

**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ
ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
ИМЕНИ И.Т. ТРУБИЛИНА»**

Факультет прикладной информатики

Кафедра Компьютерных систем и технологий

КУРСОВАЯ РАБОТА

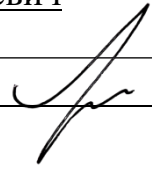
по дисциплине: Интеллектуальные системы и технологии

на тему: «АСК-анализ пригодности воды для питья»

Выполнил студент группы: ИТ2341 Польский Григорий Алексеевич

Допущен к защите _____

Руководитель проекта д.э.н., к.т.н., профессор Луценко Е.В. (


(подпись, расшифровка подписи)

Защищен _____

(дата)

Оценка _____

ОТЛИЧНО

Краснодар
2025

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ
УНИВЕРСИТЕТ ИМЕНИ И. Т. ТРУБИЛИНА»

Факультет прикладной информатики

**РЕЦЕНЗИЯ
на курсовую работу**

Студента Польского Григория Алексеевича
курса 2 очной формы обучения группы ИТ2341
Направление подготовки 09.04.02 «Информационные системы и технологии»
Наименование темы «АСК-анализ пригодности воды для питья»
Рецензент: Луценко Евгений Вениаминович, д.э.н., к.т.н., профессор
(Ф.И.О., ученое звание и степень, должность)

Оценка качества выполнения курсовой работы

№ п/п	Показатель	Оценка соответствия («неудовлетворительно», «удовлетворительно», «хорошо», «отлично»)
1	Степень полноты обзора состояния проблемы и корректности постановки цели и задач исследования	отлично
2	Уровень и корректность использования в работе различных методов исследований	отлично
3	Степень комплексности работы, применения в ней знаний общепрофессиональных и специальных дисциплин	отлично
4	Ясность, четкость, последовательность и обоснованность изложения	отлично
5	Применение современных технологий обработки информации	отлично
6	Качество оформления работы (общий уровень грамотности, стиль изложения, качество иллюстраций, соответствие требованиям по оформлению)	отлично
7	Уровень освоения компетенций в результате выполнения курсовой работы (проекта)	отлично
8	Ответы на вопросы при защите	отлично

Достоинства работы _____

Недостатки работы _____

Итоговая оценка при защите _____ отлично

Рецензент _____ (Е. В. Луценко)

«19» февраля 2025 г.

РЕФЕРАТ

Курсовая работа содержит: 89 страниц, 38 рисунков, 16 таблиц, 48 литературных источников.

Ключевые слова: СИСТЕМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, AIDOS-X, КОГНИТИВНЫЙ АНАЛИЗ, МОДЕЛИ, ШКАЛЫ, КЛАССЫ.

Целью работы является проведение автоматизированного системно-когнитивного анализа пригодности воды для питья.

Для достижения поставленной цели необходимо провести анализ методов формирования обобщенных образов классов и решения задач идентификации конкретных объектов с классами принятия решений и исследования моделируемой предметной области путем исследования модели.

УДК 004.8

4.1.1. Общее земледелие и растениеводство

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ ПРИГОДНОСТИ ВОДЫ ДЛЯ ПИТЬЯПольский Григорий Алексеевич
студент факультета ПИ, группы ИТ2341grisha.polsky@mail.ru*Кубанский Государственный Аграрный университет имени И.Т.Трубилина, Краснодар, Россия*

В условиях роста загрязнения природных водоемов и ужесточения санитарных норм важно совершенствовать методы оценки пригодности воды для питья. В данной статье для анализа данных предлагается использовать Автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – интеллектуальную систему «Эйдос». Приводится подробный численный пример. Этот пример содержит много разнообразных наглядных табличных и графических выходных форм и может использоваться для обучения применению АСК-анализа и системы «Эйдос» для научных исследований, для выработки практических рекомендаций и обоснования научных положений о механизмах действия причинно-следственных связей.

Ключевые слова: АСК-АНАЛИЗ, АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА «ЭЙДОС»

UDC 004.8 UDC 004.8

4.1.1. General agriculture and crop production

AUTOMATED SYSTEM-COGNITIVE ANALYSIS OF WATER POTABILITYPolskiy Grigoriy Alekseevich
student of the faculty of Applied Informatics, IT2341grisha.polsky@mail.ru*Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia*

In conditions of increasing pollution of natural reservoirs and stricter sanitary standards, it is important to improve methods for assessing the suitability of water for drinking. In this article, it is proposed to use Automated System Cognitive Analysis (ASK analysis) and its software toolkit, the intelligent Eidos system, for data analysis. A detailed numerical example is given. This example contains many different visual tabular and graphical output forms and can be used to teach the use of ASK analysis and the Eidos system for scientific research, to develop practical recommendations and substantiate scientific provisions on the mechanisms of causal relationships.

Keywords: ASC-ANALYSIS, AUTOMATED SYSTEM-COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS"

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ)	7
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	7
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	7
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	8
1.4. ЦЕЛЬ РАБОТЫ	8
2. METHODS (МЕТОДЫ)	8
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	8
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	9
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ.....	9
2.4. СИСТЕМА «ЭЙДОС» – ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА.....	11
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	18
3. RESULTS (РЕЗУЛЬТАТЫ)	21
3.1. Задача-1. Когнитивная структуризация предметной области. Две интерпретации классификационных и описательных шкал и градаций	21

3.1.1. Универсальная постановка задачи, не зависящая от предметной области	21
3.1.2. Конкретное решение задачи в данной работе	22
3.2. ЗАДАЧА-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	22
3.2.1. Универсальная постановка задачи, не зависящая от предметной области	22
3.2.2. Конкретное решение задачи в данной работе	23
3.3. ЗАДАЧА-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ	28
3.3.1. Универсальная постановка задачи, не зависящая от предметной области	28
3.3.2. Конкретное решение задачи в данной работе	36
3.4. ЗАДАЧА-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	39
3.4.1. Универсальная постановка задачи, не зависящая от предметной области	39
3.4.2. Конкретное решение задачи в данной работе	39
3.5. ЗАДАЧА-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	42
3.5.1. Универсальная постановка задачи, не зависящая от предметной области	42
3.5.2. Конкретное решение задачи в данной работе	42
3.6. ЗАДАЧА-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	43
3.6.1. Универсальная постановка задачи, не зависящая от предметной области	43
3.6.1.1. Интегральный критерий «Сумма знаний»	44
3.6.1.2. Интегральный критерий «Семантический резонанс знаний»	45
3.6.1.3. Важные математические свойства интегральных критериев	46
3.6.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»	47
3.7. ЗАДАЧА-7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ	50
3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ	50
3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области	50
3.7.1.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»	51
3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»	52
3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области	52
3.7.2.2. Конкретное решение задачи управления в данной работе в системе «Эйдос»	56
3.8. ЗАДАЧА-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ	56
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)	56
3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области	56
3.8.1.2. Конкретное решение задачи в данной работе	56
3.8.2. Кластерно-конструктивный анализ классов	60
3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области	60
3.8.2.2. Конкретное решение задачи в данной работе	61
3.8.3. Кластерно-конструктивный анализ значений описательных шкал	62
3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области	62
3.8.3.2. Конкретное решение задачи в данной работе	63
3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны	65
3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области	65
3.8.4.2. Конкретное решение задачи в данной работе	66
3.8.5. Нелокальная нейронная сеть	67
3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области	67
3.8.5.2. Конкретное решение задачи в данной работе	68
3.8.6. 3D-интегральные когнитивные карты	69
3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области	69
3.8.6.2. Конкретное решение задачи в данной работе	69
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения)	70
3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области	70
3.8.7.2. Конкретное решение задачи в данной работе	71
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения)	72
3.8.8.1. Универсальная постановка задачи, не зависящая от предметной области	72

3.8.8.2. Конкретное решение задачи в данной работе	73
3.8.9. Когнитивные функции	75
3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области	75
3.8.9.2. Конкретное решение задачи в данной работе	75
3.8.10. Значимость описательных шкал и их градаций.....	81
3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области	81
3.8.10.2. Конкретное решение задачи в данной работе	82
3.8.11. Степень детерминированности классов и классификационных шкал	84
3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области	84
3.8.11.2. Конкретное решение задачи в данной работе	85
4. DISCUSSION (ОБСУЖДЕНИЕ)	87
5. CONCLUSIONS (ВЫВОДЫ)	88
REFERENCES (ЛИТЕРАТУРА).....	88

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Вода – один из важнейших ресурсов, необходимых для жизни и здоровья человека. Качество питьевой воды напрямую влияет на эпидемиологическую обстановку, уровень заболеваемости и продолжительность жизни населения. Однако в современных условиях антропогенная нагрузка на водные экосистемы значительно возрастает, что приводит к ухудшению качества водных ресурсов. Загрязнение воды промышленными отходами, сельскохозяйственными удобрениями, тяжелыми металлами и патогенными микроорганизмами создает серьезные риски для здоровья человека.

Пригодность воды для питья определяется целым рядом физических, химических и биологических показателей. Согласно санитарно-гигиеническим нормам, вода должна быть прозрачной, не иметь выраженного запаха и вкуса, не содержать вредных примесей, бактерий и вирусов. В этой связи контроль качества воды приобретает стратегическое значение.

Существуют различные методы оценки пригодности воды для питья, включая химический анализ, микробиологические исследования и физико-химические тесты. Однако современные вызовы требуют более комплексного подхода к анализу данных о качестве воды. Интеграция автоматизированных систем в процессы мониторинга и оценки пригодности воды позволяет повысить точность анализа и оперативность принятия решений.

Одним из перспективных методов является Автоматизированный системно-когнитивный анализ (АСК-анализ), который позволяет выявлять причинно-следственные связи между различными факторами, влияющими на качество воды, и прогнозировать возможные изменения её состава. В сочетании с интеллектуальными информационными системами, такими как «Эйдос», этот подход может существенно повысить эффективность исследований и принятия управленческих решений в сфере водоснабжения.

Настоящее исследование посвящено применению АСК-анализа для оценки пригодности воды для питья, выявлению ключевых факторов, влияющих на её качество, и обоснованию практических рекомендаций по обеспечению безопасности водоснабжения.

1.2. Объект и предмет исследования

Объект исследования (моделирования) – вода.

Предмет исследования – выявление причинно-следственных зависимостей между составом воды и её пригодностью для питья.

1.3. Проблема, решаемая в работе и ее актуальность

В условиях растущего загрязнения водных ресурсов и изменения климатических факторов традиционные методы оценки качества питьевой воды требуют значительных временных и финансовых затрат, а также не всегда позволяют оперативно выявлять скрытые взаимосвязи между различными факторами загрязнения. Это усложняет процесс мониторинга, прогнозирования и принятия управленческих решений в сфере водоснабжения.

Обеспечение населения безопасной питьевой водой является одной из приоритетных задач в области общественного здравоохранения и экологии. Внедрение современных аналитических методов, таких как Автоматизированный системно-когнитивный анализ (АСК-анализ), позволяет не только ускорить процесс оценки качества воды, но и повысить точность прогнозирования изменений её состава. Использование интеллектуальных систем, таких как «Эйдос», способствует более глубокому пониманию причинно-следственных связей между различными параметрами воды, что делает анализ более комплексным и объективным. Это особенно важно в условиях глобальных экологических вызовов и необходимости оперативного реагирования на угрозы, связанные с качеством питьевой воды.

1.4. Цель работы

Целью работы является решение поставленной проблемы.

Достижение поставленной цели обеспечивается решением ряда *задач* и подзадач, которые являются *этапами* достижения цели. Конкретная формулировка этих задач зависит от метода решения проблемы, поэтому обоснованно мы сформулируем их в конце раздела, т.е. после обоснованного выбора метода решения проблемы и его краткого описания.

2. METHODS (МЕТОДЫ)

2.1. Обоснование требований к методу решения проблемы

Из специфики поставленной проблемы сопоставимости обработки в одной модели исходных, представленных в разных типах шкал числовых и текстовых (лингвистических) и в разных единицах измерения, вытекают следующие *требования* к методу решения проблемы:

1. Метод должен обеспечивать устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных (неточных) взаимозависимых (нелинейных) данных большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

2. Метод решения проблемы не должен предъявлять чрезмерно жестких требований к исходным данным, которые практически невозможно выполнить, а должен обеспечивать обработку тех данных, которые реально есть.

3. Метод должен не теоретически, а реально на практике решать поставленную проблему, а значит, он должен иметь поддерживающий его программный инструментарий, находящийся в полном открытом бесплатном доступе.

2.2. Литературный обзор методов решения проблемы, их характеристика и оценка степени соответствия обоснованным требованиям

Поиск в Internet математических методов и реализующих их программных систем, *одновременно* удовлетворяющих всем требованиям, обоснованным в п.2.1 данной работы показал, что альтернатив Автоматизированному системно-когнитивному анализу и его программному инструментарии – системе «Эйдос» в настоящее время практически нет [1-4].

2.3. Автоматизированный системно-когнитивный анализ (АСК-анализ) как метод решения проблемы

Автоматизированный системно-когнитивный анализ (АСК-анализ) предложен *проф.Е.В.Луценко* в 2002 году в ряде статей 1997-2001 годов¹ и фундаментальной монографии [2].

Сам термин: «Автоматизированный системно-когнитивный анализ (АСК-анализ)» был предложен проф.Е.В.Луценко в 2001 году. На тот момент этот термин вообще не встречался в Internet. Сегодня по соответствующему запросу в Гугле находится около 23000 сайтов с этим сочетанием слов².

***Примечание:** Ниже приведено очень краткое описание АСК-анализа и системе «Эйдос». Это описание может выглядеть как нескромность и самовосхваление. Но автор просит читателей понять его правильно. Это сделано исключительно для тех довольно многочисленных читателей, которые впервые слышат об этом методе и системе.*

АСК-анализ включает:

- теоретические основы, в частности базовую формализуемую когнитивную концепцию;
- математическую модель, основанную на системном обобщении теории информации (СТИ);
- методику численных расчетов (структуры баз данных и алгоритмы их обработки);

¹ <http://lc.kubagro.ru/aidos/Sprab0802.pdf> (см. с публикации № 48).

² [https://yandex.ru/search/?text=Автоматизированный%2Всистемно-когнитивный%2Ванализ%2В\(АСК-анализ\)&lr=35&clid=2327117-18&win=360](https://yandex.ru/search/?text=Автоматизированный%2Всистемно-когнитивный%2Ванализ%2В(АСК-анализ)&lr=35&clid=2327117-18&win=360)

– программный инструментарий, в качестве которого в настоящее время выступает универсальная когнитивная аналитическая система «Эйдос» (интеллектуальная система «Эйдос»).

Более подробно математический метод АСК-анализа описан в работе [3] и ряде других [5]. Около половины из 730 опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент написания данной работы автором опубликовано 50 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получено 34 патента РФ на системы искусственного интеллекта, 360 публикации в изданиях, входящих в перечень ВАК РФ и приравненных им (по данным [РИНЦ](#)), 17 публикаций в изданиях, входящих в ядро РИНЦ, 6 статей в журналах, входящих в [WoS](#), 7 публикаций в журналах, входящих в [Скопус](#)³ [5, 6, 7].

Три монографии включены в фонды библиотеки конгресса США⁴.

АСК-анализ и система "Эйдос" были успешно применены в 8 докторских диссертациях по экономическим, техническим, биологическим наукам и в 8 кандидатских диссертациях по экономическим, техническим, психологическим и медицинским наукам, еще несколько докторских и кандидатских диссертаций по техническим, экономическим, филологическим и медицинским наукам с применением АСК-анализа находятся в стадии выхода на защиту.

Автор является основателем междисциплинарной научной школы: «Автоматизированный системно-когнитивный анализ»⁵. Научная школа: "Автоматизированный системно-когнитивный анализ" является междисциплинарным научным направлением на пересечении, по крайней мере, трех научных специальностей (согласно недавно утвержденной новой номенклатуры научных специальностей ВАК РФ⁶). Основные научные специальности, которым соответствует научная школа:

- 5.12.4. Когнитивное моделирование;
- 1.2.1. Искусственный интеллект и машинное обучение;
- 2.3.1. Системный анализ, управление и обработка информации.

Научная школа: "Автоматизированный системно-когнитивный анализ" включает следующие междисциплинарные научные направления:

- Автоматизированный системно-когнитивный анализ числовых и текстовых табличных данных [1-47];
- Автоматизированный системно-когнитивный анализ текстовых данных [32];

³ <http://lc.kubagro.ru/aidos/Sprab0802.pdf>

⁴ <https://catalog.loc.gov/vwebv/search?searchArg=Lutsenko+E.V.> (и кликнуть: "Search")

⁵ <https://www.famous-scientists.ru/school/1608>

⁶ <https://www.garant.ru/products/ipo/prime/doc/400450248/>

– Спектральный и контурный автоматизированный системно-когнитивный анализ изображений [31];

– Сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов [44].

Приводить здесь ссылки на все эти работы здесь вряд ли целесообразно. Отметим лишь, что у автора есть личный сайт [5] и страничка в РесечГейт [6] и РИНЦ [7], на которых можно получить более полную информацию о методе АСК-анализа и системе «Эйдос». Краткая информация об АСК-анализе и системе «Эйдос» есть в материале: http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf.

На сайте автора размещены тематические подборки публикаций по применению АСК-анализа и системы «Эйдос» в различных предметных областях [26-47].

Решение поставленной в работе проблемы сопоставимости в АСК-анализе и системе «Эйдос» обеспечивается *путем* метризации номинальных шкал, т.е. повышения их степени формализации до уровня числовых шкал [8]. Сама метризация номинальных шкал достигается *путем* вычисления количества информации, содержащегося в градациях номинальных шкал о получении той или иной урожайности [8]. Для работы с лингвистическими переменными применяются стандартные возможности АСК-анализа [5].

2.4. Система «Эйдос» – инструментарий АСК-анализа

Конечно, на системе «Эйдос» как говорят «Свет клином не сошелся». Существует много очень достойных систем искусственного интеллекта. Чтобы лично убедиться в этом достаточно самостоятельно осуществить поиск в Internet, просто посмотреть файлы: [NCKR-1](#), [NCKR-2](#), [NCKR-3](#), [NCKR-4](#) или пройти по ссылкам: <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/how-does-ml-dotnet-work>, <http://chat.openai.com/>, <https://poe.com/>, <https://neural-university.ru/>, <https://dzen.ru/a/ZCKZRKvrlEMBWOк8>, <https://ora.ai/>, <https://ora.ai/explore?path=trending>, <https://ora.ai/eugene-lutsenko/aidos>, <https://poe.com/Aidos-X>, <https://rudalle.ru/>, <https://bard.google.com/>, <https://chatbot.theb.ai/>, <https://problembo.com/ru/services> (может пригодиться - почта на 10 минут: <https://10minutemail.net/>), <https://poe.com/GPT-3.5-Turbo-Instruct>, <https://www.seaart.ai/home>, <https://ui.chatai.com/>.

И все же Универсальная когнитивная аналитическая система «Эйдос-Х++» отличается от большинства из этих систем, по крайней мере, некоторыми из следующих своих параметров:

- является универсальной и может быть применена во многих предметных областях, т.к. разработана в универсальной постановке, не зависящей от предметной области (<http://lc.kubagro.ru/aidos/index.htm>) и имеет 6 автоматизированных программных интерфейсов (API) ввода

данных из внешних источников данных различных типов: таблиц, текстов и графики. Система «Эйдос» является автоматизированной системой, т.е. предполагает непосредственное участие человека в реальном времени в процессе создания моделей и их использования для решения задач идентификации, прогнозирования, принятия решений и исследования предметной области путем исследования ее модели (автоматические системы работают без такого участия человека);

- является одной из первых и наиболее популярных отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта и программирования: есть акт внедрения системы «Эйдос» 1987 года (<http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>);

- реально работает, обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения (т.е. не предъявляет жестких требований к данным, которые невозможно выполнить, а обрабатывает те данные, которые есть);

- имеет «нулевой порог входа»:

- содержит большое количество интеллектуальных локальных (т.е. поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 392, соответственно: http://lc.kubagro.ru/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf, http://lc.kubagro.ru/Presentation_LutsenkoEV.pdf);

- находится в полном открытом бесплатном доступе (<http://lc.kubagro.ru/aidos/Aidos-X.htm>), причем с актуальными исходными текстами (<http://lc.kubagro.ru/AidosALL.txt>): открытая лицензия: [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/) (<https://creativecommons.org/licenses/by-sa/4.0/>), и это означает, что ей могут пользоваться все, кто пожелает, без какого-либо дополнительного разрешения со стороны первичного правообладателя – автора и разработчика системы «Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 34 свидетельства РосПатента РФ);

- является «интерпретатором интеллектуальных моделей», т.е. с одной стороны является инструментальной оболочкой, позволяющей без какого-либо программирования создавать интеллектуальные приложения на основе [конфигуратора статистических и системно-когнитивных моделей](#), а с другой стороны является run-time системой или средой исполнения, обеспечивающей эксплуатацию этих интеллектуальных приложений в адаптивном режиме.

- чтобы самостоятельно освоить систему Эйдос достаточно скачать со страницы: <http://lc.kubagro.ru/aidos/index.htm> и установить полную версию систему, а затем в режиме 1.3 скачать и установить из Эйдос-облака одно из интеллектуальных облачных Эйдос-приложений (http://lc.kubagro.ru/Source_data_applications/WebAppls.htm) и выполнять его,

следуя описанию приложения. Обычно это файл readme.pdf в папке: c:\Aidos-X\AID_DATA\Inp_data. Для изучения лучше выбирать самые новые приложения, автором которых является проф.Е.В.Луценко. Кроме того на странице: [http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud Eidos-application.pdf](http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf) есть более 300 полутора-часовых видео-занятий (на русском языке) и много других учебных материалов и примеров описания интеллектуальных-Эйдос-приложений.

- поддерживает on-line среду накопления знаний и обмена ими, широко используется во всем мире (<http://lc.kubagro.ru/map5.php>);



- обеспечивает мультиязычную поддержку интерфейса на 51 языке. Языковые базы входят в инсталляцию и могут пополняться в автоматическом режиме;

- наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA, т.е. поддерживать язык OpenGL);

- обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач идентификации, прогнозирования, поддержки принятия решений и исследования предметной области путем исследования ее системно-когнитивной модели, генерируя при этом очень большое количество табличных и графических выходных форм (развития когнитивная графика), у многих из которых нет никаких аналогов в других системах (примеры форм можно посмотреть в работе: http://lc.kubagro.ru/aidos/aidos18_LLS/aidos18_LLS.pdf);

- хорошо имитирует человеческий стиль мышления и является инструментом познания: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции, если эти эксперты уже есть, а если их еще нет, то она все равно дает верные

результаты познания, что будет признано будущими экспертами, когда они появятся;

- вместо того, чтобы предъявлять к исходным данным практически неосуществимые требования (вроде нормальности распределения, абсолютной точности и полных повторностей всех сочетаний значений факторов и их полной независимости и аддитивности) автоматизированный системно-когнитивный анализ (АСК-анализ) предлагает без какой-либо предварительной обработки осмыслить те данные, которые есть, и, тем самым, преобразовать их в информацию, а затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для принятия решений и управления) и решения задач классификации, поддержки принятия решений и содержательного эмпирического исследования моделируемой предметной области.

[В чем сила подхода, реализованного в системе Эйдос?](#) В том, что она реализует подход, эффективность которого не зависит от того, что мы думаем о предметной области и думаем ли вообще. Она формирует модели непосредственно на основе эмпирических данных, а не на основе наших представлений о механизмах реализации закономерностей в этих данных. Именно поэтому Эйдос-модели эффективны даже если наши представления о предметной области ошибочны или вообще отсутствуют.

[В этом и слабость этого подхода, реализованного в системе Эйдос.](#) Модели системы Эйдос - это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах.

[We are briefly describing a new innovative method of artificial intelligence: Automated system-cognitive analysis \(ASC-analysis\), which has its own software tools – intelligent system called "Eidos" \(open source software\).](#)

В разработке системы «Эйдос» были следующие этапы:

1-й этап, «подготовительный»: 1979-1992 годы. Математическая модель системы "Эйдос" разработана в 1979 и впервые прошла экспериментальную апробацию в 1981 году (первый расчет на компьютере на основе модели). С 1981 по 1992 система "Эйдос" неоднократно реализовалась на платформе Wang (на компьютерах Wang-2200C). В 1987 году впервые получен [акт внедрения](#) на одну из ранних версий системы «Эйдос», реализованную в среде персональной технологической системы «Вега-М» разработки автора (см.2-й акт).

2-й этап, «эра IBM PC и MS DOS»: 1992-2012 годы. Для IBM-совместимых персональных компьютеров система "Эйдос" впервые реализована на языках CLIPPER-87 и CLIPPER-5.01 (5.02) в 1992 году, а в 1994 году уже были получены [свидетельства РосПатента](#), первые в

Краснодарском крае и, возможно, в России на системы искусственного интеллекта (слева приведена титульная видеogramма финальной DOS-версии системы «Эйдос-12.5», июнь 2012 года). С тех пор и до настоящего времени система непрерывно совершенствуется на IBM PC.

3-й этап, «эра MS Windows xp, 8, 7»: 2012-2020 годы. С июня 2012 по 14.12.2020 система «Эйдос» развивалась на языке [Аляска-1.9](#) + [Экспресс++](#) + библиотека для работы с Internet xb2net. Система «Эйдос-X1.9» хорошо работала на всех версиях MS Windows кроме Windows-10, которая требовала специальной настройки. Наиболее трудоемкие в вычислительном отношении операции синтеза моделей и распознавания реализует с помощью графического процессора (GPU), что на некоторых задачах обеспечивает ускорение решение этих задач в несколько тысяч раз, что реально обеспечивает интеллектуальную обработку больших данных, большой информации и больших знаний (графический процессор должен быть на чипсете NVIDIA).

4-й этап, «эра MS Windows-10»: 2020-2021 годы. С 13.12.2020 года по настоящее время система «Эйдос» развивается на языке [Аляска-2.0](#) + [Экспресс++](#). Библиотека xb2net в ней больше не используется, т.к. все возможности работы с Internet входят в [базовые возможности языка программирования](#).

5-й этап, «эра Больших данных, информации и знаний»: 2022 год. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке xBase++eXpress++Advantage Database Server (ADS), обеспечивающей обработку больших данных, информации и знаний (Big Data, Big Information, Big Knowledge).

6-й этап, «эра Больших данных, информации и знаний»: с 2023 года по настоящее время. С 2023 развитие системы «Эйдос» будет осуществляться на языках Питон (Python), а также [Аляска-2.0](#) + [Экспресс++](#).

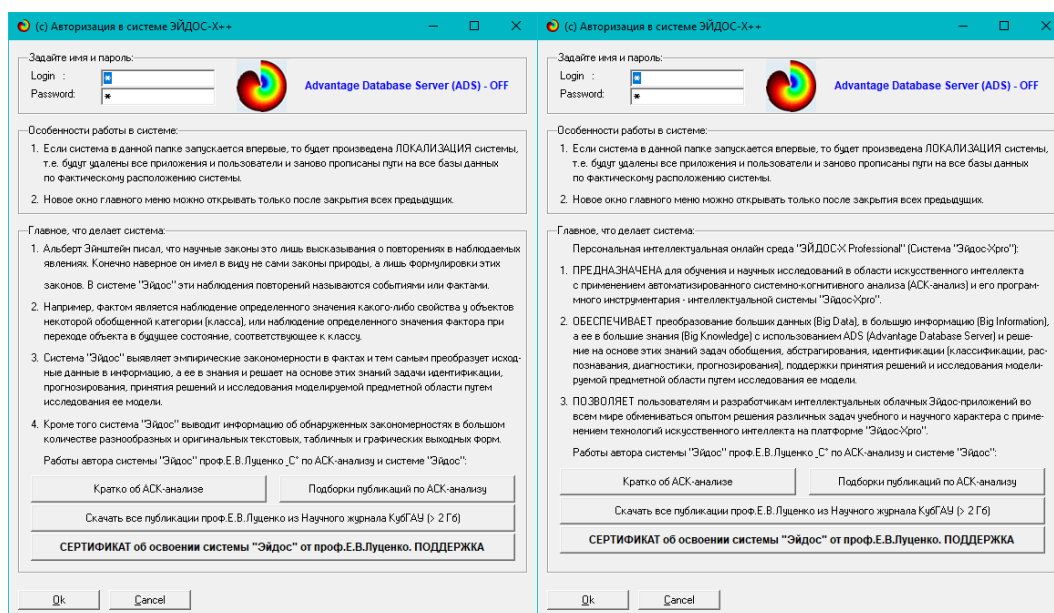
[Скачать и запустить систему «Эйдос-X++» \(самую новую на текущий момент версию\) или обновление системы до текущей версии.](#) Это наиболее полная на данный момент незащищенная от несанкционированного копирования портативная (portable) версия системы (не требующая инсталляции) с полными [исходными текстами](#) текущей версии (за исключением ключей доступа к ftp-серверу системы «Эйдос» и ключей лицензионного программного обеспечения), находящаяся в полном открытом бесплатном доступе (около 180 Мб). Обновление имеет объем около 10 Мб. [Кредо](#). Лаборатория в [ResearchGate](#) по АСК-анализу и системе «Эйдос».

[Задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения](#)⁷

На рисунке 1 приведена титульная видеोगрамма DOS-версии системы «Эйдос», а на рисунке 2 – титульные видеोगраммы текущей версии системы «Эйдос» (их в настоящее время 7 и они меняются по очереди при каждом запуске системы):

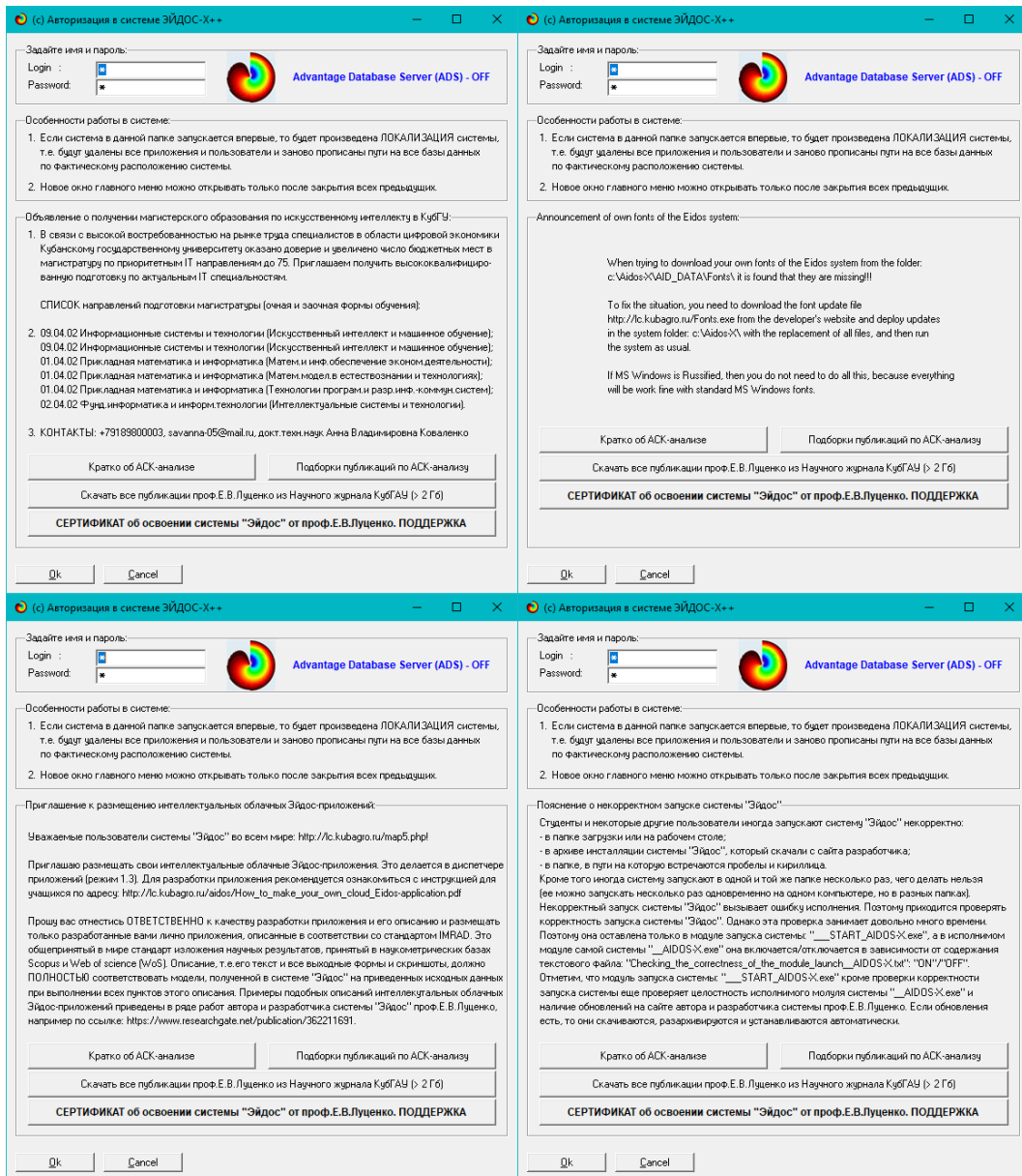


Рисунок 1. Титульная видеोगрамма DOS-версии системы «Эйдос» (до 2012 года)⁸



⁷ <http://lc.kubagro.ru/aidos/How to make your own cloud Eidos-application.pdf>

⁸ http://lc.kubagro.ru/pic/aidos_titul.jpg



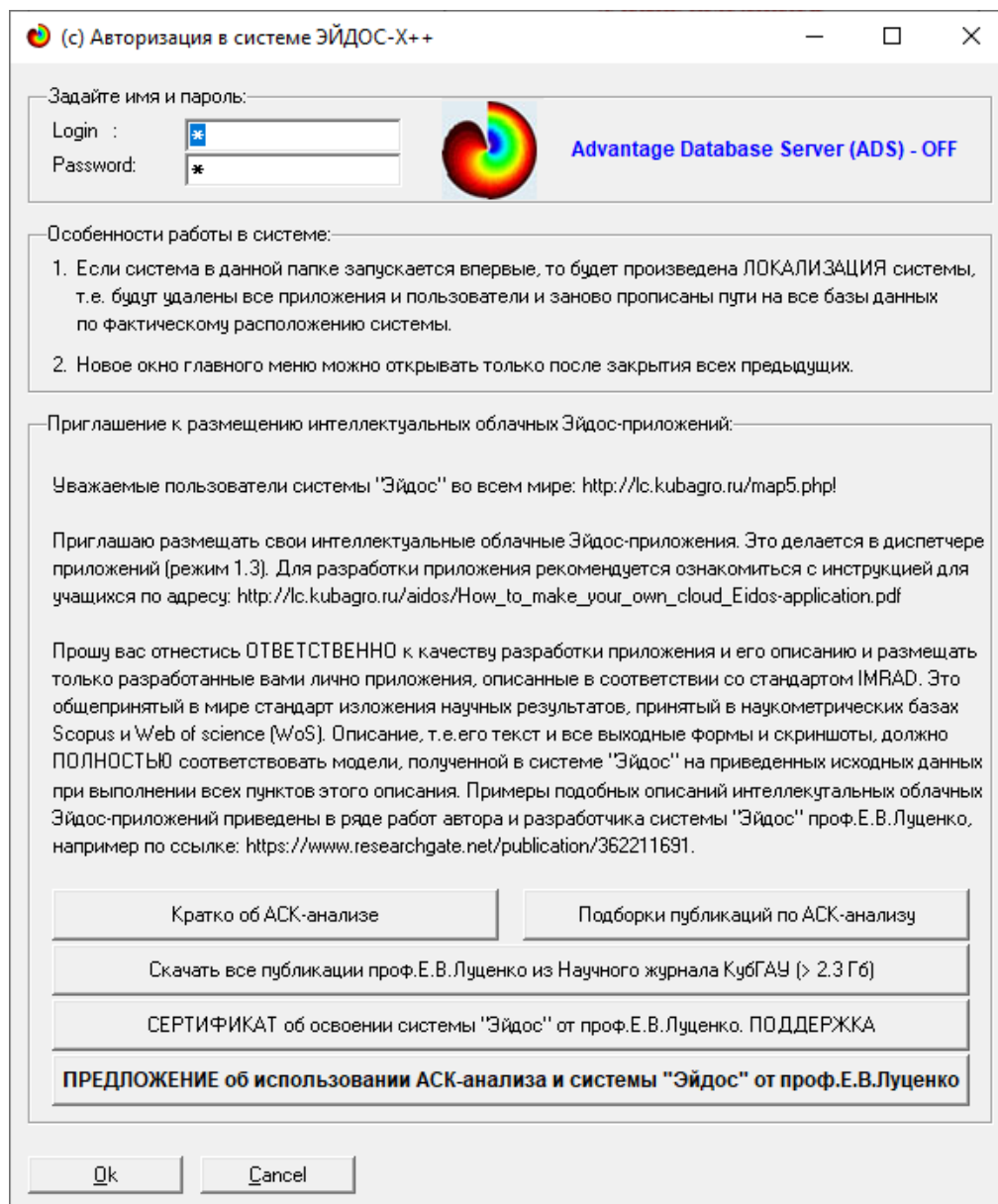


Рисунок 2. Титульные видеogramмы текущей версии системы «Эйдос»

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

Достижение поставленной цели в АСК-анализе обеспечивается решением следующих *задач* и подзадач, которые получаются в результате декомпозиции цели и являются *этапами* ее достижения:

Задача-1. Когнитивная структуризация предметной области.

Задача-2. Формализация предметной области.

Задача-3. Синтез статистических и системно-когнитивных моделей.

Многопараметрическая типизация и частные критерии знаний.

Задача-4. Верификация моделей.

Задача-5. Выбор наиболее достоверной модели.

Задача-6. Системная идентификация и прогнозирование.

Задача-7. Поддержка принятия решений (упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ; развитый алгоритм принятия решений в АСК-анализе).

Задача-8 исследование объекта моделирования путем исследования его модели, *включает ряд подзадач:*

8.1) инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы);

8.2) кластерно-конструктивный анализ классов;

8.3) кластерно-конструктивный анализ значений описательных шкал;

8.4) модель знаний системы «Эйдос» и нелокальные нейроны;

8.5) нелокальная нейронная сеть;

8.6) 3d-интегральные когнитивные карты;

8.7) 2d-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения);

8.8) 2d-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения);

8.9) когнитивные функции;

8.10) значимость описательных шкал и их градаций;

8.11) степень детерминированности классов и классификационных шкал).

Для данной работы особое значение имеет решение подзадачи 8.1, т.к. она позволяет детально исследовать влияние каждого значения каждого фактора на устойчивость сливы к температурным стрессорам, в т.ч. от обработки деревьев борной кислотой.

На рисунке 3 приведена последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос»:

**Последовательность обработки данных, информации и знаний в системе «Эйдос»,
повышение уровня системности данных, информации и знаний,
повышение уровня системности моделей**

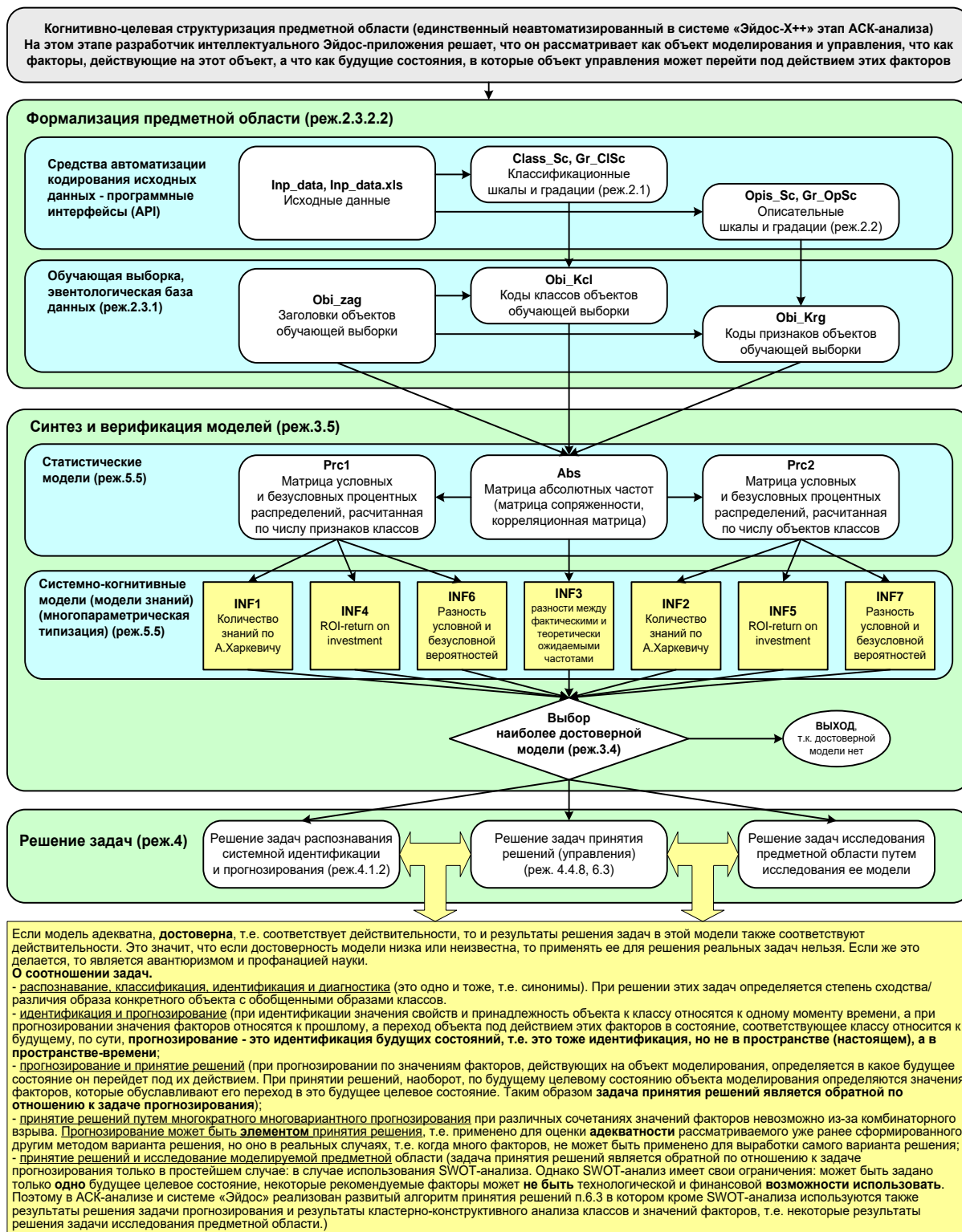


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос» (этапы АСК-анализа)

3. RESULTS (РЕЗУЛЬТАТЫ)

3.1. Задача-1. Когнитивная структуризация предметной области. Две интерпретации классификационных и описательных шкал и градаций

3.1.1. Универсальная постановка задачи, не зависящая от предметной области

На этапе когнитивно-целевой структуризации предметной области мы неформализуемым путем решаем на качественном уровне, что будем рассматривать в качестве объекта моделирования, что в качестве факторов, действующих на моделируемый объект (причин), а что в качестве результатов действия этих факторов (последствий). По сути это постановка решаемой проблемы.

Описательные шкалы служат для формального описания факторов, а классификационные – результатов их действия на объект моделирования. Шкалы могут быть числовые и текстовые. Текстовые шкалы могут быть номинальные и порядковые.

Когнитивная структуризация предметной области является первым и единственным неавтоматизированным в системе «Эйдос» этапом АСК-анализа, т.е. все последующие этапы АСК анализа в ней полностью автоматизированы.

В АСК-анализе и системе «Эйдос» применяется две интерпретации классификационных и описательных шкал и градаций: *статичная и динамичная* и соответствующая терминология (обобщающая, статичная и динамичная). Есть также обобщающая интерпретация и соответствующая ей терминология.

Статичная интерпретация и терминология:

- градации классификационных шкал – это обобщающие категории видов объектов (классы);
- описательные шкалы – свойства объектов, градации описательных шкал – значения свойств (признаки) объектов.

Динамичная интерпретация и терминология:

- градации классификационных шкал – это обобщающие категории будущих состояний объекта моделирования (классы), описывающие результаты действия факторов на объект моделирования в натуральном и стоимостном выражении: например количество и качество продукции, прибыль и рентабельность;
- описательные шкалы – факторы, действующие на объект моделирования, градации описательных шкал – значения факторов, действующие на объект моделирования.

Обобщающая терминология:

- классификационные шкалы и градации;
- описательные шкалы и градации.

3.1.2. Конкретное решение задачи в данной работе

В данной работе в качестве *объекта моделирования* выступает вода, в качестве *факторов*: рН, жесткость, твердые вещества, хлорамины, сульфаты, электропроводность, органический углерод, тригалометаны, мутность (таблица 1), а в качестве *результатов* действия этих факторов: пригодность для питья (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC	NAME_OPSC
1	pH
2	Hardness
3	Solids
4	Chloramines
5	Sulfate
6	Conductivity
7	Organic carbon
8	Trihalomethanes
9	Tubidity

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Opis_Sc.xlsx

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC	NAME_CLSC
1	Potability

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Class_Sc.xlsx

Для формирования xlsx-файлов, приведенных в таблицах 1 и 2, необходимо выполнить в системе "Эйдос" режим 5.12.

3.2. Задача-2. Формализация предметной области

3.2.1. Универсальная постановка задачи, не зависящая от предметной области

На этапе формализации предметной области разрабатываются классификационные и описательные шкалы и градации, а затем исходные данные кодируются с их использованием, в результате чего получается обучающая выборка. Обучающая выборка, по сути, представляет собой исходные данные, *нормализованные* с помощью классификационных и описательных шкал и градаций. Этим самым подготавливается все необходимое для выполнения следующего этапа АСК-анализа: синтеза и верификации статистических и системно-когнитивных моделей.

В системе «Эйдос» имеется большое количество разнообразных автоматизированных программных интерфейсов (API), обеспечивающих ввод в систему внешних данных различных типов: текстовых, табличных и графических, а также других, которые могут быть представлены в этом виде, например аудио или данные электроэнцефалограммы (ЭЦГ) или кардиограммы (ЭКГ).

Этим обеспечивается возможность комфортного для пользователя применения системы «Эйдос» для проведения научных исследований и решения практических задач в самых различных предметных областях и направлениях науки, практически везде, где человек применяет естественный интеллект [47, 48].

3.2.2. Конкретное решение задачи в данной работе

В качестве *источника исходных данных* в данной работе используем таблицу 3:

Таблица 3 – Excel-таблица исходных данных в стандарте системы «Эйдос»

observation	ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability
1	7,128062255	191,4008206	25227,54515	6,607844327	319,2508585	419,3049802	12,86536617	78,46849123	4,510595264	1
2	7,896743922	197,6020777	9407,664039	9,610055886	400,2745792	270,7940588	11,30058731	53,20184836	4,889458166	1
3	5,949518764	160,4426313	16898,8083	6,04590644	367,3285417	451,0127884	16,3599511	62,36823392	4,072197622	0
4	8,170588045	208,2791813	12806,83233	5,489258691	322,9239977	472,0282857	13,38954876	28,56636437	3,995857049	0
5	6,513016586	185,3385398	26262,54548	6,742150215	289,0143482	410,1013089	12,9246531	83,65219038	3,729960916	0
6	10,53809844	200,1398287	13867,2442	7,365015093	409,4467298	340,8088228	17,07312308	72,73063948	3,823092676	1
7	9,418296122	173,6402158	8777,686784	8,4561779	358,2920642	596,8258366	16,78521507	105,2724558	4,604633154	0
8	6,682097359	228,3031856	7705,298605	5,94735035	363,5369835	254,3923058	13,14044046	64,78395359	3,44982947	1
9	7,285371191	169,14919	17375,97074	9,404687307	394,092424	373,1902821	15,26050986	80,79598527	3,694482047	0
10	6,571099217	207,741025	23372,88265	6,280195162	313,8469725	497,9023198	12,87480498	75,4167102	4,555949236	0
11	7,191024857	151,3066736	23781,32874	9,214409135	403,7884299	465,1268383	16,88421403	77,51544432	4,270645693	0
12	7,146976296	196,5626846	16911,20377	6,89050548	320,1008753	520,1113582	12,85423898	66,81417649	4,025762216	1
13	7,14394199	153,2363644	11553,03928	8,444687185	329,4044567	278,4832659	11,26688617	81,7184127	3,792840307	0
14	8,034011917	197,6093924	25204,67408	6,5010969	330,5521729	490,8215001	15,10728865	89,80094741	2,452120806	0
15	6,013160966	218,8432561	21573,74757	9,295852001	321,1683132	444,276635	14,74434737	62,44323905	3,455622541	0
16	6,791992649	210,4079665	30471,9291	6,61443797	326,585846	493,1794166	19,98555517	61,12014349	2,91405748	1
17	8,137767883	203,8618673	24172,70331	6,844017626	305,8325797	470,8362645	19,23296141	30,7084213	5,003825786	1
18	6,970036694	182,8885886	32037,60547	7,412765411	376,9519325	325,7923996	14,35124065	70,05272018	4,27528772	0
19	10,02622381	94,09130748	25510,0973	6,614114031	385,4026595	388,467266	11,88716125	69,27397869	3,736228471	1
20	7,440825031	183,3627126	17259,8523	4,610244692	335,6264426	452,9952928	9,700906283	80,53706529	2,496342612	0
21	9,879071258	208,9056452	14673,67067	6,240247621	366,8176388	442,3374864	19,62602199	66,81658139	4,218831154	0
22	7,871961403	219,6226589	12760,43715	8,076613047	360,0866575	350,8269163	15,69872552	84,15808045	2,938536287	1
23	4,808039747	186,1605703	19118,0503	7,054443859	359,4062362	577,4913708	8,304288382	73,15308872	3,597759969	0
24	6,597292363	191,7874419	25039,3547	7,29457739	395,7395287	501,5326526	14,69539103	80,0500311	4,318305407	1
25	8,272055091	204,2253291	24196,52268	8,448405225	332,8903825	500,6848732	15,15740752	71,02274371	5,140207743	0
26	6,227826795	169,9262216	22550,49226	4,973592706	297,4044974	460,4153864	20,78959171	93,3007329	4,007347485	1
27	9,769251006	264,2656543	8471,677381	8,349709767	359,2685426	353,1964755	15,04708691	52,64545853	4,636556195	0
28	6,36729853	183,789491	11619,7097	6,035221327	343,8280153	362,7754392	16,64369604	73,25922077	2,61476606	0
29	7,965337305	151,5418889	25274,6103	7,106043486	352,316182	527,6882417	15,79262497	52,26848971	3,390982111	0
30	6,897922717	173,3909282	16616,51662	6,352933531	329,3397676	337,4360877	12,44697563	60,15941277	4,490105279	1
31	8,070477127	198,8659477	18266,61772	6,90236983	396,6195096	376,7103037	10,92446056	87,3795608	3,055790475	0
32	6,810466517	209,7355588	32602,34008	7,422781037	341,1742694	340,0300886	16,73674859	42,34946082	4,402339513	0
33	7,035036795	220,5018877	19746,49325	7,059382856	371,9126148	536,1839	15,30343324	67,90175839	4,557028866	1
34	9,894562129	128,5857327	26436,03333	8,778717494	349,7484457	480,6676165	9,105812815	65,32732965	4,61753288	0
35	4,349439999	157,5133609	24352,0661	5,96772234	287,0451104	432,642307	10,66725445	44,28901265	4,442711502	0
36	10,2684305	218,2166825	13753,99337	7,512608141	334,1780747	341,6022087	15,67156353	84,68266023	4,304689262	1

Таблица 3 имеет следующую структуру:

- каждая строка описывает одно наблюдение с определенным сочетанием значений факторов и сведениях о пригодности воды для питья;
- каждое *наблюдение* описывается одновременно *двумя* способами: с одной стороны значениями факторов, действующих на объект моделирования (числовые переменные, градации описательных шкал, бесцветный фон в таблице 3), а с другой стороны результатами действия этих факторов, т.е. жизнеспособностью пыльцы сливы, выраженными в текстовых и числовых шкалах в разных единицах измерения (желтый фон).

Такая структура описания наблюдений в технологиях искусственного интеллекта называется «*онтологией*» и модели представлений знаний Марвина Мински (1975) называется «*фрейм-экземпляр*»;

– 1-я колонка – не является шкалой и содержит номер наблюдения или другую идентифицирующую информацию о том, откуда взято описание этого наблюдения;

– колонка 11 – это классификационная шкала – это шкала *числового* типа описывающие *результаты* действия факторов в различных единицах измерения (таблица 3), в данном случае пригодность воды для питья. В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в *натуральном* и *стоимостном* выражении [8]: например *количество* и *качество* продукции, *прибыль* и *рентабельность*. В системе «Эйдос» существует не очень жесткое ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

– колонки с 2-й по 10-ю – это описательные шкалы, формализующие факторы, действующие на объект моделирования (таблица 4). Эти шкалы имеют числовой тип и их градациями являются числовые переменные;

– при вводе данных в систему «Эйдос» нули и пробелы в исходных данных могут рассматриваться как значащие или как отсутствие данных. 2-й вариант и будет использован в данной работе.

Математическая модель системы «Эйдос» обеспечивает устойчивое выявление в сопоставимой форме силы и направления причинно-следственных зависимостей в неполных (фрагментированных) зашумленных взаимозависимых (нелинейных) данных очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

Таким образом, система «Эйдос» не предъявляет чрезмерно жестких практически невыполнимых требований к исходным данным, а обрабатывает те данные, которые реально есть, например подобные представленным в таблице 3.

В системе «Эйдос» есть 6 основных автоматизированных программных интерфейсов (API), обеспечивающих ввод в систему и интеллектуальную обработку числовых, текстовых и графических данных, представленных в виде таблиц и файлов. Возможна обработка и других видов данных (например, данных землетрясений, ЭЭГ, ЭКГ, аудио и видео), которые возможно представить в этих форматах (рисунок 4).

2.3.2. Программные интерфейсы с внешними базами данных	2.3.2.1. Импорт данных из текстовых файлов
	2.3.2.2. Универсальный программный интерфейс импорта данных в систему
	2.3.2.3. Импорт данных из транспонированных внешних баз данных
	2.3.2.4. Оцифровка изображений по внешним контурам
	2.3.2.5. Оцифровка изображений по всем пикселям и спектру
	2.3.2.6. Сценарный АСК-анализ символьных и числовых рядов
	2.3.2.7. Транспонирование файлов исходных данных
	2.3.2.8. Объединение нескольких файлов исходных данных в один
	2.3.2.9. Разбиение TXT-файла на файлы-абзацы
	2.3.2.10. CSV => DBF конвертер системы "Эйдос"
	2.3.2.11. Прогноз событий по астропараметрам по Н.А.Чердниченко
	2.3.2.12. Прогнозирование землетрясений методом Н.А.Чердниченко
	2.3.2.13. Чемпионат RAIF-Challenge 2017-API-bank
	2.3.2.14. Чемпионат RAIF-Challenge 2017-API-retail
	2.3.2.15. Вставка промежуточных строк в файл исходных данных Inp_data

Рисунок 4. Автоматизированные программные интерфейсы (API) системы «Эйдос»

Для ввода исходных данных, представленных в таблице 3, в систему «Эйдос», используется один из ее автоматизированных программных интерфейсов (API), а именно универсальный автоматизированный программный интерфейс ввода данных из файлов MS Excel (API-2.3.2.2).

Требования API-2.3.2.2 к исходным данным подробно описаны в хелпе этого режима (рисунки 5):

Помощь по режиму 2.3.2.2 для случая Excel-файлов исходных данных

Режим 2.3.2.2. Универсальный программный интерфейс импорта данных из внешней базы данных "Inp_data.xls" в систему "Эйдос" и формализации предметной области.

Данный программный интерфейс обеспечивает формализацию предметной области, т.е. анализ файла исходных данных (Inp_data.xls), формирование классификационных и описательных шкал и градаций, а затем кодирование файла исходных с использованием:

- Файл исходных данных должен иметь имя: Inp_data.xls(а), а файл распознаваемой выборки имеет Inp_rasp.xls(а). Файлы Inp_data.xls(а) и Inp_rasp.xls(а) должны находиться в папке: \AIDOS\K\AID_DATA\Inp_data\
- Та строка этого файла должна содержать наименование столбца на любом языке, в т.ч. и русском. Это наименование должно быть во всех столбцах, при этом переписано по своему расширению, а объединение ячеек, разрыв строки, абзаца не допускается. Эти наименования должны быть короткими, но понятными, т.е. они будут в выводе на экран, а к ним еще будут добавляться наименования градаций. В числовых шкалах надо ОБЯЗАТЕЛЬНО указывать единицы измерения и число знаков после запятой в колонке должно быть 0 ДИНАМИЧЕСКОЕ.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длиннее, до 255 символов.
- Каждая строка этого файла, начиная со 2-й, содержит данные об одном объекте обучающей выборки или одном наблюдении. В MS Excel 2003 в листе может быть до 65536 строк, и до 256 колонок. В листе MS Excel 2010 и более поздних возможно до 1048576 строк и 16384 колонок.
- Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (именованного / порядкового) или числового типа (дискретные значения после запятой).
- Столбцы присваиваются числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. пробелом), то столбцу присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
- Столбцы со 2-го по N-й являются классификационными шкалами (возможны и параметрами) и содержат данные о классах (состоянии объектов управления), к которым принадлежит объект обучающей выборки.
- Столбцы с N+1 по последний являются описательными шкалами (свойствами или факторами) и содержат данные о признаках (т.е. значениях свойств или значений факторов), характеризующих объект обучающей выборки.
- В результате работы режима формируется файл INP_NAME.TXT стандарта MS DOS (виртуалка), в котором наименования классификационных и описательных шкал являются СТРОКАМИ. Система формирует классификационные и описательные шкалы и градации. Для этого в каждой числовой строке система находит минимальное и максимальное числовые значения и формирует заданное количество числовых интервалов, после чего числовые значения заменяются на интервальные значениями. В текстовых столбцах система находит уникальные текстовые значения. В каждой УНИКАЛЬНОЕ интервальное числовое или текстовое значение считается градацией классификационной или описательной шкалы, характеризующей объект. В каждой шкале ее градации сортируются по алфавиту. С использованием шкал и градаций кодируются исходные данные в результате чего генерируется обучающая выборка, каждый объект которой соответствует одной строке файла исходных данных INP_DATA и содержит классы классов, соответствующие фактам совпадения числовых или уникальных текстовых значений классов с градациями классификационных шкал и коды признаков, соответствующие фактам совпадения числовых или уникальных текстовых значений признаков с градациями описательных шкал.
- Распознаваемая выборка формируется на основе файла INP_RASP аналогично, за исключением того, что классификационные и описательные шкалы и градации не создаются, а используются ранее созданные в модели, и базы распознаваемой выборки могут не включать коды классов, если столбцы классов в файле INP_RASP были пустыми. Структура файла INP_RASP должна быть такой же, как INP_DATA, т.е. они должны ПОЛНОСТЬЮ совпадать по наименованиям столбцов, но могут иметь разное количество строк с разными значениями в них.

Принцип организации таблицы исходных данных:

Наименование объекта обучающей выборки	Наименование 1-й классификационной шкалы	Наименование 2-й классификационной шкалы	...	Наименование 1-й описательной шкалы	Наименование 2-й описательной шкалы	...
1-й объект обучающей выборки (1-е наблюдение)	Значение шкалы	Значение шкалы	...	Значение шкалы	Значение шкалы	...
2-й объект обучающей выборки (2-е наблюдение)	Значение шкалы	Значение шкалы	...	Значение шкалы	Значение шкалы	...
...

Определения основных терминов и профилактика типичных ошибок при подготовке Excel-файла исходных данных

Рисунок 5. Хелп API-2.3.2.2 системы «Эйдос»

Экранные формы управления API-2.3.2.2 системы «Эйдос» с *реальными параметрами*, использованными в данной работе, приведены на рисунках 6:

2.3.2.2. Универсальный программный интерфейс импорта данных в систему "ЭЙДОС-Х++"

Автоматическая формализация предметной области: генерация классификационных и описательных шкал и градаций, а также обучающей и распознаваемой выборки на основе базы исходных данных: "Inp_data"

Задайте тип файла исходных данных: "Inp_data":

XLS - MS Excel-2003 Стандарт XLS-файла
 XLSX- MS Excel-2007(2010) Стандарт XLSX-файла
 DBF - DBASE IV (DBF/NTX) Стандарт DBF-файла
 CSV - CSV => DBF конвертер Стандарт CSV-файла

Задайте параметры:

Нули и пробелы считать ОТСУТСТВИЕМ данных
 Нули и пробелы считать ЗНАЧЕНИЯМИ данных
 Создавать БД средних по классам "Inp_davr.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:

Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:

Конечный столбец описательных шкал:

Задайте режим:

Формализации предметной области (на основе "Inp_data")
 Генерации распознаваемой выборки (на основе "Inp_rasp")

Задайте способ выбора размера интервалов:

Равные интервалы с разным числом наблюдений
 Разные интервалы с равным числом наблюдений

Задание параметров формирования сценариев или способа интерпретации текстовых полей "Inp_data":

Не применять сценарный метод АСК-анализа Применить сценарный метод АСК-анализа
 Применить спец. интерпретацию текстовых полей классов Применить спец. интерпретацию текстовых полей признаков

Параметры интерпретации значений текстовых полей "Inp_data":

Интерпретация TXT-полей классов:

Значения полей текстовых классификационных шкал файла исходных данных "Inp_data" рассматриваются как целое

Интерпретация TXT-полей признаков:

Значения полей текстовых описательных шкал файла исходных данных "Inp_data" рассматриваются как целое

Какие наименования ГРАДАЦИЙ числовых шкал использовать:

Только интервальные числовые значения (например: "1/3-{59873.0000000, 178545.6666667}")
 Только наименования интервальных числовых значений (например: "Минимальное")
 И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-{59873.0000000, 178545.6666667}")

2.3.2.2. Задание размерности модели системы "ЭЙДОС-Х++"

ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (равные интервалы)

Количество градаций классификационных и описательных шкал в модели, т.е.: [2 классов x 45 признаков]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных	Среднее количество градаций на класс. шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис. шкалу
Числовые	1	2	2,00	9	45	5,00
Текстовые	0	0	0,00	0	0	0,00
ВСЕГО:	1	2	2,00	9	45	5,00

Задайте количество числовых диапазонов (интервалов, градаций) в шкале:

В классификационных шкалах: В описательных шкалах:

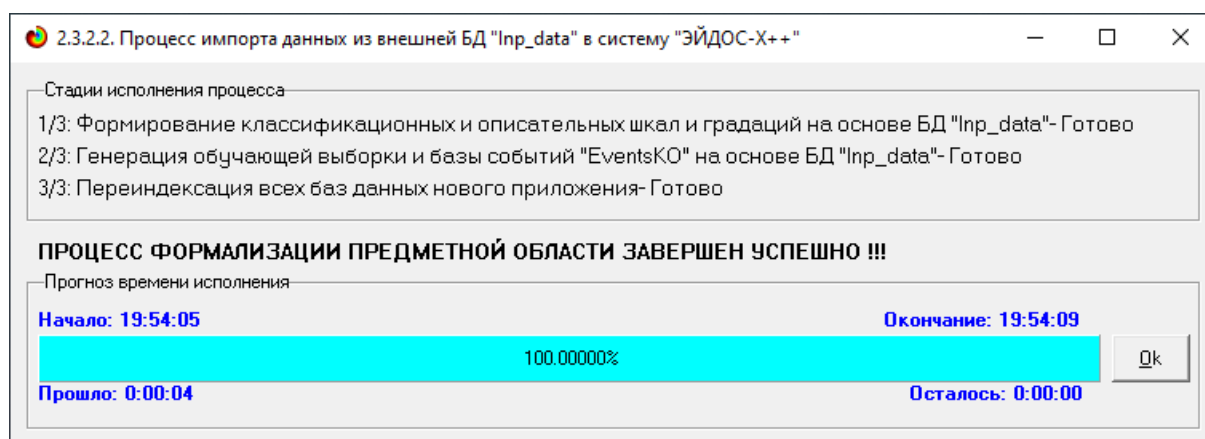


Рисунок 6. Экранные формы управления API-2.3.2.2 системы «Эйдос»

В таблицах 4, 5, 6 приведены классификационные и описательные шкалы и градации, а также обучающая выборка, сформированные API-2.3.2.2 при параметрах, показанных на рисунках 6.

Отметим, что суммарное количество градаций на 2-м рисунке 6 и в таблице 6 может не совпадать, если в некоторых описательных шкалах есть градации «Пробел» или нули, которые в соответствии с параметрами на 1-м рисунке 6 рассматриваются не как значащие, а как *отсутствие данных*.

Под несбалансированностью данных понимается неравномерность распределения значений свойств объекта моделирования или действующих на него факторов по диапазону изменения значений числовых шкал и между шкалами, как числовыми, так и текстовыми. Математическая модель АСК-анализа позволяет корректно преодолеть несбалансированность данных путем перехода от абсолютных частот к относительным и к количественным мерам знаний в системно-когнитивных моделях (мы увидим это ниже).

Таблица 4 – Классификационные шкалы и градации (полностью)

KOD_CLS	GRAD_CLS
1	Малое
2	Большое

Таблица 5 – Описательные шкалы и градации (полностью)

KOD_ATR	GRAD_ATR
1	Очень малое
2	Малое
3	Среднее
4	Большое
5	Очень большое

Таблица 6 – Обучающая выборка

NAME_O BJ	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11
1	3	8	13	18	22	28	33	38	43	2
2	3	9	11	20	24	26	32	37	44	2
3	2	8	12	18	23	28	34	38	43	1
4	3	9	12	17	23	28	33	36	43	1
5	2	8	13	18	22	28	33	39	43	1
6	5	9	12	18	24	27	34	38	43	2
7	4	8	11	19	23	30	34	40	44	1
8	3	9	11	18	23	26	33	38	42	2
9	3	8	12	20	24	27	33	39	43	1

Источник: c:\Aidos-X\AID_DATA\A0000001\System\EventsKO.xlsx

Отметим, что в системе «Эйдос» обычно используются базы данных с расширением «dbf». Они открываются в MS Excel или могут быть конвертированы в файлы xlsx с помощью онлайн-сервисов или в режиме 5.12 (этот режим системы «Эйдос» написан на Питоне).

3.3. Задача-3. Синтез статистических и системно-когнитивных моделей. Многопараметрическая типизация и частные критерии знаний

3.3.1. Универсальная постановка задачи, не зависящая от предметной области

Математические модели, на основе которых рассчитываются статистические и системно-когнитивные модели автоматизированного системно-когнитивного анализа (АСК-анализ) и интеллектуальной системы «Эйдос», подробно описаны в ряде монографий и статей автора [1-4]. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко, акцентируя внимание лишь на математической взаимосвязи коэффициента возврата инвестиций (ROI) с мерой χ -квадрат Карла Пирсона и с семантической мерой целесообразности информации Александра Харкевича.

Отметим, что модели системы «Эйдос» основаны на матрице абсолютных частот, отражающей число встреч градаций описательных шкал по градациям классификационных шкал (фактов). Но для решения всех задач используется не непосредственно сама эта матрица, а матрицы условных и безусловных процентных распределений и системно-когнитивные модели, которые рассчитываются на ее основе и отражают какое количество информации содержится в факте наблюдения определенной градации описательной шкалы о том, что объект моделирования перейдет в состояние, соответствующее определенной градации классификационной шкалы (классу).

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике [1-7] и обеспечивает сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых (нелинейных) данных, представленных в различных типах шкал (дихотомических, номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных рассчитывается матрица абсолютных частот (таблица 7):

Таблица 7 – Матрица абсолютных частот (статистическая модель ABS)

		Классы					Сумма
		1	...	j	...	W	
Значения факторов	1	N_{11}		N_{1j}		N_{1W}	
	...						
	i	N_{i1}		N_{ij}		N_{iW}	$N_{i\Sigma} = \sum_{j=1}^W N_{ij}$
	...						
	M	N_{M1}		N_{Mj}		N_{MW}	
Суммарное количество признаков по классу				$N_{\Sigma j} = \sum_{i=1}^M N_{ij}$			$N_{\Sigma\Sigma} = \sum_{i=1}^W \sum_{j=1}^M N_{ij}$
Суммарное количество объектов обучающей выборки по классу				$N_{\Sigma j}$			$N_{\Sigma\Sigma} = \sum_{j=1}^W N_{\Sigma j}$

На основе таблицы 7 рассчитываются матрицы условных и безусловных процентных распределений (таблица 8).

Отметим, что в АСК-анализе и его программном инструментарии интеллектуальной системе «Эйдос» используется два способа расчета матриц условных и безусловных процентных распределений:

1-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество признаков по классу;

2-й способ: в качестве $N_{\Sigma j}$ используется суммарное количество объектов обучающей выборки по классу.

Таблица 8 – Матрица условных и безусловных процентных распределений (статистические модели PRC1 и PRC2)

		Классы					Безусловная вероятность признака
		1	...	j	...	w	
Значения факторов	1	P_{11}		P_{1j}		P_{1w}	
	...						
	i	P_{i1}		$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$		P_{iw}	$P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
	...						
	M	P_{M1}		P_{Mj}		P_{Mw}	
Безусловная вероятность класса				$P_{\Sigma j}$			

На практике часто встречается существенная *несбалансированность* данных, под которой понимается сильно отличающееся количество наблюдений объектов обучающейся выборки, относящихся к различным градациям одной классификационной или описательной шкалы. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 7) было бы очень неразумно (за исключением меры взаимосвязи хи-квадрат) и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) (таблица 8) является весьма обоснованным и логичным.

Этот переход полностью снимает проблему *несбалансированности* данных, т.к. в последующем анализе используется не матрица абсолютных частот (таблица 7), а матрицы условных и безусловных процентных распределений (таблица 8), а также матрицы системно-когнитивных моделей, рассчитываемые на основе матрицы абсолютных частот и матрицы условных и безусловных процентных распределений. Этот подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (дихотомических, номинальных, порядковых и числовых) и в разных единицах измерения [8]. В системе «Эйдос» этот подход применяется всегда при решении любых задач.

Затем на основе таблиц 7 и 8 с использованием частных критериев, знаний приведенных таблице 9, рассчитываются матрицы 7 системно-когнитивных моделей (таблица 10).

В таблице 9 приведены формулы:

– для сравнения фактических и теоретических абсолютных частот;

– для сравнения условных и безусловных относительных частот («вероятностей»).

И это **сравнение** в таблицах 7 и 8 осуществляется двумя возможными способами: путем **вычитания** и путем **деления**.

Таблица 9– Различные аналитические формы частных критериев знаний, применяемые в АСК-анализе и системе «Эйдос»

Наименование модели знаний и частный критерий	Выражение для частного критерия	
	Через относительные частоты	Через абсолютные частоты
ABS , матрица абсолютных частот, N_{ij} - фактическое число встреч i -го признака у объектов j -го класса; \bar{N}_{ij} - теоретическое число встреч i -го признака у объектов j -го класса; N_i – суммарное количество признаков в i -й строке; N_j – суммарное количество признаков или объектов обучающей выборки в j -м классе; N – суммарное количество признаков по всей выборке (таблица 7)	N_{ij} – фактическая частота; $N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $\bar{N}_{ij} = \frac{N_i N_j}{N}$ – теоретическая частота.	
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	---	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество объектов обучающей выборки по классу	---	
INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу. Вероятность того, что если у объекта j -го класса обнаружен признак, то это i -й признак	$I_{ij} = \Psi \times \log_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \log_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \log_2 \frac{N_{ij} N}{N_i N_j}$
INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу. Вероятность того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак.		
INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами	---	$I_{ij} = N_{ij} - \bar{N}_{ij} = N_{ij} - \frac{N_i N_j}{N}$
INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	$I_{ij} = \frac{N_{ij}}{\bar{N}_{ij}} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		
INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N} = \frac{N_{ij} N - N_i N_j}{N_j N}$
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		

Обозначения к таблице:

i – значение прошлого параметра;

j – значение будущего параметра;

N_{ij} – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра;

M – суммарное число значений всех прошлых параметров;

W – суммарное число значений всех будущих параметров.

N_i – количество встреч i -м значения прошлого параметра по всей выборке;

N_j – количество встреч j -го значения будущего параметра по всей выборке;

N – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра по всей выборке.

I_{ij} – частный критерий знаний: количество знаний в факте наблюдения i -го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее j -му значению будущего параметра;

Ψ – нормировочный коэффициент (Е.В.Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;

P_i – безусловная относительная частота встречи i -го значения прошлого параметра в обучающей выборке;

P_{ij} – условная относительная частота встречи i -го значения прошлого параметра при j -м значении будущего параметра.

Таблица 10 – Матрица системно-когнитивной модели

		Классы					Значимость фактора
		1	...	j	...	W	
Значения факторов	1	I_{11}		I_{1j}		I_{1W}	$\sigma_{1\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{1j} - \bar{I}_1)^2}$
	...						
	i	I_{i1}		I_{ij}		I_{iW}	$\sigma_{i\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
	...						
M	I_{M1}		I_{Mj}		I_{MW}	$\sigma_{M\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{Mj} - \bar{I}_M)^2}$	
Степень редукции класса	$\sigma_{\Sigma 1}$		$\sigma_{\Sigma j}$		$\sigma_{\Sigma W}$	$H = \sqrt{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$	

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 9), применяемых в настоящее время в системе «Эйдос» равно 7 определяется тем, что они получаются путем всех возможных вариантов сравнения фактических и теоретических абсолютных частот, условных и безусловных относительных частот путем вычитания и путем деления, и при этом N_j рассматривается как суммарное количество или признаков, или объектов обучающей выборки в j -м классе, а **нормировка к нулю** (для аддитивных интегральных критериев), если нет связи между наличием признака и принадлежностью объекта к классу, осуществляется либо логарифмированием, либо вычитанием единицы (таблица 11).

Обратим особое внимание на то, что сравнение фактических и теоретических абсолютных частот путем деления приводит при нормировках к нулю (что нужно для применения аддитивных интегральных критериев) путем взятия логарифма и путем вычитания 1 к **тем же самым** моделям, что и сравнение условных и безусловных относительных частот путем деления с теми же самыми способами нормировки.

Таблица 11– Конфигуратор системно-когнитивных моделей АСК-анализа и интеллектуальной системы «Эйдос»

	Способ сравнения	Нормировка не требуется	Нормировка к 0 путем взятия логарифма	Нормировка к 0 путем вычитания 1
Сравнение фактических и теоретических абсолютных частот	Путем деления	---	INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
	Путем вычитания	INF3, χ -квадрат Карла Пирсона	---	---
Сравнение условных и безусловных относительных частот	Путем деления	---	INF1, INF2, Александра Харкевича	INF4, INF5, Коэффициент возврата инвестиций ROI
	Путем вычитания	INF6, INF7	---	---

Таким образом, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и провести нормировку к 0 путем взятия логарифма и путем вычитания 1, то получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Других же системно-когнитивных моделей, рассчитываемых на основе приведенных статистических моделей просто нет. Это и есть конфигуратор статистических и когнитивных моделей в смысле В.А.Лефевра. *Под конфигуратором В.А.Левефр понимал минимальный полный набор понятийных шкал или конструктов, т.е. понятий, достаточный для адекватного описания предметной области* [4]⁹. Необходимо отметить, что все эти модели рассчитываются в интеллектуальной системе «Эйдос».

Когда мы сравниваем фактические и теоретические абсолютные частоты путем вычитания у нас получается частный критерий знаний: «хи-квадрат» (СК-модель INF3), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2) или «коэффициент возврата инвестиций ROI» - Return On Investment (СК-модели INF4, INF5) в зависимости от способа нормировки.

Когда же мы сравниваем условные и безусловные относительные частоты путем вычитания у нас получается частный критерий знаний: «коэффициент взаимосвязи» (СК-модели INF6, INF7), когда же мы сравниваем их путем деления, то у нас получается частный критерий: «количество информации по А.Харкевичу» (СК-модели INF1, INF2).

9 См. 1.2.1.2.1.1. Определение понятия конфигуратора, http://lc.kubagro.ru/aidos/aidos06_lec/index.htm

Таким образом, мы видим, что все частные критерии знаний тесно взаимосвязаны друг с другом. Особенно интересна связь знаменитого критерия хи-квадрат К.Пирсона с замечательной мерой количества информации А.Харкевича и с известным в экономике коэффициентом ROI.

Вероятность рассматривается как предел, к которому стремится относительная частота (отношение количества благоприятных исходов к числу испытаний) при неограниченном увеличении количества испытаний. Ясно, что вероятность – это математическая абстракция, которая никогда не встречается на практике (также как и другие математические и физические абстракции, типа математической точки, материальной точки, бесконечно малой и т.п.). На практике встречается только относительная частота. Но она может быть весьма близкой к вероятности. Например, при 480 наблюдений различие между относительной частотой и вероятностью (погрешность) составляет около 5%, при 1250 наблюдениях – около 2.5%, при 10000 наблюдениях – 1%.

Суть этих методов в том, что вычисляется количество информации в значении фактора о том, что объект моделирования перейдет под его действием в определенное состояние, соответствующее классу. Это позволяет сопоставимо и корректно обрабатывать разнородную информацию о наблюдениях объекта моделирования, представленную в различных типах измерительных шкал и различных единицах измерения [8].

На основе системно-когнитивных моделей, представленных в таблице 10 (отличаются частыми критериями, приведенными в таблице 9), решаются задачи идентификации (классификации, распознавания, диагностики, прогнозирования), поддержки принятия решений (обратная задача прогнозирования), а также задача исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Отметим, что как значимость значения фактора, степень детерминированности класса и ценность или качество модели в АСК-анализе рассматривается вариабельность значений частных критериев этого значения фактора, класса или модели в целом (таблица 12).

Численно эта вариабельность может измеряться разными способами, например средним отклонением модулей частных критериев от среднего, дисперсией или среднеквадратичным отклонением или его квадратом. В системе «Эйдос» принят последний вариант, т.к. эта величина совпадает с мощностью сигнала, в частности мощностью информации, а в АСК-анализе все модели рассматриваются как источник информации об объекте моделирования. Поэтому есть все основания уточнить традиционную терминологию АСК-анализа (таблица 12):

Таблица 12 – Уточнение терминологии АСК-анализа

№	Традиционные термины (синонимы)	Новый термин	Формула
1	1. Значимость значения фактора (признака). 2. Дифференцирующая мощность значения фактора (признака). 3. Ценность значения фактора (признака) для решения задачи идентификации и других задач	Корень из информационной мощности значения фактора	$\sigma_{i\Sigma} = \sqrt[2]{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
2	1. Степень детерминированности класса. 2. Степень обусловленности класса.	Корень из информационной мощности класса	$\sigma_{\Sigma j} = \sqrt[2]{\frac{1}{M-1} \sum_{i=1}^M (I_{ij} - \bar{I}_j)^2}$
3	1. Качество модели. 2. Ценность модели. 3. Степень сформированности модели. 4. Количественная мера степени выраженности закономерностей в моделируемой предметной области	Корень из информационной мощности модели	$H = \sqrt[2]{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

Итак, в разделе раскрывается простая Математическая взаимосвязь меры χ -квадрат Карла Пирсона с коэффициентом возврата инвестиций (ROI) и с семантической мерой целесообразности информации Александра Харкевича. Эта взаимосвязь обнаруживается, если на основе матрицы абсолютных частот рассчитать матрицы условных и безусловных процентных распределений, а затем сравнить фактические абсолютные частоты с теоретическими путем вычитания и деления, а также сравнить условные и безусловные относительные частоты также путем вычитания и деления и выполнить нормировку к нулю путем взятия логарифма или вычитания 1. При этом получается 3 статистических модели: матрица абсолютных частот и две матрицы относительных частот, т.е. условных и безусловных процентных распределений, а также всего 7 системно-когнитивных моделей. Именно 7, а не большее количество системно-когнитивных моделей в итоге получается потому, что модели, получающиеся в результате сравнения фактических и теоретических абсолютных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1 **тождественно совпадают** с моделями, получающимися путем сравнения условных и безусловных относительных частот путем деления и нормировки к нулю путем взятия логарифма или вычитания 1. Это и есть конфигурактор статистических и когнитивных моделей в смысле В.А.Лефевра, содержащий минимальное количество моделей, позволяющих полно описать моделируемую предметную область.

Показательно, примечательно и весьма замечательно, что модель меры χ -квадрат Карла Пирсона из **статистики** оказалась математически тесно связанной с коэффициентом возврата инвестиций

(ROI), применяемой в экономике в теории управления портфелем инвестиций и с мерой информации Александра Харкевича из семантической теории информации и теории управления знаниями. Все эти модели рассчитываются в интеллектуальной системе «Эйдос».

3.3.2. Конкретное решение задачи в данной работе

В системе «Эйдос» синтез моделей производится в режиме 3.5 (рисунок 7):

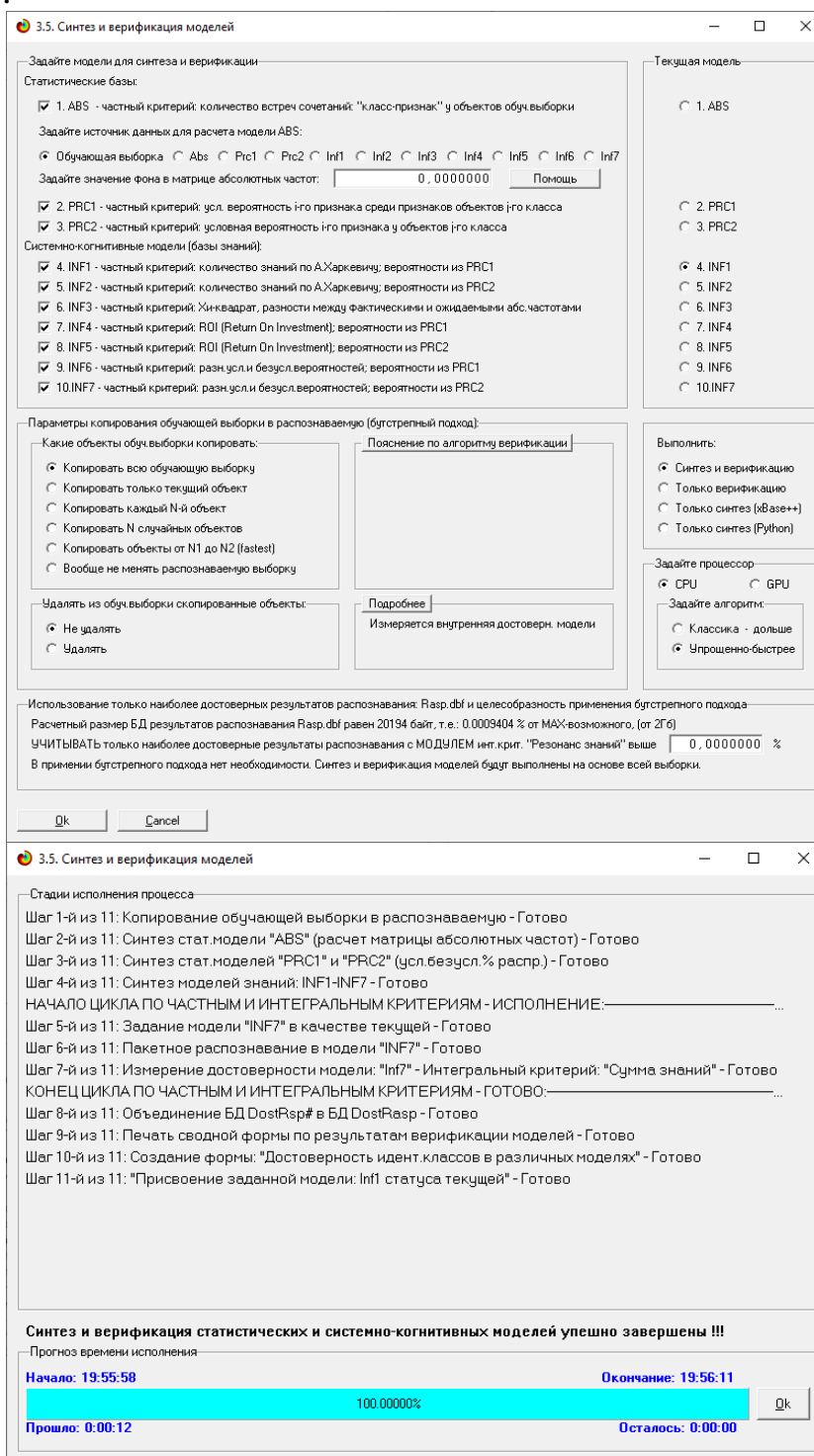


Рисунок 7. Экранные формы режима синтеза и верификации моделей

В результате работы режима 3.5 создано 3 статистических и 7 системно-когнитивных моделей, некоторые из которых приведены на рисунках 8-11:

5.5. Модели: "1. ABS - частный критерий: количество встреч сочетаний "Класс-признак" у объектов обучающей выборки"

Код признака	Наименование описательной шкалы и градации	1. ROTABILITY МАЛОЕ	2. ROTABILITY БОЛЬШОЕ	Средн.	Среднее	Средн. квадрат. откл.
1.0	RH-Очень малое	14.0	2.0	16.0	8.00	8.49
2.0	RH-Малое	36.0	21.0	57.0	28.50	10.61
3.0	RH-Среднее	50.0	47.0	97.0	48.50	2.12
4.0	RH-Большое	14.0	3.0	17.0	8.50	7.78
5.0	RH-Очень большое	7.0	6.0	13.0	6.50	0.71
6.0	HARDNESS-Очень малое	1.0	3.0	4.0	2.00	1.41
7.0	HARDNESS-Малое	14.0	7.0	21.0	10.50	4.95
8.0	HARDNESS-Среднее	46.0	24.0	70.0	35.00	15.56
9.0	HARDNESS-Большое	48.0	37.0	85.0	42.50	7.78
10.0	HARDNESS-Очень большое	12.0	8.0	20.0	10.00	2.83
11.0	SOLIDS-Очень малое	21.0	9.0	30.0	15.00	8.49
12.0	SOLIDS-Малое	51.0	34.0	85.0	42.50	12.02
13.0	SOLIDS-Среднее	33.0	25.0	58.0	29.00	5.66
14.0	SOLIDS-Большое	10.0	10.0	20.0	10.00	
15.0	SOLIDS-Очень большое	6.0	1.0	7.0	3.50	3.54
16.0	CHLORAMINES-Очень малое	2.0	2.0	4.0	2.00	
17.0	CHLORAMINES-Малое	15.0	11.0	26.0	13.00	2.83
18.0	CHLORAMINES-Среднее	67.0	36.0	103.0	51.50	21.92
19.0	CHLORAMINES-Большое	30.0	24.0	54.0	27.00	4.24
20.0	CHLORAMINES-Очень большое	7.0	6.0	13.0	6.50	0.71
21.0	SULFATE-Очень малое	5.0	3.0	8.0	4.00	1.41
22.0	SULFATE-Малое	34.0	26.0	60.0	30.00	5.66
23.0	SULFATE-Среднее	63.0	32.0	95.0	47.50	21.92
24.0	SULFATE-Большое	16.0	15.0	31.0	15.50	0.71
25.0	SULFATE-Очень большое	3.0	3.0	6.0	3.00	
26.0	CONDUCTIVITY-Очень малое	9.0	7.0	16.0	8.00	1.41
27.0	CONDUCTIVITY-Малое	37.0	27.0	64.0	32.00	7.07
28.0	CONDUCTIVITY-Среднее	41.0	24.0	65.0	32.50	12.02
29.0	CONDUCTIVITY-Большое	31.0	16.0	47.0	23.50	10.61
30.0	CONDUCTIVITY-Очень большое	2.0	5.0	7.0	3.50	2.12
31.0	ORGANIC_CARBOON-Очень малое	5.0	3.0	8.0	4.00	1.41
32.0	ORGANIC_CARBOON-Малое	25.0	15.0	40.0	20.00	7.07
33.0	ORGANIC_CARBOON-Среднее	47.0	32.0	79.0	39.50	10.61
34.0	ORGANIC_CARBOON-Большое	35.0	19.0	54.0	27.00	11.31
35.0	ORGANIC_CARBOON-Очень большое	9.0	10.0	19.0	9.50	0.71
36.0	TRIHALOMETHANES-Очень малое	6.0	2.0	8.0	4.00	2.83
37.0	TRIHALOMETHANES-Малое	29.0	19.0	48.0	24.00	7.07
38.0	TRIHALOMETHANES-Среднее	58.0	38.0	96.0	48.00	14.14
39.0	TRIHALOMETHANES-Большое	26.0	18.0	44.0	22.00	5.66
40.0	TRIHALOMETHANES-Очень большое	2.0	2.0	4.0	2.00	
41.0	TURBIDITY-Очень малое	6.0	1.0	7.0	3.50	3.54

Рисунок 8. Статистическая модель «ABS», матрица абсолютных частот (фрагмент)

5.5. Модели: "3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса"

Код признака	Наименование описательной шкалы и градации	1. ROTABILITY МАЛОЕ	2. ROTABILITY БОЛЬШОЕ	Безусл. вероятн.	Среднее	Средн. квадрат. откл.
1.0	RH-Очень малое	11.570	2.532	8.000	7.051	6.447
2.0	RH-Малое	29.752	26.582	28.500	28.167	2.297
3.0	RH-Среднее	41.322	59.494	48.500	50.408	12.904
4.0	RH-Большое	11.570	3.797	8.500	7.684	5.552
5.0	RH-Очень большое	5.785	7.595	6.500	6.690	1.334
6.0	HARDNESS-Очень малое	0.826	3.797	2.000	2.312	2.156
7.0	HARDNESS-Малое	11.570	8.861	10.500	10.216	1.971
8.0	HARDNESS-Среднее	38.017	30.380	35.000	34.198	5.456
9.0	HARDNESS-Большое	39.669	46.835	42.500	43.252	5.122
10.0	HARDNESS-Очень большое	9.917	10.127	10.000	10.022	0.194
11.0	SOLIDS-Очень малое	17.355	11.392	15.000	14.374	4.272
12.0	SOLIDS-Малое	42.149	43.038	42.500	42.593	0.680
13.0	SOLIDS-Среднее	27.273	31.646	29.000	29.459	3.147
14.0	SOLIDS-Большое	8.264	12.658	10.000	10.461	3.162
15.0	SOLIDS-Очень большое	4.959	1.266	3.500	3.112	2.666
16.0	CHLORAMINES-Очень малое	1.653	2.532	2.000	2.092	0.675
17.0	CHLORAMINES-Малое	12.397	13.924	13.000	13.160	1.134
18.0	CHLORAMINES-Среднее	55.372	45.570	51.500	50.471	6.987
19.0	CHLORAMINES-Большое	24.793	30.380	27.000	27.587	4.005
20.0	CHLORAMINES-Очень большое	5.785	7.595	6.500	6.690	1.334
21.0	SULFATE-Очень малое	4.132	3.797	4.000	3.965	0.287
22.0	SULFATE-Малое	28.099	32.911	30.000	30.505	3.458
23.0	SULFATE-Среднее	52.066	40.506	47.500	46.286	8.230
24.0	SULFATE-Большое	13.223	18.987	15.500	16.105	4.131
25.0	SULFATE-Очень большое	2.479	3.797	3.000	3.138	0.986
26.0	CONDUCTIVITY-Очень малое	7.438	8.861	8.000	8.149	1.060
27.0	CONDUCTIVITY-Малое	30.579	34.177	32.000	32.378	2.599
28.0	CONDUCTIVITY-Среднее	33.884	30.380	32.500	32.132	2.534
29.0	CONDUCTIVITY-Большое	25.620	20.253	23.500	22.936	3.850
30.0	CONDUCTIVITY-Очень большое	1.653	6.329	3.500	3.991	3.362
31.0	ORGANIC_CARBOON-Очень малое	4.132	3.797	4.000	3.965	0.287
32.0	ORGANIC_CARBOON-Малое	20.661	18.987	20.000	19.824	1.238
33.0	ORGANIC_CARBOON-Среднее	38.843	40.506	39.500	39.675	1.229
34.0	ORGANIC_CARBOON-Большое	28.926	24.051	27.000	26.488	3.503
35.0	ORGANIC_CARBOON-Очень большое	7.438	12.658	9.500	10.048	3.746
36.0	TRIHALOMETHANES-Очень малое	4.959	2.532	4.000	3.745	1.771
37.0	TRIHALOMETHANES-Малое	23.967	24.051	24.000	24.009	0.091
38.0	TRIHALOMETHANES-Среднее	47.934	48.101	48.000	48.018	0.154
39.0	TRIHALOMETHANES-Большое	21.488	22.785	22.000	22.136	0.970
40.0	TRIHALOMETHANES-Очень большое	1.653	2.532	2.000	2.092	0.675
41.0	TURBIDITY-Очень малое	4.959	1.266	3.500	3.112	2.666

Рисунок 9. Статистическая модель «PRC2», матрица условных и безусловных процентных распределений (фрагмент)

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу, вероятности из PRС1"

Код признака	Наименование описательной шкалы и градации	1. ROTABILITY МАЛОЕ	2. ROTABILITY БОЛЬШОЕ	Сумма	Среднее	Средн. квадрат. откл.
1.0	PH-Очень малое	0.049	-0.154	-0.104	-0.052	0.143
2.0	PH-Малое	0.006	-0.009	-0.004	-0.002	0.011
3.0	PH-Среднее	-0.021	0.027	0.006	0.003	0.034
4.0	PH-Большое	0.041	-0.108	-0.066	-0.033	0.105
5.0	PH-Очень большое	-0.015	0.021	0.005	0.003	0.026
6.0	HARDNESS-Очень малое	-0.118	0.095	-0.022	-0.016	0.144
7.0	HARDNESS-Малое	0.013	-0.023	-0.010	-0.005	0.025
8.0	HARDNESS-Среднее	0.011	-0.019	-0.008	-0.004	0.021
9.0	HARDNESS-Большое	-0.009	0.013	0.004	0.002	0.016
10.0	HARDNESS-Очень большое	-0.001	0.002	0.001	0.000	0.002
11.0	SOLIDS-Очень малое	0.020	-0.037	-0.017	-0.009	0.040
12.0	SOLIDS-Малое	-0.001	0.002	0.001	0.000	0.002
13.0	SOLIDS-Среднее	-0.008	0.012	0.003	0.002	0.014
14.0	SOLIDS-Большое	-0.025	0.031	0.006	0.003	0.040
15.0	SOLIDS-Очень большое	0.047	-0.136	-0.089	-0.045	0.129
16.0	CHLORAMINES-Очень малое	-0.025	0.031	0.006	0.003	0.040
17.0	CHLORAMINES-Малое	-0.006	0.009	0.003	0.001	0.011
18.0	CHLORAMINES-Среднее	0.010	-0.016	-0.007	-0.003	0.018
19.0	CHLORAMINES-Большое	-0.011	0.016	0.004	0.002	0.019
20.0	CHLORAMINES-Очень большое	-0.015	0.021	0.005	0.003	0.026
21.0	SULFATE-Очень малое	0.004	-0.007	-0.003	-0.001	0.008
22.0	SULFATE-Малое	-0.009	0.012	0.004	0.002	0.015
23.0	SULFATE-Среднее	0.012	-0.021	-0.009	-0.005	0.024
24.0	SULFATE-Большое	-0.021	0.027	0.006	0.003	0.034
25.0	SULFATE-Очень большое	-0.025	0.031	0.006	0.003	0.040
26.0	CONDUCTIVITY-Очень малое	-0.010	0.014	0.004	0.002	0.016
27.0	CONDUCTIVITY-Малое	-0.006	0.009	0.003	0.001	0.010
28.0	CONDUCTIVITY-Среднее	0.006	-0.009	-0.003	-0.002	0.010
29.0	CONDUCTIVITY-Большое	0.012	-0.020	-0.008	-0.004	0.022
30.0	CONDUCTIVITY-Очень большое	-0.100	0.079	-0.021	-0.011	0.127
31.0	ORGANIC_CARBON-Очень малое	0.004	-0.007	-0.003	-0.001	0.008
32.0	ORGANIC_CARBON-Малое	0.004	-0.007	-0.003	-0.001	0.008
33.0	ORGANIC_CARBON-Среднее	-0.002	0.003	0.001	0.001	0.004
34.0	ORGANIC_CARBON-Большое	0.009	-0.016	-0.006	-0.003	0.017
35.0	ORGANIC_CARBON-Очень большое	-0.033	0.038	0.006	0.003	0.050
36.0	TRIHALOMETHANES-Очень малое	0.029	-0.061	-0.032	-0.016	0.064
37.0	TRIHALOMETHANES-Малое	0.000	0.000	0.000	0.000	0.000
38.0	TRIHALOMETHANES-Среднее	0.000	0.000	0.000	0.000	0.000
39.0	TRIHALOMETHANES-Большое	-0.003	0.005	0.002	0.001	0.005
40.0	TRIHALOMETHANES-Очень большое	-0.025	0.031	0.006	0.003	0.040
41.0	TURBIDITY-Очень малое	0.047	-0.136	-0.089	-0.045	0.129

Рисунок 10. Системно-когнитивная модель «INF1», матрица информантов (по А.Харкевичу) (фрагмент)

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс.частотами"

Код признака	Наименование описательной шкалы и градации	1. ROTABILITY МАЛОЕ	2. ROTABILITY БОЛЬШОЕ	Сумма	Среднее	Средн. квадрат. откл.
1.0	PH-Очень малое	4.324	-4.324			6.114
2.0	PH-Малое	1.528	-1.528			2.160
3.0	PH-Среднее	-8.664	8.664			12.252
4.0	PH-Большое	3.719	-3.719			5.259
5.0	PH-Очень большое	-0.862	0.862			1.219
6.0	HARDNESS-Очень малое	-1.419	1.419			2.007
7.0	HARDNESS-Малое	1.300	-1.300			1.838
8.0	HARDNESS-Среднее	3.665	-3.665			5.184
9.0	HARDNESS-Большое	-3.406	3.406			4.817
10.0	HARDNESS-Очень большое	-0.096	0.096			0.135
11.0	SOLIDS-Очень малое	2.857	-2.857			4.040
12.0	SOLIDS-Малое	-0.406	0.406			0.575
13.0	SOLIDS-Среднее	-2.077	2.077			2.938
14.0	SOLIDS-Большое	-2.096	2.096			2.964
15.0	SOLIDS-Очень большое	1.767	-1.767			2.498
16.0	CHLORAMINES-Очень малое	-0.419	0.419			0.593
17.0	CHLORAMINES-Малое	-0.724	0.724			1.024
18.0	CHLORAMINES-Среднее	4.708	-4.708			6.658
19.0	CHLORAMINES-Большое	-2.658	2.658			3.759
20.0	CHLORAMINES-Очень большое	-0.862	0.862			1.219
21.0	SULFATE-Очень малое	0.162	-0.162			0.229
22.0	SULFATE-Малое	-2.287	2.287			3.234
23.0	SULFATE-Среднее	5.546	-5.546			7.843
24.0	SULFATE-Большое	-2.748	2.748			3.887
25.0	SULFATE-Очень большое	-0.629	0.629			0.889
26.0	CONDUCTIVITY-Очень малое	-0.676	0.676			0.957
27.0	CONDUCTIVITY-Малое	-1.706	1.706			2.413
28.0	CONDUCTIVITY-Среднее	1.689	-1.689			2.389
29.0	CONDUCTIVITY-Большое	2.575	-2.575			3.642
30.0	CONDUCTIVITY-Очень большое	-2.233	2.233			3.159
31.0	ORGANIC_CARBON-Очень малое	0.162	-0.162			0.229
32.0	ORGANIC_CARBON-Малое	0.809	-0.809			1.144
33.0	ORGANIC_CARBON-Среднее	-0.778	0.778			1.100
34.0	ORGANIC_CARBON-Большое	2.342	-2.342			3.312
35.0	ORGANIC_CARBON-Очень большое	-2.491	2.491			3.523
36.0	TRIHALOMETHANES-Очень малое	1.162	-1.162			1.643
37.0	TRIHALOMETHANES-Малое	-0.029	0.029			0.042
38.0	TRIHALOMETHANES-Среднее	-0.059	0.059			0.083
39.0	TRIHALOMETHANES-Большое	-0.610	0.610			0.863
40.0	TRIHALOMETHANES-Очень большое	-0.419	0.419			0.593
41.0	TURBIDITY-Очень малое	1.767	-1.767			2.498

Рисунок 11. Системно-когнитивная модель «INF3», матрица Хи-квадрат (по К.Пирсону) (фрагмент)

Полученные модели корректно использовать для решения задач только в том случае, если они достаточно достоверны (адекватны), т.е. верно отражают моделируемую предметную область. Поэтому в следующем разделе оценим достоверность созданных статистических и системно-когнитивных моделей.

3.4. Задача-4. Верификация моделей

3.4.1. Универсальная постановка задачи, не зависящая от предметной области

Оценка достоверности моделей в системе «Эйдос» осуществляется путем решения задачи классификации объектов обучающей выборки по обобщенным образам классов и подсчета количества истинных и ложных положительных и отрицательных решений по F-мере Ван Ризбергена, а также по критериям L1- L2-мерам проф. Е.В.Луценко, которые предложены для того, чтобы смягчить или полностью преодолеть некоторые недостатки F-меры [9].

Достоверность моделей можно оценивать и путем решения других задач, например задач прогнозирования, выработки управляющих решений, исследования объекта моделирования путем исследования его модели. Но это более трудоемко и даже всегда возможно, особенно на экономических и политических моделях.

В режиме 3.4 системы «Эйдос» и ряде других изучается достоверность каждой частной модели в соответствии с этими мерами достоверности.

3.4.2. Конкретное решение задачи в данной работе

В соответствии с критерием достоверности L1-мерой проф.Е.В.Луценко наиболее достоверной является СК-модель INF1 (хи-квадрат К.Пирсона) с интегральным критерием: «Сумма знаний»: $L1=0.766$ при максимуме 1 (рисунок 12). *Эту модель и будем использовать для решения поставленных в работе задач.*

Наименование модели и частного критерия	Интегральный критерий	но- мый N)	Точность модели	Полнота модели	F-мера Ван Ризбергена	Сумма модел. уровней сход. истинно-поло- решений (STP)	Сумма модел. уровней сход. истинно-отриц. решений (STN)	Сумма модел. уровней сход. ложно-полож. решений (SFP)	Сумма модел. уровней сход. ложно-отриц. решений (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В.Луценко	Сред- нее ист. реш.
1. ABS - частный критерий: количество встреч сометаний "Клас...	Корреляция абс частот с обр...	2	0.501	0.990	0.666	104.001	0.120	98.296	0.043	0.514	1.000	0.679	
1. ABS - частный критерий: количество встреч сометаний "Клас...	Сумма абс частот по призна...	2	0.500	1.000	0.667	127.751		114.931		0.526	1.000	0.690	
2. PRC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл отн частот с о...	2	0.501	0.990	0.666	104.001	0.120	98.296	0.043	0.514	1.000	0.679	
2. PRC1 - частный критерий: усл. вероятность по призна...	Сумма усл отн частот по при...	2	0.500	1.000	0.667	148.305		145.401		0.505	1.000	0.671	
3. PRC2 - частный критерий: условная вероятность его признака...	Корреляция усл отн частот с о...	2	0.501	0.990	0.666	104.001	0.120	98.297	0.043	0.514	1.000	0.679	
3. PRC2 - частный критерий: условная вероятность по призна...	Сумма усл отн частот по при...	2	0.500	1.000	0.667	148.360		145.482		0.505	1.000	0.671	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	43	0.602	0.785	0.681	49.796	26.609	23.281	7.179	0.681	0.874	0.766	
4. INF1 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	77	0.665	0.615	0.639	18.586	33.672	6.907	10.306	0.729	0.643	0.683	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Семантический резонанс зна...	43	0.602	0.785	0.681	49.803	26.609	23.287	7.178	0.681	0.874	0.766	
5. INF2 - частный критерий: количество знаний по А.Харкевичу; в...	Сумма знаний	77	0.665	0.615	0.639	18.644	33.658	6.983	10.347	0.728	0.643	0.683	
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Семантический резонанс зна...	73	0.635	0.635	0.635	47.684	47.684	22.790	22.790	0.677	0.677	0.677	
6. INF3 - частный критерий: Хи-квадрат, разности между факти...	Сумма знаний	73	0.635	0.635	0.635	47.684	47.684	22.772	22.772	0.677	0.677	0.677	
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	81	0.661	0.595	0.626	32.894	48.079	10.476	18.114	0.758	0.645	0.697	
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	67	0.665	0.665	0.665	36.925	37.872	14.367	12.061	0.720	0.754	0.736	
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	81	0.661	0.595	0.626	32.900	48.077	10.480	18.113	0.758	0.645	0.697	
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	67	0.665	0.665	0.665	37.066	37.864	14.523	12.126	0.718	0.753	0.736	
9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер...	Семантический резонанс зна...	73	0.638	0.635	0.637	47.297	48.149	22.447	22.955	0.678	0.673	0.676	
9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер...	Сумма знаний	73	0.635	0.635	0.635	38.557	40.216	20.889	16.765	0.649	0.697	0.672	
10. INF7 - частный критерий: разн усл и безуслов вероятностей; ве...	Семантический резонанс зна...	74	0.636	0.630	0.633	47.193	48.039	22.472	22.981	0.677	0.673	0.675	
10. INF7 - частный критерий: разн усл и безуслов вероятностей; ве...	Сумма знаний	74	0.630	0.630	0.630	38.406	39.812	20.994	16.732	0.647	0.697	0.671	

Рисунок 12. Экранные формы режима измерения достоверности моделей 3.4

На рисунках 13 приведены частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко [9] СК-модели INF1.

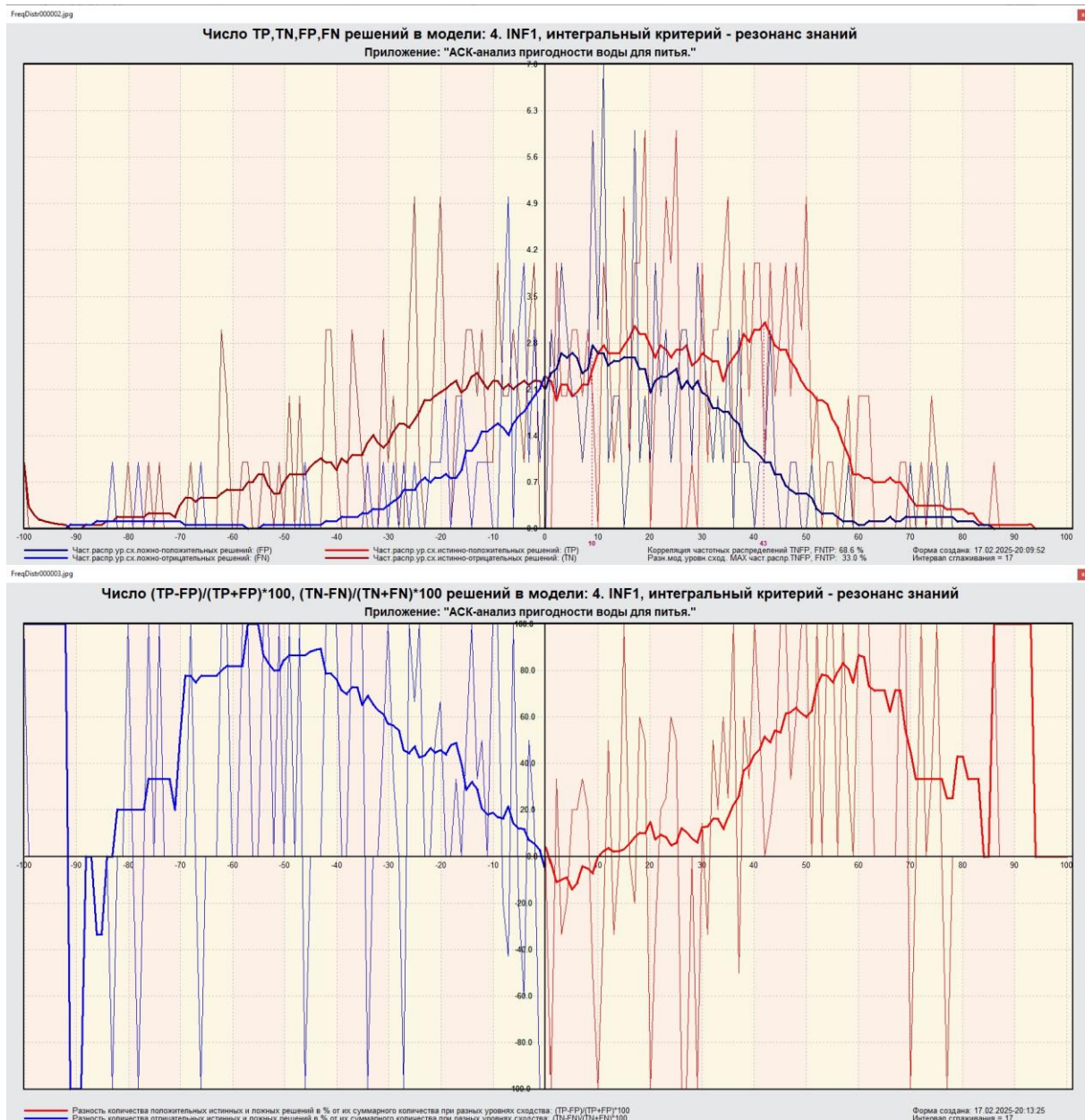


Рисунок 13. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF1 [9]

Из этих частотных распределений видно, что в наиболее достоверной по критерию достоверности L1-мерой проф.Е.В.Луценко СК-модели INF3:

– отрицательные ложные решения в данной задаче вообще не встречаются не встречаются, а в общем случае такие решения практически не встречаются за исключением очень небольшого количества случаев с очень низкими уровнями различия;

– при уровнях сходства меньше 20% в данной задаче преобладают ложные положительные решения, а при более высоких уровнях сходства – истинные положительные решения. В общем случае при уровнях сходства выше 60% ложных положительных решений практически вообще нет;

– *чем выше уровень сходства, тем больше доля истинных решений. Поэтому уровень сходства является адекватной внутренней мерой системы «Эйдос», так сказать адекватной мерой самооценки или аудита степени достоверности решений и уровня риска ошибочного решения.*

На рисунках 14 приведены экранные формы хелпов режима 3.4, в которых подробно объясняется смысл этого режима. Эти формы приводятся в данной работе вместо более детального описания данного режима.

Помощь по режимам: 3.4, 4.1.3.6, 4.1.3.7, 4.1.3.8, 4.1.3.10: Виды прогнозов и меры достоверности моделей в системе "Эйдос-Х++"

ПОЛОЖИТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.
Предположим, модель дает такой прогноз, что выпадет все: и 1, и 2, и 3, и 4, и 5, и 6. Понятно, что из всего этого выпадет лишь что-то одно. В этом случае модель не предскажет, что не выпадет, но зато она обязательно предскажет, что выпадет. Однако при этом очень много объектов будет отнесено к классам, к которым они не относятся. Тогда вероятность истинно-положительных решений у модели будет 1/6, а вероятность ложно-положительных решений - 5/6. Ясно, что такой прогноз бесполезен, поэтому он и назван мной псевдопрогнозом.

ОТРИЦАТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.
Представим себе, что мы выбрасываем кубик с 6 гранями, и модель предсказывает, что ничего не выпадет, т.е. не выпадет ни 1, ни 2, ни 3, ни 4, ни 5, ни 6, но что-то из этого, естественно, обязательно выпадет. Конечно, модель не предсказала, что выпадет, зато она очень хорошо предсказала, что не выпадет. Вероятность истинно-отрицательных решений у модели будет 5/6, а вероятность ложно-отрицательных решений - 1/6. Такой прогноз гораздо достовернее, чем положительный псевдопрогноз, но тоже бесполезен.

ИДЕАЛЬНЫЙ ПРОГНОЗ.
Если в случае с кубиком мы прогнозируем, что выпадет, например 1, и соответственно прогнозируем, что не выпадет 2, 3, 4, 5, и 6, то это идеальный прогноз, имеющий, если он осуществляется, 100% достоверность идентификации и не идентификации. Идеальный прогноз, который полностью снимает неопределенность о будущем состоянии объекта прогнозирования, на практике удается получить крайне редко и обычно мы имеем дело с реальным прогнозом.

РЕАЛЬНЫЙ ПРОГНОЗ.
На практике мы чаще всего сталкиваемся именно с этим видом прогноза. Реальный прогноз уменьшает неопределенность о будущем состоянии объекта прогнозирования, но не полностью, как идеальный прогноз, а оставляет некоторую неопределенность не снятой. Например, для игрального кубика делается такой прогноз: выпадет 1 или 2, и, соответственно, не выпадет 3, 4, 5 или 6. Понятно, что полностью на практике такой прогноз не может осуществиться, т.к. варианты выпадения кубика альтернативны, т.е. не может выпасть одновременно и 1, и 2. Поэтому у реального прогноза всегда будет определенная ошибка идентификации. Соответственно, если не осуществится один или несколько из прогнозируемых вариантов, то возникнет и ошибка не идентификации, т.к. это не прогнозировалось моделью. Теперь представьте себе, что у Вас не 1 кубик и прогноз его поведения, а тысячи. Тогда можно посчитать средневзвешенные характеристики всех этих видов прогнозов.

Таким образом, если просуммировать число верно идентифицированных и не идентифицированных объектов и вычесть число ошибочно идентифицированных и не идентифицированных объектов, а затем разделить на число всех объектов то это и будет критерий качества модели (классификатора), учитывающий как ее способность верно относить объекты к классам, которым они относятся, так и ее способность верно не относить объекты к тем классам, к которым они не относятся. Этот критерий предложен и реализован в системе "Эйдос" проф. Е.В. Луценко в 1994 году. Эта мера достоверности модели предполагает два варианта нормировки: {-1, +1} и {0, 1}:

$$L_a = \frac{TP + TN - FP - FN}{TP + TN + FP + FN} \quad (\text{нормировка: } \{-1, +1\})$$

$$L_b = \frac{1 + (TP + TN - FP - FN) / (TP + TN + FP + FN)}{2} \quad (\text{нормировка: } \{0, 1\})$$

где количество: TP - истинно-положительных решений; TN - истинно-отрицательных решений; FP - ложно-положительных решений; FN - ложно-отрицательных решений;

Классическая F-мера достоверности моделей Ван Ризбергера (колонка выделена ярко-голубым фоном):
 $F\text{-мера} = 2 * (Precision * Recall) / (Precision + Recall)$ - достоверность модели
Precision = TP / (TP + FP) - точность модели;
Recall = TP / (TP + FN) - полнота модели;

L1-мера проф. Е.В. Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СУММ уровней сходства (колонка выделена ярко-зеленым фоном):
 $L1\text{-мера} = 2 * (SPrecision * SRecall) / (SPrecision + SRecall)$
SPrecision = STP / (STP + SFP) - точность с учетом сумм уровней сходства;
SRecall = STP / (STP + SFN) - полнота с учетом сумм уровней сходства;
STP - Сумма модулей сходства истинно-положительных решений; STN - Сумма модулей сходства истинно-отрицательных решений;
SFP - Сумма модулей сходства ложно-положительных решений; SFN - Сумма модулей сходства ложно-отрицательных решений.

L2-мера проф. Е.В. Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СРЕДНИХ уровней сходства (колонка выделена желтым фоном):
 $L2\text{-мера} = 2 * (APrecision * ARecall) / (APrecision + ARecall)$
APrecision = ATP / (ATP + AFP) - точность с учетом средних уровней сходства;
ARecall = ATP / (ATP + AFN) - полнота с учетом средних уровней сходства;
ATP = STP / TP - Среднее модулей сходства истинно-положительных решений; AFN = SFN / FN - Среднее модулей сходства истинно-отрицательных решений;
AFP = SFP / FP - Среднее модулей сходства ложно-положительных решений; AFN = SFN / FN - Среднее модулей сходства ложно-отрицательных решений.

Строки с максимальными значениями F-меры, L1-меры и L2-меры выделены фоном цвета, соответствующего колонке.

Из графиков частотных распределений истинно-положительных, истинно-отрицательных, ложно-положительных и ложно-отрицательных решений видно, что чем выше модуль уровня сходства, тем больше доля истинных решений. Это значит, что модуль уровня сходства является адекватной мерой степени истинности решения и степени уверенности системы в этом решении. Поэтому система "Эйдос" имеет адекватный критерий достоверности собственных решений, с помощью которого она может отфильтровать заведомо ложные решения.

Луценко Е.В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСК-анализе и системе "Эйдос" / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета [Научный журнал КубГАУ] [Электронный ресурс]. - Краснодар: КубГАУ, 2017. - №02(126). С. 1 - 32. - IDA [article ID]: 1261702001. - Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>, 2 п.л.

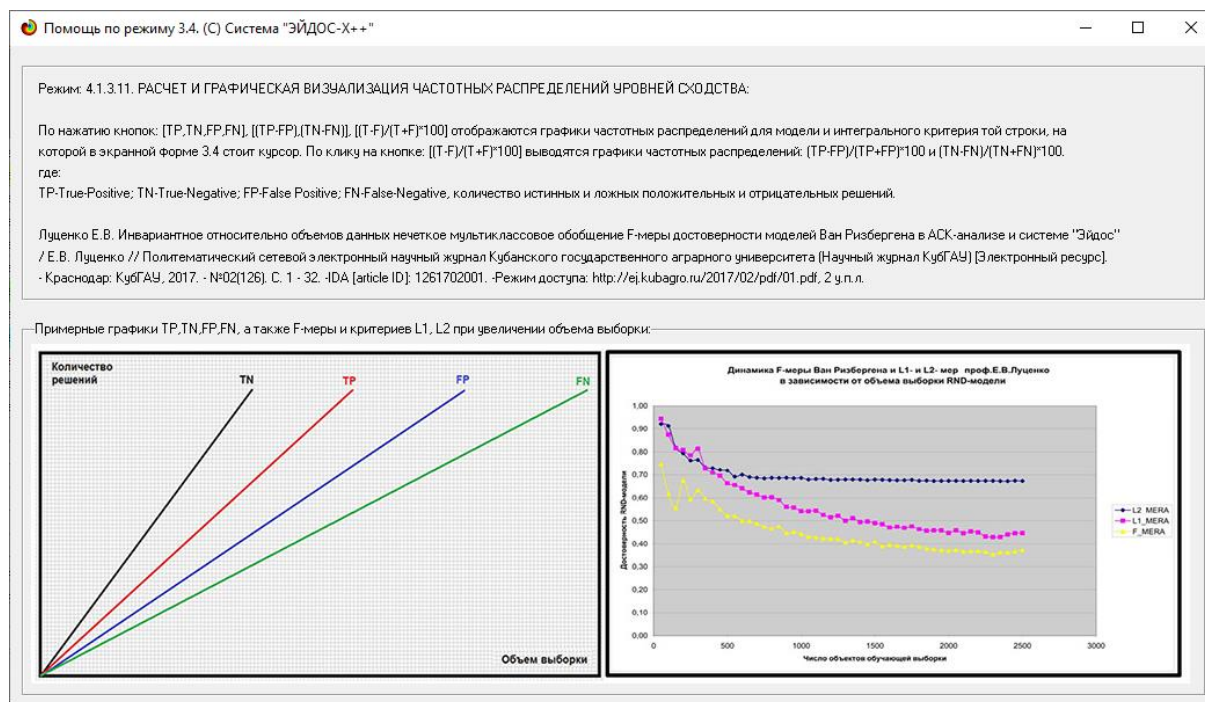


Рисунок 14. Экранные формы хелпов режима измерения достоверности моделей

3.5. Задача-5. Выбор наиболее достоверной модели

3.5.1. Универсальная постановка задачи, не зависящая от предметной области

Все последующие задачи решаются в наиболее достоверной модели.

Причины этого просты. Если модель достоверна, то:

- идентификация объекта с классом достоверна, т.е. модель относит объекты к классам, к которым они действительно принадлежат;
- прогнозирование достоверно, т.е. действительно наступают те события, которые прогнозируются;
- принятие решений адекватно (достоверно), т.е. после реализации принятых управляющих решений объект управления действительно переходит в целевые будущие состояния;
- исследование достоверно, т.е. полученные в результате исследования модели объекта моделирования выводы могут быть с полным основанием отнесены к объекту моделирования.

3.5.2. Конкретное решение задачи в данной работе

Технически сам выбор наиболее достоверной модели и задание ее текущей осуществляется в режиме 5.6 системы «Эйдос» и проходит быстро (рисунки 15). Это необходимо делать лишь для решения задачи идентификации и прогнозирования (в режиме 4.1.2), которая требует и потребляет наибольшие вычислительные ресурсы и поэтому решается только для модели, заданной в качестве текущей. Все остальные расчеты проводятся в системе «Эйдос» сразу во всех моделях.

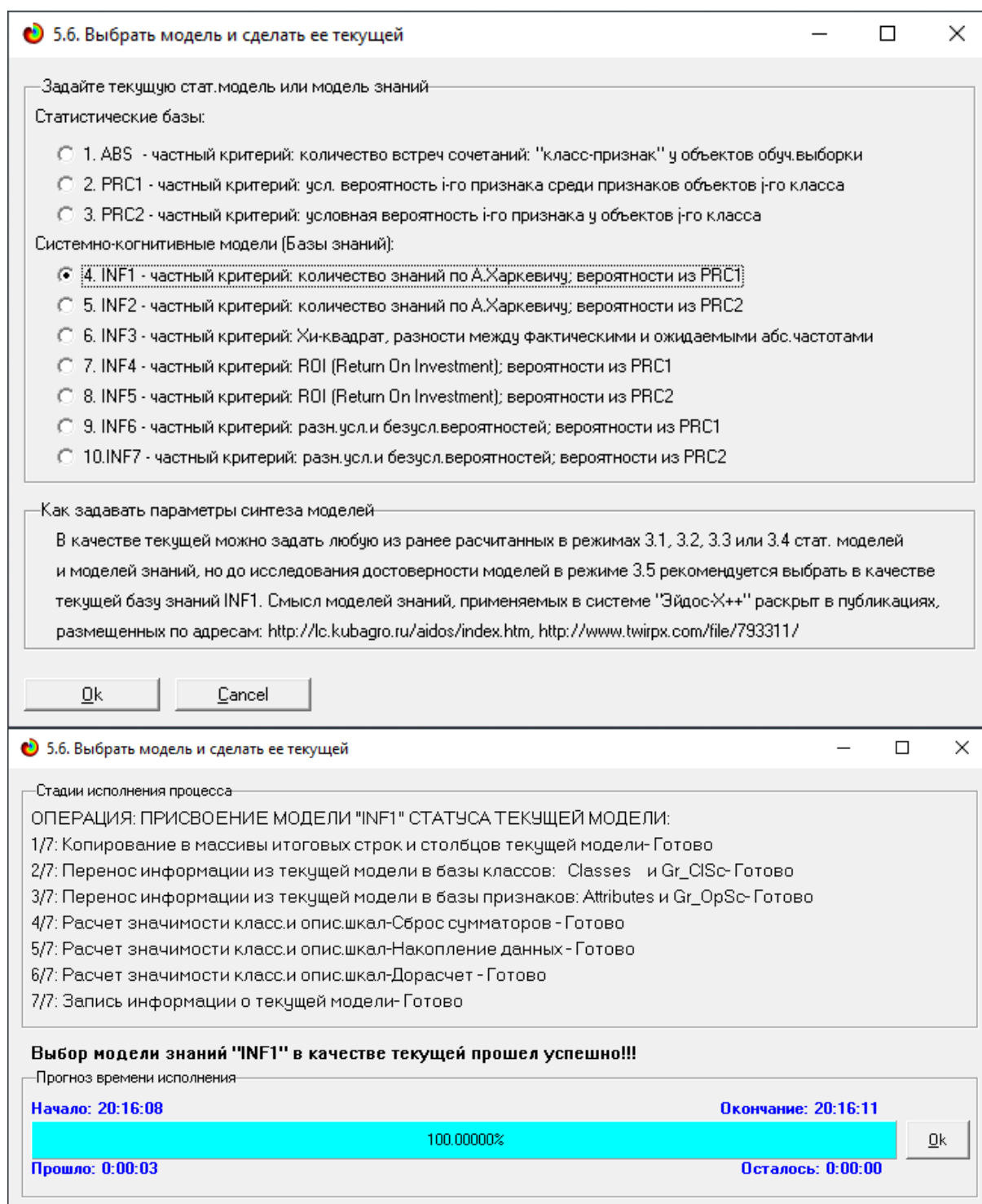


Рисунок 15. Задание СК-модели INF1 в качестве текущей

3.6. Задача-6. Системная идентификация и прогнозирование

3.6.1. Универсальная постановка задачи, не зависящая от предметной области

При решении задачи идентификации каждый объект распознаваемой выборки сравнивается по всем своим признакам с каждым из обобщенных образов классов. Смысл решения задачи идентификации заключается в

том, что при определении принадлежности конкретного объекта к обобщенному образу класса об этом конкретном объекте *по аналогии становится известно все, что известно об объектах этого класса, по крайней мере, самое существенное о них, т.е. чем они отличаются от объектов других классов.*

Задачи идентификации и прогнозирования взаимосвязаны и мало чем отличаются друг от друга. Главное различие между ними в том, что при идентификации значения свойств и принадлежность объекта к классу относятся к одному моменту времени, а при прогнозировании значения факторов относятся к прошлому, а переход объекта под действием этих факторов в состояние, соответствующее классу относится к будущему (рисунок 3).

Задача решается в модели, заданной в качестве текущей, т.к. является весьма трудоемкой в вычислительном отношении. Правда с разработкой и реализацией в системе «Эйдос» высокоэффективных алгоритмов распознавания и использованием графического процессора (GPU) для расчетов эта проблема практически снялась.

Сравнение осуществляется путем применения *неметрических интегральных критериев*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что корректны¹⁰ в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

3.6.1.1. Интегральный критерий «Сумма знаний»

Интегральный критерий «Сумма знаний» представляет собой суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help режима 5.5:

$$I_j = (\vec{I}_{ij}, \vec{L}_i).$$

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид:

$$I_j = \sum_{i=1}^M I_{ij} L_i,$$

где: M – количество градаций описательных шкал (признаков);

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j -го класса;

¹⁰ В отличие от Евклидова расстояния, которое используется для подобных целей наиболее часто

$\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, & \text{если } i\text{-й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i\text{-й фактор действует с истинностью } n; \\ 0, & \text{если } i\text{-й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-X++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n , если он присутствует у объекта с интенсивностью n , т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» – один раз).

3.6.1.2. Интегральный критерий «Семантический резонанс знаний»

Интегральный критерий «Семантический резонанс знаний»

представляет собой *нормированное* суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний, представленных в help режима 3.3 и имеет вид:

$$I_j = \frac{1}{\sigma_j \sigma_l M} \sum_{i=1}^M (I_{ij} - \bar{I}_j) (L_i - \bar{L}),$$

где:

M – количество градаций описательных шкал (признаков); \bar{I}_j – средняя информативность по вектору класса; \bar{L} – среднее по вектору объекта;

σ_j – среднеквадратичное отклонение частных критериев знаний вектора класса; σ_l – среднеквадратичное отклонение по вектору распознаваемого объекта.

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j -го класса; $\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта (состояния или явления), включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив–локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, & \text{если } i\text{-й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i\text{-й фактор действует с истинностью } n; \\ 0, & \text{если } i\text{-й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n , если он присутствует у объекта с интенсивностью n , т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

Приведенное выражение для интегрального критерия «Семантический резонанс знаний» получается непосредственно из выражения для критерия «Сумма знаний» после замены координат перемножаемых векторов их стандартизированными значениями:

$$I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}, \quad L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_l}.$$

Произведением двух стандартизированных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применения вейвлетов и сплайнов, в частности линейной интерполяции:

$$I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}, \quad L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}},$$

Это позволяет предложить неограниченное количество других видов интегральных критериев. Но результаты их применения едва ли они будут существенно отличаться от уже существующих, поэтому они не реализованы в системе «Эйдос».

3.6.1.3. Важные математические свойства интегральных критериев

Данные интегральные критерии обладают очень интересными *математическими свойствами*, которые обеспечивают ему важные достоинства:

Во-первых, интегральный критерий имеет **неметрическую** природу, т.е. он является мерой сходства векторов класса и объекта, но не расстоянием между ними, а косинусом угла между ними, т.е. это межвекторное или информационное расстояние. Поэтому его применение является корректным в **неортономрированных** пространствах, которые, как правило, и встречаются на практике и в которых применение Евклидова расстояния (теоремы Пифагора) является некорректным.

Во-вторых, данный интегральный критерий является **фильтром**, подавляющим белый шум, который всегда присутствует в эмпирических исходных данных и в моделях, созданных на их основе. Это свойство подавлять белый шум проявляется у данного критерия тем ярче, чем больше в модели градаций описательных шкал.

В-третьих, интегральный критерий сходства представляет собой количественную меру сходства/различия конкретного объекта с обобщенным образом класса и имеет тот же смысл, что и **функция принадлежности** элемента множеству в нечеткой логике Лотфи Заде. Однако в нечеткой логике эта функция задается исследователем априорно путем выбора из нескольких возможных вариантов, а в АСК-анализе и его программном инструментарии – интеллектуальной системе «Эйдос» она

рассчитывается в соответствии с хорошо обоснованной математической моделью непосредственно на основе эмпирических данных.

В-четвертых, кроме того значение интегрального критерия сходства представляет собой адекватную самооценку **степени уверенности** системы в положительном или отрицательном решении о принадлежности/непринадлежности объекта к классу или **риска ошибки** при таком решении.

В-пятых, по сути, при распознавании происходит расчет коэффициентов I_j разложения функции объекта L_i в ряд по функциям классов I_{ij} , т.е. определяется **вес** каждого обобщенного образа класса в образе объекта, что подробнее описано в работах [2, 3, 4].

На рисунках 17 приведены экранные формы режима идентификации и прогнозирования 4.1.2 системы «Эйдос»:

3.6.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»

В АСК-анализе разработаны, а в системе «Эйдос» реализованы развитые методы прогнозирования, основанные на сценарном методе АСК-анализа или сценарном АСК-анализе [11]. Но в задачи данной работы не входит их подробное рассмотрение, тем более, что они подробно описаны и на теоретическом уровне, и с детальными численными примерами в работах [4-7, 11] и в ряде других [48].

Поэтому в данной работе рассмотрим стандартный несколько упрощенный вариант решения задачи идентификации и прогнозирования. Запустим режим 4.1.2 системы «Эйдос» (рисунок 16).

По результатам решения задачи идентификации и прогнозирования в системе «Эйдос» выдается большое количество разнообразных выходных форм, которых в настоящее время 14 (рисунок 17).

Из этих выходных форм рассмотрим только две: 4.1.3.1 и 4.1.3.2 (рисунок 18).

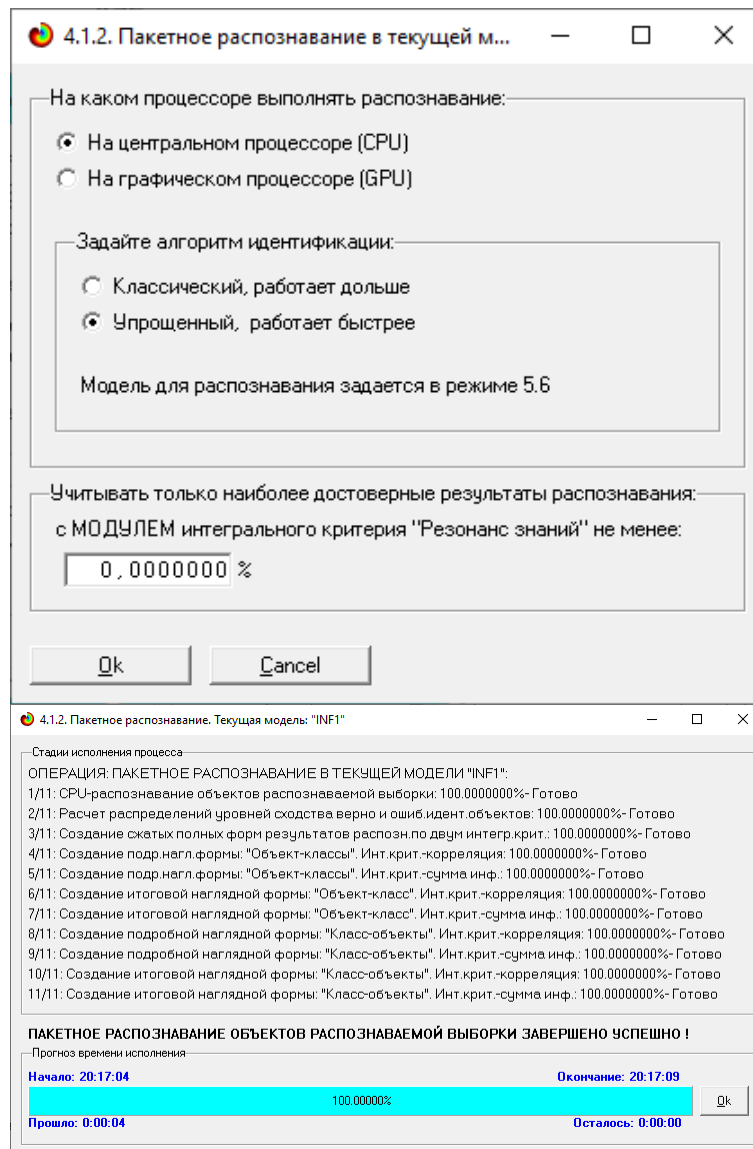


Рисунок 16. Экранные формы режима 4.1.2 идентификации и прогнозирования

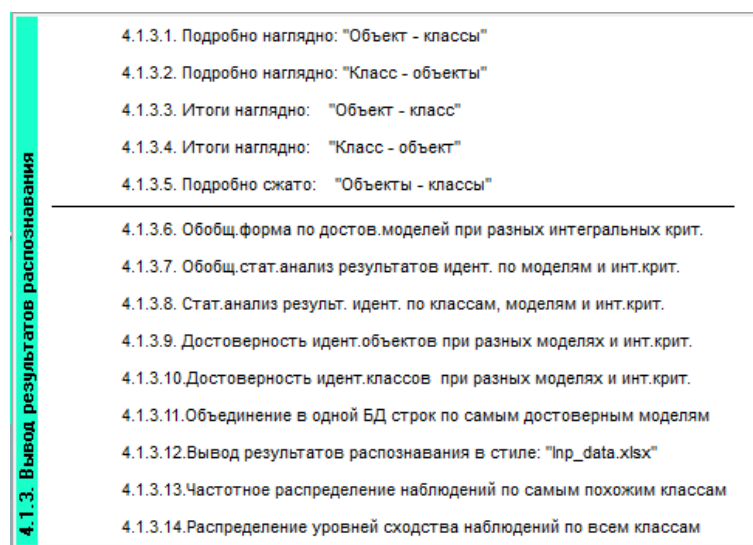


Рисунок 17. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

4.1.3.1. Визуализация результатов распознавания в отношении: "Объект-классы". Текущая модель: "INF1"

Распознаваемые объекты		Интегральный критерий сходства: "Семантический резонанс знаний"			
Код	Наим. объекта	Код	Наименование класса	Сходство	Ф... Сходство
1	1	2	POTABILITY-Большое	23,34...	v ██████████
2	2	1	POTABILITY-Малое	4,717...	█
3	3				
4	4				
5	5				
6	6				
7	7				
8	8				
9	9				
10	10				
11	11				
12	12				
13	13				
14	14	2	POTABILITY-Большое	3,278...	v █
15	15	1	POTABILITY-Малое	-4,518...	█
16	16				
17	17				
18	18				
19	19				
20	20				
21	21				
22	22				
23	23				

Интегральный критерий сходства: "Сумма знаний"

Код	Наименование класса	Сходство	Ф...	Сходство
2	POTABILITY-Большое	3,278...	v	█
1	POTABILITY-Малое	-4,518...		█

4.1.3.2. Визуализация результатов распознавания в отношении: "Класс-объекты". Текущая модель: "INF1"

Классы		Интегральный критерий сходства: "Семантический резонанс знаний"			
Код	Наим. класса	Код	Наименование объекта	Сходство	Ф... Сходство
1	POTABILITY-Малое	114	114	85,86...	v ██████████
2	POTABILITY-Большое	28	28	74,69...	v ██████████
		57	57	74,09...	v ██████████
		143	143	71,85...	v ██████████
		142	142	69,58...	v ██████████
		52	52	61,22...	v ██████████
		54	54	60,19...	v ██████████
		23	23	58,14...	v ██████████
		71	71	57,99...	v ██████████
		37	37	56,86...	v ██████████

Интегральный критерий сходства: "Сумма знаний"

Код	Наименование объекта	Сходство	Ф...	Сходство
114	114	40,57...	v	██████████
28	28	34,37...	v	██████████
57	57	34,03...	v	██████████
143	143	32,79...	v	██████████
142	142	31,52...	v	██████████
52	52	26,88...	v	██████████
54	54	26,31...	v	██████████
23	23	25,17...	v	██████████
71	71	25,09...	v	██████████
37	37	24,46...	v	██████████

Рисунок 18. Некоторые экранные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

Эти выходные формы, учитывая сказанное выше об интегральных критериях системы «Эйдос», как говорят «интуитивно понятны» и не требуют особых комментариев.

3.7. Задача-7. Поддержка принятия решений

3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ

3.7.1.1. Универсальная постановка задачи, не зависящая от предметной области

Задачи прогнозирования и принятия решений относятся друг к другу как прямая и *обратная* задачи:

– при прогнозировании по значениям факторов, действующих на объект моделирования, определяется в какое будущее состояние он перейдет под их действием;

– при принятии решений, наоборот, по будущему целевому состоянию объекта моделирования определяются значения факторов, которые обуславливают его переход в это будущее целевое состояние.

Таким образом, задача принятия решений является *обратной* по отношению к задаче прогнозирования. Но это так только в простейшем случае: в случае использования SWOT-анализа (режим 4.4.8 системы «Эйдос») [10] (рисунки 19).

На первом рисунке 19 приведена экранная форма задания в диалоге параметров отображения SWOT-диаграммы. На этой экранной форме в верхнем окне пользователь курсором выбирает исследуемый класс, внизу слева задает модель для исследования, а справа внизу задает отображать ли SWOT-диаграмму. Кроме того, пользователь может включить или выключить фильтры по факторам и посмотреть помощь по режиму. При включении фильтра по фактору, на котором стоит курсор, на экранных формах отображается влияние только значения этого фактора.

Графические выходные формы, приведенные на рисунках 19, интуитивно понятны и не требуют особых комментариев. Отметим лишь, что на SWOT-диаграммах наглядно показаны знак и сила влияния каждого значения фактора на переход объекта моделирования в состояние, соответствующее классу, выбранному в верхнем окне. Знак показан цветом, а сила влияния – толщиной линии.

Слева на SWOT-диаграмме приведены значения факторов, способствующих переходу объекта моделирования в состояние, соответствующее классу, выбранному в верхнем окне (показаны красным цветом), а справа – препятствующих этому переходу (показаны синим цветом). Сила влияния каждого значения фактора на поведение объекта моделирования показана толщиной линии.

Необходимо подчеркнуть, что в системе «Эйдос» SWOT-диаграммы формируются автоматически на основе статистических и системно-когнитивных моделей, созданных непосредственно на основе эмпирических данных, а не как обычно не формализуемым экспертным

путем на основе интуиции, опыта и профессиональной компетенции, т.е. практически «на глазок», а в некоторых случаях и вообще «от фонаря».

3.7.1.2. Конкретное решение задачи идентификации и прогнозирования в данной работе в системе «Эйдос»

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редукция клас...	N объектов (абс.)	N объектов (%)
1	POTABILITY-Малое	0,0318442	1088	60,5000000
2	POTABILITY-Большое	0,0479815	711	39,5000000

SWOT-анализ класса: 1 "POTABILITY-Малое" в модели: 4 "INF1"

Способствующие факторы и сила их влияния

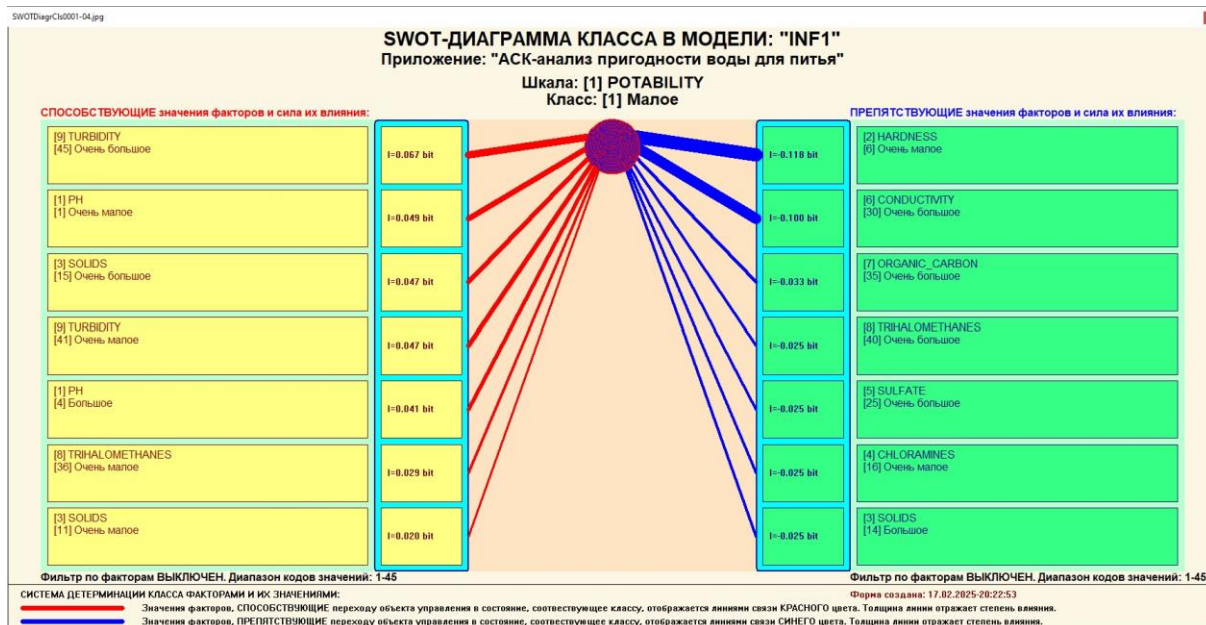
Код	Наименование фактора и его интервального значения	Сила влияния
45	TURBIDITY-Очень большое	0.067
1	PH-Очень малое	0.049
15	SOLIDS-Очень большое	0.047
41	TURBIDITY-Очень малое	0.047
4	PH-Большое	0.041
36	TRIHALOMETHANES-Очень малое	0.029
11	SOLIDS-Очень малое	0.020
7	HARDNESS-Малое	0.013
23	SULFATE-Среднее	0.012
29	CONDUCTIVITY-Большое	0.012
8	HARDNESS-Среднее	0.011
18	CHLORAMINES-Среднее	0.010
34	ORGANIC_CARBON-Большое	0.009
44	TURBIDITY-Большое	0.009

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
6	HARDNESS-Очень малое	-0.118
30	CONDUCTIVITY-Очень большое	-0.100
35	ORGANIC_CARBON-Очень большое	-0.033
40	TRIHALOMETHANES-Очень большое	-0.025
25	SULFATE-Очень большое	-0.025
16	CHLORAMINES-Очень малое	-0.025
14	SOLIDS-Большое	-0.025
3	PH-Среднее	-0.021
24	SULFATE-Большое	-0.021
42	TURBIDITY-Малое	-0.018
20	CHLORAMINES-Очень большое	-0.015
5	PH-Очень большое	-0.015
19	CHLORAMINES-Большое	-0.011
26	CONDUCTIVITY-Очень малое	-0.010

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма



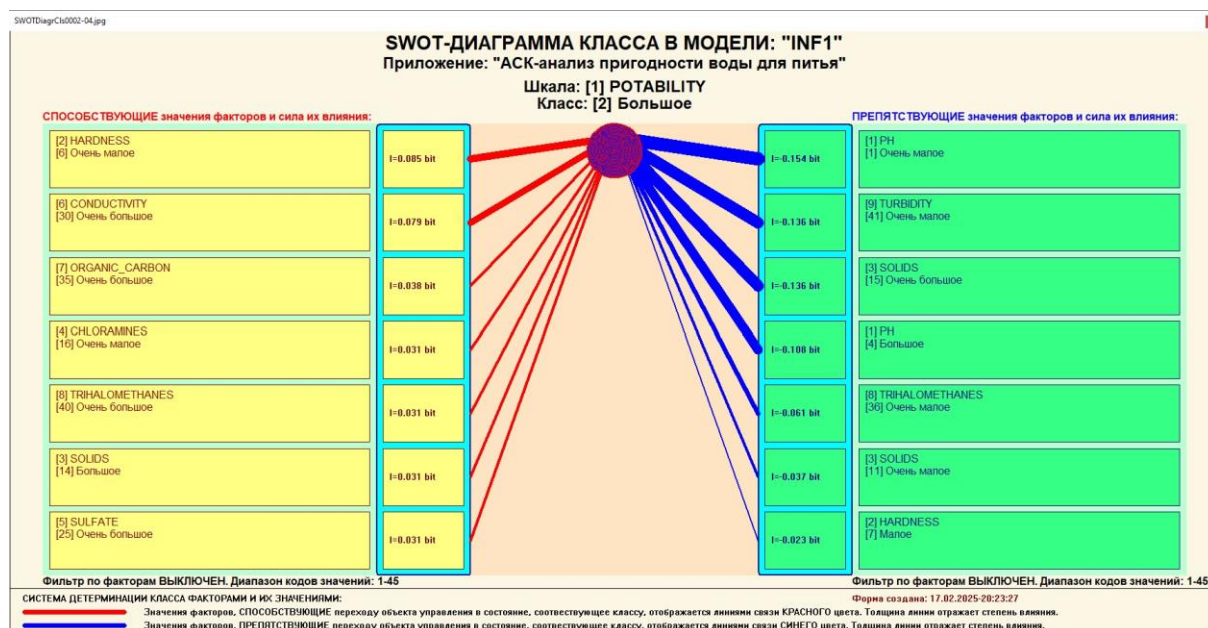


Рисунок 19. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

Из приведенных выходных форм SWOT-анализа четко видно, что очень малая жесткость способствует пригодности воды для питья, а очень малый рН – препятствует.

3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

3.7.2.1. Универсальная постановка задачи, не зависящая от предметной области

Однако SWOT-анализ (режим 4.4.8 системы «Эйдос») имеет свои ограничения: может быть задано только одно будущее целевое состояние, целевые состояния могут быть недостижимыми одновременно (альтернативными) или совместимыми по системе обуславливающих их значений факторов, некоторые рекомендуемые факторы может не быть технологической и финансовой возможности использовать и возможно надо искать им замену, примерно так же влияющую на объект моделирования.

Поэтому в АСК-анализе и системе «Эйдос» реализован развитый алгоритм принятия решений (режим 6.3) в котором кроме SWOT-анализа используются также результаты решения задачи прогнозирования и результаты кластерно-конструктивного анализа классов и значений факторов, т.е. некоторые результаты решения задачи исследования предметной области. Этот алгоритм описан в работах [11, 12, 13] и в ряде других работ.

Приведем этот алгоритм и в данной работе (рисунок 20).

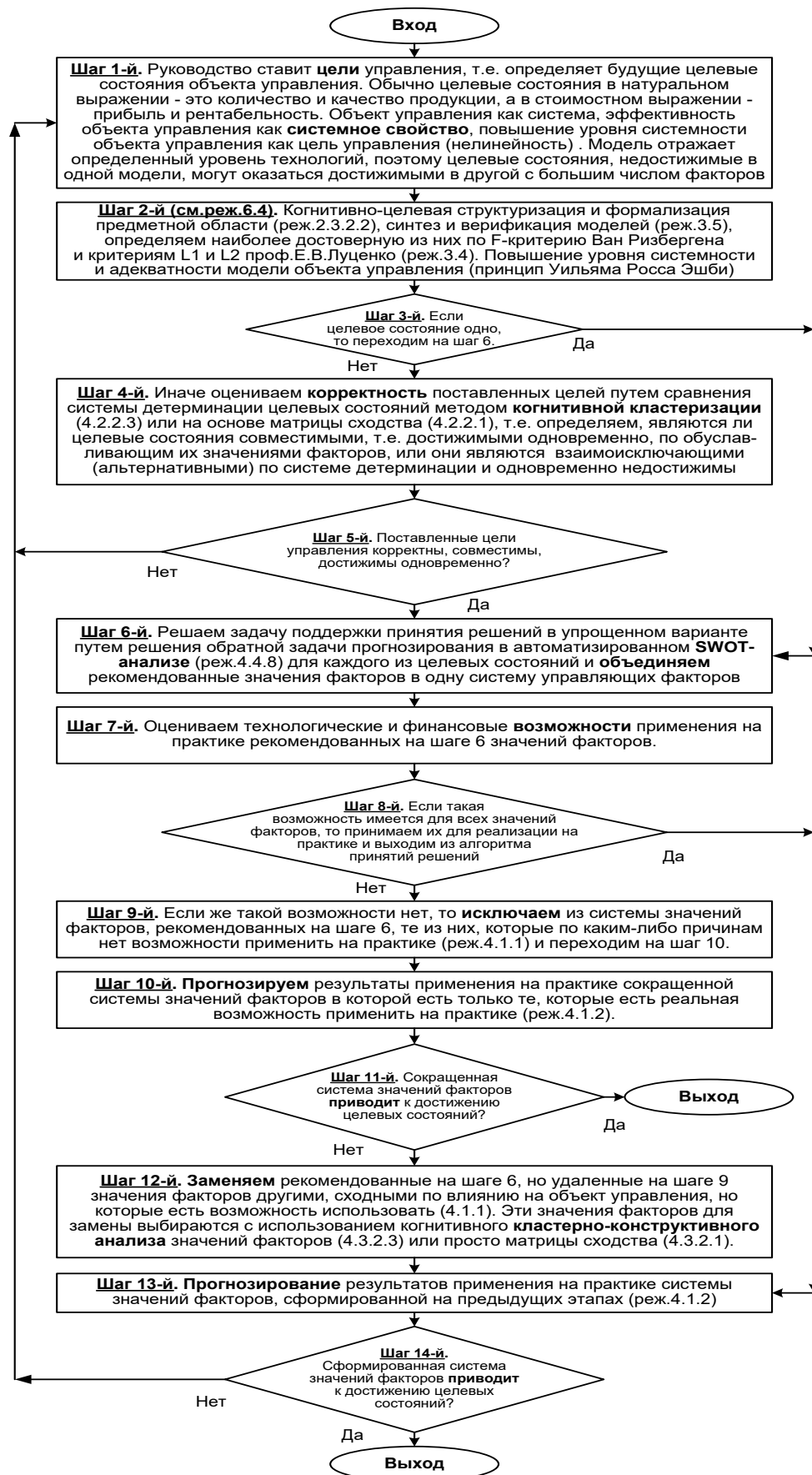


Рисунок 20. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

Шаг 1-й. Руководство ставит **цели** управления, т.е. определяет будущие целевые состояния объекта управления. Обычно целевые состояния в натуральном выражении – это количество и качество продукции, а в стоимостном выражении – прибыль и рентабельность. Объект управления как система, эффективность объекта управления как **системное свойство**, повышение уровня системности объекта управления как цель управления (нелинейность). Модель отражает определенный уровень технологий, поэтому целевые состояния, недостижимые в одной модели, могут оказаться достижимыми в другой с большим числом факторов [11, 12, 13, 14, 15, 16].

Шаг 2-й (см.реж.6.4). Когнитивно-целевая структуризация и формализация предметной области (реж.2.3.2.2), синтез и верификация моделей (реж.3.5), определяем наиболее достоверную из них по F-критерию Ван Ризбергена и критериям L1 и L2 проф.Е.В.Луценко (реж.3.4) [9]. Повышение уровня системности и адекватности модели объекта управления (принцип Уильяма Росса Эшби) [14, 15].

Шаг 3-й. Если целевое состояние одно, то переходим на шаг 6, а иначе на шаг 4.

Шаг 4-й. Иначе оцениваем **корректность** поставленных целей путем сравнения системы детерминации целевых состояний методом **когнитивной кластеризации** (4.2.2.3) или на основе матрицы сходства (4.2.2.1), т.е. определяем, являются ли целевые состояния совместимыми, т.е. достижимыми одновременно, по обуславливающим их значениями факторов, или они являются взаимоисключающими (альтернативными) по системе детерминации и одновременно недостижимы.

Шаг 5-й. Поставленные цели управления корректны, совместимы, достижимы одновременно? Если да – переход на шаг 6, иначе на шаг 1.

Шаг 6-й. Решаем задачу поддержки принятия решений в упрощенном варианте путем решения обратной задачи прогнозирования в автоматизированном **SWOT-анализе** (реж.4.4.8) для каждого из целевых состояний и **объединяем** рекомендованные значения факторов в одну систему управляющих факторов [10].

Шаг 7-й. Оцениваем технологические и финансовые **возможности** применения на практике рекомендованных на шаге 6 значений факторов.

Шаг 8-й. Если такая возможность имеется для всех значений факторов, то принимаем их для реализации на практике и переходим на шаг 13 для проверки эффективности принятых решений, а иначе переходим на шаг 9.

Шаг 9-й. Если же такой возможности нет, то **исключаем** из системы значений факторов, рекомендованных на шаге 6, те из них, которые по каким-либо причинам нет возможности применить на практике (реж.4.1.1) и переходим на шаг 10.

Шаг 10-й. Прогнозируем результаты применения на практике сокращенной системы значений факторов в которой есть только те, которые есть реальная возможность применить на практике (реж.4.1.2).

Шаг 11-й. Сокращенная система значений факторов приводит к достижению целевых состояний? Если да, то выход из алгоритма принятия решений, а иначе переход на шаг 12.

Шаг 12-й. Заменяем рекомендованные на шаге 6, но удаленные на шаге 9 значения факторов другими, сходными по влиянию на объект управления, но которые есть возможность использовать (4.1.1). Эти значения факторов для замены выбираются с использованием когнитивного кластерно-конструктивного анализа значений факторов (4.3.2.3) или просто матрицы сходства (4.3.2.1) [16].

Шаг 13-й. Прогнозирование результатов применения на практике системы значений факторов, сформированной на предыдущих этапах (реж.4.1.2)

Шаг 14-й. Сформированная система значений факторов приводит к достижению целевых состояний? Если да – то выход из алгоритма принятия решений, а иначе переход на шаг 1.

Приведенный выше развитый алгоритм принятия решений может быть применен в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос». Принципиальная схема подобной системы приведена на рисунке 21:

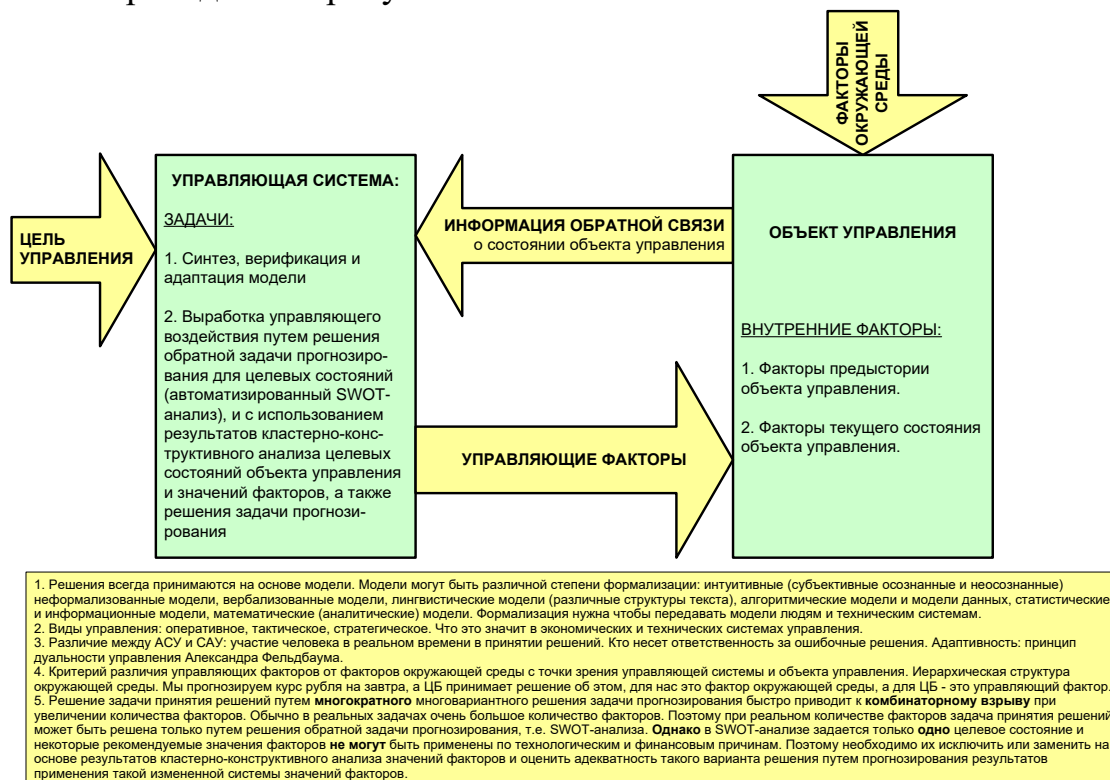


Рисунок 21. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

Как мы видим в развитом алгоритме принятия решений широко используются результаты конкретного решения различных задач: и задачи прогнозирования, и некоторых задач исследования объекта моделирования путем исследования его модели. Необходимо особо отметить, что система «Эйдос» поддерживает решение всех этих задач, которые необходимо решать в развитом алгоритме принятия решений. Потому ниже кратко рассмотрим решение этих и некоторых других задач.

3.7.2.2. Конкретное решение задачи управления в данной работе в системе «Эйдос»

Необходимо отметить, что модели системы Эйдос – это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного *механизма* детерминации, а только сам факт и характер детерминации (силу и направление влияния факторов). *Содержательное* объяснение конкретных механизмов этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [12, 17].

3.8. Задача-8. Исследование объекта моделирования путем исследования его модели

3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы)

3.8.1.1. Универсальная постановка задачи, не зависящая от предметной области

Инвертированные SWOT-диаграмм (предложены автором в работе [10]), отражают силу и направление влияния конкретной градации описательной шкалы на переход объекта моделирования в состояния, соответствующие градациям классификационных шкал (классы). Это и есть *смысл* (семантический потенциал) этой градации описательной шкалы. Инвертированные SWOT-диаграммы выводятся в режиме 4.4.9 системы «Эйдос».

3.8.1.2. Конкретное решение задачи в данной работе

Примеры инвертированных SWOT-диаграмм приведены на рисунках 22 для некоторых значений факторов:

4.4.9 Количественный автоматизированный SWOT-анализ значений факторов средствами АСК-анализа в системе "Эйдос"

Выбор значения фактора, оказывающего влияние на переход объекта управления в будущие состояния

Код	Наименование значения фактора
1	PH-Очень малое
2	PH-Малое
3	PH-Среднее
4	PH-Большое
5	PH-Очень большое
6	HARDNESS-Очень малое

SWOT-анализ значения фактора: 1 "PH-Очень малое" в модели: 4 "INF1"

СПОСОБСТВУЕТ:

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влия...
1	POTABILITY-Малое	0.049

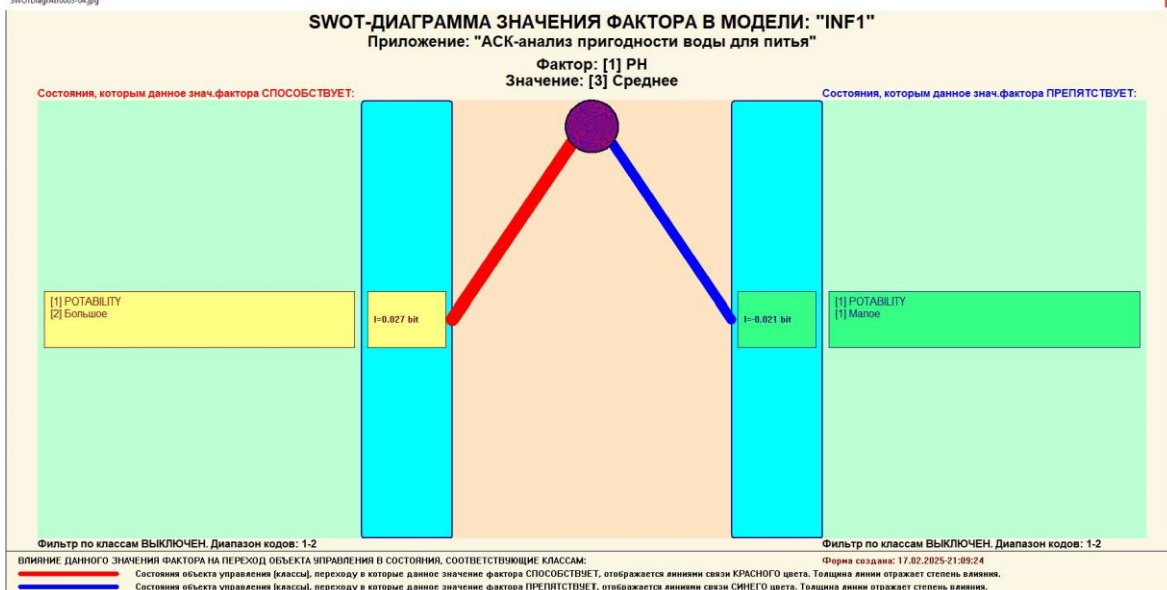
ПРЕПЯТСТВУЕТ:

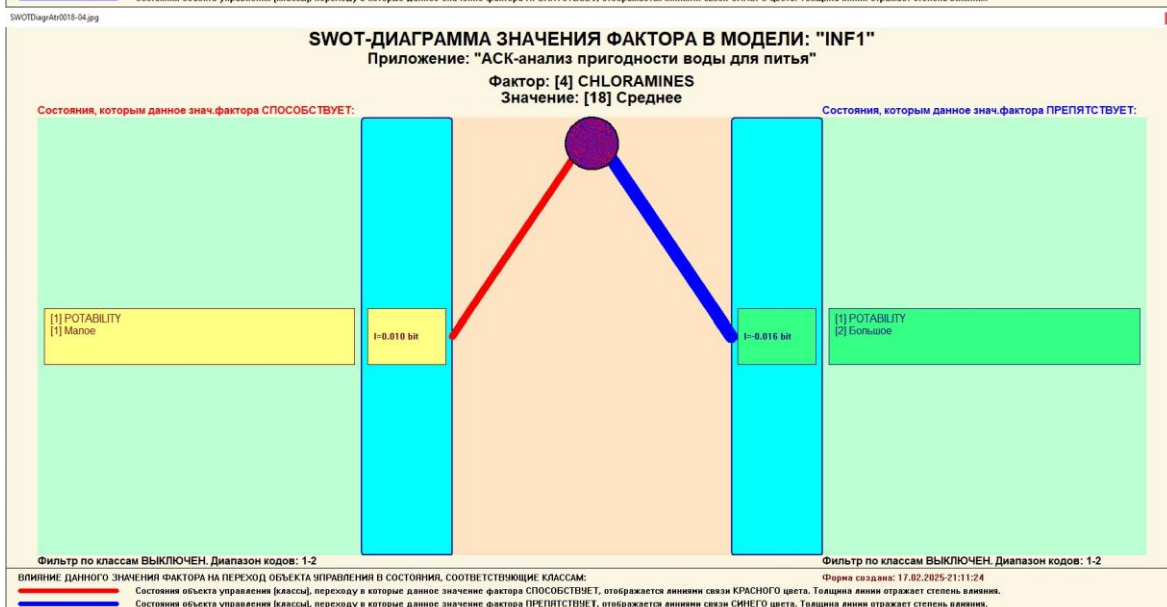
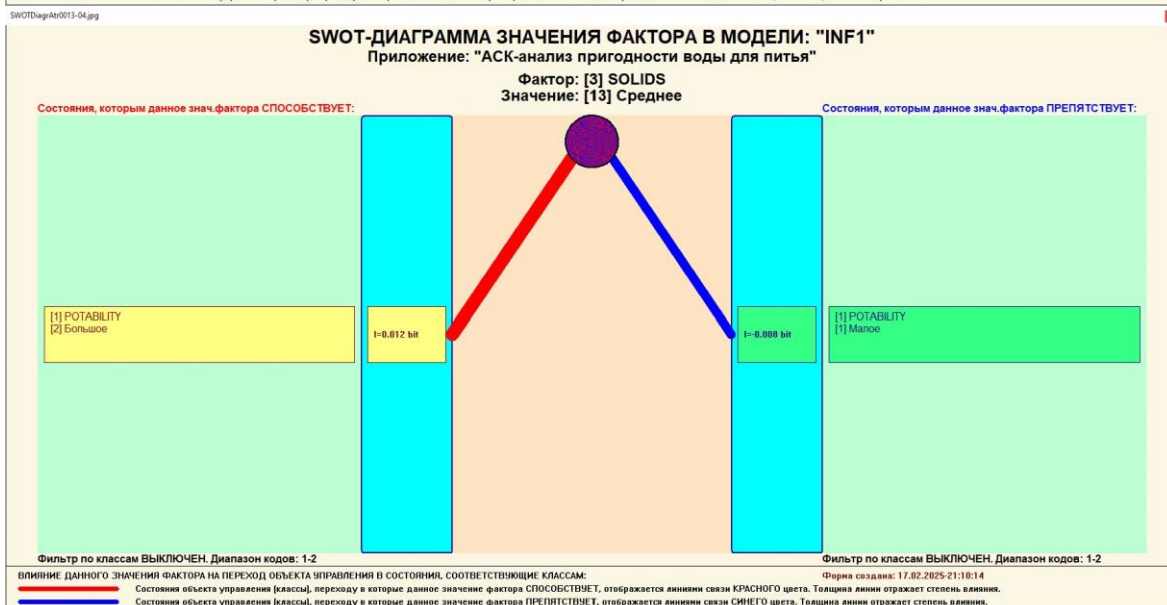
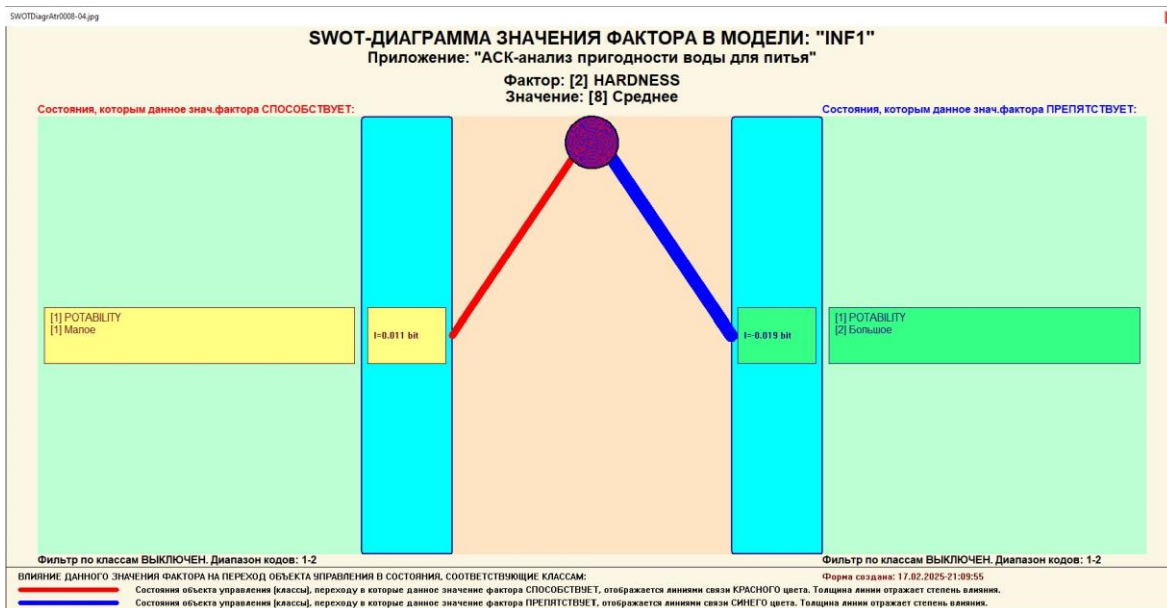
Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
2	POTABILITY-Большое	-0.154

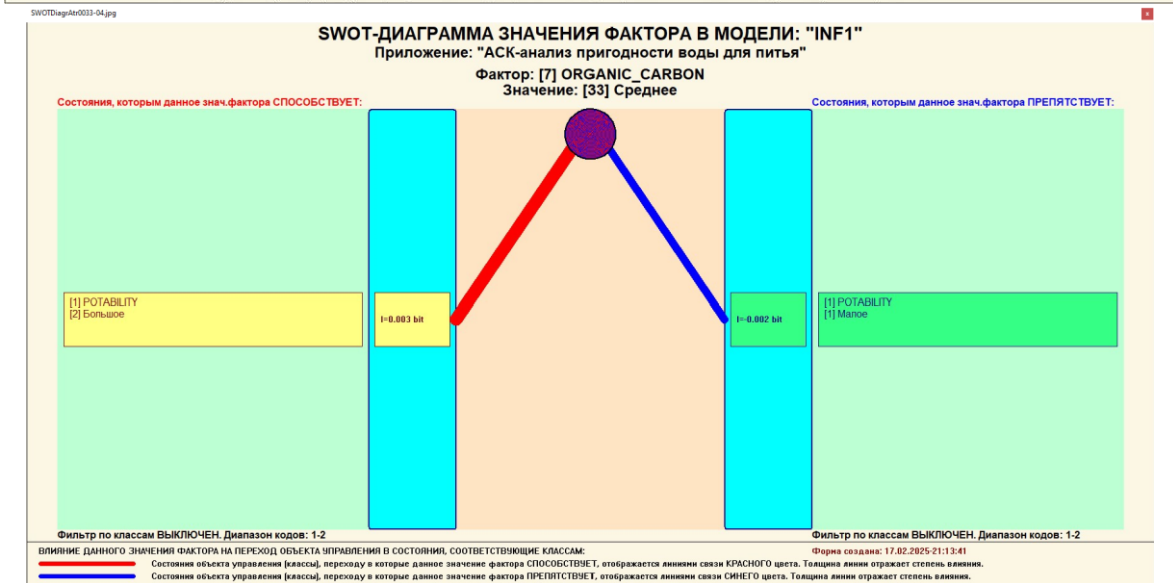
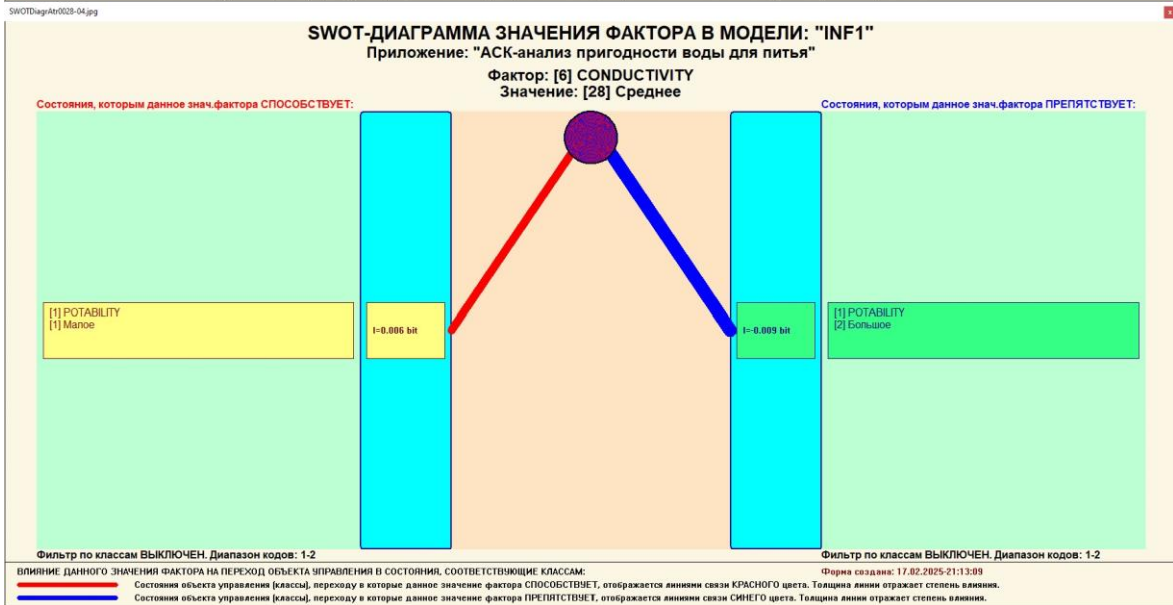
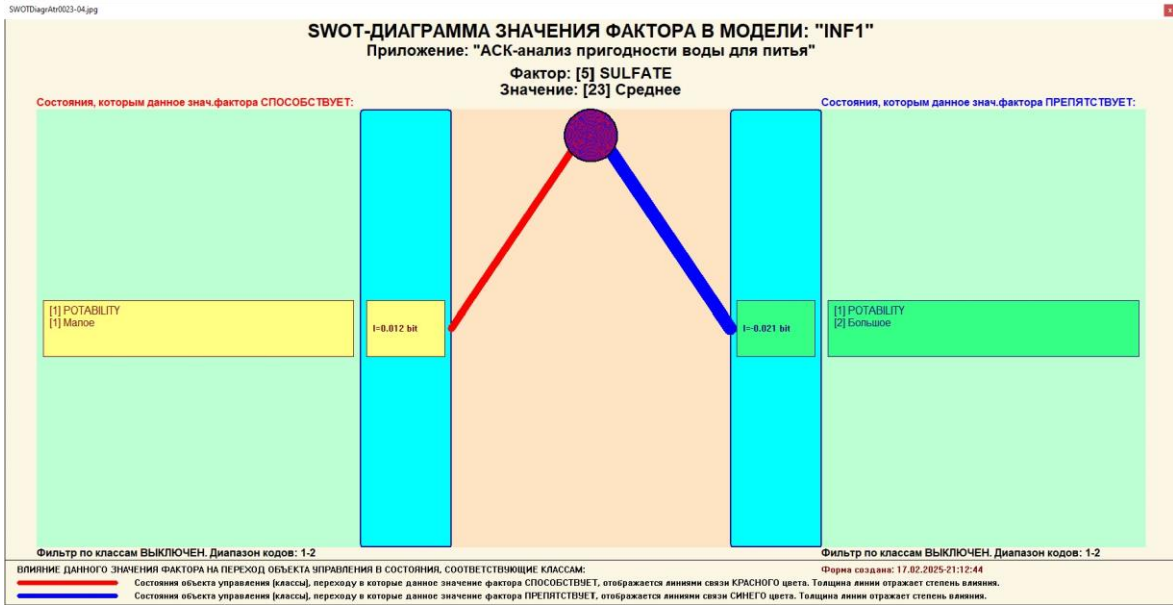
ВКЛЮЧИТЬ фильтр по клшкале ВЫКЛЮЧИТЬ фильтр по клшкале ВКЛЮЧИТЬ фильтр по клшкале ВЫКЛЮЧИТЬ фильтр по клшкале

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма

SWOTDiagram003-04.jpg







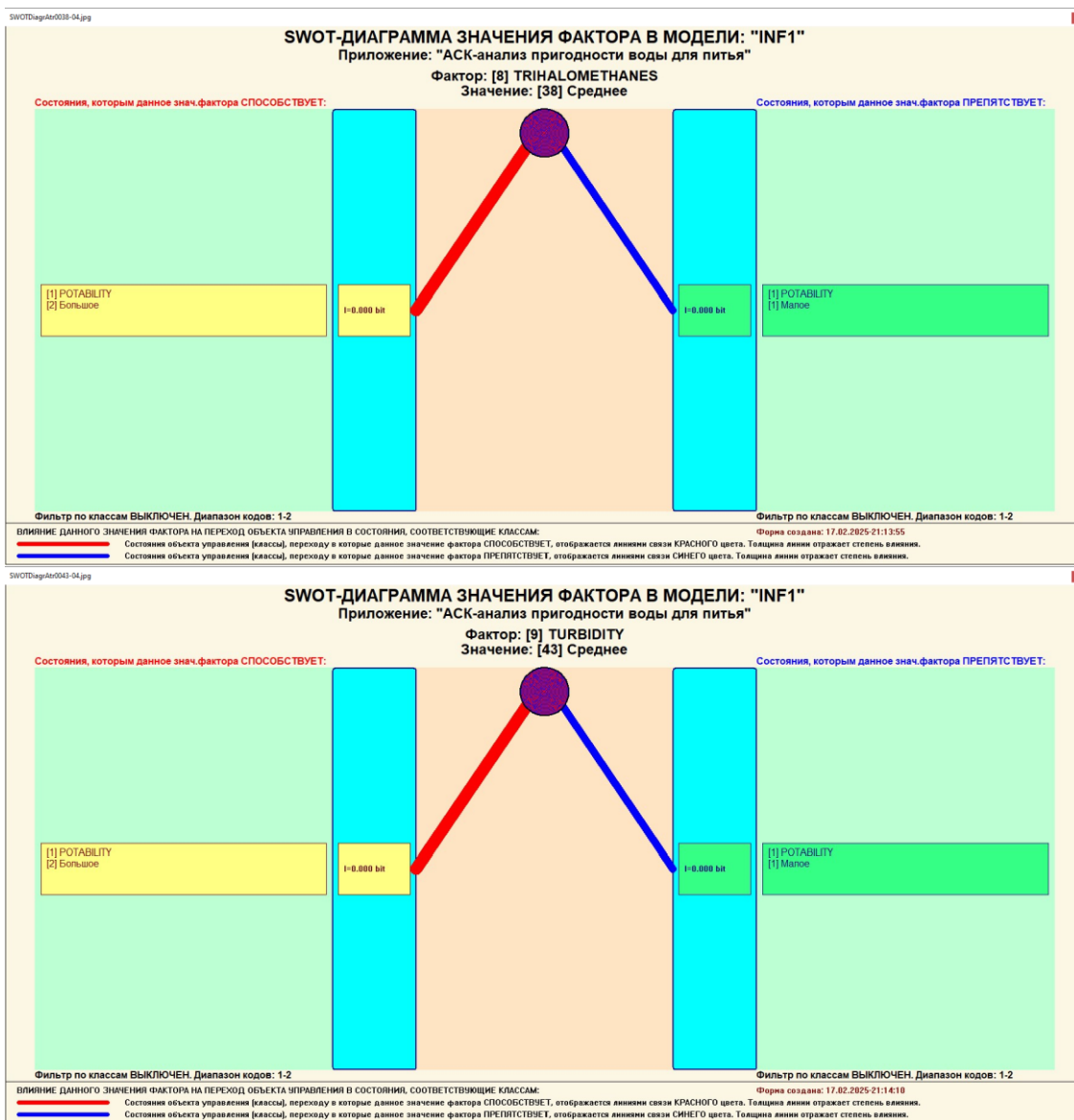


Рисунок 22. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

Приведенные на рисунке 22 инвертированные SWOT-диаграммы исчерпывающим образом отражают силу и направление влияния каждого значения каждого фактора на переход объекта моделирования в состояния, соответствующие различным классам.

3.8.2. Кластерно-конструктивный анализ классов

3.8.2.1. Универсальная постановка задачи, не зависящая от предметной области

В системе «Эйдос» (в режиме 4.2.2.1, рисунок 23) рассчитывается матрица сходства классов (таблица 13) по системе их детерминации и на основе этой матрицы рассчитывается и выводится три основных формы:

- круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2) (рисунок 24);
- агломеративные дендрограммы, полученные в результате *когнитивной (истинной) кластеризации классов* (предложена автором в 2011 году в работе [16]) (режим 4.2.2.3) (рисунок 26);
- график изменения межкластерных расстояний (режим 4.2.2.3) (рисунок 25).

Эта матрица сходства (таблица 13) используются и при расчете некоторых других выходных форм.

3.8.2.2. Конкретное решение задачи в данной работе

На рисунке 23 представлены экранные формы режима 4.2.2.1, обеспечивающего расчет матрицы сходства классов по системе их детерминации, т.е. по обуславливающим их значениям факторов:

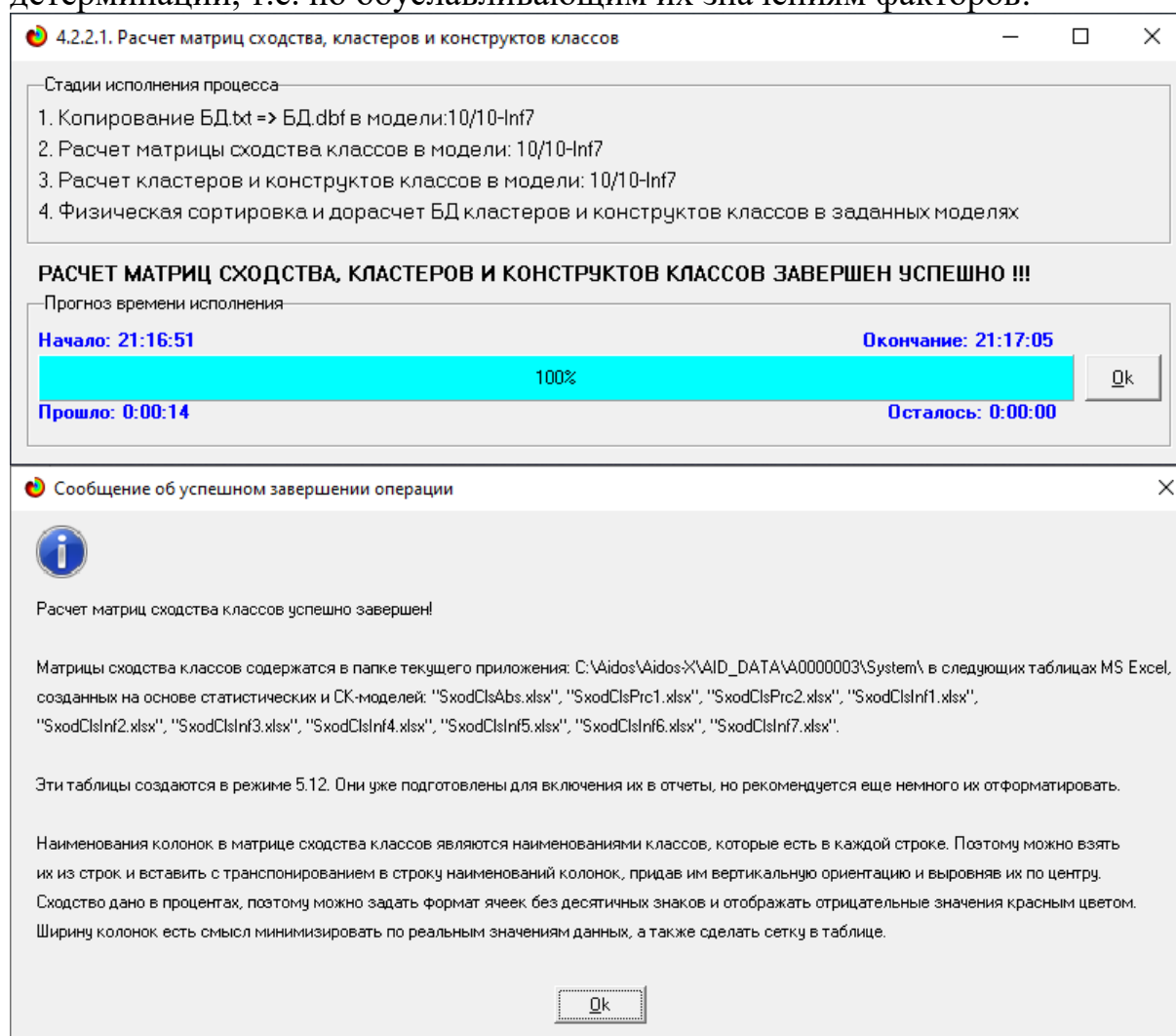


Рисунок 23. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

Таблица 13 – Матрица сходства классов в СК-модели INF3 (полностью)

	1-ROTABILITY-Малое	2-ROTABILITY-Большое
1-ROTABILITY-Малое	100	-60,1408
2-ROTABILITY-Большое	-60,1408	100

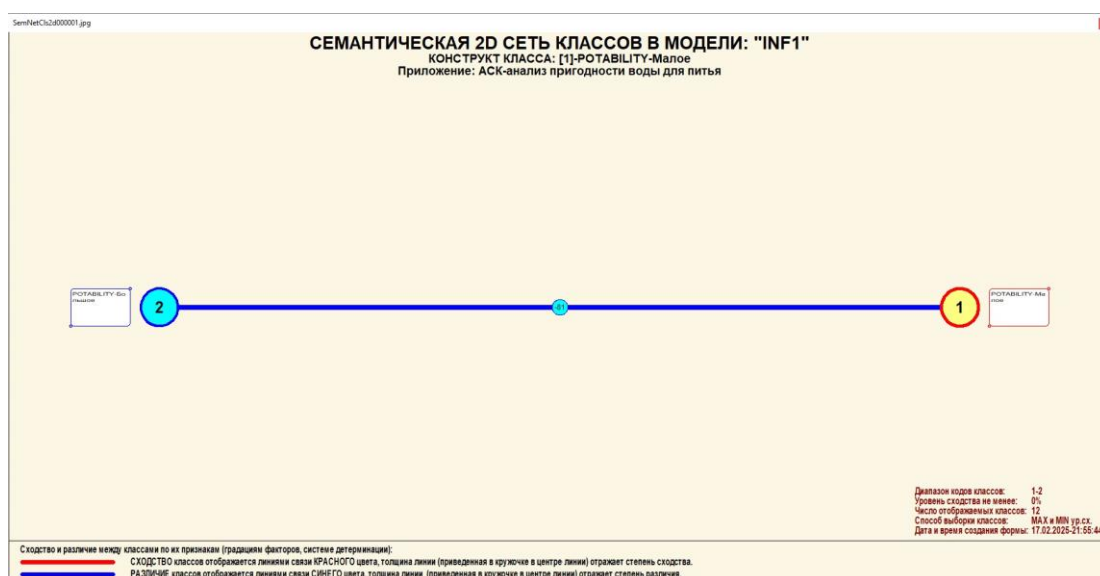


Рисунок 24. Круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2)

Из рисунка 24 видно, что классы имеют противоположную природу.

3.8.3. Кластерно-конструктивный анализ значений описательных шкал

3.8.3.1. Универсальная постановка задачи, не зависящая от предметной области

В системе «Эйдос» (в режиме 4.3.2.1, рисунок 25) рассчитывается матрица сходства признаков (таблица 15) по их смыслу и на основе этой матрицы рассчитывается и выводится четыре основных формы:

– круговая 2d-когнитивная диаграмма признаков (режим 4.3.2.2) рисунок 26);

– агломеративные дендрограммы, полученные в результате **когнитивной (истинной) кластеризации признаков** (предложена автором в 2011 году в работе [16]) (режим 4.3.2.3) рисунок 27);

– график изменения межкластерных расстояний (режим 4.3.2.3) рисунок 30).

Эта матрица сходства (таблица 15) используются и при расчете некоторых других выходных форм.

3.8.3.2. Конкретное решение задачи в данной работе

На рисунке 25 представлены экранные формы режима 4.3.2.1, обеспечивающего расчет матрицы сходства значений факторов по силе и направлению их влияния на переход объекта моделирования в различные будущие состояния, соответствующие классам:

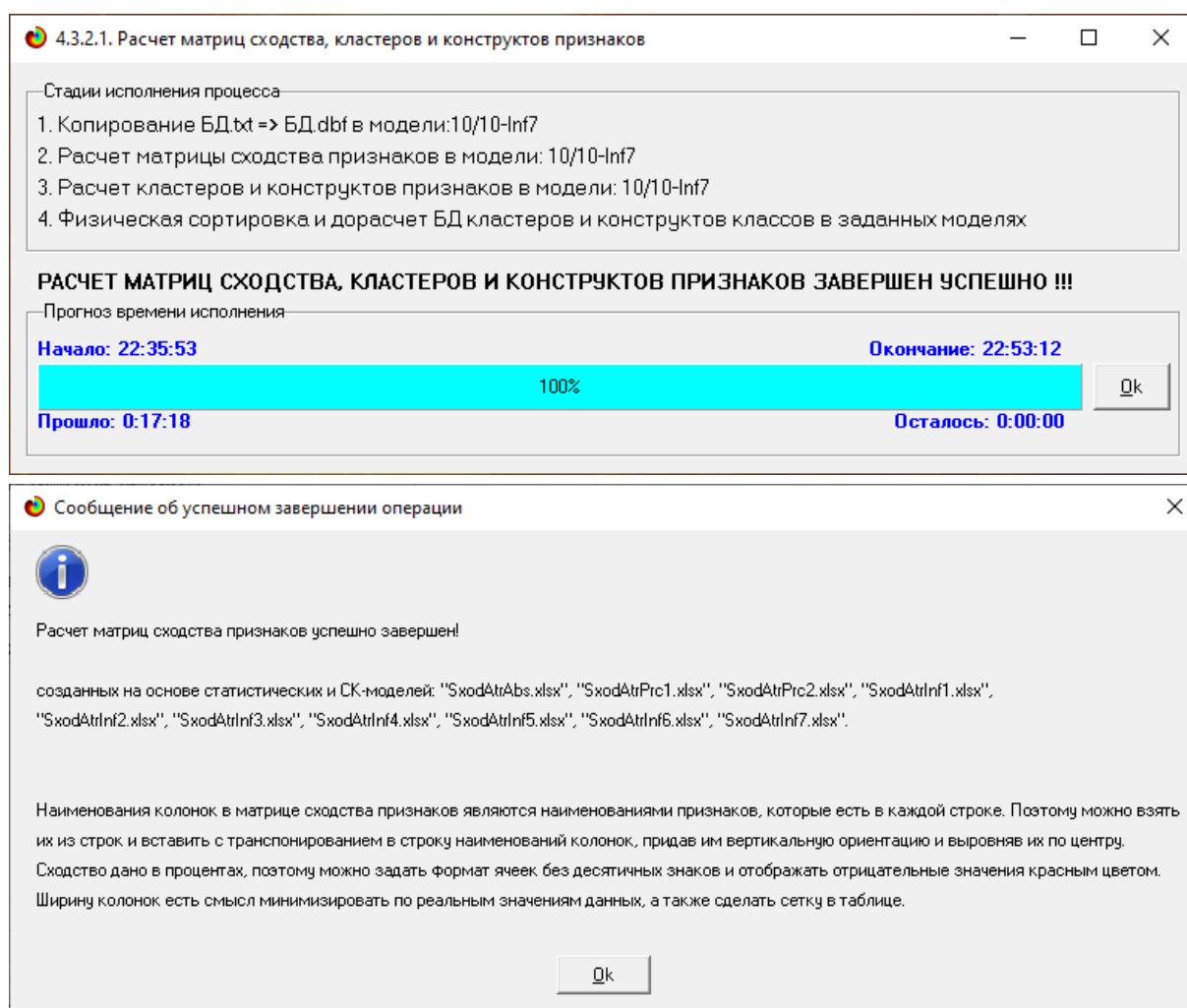


Рисунок 25 Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов

Таблица 14 – Матрица сходства значений факторов в СК-модели INF1 (фрагмент)

KOD_ATR	KOD_OPSC	NAME_ATR	N1	N2	N3
1	1	PH-Очень малое	100,000000	100,000000	-100,000000
2	1	PH-Малое	100,000000	100,000000	-100,000000
3	1	PH-Среднее	-100,000000	-100,000000	100,000000
4	1	PH-Большое	100,000000	100,000000	-100,000000
5	1	PH-Очень большое	-100,000000	-100,000000	100,000000
6	2	HARDNESS-Очень малое	-100,000000	-100,000000	100,000000
7	2	HARDNESS-Малое	100,000000	100,000000	-100,000000
8	2	HARDNESS-Среднее	100,000000	100,000000	-100,000000
9	2	HARDNESS-Большое	-100,000000	-100,000000	100,000000

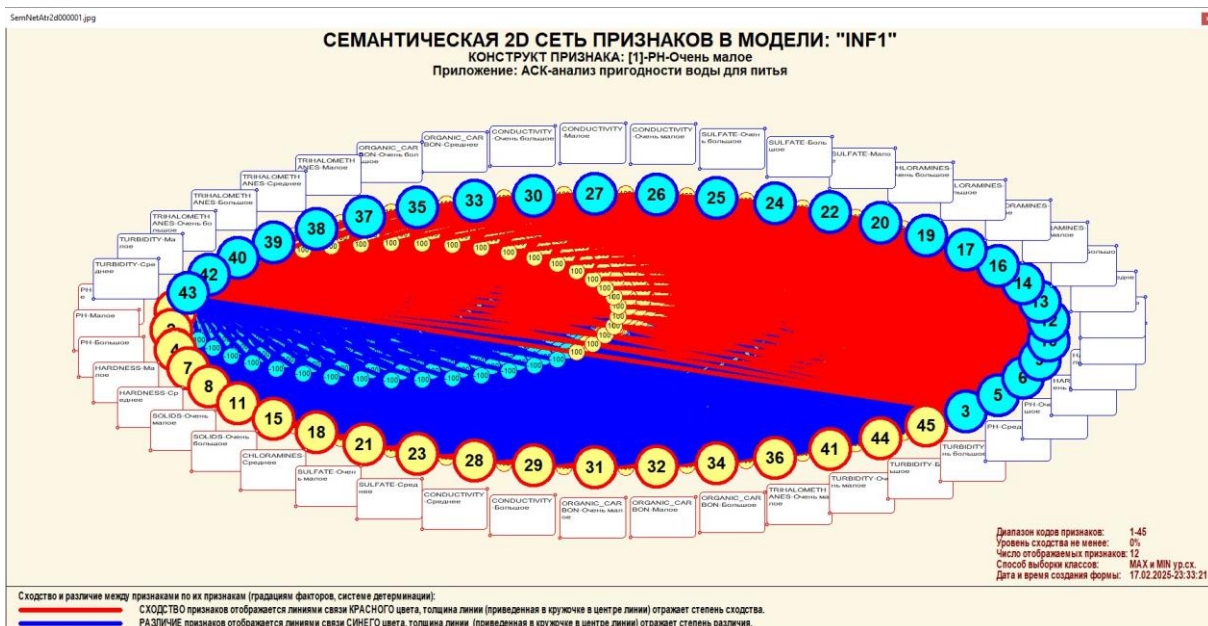


Рисунок 26. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF1 (режим 4.3.2.2)

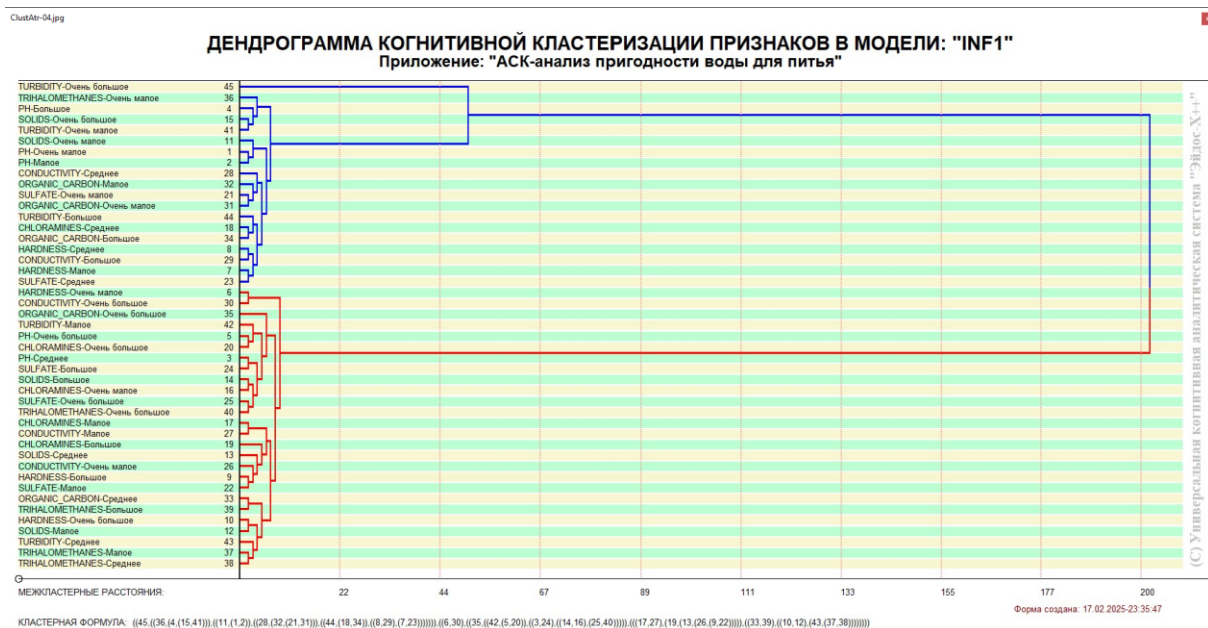


Рисунок 27. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)

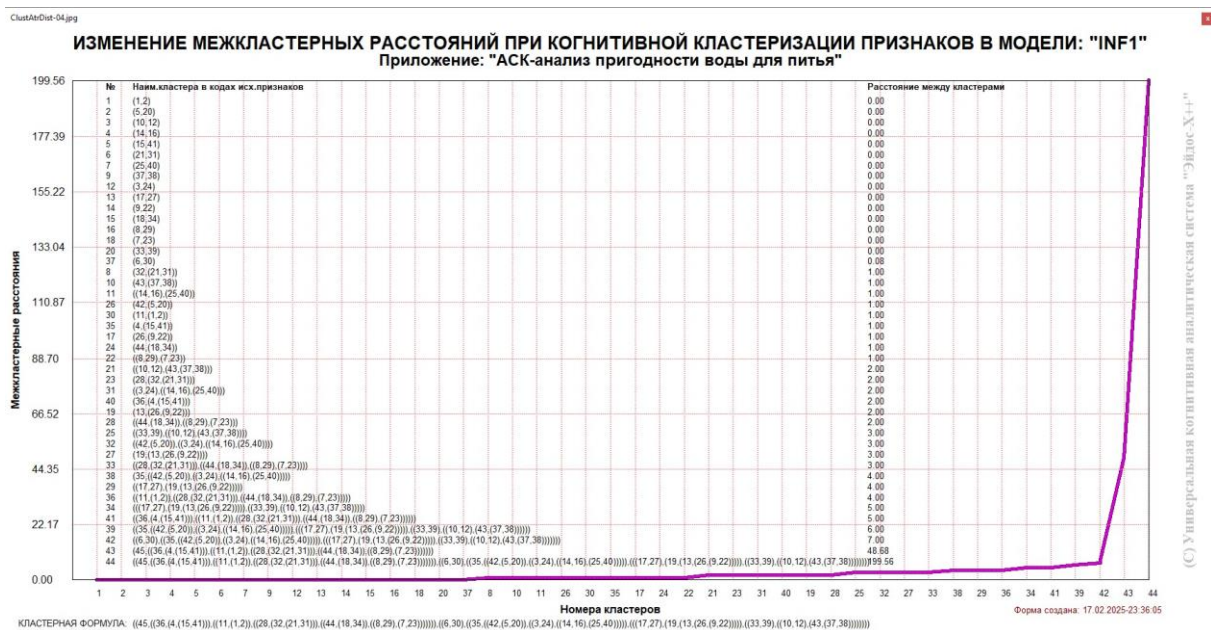


Рисунок 28. График изменения межкластерных расстояний (режим 4.3.2.3)

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

3.8.4.1. Универсальная постановка задачи, не зависящая от предметной области

Модель знаний системы «Эйдос» относится к **нечетким декларативным гибридным** моделям и объединяет в себе некоторые положительные особенности нейросетевой и фреймовой моделей представления знаний.

Классы в этой модели соответствуют нейронам и фреймам, а признаки рецепторам и шпациям (описательные шкалы – слотам).

От фреймовой модели представления знаний модель системы «Эйдос» отличается своей эффективной и простой программной реализацией, полученной за счет того, что разные фреймы отличаются друг от друга не набором слотов и шпаций, а лишь информацией в них. *Поэтому в системе «Эйдос» при увеличении числа фреймов само количество баз данных не увеличивается, а увеличивается лишь их размерность.* Это является очень важным свойством моделей системы «Эйдос», существенно облегчающим и упрощающим программную реализацию и быстрое действие системы.

От нейросетевой модели представления знаний модель системы «Эйдос» отличается тем, что [18]:

1) весовые коэффициенты на рецепторах не подбираются итерационным методом обратного распространения ошибки, а рассчитываются методом прямого счета на основе хорошо теоретически

обоснованной модели, основанной на **теории информации** (это напоминает байесовские сети);

2) весовые коэффициенты имеют хорошо теоретически обоснованную **содержательную интерпретацию**, основанную на теории информации;

3) нейросеть является **нелокальной**, как сейчас говорят «полносвязной».

В системе «Эйдос» нелокальные нейроны визуализируются (режим 4.4.10 системы «Эйдос») в виде специальных графических форм, на которых сила и направление влияния рецепторов нейрона на степень его активации/торможения, которые отображаются соответственно в форме цвета и толщины дендрита. В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.4.2. Конкретное решение задачи в данной работе

4.4.10.Графическое отображение нелокального нейрона в системе "Эйдос"

Выбор нелокального нейрона (класса) для визуализации

Код	Наименование нелокального нейрона (класса)
1	POTABILITY-Малое
2	POTABILITY-Большое

Подготовка визуализации нейрона:1 "POTABILITY-Малое" в модели:4 "INF1"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
45	TURBIDITY-Очень большое	0.067
1	PH-Очень малое	0.049
15	SOLIDS-Очень большое	0.047
41	TURBIDITY-Очень малое	0.047
4	PH-Большое	0.041
36	TRIHALOMETHANES-Очень малое	0.029
11	SOLIDS-Очень малое	0.020
7	HARDNESS-Малое	0.013
23	SULFATE-Среднее	0.012
20	CONDUCTIVITY-Большое	0.012

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
6	HARDNESS-Очень малое	-0.118
30	CONDUCTIVITY-Очень большое	-0.100
35	ORGANIC_CARBON-Очень большое	-0.033
40	TRIHALOMETHANES-Очень большое	-0.025
25	SULFATE-Очень большое	-0.025
16	CHLORAMINES-Очень малое	-0.025
14	SOLIDS-Большое	-0.025
3	PH-Среднее	-0.021
24	SULFATE-Большое	-0.021
42	TURBIDITY-Малое	-0.019

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

НЕЙРОН Максимальное количество отображаемых рецепторов: 999 Минимальный вес.коэф. отображаемых рецепторов: 0,000

Сортировать рецепторы:
 по информативности
 по модулю информативности

Отображать рецепторы:
 с наименованиями
 только с кодами

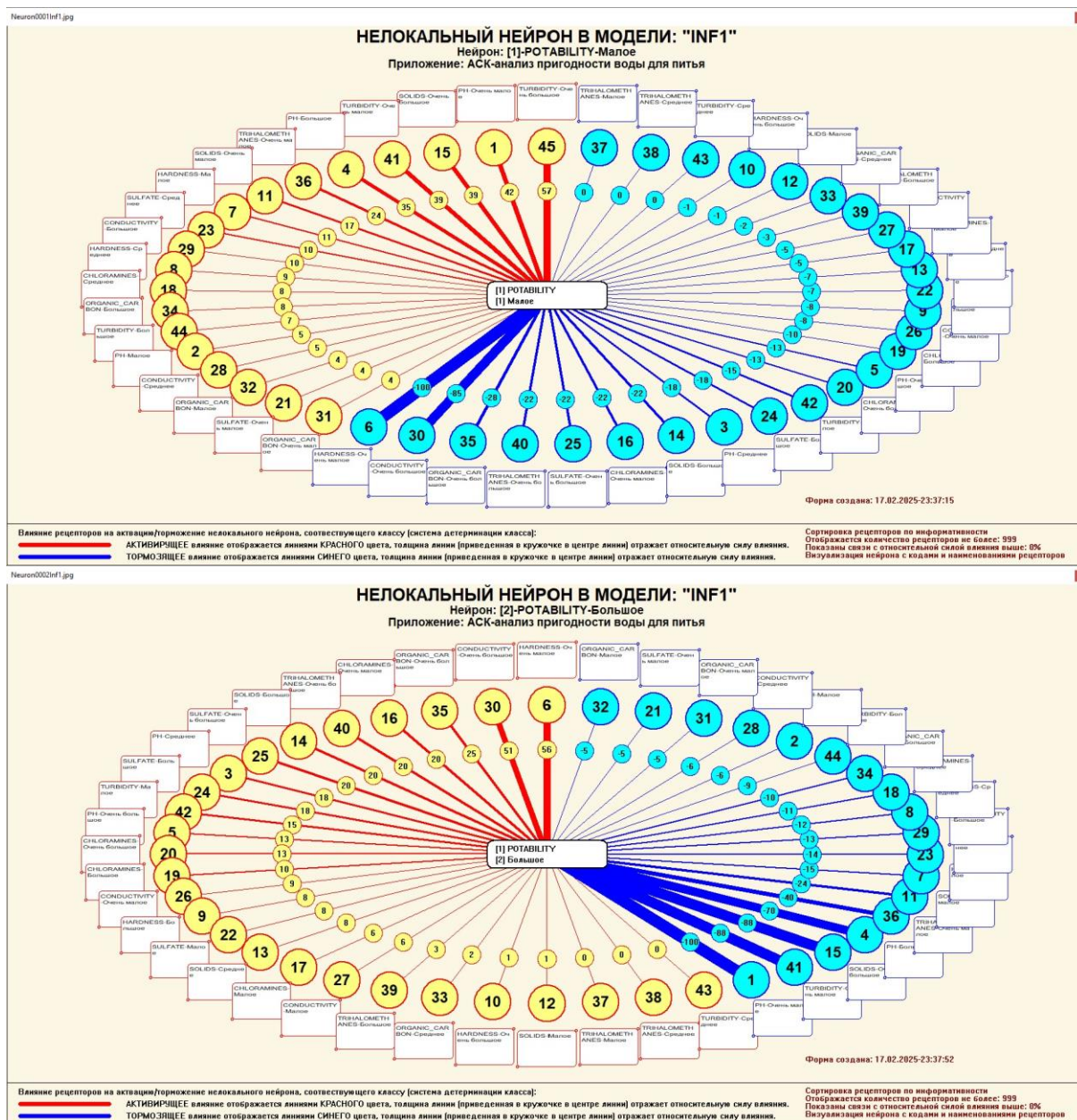


Рисунок 29. Примеры нелокальных нейронов, соответствующие классам

3.8.5. Нелокальная нейронная сеть

3.8.5.1. Универсальная постановка задачи, не зависящая от предметной области

В системе «Эйдос» есть возможность построения моделей, соответствующих многослойным нейронным сетям [18].

Есть также возможность визуализации любого одного слоя нелокальной нейронной сети (режим 4.4.11 системы «Эйдос»).

Такой слой в наглядной форме отражает силу и направление влияния рецепторов ряда нейрона на степень их активации/торможения в форме цвета и толщины дендритов.

Нейроны на изображении слоя нейронной сети расположены слева направо в порядке убывания модуля суммарной силы их детерминации

рецепторами, т.е. слева находятся результаты, наиболее жестко обусловленные действующими на них значениями факторов, а справа – менее жестко обусловленные. В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.5.2. Конкретное решение задачи в данной работе

На рисунке 30 мы видим, что пригодность воды наиболее сильно детерминирована значениями факторов Turbidity и pH. Кроме того видно, что значения Hardness и Conductivity могут сильно препятствовать пригодности воды.

4.4.11. Отображение Парето-подмножеств одного слоя нелокальной нейронной сети в системе "Эйдос"

Выбор нелокальных нейронов (классов) для визуализации в нейросети

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	POTABILITY-Малое
<input type="checkbox"/>	2	POTABILITY-Большое

Помощь Максимальное количество отображаемых нейронов: ClearSet Диапазон кодов отображаемых нейронов: -
 Максимальное количество отображаемых связей: Диапазон кодов отображаемых рецепторов: -

Подготовка визуализации нейрона:1 "POTABILITY-Малое" в модели:4 "INF1"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
45	TURBIDITY-Очень большое	0.067
1	PH-Очень малое	0.049
15	SOLIDS-Очень большое	0.047
41	TURBIDITY-Очень малое	0.047
4	PH-Большое	0.041
36	TRIHALOMETHANES-Очень малое	0.029
11	SOLIDS-Очень малое	0.020
7	HARDNESS-Малое	0.013
23	SULFATE-Среднее	0.012
30	CONDUCTIVITY-Большое	0.010

ТОРМОЗЯЩИЕ рецепторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
6	HARDNESS-Очень малое	-0.118
30	CONDUCTIVITY-Очень большое	-0.100
35	ORGANIC_CARBON-Очень большое	-0.033
40	TRIHALOMETHANES-Очень большое	-0.025
25	SULFATE-Очень большое	-0.025
16	CHLORAMINES-Очень малое	-0.025
14	SOLIDS-Большое	-0.025
3	PH-Среднее	-0.021
24	SULFATE-Большое	-0.021
42	TURBIDITY-Малое	0.010

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

НейроСеть Abs Prc1 Prc2 **Inf1** Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Максимальное количество отображаемых рецепторов:
 Отображать связи с интенсивностью >= % от макс.:

Сортировать связи:
 по модулю информативности
 по информативности и знаку

Отображать наименования:
 нейронов
 рецепторов

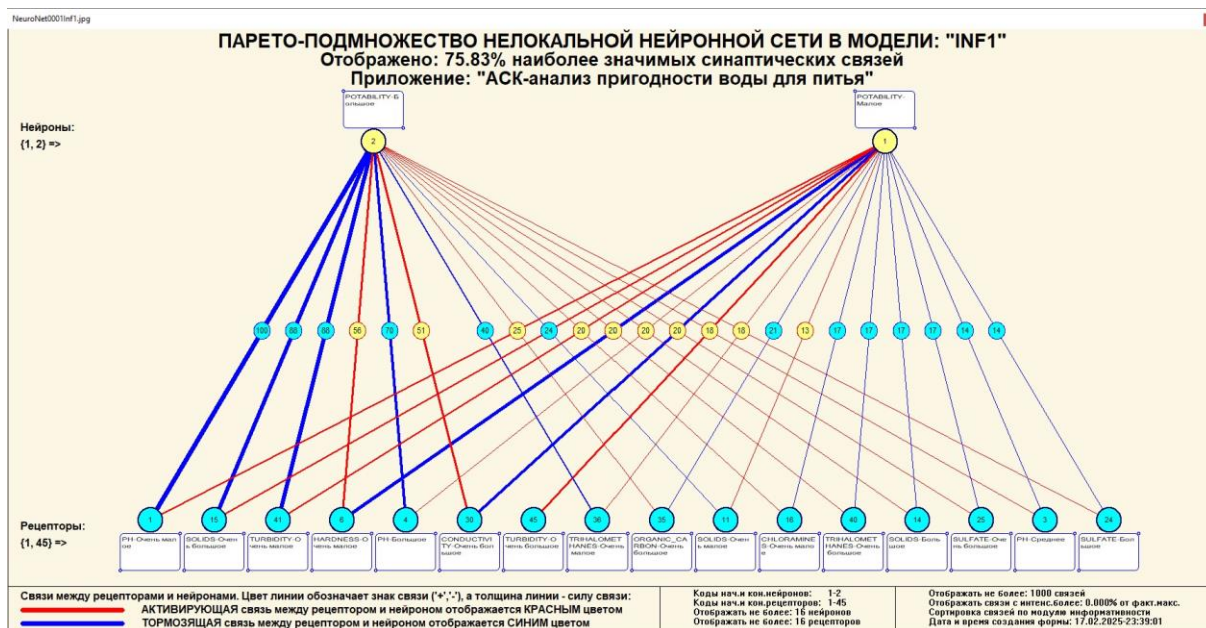


Рисунок 30. Нейронная сеть в СК-модели INF1

3.8.6. 3D-интегральные когнитивные карты

3.8.6.1. Универсальная постановка задачи, не зависящая от предметной области

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов вверху и когнитивной диаграммы значений факторов внизу и соединяющего их одного слоя нейронной сети (режим 4.4.12 системы «Эйдос»).

3.8.6.2. Конкретное решение задачи в данной работе

4.4.12. Отображение Парето-подмножеств одного слоя интегральной когнитивной карты в системе "Эйдос"

Выбор нелокальных нейронов (классов) для визуализации в когнитивной карте

Sel	Код	Наименование нелокального нейрона (класса)
<input checked="" type="checkbox"/>	1	POTABILITY-Малое
<input checked="" type="checkbox"/>	2	POTABILITY-Большое

Помощь Максимальное количество отображаемых нейронов: 16 ClearSet Диапазон кодов отображаемых нейронов: 1 2
 Максимальное количество отображаемых связей: 1000 Диапазон кодов отображаемых рецепторов: 1 45

Подготовка визуализации нейрона: 1 "POTABILITY-Малое" в модели: 4 "INF1"

АКТИВИРУЮЩИЕ рецепторы и сила их влияния **ТОРМОЗЯЩИЕ рецепторы и сила их влияния**

Код	Наименование фактора и его интервального значения	Сила влияния
45	TURBIDITY-Очень большое	0.067
1	PH-Очень малое	0.049
15	SOLIDS-Очень большое	0.047
41	TURBIDITY-Очень малое	0.047
4	PH-Большое	0.041
36	TRIHALOMETHANES-Очень малое	0.029
11	SOLIDS-Очень малое	0.020
7	HARDNESS-Малое	0.013
23	SULFATE-Среднее	0.012
30	CONDUCTIVITY-Очень большое	0.007

Код	Наименование фактора и его интервального значения	Сила влияния
6	HARDNESS-Очень малое	-0.118
30	CONDUCTIVITY-Очень большое	-0.100
35	ORGANIC_CARBON-Очень большое	-0.033
40	TRIHALOMETHANES-Очень большое	-0.025
25	SULFATE-Очень большое	-0.025
16	CHLORAMINES-Очень малое	-0.025
14	SOLIDS-Большое	-0.025
3	PH-Среднее	-0.021
24	SULFATE-Большое	-0.021
42	TURBIDITY-Малое	-0.010

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Когн. карта Максимальное количество отображаемых рецепторов: 16 Отображать связи с интенсивностью >= % от макс.: 0, 000

Сортировать связи:
 по модулю информативности
 по информативности и знаку

Отображать наименования:
 нейронов
 рецепторов

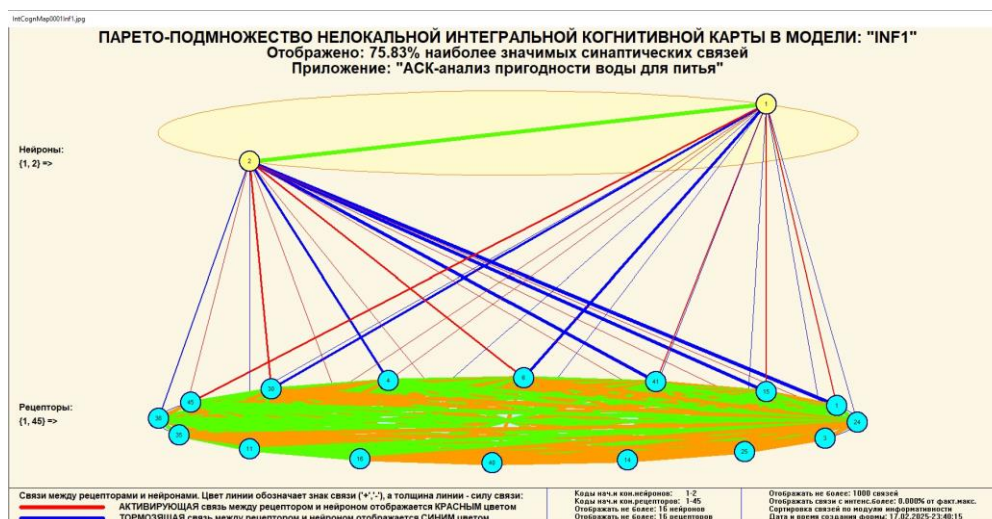


Рисунок 31. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения)

3.8.7.1. Универсальная постановка задачи, не зависящая от предметной области

В 2d-когнитивных диаграммах сравнения классов по системе их детерминации видно, насколько сходны или насколько отличаются друг от друга классы по значениям обуславливающих их факторов.

Однако мы не видим из этой диаграммы, чем именно конкретно сходны и чем именно отличаются эти классы по значениям обуславливающих их факторов.

Это мы можем увидеть из когнитивной диаграммы содержательного сравнения классов, которая отображается в режиме 4.2.3 системы «Эйдос».

2D-интегральные когнитивные карты содержательного сравнения классов являются примерами опосредованных нечетких правдоподобных логических заключений, о которых может быть одним из первых писал Дьердь Пойа [19, 20]. Впервые об автоматизированной реализации рассуждений подобного типа в интеллектуальной системе «Эйдос» написано в 2002 году в работе [2] на странице 521¹¹. Позже об этом писалось в работе [3]¹² и ряде других работ автора, поэтому здесь подробнее рассматривать этот вопрос нецелесообразно.

Пример опосредованных правдоподобных рассуждений.

Допустим нам известно, что один человек имеет голубые глаза, а другой черные волосы. Спрашивается, эти признаки вносят вклад в сходство или в различие этих двух людей? В АСК-анализе и системе «Эйдос» этот вопрос решается так. В модели на основе кластерно-

¹¹ https://www.elibrary.ru/download/elibrary_18632909_64818704.pdf, Таблица 7. 17, стр. 521

¹² <http://ej.kubagro.ru/2013/07/pdf/15.pdf>, стр.44.

конструктивного анализа классов и значений факторов (признаков) известно, насколько те или иные признаки сходны или отличаются по их влиянию на объект моделирования. Поэтому понятно, что человек с голубыми глазами вероятнее всего блондин, а брюнет, скорее всего, имеет темные глаза. Так что понятно, что эти признаки вносят вклад в различие этих двух людей.

3.8.7.2. Конкретное решение задачи в данной работе

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации приведены ниже на рисунках 32. Всего системой в данной модели генерируется 9 форм содержательного сравнения классов. Так как каждый из 3 классов сравнивается со всеми остальными, в т.ч. с собой, то всего получается $3^2=9$ подобных диаграмм. На рисунках 34 приводятся некоторые из этих диаграмм. Пользователь при желании всегда может скачать и установить систему «Эйдос» с сайта разработчика: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) скачать и установить интеллектуальное облачное Эйдос-приложение №426 и получить в нем все выходные формы, как это описано в данной статье.

4.2.3. Когнитивные диаграммы классов. Задание параметров генерации выходных форм

Выбор классов для когнитивной диаграммы

Задайте коды двух классов, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование класса
0	ВСЕ КЛАССЫ
1	POTABILITY-Малое
2	POTABILITY-Большое

Выбор кода класса левого инф.портрета Выбор кода класса правого инф.портрета

Выбор способа фильтрации признаков в информационных портретах когнитивной диаграммы

Задайте коды двух описательных шкал, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование описательной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ ОПИСАТЕЛЬНЫ	1	45
1	PH	1	5
2	HARDNESS	6	10
3	SOLIDS	11	15
4	CHLORAMINES	16	20
5	SULFATE	21	25

Выбор кода описательной шкалы левого инф.портрета Выбор кода описательной шкалы правого инф.портрета

Задайте модели, в которых проводить расчеты когнитивных диаграмм

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей: [Помощь](#)

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Класс для левого инф.портрета: [0] ВСЕ КЛАССЫ
Класс для правого инф.портрета: [0] ВСЕ КЛАССЫ
Описат.шкала для левого инф.портрета: [0] ВСЕ ОПИСАТЕЛЬНЫ
Описат.шкала для правого инф.портрета: [0] ВСЕ ОПИСАТЕЛЬНЫ
Модели, заданные для расчета: Inf1

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

Ok Cancel

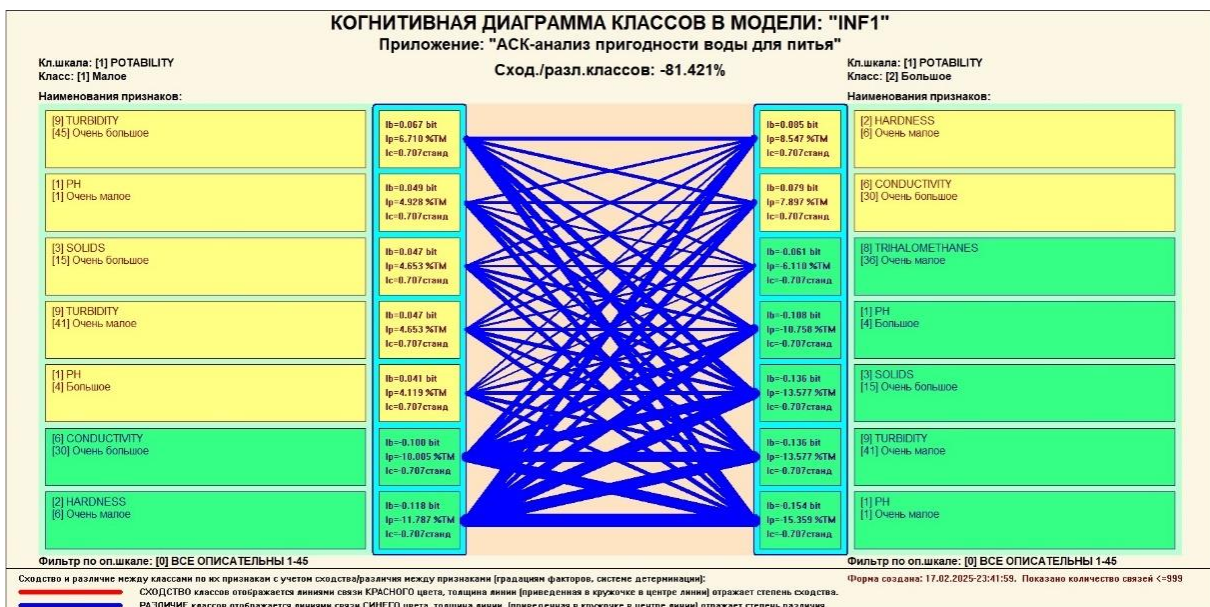
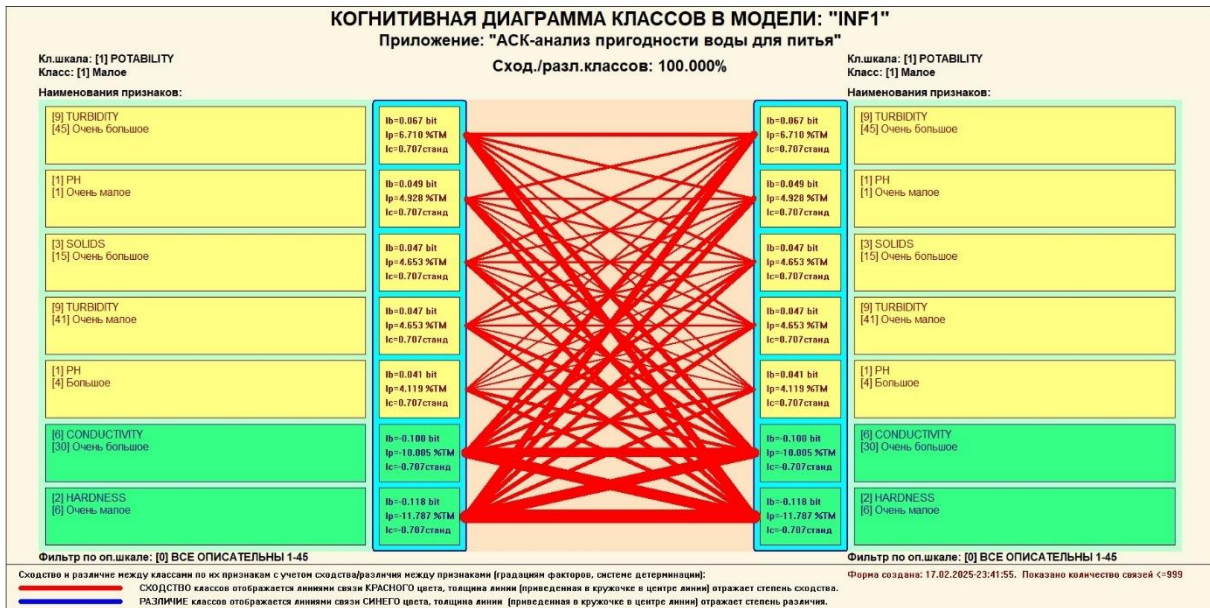


Рисунок 32. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF1

3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения)

3.8.8.1. Универсальная постановка задачи, не зависящая от предметной области

Из 2d-когнитивных диаграммах сравнения значений факторов по их влиянию на объект моделирования, т.е. на его переходы в состояния, соответствующие классам вполне понятно, насколько сходны или отличаются любые два значения факторов по их смыслу.

Напомним, что смысл событий, согласно концепции смысла Шенка-Абельсона, используемой в АСК-анализе, состоит в знании причин и последствий этих событий [21].

Однако из этой диаграммы не видно, чем именно конкретно содержательно сходны или отличаются значения факторов по их смыслу.

Это видно из когнитивных диаграмм, которые можно получить в режиме 4.3.3 системы «Эйдос».

3.8.8.2. Конкретное решение задачи в данной работе

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их силе и направлению их влияния на переход объекта моделирования в будущие состояния, соответствующие классам, приведены ниже на рисунках 33:

4.3.3. Когнитивные диаграммы признаков. Задание параметров генерации выходных форм

Выбор признаков для когнитивной диаграммы

Задайте коды двух признаков, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование признака
0	ВСЕ ПРИЗНАКИ
1	RH-Очень малое
2	RH-Малое
3	RH-Среднее
4	RH-Большое
5	RH-Очень большое
6	HARDNESS-Очень малое

Выбор кода признака левого инф.портрета Выбор кода признака правого инф.портрета

Выбор способа фильтрации классов в информационных портретах когнитивной диаграммы

Задайте коды двух классификационных шкал, для левого и правого информационных портретов когнитивной диаграммы по очереди выбирая их курсором в таблице и кликая на соответствующей кнопке ниже нее

Код	Наименование классификационной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ	1	2
1	POTABILITY	1	2

Выбор кода классификационной шкалы левого инф.портрета Выбор кода классификационной шкалы правого инф.портрета

Задайте модели, в которых проводить расчеты когнитивных диаграмм:

Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7

Задайте max количество отображаемых связей:

99999 Помощь

В диалоге заданы следующие параметры расчета когнитивных диаграмм:

Признак для левого инф.портрета: [0] ВСЕ ПРИЗНАКИ
 Признак для правого инф.портрета: [0] ВСЕ ПРИЗНАКИ
 Классиф.шкала для левого инф.портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
 Классиф.шкала для правого инф.портрета: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ
 Модели, заданные для расчета: Inf1

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа

Ok Cancel

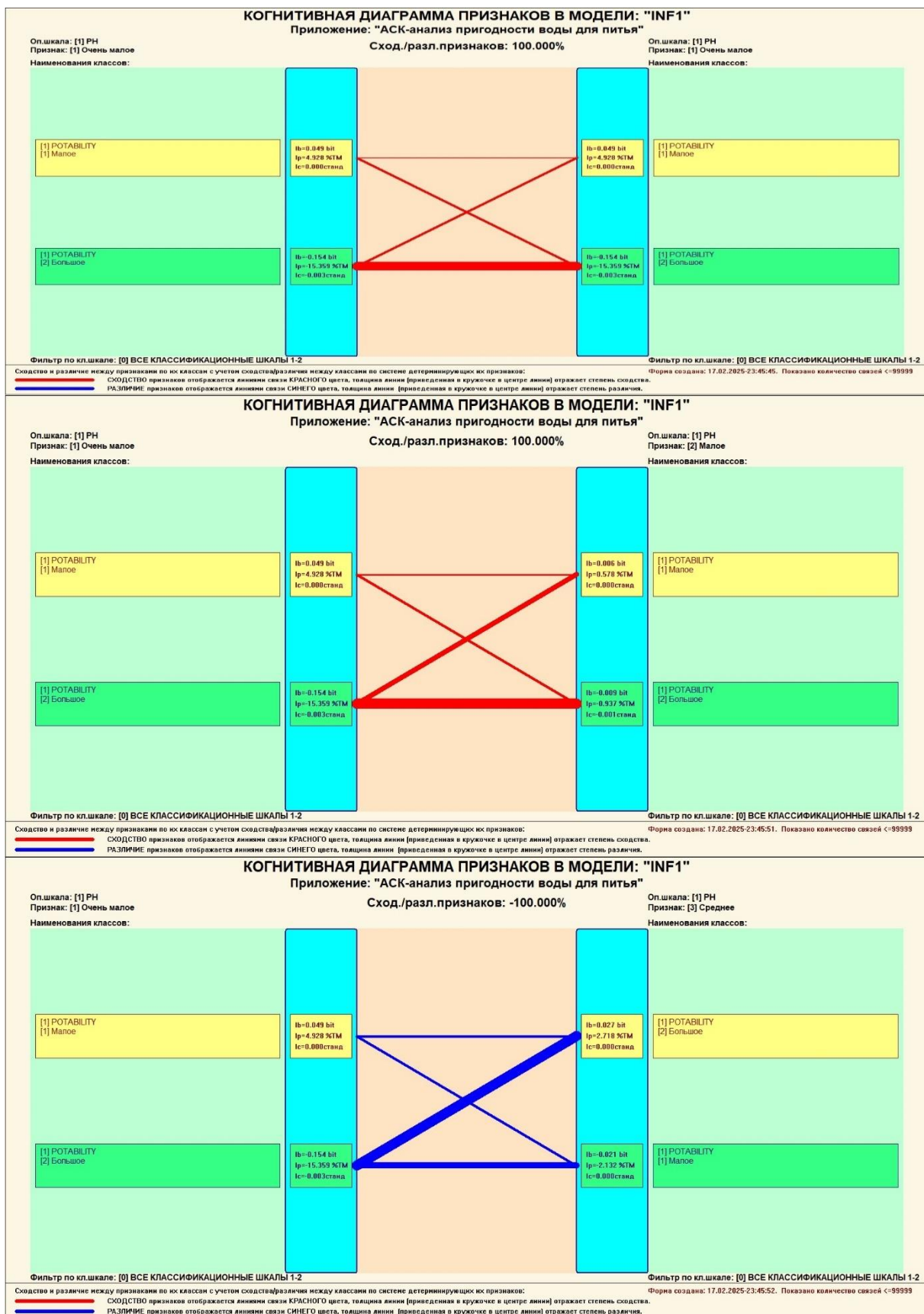


Рисунок 33. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF1

3.8.9. Когнитивные функции

3.8.9.1. Универсальная постановка задачи, не зависящая от предметной области

Когнитивные функции являются обобщением классического математического понятия функции на основе системной теории информации и предложены Е.В.Луценко в 2005 году [3, 22, 23].

Когнитивные функции отображают, какое количество информации содержится в градациях описательной шкалы о переходе объекта моделирования в состояния, соответствующие градациям классификационной шкалы. При этом в статистических и системно-когнитивных моделях в каждой градации описательной шкалы содержится информация обо всех градациях классификационной шкалы, т.е. *каждому значению аргумента соответствуют все значения функции, но соответствуют в разной степени, причем как положительной, так и отрицательной, которая отображается цветом.*

Когнитивные функции являются одним из наиболее мощных и наглядных средств когнитивной графики, имеющихся в системе «Эйдос», позволяющих отобразить силу и направление влияния каждого значения каждого фактора на переход объекта моделирования в каждое из будущих состояний, соответствующих классам.

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5. Первая экранная форма данного режима представляет собой краткий хелп, поясняющий смысл понятия «Когнитивная функция», а также позволяющий выйти на экранную форму системы «Эйдос» с действующими гиперссылками на работы по когнитивным функциям, а также страницы сайта автора со списком этих работ и работ автора по выявлению, представлению и использованию знаний, логике и методологии научного познания.

Необходимо отметить, что модели системы «Эйдос» – это *феноменологические* модели, отражающие *эмпирические* закономерности в фактах обучающей выборки, т.е. они отражают причинно-следственные связи, но не отражают *механизма детерминации*, а только сам факт и характер детерминации [17, 23, 24]. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [24].

3.8.9.2. Конкретное решение задачи в данной работе

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5. Количество когнитивных функций равно количеству сочетаний описательных и классификационных шкал. В модели, рассматриваемой в данной работе, есть 4 описательных шкалы и 1 классификационная, поэтому получается 4 когнитивных функции:

4.5. Генерация, визуализация и запись когнитивных функций системы "Эйдос"

— □ ×

— Задайте статистические и/или системно-когнитивные модели для генерации когнитивных функций:

Статистические базы:

- 1. ABS - частный критерий: количество встреч сочетаний: "класс-признак" у объектов обуч.выборки
- 2. PRC1 - частный критерий: усл. вероятность i-го признака среди признаков объектов j-го класса
- 3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса

Системно-когнитивные модели (Базы знаний):

- 4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1
- 5. INF2 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC2
- 6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абс. частотами
- 7. INF4 - частный критерий: ROI (Return On Investment); вероятности из PRC1
- 8. INF5 - частный критерий: ROI (Return On Investment); вероятности из PRC2
- 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC1
- 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; вероятности из PRC2

— Задайте виды когнитивных функций для генерации, визуализации и записи:

- 1. Сетка триангуляции Делоне без цветовой заливки.
- 2. Сглаженные изолинии триангуляции Делоне без цветовой заливки.
- 3. Сетка триангуляции Делоне с цветовой заливкой.
- 4. Сглаженные изолинии триангуляции Делоне с цветовой заливкой.
- 5. Сглаженная цветочная заливка изолиний с заданным количеством градаций цвета.

— Задайте дополнительные параметры визуализации когнитивных функций:

- Соединять ли точки с максимальным количеством информации линией КРАСНОГО цвета?
- Соединять ли точки с минимальным количеством информации линией СИНЕГО цвета?

Задайте количество градаций уровня (цвета и изолиний) когнитивных функций:

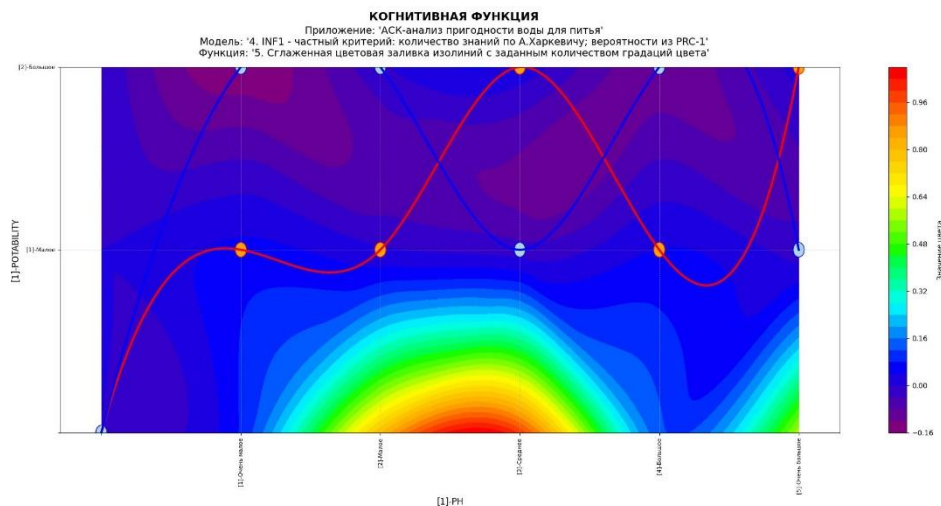
Задайте количество пикселей на дюйм в изображениях когнитивных функций:

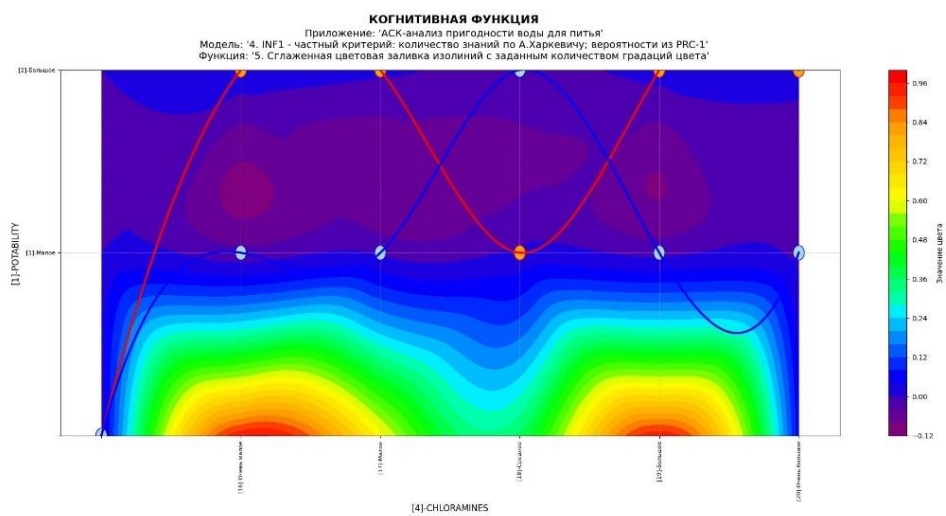
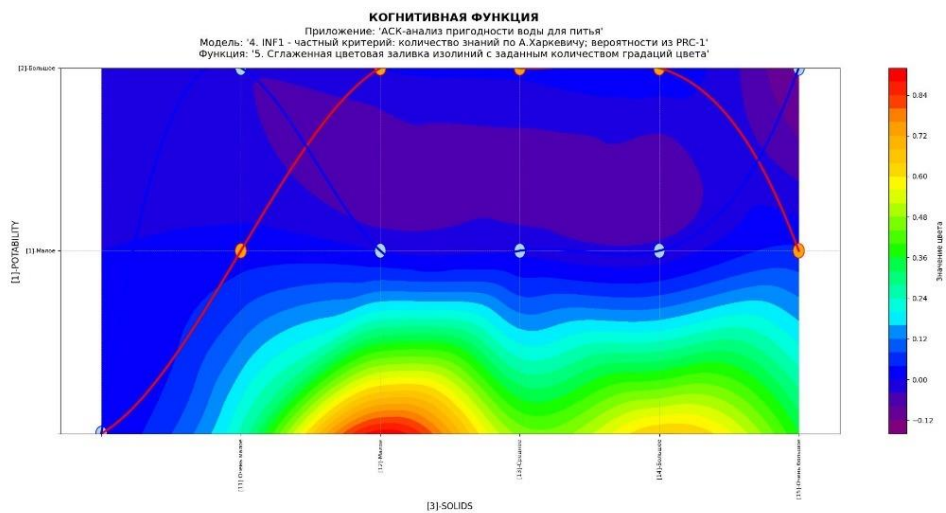
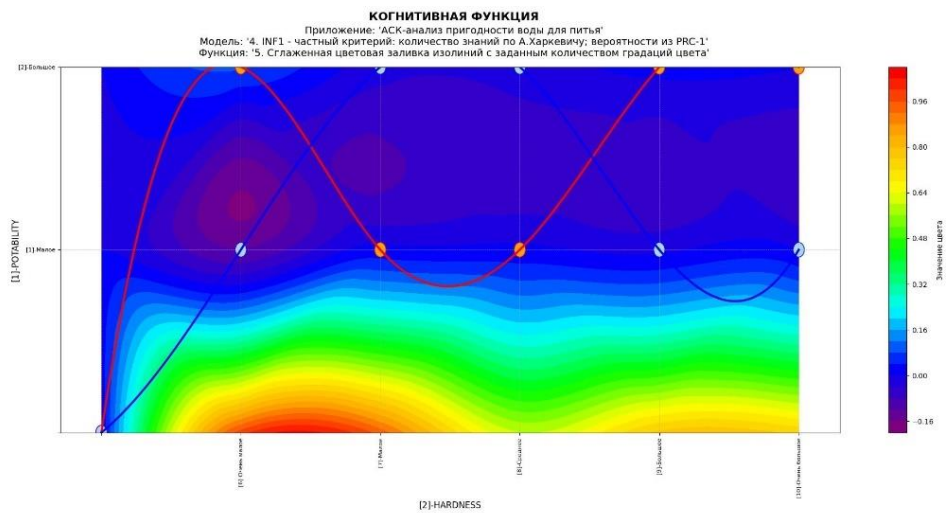
Задайте паузу в секундах между визуализациями когнитивных функций:

Задайте размер шрифта для наименований градаций шкал X и Y:

Визуализация когнитивных функций new **Визуализация когнитивных функций old**

Работы по когнитивным функциям-1 Работы по когнитивным функциям-2





4.5. Визуализация когнитивных функций

Что такое когнитивная функция:

Визуализация прямых, обратных, позитивных, негативных, полностью и частично редуцированных когнитивных функций. Когнитивная функция представляет собой графическое отображение силы и направления влияния различных значений некоторого фактора на переходы объекта управления в будущие состояния, соответствующие классам. Когнитивные функции представляют собой новый перспективный инструмент отражения и наглядной визуализации закономерностей и эмпирических законов. Разработка содержательной научной интерпретации когнитивных функций представляет собой способ познания природы, общества и человека. Когнитивные функции могут быть: прямые, отражающие зависимость классов от признаков, обобщающие информационные портреты признаков; обратные, отражающие зависимость признаков от классов, обобщающие информационные портреты классов; позитивные, показывающие чему способствуют система детерминации; негативные, отражающие чему препятствуют система детерминации; средневзвешенные, отражающие совокупное влияние всех значений факторов на поведение объекта (причем в качестве весов наблюдений используется количество информации в значении аргумента о значениях функции) различной степенью редукции или степенью детерминации, которая отражает в графической форме (в форме полосы) количество знаний в аргументе о значении функции и является аналогом и обобщением доверительного интервала. Если отобразить подматрицу матрицы знания, отображая цветом силу и направление влияния каждой градации некоторой описательной шкалы на переход объекта в состояния, соответствующие классам некоторой классификационной шкалы, то получим нередуцированную когнитивную функцию. Когнитивные функции являются наиболее развитым средством изучения причинно-следственных зависимостей в моделируемой предметной области, предоставляемым системой "Эйдос". Необходимо отметить, что на вид функций влияния математической моделью АСК-анализа не накладывается никаких ограничений, в частности, они могут быть и не дифференцируемые.

Луценко Е.В. Метод визуализации когнитивных функций - новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. - Краснодар: КубГАУ, 2011. - №03(67). С. 240 - 282. - Шифр Информрегистра: 0421100012\0077. . 2,688 у.п.л. - Режим доступа: <http://ej.kubaqo.ru/2011/03/pdf/18.pdf>

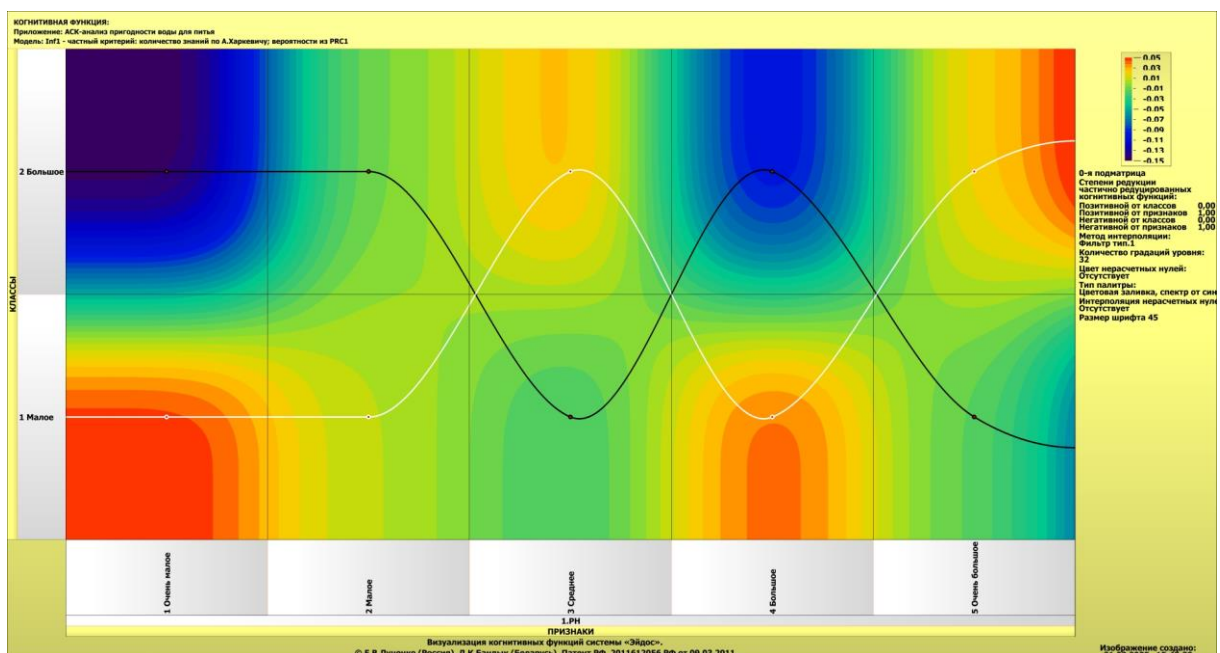
Задайте нужный режим:

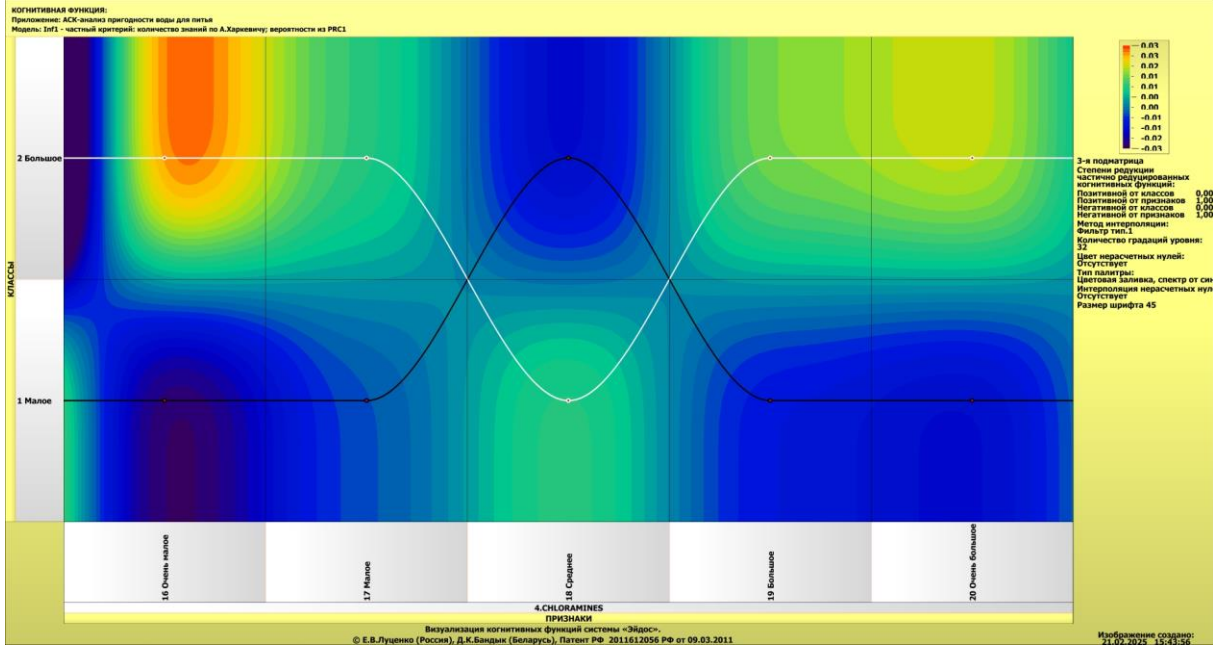
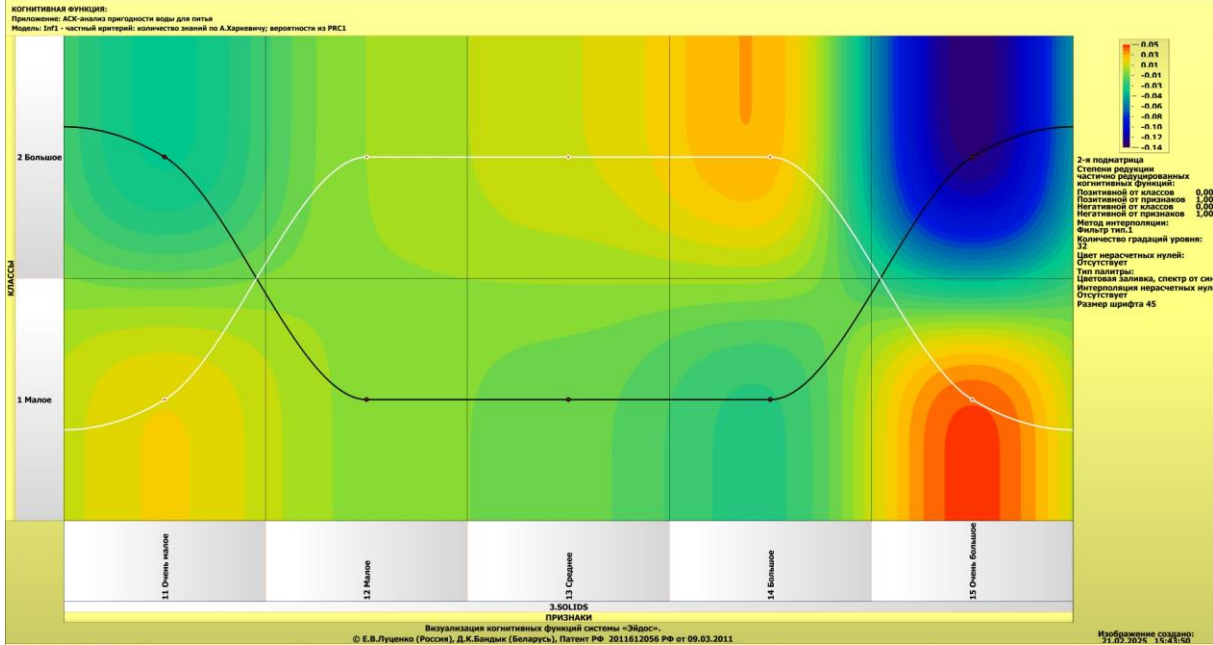
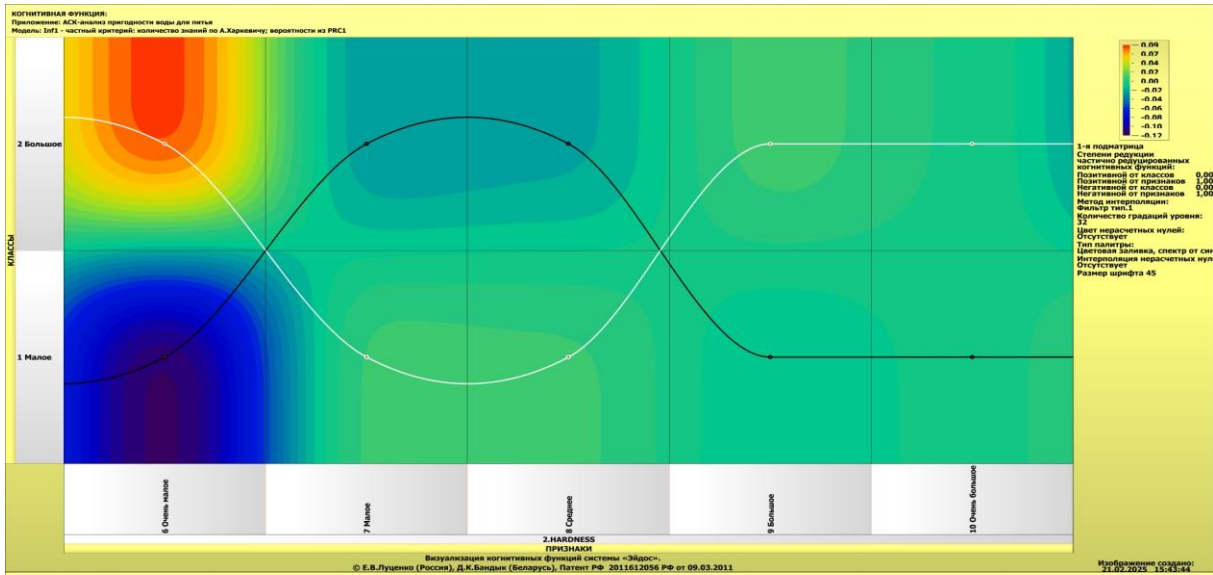
Визуализации когнитивных функций

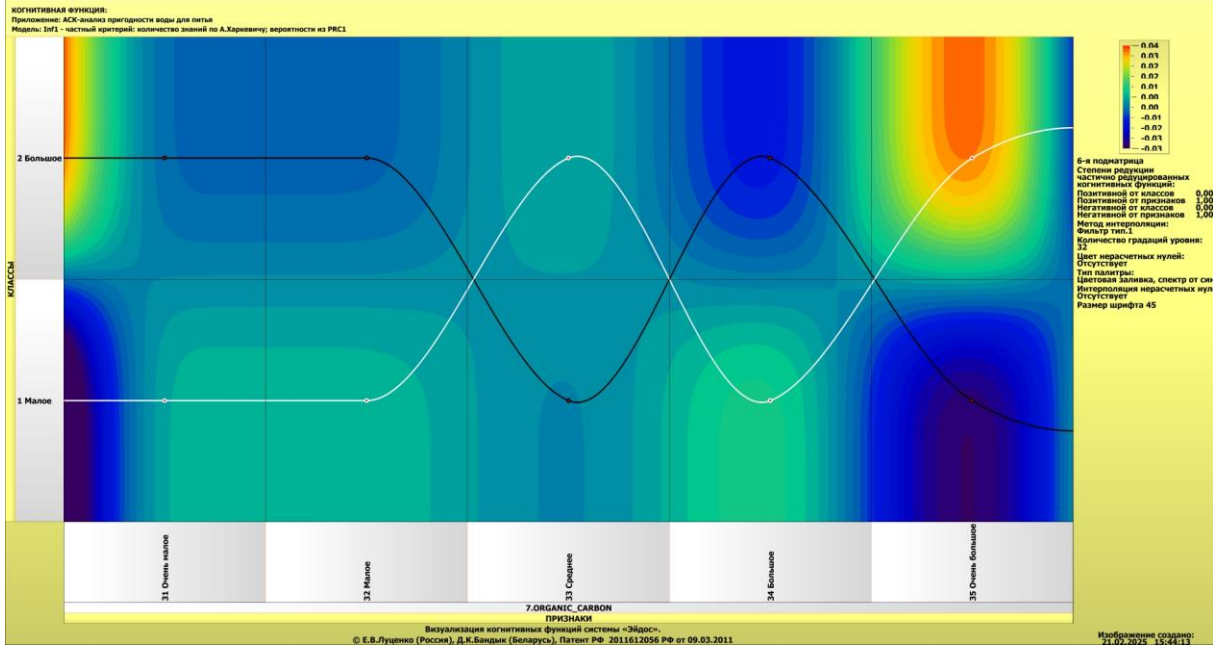
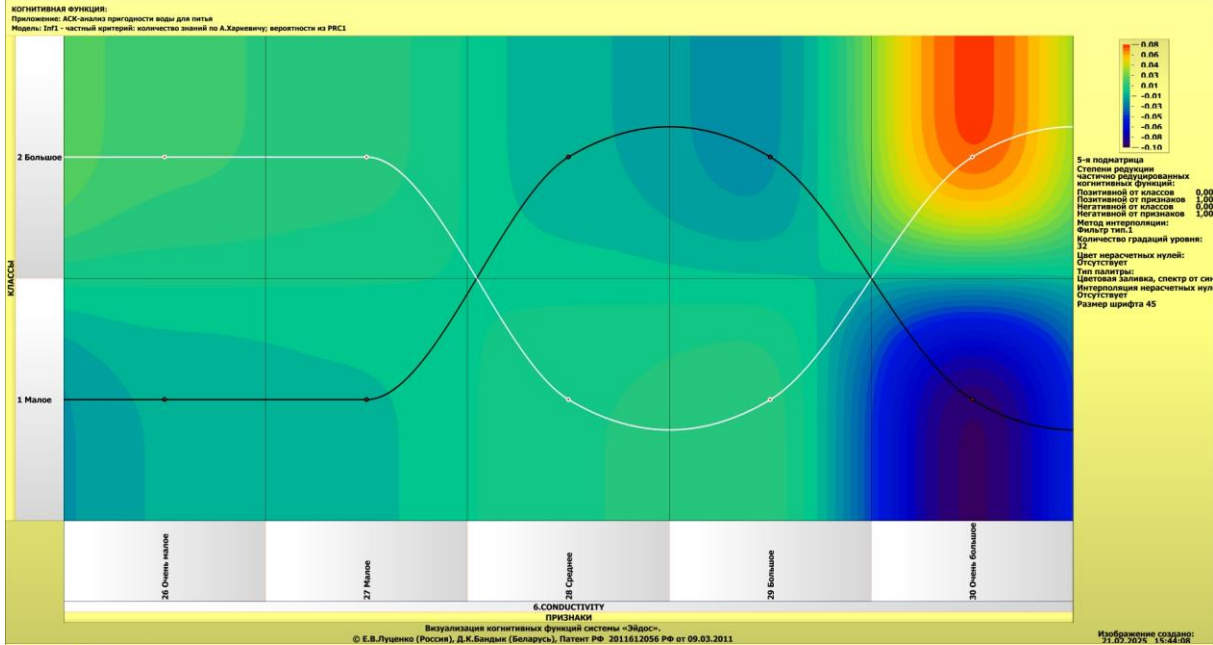
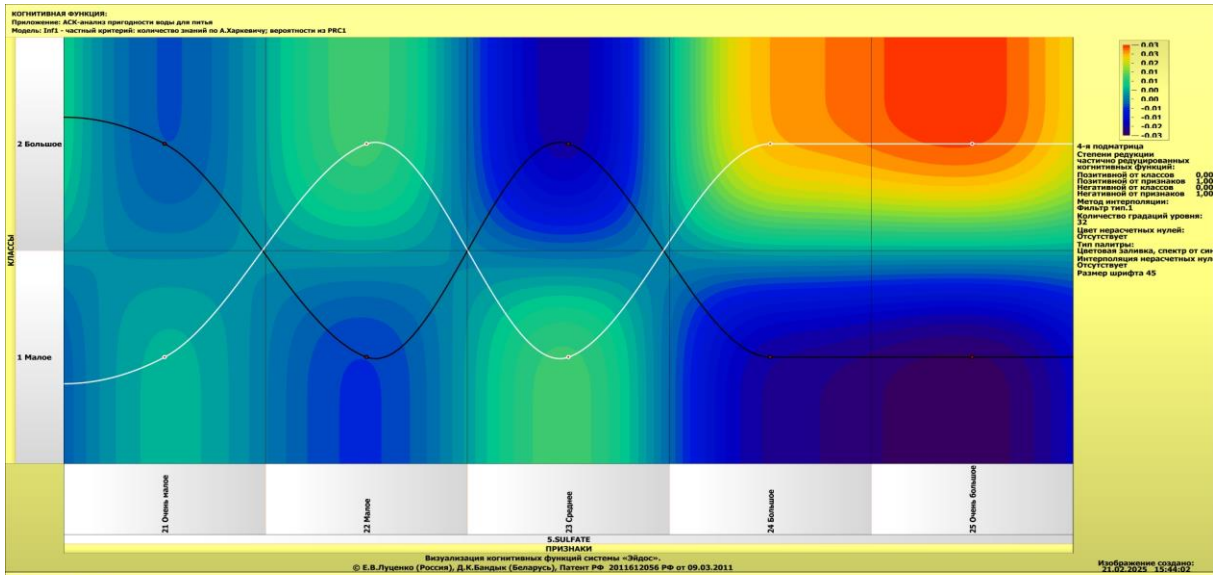
Литератур.ссылки на работы по когнитивным функциям

Литератур.ссылки на работы по когнитивным функциям

Литератур.ссылки на работы по управлению знаниями







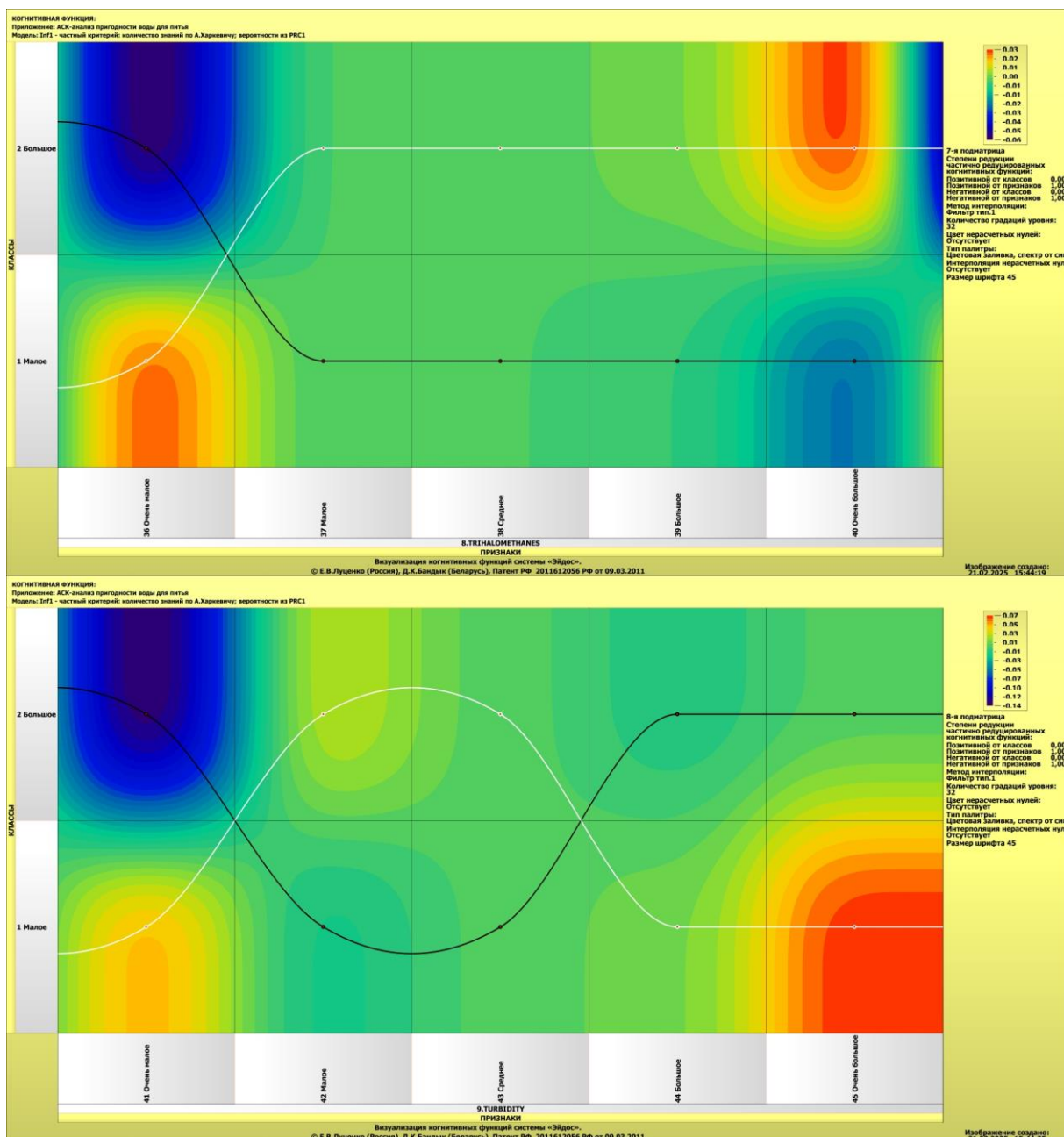


Рисунок 34. Примеры когнитивных функций в СК-модели INF1

Как уже отмечалось содержательное объяснение когнитивных функций на теоретическом уровне познания – это дело специалистов в той предметной области, к которой относится предмет моделирования [24].

3.8.10. Значимость описательных шкал и их градаций

3.8.10.1. Универсальная постановка задачи, не зависящая от предметной области

В АСК-анализе все факторы рассматриваются с одной единственной точки зрения: сколько информации содержится в их значениях о переходе объекта моделирования и управления, на который они действуют, в

определенное будущее состояние, описываемое классом (градация классификационной шкалы), и при этом сила и направление влияния всех значений факторов на объект измеряется в одних общих для всех факторов единицах измерения: единицах количества информации [6].

Значимость (селективная сила) градаций описательных шкал в АСК-анализе – это вариабельность частных критериев в статистических и системно-когнитивных моделях, например в модели Inf1, это вариабельность информативностей (режим 3.7.5 системы «Эйдос») (рисунок 37):-

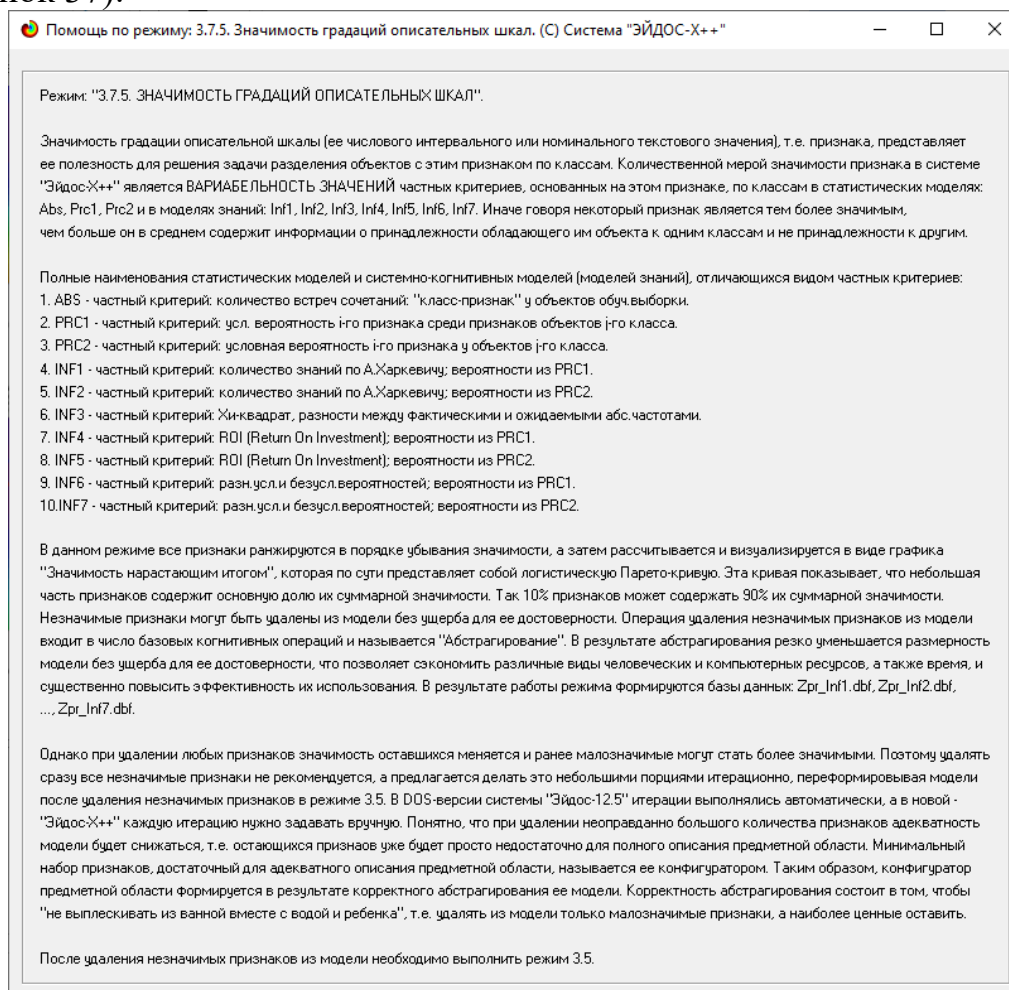


Рисунок 35. Help режима 3.7.5, поясняющий смысл значимости градаций описательных шкал

Значимость всей описательной шкалы является средним от степени значимости ее градаций (режим 3.7.4 системы «Эйдос»).

Если рассортировать все градации факторов (признаки) в порядке убывания селективной силы и получить сумму селективной силы системы значений факторов нарастающим итогом, то получим Парето-кривую.

3.8.10.2. Конкретное решение задачи в данной работе

На рисунке 36 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF1:

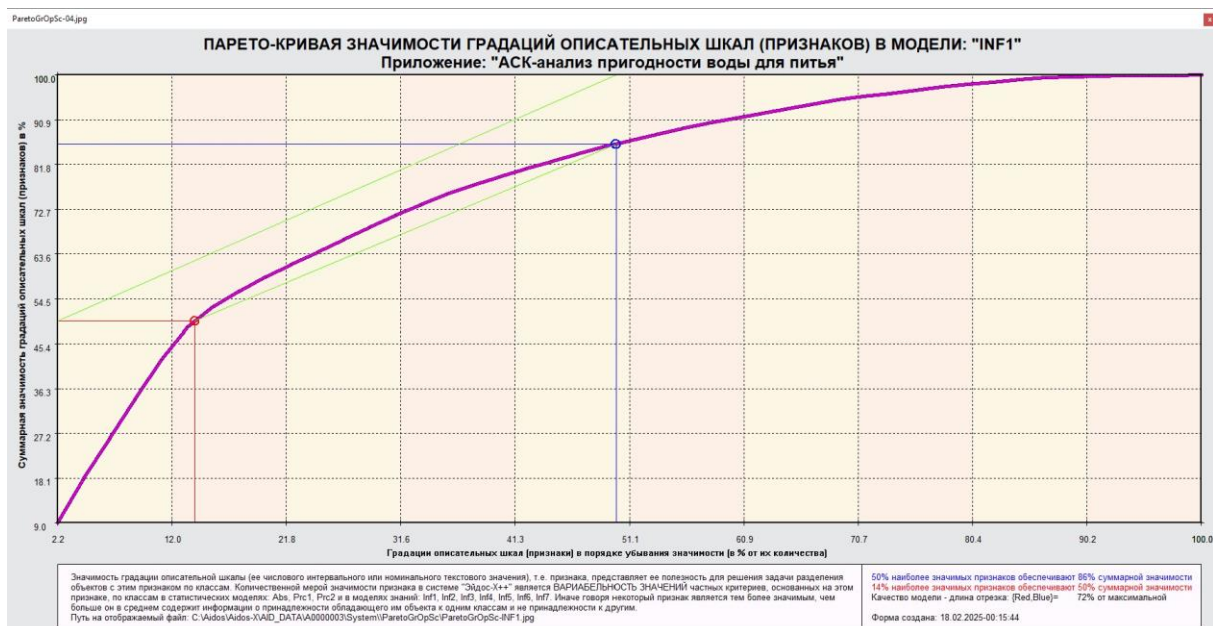


Рисунок 36. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF1

Из рисунка 36 видно, что 43% наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 58% суммарного влияния. На рисунке 37 система «Эйдос» привела рейтинг качества статистических и системно-когнитивных моделей. Кроме того на этом рисунке приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях.

В таблице 16 представлены исходные данные для построения кумулятивной кривой на рисунке 36. Из таблицы 15 видно, какую долю от суммарного влияния на переход объекта моделирования в будущее состояния, соответствующие классам, имеет каждое значение каждого фактора.

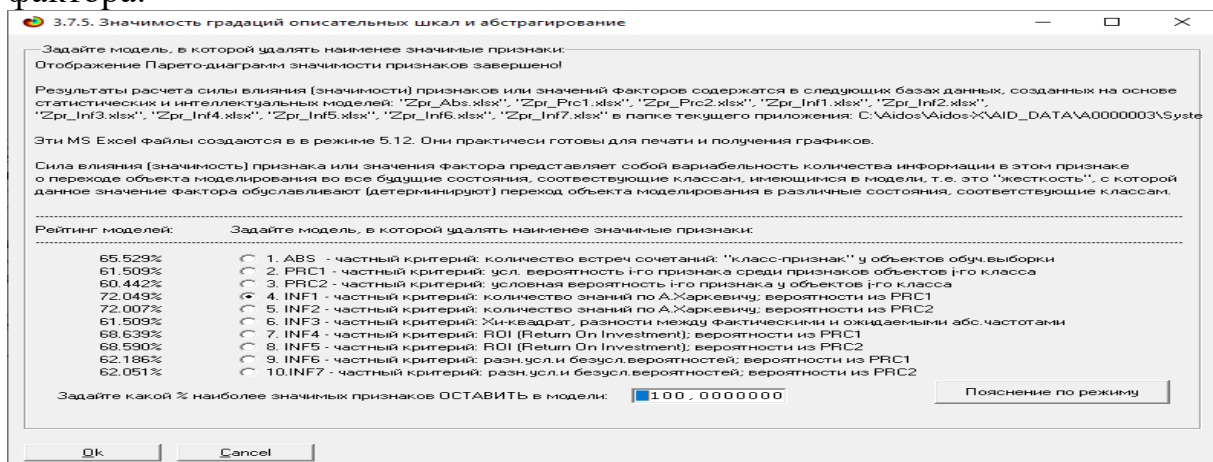


Рисунок 37. Рейтинг качества статистических и системно-когнитивных моделей и имена Excel-файлов с информацией о силе и направлении влияния значений факторов в этих моделях

Таблица 15 – Сила влияния значений факторов на поведение объекта моделирования в СК-модели INF1

NUM	NUM_P RC	KOD_A TR	NAME_ATR	KOD_OP SC	ZNACH_ ATR	ZN_ATR NIT	ZNACH_P RC	ZN_PRC NIT
1	2,22222 22	6	HARDNESS-Очень малое	2	0,1437836	0,1437836	9,0360536	9,036053 6
2	4,44444 44	1	PH-Очень малое	1	0,1434486	0,2872322	9,0150006	18,05105 41
3	6,66666 67	15	SOLIDS-Очень большое	3	0,1289055	0,4161377	8,1010421	26,15209 63
4	8,88888 89	41	TURBIDITY-Очень малое	9	0,1289055	0,5450432	8,1010421	34,25313 84
5	11,1111 111	30	CONDUCTIVITY- Очень большое	6	0,1265827	0,6716259	7,9550662	42,20820 46
6	13,3333 333	4	PH-Большое	1	0,1051955	0,7768214	6,6109916	48,81919 62
7	15,5555 556	36	TRIALOMETHANES -Очень малое	8	0,0635113	0,8403327	3,9913558	52,81055 20
8	17,7777 778	35	ORGANIC_CARBON- Очень большое	7	0,0500763	0,8904090	3,1470357	55,95758 78
9	20,0000 000	45	TURBIDITY-Очень большое	9	0,0474446	0,9378536	2,9816471	58,93923 48
10	22,2222 222	14	SOLIDS-Большое	3	0,0401362	0,9779898	2,5223520	61,46158 69
11	24,4444 444	16	CHLORAMINES- Очень малое	4	0,0401362	1,0181260	2,5223520	63,98393 89

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Zpr_Inf1.xlsx

На экранной форме рисунка 38 приведены имена Excel-файлов с информацией о силе влияния факторов в разных моделях.

В таблице 16 приведена информация о силе влияния факторов на переход объекта моделирования в различные будущие состояния, соответствующие классам, в системно-когнитивной модели INF1.

3.8.11. Степень детерминированности классов и классификационных шкал

3.8.11.1. Универсальная постановка задачи, не зависящая от предметной области

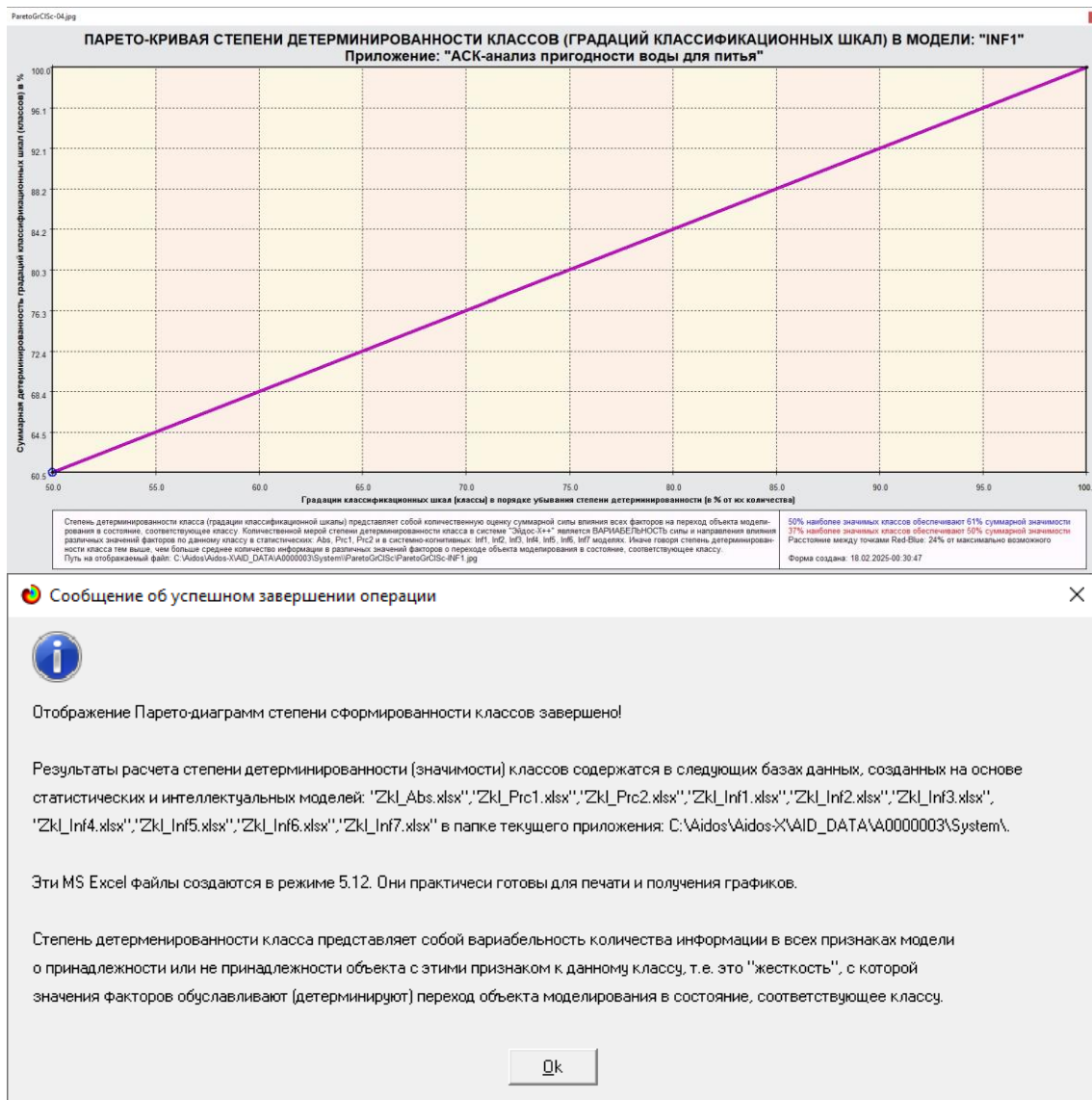
Степень детерминированности (обусловленности) класса в системе «Эйдос» количественно оценивается *степенью вариабельности значений факторов* (градаций описательных шкал) в колонке матрицы модели, соответствующей данному классу (режим 3.7.3 системы «Эйдос»).

Чем выше степень детерминированности класса, тем более достоверно он прогнозируется по значениям факторов.

Степень детерминированности (обусловленности) всей классификационной шкалы является средним от степени детерминированности ее градаций, т.е. классов (режим 3.7.2 системы «Эйдос»).

3.8.11.2. Конкретное решение задачи в данной работе

На рисунках 38 приведены экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос», содержащие информацию о степени детерминированности (обусловленности) состояний объекта моделирования действующими на него факторами:



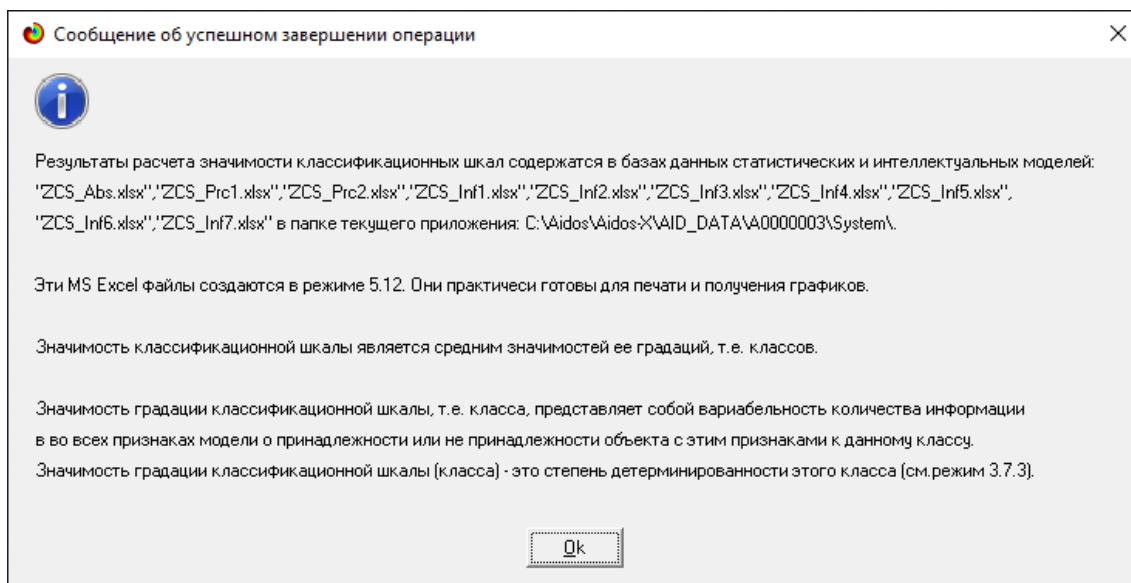


Рисунок 38. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

В таблице 17 представлены исходные данные для построения кумулятивной кривой.

Из таблицы 17 видно, какую долю от суммарной степени детерминированности всех классов имеет каждый класс. Степень обусловленности значениями факторов разных будущих состояний объекта моделирования, соответствующие классам, довольно существенно отличается друг от друга.

Например, 50% наиболее жестко детерминированных классов суммарно обеспечивают примерно 66% степень детерминированности, а 50% суммарной детерминированности обеспечивают 37% наиболее жестко детерминированных классов.

Таблица 16 – Степень детерминированности классов в СК-модели INF1

NUM	NUM_PR C	KOD_C LS	NAME_CLS	KOD_CL SC	ZNACH_C LS	ZN_CLSN IT	ZNACH_P RC	ZN_PRCN IT
1	50,000000 0	1	POTABILITY- Малое	1	121,00000 00	121,00000 00	60,500000 0	60,500000 0
2	100,00000 00	2	POTABILITY- Большое	1	79,000000 0	200,00000 00	39,500000 0	100,00000 00

Источник: c:\Aidos-X\AID_DATA\A0000001\System\Zkl_Abs.xlsx

В таблице 17 приведена информация о степени детерминированности классов значениями факторов в системно-когнитивной модели ABS. Степень детерминированности классификационных шкал является средним от степени детерминированности их градаций.

Поскольку в данном приложении одна классификационная шкала, то рейтинг шкал по силе детерминированности состоит из одной шкалы.

4. DISCUSSION (ОБСУЖДЕНИЕ)

Полученные результаты можно оценить как успешно решающие сформулированную в работе проблему и обеспечивающие достижение поставленной в работе цели. Эти результаты получены путем применения Автоматизированного системно-когнитивного анализа (АСК-анализ) и его программного инструментария – интеллектуальной системы «Эйдос».

Анализ полученных результатов, проведенный в данной работе, полностью согласуется с результатами работы [25], на исходных данных которой они основаны. С другой стороны применение АСК-анализа и системы «Эйдос» весьма существенно расширяет возможности решения задач прогнозирования, принятия решений и исследования моделируемой предметной области, по сравнению с методами, применяемыми в работе [25]. Поэтому есть все основания рекомендовать применение АСК-анализа и системы «Эйдос» для проведения дальнейших углубленных исследований.

Достижением данной работы является:

1. Возможность корректного построения сопоставимых системно-когнитивных моделей предметной области на основе исходных данных, содержащих как лингвистические переменные, так и числовые переменные в различных единицах измерения.

2. Возможность применения системно-когнитивных моделей для решения задач прогнозирования, принятия решений и исследования моделируемой предметной области.

В качестве перспективы продолжения исследований можно было бы рекомендовать существенно увеличить объем исходных данных, количество исследуемых факторов, а также количество классификационных шкал и их градаций (классов) для описания будущих состояний объекта моделирования.

Например, можно было бы исследовать в создаваемых системно-когнитивных моделях, и технологические, и природно-климатические факторы.

Рекомендуется ввести классификационные шкалы, отражающие влияние исследуемых факторов на объект моделирования не только в натуральном выражении (количество и качество различных видов продукции), но и в стоимостном выражении (прибыль и рентабельность, как общая по предприятию, так и в разрезе по видам продукции).

Перспективность и ценность результатов подобных исследований и разработок для теории и практики не вызывает особых сомнений, что подтверждается работами в этой области [1-48].

У желающих есть все возможности для изучения данной работы и для дальнейших исследований с применением АСК-анализа и системы «Эйдос» на своем компьютере.

Для этого надо скачать систему с сайта разработчика по ссылке на странице: <http://lc.kubagro.ru/aidos/Aidos-X.htm>, а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос-приложение №426. По различным аспектам применения данной технологии есть большое количество видео-занятий (около 300), с которыми можно ознакомиться по ссылкам, приведенным на странице: http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf.

5. CONCLUSIONS (ВЫВОДЫ)

В работе приводится краткое описание АСК-анализа и его программного инструментария – интеллектуальной системы «Эйдос».

В ходе исследования были рассмотрены проблемы оценки пригодности воды для питья и возможность применения Автоматизированного системно-когнитивного анализа (АСК-анализа) для решения данной задачи. Обоснована актуальность исследования, связанная с необходимостью быстрого и точного мониторинга качества питьевой воды в условиях роста загрязнения водных ресурсов. Выявлены основные факторы, влияющие на качество воды. Доказана эффективность АСК-анализа. Опробовано применение интеллектуальной системы «Эйдос». Предложены практические рекомендации по внедрению АСК-анализа и системы «Эйдос».

Результаты исследования подтверждают целесообразность использования современных автоматизированных методов анализа для оценки качества питьевой воды, что способствует повышению уровня безопасности водоснабжения и снижению рисков для здоровья населения.

Работа может быть основой для лабораторных работ и научных исследований по применению систем искусственного интеллекта.

REFERENCES (ЛИТЕРАТУРА)

1. Луценко, Е. В. Исследование влияния подсистем различных уровней иерархии на эмерджентные свойства системы в целом с применением АСК-анализа и интеллектуальной системы "Эйдос" (микроструктура системы как фактор управления ее макросвойствами) / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2012. – № 75. – С. 321-363. – EDN OOSCAV.
2. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. – 605 с. – ISBN 5-94672-020-1. – EDN OCZFHС.
3. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет, 2014. – 600 с. – ISBN 978-5-94672-757-0. – EDN RZJXZZ.

4. Работы проф.Е.В.Луценко по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики http://lc.kubagro.ru/aidos/Work_on_emergence.htm
5. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.
6. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>
7. Страница Е.В.Луценко в РИНЦ: https://elibrary.ru/author_profile.asp?id=123162.
8. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.
9. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
10. Луценко, Е. В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы "Эйдос-Х++" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2014. – № 101. – С. 1367-1409. – EDN SZVWRV.
11. Луценко, Е. В. Сценарный и спектральный автоматизированный системно-когнитивный анализ / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2021. – 288 с. – DOI 10.13140/RG.2.2.22981.37608. – EDN ZQLITW.
12. Орлов, А. И. Анализ данных, информации и знаний в системной нечеткой интервальной математике / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2022. – 405 с. – ISBN 978-5-907550-62-9. – DOI 10.13140/RG.2.2.15688.44802. – EDN OQULUW.
13. Луценко, Е. В. Автоматизация функционально-стоимостного анализа и метода "Директ-костинг" на основе АСК-анализа и системы "Эйдос" (автоматизация управления натуральной и финансовой эффективностью затрат без содержательных технологических и финансово-экономических расчетов на основе информационных и когнитивных технологий и теории управления)1 / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 131. – С. 1-18. – DOI 10.21515/1990-4665-131-001. – EDN ZRXVFN.
14. Луценко, Е. В. Системное обобщение принципа Эшби и повышение уровня системности модели объекта познания как необходимое условие адекватности процесса его познания / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2020. – № 163. – С. 100-134. – DOI 10.21515/1990-4665-163-009. – EDN SWKGWY.
15. Луценко, Е. В. Эффективность объекта управления как его эмерджентное свойство и повышение уровня системности как цель управления / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2021. – № 165. – С. 77-98. – DOI 10.13140/RG.2.2.11887.25761. – EDN UMTAMT.
16. Луценко, Е. В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной

системе "Эйдос") / Е. В. Луценко, В. Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2011. – № 71. – С. 27-74. – EDN OIGYBV.

17. Луценко, Е. В. Проблемы и перспективы теории и методологии научного познания и автоматизированный системно-когнитивный анализ как автоматизированный метод научного познания, обеспечивающий содержательное феноменологическое моделирование / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 127. – С. 1-60. – DOI 10.21515/1990-4665-127-001. – EDN YLZTMX.

18. Луценко, Е. В. Системная теория информации и нелокальные интерпретируемые нейронные сети прямого счета / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2003. – № 1. – С. 76-88. – EDN JWXLKT.

19. Работы проф.Е.В.Луценко & С^о по выявлению, представлению и использованию знаний, логике и методологии научного познания: http://lc.kubagro.ru/aidos/Works_on_identification_presentation_and_use_of_knowledge.htm

20. Пойа Дьердь. Математика и правдоподобные рассуждения. // под редакцией С.А.Яновской. Пер. с английского И.А.Вайнштейна., М., Наука, 1975 — 464 с., <http://ilib.mccme.ru/djvu/polya/rassuzhdenija.htm>

21. Луценко, Е. В. Системно-когнитивный анализ как развитие концепции смысла Шенка - Абельсона / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2004. – № 5. – С. 14-35. – EDN JWXMKX.

22. Работы проф.Е.В.Луценко & С^о по когнитивным функциям: http://lc.kubagro.ru/aidos/Works_on_cognitive_functions.htm

23. Луценко, Е. В. Системы представления и приобретения знаний / Е. В. Луценко, В. И. Лойко, В. Н. Лаптев. – Краснодар : Экоинвест, 2018. – 513 с. – ISBN 978-5-94215-415-8. – EDN UZZBLC.

24. Луценко, Е. В. Проблемы и перспективы теории и методологии научного познания и автоматизированный системно-когнитивный анализ как автоматизированный метод научного познания, обеспечивающий содержательное феноменологическое моделирование / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 127. – С. 1-60. – DOI 10.21515/1990-4665-127-001. – EDN YLZTMX.

25. Монографии по АСК-анализу: http://lc.kubagro.ru/aidos/Aidos-X.htm#_Toc128746370

26. Некоторые учебники и учебные пособия проф.Е.В.Луценко: http://lc.kubagro.ru/aidos/Aidos-X.htm#_Toc128746372.

27. Свидетельства Роспатента на систему «Эйдос» и ее подсистемы: http://lc.kubagro.ru/aidos/Aidos-X.htm#_Toc128746371.

28. Тематические подборки публикаций по применению АСК-анализа и системы «Эйдос» в различных предметных областях: <http://lc.kubagro.ru/aidos/Aidos-X.htm>

29. Работы по информационным мерам уровня системности (коэффициентам эмерджентности) и системному обобщению математики: http://lc.kubagro.ru/aidos/Work_on_emergence.htm .

30. Работы по АСК-анализу изображений: http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_images.htm

31. Работы по АСК-анализу текстов: http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_texts.htm

32. Работы по когнитивным функциям:
http://lc.kubagro.ru/aidos/Works_on_cognitive_functions.htm
33. Работы по выявлению, представлению и использованию знаний, логике и методологии научного познания:
http://lc.kubagro.ru/aidos/Work_on_identification_presentation_and_use_of_knowledge.htm
34. Работы по экологии, климатологии и изучению влияния космической среды на различные глобальные процессы на Земле:
http://lc.kubagro.ru/aidos/Work_on_the_study_of_the_influence_of_the_space_environment_on_various_processes_on_Earth.htm
35. Работы по современным информационно-коммуникационным технологиям в научно-исследовательской деятельности и образовании:
http://lc.kubagro.ru/aidos/Information_and_communication_technologies_in_research_activities_and_education.htm
36. Работы по виртуальной реальности:
http://lc.kubagro.ru/aidos/Virtual_reality_publications.htm
37. Работы по когнитивной ветеринарии:
http://lc.kubagro.ru/aidos/Publications_on_cognitive_veterinary_medicine.htm
38. Работы по когнитивной агрономии и когнитивной ампелографии:
http://lc.kubagro.ru/aidos/Works_on_cognitive_agronomy.htm
39. Работы по тематике, связанной с АПК:
http://lc.kubagro.ru/aidos/Work_with_agricultural.htm
40. Работы по наукометрии: http://lc.kubagro.ru/aidos/Works_on_scientometrics.htm
41. Работы о высших формах сознания, перспективах человека, технологии и общества: http://lc.kubagro.ru/aidos/Works_on_higher_forms_of_consciousness.htm
42. Работы по разработке и применению профессиограмм и тестов (психологических, профориентационных, медицинских и ветеринарных):
http://lc.kubagro.ru/aidos/Work_on_the_development_and_application_tests.htm
43. Работы по сценарному автоматизированному системно-когнитивному анализу (сценарный АСК-анализ): http://lc.kubagro.ru/aidos/Works_on_Scenario_ASC-analysis.htm
44. MVP-проект «Внедрение технологий АСК-анализа и системы «Эйдос» для решения задач АПК»: <http://lc.kubagro.ru/aidos/MVP-projects.htm>
45. Кратко об АСК-анализе и системе «Эйдос»:
http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf
46. Ссылки на видео-занятия и проф.Е.В.Луценко в Пермском национальном университете: <https://bigbluebutton.pstu.ru/b/w3y-2ir-ukd-bqn> (2021), <https://bigbluebutton.pstu.ru/b/3kc-n8a-gon-tjz> (2022), в Кубанском государственном университете и Кубанском государственном аграрном университете: <https://disk.yandex.ru/d/knISAD5qzV83Ng?w=1>
47. Луценко, Е. В. Революция начала XXI века в искусственном интеллекте: глубинные механизмы и перспективы / Е. В. Луценко, Н. С. Головин. – Краснодар: Кубанский государственный аграрный университет им. И.Т. Трубилина, 2024. – 394 с. – DOI 10.13140/RG.2.2.17056.56321. – EDN OMIPII.
48. Луценко Е.В. Системы искусственного интеллекта как системы автоматизации процесса научного познания и удвоение номенклатуры научных специальностей путем применения этих систем для исследований в различных направлениях науки / Е.В. Луценко, Н.С. Головин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2024. –

№01(195). С. 74 – 111. – IDA [article ID]: 1952401009. – Режим доступа:
<http://ej.kubagro.ru/2024/01/pdf/09.pdf>, 2,375 у.п.л.