

МЕТОДЫ АНАЛИЗА ДАННЫХ

Г.В. Каныгин

КОНТЕКСТНО-ОРИЕНТИРОВАННЫЙ АНАЛИЗ КАЧЕСТВЕННЫХ ДАННЫХ

В статье изложены методы контекстно-ориентированной концептуализации, представляющие собой результат структурного совершенствования функций кодирования и реконструирования данных, лежащих в основе компьютерных методов анализа качественной информации. Базовой единицей концептуализации служит аналитический код, выражающий собой конструктивные свойства социологических понятий, присущие им до прохождения процедуры операционализации. Коды связываются между собой аналитиком в виде множества их простейших изменений, имеющих одну и ту же структуру, позволяющую учитывать контекстную зависимость кодов. Коды и их контекстно зависимые изменения накапливаются в едином тезаурусе, создаваемом аналитиком в человеко-машинном режиме.

Целью концептуализации служит создание явной структуры переходов от терминов, которые вводит аналитик *ad hoc*, к общепринятым понятиям. Аналитик не имеет инструментов, позволяющих ему выразить концептуальный переход в виде единой структуры непосредственно. Единые структуры связности терминов генерируются с помощью специально разработанных алгоритмов, работающих как инструменты анализа текущего состояния тезауруса и контроля связности результатов концептуализации. Представлена концептуализация текстового свидетельства информанта, выполненная автором с помощью созданного им компьютерного пакета, воплощающего собой контекстно-ориентированные структуры и алгоритмы.

Ключевые слова: компьютеризированный анализ качественных данных, контекстно-ориентированный подход, концептуализация социологической информации, качественное социологическое исследование

Введение

В 1991 году Н. Филдинг и Р. Ли предложили термин «компьютерно-ассистируемый анализ качественных данных» (Computer Assisted Qualitative Data Analysis, CAQDAS) с целью обозначения предметной области своего сетевого проекта¹. Современный компьютерно-ассистируемый анализ качественных данных является методологической исследовательской областью, объединяющей ученых многих стран². В данной работе в качестве русскоязычного синонима используется словосочетание *анализ качественных данных* (АКД). АКД представлен множеством компьютерных пакетов, в том числе: Atlas.ti, MAXQDA, NVivo, xSight, Qualrus, Ethnograph³ и др.

Эти пакеты представляют собой класс компьютерных программ, которые, согласно У. Келле, включают в свою архитектуру специальные структуры, называемые функциями кодирования и реконструирования [качественных, нечисловых] данных (coding and retrieval functions)⁴. Функции кодирования и реконструирования данных (ФКР) представляют собой компьютерный инструмент (tool), используемый в человеко-машинном режиме и ассистирующий пользователю при изучении данных, представленных в так называемых нечисловых форматах — видео, текст, аудио и др. В основе ассистирования лежит аппарат аналитических переобозначений, называемых кодами, введение и связывание которых между собой осуществляется самим пользователем. ФКР берут свое начало в действиях социолога, которые он осуществляет в ходе полевых социологических исследований при осмыслении свидетельств информантов⁵.

¹ Gibb G.R., Friese S., Mangabeira W.C. The Use of New Technology in Qualitative Research. Introduction to Issue 3(2) of FQS [35 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*. 2002. May [On-line Journal]. Vol. 3(2). URL: <http://www.qualitative-research.net/fqs-texte/2-02/2-02hrsg-e.htm>

² См. например, http://en.wikipedia.org/wiki/Computer_Assisted_Qualitative_Data_Analysis_Software и <http://onlineqda.hud.ac.uk/index.php> (дата обращения 15.04.2011).

³ URL: <http://www.atlasti.com/>, <http://www.maxqda.com/>, http://www.qsrinternational.com/products_nvivo.aspx, http://www.qsrinternational.com/products_xsight.aspx, <http://www.ideaworks.com/qualrus/index.html>, <http://www.qualisresearch.com/> (дата обращения 15.04.2011).

⁴ Kelle U. Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data // *Sociological Research Online*. 1997. Vol. 2. No 2. URL: <http://www.socresonline.org.uk/2/2/1.html> (дата обращения 05.04.2008).

⁵ Примеры можно найти по адресу http://onlineqda.hud.ac.uk/Intro_QDA/coding_examples.php (дата считывания 15.04.2011).

С момента своего появления в виде компьютерного инструмента ФКР прошли несколько стадий своего развития⁶. Традицией АКД является сравнительный анализ этих функций, представленных в разных пакетах, с целью ознакомления социальных исследователей с многообразными современными методами, позволяющими сделать исследование качественных данных более наглядным и убедительным⁷. В наиболее совершенном виде ФКР представлены в пакетах, которые получили название компьютерных средств построения теорий (theory building software)⁸. Английская исследовательница А. Льюинс предлагает различать в ФКР отдельные составляющие. В том числе инструменты (tools) — поиска контента (content searching), связывания (linking), кодирования (coding), запросов (query), пояснений и аннотаций (writing and annotation), визуализации в виде графов (mapping or networking)⁹.

Выполненные сравнительные исследования реализаций функции кодирования в современных пакетах позволяют увидеть аналитическое своеобразие этих функций. ФКР не существуют в виде аналитической формулы или однозначно определенной структуры. Например, как синус $y = \sin(x)$ или бинарное дерево. Каждый компьютерный пакет (Atlas.ti, xSight, maxQDA и др.) предлагает свою версию функций кодирования. Исходя из общего определения функциональной зависимости, эти версии различаются в трех основных аспектах. Во-первых, областью определения, т. е. форматом нечисловых данных, которые подвергаются преобразованию с помощью ФКР. Во-вторых, областью значений функций кодирования, т. е. кодами различных типов (активными, выполняемыми, аннотациями и т. д.). В-третьих, структурными отношениями, которые устанавливаются между кодами с помощью функций кодирования. Эти отношения могут быть представлены в виде деревьев, сетей, списков и др.

⁶ Mangabeira W.C. Computer Assistance, Qualitative Analysis and Model Building // Information Technology for the Social Scientist / Raymond M. Lee (ed.). London: UCL Press, 1995. P. 129–146; Gibbs G.R., Friese S., Mangabeira W.C. The Use of New Technology in Qualitative Research. Introduction to Issue 3(2) of *FQS: Forum Qualitative Sozialforschung // Qualitative Social Research*. 2002. May. Vol. 3(2). URL: <http://www.qualitative-research.net/fqs-texte/2-02/2-02hrsg-e.htm>

⁷ Tesch R. Qualitative Research: Analysis types and software tools. N. Y. : Taylor & Francis Ltd, 1990; Lewins A. () Using Qualitative Software: A Step-by-Step Guide / Silver C. London: Sage publications, 2007.

⁸ URL: http://onlineqda.hud.ac.uk/Intro_CAQDAS/what_packages_are_available/theory_building_software.php

⁹ URL: <http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas>

Структурный аппарат АКД, т. е. ФКР, служит конструктивной основой многих методологических направлений современного качественного исследования¹⁰. Особо следует отметить, что АКД часто рассматривается в качестве инструментального аппарата исследований, основанных на методологии grounded theory¹¹. Прикладная применимость компьютерных пакетов АКД постоянно подчеркивается их разработчиками. Например, на сайте пакета Atlas.ti можно видеть длинный список организаций, которые используют этот пакет в своей работе.

Как свидетельствуют современные исследователи, занятые в прикладных проектах¹², пакеты анализа нечисловой информации успешно применяются при управлении и визуализации нечисловых данных больших объемов. Однако существуют прикладные задачи, в которых применение обсуждаемых пакетов встречает серьезные трудности. Например, в случае дискурсного анализа текстовых сообщений функции кодирования вряд ли помогут исследователю [ibid.].

Подобные трудности наталкивают на мысль о необходимости совершенствования структурного аппарата кодирования, реализованного в современных пакетах. В поисках источников такого совершенствования рассмотрим критические замечания социологов, обращающихся к указанным пакетам в своей работе. Основываясь на своем опыте работы с пакетом NUD*IST, К. МакМиллан и С. МакЛаклан утверждают, что с его помощью им не удалось удовлетворительно воспроизвести в контексте качественного исследования свое общественное представление о теории¹³. С. Барри обращает внимание на «безликость» (homogeneity) компьютерных методов как недостаток в качестве инструмента гуманитарного проекта¹⁴. С. Торн называет

¹⁰ В этом можно убедиться по адресу <http://onlineqda.hud.ac.uk/methodologies.php>

¹¹ Charmaz K. Grounded Theory: Objectivist and Constructive Methods / Norman K. Denzin & Yvonna S. Lincoln (eds.) // Handbook of Qualitative Research. 2nd edition. Thousand Oaks, Ca.: Sage, 2000. P. 509–535; Bringer J.D. Using Computer-Assisted Qualitative Data Analysis Software to Develop a Grounded Theory Project / Johnston, L.H., Brackenridge, C.H. // Field Methods. August 2006. 18. P. 245–266.

¹² Например, MacMillan K. Theory-Building with Nud*Ist: Using Computer Assisted Qualitative Analysis in a Media Case Study // Sociological Research Online. 1999. Vol. 4. №. 2. URL: http://www.socresonline.org.uk/4/2/macmillan_mclachlan.html

¹³ Ibid.

¹⁴ Barry C.A. Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared / Sociological Research Online. 1998. Vol. 3. No. 3 [2.1]. URL: <http://www.socresonline.org.uk/3/3/4.html> (дата обращения 01.06.2006).

АКД «наиболее запутанной и таинственной (complex and mysterious) частью всего качественного исследования»¹⁵. Дополнительные суждения по вопросам применения пакетов АКД в качественном социологическом исследовании могут быть найдены в Интернете¹⁶.

Пытаясь кратко резюмировать высказанные замечания, заметим, что, несмотря на их терминологические несходства, они имеют одну общую черту. Высказанная критика не позволяет понять, каким образом следует структурно усовершенствовать функции кодирования. Например, разъяснение К. МакМиллан и С. МакЛахан, в котором они обращаются к общенаучному толкованию термина «теория», не позволяет понять, как именно реализовать это толкование в составе аппарат кодирования.

Аналогичная черта может быть увидена в теоретических разработках в области качественного исследования, которые также обращаются к анализу нечисловых данных. Так, К. Чармас развивает конструктивистское направление *grounded theory*¹⁷. Создавая свои теоретические конструкции, американская исследовательница, объясняет процесс преобразования кода в категорию (*ibid.*). В контексте объяснения становится понятным, что различие между кодом и категорией состоит не только в естественноязыковых коннотациях использованных слов. Код — это структурная единица, программно в пакете. Категория — это вербальная конструкция, дистанцированная от кода и предназначенная для выражения «сугубо человеческого» понимания исследуемого случая. Таким образом код имеет структурное выражение, а категория — это вербально обозначенное кредо исследователя. К сожалению, естественно языковое понимание, в рамках которого остается термин «категория», не позволяет использовать теоретические рассуждения для целей совершенствования структурного аппарата АКД.

Можно сказать, что как прикладные исследователи, так и социальные теоретики, опирающиеся в своих рассуждениях на аппарат ФКР, не идут в аналитических предложениях дальше заимствования существующих структурных моделей и методов, предлагаемых современными пакетами АКД. Подобная сложившаяся методологическая

¹⁵ *Thorne S.* Data analysis in qualitative research // *Evidence Based Nursing.* 2000. № 3. P. 68–70.

¹⁶ Например, по адресу http://onlineqda.hud.ac.uk/Intro_CAQDAS/software_debates.php

¹⁷ *Charmaz K.* Op. cit; *Bringer J.D.* Op. cit.

практика не позволяет сформулировать оригинальные конструктивные предложения, которые могли бы быть положены в основу структурного совершенствования ФКР.

Отдельно упомянем о неформальных обоснованиях существующих методов АКД, которые формулируются разработчиками программного обеспечения. Дж. Зейдель видит программу Ethnograph как инструмент, assisting при выполнении «общенаучных» действий — «наблюдения», «обобщения» и «размышления» (noticing, collecting, thinking). А. Куккарц и У. Куккарц считают, что их пакет MAXqda построен на веберовском представлении об идеальном типе. С точки зрения Г. Хубера, пакет AQAUD¹⁸ призван выразить взгляды К. Поппера, перенесенные в контекст качественного исследования. Не вызывает сомнения, что подобные обоснования правомерны ввиду самого существования пакетов, нашедших мировое признание. Однако стоит обратить внимание, что подобные обоснования, как правило, возникают *post factum*, а их общий результат известен заранее — функции кодирования.

Суммируя сделанный краткий обзор компьютерных средств АКД и ситуаций их применения, можно высказать суждение, их структурное развитие не является ответом на исследовательские вызовы, возникшие в области прикладной или теоретической социологии. На мой взгляд, источник такого развития следует искать в области информатики, которая лежит вне области научных проблем социологии. В пользу такого суждения ясно свидетельствует, например, работа Т. Мура¹⁹, в которой достоинства XML формата предлагается использовать для целей структурного развития программ анализа качественных данных. В этой связи показательно вспомнить отличительные черты известных пакетов. MAXQDA развивается на принципах анализа качественных данных, но ориентирован на анализ сугубо текстовой информации. Разработчики Atlas.ti с самого начала своего проекта предложили и программно реализовали структурную идею герменевтического блока (hermeneutic unit), что означает совместное использование данных текстового, видео и аудио форматов. Последние

¹⁸ См. информацию разработчика на сайте <http://www.aquad.de/eng/index.html> (дата обращения 01.07.2011).

¹⁹ Muhr T. Increasing the Reusability of Qualitative Data with XML. [64 paragraphs]. Forum Qualitative Sozialforschung / Qualitative Social Research. 2000. December. Vol. 1(3). URL: <http://www.qualitative-research.net/fqs-texte/3-00/3-00muhr-e.htm> (дата обращения 05.06.2006).

версии NVivo строятся на мощном структурном языке запросов, реализованном в Microsoft SQL Server, и т. д. Информационные основания программ АКД, фактически представляющие собой их возможности воспроизводить программистские ресурсы и функции (форматы данных, средства экспорта-импорта, инструменты поиска и др.), могут быть также прослежены в их классификации²⁰. Такого сорта формально-информационная природа пакетов АКД говорит об их «методологической удаленности» от запросов исследователей, занятых в социологических проектах.

«Методологическая удаленность» пакетов АКД от предмета социологических исследований становится более наглядной, если вспомнить о существовании социологических проблем или вызовов, которые не зависят от того, какой теоретической традиции придерживается социолог, проводящий исследование с помощью «качественного» компьютерного пакета. Одним из таких вызовов является проблема оценивания качества исследований, основанных на нечисловой информации²¹. Сравнительный анализ пакетов АКД показывает, что в составе функций кодирования отсутствует какой-либо специальный инструмент, который позволял бы контролировать аналитические выкладки социолога, осуществляемые с помощью этих функций. Таким образом, оказывается, что, с одной стороны, пакеты анализа нечисловой информации общепризнанны как специфический инструмент качественного исследования. С другой стороны, применение этого инструмента в прикладном исследовании, построенном на идеях гуманитарной методологии, не способствует решению проблемы качества этого исследования.

Многочисленные методологические разработки, посвященные решению проблемы качества исследовательских подходов в гуманитарной области, ссылки на которые можно найти в работах М. Эстерби-Смит, К. Голден-Биддл, К. Локке²² и С. Трэйси²³, предлагают соответствующие критерии в виде вербальных процедур. Будучи по своей

²⁰ См. http://onlineqda.hud.ac.uk/Which_software/what_packages_are_available/index.php

²¹ Collingridge D. S., Gantt E. E. The Quality of Qualitative Research. // American Journal of Medical Quality. 2008. September/October. Vol. 23. No 5. P. 389-395.

²² Easterby-Smith M., Golden-Biddle K., Locke K. Working With Pluralism: Determining Quality in Qualitative Research // Organizational Research Methods. 2008. July. Vol. 11. No 3. P. 419-429.

²³ S. J. Tracy (2010) Qualitative Quality: Eight “Big-Tent” Criteria for Excellent Qualitative Research // Qualitative Inquiry October 1. P. 1-15.

природе продуктами естественного языка, такие процедуры, так же как и функции кодирования, не содержат в себе никаких средств, которые можно использовать для ожидаемого оценивания результатов исследования без привлечения экспертов. Например, если мы используем критерий «большого тента» (“big-tent” criteria)²⁴, мы вынуждены обратиться к эксперту, чтобы он смог ответить, скажем, на вопрос: «Правильно ли выбран образец или контекст, если принять во внимание цель исследования?» Без эксперта такие критерии в принципе не в состоянии работать.

Мои замечания показывают, что критериология (criteriology)²⁵, призванная дать средства оценивания качественных стратегий в социологии, во-первых, развивается вне компьютерных методов АКД. Во-вторых, не предлагает формально сформулированного критерия, который мог бы быть включен в состав функций кодирования. При этих обстоятельствах предлагаемые критерии качества критически зависят от лиц, которые оказываются включенными в процедуру оценивания. Тем самым, для того чтобы проводить качественное исследование, которое может быть названо профессиональным, часто основным фактором признается соответствующий опыт, передающийся из поколения в поколение²⁶. Совсем откровенно ситуацию с оцениванием характеризует С. Торн: «если результаты исследования прозвучали убедительно для тех, кому они предназначены, то и исследование признается успешным»²⁷.

Отсутствие в составе аппарата АКД какого-либо инструментального критерия, способного оценивать результаты кодирования, естественно отнести к недостаткам современного АКД. Для устранения этого недостатка сегодняшних методов качественного исследования в данной работе представлен специальный аппарат кодирования, в состав которого входит критерий связности его результатов. Оригинальные структурные модели и методы их использования, впервые предложенные автором, структурно ориентированы на выражение общенаучного тезиса о необходимости учета контекста для любого научного высказывания. Отсюда подход в целом назван контекстно-ориентированным.

²⁴ Ibid. P. 5.

²⁵ Seale C. Quality Issues in Qualitative Inquiry //Qualitative Social Work. 2002. March. Vol. 1. No 1. P. 97–110.

²⁶ Ibid. P. 107–108.

²⁷ См. Thorne S. Data analysis in qualitative research.

Все утверждения, сформулированные в данной работе, проверены путем их реализации в виде созданного автором пакета АКД. Для иллюстрации предлагаемых методов использованы результаты концептуализации текстового свидетельства информанта. Иллюстрирующие рисунки (за исключением рис. 1) представляют собой пользовательские интерфейсы указанного пакета.

Основные определения и структуры

Базовыми инструментами контекстно-ориентированного аппарата кодирования являются, во-первых, три вида аналитических обозначений (дескрипторов) — термин²⁸ (понятие, неодушевленный объект), актор (одушевленный объект) и действие (ассоциация понятия и актора).

Во-вторых, простейшее в рамках подхода структурное отношение между дескрипторами, названное контекстно-фиксированным разъяснением (КФР). КФР представляет собой обозначение *единичного изменения* дескрипторов, осуществляемого аналитиком. Моделью изменения является отношение «родитель—дети», которое трактуется как *преобразование* аналитиком дескриптора в положении «родитель» в набор других дескрипторов, составляющих список «детей».

Нотация КФР может быть представлена с помощью формулы (1).

$$t^{(c)}:t^{(p)} \rightarrow \{t_i\}, (1),$$

где $t^{(p)} \rightarrow \{t_i\}$ — структура «родитель—дети», каждым элементом которой является *термин*: $t^{(p)}$ — «родитель», а $\{t_i\}$, $i=0, \dots, n-1$ — список «детей». $t^{(c)}$ — это термин, с помощью которого обозначено *изменение* одного понятия на группу других, выраженное отношением «родитель—дети». В зависимости от положения термина в структуре КФР $t^{(c)}$ называется *контекстом*, $t^{(p)}$ — *целевым* или *разъясняемым термином*, любой t_i — *пояснением*.

КФР имеет интуитивно ясную интерпретацию. Например, проводя опрос в 2004 г., социолог обращается к представлению о социально-

²⁸ Слово «термин» употребляется как синоним слова дескриптор в том случае, если не важен вид дескриптора — понятие, актор или действие. Подобная синонимия объясняется привычностью поименования «термин» или «понятие» для обозначения простейшей структурной единицы аналитических построений.

демографических характеристиках опрашиваемых. Развиваемый подход требует, чтобы исследователь обязательно указал, в каких условиях, иначе говоря, относительно какого контекста сформировано его суждение. С учетом возможных для данного случая интуитивно понятных дескрипторов (1) может получить вид (2):

ИССЛЕДОВАНИЕ 2004²⁹: СОЦДЕМ
→ПОЛ, ВОЗРАСТ, СОЦИАЛЬНОЕ ПОЛОЖЕНИЕ (2)

Особо отметим два интуитивных основания трактовки КФР в качестве базовой структуры кодирования. Во-первых, исследователь фиксирует с помощью этой модели не «объект», существование или характеристики которого всегда могут быть оспорены другим человеком в силу известных трудностей «объективного» выражения «субъективного» смысла [16, с. 392], а *факт изменения* собственных понятий. Именно факт, т. к. термины в КФР указаны явно, в отличие от интуитивно постигаемых «объектов», предполагаемых за каждым из аналитических обозначений. Такой факт наблюдается всеми «реципиентами» процесса кодирования, в то время как соответствие «образа» его аналитическому обозначению всегда остается «внутри» автора такого действия.

Во-вторых, структура КФР означает, что факт изменения понятий соотнесен с явно указанным обозначением. Иначе говоря, структуру КФР можно интуитивно понять, как обозначение условий, при которых осуществляется изменение терминов.

Совокупность терминов, вводимых аналитиком в процессе контекстно-ориентированного кодирования и связываемых им между собой с помощью множества КФР, названа *тезаурусом*. Обратим внимание, что, во-первых, термин, представленный в тезаурусе, может не входить ни в одно КФР. Такой термин игнорируется ниже описываемыми алгоритмами. Во-вторых, связи между терминами содержатся в тезаурусе только в виде структур КФР³⁰.

Интуитивный смысл тезауруса в том, что он включает в себя *все изменения* понятий, которые аналитик счел нужным указать для сведения его вновь формируемых терминов к общеизвестным представлениям.

²⁹ Здесь и далее экземпляры терминов, используемых в составе структурных моделей развиваемого подхода, обозначены с помощью заглавных букв.

³⁰ NB! Никаких агрегированных структур тезаурус не содержит.

Термин, входящий в состав тезауруса, считается *контекстно-определенным*, если он является разъясняемым в составе хотя бы одного КФР. Термин t *контекстно определен* относительно термина t_0 , если t является целевым, а t_0 контекстом в одном и том же КФР, например, $t_0:t \rightarrow \{t_i\}$. Вышеупомянутый СОЦДЕМ в соответствии с формулой (2) является контекстно-определенным понятием относительно контекста ИССЛЕДОВАНИЕ 2004.

Один и тот же термин может быть определен относительно нескольких контекстов. Такой термин назван *поликонтекстным*. Если термин является разъясняемым только в одном КФР тезауруса, то он именуется *моноконтекстным*.

Отметим, что контекстное определение термина с помощью КФР не зависит от состава пояснений, использованных в этом КФР. В частности, множество пояснений может быть пустым. Такое множество будем обозначать $\{\dots\}$.

Алгоритм построения иерархии контекстов

Структуры связей между терминами таковы, что позволяют предложить алгоритм, способный агрегировать дескрипторы, принадлежащие различным КФР тезауруса на позиции контекста, в единую структуру соподчинения. Этот алгоритм строится на эвристическом правиле, устанавливающем наличие или отсутствие связи между двумя терминами, использованными в двух разных КФР на позиции контекста.

Если структура двух КФР такова, как показано в формулах (3) и (4), т. е. если один и тот же термин (t_1) является контекстом в одном из них (КФР_1), и он же служит разъясняемым (контекстно определен) в другом (КФР_0), то t_0 является контекстным обобщением t_1 ³¹.

$$\text{КФР}_0 = t_0:t_1 \rightarrow \{\dots\}^{32}, \quad (3)$$

$$\text{КФР}_1 = t_1:t_2 \rightarrow \{\dots\}, \quad (4)$$

Поясним на примере неформальные основания введенного правила. Допустим, что аналитик создал два интуитивно понятных КФР.

³¹ Этим же словосочетанием названа процедура, основанная на описываемых правилах.

³² Будем считать множество пояснений пустым, т. к. его состав не важен для фиксации контекстной определенности целевого термина.

В первом он предложил рассматривать ДОМ в контексте УЛИЦА (5), т. е.

$$\text{УЛИЦА: ДОМ} \rightarrow \{\dots\}. \quad (5)$$

Во втором тот же термин УЛИЦА разъяснен в контексте ГОРОД (6), т. е.

$$\text{ГОРОД: УЛИЦА} \rightarrow \{\dots\}. \quad (6)$$

Другими словами, термин, использованный как контекст, сам оказывается разъясняемым в другом контексте.

Использование одного и того же термина УЛИЦА в разных позициях двух КФР является основанием считать, что контекст ГОРОД оказывается более общим по отношению к контексту УЛИЦА. Другими словами, ГОРОД является контекстным обобщением (расширением) УЛИЦА³³.

Такому суждению есть два интуитивных обоснования. Во-первых, термины УЛИЦА и ГОРОД употреблены как контексты. Во-вторых, будучи разъясненным в контексте ГОРОД, УЛИЦА, с точки зрения выполнившего такое разъяснение аналитика, составляет только часть тех феноменов или сюжетов, которые поименованы термином ГОРОД. Это утверждение может быть пояснено тем, что относительно одного и того же контекста всегда возможно определение неограниченного числа понятий. Например, КВАРТАЛ также может быть контекстно определен относительно термина ГОРОД.

В общем случае алгоритм построения контекстного обобщения исходит из заданного пользователем КФР, например, введенного с помощью формулы (7).

$$t_0 : t \rightarrow \{\dots\}, \quad (7)$$

На исходном нулевом шаге алгоритм находит в тезаурусе множество терминов, которые контекстно обобщают (включают в себя) t_0 . Обозначим это множество $\Psi_0 = \{t_{0i}\}$, $i=0, \dots, m_0-1$, где m_0 — число контекстов, в которых введен t_0 . Если множество Ψ_0 не пусто, то для каждого его элемента алгоритм строит множество контекстных обобщений,

³³ Можно также сказать, что УЛИЦА контекстно включена в ГОРОД.

создавая тем самым совокупность множеств $\Psi_{0i} = \{t_{0ij}\}$, $i = 0, \dots, m_0 - 1$, $j = 0, \dots, m_{0i} - 1$, где m_{0i} — число контекстов, в которых введен t_{0i} .

Любой из последующих шагов приводит к разворачиванию каждого из элементов множества контекстных обобщений Ψ ³⁴, полученных на предыдущем шаге. Алгоритм прекращает свою работу, когда все множества контекстных обобщений, полученных на очередном шаге, оказываются пустыми. Результаты работы алгоритма представимы в виде графа типа «река», фрагмент формирования которого показан на рис. 1.

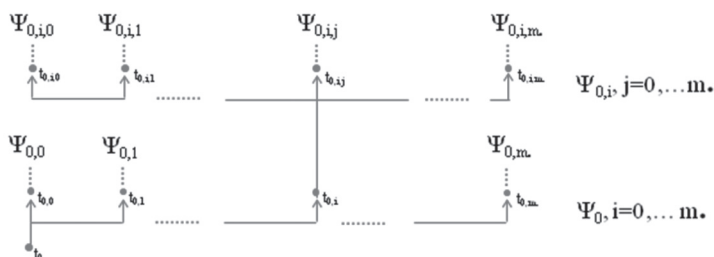


Рис. 1. Контекстное обобщение термина t_0

В пакете соподчинение контекстных включений, порождаемое заданным контекстом, выражается отношениями «родитель-дети» в составе алгоритмически создаваемого графа. Этот граф назван иерархией контекстов. Для приведенного простейшего примера иерархия контекстов показана на рис. 2 (см. также рис. 6).

³⁴ Число позиций в определении индекса контекста зависит от шага алгоритма. В тексте этот индекс обозначен как жирная точка (.) и формируется следующим индуктивным образом.

(0) — индекс контекста, для которого строится контекстное обобщение (исходный контекст, шаг 0)

(0,i) — индекс контекста, который является i -м контекстным обобщением контекста с индексом 0; $i = 0, \dots, m$.

(0,i,j) — индекс контекста, который является j -м контекстным обобщением контекста с индексом 0,i; $j = 0, \dots, m$.

...

(0,i,j,...,k) — индекс контекста, который является k -м контекстным обобщением контекста с индексом (0,i,j,...); $k = 0, \dots, m$.

m — некоторое целое число, указывающее поликонтекстность контекста с предыдущим индексом.

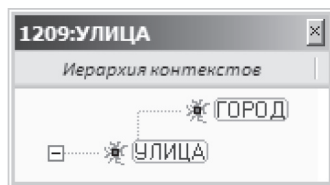


Рис. 2. ГОРОД — контекстное обобщение термина УЛИЦА

Алгоритм связывания терминов

Рассмотрим алгоритм, который позволяет на основе тезауруса и иерархии контекстов сгенерировать связи термина с другими терминами. Процесс такой генерации назван связыванием термина. Ее результат представим в виде графа типа дерево, названного терминологическим. Напомним, что любые алгоритмические операции проводятся в развиваемом подходе только с контекстно-определенными дескрипторами. Поэтому в узле графа представлена пара терминов. Один элемент пары — дескриптор (понятие, актор или действие), другой — контекст его определения, также выраженный дескриптором.

Исходным шагом алгоритма является указание аналитиком термина t_0 , определенного в контексте $t^{(c)}$ и имеющего непустое множество пояснений, как показано в формуле (8):

$$t^{(c)} : t_0 \rightarrow \{t_{0i}\}, \quad (8)$$

где $i = 0, \dots, n-1$, $n-1$ — число пояснений, $n > 1$.

При алгоритмическом построении терминологического графа иерархия контекстов термина $t^{(c)}$ предварительно автоматически генерируется с помощью вышеописанного алгоритма.

Дальнейшая работа алгоритма связывания сводится к известным методам построения древовидного графа³⁵ за счет того, что структура каждого единичного перехода «родитель—дети», автоматически извлекается из состава соответствующего КФР. Так, на исходном шаге строится корневой узел дерева, воспроизводящий «родителя» t_0 , а множество его пояснений $\{t_{0i}\}$ рассматривается как список «потенциальных детей».

³⁵ Стивенс Р. Delphi. Готовые алгоритмы / пер. с англ. П.А. Мерещук. М. : ДМК Пресс ; СПб. : Питер, 2004.

Ключевая особенность алгоритма по сравнению с указанными известными методами состоит в том, что множество пояснений, составленное аналитиком, не становится набором «детей» безусловно. Оно может быть модифицировано в зависимости от того, в каком отношении контекст каждого из пояснений находится с контекстом целевого термина.

Проблема, решаемая алгоритмом при формировании каждого перехода «родитель-дети» на основе КФР, существует уже на исходном шаге. В соответствии с высказанным выше замечанием каждый из терминов формулы (8) должен быть контекстно-определенным. Если целевой термин t_0 по (8) определен однозначно относительно $t^{(c)}$, то контекст любого из пояснений невозможно определить по этой формуле, т. к. он в ней не указан.

Для разрешения возникающей неопределенности алгоритм отыскивает по тезаурусу множество возможных контекстов каждого из пояснений t_{0i} , т. е. строит $\{t_{0ik}\}$, где $k = 0, \dots, m_{0i}-1$; $m_{0i}-1$ — число контекстов, в которых определено пояснение, т. е. термин t_{0i} . Если среди этих контекстов находится такой, который согласован (см. следующий абзац) с контекстом «родителя», т. е. $t^{(c)}$, то t_{0i} включается в автоматически генерируемый список «детей». В противном случае отклоняется.

В общем случае алгоритм связывания контекстно-определенных терминов может использовать различные правила согласования контекстов. В существующей версии пакета для отработки всей инфраструктуры подхода используются два, которые автору представились наиболее простыми. Во-первых, если пояснение не является контекстно-определенным, то оно не включается в состав «детей» «родителя», который генерируется из целевого термина при алгоритмическом «расширении» этого термина.

Во-вторых, если ни один контекст определения пояснения не входит в иерархию контекстов, предварительно построенную для контекста «родителя», то такое пояснение не включается в состав «детей», алгоритмически генерируемого отношения «родитель-дети» (результат работы алгоритма см. рис. 7).

Поясним особенность алгоритма агрегирования на интуитивно понятном примере. Пусть в уже упоминавшемся контексте УЛИЦА аналитик разъясняет термин ДОМ через четыре других термина следующим образом:

УЛИЦА: ДОМ → СТЕНА, КРЫША, ФУНДАМЕНТ, ЖИТЕЛЬ (9)

Причем контекстная определенность каждого из пояснений задается КФР, показанными в формулах 10–13, в частности, термин ЖИТЕЛЬ введен в состав тезауруса в контексте СТРАНА, который не входит в состав иерархии контекстов термина УЛИЦА (см. рис. 2).

В этом случае автоматически формируемый переход «родитель—дети» не будет включать термин ЖИТЕЛЬ (см. рис. 3).

- | | |
|------------------------|------|
| ГОРОД: СТЕНА→{...} | (10) |
| УЛИЦА: КРЫША→{...} | (11) |
| УЛИЦА: ФУНДАМЕНТ→{...} | (12) |
| СТРАНА: ЖИТЕЛЬ→{...} | (13) |

Рис. 3. Результат генерации единичного перехода «родитель—дети» для контекстно-фиксированного разьяснения термина ДОМ в контексте УЛИЦА с учетом контекстной определенности каждого из пояснений

Аналитик вправе пополнить контекстное определение термина ЖИТЕЛЬ, сделав его поликонтекстным, например, добавив еще одно КФР, выраженное формулой (14):

- | | |
|---------------------|------|
| ГОРОД: ЖИТЕЛЬ→{...} | (14) |
|---------------------|------|

В этом случае список детей в соответствующем алгоритмически генерируемом единичном переходе изменится на тот, который показан на рис. 4.

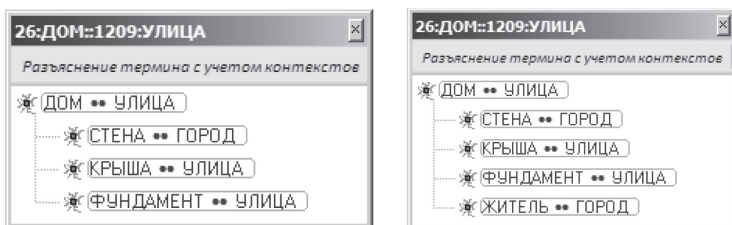


Рис. 4. Термин ЖИТЕЛЬ определен в двух контекстах, один из которых (ГОРОД) входит в иерархию контекстов термина УЛИЦА (рис. 2), а другой (СТРАНА) — нет

Концептуализация в анализе качественных данных

Термин концептуализация имеет ряд толкований³⁶. Согласно этим толкованиям, применительно к предметной области анализа качественных данных концептуализация может трактоваться как перевод потока сведений, сообщаемых информантом, в аналитические обозначения (коды), создаваемые аналитиком. В основе такой трактовки лежит методологическое допущение, возникшее в анализе нечисловой информации еще во времена «карандаша и бумаги»: фрагмент данных, полученных от информанта, считается фактологическим обоснованием кода, т. е. обозначения, вводимого социологом и не содержащимся в свидетельстве информанта. Фактологическое обоснование со стороны информанта предполагается необходимым, т. к. в отличие от социолога последний мыслится свидетелем событий или явлений, о которых рассказывает.

Указанное допущение объясняет цель концептуализации, достижение которой возлагается на ее автора — продемонстрировать связь вновь создаваемых аналитических обозначений с существующими свидетельствами.

Однако возникает вопрос, каким образом можно «структурно обосновать» код, являющийся продуктом аналитического действия социолога? С помощью функций кодирования аналитик всегда имеет возможность связать его новое обозначение с фрагментом текста информанта, что, собственно, и выражает структурно «факт обоснования». Аналитическая процедура, которая демонстрирует такого сорта обоснование, обычно называется открытым кодированием³⁷. Любой код, полученный в результате этой процедуры, представляет собой «фактическую цитату». Особенностью этой цитаты служат два обстоятельства. *De facto* один и тот же феномен, явление, событие и тому подобное обозначены дважды: в виде фрагмента свидетельства информанта и его переобозначения с помощью кода, введенного социологом. Таким образом, один и тот же «факт», по предположению, декларируется понятным, по крайней мере, двум лицам — информанту и аналитику.

³⁶ См. например, <http://ru.wikipedia.org/wiki/онтология>; <http://www.slovopedia.com/6/202/770654.html>

³⁷ Charmaz K. Grounded Theory: Objectivist and Constructive Methods / Norman K., Denzin & Yvonna S. Lincoln (Eds.) // Handbook of Qualitative Research : 2nd edition. Thousand Oaks, Ca.: Sage, 2000. P. 509–535.

Открытое кодирование структурно как бы закрепляет «взаимопонимание» двух участников концептуализации по поводу «одного и того же факта». Однако такое закрепление выполняется только одним из этих участников, причем сама связь между фрагментом качественных данных и аналитическим кодом выглядит «методологически необоснованной». С одной стороны, открытые коды являются «фактологически обоснованными». С другой, как показывает весь ход концептуализации с помощью аппарата кодирования, эти коды сами служат основой для введения новых кодов. Так возникают аналитические коды³⁸, которые уже не имеют «фактологического обоснования» в виде непосредственной связи с фрагментом данных, представленных информантом.

На мой взгляд, не следует структурно разграничивать коды в зависимости от наличия у них явной связи с фрагментом качественных данных. Любой код, появляющийся в ходе концептуализации, должен иметь один и тот же структурный механизм своего обоснования. Выше введенные контекстно-ориентированные модели позволяют ввести такой механизм в виде отношений между кодами, которые оказываются аналогичными структурам объектно-ориентированного программирования. Ниже показано, каким образом такие отношения появляются в ходе прикладной концептуализации. Теперь же попытаюсь пояснить, каким образом предлагаемые решения связаны с традицией концептуализации, характерной для современного анализа качественных данных.

Отмеченные выше «фактологические цитаты», которые обязательно возникают в результате открытого кодирования, примечательны, по крайней мере, в двух отношениях. Во-первых, аналитик, он же пользователь пакета, который выполняет кодирование, при цитировании такого рода покидает собственно аппарат кодирования, т. е. операции, которые ему доступны с кодами с помощью инструментов определенного пакета. И входит в область других операций, т. е. действий, разрешенных в том же пакете с качественными данными, т. е. со значениями типов *text*, *video*, *audio* и др., конкретный набор которых предусмотрен разработчиками соответствующего пакета. Во-вторых, указанное изменение состава операций свидетельствует о «парадигмальном разрыве», который происходит при открытом кодировании. Формально одна парадигма обслуживается аппаратом кодирования,

³⁸ См. *open coding vs. descriptive coding* по адресу http://onlineqda.hud.ac.uk/Intro_QDA/coding_examples.php

т. е. выражена функциями кодирования в части их методов связывания кодов (аннотации, списки, деревья, сети и т. д.). Другая — это машина качественных данных, структурные методы которой появились вне проблематики анализа качественных данных.

Используя это наблюдение, будем понимать под парадигмой сложившуюся область представлений (операционально определенную систему понятий), природа определенности которой не играет роли. Это может быть как «объективно» определенная предметная область (скажем, городская материальная среда), так и «конструктивистски возникшая» система мнений, например, в результате договоренности группы лиц здесь и сейчас³⁹. Тем самым, парадигма — это не традиционная предметная область, обладающая «природным содержанием», а скорее группа людей, использующая для коммуникации между собой один и тот же конструктивно определенный аппарат.

В традиции АКД таким аппаратом, выходящим в своей операциональной определенности за рамки структур кодирования, служит текст (или любая машина данных, для которых определена операция фрагментирования). С позиций КО АКД работа по концептуализации должна исходить не из наличного текста, подлежащего аналитическому переосмыслению, а в построении связной понятийной системы, которая обеспечивала бы переходы из кодирования (т. е. аппарат оперирования неоперационализируемыми метками) в другие парадигмы.

Фактически концептуализация текста, осуществляемая средствами АКД, предполагает, что, во-первых, существуют некие лица, которые обладают специальным знанием, организованным, главным образом, с помощью языка. Во-вторых, «сторонний наблюдатель» переформулирует это знание с помощью методов кодирования. В контекстно-ориентированном АКД на аналитика возлагается задача не только сформировать собственный корпус кодов в виде тезауруса, но и указать термины, через которые, по его мнению, осуществляется концептуальный переход от его терминологических построений в существующие предметные области.

Существование такого перехода не означает «фактологического» обоснования построений самого аналитика, но указывает на «парадигмальный» разрыв процесса концептуализации, осуществляемого

³⁹ Такое толкование согласуется с подразделением методов grounded theory на объективистские и конструктивистские, описанным, например, К. Чармас, см. сноску 36.

средствами кодирования. Такой разрыв, по-видимому, всегда связан с двумя обстоятельствами, отмеченными выше — изменением состава операций, которые становятся допустимыми с понятиями, и изменением требований к компетенции акторов, которые с ними работают.

В разрабатываемых методах эта идея «парадигмального» разрыва, который присущ аппарату кодирования в АКД, реализуется за счет введения в состав тезауруса *предопределенных терминов*, которые по смыслу аналогичны «фактологическому цитированию», возникающему при открытом кодировании. В общем случае за таким термином стоит некоторое множество специальных операций, которые не входят в состав аппарата кодирования. В существующей ныне программной реализации включен только один такой термин, с помощью которого пользователю пакета предоставляется возможность текстового оперирования. Этот важный частный случай широко распространен в практике применения пакетов АКД и использован для отработки всей инфраструктуры развиваемого подхода.

Контекстно-ориентированная концептуализация

Следуя сложившейся в АКД практике переосмысления свидетельств информанта с помощью вновь вводимых аналитических обозначений, я буду понимать под контекстно-ориентированной концептуализацией процедуру связывания вновь вводимых терминов, создаваемых аналитиком, с высказываниями информанта. Как обычно, осуществление этой процедуры возлагается на аналитика.

Для иллюстрации техники контекстно-ориентированных методов воспользуемся текстовым свидетельством молодого человека о плюсах и минусах совместного проживания с родителями⁴⁰.

В соответствии с требованием контекстно-ориентированного подхода поименуем весь комплекс суждений молодого человека — о его переживаниях, отношениях с родителями, обстоятельствах своего проживания с ними и т.п. — новым термином **РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА**, с учетом условий, при которых мы рассматриваем размышления информанта, обозначив их как **СВИДЕТЕЛЬСТВО ИНФОРМАНТА (СЮЖЕТ, КОНТЕКСТ)**.

⁴⁰ Пример взят на песчупсе http://onlineqda.hud.ac.uk/Intro_QDA/Examples%20of%20coding.php и переведен автором на русский язык.

Естественно считать терминологическую пару РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА в контексте СВИДЕТЕЛЬСТВО ИНФОРМАНТА (СЮЖЕТ, КОНТЕКСТ) собственными исходными представлениями аналитика. Ввиду того, что такие представления формулируются *единственным* актором⁴¹, они могут характеризоваться как несформированные или сугубо индивидуальные («субъективные») и подлежащими разъяснению.

В контекстно-ориентированном АКД на аналитика наряду с формулированием собственных терминов *ad hoc* возлагается задача указания тех понятий, которые он считает финальными в рамках его аналитических построений и означающими переход в «другую» парадигму. В случае анализируемого текста такими финальными понятиями («терминами-цитатами») концептуализации оказываются, например, НЕ С КЕМ ПОГОВОРИТЬ, САМОСТОЯТЕЛЬНОЕ ПРОЖИВАНИЕ ИНФОРМАНТА и др., с которыми аналитик связал фрагменты текстового свидетельства информанта⁴².

Теперь задачей автора концептуализации становится указание всех понятийных переходов от исходной пары терминов РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА в контексте СВИДЕТЕЛЬСТВО ИНФОРМАНТА (СЮЖЕТ, КОНТЕКСТ) к терминам-цитатам. Решение этой задачи может быть условно представлено в виде следующих взаимосвязанных процедур. Во-первых, указание социологом областей сложившихся представлений (смысловых доменов), относительно которых он в ходе концептуализации определяет как свои термины, так и суждения информанта. Во-вторых, предъявление социологом всех изменений терминов, которые он считает необходимым выполнить для цели обоснования своих суждений. Результаты выполнения аналитиком этих процедур накапливаются в тезаурусе. В-третьих, генерация единых структур: а) иерархии контекстов, т. е. соподчинения указанных аналитиком смысловых доменов и б) терминологического графа, т. е. перехода от собственных терминов аналитика к свидетельству информанта. Генерация выполняется пакетом по запросу социолога на основе текущего состояния тезауруса.

Ориентиром в решении аналитиком задачи перехода с помощью контекстно-ориентированных методов является традиционная для

⁴¹ Социологом или аналитиком, в данном случае автором статьи.

⁴² Опустим технические подробности операций связывания, осуществляемых с помощью пакета.

АКД идея покрытия «содержания» свидетельства информанта с помощью вновь вводимых обозначений⁴³. Однако при таком покрытии чрезвычайно важно обеспечить не только полную представительность мнения информанта в аналитических выкладках социолога, но и воспроизвести скрытую структуру «фактологических» суждений с помощью явных отношений между терминами концептуализации⁴⁴.

Контекстно-ориентированные методы позволяют аналитику структурировать свои теоретические выкладки с помощью аппарата, который является аналогом отношений множественного наследования (multiple inheritance) и видимости (visibility) [переменных], на которых основываются структурные связи между аналитическими обозначениями в языках высокого уровня⁴⁵. Подобные методы доказали свою эффективность в бесчисленных приложениях, благодаря которым создано современное информационное общество.

По мнению автора, подобный аппарат отсутствует в современных компьютерных методах анализа нечисловой информации. Сравнивая же предлагаемые решения с аналогами, реализованными, например, в языках программирования высокого уровня (C++, Object-Oriented Pascal и др.), можно отметить относительную простоту контекстно-ориентированного синтаксиса, с помощью которого вводятся отношения наследования и видимости.

Поясним первую процедуру контекстно-ориентированной концептуализации — указание областей сложившихся представлений. Например, молодой человек что-то рассказывает о своих переживаниях в условиях совместного проживания с родителями. Прежде чем анализировать его рассказ, следует обозначить условия, при которых этот рассказ разворачивается и концептуализируется. Таким образом в составе тезауруса появляется термин **СОВМЕСТНОЕ С РОДИТЕЛЯМИ ПРОЖИВАНИЕ**, который как раз и обозначает обстоятельства суждений информанта по поводу совместного проживания. Аналогично, на основе собственного прочтения

⁴³ Традиционно это задача решается за счет первичного (открытого) кодирования.

⁴⁴ В существующих методах АКД эта задача решается за счет концептуальных действий, «настроенных» над процедурой первичного кодирования — осевого кодирования, введения категорий, аннотирования, естественных языковых интерпретаций и др. При этом в случае применения компьютерного пакета широко используются «общейнформационные» методы — поиска, построения запросов, упорядочивания, визуализации и т. д.

⁴⁵ URL: http://en.wikipedia.org/wiki/Multiple_inheritance

свидетельства автор ввел термины САМОСТОЯТЕЛЬНОЕ ПРОЖИВАНИЕ ИНФОРМАНТА, ДЕЕСПОСОБНОСТЬ РОДИТЕЛЕЙ, МНОГО ЧИТАЮ и др. Эти понятия маркируют зоны смыслов, относительно которых определяются все другие термины, используемые при концептуализации.

Интуитивно понятно, что смысловые домены имеют некоторое соподчинение. Скажем, характеризуя ситуации, возникающие при совместном проживании с родителями, молодой человек может обращаться к общеупотребительной терминологии, используемой повсеместно и не требующей пояснений. К примеру, оценивающие высказывания «хорошо», «плохо» и др.⁴⁶

В то же время обратное следует признать неверным: то, что понимаемо в рамках конкретного свидетельства информанта о своем проживании с родителями, теряет определенность вне этого свидетельства. Скажем, упоминание о спорах с отцом представляется информативным сообщением, характеризующим условия совместного проживания, и оказывается малозначащим вне такого контекста. Потеря информативности объясняется естественным образом. Словосочетание «споры с отцом» продолжает оставаться узнаваемым и в «общих обстоятельствах», но теряются многие связи, важные для понимания анализируемых рассуждений: кто рассказывает, чей отец, какие условия проживания и многие другие коннотации, которые существуют в конкретных ситуациях рассказа молодого человека.

Аналитик выражает подобные интуитивно обоснованные соподчинения в виде контекстного обобщения пары терминов, формируя это отношение по правилам, описанным выше. Вся совокупность созданных контекстных обобщений наряду с терминами, которые их составляют, накапливается в тезаурусе. Тезаурус, созданный в результате описываемой концептуализации, показан на рис. 5. Единая иерархия контекстов, выражающая соподчинение смысловых доменов, формируется с помощью алгоритма (см. выше). Иерархия контекстов, построенная для данного случая, показана на рис. 6.

Наличие иерархии контекстов, пусть даже подлежащей дальнейшим изменениям аналитик, позволяет социологу приступить ко второй аналитической процедуре — выстраиванию совокупности

⁴⁶ В данном случае такого сорта терминология отнесена автором к контексту, обозначенному ОБЩЕПРИНЯТЫЕ ПРЕДСТАВЛЕНИЯ.

0:ТЕСТОВЫЙ. 0

Тезаурус Понятие Идентификация

Мнемоника

БЫТЬ САМОСТОЯТЕЛЬНЫМ В РЕШЕНИЯХ

ДЕеспОСОбНОСТЬ РОДИТЕЛЕЙ

ДЕЙСТВИЯ, НАРУШАЮЩИЕ ТИШИНУ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

ЗВУКИ В ДОМЕ

ЛУЧШЕ ИЗБЕГАТЬ СПОРЫ

МНОГО ЧИТАЮ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

МНОГО ЧИТАЮ: АНАЛИТИЧЕСКИЙ КОММЕНТАРИЙ

МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

НЕ С КЕМ ПОГОВОРИТЬ

НЕТ ШЕВЕЛЕНИЙ

НИКЧЕМНО

НУЖДАТЬСЯ В ПОМОЩИ РОДИТЕЛЕЙ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

ОБЩЕПРИНЯТЫЕ ПРЕДСТАВЛЕНИЯ

ОГРАНИЧИВАТЬ САМОСТОЯТЕЛЬНОСТЬ РЕШЕНИЙ

ОТДОХНУТЬ

ОТКЛИКАТЬСЯ НА ПРОСЬБЫ СЫНА//РОДИТЕЛИ

ОТСУТСТВИЕ ЛЮДЕЙ РЯДОМ

ОТСУТСТВИЕ "УКАЗЫВАЮЩИХ" ВОЗДЕЙСТВИЙ

ОЦЕНКА ПЛОХО

ОЦЕНКА СИТУАЦИИ

ОЦЕНКА СПОРОВ С ОТЦОМ

ОЦЕНКА ТЕЛЕПЕРЕДАЧ

ОЦЕНКА ХОРОШО

ПЕРЕЖИВАНИЯ ИНФОРМАНТА

ПОДВЕРГАТЬ ОПАСНОСТИ

ПОЛАГАТЬСЯ НА РОДИТЕЛЕЙ

ПРЕДМЕТНАЯ ОБЛАСТЬ АКД

ПРИНИМАТЬ СОБСТВЕННЫЕ РЕШЕНИЯ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

ПРОСМОТР ТЕЛЕПЕРЕДАЧ

РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА

РАЗЪЕХАТЬСЯ С РОДИТЕЛЯМИ

РЕШЕНИЕ

РОДИТЕЛИ

РОДИТЕЛИ СПОСОБНЫ ОТКЛИКНУТЬСЯ

САМОСТОЯТЕЛЬНОЕ ПРОЖИВАНИЕ ИНФОРМАНТА

СВИДЕТЕЛЬСТВО ИНФОРМАНТА(СЮЖЕТ,КОНТЕКСТ)

СЛУШАТЬ МУЗЫКУ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

СМОТРЕТЬ ТЕЛЕВИЗОР//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

СОВМЕСТНОЕ С РОДИТЕЛЯМИ ПРОЖИВАНИЕ

СОХРАНИТЬ САМОСТОЯТЕЛЬНОСТЬ

СПОРИТЬ С ОТЦОМ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

СТАБИЛЬНОЕ ТЕЧЕНИЕ ЖИЗНИ

СЪЕХАТЬ ОТ РОДИТЕЛЕЙ

ТЕРЯТЬ НАВЫКИ САМОСТОЯТЕЛЬНОСТИ

ТИШИНА ВОКРУГ

УЙТИ ИЗ ДОМА

УСТАЛОСТЬ

ЧЕМ БЫ СЕБЯ ЗАНЯТЬ

ЧУВСТВО ОДИНОЧЕСТВА

ЧУВСТВО ОДИНОЧЕСТВА: АНАЛИТИЧЕСКИЙ КОММЕНТАРИЙ

ЧУВСТВО ОДИНОЧЕСТВА: ФРАГМЕНТ ТЕКСТА

6. ДЕеспОСОбНОСТЬ РОДИТЕЛЕЙ

НУЖДАТЬСЯ В ПОМОЩИ РОДИТЕЛЕЙ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

ОТКЛИКАТЬСЯ НА ПРОСЬБЫ СЫНА//РОДИТЕЛИ

.0 СВИДЕТЕЛЬСТВО ИНФОРМАНТА(СЮЖЕТ,КОНТЕКСТ)

ТИШИНА ВОКРУГ

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

СМОТРЕТЬ ТЕЛЕВИЗОР/МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

МНОГО ЧИТАЮ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

СПОРИТЬ С ОТЦОМ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

1. РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА

СВИДЕТЕЛЬСТВО ИНФОРМАНТА(СЮЖЕТ,КОНТЕКСТ)

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

АНАЛИТИЧЕСКИЙ КОММЕНТАРИЙ

ПЕРЕЖИВАНИЯ ИНФОРМАНТА

УСТАЛОСТЬ

РЕШЕНИЕ

2. СМОТРЕТЬ ТЕЛЕВИЗОР

МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

ТИШИНА ВОКРУГ

СВИДЕТЕЛЬСТВО ИНФОРМАНТА(СЮЖЕТ,КОНТЕКСТ)

ЗВУКИ В ДОМЕ

ОЦЕНКА ТЕЛЕПЕРЕДАЧ

ПРОСМОТР ТЕЛЕПЕРЕДАЧ

ЧЕМ БЫ СЕБЯ ЗАНЯТЬ

3. УСТАЛОСТЬ

МНОГО ЧИТАЮ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

МНОГО ЧИТАЮ: АНАЛИТИЧЕСКИЙ КОММЕНТАРИЙ

ОТДОХНУТЬ

4. ПЕРЕЖИВАНИЯ ИНФОРМАНТА

САМОСТОЯТЕЛЬНОЕ ПРОЖИВАНИЕ ИНФОРМАНТА

КОММЕНТАРИЙ(ТЕКСТ,МЕМО)

АНАЛИТИЧЕСКИЙ КОММЕНТАРИЙ

ЧУВСТВО ОДИНОЧЕСТВА

ОТСУТСТВИЕ ЛЮДЕЙ РЯДОМ

НЕ С КЕМ ПОГОВОРИТЬ

СПОРИТЬ С ОТЦОМ//МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

ПОЛАГАТЬСЯ НА РОДИТЕЛЕЙ

5. СПОРИТЬ С ОТЦОМ

МОЛОДОЙ ЧЕЛОВЕК (ИНФОРМАНТ)

СОВМЕСТНОЕ С РОДИТЕЛЯМИ ПРОЖИВАНИЕ

ОЦЕНКА ПЛОХО

БЫТЬ САМОСТОЯТЕЛЬНЫМ В РЕШЕНИЯХ

НЕОПРЕДЕЛЕННЫЙ ТЕРМИН

ДЕеспОСОбНОСТЬ РОДИТЕЛЕЙ

ОЦЕНКА ПЛОХО

Рис. 5. Визуализация тезауруса: слева (темный фон) — совокупность терминов, справа и внизу (светлый фон) — структурно разрозненные отношения между терминами в виде контекстно-фиксированных разъяснений

изменений терминов, которые приводят от РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА в контексте СВИДЕТЕЛЬСТВО ИНФОРМАНТА (СЮЖЕТ, КОНТЕКСТ) к финальным терминам концептуализации. Каждое из таких понятийных изменений можно представить как ответ на вопрос «что такое?». Например, в качестве ответа на вопрос, «что такое РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА?» автор нашел в тексте ряд суждений, которые обозначил терминами ПЕРЕЖИВАНИЯ ИНФОРМАНТА в контексте САМОСТОЯТЕЛЬНОЕ ПРОЖИВАНИЕ ИНФОРМАНТА, ПЕРЕЖИВАНИЯ ИНФОРМАНТА в контексте СПОРИТЬ С ОТЦОМ, УСТАЛОСТЬ в контексте МНОГО ЧИТАЮ и др.

Далее, каждый из терминов, возникший в качестве ответа на предыдущий вопрос «что такое?», сам подвергается разъяснению. В соответствие с целью концептуализации этот процесс продолжается аналитиком до тех пор, пока ему не удастся связать свои исходные представления РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА в контексте СВИДЕТЕЛЬСТВО ИНФОРМАНТА (СЮЖЕТ, КОНТЕКСТ) со всеми терминами-цитатами, которые он сам же создал на основе текста информанта. Инструментально «пошаговые» ответы на вопросы «что такое?» создаются аналитиком с помощью КФР, в котором целевой термин является подлежащим, к которому обращен вопрос, а группа пояснений представляет собой ответ.

Пошаговое создание ответов на вопрос «что такое?» с помощью контекстно-ориентированных структур в точности повторяло бы построение традиционной иерархии кодов, широко используемой в классических функциях кодирования данных [3] и представляющей собой граф типа дерево, если бы не контекстная определенность каждого термина. Особенность состоит в том, что нельзя разъяснить термин, определенный в «более общем»⁴⁷ контексте через понятие, введенное в «более низком» контексте. Например, если ОЦЕНКА ПЛОХО введена в контексте ОБЩЕПРИНЯТЫЕ ПРЕДСТАВЛЕНИЯ, то ее нельзя разъяснить через термин НЕ С КЕМ ПОГОВОРИТЬ, если последний введен в «более низком» контексте САМОСТОЯТЕЛЬНОЕ ПРОЖИВАНИЕ ИНФОРМАНТА (см. рис. 6).

В практической концептуализации одно и то же КФР используется как для разбивки на смысловые зоны, так и для ответа на вопрос

⁴⁷ В смысле операции контекстного обобщения.

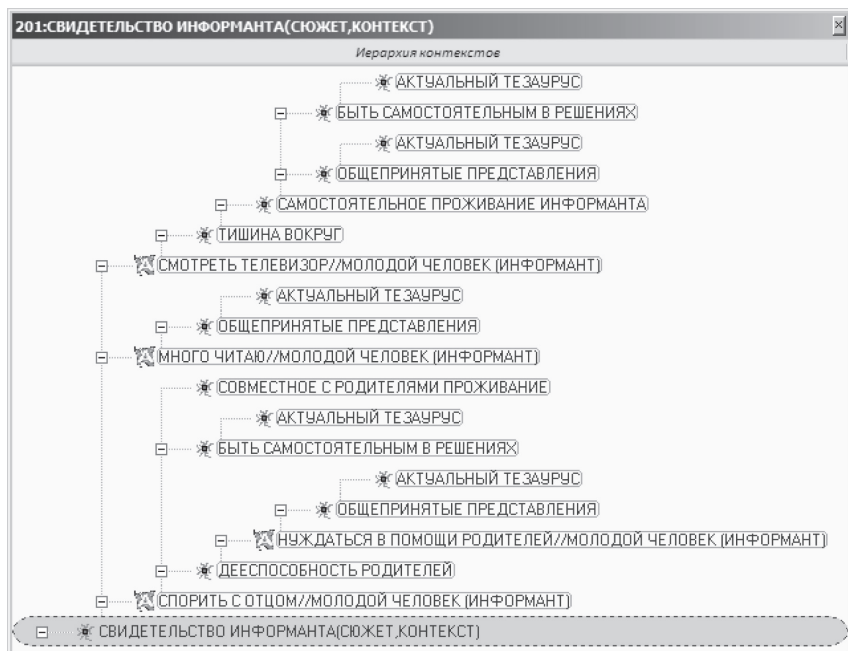


Рис. 6. Иерархия контекстов: узлами являются термины, использованные в КФР как контексты

«что такое?». В любой момент концептуализации, т. е. при любом составе формируемого тезауруса, аналитик имеет возможность контролировать, за счет каких понятий и каких глобальных связей между ними осуществляется (или не осуществляется) глобальный переход от его терминов к свидетельству информанта. Инструментами контроля выступают компьютерные средства генерации иерархии контекстов и терминологического графа.

Результатом контекстно-ориентированной концептуализации является терминологический граф, который дает единое аналитическое представление свидетельств информанта с учетом областей смыслов, в которых эти свидетельства понимаются информантом с точки зрения аналитика. Соответствующий терминологический граф, демонстрирующий концептуальные связи в анализируемом текстовом свидетельстве показан на рис. 7. Узлом графа служит пара



Рис. 7. Терминологический граф, выражающий концептуализацию понятия РАЗМЫШЛЕНИЯ И ПЕРЕЖИВАНИЯ ИНФОРМАНТА при обстоятельствах, обозначенных термином СВИДЕТЕЛЬСТВО ИНФОРМАНТА (СЮЖЕТ, КОНТЕКСТ) согласно тезаурусу, показанному на рис. 5.

терминов. Первый — соответствующее пояснение в одном из КФР тезауруса, второй⁴⁸ — понятие, которое использовано как контекст для этого пояснения.

⁴⁸ Указан через двойной знак умножения «**»

Обсуждение

Прикладное значение выполненной работы состоит в том, что разработанный контекстно-ориентированный аппарат позволяет усовершенствовать информационно-аналитические методы взаимодействия акторов в тех предметных областях, где главным средством прикладной коммуникации до сих пор остается естественный язык. Примерами таких предметных областей являются, например, многообразные задачи социального управления — составление планов развития, организация документооборота, разработка программ социальных действий и др.

Преимущество разработанных методов по сравнению с традиционным вербальным способом взаимодействия социальных акторов состоит в том, что все получаемые с помощью контекстно-ориентированных методов «текстовые документы» (планы, программы, социальные теории и т. п.) могут быть проверены на связность терминологии, использованной для их формирования. Вместе с тем предлагаемые методы основаны на работе пользователя с текстовыми обозначениями, которые понимаются человеком исключительно на основе собственного опыта и навыков владения естественным языком. Работа с такого сорта вербальными обозначениями не требует от аналитика, обращающегося к контекстно-ориентированному аппарату, никаких специальных знаний, а только практического соблюдения закона тождества вводимых им обозначений в разнообразных ситуациях их использования, отслеживаемых самим автором обозначений.

Образцами таких специальных знаний являются методы объектно-ориентированного программирования, включающего в своей состав формальный синтаксис языка, аппарат типобразования переменных и прочие приемы. Эти знания требуют трудоемкого обучения, что снижает эффективность применения такого аппарата в прикладных задачах, предполагающих вовлечение в процесс концептуализации предметной области большого числа людей. Например, в случае разработки планов социального развития [Города, территории и т. п.], учитывающих мнения широкой массы населения.

Несмотря на текстовую форму вводимых обозначений, предложенные методы дают в распоряжение пользователей изощренный аналитический аппарат, аналогичный структурным отношениям объектно-ориентированного программирования (наследование, области видимости, инкапсуляция и т. п.). Успехи современных информационных

технологий в организации совместной работы больших коллективов разработчиков интеллектуально емких проектов дают основание надеяться, что контекстно-ориентированные методы, обладающие столь же мощным аналитическим аппаратом, но не требующие от их пользователей специальных знаний, окажутся эффективным средством организации взаимодействия акторов, занятых решением прикладных социальных задач.

Теоретическое значение работы состоит в том, что она развивает научные принципы моделирования социальных явлений в тех случаях, когда позитивистские методы не применимы, а вербальные разъяснения неэффективны. Смысл сделанных теоретических предложений состоит в разработке научного аппарата, который, с одной стороны, позволяет исследователю оперировать только текстовыми описаниями (не вводя никаких мер или операциональных свойств). С другой — алгоритмически контролирует связность текстов, вводимых аналитиком.

Для пояснения полученных результатов воспользуемся аналогией с представлением о координатном (декартовом) пространстве, которое часто в области социологических дисциплин называется признаковым. Разработанный аппарат развивает не традиции вычислений, т. е. это — не «факторный анализ для качественных данных», а специфический аналог декартова пространства, который предлагается в качестве конструктивной основы всех последующих аналитических процедур. В частности, это могут быть вычислительные действия на количественных данных, определенных в результате процесса контекстно-фиксированных разъяснений. Однако в числе таких аналитических процедур могут быть действия, которые не требуют еще до появления таких количественных возможностей предложенный аппарат для, поддающихся проверке на принципах науки.

Предложенные решения названы «методом», потому что они воспроизводят основы научной методологии (оперирование понятиями, связность утверждений и т. п.), в тех ситуациях, когда измерение еще невозможно, а вербальные утверждения уже не могут быть сугубо индивидуальным мнением. Декартов метод — это процедура декомпозиции общей картины явления (вне зависимости от его природы) на отдельные составляющие. Можно сказать, что контекстно-ориентированный АКД представляет собой расширение декартова метода для случаев, когда «координаты» явления не имеют меры и формируются в процессе коммуникации акторов.

Заключение

В статье представлен новый аналитический аппарат качественного социологического исследования, основанный на текстовых описаниях изучаемых явлений и включающий в свой состав методы контроля связности исследовательской терминологии. Современным способом дальнейшей разработки предложенных методов является создание информационных технологий построения социального знания. Одной из важнейших черт такого знания является коллективный способ его создания, включающий в себя взаимный аналитический контроль утверждений участников этого процесса. В качестве таких утверждений могут выступать заявления политических лидеров, программы социального развития региона, планы совершенствования системы социального обеспечения и другие «социальные сообщения», формулируемые сегодня в текстовом виде, затрудняющем контроль их состоятельности со стороны общества.