

УДК 004.8

09.03.02 – Информационные системы и технологии

АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ РАЗЛИЧНЫХ ДАННЫХ ГРУПП, ПЕСЕН И УЧАСТНИКОВ BTS

Берёзка Александра

Комаренко Егор

Студенты КубГАУ группы ИТ2204

*Кубанский Государственный Аграрный университет
имени И.Т.Трубилина, Краснодар, Россия*

Данная работа является работой автора по применению Автоматизированного системно-когнитивного анализа (ACK-анализ) для решения широкого спектра задач в области прогнозирования и анализа. В работе решается задача выявления зависимости песен южно-корейской группы BTS и песен ее участников, в работе описаны характеристики песен, таких как: год выпуска альбома, текст, автор, наличие ремиксов и т.д. На основе знания этих зависимостей решаются задачи прогнозирования, принятия решений и исследования моделируемой предметной области путем исследования ее системно-когнитивной модели. Спецификой данной задачи является то, что все независимые переменные (факторы) являются лингвистическими (категориальными) переменными. Поэтому для решения данной задачи применяется лингвистический ACK-анализ, т.е. когнитивная математическая лингвистика. При этом зависимые переменные, т.е. результаты влияния факторов, измеряются в числовых шкалах. Таким образом, в работе строится гибридная модель, включающая как номинальные (текстовые), так и числовые шкалы. Сопоставимость обработки данных разных типов, представленных в разных типах шкал и разных единицах измерения обеспечивается путем метризации номинальных шкал, т.е. повышения их степени формализации до уровня числовых шкал. Приводится краткое описание ACK-анализа и его программного инструментария – интеллектуальной системы «Эйдос». Работа может быть основой для лабораторных работ по применению систем искусственного интеллекта, в частности лингвистического ACK-анализа для решения задач в прогноза и анализа.

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

<http://dx.doi.org/10.21515/1990-4665-181-017>

УДК 004.8

09/03/02 – Information system and technology

AUTOMATED SYSTEMIC COGNITIVE ANALYSIS OF VARIOUS DATA OF BTS GROUPS, SONGS AND PARTICIPANTS

Berezka Aleksandra

Komarenko Egor

Student of KubSAU group IT2204

*Kuban State Agrarian University named after I.T. Trubilin,
Krasnodar, Russia*

This work is the work of the author on the use of Automated System Cognitive Analysis (ASC-analysis) for solving a wide range of problems in the field of forecasting and analysis. The paper solves the problem of identifying the dependence of the songs of the South Korean group BTS and the songs of its members, the paper describes the characteristics of the songs, such as: the year the album was released, the text, the author, the presence of remixes, etc. Based on the knowledge of these dependencies, the problems of forecasting, decision making and research of the modeled subject area are solved by studying its system-cognitive model. The specificity of this task is that all independent variables (factors) are linguistic (categorical) variables. Therefore, to solve this problem, linguistic ASC analysis is used, i.e. cognitive mathematical linguistics. In this case, the dependent variables, i.e. the results of the influence of factors, measured in numerical scales. Thus, a hybrid model is built in the work, including both nominal (text) and numerical scales. The comparability of processing data of different types, presented in different types of scales and different units of measurement, is ensured by metrization of nominal scales, i.e. increasing their degree of formalization to the level of numerical scales. A brief description of the ASC-analysis and its software tools - the intellectual system "Eidos" is given. The work can be the basis for laboratory work on the use of artificial intelligence systems, in particular, linguistic ASC analysis for solving problems in forecasting and analysis.

Keywords: LINGUISTIC ASK-ANALYSIS, LINGUISTIC AUTOMATED SYSTEMIC COGNITIVE ANALYSIS, INTELLIGENT SYSTEM "EIDOS"

СОДЕРЖАНИЕ

1. INTRODUCTION (ВВЕДЕНИЕ)	3
1.1. ОПИСАНИЕ ИССЛЕДУЕМОЙ ПРЕДМЕТНОЙ ОБЛАСТИ	3
1.2. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ	3
1.3. ПРОБЛЕМА, РЕШАЕМАЯ В РАБОТЕ И ЕЕ АКТУАЛЬНОСТЬ.....	3
1.4. ЦЕЛЬ РАБОТЫ	3
2. METHODS (МЕТОДЫ).....	4
2.1. ОБОСНОВАНИЕ ТРЕБОВАНИЙ К МЕТОДУ РЕШЕНИЯ ПРОБЛЕМЫ	4
2.2. ЛИТЕРАТУРНЫЙ ОБЗОР МЕТОДОВ РЕШЕНИЯ ПРОБЛЕМЫ, ИХ ХАРАКТЕРИСТИКА И ОЦЕНКА СТЕПЕНИ СООТВЕТСТВИЯ ОБОСНОВАННЫМ ТРЕБОВАНИЯМ	4
2.3. АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ (АСК-АНАЛИЗ) КАК МЕТОД РЕШЕНИЯ ПРОБЛЕМЫ.....	4
2.4. СИСТЕМА «ЭЙДОС» - ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА	6
2.5. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ	11
3. RESULTS (РЕЗУЛЬТАТЫ).....	13
3.1. ЗАДАЧА-1. КОГНИТИВНАЯ СТРУКТУРИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ. ДВЕ ИНТЕРПРЕТАЦИИ КЛАССИФИКАЦИОННЫХ И ОПИСАТЕЛЬНЫХ ШКАЛ И ГРАДАЦИЙ	13
3.2. ЗАДАЧА-2. ФОРМАЛИЗАЦИЯ ПРЕДМЕТНОЙ ОБЛАСТИ	14
3.3. ЗАДАЧА-3. СИНТЕЗ СТАТИСТИЧЕСКИХ И СИСТЕМНО-КОГНИТИВНЫХ МОДЕЛЕЙ. МНОГОПАРАМЕТРИЧЕСКАЯ ТИПИЗАЦИЯ И ЧАСТНЫЕ КРИТЕРИИ ЗНАНИЙ	23
3.4. ЗАДАЧА-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ	31
3.5. ЗАДАЧА-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ	35
3.6. ЗАДАЧА-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ	36
3.6.1. Интегральный критерий «Сумма знаний»	37
3.6.2. Интегральный критерий «Семантический резонанс знаний».....	37
3.6.3. Важные математические свойства интегральных критериев.....	39
3.6.4. Решение задачи идентификации и прогнозирования в системе «Эйдос».....	39
3.7. ЗАДАЧА-7. ПОДДЕРЖКА ПРИНЯТИЯ РЕШЕНИЙ.....	42
3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ.....	42
3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»	44
3.8. ЗАДАЧА-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ ..	49
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические поменциалы).....	49
3.8.2. Кластерно-конструктивный анализ классов	51
3.8.3. Кластерно-конструктивный анализ значений описательных шкал	54
3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны.....	56
3.8.5. Нелокальная нейронная сеть.....	57
3.8.6. 3D-интегральные когнитивные карты	58
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	60
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	63
3.8.9. Когнитивные функции	65
3.8.10. Значимость описательных шкал и их градаций.....	68
3.8.11. Степень детерминированности классов и классификационных шкал	70
4. DISCUSSION (ОБСУЖДЕНИЕ)	72
5. CONCLUSIONS (ВЫВОДЫ)	73
REFERENCES (ЛИТЕРАТУРА).....	73

1. INTRODUCTION (ВВЕДЕНИЕ)

1.1. Описание исследуемой предметной области

Данная работа посвящена анализу песен с помощью применения Автоматизированного системно-когнитивного анализа (АСК-анализ) для решения широкого спектра задач [1, 2, 3]. В работе решается задача выявления зависимости различных показателей песен южно-корейской группы BTS и песен ее участников, в работе описаны характеристики песен, таких как: год выпуска альбома, текст, автор, наличие ремиксов и т.д. На основе знания этих зависимостей решаются разнообразные задачи прогнозирования, принятия решений и исследования моделируемой предметной области путем исследования ее системно-когнитивной модели (СК-модель).

1.2. Объект и предмет исследования

Объект исследования – выявление зависимостей показателей песни от различных факторов.

Предмет исследования – выявление зависимостей показателей песен от их обработки, тематики и авторов.

1.3. Проблема, решаемая в работе и ее актуальность

Спецификой данной задачи является то, что все независимые переменные, т.е. факторы: название альбома, текст, автор записаны в виде **лингвистических** (категориальных) переменных.

Таким образом, в работе решается проблема построения модели, включающей только номинальные (текстовые) шкалы и обеспечивающей **сопоставимость** обработки данных одного типа, представленных в одном типе шкал.

Решение поставленной **проблемы сопоставимости** при выявлении зависимостей музыковедческих показателей песен от их исполнителя, смысловой нагрузки, тематики. То, что корейская культура набирает популярность по всему миру делает эту работу **актуальной**.

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 и описания метода решения проблемы.

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 очень большой размерности числовой и не числовой природы, измеряемых в различных типах шкал (номинальных, порядковых и числовых) и в различных единицах измерения.

2. Иначе говоря, метод решения проблемы не должен предъявлять жестких требований к исходным данным, которые невозможно выполнить, а должен обеспечивать обработку тех данных, которые реально есть.

3. Метод должен 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 программных систем, *одновременно*:

- находящихся в полном открытом бесплатном доступе;
 - обеспечивающих сопоставимую обработку числовой и текстовой информации в одной модели, дает следующие результаты; показал, что альтернатив Автоматизированному системно-когнитивному анализу и его программному инструментарию – системе «Эйдос» в настоящее время здесь нет [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 (ACK-анализ)» был предложен проф. Е. В. Луценко. На тот момент он вообще не встречался в Internet. Сегодня по соответствующему запросу в Яндексе находится 9 миллионов сайтов с этим сочетанием слов².

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

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);
- 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 665 опубликованных автором научных работ посвящены теоретическим основам АСК-анализа и его практическим применениям в ряде предметных областей. На момент написания данной работы автором опубликовано более 40 монографий, 27 учебных пособий, в т.ч. 3 учебных пособия с грифами УМО и Министерства, получен 32 патент РФ на системы искусственного интеллекта, 346 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 [Scopus](#)³ [6, 7, 8].

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 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 номенклатуре 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.

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/>

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

- Автоматизированный системно-когнитивный анализ числовых и текстовых табличных данных;
- Автоматизированный системно-когнитивный анализ текстовых данных;
- Спектральный и контурный автоматизированный системно-когнитивный анализ изображений;
- Сценарный автоматизированный системно-когнитивный анализ временных и динамических рядов.

Приводить здесь ссылки на все эти работы здесь вряд ли целесообразно. Отметим лишь, что у автора есть личный сайт [7] и страничка в РесечГейт [8], 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.

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 [6]. Сама метризация номинальных шкал достигается *путем* вычисления количества информации, содержащегося в градациях номинальных шкал [5]. Для работы с лингвистическими переменными применяется лингвистический АСК-анализ [4].

2.4. Система «Эйдос» - инструментарий АСК-анализа

Существует много систем искусственного интеллекта. Универсальная когнитивная аналитическая система «Эйдос» отличается от них следующими параметрами:

- 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>). Система «Эйдос» является автоматизированной системой, т.е. предполагает непосредственное участие человека в реальном времени при решении задач идентификации, прогнозирования, принятия решений и исследования предметной области (автоматические системы работают без такого участия человека);
- 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 системы «Эйдос» проф. Е.В.Луценко (отметим, что система «Эйдос» создана

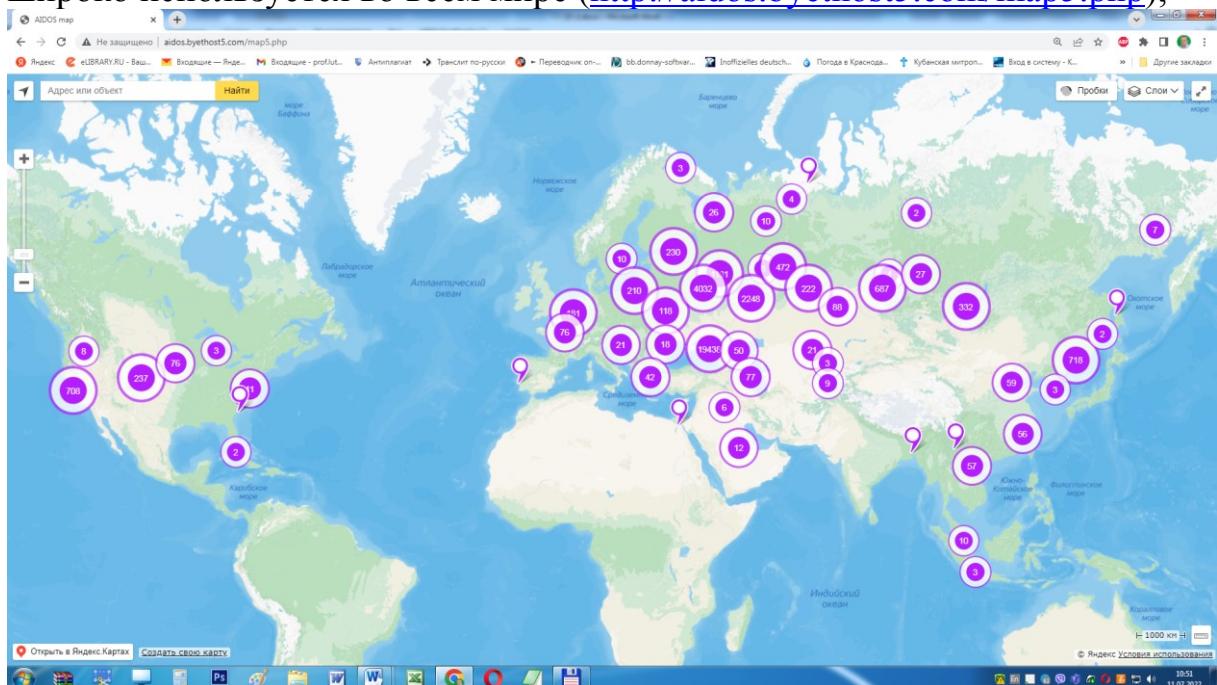
полностью с использованием только лицензионного инструментального программного обеспечения и на нее имеется 31 свидетельство РосПатента РФ);

- является одной из первых отечественных систем искусственного интеллекта персонального уровня, т.е. не требует от пользователя специальной подготовки в области технологий искусственного интеллекта: «имеет нулевой порог входа» (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 локальных (поставляемых с инсталляцией) и облачных учебных и научных Эйдос-приложений (в настоящее время их 31 и более 336, соответственно: http://aidos.byethost5.com/Source_data_applications/WebAppls.htm) (http://lc.kubagro.ru/aidos/Presentation_Aidos-online.pdf);

- поддерживает on-line среду накопления знаний и обмена ими, широко используется во всем мире (<http://aidos.byethost5.com/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 решения 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);

– обеспечивает преобразование исходных эмпирических данных в информацию, а ее в знания и решение с использованием этих знаний задач классификации, 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: дает результаты анализа, понятные экспертам на основе их опыта, интуиции и профессиональной компетенции;

– 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 эти данные и тем самым преобразовать их в информацию, а затем преобразовать эту информацию в знания путем ее применения для достижения целей (т.е. для 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 законах⁷.

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 по ссылке: <http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>).

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 видеограмма 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 решения 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»: с 2022 года по настоящее время. С 2022 года автор и разработчик системы «Эйдос» проф.Е.В.Луценко вплотную занялся разработкой профессиональной версии системы «Эйдос» на языке Аляска+Экспресс, обеспечивающей обработку больших данных, информации и знаний (Big

⁷ Ссылка на это краткое описание системы «Эйдос» на английском языке:
http://lc.kubagro.ru/aidos/The_Eidos_en.htm

Data, Big Information, Big Knowledge) с использованием ADS (Advantage Database Server), а также на языке C# (Visual Studio | C#).

На рисунке 1 приведена титульная видеограмма DOS-версии системы «Эйдос», а на рисунке 2 – текущей версии системы «Эйдос»:



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

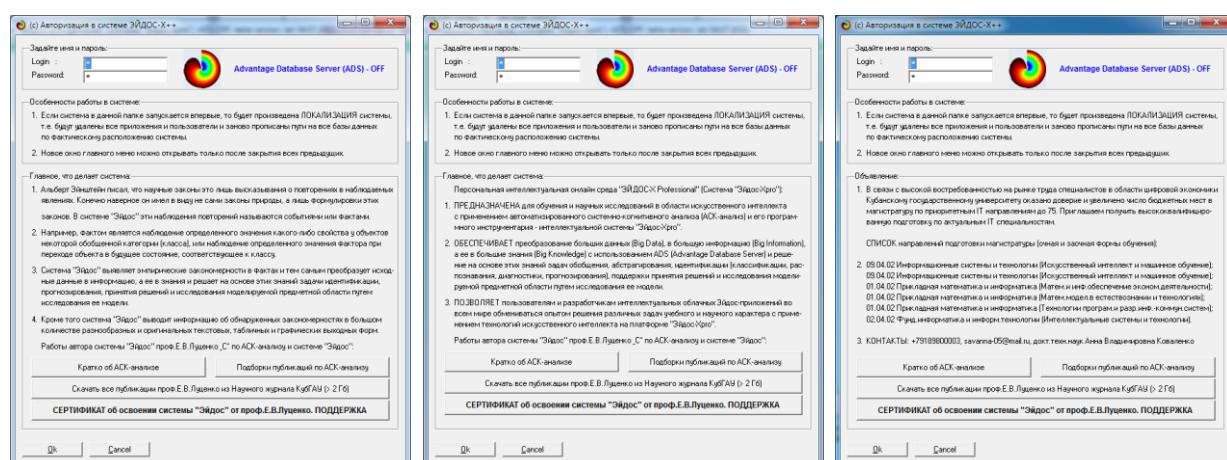


Рисунок 2. Титульные видеограммы текущей версии системы «Эйдос»

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

2.5. Цель и задачи работы

Целью работы является решение поставленной проблемы.

Как уже показано выше, для работы с лингвистическими переменными целесообразно применить лингвистический АСК-анализ [4].

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

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:

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

2) кластерно-конструктивный анализ классов;

3) кластерно-конструктивный анализ значений описательных шкал;

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

5) нелокальная нейронная сеть;

6) 3d-интегральные когнитивные карты;

7) 2d-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения);

8) 2d-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения);

9) когнитивные функции;

10) значимость описательных шкал и их градаций;

11) степень детерминированности классов и классификационных шкал).

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**

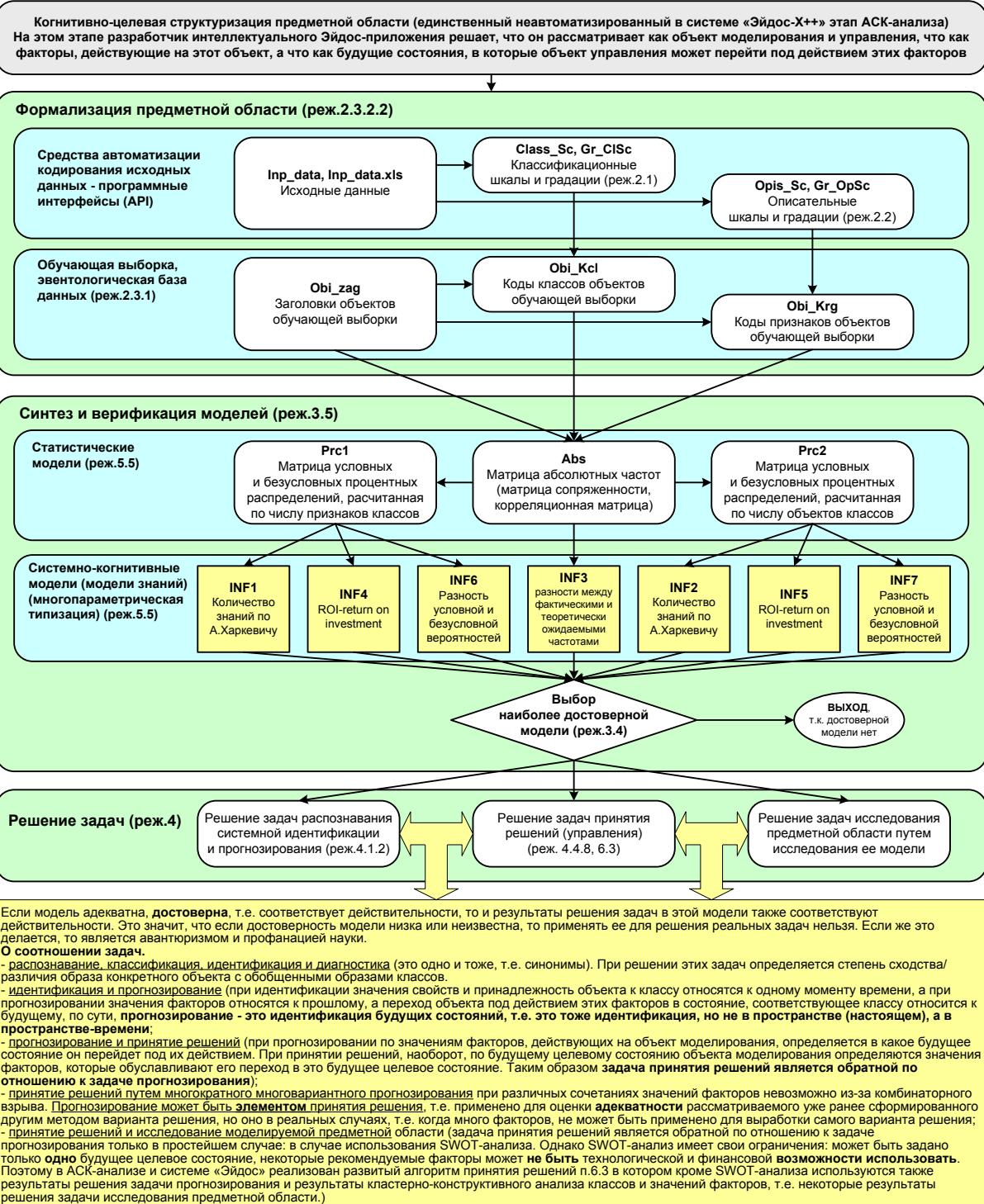


Рисунок 3. Последовательность преобразования исходных данных в информацию, а ее в знания и применения этих знаний для решения различных задач в системе «Эйдос»

3. RESULTS (РЕЗУЛЬТАТЫ)

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

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 неавтоматизированном 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.

В данной работе в качестве *объекта моделирования* выступают названия песен, в качестве *факторов* разные обстоятельства, относительно альбома, влияющие на его название (таблица 1), а в качестве *результатов* окончательное название этого альбома (таблица 2):

Таблица 1 – Описательные шкалы (факторы)

KOD_OPSC,N,15,0	NAME_OPSC,C,250
1	ENG_ALBUM_TITLE

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Opis_Sc.dbf

Таблица 2 – Классификационные шкалы (результаты действия факторов)

KOD_CLSC,N,15,0	NAME_CLSC,C,250	INSSIR/AEPSEN_GRC	CLSC,N,15,0
1	ALBUM_TITLE		23

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Class_Sc.dbf

3.2. Задача-2. Формализация 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 (ЭЭГ) или кардиограммы (ЭКГ).

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

В качестве *источника исходных данных* в данной работе используем таблицы 1 и 2 из работы.



Рисунок 4. Обложка журнала “VOGUE” с участием группы BTS.

Источник:

<https://i.pinimg.com/originals/21/4e/8d/214e8dd31c57fd1369674aec7a9791b7.png>



Рисунок 5. Фрагмент выступления в Пусане в октябре 2022

Источник:

https://lh3.googleusercontent.com/s1agVZN9Khg66kvRDnvGztLREqW8NDqYjPreOSMiThN8-Jky0kvjAFZEzXk0gA_vjwfejQIPIvcBXfCjyyGpdw3oM9MrH9BUJU84Gkkm42jDols=s0-e365-rj-l80-e365

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 6).

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.Чередниченко
2.3.2.12. Прогнозирование землетрясений методом Н.А.Чередниченко
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 6. Автоматизированные программные интерфейсы (API) системы «Эйдос»

Требования API-2.3.2.2 к исходным данным подробно описаны в хелпах этого режима (рисунки 7):

Экранные формы управления API-2.3.2.2 системы «Эйдос» с **реальными параметрами**, использованными в данной работе, приведены на рисунках 8.

В таблицах 3, 4, 5 приведены классификационные и описательные шкалы и градации, а также обучающая выборка, сформированные API-2.3.2.2 при параметрах, показанных на рисунках 8.

На 2-м рисунке 8 указано, что в описательных шкалах суммарное количество градаций 23, а в таблице 4 их приведено лишь 20. Это потому, что в шкал есть градация, которая в соответствии с параметрами на 1-м рисунке 8 рассматривается не как значащая, а как **отсутствие данных**. Для классификационных шкал на 3-м рисунке 8 приведено также количество наблюдений для каждого интервального значения (градации) и его размер. За счет того, что интервальные значения имеют разные размеры удается преодолеть **несбалансированность данных**, т.к. число наблюдений в каждом интервальном значении некоторой шкалы получается равным с точностью 1 (т.к. число наблюдений – всегда целое число).

Под несбалансированностью данных понимается неравномерность распределения значений свойств объекта моделирования или действующих на него факторов по диапазону изменения значений числовых шкал.

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-Х++" и формализации предметной области.

- Данный программный интерфейс обеспечивает формализацию предметной области, т.е. анализ файла исходных данных Inp_data.xls[x], формирование классификационных и описательных шкал и градаций, а затем кодирование файла исходных с их использованием.
- Файл исходных данных должен иметь имя: Inp_data.xls[x], а файл распознаваемой выборки имя: Inp_rasp.xls[x]. Файлы Inp_data.xls[x] и Inp_rasp.xls[x] должны находиться в папке ../AIDOS_X/AID_DATA/Inp_data/. Эти файлы имеют совершенно одинаковую структуру.
- 1-я строка этого файла должна содержать наименования колонок на любом языке, в т.ч. и русском. Эти наименования должны быть во всех колонках, при этом переносы по словам разрешены, а объединение ячеек, разрыв строки знак абзаца не допускаются. Эти наименования должны быть короткими, но понятными, т.к. они будут в выходных формах, 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 ОДИНАКОВОЕ.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длинным: 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 (номинального / порядкового) или числового типа [с десятичными знаками после запятой].
- Столбцу присваивается 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 [выходными параметрами] и содержат данные о классах (будущих состояниях объекта управления), к которым принадлежат объекты обучающей выборки.
- Столбцы с N+1 по последний являются описательными шкалами [свойствами или факторами] и содержат данные о признаках [т.е. значениях свойств или значениях факторов], характеризующих объекты обучающей выборки.
- В результате работы режима формируется файл INP_NAME.TXT стандарта MS DOS (кириллица), в котором наименования классификационных и описательных шкал являются СТРОКАМИ. Система формирует классификационные и описательные шкалы и градации. Для этого в каждом числовом столбце система находит минимальное и максимальное числовые значения и формирует заданное количество числовых интервалов, после чего числовые значения заменяются их интервальными значениями. В текстовых столбцах система находит уникальные текстовые значения. Каждое 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 NP_DATA и содержит коды классов, 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 такая 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

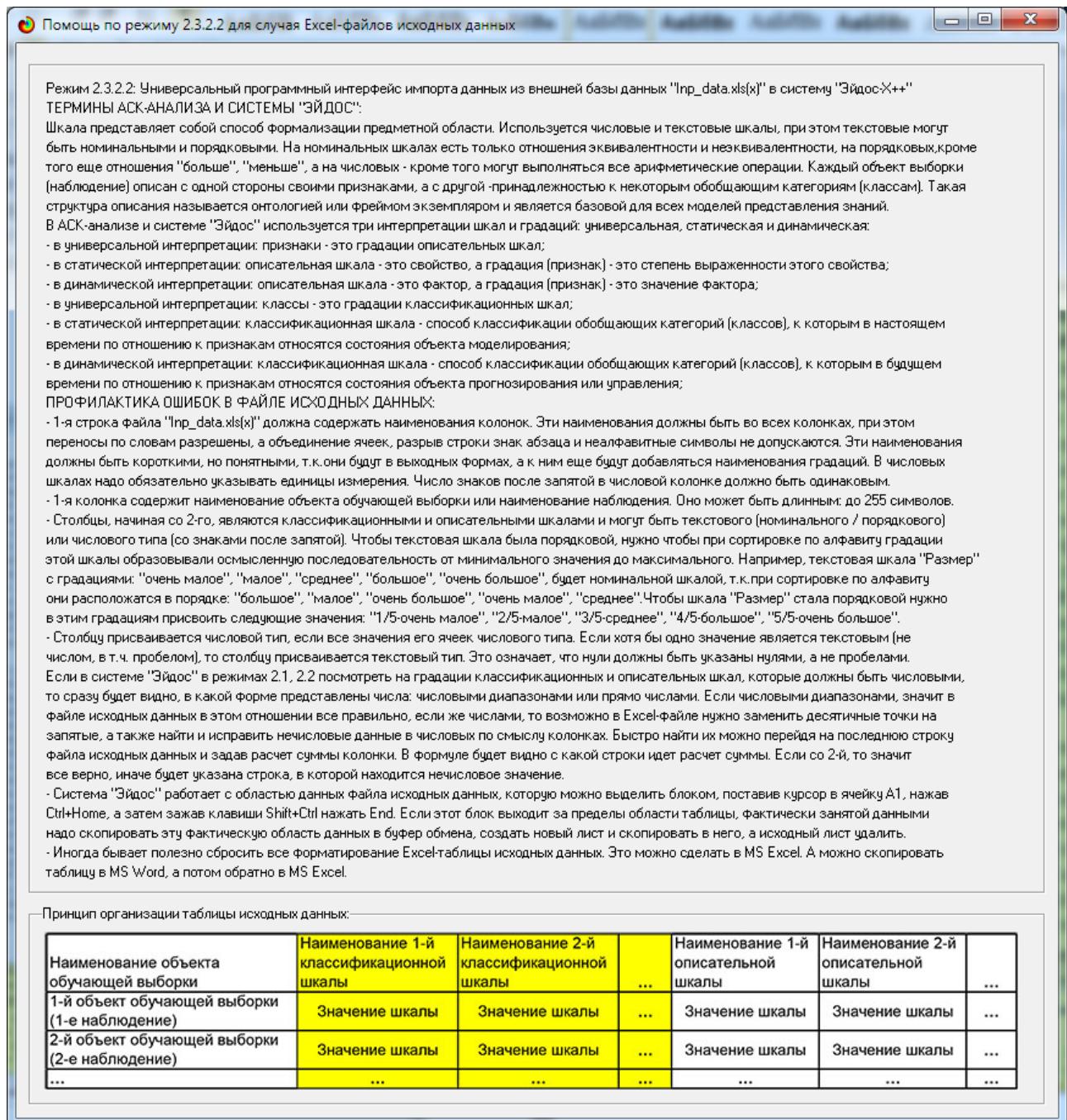


Рисунок 7. Хелпы API-2.3.2.2 системы «Эйдос»

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)
 DBF - DBASE IV (DBF/NTX)
 CSV - CSV => DBF конвертер 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 средник по классам "Inp_davt.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 спец.интерпретацию текстовых полей классов

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

Параметры интерпретации значений текстовых полей "Inp_data":

В качестве классов рассматриваются:

Значения полей целиком
 Элементы значений полей - слова > символов:
 Элементы значений полей - символы

Выделять уникальные значения и сортировать
 Не выделять уникальных значений и не сортировать

В качестве признаков рассматриваются:

Значения полей целиком
 Элементы значений полей - слова > символов:
 Элементы значений полей - символы

Проводить лемматизацию
 Не проводить лемматизацию

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)")

Ok Cancel

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

— □ ×

ИНФОРМАЦИЯ О РАЗМЕРНОСТИ МОДЕЛИ

Количество градаций классификационных и описательных шкал в модели, т.е.: [23 классов x 23 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	0	0	0,00	0	0	0,00
Текстовые	1	23	23,00	1	23	23,00
ВСЕГО:	1	23	23,00	1	23	23,00

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

Пересчитать шкалы и градации

Выйти на создание модели

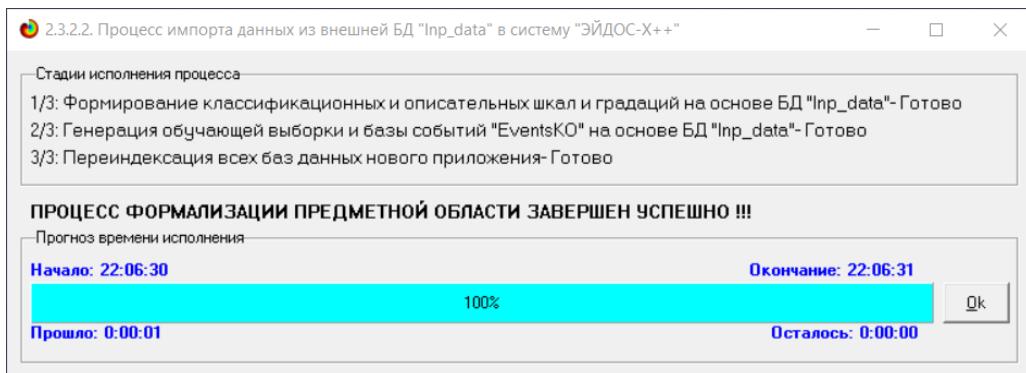


Рисунок 8. Экранные формы управления API-2.3.2.2 системы «Эйдос»

Таблица 3 – Классификационные шкалы и градации (числовые шкалы)

KOD_CLS,N,15,0	NAME_CLS,C,250	KOD_CLSC,N,15,0	N_CHRCLSC,N,3,0
1 ALBUM_TITLE-	The Most Beautiful Moment In Life Pt.1	1	11
2 ALBUM_TITLE-	The Most Beautiful Moment In Life Pt.2	1	11
3 ALBUM_TITLE-	The Most Beautiful Moment In Life: Young Forever	1	11
4 ALBUM_TITLE-2 Cool 4 Skool		1	11
5 ALBUM_TITLE-BE		1	11
6 ALBUM_TITLE-BTS WORLD: Original Soundtrack		1	11
7 ALBUM_TITLE-Butter (Hotter, Sweeter, Cooler)		1	11
8 ALBUM_TITLE-Butter / Permission to Dance		1	11
9 ALBUM_TITLE-Dark and Wild		1	11
10 ALBUM_TITLE-Dynamite (DayTime Version)		1	11
11 ALBUM_TITLE-Dynamite (NightTime Version)		1	11
12 ALBUM_TITLE-LOVE YOURSELF 'Answer'		1	11
13 ALBUM_TITLE-Love Yourself 'Her'		1	11
14 ALBUM_TITLE-LOVE YOURSELF 'Tear'		1	11
15 ALBUM_TITLE-MAP OF THE SOUL: 7		1	11
16 ALBUM_TITLE-MAP OF THE SOUL: PERSONA		1	11
17 ALBUM_TITLE-O!RUL8,2?		1	11
18 ALBUM_TITLE-Permission to Dance		1	11
19 ALBUM_TITLE-Proof		1	11
20 ALBUM_TITLE-Skool Luv Affair		1	11
21 ALBUM_TITLE-Skool Luv Affair (Special Addition)		1	11
22 ALBUM_TITLE-WINGS		1	11
23 ALBUM_TITLE-You Never Walk Alone		1	11

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Classes.dbf

Таблица 4 – Описательные шкалы и градации (лингвистические переменные)

KOD_ATR,N,15,0	NAME_ATR,C,250	KOD_OPSC,N,15,0	N_CHROPSC,N,3,0
1	ENG_ALBUM_TITLE-1	1	15
2	ENG_ALBUM_TITLE-2	1	15
3	ENG_ALBUM_TITLE-4	1	15
4	ENG_ALBUM_TITLE-7	1	15
5	ENG_ALBUM_TITLE-Addition	1	15
6	ENG_ALBUM_TITLE-Affair	1	15
7	ENG_ALBUM_TITLE-Alone	1	15
8	ENG_ALBUM_TITLE-and	1	15
9	ENG_ALBUM_TITLE-Answer	1	15
10	ENG_ALBUM_TITLE-BE	1	15
11	ENG_ALBUM_TITLE-Beautiful	1	15
12	ENG_ALBUM_TITLE-BTS	1	15
13	ENG_ALBUM_TITLE-Butter	1	15
14	ENG_ALBUM_TITLE-Cool	1	15
15	ENG_ALBUM_TITLE-Cooler	1	15
16	ENG_ALBUM_TITLE-Dance	1	15
17	ENG_ALBUM_TITLE-Dark	1	15
18	ENG_ALBUM_TITLE-DayTime	1	15
19	ENG_ALBUM_TITLE-Dynamite	1	15
20	ENG_ALBUM_TITLE-Forever	1	15
21	ENG_ALBUM_TITLE-Her	1	15
22	ENG_ALBUM_TITLE-Hotter	1	15
23	ENG_ALBUM_TITLE-In	1	15
24	ENG_ALBUM_TITLE-Life	1	15
25	ENG_ALBUM_TITLE-Love	1	15
26	ENG_ALBUM_TITLE-Luv	1	15
27	ENG_ALBUM_TITLE-MAP	1	15
28	ENG_ALBUM_TITLE-Moment	1	15
29	ENG_ALBUM_TITLE-Most	1	15
30	ENG_ALBUM_TITLE-Never	1	15
31	ENG_ALBUM_TITLE-NightTime	1	15
32	ENG_ALBUM_TITLE-O	1	15
33	ENG_ALBUM_TITLE-OF	1	15
34	ENG_ALBUM_TITLE-Original	1	15
35	ENG_ALBUM_TITLE-Permission	1	15
36	ENG_ALBUM_TITLE-PERSONA	1	15
37	ENG_ALBUM_TITLE-Proof	1	15
38	ENG_ALBUM_TITLE-Pt	1	15
39	ENG_ALBUM_TITLE-RUL8	1	15
40	ENG_ALBUM_TITLE-Skool	1	15
41	ENG_ALBUM_TITLE-SOUL	1	15
42	ENG_ALBUM_TITLE-Soundtrack	1	15
43	ENG_ALBUM_TITLE-Special	1	15
44	ENG_ALBUM_TITLE-Sweeter	1	15
45	ENG_ALBUM_TITLE-Tear	1	15
46	ENG_ALBUM_TITLE-The	1	15
47	ENG_ALBUM_TITLE-to	1	15
48	ENG_ALBUM_TITLE-Version	1	15
49	ENG_ALBUM_TITLE-Walk	1	15
50	ENG_ALBUM_TITLE-Wild	1	15
51	ENG_ALBUM_TITLE-WINGS	1	15
52	ENG_ALBUM_TITLE-WORLD	1	15
53	ENG_ALBUM_TITLE-You	1	15
54	ENG_ALBUM_TITLE-Young	1	15
55	ENG_ALBUM_TITLE-Yourself	1	15

Источник: c:\Aidos-X\AID_DATA\A0000002\System\Attributes.dbf

Таблица 5– Обучающая выборка (полностью)

NAME_OBJ	N2,N,8,0	36	21	71	2	106	22	141	14
0	4	37	21	72	2	107	22	142	14
1	4	38	21	73	3	108	22	143	14
2	4	39	21	74	3	109	22	144	14
3	4	40	21	75	3	110	22	145	14
4	4	41	9	76	3	111	23	146	14
5	4	42	9	77	3	112	23	147	14
6	4	43	9	78	3	113	23	148	14
7	4	44	9	79	3	114	23	149	14
8	4	45	9	80	3	115	23	150	14
9	17	46	9	81	3	116	23	151	12
10	17	47	9	82	3	117	23	152	12
11	17	48	9	83	3	118	23	153	12
12	17	49	9	84	3	119	23	154	12
13	17	50	9	85	3	120	23	155	12
14	17	51	9	86	3	121	23	156	12
15	17	52	9	87	3	122	23	157	12
16	17	53	9	88	3	123	23	158	12
17	17	54	9	89	3	124	23	159	12
18	20	55	1	90	3	125	23	160	12
19	20	56	1	91	3	126	23	161	12
20	20	57	1	92	3	127	23	162	12
21	20	58	1	93	3	128	23	163	12
22	20	59	1	94	3	129	13	164	12
23	20	60	1	95	3	130	13	165	12
24	20	61	1	96	22	131	13	166	12
25	20	62	1	97	22	132	13	167	12
26	20	63	1	98	22	133	13	168	12
27	20	64	2	99	22	134	13	169	12
28	21	65	2	100	22	135	13	170	12
29	21	66	2	101	22	136	13	171	12
30	21	67	2	102	22	137	13	172	12
31	21	68	2	103	22	138	13	173	12
32	21	69	2	104	22	139	13	174	12
33	21	70	2	105	22	140	14	175	12

174	12	9	205	15	4	235	5	10	265	19	37
175	12	9	206	15	4	236	5	10	266	19	37
176	12	9	207	15	4	237	5	10	267	19	37
177	12	9	208	15	4	238	5	10	268	19	37
178	16	36	209	15	4	239	5	10	269	19	37
179	16	36	210	15	4	240	7	15	270	19	37
180	16	36	211	15	4	241	7	15	271	19	37
181	16	36	212	15	4	242	7	15	272	19	37
182	16	36	213	15	4	243	7	15	273	19	37
183	16	36	214	15	4	244	7	15	274	19	37
184	16	36	215	15	4	245	8	16	275	19	37
185	6	42	216	15	4	246	8	16	276	19	37
186	6	42	217	15	4	247	8	16	277	19	37
187	6	42	218	15	4	248	8	16	278	19	37
188	6	42	219	15	4	249	18	16	279	19	37
189	6	42	220	10	48	250	18	16	280	19	37
190	6	42	221	10	48	251	18	16	281	19	37
191	6	42	222	10	48	252	19	37	282	19	37
192	6	42	223	10	48	253	19	37	283	19	37
193	6	42	224	10	48	254	19	37	284	19	37
194	6	42	225	10	48	255	19	37	285	19	37
195	6	42	226	11	48	256	19	37	286	19	37
196	6	42	227	11	48	257	19	37	287	19	37
197	6	42	228	11	48	258	19	37	288	19	37
198	15	4	229	11	48	259	19	37	289	19	37
199	15	4	230	11	48	260	19	37	290	19	37
200	15	4	231	11	48	261	19	37	291	19	37
201	15	4	232	5	10	262	19	37	292	19	37
202	15	4	233	5	10	263	19	37	293	19	37
203	15	4	234	5	10	264	19	37	294	19	37
204	15	4	235	5	10	265	19	37	295	19	37

Источник: c:\Aidos-X\AID_DATA\A0000002\System\EventsKO.dbf

Отметим, что в системе «Эйдос» обычно используются базы данных с расширением «dbf». Они открываются в MS Excel или могут быть конвертированы в файлы xls,xlsx с помощью онлайн-сервисов.

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

Синтез и верификация статистических и системно-когнитивных моделей (СК-моделей) моделей осуществляется в режиме 3.5 системы «Эйдос». Математические модели, на основе которых рассчитываются статистические и СК-модели, подробно описаны в ряде монографий и статей автора. Поэтому в данной работе мы рассмотрим эти вопросы очень кратко. Отметим лишь, 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) (рисунок 3).

Математическая модель АСК-анализа и системы «Эйдос» основана на системной нечеткой интервальной математике и обеспечивает [2, 3] сопоставимую обработку больших объемов фрагментированных и зашумленных взаимозависимых данных, представленных в различных типах шкал (номинальных, порядковых и числовых) и различных единицах измерения.

Суть математической модели АСК-анализа состоит в следующем.

Непосредственно на основе эмпирических данных (см. Help режима 2.3.2.2) рассчитывается матрица абсолютных частот (таблица 6).

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

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.

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 6) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) (таблица 7) является весьма обоснованным и логичным.

Таблица 6 – 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			$N_{\Sigma j}$				$N_{\Sigma\Sigma} = \sum_{j=1}^W N_{\Sigma j}$

Таблица 7 – 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			$P_{\Sigma j}$				

Этот переход полностью снимает проблему несбалансированности данных, т.к. в последующем анализе используется не матрица абсолютных частот (таблица 6), а матрицы условных и безусловных процентных распределений (таблица 7) и матрицы системно-когнитивных моделей (СК-модели, таблица 9), в частности матрица информативностей.

Этот подход снимает также проблему обеспечения сопоставимости обработки в одной модели исходных данных, представленных в различных видах шкал (номинальных, порядковых и числовых) и в разных единицах измерения [6].

В системе «Эйдос» этот подход применяется *всегда* при решении любых задач.

Затем на основе таблиц 9 и 10 с использованием **частных критерииев, знаний** приведенных таблице 11, рассчитываются матрицы семи системно-когнитивных моделей (таблица 9).

В таблице 11 приведены формулы:

- для сравнения **фактических и теоретических абсолютных частот**;
- для сравнения **условных и безусловных относительных частот** («вероятностей»).

И это сравнение в таблицах 9 и 10 осуществляется двумя возможными способами: путем **вычитания** и путем **деления**.

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 8), применяемых в настоящее время в системе «Эйдос» равное 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**.

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).

Таблица 8– 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 4)		$N_i = \sum_{j=1}^W N_{ij}; N_j = \sum_{i=1}^M N_{ij}; N = \sum_{i=1}^W \sum_{j=1}^M N_{ij};$ $N_{ij} - \text{фактическая частота};$ $\bar{N}_{ij} = \frac{N_i N_j}{N} - \text{теоретическая частота}.$
PRC1 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество признаков по классу	--	
PRC2 , матрица условных P_{ij} и безусловных P_i процентных распределений, в качестве N_j используется суммарное количество объектов обучающей выборки по классу	--	$P_{ij} = \frac{N_{ij}}{N_j}; P_i = \frac{N_i}{N}$
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		
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.	$I_{ij} = \Psi \times \log_2 \frac{P_{ij}}{P_i}$	$I_{ij} = \Psi \times \log_2 \frac{N_{ij}}{\bar{N}_{ij}} = \Psi \times \log_2 \frac{N_{ij} N}{N_i N_j}$
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	$I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$	
INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		$I_{ij} = \frac{N_{ij}}{\bar{N}_{ij}} - 1 = \frac{N_{ij} N}{N_i N_j} - 1$
INF6 , частный критерий: разность условной и безусловной вероятностей, 1-й вариант расчета вероятностей: N_j – суммарное количество признаков по j -му классу	$I_{ij} = P_{ij} - P_i$	
INF7 , частный критерий: разность условной и безусловной вероятностей, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по j -му классу		$I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$

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.

Таким образом, мы видим, что **все частные критерии знаний тесно взаимосвязаны друг с другом**. Особенno интересна связь знаменитого критерия хи-квадрат 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%.

Таблица 9 – Матрица системно-когнитивной модели

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

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 [6].

На основе системно-когнитивных моделей, представленных в таблице 12 (отличаются частыми критериями, приведенными в таблице 11), 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 9).

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 10):

Таблица 3 – 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	$\sigma_{i\Sigma} = \sqrt{\frac{1}{W-1} \sum_{j=1}^W (I_{ij} - \bar{I}_i)^2}$
2	1. Степень детерминированности класса. 2. Степень обусловленности класса.	Корень из информационной мощности класса	$\sigma_{\Sigma j} = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (I_{ij} - \bar{I}_j)^2}$
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	$H = \sqrt{\frac{1}{(W \cdot M - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2}$

В системе «Эйдос» синтез моделей производится в режиме 3.5 (рисунок 9):

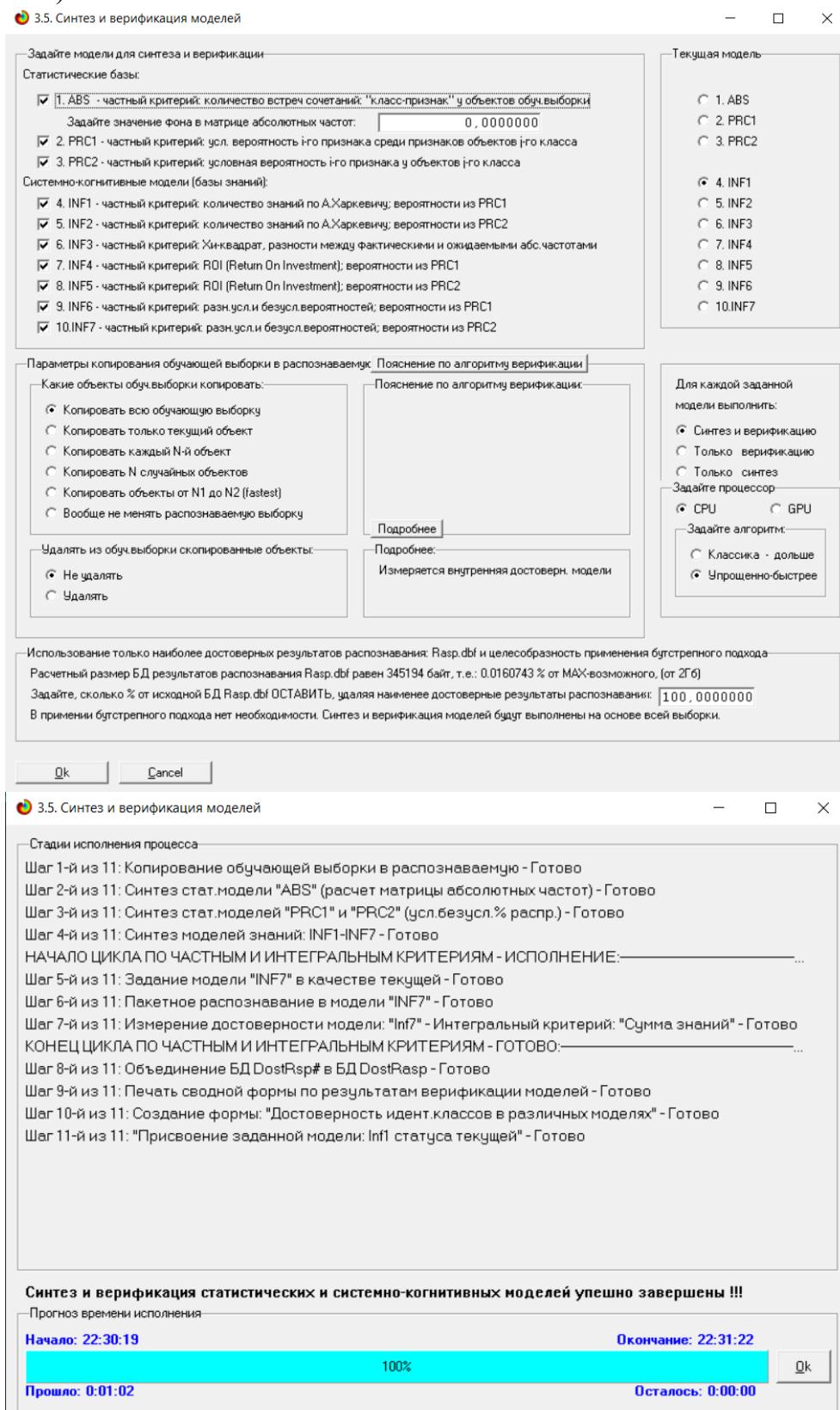


Рисунок 9. Экранные формы режима синтеза и верификации моделей

В результате работы режима 3.5 создано 3 статистических и 7 системно-когнитивных моделей, некоторые из которых приведены на рисунках 10-13:

5.5. Модель: "1. ABS - частный критерий: количество встреч сочетаний: "Класс-признак" у объектов обучвыборки"													
Код признака	Наименование описательной шкалы и градации	1. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.1	2. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.2	3. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE: YOUNG FOREVER	4. ALBUM TITLE 2 COOL 4 SKOOL	5. ALBUM TITLE BE	6. ALBUM TITLE BTS WORLD: ORIGINAL SOUNDTRACK	7. ALBUM TITLE BUTTER (HOTTER, SWEETER, COOLER)	8. ALBUM TITLE BUTTER / PERMISSION TO DANCE	9. ALBUM TITLE DARK AND WILD	10. ALBUM TITLE DYNAMITE (DAYTIME VERSION)	11. ALBUM TITLE DYNAMITE (NIGHTTIME VERSION)	12. ALBUM TITLE LOVE YOURSELF 'ANSWER'
39.0	ENG ALBUM TITLE-RUL8												
40.0	ENG ALBUM TITLE-Skool												
41.0	ENG ALBUM TITLE-SOUL												
42.0	ENG ALBUM TITLE-Soundtrack												
43.0	ENG ALBUM TITLE-Special												
44.0	ENG ALBUM TITLE-Sweeter												
45.0	ENG ALBUM TITLE-Tear												
46.0	ENG ALBUM TITLE-The	9.0	9.0	23.0									
47.0	ENG ALBUM TITLE-to												
48.0	ENG ALBUM TITLE-Version												
49.0	ENG ALBUM TITLE-Walk												
50.0	ENG ALBUM TITLE-Wild												
51.0	ENG ALBUM TITLE-WINGS												
52.0	ENG ALBUM TITLE-wORLD												
53.0	ENG ALBUM TITLE-You												
54.0	ENG ALBUM TITLE-Young			23.0									
55.0	ENG ALBUM TITLE-Yourself												
Сумма числа признаков		72.0	72.0	184.0	36.0	8.0	56.0	20.0	16.0	42.0	18.0	18.0	2
Среднее		1.3	1.3	3.3	0.7	0.1	1.0	0.4	0.3	0.8	0.3	0.3	
Среднеквадратичное отклонение		3.2	3.2	8.2	2.4	1.1	3.7	1.3	1.0	3.4	1.4	1.4	
Сумма числа объектов обучвыборки		9.0	9.0	23.0	9.0	8.0	14.0	5.0	4.0	14.0	6.0	6.0	

Рисунок 10. Статистическая модель «ABS», матрица абсолютных частот

5.5. Модель: "3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса"																
Код признака	Наименование описательной шкалы и градации	1. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.1	2. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.2	3. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE: YOUNG FOREVER	4. ALBUM TITLE 2 COOL 4 SKOOL	5. ALBUM TITLE BE	6. ALBUM TITLE BTS WORLD: ORIGINAL SOUNDTRA...	7. ALBUM TITLE BUTTER (HOTTER, SWEETER, COOLER)	8. ALBUM TITLE BUTTER / PERMISSION TO DANCE	9. ALBUM TITLE DARK AND WILD	10. ALBUM TITLE DYNAMITE (DAYTIME VERSION)	11. ALBUM TITLE DYNAMITE (NIGHTTIME VERSION)	12. ALBUM TITLE LOVE YOURSELF 'ANSWER'	13. ALBUM TITLE LOVE YOURSELF 'HER'	14. ALBUM TITLE LOVE YOURSELF 'TEAR'	15. ALBUM TITLE LOVE YOURSELF 'THE SOU
38.0	ENG ALBUM TITLE-Pt	100.000	100.000													
39.0	ENG ALBUM TITLE-RUL8															
40.0	ENG ALBUM TITLE-Skool						100.000									
41.0	ENG ALBUM TITLE-SOUL														1	
42.0	ENG ALBUM TITLE-Soundtrack							100.000								
43.0	ENG ALBUM TITLE-Special															
44.0	ENG ALBUM TITLE-Sweeter								100.000							
45.0	ENG ALBUM TITLE-Tear														100.000	
46.0	ENG ALBUM TITLE-The	100.000	100.000	100.000											1	
47.0	ENG ALBUM TITLE-to								100.000							
48.0	ENG ALBUM TITLE-Version										100.000	100.000				
49.0	ENG ALBUM TITLE-Walk															
50.0	ENG ALBUM TITLE-Wild										100.000					
51.0	ENG ALBUM TITLE-WINGS															
52.0	ENG ALBUM TITLE-wORLD							100.000								
53.0	ENG ALBUM TITLE-You															
54.0	ENG ALBUM TITLE-Young			100.000												
55.0	ENG ALBUM TITLE-Yourself															
Сумма		800.000	800.000	800.000	400.000	100.000	400.000	400.000	400.000	300.000	300.000	300.000	300.000	300.000	5	
Среднее		14.545	14.545	14.545	7.273	1.818	7.273	7.273	7.273	5.455	5.455	5.455	5.455	5.455		
Среднеквадратичное отклонение		35.582	35.582	35.582	26.210	13.493	26.210	26.210	26.210	22.921	22.921	22.921	22.921	22.921		

Рисунок 11. Статистическая модель «PRC2», матрица условных и безусловных процентных распределений

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"												
Код признака	Наименование описательной шкалы и градации	1. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.1	2. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.2	3. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE: YOUNG FOREVER	4. ALBUM TITLE 2 COOL 4 SKOOL	5. ALBUM TITLE BE	6. ALBUM TITLE BTS WORLD: ORIGINAL SOUNDTRACK	7. ALBUM TITLE BUTTER (HOTTER, SWEETER, COOLER)	8. ALBUM TITLE BUTTER / PERMISSION TO DANCE	9. ALBUM TITLE DARK AND WILD	10. ALBUM TITLE DYNAMITE (DAYTIME VERSION)	11. ALBUM TITLE DYNAMITE (NIGHTTIME VERSION)
38.0	ENG ALBUM TITLE-Pt	1. 310	1. 310									
39.0	ENG ALBUM TITLE-RUL8											
40.0	ENG ALBUM TITLE-Skool					1. 407						
41.0	ENG ALBUM TITLE-SOUL											
42.0	ENG ALBUM TITLE-Soundtrack							1. 921				
43.0	ENG ALBUM TITLE-Special								2. 587			
44.0	ENG ALBUM TITLE-Sweeter											
45.0	ENG ALBUM TITLE-Tear											
46.0	ENG ALBUM TITLE-The	0. 442	0. 442	0. 442								
47.0	ENG ALBUM TITLE-To								2. 369			
48.0	ENG ALBUM TITLE-Version										2. 207	2. 207
49.0	ENG ALBUM TITLE-Walk											
50.0	ENG ALBUM TITLE-Wild										2. 107	
51.0	ENG ALBUM TITLE-WINGS											
52.0	ENG ALBUM TITLE-WORLD							1. 921				
53.0	ENG ALBUM TITLE-You											
54.0	ENG ALBUM TITLE-Young			1. 152								
55.0	ENG ALBUM TITLE-Yourself											
Сумма		7. 402	6. 668	6. 637	7. 293	3. 179	7. 684	9. 967	9. 314	6. 321	7. 068	7. 068
Среднее		0. 135	0. 121	0. 121	0. 133	0. 058	0. 140	0. 181	0. 169	0. 145	0. 129	0. 129
Среднеквадратичное отклоне...		0. 361	0. 310	0. 307	0. 489	0. 429	0. 503	0. 655	0. 611	0. 483	0. 542	0. 542

Рисунок 12. Системно-когнитивная модель «INF1», матрица информавностей (по А.Харкевичу)

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсолютами"												
Код признака	Наименование описательной шкалы и градации	1. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.1	2. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE PT.2	3. ALBUM TITLE THE MOST BEAUTIFUL MOMENT IN LIFE: YOUNG FOREVER	4. ALBUM TITLE 2 COOL 4 SKOOL	5. ALBUM TITLE BE	6. ALBUM TITLE BTS WORLD: ORIGINAL SOUNDTRACK	7. ALBUM TITLE BUTTER (HOTTER, SWEETER, COOLER)	8. ALBUM TITLE BUTTER / PERMISSION TO DANCE	9. ALBUM TITLE DARK AND WILD	10. ALBUM TITLE DYNAMITE (DAYTIME VERSION)	11. ALBUM TITLE DYNAMITE (NIGHTTIME VERSION)
38.0	ENG ALBUM TITLE-Pt	7. 814	7. 814	-3. 030	-0. 593	-0. 132	-0. 922	-0. 329	-0. 263	-0. 692	-0. 296	-0. 296
39.0	ENG ALBUM TITLE-RUL8	-0. 659	-0. 659	-1. 683	-0. 329	-0. 073	-0. 512	-0. 183	-0. 146	-0. 384	-0. 165	-0. 165
40.0	ENG ALBUM TITLE-Skool	-2. 042	-2. 042	-5. 219	7. 979	-0. 227	-1. 588	-0. 567	-0. 454	-1. 191	-0. 511	-0. 511
41.0	ENG ALBUM TITLE-SOUL	-1. 844	-1. 844	-4. 714	-0. 922	-0. 205	-1. 435	-0. 512	-0. 410	-1. 076	-0. 461	-0. 461
42.0	ENG ALBUM TITLE-Soundtrack	-0. 922	-0. 922	-2. 357	-0. 461	-0. 102	13. 283	-0. 256	-0. 205	-0. 538	-0. 231	-0. 231
43.0	ENG ALBUM TITLE-Special	-0. 790	-0. 790	-2. 020	-0. 395	-0. 088	-0. 615	-0. 220	-0. 176	-0. 461	-0. 198	-0. 198
44.0	ENG ALBUM TITLE-Sweeter	-0. 329	-0. 329	-0. 842	-0. 165	-0. 037	-0. 256	4. 909	-0. 073	-0. 192	-0. 082	-0. 082
45.0	ENG ALBUM TITLE-Tear	-0. 725	-0. 725	-1. 852	-0. 362	-0. 081	-0. 564	-0. 201	-0. 161	-0. 423	-0. 181	-0. 181
46.0	ENG ALBUM TITLE-The	4. 455	4. 455	11. 384	-2. 273	-0. 505	-3. 535	-1. 263	-1. 010	-2. 651	-1. 136	-1. 136
47.0	ENG ALBUM TITLE-To	-0. 461	-0. 461	-1. 178	-0. 231	-0. 051	-0. 359	-0. 128	3. 898	-0. 269	-0. 115	-0. 115
48.0	ENG ALBUM TITLE-Version	-0. 790	-0. 790	-2. 020	-0. 395	-0. 088	-0. 615	-0. 220	-0. 176	-0. 461	5. 802	5. 802
49.0	ENG ALBUM TITLE-Walk	-1. 186	-1. 186	-3. 030	-0. 593	-0. 132	-0. 922	-0. 329	-0. 263	-0. 692	-0. 296	-0. 296
50.0	ENG ALBUM TITLE-Wild	-0. 922	-0. 922	-2. 357	-0. 461	-0. 102	-0. 717	-0. 256	-0. 205	13. 462	-0. 231	-0. 231
51.0	ENG ALBUM TITLE-WINGS	-0. 988	-0. 988	-2. 525	-0. 494	-0. 110	-0. 769	-0. 274	-0. 220	-0. 576	-0. 247	-0. 247
52.0	ENG ALBUM TITLE-WORLD	-0. 922	-0. 922	-2. 357	-0. 461	-0. 102	13. 283	-0. 256	-0. 205	-0. 538	-0. 231	-0. 231
53.0	ENG ALBUM TITLE-You	-1. 186	-1. 186	-3. 030	-0. 593	-0. 132	-0. 922	-0. 329	-0. 263	-0. 692	-0. 296	-0. 296
54.0	ENG ALBUM TITLE-Young	-1. 515	-1. 515	19. 128	-0. 758	-0. 168	-1. 178	-0. 421	-0. 337	-0. 884	-0. 379	-0. 379
55.0	ENG ALBUM TITLE-Yourself	-3. 228	-3. 228	-8. 249	-1. 614	-0. 359	-2. 511	-0. 897	-0. 717	-1. 883	-0. 807	-0. 807
Сумма												
Среднее		2. 829	2. 755	7. 039	2. 410	1. 095	3. 823	1. 404	1. 118	3. 308	1. 433	1. 433
Среднеквадратичное отклонение...												

Рисунок 13 . Системно-когнитивная модель «INF3», матрица Хи-квадрат (по К.Пирсону)

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

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 [12].

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.

В соответствии с критерием достоверности F-мерой Ван Ризбергена наиболее достоверной является СК-модель INF4 с интегральным критерием: «Семантический резонанс знаний»: $F=0.641$ (1-й рисунок 14).

В соответствии с критерием достоверности L1-мерой проф. Е.В.Луценко наиболее достоверной является СК-модель INF3 (хи-квадрат К.Пирсона) с интегральным критерием: «Сумма знаний»: $L1=0.855$ (2-й рисунок 14). **Эту модель и будем использовать для решения поставленных в работе задач.**

3.4. Обобщ форма по достоверности при разн. крит. Текущая модель: "INF4"												
Наименование модели и частного критерия	Интегральный критерий	отдали	F-мера Ван Ризбергена	Сумма модул... уровней сход... истинно-полож. решений (STP)	Сумма модул... уровней сход... истинно-отриц. решений (STN)	Сумма модул... уровней сход... ложного-полож. решений (SFP)	S-точность модели	S-полнота модели	L1-мера проф. Е.В.Луценко	Средний модуль уровней сход... истинно-полож. решений	Средний модуль уровней сход... истинно-отриц. решений	
1. ABS - частный критерий: количество встреч сочтений; "лас..."	Корреляция abs. частот с обр...	1. 000	0.536	285. 315	372. 172	187. 349	0. 604	1. 000	0. 753	0. 951	0. 061	
1. ABS - частный критерий: количество встреч сочтений; "лас..."	Сумма abs. частот по признак...	1. 000	0.536	98. 929	77. 870	0. 560	1. 000	0. 718	0. 330			
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл.отн частот с о...	1. 000	0.536	285. 315	372. 172	187. 349	0. 604	1. 000	0. 753	0. 951	0. 061	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл.отн частот по при...	1. 000	0.536	300. 000	236. 308	0. 559	1. 000	0. 717	1. 000			
3. PRC2 - частный критерий: условная вероятность i-го признака...	Корреляция усл.отн частот с о...	1. 000	0.536	285. 442	372. 376	187. 461	0. 604	1. 000	0. 753	0. 951	0. 061	
3. PRC2 - частный критерий: условная вероятность i-го признака...	Сумма усл.отн частот по при...	1. 000	0.536	136. 625	150. 000	0. 477	1. 000	0. 646	0. 455			
4. INF1 - частный критерий: количество знаний по АХаренбену; в...	Семантический резонанс зна...	1. 000	0.629	281. 132	370. 067	146. 504	0. 657	1. 000	0. 793	0. 937	0. 059	
4. INF1 - частный критерий: количество знаний по АХаренбену; в...	Сумма знаний	1. 000	0.536	164. 529	123. 167	0. 572	1. 000	0. 728	0. 548			
5. INF2 - частный критерий: количество знаний по АХаренбену; в...	Семантический резонанс зна...	1. 000	0.629	282. 200	369. 030	150. 470	0. 652	1. 000	0. 790	0. 941	0. 059	
5. INF2 - частный критерий: количество знаний по АХаренбену; в...	Сумма знаний	1. 000	0.536	153. 739	135. 442	0. 532	1. 000	0. 694	0. 512			
6. INF3 - частный критерий: Хи-квадрат, разности между фактами...	Семантический резонанс зна...	1. 000	0.635	281. 812	550. 232	169. 610	0. 624	1. 000	0. 769	0. 939	0. 088	
6. INF3 - частный критерий: Хи-квадрат, разности между фактами...	Сумма знаний	1. 000	0.635	119. 122	185. 180	66. 058	0. 643	1. 000	0. 783	0. 397	0. 030	
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	1. 000	0.641	246. 428	349. 824	90. 333	0. 747	1. 000	0. 855	0. 888	0. 056	
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	1. 000	0.536	84. 555	35. 490	0. 704	1. 000	0. 827	0. 282			
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	1. 000	0.641	247. 603	350. 786	93. 425	0. 741	1. 000	0. 851	0. 892	0. 056	
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	1. 000	0.536	73. 670	42. 994	0. 631	1. 000	0. 774	0. 245			
9. INF6 - частный критерий: разн. усл и без усл.вероятностей; вер...	Семантический резонанс зна...	1. 000	0.579	284. 675	371. 601	172. 921	0. 622	1. 000	0. 767	0. 949	0. 060	
9. INF6 - частный критерий: разн. усл и без усл.вероятностей; вер...	Сумма знаний	1. 000	0.536	272. 229	189. 103	0. 590	1. 000	0. 742	0. 907			
10. INF7 - частный критерий: разн. усл и без усл.вероятностей; ве...	Семантический резонанс зна...	1. 000	0.536	285. 097	371. 866	178. 229	0. 615	1. 000	0. 762	0. 950	0. 061	
10. INF7 - частный критерий: разн. усл и без усл.вероятностей; ве...	Сумма знаний	1. 000	0.536	140. 719	146. 209	0. 490	1. 000	0. 659	0. 469			

Рисунок 14. Экранные формы режима измерения достоверности моделей 3.4

На рисунках 15 приведены частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф. Е.В.Луценко СК-модели INF3.

В наиболее достоверной по критерию достоверности L1-мерой проф. Е.В.Луценко СК-модели INF3:

– при уровнях сходства меньше 40% преобладают ложные положительные решения, а при более высоких уровнях сходства – истинные положительные решения. При уровнях сходства выше 50% ложных положительных решений 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.

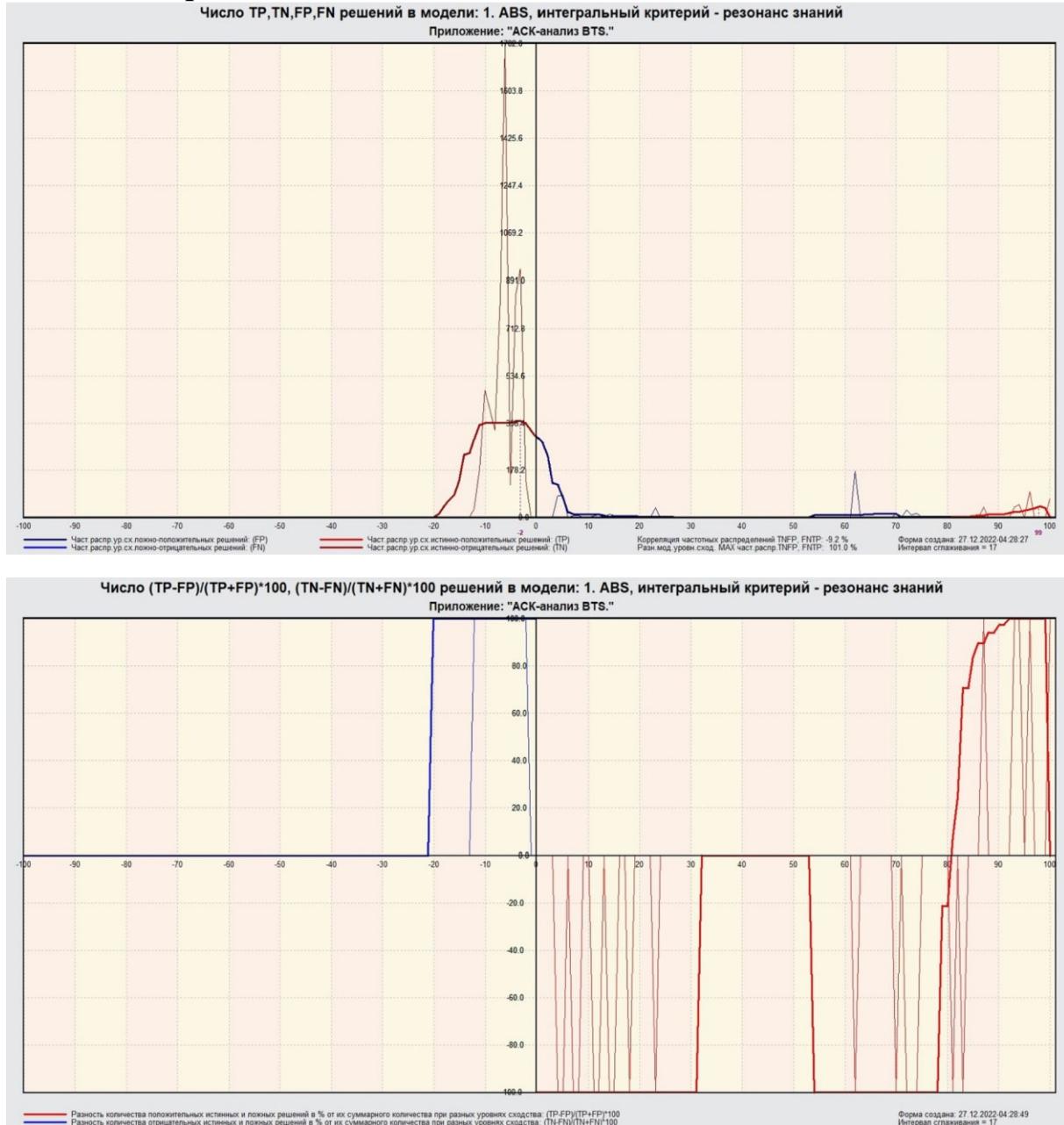


Рисунок 15. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф. Е.В.Луценко СК-модели INF3

На рисунках 16 приведены экранные формы хелпов режима 3.4, в которых подробно объясняется смысл этого режима. Эти формы приводятся в работе вместо более детального описания данного режима.

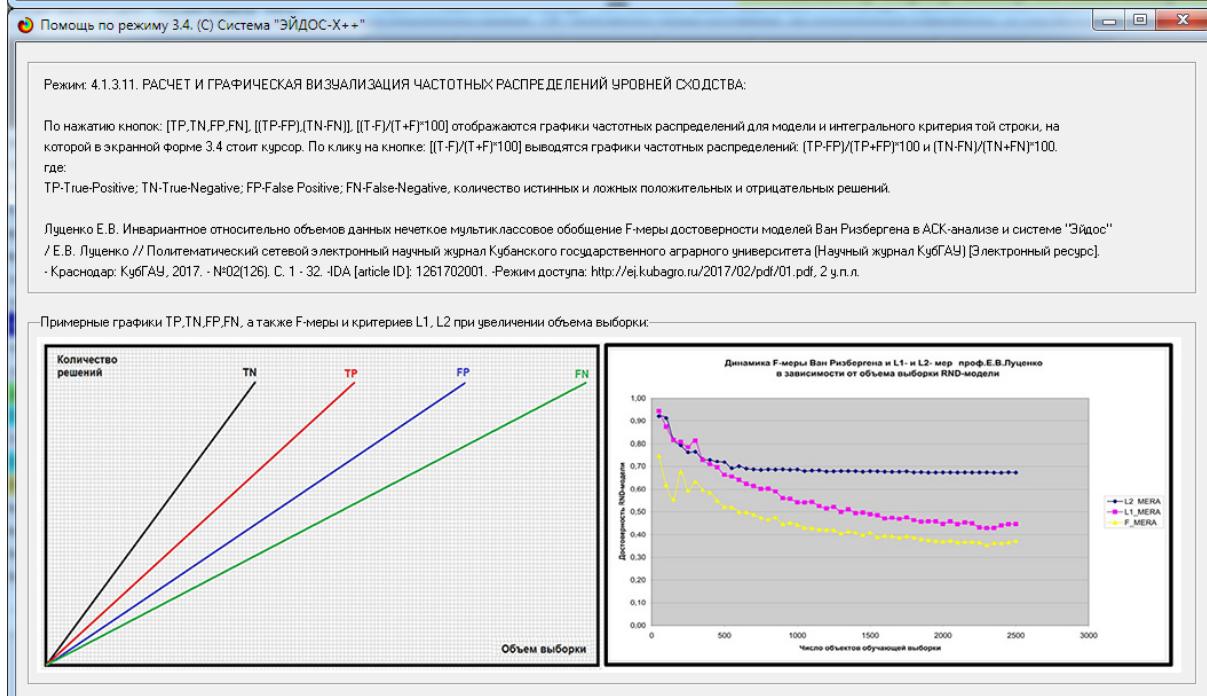
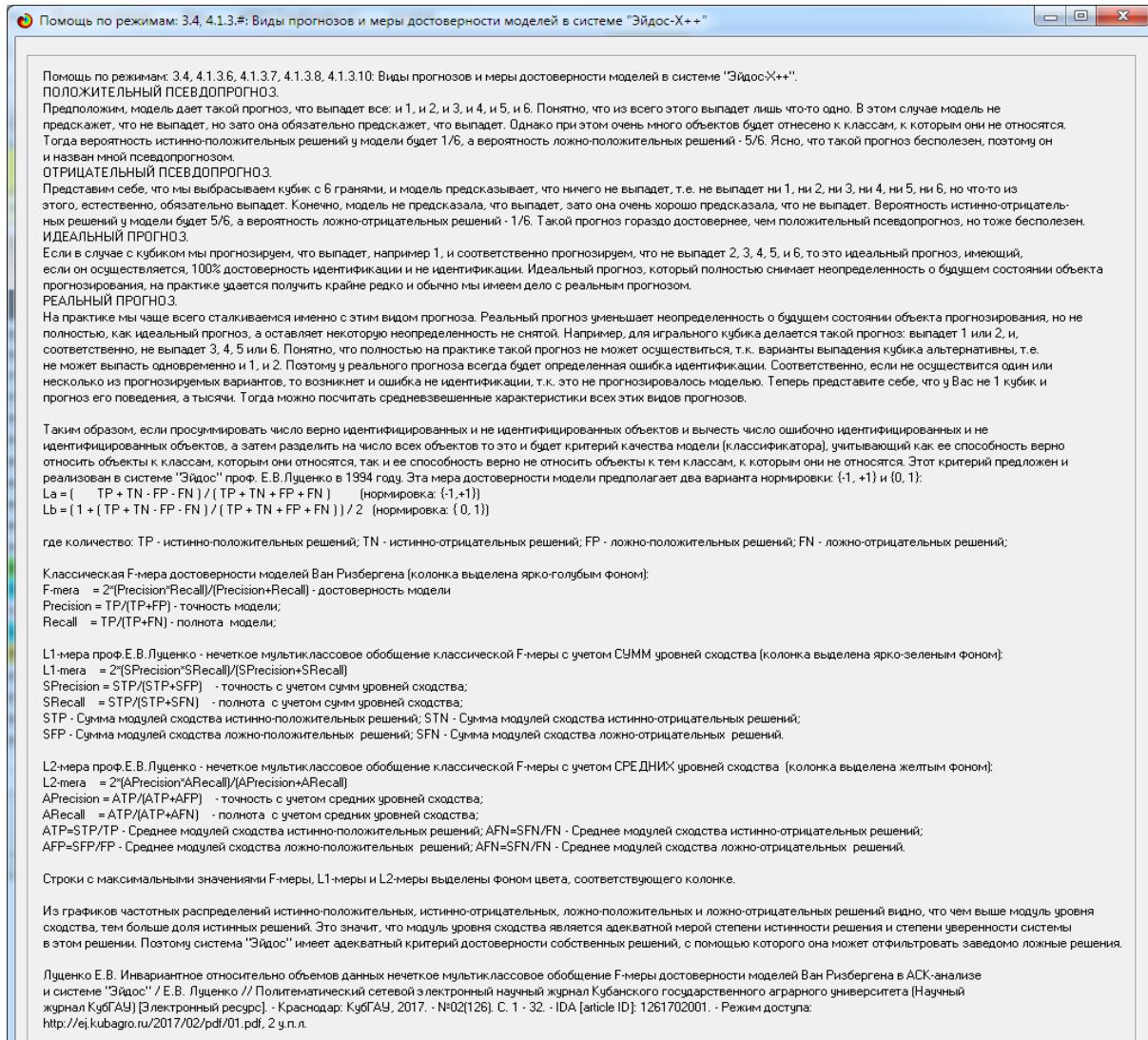


Рисунок 16. Экранные формы хелпов режима измерения достоверности моделей

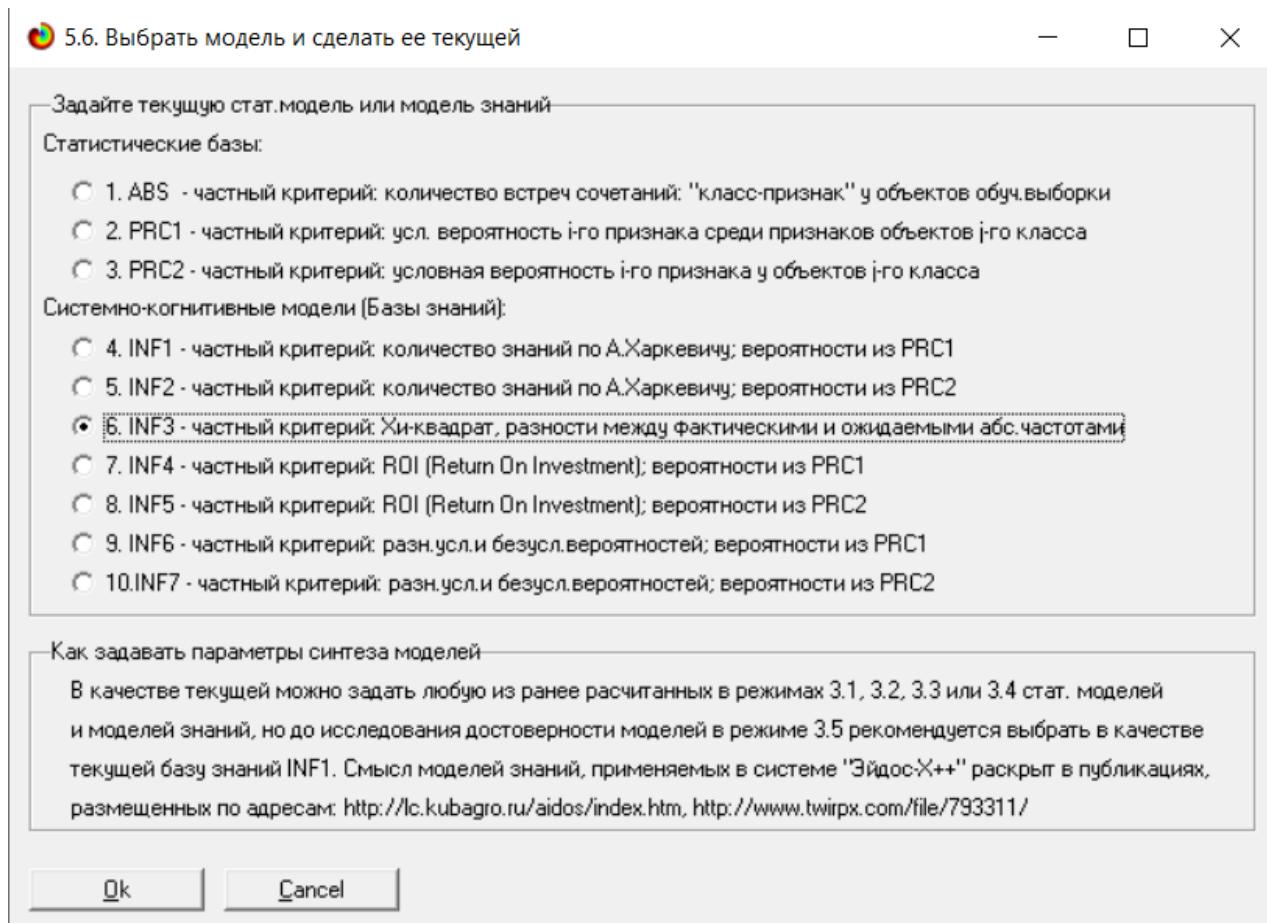
3.5. Задача-5. Выбор наиболее достоверной модели

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.

Технически сам выбор наиболее достоверной модели и задание ее текущей осуществляется в режиме 5.6 системы «Эйдос» и проходит быстро (рисунки 17). 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.



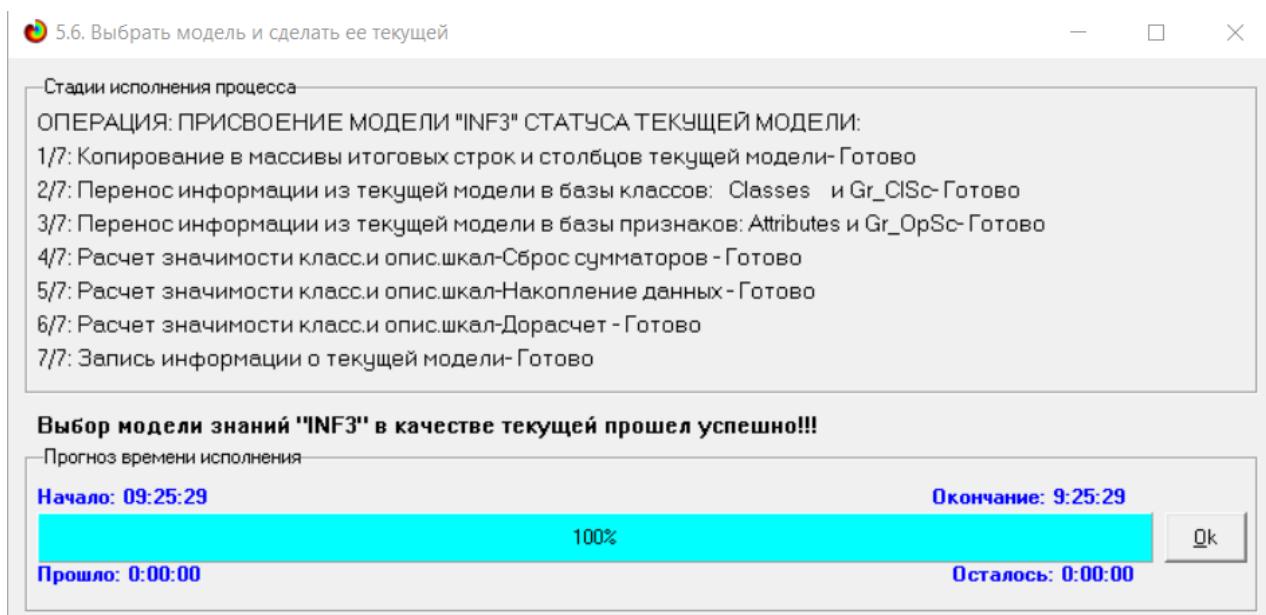


Рисунок 17. Задание СК-модели INF3 в качестве текущей

3.6. Задача-6. Системная идентификация и прогнозирование

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 классу 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» высокоеффективных алгоритмов распознавания и использованием графического процессора (GPU) для расчетов эта проблема практически снялась.

Сравнение осуществляется путем применения *неметрических интегральных критерий*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что

корректны⁹ в неортонормированных пространствах, которые всегда и встречаются на практике, и являются фильтрами подавления шума.

3.6.1. 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;

$\vec{L}_i = \{L_i\}$ – 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-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).

3.6.2. 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.

⁹ В отличие от Евклидова расстояния, которое используется для подобных целей наиболее часто

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 стандартизованными значениями:

$$I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}, \quad L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_l}. \text{ Поэтому по своей сути он также является скалярным}$$

произведением двух стандартизованных (единичных) векторов класса и объекта. Существуют и много других способов нормирования, например, путем применяя сплайнов, в частности линейной интерполяции:

$$I_{ij} \rightarrow \frac{I_{ij} - I_j^{\min}}{I_j^{\max} - I_j^{\min}}, \quad L_i \rightarrow \frac{L_i - L^{\min}}{L^{\max} - L^{\min}}, \text{ Это позволяет предложить неограниченное}$$

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».

3.6.3. Важные математические свойства интегральных критериев

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

Во-первых, интегральный критерий имеет **нemetрическую** природу, т.е. он являются 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 **неортонормированных** пространствах, которые, как правило, и встречаются на практике и в которых применение Евклидова расстояния (теоремы Пифагора) является некорректным.

Во-вторых, данный интегральный критерий являются **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 монографии [11, 12].

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

3.6.4. Решение задачи идентификации и прогнозирования 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. Но в задачи данной работы не входит их подробное рассмотрение, тем более, что они подробно

освящены и на теоретическом уровне, и с детальными численными примерами в работах [13, 14] и в ряде других¹⁰.

Поэтому в данной работе рассмотрим стандартный несколько упрощенный вариант решения задачи идентификации и прогнозирования. Запустим режим 4.1.2 системы «Эйдос» (рисунок 18):

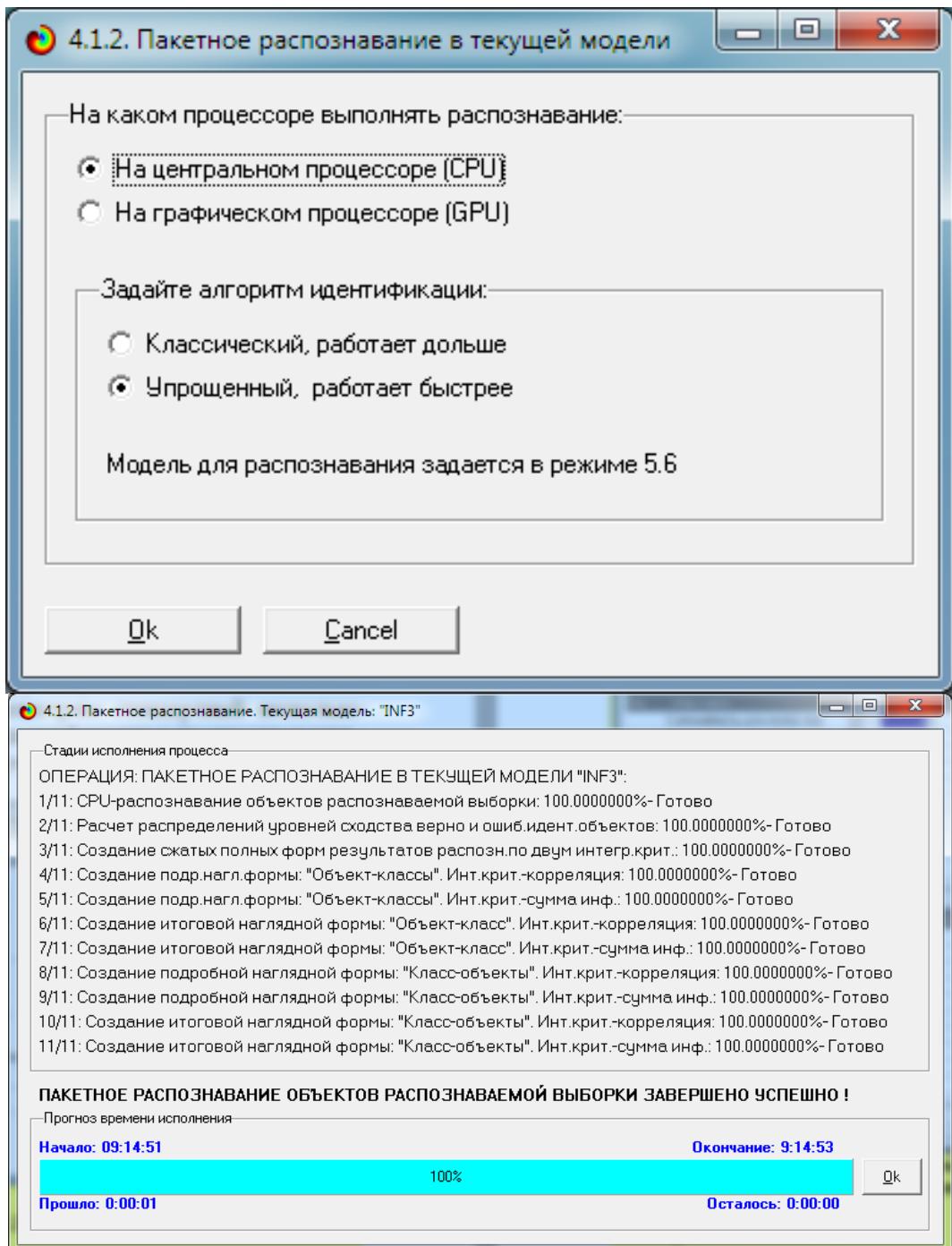


Рисунок 18. Экранные формы режима 4.1.2 идентификации и прогнозирования

¹⁰ См., например: http://lc.kubagro.ru/aidos/Works_on_Scenario_ASC-analysis.htm

По результатам решения задачи идентификации и прогнозирования в системе «Эйдос» выдается большое количество разнообразных выходных форм, которых в настоящее время 12 (рисунок 19):

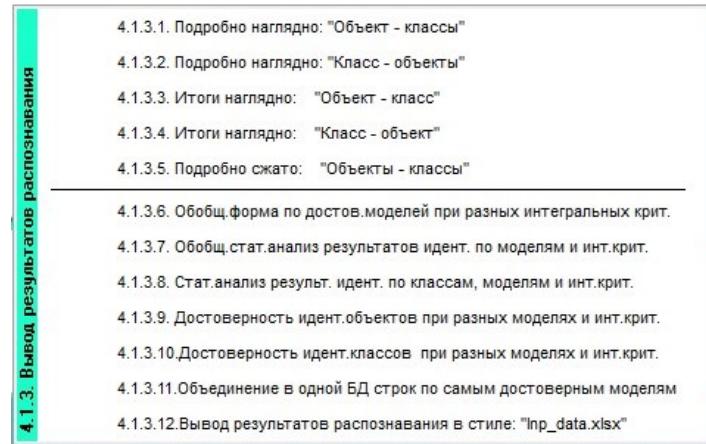
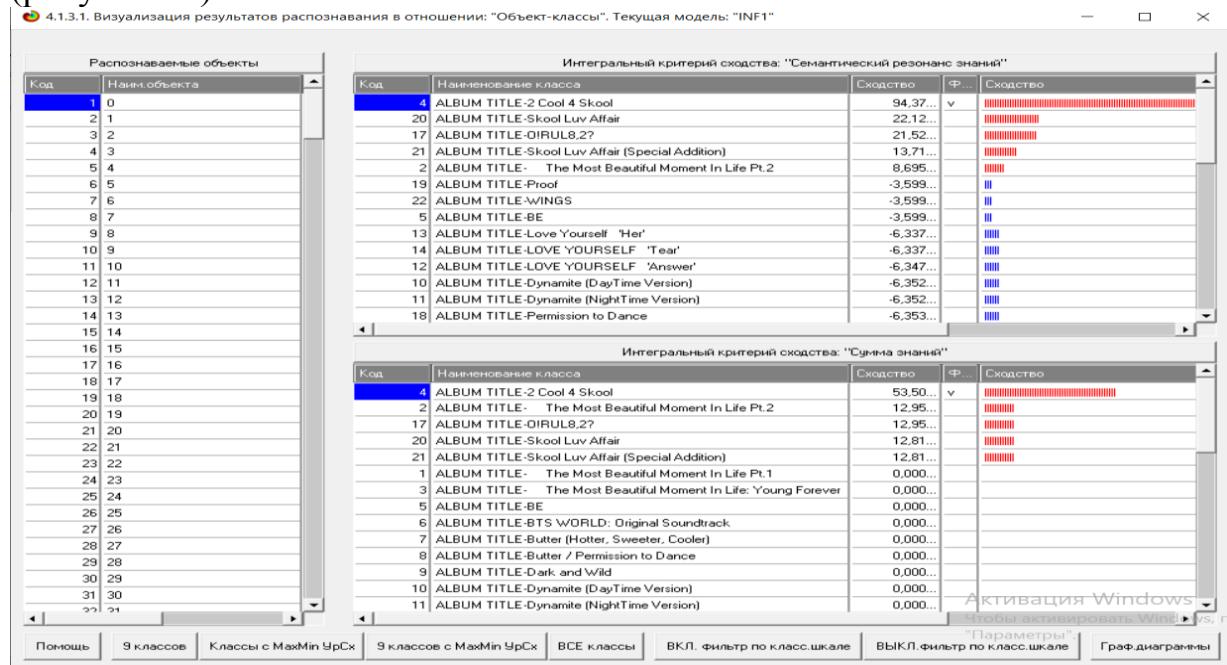


Рисунок 19. Меню подсистемы 4.1.3 системы «Эйдос», обеспечивающей вывод форм по результатам решения задачи идентификации и прогнозирования

Из этих выходных форм рассмотрим только две: 4.1.3.1 и 4.1.3.2 (рисунок 20):



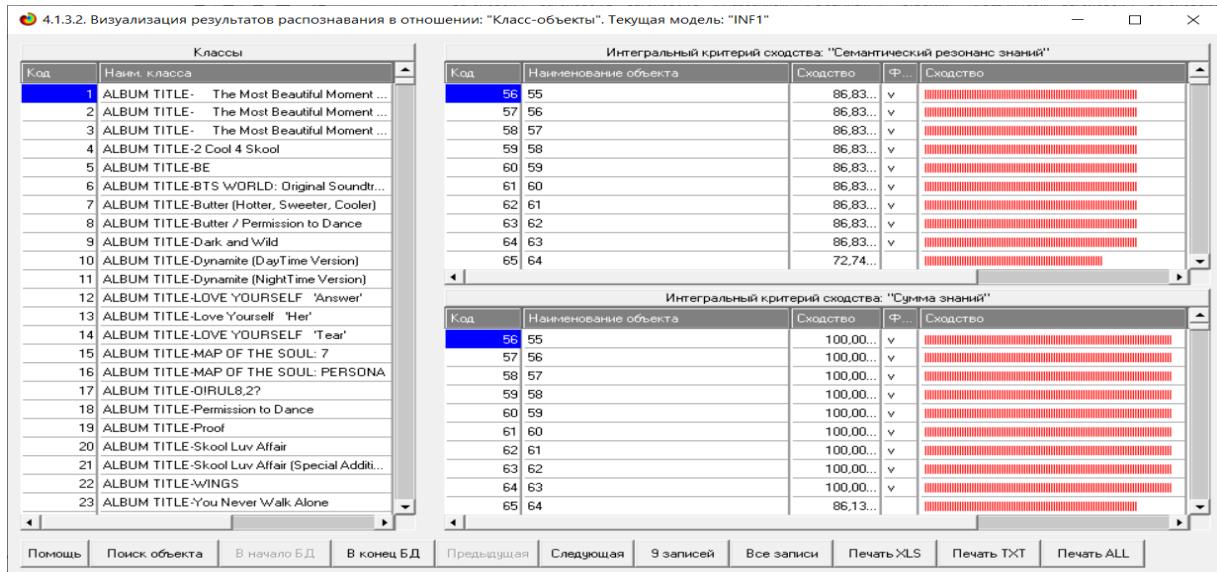


Рисунок 20. Некоторые экranные формы результатов идентификации и прогнозирования 4.1.3 системы «Эйдос»

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

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») [15] (рисунки 21).

Выходные формы, приведенные на рисунках 21, 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.

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

Выбор класса, соответствующего будущему состоянию объекта управления

Код	Наименование класса	Редакция клас...	N объектов (абс.)	N объектов (%)
1	ALBUM TITLE- The Most Beautiful Moment In Life Pt.1	0,3609864	72	3,0000000
2	ALBUM TITLE- The Most Beautiful Moment In Life Pt.2	0,3097820	72	3,0000000
3	ALBUM TITLE- The Most Beautiful Moment In Life: Young Forever	0,3066025	184	7,6666667
4	ALBUM TITLE-2 Cool 4 Skool	0,4891678	36	3,0000000
5	ALBUM TITLE-BE	0,4286753	8	2,6666667
6	ALBUM TITLE-BTS WORLD: Original Soundtrack	0,5034730	56	4,6666667

SWOT-анализ класса:1 "ALBUM TITLE- The Most Beautiful Moment In Life Pt.1" в модели:6 "INF3"

Способствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
1	ENG ALBUM TITLE-1	8.407
38	ENG ALBUM TITLE-Pt	7.814
11	ENG ALBUM TITLE-Beautiful	6.299
23	ENG ALBUM TITLE-In	6.299
24	ENG ALBUM TITLE-Life	6.299
28	ENG ALBUM TITLE-Moment	6.299
29	ENG ALBUM TITLE-Most	6.299
46	ENG ALBUM TITLE-The	4.455

Препятствующие факторы и сила их влияния

Код	Наименование фактора и его интервального значения	Сила влияния
55	ENG ALBUM TITLE-Yourself	-3.228
25	ENG ALBUM TITLE-Love	-3.228
37	ENG ALBUM TITLE-Proof	-3.162
40	ENG ALBUM TITLE-Skool	-2.042
41	ENG ALBUM TITLE-SOUL	-1.844
33	ENG ALBUM TITLE-OF	-1.844
27	ENG ALBUM TITLE-MAP	-1.844
2	ENG ALBUM TITLE-2	-1.844
9	ENG ALBUM TITLE-Answer	-1.779
54	ENG ALBUM TITLE-Young	-1.515
20	ENG ALBUM TITLE-Forever	-1.515
26	ENG ALBUM TITLE-Luv	-1.449
6	ENG ALBUM TITLE-Affair	-1.449
4	FNG AI RIIM TITI F-7	-1.393

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

SWOT-диаграмма

SWOT-ДИАГРАММА КЛАССА В МОДЕЛИ: "INF3"

Приложение: "ACK-анализ BTS"

Шкала: [1] ALBUM_TITLE
Класс: [1] НИИ D0! ` 7+M!Eal MAX_GRIN !!

СПОСОБСТВУЮЩИЕ значения факторов и сила их влияния:

- [1] ENG_ALBUM_TITLE [1] 1 I=8.407
- [1] ENG_ALBUM_TITLE [38] Pt I=7.814
- [1] ENG_ALBUM_TITLE [11] Beautiful I=6.299
- [1] ENG_ALBUM_TITLE [23] In I=6.299
- [1] ENG_ALBUM_TITLE [24] Life I=6.299
- [1] ENG_ALBUM_TITLE [28] Moment I=6.299
- [1] ENG_ALBUM_TITLE [29] Most I=6.299

Фильтр по факторам ВЫКЛЮЧЕН. Диапазон кодов значений: 1-55

ПРЕПЯТСТВУЮЩИЕ значения факторов и сила их влияния:

- [1] ENG_ALBUM_TITLE [55] Yourself I=-3.228
- [1] ENG_ALBUM_TITLE [25] Love I=-3.228
- [1] ENG_ALBUM_TITLE [37] Proof I=-3.162
- [1] ENG_ALBUM_TITLE [40] Skool I=-2.042
- [1] ENG_ALBUM_TITLE [41] SOUL I=-1.844
- [1] ENG_ALBUM_TITLE [33] OF I=-1.844
- [1] ENG_ALBUM_TITLE [27] MAP I=-1.844

Фильтр по факторам ВЫКЛЮЧЕН. Диапазон кодов значений: 1-55

СИСТЕМА ДЕТЕРМИНАЦИИ КЛАССА ФАКТОРАМИ И ИХ ЗНАЧЕНИЯМИ:
 Значения факторов, СПОСОБСТВУЮЩИЕ переходу объекта управления в состояние, соответствующее классу, отображаются линиями связи КРАСНОГО цвета. Толщина линии отражает степень влияния.
 Значения факторов, ПРЕПЯТСТВУЮЩИЕ переходу объекта управления в состояние, соответствующее классу, отображаются линиями связи СИНЕГО цвета. Толщина линии отражает степень влияния.

Форма создана: 27.12.2022-04:31:31

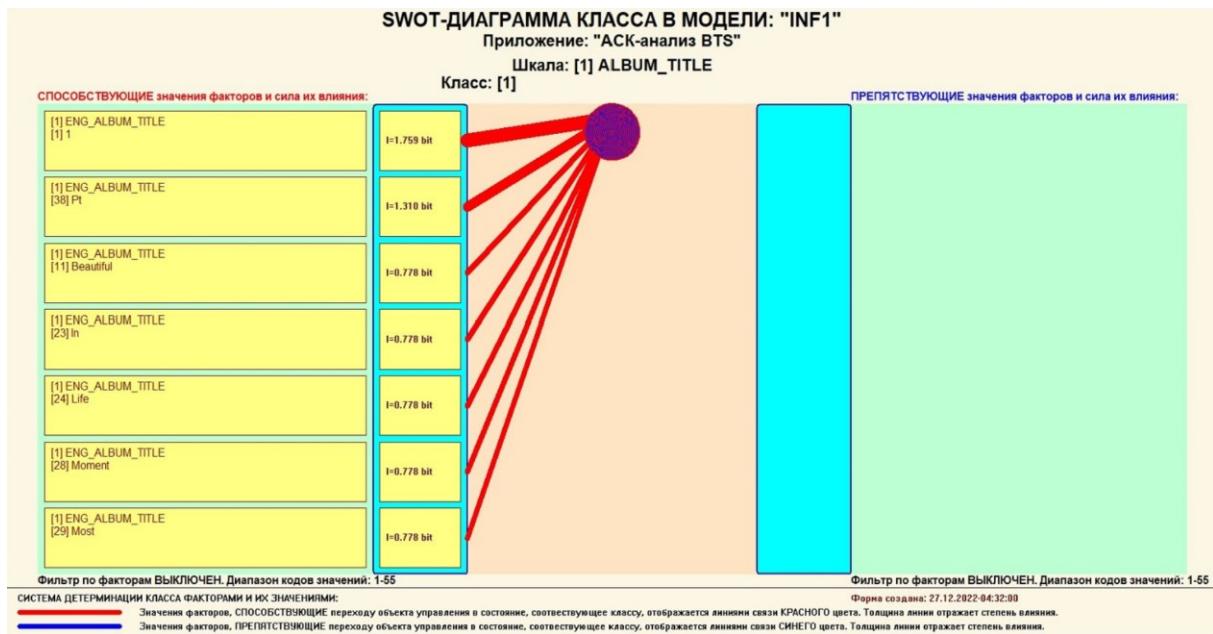


Рисунок 21. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

На первом рисунке 21 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.

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.

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

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 [13, 14, 16] и в ряде других работ.

Приведем этот алгоритм и в данной работе (рисунок 22).

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 [16, 17, 18].

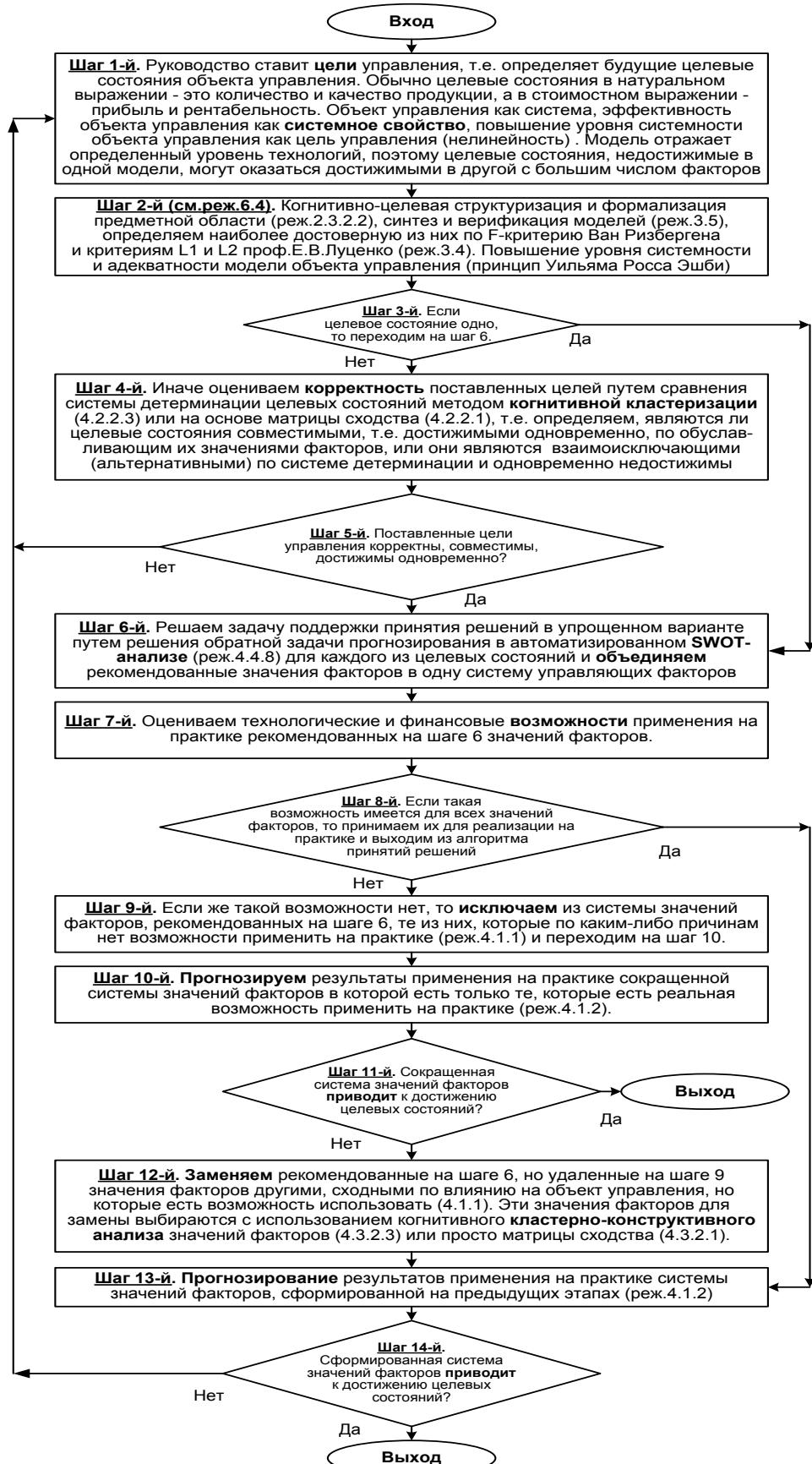


Рисунок 22. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

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) [5]. Повышение уровня системности и адекватности модели объекта управления (принцип Уильяма Росса Эшби) [17].

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 совместимы, т.е. достижимы 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 [15].

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) [19].

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 23:

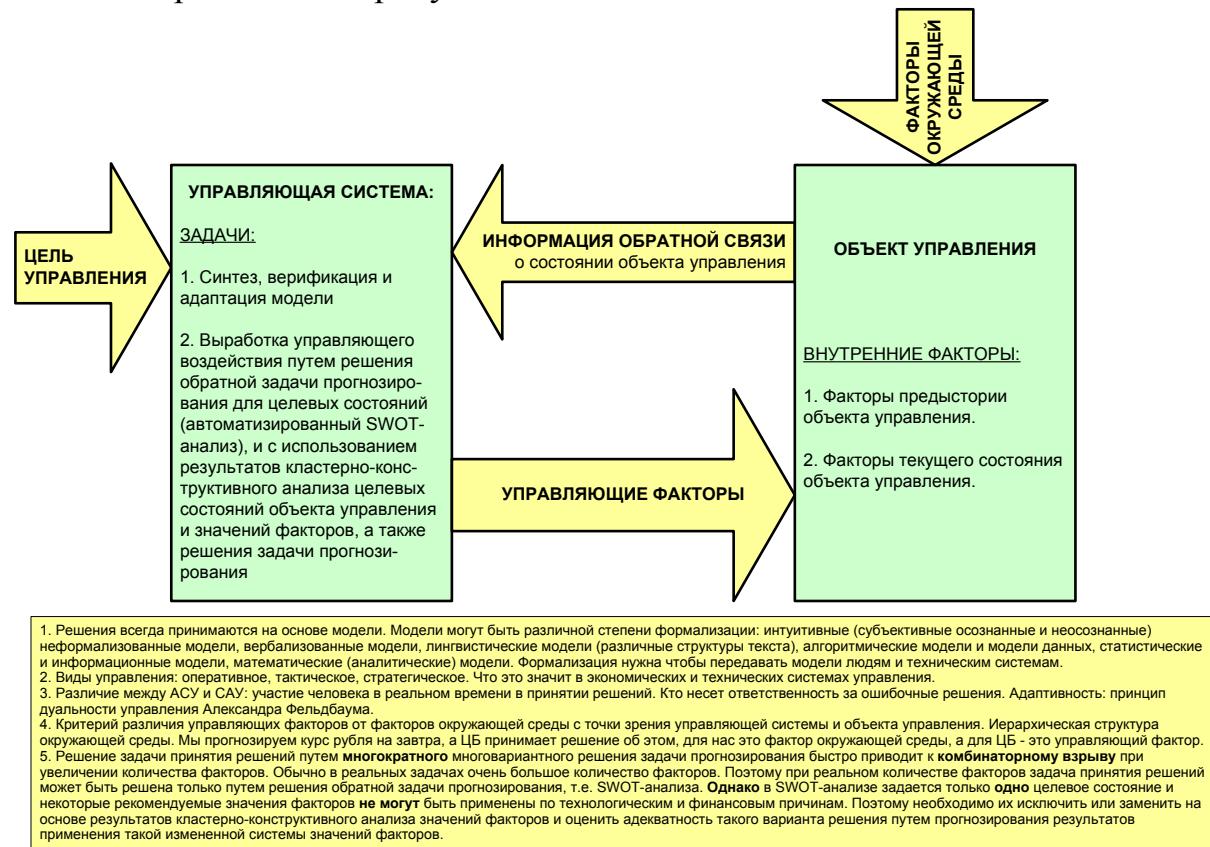


Рисунок 23. Принципиальная схема адаптивной интеллектуальной системы управления на основе АСК-анализа и системы «Эйдос»

Как мы видим в развитом алгоритме принятия решений широко используются результаты 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 задач, которые необходимо решать в развитом алгоритме принятия решений. Потому ниже кратко рассмотрим решение этих и некоторых других задач.

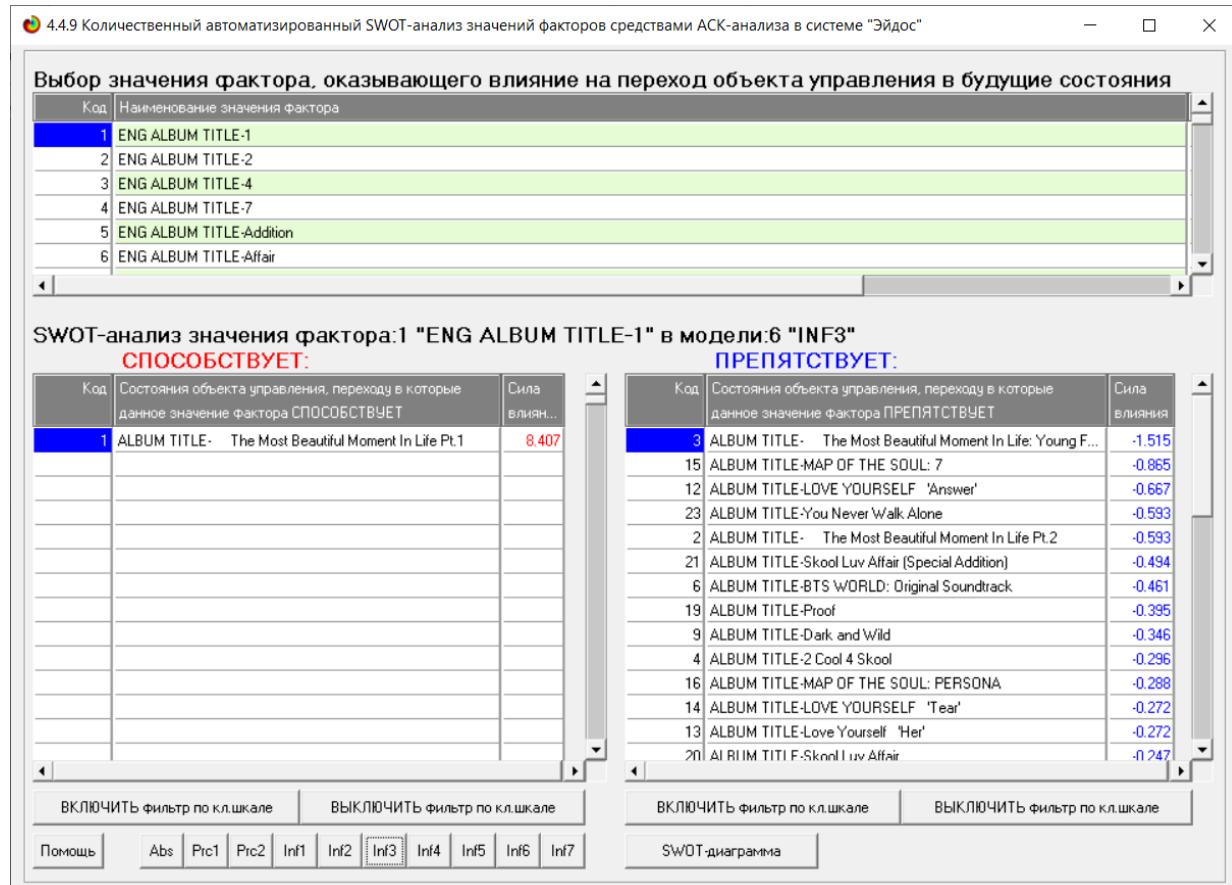
Необходимо отметить, что модели системы Эйдос – это феноменологические модели, отражающие эмпирические закономерности в фактах обучающей выборки, т.е. они не отражают причинно-следственного механизма детерминации, а только сам факт и характер детерминации. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [21, 22].

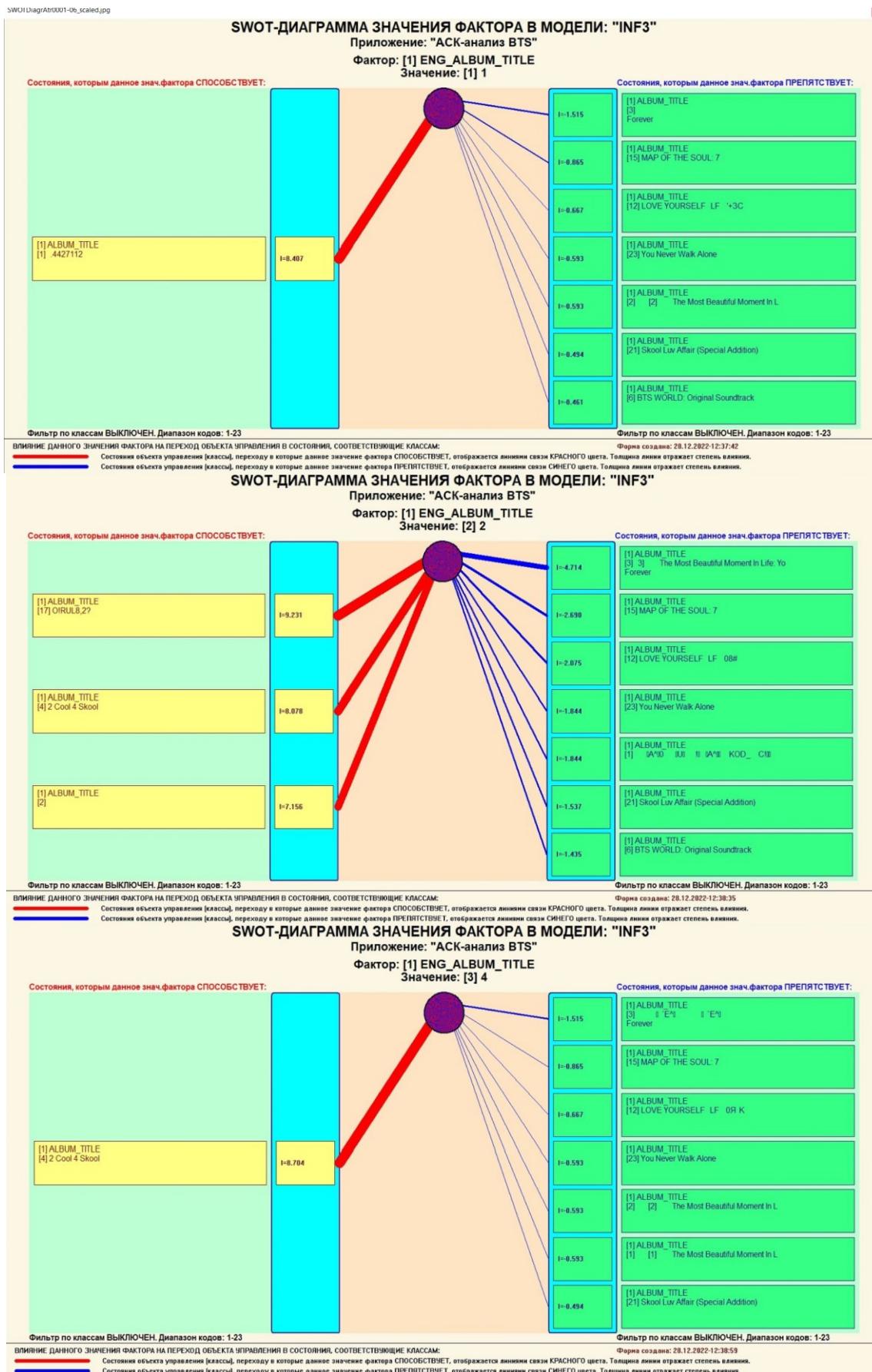
3.8. Задача-8. Исследование объекта моделирования путем исследования его модели

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

Инвертированные SWOT-диаграммы (предложены автором в работе [15]), 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».

Инвертированные SWOT-диаграммы для каждого значения фактора, которые представляют собой лингвистические переменные, приведены ниже на рисунке 24:





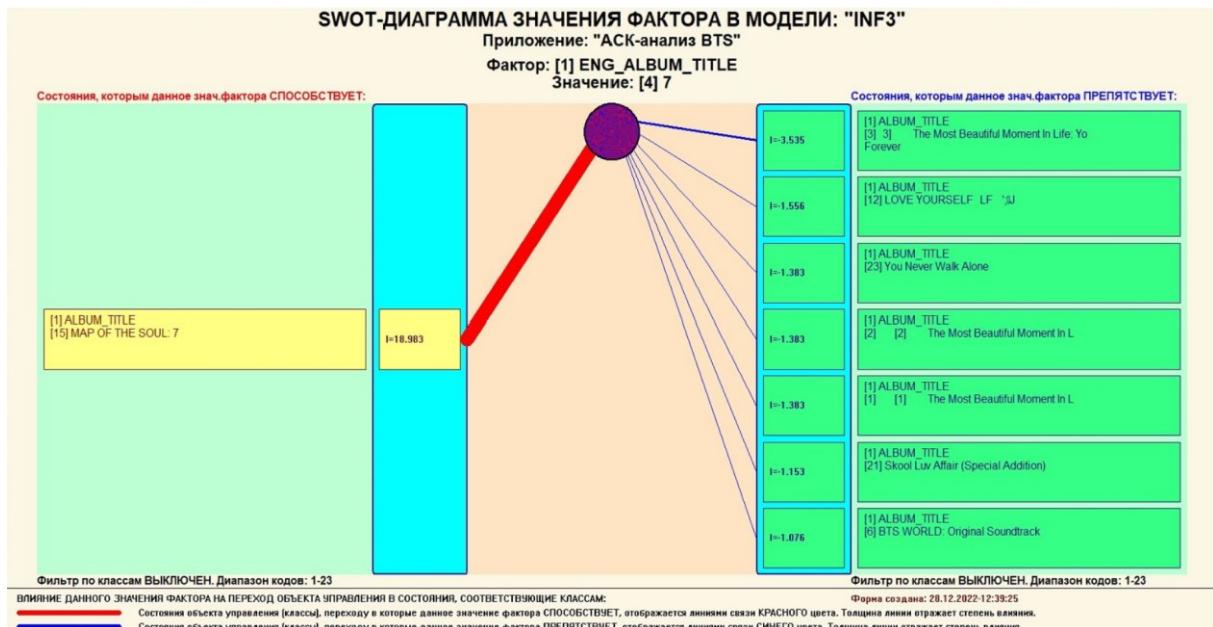


Рисунок 24. Примеры инвертированных SWOT-диаграмм влияния значений факторов на переход объекта моделирования в будущие состояния, соответствующие классам

Приведенные на рисунке 24 инвертированные SWOT-диаграммы исчерпывающим образом отражают силу и направление влияния каждого значения каждого фактора на переход объекта моделирования в состояния, соответствующие различным классам. Во многом это и есть решение проблемы, поставленной в работе.

3.8.2. Кластерно-конструктивный анализ классов

В системе «Эйдос» (в режиме 4.2.2.1, рисунок 25) рассчитывается матрица сходства классов 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 26);
- агломеративные дендрограммы, полученные в результате **когнитивной (истинной) кластеризации классов** (предложена автором в 2011 году в работе [19]) (режим 4.2.2.3) (рисунок 27);
- график изменения межкластерных расстояний (режим 4.2.2.3) (рисунок 28).

Эта матрица используется и при расчете некоторых других выходных форм.

На рисунке 25 представлены экранные формы режима 4.2.2.1, обеспечивающего расчет матрицы сходства классов по системе их детерминации, т.е. по обуславливающим их значениям факторов:

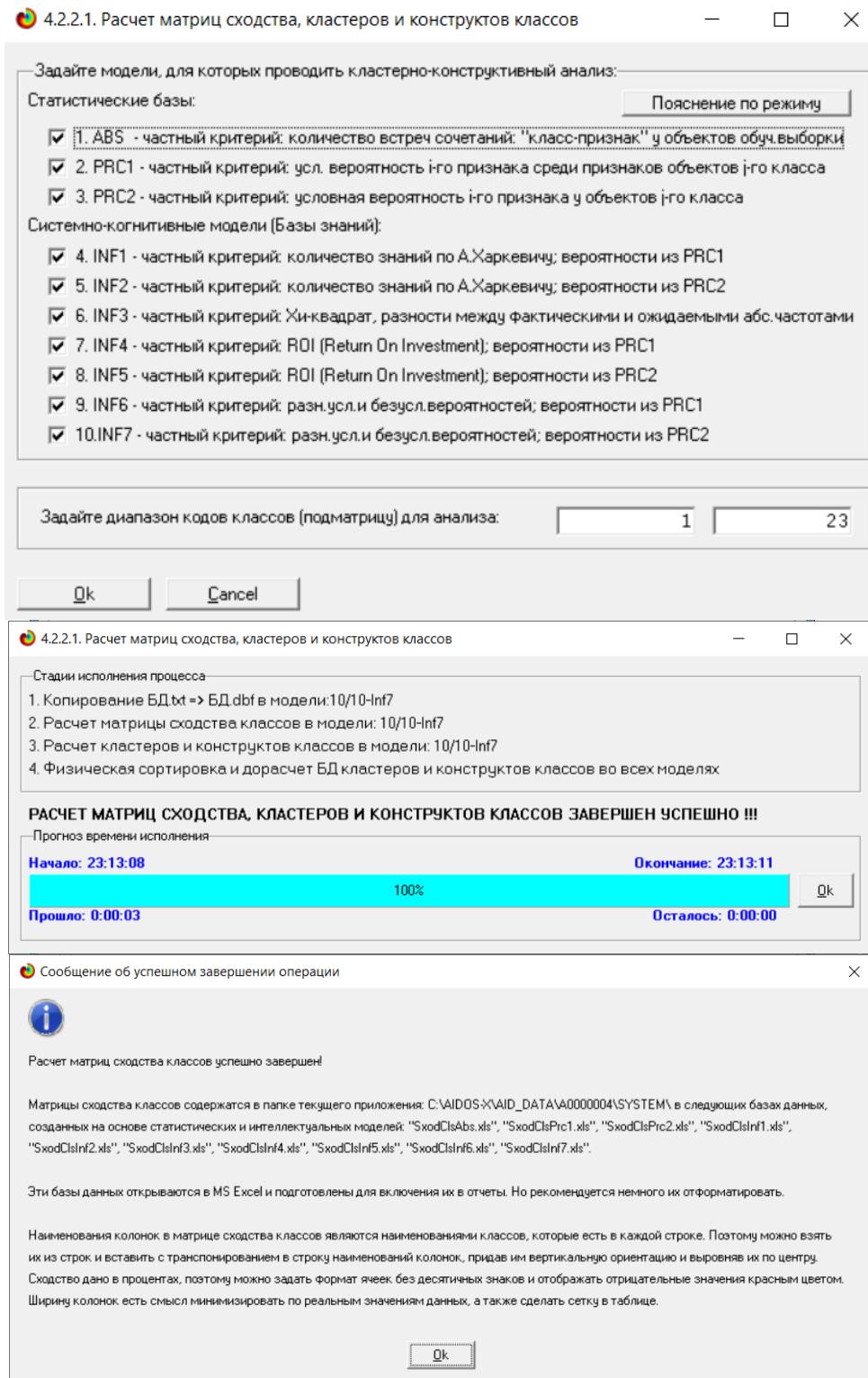


Рисунок 25. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов

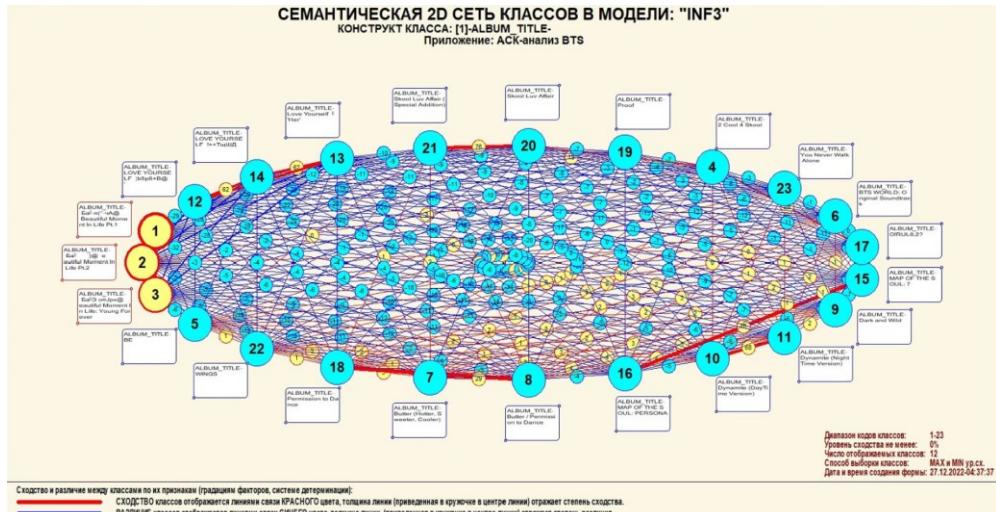


Рисунок 26. Круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2)

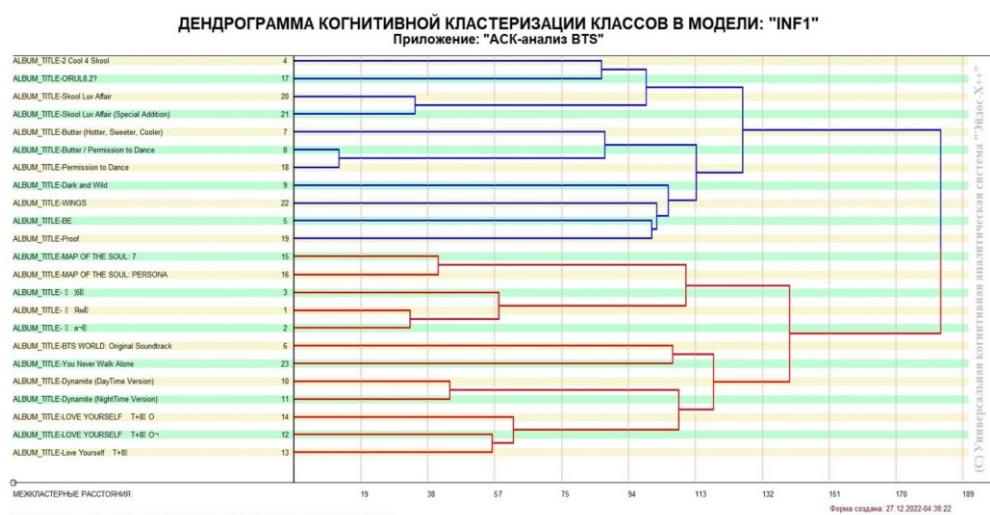


Рисунок 27. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации классов (режим 4.2.2.3)

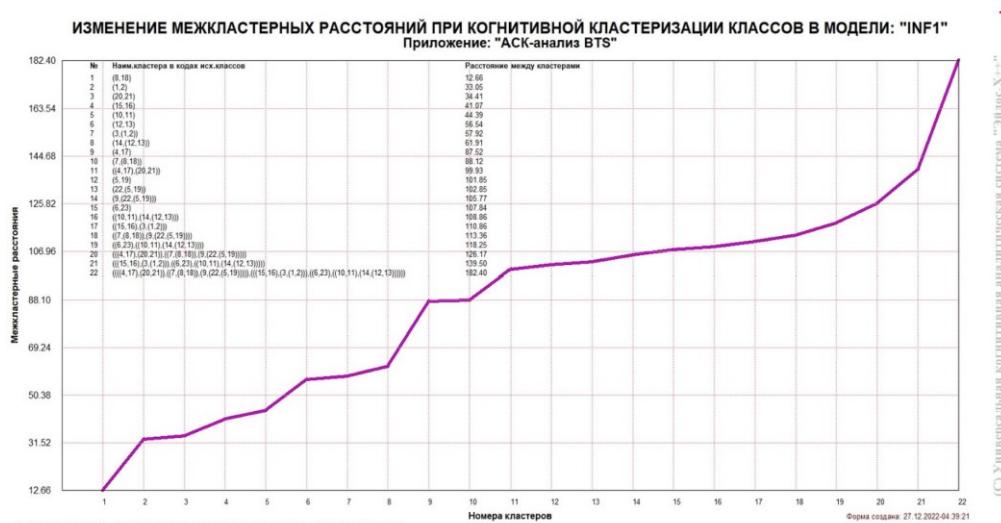


Рисунок 28. График изменения межклusterных расстояний (режим 4.2.2.3)

3.8.3. Кластерно-конструктивный анализ значений описательных шкал

В системе «Эйдос» (в режиме 4.3.2.1, рисунок 29) 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 29);
- агломеративные дендрограммы, полученные в результате **когнитивной (истинной) кластеризации признаков** (предложена автором в 2011 году в работе [19]) (режим 4.3.2.3) (рисунок 30);
- график изменения межкластерных расстояний (режим 4.3.2.3) (рисунок 31).

Матрица сходства используются и при расчете некоторых других выходных форм.

На рисунке 29 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:

4.3.2.1. Расчет матриц сходства, кластеров и конструктов

— □ ×

Задайте модели, для которых проводить кластерно-конструктивный анализ:

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 abs. 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 безусл. вероятностей; вероятности из PRC1

10. INF7 - частный критерий: разн. усл. и безусл. вероятностей; вероятности из PRC2

Задайте диапазон кодов признаков (подматрицу) для анализа: 1 55

Ok Cancel

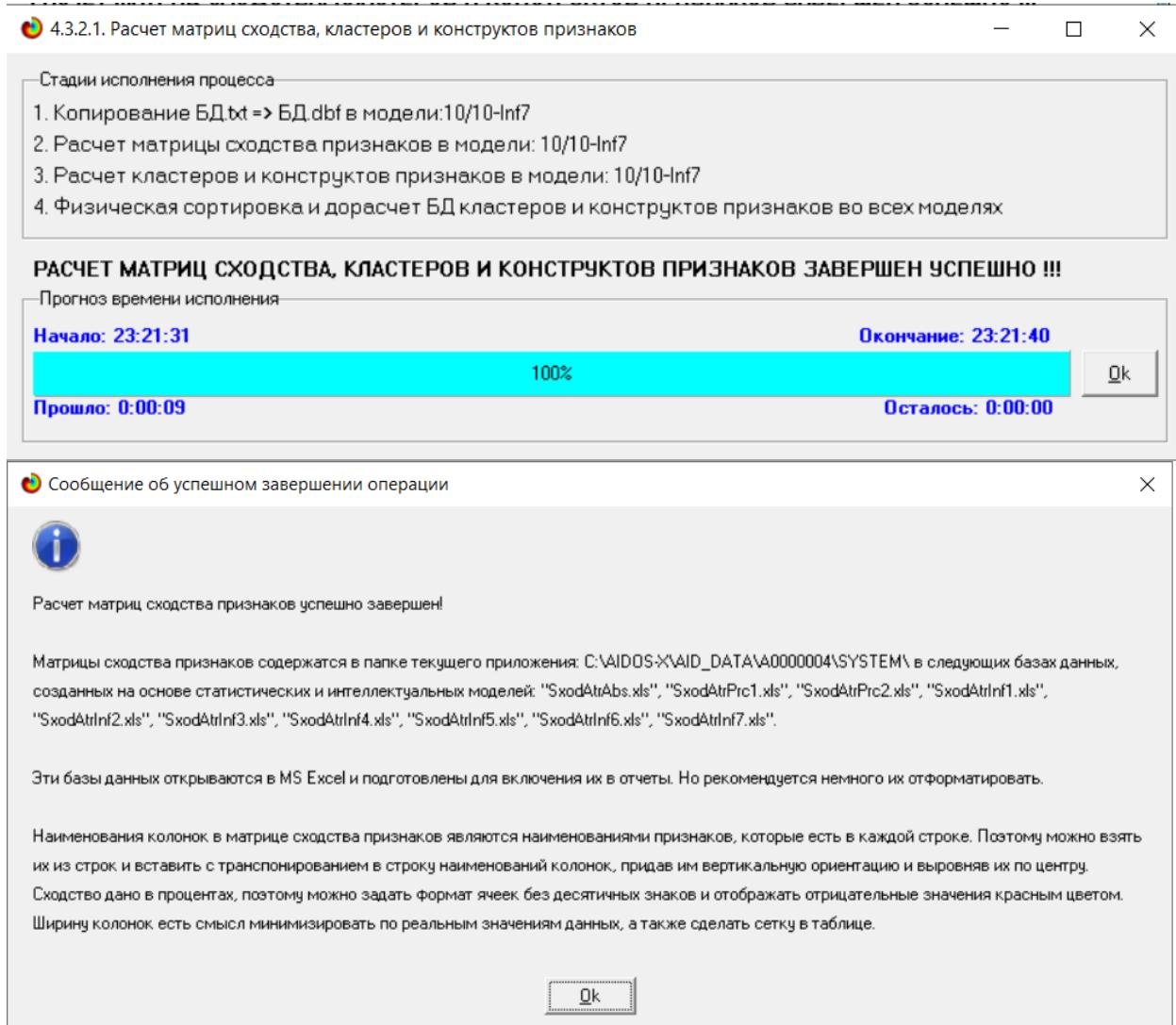


Рисунок 29. Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов

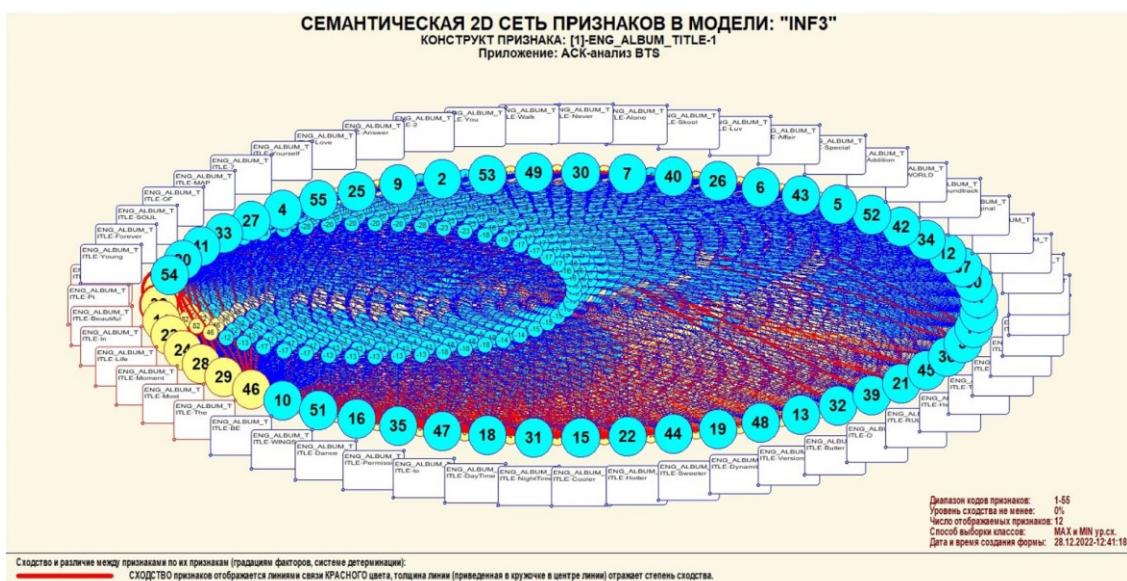


Рисунок 30. Круговая 2d-когнитивная диаграмма значений факторов в СК-модели INF1 (режим 4.3.2.2)

ДЕНДРОГРАММА КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF3"
Приложение: "ACK-анализ BTS"

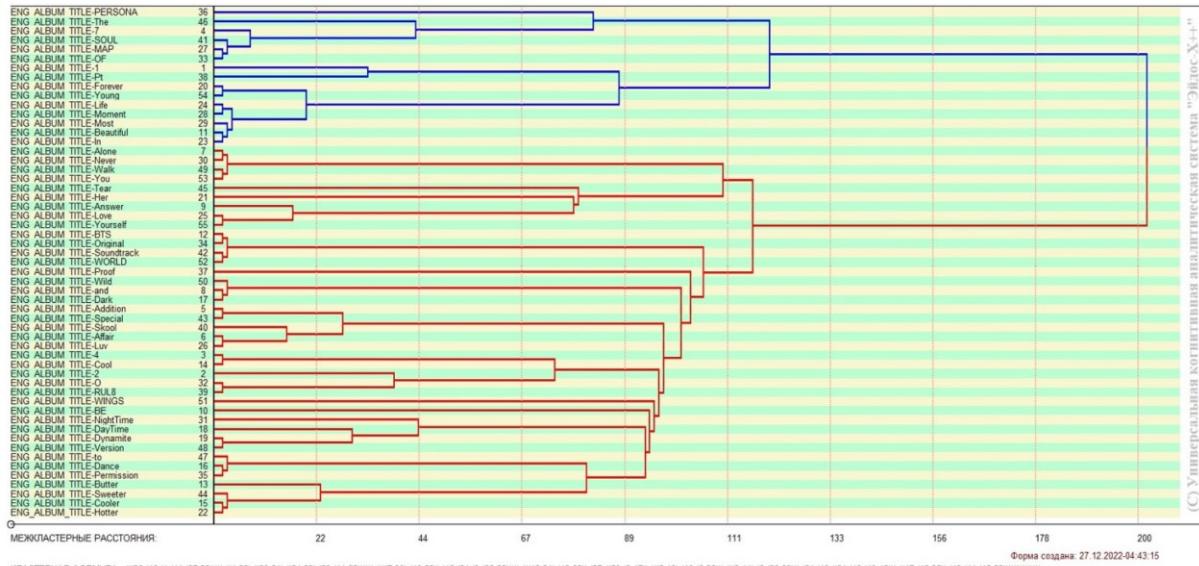


Рисунок 31. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)

ИЗМЕНЕНИЕ МЕЖКЛАСТЕРНЫХ РАССТОЯНИЙ ПРИ КОГНИТИВНОЙ КЛАСТЕРИЗАЦИИ ПРИЗНАКОВ В МОДЕЛИ: "INF3"
Приложение: "ACK-анализ BTS"

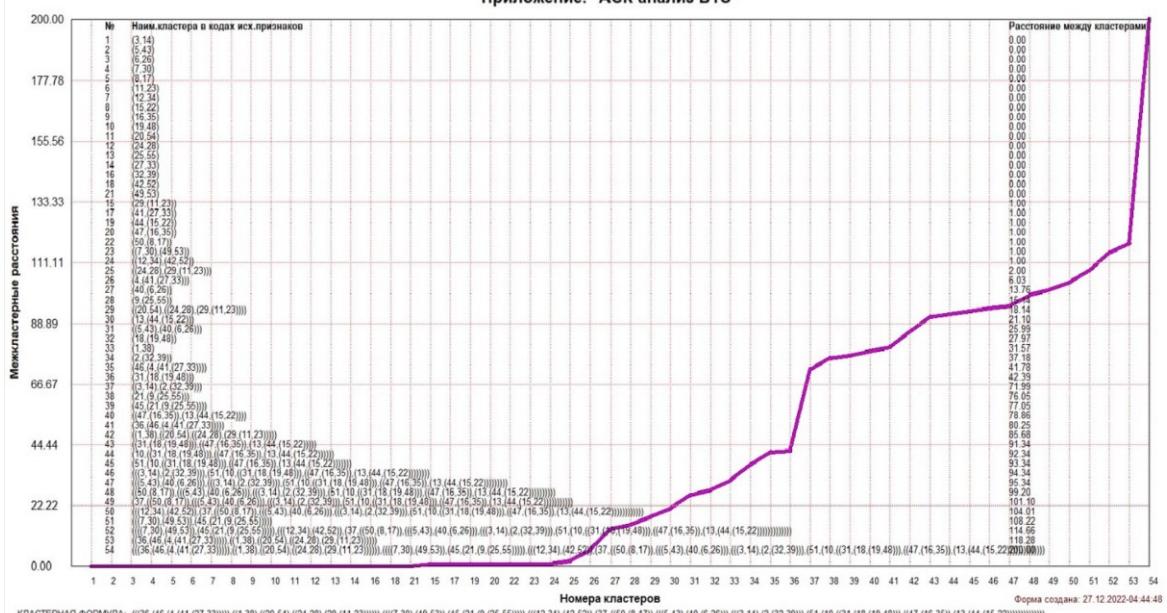


Рисунок 32. График изменения межкластерных расстояний (режим 4.3.2.3)

3.8.4. Модель знаний системы «Эйдос» и нелокальные нейроны

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 реализации.

От нейросетевой модели представления знаний модель системы «Эйдос» отличается тем, что [20]:

- 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 (рисунок 33). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

3.8.5. Нелокальная нейронная сеть

В системе «Эйдос» есть возможность построения моделей, соответствующих многослойным нейронным сетям [20].

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 (рисунок 34). В форме управления визуализацией есть возможность задавать фильтры по факторам, которые надо визуализировать.

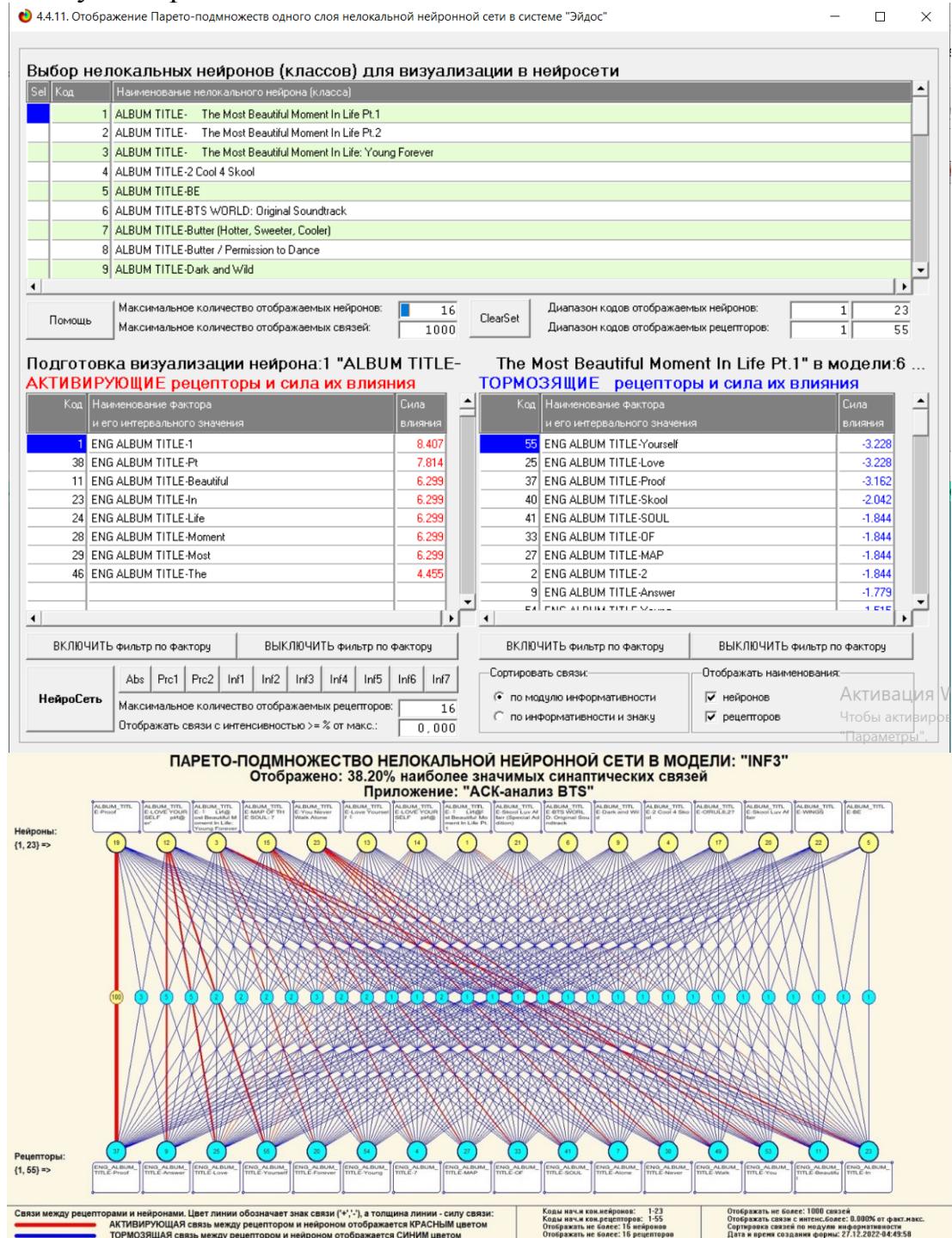


Рисунок 33. Нейронная сеть в СК-модели INF3

3.8.6. 3D-интегральные когнитивные карты

3d-интегральная когнитивная карта является отображением на одном рисунке когнитивной диаграммы классов (рисунок 26) вверху и когнитивной диаграммы значений факторов (рисунок 30) внизу и

соединяющего их одного слоя нейронной сети (рисунок 34) (режим 4.4.12 системы «Эйдос») (рисунок 35):

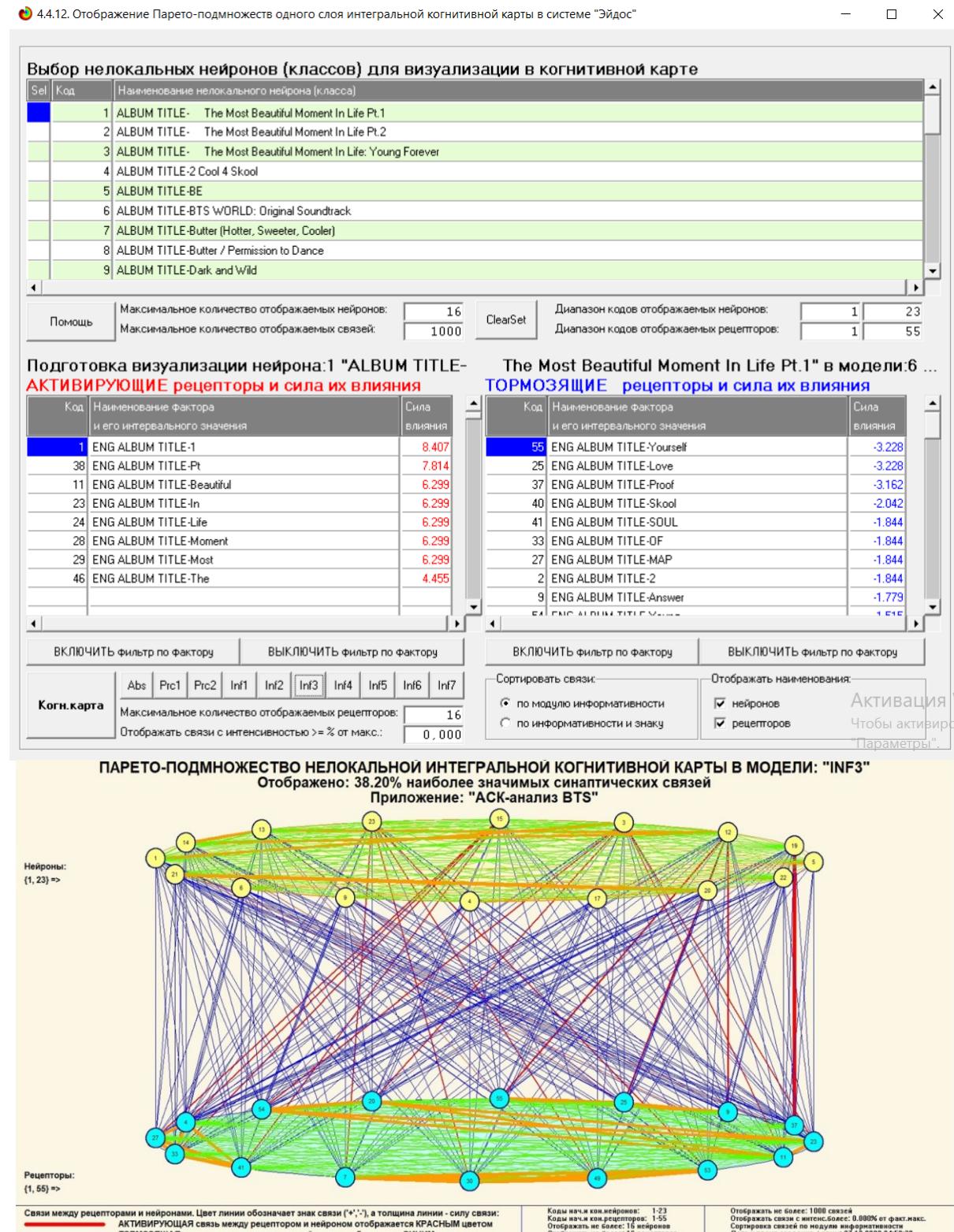


Рисунок 34. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

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

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 [23]. 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-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.

Примеры экранной формы управления и нескольких 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации приведены ниже на рисунках 36. Всего системой в данной модели генерируется 324 подобных диаграмм, поэтому, естественно, все они не приводятся. Но 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 №335 и получить в нем все выходные формы, как описано в данной статье.

¹¹ https://www.elibrary.ru/download/elibrary_18632909_64818704.pdf, Таблица 4. 17, стр. 521

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

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	ALBUM_TITLE - The Most Beautiful Moment In Life Pt.1
2	ALBUM_TITLE - The Most Beautiful Moment In Life Pt.2
3	ALBUM_TITLE - The Most Beautiful Moment In Life: Young Forever
4	ALBUM_TITLE-2 Cool 4 Skool
5	ALBUM_TITLE-BE
6	ALBUM_TITLE-BTS WORLD: Original Soundtrack

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	55
1	ENG_ALBUM_TITLE	1	55

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

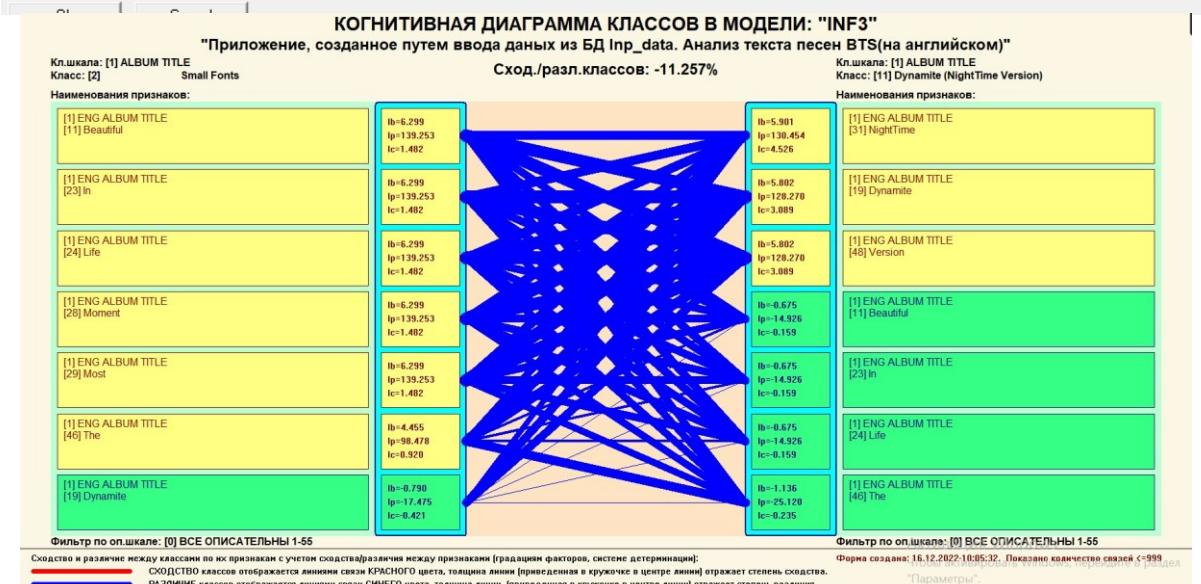
Задайте max количество отображаемых связей: Помощь

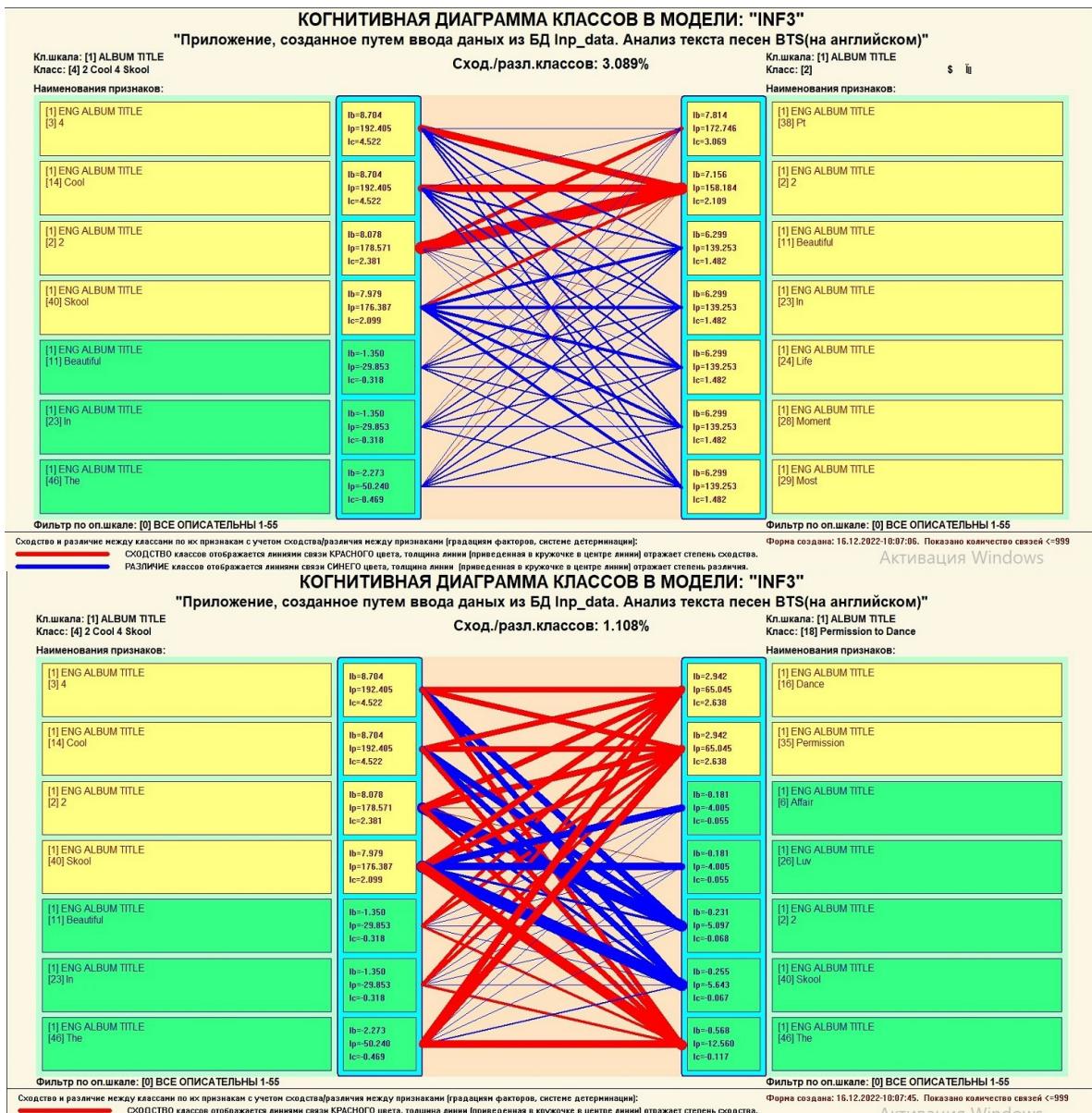
В диалоге заданы следующие параметры расчета когнитивных диаграмм:

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: Abs, Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой
 Записать все диаграммы без показа





4.2.3. Когнитивные диаграммы классов

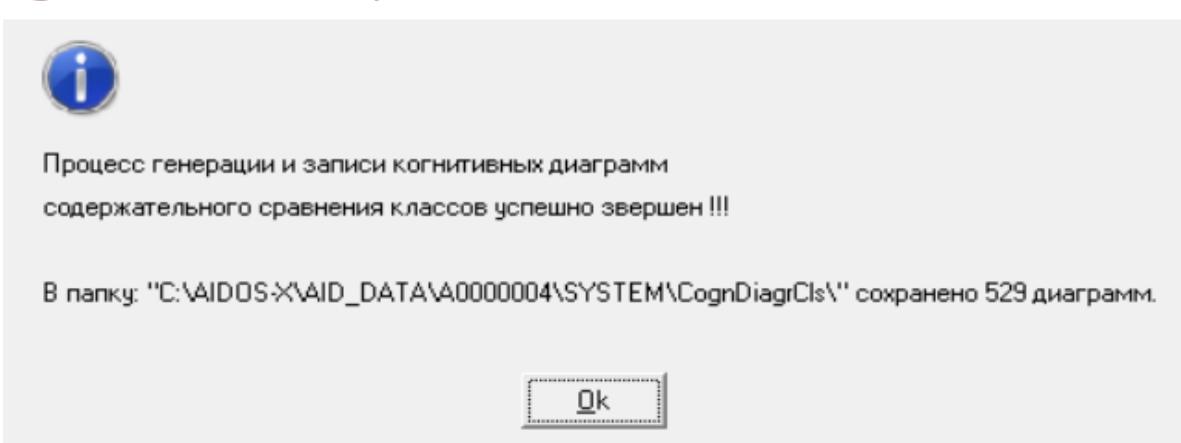


Рисунок 35. Примеры 2d-интегральных когнитивных карт содержательного сравнения классов по их системе детерминации в СК-модели INF3

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

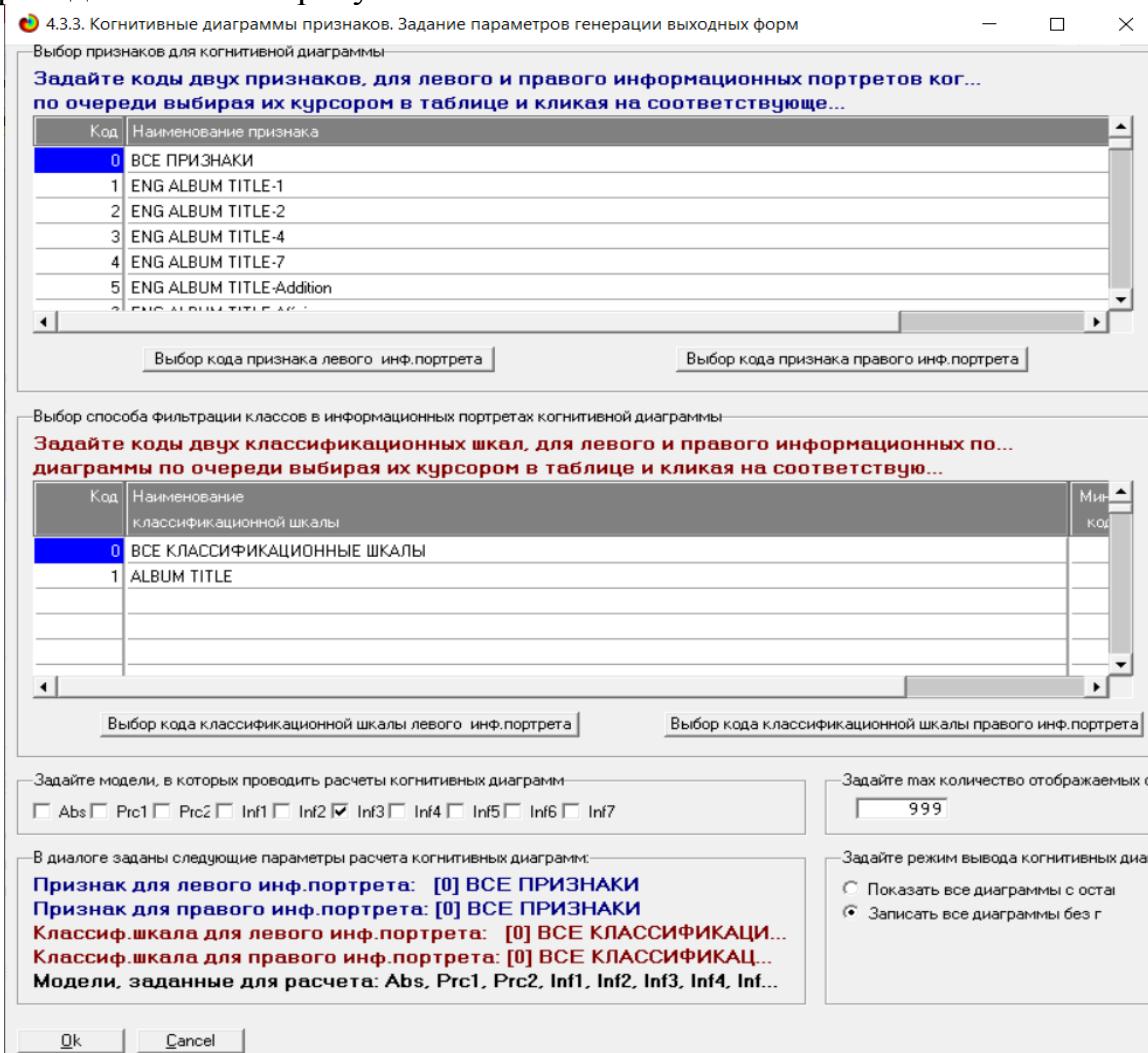
Из 2d-когнитивных диаграммах сравнения значений факторов по их влиянию на объект моделирования, т.е. на его переходы в состояния, соответствующие классам вполне понятно, насколько сходны или отличаются любые два значения факторов по их смыслу.

Напомним, что смысл, согласно концепции смысла Шенка-Абельсона, используемой в АСК-анализе, состоит в знании причин и последствий [24].

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

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

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 37.





4.3.3. Когнитивные диаграммы признаков

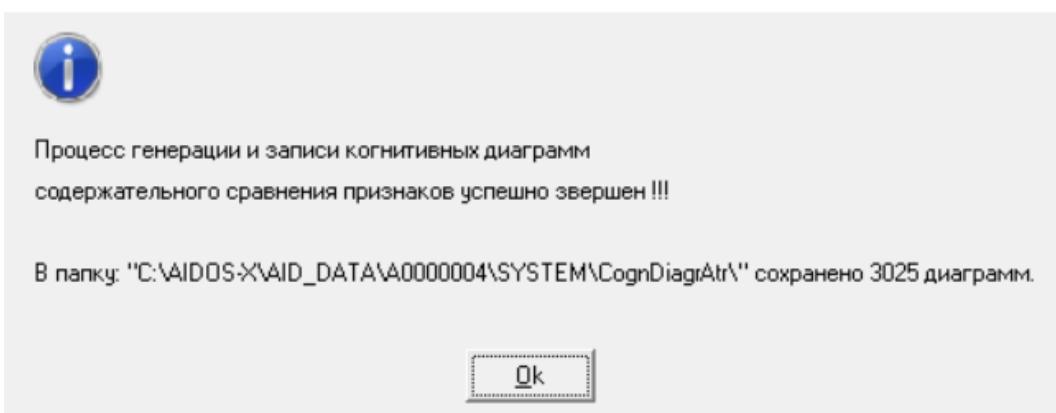


Рисунок 36. Примеры 2d-интегральных когнитивных карт содержательного сравнения значений факторов по их влиянию на переход объекта моделирования в состояния, соответствующие классам в СК-модели INF3

Всего системой в данной модели генерируется 3025 подобных диаграмм, поэтому, естественно, все они не приводятся. Но 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 №335 и получить в нем все выходные формы, как описано в данной статье.

3.8.9. Когнитивные функции

Когнитивные функции являются обобщением классического математического понятия функции на основе системной теории информации и предложены Е.В.Луценко в 2005 году [25, 26].

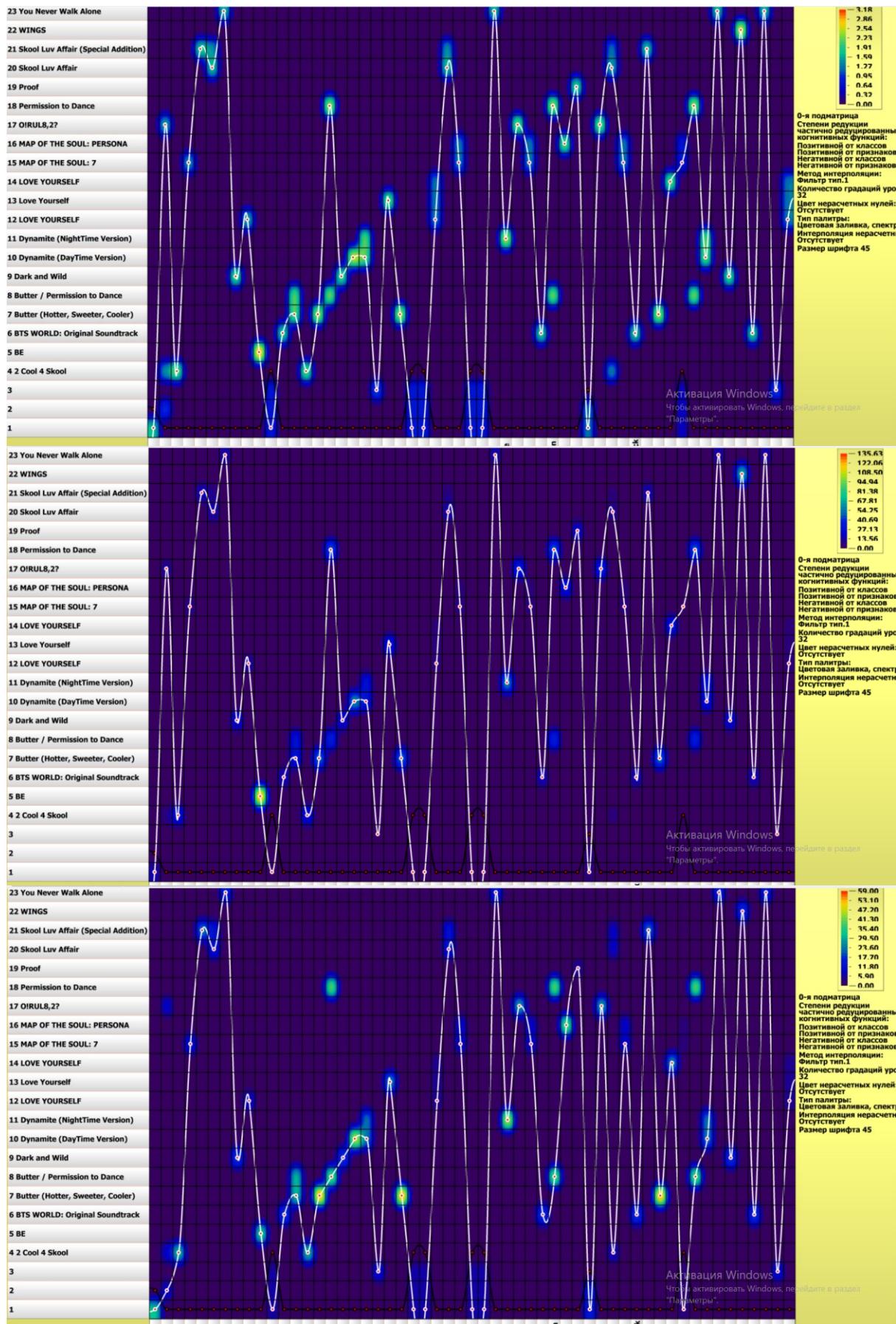
Когнитивные функции отображают, какое количество информации содержится в градациях описательной шкалы о переходе объекта моделирования в состояния, соответствующие градациям классификационной шкалы. При этом в статистических и системно-когнитивных моделях в каждой градации описательной шкалы содержится 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 фактора на переход объекта моделирования в каждое из будущих состояний.

В системе «Эйдос» когнитивные функции отображаются в режиме 4.5 (рисунки 38). 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 [21]. Содержательное объяснение этих эмпирических закономерностей формулируется уже экспертами на теоретическом уровне познания в содержательных научных законах [22].





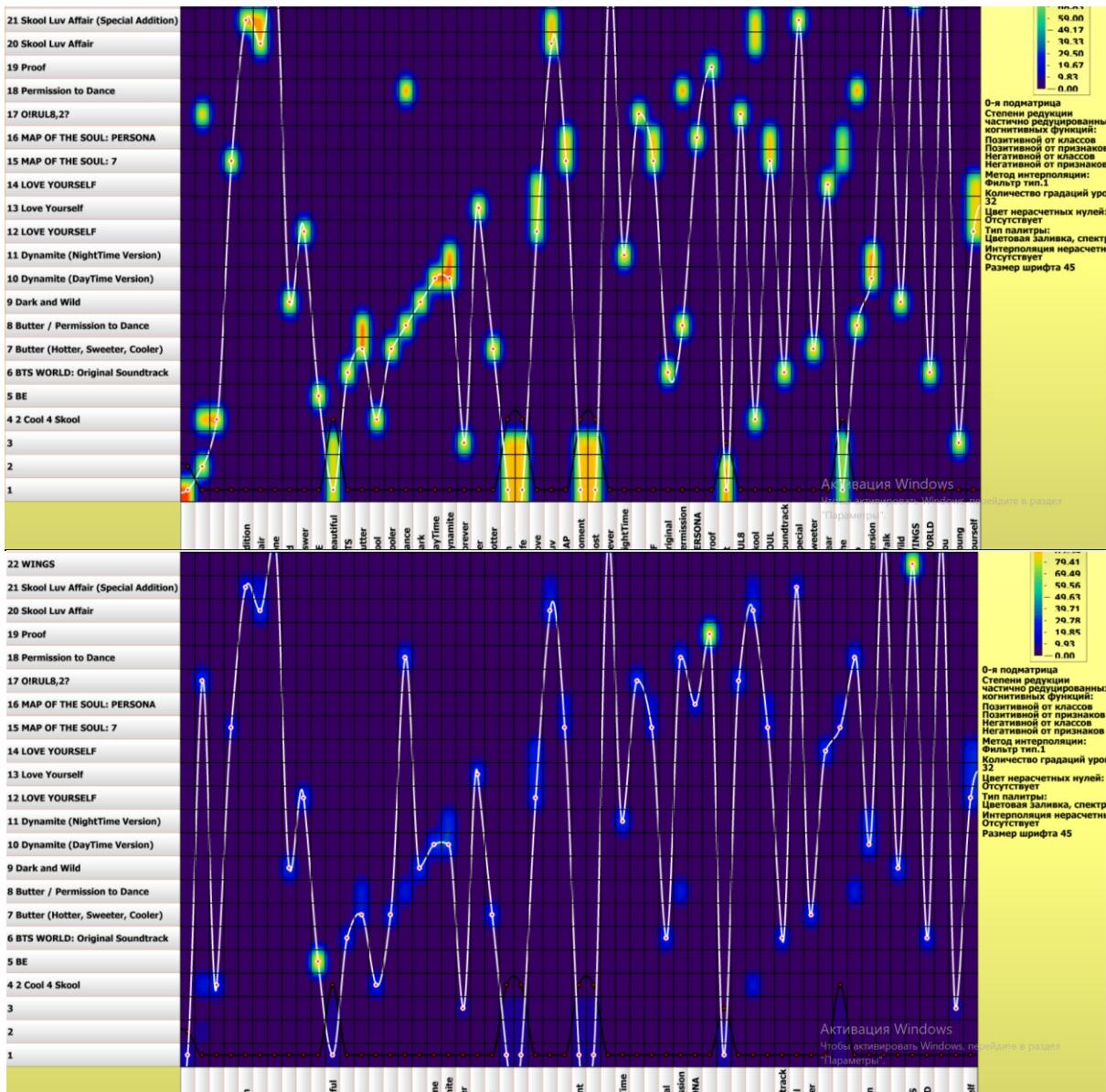


Рисунок 37. Примеры когнитивных функций в СК-модели INF1

Как уже отмечалось содержательное объяснение когнитивных функций на теоретическом уровне познания – это дело специалистов в той предметной области, к которой относится предмет моделирования [22].

3.8.10. Значимость описательных шкал и их градаций

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»).

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.

На рисунке 39 приведена Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF3:

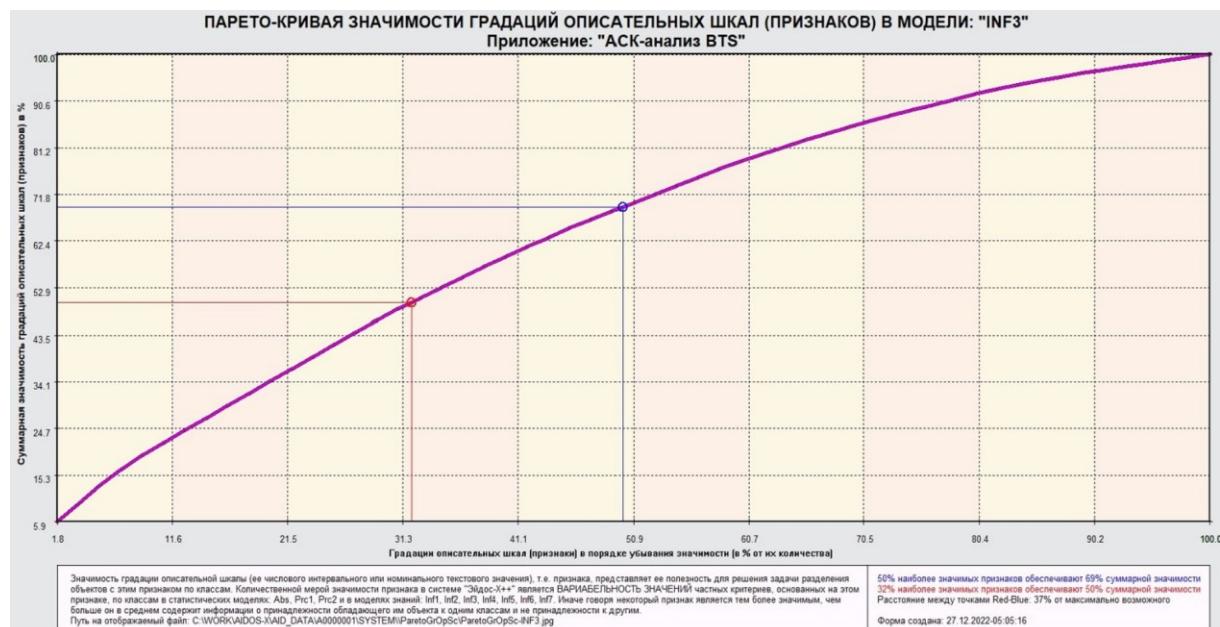


Рисунок 38. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF3

Из рисунка 39 видно, что примерно пятая часть наиболее ценных значений факторов обеспечивает половину суммарного влияния всех значений факторов, а половина наиболее ценных значений факторов обеспечивает 85% суммарного влияния.

На экранной форме рисунка 40 приведены имена Excel-файлов с информацией о силе и направлении влияния значений факторов в разных моделях:

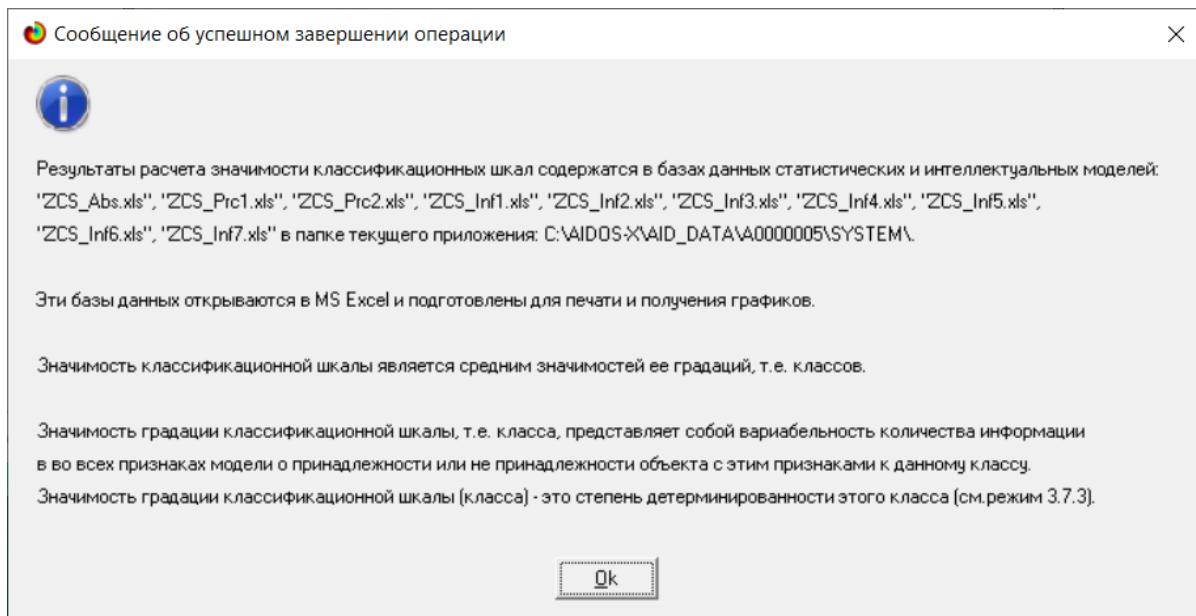


Рисунок 39. Имена Excel-файлов с информацией о силе влияния значений факторов в разных моделях

3.8.11. Степень детерминированности классов и классификационных шкал

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»).

На рисунках 42 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:

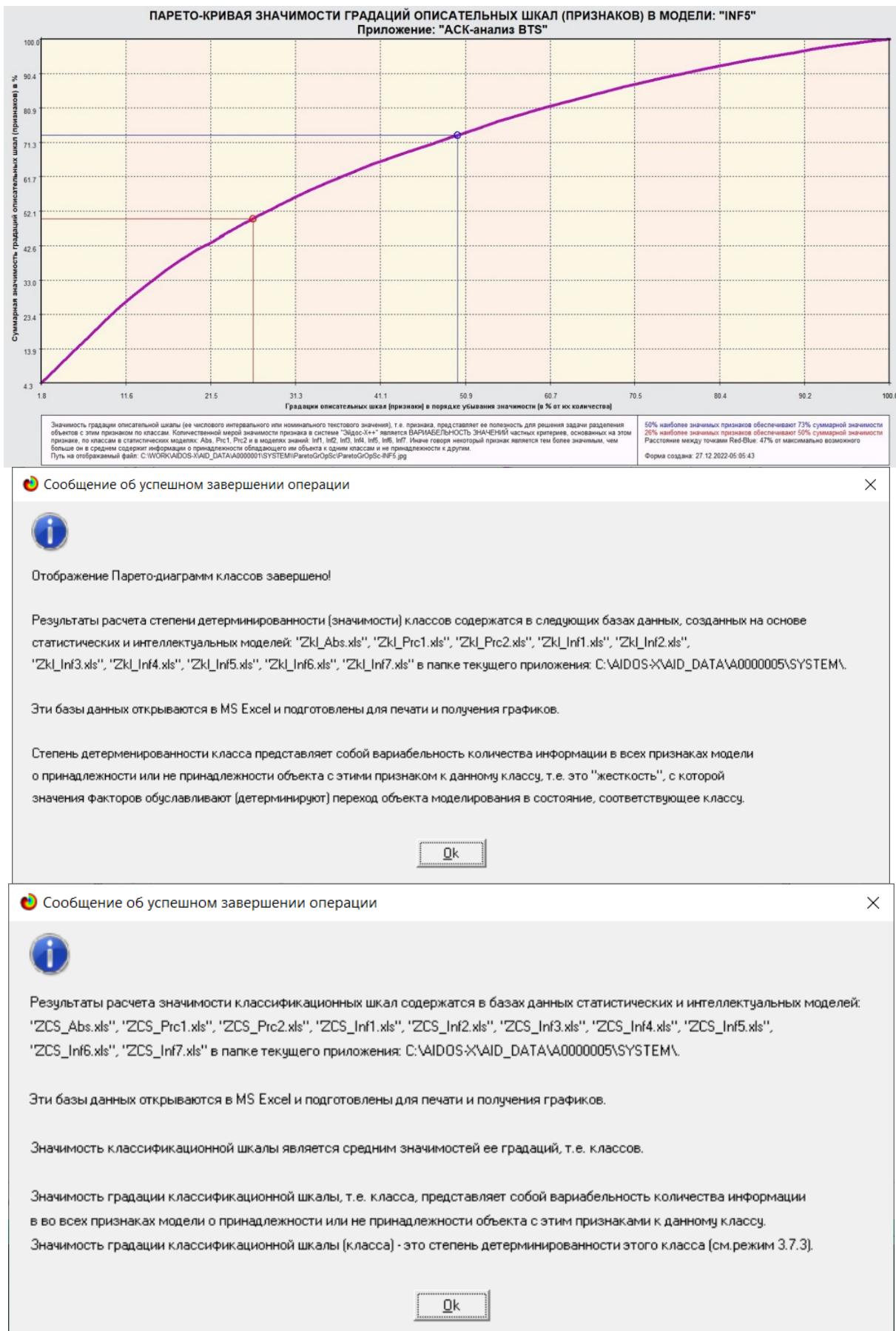


Рисунок 40. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»

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 лингвистического Автоматизированного системно-когнитивного анализа (лингвистический АСК-анализ) и его программного инструментария – интеллектуальной системы «Эйдос».

Анализ полученных результатов, проведенный в данной работе, полностью согласуется с результатами работы [10], 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 [10]. Поэтому есть все основания рекомендовать применение АСК-анализа и системы «Эйдос» для проведения дальнейших углубленных исследований.

Достижением данной работы является:

1. Возможность построения системно-когнитивных моделей предметной области на основе исходных данных, содержащих лингвистические переменные.
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, не только технологические, но субъективные 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 [1-26].

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 №335. 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 решена задача выявления зависимости времени выпуска, количества, качества и смысла песен и их текстов. На основе знания этих зависимостей наглядно показаны актуальные темы для песен, их названия и авторы, моделируемые предметной области путем исследования ее системно-когнитивной модели.

Спецификой данной задачи является то, что все переменные являются лингвистическими (категориальными) переменными. Поэтому для решения данной задачи применяется лингвистический АСК-анализ, т.е. когнитивная математическая лингвистика.

Таким образом, в работе строится модель, включающая больше номинальные (текстовые), а не числовые шкалы. Сопоставимость обработки данных одного типа. Это достигается путем вычисления количества информации, содержащегося в градациях текстовых шкал и получении той или иной популярности.

В работе приводится краткое описание АСК-анализа и его программного инструментария – интеллектуальной системы «Эйдос». Работа может быть основой для лабораторных работ и научных исследований по применению систем искусственного интеллекта, в частности лингвистического АСК-анализа для решения задач в области *когнитивного анализа текстов*.

REFERENCES (ЛИТЕРАТУРА)

- 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 OCZFHC.
- 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 RZXZ.
- Работы проф.Е.В.Луценко по АСК-анализу текстов, т.е. по когнитивной математической лингвистике: http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_texts.htm

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. 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 // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.

6. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.

7. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>

8. 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 // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.

9. 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 ZQLTW.

10. 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.

11. Луценко, Е. В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы "Эйдос-X++" / Е. В. Луценко // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2014. – № 101. – С. 1367-1409. – EDN SZVWRV.

12. 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 // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 131. – С. 1-18. – DOI 10.21515/1990-4665-131-001. – EDN ZRXVFN.

13. Луценко, Е. В. Системное обобщение принципа Эшби и повышение уровня системности модели объекта познания как необходимое условие адекватности процесса его познания / Е. В. Луценко // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2020. – № 163. – С. 100-134. – DOI 10.21515/1990-4665-163-009. – EDN SWKGWY.

14. Луценко, Е. В. Эффективность объекта управления как его эмерджентное свойство и повышение уровня системности как цель управления / Е. В. Луценко // Политехнический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2021. – № 165. – С. 77-98. – DOI 10.13140/RG.2.2.11887.25761. – EDN UMTAMT.

15. 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 // Политехнический сетевой

электронный научный журнал Кубанского государственного аграрного университета. – 2011. – № 71. – С. 27-74. – EDN OIGYBB.

16. 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.

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. Работы проф.Е.В.Луценко & С° 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

19. 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>

20. 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.

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

22. 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.