоры могут быть связаны в ассоциации, которые также являются метафорами). Процесс образования новых метафор всегда соотносен с автором, т. е. метафоротворчество осмысленно только для автора, выполняющего ассоциирование. Методологическая трудность построения метафор состоит в том, что автор метафоры и лицо, ее использующее (актор), часто вкладывают в один и тот же текст разные смыслы. Метафора, вообще говоря, неоднозначна, и для устранения этой неоднозначности необходимо выполнить определенные действия — построить такую систему ассоциаций, которая приводила бы к конвенционально принятому истолкованию.

Общая идея подхода состоит в моделировании перехода от авторской метафоры к количественным понятиям. Задачей моделирования является, в первую очередь, разработка понятий, в которых осуществляется этот переход. Исходными представлениями модели являются: во-первых, *метафорический объект* (МО), во-вторых, его «антипод» — однозначная конструкция или *примитив*, в-третьих, операции по переводу одной конструкции в другую. В ходе перевода возникают различные виды *количественных объектов* (КО), которые имитируют качественно-количественные различия социологических концептов, выражая собой различные сочетания метафорических и количественных свойств.

Процесс построения однозначных определений мы называем *дефиницией* метафор. Поскольку социологический словарь почти всегда сохраняет многозначность (полисемию), нами предлагается представление о *тезаурусе*,

объединяющие «множество» метафор, разделяемых определенной школой или коллективом. Хотя термины, реферирующие наше представление о «качестве» и «количестве», названы, но в чем состоят принципы, позволяющие выполнить их программную реализацию?

#### ПРИНЦИП РАЗДЕЛЕНИЯ ОПРЕДЕЛЕНИЙ И ЗНАЧЕНИЙ

Решение проблемы обнаруживается в нетрадиционном структурном «расщеплении» представления о данных на определения и значения. Рассмотрим стандартную процедуру определения данных. Сначала заполняется таблица описания полей (ТОП), в которой для каждого поля указываются его однозначно понимаемые параметры. Например, для данных, соответствующих анкетированию по четырем вопросам — ФИО, ПОЛ, ВОЗРАСТ, ОБРАЗОВАНИЕ — обычная ТОП выглядит следующим образом (табл. 1). Сами данные опроса представляются в виде таблицы результатов (табл. 2).

Таблица 1

ВОПРОСЫ	ТИП	РАЗМЕР
НАЗВАНИЕ ВОПРОСА		
ФИО	String <sup>1</sup>	25
ПОЛ	Short	
ВОЗРАСТ	Numeric	
ОБРАЗОВАНИЕ	Short	

Таблица 2

РЕЗУЛЬТАТЫ	ПОЛ	ВОЗРАСТ	ОБРАЗОВАНИЕ
Иванов А.А.	1	35,5	2
Петров А.Б.	1	44,0	3
Сидорова А.В.	2	25,0	2
...			

В программных системах, используемых в социологическом исследовании, указанная схема воспроизводится в статистических пакетах и системах управления базами данных. Для программ компьютерного ассистирования интервьюированию достаточно «жесткая» формализация, выражаемая таблицей 1, как правило, заменяется более понятной для социолога процедурой разработки опросника, и система генерирует требуемое определение автоматически. Основная особенность описания данных в системах ВПКМ — возникновение нетрадиционных типов полей, например, графических форматов. Кроме того, в этих пакетах создание ТОП не выглядит «чужеродным», так как осуществляется в рамках работы с деревом концептов — представлением, обусловленным запросами пользователя, а не формальными требованиями к данным [6]. В то же время сохраняется принципиальная особенность трактовки самого понятия «данные». Данные из одной таблицы (ВОПРОСЫ) оказываются определением для данных из другой таблицы (РЕЗУЛЬТАТЫ).

<sup>1</sup> String — распространенная кодировка типов полей данных в различных системах программирования.

Каким образом мы модифицируем такое представление? Каждый экземпляр данных состоит из заголовка (текстовой метки), формальной структуры и значений. Для ТОП заголовком является слово ВОПРОСЫ; структура задается набором полей НАЗВАНИЕ ВОПРОСА, ТИП, РАЗМЕР; значения же представлены записями таблицы. Аналогично представлены данные РЕЗУЛЬТАТЫ.

Обычно заголовок рассматривается как обычная текстовая переменная. Наше предложение состоит в том, чтобы считать заголовок более сложным объектом, который предполагает возможность своего связывания с другими аналогичными объектами. Сама возможность связывания заголовков друг с другом предполагает новое представление соотношений между различными компонентами понятия «данные». Между заголовками, или объектами, например, ВОПРОСЫ и РЕЗУЛЬТАТЫ, существует более существенное сходство, чем видимые структурные различия, возникающие в связи с приписанными строками значений. Сходство заключается в том, что автор описаний типа ВОПРОСЫ и РЕЗУЛЬТАТЫ считает их однозначно понятными для всех, кто использует эти определения; а строки таблицы являются реализациями однозначных определений или значениями.

Формальным выражением однозначности для объектов типа ВОПРОСЫ и РЕЗУЛЬТАТЫ может считаться указание соответствующих структур в виде наборов полей известных типов. Например, для объекта ВОПРОСЫ такими полями являются: String11, String7, Short, для объекта РЕЗУЛЬТАТЫ: String25, Short, Numeric, Short. В этом смысле как заголовки, так и соответствующие им структуры понимаемы однозначно. Если посмотреть на ВОПРОСЫ и РЕЗУЛЬТАТЫ как на метафоры, каждой из которых приписана лишь одна из возможных трактовок, то между метафорами и структурами данных всегда обнаруживается автор. Буквально ВОПРОСЫ и РЕЗУЛЬТАТЫ — это слова, которые могут иметь множество смыслов и только один из них фиксируется с помощью формальной дефиниции. Поэтому формальное выражение однозначности принадлежит только к структурам данных, а объекты, выраженные словами ВОПРОСЫ и РЕЗУЛЬТАТЫ, получают эту дефиницию через выполнение автором операции ассоциирования.

«Расщепление» данных на заголовок, структуру и значения рассмотрено нами на примере таблиц. Однако, тезис о существовании в понятии данных двух компонент, связь между которыми устанавливает автор, может быть распространен на составляющие самих таблиц. В частности, столбцы ТИП, ПОЛ, ФИО и др. также представляют собой композиции из заголовка, формальной структуры и значений. Например, ТИП состоит из заголовка «ТИП», предполагаемого формального типа String7 и набора значений String, Short, Numeric, Short. Аналогично, ПОЛ представляется как текстовая метка «ПОЛ», типа Short (в предположении, что используется целочисленная кодировка) и значений 1, 1, 2, показанных в таблице РЕЗУЛЬТАТЫ.

Структурное расчленение данных имеет визуальную интерпретацию. Пока автор не укажет для объекта, который обычно называется заголовком, однозначно понятное определение, например, тип поля, «заголовок» через ассоциации с другими «заголовками» образует набор концептов. Этот набор всегда уникален и представлен в виде сети или, в частности, дерева концептов. Как только автор указывает однозначно понятное определение, стано-

вится возможным (не только для автора, но и других лиц) воспроизводить это определение в виде набора значений или записей. Записи представляются в виде таблицы. Таким образом, представления о сети и таблице визуализируют наш подход к понятию данных.

Может показаться, что развиваемые идеи далеки от социологической проблематики. Однако конструктивная концептуализация социологического исследования, независимо от ее формы (например, анкета или дерево концептов), не выходит за рамки действий по определению данных. Социолог не имеет других средств разъяснения своих концептов кроме их сведения к однозначно понимаемым конструкциям аналогичным в этом смысле формальным типам полей. Видимым отличием является только размер «паутины» определений, приводящей исследовательскую метафору к ее формулировке, считающейся однозначно понимаемой.

#### **СТРУКТУРНАЯ МОДЕЛЬ КАЧЕСТВЕННО-КОЛИЧЕСТВЕННЫХ ПОНЯТИЙ**

##### **Метафорический объект**

«Возраст», «пол», «социальное положение» и другие социологические концепты — метафорические объекты, за которыми стоят неоднозначные ассоциации. Для их воссоздания нужны контекст, разъяснение, привязка к ситуации. Полисемические конструкции используются в речевой коммуникации с респондентом независимо от того, понятны они или нет. В качественном анализе метафорическим объектом является «качественный» концепт, противопоставленный количественному описанию в терминах «переменная», «атрибут», «параметр» и т. п. Неоднозначность метафорического объекта достигается тремя способами: специальной трактовкой его идентификационных параметров, созданием относительной однозначности посредством интерпретации любой метафоры через ссылки на другие метафоры, сведением всех метафорических объектов в единый тезаурус.

**Идентификационные параметры метафорического объекта:** мнемотекст и его автор. Мнемотекст — наблюдаемая часть метафоры, которая является ссылкой на некоторое семантическое задание автора. Поскольку это задание и другие состояния сознания автора сообщения невозможно наблюдать, мнемотекст и его автор предстают неразделимым целым.

**Ссылки метафорического объекта** на другие метафоры позволяют достигать относительно однозначной интерпретации. Любой МО может быть превращен либо в количественный объект, либо в чистую метафору в зависимости от того, какие МО связаны между собой ссылками. Ссылки собраны в списки, число которых, в принципе, не ограничено. Назначение списка определяется целью использования его ссылок в алгоритмах модели. В данном случае речь идет о трех списках: определений, ассоциированных объектов и контекстов, которые играют основную роль в моделировании неоднозначности концептов.

**Тезаурус.** Поскольку модель построена на взаимных ссылках метафор, никакая метафора не может рассматриваться изолированно, и метафорические объекты всегда образуют совокупность, целостность которой поддерживается всеми механизмами модели. Эта совокупность называется тезаурусом.

Метафорические объекты соответствуют меткам ВОПРОСЫ и РЕЗУЛЬТАТЫ в приведенном выше примере. Причем МО не имеет набора

записей, составляющих таблицу. Возможность присоединения такого набора рассматривается как проблема, решаемая самим исследователем. Метафора — это определение по ссылке, для реализации которой необходим автор, поэтому наше понимание МО является человеко-машинным. Метафорический объект инкапсулирует все действия по созданию и связыванию метафор (ТМО). Это своеобразный «класс-предок», от которого происходят классы других единичных объектов, например, алгоритмов.

Основные свойства МО, выражающие указанные выше характеристики: мнемотекст — текстовая переменная<sup>2</sup>; автор — уникальный идентификатор актора<sup>3</sup>; ссылки — список списков ссылок на МО; тезаурус МО — указатель на тезаурус, к которому приписан МО. Тезаурус — также базовый класс модели (TThesa), в котором собраны свойства и действия, необходимые для работы с совокупностью всех метафор: идентификация МО, загрузка, сохранение, поиск, переопределение «метафорический объект — количественный объект» и т. п. В тезаурусе объявлены списки примитивов и алгоритмов, т. е. МО специального вида, используемых в описанных ниже механизмах типологизации и дефинизации метафор.

#### **Количественный объект**

Ассоциации, стоящие за словами «возраст», «пол», «социальное положение», рассматривались выше как авторские конструкции, не требующие понимания другими участниками коммуникации. Теперь метафорические объекты могут быть рассмотрены как слова, относительно которых достигнуто единое понимание. Количественный объект — это соглашение об однозначности понимания метафоры в определенном круге лиц.

Однозначность достигается, во-первых, объявлением любой метафоры тезауруса однозначно понятной. Такая метафора называется примитивом. Объявление МО примитивом означает задание дополнительных параметров, которые имеют количественное или формальное выражение. В качестве таких параметров применяется тип данных, который связывается с примитивом во всех его дальнейших использованиях. Указание параметров подобного рода является *критерием однозначного понимания метафоры*<sup>4</sup>. Во-вторых, однозначность достигается определением метафорического объекта за счет указания в нем ссылок на другие метафорические объекты. Эта опе-

<sup>2</sup> Мнемотекст — одна из форм метафоры, которая может быть выражена, например, рисунком. Наш подход принципиально не меняется при использовании любого изображения метафоры. Есть основания считать, что исследователь всегда обращается к метафоре через мнемотекст.

<sup>3</sup> Актор — любое лицо, использующее МО; сведения об акторах сведены в соответствующую базу данных, поддерживаемую вне класса ТМО.

<sup>4</sup> Примитивом может быть объявлена любая метафора, и функционирование примитива в качестве инструмента, реализующего представление об однозначности, обусловлено только этим объявлением. Тем самым алгоритмы ассистирования, построенные на анализе однозначности метафор, являются условными. Программная система, основанная на этой модели, не дает ответа на вопрос, насколько верно понимается то или иное концептуальное построение социолога. Впрочем, любые исходные положения не доказываются, а декларируются. Поэтому предназначение операций с метафорическими объектами — структурное воспроизведение последствий конвенций об однозначности.

рация осуществляется путем размещения соответствующих ссылок в списках двух видов: определений и ассоциированных объектов.

Количественный объект формируется из метафоры путем преодоления ее неоднозначности, источником которой является неартикулируемое авторское представление об объекте. Переход к концепту, понятному некоторому сообществу, становится возможным, когда созданы количественные объекты. Однако неоднозначность может возникнуть и относительно однозначно понимаемых концептов. В этом случае проблема заключается не в преодолении метафорической многозначности, а в выборе возможных вариантов из нескольких однозначно понимаемых ситуаций. Мы называем такую неоднозначность количественной и различаем четыре ее типа.

*Концептуальная неоднозначность* (ветвление) возникает в том случае, когда автор целенаправленно использует различные выражения исходной метафоры. Иными словами, для одной и той же метафоры реализуется несколько однозначно понимаемых интерпретаций. Например, «возраст» в одной ситуации — число лет, в другой, — набор градаций вида: 0 — нет ответа, 1 — до 18 лет, 2 — от 19 до 33 лет и т. д. Таким образом, одна и та же метафора содержит ссылки на разные количественные объекты, выбор которых зависит от ситуации использования или контекста. Например, «возраст», выраженный в годах, может использоваться при разработке концепции, а при обработке тот же «возраст» может принять вид шкалы: 0, 1, 2...

*Информационная неоднозначность* проявляется в том, что количественный объект может определять один и тот же «информационный» объект, но в разных контекстах (экономия ресурса, традиции программирования и т. п.) используются разные части определения. Например, вариант ответа, описывающий пол, имеет в тексте анкеты структуру вида ЦЕЛОЕ, ТЕКСТ, в то время как в данных этот же вариант ответа сводится к единственному полю типа ЦЕЛОЕ. Различия между структурами, формально определенными через количественный объект ПОЛ, зависят только от контекста использования, но этот контекст, в отличие от концептуальной неоднозначности, обусловлен не автором метафоры, а формальными задачами построения программ. Для конечного пользователя эти использования укладываются в единственную структуру, заполненную значениями вида: 1-М, 2-Ж, и задача вычислительных средств — автоматически выбрать из нее требуемую информацию.

*Ситуационная неоднозначность* возникает из-за того, что метафорический объект в рамках одной и той же структуры, но в разных ситуациях может конкретизироваться с помощью различных наборов значений, которые не указаны заранее. Это означает разрешение любому актору приписывать значения количественному объекту. Примером такой неоднозначности является открытый вопрос или пополнение множества ответов по результатам опроса респондентов.

*Алгоритмическая неоднозначность* возникает вследствие возможности неограниченного числа формальных действий над однозначно определенным метафорическим объектом. Неоднозначность социологических интерпретаций характерна для этапа анализа данных и находит выражение в существовании нескольких алгоритмов, решающих одну и ту же задачу. Например, агрегированные индексы могут быть сконструированы средствами факторного анализа, шкалирования, эвристических процедур и т. д.

Метафорическая неоднозначность воспроизводится в модели за счет алгоритмов типологизации, средств объявления метафоры примитивом и событий, отслеживающих все акты ассоциирования метафор конечным пользователем. Объявление метафоры примитивом означает помещение ее в список метафор, которые декларируются акторами данного тезауруса как однозначно понимаемые. Может быть выполнено и обратное действие — изъятие метафорического объекта из списка примитивов, то есть отказ от его однозначной трактовки. Включение в список или исключение метафорического объекта из списка примитивов могут влиять на типы других метафор. Модель автоматически отслеживает эти изменения. Это отслеживание не является тривиальным, так как посредством ссылок метафорические объекты могут образовывать длинные цепи определений.

Для модельного выражения количественной неоднозначности мы используем различные приемы. Концептуальная неоднозначность моделируется за счет помещения в список ассоциируемых метафорических объектов нескольких количественных объектов, каждый из которых представляет в некотором контексте однозначное определение МО. Выбор требуемого количественного объекта осуществляется алгоритмами модели при использовании МО либо автоматически, либо в человеко-машинном режиме. Информационная неоднозначность распознается и устраняется средствами модели автоматически. Ситуационная неоднозначность моделируется с помощью регламентации прав доступа к данным. Неоднозначность алгоритмических решений моделируется через понятие алгоритмического объекта.

Количественный объект отличается от метафор наличием данных, структура которых в конечном счете определена ссылками метафорического объекта на примитивы. В приведенном выше примере количественный объект — это ВОПРОСЫ и РЕЗУЛЬТАТЫ, обладающие полями, которые могут быть заполнены значениями, указанными в соответствующих таблицах. Таким образом, количественный объект принадлежит классу TQO, который является наследником базового класса TMO. В TQO добавлены свойства и методы, позволяющие работать с данными.

#### **Алгоритмический объект**

После того, как исследователь декларировал «возраст», «пол», «социальное положение», построил для них однозначное определение и выполнил рутинную работу по сбору данных, он получает возможность формальных операций над этими понятиями, которые представляют собой обычные действия над данными: перекодировка, связывание, метризация, и т. п.

Алгоритмический объект (АО) наследует аппарат связывания концептов, введенный через понятие метафорического объекта, и всю поддержку по работе с данными, инкапсулированную в количественный объект. В данной версии модели АО реализует аппарат подключения оттранслированных процедур, значения параметров которых передаются через количественный объект. В принципе, АО может инкапсулировать и другие средства описания алгоритмов работы с данными, например, SQL. Алгоритмический объект реализуется на том этапе социологического исследования, когда концепты нашли выражение в согласованном количественном понимании. Поэтому алгоритмический объект соответствует структурному принципу размежевания определений и значений через наследование свойств метафорического

объекта и количественного объекта. АО – это объект класса ТАО — наследника ТОО. В ТАО определены ссылки на процедуры, которые могут компилироваться как вместе с классом, так и входить в состав библиотеки.

#### Основные механизмы модели

**Типология метафорических объектов.** Интуитивно понятно, что качественные или «мягкие» методы характеризуются большей неоднозначностью, чем количественные, или «жесткие». Например, полевые наблюдения предназначены в основном для их автора и соотносятся с его системой ассоциаций. Поэтому они менее понятны для читателя, чем, скажем, вопрос анкеты: «Сколько вам лет?». Как выразить эту интуитивно понятную несопоставимость подходов?

Решение проблемы обнаруживается в типологизации метафор. Каждый из предлагаемых нами типов представляет собой некоторую разновидность (но не степень) неоднозначности. В основе типологизации лежит описанное выше представление о практических источниках неоднозначности концептов социологического исследования. В зависимости от типа неоднозначности устанавливается набор возможных действий пользователя: перевод метафорического объекта в примитив, выбор одного из возможных ветвлений, задание свойств примитива и т. п.

Укажем соответствие между источниками и типами неоднозначности в мнемонической нотации нашей программной разработки. Смысл нотации пояснен в приложении.

Метафорический: mtPure, mtPrim, mtPrimAlias.

Концептуальный: mtDefined, mtDefinedAlias, mtMultiAliasOO, mtComplex.

Алгоритмический: mtProc, mtProcAlias.

При решении проблем неоднозначности в алгоритмах модели используется представление о контексте. Контекст — это дескриптор ситуации в модели, который управляет ветвлением соответствующего алгоритма, в частности, выбором одного из ассоциированных количественных объектов при неоднозначности ветвления. Особенность нашего представления о контексте — это использование метафорического объекта для описания такого дескриптора. Поэтому мы не обсуждаем вопрос, зачем нужен и как используется контекст, но в типологию включен вариант mtContext.

Модель типологизации — это алгоритм отнесения метафорического объекта к одному из типов неоднозначности, который определяет возможные действия с МО. Алгоритм вычисляет тип МО1 в предположении, что заданы типы всех МО, на которые ссылается МО1. Эти вычисления представляют собой модифицированный вариант построения сети [7, 304-345]. Отличие состоит в том, что узлом сети является МО, в котором определен не один, а несколько наборов ребер, представленных как списки ссылок на МО. Алгоритм выполняет «раскрутку» всех цепей ссылок МО1, находящихся в определяющем списке и списке ассоциированных объектов, до тех пор пока не вычислит тип в соответствии с таблицей. Алгоритм выполняется для МО1 всякий раз, когда в одной из метафор тезауруса, которому принадлежит МО1, происходит событие, меняющее тип МО, на которые ссылается МО1. Такие события связаны, как правило, с действиями пользователя, ассоциирующего метафоры между собой. Типология метафорических объектов не зависит от контекста их использования и представляет собой ответ на во-

Отформатировано: русский

прос, какими могут быть метки ВОПРОСЫ и РЕЗУЛЬТАТЫ в смысле однозначности связанных с ними определений. Алгоритм типологизации реализован как набор методов классов TMO и TThesa.

**Дефиниция метафорических объектов.** Исследователь вынужден создавать словесные описания, воспринимаемые другими людьми как понятные. Фактически этот процесс ведет к появлению новых и новых формулировок и построению связей между ними. Он часто трактуется как поиск методики сбора социологической информации, более «точно», «адекватно», «релевантно» и т. п. отражающей реальность, чем суждения в исходных вербальных формулировках. Мы не предлагаем никаких критериев «релевантности» переформулирования одних концептов через другие. Определение терминов может быть представлено как конвенциональный процесс связывания метафор между собой с целью достижения «понятности». Чтобы определить метафору (например, «возраст») исследователь должен ассоциировать ее с каким-то общеупотребительным представлением, например, с понятием «год». Указывая «год», исследователь преодолевает неоднозначность «возраста». Объект «год» должен обладать какими-то дополнительными характеристиками по сравнению с исходным концептом. Такими характеристиками могут быть, например, единицы измерения или соглашения, соблюдаемые всеми акторами, применяющими понятие (в данном случае «год»). Единственное условие введения подобных характеристик состоит в том, что они должны быть разделяемы еще кем-то, кроме их автора. Концепт «год» должен быть описан как примитив или количественный объект, приводящий к набору примитивов. Компьютерное выражение концепта «год» является следствием существования акторов, применяющих этот концепт.

Допустим, «возраст» в одном исследовании описывается через «год», а в другом через набор градаций, обозначаемых как «возраст X»: 0-нет ответа, 1-до 18 лет, 2-от 19 до 33 лет. В этом случае проблема преодоления неоднозначности уходит на второй план, так как предполагается, что «год» и «возраст» являются адекватными выражениями концепта «возраст» в различных ситуациях. Различение этих ситуаций является формальной задачей для акторов, принимающих конвенциональное определение «возраста». Чтобы модель эксплицировала эту задачу в тех же терминах, что и акторы, само предположение должно быть зафиксировано путем ассоциирования двух количественных объектов, соотносенных с метафорой «возраст».

Процесс определения метафорического объекта в общем виде состоит из двух составляющих: действий исследователя по парному связыванию метафор и отслеживанию моделью, как это связывание отражается на однозначности всех метафор тезауруса. Первую составляющую — связывание пары метафор между собой — мы называем *ассоциированием*. Ассоциирование является человеко-машинной операцией, так как, с одной стороны, связывание осуществляется на основе некоторого задания, принадлежащего исключительно компетенции исследователя, с другой стороны, для реализации этого задания исследователю необходимы интерактивные средства работы с метафорическим объектом. Формальным результатом ассоциирования является изменение числа ссылок в одном из списков МО. Однако появление или исчезновение ссылок в одной метафоре может вызвать изменение количественных свойств других метафор тезауруса. Поэтому второй состав-

ляющей определения метафорического объекта является вычисление для каждого МО всех его количественных параметров, которые отражают текущее состояние всех ссылок МО тезауруса. Такое определение количественных параметров любого МО мы называем дефинизацией.

Задачей дефинизации является ответ на вопрос, как отразится акт ассоциирования, выполняемый исследователем в ходе своей «определяющей» деятельности, на любой метафоре того тезауруса, с которым он работает. Например, если пользователь посчитал метафору М1 примитивом (однозначно понятной), то как эта конвенция скажется на всех других метафорах, которые имеют ссылки на М1? И наоборот, если метафора М1 потеряла статус примитива, то как изменится однозначность во всех других метафорах пользовательского тезауруса из-за того, что они ссылались на М1? Другой пример, если пользователь исключил ветвление в какой-либо метафоре, то есть список ассоциированных объектов перестал содержать несколько количественных объектов, то как это действие скажется на однозначности метафорических объектов, ссылающихся на отредактированный объект?

Модель дефинизации — это сетевой алгоритм вычисления значений всех параметров, необходимых для выполнения действий, разрешенных для МО с определенным типом неоднозначности. В отличие от типологизации данный алгоритм проводит полную «раскрутку» цепи ссылок, позволяющую, например, получить не только утверждение о том, что возможно создание структуры данных, но также и все типы полей этой структуры. Дефинизация проводится по текущему состоянию списков ссылок и представляет собой формальный алгоритм, сводящий это текущее состояние, как правило, к набору примитивов. Дефинизацию следует отличать от операции ассоциирования метафорических объектов, которую выполняет пользователь. С помощью ассоциирования он имеет возможность изменять однозначность метафор, то есть влиять на результаты дефинизации. Ассоциирование метафор, например, МО1 к МО2 — это человеко-машинная операция приписывания ссылки на МО1 к одному из списков ссылок МО2.

Алгоритмы дефинизации являются эвристическими. Они учитывают наличие ссылок в списках определений и ассоциированных объектов, контекст использования, соподчиненность списков и ряд других параметров. Эвристическая природа алгоритмов предполагает их совершенствование в процессе практического применения модели. Однозначность определений метафор достигается в конечном счете прослеживанием цепей их выхода на примитивы, но в качестве определителя может использоваться любой КО, не обязательно примитив. В таком случае соответствующие отсылки выполняются в системе автоматически. Дефинизация — это конструктивный ответ на вопрос, как получить данные, имея в своем распоряжении цепи определений, составленные из метафор. Иными словами, дефинизация устанавливает связи между метками типа ВОПРОСЫ и РЕЗУЛЬТАТЫ и однозначно понимаемыми данными и алгоритмами.

Модельным выражением операции ассоциирования является событие, состоящее в изменении числа ссылок в списке МО. Для пользователя ассоциирование реализовано через интерфейсный компонент, не рассматриваемый в рамках данной модели. Алгоритм дефинизации реализован как набор методов класса ТМО.

**Механизм значений.** Если пользователь определил количественный объект, то есть сделал возможной дефинизацию метафоры, такой КО используется в качестве средства определения новых метафор. В ходе исследовательской работы число определенных КО увеличивается, и они становятся основным инструментом введения новых количественных понятий.

Наличие большого числа определений ставит задачу их вычислительной поддержки. Эта задача решается систематически с помощью аппарата сопровождения соответствующих баз данных, который мы называем механизмом значений (МЗ). Алгоритмы МЗ служат для автоматизации учета и сопровождения пользовательских структур данных (задаваемых через КО), контроля ввода, управления данными через связи между МО и т. п. МЗ используется также для устранения неоднозначности информационного вида. Типичным способом разрешения неоднозначности структур, определяемых одним и тем же «информационным» объектом, является раздельное создание всех необходимых конструкций. Так, «пол» при составлении опросника — это, например, фрагмент текстового файла, который создается в текстовом редакторе, а ПОЛ для данных — это поле типа Short, возникающее с помощью утилиты создания структуры файла данных. Такой подход пригоден для квалифицированных пользователей, ориентирующихся в программных средствах, но обычно различные формы выражения одних и тех же «сущностей» представляются досадной обузой. МЗ автоматически определяет нужную структуру «информационного» объекта, создаваемую из одного и того же КО в различных ситуациях его применения. Правда, с целью контроля правильности работы этих средств результаты автоматических решений являются доступными пользователю и могут быть отредактированы. Выбор нужной интерпретации осуществляется в зависимости от контекста использования КО. Контекст определяется по эвристическим критериям, позволяющим различать разные случаи применения КО. Например, вычисляется, создана ли структура данных для рассматриваемого КО, и входит ли он в состав более развернутого определения<sup>5</sup>.

МЗ обрабатывает записи значений, рассматривая их отдельно от определений в соответствии со структурным принципом размежевания. МЗ реализован как набор методов класса TQO.

**Парадигмальное сходство.** Пользователь имеет возможность посредством создания связей между МО переводить любые метафоры в количественные объекты. Существование структур для различных МО означает возможность их сравнения, то есть нахождение идентичных или сходных типов некоторых полей. Эти формальные критерии сходства реализуются с помощью аппарата МЗ. Но между МО может существовать сходство другого рода. Допустим, что в одном исследовании качественный концепт «пол» сопоставляется со списком градаций (Г1), имеющим структуру С1 [код, метка], а в другом — со списком Г2, определенном в рамках новой структуры С2 [код, метка, длинная метка, рисунок]. Тогда Г1 и Г2 оправданно считать сходными, так как эти структуры сопоставлены с одним и тем же концептом, но по-разному трактуемым в разных проектах. Сходство такого рода мы называем парадигмальным. Оно возникает, главным образом, из-за смены ав-

<sup>5</sup> Разработка указанных эвристических критериев продолжается.

торского взгляда на концепт, и возможные структурные совпадения не играют роли. Так, в примере поля «код» и «метка» идентичны, что может служить формальным основанием для количественной оценки сходства Г1 и Г2. Однако на парадигмальное сходство (ПС) такая оценка не влияет, так как ПС является следствием только операции ассоциирования метафор.

Парадигмальное сходство между метафорическими объектами фиксируется моделью и сообщается пользователю во всех релевантных ситуациях (построение концептов, генерация отчетов, поиск и т. п.), а также используется при работе своих алгоритмов (например, типологизации метафор). Фиксация сходства MO1 и MO2 — это помещение ссылки на MO2 в список сходств<sup>6</sup> MO1 и одновременное создание ссылки на MO1 в списке сходств MO2. Иными словами, мы рассматриваем ПС как симметричное отношение между двумя метафорами, устанавливаемое человеком-машинным способом в момент ассоциирования. Парадигмальное сходство означает сходство между записями таблицы данных, рассматриваемыми как единое целое. ПС возникает, когда записи помечены одинаково, например, вместо разных меток ВОПРОСЫ и РЕЗУЛЬТАТЫ используется одна и та же — ДАННЫЕ. Реализовано ПС как событие изменения числа ссылок в списке класса TMO, одним из следствий которого является создание взаимных ссылок.

#### КАК ВЫГЛЯДИТ МОДЕЛЬ?

Обычно под моделью имеется в виду, во-первых, математическая модель (как правило, в виде задачи агрегирования данных), которая по исходным переменным продуцирует некоторый обобщенный индекс; во-вторых, готовая вычислительная система, позволяющая ввести значения исходных переменных и, выполнив те или иные формальные операции, получить все тот же обобщенный индекс, но уже как документ, включающий таблицы, формулы, сгенерированный отчет и т. п.; в-третьих, то и другое вместе взятое, но настроенное на конкретную задачу, то есть ситуацию, при которой вместо исходных переменных  $x_1, x_2, \dots$  в схеме агрегирования фигурируют метафоры ПОЛ, ВОЗРАСТ, ..., а формальный критерий обобщения, реализованный в вычислительной системе, «интуитивно понятен». Предлагаемая модель качественно-количественных представлений социологического исследования не является ни первым, ни вторым, ни третьим. Она представляет собой систему программных классов объектно-ориентированной среды программирования. Общая структура системы показана на рисунке 1, где TThesa — класс, описывающий тезаурус, TMO — метафору, TQO — количественный объект, TAO — алгоритмический объект.

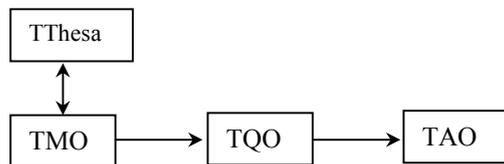


Рис.1. Программные классы модели

<sup>6</sup> Один из списков, определенных в MO.

Модель представляет собой программный аналог качественно-количественных различий в методах социологического исследования и не включает в себя компоненты доступа конечного пользователя к механизмам модели. Одним из способов использования предлагаемой схемы является включение указанных классов в программные проекты, например, создания систем анализа текстов.

#### ЗАКЛЮЧЕНИЕ

Назначение предлагаемой модели — описать соотношение количественных методов и концептуального аппарата гуманитарии. Модель устанавливает рамки, в которых находятся количественные методы. При введении этих рамок с помощью модели мы уходим от вербальных рассуждений, где построение концептов рассматривается как принципиально неформализуемый процесс. И концептуальные и количественные понятия в равной степени представлены в предложенной модели, которая позволяет конструктивным образом синтезировать качественные и количественные данные.

Нашей стратегической целью является доведение этой разработки до стадии проблемно-ориентированной вычислительной среды, предназначенной для создания гуманитарных программных приложений, которые могут использоваться, в частности, для решения задач по интервьюированию, анализу слабоструктурированных данных и текстов. Если бы удалось совместить в одном программном продукте инструментальные средства визуальной среды программирования и возможности работы с концептами предметной области, для которой разрабатывается приложение, цель была бы достигнута. Это позволило бы сделать программирование более эффективным способом работы с концептуальным аппаратом предметной области. Главным достоинством такой среды является возможность объединения двух интеллектуальных стилей — гуманитарного и технического.

*Приложение*

#### АББРЕВИАТУРЫ

ВПКМ — вычислительная поддержка качественных методов  
 КК — качественные и количественные методы и/или данные  
 КА — качественный анализ  
 МО — метафорический объект  
 КО — количественный объект  
 АО — алгоритмический объект  
 МЗ — механизм значений  
 ООП — объектно-ориентированное программирование  
 SQL — Structured Query Language, язык работы с данными в системах управления базами данных  
 ТМО — программный тип метафорического объекта  
 TThesa — программный тип тезауруса  
 TQO — программный тип количественного объекта  
 TAO — программный тип алгоритмического объекта

#### Базовые типы метафор

Тип МО	Список определений	Список ассоциированных объектов	Примечание
MtPure	Нет	Нет	Возможны ссылки в других списках

MtPrim	Нет	Нет	МО внесен в список примитивов тезауруса
MtPrimAlias	Нет	Единственная ссылка на mtPrim	Мнемоимя
MtDefined	Хотя бы один КО	Нет	КО, не зависящий от контекста
mtDefinedAlias	Нет	ровно один КО	Мнемоимя
mtMultiAliasQO	Нет	Несколько ссылок на КО	КО, зависящий от контекста
MtComplex	Хотя бы один КО	Единственная ссылка на КО	КО, частично зависящий от контекста
mtMultiComplex	Хотя бы один КО	несколько ссылок на КО	КО, частично зависящий от контекста
MtProc	Нет	Нет	В списке алгоритмов тезауруса
MtProcAlias	Нет	Единственная ссылка на MtProc	Мнемоимя
MtContext	Нет	Нет	Обязательно наличие ссылки(ок) в списке атрибутов*

\* Один из списков ссылок, определенных в МО

#### ЛИТЕРАТУРА

1. <http://www.scolari.com/> (дата считывания сведений: 21.08.2001).
2. *Tesch R.* Qualitative research: analysis types and software tools. New York: Falmer Press, 1990.
3. *Толстова Ю.Н., Масленников Е.В.* Качественная и количественная стратегии: эмпирическое исследование в широком смысле // Социологические исследования. 2000. № 10. С. 101–108.
4. *NVivo*: Demonstration software for qualitative data analysis. Version 1.1. September, 1999. CD-ROM.
5. <http://www.atlasti.de/atlastneu.html> (дата считывания сведений: 21.08.2001).
6. WinMAX: Software for qualitative data analysis (Demo version, CD-ROM), 2000.
7. *Стивенс Р.* Delphi. Готовые алгоритмы: Пер. с англ. М.: ДМК Пресс, 2001.