

## ИНЖИНИРИНГ ОНТОЛОГИЙ

УДК 004.021:004.82

DOI: 10.18287/2223-9537-2020-10-3-361-379

**Алгоритмизация формирования и прагматической трансформации ограничений существования свойств предметной области****В.А. Семенова<sup>1,2</sup>, С.В. Смирнов<sup>2</sup>**<sup>1</sup> Самарский государственный технический университет, Самара, Россия<sup>2</sup> Самарский федеральный исследовательский центр РАН,  
Институт проблем управления сложными системами РАН, Самара, Россия**Аннотация**

Областью исследований является интеллектуальный анализ данных, конкретно - развиваемое авторами направление «онтологический анализ данных», что следует понимать как анализ эмпирических данных о неизученной, неструктурированной предметной области с целью построения ее формальной онтологии. Предметом исследования статьи является формирование набора свойств, которые, как предполагается, характеризуют объекты изучаемой предметной области (и, следовательно, подлежат измерению в самом широком смысле этого слова), но с ограничениями на сочетания таких характеристик у объектов – «ограничениями существования» свойств. Задачи исследования состоят в разработке алгоритмов пошагового формирования набора измеряемых свойств с ограничениями существования, алгоритмов модификации такого набора (замещения и удаления свойств), алгоритма преобразования «естественного» описания этого набора как множества с заданными на нём отношениями в форму, удобную для последующего конструктивного, прагматического использования информации об ограничениях существования в онтологическом анализе данных. В работе используются методы теории множеств и бинарных отношений, модели и методы анализа формальных понятий, а также существующая методология применения ограничений существования для построения формальных онтологий. Отличие и новизна предложенных алгоритмов формирования набора свойств с ограничениями существования заключается в «естественном» и эффективном с точки зрения машинной реализации представлении таких наборов в форме графов и матриц инцидентности. Новизна алгоритмов модификации набора свойств с ограничениями существования – в выполненной впервые алгоритмизации уникальных методов расширения набора измеряемых свойств, непосредственно опирающихся на фундаментальные законы классической логики. Сказанное верно и для алгоритма трансформации набора измеряемых свойств в набор групп свойств, однородных по виду экзистенционального сопряжения свойств-членов. Значение полученных результатов состоит в алгоритмическом обеспечении ряда этапов онтологического анализа данных.

**Ключевые слова:** анализ формальных понятий, ограничения существования свойств, онтология, онтологический анализ данных, алгоритмизация.

**Цитирование:** Семенова, В.А. Алгоритмизация формирования и прагматической трансформации ограничений существования свойств предметной области / В.А. Семенова, С.В. Смирнов // Онтология проектирования. – 2020. – Т. 10, №3(37). – С.361-379. – DOI: 10.18287/2223-9537-2020-10-3-361-379.

**Введение**

В инжиниринге онтологий сосуществуют два различных математически хорошо обоснованных подхода, основанных на теоретико-множественной семантике.

Наибольшее признание, распространение и значение (в том числе для развития теории и практики в ряде смежных областей) получил в этом смысле *анализ формальных понятий* (АФП) [1-3]. Более двух десятилетий он служит базой для развития и совершенствования методов построения, сравнения, объединения и т.д. формальных онтологий (см., например, ретроспективную подборку [4-10]). На наш взгляд, этот успех во многом определяется тем, что концепции и модели АФП оказались адекватны чрезвычайно востребованной ныне парадигме *анализа данных*, или, как теперь принято говорить, «науки о данных»<sup>1</sup>.

Фундаментальная форма представления эмпирической информации в анализе данных – *таблица «объекты-свойства»* [11, 12] – фактически тождественна модели исходных данных АФП – *многозначному формальному контексту*

$$(G, M, V, I), \quad (1)$$

где  $G = \{g_i\}_{i=1,\dots,r}$ ,  $r = |G| \geq 1$  – множество объектов;  $M = \{m_j\}_{j=1,\dots,s}$ ,  $s = |M| \geq 1$  – множество признаков объектов («свойств»);  $V$  – совокупное множество значений разных свойств,  $V = \cup_{j=1,\dots,s} V_j$ ,  $V_j$  – область существования значений (*домен*) свойства  $m_j$ ;  $I$  – тернарное отношение между  $G$ ,  $M$  и  $V$  ( $I \subseteq G \times M \times V$ ), определённое для всех пар из  $G \times M$ . Грануляция (1) с помощью *концептуального шкалирования* [13, 14] формирует<sup>2</sup> ключевую структуру АФП – *однозначный формальный контекст*  $(G, M, I)$ , где  $I$  – *бинарное* соответствие между  $G$  и  $M$  ( $I \subseteq G \times M$ ), в котором рассматриваются операторы Галуа  $\varphi$ ,  $\omega$  (общая нотация « $\omega$ ») [1]:

- $\varphi(X) = X' = \{m_j \mid m_j \in M, \forall g_i \in X: I(g_i, m_j) = \mathbf{True}\}$  – общие свойства объектов, составляющих  $X \subseteq G$ ;
- $\omega(Y) = Y' = \{g_i \mid g_i \in G, \forall m_j \in Y: I(g_i, m_j) = \mathbf{True}\}$  – объекты, которые обладают всеми свойствами из  $Y \subseteq M$ ;
- для набора объектов  $X$  совокупность их общих признаков  $X'$  служит описанием *сходства* этого набора объектов, а замкнутое множество  $X''$  является *кластером* сходных объектов.

Анализ соответствия Галуа между булеанами  $2^G$  и  $2^M$  позволяет выявить *бикластеры*  $(X, Y)$ ,  $X \subseteq G$ ,  $Y \subseteq M$ ,  $X = Y'$ ,  $Y = X'$ , которые сохраняют информацию об объектах с *одинаковым составом* свойств и находятся *по вложенности составов* в отношении частичного порядка. Фундаментальная роль такого анализа (и в целом АФП) состоит в том, что названные его результаты соответствуют в философии и классической логике определению *понятия* и *отношению обобщения* на множестве понятий [15, 16]. Т.е. каждый выявленный бикластер  $(X, Y)$  описывает понятие (точнее – *формальное понятие*), у которого  $X$  – *объем*,  $Y$  – *содержание*, и  $(X_1, Y_1) \leq (X_2, Y_2)$ , если  $Y_1 \supseteq Y_2$  (или, равно,  $X_1 \subseteq X_2$ ). Если интерпретировать (1) как данные многомерных наблюдений и измерений в исследуемой и моделируемой ПрО [4], то совокупность выявленных формальных понятий и связывающее их бинарное отношение частичного порядка « $\leq$ » («*is a*») определяет, по меньшей мере, «скелет» *формальной онтологии* этой ПрО.

Принципиально иной подход в инжиниринге онтологий предлагают авторы [17-19]. В его основе лежит концепция *ограничений существования свойств* (ОСС) – *исходная данность* трёх бинарных отношений на множестве принимаемых во внимание свойств объектов ПрО:

- *обусловленность*  $C: M \times M \rightarrow \{\mathbf{True}, \mathbf{False}\}$ , которая устанавливает, что всякий объект  $g$  обладая свойством  $m_j$  обладает и свойством  $m_k$  (хотя обратное может быть неверно), т.е.

<sup>1</sup> «На первый план выходят области науки с интенсивным использованием данных (*data intensive sciences*). Значимый научный продукт или прорывную технологию сегодня можно получить, только если опираешься на колоссальные объемы информации и результаты их глубокой смысловой переработки» (из интервью академика И.А. Соколова газете «Поиск» от 17.10.2015).

<sup>2</sup> Роль, аналогичную концептуальному шкалированию, здесь также может играть требование и учёт полной определённости функций отображения «объект – значение свойства», осуществляемых процедурами измерения свойств объектов ПрО [4].

во введённых обозначениях  $C(m_j, m_k) \leftrightarrow \forall g \in G: m_j \in \{g\}' \rightarrow m_k \in \{g\}'$ . Обусловленность рефлексивна, несимметрична и транзитивна (рисунок 1а);

- **взаимообусловленность**  $H: M \times M \rightarrow \{\text{True}, \text{False}\}$  - рефлексивное, симметричное и транзитивное отношение, индуцируемое отношением обусловленности:  $H(m_j, m_k) \leftrightarrow C(m_j, m_k) \wedge C(m_k, m_j)$  (рисунок 1б);
- **несовместимость**  $E: M \times M \rightarrow \{\text{True}, \text{False}\}$ , когда всякий объект  $g$ , обладая свойством  $m_j$ , не может иметь  $m_k$  в качестве другого свойства, и наоборот, т.е.  $E(m_j, m_k) \leftrightarrow \forall g \in G: m_j \in \{g\}' \rightarrow m_k \notin \{g\}'$ . Отношение  $E$  антирефлексивно, симметрично и нетранзитивно, но характеризуется так называемой «транзитивностью относительно обусловленности» («ЕС-транзитивностью»), что означает  $\forall x, y, z \in M: C(x, y) \wedge E(y, z) \rightarrow E(x, z)$  (рисунок 1в).

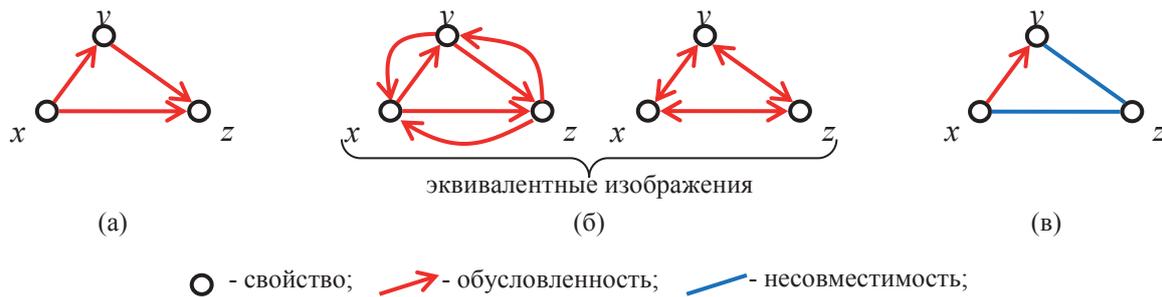


Рисунок 1 - Примеры ограничений существования свойств (самообусловленность свойств не изображена)

Тогда формирование онтологии ПрО из набора характеризующих её свойств (т.е. свойств объектов ПрО) и набора ограничений существования (т.е. ОСС как набора *сопряжённых пар свойств*, реализующих *экзистенциальные отношения* обусловленности и несовместимости свойств) состоит в выявлении всех *нормальных* подмножеств актуальных свойств ПрО, которые *замкнуты* и *совместимы*:

- $Y \subseteq M$  замкнуто, если оно содержит все свойства, обусловленные любым элементом  $Y$ , т.е.  $\forall m_j \in Y: (\exists m_k \in M, m_k \neq m_j: C(m_j, m_k)) \rightarrow m_k \in Y$ ;
- $Y \subseteq M$  совместимо, если любые два элемента  $Y$  не связаны отношением несовместимости, т.е.  $\forall m_j \in Y: (\exists m_k \in M, m_k \neq m_j: E(m_j, m_k)) \rightarrow m_k \notin Y$ .

Нормальные подмножества актуальных свойств ПрО описывают все формальные понятия искомой онтологии, которые по вложенности составов нормальных множеств свойств организуются в граф наследования «*is a*».

В контексте вышесказанного наш вклад состоит в развитии *онтологического анализа данных* (ОАД) – комплекса моделей и методов вывода формальных онтологий ПрО из характеризующих её *неполных* и *противоречивых* эмпирических данных, представленных в форме *обобщённой* таблицы «объекты-свойства» [20, 21]. При этом установлена общность очерченных формальных подходов в аспекте единства их гипотетико-дедуктивной основы и *необходимость* их *совместного* использования в ОАД [22, 23]. Последнее объясняет интерес к формированию наборов измеряемых свойств ПрО и ОСС как в аспекте их генезиса и когнитивных задач - эти вопросы рассмотрены в [22, 23], - так и чисто в структурном плане, что является предметом этой работы.

В предлагаемой статье решается задача алгоритмизации пошагового формирования и модификации субъектом ОАД набора свойств с ОСС, а также его преобразования в набор частично пересекающихся *групп сопряжённых свойств*, обеспечивающих эффективную

идентификацию нормальных множеств свойств ПрО. В отличие от техники нормализации, основанной на булевых функциях и уравнениях, которая предложена в [17, 18], используется «естественное» описание ОСС в форме графов (см. рисунок 1) и матриц инцидентности. Это по нашему убеждению предопределяет больший когнитивный и «объясняющий» потенциал такого представления для субъекта ОАД, а также эффективную машинную реализацию соответствующих алгоритмов. Кроме того, *алгоритмы модификации* реализуют не имеющие аналогов методы расширения набора актуальных свойств с ОСС, которые опираются непосредственно на фундаментальные законы классической логики [22, 23].

## 1 Пошаговое формирование набора свойств с ограничениями существования

Для обозначения свойств объектов исследуемой ПрО, или *измеряемых свойств*, в данном разделе используются идентификаторы, у которых первым символом является строчная латинская буква, например,  $x \in M$ .

Здесь и далее до определённого момента не принимается во внимание (в силу локальности действия) имманентная самообусловленность каждого свойства. Тогда нетрудно видеть, что с каждым свойством в наборе свойств с ОСС связаны три множества (рисунок 2):

- $\underline{C}(x)$  – множество свойств, обуславливающих свойство  $x$ ,  $0 \leq |\underline{C}(x)| \leq |M| - 1$ ,
- $\overline{C}(x)$  – множество свойств, обусловленных свойством  $x$ ,  $0 \leq |\overline{C}(x)| \leq |M| - 1$ ,
- $E(x)$  – множество свойств, несовместимых со свойством  $x$ ,  $0 \leq |E(x)| \leq |M| - 1$ ,

- причём лишь  $\underline{C}(x)$  и  $\overline{C}(x)$  могут пересекаться. Очевидно несложное выявление  $\underline{C}(x)$ ,  $\overline{C}(x)$  и  $E(x)$  будем относить к числу *базовых алгоритмических примитивов*, которые не детализируются.

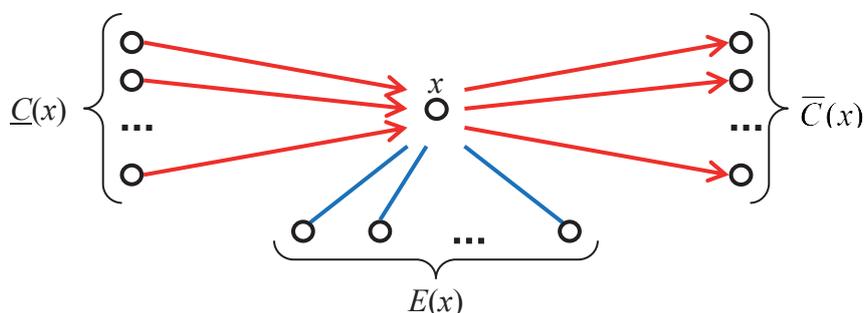


Рисунок 2 – Множества свойств, ассоциированные в наборе свойств с ограничениями существования с каждым отдельно взятым свойством (пояснение графических образов см. на рисунке 1)

В разделе рассматриваются два следующих элементарных созидательных действия:

- 1) *добавление нового свойства* в существующий набор измеряемых свойств;
- 2) *добавление новой сопряжённости пары свойств* – обусловленности или несовместимости - в набор ОСС.

Все этапы последующего изложения иллюстрирует граф на рисунке 3, который описывает матрица инцидентности, представленная на рисунке 4.

### 1.1 Расширение набора свойств

Задача добавления множества новых свойств в набор измеряемых свойств с ОСС, когда не выдвигается требование их какой-либо сопряжённости с уже существующими в этом наборе свойствами, *тривиальна*. Однако постановку этой задачи можно обобщить, считая, что в добавляемом множестве свойств могут присутствовать «свои» ОСС. Поскольку акты

добавления свойств и включения их в экзистенциальные отношения (добавления сопряжённостей) разумно разделять, то такое в целом полезное обобщение задачи целесообразно ограничить. Здесь во внимание следует принять *когнитивный смысл* расширения набора измеряемых свойств [22, 23], а также практические соображения по упрощению реализации алгоритмических примитивов, используемых при построении матрицы инцидентности, пример которой приведён на рисунке 4. С учётом этого рационально ограничиться тремя возможностями расширения набора измеряемых свойств, алгоритмы которых приводятся ниже.

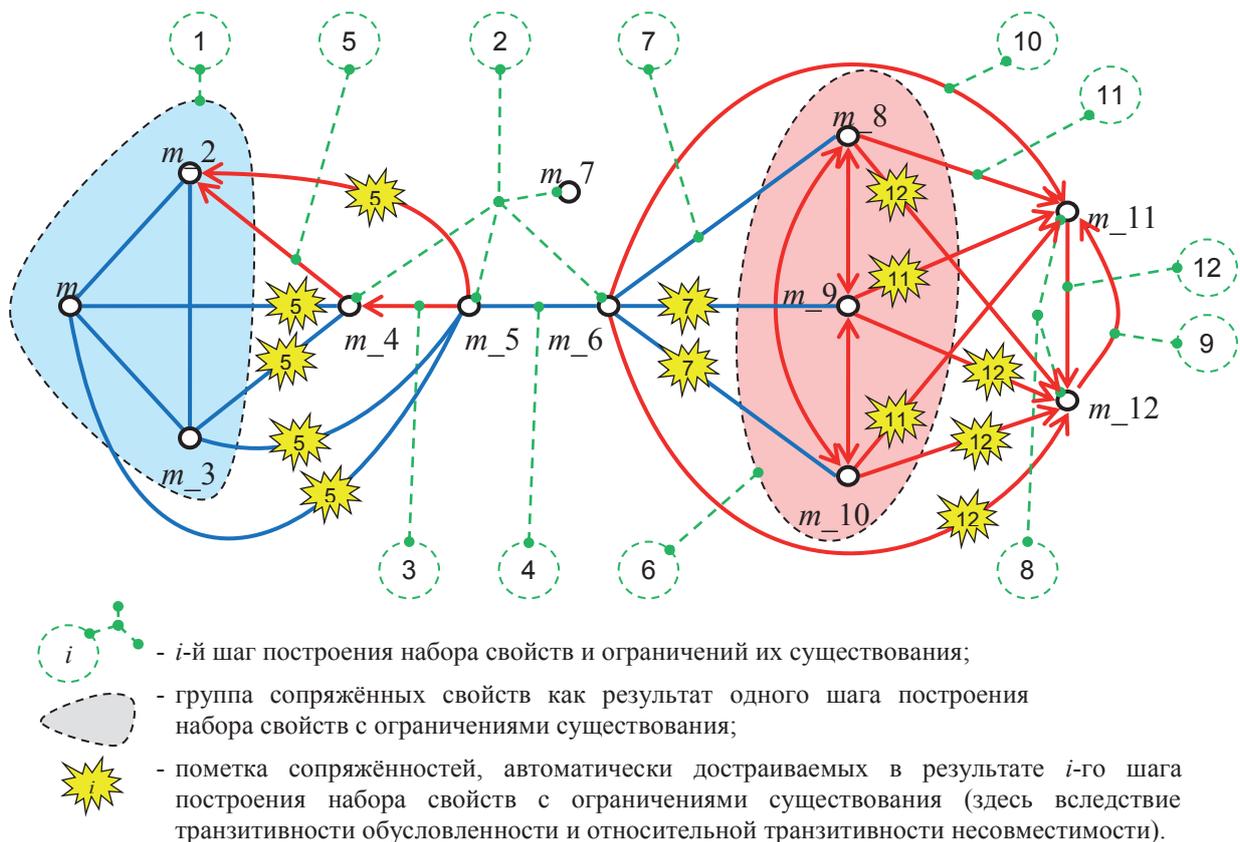


Рисунок 3 – Пример пошагового построения набора свойств с ограничениями существования (пояснение ряда графических образов см. на рисунке 1)

#### Алгоритм «Добавление $k$ несопряжённых свойств», $k \geq 1$

- 1) **for**  $i := 1$  to  $k$
- 2)      $x := \text{new } P'P$  - класс «Свойство»
- 3)      $M := M \cup \{x\}$  'добавление свойства  $x$  в набор измеряемых свойств
- 4) **next**  $i$

#### Алгоритм «Добавление $k$ взаимообусловленных свойств», $k \geq 2$

- 1)  $n := |M|$
- 2) Добавление  $k$  несопряжённых свойств 'см. описание выше
- 3) **for**  $i := 1$  to  $k$
- 4)     **for**  $j := i + 1$  to  $k$
- 5)          $C[M[n + i], M[n + j]] := \text{True}$  'сопрягаемые свойства извлекаются из набора  $M$
- 6)          $C[M[n + j], M[n + i]] := \text{True}$  'с помощью примитива  $[\cdot]$
- 7)     **next**  $j$
- 8) **next**  $i$

Алгоритм «Добавление  $k$  попарно несовместимых свойств»,  $k \geq 2$

- 1)  $n := |M|$
- 2) Добавление  $k$  несопряжённых свойств
- 3) **for**  $i := 1$  **to**  $k$
- 4)     **for**  $j := i + 1$  **to**  $k$
- 5)          $E(M[n + i], M[n + j]) := \text{True}$
- 6)     **next**  $j$
- 7) **next**  $i$

На рисунках 3, 4 шаг 1 добавляет три попарно несовместимых свойства, шаг 2 – четыре несопряжённых свойства, шаг 6 – три взаимообусловленных (ВЗО-) свойства, шаг 8 – два не-сопряжённых свойства.

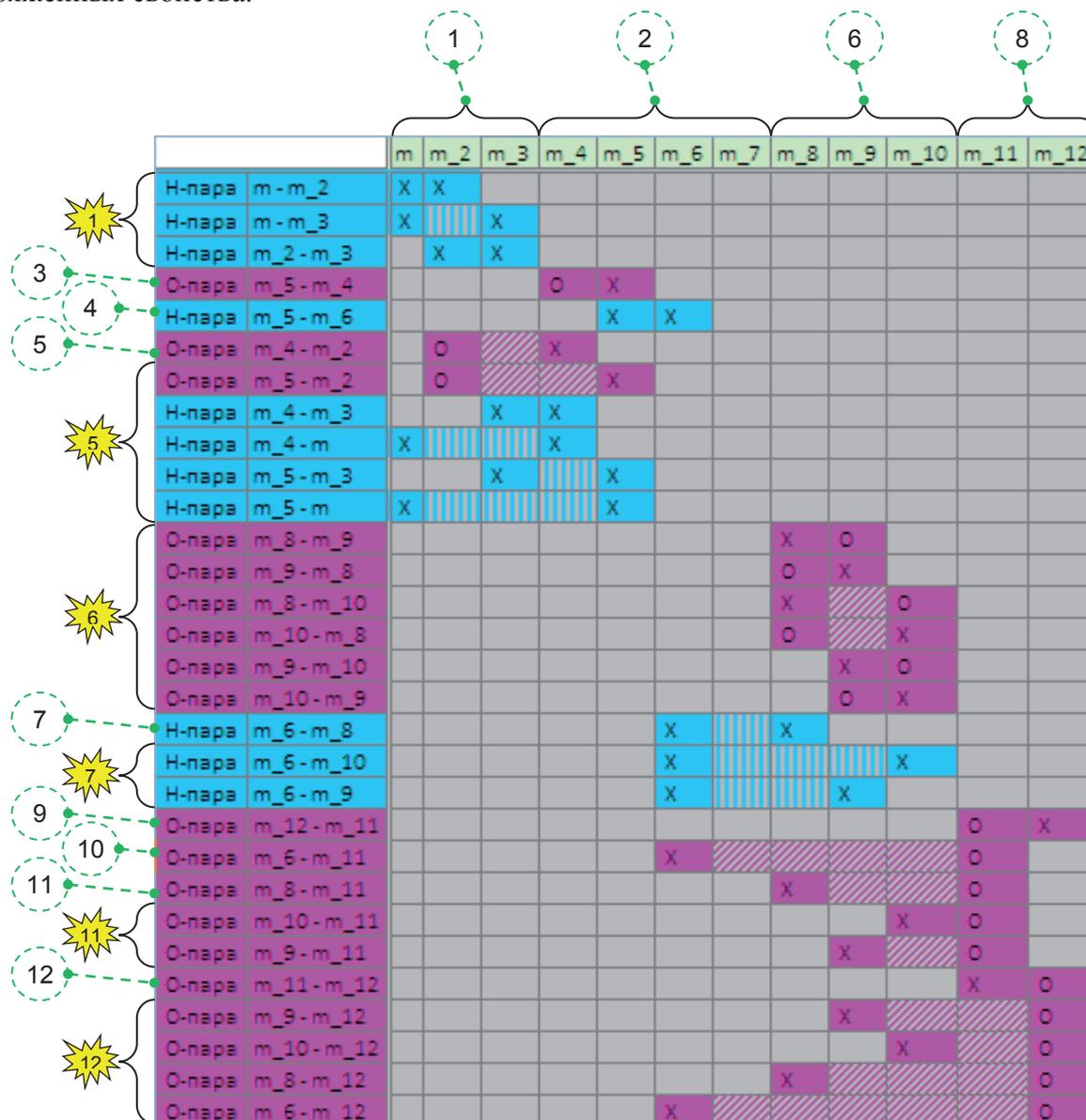


Рисунок 4 – Пример пошагового построения матрицы инцидентности «сопряжённости пар свойств – свойства», описывающей набор свойств с ограничениями существования: Н-пара – пара несовместных свойств; О-пара – пара свойств с обусловленностью, где знак инцидентности «О» указывает обусловленное свойство (пояснение других графических образов см. на рисунке 2)

## 1.2 Добавление сопряжённости пары свойств

Желание субъекта ОАД добавить в набор ОСС сопряжённость пары свойств  $x, y \in M$  – будь то обусловленность любой «направленности» (О-пара), либо несовместимость (Н-пара) – должно быть *отклонено*, если:

- указанная сопряжённость уже существует;
- указанная сопряжённость противоречит существующей сопряжённости  $x$  и  $y$ , т.е. потребована либо несовместимость  $x$  и  $y$ , а эта пара так или иначе уже состоит в отношении обусловленности, либо наоборот ( $\forall x, y \in M: E(x, y) \wedge (C(x, y) \vee C(y, x)) \equiv \mathbf{False}$ );
- добавление указанной сопряжённости нарушит законы транзитивности экзистенциальных отношений. Нетрудно показать, что нарушение ЕС-транзитивности произойдёт лишь при попытке добавления в ОСС О-пары  $(x, y)$  в случае, когда  $\exists z \in \underline{C}(x) \cap E(y)$ , а нарушение транзитивности обусловленности – при попытке добавления Н-пары  $(x, y)$  в случае, когда  $\exists z \in \underline{C}(x) \cap \underline{C}(y)$ .

Сказанное вполне определяет порядок проверки допустимости добавления сопряжённости пары свойств, а собственно добавление сопряжённости пары свойств, инициированное субъектом ОАД, в общем случае вызывает необходимость *добавления* в ОСС других новых сопряжённостей свойств в силу транзитивности экзистенциальных отношений. Эти действия составляют содержание алгоритмов добавления О- и, отдельно, Н-пары свойств.

Алгоритм «Добавление О-пары  $(x, y)$ »,  $x, y \in M$

- 1) **if not**  $C(x, y) = \mathbf{True}$  **and not**  $E(x, y) = \mathbf{True}$  **and**  $|\underline{C}(x) \cap E(y)| = 0$  **then**
- 2)      $\underline{C}(x) := \underline{C}(x) \cup \{x\}$  'x и обуславливающие его свойства
- 3)      $\bar{C}(y) := \bar{C}(y) \cup \{y\}$  'y и обуславливаемые им свойства
- 4)     'Добавление в ОСС новых О-пар (включая указанную субъектом ОАД),  
'если они там отсутствуют и речь не идёт об образовании самообусловленности;  
'тем самым транзитивность обусловленности оказывается выполненной:  
   **for each**  $z1$  **in**  $\underline{C}(x)$
- 5)         **for each**  $z2$  **in**  $\bar{C}(y)$
- 6)             **if**  $C(z1, z2) = \mathbf{False}$  **and**  $z1 \neq z2$  **then**  $C(z1, z2) := \mathbf{True}$
- 7)             **next**  $z2$
- 8)         **next**  $z1$
- 9)      $M^* := \emptyset$  'начало формирования множества свойств,  
'несовместимых с каждым свойством из  $\bar{C}(y)$
- 10)    **for each**  $z1$  **in**  $\bar{C}(y)$
- 11)         **for each**  $z2$  **in**  $E(z1)$   $M^* := M^* \cup \{z2\}$  **next**  $z2$  'бесповторное добавление свойства  $z2$
- 12)          $M^* := M^* \cup \{z2\}$  **next**  $z2$  'бесповторное добавление свойства  $z2$
- 13)         **next**  $z2$
- 14)         **next**  $z1$
- 15)     'Добавление в ОСС новых Н-пар  $(z1, z2)$  –  
'несовместимости каждого свойства из  $\underline{C}(x)$  с каждым свойством в  $M^*$ ,  
'если она в ОСС отсутствует;  
'тем самым ЕС-транзитивность оказывается выполненной:  
   **for each**  $z1$  **in**  $\underline{C}(x)$
- 16)         **for each**  $z2$  **in**  $M^*$
- 17)             **if**  $E(z1, z2) = \mathbf{False}$  **then**  $E(z1, z2) := \mathbf{True}$
- 18)             **next**  $z2$
- 19)         **next**  $z1$
- 20) **end if**

На рисунках 3, 4 шаг 5 добавляет О-пару  $(m_4, m_2)$ , и для выполнения транзитивности экзистенциальных отношений автоматически добавляются О-пара  $(m_5, m_2)$  и Н-пары  $(m_4, m_3)$ ,  $(m_4, m)$ ,  $(m_5, m_3)$ ,  $(m_5, m)$ . Аналогично при добавлении О-пары  $(m_8, m_{11})$

автоматически формируются О-пары  $(m_{10}, m_{11})$ ,  $(m_9, m_{11})$ , а при добавлении О-пары  $(m_{11}, m_{12})$  - О-пары  $(m_9, m_{12})$ ,  $(m_{10}, m_{12})$ ,  $(m_8, m_{12})$ ,  $(m_6, m_{12})$ .

Алгоритм «Добавление Н-пары  $(x, y)$ »,  $x, y \in M$

- 1) **if not**  $E(x, y) = \text{True}$  **and not**  $(C(x, y) = \text{True}$  **or**  $C(x, y) = \text{True})$  **and**  $|\underline{C}(x) \cap \underline{C}(y)| = 0$  **then**
- 2)      $\underline{C}(x) := \underline{C}(x) \cup \{x\}$  'x и обуславливающие его свойства
- 3)      $\underline{C}(y) := \underline{C}(y) \cup \{y\}$  'y и обуславливающие его свойства
- 4)     'Добавление в ОСС новых Н-пар (включая указанную субъектом ОАД),  
'если они там отсутствуют; тем самым ЕС-транзитивность оказывается выполненной:  
   **for each**  $z1$  **in**  $\underline{C}(x)$
- 5)         **for each**  $z2$  **in**  $\underline{C}(y)$
- 6)             **if**  $E(z1, z2) = \text{False}$  **then**  $E(z1, z2) := \text{True}$
- 7)             **next**  $z2$
- 8)         **next**  $z1$
- 9) **end if**

На рисунках 3, 4 шаг 7 добавляет Н-пару  $(m_6, m_8)$ , и для выполнения ЕС-транзитивности автоматически добавляются Н-пары  $(m_6, m_{10})$ ,  $(m_6, m_9)$ .

## 2 Алгоритмы модификации набора свойств с ограничениями существования

В данном разделе вначале анализируются такие элементарные модификации набора измеряемых свойств с ОСС как *удаление свойства* из набора измеряемых свойств и *удаление сопряжённости пары свойств* из ОСС. Затем рассматриваются более «нагруженные» в когнитивном смысле модификации набора измеряемых свойств и ОСС, возникающие в результате концептуального шкалирования измеряемых свойств, и, наконец, *замещение* свойства группой взаимообусловленных свойств.

### 2.1 Удаление сопряжённости пары свойств

Вне сомнения, изъятие из ОСС сопряжённости свойств  $x, y \in M$  чревато нарушением транзитивности экзистенциональных отношений измеряемых свойств:

- транзитивность обусловленности нарушится, если из ОСС удалить О-пару  $(x, y)$ , являющуюся «замыканием» двухзвенной цепочки обусловленностей, т.е.  $\exists z \in M, z \neq x, z \neq y: C(x, z) \wedge C(z, y) = \text{True}$ ;
- транзитивность несовместимости относительно обусловленности - ЕС-транзитивность – нарушится, если из ОСС удалить Н-пару  $(x, y)$ , являющуюся «замыканием» двухзвенной цепочки «обусловленность–несовместимость», конкретно  $\exists z \in M, z \neq x, z \neq y: C(x, z) \wedge E(z, y) \vee C(y, z) \wedge E(z, x) = \text{True}$  (дизъюнкция в последней формуле объясняется симметричностью несовместимости пары свойств).

Таким образом, основным содержанием алгоритмов удаления сопряжённости пар свойств из ОСС является проверка допустимости такого действия.

Алгоритм «Удаление О-пары  $(x, y)$ »,  $x, y \in M$

- 1) **if**  $|\bar{C}(y) \cap \underline{C}(y)| = 0$  **then** 'О-пара  $(x, y)$  не является «замыкающей»
- 2)      $C(x, y) := \text{False}$  'удаление обусловленности свойством  $x$  свойства  $y$
- 3) **end if**

Алгоритм «Удаления Н-пары  $(x, y)$ »,  $x, y \in M$

- 1) **if**  $|\bar{C}(y) \cap E(y)| + |\bar{C}(y) \cap E(x)| = 0$  **then** 'Н-пара  $(x, y)$  не является «замыкающей»
- 2)      $E(x, y) := \text{False}$  'удаление несовместимости свойств  $x$  и  $y$
- 3) **end if**

Например, в наборе свойств с ОСС на рисунке 3 попытка удаления О-пары ( $m_6, m_{12}$ ) или Н-пары ( $m, m_5$ ) будет отклонена, а удаление О-пары ( $m_4, m_2$ ) или Н-пары ( $m, m_3$ ) возможно.

## 2.2 Удаление свойства

Удаление свойства  $x \in M$  – абсолютно неконфликтная модификация набора измеряемых свойств с ОСС, поскольку она предполагает удаление всех сопряженностей этого свойства, причём отсутствие нарушений транзитивности экзистенциальных отношений измеряемых свойств гарантируется «автоформированием» условия

$$\exists z \in M, z \neq x, z \neq y: C(x, z) \wedge C(z, y) = \mathbf{True}, C(x, z) \wedge E(z, y) \vee C(y, z) \wedge E(z, x) = \mathbf{True},$$

а удаление несопряженного свойства из набора измеряемых свойств тривиально.

Алгоритм «Удаление свойства  $x$ »,  $x \in M$

- 1) **for each**  $y$  **in**  $\bar{C}(x)$  'удаление сопряжения с обуславливаемыми свойствами
- 2)      $C(x, y) := \mathbf{False}$
- 3) **next**  $y$
- 4) **for each**  $y$  **in**  $\underline{C}(x)$  'удаление сопряжения с обуславливающими свойствами
- 5)      $C(y, x) := \mathbf{False}$
- 6) **next**  $y$
- 7) **for each**  $y$  **in**  $E(x)$  'удаление сопряжения с несовместимыми свойствами
- 8)      $E(y, x) := \mathbf{False}$
- 9) **next**  $y$
- 10)  $M := M \setminus \{x\}$

## 2.3 Когнитивное шкалирование свойств

В [22, 23] обоснована концепция, согласно которой набор измеряемых свойств с ОСС – продукт, возникающий в результате выдвижения субъектом ОАД *гипотез о понятийной структуре*, описывающей исследуемую ПрО.

Например, добавление в набор измеряемых свойств одного нового свойства знаменует образование множества новых гипотетических понятий, содержания которых возникают как определенные комбинаторные объекты из нового и существовавших измеряемых свойств, а также новых гипотетических понятий, обобщающих такие комбинаторные объекты.

Два (и только два) других способа формирования новых гипотетических понятий - известные в классической логике *деление* и *ограничение* существующего понятия [15, 16] - вызывают *модификации* набора измеряемых свойств, которые следует расценивать как соответственно *дизъюнктивное* и *уточняющее* концептуальное шкалирование свойств [22, 23]:

- дизъюнктивное (номинальное) шкалирование свойства выполняется путём деления его домена на  $k \geq 2$  непересекающихся частей и приводит к *замещению* шкалируемого свойства набором  $k$  новых *несовместимых* свойств;
- уточняющее (порядковое) шкалирование свойства устанавливается покрытием его домена двумя областями, первая из которых есть домен целиком, а вторая – какая-то его часть («строгая часть»). Тогда у объектов исследуемой ПрО *добавляется* новое свойство, которое некоторым образом определяет выделенную часть домена шкалируемого свойства и *обуславливает* последнее. Обобщением такого приема является уточняющее шкалирование одновременно  $k \geq 2$  свойств, когда добавляемое новое свойство описывает одновременно *все* выделенные части (среди которых хотя бы одна строгая) доменов шкалируемых свойств и, следовательно, обуславливает все  $k$  шкалируемых свойств.

### 2.3.1 Дизъюнктивное шкалирование свойства

Анализ возможности дизъюнктивного шкалирования свойства  $x \in M$  иллюстрирует рисунок 5, где, в частности, демонстрируется *неопределённость* результата шкалирования обусловленного свойства ( $|\underline{C}(x)| > 0$ ), возникающая в силу ЕС-транзитивности отношения несовместимости свойств. Поэтому предлагается постулировать невозможность дизъюнктивного шкалирования обусловленного измеряемого свойства, и алгоритм такого шкалирования начинать с соответствующей проверки.

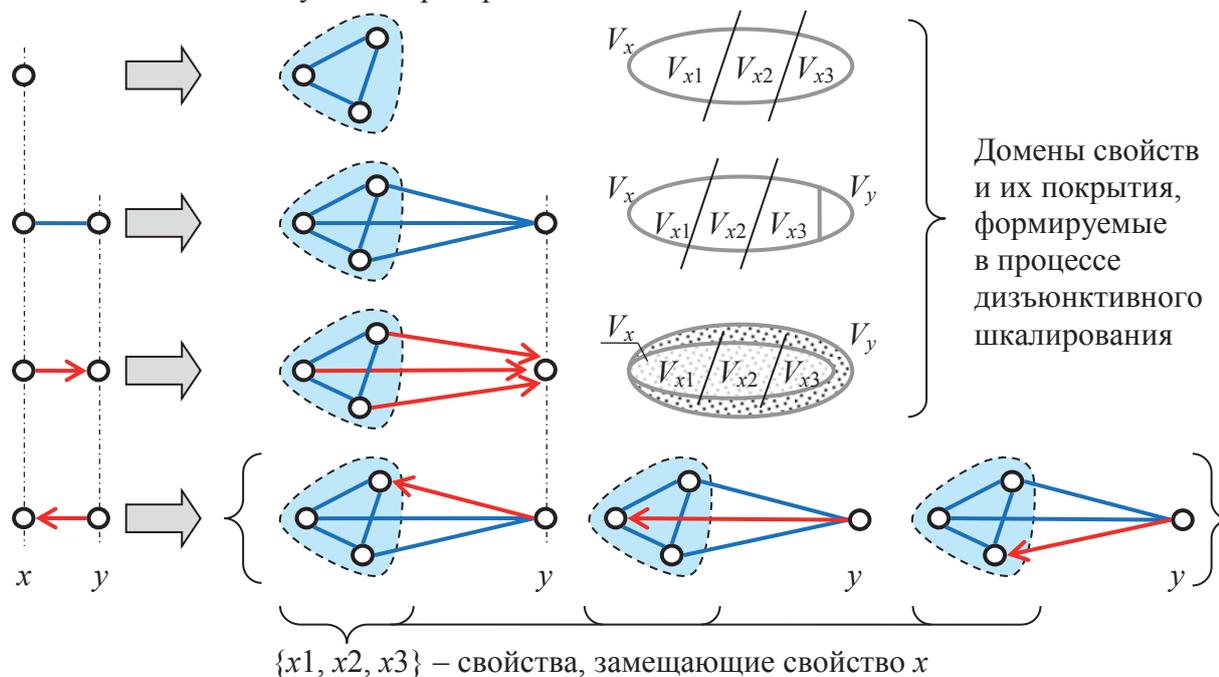


Рисунок 5 – Пример дизъюнктивного шкалирования: замещение тремя несовместимыми свойствами  $\{x_1, x_2, x_3\}$  свойства  $x$ , сопряжённого со свойством  $y$ ; приведены случаи:  $x$  без сопряжения,  $y \in E(x)$ ,  $y \in \bar{C}(x)$ ,  $y \in \underline{C}(x)$  (пояснение ряда графических образов см. на рисунке 1)

#### Алгоритм «Дизъюнктивное шкалирование свойства $x$ с замещением шкалируемого свойства $k$ новыми свойствами», $x \in M, k \geq 2$

- 1) **if**  $|\underline{C}(x)| = 0$  **then** 'дизъюнктивное шкалирование свойства  $x$  возможно
- 2)      $n := |M|$
- 3)     Добавление  $k$  несопряжённых свойств 'см. описание выше
- 4)     **for**  $i := 1$  **to**  $k$
- 5)         **for each**  $y$  **in**  $E(x)$
- 6)             Добавление  $N$ -пары  $(M[n + i], y)$  'см. описание выше
- 7)         **next**  $y$
- 8)         **for each**  $y$  **in**  $\underline{C}(x)$
- 9)             Добавление  $O$ -пары  $(M[n + i], y)$  'см. описание выше
- 10)         **next**  $y$
- 11)     **next**  $i$
- 12)     Удаление свойства  $x$  'см. описание выше
- 13) **end if**

### 2.3.2 Уточняющее шкалирование свойства

Анализ, аналогичный тому, что очерчен для дизъюнктивного шкалирования, позволяет установить, что невозможно реализовать обобщённую версию уточняющего шкалирования множества  $M^* \subseteq M, |M^*| \geq 2$  измеряемых свойств, если хотя бы два из них несовместимы

(естественно, для уточняющего шкалирования одного свойства, т.е. при  $|M^*| = 1$ , последнее условие не возникает; формально это следует из антирефлексивности отношения несовместимости свойств).

Например, присутствующее на рисунке 3 множество свойств  $\{m_4, m_5, m_6\}$  не может быть подвергнуто уточняющему шкалированию, а множество  $\{m_4, m_5, m_7\}$  шкалировать таким образом допустимо.

В приводимом далее алгоритме смысл проверки возможности уточняющего шкалирования заданного множества  $M^* \subseteq M$  свойств передан, но её реализация может быть эффективнее, т.к. первое же обнаружение несовместимости пары свойств в  $M^*$  проясняет ситуацию.

Алгоритм «Уточняющее шкалирование множества свойств  $M^*$  с добавлением нового свойства»,  $M^* \subseteq M, |M^*| \geq 1$

```

1) b := True 'флажок «Шкалирование возможно»
2) if  $|M^*| \geq 2$  then
3)   for each x in  $M^*$ 
4)     for each  $y \in M^* \setminus \{x\}$ 
5)       if  $E(x, y) = \text{True}$  then b := False
6)     next y
7)   next x
8) end if
9) if b = True then 'уточняющее шкалирование множества  $M^*$  возможно
10)  n :=  $|M|$ 
11)  Добавление_k_несопряжённых_свойств, k = 1
12)  for each x in  $M^*$ 
13)    Добавление_О-пары_( $M[n + 1], x$ )
14)  next x
15) end if

```

## 2.4 Замещение свойства группой взаимообусловленных свойств

По определению во множестве гипотетических понятий, отвечающих формируемому субъектом ОАД набору измеряемых свойств с ОСС, содержание отдельно взятого понятия отличается от содержаний других понятий, по меньшей мере, одним свойством. Когда таких *отличительных* свойств в содержании какого-то понятия несколько, то в аспекте ОСС они *взаимообусловлены* [22]. Действительно, у всех объектов из объёма этого гипотетического понятия (и только у них среди всех объектов, гипотетически присутствующих в исследуемой ПрО) кроме, возможно, других свойств наличествуют *все* рассматриваемые отличительные свойства. Следовательно, устанавливая, что наблюдаемый в ПрО объект обладает *любым одним* таким отличительным свойством, согласно выдвинутой гипотетической картине следует констатировать, что он характеризуется и всеми остальными отличительными свойствами, и, таким образом, любое одно такое отличительное свойство обуславливает все другие, а это означает, что все обсуждаемые отличительные свойства взаимообусловлены.

Опыт показывает, что формируя гипотезы о понятийной структуре, описывающей исследуемую ПрО, субъект ОАД нередко решает, что вместо одного отличительного свойства некоторое гипотетическое понятие о ПрО должно характеризоваться несколькими, например, уточняя, что интересующие его люди должны характеризоваться не только фамилией, но и именем. Поэтому целесообразно при формировании набора измеряемых свойств с ОСС предполагать возможностью *замещения* одного свойства группой ВЗО-свойств.

Очевидно, что в силу транзитивности экзистенциональных отношений между свойствами *каждое* из ВЗО-свойств, замещающих выбранное измеряемое свойство, «наследует» все его сопряжённости в наборе измеряемых свойств (это также следует из того, что отношение вза-

имообусловленности  $H$  в силу характеризующих его свойств симметрии – см. введение - разбивает множество измеряемых свойств на классы эквивалентности). Реализация наследования сопряжённостей замещаемого свойства замещающими его ВЗО-свойствами составляет основное содержание рассматриваемого алгоритма замещения.

#### Алгоритм «Замещение свойства $x$ группой $k$ ВЗО-свойств», $x \in M, k \geq 2$

- 1) Добавление  $k$  взаимобусловленных свойств 'см. описание выше
- 2)  $y := M[n + k]$  'последнее из замещающих ВЗО-свойств,  
'которое далее наследует все сопряжённости замещаемого свойства  $x$ ;  
'при этом остальные замещающие ВЗО-свойства получают все эти сопряжённости  
'автоматически – см. алгоритмы добавления сопряжённостей пар свойств
- 3) **for each**  $z$  **in**  $\bar{C}(x)$
- 4)     Добавление\_О-пары\_( $y, z$ )
- 5) **next**  $z$
- 6) **for each**  $z$  **in**  $\underline{C}(x)$
- 7)     Добавление\_О-пары\_( $z, y$ )
- 8) **next**  $z$
- 9) **for each**  $z$  **in**  $E(x)$
- 10)    Добавление\_Н-пары\_( $z, y$ )
- 11) **next**  $z$
- 12) Удаление\_свойства\_ $x$

Иллюстрацию замещения измеряемого свойства группой ВЗО-свойств можно найти на рисунке 3, если предположить, что группа ВЗО-свойств  $\{m_8, m_9, m_{10}\}$  - результат замещения некоторого «протосвойства»  $m_{8910}$ , которое было несовместимо со свойством  $m_6$  и обуславливало свойства  $m_{11}, m_{12}$ .

### 3 Трансформация ограничений существования свойств

В отличие от метода построения онтологий, прямо основанного на анализе ОСС [17-19], в ОАД информация об ОСС используется при дефазификации нестроого формального контекста, который возникает вследствие неполноты и противоречивости имеющихся эмпирических данных, причём локально для каждого объекта, попавшего в поле зрения исследователя [20, 21]. Однако и та и другая методики связаны выявлением подмножеств свойств, *совместимых* с ОСС, т.е. нормальных подмножеств измеряемых свойств. Разбор формального определения нормального подмножества свойств (см. введение) позволяет указать в наборе свойств с ОСС *структуры*, которые будучи выявлены, открывают путь к быстрой констатации «нормальности» *актуальных* подмножеств измеряемых свойств<sup>3</sup>:

- *группа ВЗО-свойств* (ВЗО-группа) принадлежит нормальному подмножеству *только целиком* (что естественно для отдельно взятого свойства, которое формально в силу самообусловленности следует считать частным случаем ВЗО-группы);
- *группа с односторонней обусловленностью* (О-группа), в которой одно свойство или, эквивалентно, группа ВЗО-свойств обуславливает другое свойство или, эквивалентно, группу ВЗО-свойств, либо входит в нормальное подмножество *целиком*, либо своей *обусловливаемой частью*. Отношение обусловленности между ВЗО-группами обобщает отношение между свойствами в том смысле, что каждое свойство-член одной группы обуславливает все свойства-члены другой (формальный анализ этого отношения можно найти в [19]);

<sup>3</sup> Именно это – выяснение нормально или нет некоторое *данное* подмножество свойств, - а не выявление *всех* нормальных подмножеств набора свойств с ОСС, необходимо при дефазификации формального контекста в ОАД [20].

- группа попарно несовместимых свойств (Н-группа) может быть представлена в нормальном подмножестве только одним своим членом.

Таким образом, актуальна задача *трансформации* «естественного» представления набора измеряемых свойств с ОСС как конечного множества с заданными на нём бинарными отношениями в набор групп свойств, однородных по виду экзистенционального сопряжения свойств-членов. Решение этой задачи в форме матрицы инцидентности «однородные группы свойств – свойства» для набора свойств с ОСС, представленного на рисунках 3, 4, демонстрирует рисунок 6.

	m	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12
ВЗО-группа	o											
ВЗО-группа		o										
ВЗО-группа			o									
ВЗО-группа				o								
ВЗО-группа					o							
ВЗО-группа						o						
ВЗО-группа							o					
ВЗО-группа								o	o	o		
ВЗО-группа											o	o
О-группа		o		x								
О-группа		o			x							
О-группа				o	x							
О-группа						x					o	o
О-группа							x	x	x		o	o
Н-группа	x		x		x							
Н-группа	x		x	x								
Н-группа	x	x	x									
Н-группа					x	x						
Н-группа						x			x			
Н-группа						x				x		
Н-группа						x		x				

Рисунок 6 - Пример описания набора свойств с ограничениями существования в виде матрицы инцидентности «однородные группы свойств – свойства»

### 3.1 Выявление групп взаимообусловленных свойств

Во-первых, каждое в отдельности измеряемое свойство является самостоятельной ВЗО-группой. Во-вторых, ясно, что признаком ВЗО-группы является не только взаимообусловленность входящих в неё свойств, но и отсутствие пересечения с другими ВЗО-группами (исключая, разумеется, некоторые ВЗО-группы, состоящие из одного свойства). Поэтому после выявления очередной ВЗО-группы её свойства-члены следует исключить из дальнейшего анализа, что сокращает трудоёмкость выявления каждой следующей такой группы.

#### Алгоритм «Выявление ВЗО-групп»

- 1)  $M^* := M$  'получение копии множества свойств
- 2) **do while**  $|M^*| > 0$
- 3)  $x := M^*[1]$  'очередное свойство, не состоящее в ранее выявленных ВЗО-группах
- 4)  $G := \text{new } VZO$  'VZO - класс «ВЗО-группа»;  
'G используется здесь лишь как множество свойств-членов группы G
- 5)  $G := G \cup \{x\}$  'x – ядро и, возможно, единственный член новой ВЗО-группы

```

6)   for each y in C(x)
7)     for each z in C(y)
8)       if z = x then G := G ∪ {y}
9)     next z
10)  next y
11)  M* := M* \ G
12) loop 'множество свойств конечно
    
```

### 3.2 Выделение групп с односторонней обусловленностью

Поскольку по определению О-группа - это две ВЗО-группы, одна из которых обуславливает другую, то алгоритм выделения О-групп строится в предположении, что все ВЗО-группы в наборе измеряемых свойств с ОСС уже выявлены.

#### Алгоритм «Выявление О-групп»

```

1) for each G2 in VZO's 'перебор всех построенных ВЗО-групп
   'G2 используется здесь лишь как множество свойств-членов группы G2
2)  M* := C(G2[1]) \ G2
   'множество свойств-членов ВЗО-групп, отличных от ВЗО-группы G2;
   'каждое из этих свойств обуславливает каждый член ВЗО-группы G2
3)  for each G1 in VZO's 'перебор всех построенных ВЗО-групп
   'G1 используется здесь лишь как множество свойств-членов группы G1
4)    if |G1 ∩ M*| > 0 then 'ВЗО-группа G1 обуславливает ВЗО-группу G2
5)      G := new O 'O - класс «О-группа»
6)      G.C := G2 'обуславливаемые свойства-члены О-группы
7)      G.C := G1 'обуславливающие свойства-члены О-группы
8)    end if
9)  next G1
10) next G2
    
```

### 3.3 Выявление групп попарно несовместимых свойств

Проблемой выявления Н-групп в наборе измеряемых свойств с ОСС является не только их пересечение между собой, но, прежде всего, вложенности одних таких групп в другие - как правило, подобные структуры эффективно обрабатываются лишь путём рекурсии. Например, на рисунке 3 Н-группы  $\{m, m_2, m_3\}$  и  $\{m, m_3, m_5\}$  пересекаются, а в Н-группу  $\{m, m_2, m_3\}$  вложены Н-группы  $\{m, m_2\}$ ,  $\{m, m_3\}$ ,  $\{m_2, m_3\}$ .

#### Алгоритм «Выявление Н-групп»

```

1) for each x in M
2)   G := Ассоциированные_с_x_Н-группы
   'множество Н-групп, каждая из которых включает свойство x и
   'определённое подмножество попарно несовместимых с ним свойств
3)   Ns := Ns ∪ G 'бесповторное (по составу свойств-членов) накопление Н-групп
4) next x
    
```

Алгоритм рекурсивной функции «Ассоциированные\_с\_x\_Н-группы» выявляет как все пересекающиеся, так и вложенные одна в другую Н-группы, включающие в себя свойство x.

#### Алгоритм «Ассоциированные с x Н-группы», $x \in M$

```

1) Gs := ∅ 'множество Н-групп, ассоциированных со свойством x
2) for each y in E(x)
3)   M* := E(y) ∩ E(x) 'свойства, несовместимые как со свойством y, так и со свойством x
4)   if |M*| ≤ 1 then 'M* пусто или содержит одно свойство
5)     G := new N 'N - класс «Н-группа»
   'G используется здесь и как множество свойств-членов группы G
6)     G := G ∪ {y}
    
```

```

7)      if  $|M^*| = 1$  then  $G := G \cup M^*$ 
8)      else
9)       $G :=$  Ассоциированные_с_у_Н-группы 'рекурсия для свойства у
10)     end if
11)      $G_s := G_s \cup G$  'бесповторное накопление Н-групп
12) next z
13) for each G in  $G_s$ 
14)    $G := G \cup \{x\}$ 
15) next G
16) Ассоциированные_с_x_Н-группы :=  $G_s$ 

```

## Заключение

Представленные в статье алгоритмы формирования набора свойств, подлежащих измерению у объектов исследуемой ПрО, с набором ограничений существования этих свойств призваны обеспечить поддержку и сохранение результатов *априорных когнитивных актов* субъекта, намеревающегося получить данные измерений, а затем формально вывести из этой эмпирической информации онтологическую модель интересующей его ПрО. Полнота, достаточность и эффективность алгоритмической поддержки этой деятельности субъекта-исследователя достигнута как за счёт теоретического обоснования логического смысла рассматриваемых актов, так и благодаря целесообразному ограничению мыслимых возможностей такой поддержки.

Алгоритмы трансформации исходного «естественного» представления набора измеряемых свойств с ограничениями существования в набор субструктур, однородных по виду экзистенционального сопряжения свойств-членов, обеспечивают возможность эффективного практического использования априорных соображений субъекта онтологического анализа эмпирических данных. Основанием для такого вывода является возможность *прямого соотнесения* внутреннего устройства названных субструктур с прагматически важным и теоретически строгим понятием нормального подмножества измеряемых свойств.

Представленный в статье комплекс алгоритмов предполагается реализовать в разрабатываемой в Институте проблем управления сложными системами РАН системе онтологического анализа данных на массовой программной платформе.

## Благодарности

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (регистрационный номер НИОКТР АААА-А19-119030190053-2).

## Список источников

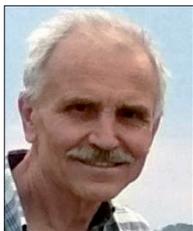
- [1] **Ganter, B.** Formal Concept Analysis. Mathematical foundations / B. Ganter, R. Wille. – SpringerVerlag Berlin Heidelberg, 1999.
- [2] **Ferré, S.** Formal Concept Analysis: From Knowledge Discovery to Knowledge Processing / S. Ferré, M. Huchard, M. Kaytoue, S.O. Kuznetsov, A. Napoli // In: P. Marquis, O. Papini, H. Prade (eds.): A Guided Tour of Artificial Intelligence Research. Vol. II: AI Algorithms. - Springer International Publishing, 2020 - P.411-445.
- [3] Formal Concept Analysis Homepage - <http://www.upriss.org.uk/fca/fca.html>
- [4] **Смирнов, С.В.** Онтологический анализ предметных областей моделирования / С.В. Смирнов // Известия Самарского научного центра РАН. – 2001. – Т. 3, № 1. – С.62-70.
- [5] **Obitko, M.** Ontology Design with Formal Concept Analysis / M. Obitko, V. Snasel, J. Smid // In: V. Snasel, R. Belohlavek (eds.): Proc. of the CLA 2004 Int. Workshop on Concept Lattices and their Applications (Ostrava, Czech Republic, September 23-24, 2004). - VŠB–Technical University of Ostrava, Dept. of Computer Science, 2004. - P.111-119.

- [6] **Sertkaya, B.** A survey on how description logic ontologies benefit from FCA / B. Sertkaya // In: Proc. of the 7th Int. Conf. on Concept Lattices and Their Applications (October 19-21, 2010). - University of Sevilla, 2010. P.2–21.
- [7] **Priya, M.** A Survey of State of the Art of Ontology Construction and Merging using Formal Concept Analysis / M. Priya, Ch.A. Kumar // Indian journal of science and technology. – 2015. – Vol. 8, Issue 24. - P.1-7.
- [8] **Ganter, B.** Conceptual Exploration / B. Ganter, S. Obiedkov. – Springer-Verlag Berlin Heidelberg, 2016. – 315 p.
- [9] **Bhuyan, Z.** FCA Based Ontology Revision / Z. Bhuyan, S.M. Hazarika // In: J.K. Mandal, G. Saha, D. Kandar, A.K. Maji (eds.): Proc. of the Int. Conf. on Computing and Communication Systems I3CS 2016 (NEHU, Shillong, India 2016). Lecture Notes in Networks and Systems, vol 24. – Springer Singapore, 2018. – P.745-755.
- [10] **Zhao, M.** Matching biomedical ontologies based on formal concept analysis / M. Zhao, S. Zhang, W. Li, G. Chen // Journal of Biomedical Semantics. - 2018. – Vol. 9:11. - 27 p.
- [11] **Загоруйко, Н.Г.** Когнитивный анализ данных / Н.Г. Загоруйко. – Новосибирск: Академическое изд-во «Гео», 2013.
- [12] **Барсегян, А.А.** Анализ данных и процессов / А.А. Барсегян, М.С. Куприянов, И.И. Холод, М.Д. Тесс, С.И. Елизаров. – СПб.: БХВ-Петербург, 2009. – 512 с.
- [13] **Ganter, B.** Conceptual scaling / B. Ganter, R. Wille // In: F. Roberts (ed.): Applications of Combinatorics and Graph Theory to the Biological and Social Sciences. – New York Springer-Verlag, 1989. – P.139-167.
- [14] **Ignatov, D.I.** Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields / D.I. Ignatov // In: P. Braslavski, N. Karpov, M. Worring, Y. Volkovich, D.I. Ignatov (eds.): Information Retrieval. Revised Selected Papers 8<sup>th</sup> Russian Summer School, RuSSIR 2014 (Nizhniy Novgorod, Russia, August 18-22, 2014). – Springer International Publishing, 2015. – P.42-141.
- [15] **Гетманова, А.Д.** Логика. Углубленный курс / А.Д. Гетманова - М.: КНОРУС, 2016.
- [16] **Ивин, А.А.** Словарь по логике / А.А. Ивин, А.Л. Никифоров. – М.: Гуманит. изд. центр ВЛАДОС, 1997.
- [17] **Lammari, N.** Building and maintaining ontologies: a set of algorithms / N. Lammari, E. Metais // Data & Knowledge Engineering. – 2004. – Vol. 48, No. 2. – P.155-176.
- [18] **Lammari, N.** POEM: an Ontology Manager based on Existence Constraints / N. Lammari, C. du Mouza, E. Metais // In: S.S. Bhowmick, J. Küng, R. Wagner (eds.): Database and Expert Systems Applications. Proc. 19th Int. Conf. DEXA 2008 (Turin, Italy, September 1-5, 2008). Lecture Notes in Computer Science, vol. 5181. - Springer-Verlag Berlin Heidelberg, 2008. – P.81-88.
- [19] **Пронина, В.А.** Использование отношений между атрибутами для построения онтологии предметной области / В.А. Пронина, Л.Б. Шипилина // Проблемы управления. – 2009. – №1. – С.27-32.
- [20] **Самойлов, Д.Е.** Анализ неполных данных в задачах построения формальных онтологий / Д.Е. Самойлов, В.А. Семенова, С.В. Смирнов // Онтология проектирования. – 2016. – Т. 6, №3(21). – С.317-339. DOI: 10.18287/2223-9537-2016-6-3-317-339.
- [21] **Samoylov, D.E.** Defuzzification of the initial context in Formal Concept Analysis / D.E. Samoylov, V.A. Semenova, S.V. Smirnov // In: V Fursov, Y Goshin, D Kudryashov (eds.): Information Technology and Nanotechnology ITNT 2019: Data Science. Proc. of the Data Science Session at the V Int. Conf. on Information Technology and Nanotechnology (Samara, Russia, May 21-24, 2019). - CEUR Workshop Proceedings, 2019. - Vol. 2416. - P.1-9. - DOI: 10.18287/1613-0073-2019-2416-1-9.
- [22] **Samoylov, D.E.** Multilevel recursive model of properties existence constraints in machine learning / D.E. Samoylov, V.A. Semenova, S.V. Smirnov // Journal of Physics: Conf. Series 1096 (2018) 012096. DOI: 10.1088/1742-6596/1096/1/012096
- [23] **Смирнов, С.В.** Две методологии вывода формальных понятий: когда и как они должны работать вместе / С.В. Смирнов // Знания – Онтологии – Теории: Материалы VII международной конференции. – Новосибирск: Институт математики СО РАН, Новосибирский государственный ун-т, 2019. - С.355-363.

## Сведения об авторах



**Семенова Валентина Андреевна**, 1994 г. рождения. Бакалавр Самарского государственного аэрокосмического университета им. С.П. Королёва по направлению «Прикладная математика и информатика» (2015), магистр Самарского университета по направлению «Механика и математическое моделирование» (2017). Аспирантка Самарского государственного технического университета по направлению «Информатика и вычислительная техника», младший научный сотрудник Института проблем управления сложными системами Самарского федерального исследовательского центра РАН. Автор и соавтор 19 научных трудов в области интеллектуального анализа данных и проектирования программных систем. ORCID: 0000-0002-0557-3890; Author ID (Scopus): 57204366624. [queenbfjr@gmail.com](mailto:queenbfjr@gmail.com)



**Смирнов Сергей Викторович**, 1952 г. рождения. Окончил Куйбышевский авиационный институт им. С.П. Королёва в 1975 г., д.т.н. (2002). Главный научный сотрудник Института проблем управления сложными системами Самарского федерального исследовательского центра РАН, профессор Поволжского государственного университета телекоммуникаций и информатики. Член Российской ассоциации искусственного интеллекта, Международной ассоциации по онтологиям и их приложениям (IAOA). В списке научных трудов более 170 работ в области прикладной математики, компьютерного моделирования сложных систем, создания интеллектуальных систем поддержки принятия решений. ORCID: 0000-0002-3332-5261; Author ID (РИНЦ): 17628; Author ID (Scopus): 57205017287. [smirnov@iccs.ru](mailto:smirnov@iccs.ru)

Поступила в редакцию 10.08.2020, после рецензирования 01.09.2020. Принята к публикации 14.09.2020.

## Algorithms for the formation and pragmatic transformation of Existence Constraints

V.A. Semenova<sup>1,2</sup>, S.V. Smirnov<sup>2</sup>

<sup>1</sup> Samara State Technical University, Samara, Russia

<sup>2</sup> Samara Federal Research Scientific Center RAS, Institute for the Control of Complex Systems RAS, Samara, Russia

### Abstract

The research area of this article is data mining, specifically the direction “ontological data analysis” developed by the authors, which should be understood as the analysis of empirical data on an unexplored, unstructured knowledge domain in order to build its formal ontology. The research subject of this article is the formation of a set of properties which are supposed to characterize objects of the studied knowledge domain (and, therefore, are subject to measurement in the broadest sense of the word), but with restrictions on the combination of such characteristics in objects — properties “existence constraints”. The objectives of the research are to develop algorithms for the step-by-step formation of a set of measured properties with existence constraints, algorithms for modifying such a set (replacing and deleting properties), an algorithm for transforming the “natural” description of this set as a set with relations specified on it into a form convenient for the subsequent constructive use of information about existence constraints in ontological data analysis. In the paper we use the methods of set theory, binary relations, models and methods of formal concept analysis and the methodology for applying the existence constraints to construct formal ontologies. The difference and novelty of the proposed algorithms for the formation of a set of properties with existence constraints lies, first of all, in the “natural” and efficient representation of such sets in the form of graphs and incidence matrices (from the point of view of machine implementation). The novelty of the algorithms for modifying a set of properties with existence constraints lies in performed for the first time algorithmization of unique methods for expanding the set of measured properties that address directly the fundamental laws of classical logic. The same is true for the algorithm for transforming a set of measured properties into a set of properties groups that are homogeneous in the form of existential conjugation of member properties. The significance of the results obtained lies in the algorithmic support of a number of stages of ontological data analysis.

**Key words:** Formal Concept Analysis, Existence Constraints, ontology, ontological data analysis, algorithms.

**Citation:** Semenova VA, Smirnov SV. Algorithms for the formation and pragmatic transformation of Existence Constraints [In Russian]. *Ontology of designing*. 2020; 10(3): 361-379. DOI: 10.18287/2223-9537-2020-10-3-361-379.

**Acknowledgment:** This research was funded by Ministry of Science and Higher Education of the Russian Federation, R&D registration numbers AAAA-A19-119030190053-2.

### List of figures

Figure 1 - Examples of Property Existence Constraints (self-conditionality of properties is not shown)

Figure 2 - Sets of properties associated with each individual property in a set of properties and a set of existence constraints (for an explanation of graphical images, see Figure 1)

- Figure 3 - An example of a step-by-step formation of a set of properties and a set of existence constraints (for an explanation of graphical images, see Figure 1)
- Figure 4 - An example of a step-by-step formation of incidence matrix of “conjugation of pairs of properties – properties”, which describes a set of properties and a set of existence constraints:  
“Н-пара” - a pair of incompatible properties; “О-пара” - a pair of properties with conditionality, where the incidence “O” indicates the conditioned property (for an explanation of other graphical images, see Figure 2)
- Figure 5 - Example of disjunctive scaling: substitution of property  $x$  (associated with property  $y$ ) with three incompatible properties  $\{x_1, x_2, x_3\}$ ; the following situations are given:  
 $x$  not related,  $y \in E(x)$ ,  $y \in \bar{C}(x)$ ,  $y \in \underline{C}(x)$  (for an explanation of some graphical images, see Figure 1)
- Figure 6 - An example of describing a set of properties and a set of existence constraints in the form of an incidence matrix “homogeneous groups of properties – properties”

## References

- [1] **Ganter B, Wille R.** Formal Concept Analysis. Mathematical foundations. Springer Berlin-Heidelberg, 1999.
- [2] **Ferré S, Huchard M, Kaytoue M, Kuznetsov SO, Napoli A.** Formal Concept Analysis: From Knowledge Discovery to Knowledge Processing. In: P Marquis, O Papini, H Prade (eds.): A Guided Tour of Artificial Intelligence Research. Vol. II: AI Algorithms. Springer International Publishing, 2020: 411-445.
- [3] Formal Concept Analysis Homepage. Source: <http://www.upriss.org.uk/fca/fca.html>
- [4] **Smirnov SV.** Ontological analysis of modeling domain [In Russian]. Bulletin of the Samara Scientific Center of RAS, 2001; 3(1): 62-70.
- [5] **Obitko M, Snasel V, Smid J.** Ontology Design with Formal Concept Analysis. In.: V Snasel, R Belohlavek (eds.): Proc. of the CLA 2004 Int. Workshop on Concept Lattices and their Applications (Ostrava, Czech Republic, September 23-24, 2004). VŠB–Technical University of Ostrava, Dept. of Computer Science, 2004: 111-119.
- [6] **Sertkaya B.** A survey on how description logic ontologies benefit from FCA. In: Concept Lattices and Their Applications: Proc. of the 7th Int. Conf. (Sevilla, Spain, 2010, October 19-21). University of Sevilla, 2010: 2–21.
- [7] **Priya M, Kumar ChA.** A Survey of State of the Art of Ontology Construction and Merging using Formal Concept Analysis. Indian journal of science and technology, 2015; 8(24): 1-7.
- [8] **Ganter B, Obiedkov S.** Conceptual Exploration. Springer, 2016.
- [9] **Bhuyan Z, Hazarika SM.** FCA Based Ontology Revision. In: JK Mandal, G Saha, D Kandar, AK Maji (eds.): Proc. of the Int. Conf. on Computing and Communication Systems. Lecture Notes in Networks and Systems, vol 24. Springer Singapore, 2018: 745-755.
- [10] **Zhao M, Zhang S, Li W, Chen G.** Matching biomedical ontologies based on formal concept analysis. Journal of Biomedical Semantics, 2018; 9:11.
- [11] **Zagoruyko NG.** Cognitive data analysis [In Russian]. Novosibirsk: “Geo” Publisher, 2013.
- [12] **Barsegyan AA, Kupriyanov MS, Holod II, Tess MD, Elizarov SI.** Data and Process Analysis [In Russian]. St. Petersburg: BHV-Petersburg, 2009.
- [13] **Ganter B, Wille R.** Conceptual scaling. In: F Roberts (ed.): Applications of Combinatorics and Graph Theory to the Biological and Social Sciences. Springer-Verlag New York, 1989: 139-167.
- [14] **Ignatov DI.** Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields. In: P Braslavski, N Karpov, M Worrington, Y Volkovich, DI Ignatov (eds.): Information Retrieval (Revised Selected Papers 8th Russian Summer School, RuSSIR 2014 (Nizhniy Novgorod, Russia, 2014, August 18-22, 2014). Springer International Publishing, 2015: 42-141.
- [15] **Getmanova AD.** Logic. Advanced course [In Russian]. Moscow: KNORUS Publisher, 2016.
- [16] **Ivin AA, Nikiforov AL.** Logic dictionary [In Russian]. Moscow: VLADOS Publisher, 1997.
- [17] **Lammari N, Metais E.** Building and maintaining ontologies: a set of algorithms. Data & Knowledge Engineering, 2004; 48(2): 155-176.
- [18] **Lammari N, du Mouza C, Metais E.** POEM: an Ontology Manager based on Existence Constraints. In: S.S. Bhowmick, J. Küng, R. Wagner (eds.): Database and Expert Systems Applications. Proc. 19th Int. Conf. DEXA 2008 (Turin, Italy, September 1-5, 2008). Lecture Notes in Computer Science, vol. 5181. Springer-Verlag Berlin Heidelberg, 2008: 81-88.
- [19] **Pronina VA, Shipilina LB.** Using the relationships between attributes to build domain ontology [In Russian]. Control Science, 2009; 1: 27-32.
- [20] **Samoilov DE, Semenova VA, Smirnov SV.** Incomplete data analysis of for building formal ontologies [In Russian]. Ontology of designing. 2016; 6(3): 317-339. DOI: 10.18287/2223-9537-2016-6-3-317-339.

- [21] **Samoilov DE., Semenova VA, Smirnov SV.** Defuzzification of the initial context in Formal Concept Analysis. In: V Fursov, Y Goshin, D Kudryashov (eds.): Information Technology and Nanotechnology ITNT 2019: Data Science. Proc. of the Data Science Session at the V Int. Conf. on Information Technology and Nanotechnology (Samara, Russia, May 21-24, 2019). CEUR Workshop Proceedings, 2019; 2416: 1-9. DOI: 10.18287/1613-0073-2019-2416-1-9.
- [22] **Samoilov DE., Semenova VA, Smirnov SV.** Multilevel recursive model of properties existence constraints in machine learning. Journal of Physics: Conf. Series 1096 (2018) 012096. DOI: 10.1088/1742-6596/1096/1/012096
- [23] **Smirnov SV.** Two methodologies for formal concepts inference: when and how they should work together [In Russian]. In: Znaniya – Ontologii – Teorii: Materials of VII int. conf. – Novosibirsk: Sobolev Institute of mathematics of the Siberian Branch of RAS, Novosibirsk State University, 2019: 355-363.
- 

## About the authors

**Semenova Valentina Andreevna** (b. 1994). Bachelor in Applied Mathematics and Computer Science (Samara State Aerospace University, 2015) and Master in Mechanics and Mathematical Modeling (Samara University, 2017). Post-graduate student of Samara State Technical University in the direction of Informatics and Computer Engineering, junior researcher at Institute for the Control of Complex Systems of Russian Academy of Sciences. Author and co-author of 19 scientific papers in the field of data mining and software engineering. ORCID: 0000-0002-0557-3890; Author ID (Scopus): 57204366624. [queenbfjr@gmail.com](mailto:queenbfjr@gmail.com)

**Sergey Victorovich Smirnov** (b. 1952) graduated from the Kuybyshev Aerospace Institute (Kuibyshev-city) in 1975, D. Sc. Eng. (2002). Chief Researcher at Institute for the Control of Complex Systems of Russian Academy of Sciences, professor at Povolzhskiy State University of Telecommunication and Informatics. He is a member of Russian Association of Artificial Intelligence and International Association for Ontology and its Applications. He is a co-author of more than 170 publications in the field of applied mathematics, complex systems simulation and development knowledge based decision support systems in control and management. ORCID: 0000-0002-3332-5261; Author ID (RSCI): 17628; Author ID (Scopus): 57205017287. [smirnov@iccs.ru](mailto:smirnov@iccs.ru)

---

*Received August 10, 2020. Revised September 01, 2020. Accepted September 14, 2020.*

---