

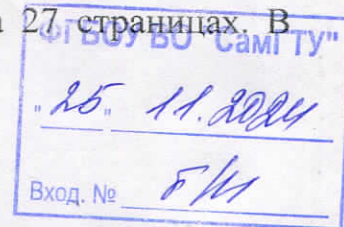
## ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

официального оппонента Семеновой Дарьи Владиславовны на диссертацию Семеновой Валентины Андреевны «Методическое и алгоритмическое обеспечение подготовки контекста для вывода формальных понятий в онтологическом анализе данных», представленную на соискание учёной степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика.

**Актуальность темы исследования.** Стремительный рост и разнообразие цифровых данных обуславливают необходимость использования онтологического подхода для структурирования информации и обеспечения её семантической интерпретируемости. Однако, ручное построение онтологий экспертами — дорогостоящий и длительный процесс, приводящий к субъективности и несогласованности результатов. Более того, при описании одной и той же предметной области разными экспертами могут быть созданы разные онтологии. На практике подобные задачи имеют сложную информационную ситуацию, для которой характерна неопределенность знаний о свойствах объекта в связи с неточностью и неполнотой экспериментальной информации. Поэтому актуальной задачей является разработка методов обработки исходных данных для автоматизированного построения качественных онтологий для описания понятий исследуемой предметной области. Диссертационная работа Семеновой В.А. посвящена разработке алгоритмического и программного аппарата для формализации объектно-признакового представления в виде однозначного контекста об изучаемой предметной области на основе имеющихся эмпирических данных из различных, несогласованных источников с учетом априорных знаний об измеряемых свойствах.

**Характеристика диссертационной работы по главам.** Диссертационная работа Семеновой В.А. состоит из введения, пяти разделов, заключения, списка сокращений и обозначений, списка используемой литературы и приложений. Объём текста составляет 150 страниц и приложений, приведенных на 27-страницах. В

С отзывом ознакомлена Семенова В.А.  
Секр 25.11.24



тексте содержится 57 рисунков, библиографический список состоит из 101 источника.

*Во введении* отражена актуальность темы исследования, сформулированы цель и задачи диссертации, показаны научная новизна, практическая ценность, достоверность и обоснованность результатов диссертации, приведены основные положения, выносимые на защиту, указана степень апробации и реализации результатов диссертационного исследования.

*В первой главе* рассматриваются теоретические основы двух, используемых в дальнейшем, методов «автоматического» построения онтологий: анализа формальных понятий (АФП) и анализа ограничений существования свойств (ОСС). Излагаются необходимые определения и положения, а также вводятся обозначения, используемые далее. Идея АФП основана на построении частично упорядоченной решетки по произвольному бинарному или нечёткому отношению и формализации описания понятия в виде пары (объём, содержание). Основное направление использования методов анализа формальных понятий для формальных контекстов состоит в поиске взаимосвязей между объектами и их атрибутами/признаками в условиях неполной или неточной информации. Таким образом, этот подход предлагает средства моделирования онтологий и таксономий предметных областей на основе решеток понятий, а также модели точных и приближенных зависимостей в данных. Идея ОСС основана на выявлении всех нормальных (замкнутых и совместимых) подмножеств в множестве свойств. Данный подход использует методы реляционной алгебры и позволяет строить нормализованные иерархии, начиная с набора понятий, описанных ключевыми словами. Основным результатом главы заключается в обосновании необходимости синтеза рассмотренных методов для получения новых способов построения и предобработки формальных контекстов.

*Во второй главе* предложена модификация модели системы измеряемых свойств в ОСС и метод построения «мягкого» формального контекста для АФП. Автор обосновывает использование концептуального шкалирования и векторной логики  $V^{TF}$  для моделирования неполноты и противоречивости разнотипных

эмпирических данных. Предложено обобщение стандартного протокола измерений для таблицы «объект-свойство».

*В третьей главе* непосредственно описана методика подготовки однозначного контекста для вывода формальных понятий в онтологическом анализе данных. Данная методика предусматривает дефаззификацию и нормализацию исходного формального контекста. Предложена новая двухэтапная схема пороговой дефаззификации нестрогого формального контекста, построенного с использованием векторной логики  $V^{TF}$  для оценки базовых семантических суждений о предметной области. Предложен новый эвристический метод нормализации контекста, согласно которому из формального контекста исключаются наименее достоверные базовые семантические суждения. Метод позволяет привести наборов свойств объектов в соответствие с априорными знаниями субъекта о зависимостях между измеряемыми свойствами объектов изучаемой предметной области. Для ускорения метода нормализации контекста предложен метод редуцирования модели системы измеряемых свойств на основе выделения в ней классов эквивалентности измеряемых свойств и расширенных отношений на множестве этих классов. Доказаны необходимые и достаточные условия существования нормального подмножества в СИС.

*Четвертая глава* посвящена описанию алгоритмов и комплекса проблемно-ориентированных программ, направленных на подготовку контекста для вывода формальных понятий в онтологическом анализе данных.

*В пятой главе* приведены результаты апробации разработанных методов и программных средств к решению ряда прикладных задач.

*В заключении* приведены основные результаты диссертационной работы.

**Новизна результатов проведенного исследования.** Основные пункты научной новизны результатов, представленных в работе, состоят в следующем:

1. Предложен новый метод построения нестрогого формального контекста из неполных противоречивых разнотипных эмпирических данных. Метод использует операции концептуального шкалирования, а также векторную логику  $V^{TF}$  для оценки базовых семантических суждений о предметной области. Предложенный метод позволил модифицировать модель системы измеряемых

свойств для таблицы «сопряженная пара свойств–свойство» и обобщить стандартный протокол измерений для таблицы «объект–свойство».

2. Разработан новый двухэтапный метод построения однозначного формального контекста для вывода формальных понятий исследуемой предметной области. Первый этап включает дефаззификацию нестрогого формального контекста. Для второго этапа предложен и теоретически обоснован (Утверждения 3.1–3.4) новый эвристический метод нормализации формального контекста, учитывающий ограничения существования свойств. Это является наиболее значимым результатом диссертации.

3. Разработан комплекс алгоритмов и проблемно-ориентированных программ подготовки контекста для вывода формальных понятий, реализующий предложенные модели и методы и позволяющий проводить предобработку данных для реальных прикладных задач и вычислительных экспериментов с целью построения онтологии предметной области.

**Практическая ценность полученных результатов.** Разработанный комплекс алгоритмов и проблемно-ориентированных программ может быть достаточно эффективно использован для построения однозначных формальных контекстов на основе неполных и противоречивых данных с учётом априорных знаний субъекта о зависимостях между измеряемыми свойствами объектов изучаемой предметной области. Таким образом, полученные в диссертационной работе результаты применимы на практике при проектировании, построении и оптимизации онтологий предметных областей.

**Степень обоснованности и достоверности полученных результатов.** Все научные положения, выводы и рекомендации, сформулированные в диссертации, подтверждаются корректными применением моделей и методов следующих математических теорий: анализ формальных понятий, анализ ограничений существования свойств, векторная логика  $V^{TF}$ , статистическое моделирование, объектно-ориентированное проектирование. Все теоретические результаты подтверждаются результатами, полученными на основе имитационного моделирования и численных расчетов для реальных прикладных задач.

Кроме того, результаты диссертационного исследования апробированы на многочисленных отечественных и международных конференциях и опубликованы в отечественных и зарубежных изданиях. По материалам диссертаций опубликовано 26 работ, из них 8 статей в журналах, включенных в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, 17 статей в сборниках материалов международных и всероссийских научных конференций, школы-семинара, получено одно свидетельство о государственной регистрации программы для ЭВМ, имеются 3 акта о внедрении и 1 акт об использовании результатов диссертационного исследования.

### *Замечания и рекомендации.*

1. Литературный обзор по теме диссертационной работы выполнен в объеме, достаточном для обоснования выбора векторной логики  $V^{TF}$  в качестве инструмента для оценки базовых семантических суждений о предметной области. Однако, утверждение в выводах по второй главе на стр. 40 о том, что векторная  $V^{TF}$ -логика является обобщением нечёткой логики, весьма спорно. В этой ситуации можно говорить об ещё одном способе представления нечёткой меры. Анализ литературы показал, что в обзоре не освещены современные работы по представлению нечётких мер и «мягких» измерений таких авторов как K. Atanassov, M. Grabisch, G. Beliakov, L. Garmendia, и др. В частности, было бы полезно ознакомиться с публикациями, указанными в обзорной статье Gleb Beliakov, Jian-Zhang Wu, Weiping Ding, Representation, optimization and generation of fuzzy measures, Information Fusion, Volume 106, 2024. В качестве рекомендаций для дальнейших исследований можно посоветовать рассмотреть подход, предложенный в работе G. Beliakov, M. Pagola, T. Wilkin, Vector valued similarity measures for Atanassov's intuitionistic fuzzy sets, Information Sciences, Volume 280, 2014, Pages 352-367.
2. Утверждения 3.1, 3.3 носят неформальный характер и больше похожи на пояснения. Утверждения лучше было сформулировать в традиционной форме: разъяснительная часть, условие, требование или заключение.

Например, утверждение 3.1 можно переформулировать следующим образом: Пусть  $X \subseteq M$  – ВЗО-группа,  $N \subseteq M$  – нормальное подмножество измеряемых свойств и  $X \cap N \neq \emptyset$ . Тогда  $X \subseteq N$ .

3. Исследования выигрышей от реструктуризации системы измеряемых свойств проводятся при заметно малом количестве измеряемых свойств.
4. В выводах по третьей главе в пункте 3 (стр. 72) говорится об эффективном решении задачи комбинаторной оптимизации, включенной в метод нормализации. Интуитивно понятно о чём идёт речь, однако, в явном виде постановка задачи не представлена в тексте диссертации.
5. Алгоритмы реструктуризации системы измеряемых свойств (раздел 4.3.2) носят переборный характер. В связи с этим представляет интерес оценка вычислительной сложности разработанных алгоритмов.
6. В работе пятая глава посвящена построению и предобработке формальных контекстов для различных прикладных задач. Для практических целей интерес представляет визуализация решётки понятий. Однако, ни в одном из примеров не приведена решётка формальных понятий, построенная по полученному формальному контексту.

В целом, несмотря на высказанные замечания, представленная диссертация выполнена на высоком научно-техническом уровне и представляет собой законченную научно-квалификационную работу.

#### ***Общее заключение.***

Диссертационная работа Семеновой Валентины Андреевны соответствует специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика имеет внутреннее единство и является завершённой научно-квалификационной работой, в которой на основании выполненных автором исследований содержится решение задачи обработки эмпирических объектно-признаковых данных для построения и подготовки однозначного формального контекста для вывода формальных понятий, имеющей значение для развития онтологического анализа данных. Материалы исследования в достаточно полном объеме отражены в публикациях автора и прошли апробацию на международных и

всероссийских научных конференциях. Автореферат полностью и точно отражает содержание диссертации.

Считаю, что диссертационная работа «Методическое и алгоритмическое обеспечение подготовки контекста для вывода формальных понятий в онтологическом анализе данных» соответствует требованиям пп. 9–11, 13, 14 «Положения о присуждении учёных степеней» постановления Правительства Российской Федерации от 24.09.2013 г. № 842, а её автор Семенова Валентина Андреевна заслуживает присуждения учёной степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Официальный оппонент:

доцент кафедры высшей и прикладной  
математики федерального  
государственного автономного  
образовательного учреждения высшего  
образования «Сибирский федеральный  
университет», кандидат физико-  
математических наук (05.13.01 –  
Системный анализ, управление и обработка  
информации),

доцент

Семенова Дарья Владиславовна

18 ноября 2024

Федеральное государственное автономное образовательное учреждение высшего образования «Сибирский федеральный университет». Адрес: 660041, г. Красноярск, пр. Свободный, д. 79. Тел: +7 (391) 206-21-18. E-mail: [DVSeменова@sfu-kras.ru](mailto:DVSeменова@sfu-kras.ru)

