

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Самарский государственный технический университет»,
Федеральное государственное бюджетное учреждение науки
Самарский федеральный исследовательский центр
Российской академии наук

На правах рукописи

Семенова Валентина Андреевна

**МЕТОДИЧЕСКОЕ И АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ
ПОДГОТОВКИ КОНТЕКСТА ДЛЯ ВЫВОДА ФОРМАЛЬНЫХ
ПОНЯТИЙ В ОНТОЛОГИЧЕСКОМ АНАЛИЗЕ ДАННЫХ**

2.3.1 – Системный анализ, управление и обработка информации, статистика

Диссертация на соискание учёной степени
кандидата технических наук

Научный руководитель:
доктор технических наук, доцент
Смирнов Сергей Викторович

Самара 2024

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1 ИССЛЕДОВАНИЕ ОСНОВАНИЙ	
ОНТОЛОГИЧЕСКОГО АНАЛИЗА ДАННЫХ	12
1.1 Формальные онтологии для представления знаний	12
1.2 Критический разбор методик вывода формальных онтологий	14
1.2.1 Анализ формальных понятий	14
1.2.2 Анализ ограничений существования свойств	17
1.3 Онтологический анализ данных как комбинация методик вывода формальных онтологий	20
1.3.1 Обоснование единства логической природы существующих методик	21
1.3.2 Анализ необходимости и возможности комбинирования методик	23
1.4 Выводы по разделу 1	25
2 СТРУКТУРИРОВАНИЕ ИСХОДНОЙ ИНФОРМАЦИИ	
ДЛЯ ОНТОЛОГИЧЕСКОГО АНАЛИЗА	26
2.1 Формирование системы измеряемых свойств	26
2.1.1 Таблица «пара свойств – свойства»	26
2.1.2 Концептуальное шкалирование при построении системы измеряемых свойств	27
2.2 Обобщение стандартного протокола измерений – таблицы «объекты-свойства»	32
2.2.1 Морфологический анализ протокола измерений предметной области	32
2.2.2 Концептуальное шкалирование в обобщенной таблице «объекты-свойства»	34
2.2.3 Обоснование выбора логической платформы для построения нестроого формального контекста	35

2.2.4	Построение исходного нестроого контекста на базе V^{TF} -логики.	38
2.3	Выводы по разделу 2	40
3	ФОРМИРОВАНИЕ РАБОЧЕГО КОНТЕКСТА ДЛЯ ВЫВОДА ФОРМАЛЬНЫХ ПОНЯТИЙ	41
3.1	Дефаззификация эмпирического контекста	41
3.2	Разработка метода нормализации рабочего контекста	44
3.2.1	Критический анализ существующих подходов к нормализации контекста	45
3.2.2	Эвристика и механизм нормализации контекста	48
3.2.3	Метод нормализации контекста	52
3.2.4	Сравнительное исследование вариантов определения агрегированного показателя достоверности у денонсируемых базовых семантических суждений	58
3.3	Реструктуризация системы измеряемых свойств для ускорения нормализации рабочего контекста	61
3.3.1	Выявление классов эквивалентности и расширенных ограничений существования на множестве измеряемых свойств	61
3.3.2	Исследование эффективности реструктуризации системы измеряемых свойств	65
3.4	Выводы по разделу 3	72
4	РАЗРАБОТКА ПРОГРАММНОЙ ЛАБОРАТОРИИ ДЛЯ ОНТОЛОГИЧЕСКОГО АНАЛИЗА ДАННЫХ	73
4.1	Выбор программной платформы	77
4.2	Архитектурное проектирование лаборатории	78
4.2.1	Функциональность и архитектура	78

4.2.2	Концептуальная модель предметной области задачи онтологического анализа данных	81
4.2.3	Организация интерфейса с пользователем	83
4.3	Разработка ключевых алгоритмов обработки данных	84
4.3.1	Поддержка формирования и редактирования системы измеряемых свойств	85
4.3.2	Алгоритмы реструктуризации системы измеряемых свойств . .	92
4.3.3	Алгоритмизация работы с таблицей «объекты-свойства» и формирование исходного контекста	94
4.3.4	Опорный алгоритм нормализации рабочего контекста	101
4.4	Демонстрация работы программного комплекса	104
4.5	Выводы по разделу 4	107
5	АПРОБАЦИЯ РАЗРАБОТАННЫХ МЕТОДОВ И СРЕДСТВ	115
5.1	Поддержка структурного анализа и синтеза технических решений . .	115
5.1.1	Ограниченная сочетаемость вариантов реализации различных функционально-структурных подсистем	118
5.1.2	Апология запретов и преимущества ограничений существования	124
5.1.3	Структурный синтез технического решения как семантическая идентификация объекта проектирования . . .	126
5.2	Поддержка коллективного построения когнитивных карт	128
5.3	Особенности задачи портретной идентификации человека в криминалистике	132
5.4	Выводы по разделу 5	134
	ЗАКЛЮЧЕНИЕ	135
	СПИСОК СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ	136
	СПИСОК ЛИТЕРАТУРЫ	138

Приложение А	Эксперимент в области портретной идентификации людей	151
Приложение Б	Использование неравенства Хёфдинга для определения количества статистических испытаний, обеспечивающих качественную оценку среднего исследуемого показателя	156
Приложение В	Результаты имитационных статистических экспериментов для определения среднего относительного уменьшения количества сущностей в редуцированном представлении ограничений существования свойств	159
Приложение Г	Свидетельство о государственной регистрации программы для ЭВМ	162
Приложение Д	Контекстные меню рабочих листов программной лаборатории для онтологического анализа данных	165
Приложение Е	Акты внедрения результатов исследования	173

ВВЕДЕНИЕ

Актуальность работы. Практика показывает, что сложные информационные системы результативны лишь при надежном и согласованном представлении их предмета. Систематизация, разработка и использование таких информационных моделей составляют современное содержание онтологического подхода в вычислительных науках [1, 2], который развивается в русле парадигмы представления и использования знаний в искусственном интеллекте [3-5].

Онтологии являются одной из ведущих парадигм структурирования информационного контента, преимуществом их использования в качестве средства познания является системный подход к изучению предметной области. Всё более широкое применение термин онтологии находит сейчас в «мягких» науках – менеджменте, экономике, электронном бизнесе, где структурирование информации представляет значительную сложность. Также онтологии могут явиться эффективным инструментом формирования корпоративного информационного поля.

В процессе построения онтологий применяют методы инженерии знаний, или онтологический инжиниринг, получивший развитие в работах В.Ф. Хорошевского, Т.А. Гавриловой, В.В. Грибовой, Ю.А. Загорулько, Л.В. Массель, П.О. Скобелева, С.В. Смирнова, Т. Gruber, N. Guarino, J. Sowa и др.

Одним из путей построения формальных онтологий в информатике является их «вывод» из доступных эмпирических данных о предметной области (ПрО), или онтологический анализ данных (ОАД) [2]. Данные рассматриваются при этом как результат измерений (в самом широком понимании) свойств объектов актуальной ПрО и сводятся в стандартизованные таблицы «объект-свойство».

Наиболее результативные методы этого направления опираются на анализ формальных понятий (АФП) [6-8], основоположником которого является R. Wille, развиваемый в трудах Е.Е. Витяева, С.О. Кузнецова, Д.И. Игнатова, В.В. Быковой, С.А. Евтушенко, Д.В. Семёновой, а также зарубежных авторов В. Ganter, S. Ferré, R. Belohlávek и др. Менее известна содержательно иная методика онтологического инжиниринга, основанная на концепции ограничений существования свойств (ОСС) [9, 10], описываемая в работах N. Lammari, E. Metais, В.А. Прониной и др.

На основе АФП можно построить только «скелет» онтологии без учета множества факторов: структурных отношений, реалий накопления эмпирической информации, априорной информации, известной исследователю. АФП исходит из наличия однозначного формального контекста (ФК), описывающего ПрО, формирование которого по реальным данным представляет из себя отдельную задачу. Предпринятая в работах В.А. Прониной и Л.Б. Шипилиной попытка совместить указанные подходы малопродуктивна, поскольку сведена лишь к рекомендации проверять ФК насчет соблюдения ОСС.

Актуальность проблемы гибридизации дедуктивных методов вывода ФП при учете реалий накопления эмпирической информации определила цель и задачи диссертации.

Цель диссертационного исследования: разработка моделей, методов, алгоритмического обеспечения и программных средств подготовки формального контекста для вывода множества формальных понятий, формирующих остов онтологического описания интересующей субъекта предметной области, на основе неполных и противоречивых эмпирических данных и учёта ограничений существования свойств у объектов предметной области.

Диссертационная работа содержит решение следующих задач:

1. Анализ существующих методик вывода формальных понятий, необходимости и возможности их гибридизации.
2. Разработка моделей и методов структурирования гипотетических представлений и эмпирической информации о ПрО, включая метод интерпретации и консолидации объектно-признаковых данных с использованием многозначной логики для построения нестроого *исходного* ФК.
3. Разработка методики построения логически однозначного *рабочего* ФК, включающей метод учёта ОСС.
4. Разработка алгоритмического обеспечения подготовки ФК для вывода ФП и программных средств онтологического анализа данных.
5. Апробация разработанных методов и средств при решении технических и других задач.

Объект исследования – процесс первичной обработки объектно-признаковых данных, формирующий ФК для вывода ФП.

Предмет исследования – способы учёта при формировании ФК реалий накопления эмпирической информации и априорных гипотез субъекта о понятийном описании интересующей его ПрО.

Методы исследования. Для решения поставленных задач были использованы методы теории множеств, бинарных отношений, классической и математической многозначной логики, имитационного статистического моделирования, многокритериальной оптимизации, объектно-ориентированного проектирования.

Достоверность результатов диссертационного исследования подтверждается корректным применением теоретического аппарата использованных методов исследования, подтверждается математически корректными формулировками и выводами, большим количеством компьютерных экспериментов и апробацией предложенных в диссертации моделей, методов и алгоритмов для решения практически важных задач.

Научная новизна заключается в следующем:

1. Доказано единство логической природы существующих подходов к выводу формальных понятий, отличающееся обращением к их гипотетико-дедуктивным схемам, и предложен механизм их совмещения, отличающийся использованием нестрогих формальных контекстов, что в совокупности открыло возможность гибридизации дедуктивных методик вывода формальных понятий.

2. Предложены модели и методы структурирования эмпирической информации, отличающиеся использованием векторной логики, что дало возможность отразить реалии накопления эмпирических данных о предметной области.

3. Для нормализации множества свойств каждого объекта в контексте предложена эвристика, отличающаяся использованием сведений о достоверности базовых семантических суждений, образующих контекст, и последовательным приближением к нормальному множеству свойств объекта, что позволило разработать результативный метод нормализации контекста.

4. Предложен метод реструктуризации системы измеряемых свойств, отличающийся выявлением расширенных экзистенциальных отношений на классах эквивалентности множества измеряемых свойств, что обеспечило существенное сокращение размерности его описания и, как следствие, ускорение процесса нормализации контекста.

5. Разработано алгоритмическое обеспечение подготовки контекста для вывода формальных понятий, отличающееся использованием предложенных моделей и методов, что обеспечило корректную обработку неполных и противоречивых эмпирических данных многомерных измерений и наблюдений в предметной области с учётом ограничений существования свойств.

Теоретическая значимость работы заключается в обосновании единства гипотетико-дедуктивных схем сосуществующих подходов к выводу формальных понятий и разработке гибридной методологии подготовки контекста для вывода формальных понятий в онтологическом анализе данных.

Практическая значимость определяется реализацией программных инструментов первичной обработки эмпирических объектно-признаковых данных для построения формального контекста.

Результаты работы применены при выполнении следующих НИР по госзаданиям ИПУСС РАН – СамНЦ РАН:

- «Интерсубъективное управление инновационным развитием социотехнических объектов с применением онтологических моделей ситуаций» гос. рег. № АААА-А16-116040410061-0;

- «Разработка и исследование методов и средств аналитического конструирования, компьютерного представления знаний, вычислительных алгоритмов и мультиагентных технологий в задачах оптимизации процессов управления сложными системами» рег. № АААА-А19-119030190053-2.

Результаты диссертационной работы использованы при выполнении научно-исследовательских и выпускных квалификационных работ по образовательным программам бакалавриата и магистратуры Самарского национального исследовательского университета имени академика С.П. Королёва на кафедре программных

систем. Практическое использование подтверждено актами ООО «Открытый код» (г. Самара) и ООО «НПК «Сетецентрические платформы» (г. Самара).

Соответствие диссертации паспорту научной специальности 2.3.1: п. 1. Теоретические основы и методы системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта (пп. 1-3 новизны); п. 3. Разработка критериев и моделей описания и оценки эффективности решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта (пп. 2, 3 новизны); п. 4. Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта (пп. 2, 3 новизны); п. 5. Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта (пп. 4 новизны); п. 13. Методы получения, анализа и обработки экспертной информации, в том числе на основе статистических показателей (пп. 2, 3 новизны).

Личный вклад соискателя. Постановка и обсуждение возможных путей решения задач исследования осуществлялась совместно с научным руководителем. Результаты, составляющие научную новизну и выносимые на защиту, получены лично автором, либо в неделимом соавторстве. Конфликт интересов с соавторами отсутствует. Автором выполнено проектирование архитектуры и функциональности программной лаборатории для онтологического анализа данных, проведена апробация программных средств при решении задач в различных научных и технических областях.

Апробация результатов. Основные положения, результаты и выводы диссертационного исследования докладывались на Всероссийской конф. с международным участием «Информационные и математические технологии в науке и управлении» (Иркутск, 2021, 2023 гг.), Международной конф. «Информационные технологии и нанотехнологии» (Самара, 2015, 2018, 2019, 2022 гг.), XI Международной н.-т. и н.-метод. конф. «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (Санкт-Петербург, 2022 г.), Международной конф. «Марчу-

ковские научные чтения» (Новосибирск, 2021 г.), IX Всероссийской н.-т. конф. с международным участием «Информационные технологии интеллектуальной поддержки принятия решений» (Уфа, 2021 г.), XII Всероссийской н.-т. конф. аспирантов, студентов и молодых ученых «Информатика и вычислительная техника», (Ульяновск, 2020 г.), Международной н.-т. конф. «Перспективные информационные технологии» (Самара, 2016, 2017, 2018, 2020 гг.), Международной конф. «Проблемы управления и моделирования в сложных системах» (Самара, 2017, 2018, 2019 гг.), VI Международной н.-т. конф. «Открытые семантические технологии проектирования интеллектуальных систем» (Минск, 2016 г.).

Положения, выносимые на защиту:

1. Гибридизация существующих дедуктивных методик вывода формальных понятий, основанная на единой логической природе данных методик и отказе от части эмпирической информации.

2. Прагматические модели системы измеряемых свойств и результатов многомерных измерений предметной области. Методика использования векторной логики для интерпретации и консолидации эмпирической информации.

3. Эвристический метод нормализации нестроого формального контекста.

4. Метод реструктуризации системы измеряемых свойств.

5. Архитектурные решения и алгоритмическое обеспечение программной лаборатории для онтологического анализа данных.

Публикации. Результаты исследования опубликованы в 8 статьях в журналах из Перечня ВАК. 5 статей опубликованы в изданиях, индексируемых в международной информационной базе Scopus, 12 в других изданиях. Получено свидетельство о государственной регистрации программы для ЭВМ.

Структура и объём диссертации. Диссертационная работа состоит из введения, пяти разделов, заключения, списка сокращений и обозначений, списка литературы и приложений. Текст диссертации изложен на 177 страницах машинописного текста, включающих 57 рисунков, 6 таблиц; 27 страниц занимают 6 приложений.

1 ИССЛЕДОВАНИЕ ОСНОВАНИЙ ОНТОЛОГИЧЕСКОГО АНАЛИЗА ДАННЫХ

В разделе описываются существующие методики вывода онтологий, обосновывается единство их логической природы и проводится анализ возможности их комбинирования в рамках онтологического анализа данных. Приведенные результаты отражены в публикациях автора [11-16].

1.1 Формальные онтологии для представления знаний

В философии онтология не что иное, как составная часть данной науки, посвященная анализу и классификации того, что есть. В информатике онтология описывает некоторую отрасль науки, ПрО. Поэтому в силу множественности наук и ПрО, когда каждая из них имеет свою собственную или даже несколько конкурирующих терминологий, здесь в противоположность философии приобретает смысл употребление множественного числа для термина, т.е. говорят об «онтологиях». Более того, различают лингвистические и формальные онтологии, где последние наследуют парадигму моделей и методов представления знаний, разработанных в искусственном интеллекте [5].

Формальная онтология описывает ПрО с помощью стандартизирующей терминологии – словаря, а также неоднородных связей между определенными в нем понятиями. Моделирующими примитивами онтологической спецификации служат классы, отношения, функции и аксиомы, что в некотором смысле сближает эти структуры представления знаний и соответствующие компьютерные ресурсы с алгебраическими системами А.И. Мальцева [17].

[Канту приписывается фраза: «*Anschauung ohne Begriffe ist blind*» – «воззрение без понятий слепое». В этом смысле онтологии как информационные модели понятийных структур субъекта определяют пределы его ориентирования в пространствах бытия, горизонты истолкования действительности и, следовательно, определяют его предпонимание».

В сжатой трактовке онтология – это общая, разделяемая коллективом субъектов концептуальная информационная модель ПрО, в «освоение» (проектирова-

ние, управление и т.п.) которой эти субъекты вовлечены. Для этих субъектов онтология действует как средство коммуникации и обладает нормативной ролью.]¹

На рисунке 1.1 представлены три основных пути разработки онтологий, существующих на сегодняшний день [18]:

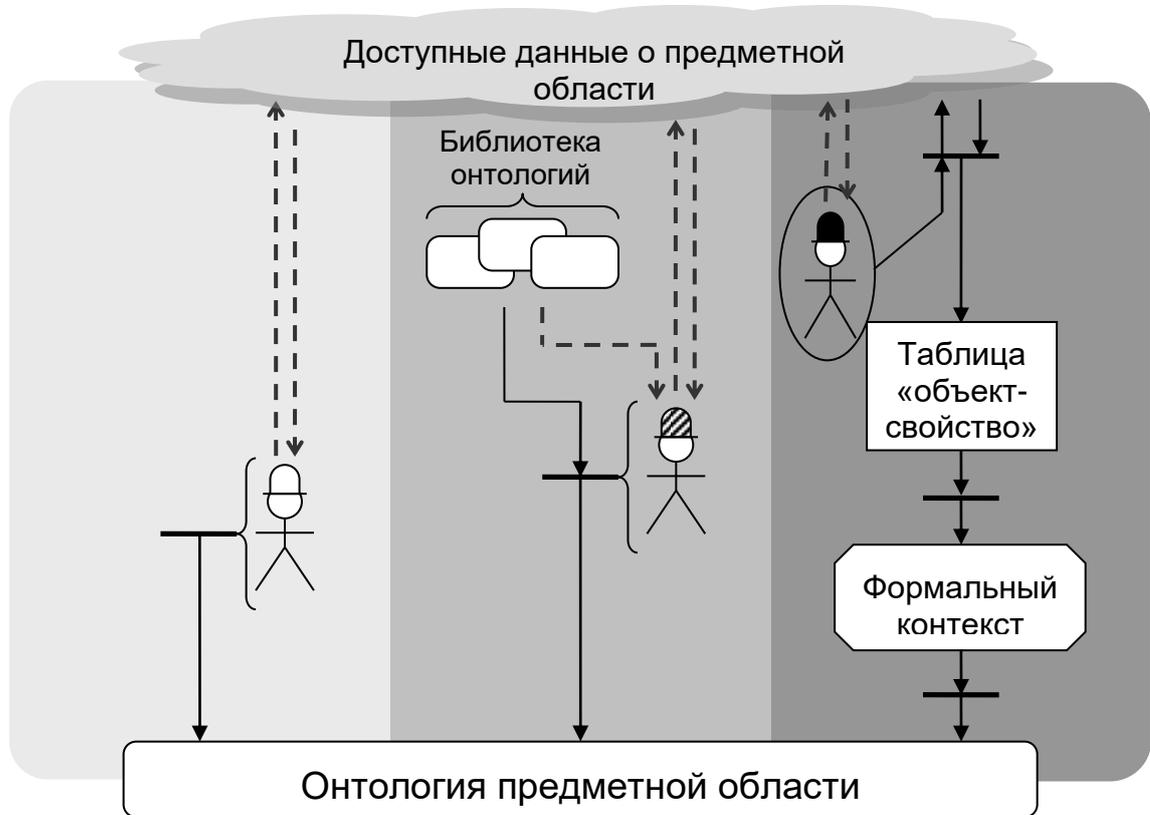


Рисунок 1.1 – Способы построения онтологий

- наиболее используемый связан с прямой формализацией опыта и знаний экспертов, которые с помощью компьютерных языковых средств либо автоформализуют свои представления о предметной области (ПрО), либо фиксируют их с помощью инженера по знаниям;
- актуальные онтологии могут синтезироваться в результате человеко-машинных процедур композиции/декомпозиции апробированных формальных онтологий разного уровня и направленности [19];
- третий путь связан с автоматическим «выводом» формальной онтологии из доступных данных. Эти данные рассматриваются как результат измерений объектов исследуемой ПрО и сводятся в стандартизованные таблицы «объекты-

¹ Цитировано по источнику [18, стр. 19]

свойства» [20], анализ которых приводит к выявлению понятийной структуры ПрО.

В данной работе изучается и развивается третье направление построения онтологий – интеллектуальный ОАД, – наиболее результативные методы которого опираются на ветвь теории решёток – анализ формальных понятий (АФП) [6], широко использующийся при работе с онтологиями.

1.2 Критический разбор методик вывода формальных понятий

1.2.1 Анализ формальных понятий

Как отмечено в [21], АФП – один из теоретически хорошо обоснованных и активно развиваемых методов анализа данных, в котором отражен классический подход к понятию как к фундаментальному эпистемологическому элементу, определяемому объёмом и содержанием. Имеется множество систем информационного моделирования, имеющих в своей основе или интегрирующих АФП, например: ConExp, Galicia, ToscanaJ, Python FCA Tool, gV и др. В [21] также отмечается, что метод широко применяется в различных областях науки и техники: медицине, генетике, психологии, анализе литературных источников, реинжиниринге программного обеспечения, проектировании и т.д. Во многом успех АФП обусловлен его пригодностью для вывода формальных онтологий из экспериментальных данных, представляющих ПрО, или, проще, из данных измерения ПрО.

Моделью исходных данных в АФП служит многозначный *формальный контекст* (ФК), который фактически является стандартным протоколом представления эмпирической информации известным как таблица «объекты-свойства» (ТОС) [20, 22]:

$$(G, M, D, I), \quad (1.1)$$

где $G = \{g_i\}_{i=1,\dots,r}$, $r = |G| \geq 1$ – конечное множество наблюдавшихся объектов: $G \subseteq G^*$, G^* – всё гипотетически мыслимое множество объектов зондируемой ПрО; $M = \{m_j\}_{j=1,\dots,s}$, $s = |M| \geq 1$ – конечное множество измеренных у объектов свойств (признаков); $D \subseteq D^* = \cup_{j=1,\dots,s} D_j$, D_j – домен свойства m_j (в технике ему сопостав-

ляется динамический диапазон *процедуры измерения* j -го свойства); I – тернарное отношение между G , M и D ($I \subseteq G \times M \times D$), определённое для всех пар из $G \times M$.

Путём *концептуального шкалирования* [23] и/или учета ограниченности динамических диапазонов процедур измерения модель (1.1) преобразуется в *логически однозначный* ФК, ключевую структуру АФП:

$$(G, M, I), \quad (1.2)$$

где I – *бинарное* соответствие между G и M , т.е. $I \subseteq G \times M$, или $I: G \times M \rightarrow \{\text{True}, \text{False}\}$.

Соответствие I фактически представляет собой совокупность оценок истинности *базовых семантических суждений* (БСС) о ПрО: «объекту g_i присуще свойство m_j ». Классический АФП исходит из того, что $b_{ij} \in \{\text{Истина}, \text{Ложь}\}$, и формальная онтология выводится из ФК в результате специального анализа соответствия Галуа, индуцируемого таким однозначным соответствием I между булеанами множеств G^* и M . В нечетком АФП $b_{ij} \in [0, 1]$.

АФП определяет, что для любого подмножества $X \subseteq G$ объектов множество $X' = \{m \mid m \in M, \forall g \in X : gIm\}$ является набором их общих атрибутов, а для $Y \subseteq M$ множество $Y' = \{g \mid g \in G, \forall m \in Y : gIm\}$ составляют объекты, которые обладают всеми атрибутами из Y . Анализ соответствия Галуа между булеанами 2^G и 2^M позволяет выявить *формальные понятия* (ФП) – *объектно-признаковые бикластеры* вида (X, Y) , где $X = Y'$, $Y = X'$, $X \subseteq G$ – *экстенционал*, или *объем* формального понятия, а $Y \subseteq M$ – *его интенционал*, или *содержание*.

Множество всех понятий контекста K обозначается как $B(K) = B(G, M, I) = \{(X, Y) \mid X \subseteq G, Y \subseteq M, X = Y', Y = X'\}$.

Как отмечено в [2], каждому понятию из $B(K)$ соответствует максимальная по вложению единичная подматрица матрицы инцидентности, описывающей соответствие I . Возможна и графическая интерпретация формального понятия: если формальный контекст K представлен двудольным графом, то одному понятию в этом контексте соответствует максимальный по вложению полный подграф двудольного графа. На рисунке 1.2 приведен пример формального контекста.

Реки	Горная_система_ в истоке	Впадает_ в реку	Впадает_ в море	Суммарная_мощ- ность_ГЭС	Год_начала_эко мониторинга
	<i>String</i>	<i>String</i>	<i>String</i>	<i>Double</i>	<i>Integer</i>
Ангара	X	X		X	
Белая	X	X			X
Волга			X	X	
Москва		X			X

Рисунок 1.2 – Матричное представление формального контекста, где вложенная единичная подматрица указывает одно из понятий в ФК «Реки»

Согласно [2], АФП предлагает формализацию того, что рассматривают в аналитической философии и логике, исследуя смысл имен (символов), и отражают в известной схеме, получившей название «семантического треугольника». Формально-математическое содержание АФП можно охарактеризовать как прикладной анализ соответствия Галуа, устанавливаемого отношением I между подмножествами множеств G и M (рис. 1.3). Интересующее нас конструирование структуры обобщающих абстрактных понятий «над» объектно-признаковой моделью связано с изучением отношения порядка на множестве всех понятий данного ФК.

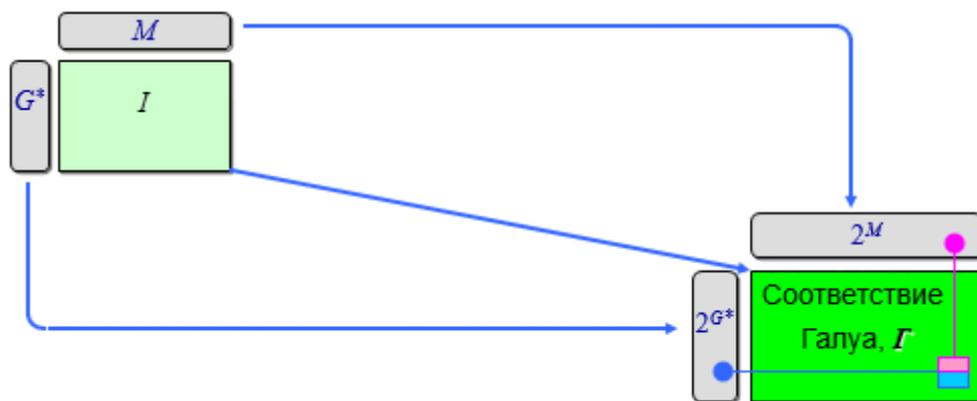


Рисунок 1.3 – Переход от ФК к соответствию Галуа

ФП консолидируют информацию обо всех объектах с одинаковым составом свойств и состоят в отношении частичного порядка по вложенности составов: $(X_1, Y_1) \leq (X_2, Y_2)$, если $Y_1 \supseteq Y_2$ (или, равно, $X_1 \subseteq X_2$), – образуя на множестве всех понятий ФК замкнутую решётку $(B(K); \leq)$, называемую *решёткой формальных понятий* (РФП) [6, 24].

Фундаментальная роль такого анализа (и в целом АФП) состоит в том, что указанные его результаты соответствуют в философии и классической логике определению *понятия* и *отношению обобщения* на множестве понятий [25, 26]. Поэтому совокупность выявленных формальных понятий и связывающее их бинарное отношение частичного порядка « \leq » («*is a*») определяет, по меньшей мере, «скелет» *формальной онтологии* исследуемой ПрО.

Важно отметить, что минимальное понятие, содержащее объект $g \in G$ в своем объеме, есть $\gamma g = (\{g\}'', \{g\}')$, а $\mu t = (\{t\}', \{t\}'')$ – максимальное понятие включающее атрибут $t \in M$ в своем содержании.

Согласно [2], решетка понятий может быть визуализирована с помощью ациклического размеченного орграфа – диаграммы Хассе, узлы которого представляют понятия, а дуги, изображающие отношение порядка, направлены из каждого узла лишь к его «соседям снизу». Таким образом, пути, исходящие из узла-понятия ведут ко всем узлам-подпонятиям. Для каждого $g \in G$ узел диаграммы, представляющий понятие γg , маркируется идентифицирующим g выражением (именем объекта) «снизу», а $\forall t \in M$ узел-понятие μt помечается именем атрибута t «сверху» (рис. 1.4).



Рисунок 1.4 – Диаграмма Хассе для формального контекста «Реки»

1.2.2 Анализ ограничений существования свойств

С другой стороны, существует подход к выводу онтологий, где у исследователя в качестве исходных данных есть только ограничения на появление объектов

с определённым набором свойств, измеряемых у объектов рассматриваемой ПрО [9, 10]. То есть, если у исследователя имеется точка зрения на то, какие объекты есть в реальной действительности, выражающаяся в описании того, как могут сочетаться свойства объекта, который существует в действительности, то из этой информации можно построить онтологию. Другими словами, если исходить из того, что в определённой ПрО возможны только определённые отношения между свойствами, то тогда можно сказать, какие понятия могут описывать реальность, где такие зависимости есть.

Суть такой концепции в том, что исходно дано два бинарных отношения 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. во введенных обозначениях $C(m_j, m_k) \leftrightarrow \forall g \in G: m_j \in \{g\}' \rightarrow m_k \in \{g\}'$. Отношение обусловленности рефлексивно, несимметрично и транзитивно (рисунок 1.5а);

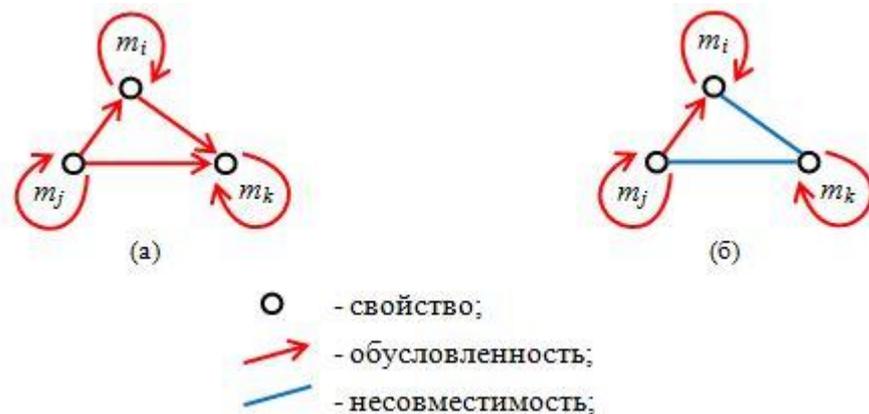


Рисунок 1.5 – Примеры экзистенциальных зависимостей между свойствами $m_i, m_j, m_k \in M$ объекта $g \in G$

- *несовместимость* $E: M \times M \rightarrow \{\mathbf{True}, \mathbf{False}\}$, 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 (рисунок 1.5б).

Сосуществование экзистенциальных отношений E и C регулируется следующими двумя законами:

- естественный *запрет* одновременного включения в E и C любой пары элементов множества-основы M :

$$\forall x, y \in M, x \neq y: \neg((E(x, y) = C(x, y)) \vee (E(x, y) = C(y, x)));$$

- несовместимость характеризуется «*транзитивностью относительно обусловленности*» (E -транзитивностью) (рисунок 1.5б):

$$\forall x, y, z \in M, x \neq z, y \neq z: C(x, y) \wedge E(y, z) \rightarrow E(x, z). \quad (1.3)$$

Совокупность измеряемых свойств и отношений между ними, которая для краткости названа системой измеряемых свойств (СИС), по сути кодирует гипотетическое представление субъекта-исследователя о ПрО и представляется кортежем-тройкой (M, C, E) . Естественной моделью и наглядным представлением СИС (M, C, E) является граф (см. рисунок 1.6).

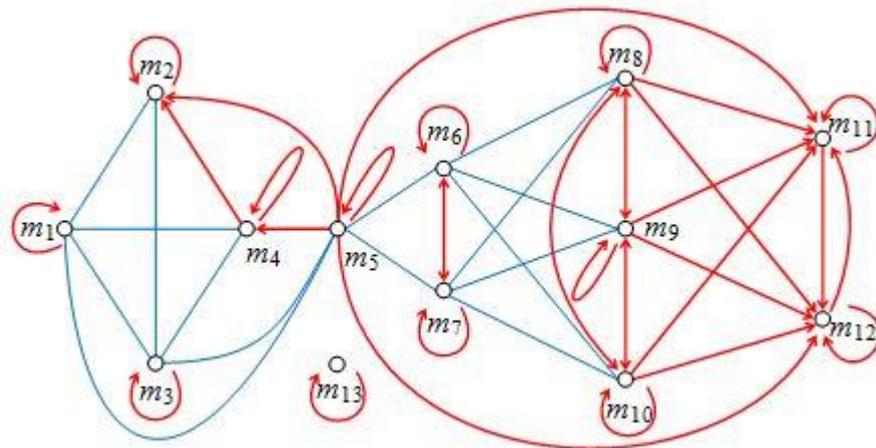


Рисунок 1.6 – Теоретический пример системы измеряемых свойств в виде графа

Тогда формирование онтологии ПрО из набора характеризующих ее свойств (т.е. свойств объектов ПрО) и набора ограничений существования (т.е. ОСС как набора *сопряженных пар свойств*, реализующих экзистенциальные отношения обусловленности и несовместимости свойств) состоит в выявлении всех *нормальных* подмножеств актуальных свойств предметной области, 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*», что, по сути, является решёткой формальных понятий.

Такой подход вызывает интерес с точки зрения того, что, в общем случае, обработка экспериментальных данных средствами ОАД может привести к некорректному результату по причине того, что исследователь обладает априорными знаниями об ОСС, но существующие методы вывода ФП не позволяют их учитывать. Так приходим к задаче объединения рассмотренных методик в рамках ОАД.

1.3 Онтологический анализ данных как комбинация методик вывода формальных понятий

ОАД – методика обработки данных, которая применяется для автоматизации формирования понятийной структуры, описывающей представленную данными ПрО в виде неоднородных семантических сетей, которые соответствуют современным представлениям о формальных онтологиях.

ОАД базируется на АФП и предлагает новые модели и методы представления и обработки эмпирических данных о ПрО, где основной задачей исследователя становится выдвижение гипотез о свойствах объектов и затем априорное комплектование арсенала измерительных процедур (органов чувств, вербальных возможностей экспертов, искусственных сенсоров, приборов, систем и т.д.), с помощью которых интересующая его ПрО будет зондироваться. Принятая в анализе данных форма структуризации экспериментального материала, представляющего ПрО, в ОАД обобщается (см. подраздел 2.2).

ОАД ограничивает естественное стремление употреблять для оценки БСС истинностных значений, вводимых многозначными логиками лишь в том плане, что для определенного по законам выбранной логики неклассического множества должен существовать механизм аппроксимации обычным множеством. Если для

этого нет препятствий, то ОАД в своем поэтапном развертывании требует замечать соответствие I , эмпирически установленное согласно подходящей многозначной логике, его строгим приближением, предусматривая применение классического АФП на завершающей стадии обработки данных.

1.3.1 Обоснование единства логической природы существующих методик

Подход АФП подразумевает, что на объекты не накладываются какие-либо ограничения, т.е. выдвигаются гипотезы о существовании понятий с любым сочетанием свойств, и, по сути, в процессе вывода ФП проверяется, какие именно гипотезы оказались истинными согласно эмпирическому материалу. Таким образом, мы квалифицируем АФП как гипотетико-дедуктивный метод. В случае, если у объекта не обнаруживается ни одного свойства из набора измеряемых свойств, то такой объект вообще не находит отражения в итоговой РФП.

Можно утверждать, что если формирование множества измеряемых свойств M можно рассматривать как *априорный когнитивный акт* субъекта, определяющий арсенал процедур измерения исследуемой ПрО, то объекты G *независимы* от субъекта в том смысле, что отсутствуют какие-либо допущения относительно их появления в ФК (в [10] эта черта АФП охарактеризована как ориентация «на объекты»).

В случае анализа ОСС очевидно, что основанная на нём методика вывода ФП ориентирована «на свойства» [10], и поскольку за ее рамками остается генезис СИС – совокупности измеряемых свойств и отношений между ними, – то естественно выдвинуть следующее утверждение: формирование СИС можно рассматривать как *априорный когнитивный акт*, а принятие решения о выявлении всех нормальных подмножеств множества измеряемых свойств M – как *апостериорный когнитивный акт* субъекта-исследователя ПрО.

Для исследования механизма возникновения ОСС обратимся к когнитивной деятельности субъекта.

Решение об измерении некоторого свойства t_x у объектов изучаемой ПрО является рациональным следствием субъективного предположения о *наличии* в ПрО объектов, обладающих этим свойством [2, 11]. В концептуальном плане это означает выдвижение *гипотезы* о том, что рассматриваемая ПрО характеризуется множеством *гипотетических формальных понятий*, имеющих в своём содержании свойство t_x . Однако лишь для *одного* из этих гипотетических понятий t_x будет *отличительным* свойством [6, 14]. Формально это понятие может быть описано кортежем $(\{t_x\}', \{t_x\})$ (гипотеза об описании ПрО формальным понятием $(\{\emptyset\}', \{\emptyset\})$ – «ничто характеризует все» – всегда выдвигается по умолчанию). Добавление нового свойства t_y , $t_y \neq t_x$, к числу измеряемых доводит число таких гипотез до 4: $(\{\emptyset\}', \{\emptyset\})$, $(\{t_x\}', \{t_x\})$, $(\{t_y\}', \{t_y\})$, $(\{t_x, t_y\}', \{t_x, t_y\})$. При M измеряемых свойствах это количество равно $2^{|M|}$.

Именно эти $2^{|M|}$ различных *гипотетических формальных понятий* изучаются в классической методике вывода формальных понятий из контекста (1.2).

Очевидно, что каждое измеряемое свойство обуславливает *само себя* (бинарное отношение обусловленности рефлексивно [9]), а отличительные свойства $t_i \in M$ формального понятия *взаимно обусловлены*.

Таким образом, расширение множества измеряемых свойств M означает либо пополнение отличительных свойств имеющихся гипотетических понятий об изучаемой ПрО, либо появление новых подобных понятий, т.е. *изменение априорных концептуальных представлений* о ПрО. В аспекте ОСС в первом случае лишь пополняются множества взаимно обусловленных свойств. Второй случай требует специального анализа.

Априорное формирование СИС следует рассматривать как выдвижение гипотезы о концептуальной структуре изучаемой ПрО. Во-первых, формирование гипотетических понятий происходит по указанному выше правилу с обычным пополнением множества M измеряемыми свойствами. Во-вторых, введение экзистенциальных отношений на множестве измеряемых свойств M ограничивает допустимое концептуальное описание ПрО.

При этом в аспекте априорного формирования субъектом гипотетических понятий экзистенциальные отношения измеряемых свойств являются *не причиной, а следствием*.

Действительно, в классической логике имеются две и только две возможности образования новых понятий на основе уже существующих [26, 27]:

- *деление* существующего понятия с образованием замещающих его новых понятий связано с перечислением всех *непересекающихся частей* его объема на *единой основе* – признаке (измеряемом свойстве), *изменяющемся* в пределах делимого объёма (вместо деления можно рассматривать противоположное действие – *обобщение* существующего понятия [26]);
- *ограничение* понятия означает введение *нового* понятия, объем которого составляет *часть объема* исходного понятия, а содержание выделяется *дополнительным признаком* (новым измеряемым свойством), присущим только той части объектов, мыслимых в исходном понятии, которая составляет объем вводимого нового понятия.

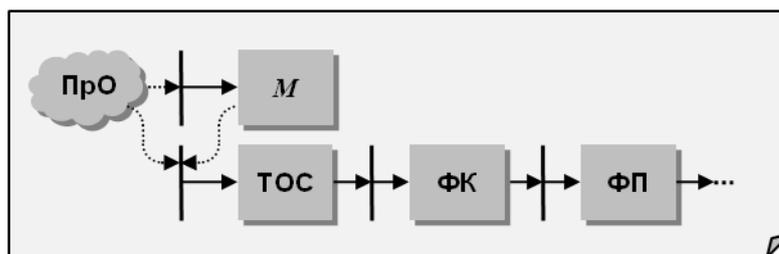
С учётом вышесказанного, констатируем, что рассмотренные методики вывода ФП имеют единую логическую основу.

1.3.2 Анализ необходимости и возможности комбинирования методик

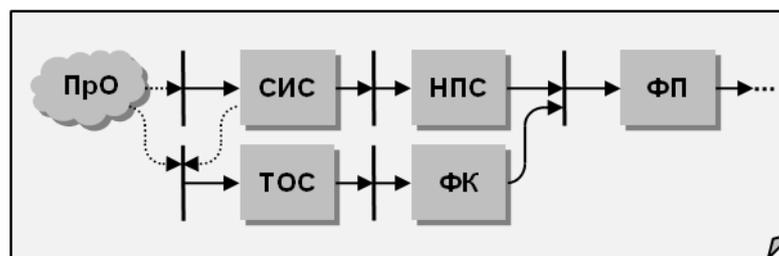
Попытки объединения двух сосуществующих подходов к выводу ФП (краткие схемы приведены на рисунке 1.7) ранее сводились к следующему:

- со стороны концепции ОСС данные о ПрО вида (1.1) и (1.2) предлагалось использовать для «фильтрации» множества нормальных подмножеств измеряемых свойств M с целью отбора таких нормальных подмножеств, которые действительно описывают состав свойств эмпирически наблюдаемых объектов ПрО [9]. Нормальные подмножества и, следовательно, соответствующие ФП, не получающие такого эмпирического подтверждения, рекомендовалось исключать из формируемой онтологии ПрО;

- со стороны АФП интеграционные моменты выглядели еще менее значимыми. Вывод ФП из (1.2) предлагалось производить лишь при условии, что каждая строка соответствия I определяет нормальное подмножество измеряемых свойств M ; конструктивным такое предложение оказывалось лишь для реинжиниринга апробированных онтологий при поступлении новых неполных эмпирических данных о ПрО [10].



а) схема анализа формальных понятий



НПС – нормальное подмножество

б) схема анализа ограничений существования

Рисунок 1.7 – Схемы методик вывода формальных понятий

При этом потребность действенного учета ОСС при выводе ФП для построения онтологий на основе многомерных измерений и наблюдений ПрО отчетливо проявилась в задачах *нечеткого* АФП, где реалии накопления эмпирической информации допускают адекватную оценку истинности БСС лишь в рамках подходящих многозначных логик, и соответствие I в ФК (1.2) становится «мягким» [28-30]. В [31] показано, что обычно используемый в подобных случаях подход – *апостериорный когнитивный акт* установки субъективного порога доверия (СПД) α к нечетким эмпирическим данным и замена нечеткой модели четкой (в данном случае – замена нечеткого соответствия I его α -сечением) – способен привести к выводу ложных понятий, катастрофически искажающих имеющиеся представления исследователя (см. Приложение А) – назовём их *априорными знаниями* – об исследуемой ПрО.

Вместе с тем, постановка задачи вывода ФП из неполной, нечеткой и даже противоречивой эмпирической информации о ПрО открывает ранее не рассматривавшуюся возможность [32, 33]:

- выделить часть эмпирической информации о ПрО, отказ от которой приведет к ФК вида (1.2), удовлетворяющему ОСС;
- среди всех возможных таких частей для конкретного ФК, выбрать и исключить из рассмотрения наименее «достоверную» (понятие достоверности раскрывается в пункте 3.2.2).

Именно в контексте реализации этого тезиса проводились исследования по развитию ОАД, как расширенной методики вывода онтологий.

1.4 Выводы по разделу 1

- 1) Обзор существующих методик по выводу онтологий из имеющихся данных позволяет сделать следующие заключения:
 - АФП ориентирован на объекты, и отсутствуют какие-либо допущения относительно их появления в формальном контексте;
 - концепция ОСС ориентирована на свойства и говорит о том, что любой объект изучаемой ПрО может обладать лишь «нормальным» подмножеством измеряемых свойств, а ФК уже только конкретизирует набор ФП о ПрО;
 - рассмотренные методики имеют единую логическую основу.
- 2) Целесообразность гибридизации методик состоит в том, что это позволит избежать ошибочных результатов в построении онтологии при имеющихся у исследователя ПрО априорных знаниях о зависимостях между свойствами объектов изучаемой ПрО.
- 3) Состоявшиеся ранее попытки объединения рассмотренных методик имеют малозначимый потенциал и требуют активного участия исследователя. Условием реализации выдвинутой идеи гибридизации методик является «мягкий» характер базовых семантических суждений о ПрО.

2 СТРУКТУРИРОВАНИЕ ИСХОДНОЙ ИНФОРМАЦИИ ДЛЯ ОНТОЛОГИЧЕСКОГО АНАЛИЗА

В данном разделе строится прагматическая модель СИС и описывается модель ОТОС, используемая для структурирования эмпирической информации. Поясняется принцип механизма концептуального шкалирования при их построении.

В подразделе 2.1 на основе ОСС формируется модель СИС, достаточная для нормализации нестроого ФК. В подразделе 2.2 на основе работ научного руководителя приводится обобщение стандартного протокола ТОС, которое дополняется обоснованием необходимости и достаточности использования для интерпретации и консолидации эмпирических объектно-признаковых данных многозначной двухаспектной векторной V^{TF} -логики. Приведенные результаты отражены в публикациях автора [11, 13, 15, 32-36].

2.1 Формирование системы измеряемых свойств

2.1.1 Таблица «сопряженная пара свойств – свойства»

Прагматическая модель СИС формируется на основе инцидентности «сопряженная пара свойств – свойства» (см. рисунок 2.1); такая структура достаточна для описания обусловленности и несовместимости пар *различных* измеряемых свойств, а, казалось бы, теряемая информация о рефлексивности обусловленности в действительности сохраняется по принципу «умолчания», что допустимо в силу «тотальной» самообусловленности свойств. Это обеспечивает достаточный когнитивный и «объясняющий» потенциал такого представления для субъекта ОАД, а также эффективную реализацию соответствующих компьютерных ресурсов.

Таким образом, СИС представляется кортежем (M, C_{NR}, E, H) , где:

- $C_{NR} = \{c_k\}_{k=1, \dots, t}$, $t \geq 0$ – множество пар свойств с обусловленностью, $C_{NR} \subseteq M \times M$, $|C_{NR}| \leq C_{|M|}^2$ (число сочетаний);
- $E = \{e_l\}_{l=1, \dots, u}$, $u \geq 0$ – множество пар несовместимых свойств, $E \subseteq M \times M$, $|E| \leq A_{|M|}^2$ (число размещений);

- $H = (h_{kl})_{k=1,\dots,t; l=1,\dots,u}$ – матрица инцидентности «сопряженные пары разных свойств – свойства».

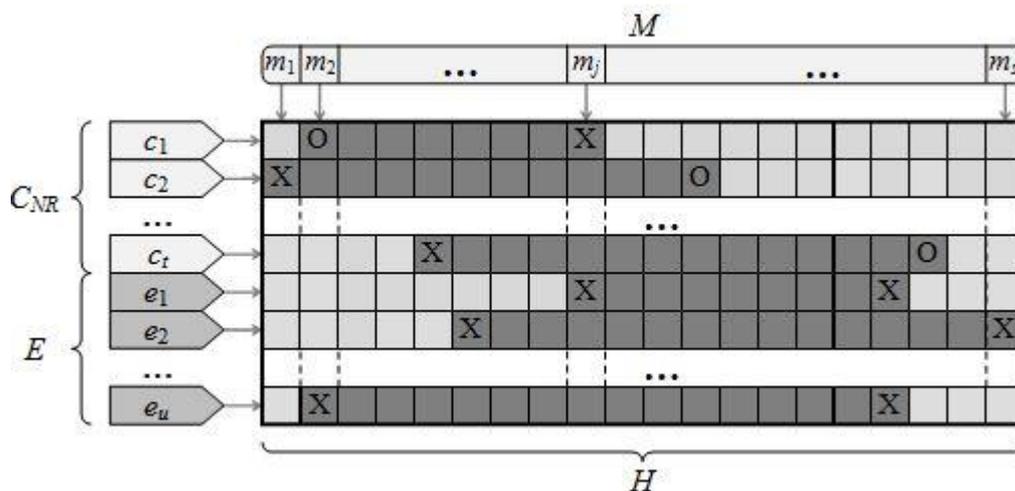


Рисунок 2.1 – Структура системы измеряемых свойств на основе двумерной матрицы:
 «X» – указатель инцидентности для обуславливающего свойства,
 «O» – указатель инцидентности для обусловленного свойства

2.1.2 Концептуальное шкалирование

при построении системы измеряемых свойств

Неформально концептуальное шкалирование измеряемого свойства представляет собой субъективное конструирование «покрытия» динамического диапазона (домена доставляемых значений) соответствующей процедуры измерения с образованием *новых* измеряемых свойств у объектов ПрО. Такое покрытие именуется *концептуальной шкалой* [37].

Пример концептуальной шкалы с дизъюнктивным «расщеплением» динамического диапазона процедуры измерения даёт таблица 2.1, являющаяся примером номинальной шкалы, или шкалы наименований, устанавливающей строго дизъюнктивное покрытие исходного домена значений шкалируемого свойства «Рост мужчины, см». Дизъюнктивное (номинальное) шкалирование свойства предполагает разделение его домена на две или, в общем случае, $n \geq 2$ непересекающихся частей и приводит к *замещению* шкалируемого свойства набором *новых несовместимых* свойств. В этом смысле результат номинального шкалирования согласно таблице 2.1 иллюстрирует рисунок 2.2а.

Таблица 2.1 – Концептуальная шкала свойства «Рост мужчины» («х» – **Истина**, пробел – **Ложь**)

Рост мужчины, см	Низкий	Средний	Высокий
< 168	х		
168-175		х	
> 175			х

Уточняющее (порядковое) шкалирование свойства устанавливается покрытием его домена двумя областями, первая из которых покрывает домен целиком, а вторая – некоторую его часть («строгая часть»). При этом у объектов исследуемой ПрО субъективно *добавляется* новое свойство, которое определяет выделенную часть домена шкалируемого свойства и *обуславливает* последнее. Обобщением такого приема является уточняющее шкалирование одновременно $n \geq 2$ свойств [10]. *Порядковую шкалу* целесообразно использовать для сохранения упорядоченности значений в домене измеряемого свойства.

Например, домен свойства «Материальное положение (МП)» может быть описан следующими выражениями (от тяжелого до благополучного) [37]:

- 1) денег не хватает даже на питание;
- 2) на питание денег хватает, но не хватает на покупку одежды и обуви;
- 3) денег хватает на одежду и обувь, но не на приобретение бытовой техники;
- 4) денег хватает на приобретение бытовой техники, но не можем купить машину;
- 5) денег хватает на всё, кроме таких дорогих приобретений как квартира, дом;
- 6) материальных затруднений не испытываем, можем приобрести квартиру, дом.

Для субъекта принятия решений естественной шкалой для такого свойства будет таблица 2.2, которая устанавливает между вновь введенными свойствами бинарное отношение *обусловленности* $C: i < k \leftrightarrow C(МП_k, МП_i)$ – см. рисунок 2.2б.

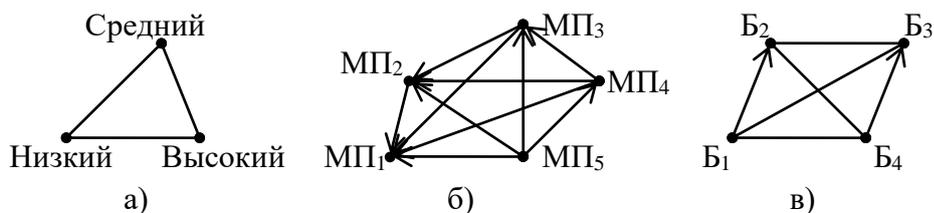


Рисунок 2.2 – Примеры отношений несовместимости и обусловленности свойств:

• - свойство; / - несовместимость; → - обусловленность

Таблица 2.2 - Шкала материального положения

\leq	МП ₁	МП ₂	МП ₃	МП ₄	МП ₅	МП ₆
1	×					
2	×	×				
3	×	×	×			
4	×	×	×	×		
5	×	×	×	×	×	
6	×	×	×	×	×	×

В настоящее время весьма популярны шкалы с разделением и порядком, описанные в [37] на примере закрытого вопроса «Чувствуете ли Вы себя в безопасности? (Б)» с вариантами ответа:

- 1) безусловно да;
- 2) скорее да;
- 3) скорее нет;
- 4) безусловно нет.

Субъективное понимание значений этого домена значений может быть выражено *двупорядковой шкалой*, доставляемой таблицей 2.3.

Таблица 2.3 - Шкала безопасности

	Б ₁	Б ₂	Б ₃	Б ₄
1	×	×		
2		×		
3			×	
4			×	×

В этом примере субъект фактически расширяет имеющиеся наряду с ТОС данные о ПрО, вводя следующие бинарные отношения между вновь вводимыми свойствами (рисунок 2.2в):

- $E = \{(B_1, B_3), (B_1, B_4), (B_2, B_3), (B_2, B_4)\}$;
- $C = \{(B_1, B_2), (B_4, B_3)\}$.

Нетрудно установить, что номинальная шкала является последствием деления понятия по объёму домена, а порядковая шкала – следствием ограничения понятия [13].

Рассмотрим подробнее эти классические операции с понятиями и покажем верность вышеприведённого утверждения.

Деление гипотетического понятия. Если свойство m_x является основанием деления гипотетического понятия, то его «изменение» в пределах делимого объема $\{\dots, m_x\}'$ («...» – замещает совокупность унаследованных свойств от более общих гипотетических формальных понятий) не может быть истолковано иначе, чем как переход в области D_x от одного подмножества значений свойств к другому при условии, что эти подмножества *не пересекаются*. Легко видеть, что субъективно построенное дизъюнктивное покрытие домена D_x : $D_x = \cup_{k=1, \dots, q} D_{xk}$, $q \geq 2$, $D_{xi} \cap D_{xj} = \emptyset$, $i \neq j$, соответствует номинальному концептуальному шкалированию свойства m_x , когда новые измеряемые свойства $\{m_{xk}\}_{k=1, \dots, q}$ с доменами $\{D_{xk}\}_{k=1, \dots, q}$ *несовместимы*.

В априорных гипотезах об изучаемой ПрО делимое понятие $(\{m\}', \{m\})$ *замещается* набором новых понятий $(\{m_i\}', \{m_i\})$, $i = 1, \dots, n$, для которых справедливо $\{m_i\}' \cap \{m_j\}' = \emptyset$ при $i \neq j$ и $\cup_{i=1}^n \{m_i\}' = \{m\}'$. Это приведет к расширению отношения обобщения на множестве гипотетических понятий. Расширение коснётся новых понятий, прямых «предков» замещенного понятия и, вообще говоря, всех «наследников» замещенного понятия.

Ограничение гипотетического понятия. Обратимся к понятию операции ограничения. В наших обозначениях ограничением гипотетического понятия $(\{m_k\}', \{m_k\})$ является ФП $(\{m_k, m_x\}', \{m_k, m_x\})$, где $m_k \in M$, m_x новое измеряемое свойство, добавляемое к множеству M : $m_x \notin M$ и $\{m_k, m_x\}' \subseteq \{m_k\}'$.

Согласно определению операции ограничения, обнаружение свойства m_x у объекта ПрО является достоверным признаком принадлежности этого объекта объему $\{m_k\}'$. Но так как все объекты в $\{m_k\}'$ обладают свойством m_k , то наличие свойства m_x у объекта обуславливает наличие у него свойства m_k .

«Ограничивающее» свойство m_x не может быть истолковано иначе, как указание на принцип выделения части объема $\{m_k\}'$ на основе значений свойства m_k . Поэтому для доменов свойств ограничиваемого понятия $D_k = D_k \cup D_k(m_x)$, где $D_k(m_x)$ – область домена измеряемого свойства m_k , значения из которой характеризуют объекты, составляющие часть объёма исходного понятия, выделяемую при

его ограничении. Это означает, что ограничение понятия $(\{m_k\}', \{m_k\})$ реализуется путём простейшего порядкового концептуального шкалирования [23] свойства $m_k \in M$, когда его домен D_k субъективно покрывается всего двумя множествами: собственно доменом D_k и его строгой частью $D_k(m_x) \subset D_k$.

Вообще говоря, можно построить *произвольное* число ограничений понятия $(\{m\}', \{m\})$, т.е. понятий вида $(\{m, m_i\}', \{m, m_i\})$, где $m_i, i = 1, \dots, n, n \geq 1$ – новые измеряемые свойства, добавляемые во множество $M, \{m, m_i\}' \subseteq \{m\}'$. Заметим, что $\bigcap_{i=1}^n D(m_i) = \emptyset$, т.к. для любого пересечения доменов добавляемых свойств может и должно быть «сконструировано» новое ограничивающее свойство.

Ограничиваемое понятие $(\{m\}', \{m\})$ *сохраняется* как часть априорных предположений об изучаемой ПрО и становится *родовым* (обобщающим) для всех его *видовых* ограничений вида $(\{m, m_i\}', \{m, m_i\})$, которые дополняют гипотетические представления об изучаемой ПрО. Расширение отношения обобщения коснётся и, вообще говоря, всех «наследников» понятия, ставшего родовым для вновь введенных гипотетических понятий.

В контексте основной цели диссертационного исследования следует отметить, что обе подлежащие гибридизации методики вывода ФП предполагают априорную познавательную деятельность субъекта, связанную с формированием гипотетических понятий, описывающих неизвестную ПрО. В «объектно-ориентированной» методике субъект предопределяет наибольшую вариативность возможных результатов вывода ФП. «Свойство-ориентированная» методика связана с субъективным сужением этой вариативности. В этом случае исследователь осуществляет детальную проработку первичных априорных гипотез посредством двух (и только двух) видов *априорного концептуального шкалирования* измеряемых свойств – номинального и порядкового [12, 39].

Как итог, построение СИС завязано на комплектовании процедур и дальнейшем концептуальном шкалировании измеряемых свойств, что *влечёт* возникновение ОСС.

2.2 Обобщение стандартного протокола измерений – таблицы «объекты-свойства»

2.2.1 Морфологический анализ протокола измерений предметной области

Общепринято, что базовой формой эмпирической информации о ПрО служит ТОС (1.1). Однако попытка ответить на вопрос о генезисе «мягкого» ФК, открывающего путь к гибридизации методик вывода ФП, [показывает, что для построения такого контекста нужна более архетипическая основа, чем обычная ТОС, т.е. такие исходные данные, которые включают информацию, с необходимостью порождающую многозначные заключения об истинности БСС ПрО. В отмечено, что конструктивный ответ о составе подобных данных дает морфологический анализ ТОС, который вскрывает реалии накопления эмпирической информации, элиминируемые стандартной структурой ТОС]²:

- многократные независимые измерения свойств у объекта из обучающей выборки. Т.е. усматривается, что каждый объект $g_i \in G$ подвергается, вообще говоря, нескольким сериям измерений с помощью всего арсенала измерительных процедур, используемых для исследования ПрО, причем всякая серия имеет ограниченную достоверность;
- использование нескольких независимых источников для оценки одного и того же свойства объекта. Т.е. принимается во внимание, что каждое актуальное для исследователя свойство $m_j \in M$ измеряется, вообще говоря, с помощью множества измерительных процедур (конгруэнтных по назначению), причем всякая процедура характеризуется степенью доверия к ее результатам;
- безусловная результативность выполнения измерительной процедуры (в отличие от анализа данных [22]). При этом учитывается, что всякий акт измерения кроме ожидаемого результата из динамического диапазона измерительной

² Цитировано по источнику [21, стр. 229]

процедуры (собираетельно всякий такой результат у любой процедуры измерения можно обозначить как «**X**»), может дать два особых:

- «**None**», который свидетельствует о нахождении значения измеряемого свойства за порогами чувствительности, вне динамического диапазона средства измерений, о «семантическом несоответствии» исследуемого объекта и измерительной процедуры и др.;
- «**Failure**», который фиксирует невыполнение задачи измерений (отказ, сбой измерительного средства, воздержание при голосовании и т.п.).

Инспирировать появление результата **None** может и сам субъект, который в процессе восприятия действительности обычно применяет концептуальное шкалирование. После такого шкалирования вновь введенные свойства фактически измеряются в бинарной шкале наименований $\{X, \text{None}\}$.

- неполнота серий измерений, выражающаяся в пропуске выполнения некоторых процедур измерения (формально пропуск может фиксироваться как «получение результата» **NM** «измерение не производилось»).

Таким образом, *обобщенная* ТОС (ОТОС) представляется кортежем

$$(G, M, Se, Pr, A), \quad (2.1)$$

где:

- $Se = \{Se_{(i)}\}_{i=1, \dots, r}$, $|Se| = \sum_{i=1}^r |Se_{(i)}| = m$ – множество всех выполненных для изучаемой ПРО серий измерений, $Se_{(i)} = \{se_{(i)k}\}_{k=1, \dots, q(i)}$, $q(i) \geq 1$, $i = 1, \dots, r$ – множество серий измерений, выполненных для объекта $g_i \in G$, причем каждая серия $se_{(i)k}$ характеризуется степенью достоверности ее результатов $st_{(i)k}$;
- $Pr = \{Pr_{(j)}\}_{j=1, \dots, s}$, $|Pr| = \sum_{j=1}^s |Pr_{(j)}| = n$ – арсенал всех используемых процедур измерения, $Pr_{(j)} = \{pr_{(j)k}\}_{k=1, \dots, p(j)}$, $p(j) \geq 1$, $j = 1, \dots, s$ – множество процедур измерения свойства $m_j \in M$, причем всякая процедура $pr_{(j)k}$ характеризуется степенью доверия к ее результатам $pt_{(j)k}$;
- $A = (a_{ij})_{i=1, \dots, m; j=1, \dots, n}$ – матрица результатов серий измерений Se свойств M у объектов из выборки G , выполненных с помощью процедур измерения Pr . В качестве элементов этой матрицы выступают значения, принадлежащие дина-

мическим диапазонам процедур измерения, *лингвоконстанта X*, свидетельствующая о получении подобного результата, а также *лингвоконстанты NM, None, Failure* (иначе говоря, диапазон измерения всякой процедуры измерения *пополняется* лингвоконстантами **NM, None, Failure**).

В ОТОС, представленной на рисунке 2.3, темные ячейки соответствуют результатам измерения **None, Failure** и символам шкал динамических диапазонов измерительных процедур (возможно, замещенными лингвоконстантой **X**), а светлые – результату **NM**.

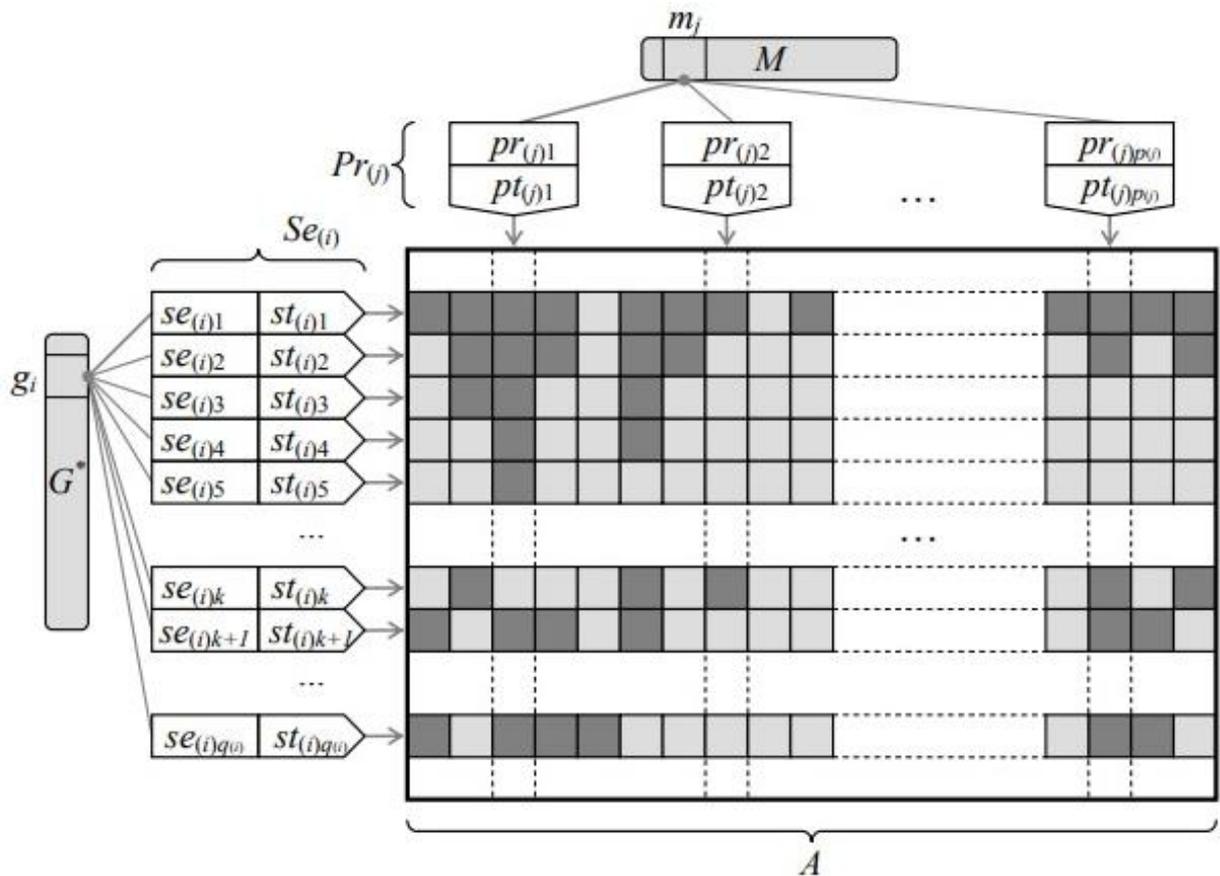


Рисунок 2.3 – Структура обобщенной таблицы «объекты-свойства» на основе двумерной матрицы

2.2.2 Концептуальное шкалирование в обобщённой таблице «объекты-свойства»

Как в СИС, так и в ОТОС, по отношению к свойствам можно применять концептуальное шкалирование, только в первом случае оно является *априорным* когнитивным актом, а во втором – *апостериорным*. Это влечёт за собой некоторые ограничения в его применении в ОТОС. Благодаря использованию лишь двух

фундаментальных способов шкалирования, обоснован ограниченный набор семантических правил манипуляции с лингвистическими константами, которые могут присутствовать в столбцах ОТОС наряду с «натуральными» значениями результатов измерения свойств.

Апостериорно возможно шкалировать только те свойства, измеренные значения которых выражаются либо «натуральным» значением (15 кг, 134 см, 28 м³ и т.д.), либо константой **None**. В случае со шкалированием «натурального» значения получаем результаты, аналогичные приведённым в таблицах 2.1 и 2.2. Для константы **None** шкалирование проводится тривиально – всем новым свойствам будет присвоено значение **None**. Тогда как концептуальное шкалирование–свойства, результаты измерения которого включают константы **X**, **NM** и **Failure** не представляется возможным, т.к. нет достаточных оснований для присвоения определённых значений свойствам, которые замещали/добавлялись бы вследствие разделения/уточнения домена шкалируемого свойства.

Следует отметить, что в АФП концептуальное шкалирование используется для перехода от многозначного ФК (1.1) к однозначному (1.2) [7, 23]. Здесь же оно имеет более широкое применение: как один из механизмов формирования СИС, т.е. изменения взгляда исследователя на ПрО, и как способ структуризации полученного эмпирического материала – ОТОС (2.1).

2.2.3 Обоснование выбора логической платформы для построения нестрогого формального контекста

Извлечение из ОТОС (2.1) совокупности оценок истинности БСС, составляющих основу ФК, невозможно в рамках классической логики: двух истинностных констант недостаточно для интерпретации результатов процедур измерений с расширенными диапазонами измерений; необходимо использование некоторой многозначной логики. Кроме того, требуется подбор такой многозначной логики, которая имела бы развитый математический аппарат для консолидации подобных результатов измерений, спектр результатов которой был бы достаточно широк для дифференциации получаемых оценок истинности БСС.

Рассмотрим некоторые известные многозначные логики.

- 1) *Логика Я. Лукасевича \mathcal{L}_3* [41] характеризуется тремя значениями истинности: 0, 1 и $\frac{1}{2}$. Классические **True** и **False** - это 1 и 0 соответственно, а $\frac{1}{2}$ интерпретировалось Лукасевичем как «возможность». Потребность описать константы **Failure** и **NM** закрывается в данной логике наличием истинностного значения $\frac{1}{2}$. Но поскольку операционный аппарат логики Лукасевича позволяет получить с помощью разных связок только эти же 3 константы, результат совмещения оценок БСС, смоделированных данной логикой, будет далек от интуитивного ожидания исследователя и чреват потерей информации. Таким образом, в случае задачи ОАД, трехзначная логика Лукасевича не может являться адекватной моделью эмпирических данных, представленных в ОТОС.
- 2) *Четырехзначная логика Дж.М. Данна* [42] была разработана применительно к обработке информации Н. Белнапом и в её основу положена идея сопоставления суждениям подмножеств множества истинностных значений {1-Истина, 0-Ложь}: $\{\{1, 0\}, \{1\}, \{0\}, \{\emptyset\}\}$. Подмножество $\{1, 0\}$ описывает ситуацию противоречия данных (пресыщенная оценка), а пустое множество \emptyset – их отсутствия (истиннозначный провал). В данном случае значению **NM** может соответствовать истиннозначный провал, а для **Failure** подходящей истинностной константы нет. Поэтому четырехзначная логика Данна также не может использоваться для представления экспериментальных данных в ОАД.
- 3) *Нечеткая логика Заде* [43], получившая широкое распространение в самых разных областях, характеризуется интервалом истинностных значений $[0,1]$, вполне пригодным для тонкой дифференциации истинностных значений БСС. Естественной интерпретацией результатов измерения **X** и **None** здесь будут истинностные константы 0 и 1. Однако корректная интерпретация **Failure** и **NM** невозможна: в обоих случаях есть основания использовать истинностное значение 0.5, а другие альтернативы для данных констант отсутствуют.

Так как при компьютерной обработке информации достаточно типична ситуация внесения в ОТОС взаимоисключающих сведений, применение ОАД ставит нас перед необходимостью выбора адекватного способа оценивания истинности

БСС для ПрО. Пожалуй, на сегодня наибольшей мощностью моделирования «человеческого подхода» к подобным оценкам обладают векторные логики, для которых хорошо развит понятийный и аналитический аппарат, обобщающий положения и классической, и нечеткой, и многих многозначных логик. Векторная логика – разновидность логического исчисления, работающего с неполной и противоречивой информацией, где истинность рассматривается как свойство суждения, характеризующее соответствие суждения отраженному в нем миру [44]. Формальное описание истинностных значений состоит в упорядоченном указании аспектов, характеризующих отдельно степени истинности и ложности такого соответствия (т.е. Ложь полагается такой же конструктивно определяемой величиной как Истина, и принцип исключенного не сохраняется). Истинность суждения a оценивается вектором

$$\mathbf{a} = \langle a_1^+, a_2^+, \dots, a_n^+, a_1^-, \dots, a_{n-1}^-, a_n^- \rangle,$$

где $a_1^+, a_2^+, \dots, a_{n-1}^-, a_n^- \in [0, 1]$, a_i^+, a_j^- – аспекты Истины и Лжи соответственно.

В частности, в логике класса V^{TF} истинность суждения b будет оцениваться вектором $\langle \text{Истина}, \text{Ложь} \rangle$ (рисунок 2.4):

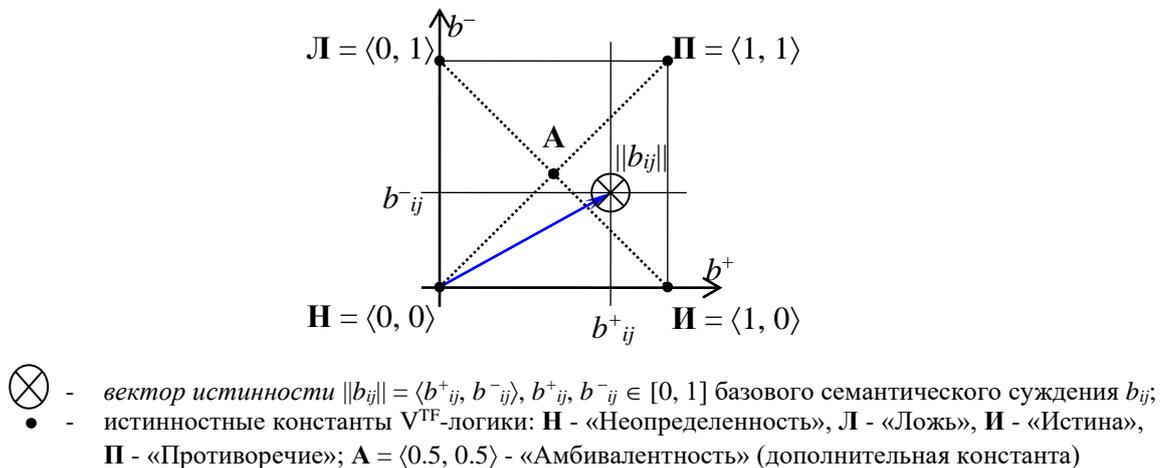


Рисунок 2.4 – Оценка истинности базового семантического суждения в векторной V^{TF} -логике

$$\mathbf{b}_{ij} = \langle b_{ij}^+, b_{ij}^- \rangle, b_{ij}^+, b_{ij}^- \in [0, 1],$$

где компонент, или аспект истинности, b_{ij}^+ – Истина, формируется свидетельствами, подтверждающими БСС, а компонент (аспект) b_{ij}^- – Ложь, – отрицающими

ми БСС [44]. Ожидание большей адекватности соответствия I_A , заданного матрицей векторов \bar{b}_{ij} , реальному процессу осмысления ПрО объяснимо тем, что традиционные константы *Истина* и *Ложь*, которыми оценивается истинность БСС, зачастую определяются независимым комплексом свидетельств так, что *Ложь* не выводима из отсутствия (дефицита) *Истины*, а *Истина* – из отсутствия *Лжи*.

2.2.4 Построение исходного нестрогого контекста на базе V^{TF} -логики

Интерпретация исходных данных. Первым шагом является переход от первичных данных, структурированных в виде матрицы A , к их смысловой интерпретации в виде нестрогого соответствия «объекты-свойства» I_A . В выбранной для представления исходных экспериментальных данных векторной логике присутствуют истинностные константы «*Истина*» $\{1, 0\}$, «*Ложь*» $\{0, 1\}$, «*Неопределенность*» $\{0, 0\}$ и «*Противоречие*» $\{1, 1\}$. Для константы **NM** в данном случае подходит значение истинности $\{0, 0\}$, т. к. при отсутствии измерения не имеется свидетельств ни в пользу Истины, ни в пользу Лжи, и нет противоречивости соответствующих свидетельств. Учет же значения **Failure** потребовал введения в анализ значения истинности «*Амбивалентность*» $\{0.5, 0.5\}$, которая имеет аналоги в прагматически ориентированных многозначных логиках (см., например, значение истинности «круглый информационный ноль» в [45]) и равноудалена от перечисленных выше истинностных констант. Тогда нестрогое соответствие $I_A = (b_{ij})_{i=1, \dots, m; j=1, \dots, n}$ будет содержать следующие возможные оценки БСС:

$$b_{ij} = \left. \begin{array}{ll} \langle 1, 0 \rangle, & \text{если } a_{ij} \in D^* \text{ или } a_{ij} = \mathbf{X}; \\ \langle 0, 1 \rangle, & \text{если } a_{ij} = \mathbf{None}; \\ \langle 0.5, 0.5 \rangle, & \text{если } a_{ij} = \mathbf{Failure}; \\ \langle 0, 0 \rangle, & \text{если } a_{ij} = \mathbf{NM}. \end{array} \right\}$$

где $D^* = \cup_{j=1, \dots, s} D_j$, D_j – домен свойства m_j .

Консолидация промежуточных данных. Далее соответствие I_A трансформируется в следующем порядке.

Во-первых, каждая оценка истинности БСС в I_A совмещается со степенью доверия к источнику данных, которые были положены в основу этой оценки. Используется простейшая модель учета степени доверия к измерительным процедурам в виде скалярных мер $pt_{(j)} \in [0, 1]$ и элементарная формула пересчета вектора истинности БСС b_{ij} : $b_{ij} = t_j \cdot b_{ij}$ (фактически это соответствует применению для моделирования дифференциации доверия к источникам информации вектора истинности V^{TF} -логики (t_j^+, t_j^-) , где $t_j^+ = t_j^- = t_j$, и объединению свидетельств об истинности каждого БСС и уровня доверия к источнику этого БСС по схеме так называемого «00-композиционного совмещения» на основе композиционного умножения согласно t -норме $x \bullet y = xy$ [44]).

Во-вторых, согласно полученному соответствию I_A определяются оценки истинности БСС по результатам всех серий измерений, выполненных для отдельно взятого объекта каждой измерительной процедурой.

Когда для объекта $g_i \in G$ произведена лишь одна серия измерений, т.е. $|Se_{(i)}| = 1$, то требуемая оценка де-факто уже имеется. Иначе оценки истинности во всех строках соответствия I_A , соответствующих элементам серии $Se_{(i)}$, должны быть совмещены. Другими словами, обработка данных многократного измерения свойства объекта приводит к необходимости интеграции независимых свидетельств истинности отдельного БСС.

Для совмещения истинностных значений в векторной логике имеется несколько допустимых схем. Наиболее адекватной в нашем случае следует признать схему «усиления-усреднения» – частный случай так называемого «11-композиционного совмещения» на основе композиционного сложения согласно триангулярной s -нормы $x \oplus y = \mathbf{min}(I, x + y)$ с весами для поступающих аспектов истинности, равными $1/N$, где N – число поступивших значений [44].

Важно отметить, что, совмещая оценки, полученные на основе работы отдельно взятой измерительной процедуры применительно к определенному объекту, следует, безусловно, отбрасывать оценки, которые были помещены в I_A на основе «не-измерений» **NM**, т.к. как их появление вызвано исключительно формальными потребностями обобщенной структуризации ТОС, их учет приведет к

некорректному результату совмещения. Единственным изъятием из этого правила будет случай, когда результат NM был «получен» во всех без исключения сериях измерений $Se_{(i)}$.

Наконец, в-третьих, для каждого свойства $m_j \in M$ в формируемом соответствии I_A «свертывается» информация конгруэнтных столбцов, соответствующих процедурам $Pr_{(j)}$. И эту интегрирующую операцию естественно производить по схеме «усиления-усреднения» при совмещении свидетельств векторной логики.

В результате получаем нестрогий исходный ФК (ИФК)

$$(G, M, I'), \quad (2.1)$$

где итоговое нестрогое соответствие I' имеет размерность $r \times s$, а его элемент b_{ij} можно рассматривать как векторный показатель принадлежности свойства m_j априори неизвестному множеству M_i свойств объекта g_i . В терминах V^{TF} -логики M_i есть нестрогое подмножество универсума M .

2.3 Выводы по разделу 2

- 1) Исходная информация для онтологического анализа данных представляется в виде СИС, как продукта априорного гипотезирования субъекта, и ОТОС, как продукта взаимодействия, чаще опосредованного приборами, субъекта с ПрО и апостериорного гипотезирования субъекта.
- 2) Концептуальное шкалирование выступает в роли источника возникновения ОСС, и способа реструктуризации полученных в ходе измерений эмпирических данных.
- 3) Для интерпретации результатов измерений в виде оценок истинности БСС избрана векторная V^{TF} -логика, являющаяся обобщением распространённой нечёткой логики, и обладающая достаточным набором констант, а также развитым математическим аппаратом для консолидации неполных данных.

3 ФОРМИРОВАНИЕ РАБОЧЕГО КОНТЕКСТА ДЛЯ ВЫВОДА ФОРМАЛЬНЫХ ПОНЯТИЙ

В разделе описывается методическое обеспечение апостериорных когнитивных актов субъекта-исследователя, направленных на формирование *рабочего* ФК (РФК). Приведенные результаты отражены в публикациях [11, 46-54].

3.1 Дефаззификация эмпирического контекста

Эмпирический ИФК (2.1), интерпретирующий и консолидирующий данные многомерных наблюдений и измерений в исследуемой ПрО в виде нестрогих оценок истинности БСС, вообще говоря, *непригоден* для онтологического анализа:

- во-первых, апробированные методы вывода из эмпирических данных формальных понятий, соответствующих определению понятия в классической логике и аналитической философии (см. пункт 1.2.1), нуждаются в *однозначном* ФК³, $I^*: G \times M \rightarrow \{\mathbf{True}, \mathbf{False}\}$;
- во-вторых, поскольку *естественным условием* выполнения ОСС для ФП, выводимых из эмпирических данных, является *нормальность* наборов свойств у каждого объекта в используемом контексте [10], то для ИФК, где объектам присущи нестрогие подмножества измеряемых свойств, вопрос о выполнении ОСС *не имеет смысла* т.к. понятие «нормальности» определено для *обычных* множеств. Следовательно, требование соблюдения ОСС можно предъявлять лишь по отношению к однозначному ФК.

Поэтому для формирования РФК

$$(G, M, I^*), \quad (3.1)$$

т.е. контекста, который применим для вывода ФП и контроля ОСС в рамках ОАД, эмпирический ИФК должен быть *дефаззифицирован*.

Следуя общепринятому подходу для дефаззификации ИФК (2.1) следует произвести *пороговое сечение* нестрогости соответствия I' [44] (рисунок 3.1):

³ Сложный в теоретическом и вычислительном плане метод нечеткого замыкания [55-57], который позволяет выводить все нечеткие ФП с нечеткими объемами и с нечеткими содержаниями из *нечеткого* ФК, на сегодня вызывает лишь академический интерес, поскольку генерирует гигантское количество понятий даже для малоразмерных «разреженных» контекстов, затрудняя извлечение из данных нового полезного знания, и к тому же требует колоссальных затрат компьютерных ресурсов.

$$I^*(g, m) = \begin{cases} \mathbf{True}, & \text{если } I'(g, m) = \langle b_{gm}^+, b_{gm}^- \rangle, b_{gm}^+ \geq \alpha^+, b_{gm}^- \leq \alpha^-; \\ \mathbf{False} & \text{в противоположном случае.} \end{cases} \quad (3.2)$$

где вектор истинности $\alpha = \langle \alpha^+, \alpha^- \rangle$, – субъективный порог доверия (СПД, α -порог) к эмпирическим данным, представленным в ИФК в виде векторных оценок истинности БСС, а b_{gm} – оценка истинности БСС «объекту g присуще свойство m ».

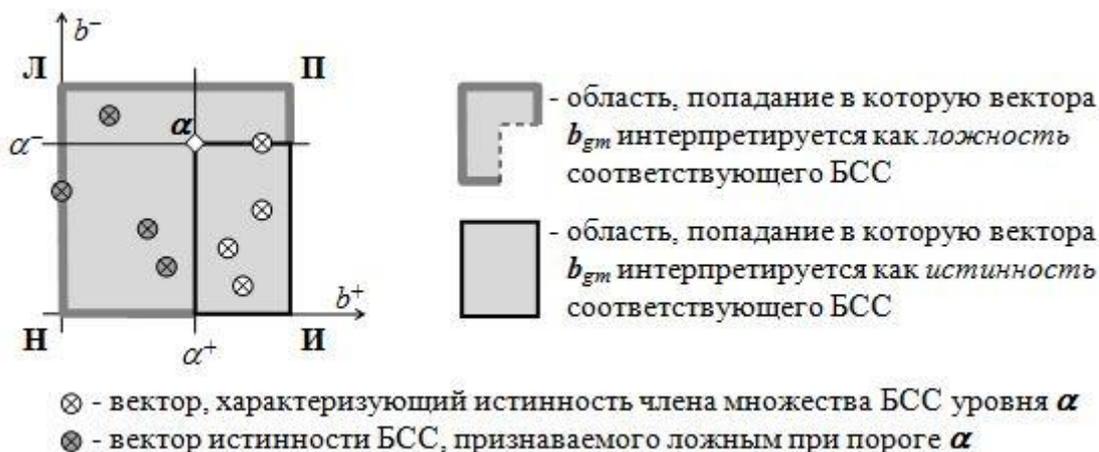


Рисунок 3.1 – Определение множества уровня $\alpha = \langle \alpha^+, \alpha^- \rangle$ (или α -сечения) нестроеного множества через отношение логического порядка для векторов истинности в V^{TF} -логике: $b_{gm} \succ \alpha$, если $b_{gm}^+ \geq \alpha^+, b_{gm}^- \leq \alpha^-$

Субъект выбирает порог доверия к эмпирическим данным исключительно из собственных соображений. Вместе с тем, по логическим мотивам установку некоторых СПД при использовании для дефаззификации ИФК формулы (3.2) нельзя признать корректной:

- порог $\alpha = \langle \alpha^+, 1 \rangle$, $\alpha^+ \in [0, 1]$ отнесет к истинным в частности и те БСС, которые эмпирически оценены как *противоречивые* или «*ложные с долей правды*»;
- порог $\alpha = \langle 0, \alpha^- \rangle$, $\alpha^- \in [0, 1]$ отнесет к истинным в частности эмпирически *неопределенные* и «*неопределенные с долей лжи*» БСС.

При дефаззификации ИФК все эти БСС разумно (пусть одни с большим, другие с меньшим основанием) отнести к *ложным*.

Пороговую дефаззификацию эмпирического ИФК безусловно следует отнести к *когнитивным актам* субъекта при исследовании ПрО. Возможность экспериментирования со значением СПД в ОАД позволяет субъекту существенно изменять характер фактической информации, из которой извлекается концептуаль-

ная картина ПрО. В [30] показано, что варьирование СПД при дефаззификации нечеткого ФК приводит к извлечению *конечного* множества *различных* формальных онтологий. При этом и ужесточение, и смягчение СПД относительно амбивалентной константы истинности – в V^{TF} -логике такой константой является вектор истинности $\mathbf{O} = \langle 0.5, 0.5 \rangle$ (см. рисунок 2.5) – ведут в целом к постепенной деградации извлекаемых вариантов онтологии ПрО и в части количества понятий, и в структуре отношения обобщения между ними, и в количестве интенциональных отношений между понятиями.

Подводя итог сказанному можно констатировать, что дефаззификацию ИФК в ОАД надлежит осуществлять в *два этапа*:

- на первом *единожды* выполняемом этапе все элементы соответствия I' в ИФК (2.1), которые оценены векторами истинности $\langle b^+, 1 \rangle$ и $\langle 0, b^- \rangle$, где $b^+, b^- \in [0, 1]$, означиваются специальной лингвистической константой **FP** (от англ. *false proposition* – ложное суждение);
- ко второму этапу в ходе онтологического анализа субъект должен иметь возможность возвращаться неоднократно для экспериментирования с различными значениями СПД. На этом этапе строится собственно РФК (3.1), однозначное соответствие I^* которого определяется как пороговое α -сечение модифицированного на первом этапе соответствия I' :

$$\left\{ \begin{array}{l} \forall g \in G, \forall m \in M : \\ I^*(g, m) = \left\{ \begin{array}{ll} \mathbf{True} & \text{если } I'(g, m) = \langle b_{gm}^+, b_{gm}^- \rangle, b_{gm}^+ \geq \alpha^+, b_{gm}^- \leq \alpha^-; \\ \mathbf{False} & \text{если } I'(g, m) = \mathbf{FP}; \\ \mathbf{False} & \text{в иных случаях.} \end{array} \right. \end{array} \right\}, \quad (3.3)$$

где порог $\alpha = \langle \alpha^+, \alpha^- \rangle$ либо устанавливается субъектом с соблюдением ограничений

$$\alpha^+ \in (0, 1], \alpha^- \in [0, 1), \quad (3.4)$$

либо, если исследователь игнорирует когнитивную возможность указания СПД, принимается по умолчанию равным $\langle 0^+, 1^- \rangle$, где знаки «+» и «-» означа-

ют соответственно сколь угодно малое приращение и уменьшение помечаемых значений.

3.2 Разработка метода нормализации рабочего контекста

Очерченный выше процесс формирования РФК (3.1) не исключает присутствия в нём объектов, набор свойств которых не удовлетворяет ОСС. Поэтому в общем случае, несмотря на то, что этот контекст представляет эмпирические данные, признанные субъектом достоверными, выводимая из него понятийная структура исследуемой ПрО будет противоречить соответствующим априорным представлениям (гипотезам) субъекта (см. пункт 1.3.2).

В предыдущем подразделе отмечено, что условием выполнения ОСС для ФП, выводимых из эмпирических данных, является нормальность наборов свойств у каждого объекта в используемом контексте. Вследствие этого тривиальной рекомендацией для соблюдения ОСС при выводе ФП (и, следовательно, элементарным приемом комбинирования двух существующих методик ФП – см. пункт 1.3.2) является сепарабельная *проверка* выполнения этого условия для каждого объекта в РФК (3.1).

Отрицательный результат такой проверки будет означать, что эмпирические данные противоречат, по меньшей мере, части априорных гипотез субъекта о понятийной структуре ПрО. В этой ситуации для продолжения онтологического анализа имеющихся данных субъект-исследователь располагает всего тремя вариантами действий:

- *продолжить* анализ, невзирая на нарушение ОСС в РФК, фактически отказываясь от некоторой части своих априорных предположений о ПрО, зафиксированных в СИС;
- *изменить* СПД и следуя (3.3) получить иную версию РФК (3.1), в котором, *возможно*, ОСС будут удовлетворены;
- за счёт отказа от части эмпирической информации о ПрО *трансформировать* соответствие I^* в РФК (3.1): $I^* \rightarrow I^\circ$, – таким образом, чтобы в преобразованном РФК

$$(G, M, I^\circ), \quad (3.5)$$

наборы свойств у всех объектов стали нормальными множествами:

$$\left\{ \begin{array}{l} \forall g_i : m_j \cup \{g\}' \rightarrow ((\exists m_k \in M, k \neq j : C(m_j, m_k)) \rightarrow m_k \in \{g\}' \wedge) \\ ((\exists m_k \in M, k \neq j : E(m_j, m_k)) \rightarrow m_k \neq \{g\}') \end{array} \right\}. \quad (3.6)$$

Строго говоря, актом *нормализации* РФК следует считать лишь последний вариант действий, но с позиции достижения искомой цели к методам нормализации условно можно отнести и *поиск СПД, формирующего* нормализованный РФК (3.5) при дефаззификации ИФК (2.1). В любом случае принятие субъектом-исследователем решения добиваться или нет нормализации контекста следует квалифицировать как еще один (наряду с выбором СПД) *когнитивный акт* в процессе обработки эмпирических данных о ПрО.

3.2.1 Критический анализ существующих подходов к нормализации контекста

Метод выявления области существования *значений* СПД, которые при дефаззификации ИФК согласно (3.3) обеспечивают нормальность наборов свойств у каждого объекта формируемого РФК, предложен в [31].

В основу метода положена идея составления и решения относительно вектора-порога $\alpha = \langle \alpha^+, \alpha^- \rangle$ уравнения

$$T(\alpha) = \mathbf{True},$$

в котором $T(\alpha)$ – логическое выражение, описывающее условия нормальности набора свойств у каждого объекта РФК при заданных ОСС и всех возможных соотношениях между порогом α и оценками истинности БСС в дефаззифицируемом ИФК. Охват всего ИФК обеспечивается тем, что

$$T(\alpha) = \bigwedge_{i=1, \dots, |G|} T_i(\alpha),$$

где $T_i(\alpha)$ – условие нормальности набора свойств у i -го объекта РФК (3.1), которое определяется *конъюнкцией* определенного множества выражений, описывающих замкнутость и совместимость указанного набора на основе отношения логи-

ческого порядка для векторов истинности в V^{TF} -логике (см. рисунок 3.1) и условия (3.6).

Например, если $m_j, m_k \in M, j \neq k$ и $C(m_j, m_k) = \mathbf{True}$ (т.е. согласно ОСС присутствие у любого объекта $g_i \in G$ свойства m_j обуславливает наличие у него свойства m_k), то порог α должен обеспечивать истинность импликации

$$(b_{ij}^+ \geq \alpha^+) \wedge (b_{ij}^- \leq \alpha^-) \rightarrow (b_{ik}^+ \geq \alpha^+) \wedge (b_{ik}^- \leq \alpha^-),$$

где $\langle b_{ix}^+, b_{ix}^- \rangle$ оценка истинности БСС «объекту g_i присуще свойство m_x ».

Значительно сложнее описывается условие совместимости множества свойств объекта в РФК. В частности, в ОСС для этого требуется предварительно выделить пересекающиеся подмножества попарно несовместимых свойств, а утверждения, устанавливающие истинность искомой совместимости свойств, конструируются как разделяющие дизъюнкции допустимых вариантов существования свойств объекта.

В [31] отмечается возможность автоматизации построения весьма громоздкого выражения $T(\alpha)$, но обходится вопрос об эффективном способе отыскания области существования релевантных порогов доверия, и лишь подчеркивается, что она может оказаться несвязной или даже пустой. В целом авторы [31] сами критически оценивают предложенный метод, считая его практически малополезным.

Соглашаясь с этой оценкой, следует, тем не менее, подчеркнуть, что с позиции данной работы фундаментальным недостатком рассматриваемого метода является *ограничение* допустимых значений СПД. Этот существенный *концептуальный изъян* означает деградацию когнитивного потенциала, обусловленного свободой выбора порога доверия к эмпирическим данным с целью дефаззификации ИФК.

Более приемлемым является метод нормализации РФК (3.1), идея которого высказана в [58], а одна из возможных реализаций представлена в [59]. В указанных работах ставилась задача такого преобразования *нечеткого* ИФК, чтобы для *любого* его порогового сечения выполнялись ОСС.

В основу метода положена известная *теорема о декомпозиции* нечетких множеств [60]; аналогичное утверждение легко доказывается для случая простейшего обобщения нечеткой логики – векторной логики V^{TF} [44].

В используемых обозначениях для каждого $g \in G$ в ИФК (2.1) имеем

$$\{g\}' = \cup_{\forall \alpha} \alpha \cdot (\{g\}')_{\alpha}, \quad (3.7)$$

где $\alpha = \langle \alpha^+, \alpha^- \rangle$ – вектор-порог доверия к эмпирическим данным (допустимые значения порога определяет формула (3.4)), $\{g\}'$ – нестрогое множество свойств объекта g в ИФК, $(\{g\}')_{\alpha}$ – α -сечение нестрогого множества $\{g\}'$ (обычное подмножество множества измеряемых свойств M). Элементы нестрогого подмножества свойств $\alpha \cdot (\{g\}')_{\alpha}$ оцениваются вектором истинности α , если рассматриваемое свойство принадлежит α -сечению $(\{g\}')_{\alpha}$, и константой $\mathbf{1} = \langle 0, 1 \rangle$ в противоположном случае.

Предложенный в [58] метод трансформации ИФК учитывает, что в реальных контекстах для каждого $g \in G$ количество различных «слагаемых» в (3.7) конечно, и может быть сведен к следующему:

- согласно имеющимся ОСС формируется множество-репозиторий всех нормальных подмножеств измеряемых свойств M (создание такого репозитория призвано ускорить целевые действия на последующем этапе);
- сепарабельно трансформируется нестрогое подмножество свойств каждого объекта в ИФК: в качестве нового нестрогого подмножества свойств принимается правая часть (3.7) за исключением «слагаемых», где множества свойств $(\{g\}')_{\alpha}$ не являются нормальными.

Регламентируемый таким образом отказ от части эмпирических данных в ИФК гарантирует, что дефаззификация преобразованного ИФК при любом СПД даст однозначный РФК, в котором каждый объект характеризуется нормальным множеством измеряемых свойств [59].

Концептуальным недостатком описанного метода нормализации РФК является избыточность действий, которые опережают проявление когнитивных потребностей субъекта в части варьирования порогом доверия к данным, и, тем са-

мым, ведут к возможно неоправданному отказу от части эмпирической информации, потенциально полезной для онтологического анализа. Кроме того, для контекстов большой размерности очевидна высокая трудоемкость данного метода, связанная с формированием и анализом композиций вида (3.5).

3.2.2 Эвристика и механизм нормализации контекста

Представляется, что разумным подходом к нормализации эмпирически и субъективно (вследствие установки СПД) выявленных наборов свойств объектов в РФК (3.1) будет *денонсирование* некоторых БСС, которые описываются этим контекстом как истинные, но *нарушают* «нормальность» множеств свойств объектов (3.6).

Элементарным актом этой *эвристики* является *денонсация* в РФК (3.1) отдельно взятого БСС, оцениваемого в эмпирическом ИФК вектором истинности $\langle b_{gm}^+, b_{gm}^- \rangle$ [48, 52].

«Физическая» интерпретация этого акта состоит в «отсечении» вектора истинности этого БСС с помощью порога доверия $\beta = \langle \beta^+, \beta^- \rangle$ (рисунок 3.2):

$$\beta^+ \geq b_{gm}^+ \vee \beta^- \leq b_{gm}^-,$$

откуда, в частности, ясно, что *абсолютно достоверное* БСС – суждение, истинность которого оценивается вектором истинности $\langle 1, 0 \rangle$, – не может быть денонсировано.

Примером порога β , денонсирующего БСС, истинность которого не абсолютна и оценивается вектором $\langle b_{gm}^+, b_{gm}^- \rangle$, является

$$\beta = \langle b_{gm}^+ + \varepsilon^+, b_{gm}^- - \varepsilon^- \rangle, \quad (3.8)$$

где сколь угодно малые $\varepsilon^+ \geq 0$ и $\varepsilon^- \geq 0$ одновременно не равны нулю. Поэтому для определенности при упоминании порога доверия, денонсирующего отдельно взятое БСС, далее будет подразумеваться вектор-порог (3.8) – результат *предельно малого ужесточения* векторной оценки истинности денонсируемого БСС, – а для указания этого порога для краткости будет использоваться непосредственно вектор-оценка истинности денонсируемого БСС.

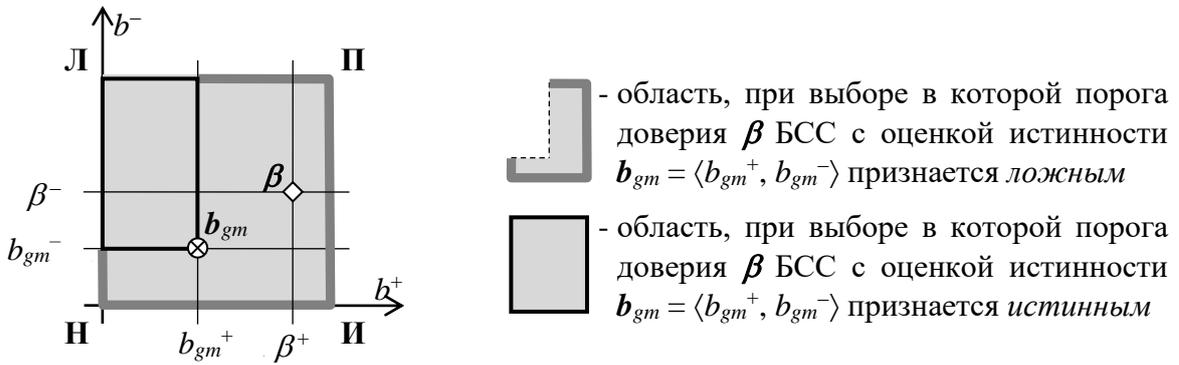


Рисунок 3.2 – Области выбора порога доверия к исходным данным $\beta = \langle \beta^+, \beta^- \rangle$, определяющие отнесение БСС с оценкой истинности $\mathbf{b}_{gm} = \langle b_{gm}^+, b_{gm}^- \rangle$ либо к истинным, либо к ложным (отображено последнее)

Следуя этому соглашению, констатируем, что для денонсации БСС с оценкой истинности $\mathbf{b}_{gm} = \langle b_{gm}^+, b_{gm}^- \rangle$, которое в РФК (3.1) при СПД $\alpha = \langle \alpha^+, \alpha^- \rangle$ признано достоверным, необходимо установить порог доверия к эмпирическим данным, определяемый суммой вектора α и вектора ужесточения порога TTV (от англ. *Threshold Tightening Vector*) при условии

$$\alpha + TTV = \mathbf{b}_{gm}.$$

Это положение иллюстрирует рисунок 3.3.

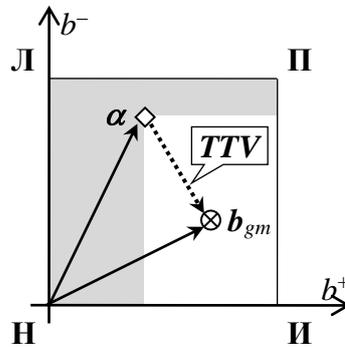


Рисунок 3.3 – Ужесточение TTV субъективного порога доверия α для денонсации базового семантического суждения с оценкой истинности \mathbf{b}_{gm}

Рассмотрим ситуацию, когда для реализации сформулированной выше эвристики для конкретного объекта $g \in G$ в РФК (3.1) необходимо денонсировать *одновременно* (одноактно) несколько БСС, которые признаны истинными при пороге α , а в эмпирическом ИФК оцениваются векторами истинности вида $\mathbf{b}_{gx} = \langle b_{gx}^+, b_{gx}^- \rangle$, где $x \in M_g^\bullet, M_g^\bullet \subseteq M, |M_g^\bullet| \geq 1$.

Тогда можно предложить несколько вариантов определения вектора ужесточения, обеспечивающего «отсечение» сразу всех векторов \mathbf{b}_{gx} (рисунок 3.4):

- $\mathbf{TTV}_{maxmin} = \{\mathbf{TTV}_{maxmin}\}$ – единичное множество, содержащее вектор ужесточения, координаты стока которого задает вектор

$$\langle \max_{x \in M} \{b_{gx}^+\}, \min_{x \in M} \{b_{gx}^-\} \rangle; \quad (3.9)$$

- $\mathbf{TTV}_{Ps} = \{\mathbf{TTV}_{Ps,i}\}$ – вектора ужесточения, координаты стоков которых определяются $\max\min$ -множеством Парето (*Pareto set* – Ps) векторов истинности денонсируемых БСС, $i \in IxPs$ – множество индексов элементов этого Парето-множества;
- $\mathbf{TTV}_{JPs} = \{\mathbf{TTV}_{JPs,i}\}$ – вектора ужесточения, координаты стоков которых определяются «стыками» членов $\max\min$ -множества Парето (“*Joints*” of *Pareto set members* – JPs) векторов истинности денонсируемых БСС, $i \in IxJPs$ – множество индексов элементов этих «стыков».

Для сравнения альтернативных \mathbf{TTV} -векторов в определенной степени пригоден показатель *достоверности* векторов истинности V^{FT} -логики, где достоверность вектора истинности $\mathbf{b} = \langle b^+, b^- \rangle$ определяется скаляром [44]

$$\mu_b = b^+ - b^-, \quad (3.10)$$

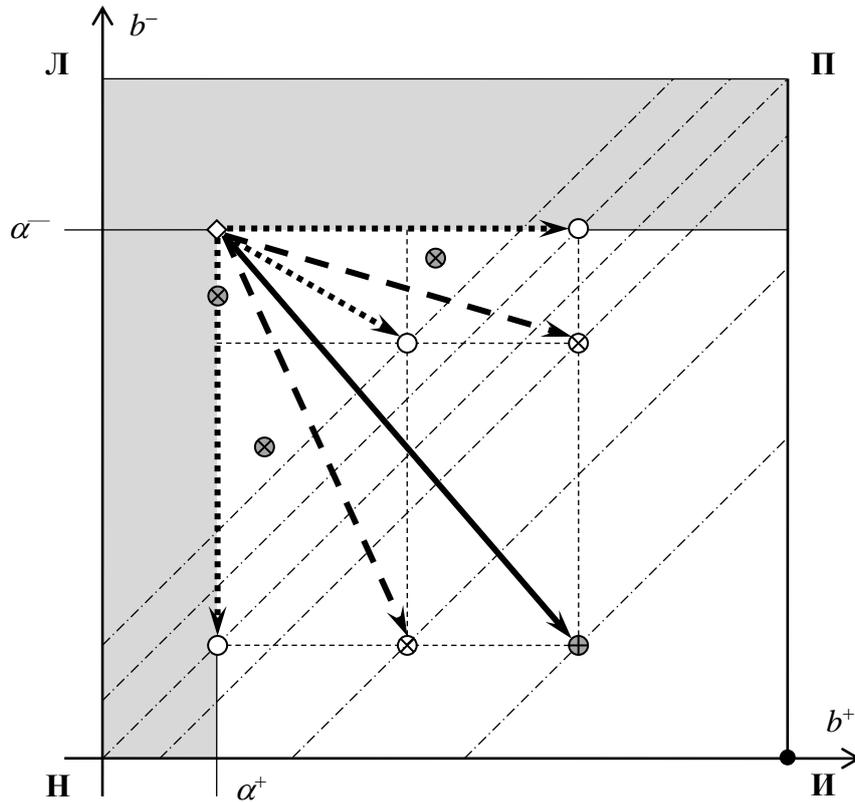
который принимает значение -1 для строго ложных и 1 для строго истинных суждений, в целом $\mu_b \in [-1, 1]$. В привычном диапазоне $[0, 1]$ достоверности (3.10) изоморфна величина

$$(b^+ - b^- + 1)/2. \quad (3.11)$$

Применительно к оцениванию \mathbf{TTV} -векторов очевидно значима лишь достоверность их *стоков* (см. рисунок 3.4): для вектора (3.9), векторов-членов и векторов-стыков $\max\min$ -множества Парето соответственно приведенному выше перечислению множеств \mathbf{TTV}_{maxmin} , \mathbf{TTV}_{Ps} и \mathbf{TTV}_{JPs} .

Тогда нетрудно показать, что для такой характеристики \mathbf{TTV} -вектора – «*достоверности стока*» – верно

$$\mu_{\mathbf{TTV}} = |\mathbf{TTV}| + \mu_\alpha, \quad (3.12)$$



- ◇ - субъективный порог доверия к исходным данным (порог α);
- ⊗ - вектор истинности одноактно денонсируемых базовых семантических суждений (БСС);
- ⊗ - член $\max\min$ -множества Парето векторов истинности денонсируемых БСС;
- - «стык» членов $\max\min$ -множества Парето векторов истинности денонсируемых БСС;
- ⊕ - сток $\max\min$ -вектора ужесточения субъективного порога доверия к исходным данным;
- ↗ - $\max\min$ -вектор ужесточения субъективного порога доверия к исходным данным;
- ↘ - вектор ужесточения субъективного порога доверия к исходным данным, сток которого – член $\max\min$ -множества Парето векторов истинности денонсируемых БСС;
- ⋯ - вектор ужесточения субъективного порога доверия к исходным данным, сток которого – «стык» $\max\min$ -множества Парето векторов истинности денонсируемых БСС;
- - линия уровня достоверности векторов истинности

Рисунок 3.4 – Варианты ужесточения субъективного порога доверия $\alpha = \langle \alpha^+, \alpha^- \rangle$ для отсекаания векторов истинности множества одноактно денонсируемых БСС

где TTV – вектор ужесточения порога α , примененного для построения однозначного РФК (3.1), μ_{TTV} , μ_α - достоверность стока вектора и порога α соответственно, TTV , $|TTV|$ – длина вектора TTV в метрике «расстояния городских кварталов» L1.

Таким образом, достоверность стока любого TTV-вектора равна его длине с точностью до величины μ_α , которая в задаче нормализации РФК (3.1) является *константой*.

Это не только вполне оправдывает пригодность достоверности (3.10) или (3.11) для измерения векторов ужесточения, но и позволяет выдвинуть *гипотезу*, что лучшим TTV-вектором для случая денонсации множества БСС в РФК является член множества TTV_{JP_s} , характеризуемый *минимальной* достоверностью стока, поскольку по определению справедливо

$$\min_{i \in IxJP_s} \{|TTV_{JP_s,i}|\} < \min_{i \in IxPs} \{|TTV_{Ps,i}|\} \leq |TTV_{maxmin}|. \quad (3.13)$$

Гипотеза основывается на том, что при нормализации РФК путем денонсации некоторого множества БСС *интуитивно предпочтительнее* как можно менее ужесточать порог α , обеспечивая тем самым меньшее искажение принятой субъектом эмпирической картины ПрО.

Вместе с тем это предпочтение противоречит требованию минимизации затрат вычислительных ресурсов, т.к. очевидно соотношение

$$B(TTV_{JP_s}) \geq B(TTV_{Ps}) \geq B(TTV_{maxmin}), \quad (3.14)$$

где B – функция затрат.

Наконец, легко показать, что длина (и, равно, достоверность стока) любого TTV-вектора *не будет убывать* с увеличением количества БСС, подлежащих одноактной денонсации.

3.2.3 Метод нормализации контекста

Предложенная эвристика и механизм денонсации БСС положены в основу метода нормализации РФК, который лишён недостатков ранее разработанных подходов к решению этой задачи [46-51, 53].

Стратегически предлагаемый метод предусматривает циклическое выполнение следующих двух этапов обработки данных сепарабельно для каждого объекта однозначного РФК (3.1):

- поиск нарушителей ОСС во множестве свойств $M_g^* \subseteq M$ объекта $g \in G$, которое определяется *текущим состоянием* нормализуемого контекста;
- анализ множества найденных нарушителей и денонсация в обрабатываемом контексте одного или одновременно нескольких *рационально* выбранных БСС, свидетельствующих об обладании рассматриваемым объектом некоторого множества свойств-нарушителей $M_g^\bullet \subseteq M_g^*$, или, для краткости, «БСС о свойствах-нарушителях M_g^\bullet ».

Эти действия повторяются *до тех пор*, пока будут обнаруживаться свойства-нарушители ОСС (при необходимости вплоть до $M_g^* = \emptyset$). Вместе с тем, обнаружение свойств-нарушителей, которые присущи объекту согласно абсолютно достоверным БСС, может привести к *прерыванию и завершению* этого процесса с вердиктом о невозможности нормализовать РФК. Таким образом, разработанный метод в своей основе *результативен*.

На очередном цикле процесса нормализации *рациональным* разумно считать такой выбор подлежащего денонсации множества БСС о свойствах-нарушителях (возможно, единичного), члены которого *менее достоверны*, чем члены других обнаруженных множеств БСС о свойствах-нарушителях. Такая рационализация означает, что нормализация контекста будет достигаться за счет *элиминирования на каждом цикле относительно менее достоверной* эмпирической информации.

Однако такой выбор подлежащего денонсации множества БСС о свойствах-нарушителях *неоднозначен*, т.к. достоверность членов у нескольких таких множеств БСС может быть одинаковой. Поэтому в качестве дополнительного критерия рациональности разумно предложить *мощность* множества БСС, выбранного для денонсации на очередном цикле нормализации свойств объекта в РФК: меньшее количество денонсируемых БСС – это и меньшее изменение эмпирически и субъективно (путем задания СПД) обусловленной мощности множества свойств рассматриваемого объекта.

При обнаружении нескольких множеств БСС о свойствах-нарушителях, которые эквивалентны согласно вышеуказанным критериям, следует реализовать либо субъективный, либо случайный выбор одного из найденных множеств БСС.

Достоверность одного БСС определяет формула (3.10). Достоверность множества БСС естественно считать функцией от достоверностей его членов, значение которой определит *агрегированный показатель достоверности* (АПД) множества БСС. Например, в таком качестве может выступить среднее арифметическое (или любая функция среднего Колмогорова [61]). Однако подобные средние не демонстрируют монотонность при изменении *количества* аргументов, что полезно при формировании множества денонсируемых БСС с точки зрения критерия мощности этого множества.

Можно обозначить несколько весомых аргументов в пользу выбора в качестве АПД множества БСС *достоверности стока* соответствующего TTV-вектора:

- АПД единичного множества БСС будет в точности *равен* достоверности (3.10) отдельно взятого БСС;
- логично полагать, что АПД множества БСС, каждое из которых признано истинным в РФК (3.1), *превышает* достоверность порога α . Это положение для предлагаемого АПД справедливо согласно (3.12);
- *достоверность стока* TTV-вектора при любом варианте его определения *не убывает* с увеличением количества одноактно денонсируемых БСС.

Такое определение АПД оставляет неясным вопрос о конкретном варианте его вычисления, тем более что интуитивно более предпочтительный способ, связанный с отысканием множества векторов TTV_{JP_s} , в вычислительном плане наиболее трудоемок (см. соотношения (3.13) и (3.14)). Это наблюдение требует специального исследования для рационального выбора способа вычисления АПД, результаты которого представлены далее в пункте 3.2.4.

Определившись с результативностью и рабочими критериями предлагаемого метода, обратимся к детализации выполнения отдельного цикла нормализации для определенного объекта $g \in G$ в РФК. Это следующая последовательность действий [49]:

- 1) Безусловная *денонсация* БСС для свойств из M_g^* , которые *обуславливают* свойства, отсутствующие в M_g^* .

2) *Выявление* в M_g^* групп *попарно несовместимых* свойств (Н-групп), включающих более одного члена, т.е. Н-групп - нарушителей ОСС:

$$IGr_{g,1}, IGr_{g,2}, \dots, IGr_{g,n} \subseteq M_g^*, n \geq 0.$$

Если $n = 0$, выход из цикла нормализации объекта g .

3) *Выделение* $q \leq n$ субструктур множества M_g^* , каждая из которых есть совокупность связанных Н-групп – нарушителей, и подмножеств свойств объекта g , составляющих каждую такую субструктуру: $M_{g,1}, M_{g,2}, \dots, M_{g,q} \subseteq M_g^*$, $\forall j, k \in \{1, 2, \dots, q\}: j \neq k \rightarrow M_{g,j} \cap M_{g,k} = \emptyset$.

4) *Перебор* всех версий надлежащей редукции *каждого* связанного множества свойств $M_{g,i}$ путем изъятия из него множества свойств $M_{g,i}^\bullet$, варианты которого формируются с помощью рекурсивной процедуры сокращенного обхода конечного *дерева решений*, соответствующего рассматриваемому связанному множеству свойств $M_{g,i}$.

Обход такого дерева *сокращается* за счет упоминавшейся монотонности используемого критерия эффективности – АПД тех БСС, денонсация которых приведёт к исключению множества свойств $M_{g,i}^\bullet$ из $M_{g,i}$ (т.е. к некоторой редукции $M_{g,i}$).

Отбор вариантов $M_{g,i}^\bullet$, обеспечивающих *надлежащую* редукцию $M_{g,i}$ производится в результате серии блокирующих проверок:

- во-первых, проверяется *допустимость* формируемого варианта, когда устанавливается, что очередное свойство $x \in M_{g,i}$, предлагаемое для пополнения текущего варианта $M_{g,i}^\bullet$, либо не является обусловленным, либо обуславливающее его свойство $y \notin M_g^*$;
- во-вторых, устанавливается *пригодность* варианта, т.е. сохранение в результате предлагаемого изъятия множества свойств $M_{g,i}^\bullet$ из $M_{g,i}$ в каждой из связанных Н-групп, включающих свойства $M_{g,i}$, не более чем одного свойства;
- в-третьих, исследуется *эффективность* варианта $M_{g,i}^\bullet$ с точки зрения величины АПД денонсируемых БСС. Анализ прерывается, и вариант отвергается.

ся, если вычисленный АПД превышает достигнутый минимум этого критерия у исследованных и сохраненных вариантов $M_{g,i}^{\bullet}$;

- в-четвертых, оценивается *рациональность* варианта $M_{g,i}^{\bullet}$. Если вычисленный АПД равен ранее достигнутому минимуму этого критерия у сохраненных вариантов $M_{g,i}^{\bullet}$, то сравнивается количество подлежащих денонсации БСС согласно рассматриваемому варианту (т.е. $|M_{g,i}^{\bullet}|$) и согласно сохраненным вариантам. Рассматриваемый вариант отвергается, когда $|M_{g,i}^{\bullet}|$ превышает количество подлежащих денонсации БСС согласно сохраненным вариантам этого множества.

Если блокирующие проверки успешно преодолены, то вариант $M_{g,i}^{\bullet}$ для редукции M_g^* признается лучшим и *сохраняется*. При этом ранее сохраненные варианты этого множества *уничтожаются*, если у них количество денонсированных БСС больше.

Если в результате перебора версий этой редукции сохранены несколько вариантов множества подлежащих изъятию свойств $M_{g,i}^{\bullet}$, то производится обращение к субъекту-исследователю для выбора одного из вариантов редукции связного множества свойств $M_{g,i}$. Все варианты за исключением выбранного уничтожаются.

Анализ $M_{g,i}$ завершается денонсацией БСС для свойств $M_{g,i}^{\bullet}$.

На рисунке 3.5 приведен пример, иллюстрирующий решение задачи нормализации контекста с помощью описанного алгоритма: (а) ИФК, где $G = \{g\}$, $M = \{m_1, m_2, \dots, m_7\}$, I' – нестрогое соответствие множеств G и M ; (б) – РФК при пороге доверия к эмпирическим данным $\langle 0+, 1- \rangle$; (в) – множество измеряемых свойств с заданными на нём экзистенциальными зависимостями т.е. СИС (самообусловленность свойств опущена); (г) – нормализованный РФК.

Нужно отметить следующие особенности применения метода нормализации контекста в рассматриваемом примере:

- денонсация БСС «объекту g присуще свойство m_1 » не потребовалась, поскольку объект g обладает свойством m_2 , обусловливаемым свойством m_1 ;

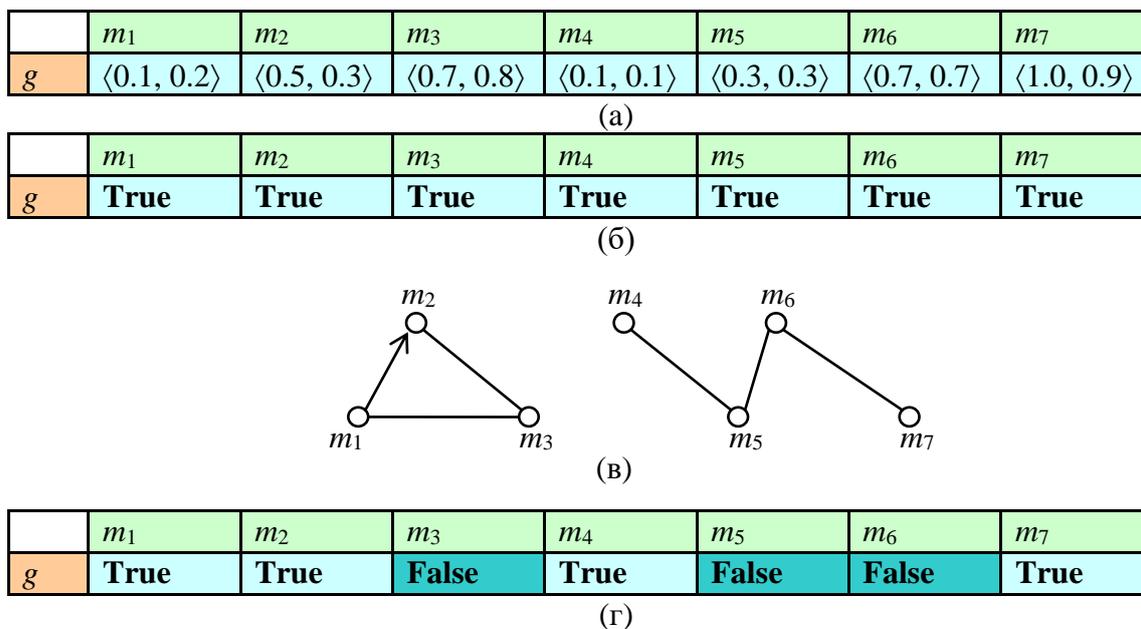


Рисунок 3.5 – Иллюстрация применения метода нормализации формального контекста

- выявлены следующие Н-группы измеряемых свойств, нарушающие ОСС: $\{m_2, m_3\}$, $\{m_1, m_3\}$, $\{m_4, m_5\}$, $\{m_5, m_6\}$, $\{m_6, m_7\}$;
- В Н-группах - нарушителях выделены следующие связные подмножества измеряемых свойств: $\{m_1, m_2, m_3\}$, $\{m_4, m_5, m_6, m_7\}$. При нормализации первого из них требуется, вообще говоря, перебрать 2^3 вариантов решения, второго – 2^4 . Применение предложенного метода нормализации уменьшило в первом случае перебор в два раза, во втором – на треть;
- АПД денонсированных БСС, вычисленный согласно варианта (3.10) по формуле (3.11), при нормализации первого связного подмножества свойств составил 0.45, второго – 0,7.

По завершению нормализации РФК применительно к объекту g можно вычислить АПД всех БСС, благодаря денонсации которых множество M_g^* редуцировано до некоторого нормального подмножества, принимаемого в контексте в качестве искомого признакового описания объекта g . Для субъекта значение этого АПД играет роль *цены* нормализации признакового описания объекта в РФК. Подсчет этого показателя в рассмотренном примере дает при использовании (3.10) и (3.11) величину, равную 0,7.

3.2.4 Сравнительное исследование вариантов определения агрегированного показателя достоверности у денонсируемых базовых семантических суждений

АПД множества денонсируемых БСС, который является основным рабочим критерием в предложенном методе нормализации контекста, предложено вычислять как достоверности стока TTV-вектора, «отсекающего» вектора истинности денонсируемого множества БСС.

Однако в пункте 3.2.2 были представлены три различных варианта определения TTV-вектора, которые эквивалентны с точки зрения задачи денонсации заданного множества БСС. Хотя один из них в силу (3.13) интуитивно выглядит более предпочтительным, контрапозиция характеристик (3.13) и (3.14) этих вариантов потребовала провести специальное сравнительное исследование для выяснения, какой из предложенных вариантов определения TTV-вектора более пригоден как основа для вычисления АПД денонсируемых БСС. В качестве критерия такой пригодности резонно использовать величину, связанную со вспомогательным критерием рациональности нормализации – длиной получаемого нормального множества свойств, – т.к. нормализации множества свойств любого объекта в РФК желательно достичь, денонсируя возможно меньшее количество БСС, которые согласно эмпирическим данным и СПД признаны истинными. Этому условию удовлетворяет величина *относительной мощности* Q_{norm} результирующего нормального множества свойств объекта $g \in G$ в РФК:

$$Q_{norm} = |M_g^* \setminus M_g^\bullet| / |M|,$$

где $M_g^* \subseteq M$ – множество свойств объекта g в РФК до его нормализации, $M_g^\bullet \subseteq M_g^*$ – множество свойств объекта g , которых он был лишен для нормализации его признакового описания.

Величину Q_{norm} для разных вариантов определения TTV-вектора достаточно исследовать лишь в зависимости от доли Н-пар $q_{Н-п}$ в СИС (по отношению к максимально возможному количеству таких пар), т.к. вычисление АПД имеет значение только при нормализации связанных подмножеств Н-групп (см. пункт 3.2.3).

Поэтому эксперименты для количественной оценки Q_{norm} выполнялись для ОСС, характеризуемых параметрами

$$N_{O-II} = 0, N_{H-II} \in \{0, 1, 2, \dots, |M| \cdot (|M| - 1) / 2\},$$

где N_{O-II} и N_{H-II} – количество соответственно О- и Н-пар в СИС.

Заранее очевидно, что при $q_{H-II} = 0$ длина нормального подмножества будет равна мощности изначального множества свойств у объекта, т.е. максимальна – Q_{norm}^{max} , а при $q_{H-II} = |M| \cdot (|M| - 1) / 2$, она будет минимальна – Q_{norm}^{min} . Таким образом,

$$0 \leq Q_{norm}^{min} \leq Q_{norm} \leq Q_{norm}^{max} \leq 1,$$

и область существования значений Q_{norm} на графике зависимости этого показателя от доли $q_{H-II} = N_{H-II} / (|M| \cdot (|M| - 1) / 2)$ вполне определена, что иллюстрирует рисунок 3.6 [54].

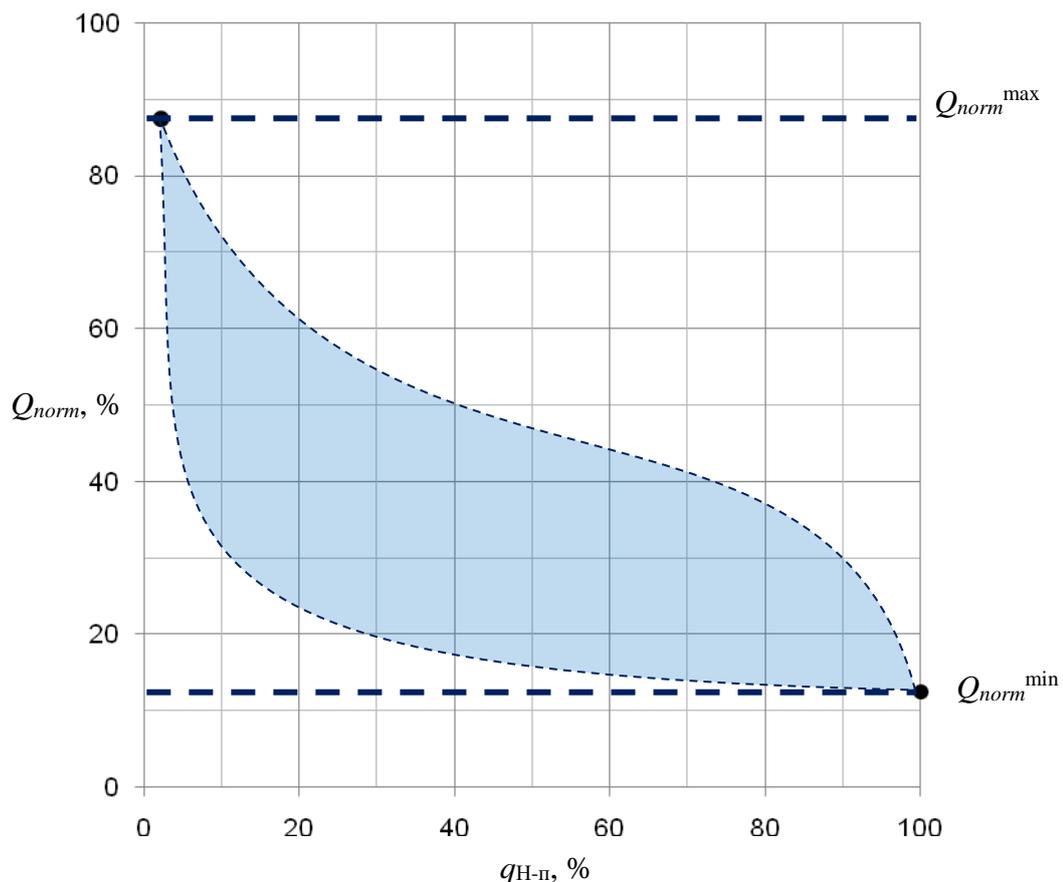


Рисунок 3.6 – Область существования значений показателя Q_{norm}

Оценка Q_{norm} для разных вариантов определения ТТВ-вектора производилась путём имитационного моделирования. Учитывая, что нормализация произ-

водится для каждого объекта в ФК сепарабельно, имитировался контекст с $|G| = 1$. В целом РФК и СИС формировались при заданных M и $q_{H-п}$ и рандомизации множества M_g^* и отношения несовместимости E на M_g^* .

Важный в подобных исследованиях вопрос о надлежащем количестве статистических испытаний решался на основе неравенства Хёфдинга (см. приложение Б), которое даёт экспоненциально убывающие оценки вероятности отклонений сумм независимых и ограниченных случайных величин от среднего этих сумм. Эти оценки точнее, чем аналогичные, полученные с использованием первых или вторых моментов (например, неравенства Маркова, Чебышёва), обеспечивающие лишь степенной закон убывания указанных отклонений.

В качестве примера на рисунке 3.7 представлена экспериментально установленная зависимость средней величины Q_{norm} от доли $q_{H-п}$ при $|M| = 10$ и для разных вариантов определения TTV-вектора [54].

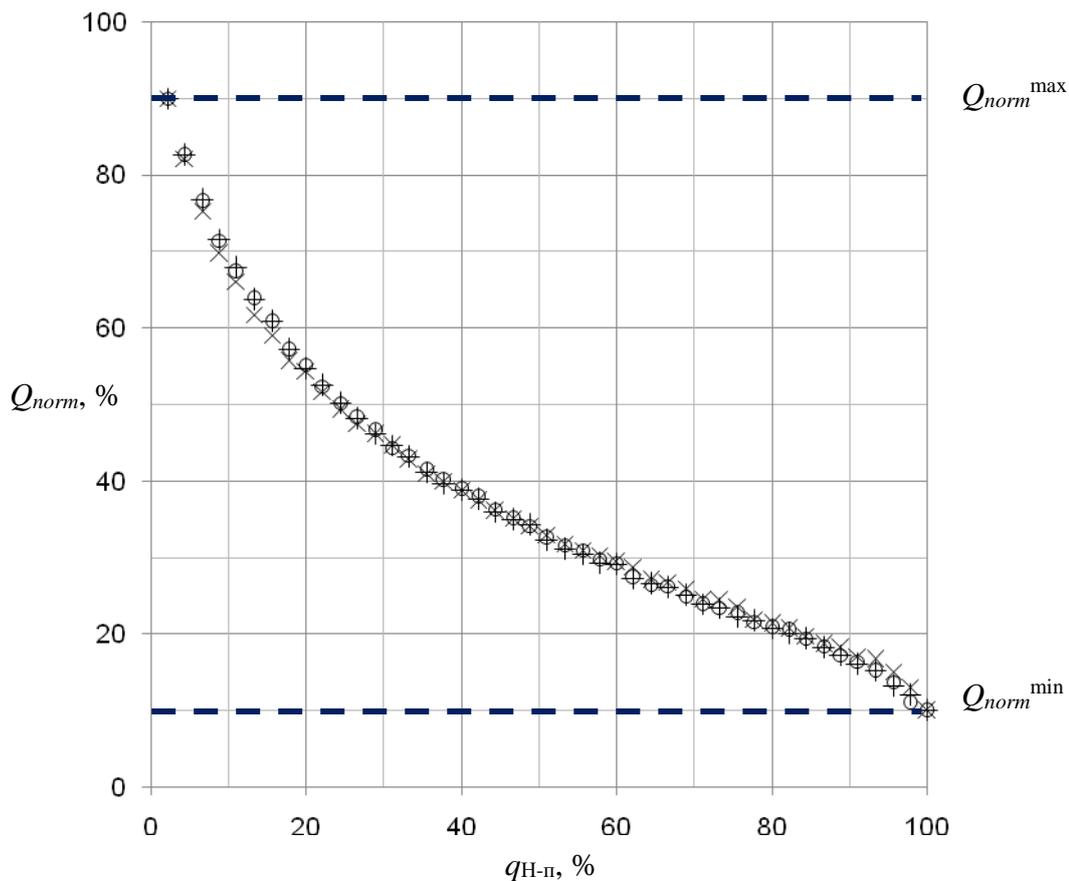


Рисунок 3.7 – Пример зависимости величины мощности нормализованного множества свойств объекта Q_{norm} от доли H-пар $q_{H-п}$ для системы $|M| = 10$ измеряемых свойств.

Варианты определения вектора ужесточения субъективного порога доверия:

+ - $|TTV_{maxmin}|$; × - $\langle \min_{i \in IxPs} \{ |TTV_{Ps,i}| \} \rangle$; ○ - $\langle \min_{i \in IxJPs} \{ |TTV_{JP_s,i}| \} \rangle$

В целом проведённые исследования показали, что с увеличением в ОСС доли Н-пар наблюдается уменьшение показателя Q_{norm} , но сколько-нибудь значимой разницы между разными вариантами определения ТТВ-вектора, и, следовательно, между вариантами вычисления АПД нет, т.е. все варианты обеспечивают получение у объектов в РФК нормальных множеств $M_g^* \setminus M_g^\bullet$ свойств равной мощности. Тогда, для разработанного метода нормализации наилучший вариант определения ТТВ-вектора задается выражением (3.9), т.е. вариант, который согласно (3.14) требует минимальных затрат вычислительных ресурсов.

3.3 Реструктуризация системы измеряемых свойств для ускорения нормализации рабочего контекста

3.3.1 Выявление классов эквивалентности и расширенных ограничений существования на множестве измеряемых свойств

Легко видеть, что отношение обусловленности $C(x, y)$ индуцирует на множестве измеряемых свойств M бинарное отношение взаимообусловленности MC : $MC(x, y) \leftrightarrow C(x, y) \wedge C(y, x)$.

Поскольку MC рефлексивно, симметрично и транзитивно, то оно разбивает множество измеряемых свойств M на классы эквивалентности M° – группы *взаимообусловленных* свойств, или ВЗО-группы. Опираясь на транзитивность отношения C и E -транзитивность, нетрудно определить и выявить важнейшие характеристики расширенных отношений обусловленности C° и несовместимости E° на множестве M° . Относительно C° это произведено в [10]; здесь определим и исследуем расширенное отношение несовместимости E° .

Считаем, что отношение несовместимости E° на множестве M° классов эквивалентности измеряемых свойств есть расширение отношения несовместимости на множестве свойств:

$$E^\circ(X, Y) \leftrightarrow \exists (x \in X, y \in Y) \mid E(x, y), \text{ где } X, Y \in M^\circ, X \cap Y = \emptyset.$$

Поскольку согласно E -транзитивности верны импликации

$$E(x, y) \& MC(x, x_1) \rightarrow E(x_1, y),$$

$$E(x, y) \& MC(y, y_1) \rightarrow E(x, y_1),$$

то с учетом определения расширения отношения несовместимости получаем

$$E^\circ(X, Y) \leftrightarrow \forall(x \in X, y \in Y) E(x, y), X, Y \in M^\circ, X \cap Y = \emptyset.$$

С учётом существования отношений E° и C° под *H-группой* будем понимать набор попарно несовместимых ВЗО-групп, а под *O-группой* – пару ВЗО-групп, связанных расширенным отношением C° . Для каждой группы легко обосновывается связь с нормальностью содержащего их подмножества измеряемых свойств.

Утверждение 3.1. ВЗО-группа может принадлежать нормальному подмножеству измеряемых свойств только целиком.

Доказательство. Пусть $X \subseteq M$ – ВЗО-группа, $N \subseteq M$ – нормальное подмножество измеряемых свойств и $X \cap N \neq \emptyset$. Предположим, что $x, y \in X$, $x \neq y$ и $x \in N$, $y \notin N$. Свойства в X взаимообусловлены, т.е. $(x, y) \in MC$ и, следовательно, $(x, y) \in C$, что согласно условию замкнутости нормальных подмножеств свойств требует $y \in N$. Полученное противоречие доказывает истинность утверждения 3.1.

Утверждение 3.2. O-группа представлена в нормальном подмножестве измеряемых свойств либо целиком, либо своей обуславливаемой частью.

Доказательство. Пусть $X, Y \subseteq M$ – ВЗО-группы, $(X, Y) \in C^\circ$ (т.е. X обуславливающая, а Y обуславливаемая ВЗО-группа O-группы), $N \subseteq M$ – нормальное подмножество измеряемых свойств, в котором представлена O-группа (X, Y) . Допустим, что вариантом такого представления может быть любое подмножество $Z \subseteq X \cup Y$, и покажем, что корректными из них являются только те, что указывает утверждение 3.2, т.е. $Z = X \cup Y$ и $Z = Y$. Действительно, согласно утверждению 3.1 возможны лишь варианты $Z = X$, $Z = Y$ и $Z = X \cup Y$, первый из которых исключается, т.к. при его реализации подмножество N оказалось бы незамкнутым: $\forall y \in Y y \notin N$.

Утверждение 3.3. H-группа может быть представлена в нормальном подмножестве не более чем одним своим членом.

Доказательство. Пусть $X \subseteq M$ – ВЗО-группа, $N \subseteq M$ – нормальное подмножество измеряемых свойств. Предположим, что существует H-группа, включающая X и

некоторое множество X_Y других ВЗО-групп, $|X_Y| \geq 1$. Тогда $\forall Y \in X_Y (X, Y) \in E^\circ$ и, следовательно, $\forall (x \in X, y \in Y) E(x, y)$. Поэтому условием совместимости N будет $X \not\subset N \wedge Y \not\subset N \vee X \subseteq N \wedge Y \not\subset N \vee X \not\subset N \wedge Y \subseteq N$, что и доказывает справедливость утверждения 3.3.

Ключевым для констатации «нормальности» подмножества СИС является

Утверждение 3.4: подмножество СИС нормально тогда и только тогда, когда для него и пересекающихся с ним ВЗО-, О- и Н-групп справедливы утверждения 3.1-3.3.

Доказательство. Действительно, пусть $N \subseteq M$ – нормальное подмножество измеряемых свойств. Тогда для пересекающихся с N ВЗО- и О-групп нарушение утверждений 3.1 и 3.2 означает, что подмножество N незамкнуто, а для пересекающихся с N Н-групп нарушение утверждения 3.3 свидетельствует, что подмножество N несовместимо. Вместе и по отдельности и то, и другое противоречит исходному допущению. С другой стороны, когда для подмножества $N \subseteq M$ выполняются утверждения 3.1-3.3, оно будет замкнуто и совместимо, т.е. нормально.

Теперь нормализацию признакового описания свойств можно существенно ускорить, осуществляя нормализацию признакового описания объекта в РФК «крупными мазками», когда манипуляции со свойствами и связывающими их отношениями обусловленности и несовместимости заменяются оперированием с ВЗО-группами и отношениями C° и E° .

Предложенный метод нормализации свойств объекта в РФК реализуется практически одинаково для описания СИС в формах (M, E, C) и $(M^\circ, E^\circ, C^\circ)$. Дополнительно необходимо лишь предварить циклический процесс нормализации (см. пункт 3.2.3) денонсацией БСС для свойств, входящих в *некомплектные* ВЗО-группы, т.е. такие, где хотя бы одно свойство не входит во множество M° .

Полезный эффект перехода к редуцированному представлению СИС $(M^\circ, E^\circ, C^\circ)$ выражается в сокращении размерности задачи. Механизм такой редукции иллюстрирует рисунок 3.8. Во-первых, уменьшается количество сущностей, состоящих в расширенных отношениях обусловленности и несовместимо-

сти, – вместо свойств рассматриваются их ВЗО-группы: $|M^o| \leq |M|$ (см рисунок 3.8а). Во-вторых, уменьшается количество пар, состоящих в экзистенциальных отношениях, – вместо О- и Н-пар свойств анализируются возникающие О- и Н-пары ВЗО-групп [32] (см. рисунки 3.8б и 3.8в).

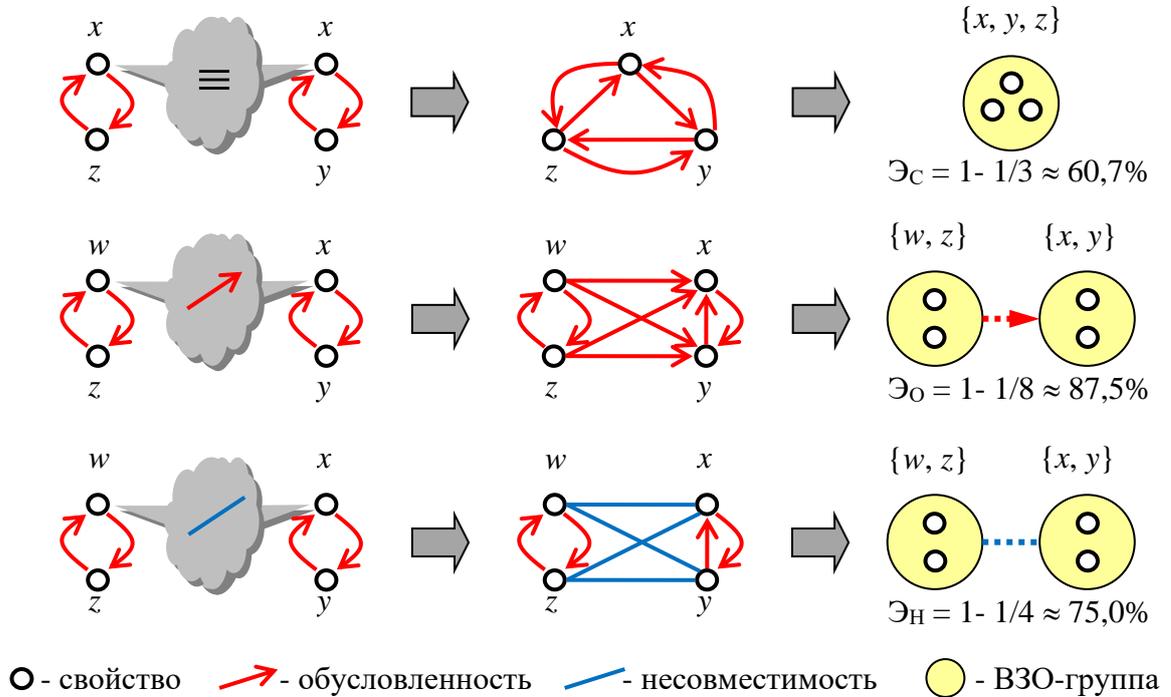


Рисунок 3.8 - Примеры редукции описания системы измеряемых свойств и достигаемый при этом выигрыш (самообусловленность свойств опущена, расширенные экзистенциальные связи показаны пунктирной линией)

На рисунке 3.9 приведено редуцированное описание СИС, исходный вид которой был представлен в подразделе 2.1.2 (рисунок 2.1).

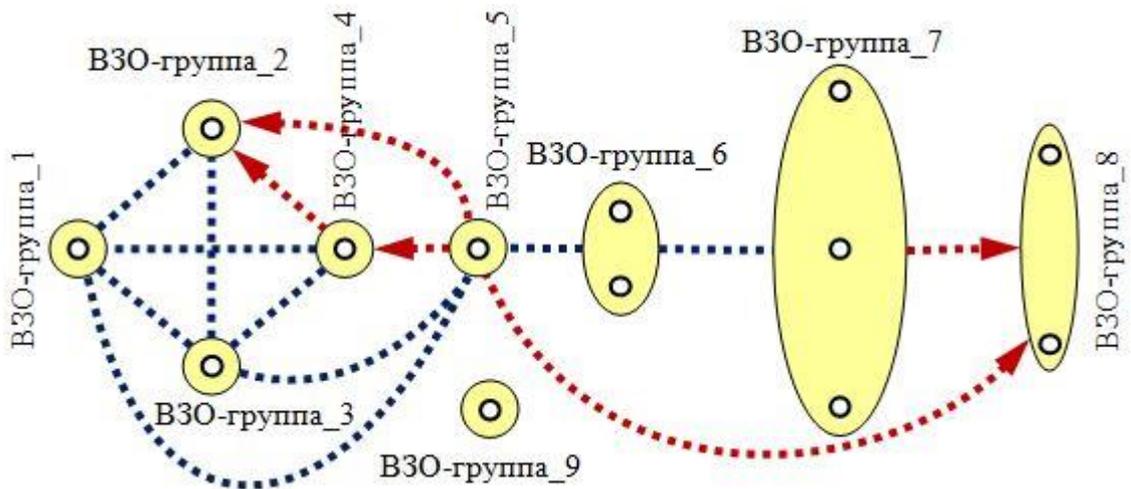


Рисунок 3.9 - Пример описания ограничений существования свойств в форме расширенных отношений обусловленности и несовместимости на множестве групп взаимообусловленных измеряемых свойств

Очевидно, что масштабы редукции описания СИС определяются характеристиками ее исходной модели и, прежде всего, мощностью ее множества ВЗО-пар.

3.3.2 Исследование эффективности реструктуризации системы измеряемых свойств

Количественно эффект редукции ОСС оценивался в имитационных статистических экспериментах, когда при заданной величине $|M|$ структура СИС рандомизировалась при варьировании долями ВЗО-, О-, Н-пар, и исследовался выигрыш от редукции описания СИС – среднее относительное уменьшение количества сущностей $\mathcal{E}_C = 1 - |M^\circ|/|M|$ (см. рисунок 3.8а) и их экзистенциальных связей – обусловленностей $\mathcal{E}_O = 1 - |C^\circ|/|C|$ (см. рисунок 3.8б) и несовместимостей $\mathcal{E}_H = 1 - |E^\circ|/|E|$ (см. рисунок 3.8в). При этом вопрос о необходимом количестве статистических испытаний решался на основе неравенства Хёфдинга (см. приложение Б).

Особенность имитации СИС в исходной форме (M, E, C) заключается в естественных ограничениях на количество одновременно присутствующих ВЗО-, О- и Н-пар – $N_{\text{ВЗО-п}}, N_{\text{О-п}} = |C|, N_{\text{Н-п}} = |E|$ соответственно:

$$2N_{\text{ВЗО-п}} + N_{\text{О-п}} + 2N_{\text{Н-п}} \leq |M| \cdot (|M| - 1),$$

где $N_{\text{ВЗО-п}}, N_{\text{Н-п}} \in \{0, 1, 2, \dots, |M| \cdot (|M| - 1)/2\}, N_{\text{О-п}} \in \{0, 1, 2, \dots, |M| \cdot (|M| - 1)\}$.

Кроме того, некоторые конфигурации СИС (здесь и далее «конфигурация» – это сочетание параметров $|M|, N_{\text{ВЗО-п}}, N_{\text{О-п}}, N_{\text{Н-п}}$) в указанном смысле могут не существовать. Например, легко проверить, что при $|M| = 3$ имеем $N_{\text{ВЗО-п}} \in \{0, 1, 3\}$, т.к. при попытке построить ОСС с $N_{\text{ВЗО-п}} = 2$ вследствие транзитивности обусловленности количество ВЗО-пар закономерно становится равным 3. В присутствии обусловленностей подобные ограничения возникают и для Н-пар.

Отмеченная феноменология формирования СИС потребовала выработки следующей представляющей самостоятельный интерес тактики имитационного статистического моделирования:

- имитация СИС с заданными количествами ВЗО-, О- и Н-пар одновременно рассматривается как *испытание Бернулли*, успех которого констатирует существо-

вание требуемой конфигурации СИС. Если N таких испытаний неудачны (здесь можно воспользоваться тем же значением необходимого числа испытаний, которое определено для «целевого» эксперимента с задачей наблюдения эффекта редукции СИС, ибо результат испытания Бернулли можно интерпретировать как реализацию дискретной случайной величины с двумя возможными значениями – «0» и «1», которые, равно как и оцениваемые показатели эффективности редукции, принадлежат ограниченному диапазону $[0, 1]$), то эксперимент с СИС, определяемой данными параметрами, прекращается;

- при возникновении в результате имитации СИС с заданным количеством ВЗО-, О- и Н-пар закономерного перехода СИС в «состояние», характеризующиеся *иными значениями* $N_{\text{ВЗО-п}}$, $N_{\text{О-п}}$, $N_{\text{Н-п}}$, разумно не считать эту фазу вычислительного эксперимента неудачей, которая заставляет повторить попытку имитации СИС с заданными параметрами, а продолжить имитационный эксперимент (т.е. редуцировать «возникшую» СИС с регистрацией соответствующих эффектов), накапливая статистику для ставшей актуальной конфигурации СИС, и лишь затем обратиться к новой попытке имитации СИС с целевыми параметрами. Фактически это означает «параллельное» выполнение экспериментов для всех актуальных конфигураций СИС, требуя подготовки и использования в общем случае $N_{\text{ВЗО-п}} \times N_{\text{О-п}} \times N_{\text{Н-п}}$ организующих и регистрирующих структур, что, тем не менее, компенсируется сокращением вычислительных затрат на имитационное моделирование.

Эффект уменьшения количества сущностей в редуцированной СИС возникает благодаря наличию в исходном описании СИС ВЗО-пар и усиливается вследствие их пересечений. Поэтому эксперименты для количественной оценки этого эффекта, т.е. определения показателя \mathcal{E}_C , выполнялись для СИС, характеризующихся параметрами $N_{\text{ВЗО-п}} \in \{0, 1, 2, \dots, |M| \cdot (|M| - 1)/2\}$, $N_{\text{О-п}} = 2N_{\text{ВЗО-п}}$ и $N_{\text{Н-п}} = 0$. При этом априори очевидно, что при $N_{\text{ВЗО-п}} = 0$ эффект отсутствует ($\mathcal{E}_C = 0$), а при $N_{\text{ВЗО-п}} = |M| \cdot (|M| - 1)/2$ максимален: $\mathcal{E}_C^{\max} = 1 - 1/|M|$. Таким образом, область существования значений \mathcal{E}_C на графике зависимости этого показателя от доли $q_{\text{ВЗО-п}} = N_{\text{ВЗО-п}} / (|M| \cdot (|M| - 1)/2)$ вполне определена, что иллюстрирует рисунок 3.10.

Результаты конкретных исследований по экспериментальному вычислению \mathcal{E}_C для $|M| = 10$ и $|M| = 20$ представлены на рисунке 3.11. Исследования показали, что с увеличением в СИС доли ВЗО-пар наблюдается почти экспоненциальный с насыщением рост выигрыша, оцениваемого показателем \mathcal{E}_C . Подобная затухающая экспонента отклика характерна для систем с отрицательной обратной связью, хотя в области насыщения наблюдается множество «разрывов» – отсутствие конфигураций СИС с определенным количеством ВЗО-пар – и существенный разброс величины выигрыша. Такая картина выигрыша \mathcal{E}_C принципиально не меняется при варьировании количеством измеряемых свойств.

В приложении В для графиков, показанных на рисунке 3.11, приведены табличные представления, где явно отмечены те количества ВЗО-пар (и доли $q_{\text{ВЗО-п}}$), для которых невозможно построение СИС требуемой конфигурации, и возникают «разрывы» исследуемых зависимостей.

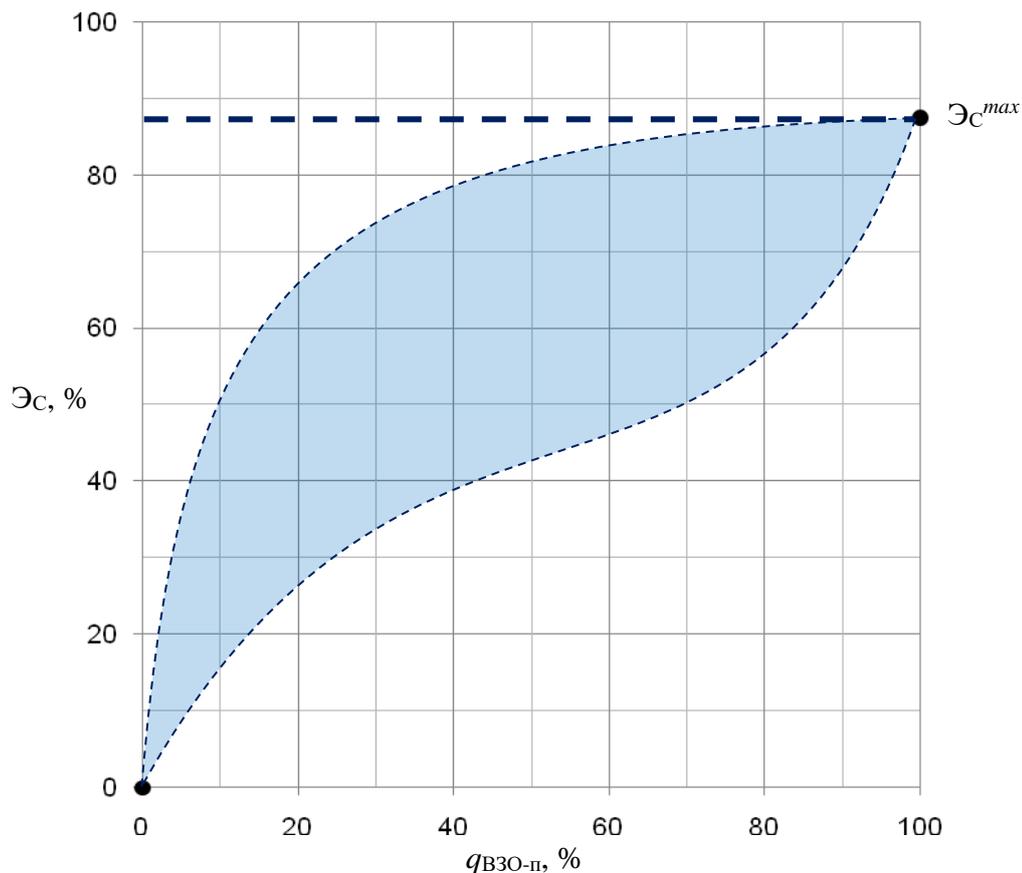


Рисунок 3.10 – Область существования значений показателя \mathcal{E}_C – среднего относительного сокращения количества сущностей в обобщенном представлении системы измеряемых свойств в зависимости от доли ВЗО-пар $q_{\text{ВЗО-п}}$ в исходном представлении этой системы

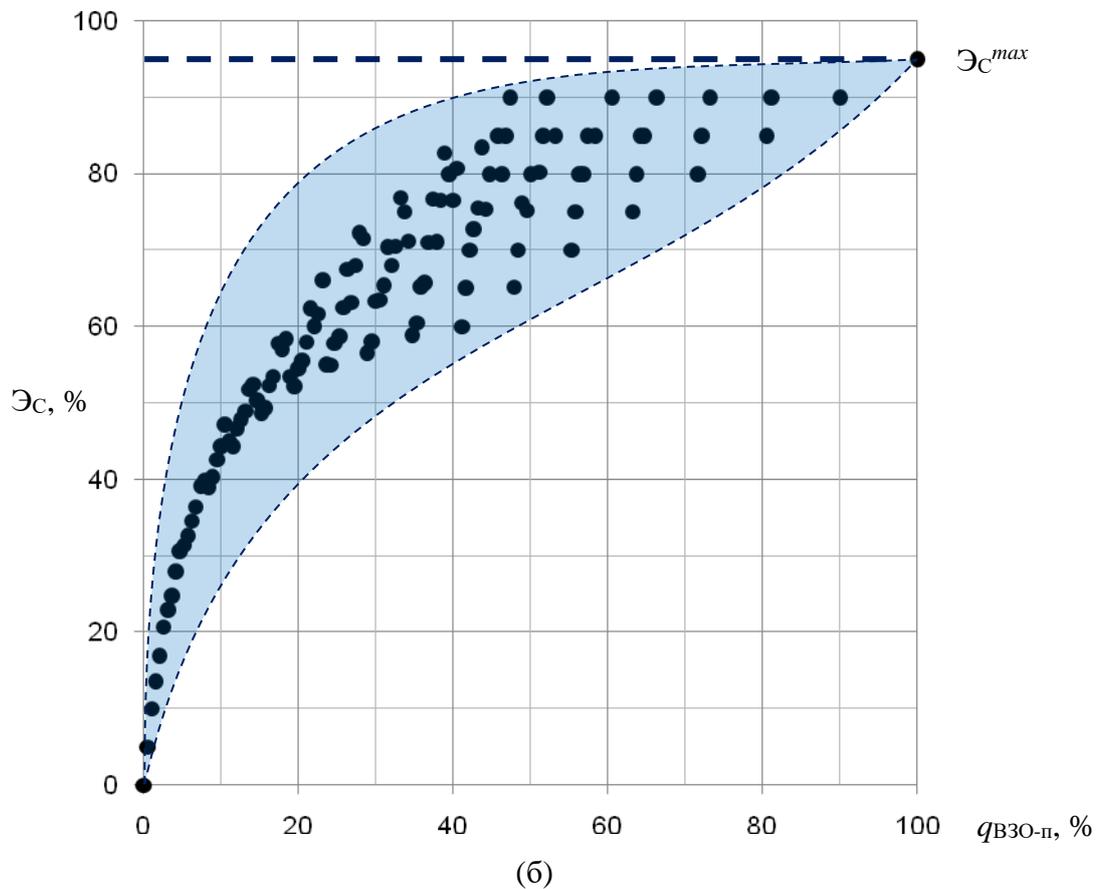
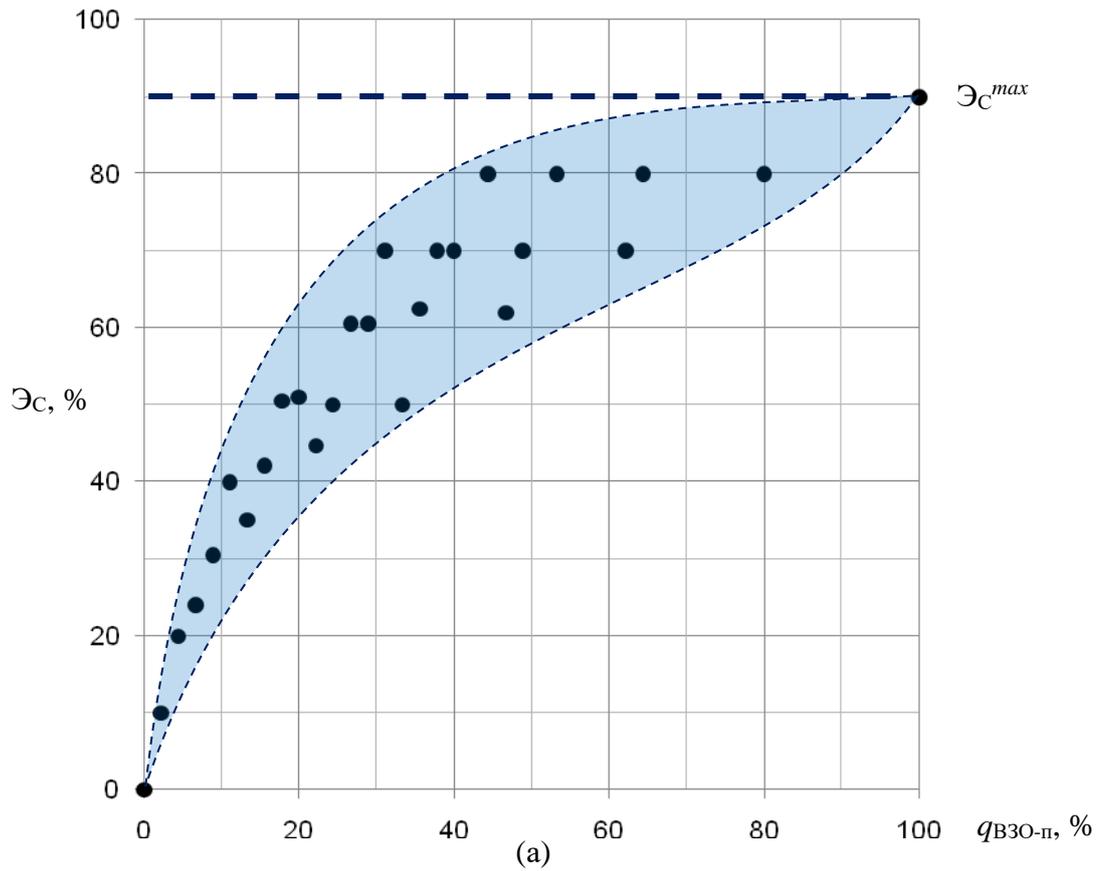


Рисунок 3.11 – Значения показателя \mathcal{E}_c – среднего относительного сокращения количества сущностей в редукции представлении системы $|M| = 10$ (а) и $|M| = 20$ (б) измеряемых свойств - в зависимости от доли ВЗО-пар $q_{\text{ВЗО-п}}$ в исходном представлении этой системы

На рисунке 3.12 приведен пример экспериментально установленного среднего относительного сокращения количества обусловленностей \mathcal{E}_O в редуцированной СИС для $|M| = 10$ (рисунок 3.12а) и $|M| = 20$ (рисунок 3.12б) в зависимости от количества ВЗО-пар $N_{\text{ВЗО-п}}$ и доли О-пар $q_{\text{О-п}} = N_{\text{О-п}} / (|M| \cdot (|M| - 1))$ при $N_{\text{Н-п}} = 0$ в исходном представлении СИС. Выигрыш \mathcal{E}_O сначала значительный (и тем выше, чем больше ВЗО-пар в (M, E, C)) довольно быстро снижается. Это можно объяснить тем, что при фиксированном $N_{\text{ВЗО-п}}$ рост $q_{\text{О-п}}$ происходит за счет О-пар, не пересекающихся с существующими ВЗО-парами (пересечение вызвало бы образование дополнительных ВЗО-пар, а их количество фиксировано) и не образующих новых ВЗО-пар, а обусловленности, определяемые такими «чистыми» О-парами, сохраняются в редуцированной СИС в неизменном виде, сокращая выигрыш.

Правую границу области существования графиков \mathcal{E}_O определяет максимум количества «чистых» О-пар — $|M| \cdot (|M| - 1) / 2 - N_{\text{ВЗО-п}}$; левую — количество О-пар, образующих исходное количество ВЗО-пар в СИС при отсутствии «чистых» О-пар; доля таких О-пар составляет $2N_{\text{ВЗО-п}} / (|M| \cdot (|M| - 1))$.

Рисунок 3.13 показывает пример среднего относительного сокращения количества несовместимостей \mathcal{E}_H в редуцированной СИС для $|M| = 10$ (рисунок 3.13а) и $|M| = 20$ (рисунок 3.13б) в зависимости от количества ВЗО-пар $N_{\text{ВЗО-п}}$ и доли Н-пар $q_{\text{Н-п}} = N_{\text{Н-п}} / (|M| \cdot (|M| - 1) / 2)$ при $N_{\text{О-п}} = 0$ в исходном представлении СИС. Наблюдаются следующие характерные черты зависимостей такого вида: рост выигрыша \mathcal{E}_H при увеличении $N_{\text{ВЗО-п}}$; очень быстрое достижение с ростом $q_{\text{Н-п}}$ максимальной величины выигрыша, который медленно и почти линейно уменьшается до достижения величиной $N_{\text{Н-п}}$ своего максимума $|M| \cdot (|M| - 1) / 2 - N_{\text{ВЗО-п}}$, который определяет правую границу области существования графиков \mathcal{E}_H ; нулевой выигрыш в начале всех графиков \mathcal{E}_H — при $N_{\text{Н-п}} = 1$, — поскольку существование в СИС единственной Н-пары возможно лишь при отсутствии ее пересечения с существующими в СИС ВЗО-парами, а несовместимость, определяемая такими Н-парами, сохраняется в редуцированном описании СИС в неизменном виде [54].

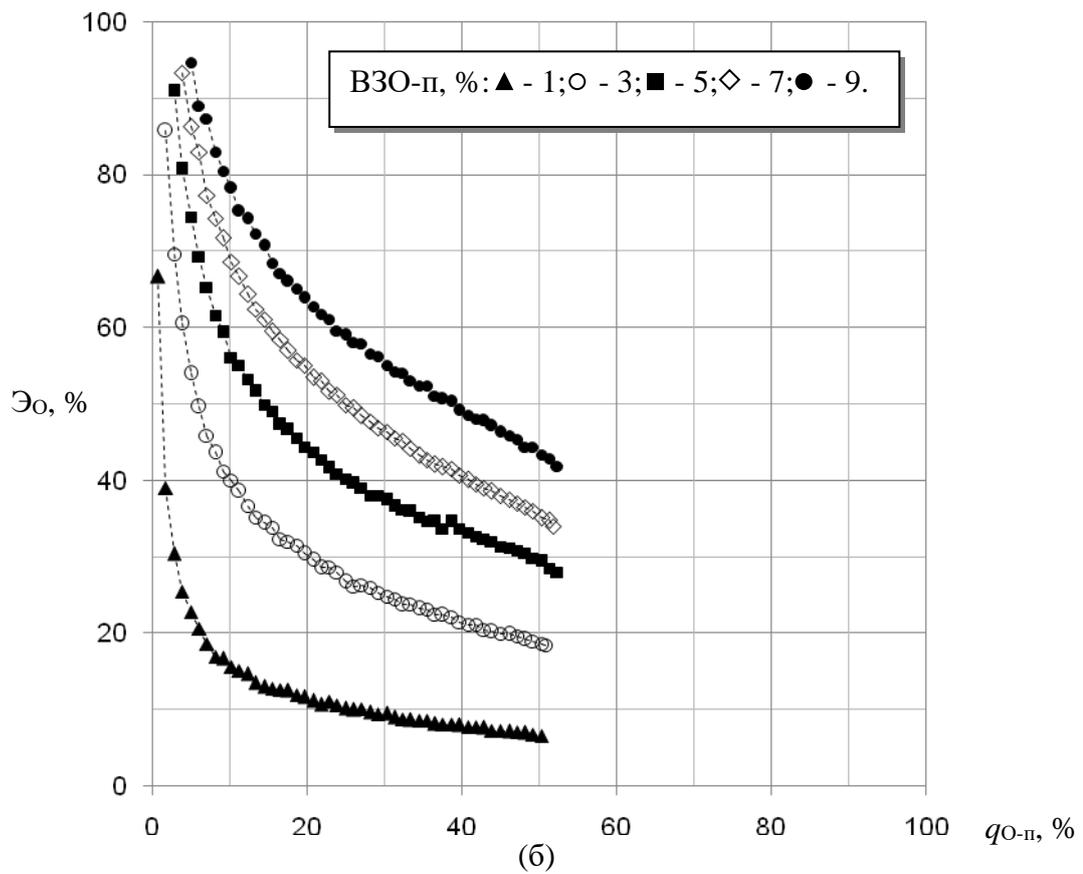
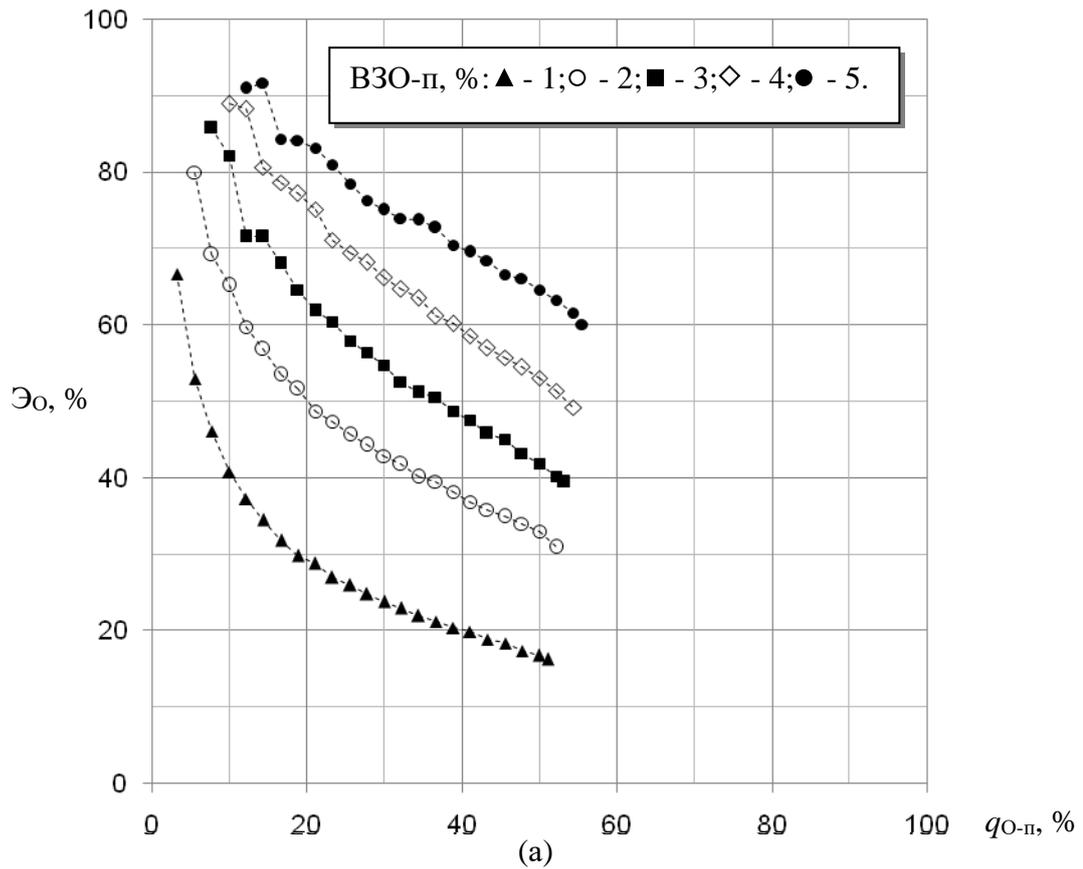


Рисунок 3.12 – Значения показателя \mathcal{E}_0 – среднего относительного сокращения количества обусловленностей в обобщенном представлении системы $|M| = 10$ (а) и $|M| = 20$ (б) измеряемых свойств - в зависимости от долей ВЗО-пар $q_{ВЗО-п}$ и чистых О-пар $q_{O-п}$ в исходном представлении этой системы

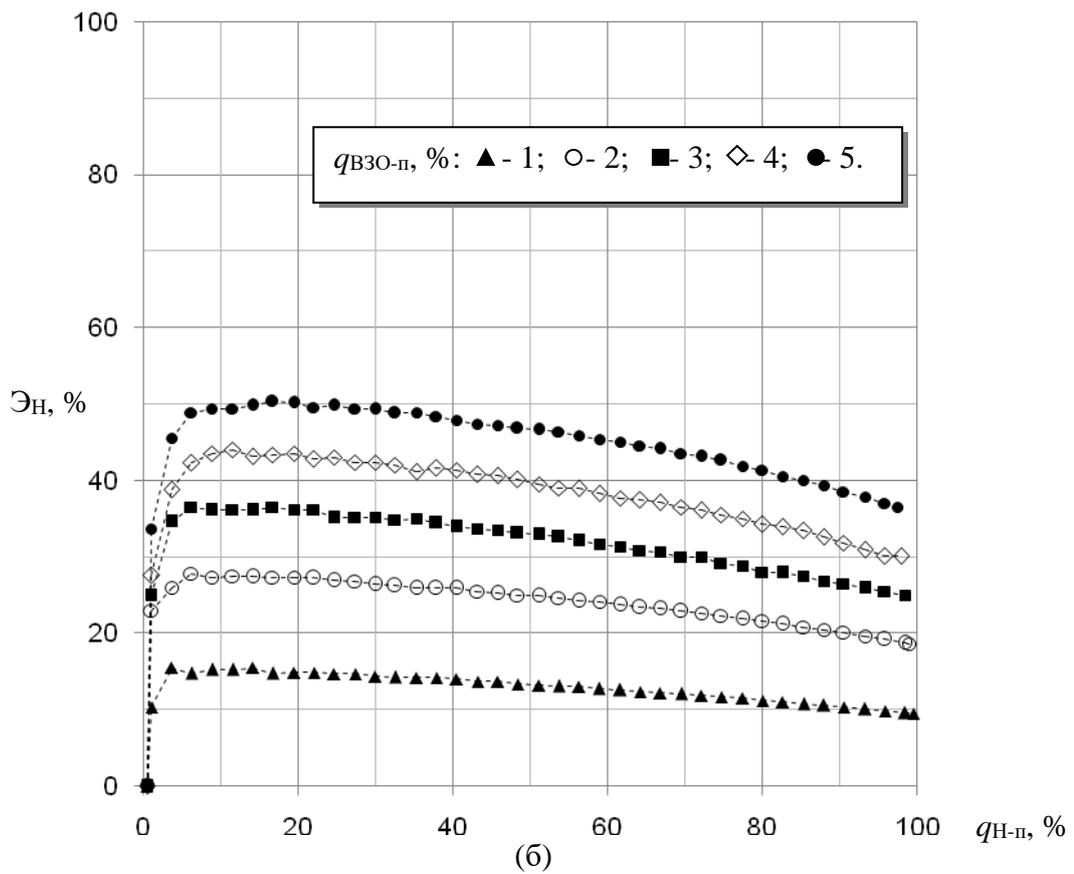
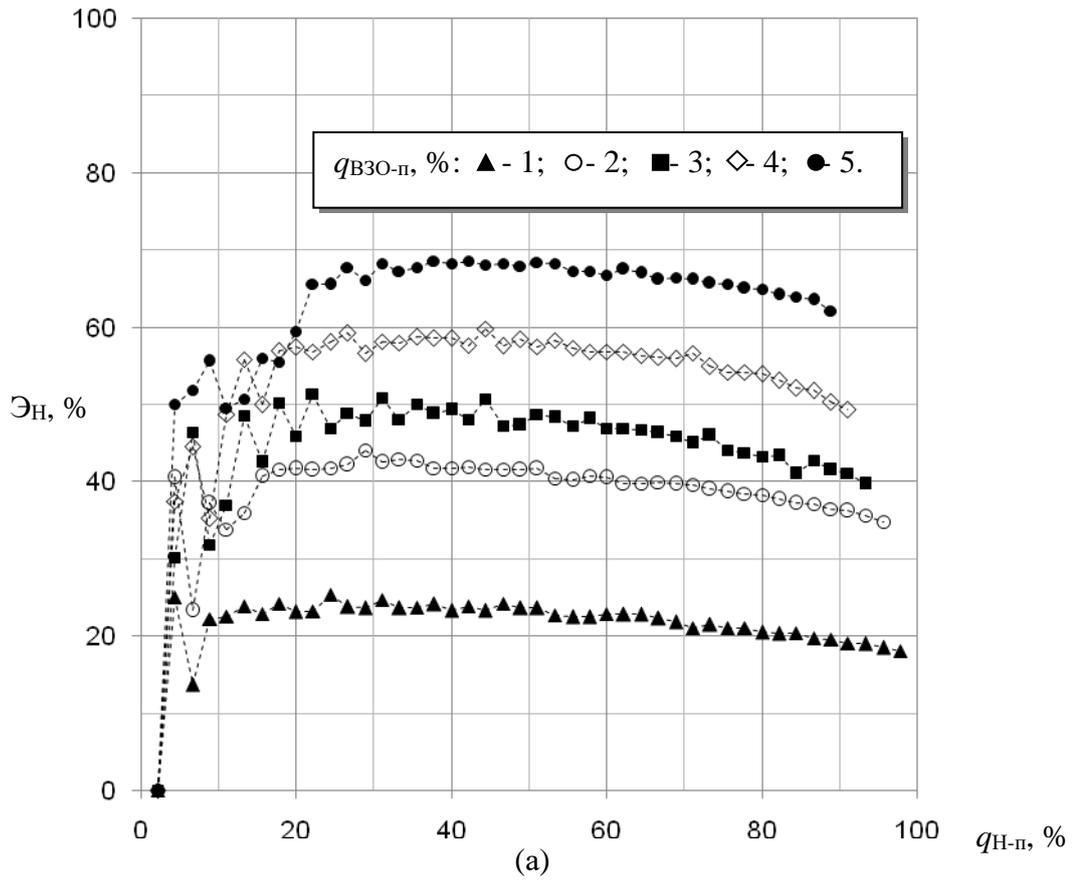


Рисунок 3.13 – Значения показателя \mathcal{E}_H – среднего относительного сокращения количества несовместимостей в обобщенном представлении системы $|M| = 10$ (а) и $|M| = 20$ (б) измеряемых свойств - в зависимости от долей ВЗО-пар $q_{ВЗО-п}$ и Н-пар $q_{Н-п}$ в исходном представлении этой системы

В целом проведенное имитационное исследование позволило количественно оценить потенциал снижения размерности ОСС при редукции описания СИС на основе выявления классов эквивалентности измеряемых свойств и выявить характер зависимости получаемого выигрыша от параметров СИС.

5.4 Выводы по главе 3

- 1) Формирование рабочего контекста для вывода формальных понятий предусматривает обязательную дефаззификацию и, при необходимости, нормализацию исходного эмпирического формального контекста.
- 2) Дефаззификация исходного эмпирического формального контекста, где оценка базовых семантических суждений об исследуемой предметной области дана в рамках V^{TF} -логики, производится либо путем исключения заведомо ложных суждений, либо реализуется как когнитивный акт субъекта-исследователя путем задания субъективного порога доверия к эмпирическим данным.
- 3) Идея предложенного подхода к нормализации рабочего формального контекста состоит в том, что задачу получения признакового описания объекта, которое удовлетворяет ограничениям существования, можно решить путем отказа от части эмпирической информации, имеющей при прочих равных условиях меньшую достоверность. Разработанный метод нормализации рабочего контекста опирается на совокупность хорошо обоснованных математических моделей и рассматривается как когнитивный инструмент субъекта-исследователя. Включенная в метод задача комбинаторной оптимизации эффективно решается за счет сокращенного обхода конечного дерева возможных решений.
- 4) Для повышения вычислительной эффективности разработанного метода нормализации теоретически и экспериментально подтверждена целесообразность редукционирования описания системы измеряемых свойств на основе выделения в ней классов эквивалентности измеряемых свойств и расширенных экзистенциальных отношений на множестве этих классов. Соответствующую трансформацию исходного описания системы измеряемых свойств целесообразно считать отдельным обязательным этапом онтологического анализа данных.

4 РАЗРАБОТКА ПРОГРАММНОЙ ЛАБОРАТОРИИ ДЛЯ ОНТОЛОГИЧЕСКОГО АНАЛИЗА ДАННЫХ

В разделе описывается алгоритмическое и программное обеспечение разработанных методов, направленных на подготовку контекста для вывода формальных понятий в рамках онтологического анализа данных. Приведенные результаты отражены в публикациях [11, 32, 62-67].

При программной реализации разработанных моделей и методов сложилась следующая пошаговая схема ОАД (рис. 4.1 и рис. 4.2). На схеме выделены когнитивные акты субъекта-исследователя ПрО и отражены этапы (шаги) и продукты ОАД:

- 1, 4, 8, 17 — осмысление субъектом, находящимся на определенном этапе решения ОАД-задачи, либо результата *предшествующего*, либо результата *любого из последующих* шагов анализа. Например, 4-й шаг может быть инициирован восприятием субъектом как ОТОС, так и СИС/ГСС, ИФК, РФК или ФП. Для исключения нагромождения связей такие переходы изображены лишь в пределах выделенных когнитивных актов;
- 2, 5 — формирование/редактирование соответственно СИС и ОТОС, включая концептуальное шкалирование измеряемых свойств. Обратные по отношению к субъекту воздействия связаны с информированием о невозможности выполнить выбранное действие;
- 3 — измерение (в широком смысле) свойств объектов ПрО;
- 6 — построение СИС/ГСС и ИФК:
 - СИС/ГСС – представление СИС в виде множества пересекающихся ГСС, однородных по виду экзистенциального сопряжения;
 - ИФК – продукт трех последовательных действий:
 - (1) логической интерпретации каждого результата измерения в ОТОС с учетом степени доверия к процедуре измерения и достоверности серии измерений, что в целом доставляет оценку истинности БСС, соответствующего рассматриваемому результату измерения;

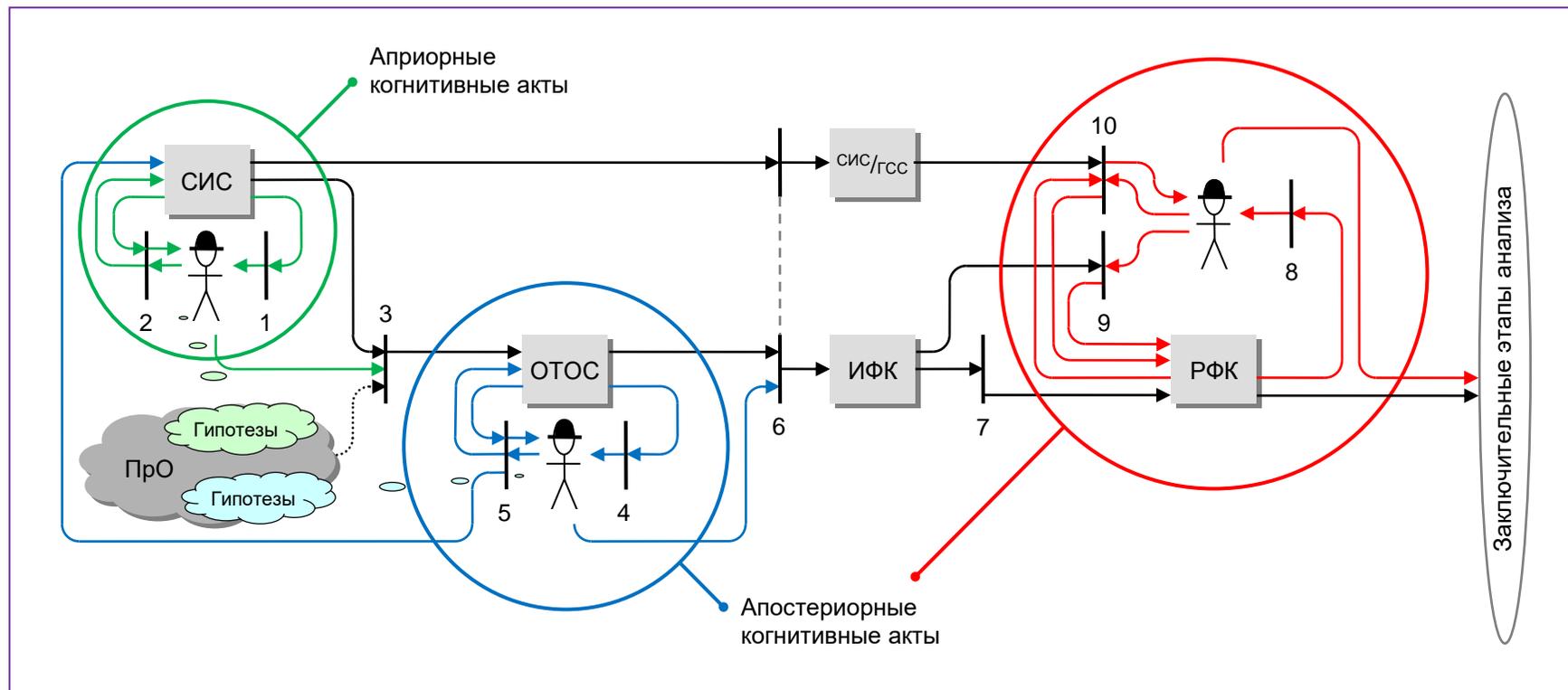


Рисунок 4.1 – Пошаговая схема онтологического анализа данных: подготовка данных и выполняемые при этом когнитивные акты

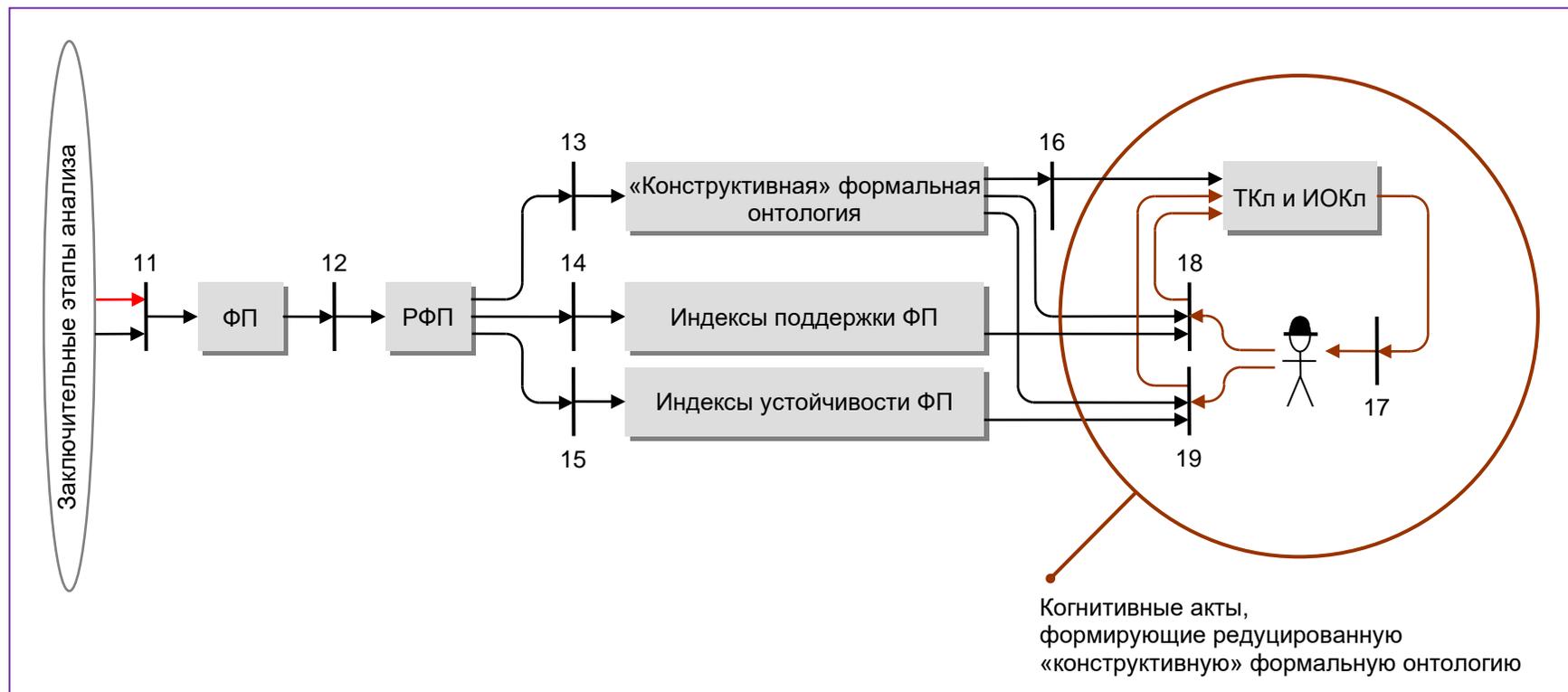


Рисунок 4.2 – Пошаговая схема онтологического анализа данных: заключительные этапы анализа и когнитивные акты, редуцирующие его основной продукт

- (2) консолидации (в смысле совмещения, накопления свидетельств) полученных в первом действии оценок истинности для каждого множества однородных БСС (БСС относительно одной и той же пары «объект-свойство»);
- (3) отнесения в образовавшемся нестрогом ФК всех неопределенных, противоречивых, «неопределенных с долей лжи» и «лживых с долей правды» БСС к ложным, а остальных – к истинным (это определяет сечение нестрогого ФК с порогом доверия $\langle 0+; 100- \rangle\%$, где «+» и «-» обозначают соответственно увеличение и уменьшение чисел на предельно малую величину $\varepsilon > 0$);
- 7 — исходное отождествление РФК с ИФК;
 - 9 — дефаззификация ИФК: установка СПД к БСС и обновление РФК, путем преобразования ИФК в РФК с отнесением к ложным всех БСС, оценка истинности которых не достигает СПД, а остальных – к истинным (т.е. построение РФК как стандартного α -сечения ИФК с порогом, равным СПД);
 - 10 — нормализация РФК: преобразование РФК, в ходе которого снимаются противоречия между эмпирической информацией и априорными гипотезам об искомым ФП — множество свойств каждого отдельно взятого объекта измерения нормализуется путём отнесения части БСС, до этого считавшихся истинными, к ложным при минимизации АПД и количества денонсируемых БСС. Обратное по отношению к субъекту воздействие связано с информированием о невозможности выполнить данное действие;
 - 11 — извлечение из РФК всех ФП;
 - 12 — построение РФП;
 - 13 — извлечение «конструктивной» формальной онтологии, описывающей ПрО, объектно-признаковые данные о которой предоставляет РФК;
 - 14, 15 — вычисление поддержки и устойчивости ФП;
 - 16 — показ извлеченной «конструктивной» формальной онтологии с построением таксономии классов (ТКл) объектов ПрО и с идентификацией всех интен-

сиональных отношений классов (ИОКл).

Замечание: ФП, включаемые в онтологию, именуются классами;

ТКл – экстенциональная сеть ФП, используемая для представления формальной онтологии ПрО. ИОКл – интенциональное отношение классов.

- 18 — показ «конструктивной» формальной онтологии с построением ТКл, при условии, что поддержка включаемых в неё классов (возможно, модифицированных в части отличительного содержания) составляет не менее субъективно указываемого порога, и с идентификацией всех ИОКл при указанном пороге поддержки;
- 19 — показ «конструктивной» формальной онтологии с построением ТКл, при условии, что устойчивость включаемых в неё data-классов (возможно, модифицированных в части отличительного содержания) составляет не менее субъективно указываемого порога, и с идентификацией всех ИОКл при указанном пороге устойчивости.

4.1 Выбор программной платформы

Разработанная методика в полной мере реализована в программной лаборатории для апробации разработанных моделей, методов и алгоритмов ОАД *Onto-Worker* (OW) [67], о государственной регистрации которой имеется свидетельство (см. Приложение Г). В качестве программной платформы был избран MS Excel.

Достоинства Excel общеизвестны. Это широчайшее распространение и признание; действительно богатейшая функциональность; простой, но достаточно мощный VBA – встроенный язык программирования с развитой интегрированной средой разработки приложений, включающей средства визуального проектирования пользовательских интерфейсов и инструменты автоматизации программирования [68]; это возможность использования стандартных для MS интерфейсов межзадачных коммуникаций для общения с другими приложениями. И, несомненно, для быстрого прототипирования программных продуктов наиболее ценным качеством Excel является редкое *единство* среды разработки пользовательских приложений и хранения данных.

4.2 Архитектурное проектирование лаборатории

4.2.1 Функциональность и архитектура

Для описания и структурирования функциональности, предоставляемой пользователю, общепризнано удобство диаграмм вариантов использования (use case diagram) Джекобсона [69], вошедших без существенных модификаций в де факто стандарт проектирования UML.

Диаграмма вариантов использования OW представлена на рисунке 4.3.

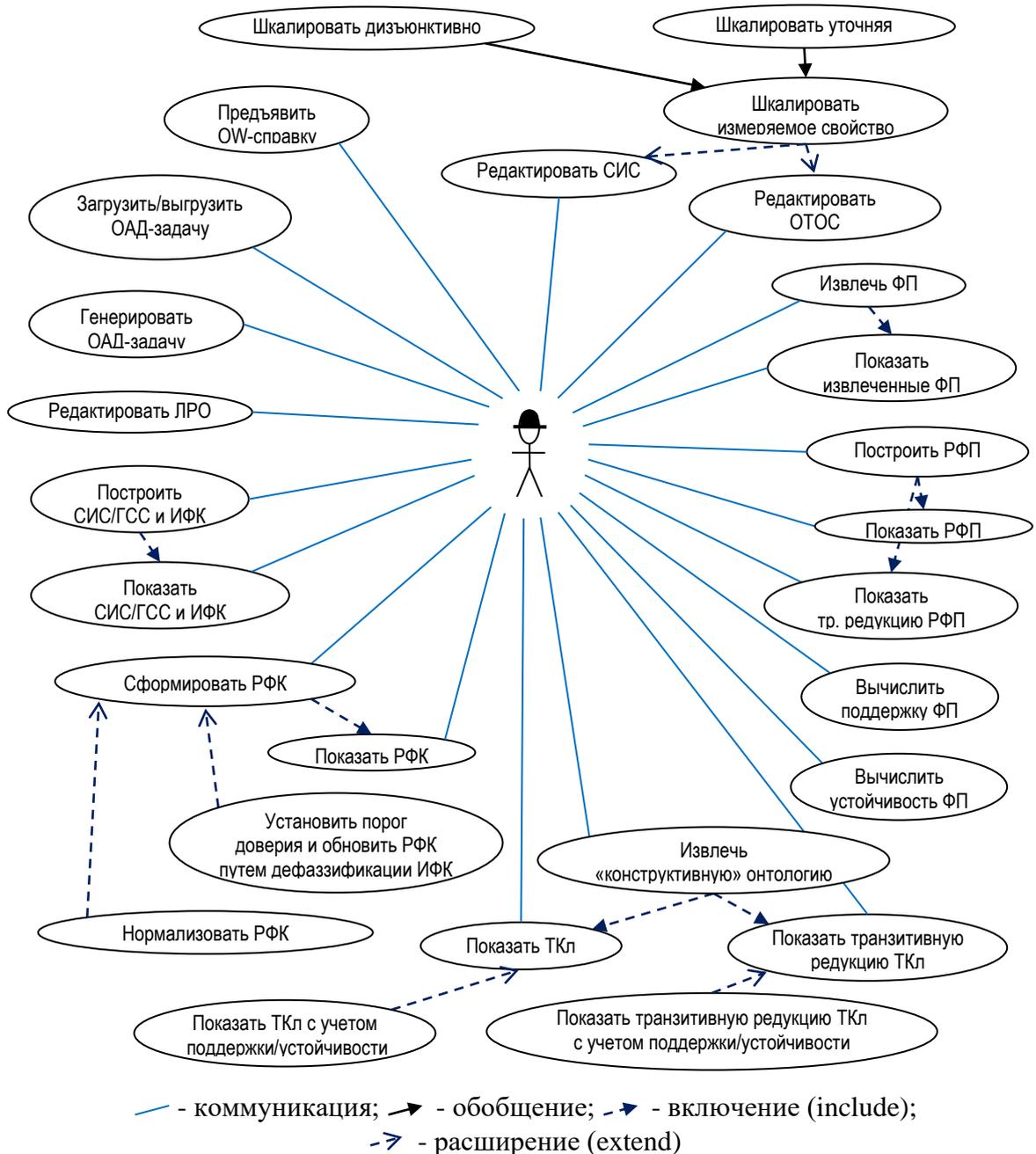


Рисунок 4.3 – Диаграмма вариантов использования программной лаборатории для онтологического анализа данных OntoWorker:

Более подробную информацию о функциях, интерфейсе, уровне автоматизации обработки данных в OW можно найти в подразделе 4.3.

OW представляет собой Excel-книгу, включающую набор структурированных рабочих листов для представления исходных данных, результатов решения и различной сопровождающей информации задачи вывода формальных понятий и онтологий, или ОАД-задачи. ОАД-задачи хранятся в отдельных Excel-книгах с фиксированной структурой (OW-книгах), которые могут открываться и обрабатываться *OntoWorker*’ом по отдельности [67].

Архитектура разработанного программного комплекса представлена в виде интеллект-карты на рисунке 4.4, её составляющие перечислены далее:

- для ввода, редактирования и отображения СИС и ОТОС используется одноименные рабочие листы Excel, которые имеют зонную организацию. Эта организация определяет табличное представление СИС как соответствия «сопряженные пары свойств – свойства» (сопряженная пара свойств – это либо пара несовместимых свойств – Н-пара, либо пара свойств с обусловленностью – О-пара) и табличное представление ОТОС как соответствия «серии измерений – процедуры измерения». Содержание этих листов защищено, и его изменение возможно только специализированными сервисами, интерфейс к которым осуществляется с помощью различных контекстных меню, локализованных в специализированных зонах этих листов. Наконец, процедура обработки события BRC («Before Right Click») служит здесь и в рамках других OW-листов для вызова контекстных меню, а в случае зонной организации листов – еще и для идентификации зоны, где пользователь кликнул правой кнопкой мыши;
- лист ЛРО служит для размещения любой информации (тексты, рисунки, мультимедиа и т.п.) сопровождающей загруженную ОАД-задачу. Контекстное меню листа позволяет в частности временно снимать его защиту, открывая пользователю возможность редактирования содержания листа ЛРО с помощью всей инструментальной палитры Excel;

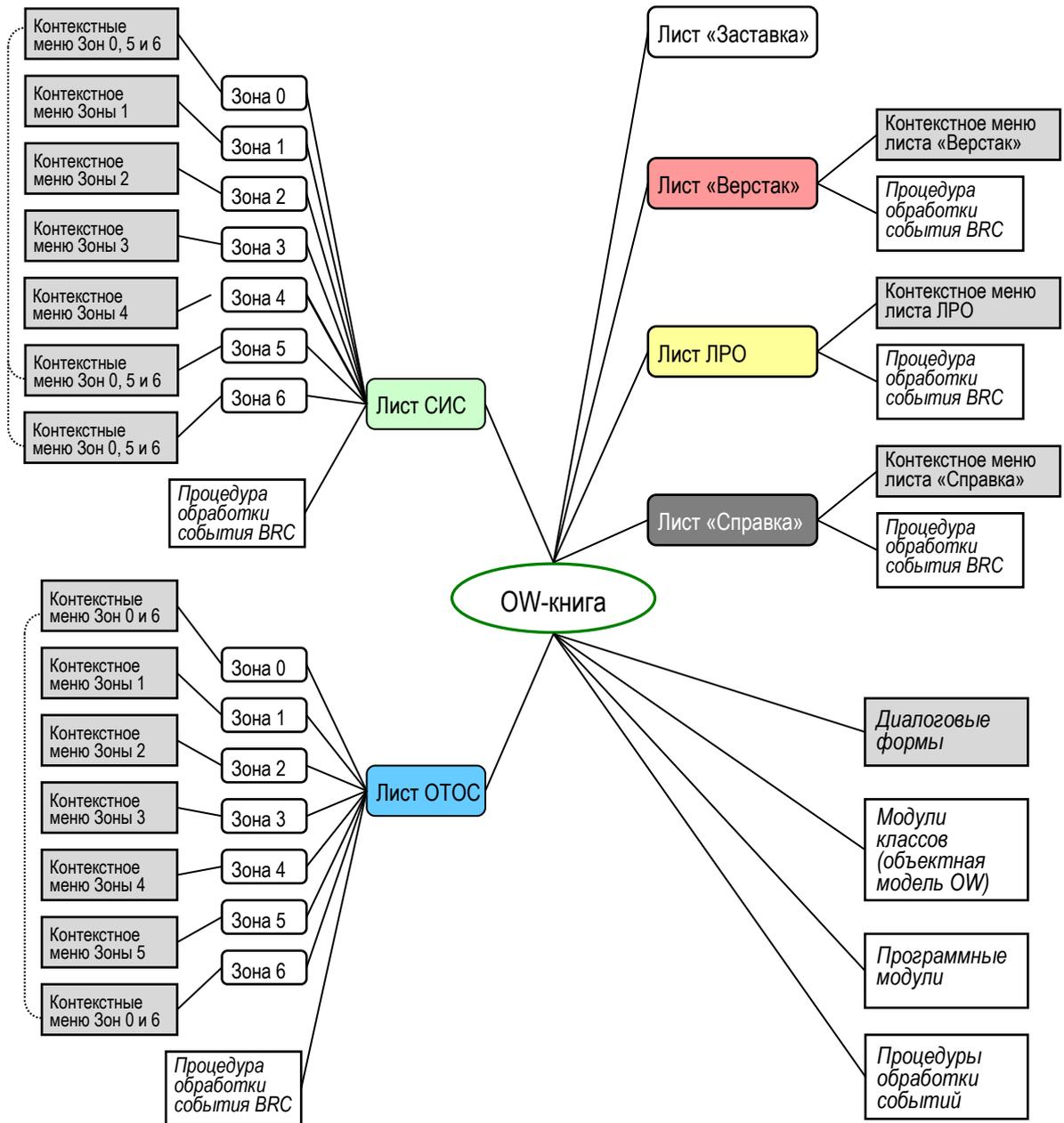


Рисунок 4.4 – Архитектура программной лаборатории для онтологического анализа данных OW

- лист «Справка» содержит справочную информацию о программной лаборатории, включая используемые термины и определения, описание функциональных возможностей и архитектуры OW. Контекстное меню листа ограничивается навигационными задачами;
- на листе «Верстак» отображаются результаты работы OW. Контекстное меню листа ориентировано, прежде всего, на предоставление пользователю основных аналитических функций программного приложения, но, в частности, поз-

воляет временно снимать защиту листа для его произвольного редактирования стандартными средствами Excel;

- лист «Заставка» играет исключительно декоративную роль и появляется при открытии OW-книги прежде, чем будет развернут формально основной в программной лаборатории лист СИС;
- наряду с контекстными меню интерактивное взаимодействие пользователя и OW поддерживается с помощью диалоговых форм, а одностороннее – от программных компонентов OW к пользователю – с помощью сообщений;
- объектная модель OW насчитывает почти полтора десятка классов, связанных помимо отношения наследования дюжиной различных интенциональных бинарных отношений (тех, что при описании классов представляют бинарные отношения, в которых состоят экземпляры соответствующих классов, образуя экстенциональную сеть экземпляров);
- функциональные сервисы OW-листов объединены в одноименные программные модули OW-книги. К ним добавлен модуль, содержащий описания разделяемых переменных и сервисов;
- процедуры обработки событий OW-книги служат для реализации надлежащих реакций на ее открытие, закрытие и попытку сохранения (последнее действие некорректно, и OW прекращает работу, прежде предоставляя пользователю возможность сохранить ее результаты).

Контекстные меню всех листов OW-книги показаны в Приложении Д.

4.2.2 Концептуальная модель предметной области задачи онтологического анализа данных

Представление предметной области OW – в современных программных приложениях его принято называть «объектной моделью» – естественно является формальной онтологией. В этом смысле в целом вполне пригодная диаграмма классов UML проигрывает, т.к. «отягощена» символикой, более ориентированной на программистов. Как отмечено в [70], [семантические сети признаны как наиболее удобная форма концептуализации знаний. Семантические сети состоят из уз-

лов и упорядоченных отношений (связей), соединяющих эти узлы. Узлы выражают понятия или предположения, а связи описывают взаимоотношения между этими узлами.

Для представления знаний обычно используют неоднородные семантические сети, в которых используют разные типы отношений.] Диаграмма классов (рисунок 4.5) была разработана в виде интенциональной семантической сети.

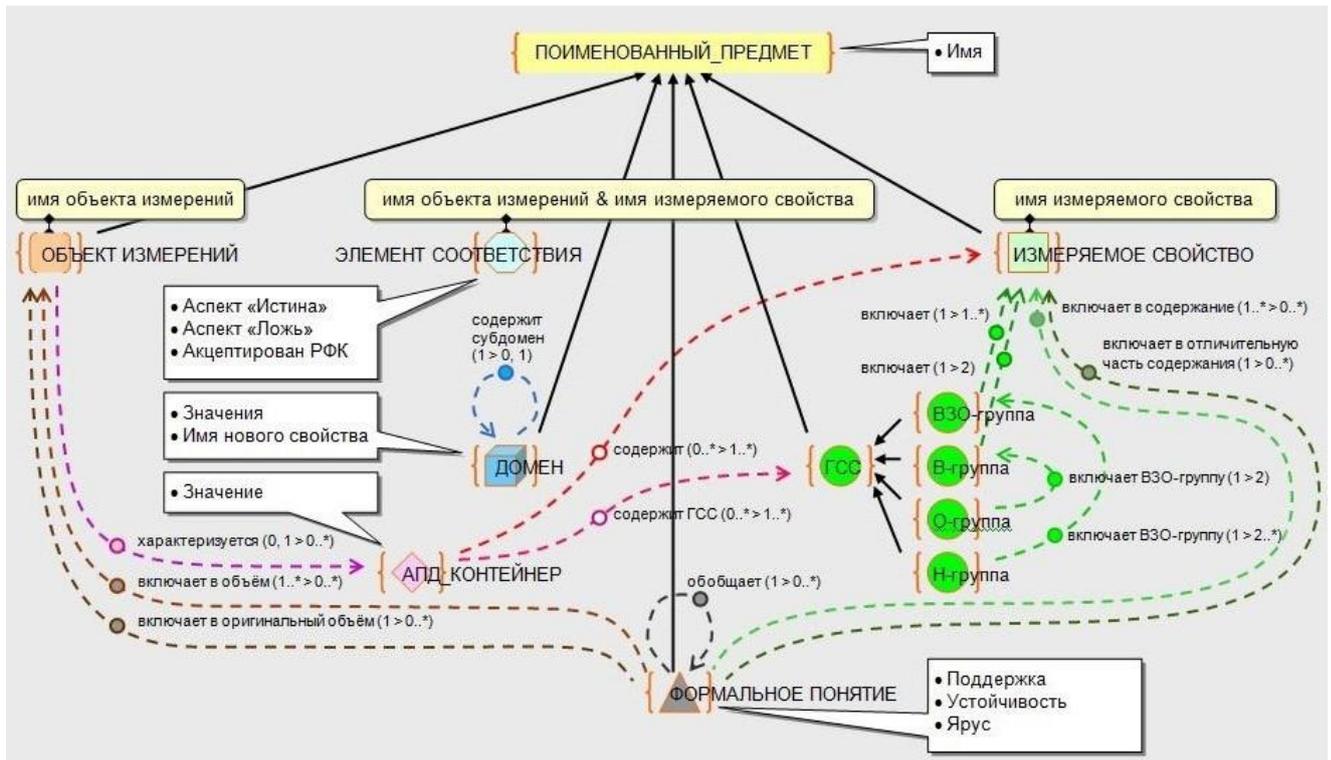


Рисунок 4.5 – Диаграмма классов

Корневые вершины графа транзитивной редукции отношения наследования – терминальные вершины-истоки непомеченных маркером-кружком сплошных дуг на рисунке 4.5 – определяют классы объектов объектно-ориентированной модели данных программной лаборатории OW.

Пунктирные маркированные дуги указывают отношения между объектами классов. Кардинальность этих отношений различна, и встречаются все ее разновидности: например, отношение «представлен серией» имеет кардинальность «один ко многим», «принадлежит серии» - «один к одному», а «обобщает» - «многие ко многим».

В целом рассматриваемая модель данных отчетливо воспроизводит «физические» модели ПрО ОАД, представленные на рисунке 4.1.

4.2.3 Организация интерфейса с пользователем

Пользовательский интерфейс разработанной программной лаборатории содержит контекстное меню, формы и интерактивные диаграммы.

Для ввода, редактирования и просмотра исходных данных, промежуточных и итоговых результатов OW оснащен современными средствами графического WIMP (*Windows, Icons, Menus, Pointers*)-интерфейса. Как обычно в Excel-приложениях в OW множество необходимых рабочих окон заменяется совокупностью стандартно управляемых рабочих листов книги-носителя приложения (или OW-книги). На рабочих листах OW-книги интерфейс пользователя составляют интерактивные таблицы, специализированные контекстные меню и формы. Кроме того, возможности основного меню Excel в полном объеме доступны на незащищенных листах «Верстак» и «ЛРО» [65].

На рисунках 4.6-4.8 приведены скриншоты листов «СИС» и «ОТОС», иллюстрирующие моменты когнитивной деятельности субъекта при решении задач ОАД.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1			m	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12
2	Н-пара	m_4 - m_5				x	x							
3	Н-пара	m_4 - m_6				x		x						
4	Н-пара	m_5 - m_6					x	x						
5	О-пара	m_7 - m_8							x	o				
6	О-пара	m_8 - m_7							o	x				
7	О-пара	m_2 - m_9	x								o			
8	Н-пара	m_11 - m_12											x	x

Рисунок 4.6 – Пример описания системы измеряемых свойств в *OntoWorker*: m, m_2, \dots, m_{12} – измеряемые свойства; О-, Н-пары сопряженных свойств с обусловленностью и несовместимостью соответственно

	1	2	3	7	8	9	10	11	12	13	14	15
1				p_4	p_5	p_6	p_7	p_8	p_9	p_10	p_11	p_12
2				100	100	100	100	100	100	100	100	100
3				m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12
4												
5												
6												

Рисунок 4.7 – Пример таблицы «объекты-свойства» (заготовки таблицы), автоматически создаваемой при априорном формировании системы измеряемых свойств в *OntoWorker*

	1	2	3	7	8	9	10	11	12	13	14	15
1				p_4	p_5	p_6	p_7	p_8	p_9	p_10	p_11	p_12
2				100	100	85	100	100	100	60	100	100
3				m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12
4	o	90	O	X	NM		Сидоров	Failure	NM	с.н.с.	5	NM
5	o_1	100	O	X	X	NM	Иванов	NM	NM	г.н.с.	10	X
6	p	100	P	X	X		Самойлов		Failure		9	X
7	p_2	100	P	X	Failure		Игнатьев	Failure	Failure	м.н.с.	6,5	X
8	p_3	80	P	NM	X	NM	Костенко	NM	NM	м.н.с.	7	NM
9	s	100	S	X	X	Failure	Куприянов	NM	NM		Failure	X

Рисунок 4.8 – Пример обобщенной таблицы «объекты-свойства» в OntoWorker

4.3 Разработка ключевых алгоритмов обработки данных

В подразделе представлены ряд системообразующих алгоритмов программной лаборатории OW. Одни из них детально описаны на интуитивно понятном псевдокоде с элементами символики теории множеств. Для других порядок работы раскрывается неформально с помощью последовательностей скриншотов, что одновременно имеет целью проиллюстрировать элементы разработанного WIMP-интерфейса с пользователем.

Здесь и далее до определенного 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 4.9):

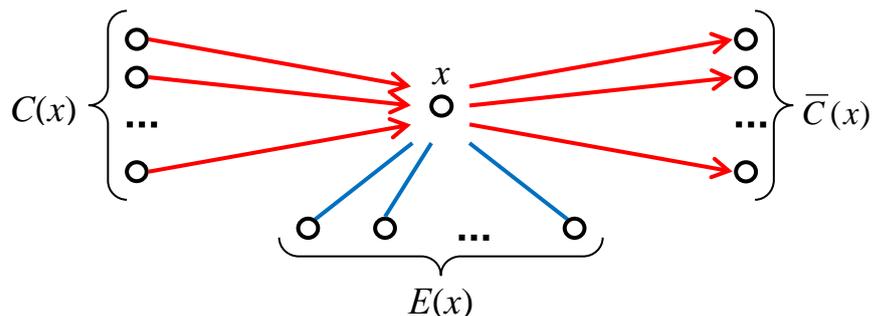


Рисунок 4.9 – Множества свойств, ассоциированные в системе измеряемых свойств с ограничениями существования с каждым отдельно взятым свойством

$\underline{C}(x)$ – множество свойств, обуславливающих свойство x , $0 \leq |\underline{C}(x)| \leq |M|-1$,

$\bar{C}(x)$ – множество свойств, обусловленных свойством x , $0 \leq |\bar{C}(x)| \leq |M| - 1$,

$E(x)$ – множество свойств, несовместимых со свойством x , $0 \leq |E(x)| \leq |M|-1$,

– причем лишь $\underline{C}(x)$ и $\bar{C}(x)$ могут пересекаться. Относительно несложное выявление $\underline{C}(x)$, $\bar{C}(x)$ и $E(x)$ будем относить к числу *базовых алгоритмических примитивов*, которые не детализируются.

4.3.1 Поддержка формирования и редактирования системы измеряемых свойств

Очертим реализацию поддержки когнитивных актов субъекта-исследователя в настольном Excel-приложении OW [65, 67].

Как отмечено в первой главе, *apriori*, при подготовке к наблюдениям или измерительным экспериментам субъект, осознанно или нет, формирует гипотезы о понятийной структуре интересующей его ПрО. Устанавливая набор свойств, подлежащих фиксации у объектов исследуемой ПрО, исследователь фактически предопределяет наибольшую вариативность потенциальных результатов ОАД. Осознанное гипотезирование, выражающееся в построении СИС с ограничениями существования свойств, сужает эту вариативность, включая в анализ априорные знания субъекта.

OW поддерживает априорное построение СИС любой сложности в форме сложно структурированного множества ГСС на основе ограниченного числа *шаблонов* [13, 34]. Скриншот на рис. 4.6 демонстрирует OW-представление примера СИС в легко создаваемой и управляемой в Excel табличной форме, в виде матрицы инцидентности «сопряженные пары свойств – свойства».

Априорное создание и редактирование СИС автоматически сопровождается построением и редактированием заготовки таблицы «объекты-свойства» («пустой» ОТОС – см. рис. 4.7). Последующая работа с этой таблицей связана с выполнением следующих действий:

- введение дополнительных столбцов для результатов измерения свойств конгруэнтными процедурами измерения (по-разному и/или с различной достоверностью измеряющими одно и то же свойство – см. на рис. 4.8 повторяющиеся столбцы для $m1$ и $m2$);

- добавлением строк, фиксирующих различные и отличающиеся достоверностью серии измерений для объектов зондируемой ПрО, измерения которых формируют обучающую выборку (см. на рис. 4.8 строки, идентифицируемые именами объектов обучающей выборки $O1, O2, \dots, O4$);
- наполнение ячеек таблицы результатами измерений (рис. 4.8).

Все этапы последующего изложения иллюстрирует граф на рисунке 4.10, который описывает матрицу инцидентности «сопряженные пары свойств – свойства», представленную на рисунке 4.11.

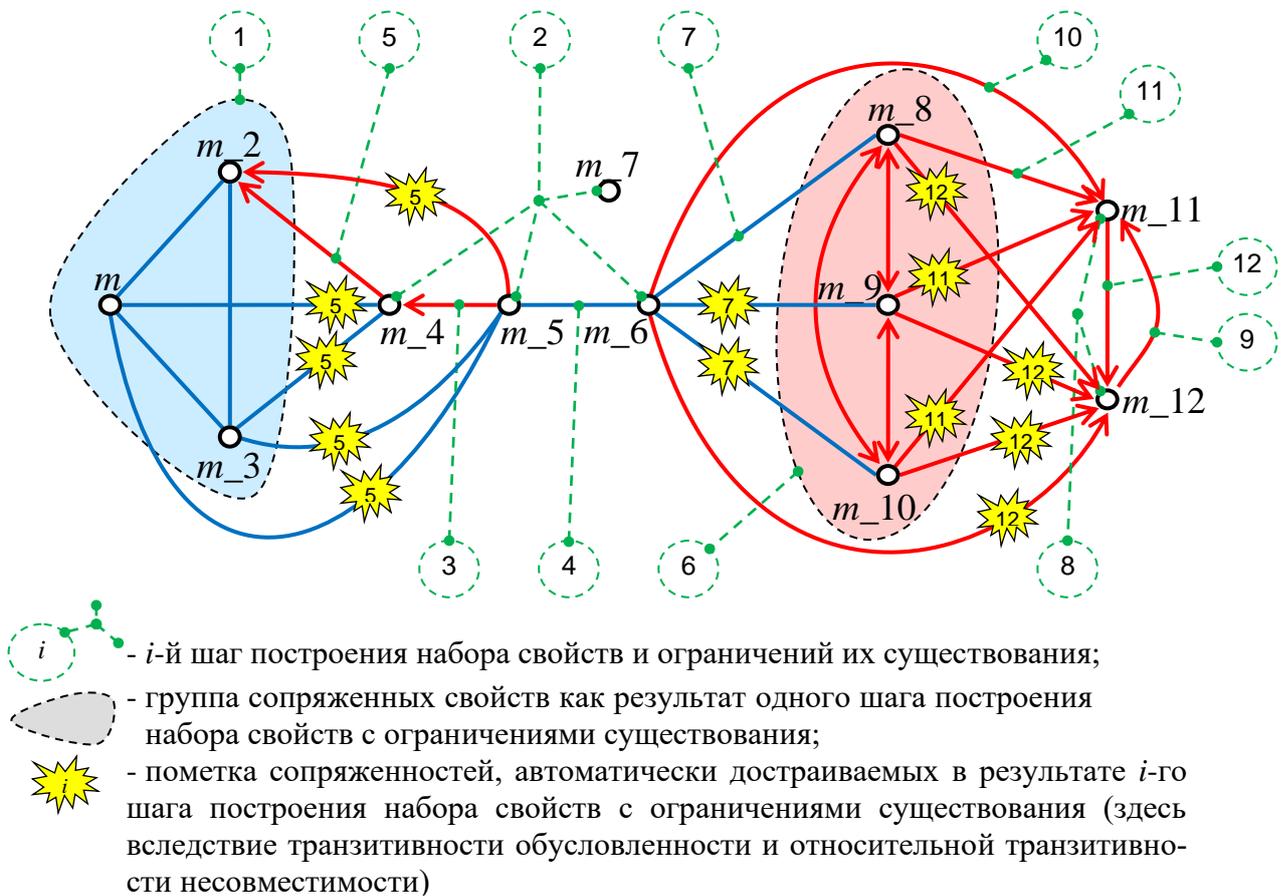


Рисунок 4.10 – Пример пошагового построения набора свойств с ограничениями существования (пояснение ряда графических образов см. на рисунке 2.1)

Задача добавления множества новых свойств в набор измеряемых свойств с ОСС, когда не выдвигается требование их какой-либо сопряженности 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 эк-

зистенциальные отношения (добавления сопряженностей) разумно разделять, то такое в целом полезное обобщение задачи целесообразно ограничить. Здесь во внимание следует принять *когнитивный смысл* расширения набора измеряемых свойств [12, 39], а также практические соображения по упрощению реализации алгоритмических примитивов, используемых при построении матрицы инцидентности, пример которой приведен на рисунке 4.11. С учетом этого рационально ограничиться тремя возможностями расширения набора измеряемых свойств, алгоритмы которых (на псевдокоде) приводятся ниже.

Алгоритм «Добавление k несопряженных 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 несопряженных 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 несопряженных 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

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

- указанная сопряженность уже существует;
- указанная сопряженность противоречит существующей сопряженности x и y , т.е. потребована либо несовместимость x и y , а эта пара так или иначе уже со-

стоит в отношении обусловленности, либо наоборот ($\forall x, y \in M$:
 $E(x, y) \wedge (C(x, y) \vee C(y, x)) \equiv \text{False}$);

		m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12
1-пара	m_1 - m_2	X	X										
1-пара	m_1 - m_3	X		X									
1-пара	m_2 - m_3		X	X									
0-пара	m_5 - m_4				O	X							
1-пара	m_5 - m_6					X	X						
0-пара	m_4 - m_2		O		X								
0-пара	m_5 - m_2		O		X	X							
1-пара	m_4 - m_3			X	X								
1-пара	m_4 - m_1	X			X								
1-пара	m_5 - m_3			X		X							
1-пара	m_5 - m_1	X				X							
0-пара	m_8 - m_9							X	O				
0-пара	m_9 - m_8							O	X				
0-пара	m_8 - m_10							X		O			
0-пара	m_10 - m_8							O		X			
0-пара	m_9 - m_10								X	O			
0-пара	m_10 - m_9								O	X			
1-пара	m_6 - m_8					X		X					
1-пара	m_6 - m_10					X				X			
1-пара	m_6 - m_9					X			X				
0-пара	m_12 - m_11										O	X	
0-пара	m_6 - m_11					X					O		
0-пара	m_8 - m_11						X				O		
0-пара	m_10 - m_11							X		O			
0-пара	m_9 - m_11							X		O			
0-пара	m_11 - m_12								X		O		
0-пара	m_9 - m_12							X			O		
0-пара	m_10 - m_12								X		O		
0-пара	m_8 - m_12							X			O		
0-пара	m_6 - m_12					X					O		

Рисунок 4.11 – Пример пошагового построения матрицы инцидентности «сопряженности 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

- добавление указанной сопряженности приведет к противоречию с существующими ОСС опосредованно вследствие законов транзитивности экзистенциальных отношений. Нетрудно показать, что такое нарушение произойдет при попытке добавления в ОСС О-пары (x, y) в случае, когда $\exists z \in \underline{C}(x) \cap E(y)$, или 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 сопряженности пары свойств, а собственно добавление сопряженности пары свойств, инициированное субъектом ОАД, в общем случае вызывает необходимость *добавления* в ОСС других новых сопряженностей 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** $C(x, y) = \text{True}$ **and not** $E(x, y) = \text{True}$ **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 идет об образовании самообусловленности; тем самым законы 'транзитивности экзистенциальных отношений не приведут к противоречиям в ОСС:
- 5) **for each** $z1$ **in** $\underline{C}(x)$
- 6) **for each** $z2$ **in** $\bar{C}(y)$
- 7) **if** $C(z1, z2) = \text{False}$ **and** $z1 \neq z2$ **then** $C(z1, z2) := \text{True}$
- 8) **next** $z2$
- 9) **next** $z1$
- 10) $M^* := \emptyset$ 'начало формирования множества свойств, 'несовместимых с каждым свойством из $\bar{C}(y)$
- 11) **for each** $z1$ **in** $\bar{C}(y)$
- 12) **for each** $z2$ **in** $E(z1)$
- 13) $M^* := M^* \cup \{z2\}$ **next** $z2$ 'бесповторное добавление свойства $z2$
- 14) **next** $z2$
- 15) **next** $z1$
- 16) '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:
- 17) **for each** $z1$ **in** $\underline{C}(x)$
- 18) **for each** $z2$ **in** M^*
- 19) **if** $E(z1, z2) = \text{False}$ **then** $E(z1, z2) := \text{True}$
- 20) **next** $z2$
- 21) **next** $z1$
- 21) **end if**

На рисунках 4.10, 4.11 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** $E(x, y) = \mathbf{True}$ **and not** $(C(x, y) = \mathbf{True} \text{ or } C(x, y) = \mathbf{True})$ **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 не приведет к противоречиям в ОСС:
 for each $z1$ **in** $\underline{C}(x)$
- 5) **for each** $z2$ **in** $\underline{C}(y)$
- 6) **if** $E(z1, z2) = \mathbf{False}$ **then** $E(z1, z2) := \mathbf{True}$
- 7) **next** $z2$
- 8) **next** $z1$
- 9) **end if**

На рисунках 4.10, 4.11 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) .

Далее анализируются такие элементарные модификации набора измеряемых свойств с ОСС как *удаление свойства* из набора измеряемых свойств и *удаление сопряженности пары свойств* из ОСС [32].

Изъятие из ОСС сопряженности свойств $x, y \in M$ чревато нарушением транзитивности экзистенциальных отношений измеряемых свойств:

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

Таким образом, основным содержанием алгоритмов удаления сопряженностей пар свойств 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) $C(x, y) := \mathbf{False}$ '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 «закрывающей»
- 2) $E(x, y) := \text{False}$ 'удаление несовместимости свойств x и y
- 3) **end if**

Например, в наборе свойств с ОСС на рисунке 4.10 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.

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

$$\begin{aligned} \exists z \in M, z \neq x, z \neq y: C(x, z) \wedge C(z, y) &= \text{True}, \\ C(x, z) \wedge E(z, y) \vee C(y, z) \wedge E(z, x) &= \text{True}, \end{aligned}$$

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 обусловливаемыми свойствами
- 2) $C(x, y) := \text{False}$
- 3) **next** y
- 4) **for each** y **in** $C(x)$ 'удаление сопряжения с обусловливающими свойствами
- 5) $C(y, x) := \text{False}$
- 6) **next** y
- 7) **for each** y **in** $E(x)$ 'удаление сопряжения с несовместимыми свойствами
- 8) $E(y, x) := \text{False}$
- 9) **Next** y
- 10) $M := M \setminus \{x\}$

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 располагать возможностью *замещения* одного свойства группой ВЗО-свойств.

Очевидно, что в силу транзитивности экзистенциальных отношений между свойствами *каждое* из ВЗО-свойств, замещающих выбранное измеряемое свойство, «наследует» все его сопряженности в наборе измеряемых свойств (это также следует из того, что отношение взаимообусловленности H в силу характеризующих его свойств симметрии – см. введение – разбивает множество измеряемых свойств на классы эквивалентности). Реализация наследования сопряженностей 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 взаимообусловленных свойств 'см. описание выше
- 2) $y := M[n + k]$ 'последнее из замещающих ВЗО-свойств, которое далее наследует
'все сопряженности замещаемого свойства x ; при этом остальные
'замещающие ВЗО-свойства получают все эти сопряженности автоматически
- 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 4.10, 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} .

4.3.2 Алгоритмы реструктуризации системы измеряемых свойств

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 нее свойств, но и отсутствие пересечения с другими ВЗО-группами (исключая, разумеется, некоторые ВЗО-группы, состоящие из одного свойства). Поэтому после выявления очередной ВЗО-группы ее свойства-члены

следует исключить из дальнейшего анализа, что сокращает трудоемкость выявления каждой следующей такой группы [32].

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** $\underline{C}(x)$
- 7) **for each** z **in** $\underline{C}(y)$
- 8) **if** $z = x$ **then** $G := G \cup \{y\}$
- 9) **next** z
- 10) **next** y
- 11) $M^* := M^* \setminus G$
- 12) **loop** 'множество свойств конечно

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) $M^* := \underline{C}(G2[1]) \setminus G2$
'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** $|G1 \cap M^*| > 0$ **then** 'ВЗО-группа G1 обуславливает ВЗО-группу G2
- 5) $G := \text{new } O$ 'O - класс «О-группа»
- 6) $G.C := G2$ 'обуславливаемые свойства-члены О-группы
- 7) $G.\underline{C} := G1$ 'обуславливающие свойства-члены О-группы
- 8) **end if**
- 9) **next** $G1$
- 10) **next** $G2$

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 путем рекурсии. Например, на рисунке 4.11 Н-группы $\{m, m_2, m_3\}$ и $\{m, m_3, m_5\}$ пересекают-

ся, а в H -группу $\{m, m_2, m_3\}$ вложены H -группы $\{m, m_2\}$, $\{m, m_3\}$, $\{m_2, m_3\}$.

Алгоритм «Выявление H -групп»

- 1) **for each** x **in** M
- 2) $GrS := H\text{-группы, включающие } x \text{ и подмножества } Ex$ $Ex = E(x)$
'множество H -групп, каждая из которых включает свойство x и
'определенное подмножество попарно несовместимых с ним свойств из множества Ex
- 3) $Ns := Ns \cup GrS$ 'бесповторное (по составу свойств-членов) накопление H -групп
- 4) **next** x

Алгоритм рекурсивной функции « H -группы, включающие x и подмножества Ex » выявляет как все пересекающиеся, так и вложенные одна в другую H -группы, включающие в себя свойство x .

Алгоритм « H -группы, включающие x и подмножества Ex »

- 1) $GrS := \emptyset$ 'искомое множество H -групп
- 2) **for each** y **in** $E(x)$
- 3) $M^* := E(y) \cap E(x)$ 'свойства, несовместимые как со свойством y , так и со свойством x
- 4) **if** $|M^*| \leq 1$ **then** ' M^* пусто или содержит одно свойство
- 5) $Gr := \text{new } N$ ' N - класс « H -группа»
' Gr используется здесь и как множество свойств-членов группы Gr
- 6) $Gr := Gr \cup \{y\}$
- 7) **if** $|M^*| = 1$ **then** $Gr := Gr \cup M^*$
- 8) **else**
- 9) $Gr := H\text{-группы, включающие } x \text{ и подмножества } Ex$
' $x = y, Ex = M^*$; рекурсия для свойства y и M^*
- 10) **end if**
- 11) $GrS := GrS \cup Gr$ 'бесповторное накопление искомым H -групп
- 12) **next** y
- 13) **for each** Gr **in** GrS
- 14) $Gr := Gr \cup \{x\}$
- 15) **next** Gr
- 16) $H\text{-группы, включающие } x \text{ и подмножества } Ex := GrS$

4.3.3 Алгоритмизация работы с таблицей «объекты-свойства» и формирование исходного контекста

Здесь рассматриваются более «нагруженные» в когнитивном смысле модификации СИС и ОСС, возникающие в результате концептуального шкалирования измеряемых свойств.

Поддержка когнитивных актов субъекта, выражающихся в редактировании СИС при непустой ОТОС, существенно усложняется, поскольку требует надлежащей автоматической трансформации в общем случае неоднородного содержи-

мого столбцов ОТОС. В целом методы необходимых преобразований ОТОС могут быть основаны на переходе к интерпретации результатов измерений в векторной логике (т.е. переходу к нестрогому соответствию «объекты-свойства» – см. [15]) и применению соответствующего аппарата вычисления векторов истинности связок «И» и «ИЛИ» [44].

Концептуальное шкалирование измеряемых свойств в ОТОС – традиционная точка приложения когнитивных усилий субъекта в системах анализа данных, основанных на АФП, – связано, прежде всего, с осуществлением отбора в столбце ОТОС элементов, удовлетворяющих условиям, задаваемых субъектом. Этот отбор в общем случае может быть эффективно реализован на основе использования Excel-функции «Расширенный фильтр» (*Advanced Filter*) [71].

Формирование дизъюнктивной шкалы осуществляется в следующем порядке:

- *Бесповторное* извлечение из ОТОС и лексикографическое упорядочение множества всех зафиксированных значений шкалируемого свойства. Разумеется, полученное множество значений будет лишь *представлять* домен шкалируемого свойства, т.к. в общем случае это представление - «*эмпирический*» домен - не исчерпывает всё реальное содержание домена шкалируемого свойства. Тем не менее, это вся доступная информация об интересующем нас домене.
- Отнесение построенного эмпирического домена к поддомену n -го нового свойства.
- Интерактивное в произвольном порядке и возможностью возвратов формирование поддомена, отходящего к i -му новому свойству, $i = 1, 2, \dots, n$, путем обмена значениями с n -м поддоменом шкалируемого свойства (см. рисунок 4.12). Одновременно с формированием каждого поддомена определяется имя соответствующего нового свойства.

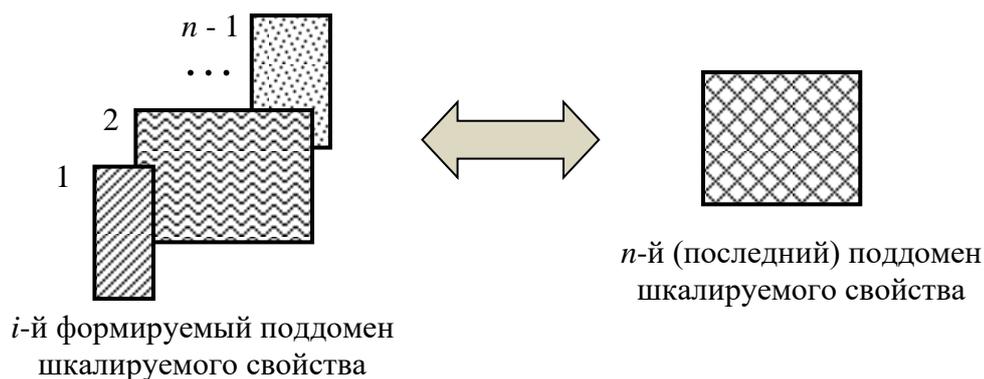


Рисунок 4.12 – Формирование доменов новых свойств как поддоменов дизъюнктивно шкалируемого свойства

Заметим, что последний пункт составляет одно из важнейших звеньев поддержки когнитивных действий субъекта при ОАД.

Сформированные поддомены шкалируемого свойства служат основой для трансформации ОТОС, которую можно описать следующим образом:

- 1) В таблице отбираются объекты измерения со значениями шкалируемого свойства, принадлежащими i -му поддомену.
- 2) В таблицу добавляется столбец, содержащий результаты измерения i -го нового свойства, а именно знаком **X** отмечаются отобранные объекты и знаком **None** все остальные.
- 3) После выполнения пунктов 1 и 2 для всех $i = 1, 2, \dots, n$ из таблицы исключается столбец свойства, подвергнутого дизъюнктивному шкалированию.

Анализ возможности дизъюнктивного шкалирования свойства $x \in M$ иллюстрирует рисунок 4.13, где, в частности, демонстрируется *неопределенность* 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.

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

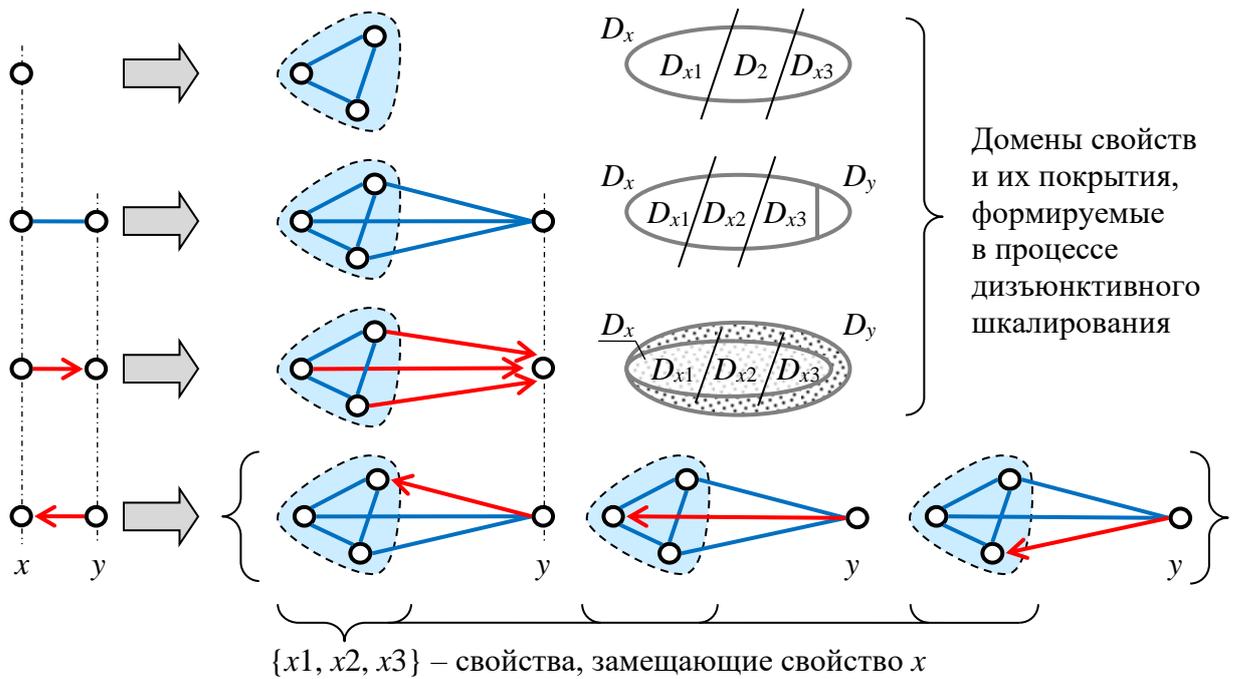


Рисунок 4.13 – Пример дизъюнктивного шкалирования: замещение тремя несовместимыми свойствами $\{x_1, x_2, x_3\}$ свойства x , сопряженного со свойством y ; приведены случаи: x без сопряжения, $y \in E(x)$, $y \in \bar{C}(x)$, $y \in \underline{C}(x)$ (пояснение ряда графических образов см. на рисунке 1)

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

Анализ, аналогичный тому, что очерчен для дизъюнктивного шкалирования, позволяет установить, что невозможно реализовать обобщенную 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 4.10 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 ее 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) $b := \mathbf{True}$ '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** $E(x, y) = \mathbf{True}$ **then** $b := \mathbf{False}$
- 6) **next** y
- 7) **next** x
- 8) **end if**
- 9) **if** $b = \mathbf{True}$ **then** 'уточняющее шкалирование множества M^* возможно
- 10) $n := |M|$
- 11) Добавление k -несопряженных свойств, $k = 1$
- 12) **for each** x **in** M^*
- 13) Добавление O -пары $(M[n + 1], x)$ **next** x
- 14) **next** x
- 15) **end if**

Для реализации управления данными при дизъюнктивном шкалировании в программной лаборатории OW реализован комплекс действий по команде «Шкалировать измеряемое свойство» контекстного меню, «привязанного» к зоне ОТОС, где представлены процедуры измерения. При этом для интерактивного ввода управляющих параметров и манипуляции данными используются формы «Шкалирование1» и «Шкалирование2». Общий вид этих форм представлен на рисунках 4.14 и 4.15.

На рисунке 4.16 приведен скриншот ситуации выбора команды «Шкалировать измеряемое свойство» применительно к измеряемому свойству «Оценка возможности добавления новой сопряженной пары свойств в СИС» в реальной задаче онтологического анализа данных.

Выполнение команды «Шкалировать измеряемое свойство» начинается

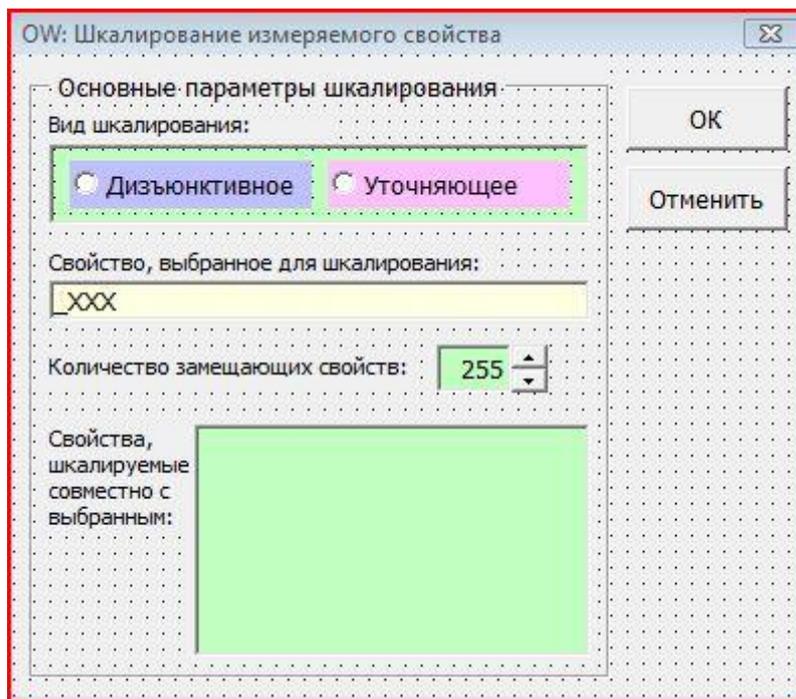


Рисунок 4.14 – Общий вид Excel-формы «Шкалирование1» для выбора вида и основных параметров когнитивного шкалирования в программной лаборатории *OntoWorker*

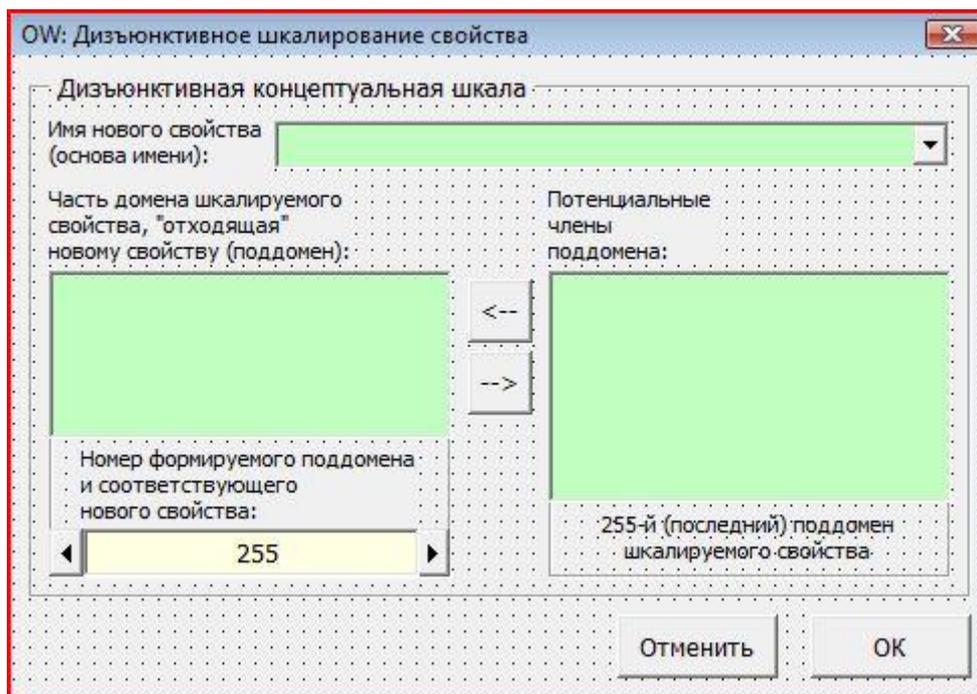


Рисунок 4.15 – Общий вид Excel-формы «Шкалирование2» для формирования когнитивных шкал в программной лаборатории *OntoWorker*

с интерактивного выбора вида когнитивного шкалирования и его основных параметров с помощью формы «Шкалирование1». Скриншот на рисунке 4.17 демонстрирует вид этой формы, предъявляемый пользователю непосредственно после выбора команды «Шкалировать измеряемое свойство».

OntoWorker-43 - Microsoft Excel

Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид

Буфер обмена Вставить Шрифт: Calibri, Размер: 11, Цвет: Число: %0,000,00

	1	2	3	4	5	6	7	8	9	10	11	12	13
1					р_2	р_3	р_4	р_5	р_6	р_7	р_8		
2				100	100	100	100	100	100	100	100		
3				Новая В-пара	Новая О-пара	Новая Н-пара	Существующая В-пара	Существующая О-пара	Существующая Н-пара	Пересечение составов новой и существующей пар	Оценка возможности добавления новой сопряженной пары свойств в СИС		
4	1	100	Ситуация 1	X	None	None	X	None	None	None	None	None	None
5	1_2	100	Ситуация 2	None	X	None	X	None	None	None	None	None	None
6	1_3	100	Ситуация 3	None	None	X	X	None	None	None	None	None	None
7	1_4	100	Ситуация 4	X	None	None	None	X	None	None	None	None	None
8	1_5	100	Ситуация 5	None	X	None	None	X	None	None	None	None	None
9	1_6	100	Ситуация 6	None	None	X	None	X	None	None	None	None	None
10	1_7	100	Ситуация 7	X	None	None	None	None	X	None	None	None	None
11	1_8	100	Ситуация 8	None	X	None	None	None	X	None	None	None	None
12	1_9	100	Ситуация 9	None	None	X	None	None	X	None	None	None	None
13	1_10	100	Ситуация 10	X	None	None	None	None	None	None	None	None	None
14	1_11	100	Ситуация 11	None	X	None	None	None	None	None	None	None	None
15	1_12	100	Ситуация 12	None	None	X	None	None	None	None	None	None	None
16	1_13	100	Ситуация 13	X	None	None	None	X	None	None	None	None	None
17	1_14	100	Ситуация 14	None	X	None	None	None	None	None	None	None	None
18	1_15	100	Ситуация 15	None	None	X	None	None	None	None	None	None	None
19	1_16	100	Ситуация 16	X	None	None	None	None	X	None	None	None	None
20	1_17	100	Ситуация 17	None	X	None	None	None	X	None	None	None	None
21	1_18	100	Ситуация 18	None	None	X	None	None	X	None	None	None	None
22													

Контекстное меню «Шкалировать измеряемое свойство...»

- Редактировать процедуру измерения...
- Удалить процедуру
- Шкалировать измеряемое свойство...
- Ширина столбца...
- Переместить столбец...
- Невозможно
- Избыточно
- Невозможно
- Возможно

Рисунок 4.16 – Выбор команды «Шкалировать измеряемое свойство...» в контекстном меню листа OntoWorker’ а

Избрание с помощью кнопочного переключателя вида шкалирования приводит к разворачиванию формы «Шкалирование1» с целью предоставления элементов управления, позволяющих в случае выбора дизъюнктивного шкалирования указать количество свойств, которые в результате шкалирования должны заместить шкалируемое измеряемое свойство (рисунок 4.18). Количество замещающих свойств ограничено диапазоном от 2-х до 255.

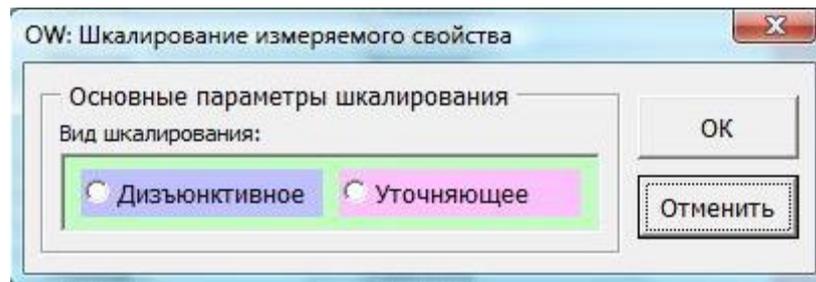


Рисунок 4.17 – Начальный вид формы «Шкалирование1» для выбора вида и основных параметров шкалирования измеряемого свойства

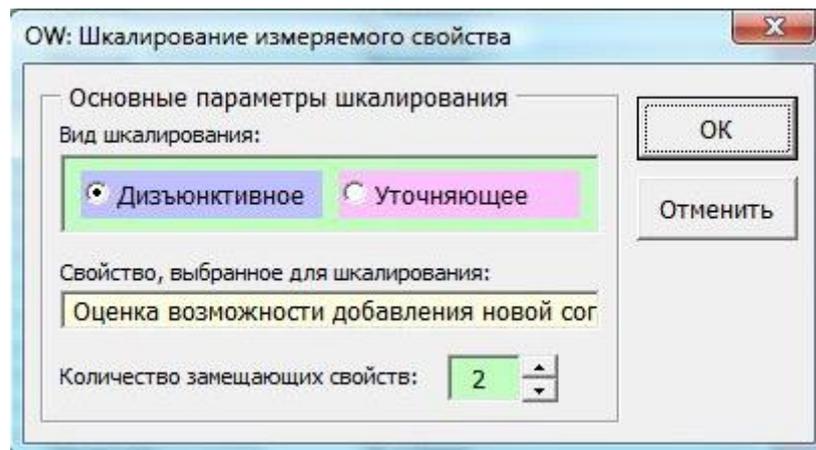


Рисунок 4.18 – Расширенный вид формы «Шкалирование1» для выбора количества свойств, замещающих дизъюнктивно шкалируемое измеряемое свойство

После выбора дизъюнктивного шкалирования и количества замещающих свойств, производится интерактивное конструирование соответствующей концептуальной шкалы с помощью формы «Шкалирование2», использование управляющих элементов которой позволяет реализовать алгоритм формирования поддоменов новых свойств, описанный выше.

4.3.4 Опорный алгоритм нормализации рабочего контекста

Опорный алгоритм нормализации [48] устанавливает для каждого отдельно взятого объекта $g \in G$ статус любой субструктурной группы свойств в СИС.

Группа получает статус «0» либо при отсутствии пересечения с множеством свойств $M^* \in M$ объекта g , либо при выполнении применительно к M^* условия представления группы сопряженных свойств в нормальном подмножестве измеряемых свойств (сообразно виду группы – см. предыдущий раздел). Иначе группе присваивается статус «1» («проблемная группа»). Статус «1» переходит в «2» («критичная группа»), если группе невозможно вернуть статус «0» за счет надлежащего исключения свойств-членов этой группы, БСС которых применительно к объекту g наименее достоверны.

Алгоритм «Определение статуса st субструктурной группы свойств Gr для объекта g »

Input: (G^*, M, I) – формальный контекст

$g \in G^*$ – отдельно взятый объект измерения;

Gr – субструктурная группа свойств в системе (M, C, E) ;

MCs, Cs, Es – множество ВЗО-, О- и Н-групп свойств во множестве M

Output: st – статус группы Gr для объекта g ,

$st \in \{0, 1, 2\}$

- 1) $st := 0$ 'исходное предположение о статусе Gr
- 2) $M^* := \{g\}$ 'множество свойств объекта g («'» - оператор Галуа)
- 3) **if** $Gr \in MCs$ **And** $|Gr| > 1$ **then** ' Gr – ВЗО-группа с несколькими членами
- 4) $IS := |Gr \cap M^*|$ 'множество свойств объекта g , входящих в ВЗО-группу Gr
- 5) **if** $0 < |IS| < |Gr|$ **then** ' Gr – проблемная ВЗО-группа,
'т.к. она пересекается с M^* , но не вложена в него
- 6) $st :=$ **Ненулевой_статус** (g, IS) ' Gr – проблемная или критичная группа
- 7) **end if**
- 8) **else if** $Gr \in Cs$ **then** ' Gr – О-группа
- 9) $IS := |Gr.C_o \cap M^*|$ 'множество свойств объекта g , входящих в обусловливаемую
'ВЗО-группу О-группы Gr ($Gr.C_o$ суть эта ВЗО-группа)
- 10) **if** $|IS| = 0$ **then**
- 11) $IS := |Gr.C_x \cap M^*|$ 'множество свойств объекта g , входящих в обусловливающую
'ВЗО-группу О-группы Gr ($Gr.C_x$ суть эта ВЗО-группа)
- 12) **if** $|IS| = |Gr.C_x|$ **then** ' Gr проблемная О-группа, т.к. её обусловливающая ВЗО-группа
'вложена в неё, а её обусловливаемая ВЗО-группа с ней не пересекается
- 13) $st :=$ **Ненулевой_статус** (g, IS) ' Gr – проблемная или критичная группа
- 14) **end if**
- 15) **end if**
- 16) **elseif** $Gr \in Es$ **then** ' Gr – Н-группа
- 17) $IS := |Gr \cap M^*|$ 'множество свойств объекта g , входящих в Н-группу Gr
- 18) $k := |IS|$
- 19) **if** $k > 1$ **then** 'мощность $Gr \cap M^*$ больше 1
- 20) $st := 1$ ' Gr – проблемная Н-группа
- 21) **for each** x **in** IS 'анализ вариантов исключения свойств из IS
- 22) **if** **Ненулевой_статус** ($g, IS \setminus \{x\}$) = 2 **then**
- 23) $k := k - 1$
- 24) **end if**
- 25) **next** x

- 26) **end if**
 27) **if** $k = 0$ **then** $st := 2$ 'Gr – критичная H-группа
 28) **end if**

Алгоритм функции «Ненулевой_статус» (g, IS) выявляет возможность исключения множества свойств IS из M^* – множества свойств объекта g , за счет локального – для объекта g и его проблемной группы свойств – ужесточения порога доверия к эмпирическим данным. Очевидно, что для логически однозначного формального контекста функция всегда будет доставлять значение «2» (рассматриваемая проблемная группа свойств «критична» для объекта g). В случае логической многозначности эмпирического соответствия «объекты-свойства» такой результат будет получен лишь тогда, когда исключение IS из M^* потребует исключения из формального контекста абсолютно достоверных оценок БСС.

С учетом утверждения, констатирующего нормальность любого подмножества измеряемых свойств (см. п. 3.1), выявление для $g \in G^*$ статуса каждой субструктурной группы свойств в СИС позволяет дать следующее заключение: множество свойств объекта g нормально, тогда и только тогда, если по отношению к нему все субструктурные группы свойств имеют статус «0».

Что касается принудительной нормализации логически многозначного эмпирического контекста, то соответствующий алгоритм в цикле по всем объектам $g \in G^*$ и цикле по всем субструктурным группам в целом воспроизводит рассмотренный опорный алгоритм нормализации и дополняется:

- в случае пороговых методов вывода четких и нечетких понятий [72-74] апробирован метод рационального выбора варианта исключения подмножества свойств объекта g до тех пор, пока множество свойств не станет нормальным, либо до обнаружения неразрешимого противоречия между априорными гипотезами и эмпирическими данными о ПрО [15, 72];
- при использовании оператора нечеткого замыкания для вывода нечетких понятий [55, 73, 74] может быть применен метод фильтрации декомпозиции нечеткого соответствия «объекты-свойства» с исключением четких слагаемых композиции, не являющимися нормальными множествами [75].

4.4 Демонстрация работы программного комплекса

Для демонстрации подготовки контекста для вывода формальных понятий, реализованной в программной лаборатории OW, использован материал из в каталога инновационных разработок Самарского НИИ сельского хозяйства имени Н.М. Тулайкова [76].

В каталоге представлен список селекционных сортов, включенных в Государственный реестр, и сведения об инновационных технологиях возделывания сельскохозяйственных культур. В качестве объектов обследуемой ПрО выступили следующие сорта мягкой яровой пшеницы:

- Жигулевская;
- Самсар;
- Тулайковская 5;
- Тулайковская 10;
- Тулайковская 100;
- Тулайковская золотистая;
- Экада 70;
- Экада 66;
- Тулайковская 110;
- Тулайковская победа;
- Тулайковская надежда;
- Тулайковская 116;
- Экада 214;
- Зауральская волна.

Признаки этих объектов отыскивались в описании каждого сорта, содержащем, как качественную, так и количественную информацию. Некоторые словесные характеристики сортов, представленные в [76], трудно формализуемы без помощи эксперта в данной ПрО. Например, сложно структурировать сведения об урожайности: в каталоге описывается сравнение урожаев в разные годы, в разных

регионах, иногда отмечаются численные показатели урожайности. Однако некоторая часть характеристик сортов может быть формализована и неспециалистом.

С учетом подобных замечаний и стремления упростить демонстрационный пример были избраны следующие свойства сортов:

- показатели засухоустойчивости;
- процент содержания белка в зерне;
- характеристика устойчивости к одному из биологических вредителей – «бурой ржавчине».

Зачастую людей интересует не сам факт наличия свойства у объекта, а именно его значение или диапазон значений. Именно поэтому в тексте каталога [76] присутствуют такие формулировки, как «засухоустойчивость», «высокая засухоустойчивость» и «формирует высокие урожаи зерна в экстремальных условиях высоких температур» (что было интерпретировано как «*экстремальная засухоустойчивость*»). Ясно, что если сорт экстремально засухоустойчив, то он заведомо и *высоко засухоустойчив*, и *засухоустойчив* (аналогично с высокой засухоустойчивостью и просто засухоустойчивостью). И в случае, если сорт имеет иммунитет к бурой ржавчине, будем считать, что он также *устойчив* и *достаточно устойчив* к бурой ржавчине.

Несколько другую ситуацию обнаруживаем с характеристикой содержания белка в зерне: здесь наблюдаются изменения в показателях от 12 до 18 % в пределах рассматриваемых сортов. Поскольку для нас не очевидно, на какие диапазоны следует разделять данное свойство, мы обратились к ГОСТ 9353-2016 для пшеницы [77]. В нем перечислены требования к классам пшеницы, в зависимости от качества зерна, к которому и относится, в том числе, процент белка в зерне.

По сути дела, в ГОСТе осуществлено *номинальное концептуальное шкалирование* потому, что диапазон процента содержания белка важен для определения класса пшеницы в зависимости от этого показателя зерна. Заметим, что для свойств засухоустойчивости и устойчивости к болезням авторами каталога, по сути, было осуществлено *порядковое шкалирование* [23].

В соответствии с требованиями ГОСТ 9353-2016 были сформированы свойства «12-13,5% белка», «13,5-14,5% белка» и «>14,5% белка», которые соответствуют III, II и I классам зерна. В этом смысле, сведения каталога фактически указывают, что зачастую сорт относится к двум классам, а детальной информации о разделении зерна пшеницы между этими классами в каталоге нет, т.е. возникает неопределенность отнесения рассматриваемых сортов пшеницы к гостовским классам. При этом будем придерживаться той точки зрения, что в действительности сорт пшеницы должен быть отнесен к одному и только одному классу. Т.е. свойства, описывающие характеризующий класс, считаем *несовместимыми*, представляя номинальную шкалу для измерения содержания в зерне белка. Информацию о принадлежности сортов мягкой пшеницы к различным классам, представленную в каталоге [76] детализируем в форме различных степеней принадлежности сортов пшеницы к классам ГОСТ. Поскольку в каталоге приведен (и то не для всех рассматриваемых сортов) лишь диапазон величины содержания белка в зерне, то разумно рассчитать искомую степень принадлежности при условии равномерного распределения зерновой массы по указанным диапазонам.

Таким образом, рассматриваются данные о 15 объектах в разрезе 11 свойств.

В результате формируется СИС, скриншот которой приводится на рисунке 4.19, а эмпирические данные о соответствии объектов ПрО множеству измеряемых свойств представляются в OW в виде ОТОС, скриншот которой дан на рисунке 4.20.

Отметим некоторую условность характеристики сорта Самсар относительно принадлежности гостовским классам пшеницы. В ГОСТе требуется, чтобы пшеница 1-го класса содержала процент белка *не менее* 14%, т.е. в минимальной степени предпочтительно отнести Самсар к 1-му классу. Поэтому условно Самсар отнесен к 1-му классу с достоверностью $\langle 0,51; 0 \rangle$, а ко 2-му – с $\langle 0,49; 0 \rangle$.

Далее OW преобразует ОТОС в ИФК, где результаты измерений заменяются совокупностью оценок истинности БСС (скриншот на рисунке 4.21), а СИС представляется в виде совокупности пересекающихся субструктур – групп

свойств, однородных по характеру экзистенциального сопряжения свойств-членов (скриншот на рисунке 4.22).

На следующем шаге формируется РФК, где исключатся заведомо ложные БСС, а все оставшиеся признаются истинными (в скриншоте на рисунке 5.3 инцидентности, соответствующие заведомо ложным БСС, затенены, а в РФК они исключены – см. скриншот на рисунке 4.23).

Можно видеть, что в исходном РФК ОСС нарушены, т.к. некоторые сорта пшеницы оказались отнесены одновременно к двум классам мягкой пшеницы по ГОСТу. Для каждого исследуемого сорта пшеницы OW идентифицирует так называемые «проблемные» группы сопряженных свойств – группы, состав которых у рассматриваемого сорта свидетельствуют о нарушении ОСС (рисунок 4.24). Собственно, на этом моменте и проводится *нормализация* РФК, когда удовлетворение ОСС достигается надлежащим исключением из РФК наименее достоверной эмпирической информации [49].

На рисунке 4.25 приведен скриншот нормализованного РФК, из которого возможно выводить формальные понятия с помощью методов классического АФП.

5.5 Выводы по разделу 4

- 1) Выполнено архитектурное проектирование программной лаборатории для онтологического анализа данных, которая в частности поддерживает автоматизацию всех составляющих методики подготовки контекста для вывода формальных понятий.
- 2) Разработаны базовые алгоритмы формирования и редактирования СИС и ОТОС а также алгоритмическое обеспечение редукции СИС на основе выделения классов эквивалентности измеряемых свойств, индуцируемой отношением взаимобусловленности свойств.
- 3) Разработан опорный алгоритм нормализации рабочего контекста.
- 4) Работа программной лаборатории продемонстрирована на реальных данных, описывающих сорта сельскохозяйственных культур.

Заданная система измеряемых свойств - СИС:										
злустой чив	высоко злустой чив	экстремально злустой чив	достаточно устойчив к БР	устойчив к БР	иммунитет к БР	5-й класс (белок < 10)	4-й класс (10 <= белок < 12)	3-й класс (12 <= белок < 13,5)	2-й класс (13,5 <= белок < 14,5)	1-й класс (белок >= 14,5)
О	Х									
О-пара высоко злустойчив - злустойчив	Х									
О-пара экстремально злустойчив - высоко злустойчив	О	Х								
О-пара экстремально злустойчив - злустойчив	Х									
О-пара устойчив к БР - достаточно устойчив к БР			О	Х						
О-пара иммунитет к БР - устойчив к БР			О	О	Х					
О-пара иммунитет к БР - достаточно устойчив к БР			О	Х	Х					
Н-пара m - m_2						Х	Х			
Н-пара m - m_3						Х	Х			
Н-пара m - m_4						Х	Х		Х	
Н-пара m - m_5						Х	Х		Х	
Н-пара m_2 - m_3							Х	Х		
Н-пара m_2 - m_4							Х	Х	Х	
Н-пара m_2 - m_5							Х	Х	Х	
Н-пара m_3 - m_4								Х	Х	Х
Н-пара m_3 - m_5								Х	Х	Х
Н-пара m_4 - m_5									Х	Х

Рисунок 4.19 – Система измеряемых свойств, принятая при онтологическом анализе сортов мягкой яровой пшеницы

Заданная обобщенная таблица "объекты-свойства" - ОТОС:												
	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_10	p_11	p_12	
	100	100	100	100	100	100	100	100	100	100	100	
	зустойчив	высоко	экстрема	достаточно	устойчив	иммунитет	5-й класс	4-й класс	3-й класс	2-й класс	1-й класс	
	зустойчив	зустойчив	льно	зустойчив	к БР	к БР	(белок < 10)	(белок < 12)	(12 <= белок < 13,5)	(13,5 <= белок < 14,5)	(белок < 14,5)	
s_100	X	X	None	X	None	None	None	None	None	<20; 0>	<80; 0>	
s_100	X	None	None	X	None	None	None	None	None	<49; 0>	<51; 0>	
s_100	X	None	None	X	X	X	None	None	None	None	<100; 0>	
s_100	X	X	None	X	X	X	None	None	None	None	<100; 0>	
s_100	X	X	None	X	X	X	None	None	None	<20; 0>	<80; 0>	
s_100	X	X	X	X	X	X	None	None	None	None	<100; 0>	
s_100	X	None	None	X	None	None	None	None	None	None	<100; 0>	
s_100	X	None	None	X	None	None	None	None	None	None	<100; 0>	
s_100	X	X	None	X	X	None	None	None	<75; 0>	<25; 0>	None	
s_100	X	X	None	X	X	None	None	None	None	<25; 0>	<75; 0>	
s_100	X	X	X	X	X	None	None	None	None	<17; 0>	<83; 0>	
s_100	X	None	None	X	X	None	None	None	<55; 0>	<45; 0>	None	
s_100	X	None	None	X	None	None	None	None	None	None	<100; 0>	

Рисунок 4.20 – Обобщённая таблица «объекты-свойства», представляющая известные данные о сортах мягкой яровой пшеницы

Исходный формальный контекст задачи - ИФК:											
	з/устойч ИВ	высоко з/устойч ИВ	экстрем ально з/устойч ИВ	достаточ но устойчив к БР	устойчив к БР	иммунит ет к БР	5-й класс (белок < 10)	4-й класс (10 <= белок < 12)	3-й класс (12 <= белок < 13,5)	2-й класс (13,5 <= белок < 14,5)	1-й класс (белок >= 14,5)
Жигулевская	<100; 0>	<100; 0>	<0; 100>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<20; 0>	<80; 0>
Самсар	<100; 0>	<0; 100>	<0; 100>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<49; 0>	<51; 0>
Тулайковская 5	<100; 0>	<0; 100>	<0; 100>	<100; 0>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>
Тулайковская 10	<100; 0>	<100; 0>	<0; 100>	<100; 0>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>
Тулайковская 100	<100; 0>	<100; 0>	<0; 100>	<100; 0>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<20; 0>	<80; 0>
Тулайковская золотистая	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>
Экада 70	<100; 0>	<0; 100>	<0; 100>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>
Экада 66	<100; 0>	<0; 100>	<0; 100>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>
Тулайковская 110	<100; 0>	<100; 0>	<0; 100>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>
Тулайковская надежда	<100; 0>	<100; 0>	<0; 100>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<75; 0>	<25; 0>	<0; 100>
Тулайковская победа	<100; 0>	<100; 0>	<0; 100>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<25; 0>	<75; 0>
Тулайковская 116	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<17; 0>	<83; 0>
Экада 214	<100; 0>	<0; 100>	<0; 100>	<100; 0>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<55; 0>	<45; 0>	<0; 100>
Зауральская волна	<100; 0>	<0; 100>	<0; 100>	<100; 0>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<0; 100>	<100; 0>

Рисунок 4.21 – Исходный формальный контекст задачи онтологического анализа сортов мягкой яровой пшеницы

Модель системы измеряемых свойств - СИС:												
	з/устойч ИВ	высоко з/устойч ИВ	экстрема льно з/устойчив	достаточно устойчив к БР	устойчив к БР	иммунит ет к БР	5-й класс (белок < 10)	4-й класс (10 <= белок < 12)	3-й класс (12 <= белок < 13,5)	2-й класс (13,5 <= белок < 14,5)	1-й класс (белок >= 14,5)	
В30-группа_1	0											
В30-группа_2		0										
В30-группа_3			0									
В30-группа_4				0								
В30-группа_5					0							
В30-группа_6						0						
В30-группа_7							0					
В30-группа_8								0				
В30-группа_9									0			
В30-группа_10										0		
В30-группа_11												
В30-группа_12												
В30-группа_13												
О-группа_1	0	х										
О-группа_2	0	х	х									
О-группа_3		0	х									
О-группа_4				0	х							
О-группа_5				0	х	х						
О-группа_6				0	х	х						
Н-группа_1							х	х	х	х	х	

Рисунок 4.22 – Система измеряемых свойств, построенная для онтологического анализа сортов мягкой яровой пшеницы, в виде совокупности пересекающихся групп сопряжённых свойств, однородных по характеру экзистенциального сопряжения свойств-членов

Рабочий формальный контекст задачи - РФК, сформированный при СПД <0+; 100> и имеющий статус "Противоречит априорным гипотезам об											
	з/устойч ив	высоко з/устойч ив	экстрема льно з/устойчив	достаточно устойчив к БР	устойчив к БР	иммунит ет к БР	5-й класс (белок < 10)	4-й класс (10 <= белок < 12)	3-й класс (12 <= белок < 13,5)	2-й класс (13,5 <= белок < 14,5)	1-й класс (белок >= 14,5)
Жигулевская	<100; 0>	<100; 0>		<100; 0>						<20; 0>	<80; 0>
Самсар	<100; 0>			<100; 0>						<49; 0>	<51; 0>
Тулайковская 5	<100; 0>			<100; 0>	<100; 0>	<100; 0>					<100; 0>
Тулайковская 10	<100; 0>	<100; 0>		<100; 0>	<100; 0>	<100; 0>					<100; 0>
Тулайковская 100	<100; 0>	<100; 0>		<100; 0>	<100; 0>	<100; 0>				<20; 0>	<80; 0>
Тулайковская золотистая	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>					<100; 0>
Экада 70	<100; 0>			<100; 0>							<100; 0>
Экада 66	<100; 0>			<100; 0>							<100; 0>
Тулайковская 110	<100; 0>	<100; 0>		<100; 0>	<100; 0>						<100; 0>
Тулайковская надежда	<100; 0>	<100; 0>		<100; 0>	<100; 0>				<75; 0>	<25; 0>	
Тулайковская победа	<100; 0>	<100; 0>		<100; 0>	<100; 0>					<25; 0>	<75; 0>
Тулайковская 116	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>					<17; 0>	<83; 0>
Экада 214	<100; 0>			<100; 0>	<100; 0>				<55; 0>	<45; 0>	
Зауральская волна	<100; 0>			<100; 0>							<100; 0>

Рисунок 4.23 – Первоначально формируемый рабочий формальный контекст задачи онтологического анализа сортов мягкой яровой пшеницы

Рабочий формальный контекст задачи - РФК, сформированный при СПД <0+; 100-> и имеющий статус "Противоречит априорным гипотезам об искоми											
Характерные (проблемные и, если имеются, потенциально критические и критические) ГСС объектов измерения:											
	з/устойчив	Высоко з/устойчив	экстремально з/устойчив	Достаточно устойчив к БР	устойчив к БР	иммунитет к БР	5-й класс (белок < 10)	4-й класс (10 <= белок < 12)	3-й класс (12 <= белок < 13,5)	2-й класс (13,5 <= белок < 14,5)	1-й класс (белок >= 14,5)
Жигулевская	<100; 0>	<100; 0>		<100; 0>						<20; 0>	<80; 0>
В30-группа_1	<100; 0>	<100; 0>									
В30-группа_2		<100; 0>									
В30-группа_4				<100; 0>							
В30-группа_12											
О-группа_1	<100; 0>	<100; 0>									
Н-группа_1										<20; 0>	<80; 0>
Самсар	<100; 0>	<100; 0>		<100; 0>						<49; 0>	<51; 0>
В30-группа_1	<100; 0>	<100; 0>									
В30-группа_4				<100; 0>							
В30-группа_12											
Н-группа_1											
Тулайковская 5	<100; 0>	<100; 0>		<100; 0>	<100; 0>	<100; 0>				<49; 0>	<51; 0>
В30-группа_1	<100; 0>	<100; 0>									
В30-группа_4				<100; 0>							
В30-группа_5					<100; 0>						
В30-группа_6						<100; 0>					
В30-группа_11											
В30-группа_12											<100; 0>
О-группа_4				<100; 0>	<100; 0>	<100; 0>					
О-группа_5				<100; 0>	<100; 0>	<100; 0>					
О-группа_6				<100; 0>	<100; 0>	<100; 0>					

Рисунок 4.24 – Фрагмент скриншота таблицы представляющей для каждого исследуемого сорта мягкой яровой пшеницы проблемные группы сопряженных свойств

Рабочий формальный контекст задачи - РФК, сформированный при СПД <0>; 100-> и имеющий статус "Нормализован":												
	злустойч ив	высоко злустойч ив	экстрем ально злустойч ив	достаточ но злустойч к БР	устойчив к БР	иммунит ет к БР	5-й класс (белок < 10)	4-й класс (10 <= белок < 12)	3-й класс (12 <= белок < 13,5)	2-й класс (13,5 <= белок < 14,5)	1-й класс (белок >= 14,5)	
Жигулевская	<100; 0>	<100; 0>		<100; 0>							<80; 0>	
Самсар	<100; 0>			<100; 0>							<51; 0>	
Тулайковская 5	<100; 0>			<100; 0>	<100; 0>	<100; 0>					<100; 0>	
Тулайковская	<100; 0>	<100; 0>		<100; 0>	<100; 0>	<100; 0>					<100; 0>	
Тулайковская	<100; 0>	<100; 0>		<100; 0>	<100; 0>	<100; 0>					<80; 0>	
Тулайковская золотистая	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>					<100; 0>	
Экада 70	<100; 0>			<100; 0>							<100; 0>	
Экада 66	<100; 0>			<100; 0>							<100; 0>	
Тулайковская	<100; 0>	<100; 0>		<100; 0>	<100; 0>						<100; 0>	
Тулайковская	<100; 0>	<100; 0>		<100; 0>	<100; 0>				<75; 0>			
Тулайковская	<100; 0>	<100; 0>		<100; 0>	<100; 0>						<75; 0>	
Тулайковская	<100; 0>	<100; 0>	<100; 0>	<100; 0>	<100; 0>						<83; 0>	
Экада 214	<100; 0>			<100; 0>	<100; 0>				<55; 0>			
Зауральская	<100; 0>			<100; 0>							<100; 0>	

Рисунок 4.25 – Нормализованный рабочий формальный контекст задачи онтологического анализа сортов мягкой яровой пшеницы

5 АПРОБАЦИЯ РАЗРАБОТАННЫХ МЕТОДОВ И СРЕДСТВ

5.1 Поддержка структурного анализа и синтеза технических решений

В данном подразделе рассмотренные модели ОАД получают применение и развитие для *структурного анализа и синтеза технических решений* (ТР) [78]. Конкретно речь идет об использовании для структурного анализа и синтеза ТР моделей и методов, применяемых в задаче *семантической идентификации объекта*, которая занимает одно из важнейших мест в разработанной методике формирования контекста для вывода формальных понятий в ОАД [47, 54].

В часто цитируемой статье [79] отмечается, что большая часть работ по теории автоматизированного проектирования посвящена вопросам параметрического синтеза и геометрического моделирования технических систем, а структурному синтезу технических решений уделяется несопоставимо меньшее внимание. Вместе с тем, нельзя сказать, что соответствующий этап проектирования, задачей которого является построение некоторой упрощенной функционально-структурной модели будущего артефакта, формирование его *облика* (*компоновочной схемы, эскиза* и т.п.), с позиций инженерии знаний «слабоструктурирован» или «слабодокументирован» [1]. Напротив, эти действия признаются актуальными для инженерного творчества вообще, и существует апробированный формализованный подход для их поддержки – *морфологический анализ и синтез* ТР.

Разъяснению, развитию и применению морфологического анализа и синтеза ТР посвящена обширная литература [80-85].

В предельно сжатом изложении *морфологический анализ* области, где предстоит выбрать решение, состоит в формировании таблицы, описывающей «*морфологическое пространство*» [86] выбора ТР как множество дискретных точек, каждая из которых определяет комбинацию-*трансверсаль* [87] *вариантов реализаций* (ВР) каждой из функций (шире – *функционально-структурных подсистем, функциональных подсистем*, – далее ФСП) проектируемого артефакта. А *морфо-*

логический синтез ТР сводится к выполнению некоторой *процедуры рационального выбора* определённой точки морфологического пространства (рисунок 5.1).

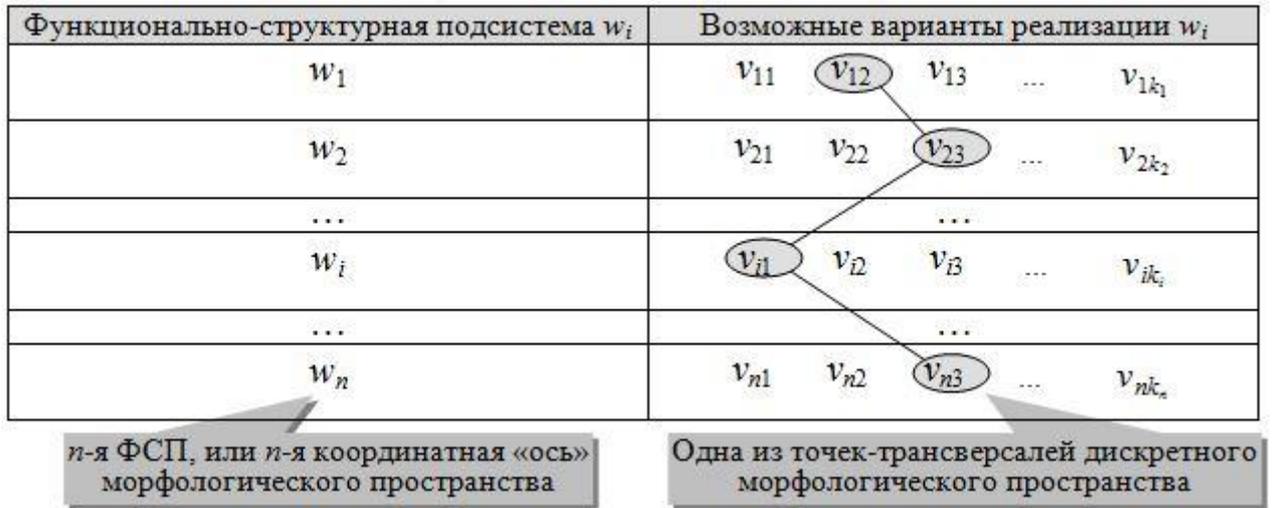


Рисунок 5.1 – Табличное описание морфологического пространства решений

Используя, в частности, введенные модели экзистенциальных отношений можно предложить следующее формальное описание морфологического пространства:

- множество ФСП $W = \{w_i\}_{i=1, \dots, n}$, где $n \geq 1$;
- множество ВР ФСП V такое, что $V = \cup_{i=1, \dots, n} V_i$, $V_i = \{v_{ij}\}_{j=1, 2, \dots, k_i}$, $k_i \geq 1$, и $\forall i, j$, $i \neq j$: $V_i \cap V_j = \emptyset$ (пример подхода к моделированию морфологического пространства, когда $\exists i, j, i = j$: $V_i \cap V_j \neq \emptyset$, есть в [86]);
- множество потенциальных ТР (т.е. множество точек-трансверсалей морфологического пространства) T , у которого $|T| = \prod_{i=1, 2, \dots, n} k_i$;
- бинарное отношение *взаимообусловленности*, которое полно на множестве W – $MC: W \times W \rightarrow \mathbf{True}$ (в морфологическом анализе и синтезе обычно это явно не фиксируется, но *фактически предполагается*, что ТР состоятельно тогда и только тогда, когда реализованы все ФСП проектируемого артефакта);
- бинарное отношение *несовместимости* (попарной несовместимости) $E: V \times V \rightarrow \{\mathbf{True}, \mathbf{False}\}$, – которое здесь конфигурируется так, что

$$\forall i \forall x \in V_i: E(x) = V_i \setminus \{x\} \quad (5.1)$$

(именно такая несовместимость ВР каждой отдельно взятой ФСП характерна для классической версии морфологического анализа и синтеза ТР);

- бинарное отношение «является вариантом реализации» $R: V \times W \rightarrow \{\mathbf{True}, \mathbf{False}\}$, – характерная особенность которого состоит в том, что $\forall i \forall x \in V_i: R(x) = w_i$ (т.е. R – отношение между каждым ВР некоторой ФСП и этой ФСП);
- бинарное отношение «ТР включает ВР ФСП» $S: T \times V \rightarrow \{\mathbf{True}, \mathbf{False}\}$, – причем $\forall x \in T, \forall i: (|S(x)| = n) \wedge (|S(x) \cap V_i|) = 1$.

Пример, представленный таблицей 5.1, упрощенно определяет возможные компоновки холодильной машины⁴. Строка «Испаритель» наводит на мысль о потенциальной полезности введения *многоуровневого описания* морфологического пространства: на нулевом (базовом) уровне – морфологической таблице холодильного устройства – в строке «Испаритель» можно оставить варианты его формы; на первом уровне описать одномерное в данном случае морфологическое пространство материалов, пригодных для изготовления испарителя.

Таблица 5.1 – Морфологическое пространство компоновок холодильной машины

Функциональная подсистема	Возможные варианты реализации подсистемы					
1. <i>Хладагент</i>	1.1 Аммиак	1.2 Пропан	1.3 Пропилен	1.4 Изобутан	1.5 Тетрафторэтан	
2. <i>Компрессор</i>	2.1 Поршневой	2.2 Ротационный	2.3 Центробежный		2.4 Тепловой (абсорбер)	
3. <i>Конденсатор</i>	3.1 Жидкостного испарения		3.2 Воздушного испарения		3.3 Испарительного охлаждения	
4. <i>Дроссель</i>	4.1 Капиллярная трубка	4.2 Автоматический клапан	4.3 Терморегулируемый вентиль		4.4 Электронно-управляемый вентиль	
5. <i>Испаритель</i>	5.1 Трубчатый медный	5.2 Трубчатый стальной	5.3 Пластинчатый медный	5.4 Пластинчатый стальной	5.5 Ребристый медный	5.6 Ребристый стальной

Развитие морфологического анализа, осуществляющее идею многоуровневых морфологических описаний, достигается путём введения бинарного отношения между множеством ВР ФСП, описанным на уровне q , и множеством ФСП на уровне $q + 1$, $q \geq 0$. Это отношение «содержит подсистему» $P: V^{(q)} \times W^{(q+1)} \rightarrow \{\mathbf{True}, \mathbf{False}\}$, которое отличается тем, что $\forall x, y \in V^{(q)}: P(x) \neq P(y)$,

⁴ https://eti.su/articles/over/over_1534.html; <https://studfile.net/preview/5125548/page:11>; <https://ikuhnya.com/holodilnik/chto-takoe-ispavitel-holodilnika/?ysclid=ll3nj6fuun116740709>; <https://extxe.com/25057/kondensatory-holodilnoj-tehniki-vidy-ustrojstvo-kondensatorov/?ysclid=ll3mndbocht508977940#2>; <https://bon-nix.ru/основы-холодильной-техники.html>

и $\forall x \in W^{(q+1)}: |P(x)| = 1$. В совокупности с ранее введёнными формализмами получаем представление морфологического пространства в виде широко известной математической модели И/ИЛИ-дерева (графа) [83, 88, 89]. Подобная модель для холодильной машины приведена на рисунке 5.2 с явной графической нотацией всех определенных выше элементов морфологического анализа (но без представления точек-трансверселей).

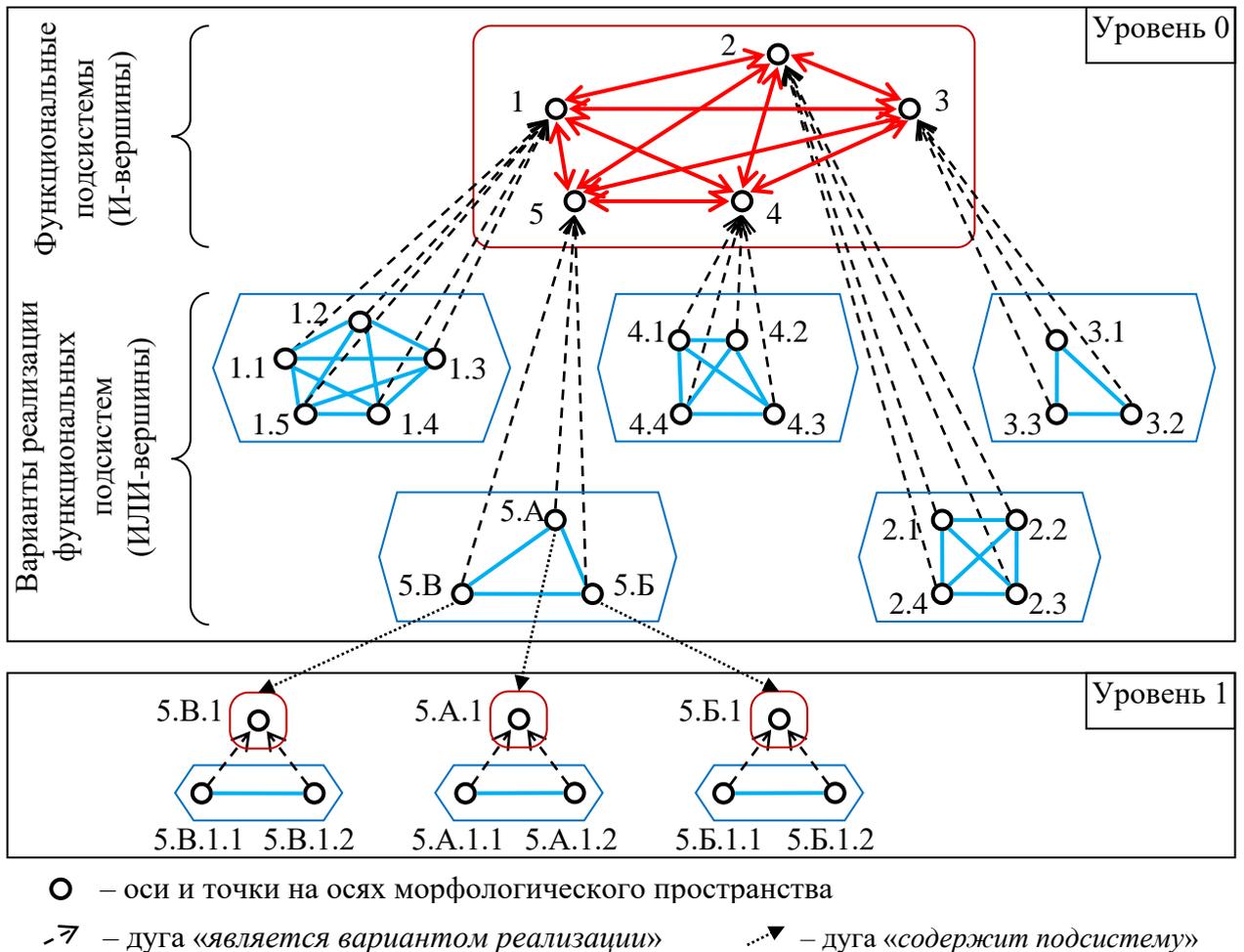


Рисунок 5.2 – Многоуровневая морфологическая модель холодильной машины в форме И/ИЛИ графа с детализацией отношений между элементами морфологического пространства (см. таблицу 5.1; здесь: 5.А, 5.Б, 5.В – соответственно трубчатая, пластинчатая и ребристая конструкции испарителя; 5.*.1 – материал, пригодный для изготовления испарителя «*»; 5.*.1.1 – медь; 5.*.1.2 – сталь)

5.1.1 Ограниченная сочетаемость вариантов реализации различных функционально-структурных подсистем

Обычно на ВР ФСП накладываются различные – физические, технические, технологические, гуманитарные и др. – ограничения. В морфологическом плане они могут выражаться в том, что некоторые точки-трансверсели морфологическо-

го пространства определяют ТР, которые в разных смыслах нежизненны, утопичны или, например, охраняются авторскими правами. Подобные обстоятельства в большинстве практических случаев могут быть интерпретированы как *ограниченная сочетаемость* ВР различных ФСП в одном ТР, которая проявляется либо как предусловие, либо как постусловие выбора ВР одной ФСП – $x \in V_i$, – на выбор ВР другой ФСП – $y \in V_j, j \neq i$ [79]:

- *принуждение*, когда выбор x влечет, обуславливает выбор y , т.е. $C(x, y)$;
- *необходимость*, когда условием выбора x служит выбор y , т.е. $C(y, x)$;
- *запрет на сочетание*, когда x и y не могут входить в одно ТР, т.е. $E(x, y)$;
- *двойное принуждение*, когда x и y могут входить в ТР лишь одновременно, т.е. x и y *взаимобусловлены* – $C(x, y) \wedge C(y, x) = MC(x, y)$.

Рисунки 5.3а и 5.3б демонстрируют две версии описания ограниченной сочетаемости реализаций холодильной машины.

Каждая версия одинаково учитывает знание о разрушительном действии аммиака на медную арматуру, но по-разному описывает то, что из хладагентов, указанных в таблице 5.1, только аммиак пригоден для постройки абсорбционного холодильника⁵. Для упрощения восприятия на этих рисунках взаимобусловленность ФСП отражена лишь охватывающим их контуром – скругленным прямоугольником. Аналогично – но шестиугольным контуром – показывается несовместимость ВР каждой ФСП, а связь этих вариантов с ФСП заменена соответствующей единичной связью контура и ФСП.

Предлагается следующее обобщение модели морфологического пространства для возможности описания ограниченной сочетаемости ВР различных ФСП.

Во-первых, для отражения возможной несовместимости ВР различных ФСП уточняется условие (5.1):

$$\forall i \forall x \in V_i: E(x) = V_i \setminus \{x\} \cup Y, \quad (5.2)$$

где $Y \subseteq V \setminus V_i$, т.е. ВР некоторой ФСП несовместим не только с другими ВР этой ФСП, но, возможно, и с некоторыми ВР других ФСП.

⁵ <https://habr.com/ru/companies/ruvds/articles/749776/>; <https://wihometals.com/ru/why-ammonia-is-not-compatible-with-copper>

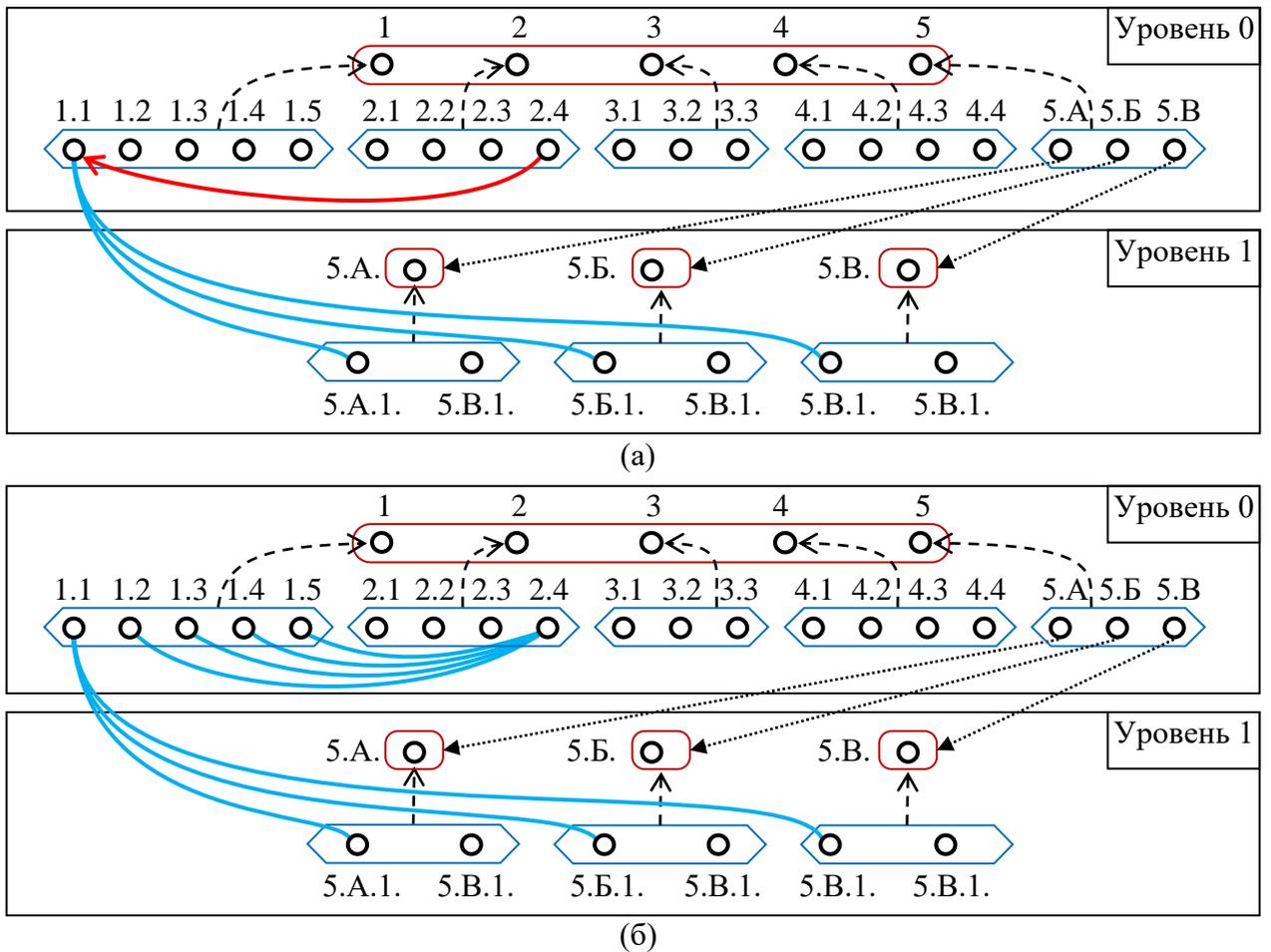


Рисунок 5.3 – Две версии многоуровневой морфологической модели холодильной машины при ограниченной сочетаемости вариантов реализации её различных функционально-структурных подсистем

Во-вторых, для описания потенциальной обусловленности ВР различных ФСП в модель вводится соответствующее отношение – $C: V \times V \rightarrow \{\mathbf{True}, \mathbf{False}\}$. Это знаковый акт, т.к. определение на одной основе и отношения E , и отношения C устанавливает *ограничения существования* элементов этой основы – именно так описываются ОСС у объектов исследуемой ПрО в ОАД [9, 10, 15].

На основании (1.3) из пункта 1.2.2 следует констатировать, что пример на рисунке 5.3а некорректен. Формально верная морфологическая модель холодильной машины, учитывающая обнаруженную ограниченную сочетаемость ВР её ФСП получится при *совмещении* рисунков 5.3а и 5.3б.

Математически корректна в этом смысле морфологическая модель, показанная на рисунке 5.4. Рисунок 5.4, информационно поддерживаемый таблицей 5.2, демонстрирует морфологическое пространство с ограничениями существования ВР ФСП, сформированное для выбора облика реактивного самолёта.

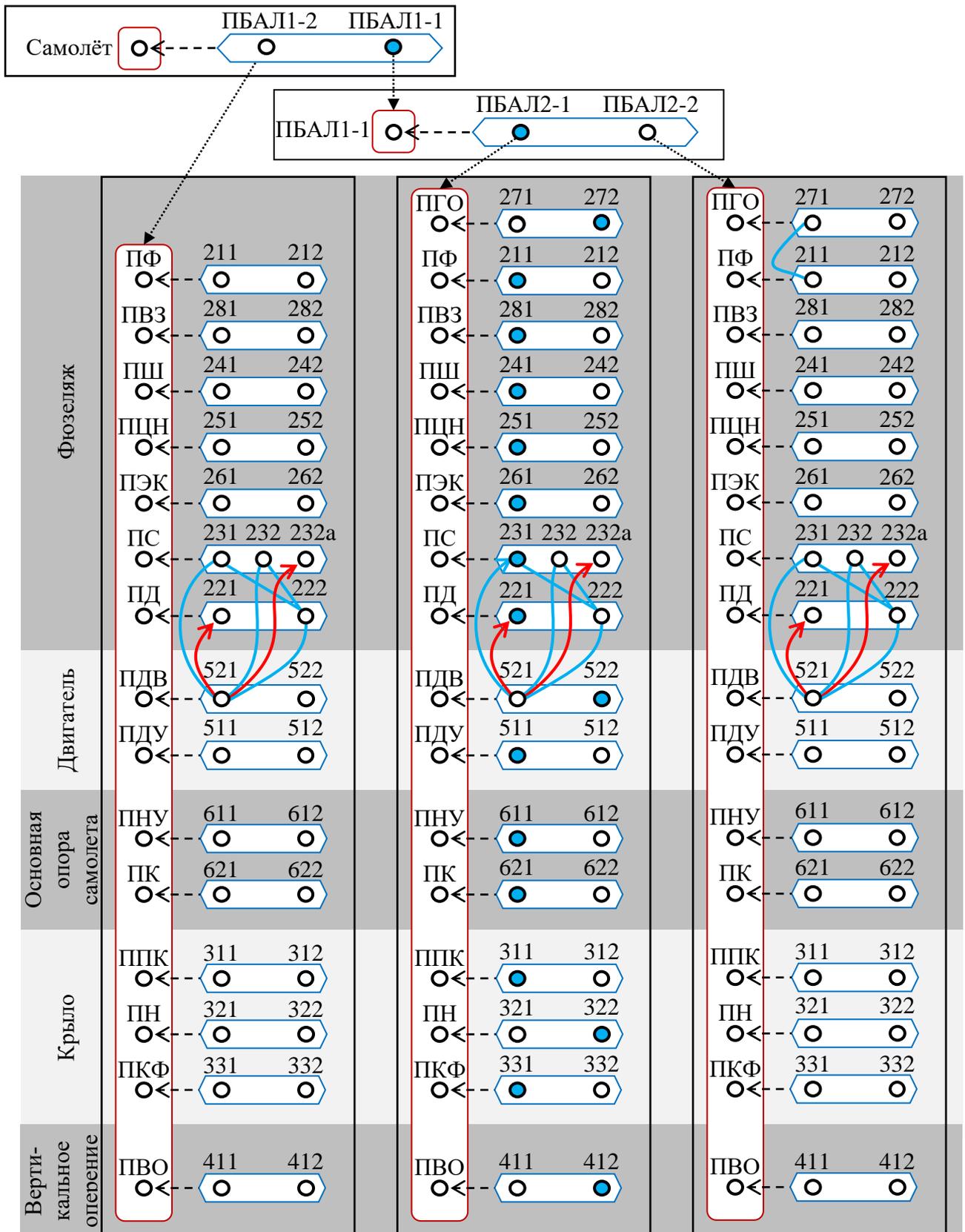


Рисунок 5.4 – Морфологическая модель признаков компоновочной схемы самолёта (по [91]; расшифровка оригинальных аббревиатур и цифровых кодов признаков представлена в таблице 5.2); ● - признаки компоновочной схемы истребителя F-15A (см. рисунок 5.5)

Таблица 5.2 – Признаки компоновочной схемы самолёта согласно [91]

Аббревиатура, код	Расшифровка
ПБАЛ1-1	Горизонтальное оперение имеется
ПБАЛ1-2	Горизонтальное оперение отсутствует
ПБАЛ2-1	Горизонтальное оперение располагается по оси X позади крыла
ПБАЛ2-2	Горизонтальное оперение располагается по оси X впереди крыла
ПГО	Размещение горизонтального оперения
271	На фюзеляже
272	На агрегатах, примыкающих к фюзеляжу
ПФ	Расположение воздухозаборников (ВЗ)
211	Сбоку фюзеляжа
212	Под фюзеляжем
ПВЗ	Схема воздухозаборников
281	Плоский воздухозаборник с горизонтальным клином торможения
282	Плоский воздухозаборник с вертикальным клином торможения
ПШ	Размещение колеса главной опоры относительно канала подвода воздуха к двигателю
241	Снизу-сбоку канала подвода воздуха к двигателю
242	Сверху-сбоку канала подвода воздуха к двигателю
ПЦН	Схема фюзеляжа по размещению целевой нагрузки
251	Наружная подвеска целевой нагрузки
252	Конформная подвеска целевой нагрузки
ПЭК	Размещение экипажа
261	Тандем
262	Рядом
ПС	Схема донной части фюзеляжа
231	Сопла двигателей разнесены
232	Сопла двигателей сомкнуты
232а	Сопло одного двигателя
ПД	Размещение двигателей
221	В едином отсеке фюзеляжа
222	В изолированных отсеках фюзеляжа
ПДВ	Число двигателей
521	Один
522	Два
ПДУ	Тип двигателей
511	ДТРДФ
512	ТРДФ
ПНУ	Направление уборки основной опоры самолёта
611	Вперёд (против полёта)
612	Назад (по полёту)
ПК	Число колёс
621	Одно
622	Два
ППК	Стреловидность передней кромки крыла
311	Прямая стреловидность
312	Обратная стреловидность
ПН	Конфигурация передней кромки крыла
321	С наплывом
322	Без наплыва

Продолжение таблицы 5.2

ПКФ	Расположение крыла относительно строительной горизонтали фюзеляжа
331	Высокоплан
332	Низкоплан
ПВО	Число килей вертикального оперения
411	Один
412	Два

Пример построен по материалам [90], где описание морфологического пространства выбора ТР, сделанное в значительной степени неформально, уточнено введением варианта 232а для схемы донной части фюзеляжа самолёта. В [90] достоверность этого описания подтверждается, в частности, идентификацией на его основе облика реального выпускавшегося самолёта-истребителя (см. рисунки 5.4 и 5.5).

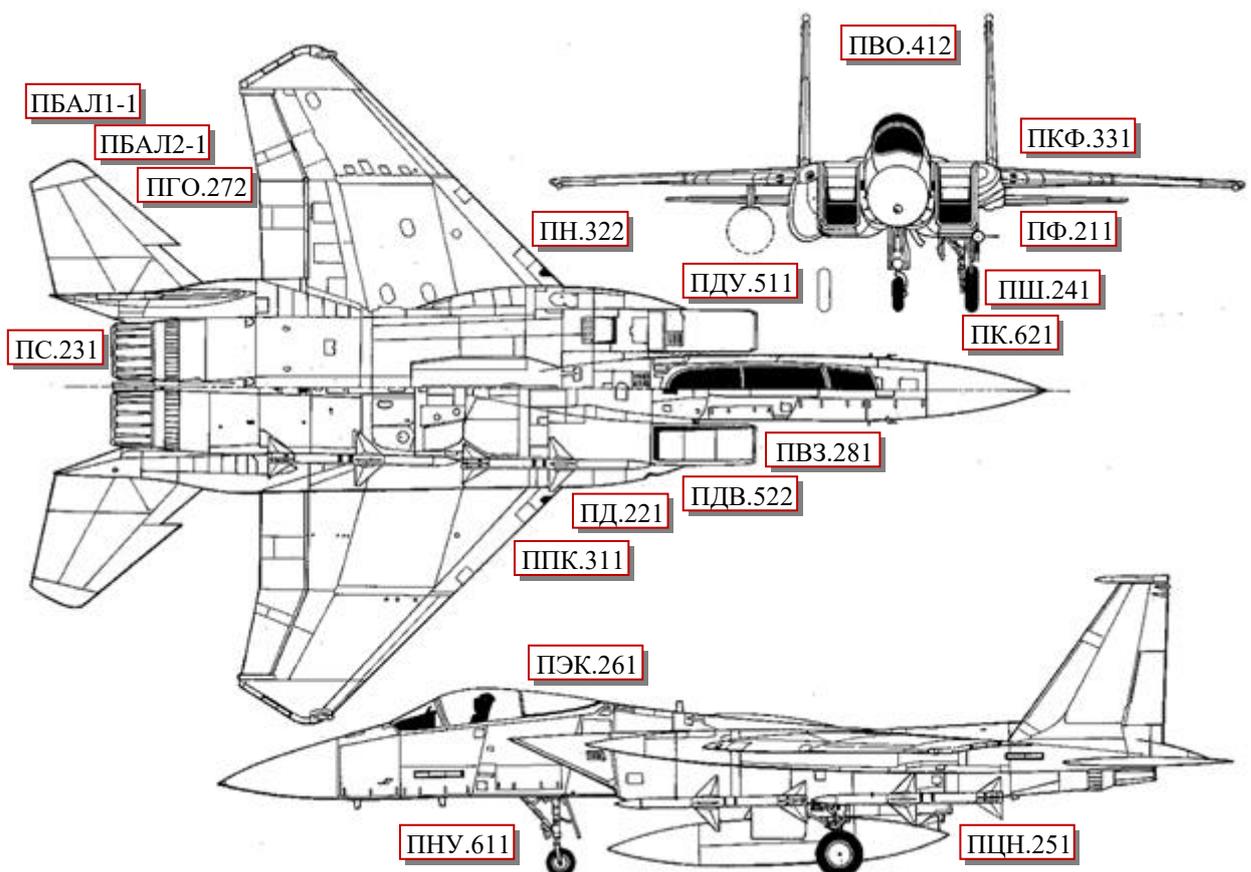


Рисунок 5.5 – Компоновочная схема истребителя F-15A Eagle (корпорация «Макдоннел-Дуглас», США)⁶

⁶ <http://www.airwar.ru/image/idop/fighter/f15a/f15a-1.gif>

5.1.2 Апология запретов и преимущества ограничений существования

Обычно при необходимости введения в морфологическую модель ограниченной сочетаемости ВР различных ФСП проектируемого артефакта стараются обойтись «запретами на сочетания», т.е. ограничиться расширением (5.2) базовой модели морфологического пространства. Это объясняется тем, что в этом случае морфологический синтез ТР получает хорошую теоретическую поддержку, связанную с моделированием и анализом *многодольных* графов, а также методами и алгоритмами решения задачи о *независимом множестве* вершин графа [91].

Аргументация такого подхода, как правило, ограничивается экспертным мнением, что «запреты на сочетания» *чаще всего* встречаются в проектной практике инженеров различного профиля [79], что *во многих случаях* с помощью одних только запретов можно обосновывать различные закономерности в природе и технике [92, 93].

Покажем, что моделирование ограниченной сочетаемости ВР различных ФСП с помощью *ограничений существования* – используемой в ОАД фундаментальной модели сосуществования элементов некоего целого – позволяет строго обосновать *апологию запретов* в морфологическом анализе и синтезе ТР.

Прежде всего, следует отметить, что за исключением запретов *все виды* ограниченной сочетаемости ВР различных ФСП описываются в рамках бинарного отношения обусловленности C (см. пункт 5.1.1). Поэтому достаточно установить, что на этапе синтеза ТР обусловленность одного отдельно взятого ВР x некоторой ФСП отдельно взятым ВР y другой ФСП *неизбежно* при включении в ТР y приведёт к включению в ТР x независимо от того, есть или отсутствует информация о рассматриваемой обусловленности (т.е. включение x будет осуществлено *независимо от знания* об этой в действительности существующей обусловленности благодаря другому механизму, и, забегая вперёд, его реализуют запреты).

Пусть $x \in V_i$, $y \in V_j$, $i \neq j$ и $C(y, x)$. Тогда согласно (1.3) $E(y) \supseteq V_i \setminus \{x\}$, т.е. в морфологической модели по меньшей мере имеют место несовместимости вида

$$\forall z \in V_i \setminus \{x\}: E(y, z). \quad (5.3)$$

Следовательно, при включении в ТР y для реализации j -й ФСП для реализации i -й ФСП не остаётся никакого другого ВР кроме x . Если обусловленность $C(y, x)$ отсутствует (явно не зафиксирована), а несовместимости (5.3) *определены*, произойдет то же самое. Поэтому *принципиально возможно* при описании ограниченной сочетаемости ВР различных ФСП заместить обусловленность запретами.

В этом смысле морфологическая модель холодильной машины на рисунке 5.3б также корректна, как и модель, возникающая при совмещении схем на рисунках 5.3а и 5.3б.

Тем не менее описание морфологического пространства в задаче структурного анализа и синтеза ТР с использованием *ограничений существования* имеет ряд неоспоримых преимуществ:

- *естественность* отображения всех разновидностей ограниченной сочетаемости ВР различных ФСП, в частности, непосредственное описание 3-х из 4-х их видов с помощью отношения обусловленности;
- «автоматизм» *вскрытия закономерностей* ограниченной сочетаемости ВР различных ФСП в форме запретов (несовместимостей) при фиксации в модели обусловленности некоторых таких ВР, который возникает при формировании морфологического пространства благодаря транзитивности несовместимости относительно обусловленности (1.3);
- возможность эффективной *редукции ограничений существования* [32, 54], которая здесь означает выявление классов эквивалентности во множестве точек-трансверселей морфологического пространства, что снижает размерность задачи синтеза ТР.

Таким образом, модель ограничений существования, получившая важное применение и развитие в ОАД, следует признать пригодной и предпочтительной для описания ограниченной сочетаемости ВР различных ФСП в задаче структурного анализа и синтеза ТР.

5.1.3 Структурный синтез технического решения как семантическая идентификация объекта проектирования

Преимущества использования ограничений существования в задаче структурного анализа и синтеза ТР будут состоятельны, если при этом способе описания в морфологическом пространстве можно осуществить эффективный поиск оптимальной точки-трансверсали (далее для определенности – *максимальной* по «весу» ВР ФСП, «включаемых» в ТР).

Поскольку, во-первых, задача поиска максимального по весу независимого множества вершин в графе, где вершины – компоненты решения, а ребра – свидетельства их несовместимости, является *NP*-полной [94], а, во-вторых, в п. 5.1.2 показана возможность *приведения* к такой постановке рассматриваемой поисковой задачи, где связи вершин описываются ограничениями существования, то последняя также *NP*-полна. Это означает, что трудоемкость решения задачи структурного синтеза ТР независимо от способа описания морфологического пространства – с помощью запретов или ограничений существования – экспоненциально зависит от параметров размерности исходных данных, и улучшить это положение можно лишь за счет эвристик и/или организации поиска приближенных решений.

Монотонность АПД определяет преимущество разработанного в разделе 3 эвристического метода, т.к. обеспечивает возможность отсекаания бесперспективных ветвей конечного дерева решений, формируемого для каждого набора связанных клик [91] в M , или, другими словами, для каждой связной группы попарно несовместимых свойств объекта g .

Вспомогательным критерием качества рассматриваемого нормального множества служит его мощность: при прочих равных условиях желательно уменьшить отклонение от эмпирического признакового описания объекта и, следовательно, сохранить в этом описании как можно больше свойств, эмпирические свидетельства о наличии которых признаны в исходной постановке задачи достаточными. Явно выраженная разница важности частных критериев оправдывает применение для многокритериальной оптимизации *лексикографического метода* [95].

Соответствие элементов задачи семантической идентификации объекта и задачи структурного синтеза ТР состоит в следующем:

- объект \leftrightarrow ТР;
- измеряемое свойство объекта \leftrightarrow отдельно взятый ВР некоторой ФСП проектируемого артефакта;
- нормальное подмножество свойств объекта \leftrightarrow допустимое ТР;
- оценка истинности БСС «объекту g присуще свойство m » $\langle b_{gm}^+, b_{gm}^- \rangle \leftrightarrow$ оценка эффекта включения в ТР отдельно взятого ВР $x \in V$ некоторой ФСП, которая на практике часто имеет векторный вид $\langle b_x^+, b_x^- \rangle$, где b_x^+ – выгода, а b_x^- – издержки решения о включении x в ТР [86]. Величины выгод и издержек можно нормировать после их выявления для всех ВР из V , поэтому $b_x^+, b_x^- \in [0, 1]$. В случае нормированной скалярной оценки b_x эффекта включения ВР x в ТР (например, в [86] такой эффект оценивается величиной $b_x = b_x^+ / b_x^-$) можно оперировать векторной оценкой $\langle b_x, 0 \rangle$;
- достоверность БСС \leftrightarrow абсолютный «выигрыш» $(b_x^+ - b_x^-) \in [-1, 1]$ от включения в ТР отдельно взятого ВР $x \in V$ некоторой ФСП, который аналогично достоверности БСС может быть трансформирован в «приведенный выигрыш» в диапазоне $[0, 1]$;
- АПД (5) \leftrightarrow аналогично вычисляемый критерий эффективности ТР. Физический смысл этого критерия состоит в указании верхней оценки приведенного выигрыша, «теряемого» (как плата за корректность решения) с каждым отдельно взятым ВР, исключенным из ТР.

Вспомогательный критерий качества нормального подмножества свойств объекта, используемый в задаче семантической идентификации объекта, в задаче структурного синтеза ТР не имеет смысла, т.к. каждое допустимое ТР, если такое найдётся, есть множество ВР по одному для каждой ФСП, которое всегда имеет мощность $n = |W|$.

То, что при заданных ограничениях существования решение задачи синтеза ТР *может не существовать*, отличает её от задачи семантической идентифика-

ции свойств объекта, решение которой всегда существует [53]. Содержательно это объясняется тем, что в первой задаче не допускается исключение из ТР всех ВР некоторой ФСП. Формально же отличие заключается лишь в том, считается ли *пустое* множество нормальным. Опционный запрет квалифицировать пустое множество как нормальное делает метод семантической идентификации объекта более гибким и пригодным для решения задачи структурного синтеза ТР.

Рисунок 5.6 иллюстрирует решение задачи синтеза ТР на примере холодильной машины. Слева представлены ВР ФСП и ОСС: «линейная» форма морфологической таблицы 1 устанавливает несовместимость ВР каждой ФСП, а дуга и ребра описывают ограниченную сочетаемость ВР разных ФСП. В средней части рисунка 5.6 приведены нормированные значения выгод и издержек, связанных с включением в ТР отдельных ВР ФСП (в процентах), справа демонстрируются ВР ФСП – компоненты ТР, найденного методом семантической идентификации объекта при запрете считать пустое множество нормальным.

5.2 Поддержка коллективного построения когнитивных карт

Когнитивные карты – часто применяющийся на практике и активно развивающийся комплекс моделей, методов и компьютерных средств для формализации экспертных знаний при управлении слабоструктурированными ситуациями (см., например, обзоры [96, 97]).

При этом специалистами отмечается факт слабой поддержанности формальными моделями и методами *этапа описания* представлений о слабоструктурированной динамической ситуации – этапа построения когнитивных карт экспертами [96, 98]. В основном данная проблема возникает при согласовании понимания разными экспертами анализируемой проблемной ситуации, когда формируемая когнитивная карта является продуктом коллективного труда. Увязка знаний и мнений экспертов обеспечивается в таком случае неформальными действиями выделенного координатора работы, в некоторых случаях с поддержкой современных программно-технических средств удаленного взаимодействия работников [99].

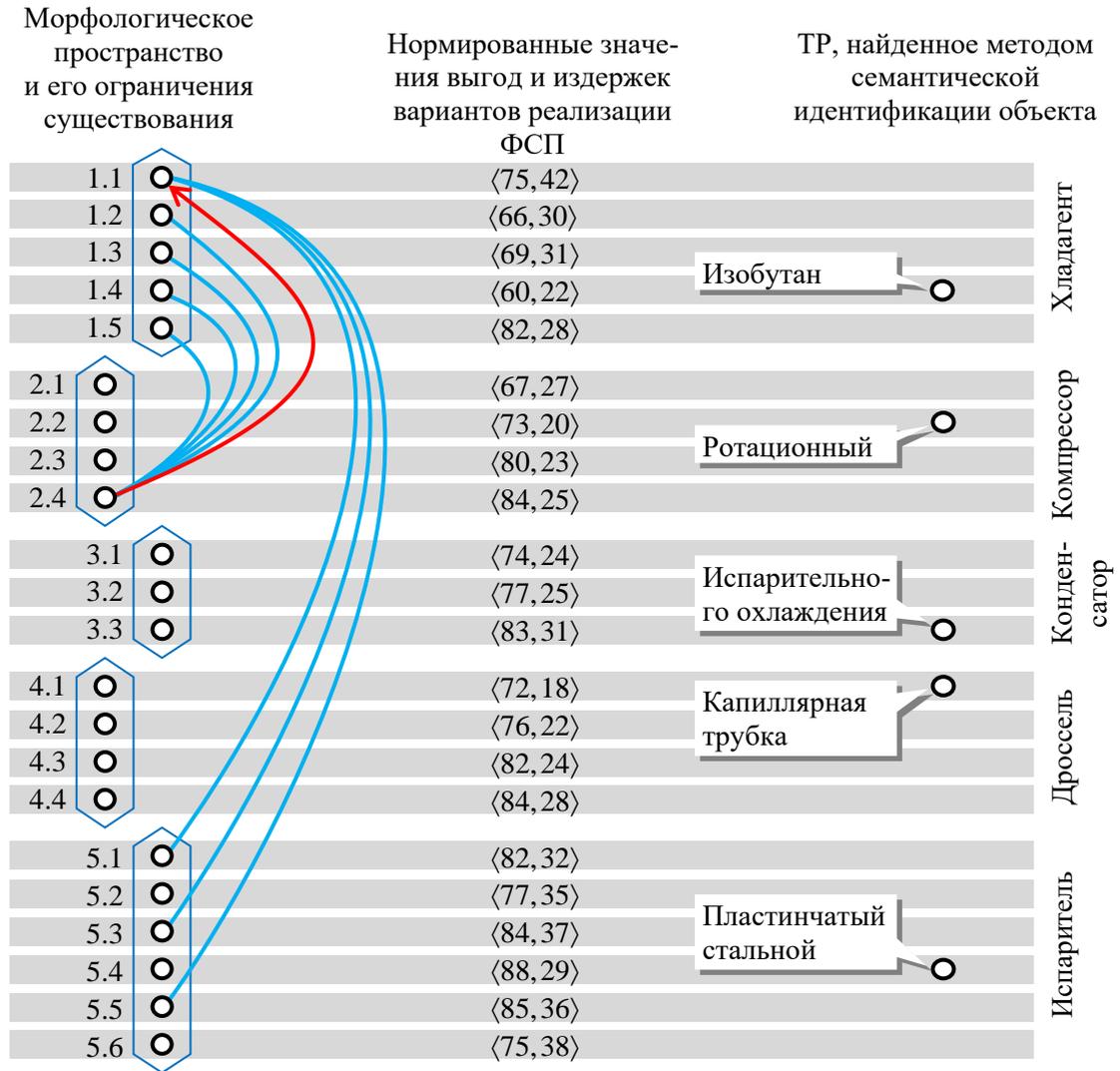


Рисунок 5.6 – Пример решения задачи синтеза облика артефакта, полученного методом семантической идентификации объекта на основе описания морфологического пространства выбора решений и учета ограниченной сочетаемости вариантов реализации функциональных подсистем в форме ограничений существования

Представляется, что в коллективном построении когнитивных карт могут быть применены подходы, используемые на отдельных этапах онтологического анализа данных, описанных ранее в разделах 2 и 3. Рассмотрим соответствующие модели и методы на примере формирования знакового графа – исторически первой и фундаментальной когнитивной модели динамической ситуации.

Знаковый граф может быть описан кортежем (F, W) , где $F = \{f_i\}_{i=1, \dots, n}$ – множество вершин-факторов ситуации, $W = (w_{ij})_{i=1, \dots, n; j=1, \dots, n}$ – матрица смежности графа: $w_{ij} \in \{+1, 0, -1\}$, $w_{ij} \neq 0$ свидетельствует о существовании в графе ребра (i, j) , фиксирующего влияние i -го фактора на j -й – «положительное» при $w_{ij} = +1$ и «отрицательное» при $w_{ij} = -1$, а $w_{ii} = 0$ свидетельствует об отсутствии влияния.

Результат коллективной работы k экспертов над формированием подобной когнитивной карты можно выразить следующим образом:

$$F = \cup_{i=1, \dots, k} F_i,$$

$$W = \oplus_{i=1, \dots, k} W_i.$$

Это означает учет в модели *по возможности всех* факторов, выделяемых каждым экспертом (здесь не касаемся лингвистического аспекта осуществления подобных действий на практике, а также куда более сложной проблемы соразмерности объемов понятий различных факторов [98]), и формирование *некоторой композиции* описаний межфакторных влияний W_i , указанных для F_i каждым отдельным экспертом.

Покажем применение моделей и методов формирования контекста для вывода ФП для построения корректной когнитивной модели на основе коллективной экспертизы.

Воспользуемся базовым положением работы [100], где в качестве объектов исследуемой ПрО рассматриваются выделенные экспертами упорядоченные пары факторов (f_i, f_j) , $i \neq j$, $f_i, f_j \in \cup_{i=1, \dots, k} F_i$. БСС в этой ПрО являются независимые заключения экспертов не только о «положительном» («+1») и «отрицательном» («-1») влиянии i -го фактора на j -й (как в [100]), но и об отсутствии влияния («0»).

Данные, поступившие от экспертов, структурируются в виде ОТОС (таблица 5.3), где константа **X** указывает на наличие, а **None** – на отсутствие у объекта того или иного свойства, т.е. соответствующего влияния f_i на f_j . Константа **NM** указывает, что экспертом оценка соответствующего влияния не проводилась. Константа **Failure** фиксирует иногда наблюдаемый на практике случай сомнения эксперта в существовании влияния i -го фактора на j -й или, наоборот, уверенности в наличии такого влияния, но сомнения в его «знаке».

Считаем, что каждая отдельная экспертиза корректна:

- если отдельный эксперт выделил некоторых два фактора f_i и f_j , $i \neq j$, то в его ответе однозначно фиксируется наличие только одного вида влияния f_i на f_j – либо «0», либо «+1», либо «-1» (см. таблицу 5.3);

- каждый фактор, выделяемый экспертом, связан хотя бы с одним другим выделенным экспертом фактором – либо влияет на него, либо подвержен влиянию со стороны этого фактора;
- если хотя бы один фактор в паре (f_i, f_j) , $i \neq j$ не был выделен экспертом, то все позиции в его ответе занимают константы **NM** (см. таблицу 5.3);
- если эксперт отказывается от определенной характеристики влияния одного из выделенных им факторов на другой указанный им фактор, то все позиции в его ответе занимают константы **Failure** (см. таблицу 5.3).

Дополнительно ОТОС может отражать степени доверия к отдельным экспертам, и независимые серии оценок одного и того же эксперта (возникающие, например, в результате нескольких его «подходов» к задаче), причем при различном количестве серий у каждого эксперта и отличающейся достоверности данных, фиксируемых в каждой серии [15, 101].

Таблица 5.3. Пример данных экспертизы влияний факторов ситуации в виде обобщённой таблицы «объекты-свойства»

	Эксперт 1			...	Эксперт k		
	«0»	«+1»	«-1»		«0»	«+1»	«-1»
(f_1, f_2)	X	None	None	...	NM	NM	NM
(f_1, f_3)	None	None	X		None	None	X
...
(f_n, f_{n-1})	Failure	Failure	Failure		X	None	None

Нетрудно видеть, что данные, собранные с разных экспертов, потенциально характеризуются неполнотой и противоречивостью, что вписывается в рамки предложенной в работе методики формирования контекста для вывода ФП. Что же насчёт *ограничений существования* межфакторных влияний в знаковом орграфе, то в литературе не удалось найти явных формулировок указанных ограничений, но, по крайней мере, некоторые из них непосредственно следуют из фундаментальных определений знаковых моделей динамической ситуации [96]:

- в упорядоченной паре факторов (f_i, f_j) , $i \neq j$ фактор f_i влияет на фактор f_j либо «положительно», либо «отрицательно» (т.е. свойства «+1» и «-1» упорядоченной пары факторов несовместимы);

- в ориентированном знаковом графе отсутствуют петли, т.е. пара факторов (f_i, f_i) заведомо исключается из анализа влияний;
- ориентированный знаковый граф описывает совокупность прямых влияний одного фактора ситуации на другой (а не прямые влияния вычисляются по известным правилам), другими словами, фактор f_i может влиять на фактор f_j либо только прямо, либо только косвенно.

Путем интерпретации и консолидации данных экспертизы, структурированных как ОТОС (таблица 5.3), с помощью констант и правил многозначной логики V^{TF} получаем нестрогий ФК, включающий множество объектов в виде пар факторов (f_i, f_j) , $i \neq j$, множество свойств $\{\langle 0 \rangle, \langle +1 \rangle, \langle -1 \rangle\}$ и нестрогое соответствие «упорядоченная пара факторов – влияние по мнению разных экспертов». Ограничением существования свойств для указанных объектов-пар является попарная несовместимость измеряемых свойств.

Тогда нормализация полученного нестрогого ФК даст согласованной результат работы коллектива экспертов:

- множество F составят лишь те из факторов, входящих в объекты-пары нормализованного ФК, которые обладают свойством $\langle +1 \rangle$ или $\langle -1 \rangle$;
- противоречия экспертов относительно взаимовлияния факторов исключаются путем отказа от наименее достоверных (редких, маргинальных) мнений, что обеспечивает корректное построение матрицы W .

5.3 Особенности задачи портретной идентификации человека в криминалистике

В Приложении А был продемонстрирован некорректный результат обработки многомерных наблюдений и измерений методами классического АФП. Ошибочность результата является следствием игнорирования ОСС, явно заданных в габитологии в классификации элементов и характеризующих их признаков.

Попытка использования рассматриваемой в работе модели ОСС наталкивается на противоречия. Согласно рисунку 4.13 из пункта 4.3.3, при попытке провести обусловленность от вида «Мужской пол» к роду «мужской рост», в силу тран-

зативности несовместимости относительно обусловленности неизбежно возникнет несовместимость между «мужским полом» и любыми двумя видами роста из трех (см. рисунок 5.7), что не соответствует реальности.

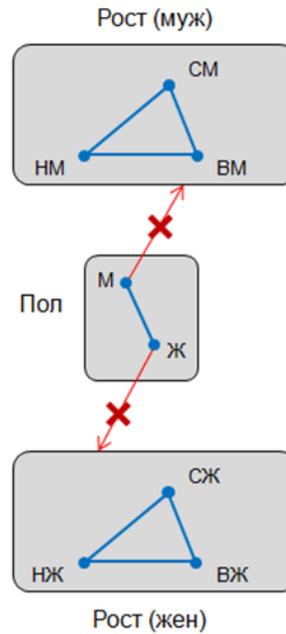


Рисунок 5.7 – Пример возникающих противоречий при использовании одноуровневых ограничений существования в габитологии

Подобный опыт привел к выводу, что в общем случае ОСС как и в рассматриваемой задаче следует представлять в многоуровневой форме, в виде И/ИЛИ графа (см. подраздел 5.1).

В случае габитологии на первом уровне моделируется взаимообусловленность характеристик, которые присущи каждому человеку и несовместимые разновидности каждой характеристики возможно с ограниченной сочетаемостью характеристиками другой природы (см. фрагмент таких ограничений на рисунке 5.7). На втором уровне детализируется описание альтернативных признаков, представленных на первом уровне.

Задачу нормализации семантического описания объекта в таком случае – здесь человека – следует решать на каждом ИЛИ-уровне отдельно, начиная с самого низшего в иерархии (см. п. 5.1.3).

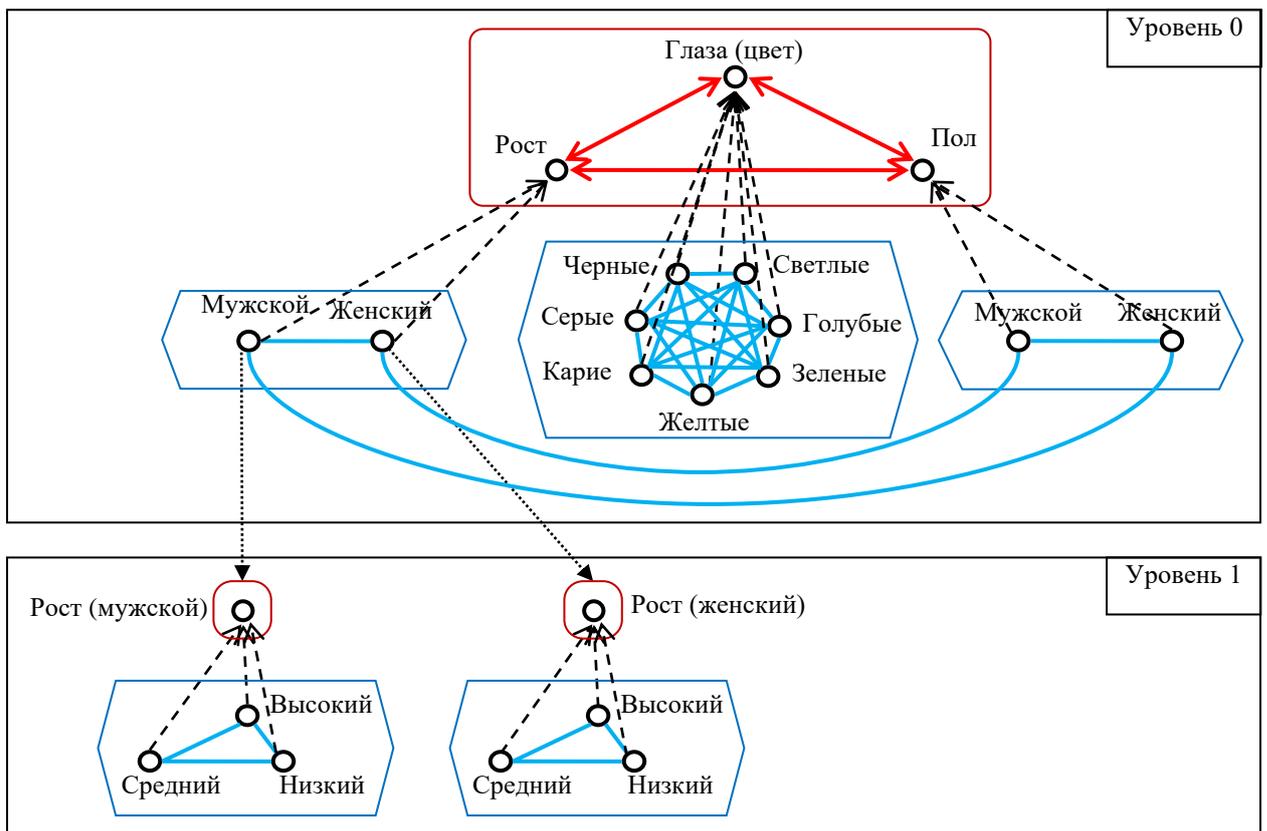


Рисунок 5.8 – Моделирование ограничений существования свойств в габитологии с нейтрализацией эффекта транзитивности несовместимости относительно обусловленности

5.4 Выводы по разделу 5

- 1) Продемонстрирована пригодность разработанных методов и программных средств для эффективного решения задачи структурного анализа и синтеза технических решений, а также актуальных задач в других областях.
- 2) Используемые в развиваемой методике модели и методы позволяют описывать априорные знания о предметной области и справляться с неполнотой и противоречивостью эмпирических данных, имеющими различный генезис. При этом необходим тщательный анализ априорных ограничений существования свойств у объектов исследуемой предметной области и адекватное представление таких ограничений, включая по необходимости иерархические конструкции в форме И/ИЛИ графов с ограниченной сочетаемостью ИЛИ-вершин.

ЗАКЛЮЧЕНИЕ

В результате исследований в диссертационной работе были разработаны модели, методы и программные средства, обеспечивающие подготовку формального контекста для вывода понятийной структуры предметной области на основе неполных и противоречивых эмпирических данных многомерных измерений и наблюдений и учёта априорных ограничений существования свойств у объектов. Все разработки внедрены в программную лабораторию онтологического анализа данных *OntoWorker*.

Основные научные результаты и выводы состоят в следующем:

- 1) Проанализированы существующие методики вывода формальных понятий, обоснована необходимость и возможность их комбинирования. Выявлено единство логической природы рассмотренных методик и выдвинута идея осуществления их гибридизации за счет отказа отчасти наименее достоверной эмпирической информации.
- 2) Разработаны прагматические модели системы измеряемых свойств и протокола измерений «объект-свойство», структурирующие соответственно гипотетические представления и эмпирическую информацию о предметной области, а также метод интерпретации и консолидации эмпирических данных с использованием многозначной векторной логики V^{TF} с целью построения нестрогого формального контекста.
- 3) Разработана методика построения рабочего однозначного формального контекста, включающая дефаззификацию нестрогого формального контекста и эвристический метод его нормализации, учитывающий ограничения существования свойств.
- 4) Разработано алгоритмическое обеспечение и программные средства подготовки контекста для вывода формальных понятий.
- 5) Разработанные методы и средства апробированы при решении различных научно-технических системологических задач и, в частности, задачи структурного анализа и синтеза технических решений.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

АПД	- агрегированный показатель достоверности
АФП	- анализ формальных понятий
БСС	- базовое семантическое суждение
ВЗО-группа	- группа взаимообусловленных свойств
ИОКл	- интенциональное отношение классов
ИФК	- исходный ФК
ЛРО	- лист расширенного описания
Н-группа	- группа попарно несовместимых свойств
НФК	- нестрогий формальный контекст
О-группа	- группа свойств, связанных отношением обусловленности
ОАД	- онтологический анализ данных
ОСС	- ограничения существования свойств
ОТОС	- обобщенная таблица «объекты-свойства»
ПрО	- предметная область.
РФК	- рабочий <u>ФК</u>
РФП	- решётка формальных понятий
СИС	- система измеряемых свойств
СИС/ГСС	- представление <u>СИС</u> как совокупности <u>ГСС</u> .
СПД	- субъективный порог доверия
ТОС	- таблица «объекты-свойства»
ТКл	- таксономия классов объектов
ФК	- формальный контекст
ФП	- формальное понятие
Failure	- лингвистическая константа, указывающая особый результат акта измерения свойства объекта, квалифицируемый как отказ от выполнения измерения (от англ. <i>failure</i> - отказ).
FP	- лингвистическая константа, указывающая оценку истинности <u>БСС</u> , которое является либо неопределённым, либо противоречивым, либо «неопределённым с долей лжи»,

либо «лживым с долей правды»
(от англ. *false proposition* – ложное суждение).

NM

- лингвистическая константа, указывающая особый результат акта измерения свойства объекта, означающий, что в действительности измерение не проводилось (от англ. *not measured* - не измерялось).

None

- лингвистическая константа, указывающая особый результат акта измерения свойства объекта, свидетельствующий о нахождении значения измеряемого свойства за порогом чувствительности, вне динамического диапазона средства (процедуры) измерения (от англ. *none* – нисколько, ничто).

OW

- OntoWorker.

TTV

- вектор ужесточения порога α , примененного для построения однозначного РФК

X

- лингвистическая константа, обозначающая любое натуральное значение.

СПИСОК ЛИТЕРАТУРЫ

- 1 Гаврилова, Т.А. Инженерия знаний. Модели и методы / Т.А. Гаврилова, Д.В. Кудрявцев, Д.И. Муромцев. - СПб.: Лань, 2016. - 324 с.
- 2 Смирнов, С.В. Онтологический анализ предметных областей моделирования / С.В. Смирнов // Известия Самарского научного центра РАН. - 2001. - Т. 3, № 1. - С. 62-70.
- 3 Боргест, Н.М. Стратегии интеллекта и его онтологии: попытка разобраться / Н.М. Боргест // Онтология проектирования. - 2019. - Т. 9, № 4(34). - С. 407-428.
- 4 Russel, S.J. Artificial Intelligence. A modern approach / S.J. Russel, P. Norvig. - New Jersey: Prentice Hall Press, 2003. - 1152 p.
- 5 Искусственный интеллект. В 3-х кн. Кн. 2. Модели и методы. Справочник / под ред. Д.А. Поспелова. - М.: Радио и связь, 1990. - 304 с.
- 6 Ganter, B. Formal Concept Analysis. Mathematical foundations / B. Ganter, R. Wille. - Berlin-Heidelberg: Springer-Verlag, 1999. - 290 p.
- 7 Ignatov, D.I. Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields // In: P. Braslavski, N. Karpov, M. Worring, Y. Volkovich, D.I. Ignatov (eds.): Information Retrieval. Revised Selected Papers 8th Russian Summer School, RuSSIR 2014 (August 18-22, 2014, Nizhniy Novgorod, Russia). - Springer Int. Publishing, 2015. - P. 42-141.
- 8 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 Int. Publishing, 2020 - P. 411-445.
- 9 Lammari, N. Building and maintaining ontologies: a set of algorithms / N. Lammari, E. Metais // Data & Knowledge Engineering. - 2004. - Vol. 48(2). - P. 155-176.

- 10 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.
- 11 Semenova, V.A. Extended methodology for deriving formal concepts / V.A. Semenova, S.V. Smirnov // J. of Physics: Conf. Series. - Int. Conf. «Marchuk Scientific Readings 2021» MSR-2021 (4-8 October 2021, Novosibirsk, Russia). - 2021. - Vol. 2099, 012026. - P. 1-9. DOI: 10.1088/1742-6596/2099/1/012026.
- 12 Samoylov, D.E. Multilevel recursive model of properties existence constraints in machine learning / D.E. Samoylov, V.A. Semenova, S.V. Smirnov // J. of Physics: Conf. Series 1096 (2018) 012096. DOI: 10.1088/1742-6596/1096/1/012096.
- 13 Самойлов, Д.Е. Фрактальность ограничений сосуществования свойств в задачах машинного обучения / Д.Е. Самойлов, В.А. Семенова, С.В. Смирнов // Сб. трудов IV междунар. конф. и молодеж. школы «Информационные технологии и нанотехнологии» ИТНТ-2018 (24-27 апреля 2018 г., Самара, Россия). - Самара: Новая техника, 2018. - С. 2512-2518.
- 14 Самойлов, Д.Е. Паттерны структурной организации системы измеряемых свойств в онтологическом анализе данных / Д.Е. Самойлов, В.А. Семенова, С.В. Смирнов // Проблемы управления и моделирования в сложных системах: труды XX междунар. конф. (3-6 сентября 2018 г., Самара, Россия). - Самара: ООО «Офорт», 2018. - С. 358-366.
- 15 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.
- 16 Семенова, В.А. Модели и методы интеллектуального анализа неполных данных для построения формальных онтологий / В.А. Семенова, С.В. Смирнов // Информационные технологии и нанотехнологии (ИТНТ-2015): материалы междунар. конф. и молодеж. школы (29 июня-1 июля 2015 г., Самара, Россия). - Самара: Изд-во СамНЦ РАН, 2015. - С. 192-196.

- 17 Мальцев, А.И. Алгебраические системы / А.И. Мальцев. - М.: Наука, 1970. - 392 с.
- 18 Смирнов, С.В. Онтологическое моделирование в ситуационном управлении / С.В. Смирнов // Онтология проектирования. - 2012. - №2(4). - С. 16-24.
- 19 Онтологии в системах искусственного интеллекта: способы построения и организации / А.В. Смирнов, М.П. Пашкин, Н.Г. Шилов и др. // Новости искусственного интеллекта. - 2002. - № 1. - С. 3-13 (Часть 1); № 2. - С. 3-9 (Часть 2).
- 20 Барсегян, А.А. Анализ данных и процессов / А.А. Барсегян, М.С. Куприянов, И.И. Холод, М.Д. Тесс, С.И. Елизаров. - 3-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2009. - 512 с.
- 21 Смирнов С.В. Построение онтологий предметных областей со структурными отношениями на основе анализа формальных понятий / С.В. Смирнов // Знания – Онтологии – Теории: материалы Всероссийской конф. с междунар. участием (3-5 октября 2011 г., Новосибирск, Россия). Т. 2. - Новосибирск: Институт математики СО РАН, 2011. - С. 103-112.
- 22 Загоруйко, Н.Г. Когнитивный анализ данных / Н.Г. Загоруйко. - Новосибирск: Академическое изд-во «Гео», 2013. - 186 с.
- 23 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.
- 24 Кузнецов, С.О. Алгоритмы построения множества всех понятий формального контекста и его диаграммы Хассе / С.О. Кузнецов, С.А. Обьедков // Известия академии наук. Теория и системы управления. - 2001. - №1. - С. 120-129.
- 25 Гетманова, А.Д. Логика. Углубленный курс / А.Д. Гетманова. - М.: КНОРУС, 2016 - 192 с.
- 26 Ивин, А.А. Словарь по логике / А.А. Ивин, А.Л. Никифоров. - М.: Гуманитарный изд. центр ВЛАДОС, 1997. - 384 с.

- 27 Бочаров, В.А. Введение в логику: учебник / В.А. Бочаров, В.И. Маркин. - М.: ИД «Форум»: Инфра-М, 2008. - 560 с.
- 28 Pollandt, S. Fuzzy-Begriffe: Formale Begriffsanalyse unscharfer Daten. - Berlin-Heidelberg: Springer-Verlag, 1997.
- 29 Belohlavek, R. Scaling, Granulation, and Fuzzy Attributes in Formal Concept Analysis / R. Belohlavek, J. Konecny // The IEEE Int. Conf. on Fuzzy Systems (London, UK, July 23-26, 2007). - P. 918-923.
- 30 Смирнов, С.В. Нечеткие формальные контексты при построении онтологий на основе анализа формальных понятий: происхождение и использование / С.В. Смирнов // Знания – Онтологии – Теории: труды всероссийской конф. с междунар. участием ЗОНТ-2007 (14-16 сентября 2007 г., Новосибирск, Россия). Т. 2. - Новосибирск: Институт математики СО РАН, 2007. - С. 17-25.
- 31 Офицеров, В.П. Метод альфа-сечения нестрогих формальных контекстов в анализе формальных понятий / В.П. Офицеров, В.С. Смирнов, С.В. Смирнов // Проблемы управления и моделирования в сложных системах: труды XVI междунар. конф. - Самара: СамНЦ РАН, 2014. - С. 228-244.
- 32 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.
- 33 Семенова, В.А. Выбор логической модели для представления эмпирической информации в онтологическом анализе данных / В.А. Семенова // Информатика и вычислительная техника. XII Всероссийская науч.-тех. конф. аспирантов, студентов и молодых ученых ИВТ-2020 (15-16 июня 2020г., Ульяновск, Россия): сб. научных трудов / под общей ред. В.Н. Негоды. - Ульяновск: УлГТУ, 2020. - С. 205-210.
- 34 Самойлов, Д.Е. Структура системы измеряемых свойств у объектов многомерного наблюдения и экспериментов / Д.Е. Самойлов, В.А. Семенова,

- С.В. Смирнов // Вестник Самарского государственного технического университета. Серия: Технические науки. - 2018. - №3(59). - С. 56-71.
- 35 Самойлов, Д.Е. Модель ограничений существования свойств в онтологическом моделировании / Д.Е. Самойлов, В.А. Семенова, С.В. Смирнов // Проблемы управления и моделирования в сложных системах (12-15 сентября 2020г., Самара, Россия): труды XIX междунар. конф. - Самара: ООО «Офорт», 2017. - С. 468-473.
- 36 Semenova, VA Intelligent analysis of incomplete data to building formal ontologies / VA Semenova, SV Smirnov // CEUR Workshop Proceedings, 2016; 1638: 796-805. DOI: 10.18287/1613-0073-2016-1638-796-805.
- 37 Игнатов, Д.И. Решетки формальных понятий для анализа данных социологических опросов / Д.И. Игнатов, О.Н. Кононыхина // Интегрированные модели и мягкие вычисления в искусственном интеллекте: труды V-й междунар. науч.-тех. конф. (20-30 мая 2009 г., Коломна, Россия). Т 1. - М.: Физматлит, 2009. - С. 230-240.
- 38 Самойлов, Д.Е. Субъективные аспекты формирования и обработки данных в анализе формальных понятий / Д.Е. Самойлов, С.В. Смирнов // Информационные технологии и нанотехнологии (ИТНТ-2016): материалы междунар. конф. и молодеж. школы (17-19 мая 2016 г., Самара, Россия). - Самара: Изд-во СамНЦ РАН, 2016. - С. 1038-1043.
- 39 Смирнов, С.В. Две методологии вывода формальных понятий: когда и как они должны работать вместе / С.В. Смирнов // Знания – Онтологии – Теории: материалы VII междунар. конф. - Новосибирск: Институт математики СО РАН, Новосибирский государственный ун-т, 2019. - С. 355-363.
- 40 Smirnov, S.V. Biconstituent phenomenon of information and cognitive data analysis / S.V. Smirnov // Procedia Engineering. - 2017. - Vol. 201. - P. 773-778.
- 41 Lukasiewicz, J. On three-valued logic / J. Lukasiewicz // Selected Works. Oxford, 1970. - P. 87-88.

- 42 Dunn, J.M. Intuitive semantics for first-degree entailment and ‘coupled trees’ / J.M. Dunn // *Philosophical Studies*. - 1976. - Vol. 29. - P. 149-168.
- 43 Заде, Л. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. Заде. - М.: Мир, 1976. - 166 с.
- 44 Аршинский, Л.В. Векторные логики: основания, концепции, модели / Л.В. Аршинский. - Иркутск: Иркутский государственный ун-т, 2007. - 228 с.
- 45 Аншаков, О.М. Многозначные логики и их применения. Том 1. Логические исчисления, алгебры и функциональные свойства/ О.М. Аншаков, Д.В. Виноградов, В.К. Финн. - Изд. стереотип. URSS, 2020. - 502 с.
- 46 Semenova, V Revealing attributes of an object based on incomplete and inconsistent empirical data / V Semenova, S Smirnov // 2022 VIII Int. Conf. on Information Technology & Nanotechnology (ITNT). IEEE Xplore, 2022. - P. 1-5. DOI: 10.1109/ITNT55410.2022.9848601.
- 47 Семенова, В.А. Выявление признаков объекта на основе неполных и противоречивых эмпирических данных / В.А. Семенова, С.В. Смирнов // Информационные технологии и нанотехнологии (ИТНТ-2022). Сб. трудов VIII Междунар. конф. и молодеж. школы. В 5-ти томах. Том 5. Под ред. А.В. Куприянова. - Самара: Самарский национальный исследовательский ун-т им. академика С.П. Королева, 2022. - С. 51122.
- 48 Семенова, В.А. Механизм нормализации эмпирического контекста в онтологическом анализе данных / В.А. Семенова, С.В. Смирнов // Системная инженерия и информационные технологии. - 2021. - Т. 3, № 3(7). - С. 45-52.
- 49 Семенова, В.А. Эвристика и численный метод нормализации эмпирического V^{TF} -контекста в задаче онтологического анализа данных / В.А. Семенова // Информационные и математические технологии в науке и управлении. - 2021. - № 2(22). - С. 61-69.
- 50 Samoylov, DE Defuzzification of the initial context in Formal Concept Analysis / DE Samoylov, VA Semenova, SV Smirnov // DS-ITNT 2019 Information Technology and Nanotechnology: Data Science: Proceedings of the Data Science Ses-

- sion at the V Int. Conf. on Information Technology and Nanotechnology (Samara, Russia, May 21-24, 2019). Eds.: V. Fursov, Y. Goshin, D. Kudryashov. - CEUR Workshop Proceedings, 2019. - Vol. 2416. - P. 1-9. DOI: 10.18287/1613-0073-2019-2416-1-9.
- 51 Самойлов, Д.Е. Эвристический алгоритм дефаззификации исходного контекста в анализе формальных понятий / Д.Е. Самойлов, В.А. Семенова, С.В. Смирнов // Сб. трудов V междунар. конф. и молодеж. школы «Информационные технологии и нанотехнологии» ИТНТ-2019 (21-24 мая 2019 г., Самара, Россия). Т. 4. «Науки о данных». - Самара: Новая техника, 2019. - С. 9-16.
- 52 Семенова, В.А. Построение формальных онтологий на основе противоречивых эмпирических данных / В.А. Семенова // Перспективные информационные технологии (ПИТ 2017): труды междунар. науч.-тех. конф. (14-16 марта 2017 г., Самара, Россия) / Под ред. С.А. Прохорова. - Самара: Изд-во Самарского научного центра РАН, 2017. - С. 164-168.
- 53 Семенова, В.А. Семантическая идентификация объекта в задачах концептуального моделирования / В.А. Семенова, С.В. Смирнов // Перспективные информационные технологии (ПИТ 2016): Труды междунар. науч.-тех. конф. / Под ред. С.А. Прохорова. - Самара: Изд-во Самарского научного центра РАН, 2016. - С. 330-333.
- 54 Семенова, В.А. Эффективность редукции ограничений существования свойств в задаче идентификации признаков объекта / В.А. Семенова, С.В. Смирнов // Информационные и математические технологии в науке и управлении. 2023. №2(30). С. 5-13. DOI: 10.25729/ESI.2023.30.2.001.
- 55 Belohlavek, R. Computing the lattice of all fixpoints of a fuzzy closure operator / R. Belohlavek, B. De Baets, B. Outrata, J. Vychodil // IEEE Trans. on Fuzzy systems. - 2010. - I. 3. - Vol. 18. - P. 546-557.
- 56 Belohlavek, R. What is a Fuzzy Concept Lattice? / R. Belohlavek, V. Vychodil // Proc. of the Int. Workshop on Concept Lattices and their Application CLA 2005

- (September 7-9, 2005 Olomouc, Czech Republic) CEUR Workshop Proceedings, 2005. - Vol. 162. - P. 34-45.
- 57 Belohlavek, R. What is a Fuzzy Concept Lattice? II / R. Belohlavek // In: S.O. Kuznetsov, D. Ślęzak, D.H. Hepting, B.G. Mirkin (eds.). Rough Sets, Fuzzy Sets, Data Mining and Granular Computing RSFDGrC 2011. LNAI 6743. - Berlin-Heidelberg: Springer, 2011. - P. 19-26.
- 58 Самойлов, А.Е. Ограничения существования свойств в нечетком анализе формальных понятий / А.Е. Самойлов, С.В. Смирнов // Сб. трудов VI междунар. конф. и молодеж. школы «Информационные технологии и нанотехнологии» ИТНТ-2020 (26-29 мая 2020 г., Самара, Россия). Т. 4 «Науки о данных» / Ред.: В.А. Фурсов, Д.В. Кудряшов. - Самара: Самарский национальный исследовательский ун-т им. акад. С.П. Королева, 2020. - С. 779-785.
- 59 Самойлов, А.Е. Универсальный метод учета ограничений существования свойств в задачах вывода формальных понятий из нечетких объектно-признаковых данных / А.Е. Самойлов, С.В. Смирнов // Сб. трудов VII междунар. конф. и молодеж. школы (20-24 сентября 2021 г., Самара, Россия) Т. 3. Ред.: Куприянов А.В., Кудряшов Д.В. - Самара: Самарский нац. исследовательский ун-т им. академика С.П. Королева, 2021. С. 33862.
- 60 Кофман, А. Введение в теорию нечетких множеств: Пер. с франц. - М.: Радио и связь, 1982. - 432 с.
- 61 Колмогоров, А.Н. Избранные труды: математика и механика / А.Н. Колмогоров. - М.: Наука, 1985. - 470 с.
- 62 Семенова, В.А. Функциональное наполнение и архитектура программной лаборатории для онтологического анализа данных / В.А. Семенова, С.В. Смирнов // Вестник Самарского государственного технического университета. Серия: Технические науки. - 2023. - Т. 31, №2. - С. 85-100. DOI: 10.14498/tech.2023.2.7.
- 63 Лещева, Д.В. Управление данными при кластеризации объектов многомерных наблюдений и экспериментов / Д.В. Лещева, В.А. Семенова // Перспек-

- тивные информационные технологии (ПИТ 2020) [Электр. ресурс]: труды междунар. науч.-тех. конф. / под ред. С.А. Прохорова. - Электр. текстовые и граф. дан. (28,4 Мбайт). - Самара: Изд-во Самарского научного центра РАН, 2020. - 523 с. - 1 эл. опт. диск (CD-ROM).
- 64 Лещева, Д.В. О поддержке когнитивных актов в задачах онтологического анализа данных / Д.В. Лещева, В.А. Семенова, С.В. Смирнов // Проблемы управления и моделирования в сложных системах: труды XXI междунар. конф. (3-6 сентября 2019 г., Самара, Россия). - Самара: ООО «Офорт», 2019. - Т.2. - С. 260-263.
- 65 Лещева, Д.В. Пользовательский интерфейс программной лаборатории для онтологического анализа данных / Д.В. Лещева, В.А. Семенова, С.В. Смирнов // Перспективные информационные технологии (ПИТ 2018): труды междунар. науч.-тех. конф. / под ред. С.А. Прохорова. - Самара: СамНЦ РАН, 2018. - С. 479-482.
- 66 Зубцов, Р.О. Алгоритмическое и программное обеспечение онтологического анализа данных / Р.О. Зубцов, В.А. Семенова, С.В. Смирнов // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2016): Материалы VI междунар. науч.-тех. конф. (18-20 февраля 2016 г. Минск, Беларусь) / Редкол.: В.В. Голенков (отв. ред.) [и др.]. - Минск: БГУИР, 2016. - С. 83-88.
- 67 Семенова, В.А. OntoWorker: программная лаборатория для онтологического анализа данных / В.А. Семенова, В.С. Смирнов, С.В. Смирнов // Проблемы управления и моделирования в сложных системах: труды XVII междунар. конф. (22-25 июня 2015 г., Самара, Россия). - Самара: СамНЦ РАН, 2015. - С. 382-393.
- 68 Уокенбах, Дж. Excel 2010: профессиональное программирование на VBA / Дж. Уокенбах. - М.: Изд-во «Вильямс», 2011. - 994 с.
- 69 Иванов, Д.Ю. Моделирование на UML / Д.Ю. Иванов, Ф.А. Новиков. - СПб.: Наука и техника, 2010. - 640 с.

- 70 Скобелев П.О. Онтология деятельности для ситуационного управления // Онтология проектирования. 2012. №1. С. 7-38.
- 71 Киммел, П. Excel 2003 и VBA. Справочник программиста / П. Киммел, Дж. Грин, С. Буллен и др: пер. с англ. - М.: Изд-во «Вильямс», 2006. - 1088 с.
- 72 Oficerov, V.P. Fuzzy Formal Concept Analysis in the construction of ontologies / V.P. Oficerov, S.V. Smirnov // Ontology of designing. - 2017. - Vol. 7. - No 4. - P. 487-495.
- 73 Cross, V. Creating Fuzzy Concepts: The One-Sided Threshold, Fuzzy Closure and Factor Analysis Methods / V. Cross, M. Kandasamy // Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. - 2011. - LNAI 6743. - P. 127-134.
- 74 Yang, K.M. Fuzzy Concept Mining based on Formal Concept Analysis / K.M. Yang, E.H. Kim, S.H. Hwang, S.H. Choi // Int. J. of Computers. - 2008. - Vol. 2, No. 3. - P. 279-290.
- 75 Samoilov, A.E. Ways to accept the properties existence constraints in fuzzy formal concept analysis / A.E. Samoilov, S.V. Smirnov // in: Fursov V., Goshin Y., Kudryashov D. CEUR Workshop Proc. of the 6th Int. Conf. on Information Technology and Nanotechnology. Session Data Science. ITNT 2020. - Vol. 2667. - P. 269-272.
- 76 Каталог инновационных разработок Самарского НИИ сельского хозяйства имени Н.М. Тулайкова на 2018 год / Под ред. С.Н. Шевченко - Самара: Изд-во СамНЦ РАН, 2018. - 92 с.
- 77 ГОСТ 9353-2016 Пшеница. Технические условия: [Электронный ресурс]. URL: <https://pdf.standartgost.ru/catalog/Data2/1/4293751/4293751950.pdf> (Дата обращения: 10.02.2023).
- 78 Семенова В.А., Смирнов С.В. Модели и методы онтологического анализа данных в задаче структурного анализа и синтеза технических решений // Онтология проектирования. 2023. Т. 13, №4(50). С. 531-547. DOI: 10.18287/2223-9537-2023-13-4-531-547.

- 79 Божко, А.Н. Структурный синтез на элементах с ограниченной сочетаемостью / А.Н. Божко, А.Ч. Толпаров // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. 2004. № 5. <http://technomag.edu.ru/doc/44191.html> (Дата обращения: 11.09.2023).
- 80 Zwicky, F. Discovery Invention, Research Through the Morphological Approach / F. Zwicky. - McMillan, 1969. - 276 p.
- 81 Одрин, В.М. Метод морфологического анализа технических систем / В.М. Одрин. - М.: ВНИИПИ, 1989. - 312 с.
- 82 Закревский, А.Д. Алгоритмы синтеза дискретных автоматов / А.Д. Закревский. - М.: Наука, 1971. - 512 с.
- 83 Половинкин, А.И. Основы инженерного творчества / А.И. Половинкин. - М.: Машиностроение, 1988. - 368 с.
- 84 Андрейчиков, А.В. Анализ, синтез, планирование решений в экономике / А.В. Андрейчиков, О.Н. Андрейчикова. - М.: «Финансы и статистика», 2000. - 368 с.
- 85 Андрейчиков, А.В. Наука и искусство принятия решений. Кн. 3: Подходы к решению задач коллективного выбора. Методы генерации решений / А.В. Андрейчиков, О.Н. Андрейчикова. - М.: ЛЕНАНД, 2021. - 248 с.
- 86 Дубов, Ю.А. Многокритериальные модели формирования и выбора вариантов систем / Ю.А. Дубов, С.И. Травкин, В.Н. Якимец. - М.: Наука, 1986. - 296 с.
- 87 Каазик, Ю.А. Математический словарь / Ю.А. Каазик. - М.: Физматлит, 2007. - 334 с.
- 88 Райншке, К. Оценка надежности систем с использованием графов / К. Райншке, И.А. Ушаков. - М.: Радио и связь, 1988. - 208 с.
- 89 Люгер, Дж.Ф. Искусственный интеллект. Стратегии и методы решения сложных проблем / Дж.Ф. Люгер. - М.: Вильямс, 2003. - 864 с.

- 90 Егер, С.М. Основы автоматизированного проектирования самолетов / С.М. Егер, Н.К. Лисейцев, О.С. Самойлович. - М.: Машиностроение, 1986. - 232 с.
- 91 Емеличев, В.А. Лекции по теории графов / В.А. Емеличев, О.И. Мельников, В.И. Сарванов, Р.И. Тышкевич. - М.: Наука, Физматлит, 1990. - 384 с.
- 92 Закревский, А.Д. Логика распознавания / А.Д. Закревский. - Мн.: Наука и техника, 1988. - 118 с.
- 93 Закревский, А.Д. Логические основы проектирования дискретных устройств / А.Д. Закревский, Ю.В. Поттосин, Л.Д. Черемисинова. - М.: Физматлит, 2007. - 592 с.
- 94 Клейнберг, Дж. Алгоритмы: разработка и применение. Классика Computers Science / Дж. Клейнберг, Е. Тардос. - СПб.: Питер, 2016. - 800 с.
- 95 Лотов, А.В. Многокритериальные задачи принятия решений / А.В. Лотов, И.И. Поспелова. - М.: МАКС Пресс, 2008. - 197 с.
- 96 Кузнецов, О.П. Анализ влияний при управлении слабоструктурированными ситуациями на основе когнитивных карт / О.П. Кузнецов, А.А. Кулинич, А.В. Марковский // Человеческий фактор в управлении / Под ред. Н.А. Абрамовой, К.С. Гинсберга, Д.А. Новикова. - М.: КомКнига, 2006. - С. 313-344.
- 97 Кулинич, А.А. Компьютерные системы моделирования когнитивных карт: подходы и методы / А.А. Кулинич // Проблемы управления. - 2010. - № 3. - С. 2-16.
- 98 Абрамова, Н.А. Некоторые критерии достоверности моделей на основе когнитивных карт / Н.А. Абрамова, С.В. Коврига // Проблемы управления. - 2008. - № 6. - С. 23-33.
- 99 Коростелёв, Д.А. Система поддержки принятия решений на основе нечетких когнитивных моделей «ИГЛА» / Д.А. Коростелёв, Д.Г. Лагереv, А.Г. Подвесовский // Одиннадцатая нац. конф. по искусственному интеллек-

ту с междунар. участием КИИ-2008 (28 сентября-03 октября 2008 г., Дубна, Россия): Труды конф., Т. 3. - М.: ЛЕНАНД, 2008. - С. 329-336.

- 100 Смирнов, С.В. Модели и методы формирования когнитивных карт при их коллективной разработке / С.В. Смирнов // Информационные технологии и системы: Труды Шестой междунар. науч. конф. ИТиС-2017 (1-5 марта 2017 г., Банное, Россия) / Отв. ред.: Ю.С. Попков, А.В. Мельников. - Челябинск: Изд-во Челяб. гос. ун-та, 2017. – С. 281-283.
- 101 Смирнов, С.В. Многозначная и нечеткая логики в онтологическом анализе данных / С.В. Смирнов // Информационные технологии и системы: Труды Третьей междунар. науч. конф. ИТиС-2014 (26 февраля - 2 марта 2014 г., Банное, Россия) / Отв. ред.: Ю.С. Попков, А.В. Мельников. - Челябинск: Изд-во Челяб. гос. ун-та, 2014. - С. 90-91.

ПРИЛОЖЕНИЕ А

Эксперимент в области портретной идентификации людей

В криминалистике существенную роль играет портретная идентификация людей, когда требуется составить портрет преступника на основании показаний одного или нескольких свидетелей.

Эксперимент, воспроизводящий подобную идентификацию на основе опроса группы людей, был проведён в Институте проблем управления сложными системами РАН (ИПУСС РАН) с помощью средств ОАД. Интерес к данной Про был обусловлен, в первую очередь, тем фактом, что исходные данные, используемые при описании внешности человека, характеризуются существенной неопределенностью и противоречивостью, поскольку, как правило, поступают от других людей (независимых экспертов, «свидетелей»). С другой стороны, учение о признаках внешности человека – *габитология* – является хорошо апробированным результатом концептуального шкалирования в своей области, и номинальные значения прошкалированных признаков (свойств) элементов внешности человека для удобства сбора информации предлагаются в форме ответов на вопросы в закрытой форме (рисунок А.1).

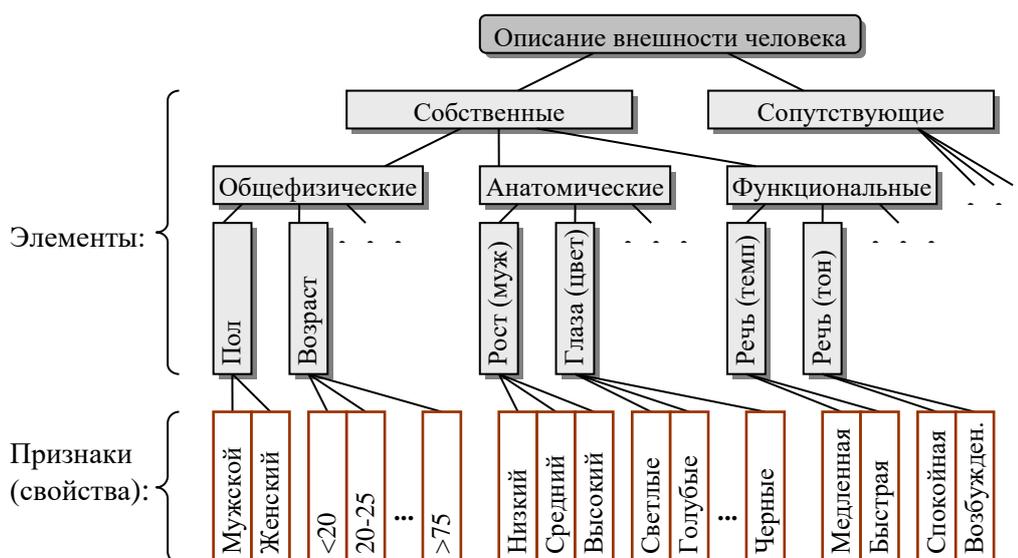


Рисунок А.1 – Фрагмент классификация элементов и характеризующих их признаков в портретной идентификации человека

Участниками эксперимента – *респондентами* – стали практически все сотрудники ИПУСС РАН, каждому из которых было предложено *быстро, особо не задумываясь* и используя заранее подготовленную *таблицу-анкету* охарактеризовать ряд внешних портретных признаков товарищей по работе.

Строки предложенной таблицы представляли сотрудников Института, а столбцы – запрашиваемые признаки (свойства) внешности сотрудников, сгруппированные по элементам портретной идентификации (были использованы не более 30% наиболее простых анатомических и функциональных признаков из габитологии). Каждая такая группа столбцов, по построению представляющая содержательно замкнутую совокупность несовместимых свойств, дополнялась двумя специальными столбцами: «Иное» и «Отказ».

- «Иное» следовало использовать в случаях, если, по мнению респондента, значение рассматриваемой характеристики у данного сотрудника отличается от приведенных вариантов ее значения⁷;
- «Отказ» предлагалось отметить, когда респондент по какой-либо причине не рисковал выбрать содержательный вариант ответа о некотором элементе внешности человека, или, просто, не желал освещать данный вопрос⁸.

Описание респондентом внешности сотрудника Института сводилось к постановке «галочек» в соответствующей строке ТОС-анкеты. Это описание признавалось корректным, если в группе несовместимых свойств (включая «Иное» и «Отказ») отмечалось одно и только одно свойство. Выполнение априори задаваемых несовместимостей «второго порядка», к которым в габитологии относится, например, требование отдельного описания роста мужчин и женщин – если описывается внешность мужчины, то не должно быть отмечено ни одно свойство, характеризующее рост женщин, и наоборот, – в эксперименте не контролировалось.

⁷ Тем самым формально допускалась неполнота концептуального шкалирования признаков внешности человека в габитологии. Эксперимент подтвердил необоснованность этого опасения, но в общем случае наличие такого варианта ответа обязательно.

⁸ Возможность выбора респондентом ответа «Отказ» соответствует одному из ключевых положений формирования ТОС: всякая измерительная процедура может дать специальный результат **Failure**, указывающий на *невыполнение* задачи измерений. Вместе с тем, в рассматриваемой ТОС-анкете это положение реализовано в упрощенном варианте, когда результат **Failure** связывается с группой сопряженных свойств, характеризующей отдельный элемент внешности человека. Этот вариант лучше воспринимается респондентами и принципиальных искажений в формировании ТОС и ФК не вносит.

К каждому сотруднику-респонденту была обращена просьба охарактеризовать не менее пяти сотрудников по его выбору. Данные собирались анонимно, но анкеты респондентов-женщин, обычно более внимательных к внешности окружающих, отмечались как заслуживающие достаточно высокой степени доверия, уровень доверия к ответам мужчин был ниже, а доверие к экспертизе сотрудников-совместителей оценивалось еще меньше.

После сведения собранных результатов опросов в единую таблицу была получена ОТОС, которая в свою очередь была обработана согласно методам, описанным в параграфе 2.2.2, и трансформирована в НФК, фрагмент которого представлен в таблице А.1.

Таблица А.1 – Фрагмент нестроого формального контекста предметной области «Внешность сотрудников Института»

Сотрудник	Рост (мужской)			Рост (женский)			Глаза (цвет)						
	Низкий	Средний	Высокий	Низкий	Средний	Высокий	Светлые	Голубые	Серые	Зеленые	Желтые	Карие	Черные
Виттих В.А.	.31, .36	.36, .31	0, .68	0, .68	0, .68	0, .68	.14, .53	.03, .64	.36, .32	.14, .53	.03, .64	.09, .58	.03, .64
Смирнов С.В.	0, .68	.12, .57	.57, .12	0, .68	0, .68	0, .68	.04, .64	.16, .52	.04, .64	.04, .64	.04, .64	.52, .16	.04, .64
Кузнецов С.В.	0, .72	.60, .12	.12, .60	0, .72	0, .72	0, .72	0, .72	0, .72	0, .72	.32, .40	0, .72	.40, .32	0, .72
Скобелев О.П.	.05, .56	.49, .12	.12, .49	0, .61	.07, .54	0, .61	.08, .53	.08, .53	.18, .43	.08, .53	.08, .53	.36, .26	.08, .53
Боровик С.Ю.	.08, .64	.31, .40	.40, .31	0, .71	0, .71	0, .71	.08, .64	.21, .50	.50, .21	.08, .64	.08, .64	.08, .64	.08, .64
Андрианов М.П.	.04, .72	.04, .72	.72, .04	0, .77	0, .77	0, .77	.19, .58	.16, .61	.46, .31	.04, .72	.04, .72	.04, .72	.04, .72
Кутейникова М.М.	0, .70	0, .70	0, .70	.06, .64	.29, .41	.41, .29	.14, .56	.18, .52	.44, .26	.06, .64	.06, .64	.06, .64	.06, .64
Морозова Л.Ф.	0, .71	0, .71	0, .71	0, .71	.54, .17	.17, .54	0, .71	0, .71	.07, .64	0, .71	0, .71	.54, .17	.10, .61
Сергеева Н.А.	0, .74	0, .74	0, .74	0, .74	.74, 0	0, .74	.28, .46	0, .74	.46, .28	0, .74	0, .74	0, .74	0, .74

Для получения стандартного ФК требовалось применить к НФК альфа-сечение, однако экспериментирование с величиной СПД к данным, содержащимся в нестрогом ФК с несовместимыми свойствами, быстро убеждает, что получаемые сечения зачастую оказываются *семантически некорректными*, т.к. в них об-

наруживаются нарушения необходимых и достаточных условий соблюдения нормальности выводимых подмножеств. Например, в альфа-сечении НФК при СПД к исходным данным $\alpha = (0.3, 0.4)$ объекты «Виттих В.А.» и «Боровик С.Ю.» получают противоречивое описание роста (см. таблицу А.2). В (0.4, 0.3)-сечении НФК (таблица А.3) первый из указанных объектов оказывается уже полностью неопределенным, второй – неопределенным в части роста, а описания объектов «Кузнецов С.В.» «Скобелев О.П.» и «Андрианов М.П.» становятся неопределенными в части цвета глаз.

Таблица А.2 – Альфа-сечение нестрогого формального контекста, заданного таблицей А.1 при субъективном пороге доверия к исходным данным $\alpha = (0.3, 0.4)$

Сотрудник	Рост (мужской)			Рост (женский)			Глаза (цвет)						
	Низкий	Средний	Высокий	Низкий	Средний	Высокий	Светлые	Голубые	Серые	Зеленые	Желтые	Карие	Черные
Виттих В.А.	×	×							×				
Смирнов С.В.			×									×	
Кузнецов С.В.		×										×	
Скобелев О.П.		×										×	
Боровик С.Ю.		×	×						×				
Андрианов М.П.			×						×				
Кутейникова М.М.						×			×				
Морозова Л.Ф.					×							×	
Сергеева Н.А.					×				×				

Таким образом, поставленный эксперимент доказал, что произвольное альфа-сечение НФК со свойствами, полученными в результате концептуального шкалирования, представляет собой в общем случае *неопределенный* и/или *противоречивый* ФК ПрО, и учёт априорных ОСС имеет смысл.

Таблица А.3 – Альфа-сечение нестроого формального контекста, заданного таблицей А.1 при пороге доверия к исходным данным $\alpha = (0.4, 0.3)$

Сотрудник	Рост (мужской)			Рост (женский)			Глаза (цвет)						
	Низкий	Средний	Высокий	Низкий	Средний	Высокий	Светлые	Голубые	Серые	Зеленые	Желтые	Карие	Черные
Виттих В.А.													
Смирнов С.В.			×									×	
Кузнецов С.В.		×											
Скобелев О.П.		×											
Боровик С.Ю.									×				
Анрианов М.П.			×										
Кутейникова М.М.						×			×				
Морозова Л.Ф.					×							×	
Сергеева Н.А.					×				×				

ПРИЛОЖЕНИЕ Б

Использование неравенства Хёфдинга для определения количества статистических испытаний, обеспечивающих качественную оценку среднего исследуемого показателя

Искомым результатом эксперимента с имитационной статистической моделью чаще всего является *среднее* $E(X)$ вероятностного распределения исследуемого показателя X . Показатель X трактуется как случайная величина, и каждое статистическое испытание модели (т.е. прогон программной реализации имитационной статистической модели) доставляет его конкретное значение. В качестве оценки среднего (математического ожидания) $E(X)$ обычно используют *среднее арифметическое* $E^*(X)$ значений показателя x_1, x_2, \dots , получаемых в результате многократных статистических испытаний.

В литературе по теории и практике статистического моделирования определение необходимого в таком случае *количества* статистических испытаний N модели обычно основывается на анализе некоторых предельных неравенств теории вероятностей, в которых устанавливается взаимосвязь N , требуемой *абсолютной точности* ε оценки $E^*(X)$ и *достоверности* результата испытаний $1 - \delta$, $\delta \in (0, 1)$. При этом, как правило, предполагается, что прогоны модели независимы друг от друга, т.е. фиксируемые значения x_1, x_2, \dots, x_N показателя X являются реализациями *независимых* случайных величин X_1, X_2, \dots, X_N , распределения которых считаются *одинаковыми* и *равными* распределению случайной величины исследуемого показателя X , имеющего по предположению ограниченные среднее и дисперсию: $E(X) < \infty$, $D^2(X) < \infty$.

Если можно указать границы области существования случайной величины X , то для определения количества статистических испытаний N целесообразно воспользоваться известным неравенством Хёфдинга^{9,10}, которое даёт экспоненци-

⁹ Hoeffding W. Probability inequalities for sums of bounded random variables // J. of the American Statistical Association. 1963. Т. 58, №301. С. 13–30.

¹⁰ Петров В.В. Предельные теоремы для сумм независимых случайных величин. Изд. 2. – М.: URSS, 2022. 320 с.

ально убывающие оценки вероятности отклонений сумм независимых случайных величин от среднего этих сумм. Эти оценки точнее, чем аналогичные, полученные с использованием первых или вторых моментов (например, неравенства Маркова, Чебышёва), обеспечивающие лишь степенной закон убывания указанных отклонений.

Хёфдинг рассматривает независимые случайные величины X_1, X_2, \dots, X_N , такие, что $a_i \leq X_i \leq b_i$ и сумму этих случайных величин

$$S_N = X_1 + X_2 + \dots + X_N.$$

Доказанная Хёфдингом теорема гласит, что для всех $\varepsilon > 0$

$$Pr\{|S_N - E(S_N)| \geq N\varepsilon\} \leq 2\exp\left(-\frac{2N^2\varepsilon^2}{\sum_{i=1}^N (b_i - a_i)^2}\right),$$

где $E(S_N)$ – математическое ожидание величины S_N .

Использование этого результата для характеристики отклонения оценки $E^*(X)$ неизвестного среднего значения исследуемого показателя $E(X)$ основывается, во-первых, на равенстве следующих вероятностей:

$$Pr\{|S_N - E(S_N)| \geq N\varepsilon\} = Pr\left\{\left|\frac{S_N - E(S_N)}{N}\right| \geq \varepsilon\right\} = Pr\left\{\left|\frac{S_N}{N} - \frac{E(S_N)}{N}\right| \geq \varepsilon\right\},$$

где согласно сделанным предположениям об условиях оценки среднего исследуемого показателя в статистическом эксперименте имеем $\frac{S_N}{N} = E^*(X)$,

$$\frac{E(S_N)}{N} = \frac{E(X_1) + E(X_2) + \dots + E(X_N)}{N} = \frac{N \cdot E(X)}{N} = E(X).$$

Во-вторых, требуя чтобы вероятность абсолютной ошибки равной или большей ε должна не превышать величины $\delta \in (0, 1)$, получаем

$$Pr\{|E^*(X) - E(X)| \geq \varepsilon\} \leq 2\exp\left(-\frac{2N^2\varepsilon^2}{\sum_{i=1}^N (b_i - a_i)^2}\right) \leq \delta.$$

Если априори известно, что $0 \leq X \leq 1$ (т.е. $a_i = 0, b_i = 1$), то последнее неравенство принимает вид

$$2\exp\left(-\frac{2N^2\varepsilon^2}{N}\right) \leq \delta.$$

Тогда количество статистических испытаний N , которое с вероятностью $1 - \delta$ гарантирует, что абсолютная ошибка (погрешность) оценки $E^*(X)$ среднего $E(X)$ не превысит величину ε , определяется соотношением

$$N \geq \frac{1}{2\varepsilon^2} \ln \frac{2}{\delta},$$

где $\varepsilon, \delta \in (0, 1)$.

В таблице Б.1 приведены варианты соответствующего расчета, демонстрирующие чувствительность необходимого количества статистических испытаний к изменению параметров достоверности и погрешности оценок средней величины исследуемого показателя.

Таблица Б.1 – Количество статистических испытаний для оценки среднего исследуемого показателя

Достоверность оценки $1 - \delta$	Абсолютная погрешность оценки ε	Нижняя граница минимально необходимого количества испытаний	Рекомендуемое количество испытаний для достижения надлежащей точности и достоверности оценки среднего исследуемого показателя
0,90	0,05	$200 \cdot \text{LN}20 \approx 599,1$	600
0,90	0,01	$5000 \cdot \text{LN}20 \approx 14978,7$	15000
0,95	0,01	$5000 \cdot \text{LN}40 \approx 18444,4$	18450

ПРИЛОЖЕНИЕ В

**Результаты имитационных статистических экспериментов
для определения среднего относительного уменьшения
количества сущностей в редуцированном представлении
ограничений существования свойств**

Таблица В.1 - Значения показателя \mathcal{E}_c - величины сокращения количества сущностей в обобщенном представлении системы $|M| = 20$ измеряемых свойств в зависимости от количества (и доли $q_{\text{ВЗО-П}}$) ВЗО-пар в исходном представлении этой системы (начало таблицы)

Кол. Взо-пар	$Q_{\text{ВЗО-П}}$, %	\mathcal{E}_c , %	Кол. Взо-пар*	$Q_{\text{ВЗО-П}}$, %	\mathcal{E}_c , %	Кол. Взо-пар*	$Q_{\text{ВЗО-П}}$, %	\mathcal{E}_c , %
0	0,0	0,0	44	23.2	66.1	88	46.3	80
1	0.5	5	45	23.7	55.1	89	46.8	85
2	1.1	10	46	24.2	55	90	47.4	90
3	1.6	13.6	47	24.7	57.8	91	47.9	65.2
4	2.1	17	48	25.3	58.7	92	48.4	70
5	2.6	20.7	49	25.8	62.5	93	48.9	76.2
6	3.2	23	50	26.3	67.5	94	49.5	75.2
7	3.7	24.8	51	26.8	63.2	95	50	80
8	4.2	28	52	27.4	68	#96		
9	4.7	30.7	53	27.9	72.3	97	51.1	80.2
10	5.3	31.4	54	28.4	71.5	98	51.6	85
11	5.8	32.7	55	28.9	56.6	99	52.1	90
12	6.3	34.6	56	29.5	58.1	#100		
13	6.8	36.4	57	30	63.3	101	53.2	85
14	7.4	39.2	58	30.5	63.5	#102		
15	7.9	39.9	59	31.1	65.4	#103		
16	8.4	39	60	31.6	70.4	#104		
17	8.9	40.4	61	32.1	68	105	55.3	70
18	9.5	42.6	62	32.6	70.5	106	55.8	75
19	10	44.4	63	33.2	76.9	107	56.3	80
20	10.5	47.2	64	33.7	75	108	56.8	80
21	11.1	45	65	34.2	71.2	109	57.4	85
22	11.6	44.4	66	34.7	58.9	#110		
23	12.1	46.7	67	35.3	60.5	111	58.4	85
24	12.6	47.9	68	35.8	65.3	#112		
25	13.2	49	69	36.3	65.8	#113		
26	13.7	51.8	70	36.8	71	#114		
27	14.2	52.5	71	37.4	76.7	115	60.5	90
28	14.7	50.4	72	37.9	71.1	#116		
29	15.3	48.7	73	38.4	76.5	#117		
30	15.8	49.4	74	38.9	82.7	#118		
31	16.3	52.3	75	39.5	80	#119		
32	16.8	53.5	76	40	76.5	120	63.2	75
33	17.4	57.8	77	40.5	80.7	121	63.7	80
34	17.9	57.1	78	41.1	60	122	64.2	85
35	18.4	58.4	79	41.6	65.1	123	64.7	85
36	18.9	53.5	80	42.1	70	#124		
37	19.5	52.2	81	42.6	72.8	#125		
38	20	54.6	82	43.2	75.5	126	66.3	90
39	20.5	55.6	83	43.7	83.5	#127		
40	21.1	58	84	44.2	75.4	#128		
41	21.6	62.4	85	44.7	80	#129		
42	22.1	60.1	#86			#130		
43	22.6	61.7	87	45.8	85	#131		

* СИС с количеством ВЗО-пар, отмеченных знаком «#», не существует.

Таблица В.2 - Значения показателя \mathcal{E}_c - величины сокращения количества сущностей в обобщенном представлении системы $|M| = 20$ измеряемых свойств в зависимости от количества (и доли $q_{\text{ВЗО-п}}$) ВЗО-пар в исходном представлении этой системы (окончание таблицы)

Кол. Взо-пар*	$Q_{\text{ВЗО-п}}, \%$	$\mathcal{E}_c, \%$	Кол. Взо-пар*	$Q_{\text{ВЗО-п}}, \%$	$\mathcal{E}_c, \%$	Кол. Взо-пар*	$Q_{\text{ВЗО-п}}, \%$	$\mathcal{E}_c, \%$
#132			#152			172	90	90
#133			153	80.5	85	#173		
#134			154	81.1	90	#174		
#135			#155			#175		
136	71.6	80	#156			#176		
137	72.1	85	#157			#177		
#138			#158			#178		
139	73.2	90	#159			#179		
#140			#160			#180		
#141			#161			#181		
#142			#162			#182		
#143			#163			#183		
#144			#164			#184		
#145			#165			#185		
#146			#166			#186		
#147			#167			#187		
#148			#168			#188		
#149			#169			#189		
#150			#170			#190		
#151			#171			191	100	95

* СИС с количеством ВЗО-пар, отмеченных знаком «#», не существует.

ПРИЛОЖЕНИЕ Г

**Свидетельство
о государственной регистрации программы для ЭВМ
«Программная лаборатория для онтологического анализа данных
OntoWorker»**

РОССИЙСКАЯ ФЕДЕРАЦИЯ

**RU2022669423**

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ

Номер регистрации (свидетельства): 2022669423 Дата регистрации: 19.10.2022 Номер и дата поступления заявки: 2022666316 06.09.2022 Дата публикации и номер бюллетеня: 19.10.2022 Бюл. № 10 Контактные реквизиты: нет	Автор(ы): Семенова Валентина Андреевна (RU), Смирнов Сергей Викторович (RU) Правообладатель(и): Федеральное государственное бюджетное учреждение науки Самарский федеральный исследовательский центр Российской академии наук (СамНИЦ РАН) (RU)
---	--

Название программы для ЭВМ:
«Программная лаборатория для онтологического анализа данных OntoWorker»

Реферат:

Назначение программы - обработка возможно неполных и противоречивых объектно-признаковых данных с целью вывода формальной онтологии исследуемой предметной области. В интерактивном режиме производятся: ввод/импорт и редактирование (включая концептуальное шкалирование) объектно-признаковых данных и априорных ограничений сосуществования признаков объектов; интерпретация и консолидация объектно-признаковых данных с получением нестрогого соответствия «объекты-свойства»; прагматическая реструктуризация ограничений сосуществования признаков; трансформация нестрогого соответствия «объекты-свойства» в бинарное с возможностью учета порога доверия к данным и/или ограничений сосуществования признаков; извлечение формальных понятий с отношением обобщения между ними; вывод формальной онтологии. Тип ЭВМ: персональный компьютер. ОС: MS Windows 7.

Язык программирования: VBA
Объем программы для ЭВМ: 651 КБ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2022669423

**«Программная лаборатория для онтологического
анализа данных OntoWorker»**

Правообладатель: *Федеральное государственное бюджетное
учреждение науки Самарский федеральный
исследовательский центр Российской академии наук
(СамНЦ РАН) (RU)*

Авторы: *Семенова Валентина Андреевна (RU), Смирнов
Сергей Викторович (RU)*

Заявка № 2022666316

Дата поступления 06 сентября 2022 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 19 октября 2022 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов

ПРИЛОЖЕНИЕ Д

**Контекстные меню рабочих листов
программной лаборатории
для онтологического анализа данных**

Таблица Д.1 – Контекстное меню зон 0, 5 и 6 листа СИС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Предъявить лист ОТОС	Делает активным лист ОТОС.	1
 Предъявить лист Верстак	Делает активным лист Верстак.	2
 Предъявить ЛРО	Делает активным лист ЛРО.	3
 Предъявить OW-справку	Делает активным лист OW-справка.	4
 Создать	Создает новую ОАД-задачу, которая приобретает статус актуальной.	5
 Открыть...	Иницирует Excel-диалог «Открыть книгу» для загрузки Excel-книги, предположительно содержащей данные ОАД-задачи (диалог для загрузки OWD-книги). Если имя выбранной пользователем Excel-книги совпадает с именем актуальной OWD-книги, то отменяет загрузку, информируя об указанном совпадении. Отменяет загрузку, когда выбранная Excel-книга не является OWD-книгой. Иначе открывает выбранную пользователем OWD-книгу, присваивая ей статус актуальной, и загружает содержащиеся в ней данные новой актуальной ОАД-задачи.	6
Закр <u>ы</u> ть	Удаляет актуальную ОАД-задачу, очищая все OW-листы за исключением листа OW-справка. Если данные удаляемой ОАД-задачи изменялись, то по согласованию с пользователем сохраняет изменённую ОАД-задачу в актуальной (а при отсутствии таковой - в новой) OWD-книге, которую затем закрывает.	7
 Сохранить	Сохраняет данные актуальной ОАД-задачи в её OWD-книге.	8
Сохранить к <u>а</u> к...	Иницирует Excel-диалог «Сохранить как...» для сохранения актуальной ОАД-задачи в OWD-книге.	9
 Генерировать задачу...	Иницирует диалог, в котором согласовываются и корректируются случайно выбранные в предопределенных диапазонах параметры СИС и ОТОС, и создает с согласованными значениями этих параметров новую актуальную ОАД-задачу.	10
 Сведения о задаче...	Предъявляет краткую справку об актуальной ОАД-задаче.	11
Масш <u>т</u> аб...	Иницирует Excel-диалог «Масштаб».	12
 Печ <u>а</u> ть...	Иницирует Excel-диалог «Печать».	13

Таблица Д.2 – Контекстное меню зоны 1 листа СИС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Редактировать имя сопряженности...	Иницирует диалог для редактирования имени выбранной СПС.	1
 Удалить сопряженность	Удаляет выбранную СПС.	2
 Ширина столбца...	Иницирует диалог для изменения ширины столбца, размещающего выбранную часть дескриптора СПС.	3
 Переместить строку...	Иницирует диалог для перемещения строки, описывающей выбранную СПС и символы её инцидентности со свойствами.	4

Таблица Д.3 – Контекстное меню зоны 2 листа СИС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Редактировать имя свойства...	Иницирует диалог для редактирования имени выбранного свойства.	1
 Удалить свойство	Удаляет выбранное свойство и все СПС, в которые оно включено.	2
 Шкалировать свойство...	Иницирует диалог для концептуального шкалирования выбранного свойства.	3
 Заместить ВЗО-свойствами...	Иницирует диалог для замещения выбранного свойства ВЗО-группой свойств.	4
 Ширина столбца...	Иницирует диалог для изменения ширины столбца, размещающего выбранное свойство и символы его инцидентности с СПС.	5
 Переместить столбец...	Иницирует диалог для перемещения столбца, размещающего выбранное свойство и символы его инцидентности с СПС.	6

Таблица Д.4 – Контекстное меню зоны 3 листа СИС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Добавить сопряженность пары свойств...	Иницирует диалог для добавления СПС.	1

Таблица Д.5 – Контекстное меню зоны 4 СИС-листа

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Добавить свойства...	Иницирует диалог для добавления свойств.	1

Таблица Д.6 – Контекстное меню зон 0 и 6 листа ОТОС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Предъявить <u>С</u> ИС	Делает активным лист СИС.	1
 Предъявить <u>В</u> ерстак	Делает активным лист Верстак.	2
 Предъявить <u>Л</u> РО	Делает активным лист ЛРО.	3
 Предъявить OW-справку	Делает активным лист OW-справка.	4
 Сохранить	Сохраняет данные актуальной ОАД-задачи в её OWD-книге.	5
 Сведения о задаче...	Предъявляет краткую справку об актуальной ОАД-задаче.	6
Масштаб...	Иницирует Excel-диалог «Масштаб».	7
 Печать...	Иницирует Excel-диалог «Печать».	8

Таблица Д.7 – Контекстное меню зоны 1 листа ОТОС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Редактировать дескриптор серии измерений...	Иницирует диалог для редактирования дескриптора выбранной серии измерений.	1
 Удалить серию...	Иницирует диалог для удаления выбранной серии измерений или всех серий измерений объекта, указанного в дескрипторе выбранной серии измерений.	2
 Ширина столбца...	Иницирует диалог для изменения ширины столбца, размещающего выбранную часть дескриптора серии измерений.	3
 Переместить строку...	Иницирует диалог для перемещения строки, описывающей выбранную серию измерений и их результаты.	4

Таблица Д.8 – Контекстное меню зоны 2 листа ОТОС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Редактировать процедуру измерения...	Иницирует диалог для редактирования дескриптора выбранной процедуры измерения.	1
 Удалить процедуру...	Иницирует диалог для удаления всех процедур измерения свойства, указанного в дескрипторе выбранной процедуры, кроме выбранной или удаления лишь выбранной процедуры измерения.	2
 Шкалировать свойство...	Иницирует диалог для концептуального шкалирования свойства, указанного в дескрипторе выбранной процедуры измерения.	3
 Ширина столбца...	Иницирует диалог для изменения ширины столбца, размещающего дескриптор выбранной процедуры измерения и полученные с её помощью результаты измерений.	4
 Переместить столбец...	Иницирует диалог для перемещения столбца, размещающего дескриптор выбранной процедуры измерения и полученные с её помощью результаты измерений.	5

Таблица Д.9 – Контекстное меню зоны 3 листа ОТОС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Добавить серии измерений...	Иницирует диалог для добавления серий измерений.	1

Таблица Д.10 – Контекстное меню зоны 4 листа ОТОС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Добавить процедуры измерения...	Иницирует диалог для добавления процедур измерения.	1

Таблица Д.11 – Контекстное меню зоны 5 листа ОТОС

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Редактировать результат измерения...	Иницирует диалог для редактирования результата измерения (в качестве вариантов значения результата измерения предлагает использовать лингвистическую константу из { X , None , Failure , NM }, произвольное «натурзначение», любой вектор истинности $\langle b^+, b^- \rangle$ VTF-логики или импортировать подобное значение из какой-либо открытой Excel-книги).	1

Таблица Д.12 – Контекстное меню листа Верстак (начало таблицы)

Пункт меню*	Действие OntoWorker при выборе пункта меню	№
 Предъявить <u>С</u> ИС	Делает активным лист СИС.	1
 Предъявить О <u>Т</u> ОС	Делает активным лист ОТОС.	2
 Предъявить Л <u>П</u> О	Делает активным лист ЛПО.	3
 Предъявить OW-справку	Делает активным лист OW-справка.	4
 Очистить верстак	Очищает лист Верстак.	5
 Копировать исходные СИС и ОТОС	Копирует СИС и ОТОС на лист Верстак.	6
<input checked="" type="checkbox"/> Сопроводить действия показом результата	Инвертирует режим показа результатов выполнения работ при решении актуальной ОАД-задачи.	7
 Построить СИС/ГСС и ИФК	Преобразует СИС, заменяя описывающую её матрицу инцидентности «СПС-свойства» матрицей инцидентности «ГСС-свойства». Строит ИФК, интерпретируя и консолидируя содержащиеся в ОТОС данные.	8
 Показать СИС/ГСС и ИФК	Печатает СИС/ГСС и ИФК на листе Верстак.	9
 Сформировать Р <u>Ф</u> К	Производит исходное отождествление РФК с ИФК.	10
 Показать Р <u>Ф</u> К...	Иницирует диалог, в котором устанавливается объём сведений о РФК, подлежащих выводу, и печатает отобранные сведения на листе Верстак.	11
 Установить порог доверия и обновить Р <u>Ф</u> К...	Иницирует диалог, в котором устанавливается СПД к БСС, и обновляет РФК, формируя его путем преобразования ИФК в РФК с отнесением к ложным всех БСС, оценка истинности которых не достигает СПД, а остальных – к истинным.	12
 Нормализовать Р <u>Ф</u> К	Преобразует РФК, снимая противоречие между эмпирическими данными и априорными гипотезами об искомым ФП, индуцирующими ОСС. А именно: нормализует множества свойств каждого объекта измерения в РФК за счет изъятия части присущих ему свойств, которая при прочих равных условиях имеет среди других таких частей, наименьшие АПД и количество денонсируемых БСС (т.е. БСС, считавшихся истинными, но в результате нормализации относимых к ложным).	13
 Извлечь Ф <u>П</u>	Извлекает из РФК все ФП.	14
 Показать извлечённые Ф <u>П</u>	Печатает все ФП на листе Верстак.	15
 Построить решётку Ф <u>П</u>	Строит решётку формальных понятий (РФП).	16

* Префикс пункта меню в виде знака «→» указывает, что данный пункт замещает в меню предшествующий пункт по завершению последним своих действий.

Таблица Д.12 – Контекстное меню листа Верстак (окончание таблицы)

Пункт меню*	Действие OntoWorker при выборе пункта меню	№
→ Показать <u>Р</u> ФП...	Иницирует диалог, в котором выбираются подлежащие отображению сведения о РФП: экстенциональная сеть ФП и/или транзитивная редукция РФП, - и печатает эти сведения на листе Верстак. При этом обеспечивает возможность учета поддержки и/или устойчивости ФП	17
 Извлечь "конструктивную" он-тологию	Извлекает "конструктивную" формальную онтологию ПрО, исходные данные о которой предоставляет РФК.	18
→ Показать "конструктивную" онтологию...	Иницирует диалог, в котором выбираются подлежащие отображению сведения об онтологии: ТКл - экстенциональная сеть ФП, именуемых классам, и/или транзитивная редукция ТКл, а также идентифицирует ИОКл – и печатает эти сведения на листе Верстак. При этом обеспечивает возможность учета поддержки и/или устойчивости ФП (если эти индексы вычислены).	19
Вычислить под <u>д</u> ержку ФП	Вычисляет величины индекса поддержки всех ФП.	20
→ Показать под <u>д</u> ержку ФП...	Иницирует диалог, в котором устанавливается порог поддержки ФП, и печатает на листе Верстак все ФП, поддержка которых имеет значение не ниже установленного для неё порога.	21
Вычислить уст <u>о</u> йчивость ФП	Иницирует диалог для выбора варианта вычисления величин индекса интенциональной устойчивости либо всех ФП, либо ФП, обладающих собственным объёмом.	22
→ Показать уст <u>о</u> йчивость ФП...	Иницирует диалог, в котором устанавливается порог устойчивости ФП, и печатает на листе Верстак ФП, устойчивость которых вычислена и имеет значение не ниже установленного для неё порога.	23
 Сохр <u>ан</u> ить	Сохраняет данные актуальной ОАД-задачи в её OWD-книге.	26
 Сведения о зада <u>че</u> ...	Предъявляет краткую справку об актуальной ОАД-задаче.	27
Мас <u>ш</u> таб...	Иницирует Excel-диалог «Масштаб».	28
 Печ <u>ат</u> ать...	Иницирует Excel-диалог «Печать».	29

* Префикс пункта меню в виде знака «→» указывает, что данный пункт замещает в меню предшествующий пункт по завершению последним своих действий.

Таблица Д.13 – Контекстное меню листа ЛРО

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Предъявить <u>С</u> ИС	Делает активным лист СИС.	1
 Предъявить <u>О</u> ТОС	Делает активным лист ОТОС.	2
 Предъявить <u>В</u> ерстак	Делает активным лист Верстак.	3
 Предъявить <u>OW</u> -справку	Делает активным лист OW-справка.	4
 Снять защиту ЛРО	Открывает возможность редактирования листа ЛРО.	5
 → Очистить ЛРО	Очищает лист ЛРО.	6
 Сохранить	Сохраняет данные актуальной ОАД-задачи в её OWD-книге.	7
 Сведения о задаче...	Предъявляет краткую справку об актуальной ОАД-задаче.	8
Масштаб...	Иницирует Excel-диалог «Масштаб».	9
 Печать...	Иницирует Excel-диалог «Печать».	10

* Префикс пункта меню в виде знака «→» указывает, что данный пункт замещает в меню предшествующий пункт по завершению последним своих действий.

Таблица Д.14 – Контекстное меню листа OW-справка

Пункт меню	Действие OntoWorker при выборе пункта меню	№
 Предъявить <u>С</u> ИС	Делает активным лист СИС.	1
 Предъявить <u>О</u> ТОС	Делает активным лист ОТОС.	2
 Предъявить <u>В</u> ерстак	Делает активным лист Верстак.	3
 Предъявить <u>Л</u> РО	Делает активным лист ЛРО.	4
 Показать содержание справки	Выполняет на листе справки переход к разделу «Содержание OW-справки».	5
 О программе	Предъявляет сведения о реализации OntoWorker.	6
Масштаб...	Иницирует Excel-диалог «Масштаб».	7
 Печать...	Иницирует Excel-диалог «Печать».	8

ПРИЛОЖЕНИЕ Е

Акты внедрения результатов исследования

УТВЕРЖДАЮ
 Генеральный директор
 ООО «НПК «Сетецентрические
 платформы»
 А.Н. Седов
 « 06 » марта 2024 г.

АКТ О ВНЕДРЕНИИ

результатов диссертационной работы Семеновой Валентины Андреевны
 «Методическое и алгоритмическое обеспечение подготовки контекста
 для вывода формальных понятий в онтологическом анализе данных»

Комиссия в составе:

Председатель

Седов А.Н. – генеральный директор ООО «НПК «Сетецентрические платформы»;

Члены комиссии:

Новичков Д.Е. – заместитель генерального директора по разработкам ООО «НПК «Сетецентрические платформы»;

Панов А.О. – аналитик ООО «НПК «Сетецентрические платформы» составили настоящий акт о том, что при разработке в ООО «НПК «Сетецентрические платформы» конфигураций программно-аппаратного обеспечения спутников ДЗЗ, ориентированных на групповую работу, были использованы следующие результаты диссертационной работы Семеновой Валентины Андреевны «Методическое и алгоритмическое обеспечение подготовки контекста для вывода формальных понятий в онтологическом анализе данных»:

- 1) методы описания морфологического пространства выбора конфигурации программно-аппаратного обеспечения спутника с помощью И/ИЛИ дерева, построенного на основе ограничений существования свойств у элементов программно-аппаратных составляющих;
- 2) метод и программные средства синтеза оптимальной конфигурации программно-аппаратного обеспечения спутника по критериям выгоды и издержек.

Председатель комиссии:

Члены комиссии:



Седов А.Н.

Новичков Д.Е.

Панов А.О.

open{code}

ОТКРЫТЫЙ КОД

Общество с ограниченной ответственностью "Открытый код"

Россия, 443001, г. Самара, ул. Ульяновская, 52/55, этаж 15, ком.14
Тел./факс: (846) 331-11-11, 331-21-01(02,03,04)
E-mail: info@o-code.ru | www.o-code.ru
ОГРН 1036300222100 | ИНН 6313007301 | КПП 63150100113.03.2024 № 171

Акт

о внедрении результатов диссертационного исследования

Результаты диссертационного исследования **Семеновы Валентины Андреевны** на тему «*Методическое и алгоритмическое обеспечение подготовки контекста для вывода формальных понятий в онтологическом анализе данных*» в части, связанной с

разработанной в диссертации методикой семантической идентификации объекта на основе неполной и противоречивой информации при наличии ограничений существования свойств объекта, которая позволяет в полной мере формализовать и обеспечить вычислительные составляющие методологии морфологического анализа и синтеза технических решений,

использованы

для анализа и эскизного синтеза проектных решений при разработке архитектур программно-аппаратных комплексов систем поддержки принятия решений в органах власти, государственных корпорациях, промышленных предприятиях и систем контроля качества продукции промышленных предприятий.

Директор по управлению проектами –
исполнительный директор ООО
«Открытый код»,
доктор технических наук



Ситников П. В.

УТВЕРЖДАЮ

Директор СамНЦ РАН

академик


 С.Н. Шевченко
 «21» марта 2024 г.



АКТ

о внедрении результатов диссертационной работы
 на соискание ученой степени кандидата технических наук,
 подготовленной младшим научным сотрудником Семеновой Валентиной Андреевной
 на тему «Методическое и алгоритмическое обеспечение подготовки контекста
 для вывода формальных понятий в онтологическом анализе данных»,
 в Институте проблем управления сложными системами
 Российской академии наук – обособленном подразделении
 Федерального государственного бюджетного учреждения науки
 Самарского федерального исследовательского центра Российской академии наук
 (ИПУСС РАН – СамНЦ РАН)

Настоящим подтверждаем, что результаты диссертационной работы, подготовленной В.А. Семеновой на тему «Методическое и алгоритмическое обеспечение подготовки контекста для вывода формальных понятий в онтологическом анализе данных», составляют часть итогов научно-исследовательских работ (НИР) ИПУСС РАН – СамНЦ РАН. Результаты диссертационной работы, связанные моделированием и учетом ограничений существования свойств у объектов при интеллектуальном анализе неполных и противоречивых данных об исследуемых предметных областях, представлены в отчётах о НИР ИПУСС РАН – СамНЦ РАН, в которых В.А. Семенова принимала участие в 2015-2024 гг.:

- Тема НИР: «Интерсубъективное управление инновационным развитием социотехнических объектов с применением онтологических моделей ситуаций»; гос. рег. № АААА-А16-116040410061-0. Руководитель темы С.В. Смирнов. Итоговый отчет. – Самара: ИПУСС РАН – СамНЦ РАН, 2019. - 111 с.
- Тема НИР: «Разработка и исследование методов и средств аналитического конструирования, компьютерного представления знаний, вычислительных алгоритмов и мультиагентных технологий в задачах оптимизации процессов управления сложными системами»; гос. рег. № АААА-А19-119030190053-2. Руководитель темы С.В. Смирнов. Итоговый отчет. – Самара: ИПУСС РАН – СамНЦ РАН, 2022. - 248 с.
- Тема НИР: «Разработка и исследование теории и методов цифровизации управления сложными системами на основе аналитического и численного моделирования, компьютерных онтологий и многоагентных сетей»; код темы FMRW-2022-0030 (Минобрнауки России). Руководитель темы С.В. Смирнов. Промежуточный отчет. – Самара: ИПУСС РАН – СамНЦ РАН, 2024. – 130 с.

Директор ИПУСС РАН – СамНЦ РАН
 д.т.н.



С.Ю. Боровик

Ученый секретарь ИПУСС РАН – СамНЦ РАН
 д.т.н.



Т.В. Моисеева

УТВЕРЖДАЮ
 Проректор по учебной работе
 Гаврилов А.В.
 « 26 » 06 2023 г.



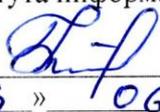
АКТ
 об использовании результатов
 диссертационного исследования в учебном процессе

Результаты диссертационного исследования **Семеновой Валентины Андреевны** по теме «**Методическое и алгоритмическое обеспечение подготовки контекста для вывода формальных понятий в онтологическом анализе данных**», выполненного в ИПУСС РАН и аспирантуре СамГТУ, а именно, методика использования многозначной векторной логики для адекватного моделирования неполных и противоречивых суждений о свойствах, присущих объектам изучаемой предметной области; модель системы свойств, измеряемых у объектов изучаемой предметной области, с ограничениями существования и метод её редукции; методика использования неравенства Хёвдинга для оценки необходимого количества испытаний в статистических имитационных экспериментах; программная лаборатория для онтологического анализа данных, используются при выполнении научно-исследовательских работ и выпускных квалификационных работ по образовательным программам бакалавриата («Информационные технологии», направление подготовки 02.03.02 Фундаментальная информатика и информационные технологии) и магистратуры («Инженерия программного обеспечения», направление подготовки 02.04.02 Фундаментальная информатика и информационные технологии, магистерская программа). Акт составлен на основании решения кафедры программных систем, протокол №10 от «22» июня 2023 г.

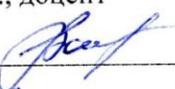
Заведующий кафедрой
 программных систем

 /Востокин С.В.
 « 26 » 06 2023 г.

Директор департамента
 образовательных программ
 Института информатики и кибернетики

 /Коломиец Э.И.
 « 26 » 06 2023 г.

Начальник методического отдела
 учебно-методического управления
 к.т.н., доцент

 /Вдовин Р.А.
 « 26 » 06 2023 г.

Начальник отдела сопровождения
 научных исследований
 к.т.н., доцент

 /Родионов Л.В.
 « 26 » 06 2023 г.