

УДК 519.711.3

СТРУКТУРА СИСТЕМЫ ИЗМЕРЯЕМЫХ СВОЙСТВ У ОБЪЕКТОВ МНОГОМЕРНОГО НАБЛЮДЕНИЯ И ЭКСПЕРИМЕНТОВ

Д.Е. Самойлов^{1,2}, В.А. Семенова^{2,3}, С.В. Смирнов²

¹ Самарский национальный исследовательский университет им. академика С.П. Королева
Россия, 443086, г. Самара, Московское шоссе, 34А

² Институт проблем управления сложными системами РАН
Россия, 443020, г. Самара, ул. Садовая, 61

³ Самарская государственная областная академия (Наяновой)
Россия, 443001, г. Самара, ул. Молодогвардейская, 196

Аннотация. Областью исследований, которой посвящена предлагаемая статья, является онтологический анализ данных. Это комплекс моделей и методов, предназначенный для извлечения из исходного эмпирического материала формальной онтологии исследуемой предметной области – описывающей предметную область совокупности понятий и отношений. Исходному эмпирическому материалу, т. е. данным наблюдений и экспериментов, как правило, свойственна неполнота и противоречивость, обусловленная реалиями накопления эмпирической информации. Это приводит к тому, что необходимый для построения онтологии формальный контекст задачи онтологического анализа данных – соответствие «объекты-свойства» – предварительно может быть представлен лишь в рамках какой-либо многозначной логики. Его требуется аппроксимировать в обычной бинарной логике, поскольку эффективные методы вывода формальных понятий разработаны лишь для однозначных (бинарных) формальных контекстов. Для корректного решения этой задачи в общем случае необходимо знать ограничения существования свойств у объектов исследуемой предметной области. Откуда же априори могут быть известны столь тонкие сведения об еще лишь исследуемой предметной области? Основная идея авторов статьи состоит в том, что эти знания представляют собой продукт априорного гипотезирования субъекта исследования. Выбирая для измерения у объектов многомерного наблюдения и экспериментов определенные свойства, субъект фактически формулирует гипотетические понятия об исследуемой предметной области, а используя фундаментальную когнитивную процедуру – концептуальное шкалирование измеряемых свойств – эти гипотетические представления развивает. При таком формировании системы измеряемых свойств фиксируются вполне определенные ограничения существования свойств, которые определяют структуру этой системы. Апостериори подтверждается лишь часть гипотез субъекта, и, следовательно, верными оказываются некоторые ограничения существования свойств. В статье выявляются и анализируются фундаментальные паттерны структурной организации системы измеряемых свойств, и тем самым предлагается исчерпывающая модель ограничений существования свойств в онтологическом анализе данных.

Ключевые слова: система измеряемых свойств, ограничения существования свойств, паттерн структурной организации, онтологический анализ данных, формальный контекст.

Дмитрий Евгеньевич Самойлов, студент, лаборант ИПУСС РАН.

Валентина Андреевна Семенова, аспирант Самарской государственной областной академии (Наяновой), м.н.с. ИПУСС РАН.

Сергей Владимирович Смирнов (д.т.н., доц.), главный научный сотрудник.

Введение

Изучение субъектом ранее неизвестной ему предметной области (ПрО), синтез проектировщиком инновационного продукта, конструирование маркетологом новой потребности и т. п. связаны с *наделением* всякого реального или ментального объекта в *terra incognita* определенными *свойствами* (см., например, [1–4]). Ключевыми здесь оказываются два акта: *гранулирование* действительности, т. е. выделение объектов, отделение объектов друг от друга (сочтем это имманентной способностью нашего сознания и далее рассматривать не будем), и *измерение* у объектов некоторого ограниченного набора характеристик (отсюда «*наделение*» гранул-объектов свойствами; заметим, что измерять можно и связи, *отношения* между объектами). Таким образом, *первичной* картиной *terra incognita* оказывается соответствие «*объекты – свойства*» – либо протокол наблюдений и/или экспериментов, либо тактико-технические характеристики необходимой искусственной системы.

Одним из методов дальнейшего «проникновения» в *terra incognita* является *онтологический анализ данных* (ОАД) – вывод *формальной онтологии*, интересующей субъекта ПрО исходя из соответствия «объекты – свойства» [5–7]. Кроме формальной онтологии – описательной модели ПрО – к ОАД можно отнести и извлечение из данных *множества импликаций на признаках объектов* – предсказательную модель ПрО, более известную как целевой результат задачи *поиска ассоциативных правил* [8–10].

Важнейшим этапом ОАД является формирование *однозначного* соответствия «объекты – свойства» – *формального контекста* (ФК) ПрО. Речь идет о кортеже $K = (G^*, M, I)$, где $G^* = \{g_i\}_{i=1, \dots, r}$, $r = |G^*| \geq 1$ – множество наблюдавшихся объектов; $G^* \subseteq G$, G – все гипотетически мыслимое множество объектов исследуемой ПрО; $M = \{m_j\}_{j=1, \dots, s}$, $s = |M| \geq 1$ – априори заданный набор измеряемых у объектов свойств; $|M| \geq 1$; I – бинарное соответствие «объекты – свойства», т. е. совокупность оценок истинности $\|b_{xy}\| \in \{\text{Истина}, \text{Ложь}\}$ базовых семантических суждений (БСС) об изучаемой ПрО вида b_{xy} = «объект x обладает свойством y ».

Формирование ФК в общем случае осложнено *неполнотой и противоречивостью* эмпирической информации о ПрО. Это приводит к «размытости» исходного соответствия «объекты – свойства» [7, 11]. Построение *корректного однозначного соответствия* I возможно лишь при учете экзистенциальных зависимостей между измеряемыми свойствами объектов, или *ограничений существования свойств* (ОСС) [7, 12, 13]:

– *обусловленности* $C: M \times M \rightarrow \{\text{Истина}, \text{Ложь}\}$, когда, обладая свойством m_j , объект g_i непременно обладает свойством m_k (хотя обратное может быть неверно), т. е. $C(m_j, m_k) = \text{Истина} \leftrightarrow \forall g_i \in G^*: m_j \in \{g_i\}' \rightarrow m_k \in \{g_i\}'$ (здесь и далее «'» есть оператор соответствия Галуа между множествами G^* и M). Отношение C рефлексивно, несимметрично и транзитивно, т. е. $\forall a, b, c \in M: C(a, b) \& C(b, c) \rightarrow C(a, c)$ – рис. 1;

– *несовместимости* $E: M \times M \rightarrow \{\text{Истина}, \text{Ложь}\}$, когда, обладая свойством m_j , объект g_i заведомо не обладает свойством m_k , и наоборот, т. е. $E(m_j, m_k) = \text{Истина} \leftrightarrow \forall g_i \in G^*: m_j \in \{g_i\}' \rightarrow m_k \notin \{g_i\}'$. Отношение E антирефлексивно, симметрично и нетранзитивно, но характеризуется так называемой *транзитивностью относительно обусловленности*, т. е. верна импликация $\forall d, e, f \in M: C(d, e) \& E(e, f) \rightarrow E(d, f)$ – см. рис. 1.

Таким образом, ОАД предполагает *априорное* построение исследователем ПрО *системы измеряемых свойств* (СИС) – множества измеряемых свойств

с заданными на нем бинарными отношениями несовместимости и обусловленности, или (M, E, C) -кортежа.

В общей постановке вопросы подобной *структуризации* СИС рассматривались в [14, 15], где установлен ее многоуровневый рекурсивный характер, обусловленный возможностью неоднократно раз за разом повторять процедуру *дизъюнктивного* (номинального) и *уточняющего* (порядкового) *концептуального* *шкалирования* [14–16] измеряемых свойств, расширяя и усложняя СИС. В завершенном виде *стратегия* априорного формирования СИС обоснована в [15] путем реконструкции *логико-понятийного* *генезиса* ОСС.

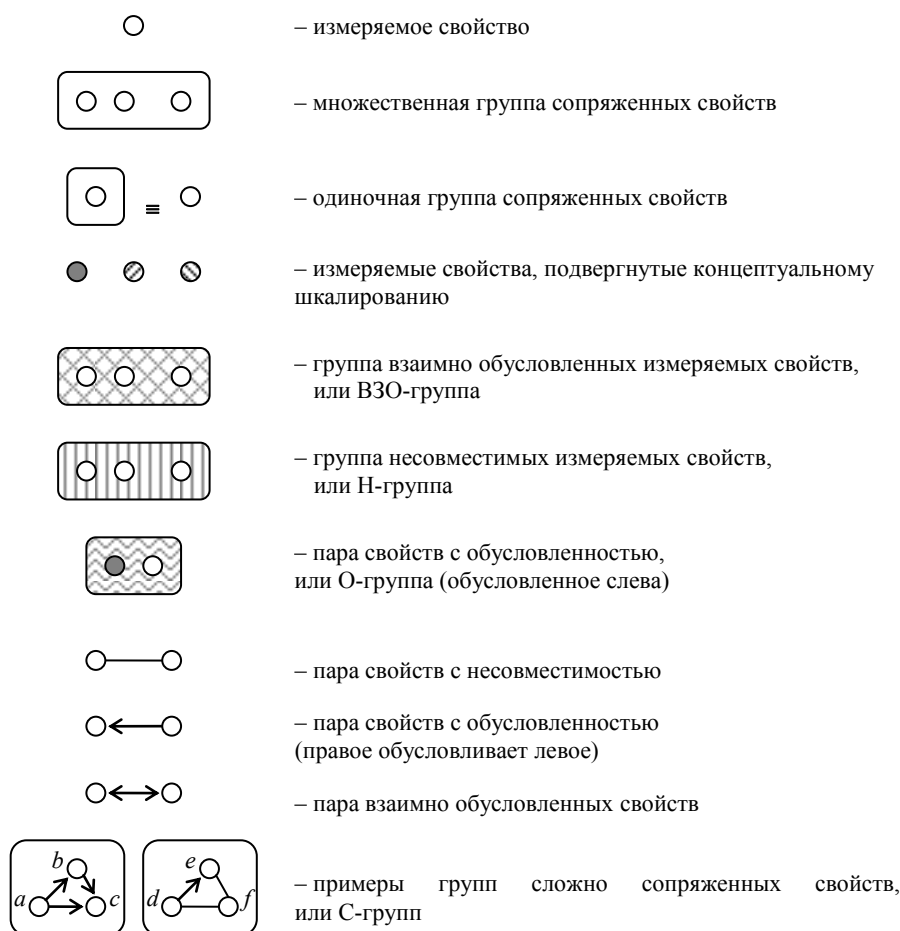


Рис. 1. Основные сущности задачи формирования структуры системы измеряемых свойств и их обозначения (знак « $\equiv$ » здесь и далее на рисунках используется для указания эквивалентности изображений структурных элементов)

Данная статья посвящена *тактике* формирования СИС, т. е. выявлению всех возможных *шаблонов* выполнения рекурсивной процедуры пополнения множества измеряемых свойств и получаемых при этом результатов – структурных *паттернов* *сопряжения* *измеряемых* *свойств* (ПСС).

1. Группа сопряженных свойств

Рис. 1 вводит графическую нотацию сущностей, необходимых для описания структуры СИС. Среди них стержневым концептом является «*группа сопряженных свойств*» (ГСС):

- измеряемое свойство составляет *одинокую* ГСС, единственный член которой самообусловлен. В качестве графического образа одиночной ГСС может использоваться собственно значок свойства, но в любом случае «петлю самообусловленности» на диаграммах будем только подразумевать;

- ГСС образуют все *отличительные свойства* [1] *гипотетического формального понятия*, вводимого субъектом при формировании СИС (подробно это априорное действие субъекта рассматривается в [15]). Члены такой ГСС *взаимно обусловлены* (ВЗО-группа);

- множество измеряемых свойств, вводимых в *M* в результате *дизъюнктивного шкалирования* уже имеющегося в *M* измеряемого свойства, определяет ГСС, члены которой несовместимы (Н-группа);

- измеряемое свойство, подвергнутое *уточняющему шкалированию*, вместе с измеряемым свойством, вводимым в *M* в результате этого шкалирования, образует ГСС, в которой второе из указанных свойств обуславливает первое (О-группа);

- в сложно сопряженной группе (С-группе) зависимости между измеряемыми свойствами могут быть неоднородными и экзистенциальные отношения между свойствами представлены парами соответствующих связей.

Нетрудно видеть, что такое определение ГСС устанавливает частично упорядоченное отношение *вложенности/замещения* на множестве ГСС. В частности:

- ГСС, включающая *протосвойство* – свойство, измерение которого у наблюдаемых объектов производится в соответствии с *исходным гипотетическим понятием* субъекта об исследуемой ПрО [15], – является корнем дерева вложенности (рис. 2, а);

- ГСС, включающая множество измеряемых свойств, вводимых в *M* в результате *дизъюнктивного шкалирования*, вложена в ГСС, содержащую подвергнутое шкалированию свойство, и его *замещает* (рис. 2, б);

- ГСС, включающая измеряемое свойство, подвергнутое *уточняющему шкалированию*, вместе с измеряемым свойством, вводимым в *M* в результате этого шкалирования, вложена в ГСС, содержащую подвергнутое шкалированию свойство, и его замещает (рис. 2, в);

- ГСС, включающая *множество отличительных свойств* гипотетического формального понятия вместо *одного* такого свойства¹, вложена в ГСС, которая содержит замещаемое отличительное свойство (рис. 2, г);

- ситуация обратная к предыдущей: ГСС, включающая *некоторую часть множества отличительных свойств* гипотетического формального понятия, вложена в ГСС всех этих отличительных свойств и замещает в ней указанную часть свойств (рис. 2, д).

¹ С логической точки зрения количество отличительных свойств гипотетического понятия не влияет ни на гипотетическое понимание ПрО в целом, ни на положение в этой гипотетической картине данного понятия в частности. Количество характеризует лишь «*широту*» описания объектов ПрО из объема рассматриваемого гипотетического понятия. Поэтому отличающее свойство *всегда* может быть замещено набором таких свойств.

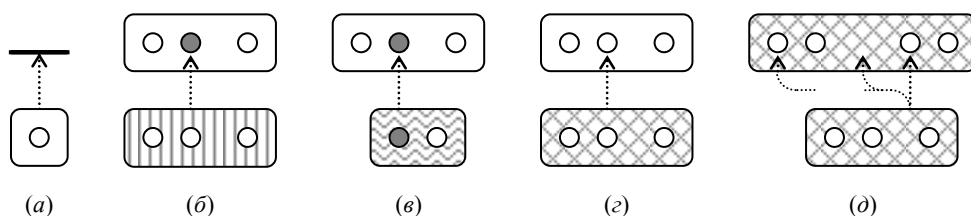


Рис. 2. Вложенность/замещение групп сопряженных свойств

Термин «замещение» применительно к ГСС подчеркивает, что на множестве ГСС, как и на множестве измеряемых свойств, заданы отношения несовместимости и обусловленности.

2. Возникновение элементарных групп сопряженных свойств

2.1. Шкалирование протосвойства

Дизъюнктивное шкалирование. ПСС дизъюнктивного шкалирования протосвойства недвусмысленно следует из определения операции *деления понятия* (фактически это определение интерпретирует деление понятия как результат такого шкалирования) [17, 18]. Речь идет об *исходном гипотетическом понятии* о ПрО, которое выдвигается субъектом при решении измерять протосвойство и содержание которого исчерпывается протосвойством.

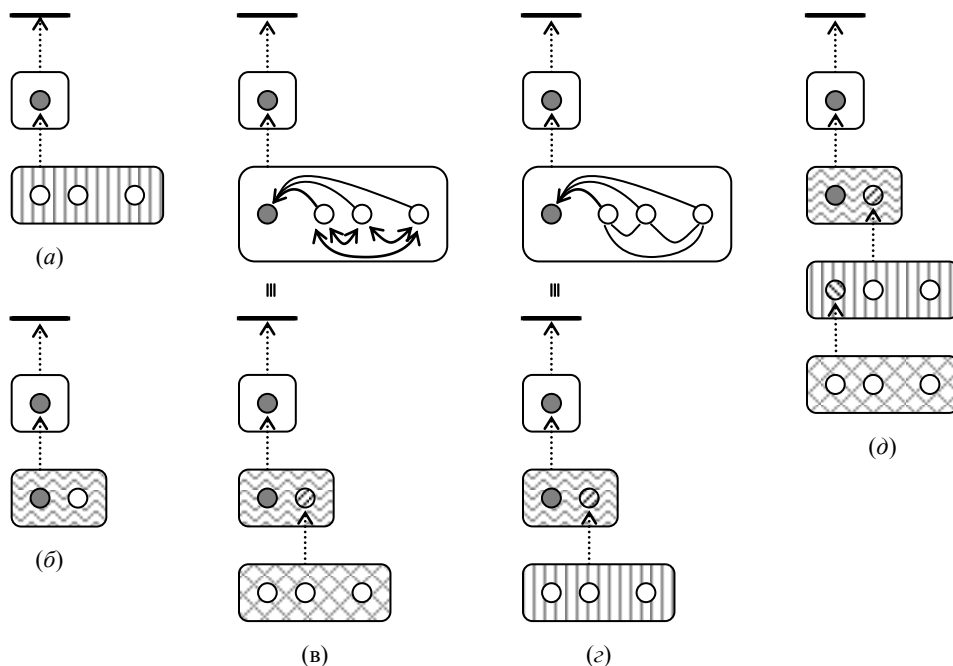


Рис. 3. Структурные паттерны сопряжения измеряемых свойств при шкалировании протосвойства

На рис. 3, *а* в одиночной ГСС протосвойство *замещается* группой из $n \geq 2$ новых несовместимых измеряемых свойств. При этом исходное гипотетическое понятие о ПрО замещается n гипотетическими понятиями, каждое из которых в качестве содержания получает одно из новых измеряемых свойств, замещающих протосвойство.

Уточняющее шкалирование. В данном случае конфигурация ПСС непосредственно следует из определения операции *ограничения понятия* [17, 18], содержание которого составляет протосвойство, и иллюстрируется рис. 3, *б*.

ПСС на рис. 3, *б* соответствует простейшему случаю, когда ограничение понятия производится с помощью добавления *одного* уточняющего свойства – нового измеряемого свойства, которое становится *единственным отличительным* свойством нового формального понятия (подпонятия исходного гипотетического понятия о ПрО). Конечно, в общем случае образуемое подпонятие может быть охарактеризовано $m > 1$ отличительными свойствами, которые, естественно, будут *взаимно обусловлены*. Тогда ПСС, соответствующий такой ситуации, приобретает *два* уровня, как показано на нижней части рис. 3, *в*.

Еще один результирующий ПСС уточняющего шкалирования протосвойства связан с тем, что исходное понятие может ограничиваться с образованием $n \geq 2$ подпонятий. Необходимые для этого n новых уточняющих измеряемых свойств будут *несовместимыми*. Поэтому третий ПСС также имеет два уровня (см. нижнюю часть рис. 3, *г*).

Наконец, учитывая возможность композиции рассмотренных выше случаев, ПСС при уточняющем шкалировании протосвойства приобретает трехуровневый вид, что иллюстрирует рис. 3, *д*.

2.2. Шкалирование безусловленного измеряемого свойства

Необусловленные измеряемые свойства образуются в СИС, например, при дизъюнктивном шкалировании протосвойства (рис. 3, *а*) или в результате уточняющего шкалирования такого свойства (рис. 3, *б*).

Дизъюнктивное шкалирование. По определению несовместимости очевидно, что при замещении одного из m свойств Н-группы n новыми измеряемыми свойствами нет необходимости создавать новые ГСС. ПСС фиксирует в данном случае расширение состава существующей Н-группы до $(m - 1) + n$ свойств (рис. 4).

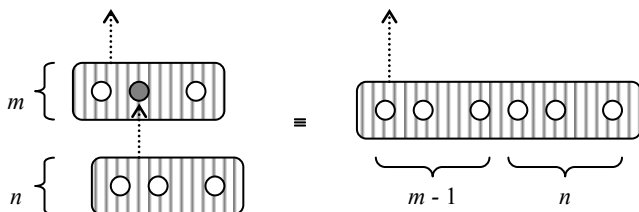


Рис. 4. Структурный паттерн сопряжения измеряемых свойств при дизъюнктивном шкалировании безусловленного измеряемого свойства

Дизъюнктивное шкалирование безусловленного измеряемого свойства, когда оно само является обуславливающим свойством (рис. 3, *б*), фактически уже

рассмотрено. Возникающий ПСС составляет легко идентифицируемую часть структурных паттернов на рис. 3, *з* и 3, *д*.

Уточняющее шкалирование. Необусловленность шкалируемого измеряемого свойства означает отсутствие каких-либо сдерживающих экзистенциальных ограничений для уточняющего шкалирования. Поэтому ПСС, возникающие при шкалировании необусловленного измеряемого свойства, эквивалентны паттернам, возникающим при уточняющем шкалировании протосвойства. Таким образом, эти ПСС составляют легко идентифицируемые части структурных паттернов, представленных на рис. 3, *б* – 3, *д*, где вместо протосвойства фигурирует необусловленное измеряемое свойство из некоторой Н-группы.

2.3. Шкалирование обусловленного измеряемого свойства

Исключим пока из рассмотрения случай, когда шкалируемое свойство *взаимно обусловлено* с другими свойствами. Тогда можно утверждать, что обусловленность измеряемого свойства свидетельствует о том, что оно ранее было подвергнуто уточняющему шкалированию (см. рис. 3, *б* – 3, *д*).

Дизъюнктивное шкалирование. Вообще говоря, в данном случае требуется заместить свойство, обусловленное *m* измеряемыми свойствами, *n* новыми измеряемыми свойствами. Это порождает *комбинаторное число* вариантов сопряжения свойств, уже «уточняющих» делимое гипотетическое понятие, отличительным свойством которого является подвергаемое дизъюнктивному шкалированию измеряемое свойство, и новых свойств, определяющих деление этого понятия. Проблема состоит в конструировании всех таких вариантов с целью выбора единственного, отвечающего совокупности определенных, предпочитаемых субъектом гипотез об изучаемой ПрО. Однако нетрудно понять, что какова бы ни была соответствующая такому предпочтению конфигурация свойств и зависимостей между ними, она может быть получена при иной, легко контролируемой *последовательности* шкалирования рассматриваемого измеряемого свойства, когда сначала выполняется дизъюнктивное шкалирование, а затем – надлежащее уточняющее шкалирование измеряемых свойств, заместивших рассматриваемое свойство.

На этом основании дизъюнктивное шкалирование обусловленного измеряемого свойства можно исключить из числа механизмов формирования СИС.

Уточняющее шкалирование. Естественнo исходить из того, что к моменту такого шкалирования конструирование СИС приходило с использованием рассмотренных ПСС (см. рис. 3, *б* – 3, *д*). Это вполне определяет условия данной задачи шкалирования. Для вводимых при уточняющем шкалировании *n* новых измеряемых свойств существуют лишь следующие альтернативы:

- свойства попадут в состав отличительных свойств уже существующего гипотетического формального понятия и войдут в соответствующую ВЗО-группу (рис. 3, *в*);

- свойства пополнят множество несовместимых обусловливающих свойств и соответствующую Н-группу (рис. 3, *з*), причем с учетом пересечений доменов значений у старых и новых «уточняющих» свойств (и, следовательно, введения *дополнительных новых* «уточняющих» свойств, характеризующих такие пересечения) состав Н-группы возрастет не менее чем на *n* единиц;

- произойдет и первое, и второе, т. е. реализуется соответствующий фрагмент структурного паттерна, представленного на рис. 3, *д*. При этом количе-

ственный состав ВЗО- и Н-групп будет отвечать конкретным гипотетическим представлениям субъекта об изучаемой ПрО.

2.4. Базовые паттерны сопряжения измеряемых свойств

Анализ, проведенной в подразделах 2.1–2.3, показывает, что структура СИС во многих случаях формируется на основе всего лишь четырех базовых ПСС, приведенных на рис. 2, а – 2, г.

Рис. 5 иллюстрирует применение базовых ППС для модельной ситуации – формирования начальной СИС дачника, намерившегося изучить рынок емкостей для поливной воды.

Во-первых, дачник исходит из предположения, что объекты-емкости на рынке имеются (булевское протосвойство «Емкость»). Но фактически на первых порах представление о емкости наш исследователь сводит к ее объему, одновременно понимая, что она изготовлена из определенного материала. Это ведет к замещению протосвойства двумя взаимно обусловленными свойствами «Объем» и «Материал». Об объеме дачник судит дифференцированно: малый, средний, большой – и априори понимает, что на рынке могут существовать емкости соответствующих объемов (свойство «Объем» дизъюнктивно шкалируется). Аналогично оформляется представление о материале: в СИС вводятся свойства «Металл» и «Пластик». Но где-то наш маркетолог слышал, что емкости из пластика могут обвязываться металлическим каркасом. Поэтому свойство «Пластик» надлежащим образом уточняется. Теперь дачник может собирать сведения о емкостях у «коллег», в интернете, в магазинах «Садовод» и т. д., формируя таблицу «емкости – свойства», где последние взяты из сформированной СИС. Априори зафиксированные отношения на множестве «измеряемых» свойств емкостей – структура СИС – должны защитить его от получения в результате онтологического анализа собранных и, скорее всего, неполных и противоречивых данных такого, например, знания: «на рынке имеются металлические емкости в металлическом каркасе, объем которых одновременно и мал, и велик».

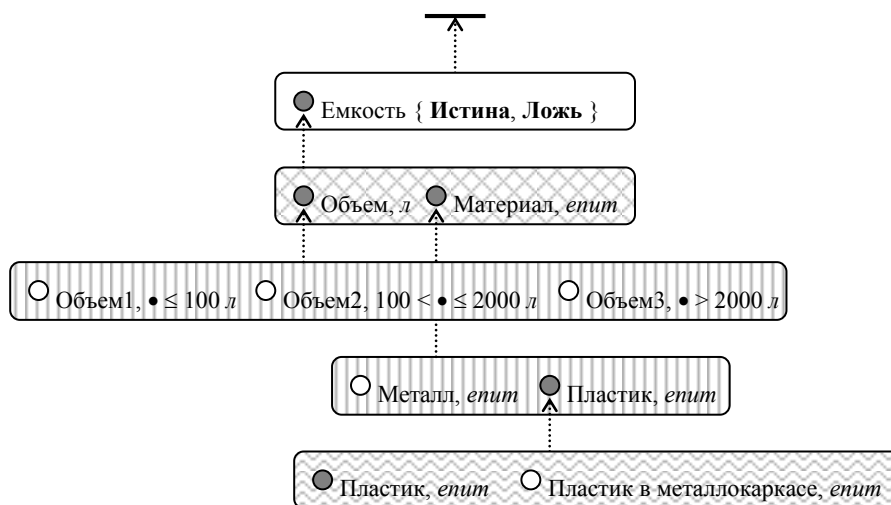


Рис. 5. Система измеряемых свойств – {Объем1, Объем2, Объем3, Металл, Пластик, Пластик в металлокаркасе} – при изучении рынка емкостей для поливной воды (епит – перечисляемый тип, или номинальная шкала значений)

3. Шкалирование измеряемого свойства, являющегося членом группы взаимно-обусловленных свойств

ВЗО-группы – часто встречающийся элемент структурной организации СИС (см. рис. 2, а; 3, в; 3, д; 4). Например, задачи классического анализа данных *фактически* по умолчанию исходят из предположения о взаимной обусловленности *всех* свойств, измеряемых у объектов изучаемой ПрО [19, 20].

Выделение части объема гипотетического формального понятия, чьи отличительные свойства составляют ВЗО-группу, вызывают специфическую проблему, определяемую *потенциальной согласованностью, когерентностью* результатов шкалирования разных измеряемых свойств, являющихся членами ВЗО-группы. Рассмотрим этот вопрос подробнее.

Пусть $\{z_1, z_2, \dots, z_n\}$ – множество отличительных и, следовательно, взаимно-обусловленных свойств гипотетического понятия субъекта об исследуемой ПрО, т. е. понятия

$$(\{\dots, z_1, z_2, \dots, z_n\}', \{\dots, z_1, z_2, \dots, z_n\}),$$

где многоточие «...» замещает свойства, унаследованные от более общих понятий. Тогда согласно определению формального понятия имеем

$$\{\dots, z_1, z_2, \dots, z_n\}' = \{\dots, z_1\}' = \{\dots, z_2\}' = \dots = \{\dots, z_n\}',$$

т. е. как все ВЗО-свойства вместе, так и каждое из них по отдельности выделяют в ПрО одно и то же гипотетическое множество объектов. Однако области значений у ВЗО-свойств – домены $D(z_1), D(z_2), \dots, D(z_n)$ – вообще говоря, различны (рис. 6, а).

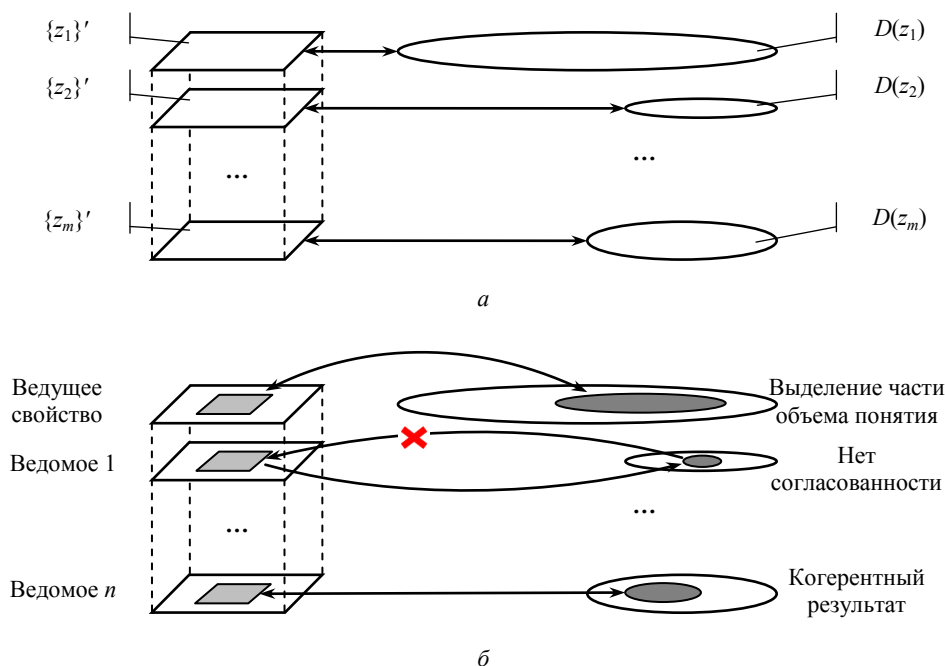


Рис. 6. Возникновение когерентности результата концептуального шкалирования члена группы взаимно обусловленных измеряемых свойств

Части объема понятия, выделенной при его делении или ограничении путем шкалирования одного из свойств ВЗО-группы («ведущего» свойства), соответствуют некоторые части доменов *остальных* («ведомых») свойств этой группы. Возникающие при этом ситуации иллюстрирует рис. 6, б.

Опираясь на определения деления и ограничения понятия, учитывая сущность реализующих эти логические операции процедур дизъюнктивного и уточняющего шкалирования, нетрудно показать, что:

- если части объема, выделяемой при уточняющем шкалировании ведомого свойства, соответствует *вполне определенная и только эта* часть домена ведомого свойства, то это означает, что данное шкалирование ведомого свойства автоматически приводит к уточняющему шкалированию рассматриваемого ведомого свойства, и наоборот – см. рис. 6, б;

- если каждой части объема, выделяемой при дизъюнктивном шкалировании ведомого свойства, соответствует *вполне определенная и только эта* часть домена ведомого свойства, то это означает, что данное шкалирование ведомого свойства автоматически приводит к дизъюнктивному шкалированию рассматриваемого ведомого свойства, и наоборот.

Из сказанного следует, что если при *уточняющем шкалировании* ВЗО-свойства субъект *не указывает* среди других ВЗО-свойств набора когерентно шкалируемых, то расширение и структуризация СИС идет по соответствующему базовому ПСС.

Иначе, когда указываются $n > 0$ когерентных ведомых ВЗО-свойств, ПСС предполагает следующее:

- все когерентно шкалируемые ВЗО-свойства замещаются группой сложно-сопряженных свойств (С-группой), включающей $2 \times (n + 1)$ измеряемых свойств, половина из которых являются новыми;

- каждое когерентно шкалируемое ВЗО-свойство в С-группе представляет *подгруппу*, включающая два свойства, где первое обусловлено вторым, и каждый по порядку член подгруппы взаимно обусловлен со всеми членами других подгрупп с таким же порядковым номером;

- связи между свойствами в С-группе достраиваются согласно правилу транзитивности обусловленности свойств.

Модельный пример применения такого шаблона приведен на рис. 7. В ВЗО-группе уточняющему шкалированию подвергнуто первое из 10 включенных в нее свойств. Свойства 3 и 9 были указаны как когерентно шкалируемые. В ВЗО-группу вложена С-группа, заместившая свойства 1, 3 и 9. В С-группе, включившей когерентно уточняемые свойства и уточняющие их новые измеряемые свойства, построены (явно) все экзистенциональные связи между свойствами.

Аналогично СИС структурируется при дизъюнктивном шкалировании ВЗО-свойства.

Когда при дизъюнктивном шкалировании ВЗО-свойства субъект не подтверждает наличие других когерентно замещаемых свойств в ВЗО-группе, то расширение и структуризация СИС идет по базовому ПСС при замещении свойства набором несовместимых свойств.

Если дизъюнктивное шкалирование ВЗО-свойства предполагает его замещение $m > 1$ несовместимыми свойствами и указываются $n > 0$ когерентных ведомых ВЗО-свойств, то ПСС предполагает следующее:

- все когерентно шкалируемые ВЗО-свойства замещаются С-группой, включающей $m \times n$ новых измеряемых свойств;
- каждое когерентно шкалируемое ВЗО-свойство в С-группе представляет подгруппа, включающая m несовместимых свойств, и каждый по порядку член подгруппы взаимно обусловлен со всеми членами других подгрупп с таким же порядковым номером;
- связи между свойствами в С-группе достраиваются согласно правилу транзитивности несовместимости относительно обусловленности.

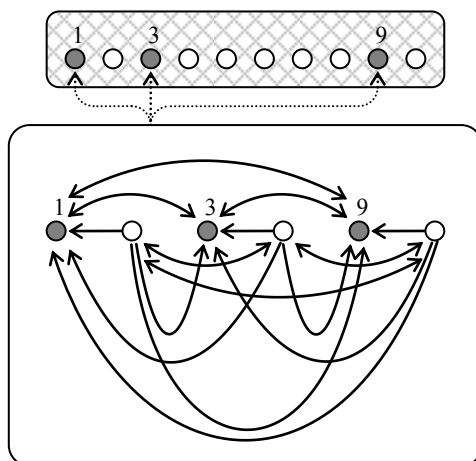


Рис. 7. Пример структуризации системы измеряемых свойств при расширении гипотетического понимания исследуемой предметной области путем уточняющего когерентного шкалирования взаимно обусловленных измеряемых свойств

Иллюстрирующий пример приведен на рис. 8. Материал взят из габитоскопии – теории портретной идентификации людей, которая широко используется в криминалистике [21]. Габитоскопия исследует различные аспекты внешности людей, устанавливая набор признаков, пригодных для идентификации. В наших терминах у объекта-человека с целью его портретной идентификации «измеряется» обширная совокупность разнообразно шкалированных ВЗО-свойств. Рис. 8 описывает структуризацию фрагмента соответствующей СИС, касающегося так называемых *собственных* составляющих внешности человека, и демонстрирует реализацию как базового, так и специального ПСС при дизъюнктивном шкалировании ВЗО-свойств.

К сожалению, рассмотренные ПСС не являются исчерпывающими для теоретически возможных случаев шкалирования ВЗО-свойств. Достаточно обратить внимание на то, что шкалируемые свойства при применении этих шаблонов изымаются из ВЗО-группы, а следовательно, в следующем акте шкалирования база для анализа когерентности оказывается некорректно суженной. В некотором смысле «одновременное» шкалирование нескольких свойств в ВЗО-группе теоретически не исключает «интерференцию» результатов такого шкалирования. Единственным способом поддержать субъекта, который так изощренно выстраивает гипотетическое представление об исследуемой ПрО, будут средства «ручного» конструирования необходимой С-группы свойств, замещающих все вовлеченные в шкалирование свойства ВЗО-группы.

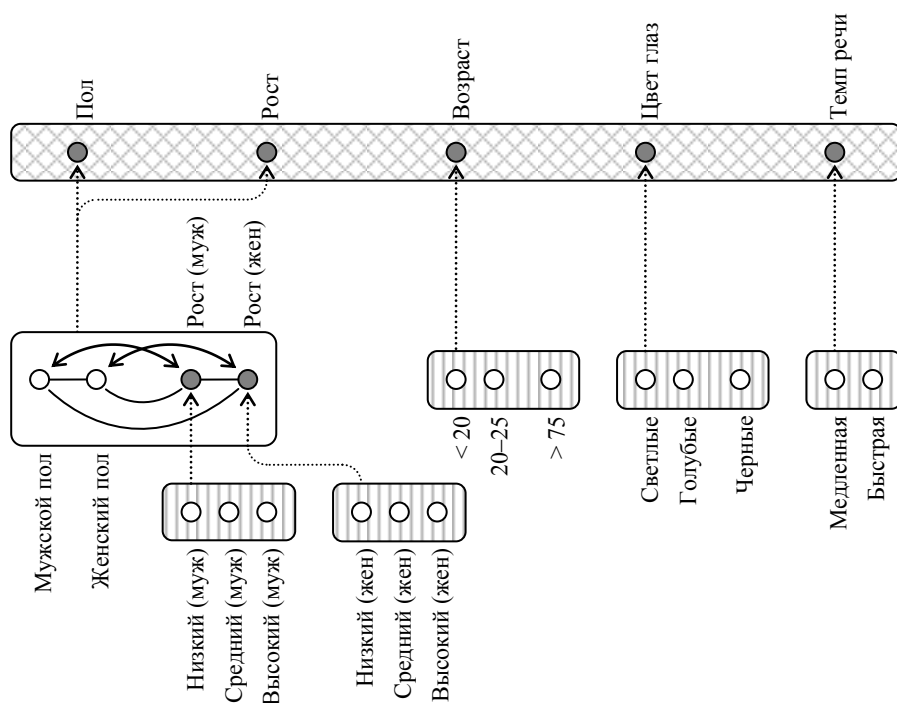


Рис.8. Практический пример структуризации системы измеряемых свойств при расширении гипотетического понимания исследуемой предметной области путем дизъюнктивного шкалирования по базовому и специальному шаблону

Закключение

В целом предложенная концепция системы измеряемых свойств (СИС) в онтологическом анализе данных (ОАД) обобщает формальные схемы описания ограничений существования свойств (ОСС), ранее построенные в работах авторов, и по-прежнему базируется на элементарных ограничениях сосуществования любых сущностей – *несовместимости* и *обусловленности* (*нейтральность*, *безразличие* не рассматривается как ограничение, но структурно всегда проявляется; наша модель в этом смысле также не является исключением).

Количество выявленных паттернов структурной организации СИС *невелико* (это следствие независимости результата концептуального шкалирования измеряемого свойства от «предыстории» появления свойства в СИС; определенное отклонение от этого возможно лишь при шкалировании взаимно обусловленных свойств), генерацию соответствующих структур легко алгоритмизировать. Поэтому на этой основе возможна *эффективная компьютерная поддержка* субъекта, формирующего СИС для онтологического исследования.

Тем не менее представленное в данной статье видение ОСС в ОАД как многоуровневой и ограниченной правилами формирования уровней СИС существенно сложнее использовавшегося авторами ранее. Однако очевидно, что это усложнение не повлечет пересмотр эвристики, предложенный авторами для формирования корректного однозначного формального контекста на основе неполной и противоречивой информации о ПрО [3]:

– выбор порога доверия к исходным данным *произволен*;

– соответствующий порогу и в общем случае недопустимый состав свойств каждого объекта искомого однозначного формального контекста следует последовательно сокращать за счет *отсечения* на каждом шаге свойства, нарушающего ОСС;

– механизм отсечения состоит в *локальном ужесточении* порога доверия в пределах каждой группы сопряженных свойств в СИС, когда основным критерием выбора измеряемого свойства для отсечения служит *минимальное ужесточение* выбранного порога доверия.

Разумеется, алгоритм, реализующий эвристику, при новом видении ОСС должен быть существенно модифицирован. Прежде всего задачей разработки и исследования здесь должны стать его рекурсивные составляющие, предназначенные для обработки многоуровневой структуры СИС.

Благодарности

Статья подготовлена по материалам научных исследований в рамках субсидированного государственного задания Института проблем управления сложными системами РАН на НИР по теме «Интерсубъективное управление инновационным развитием социотехнических объектов с применением онтологических моделей ситуаций».

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Божко А.Н., Толпаров А.Ч. Структурный синтез на элементах с ограниченной сочетаемостью. – Центр «Методолог». <http://www.metodolog.ru/00562/00562.html> (дата обращения 10.09.2018).
2. Крюни Н.К., Гвоздев В.Е., Ильясов Б.Г., Бежаева О.Я., Блинова Д.В. Элементы системной инженерии. Технологии формирования требований к аппаратно-программным комплексам на основе экспертно-статистических методов. – М.: Инновационное машиностроение, 2017. – 295 с.
3. Боргест Н.М. Распознавание образов при создании артефактов как метафора и как прикладные технологии онтологии проектирования // Онтология проектирования. – 2015. – Т. 5. – № 1 (15). – С. 19–29.
4. Половинкин А.И. Основы инженерного творчества. – М.: Машиностроение, 1988. – 368 с.
5. Смирнов С.В. Онтологический анализ предметных областей моделирования // Изв. Самарского научного центра РАН. – 2001. – Т. 3. – № 1. – С. 62–70.
6. Смирнов С.В. Построение онтологий предметных областей со структурными отношениями на основе анализа формальных понятий // Знания – онтологии – теории: Материалы Всероссийской конференции с международным участием (3–5 октября 2011 г., Новосибирск, Россия). Т. 2. – Новосибирск: Институт математики СО РАН, 2011. – С. 103–112.
7. Самойлов Д.Е., Семенова В.А., Смирнов С.В. Анализ неполных данных в задачах построения формальных онтологий // Онтология проектирования. – 2016. – Т. 6. – № 3(21). – С. 317–339.
8. Быкова В.В., Катаева А.В. Сжатое представление строгих ассоциативных правил в анализе данных // Программные продукты и системы. – 2017. – Т. 30. – № 2. – С. 187–195.
9. Ignatov D.I. Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields. In: P. Braslavski, N. Karpov, M. Worrning, Y. Volkovich, D.I. Ignatov (Eds.): Information Retrieval. Revised Selected Papers 8th Russian Summer School, RuSSIR 2014. Springer International Publishing, 2015. P. 42–141.
10. Agrawal R., Imielinski T., Swami A. Mining Associations between Sets of Items in Massive Databases // Proc. of the 1993 ACM-SIGMOD Int'l Conf. on Management of Data. P. 207–216.
11. Офicerov V.P., Smirnov S.V. Fuzzy Formal Concept Analysis in the construction of ontologies // Онтология проектирования. – 2017. – Т. 7. – № 4(26). – С. 487–495.
12. Lammari N., Metais E. Building and maintaining ontologies: a set of algorithms // Data & Knowledge Engineering. 2004. Vol. 48(2). P. 155–176.
13. Пронина В.А., Шитилина Л.Б. Использование отношений между атрибутами для построения онтологии предметной области // Проблемы управления. – 2009. – № 1. – С. 27–32.

14. Самойлов Д.Е., Семенова В.А., Смирнов С.В. Модель ограничений существования свойств в онтологическом моделировании // Проблемы управления и моделирования в сложных системах: Тр. XIX Международной конференции. – Самара: Офорт, 2017. – С. 468–473.
15. Самойлов Д.Е., Семенова В.А., Смирнов С.В. Фрактальность ограничений сосуществования свойств в задачах машинного обучения // Сборник трудов IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» ИТНТ-2018. Самара: Новая техника, 2018. – С. 2512–2518.
16. Ganter B., Wille R. Conceptual scaling // In: F. Roberts (Ed.): Applications of Combinatorics and Graph Theory to the Biological and Social Sciences. New York Springer-Verlag, 1989. P. 139–167.
17. Бочаров В.А., Маркин В.И. Основы логики. – М.: ИНФРА-М, 1998. – 296 с.
18. Ивин А.А., Никифоров А.Л. Словарь по логике. – М.: Гуманит. изд. центр ВЛАДОС, 1997. – 384 с.
19. Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д., Елизаров С.И. Анализ данных и процессов. 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2009. – 512 с.
20. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: Институт математики СО РАН, 1999. – 270 с.
21. Зинин А.М., Подволоцкий И.Н. Габитоскопия и портретная экспертиза. – М.: Норма, 2018. – 288 с.

Статья поступила в редакцию 10 сентября 2018 г.

STRUCTURE OF THE SYSTEM OF MEASURABLE PROPERTIES FOR OBJECTS EXPOSED TO MULTIDIMENSIONAL OBSERVATION AND EXPERIMENTS

D.E. Samoylov^{1,2}, V.A. Semenova^{2,3}, S.V. Smirnov²

¹ Samara University

34, Moskovskoye sh., Samara, 443086, Russian Federation

² Institute for the Control of Complex Systems of Russian Academy of Sciences

61, Sadovaya st., 443020, Samara, Russian Federation

³ Samara State Region Academiya (Nayanova)

196, Molodogvardeyskaya st., Samara, 443001, Russian Federation

Abstract. *The is devoted to ontological data analysis. This is a complex of models and methods designed to extract the formal ontology of the studied knowledge domain from the initial empirical material. Such ontologies are describing the knowledge domain of the aggregate of concepts and relationships. According to observational data and experiments, the initial empirical materials are characterized by incompleteness and inconsistency, which are conditioned by the realities of the accumulation of empirical information. This leads to the fact that the required for ontology's building formal context — the correspondence “objects-properties” — can be previously presented only within the framework of some multi-valued logic. It needs to be approximated in ordinary binary logic, since effective methods for the derivation of formal concepts are developed only for unambiguous (binary) formal contexts. In order to solve this problem correctly, in the general case it is necessary to know the properties existence constraints of the objects in the studied knowledge domain. But from where can be a priori known such subtle information about the knowledge domain that is only being studied? The main idea of the authors is that this*

Dmitry E. Samoylov, Student.

Valentina A. Semenova, Postgraduate Student, Younger Scientific Researcher.

Sergey V. Smirnov (Dr. Sci. (Techn.)), Chief Research Scientist.

knowledge is a product of a priori hypothesizing of the study's subject. Choosing certain properties to be measured from objects of multidimensional observation and experiments, the subject actually formulates hypothetical concepts about the studied knowledge domain, and using a fundamental cognitive procedure - conceptual scaling of measurable properties - he develops these hypothetical ideas. With such a formation of a system of measurable properties, well-defined properties existence constraints that determine the structure of this system are fixed. Only part of the subject's hypotheses is confirmed a posteriori, and, therefore, certain properties existence constraints are true. The paper identifies and analyzes the fundamental patterns of the structural organization of the system of measurable properties, and thus proposes an exhaustive model of properties existence constraints in the ontological data analysis.

Keywords: system of measurable properties, properties existence constraints, pattern of the structural organization, ontological data analysis, formal context.

REFERENCES

1. Bozhko A.N., Tolparov A. Ch. Structurnyy sintez na elementah s ograniichennoy sochetaemost'u [Structural synthesis on elements with limited compatibility] <http://www.metodolog.ru/00562/00562.html> (accessed September 10, 2018). (In Russian).
2. Krioni N.K., Gvozdev V.E., Il'yasov B.G., Bezhaeva O.Ya., Blinova D.V. Elementy sistemnoy inzhenerii. Tehnologii formirovaniya trebovaniy k apparatno-programmnyim kompleksam na osnove ekspertno-statisticheskikh metodov [Elements of system engineering. Technologies for the formation of requirements for hardware and software systems based on expert-statistical methods]. Innovatsionnoe mashinostroeniye, 2017. 295 p. (In Russian).
3. Borgest N.M. Pattern recognition in designing artifacts is as metaphor and as an applied technology of ontology of designing // Ontologiya proyektirovaniya [Ontology of Designing]. 2015. Vol. 5, no. 1(15). P. 19–29. (In Russian).
4. Polovinkin A.I. Osnovy inzhenernogo tvorchestva [Fundamentals of engineering creativity]. Mashinostroeniye, 1988. 368 p. (In Russian).
5. Smirnov S.V. Ontological analysis of knowledge domains of modeling // Izv. Samarskogo nauchnogo tsentra RAN. 2001, vol. 3, no.1. P. 62–70. (In Russian).
6. Smirnov S.V. Building ontologies of knowledge domains with structural relationships based on the formal concept analysis. // Znaniya – Ontologii – Teorii. Vol. 2. Novosibirsk, 2011. P. 103–112. (In Russian).
7. Samoylov D.E., Semenova V.A., Smirnov S.V. Incomplete data analysis for building formal ontologies // Ontologiya proyektirovaniya [Ontology of Designing]. 2016. Vol. 6, no. 3 (21). P. 317–339. (In Russian).
8. Bykova V.V., Kataeva A.V. A contracted representation of strong associative rules in data analysis. Programmnye produkty i sistemy [Software & Systems]. 2017. Vol. 30, no. 2. P. 187–195. (In Russian).
9. Ignatov D.I. Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields. In: P. Braslavski, N. Karpov, M. Worrington, Y. Volkovich, D.I. Ignatov (Eds.): Information Retrieval. Revised Selected Papers 8th Russian Summer School, RuSSIR 2014. Springer International Publishing, 2015. P. 42–141.
10. Agrawal R., Imielinski T., Swami A. Mining Associations between Sets of Items in Massive Databases // Proc. of the 1993 ACM-SIGMOD Int'l Conf. on Management of Data. P. 207–216.
11. Oficerov V.P., Smirnov S.V. Fuzzy Formal Concept Analysis in the construction of ontologies // Ontologiya proyektirovaniya [Ontology of Designing]. 2017. Vol. 7, no. 4(26). P. 487–495. (In Russian).
12. Lammari N., Metals E. Building and maintaining ontologies: a set of algorithms // Data & Knowledge Engineering. 2004. Vol. 48(2). P. 155–176.
13. Pronina V.A., Shipilina L.B. Using relationships between attributes to build an ontology of knowledge domain // Problemy upravleniya [Control Sciences]. 2009, no. 1. P. 27–32. (In Russian).
14. Samoylov D.E., Semenova V.A., Smirnov S.V. Model of properties existence constraints in ontological modeling // Proceedings of International Conference on Complex Systems: Control and Modeling Problems. Samara, 2017. P. 468–473. (In Russian).
15. Samoylov D.E., Semenova V.A., Smirnov S.V. Fractality of the object's properties existence constraints in machine learning // Proceedings of International Conference on Information technologies and nanotechnologies. Samara, 2018. P. 2512–2518. (In Russian).

16. *Ganter B., Wille R.* Conceptual scaling // In: F. Roberts (Ed.): Applications of Combinatorics and Graph Theory to the Biological and Social Sciences. New York Springer-Verlag, 1989. P. 139–167.
17. *Bocharov V.A., Markin V.I.* Osnovy logiki [The basics of logic]. INFRA-M, 1998. 296 p. (In Russian).
18. *Ivin A.A., Nikiforov A.L.* Slovar' po logike [Dictionary of logic]. Gumanit. izd. tsentr VLADOS, 1997. 384 p. (In Russian).
19. *Barsegyan A.A., Kupriyanov M.S., Holod I.I., Tess M.D., Yelizarov S.I.* Analiz dannyh i protsessov [Data and process analysis]. 3-e izd., pererab. i dop. SPb.: BHV-Peterburg, 2009. 512 p. (In Russian).
20. *Zagoruyko N.G.* Prikladnyye metody analiza dannyh i znaniy [Applied methods of data and knowledge analysis]. Novosibirsk: Sobolev Institute of Mathematics, 1999. 270 p. (In Russian).
21. *Zinin A.M., Podvolotskiy I.N.* Gabitoskopiya i portretnaya ekspertiza [Habitoscopy and portrait expertise]. Norma, 2018. 288 p. (In Russian).