

09.03.02 Информационные системы и технологии

## **СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ СРЕДНЕГО ВОЗРАСТА ПРИ ПЕРВОМ БРАКЕ В АМЕРИКЕ**

Пахтусов Егор Петрович

Студент прикладной информатики, Информационные системы и технологии. Группа ИТ2204  
*Кубанский Государственный Аграрный университет имени И.Т.Трубилина, Краснодар, Россия*

Данная работа является продолжением серии работ автора по применению Автоматизированного системно-когнитивного анализа (ACK-анализ) для решения широкого спектра задач в области анализа различных данных. В работе решается задача выявления зависимости первого брачного возраста от места проживания в США. 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. Это достигается путем вычисления количества информации, содержащегося в градациях номинальных шкал и получении той или иной урожайности. 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 определении первого брачного возраста.

Ключевые слова: ЛИНГВИСТИЧЕСКИЙ ACK-АНАЛИЗ, ЛИНГВИСТИЧЕСКИЙ АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, СРЕДНИЙ ВОЗРАСТ ПЕРВОГО БРАКА, МЕСТО ПРОЖИВАНИЯ

<https://www.kaggle.com/datasets/thedevastator/median-age-at-first-marriage-in-america>

09.03.02 Information systems and technologies

## **SYSTEMIC COGNITIVE ANALYSIS OF AVERAGE AGE AT FIRST MARRIAGE IN AMERICAN**

Pakhtusov Egor Petrovich

Student of the Faculty of Applied Informatics, Information Systems direction. Group IT2204  
*Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia*

This work is a continuation of the author's series of works on the use of Automated System Cognitive Analysis (ASC-analysis) for solving a wide range of problems in the field of various data analysis. The paper solves the problem of identifying the dependence of the first marriageable age on the place of residence in the United States. 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. This is achieved by calculating the amount of information contained in the gradations of nominal scales and obtaining one or another yield. 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 of determining the first marriageable age.

Keywords: LINGUISTIC ASC ANALYSIS, AUTOMATED LINGUISTIC COGNITIVE ANALYSIS, AVERAGE AGE OF FIRST MARRIAGE, PLACE OF RESIDENCE

## СОДЕРЖАНИЕ

<b>1. INTRODUCTION(ВВЕДЕНИЕ) .....</b>	<b>3</b>
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 .....	4
<b>2. METHODS (МЕТОДЫ).....</b>	<b>4</b>
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.....	5
2.4. СИСТЕМА «ЭЙДОС» - ИНСТРУМЕНТАРИЙ АСК-АНАЛИЗА .....	6
<b>3. RESULTS (РЕЗУЛЬТАТЫ).....</b>	<b>13</b>
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.....	24
3.4. ЗАДАЧА-4. ВЕРИФИКАЦИЯ МОДЕЛЕЙ .....	33
3.5. ЗАДАЧА-5. ВЫБОР НАИБОЛЕЕ ДОСТОВЕРНОЙ МОДЕЛИ .....	37
3.6. ЗАДАЧА-6. СИСТЕМНАЯ ИДЕНТИФИКАЦИЯ И ПРОГНОЗИРОВАНИЕ .....	39
3.6.1. Интегральный критерий «Сумма знаний» .....	39
3.6.2. Интегральный критерий «Семантический резонанс знаний».....	40
3.6.3. Важные математические свойства интегральных критериев.....	41
3.6.4. Решение задачи идентификации и прогнозирования в системе «Эйдос».....	42
3.7.1. Упрощенный вариант принятия решений как обратная задача прогнозирования, позитивный и негативный информационные портреты классов, SWOT-анализ.....	45
3.7.2. Развитый алгоритм принятия решений в адаптивных интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос» .....	47
3.8. ЗАДАЧА-8. ИССЛЕДОВАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ ПУТЕМ ИССЛЕДОВАНИЯ ЕГО МОДЕЛИ ....	51
3.8.1. Инвертированные SWOT-диаграммы значений описательных шкал (семантические потенциалы) .....	51
3.8.2. Кластерно-конструктивный анализ классов .....	54
3.8.3. Кластерно-конструктивный анализ значений описательных шкал .....	57
3.8.4. Модель знаний системы «Эйдос» и нелокальные .....	61
нейроны .....	61
3.8.6. 3D-интегральные когнитивные карты .....	65
3.8.7. 2D-интегральные когнитивные карты содержательного сравнения классов (опосредованные нечеткие правдоподобные рассуждения).....	66
3.8.8. 2D-интегральные когнитивные карты содержательного сравнения значений факторов (опосредованные нечеткие правдоподобные рассуждения).....	69
3.8.9. Когнитивные функции .....	71
3.8.9. Значимость описательных шкал и их градаций.....	77
<b>4. DISCUSSION (ОБСУЖДЕНИЕ) .....</b>	<b>82</b>
<b>5. CONCLUSIONS (ВЫВОДЫ) .....</b>	<b>83</b>
<b>REFERENCES (ЛИТЕРАТУРА) .....</b>	<b>84</b>

## **1.INTRODUCTION(ВВЕДЕНИЕ)**

### **1.1. Описание исследуемой предметной области**

Данная работа является исследованием различных наборов данных по брачному возрасту при первой свадьбе на основе Автоматизированного системно-когнитивного анализа (АСК-анализ).

На сегодняшний день создание систем искусственного интеллекта – это перспективный шаг к расширению возможностей компьютерных наук и автоматизации разумного поведения, опирающийся на теоретические и прикладные принципы. Однако возникает необходимость оценки качества математических моделей этих систем. В данной курсовой работе рассмотрено решение, задачи, прогнозирования, дисциплины, района жилья и состава семьи.

Целью работы является разработка системно-когнитивной модели, основывающейся на следующих признаках: плотность населения, возраст супругов, тип района, индекс штата.

Задачами работы являются: 1) систематизация, закрепление и расширение теоретических и практических знаний по дисциплине "Интеллектуальные информационные системы и технологии"; 2) изучение интеллектуальной информационной системы "Эйдос"; 3) решение поставленной цели с помощью интеллектуальной информационной системы "Эйдос". Объектом исследования является база данных "Kaggle". Для решения задачи используем стандартные возможности Microsoft Office Word и Excel, блокнот, а также систему искусственного интеллекта "Эйдос-X++".

### **1.2. Объект и предмет исследования**

Объект исследования – выявление зависимостей показателей от места проживания

Предмет исследования – выявление зависимостей показателей возраста от плотности населения определённого района

### **1.3. Проблема, решаемая в работе и ее актуальность**

Спецификой данной работы является то, что все её независимые переменные, такие как возраст супругов, индекс района, имеют очень низкую формализацию, тогда как результаты влияния этих факторов, то есть различные показатели, измеряются в числовых шкалах.

Таким образом, в работе решается **проблема** построения гибридной модели, включающей как номинальные (текстовые), так и числовые шкалы, и обеспечивающей **сопоставимость** обработки данных разных типов, представленных в разных типах шкал и разных единицах измерения, соотношении, содержания понятий «Проблема» и «Задача».

## **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 годов<sup>1</sup> 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 слов<sup>2</sup>.

**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 применением 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<sup>3</sup>](#) [6, 7, 8].

Три монографии включены в фонды библиотеки конгресса США<sup>4</sup>.

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, психологическими и медицинским наукам, еще несколько докторских и кандидатских диссертаций с применением АСК-анализа находятся в стадии выхода на защиту.

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

<sup>2</sup> [https://yandex.ru/search/?text=Автоматизированный%2Bсистемно-когнитивный%2Bанализ%2B\(ACK-анализ\)&l=35&clid=2327117-18&win=360](https://yandex.ru/search/?text=Автоматизированный%2Bсистемно-когнитивный%2Bанализ%2B(ACK-анализ)&l=35&clid=2327117-18&win=360)

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

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

Автор является основателем междисциплинарной научной школы: «Автоматизированный системно-когнитивный анализ»<sup>5</sup>. Научная школа: "Автоматизированный системно-когнитивный анализ" является междисциплинарным научным направлением на пересечении по крайней мере трех научных специальностей (согласно недавно утвержденной новой номенклатуры научных специальностей ВАК РФ<sup>6</sup>). Основные научные специальности, которым соответствует научная школа:

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

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

- Автоматизированный системно-когнитивный анализ числовых и текстовых табличных данных;
- 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](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 о получении той или иной урожайности [5]. Для работы с лингвистическими переменными применяется лингвистический АСК-анализ [4].

## **2.4. Система «Эйдос» - инструментарий АСК-анализа**

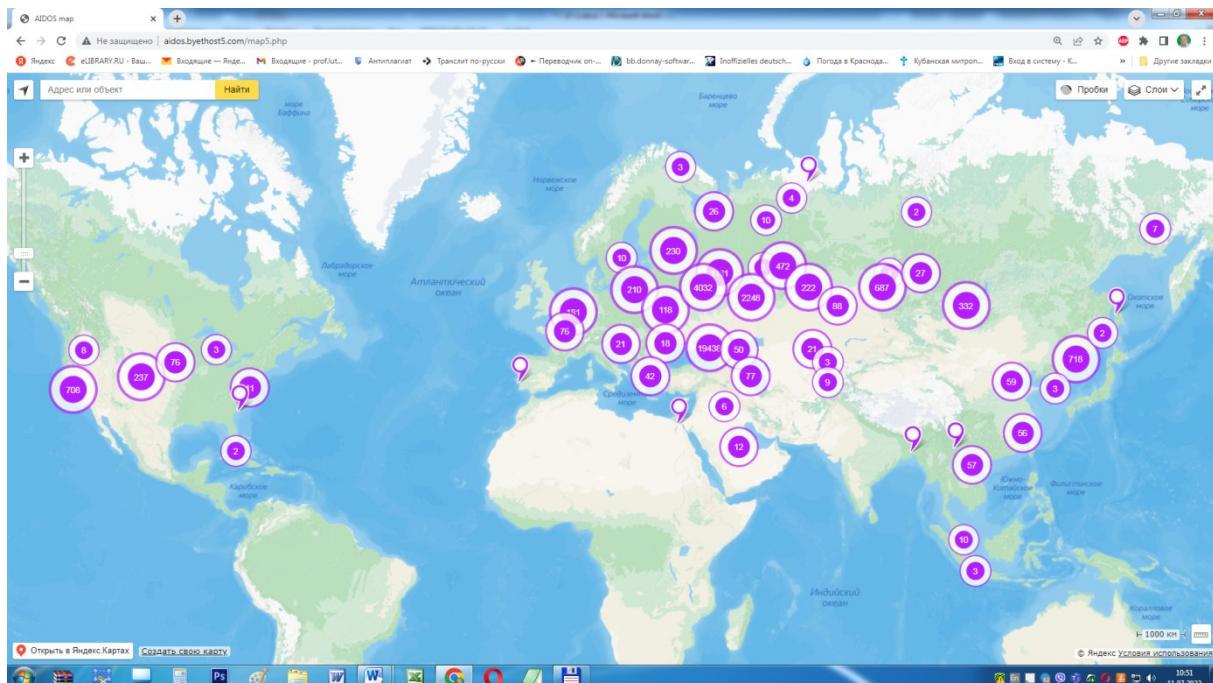
Существует много систем искусственного интеллекта. Универсальная когнитивная аналитическая система «Эйдос» отличается от них следующими параметрами:

---

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

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

- 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](http://lc.kubagro.ru/aidos/_Aidos-X.htm)), причем с актуальными исходными текстами ([http://lc.kubagro.ru/\\_AidosALL.txt](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://aidos.byethost5.com/Source_data_applications/WebAppls.htm)) ([http://lc.kubagro.ru/aidos/Presentation\\_Aidos-online.pdf](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](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 законах<sup>7</sup>.

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,

---

<sup>7</sup> Ссылка на это краткое описание системы «Эйдос» на английском языке:  
[http://lc.kubagro.ru/aidos/The\\_Eidos\\_en.htm](http://lc.kubagro.ru/aidos/The_Eidos_en.htm)

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 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»

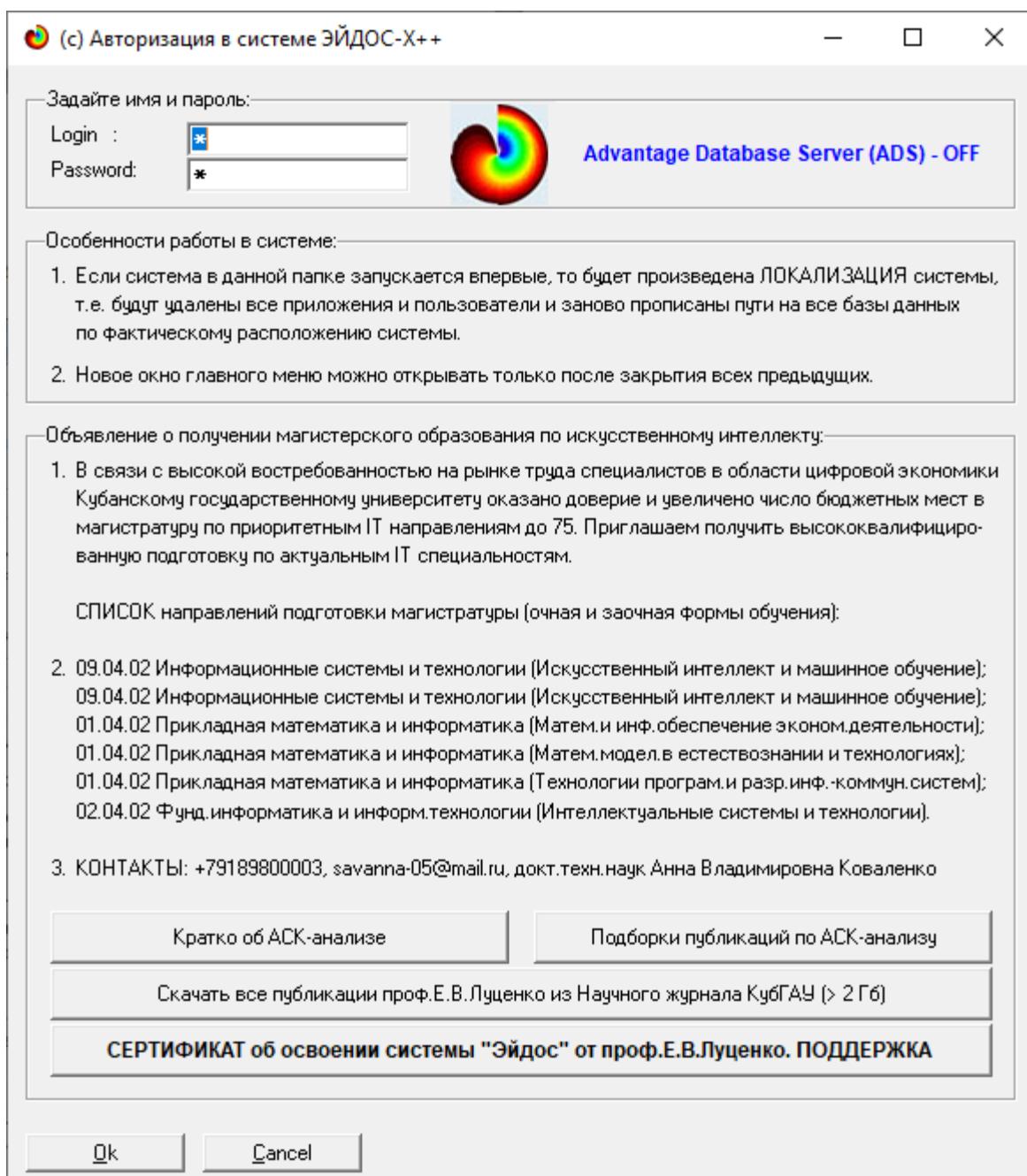
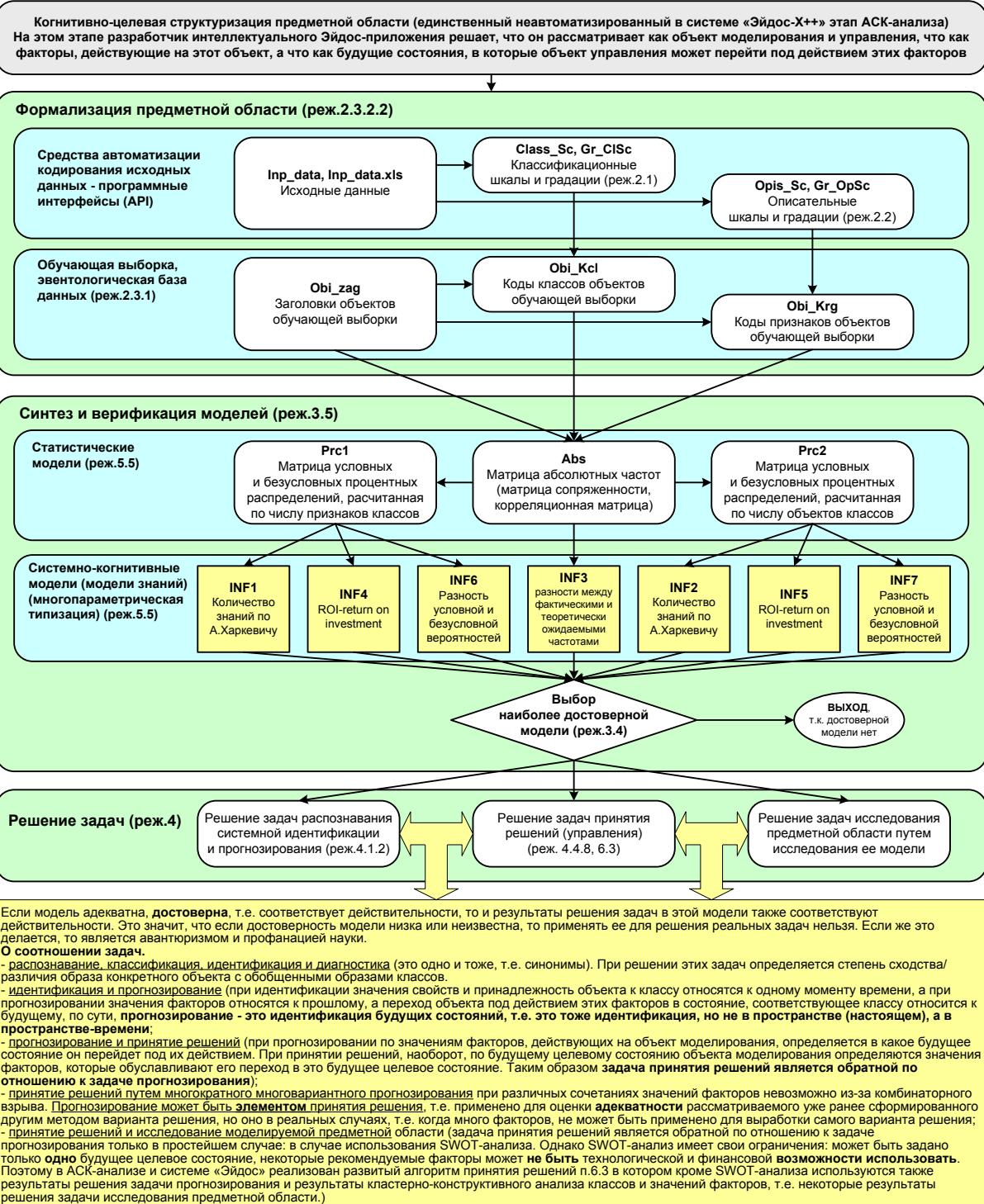


Рисунок 2. Титульные видеограммы текущей версии системы «Эйдос»

**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** выступает место проживания, в качестве **факторов** название штата, тип, временное ограничение, средний возраст, плотность населения штатов, а в качестве **результатов** обработанная информация, на основе вышеперечисленных факторов.

**Таблица 1 – Описательные шкалы (факторы)**

2.2. Описательные шкалы и градации. Текущая модель: "INF1"	
Код шкалы	Наименование описательной шкалы
1	NAME
2	TYPE
3	TIMEFRAME
4	POPULATION DENSITY PER SQ MI

Код градации	Наименование градации описательной шкалы
1	Alabama
2	Alaska
3	Arizona
4	Arkansas
5	California
6	Carolina
7	Colorado
8	Columbia
9	Connecticut
10	Dakota
11	Delaware
12	District
13	Florida
14	Georgia
15	Hampshire
16	Hawaii
17	Idaho
18	Illinois
19	Indiana
20	Iowa
21	Island
22	Jersey
23	Kansas
24	Kentucky
25	Louisiana
26	Maine
27	Maryland

**Таблица 2 – Классификационные шкалы (результаты действия факторов)**

2.1. Классификационные шкалы и градации. Текущая модель: "INF1"	
Код шкалы	Наименование классификационной шкалы
1	MEDIAN AGE

Код градации	Наименование градации классификационной шкалы
1	1/3-(23.000000, 25.533333)
2	2/3-(25.533333, 28.066667)
3	3/3-(28.066667, 30.600000)

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 (ЭЦГ) 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 [10] (см. таблицы 3 и 4):

Таблица 3 имеет следующую структуру:

- каждая строка описывает одно наблюдение, всего 40 наблюдений;
- каждое **наблюдение** описывается одновременно двумя способами: с одной стороны значениями факторов, действующих на объект моделирования (лингвистические переменные, градации описательных шкал, бесцветный фон), а с другой стороны результатами действия этих факторов, т.е. фактическими показателями такими как, название штата, плотность населения штата и тд. Такая структура описания наблюдений в технологиях искусственного интеллекта называется «онтологией» и модели представлений знаний Марвина Мински (1975) называется «фрейм-экземпляр»;
- 1-я колонка – номер наблюдения (не является шкалой);
- колонки со 2-й по 7-ю – это классификационные шкалы – это шкалы **числового** типа, описывающие *результаты* действия факторов, в данном случае показатели выраженные в словах и числах: название штата, тип, возраст и тд. В общем случае в исходных данных может быть значительно больше классификационных шкал, описывающих результаты действия факторов на объект моделирования в натуральном и стоимостном выражении. В системе «Эйдос» существует не очень жестко ограничение на суммарное количество градаций классификационных шкал: их должно быть не более 2032;

**Таблица 3 – Таблица исходных данных в стандарте системы «Эйдос»**

index	Median Age	Name	Type	TimeFrame	Population Density per Sq Mi
0	25,5	Alabama	State	2006-2010	94,4
1	24,8	Alaska	State	2006-2010	1,2
2	25,6	Arizona	State	2006-2010	56,3
3	24,5	Arkansas	State	2006-2010	56
4	26,9	California	State	2006-2010	239,1
5	25,7	Colorado	State	2006-2010	48,5
6	27,6	Connecticut	State	2006-2010	738,1
7	26,4	Delaware	State	2006-2010	460,8
8	29,6	District of Columbia	State	2006-2010	9856,5
9	26,6	Florida	State	2006-2010	350,6
10	25,9	Georgia	State	2006-2010	168,4
11	27	Hawaii	State	2006-2010	211,8
12	23,3	Idaho	State	2006-2010	19
13	27,1	Illinois	State	2006-2010	231,1
14	25,9	Indiana	State	2006-2010	181
15	25,5	Iowa	State	2006-2010	54,5
16	25,1	Kansas	State	2006-2010	34,9
17	25	Kentucky	State	2006-2010	109,9
18	26,2	Louisiana	State	2006-2010	104,9
19	26,5	Maine	State	2006-2010	43,1
20	27,4	Maryland	State	2006-2010	594,8
21	28,5	Massachusetts	State	2006-2010	839,4
22	26,6	Michigan	State	2006-2010	174,8
23	26,3	Minnesota	State	2006-2010	66,6
24	25,9	Mississippi	State	2006-2010	63,2
25	25,8	Missouri	State	2006-2010	87,1
26	25,4	Montana	State	2006-2010	6,8
27	25,5	Nebraska	State	2006-2010	23,8
28	25,7	Nevada	State	2006-2010	24,6
29	26,8	New Hampshire	State	2006-2010	147
30	27,8	New Jersey	State	2006-2010	1195,5
31	25,9	New Mexico	State	2006-2010	17
32	28,3	New York	State	2006-2010	411,2
33	25,9	North Carolina	State	2006-2010	196,1
34	25,7	North Dakota	State	2006-2010	9,7
35	26,3	Ohio	State	2006-2010	282,3
36	24,5	Oklahoma	State	2006-2010	54,7
37	26,1	Oregon	State	2006-2010	39,9
38	27,2	Pennsylvania	State	2006-2010	283,9
39	27,9	Rhode Island	State	2006-2010	1018,1
40	26,4	South Carolina	State	2006-2010	153,9
41	25,5	South Dakota	State	2006-2010	10,7
42	25,4	Tennessee	State	2006-2010	153,9
43	25,3	Texas	State	2006-2010	96,3
44	23	Utah	State	2006-2010	33,6
45	27	Vermont	State	2006-2010	67,9
46	26,5	Virginia	State	2006-2010	202,6
47	25,8	Washington	State	2006-2010	101,2
48	25,4	West Virginia	State	2006-2010	77,1
49	26,4	Wisconsin	State	2006-2010	105
50	24,2	Wyoming	State	2006-2010	5,8
51	26,3	United States	Country	2006-2010	87,4
52	26,7	Alabama	State	2015-2019	99,2
53	26,4	Alaska	State	2015-2019	1,3
54	27,8	Arizona	State	2015-2019	63
55	25,7	Arkansas	State	2015-2019	57,9

56	29	California	State	2015-2019	254
57	27,7	Colorado	State	2015-2019	54,3
58	29,5	Connecticut	State	2015-2019	745
59	29	Delaware	State	2015-2019	508
60	30,6	District of Columbia	State	2015-2019	11295
61	28,9	Florida	State	2015-2019	402
62	27,7	Georgia	State	2015-2019	186
63	28,4	Hawaii	State	2015-2019	227
64	25,4	Idaho	State	2015-2019	22,3
65	28,9	Illinois	State	2015-2019	231
66	27	Indiana	State	2015-2019	189
67	26,6	Iowa	State	2015-2019	57,1
68	26,1	Kansas	State	2015-2019	35,9
69	26,4	Kentucky	State	2015-2019	114
70	27,9	Louisiana	State	2015-2019	108
71	27,8	Maine	State	2015-2019	44,2
72	29,1	Maryland	State	2015-2019	636
73	29,9	Massachusetts	State	2015-2019	901
74	28,1	Michigan	State	2015-2019	178
75	27,8	Minnesota	State	2015-2019	71,7
76	26,9	Mississippi	State	2015-2019	63,1
77	26,9	Missouri	State	2015-2019	89,5
78	26,4	Montana	State	2015-2019	7,5
79	26,4	Nebraska	State	2015-2019	25,5
80	27,8	Nevada	State	2015-2019	28,3
81	28,7	New Hampshire	State	2015-2019	154
82	29,3	New Jersey	State	2015-2019	1263
83	27,7	New Mexico	State	2015-2019	17,5
84	29,7	New York	State	2015-2019	429
85	27,6	North Carolina	State	2015-2019	215
86	26	North Dakota	State	2015-2019	11,3
87	27,9	Ohio	State	2015-2019	289
88	25,9	Oklahoma	State	2015-2019	57,7
89	27,9	Oregon	State	2015-2019	44,1
90	28,7	Pennsylvania	State	2015-2019	291
91	29,7	Rhode Island	State	2015-2019	1061
92	28	South Carolina	State	2015-2019	170
93	26,4	South Dakota	State	2015-2019	11,7
94	26,9	Tennessee	State	2015-2019	168
95	27	Texas	State	2015-2019	112
96	24,6	Utah	State	2015-2019	39,8
97	28,4	Vermont	State	2015-2019	69,8
98	27,8	Virginia	State	2015-2019	219
99	27,3	Washington	State	2015-2019	116
100	26,7	West Virginia	State	2015-2019	74,6
101	27,8	Wisconsin	State	2015-2019	109
102	25,7	Wyoming	State	2015-2019	5,9
103	28	United States	Country	2015-2019	94,8

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

единицах измерения. Таким образом, система «Эйдос» не предъявляет жестких практически невыполнимых требований к исходным данным, а обрабатывает те данные, которые реально есть, например подобные представленным в таблице 5.

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

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

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

Для классификационных шкал на 3-м рисунке 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(x)" в систему "Эйдос-X++" 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 допускается. Эти наименования должны быть короткими, но понятными, т.к. они будут в выхodных формах, а к ним еще будут добавляться наименования градаций. В числовых шкалах надо ОБЯЗАТЕЛЬНО указывать единицы измерения 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-го, 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).
- Столбец присваивается числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. проблем), то столбец присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
- Столбцы со 2-го по N-й являются классификационными шкалами (выходные параметры) и содержат данные о классах (будущих состояниях объекта управления), к которым принадлежат объекты обучающей выборки.
- Столбцы с N+1 по последний являются описательными шкалами (свойстваами 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 число или текстовое значение считается градацией классификационной или описательной шкалы, характеризующей объект. В каждой шкале ее градации сортируются по алфавиту. С использованием шкал и градаций сортируются 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 значений классов с градациями классификационных шкал и коды признаков, соответствующие фактам совпадения числовых или уникальных текстовых значений признаков с градациями описательных шкал.
- Распознаваемая выборка формируется на основе файла INP\_RASP аналогично, за исключением того, что классификационные и описательные шкалы и градации не создаются, а используются ранее созданные в модели, и базы распознаваемой выборки могут не включать коды классов, если столбцы классов в файле INP\_RASP были пустыми. Структура файла INP\_RASP должна быть 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

**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(x)" в систему "Эйдос-X++". ТЕРИНЫ АСК-АНАЛИЗА И СИСТЕМЫ "ЭЙДОС":**

Шкала представляет собой способ формализации предметной области. Используются числовые и текстовые шкалы, при этом текстовые могут быть номинальными и порядковыми. На номинальных шкалах есть только отношения эквивалентности и нежективантности, на порядковых, кроме того еще отношения "больше", "меньше", а на числовых - кроме того могут выполняться все арифметические операции. Каждый объект выборки (наблюдение) описан с одной стороны своими признаками, а с другой -принадлежностью к некоторым обобщающим категориям (классам). Такая структура описания называется онтологией или фреймом экспземпляром и является базовой для всех моделей представления знаний.

В АСК-анализе и системе "Эйдос" используется три интерпретации шкал и градаций: универсальная, статическая и динамическая:

- в универсальной интерпретации: признаки - это градации описательных шкал;
- в статической интерпретации: описательная шкала - это свойство, а градация (признак) - это степень выраженности этого свойства;
- в динамической интерпретации: описательная шкала - это фактор, а градация (признак) - это значение фактора;
- в универсальной интерпретации: классификационная шкала - это градации классификационных шкал;
- в статической интерпретации: классификационная шкала - способ классификации обобщающих категорий (классов), к которым в настоящем времени по отношению к признакам относятся состояния объекта прогнозирования или управления;
- в динамической интерпретации: классификационная шкала - способ классификации обобщающих категорий (классов), к которым в будущем времени по отношению к признакам относятся состояния объекта прогнозирования или управления;

ПРОФИЛАКТИКА ОШИБОК В ФАЙЛЕ ИСХОДНЫХ ДАННЫХ:

- 1-я строка файла "Inp\_data.xls(x)" должна содержать наименование колонок. Эти наименования должны быть во всех колонках, при этом переносы по словам разрешены, а объединение ячеек, разрыв строки знак абзаца и неавтфавитные символы не допускаются. Эти наименования должны быть короткими, но понятными, т.к. они будут в выхodных формах, а к ним еще будут добавляться наименования градаций. В числовых шкалах надо обязательно указывать единицы измерения. Число знаков после запятой в числовой колонке должно быть одинаковым.
- 1-я колонка содержит наименование объекта обучающей выборки или наименование наблюдения. Оно может быть длинным до 255 символов.
- Столбцы, начиная со 2-го, являются классификационными и описательными шкалами и могут быть текстового (номинального / порядкового) или числового типа (со знаками после запятой). Чтобы текстовая шкала была порядковой, нужно чтобы при сортировке по алфавиту градации этой шкалы образовывали осмысленную последовательность от минимального значения до максимального. Например, текстовая шкала "Размер" с градациями: "очень малое", "малое", "среднее", "большое", "очень большое", будет номинальной шкалой, т.к. при сортировке по алфавиту они расположятся в порядке: "большое", "малое", "очень большое", "очень малое", "среднее". Чтобы шкала "Размер" стала порядковой нужно в этом градации присвоить следующие значения: "1/5-очень малое", "2/5-малое", "3/5-среднее", "4/5-большое", "5/5-очень большое".
- Столбцу присваивается числовой тип, если все значения его ячеек числового типа. Если хотя бы одно значение является текстовым (не числом, в т.ч. проблем), то столбец присваивается текстовый тип. Это означает, что нули должны быть указаны нулями, а не пробелами.
- Если в системе "Эйдос" в режимах 2.1, 2.2 посмотреть на градации классификационных и описательных шкал, которые должны быть числовыми, то сразу будет видно, в какой форме представлены числа: числовыми диапазонами или прямо числами. Если числовыми диапазонами, значит в файле исходных данных в этом отношении все правильно, если же числами, то возможно в Excel-файле нужно заменить десятичные точки на запятые, а также найти и исправить нечисловые данные в числовых по смыслу колонках. Быстро найти их можно перейдя на последнюю строку файла исходных данных и зайдя расчет суммы колонки. В формуле будет видно с какой строкой идет расчет суммы. Если со 2-й, то значит все верно, иначе будет указана строка, в которой находится нечисловое значение.
- Система "Эйдос" работает с областью данных файла исходных данных, которую можно выделить блоком, поставив курсор в ячейку A1, нажав Ctrl+Home, а затем нажав клавиши Shift+Ctrl нажать End. Если этот блок выходит за пределы области таблицы, фактически занятой данными надо скопировать эту фактическую область данных в буфер обмена, создать новый лист и скопировать в него, а исходный лист удалить.
- Иногда бывает полезно сбросить все форматирование Excel-таблицы исходных данных. Это можно сделать в MS Excel. А можно скопировать таблицу в MS Word, а потом обратно в MS Excel.

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	...
...	...	...	...	...	...	...

Рисунок 5. 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
- XLSX - MS Excel-2007(2010)
- DBF - DBASE IV (DBF/NTX)
- CSV - CSV => DBF конвертер

Стандарт XLS-файла  
Стандарт DBF-файла  
Стандарт CSV-файла

Задайте параметры:

- Нули и пробелы считать ОТСУТСТВИЕМ данных
- Нули и пробелы считать ЗНАЧЕНИЯМИ данных
- Создавать БД средних по классам "Inp\_davr.dbf"?

Требования к файлу исходных данных

Задайте диапазон столбцов классификационных шкал:

Начальный столбец классификационных шкал:   
 Конечный столбец классификационных шкал:

Задайте диапазон столбцов описательных шкал:

Начальный столбец описательных шкал:   
 Конечный столбец описательных шкал:

Задайте режим:

- Формализации предметной области (на основе "Inp\_data")
- Генерации распознаваемой выборки (на основе "Inp\_rasp")

Задание параметров формирования сценариев или способа интерпретации текстовых полей "Inp\_data":

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

**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-(59873.000000, 178545.6666667)")
- Только наименования интервальных числовых значений (например: "Минимальное")
- И интервальные числовые значения, и их наименования (например: "Минимальное: 1/3-(59873.000000, 178545.6666667)")

**Ok** **Cancel**

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

**ЗАДАНИЕ В ДИАЛОГЕ РАЗМЕРНОСТИ МОДЕЛИ: (равные интервалы)**

Количество градаций классификационных и описательных шкал в модели, т.е.: [3 классов x 68 признаков]

Тип шкалы	Количество классификационных шкал	Количество градаций классификационных шкал	Среднее количество градаций на класс.шкалу	Количество описательных шкал	Количество градаций описательных шкал	Среднее количество градаций на опис.шкалу
Числовые	1	3	3,00	1	12	12,00
Текстовые	0	0	0,00	3	56	18,67
ВСЕГО:	1	3	3,00	4	68	17,00

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

В классификационных шкалах:  В описательных шкалах:

**Пересчитать шкалы и градации** **Выйти на создание модели**

**Таблица 4 – Классификационные шкалы и градации (числовые шкалы)**

2.1. Классификационные шкалы и градации. Текущая модель: "INF1"

Код шкалы	Наименование классификационной шкалы	Инф.	Код градации	Наименование градации классификационной шкалы	DEL	Инф.
1	MEDIAN AGE		1	1/3-[23.0000000, 25.5333333]		
			2	2/3-(25.5333333, 28.0666667)		
			3	3/3-[28.0666667, 30.6000000]		

Помощь Доб.шкалу Доб.град.шкалы Копир.шкалу Копир.град.шкалы Копир.шкалу с град. Удал.шкалу с град. Удал.град.шкалы Удаление и перекодирование Графики будущих сценариев

Источник: c:\Aidos-X\AID\_DATA\A0000001\System\Classes.dbf

**Таблица 5 – Описательные шкалы и градации (лингвистические переменные)**

KOD_ATR	NAME_ATR
1	NAME-Alabama
2	NAME-Alaska
3	NAME-Arizona
4	NAME-Arkansas
5	NAME-California
6	NAME-Carolina
7	NAME-Colorado
8	NAME-Columbia
9	NAME-Connecticut
10	NAME-Dakota
11	NAME-Delaware
12	NAME-District
13	NAME-Florida
14	NAME-Georgia
15	NAME-Hampshire
16	NAME-Hawaii
17	NAME-Idaho
18	NAME-Illinois
19	NAME-Indiana
20	NAME-Iowa
21	NAME-Island
22	NAME-Jersey
23	NAME-Kansas
24	NAME-Kentucky
25	NAME-Louisiana
26	NAME-Maine
27	NAME-Maryland
28	NAME-Massachusetts
29	NAME-Mexico
30	NAME-Michigan
31	NAME-Minnesota
32	NAME-Mississippi
33	NAME-Missouri

34	NAME-Montana
35	NAME-Nebraska
36	NAME-Nevada
37	NAME-New
38	NAME-North
39	NAME-of
40	NAME-Ohio
41	NAME-Oklahoma
42	NAME-Oregon
43	NAME-Pennsylvania
44	NAME-Rhode
45	NAME-South
46	NAME-States
47	NAME-Tennessee
48	NAME-Texas
49	NAME-United
50	NAME-Utah
51	NAME-Vermont
52	NAME-Virginia
53	NAME-Washington
54	NAME-West
55	NAME-Wisconsin
56	NAME-Wyoming
57	NAME-York
58	TYPE-Country
59	TYPE-State
60	TIMEFRAME-2006
61	TIMEFRAME-2010
62	TIMEFRAME-2015
63	TIMEFRAME-2019
64	POPULATION DENSITY PER SQ MI-1/12-{1.2000000, 942.3500000}
65	POPULATION DENSITY PER SQ MI-2/12-{942.3500000, 1883.5000000}
66	POPULATION DENSITY PER SQ MI-3/12-{1883.5000000, 2824.6500000}
67	POPULATION DENSITY PER SQ MI-4/12-{2824.6500000, 3765.8000000}
68	POPULATION DENSITY PER SQ MI-5/12-{3765.8000000, 4706.9500000}
69	POPULATION DENSITY PER SQ MI-6/12-{4706.9500000, 5648.1000000}
70	POPULATION DENSITY PER SQ MI-7/12-{5648.1000000, 6589.2500000}
71	POPULATION DENSITY PER SQ MI-8/12-{6589.2500000, 7530.4000000}
72	POPULATION DENSITY PER SQ MI-9/12-{7530.4000000, 8471.5500000}
73	POPULATION DENSITY PER SQ MI-10/12-{8471.5500000, 9412.7000000}
74	POPULATION DENSITY PER SQ MI-11/12-{9412.7000000, 10353.8500000}
75	POPULATION DENSITY PER SQ MI-12/12-{10353.8500000, 11295.0000000}

Источник: c:\Aidos-X\AID\_DATA\A0000001\System\Attributes.dbf

**Таблица 6 – Обучающая выборка (полностью)**

NAME_OBJ	N2	N3	N4	N5	N6
0	1	1	59	61	64
1	1	2	59	61	64
2	2	3	59	61	64
3	1	4	59	61	64
4	2	5	59	61	64
5	2	7	59	61	64
6	2	9	59	61	64
7	2	11	59	61	64
8	3	8	59	61	74
9	2	13	59	61	64
10	2	14	59	61	64
11	2	16	59	61	64
12	1	17	59	61	64
13	2	18	59	61	64
14	2	19	59	61	64
15	1	20	59	61	64
16	1	23	59	61	64
17	1	24	59	61	64
18	2	25	59	61	64
19	2	26	59	61	64
20	2	27	59	61	64

21	3	28	59	61	64
22	2	30	59	61	64
23	2	31	59	61	64
24	2	32	59	61	64
25	2	33	59	61	64
26	1	34	59	61	64
27	1	35	59	61	64
28	2	36	59	61	64
29	2	15	59	61	64
30	2	22	59	61	65
31	2	29	59	61	64
32	3	57	59	61	64
33	2	6	59	61	64
34	2	10	59	61	64
35	2	40	59	61	64
36	1	41	59	61	64
37	2	42	59	61	64
38	2	43	59	61	64
39	2	21	59	61	65
40	2	6	59	61	64
41	1	10	59	61	64
42	1	47	59	61	64
43	1	48	59	61	64
44	1	50	59	61	64
45	2	51	59	61	64
46	2	52	59	61	64
47	2	53	59	61	64
48	1	52	59	61	64
49	2	55	59	61	64
50	1	56	59	61	64
51	2	46	58	61	64
52	2	1	59	63	64
53	2	2	59	63	64
54	2	3	59	63	64
55	2	4	59	63	64
56	3	5	59	63	64
57	2	7	59	63	64
58	3	9	59	63	64
59	3	11	59	63	64
60	3	8	59	63	75
61	3	13	59	63	64
62	2	14	59	63	64
63	3	16	59	63	64
64	1	17	59	63	64
65	3	18	59	63	64
66	2	19	59	63	64
67	2	20	59	63	64
68	2	23	59	63	64
69	2	24	59	63	64
70	2	25	59	63	64
71	2	26	59	63	64
72	3	27	59	63	64
73	3	28	59	63	64
74	3	30	59	63	64
75	2	31	59	63	64
76	2	32	59	63	64
77	2	33	59	63	64
78	2	34	59	63	64
79	2	35	59	63	64
80	2	36	59	63	64
81	3	15	59	63	64
82	3	22	59	63	65
83	2	29	59	63	64
84	3	57	59	63	64
85	2	6	59	63	64
86	2	10	59	63	64
87	2	40	59	63	64
88	2	41	59	63	64
89	2	42	59	63	64
90	3	43	59	63	64
91	3	21	59	63	65
92	2	6	59	63	64

93	2	10	59	63	64
94	2	47	59	63	64
95	2	48	59	63	64
96	1	50	59	63	64
97	3	51	59	63	64
98	2	52	59	63	64
99	2	53	59	63	64
100	2	52	59	63	64
101	2	55	59	63	64
102	2	56	59	63	64
103	2	46	58	63	64

*Источник:* c:\Aidos-X\AID\_DATA\A0000001\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).

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 9).

На ее основе рассчитываются матрицы условных и безусловных процентных распределений (таблица 10).

Таблица 7 – Матрица абсолютных частот (статистическая модель ABS)

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

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 обучающей выборки, относящихся к различным градациям одной классификационной или описательной шкалы. Поэтому решать задачу на основе непосредственно матрицы абсолютных частот (таблица 9) было бы очень неразумно и переход от абсолютных частот к условным и безусловным относительным частотам (частостям) (таблица 10) является весьма обоснованным и логичным.

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

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

		Классы					Безусловная вероятность признака
		1	...	j	...	w	
Значения факторов	1	$P_{11}$		$P_{1j}$		$P_{1W}$	
	...						
	i	$P_{i1}$		$P_{ij} = \frac{N_{ij}}{N_{\Sigma j}}$		$P_{iW}$	$P_{i\Sigma} = \frac{N_{i\Sigma}}{N_{\Sigma\Sigma}}$
	...						
	m	$P_{M1}$		$P_{Mj}$		$P_{MW}$	
Безусловная вероятность класса				$P_{\Sigma j}$			

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 7 и 8 с использованием *частных критериев, знаний* приведенных таблице 11, рассчитываются матрицы семи системно-когнитивных моделей (таблица 12).

В таблице 9 приведены формулы:

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

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

Количество частных критериев знаний и основанных на них системно-когнитивных моделей (таблица 11), 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.

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

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

2-й вариант расчета вероятностей: $N_j$ – суммарное количество объектов по $j$ -му классу		
---	--	--

**Обозначения к таблице:**

$i$  – значение прошлого параметра;

$j$  - значение будущего параметра;

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

$M$  – суммарное число значений всех прошлых параметров;

$W$  - суммарное число значений всех будущих параметров.

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

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

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

$I_{ij}$  – частный критерий знаний: количество знаний в факте наблюдения  $i$ -го значения прошлого параметра о том,

что объект перейдет в состояние, соответствующее  $j$ -му значению будущего параметра;

$\Psi$  – 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.

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).

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 3 – Матрица системно-когнитивной модели

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

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 10, 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.**

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 13):

**Таблица 4 – Уточнение терминологии АСК-анализа**

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

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

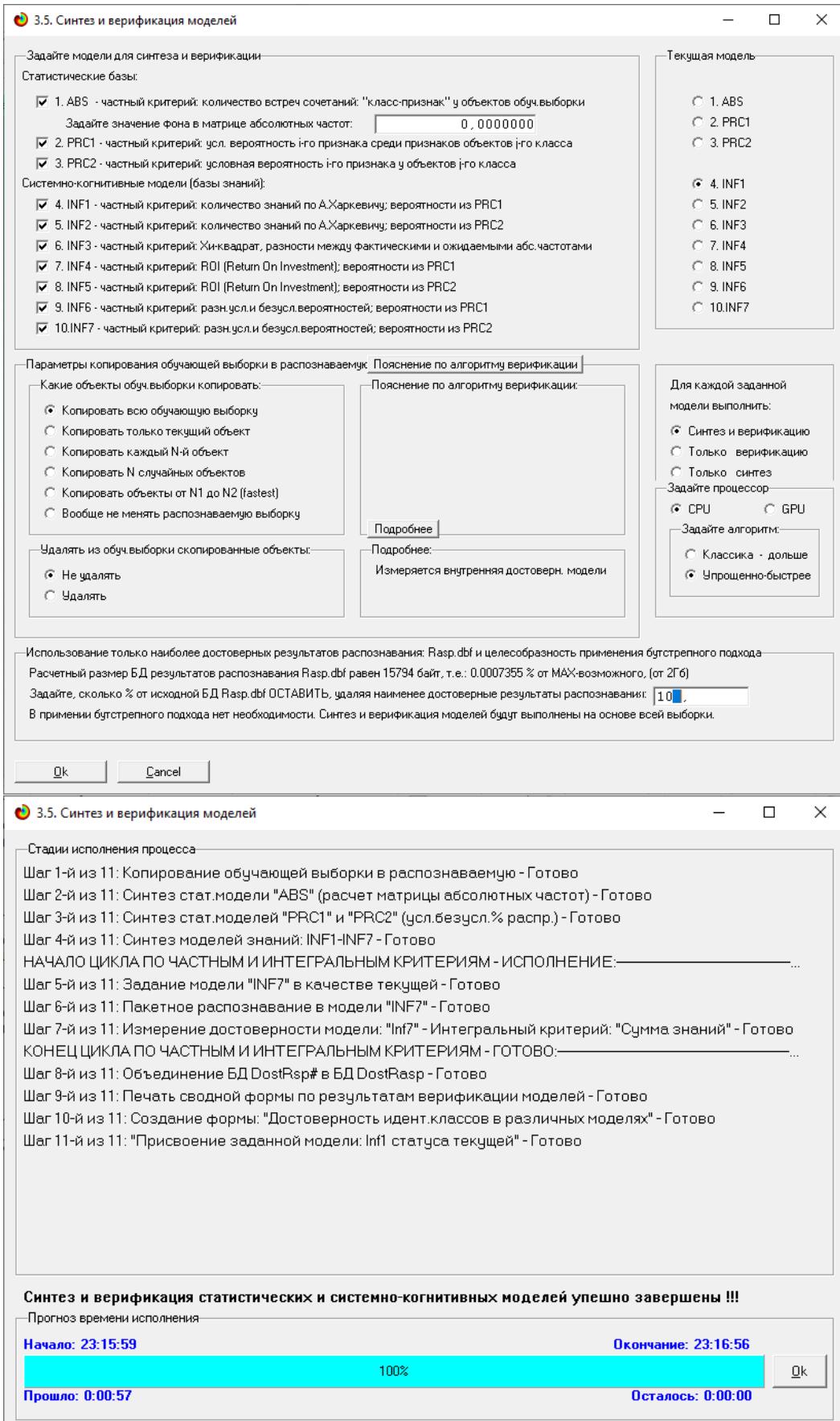


Рисунок 6. Экранные формы режима синтеза и верификации моделей

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 обуч.выборки"							
Код признака	Наименование описательной шкалы и градации	1. MEDIAN AGE 1/3 (23.0000000, 25.5333333)	2. MEDIAN AGE 2/3 (25.5333333, 28.0666667)	3. MEDIAN AGE 3/3 (28.0666667, 30.6000000)	Сумма	Среднее	Средн. квадр. откл.
1.0	NAME-Alabama	1.0	1.0		2.0	0.67	0.58
2.0	NAME-Alaska	1.0	1.0		2.0	0.67	0.58
3.0	NAME-Arizona		2.0		2.0	0.67	1.15
4.0	NAME-Arkansas	1.0	1.0		2.0	0.67	0.58
5.0	NAME-California		1.0	1.0	2.0	0.67	0.58
6.0	NAME-Colorado		2.0		2.0	0.67	1.15
7.0	NAME-Connecticut		1.0	1.0	2.0	0.67	0.58
8.0	NAME-Delaware		1.0	1.0	2.0	0.67	0.58
9.0	NAME-District of Columbia			2.0	2.0	0.67	1.15
10.0	NAME-Florida		1.0	1.0	2.0	0.67	0.58
11.0	NAME-Georgia		2.0		2.0	0.67	1.15
12.0	NAME-Hawaii			1.0	2.0	0.67	0.58
13.0	NAME-Idaho	2.0			2.0	0.67	1.15
14.0	NAME-Illinois		1.0	1.0	2.0	0.67	0.58
15.0	NAME-Indiana		2.0		2.0	0.67	1.15
16.0	NAME-Iowa	1.0	1.0		2.0	0.67	0.58
17.0	NAME-Kansas	1.0	1.0		2.0	0.67	0.58
18.0	NAME-Kentucky	1.0	1.0		2.0	0.67	0.58
19.0	NAME-Louisiana		2.0		2.0	0.67	1.15
20.0	NAME-Maine		2.0		2.0	0.67	1.15
21.0	NAME-Maryland		1.0	1.0	2.0	0.67	0.58
22.0	NAME-Massachusetts			2.0	2.0	0.67	1.15
23.0	NAME-Michigan		1.0	1.0	2.0	0.67	0.58
24.0	NAME-Minnesota		2.0		2.0	0.67	1.15

Рисунок 7. 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. MEDIAN AGE 1/3 (23.0000000, 25.5333333)	2. MEDIAN AGE 2/3 (25.5333333, 28.0666667)	3. MEDIAN AGE 3/3 (28.0666667, 30.6000000)	Безусл. вероятн.	Среднее	Средн. квадр. откл.
1.0	NAME-Alabama	5.556	1.493		1.923	2.349	2.924
2.0	NAME-Alaska	5.556	1.493		1.923	2.349	2.924
3.0	NAME-Arizona		2.985		1.923	0.995	1.771
4.0	NAME-Arkansas	5.556	1.493		1.923	2.349	2.924
5.0	NAME-California		1.493	5.263	1.923	2.252	2.758
6.0	NAME-Carolina		5.970		3.846	1.990	3.494
7.0	NAME-Colorado		2.985		1.923	0.995	1.771
8.0	NAME-Columbia			10.526	1.923	3.509	6.123
9.0	NAME-Connecticut		1.493	5.263	1.923	2.252	2.758
10.0	NAME-Dakota	5.556	4.478		3.846	3.344	2.994
11.0	NAME-Delaware		1.493	5.263	1.923	2.252	2.758
12.0	NAME-District			10.526	1.923	3.509	6.123
13.0	NAME-Florida		1.493	5.263	1.923	2.252	2.758
14.0	NAME-Georgia		2.985		1.923	0.995	1.771
15.0	NAME-Hampshire		1.493	5.263	1.923	2.252	2.758
16.0	NAME-Hawaii		1.493	5.263	1.923	2.252	2.758
17.0	NAME-Idaho	11.111			1.923	3.704	6.464
18.0	NAME-Illinois		1.493	5.263	1.923	2.252	2.758
19.0	NAME-Indiana		2.985		1.923	0.995	1.771
20.0	NAME-Iowa	5.556	1.493		1.923	2.349	2.924
21.0	NAME-Island		1.493	5.263	1.923	2.252	2.758
22.0	NAME-Jersey		1.493	5.263	1.923	2.252	2.758
23.0	NAME-Kansas	5.556	1.493		1.923	2.349	2.924
24.0	NAME-Kentucky	5.556	1.493		1.923	2.349	2.924

Рисунок 8. Статистическая модель «PRC2», матрица условных и безусловных процентных распределений

Код признака	Наименование описательной шкалы и градации	1.	2.	3.	Сумма	Среднее	Средн. квадр. откл.
		MEDIAN AGE 1/3 (23.000000, 25.533333)	MEDIAN AGE 2/3 (25.533333, 28.066667)	MEDIAN AGE 3/3 (28.066667, 30.600000)			
1.0	NAME-Alabama	0.279	-0.067		0.212	0.071	0.183
2.0	NAME-Alaska	0.279	-0.067		0.212	0.071	0.183
3.0	NAME-Arizona		0.116		0.116	0.039	0.067
4.0	NAME-Arkansas	0.279	-0.067		0.212	0.071	0.183
5.0	NAME-California		-0.067	0.265	0.198	0.066	0.175
6.0	NAME-Colorado		0.116		0.116	0.039	0.067
7.0	NAME-Connecticut		-0.067	0.265	0.198	0.066	0.175
8.0	NAME-Delaware		-0.067	0.265	0.198	0.066	0.175
9.0	NAME-District of Columbia			0.447	0.447	0.149	0.258
10.0	NAME-Florida		-0.067	0.265	0.198	0.066	0.175
11.0	NAME-Georgia		0.116		0.116	0.039	0.067
12.0	NAME-Hawaii		-0.067	0.265	0.198	0.066	0.175
13.0	NAME-Idaho	0.461			0.461	0.154	0.266
14.0	NAME-Illinois		-0.067	0.265	0.198	0.066	0.175
15.0	NAME-Indiana		0.116		0.116	0.039	0.067
16.0	NAME-Iowa	0.279	-0.067		0.212	0.071	0.183
17.0	NAME-Kansas	0.279	-0.067		0.212	0.071	0.183
18.0	NAME-Kentucky	0.279	-0.067		0.212	0.071	0.183
19.0	NAME-Louisiana		0.116		0.116	0.039	0.067
20.0	NAME-Maine		0.116		0.116	0.039	0.067
21.0	NAME-Maryland		-0.067	0.265	0.198	0.066	0.175
22.0	NAME-Massachusetts			0.447	0.447	0.149	0.258
23.0	NAME-Michigan		-0.067	0.265	0.198	0.066	0.175
24.0	NAME-Minnesota			0.116	0.116	0.039	0.067

Рисунок 9. Системно-когнитивная модель «INF1», матрица информавностей (по А.Харкевичу)

Код признака	Наименование описательной шкалы и градации	1.	2.	3.	Сумма	Среднее	Средн. квадр. откл.
		MEDIAN AGE 1/3 (23.000000, 25.533333)	MEDIAN AGE 2/3 (25.533333, 28.066667)	MEDIAN AGE 3/3 (28.066667, 30.600000)			
1.0	NAME-Alabama	0.663	-0.282	-0.381		0.576	
2.0	NAME-Alaska	0.663	-0.282	-0.381		0.576	
3.0	NAME-Arizona	-0.337	0.718	-0.381		0.622	
4.0	NAME-Arkansas	0.663	-0.282	-0.381		0.576	
5.0	NAME-California	-0.337	-0.282	0.619		0.537	
6.0	NAME-Carolina	-0.674	1.436	-0.762		1.244	
7.0	NAME-Colorado	-0.337	0.718	-0.381		0.622	
8.0	NAME-Columbia	-0.337	-1.282	1.619		1.480	
9.0	NAME-Connecticut	-0.337	-0.282	0.619		0.537	
10.0	NAME-Dakota	0.326	0.436	-0.762		0.662	
11.0	NAME-Delaware	-0.337	-0.282	0.619		0.537	
12.0	NAME-District	-0.337	-1.282	1.619		1.480	
13.0	NAME-Florida	-0.337	-0.282	0.619		0.537	
14.0	NAME-Georgia	-0.337	0.718	-0.381		0.622	
15.0	NAME-Hampshire	-0.337	-0.282	0.619		0.537	
16.0	NAME-Hawaii	-0.337	-0.282	0.619		0.537	
17.0	NAME-Idaho	1.663	-1.282	-0.381		1.509	
18.0	NAME-Illinois	-0.337	-0.282	0.619		0.537	
19.0	NAME-Indiana	-0.337	0.718	-0.381		0.622	
20.0	NAME-Iowa	0.663	-0.282	-0.381		0.576	
21.0	NAME-Island	-0.337	-0.282	0.619		0.537	
22.0	NAME-Jersey	-0.337	-0.282	0.619		0.537	
23.0	NAME-Kansas	0.663	-0.282	-0.381		0.576	
24.0	NAME-Kentucky	0.663	-0.282	-0.381		0.576	

Рисунок 10 . 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»:  $F=0.867$  (1-й рисунок 14).

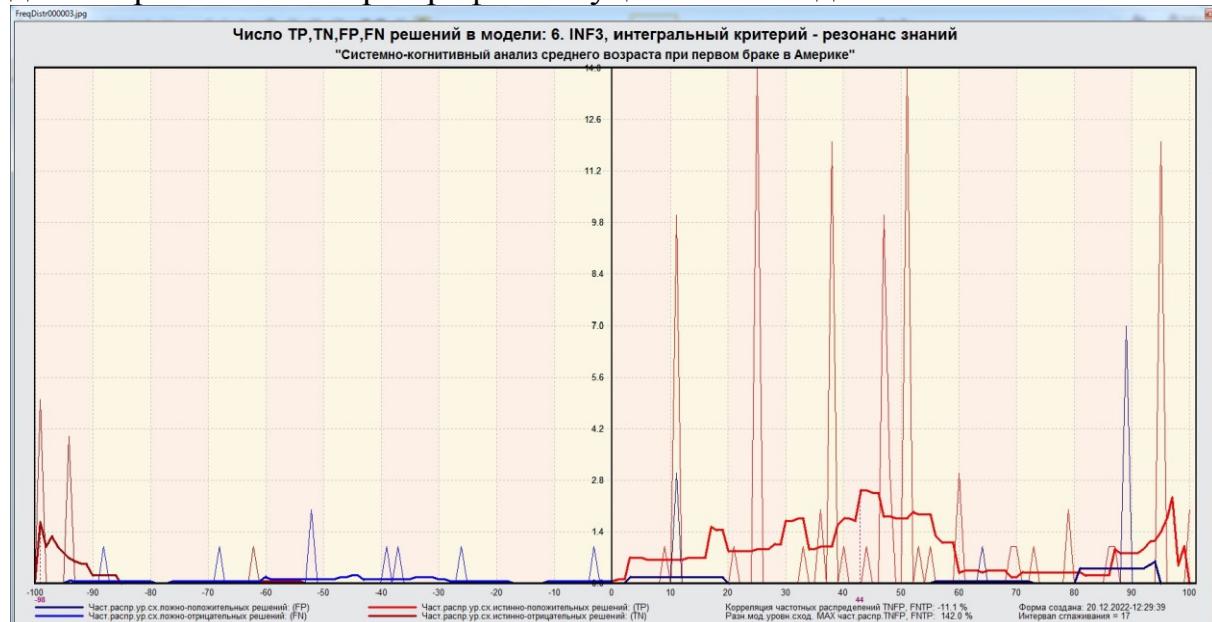
В соответствии с критерием достоверности L1-мерой проф. Е. В. Луценко наиболее достоверной является СК-модель INF3 (хи-квадрат К. Пирсона) с интегральным критерием: «Сумма знаний»:  $L1=0.961$  (2-й рисунок 14). *Эту модель и будем использовать для решения поставленных в работе задач.*

Нажмите на название модели и частного критерия	Интегральный критерий	Всего логических объектов выборки	Число истинных положительных решений (TP)	Число истинных отрицательных решений (TN)	Число ложных положительных решений (FP)	Число ложных отрицательных решений (FN)	Точность	Полнота	F-мера	Сумма моделей, усредненной по истинно-положительным решениям (STP)	Сумма моделей, усредненной по истинно-отрицательным решениям (STN)	Сумма моделей, усредненной по ложноположительным решениям
1. ABS - частный критерий: количество встроенных синонимов "кис"	Корреляция abs.частот с обр...	104	104	21		0.832	1.000	0.908	86.680		15	
1. ABS - частный критерий: количество встроенных синонимов "кис"	Сумма abs.частот по признакам	104	104	21		0.832	1.000	0.908	75.155		7	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл. частот с ...	104	104	21		0.832	1.000	0.908	86.680		15	
2. PRC1 - частный критерий: усл. вероятность i-го признака сред...	Сумма усл. частот по приз...	104	104	21		0.832	1.000	0.908	81.099		14	
3. PRC2 - частный критерий: условная вероятность i-го признака	Корреляция усл. частот с ...	104	104	21		0.832	1.000	0.908	86.679		15	
3. PRC2 - частный критерий: условная вероятность i-го признака	Сумма усл. частот по приз...	104	104	21		0.832	1.000	0.908	83.236		15	
4. INF1 - частный критерий: количество знаний по А.Хархавину	Семантический резонанс зна...	104	64	14	7	0.901	0.615	0.731	15.879		8.865	0
4. INF1 - частный критерий: количество знаний по А.Хархавину	Сумма знаний	104	74	13	8	0.902	0.712	0.796	14.116		2.715	1
5. INF2 - частный критерий: количество знаний по А.Хархавину	Семантический резонанс зна...	104	74	14	7	0.914	0.712	0.800	15.701		8.386	0
5. INF2 - частный критерий: количество знаний по А.Хархавину	Сумма знаний	104	75	13	8	0.904	0.721	0.802	13.699		2.414	1
6. INF3 - частный критерий: Хиггерад, различия между фактами	Семантический резонанс зна...	104	96	10	11	0.897	0.923	0.910	46.320		9.317	7
6. INF3 - частный критерий: Хиггерад, различия между фактами	Сумма знаний	104	96	10	11	0.897	0.923	0.910	33.777		9.008	7
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	104	59	21		0.866	1.000	0.567	0.724		10.010	5.343
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	104	79	9	12	0.868	0.760	0.810	10.733		0.479	0
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	104	59	21		0.866	1.000	0.567	0.724		9.800	5.269
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	104	78	9	12	0.867	0.750	0.804	10.287		0.430	0
9. INF6 - частный критерий: разн. усл. и безузл.вероятностей; вер...	Семантический резонанс зна...	104	95	10	11	0.896	0.913	0.905	41.207		9.139	6
9. INF6 - частный критерий: разн. усл. и безузл.вероятностей; вер...	Сумма знаний	104	96	10	11	0.897	0.923	0.910	25.791		7.627	7
10. INF7 - частный критерий: разн. усл. и безузл.вероятностей; вер...	Семантический резонанс зна...	104	85	13	8	0.914	0.817	0.863	38.560		9.609	6
10. INF7 - частный критерий: разн. усл. и безузл.вероятностей; вер...	Сумма знаний	104	85	13	8	0.914	0.817	0.863	26.687		7.225	6

Наименование модели и частного критерия	Интегральный критерий	Иногда	Гамма Ван Ризбергена	Симма модул. зонной схемы исполн.-стр. решений (STN)	Симма модул. зонной схемы исполн.-стр. решения (SFR)	Симма модул. зонной схемы исполн.-стр. решения (SFN)	S-Точность модели	S-Полнота модели	L1-мера проф. Е.В. Лиденко	Средний индекс зонной схемы исполн.-стр. решений	Средний индекс зонной схемы исполн.-стр. решений
1. ABS - частный критерий: количество встреч Сочетаний: "клас..."	Корреляция abs. частот с обр...	1. 000	0. 908	86. 680	15. 054		0. 852	1. 000	0. 920	0. 833	
1. ABS - частный критерий: количество встреч Сочетаний: "клас..."	Сумма abs. частот по признаку:	1. 000	0. 908	75. 155	7. 825		0. 906	1. 000	0. 951	0. 723	
2. PRC1 - частный критерий: усл. вероятность i-го признака с предиктором	Корреляция усл. отн частот с о...	1. 000	0. 908	86. 680	15. 054		0. 852	1. 000	0. 920	0. 833	
2. PRC1 - частный критерий: усл. вероятность i-го признака с предиктором	Сумма усл. отн частот по приз...	1. 000	0. 908	81. 099	14. 825		0. 845	1. 000	0. 916	0. 760	
3. PRC2 - частный критерий: условная вероятность i-го признака.	Корреляция усл. отн частот с о...	1. 000	0. 908	86. 679	15. 054		0. 852	1. 000	0. 920	0. 833	
3. PRC2 - частный критерий: условная вероятность i-го признака.	Сумма усл. отн частот по приз...	1. 000	0. 908	83. 236	15. 212		0. 845	1. 000	0. 916	0. 800	
4. INF1 - частный критерий: количество знаний по АХаркевичу, в...	Семантический резонанс зна...	0. 615	0. 731	15. 879	8. 865	0. 334	7. 523	0. 979	0. 679	0. 802	0. 248
4. INF1 - частный критерий: количество знаний по АХаркевичу, в...	Сумма знаний	0. 712	0. 796	14. 116	2. 715	1. 401	1. 280	0. 910	0. 917	0. 913	0. 191
5. INF2 - частный критерий: количество знаний по АХаркевичу, в...	Семантический резонанс зна...	0. 712	0. 800	15. 701	8. 386	0. 219	7. 476	0. 986	0. 677	0. 803	0. 212
5. INF2 - частный критерий: количество знаний по АХаркевичу, в...	Сумма знаний	0. 721	0. 802	13. 699	2. 414	1. 244	1. 324	0. 917	0. 912	0. 914	0. 183
6. INF3 - частный критерий: Квадрат разности между фактическим...	Семантический резонанс зна...	0. 923	0. 910	46. 320	9. 317	7. 217	3. 648	0. 865	0. 927	0. 895	0. 482
6. INF3 - частный критерий: Квадрат разности между фактическим...	Сумма знаний	0. 923	0. 910	33. 777	9. 008	7. 144	3. 431	0. 825	0. 908	0. 865	0. 352
7. INF4 - частный критерий ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0. 567	0. 724	10. 010	5. 343		7. 060	1. 000	0. 586	0. 739	0. 170
7. INF4 - частный критерий ROI (Return On Investment); вероятно...	Сумма знаний	0. 760	0. 810	10. 733	0. 479	0. 784	0. 315	0. 932	0. 972	0. 951	0. 136
8. INF5 - частный критерий ROI (Return On Investment); вероятно...	Семантический резонанс зна...	0. 567	0. 724	9. 800	5. 269		6. 965	1. 000	0. 585	0. 738	0. 166
8. INF5 - частный критерий ROI (Return On Investment); вероятно...	Сумма знаний	0. 750	0. 804	10. 287	0. 438	0. 723	0. 323	0. 934	0. 563	0. 951	0. 132
9. INF6 - частный критерий: разн. усл и безсл.вероятностей, вер...	Семантический резонанс зна...	0. 913	0. 905	41. 207	9. 138	6. 467	4. 369	0. 864	0. 904	0. 884	0. 434
9. INF6 - частный критерий: разн. усл и безсл.вероятностей, вер...	Сумма знаний	0. 923	0. 910	25. 791	7. 627	7. 234	2. 970	0. 783	0. 897	0. 895	0. 269
10. INF7 - частный критерий: разн. усл и безсл.вероятностей, вер...	Семантический резонанс зна...	0. 817	0. 863	38. 560	9. 609	6. 552	6. 485	0. 855	0. 856	0. 855	0. 454
10. INF7 - частный критерий: разн. усл и безсл.вероятностей, вер...	Сумма знаний	0. 817	0. 863	26. 687	7. 225	6. 749	3. 040	0. 798	0. 898	0. 845	0. 314

**Рисунок 11. Экранные формы режима измерения достоверности моделей 3.4**

На рисунке 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.



**Рисунок 12. Частотные распределения количества истинных и ложных, положительных и отрицательных решений в наиболее достоверной по L1-мере проф.Е.В.Луценко СК-модели INF3**

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

- отрицательные ложные решения вообще практически не встречаются не встречаются, за исключением 1-го случая при уровне различия –10%;
- 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 большее доля истинных решений. Поэтому уровень сходства является адекватной внутренней мерой системы «Эйдос», так сказать адекватной мерой самооценки или аудита степени достоверности решений и уровня риска ошибочного решения.***

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

Помощь по режимам: 3.4, 4.1.3.#: Виды прогнозов и меры достоверности моделей в системе "Эйдос-X++"

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

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

**ОТРИЦАТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.**

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

**ИДЕАЛЬНЫЙ ПРОГНОЗ.**

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

**РЕАЛЬНЫЙ ПРОГНОЗ.**

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

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

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

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

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

Классическая F-мера достоверности моделей Ван Ризбергена (колонка выделена ярко-голубым фоном):

F-мера =  $2 \cdot [Precision \cdot Recall] / [Precision + Recall]$  - достоверность модели

Precision =  $TP / (TP + FP)$  - точность модели;

Recall =  $TP / (TP + FN)$  - полнота модели;

L1-мера проф. Е.В.Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СУММ уровней сходства (колонка выделена ярко-зеленым фоном):

L1-мера =  $2 \cdot (SPrecision \cdot SRecall) / (SPrecision + SRecall)$

SPrecision =  $STP / (STP + SFP)$  - точность с учетом сумм уровней сходства;

SRecall =  $STP / (STP + SFN)$  - полнота с учетом сумм уровней сходства;

STP - Сумма модулей сходства истинно-положительных решений; STN - Сумма модулей сходства истинно-отрицательных решений;

SFP - Сумма модулей сходства ложно-положительных решений; SFN - Сумма модулей сходства ложно-отрицательных решений.

L2-мера проф. Е.В.Луценко - нечеткое мультиклассовое обобщение классической F-меры с учетом СРЕДНИХ уровней сходства (колонка выделена желтым фоном):

L2-мера =  $2 \cdot (APrecision \cdot ARcall) / (APrecision + ARcall)$

APrecision =  $ATP / (ATP + AFP)$  - точность с учетом средних уровней сходства;

ARcall =  $ATP / (ATP + AFN)$  - полнота с учетом средних уровней сходства;

ATP=STP/TP - Среднее модулей сходства истинно-положительных решений; AFN=SFN/FN - Среднее модулей сходства истинно-отрицательных решений;

AFP=SFP/FP - Среднее модулей сходства ложно-положительных решений; AFN=SFN/FN - Среднее модулей сходства ложно-отрицательных решений.

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

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

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

Активай  
Чтобы ак

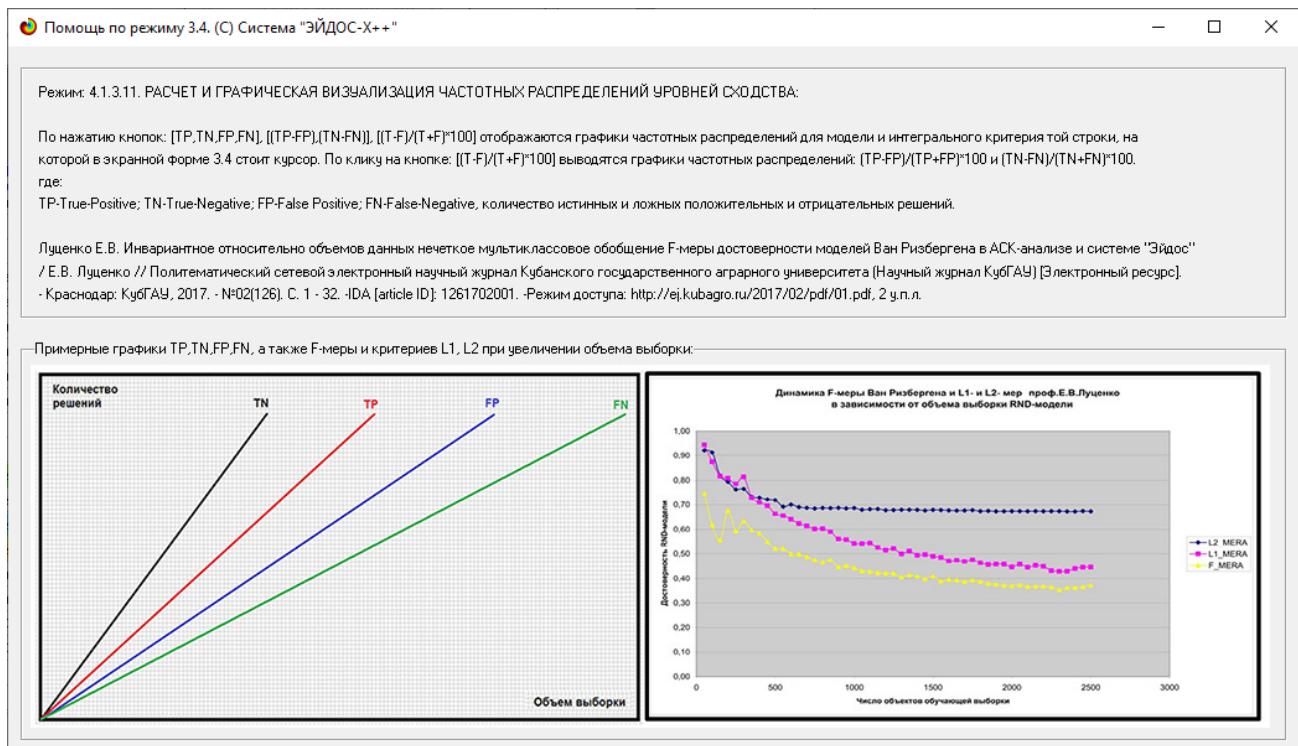


Рисунок 13. Экранные формы хелпов режима измерения достоверности моделей

### 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.

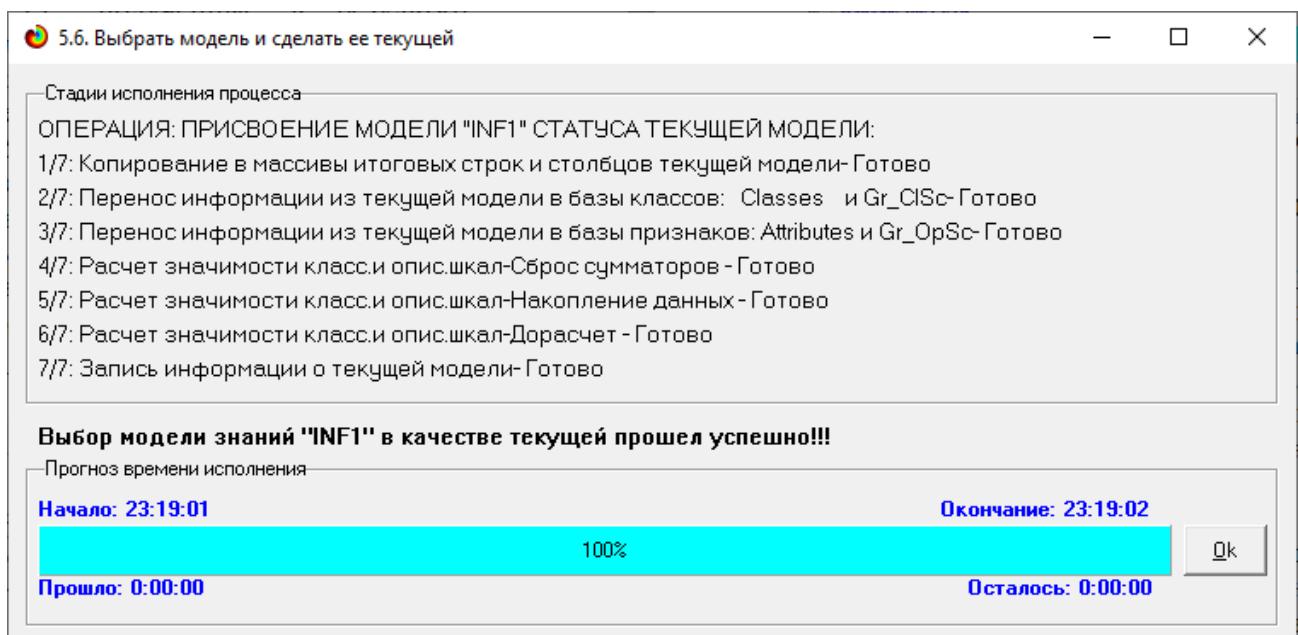
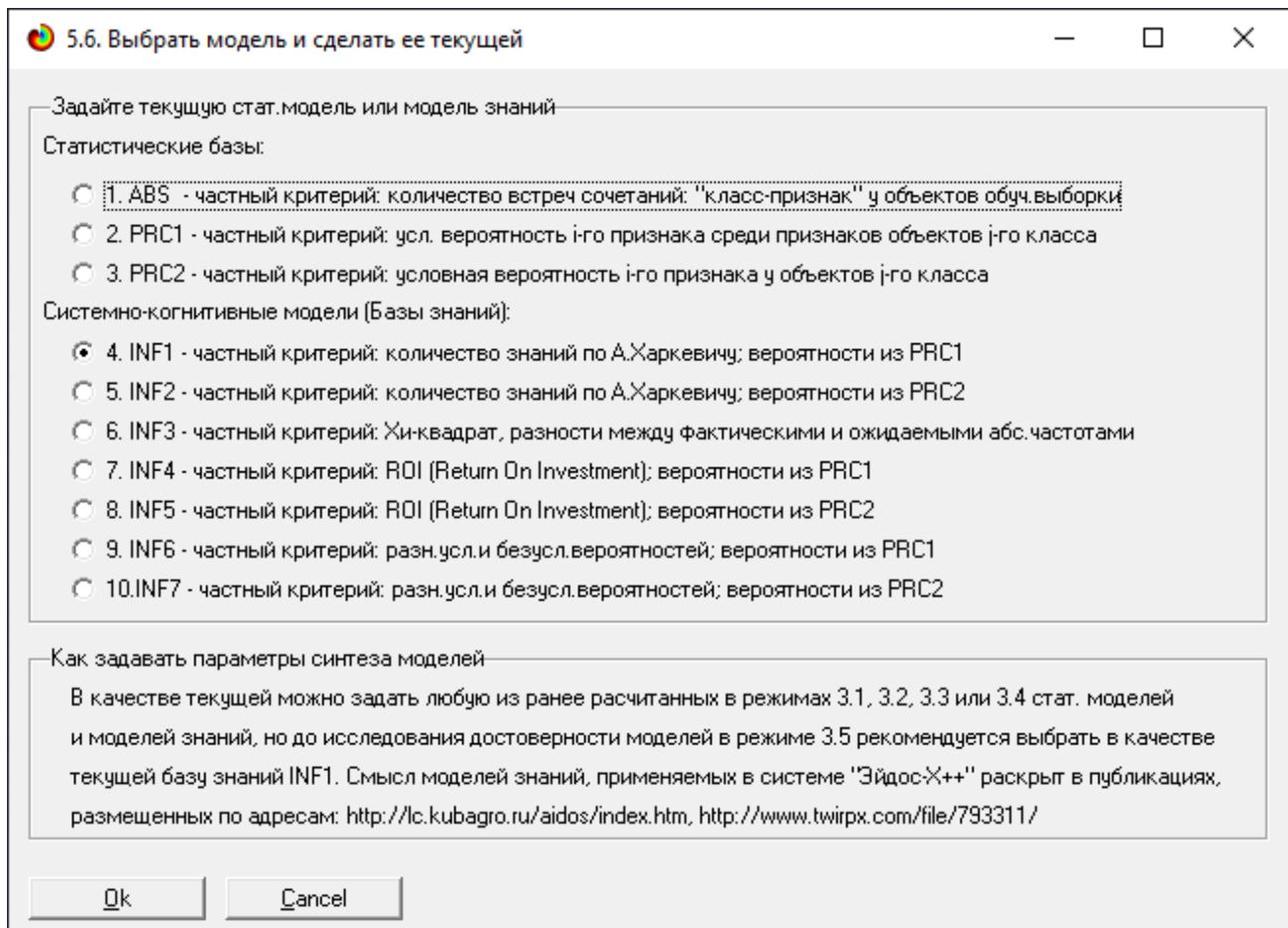


Рисунок 14. Задание СК-модели INF3 в качестве текущей

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

Сравнение осуществляется путем применения *неметрических интегральных критериев*, которых в настоящее время в системе «Эйдос» используется два. Эти интегральные критерии интересны тем, что корректны<sup>8</sup> 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:

---

<sup>8</sup> В отличие от Евклидова расстояния, которое используется для подобных целей наиболее часто

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

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.

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;

$\sigma_j$  – среднеквадратичное отклонение частных критериев знаний вектора класса;  $\sigma_l$  – 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.:

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

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:

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

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:

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

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 **неметрическую** 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 экранные формы режима идентификации и прогнозирования 4.1.2 системы «Эйдос»:

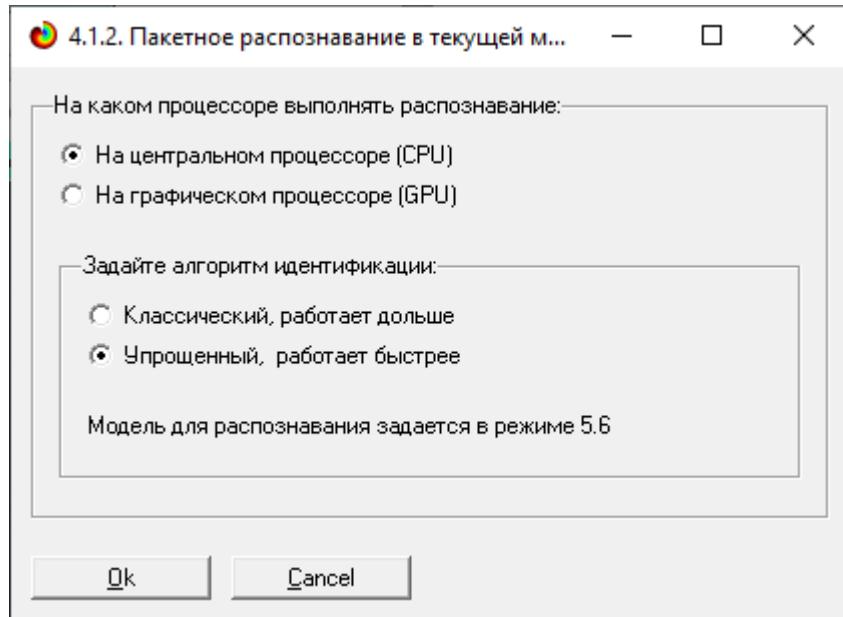
### 3.6.4. Решение задачи идентификации и прогнозирования в системе «Эйдос»

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 освещены и на теоретическом уровне, и с детальными численными примерами в работах [13, 14] и в ряде других<sup>9</sup>.

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

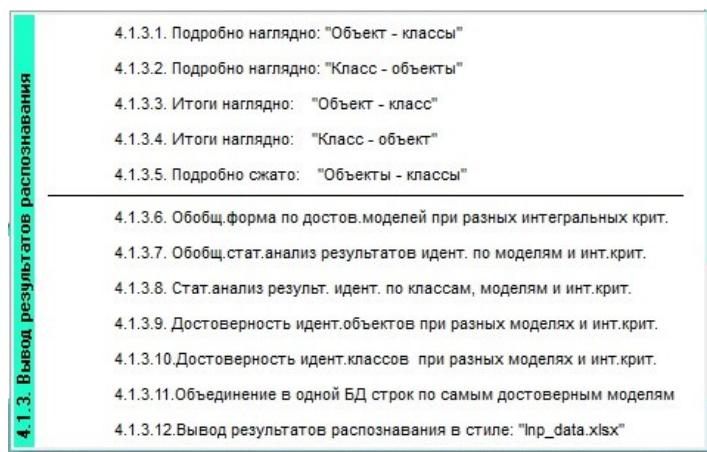
---

<sup>9</sup> См., например: [http://lc.kubagro.ru/aidos/Works\\_on\\_Scenario\\_ASC-analysis.htm](http://lc.kubagro.ru/aidos/Works_on_Scenario_ASC-analysis.htm)



**Рисунок 15. Экранная форма режима 4.1.2 идентификации и прогнозирования**

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



**Рисунок 16. 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):

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

Распознаваемые объекты	
Код	Наим.объекта
1	0
2	1
3	2
4	3
5	4
6	5
7	6
8	7
9	8
10	9
11	10
12	11
13	12
14	13
15	14
16	15
17	16
18	17
19	18
20	19
21	20
22	21
23	22

**Интегральный критерий сходства: "Семантический резонанс знаний"**

Код	Наименование класса	Сходство	Ф...	Сходство
1	MEDIAN AGE-1/3-(23.0, 25.5)	88,35...	v	
2	MEDIAN AGE-2/3-(25.5, 28.1)	-12,73...		
3	MEDIAN AGE-3/3-(28.1, 30.6)	-100,00...		

**Интегральный критерий сходства: "Сумма знаний"**

Код	Наименование класса	Сходство	Ф...	Сходство
1	MEDIAN AGE-1/3-(23.0, 25.5)	88,69...	v	
2	MEDIAN AGE-2/3-(25.5, 28.1)	-0,200...		
3	MEDIAN AGE-3/3-(28.1, 30.6)	-81,41...		

**Помощь**   **9 классов**   **Классы с MaxMin УрCx**   **9 классов с MaxMin УрCx**   **Все классы**   **ВКЛ. фильтр по класс.шкале**   **ВыКЛ.фильтр по класс.шкале**   **Граф.диаграммы**

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

Классы	
Код	Наим. класса
1	MEDIAN AGE-1/3-(23.0000000, 25.5333333)
2	MEDIAN AGE-2/3-(25.5333333, 28.0666667)
3	MEDIAN AGE-3/3-(28.0666667, 30.6000000)

**Интегральный критерий сходства: "Семантический резонанс знаний"**

Код	Наименование объекта	Сходство	Ф...	Сходство
13	12	94,20...	v	
45	44	94,20...	v	
1	0	88,35...	v	
2	1	88,35...	v	
4	3	88,35...	v	
16	15	88,35...	v	
17	16	88,35...	v	
18	17	88,35...	v	
27	26	88,35...	v	
28	27	88,35...	v	

**Интегральный критерий сходства: "Сумма знаний"**

Код	Наименование объекта	Сходство	Ф...	Сходство
13	12	94,22...	v	
45	44	94,22...	v	
49	48	90,40...	v	
1	0	88,69...	v	
2	1	88,69...	v	
4	3	88,69...	v	
16	15	88,69...	v	
17	16	88,69...	v	
18	17	88,69...	v	
27	26	88,69...	v	

**Помощь**   **Поиск объекта**   **В начало БД**   **В конец БД**   **Предыдущая**   **Следующая**   **9 записей**   **Все записи**   **Печать XLS**   **Печать TXT**   **Печать ALL**

**Рисунок 17. 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

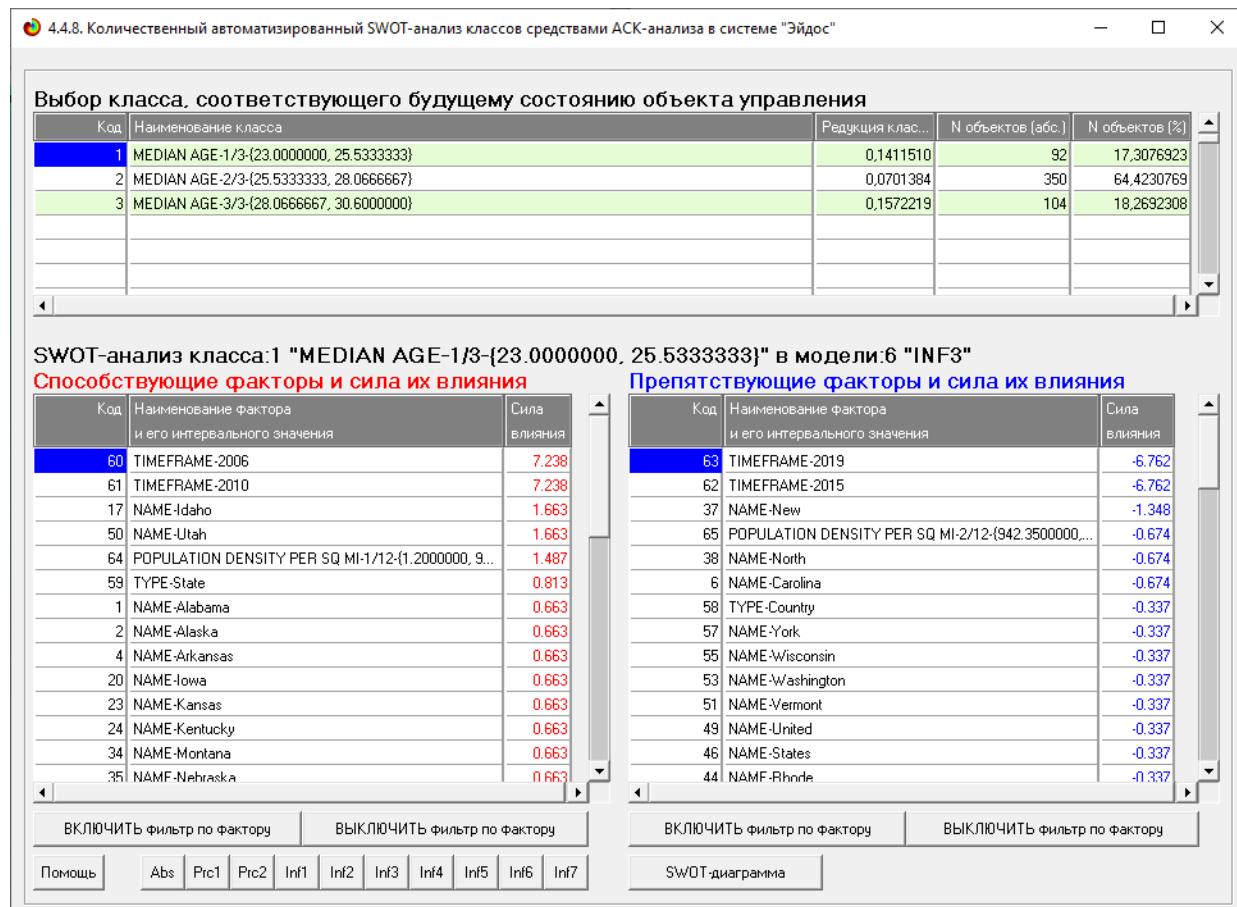
Задачи прогнозирования и принятия решений относятся друг к другу как прямая и *обратная* задачи:

- 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.



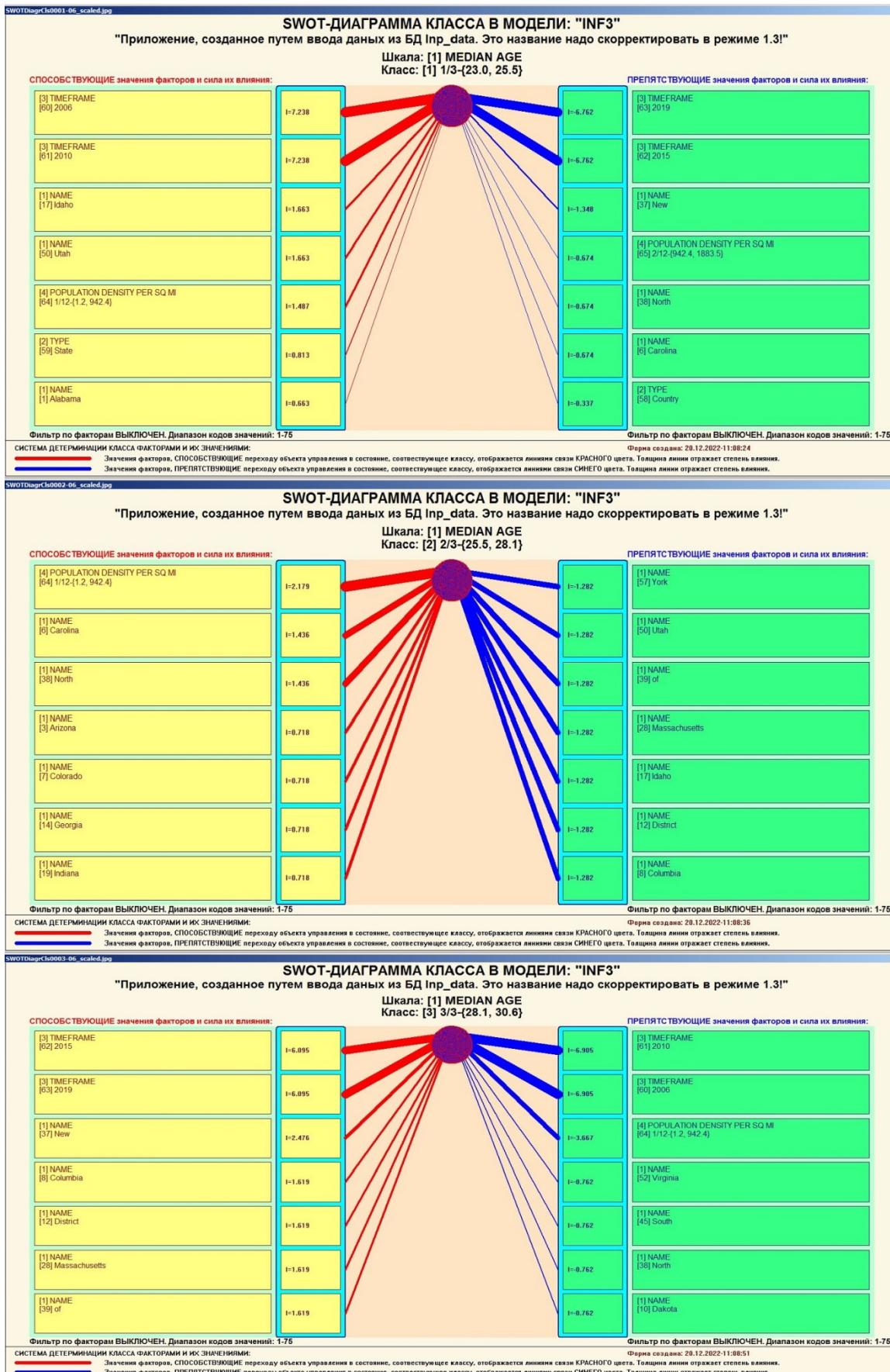


Рисунок 18. Примеры экранной формы режима автоматизированного SWOT-анализа (режим 4.4.8 системы «Эйдос»)

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].

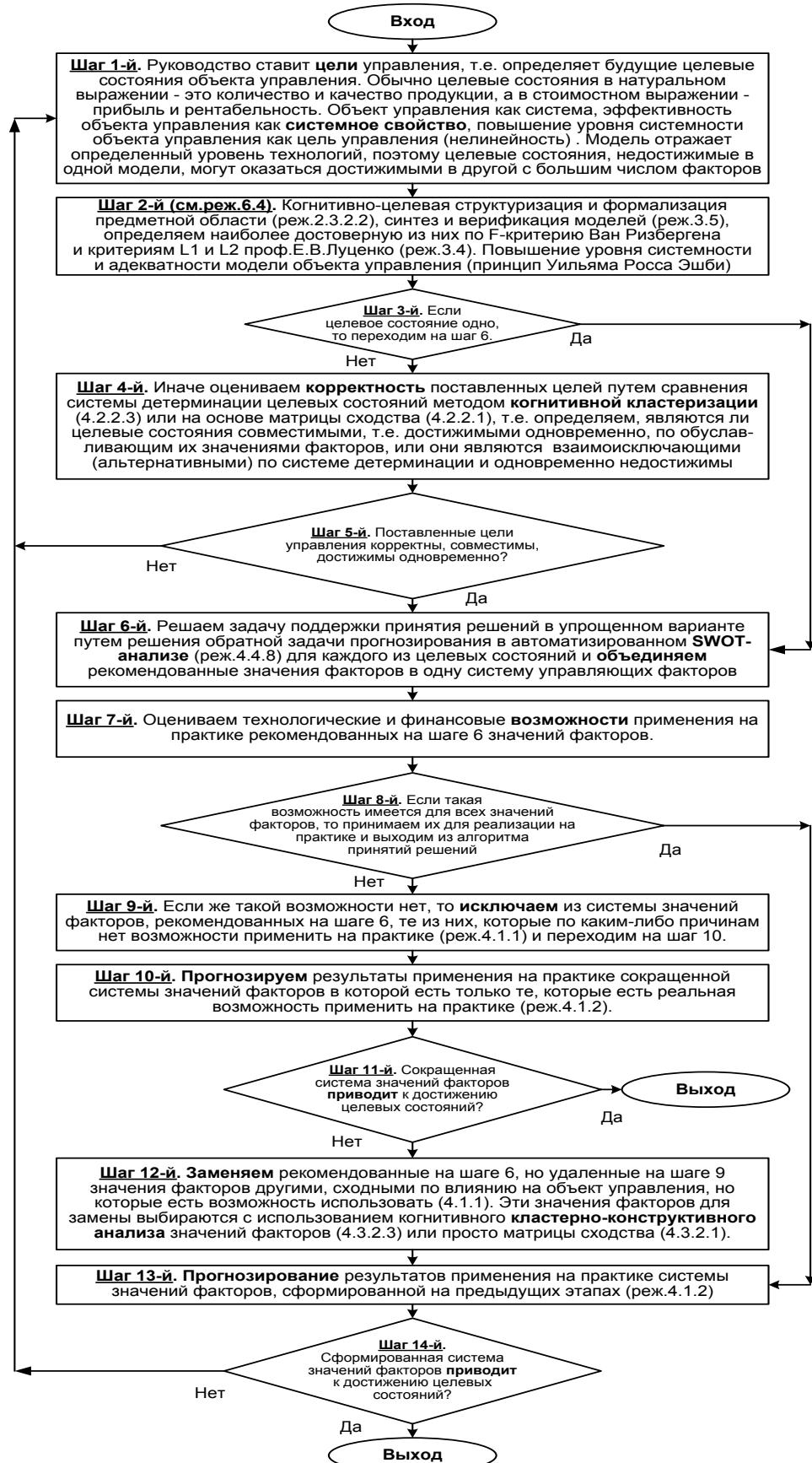


Рисунок 19. Развитый алгоритм принятия решений в интеллектуальных системах управления на основе АСК-анализа и системы «Эйдос»

**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.

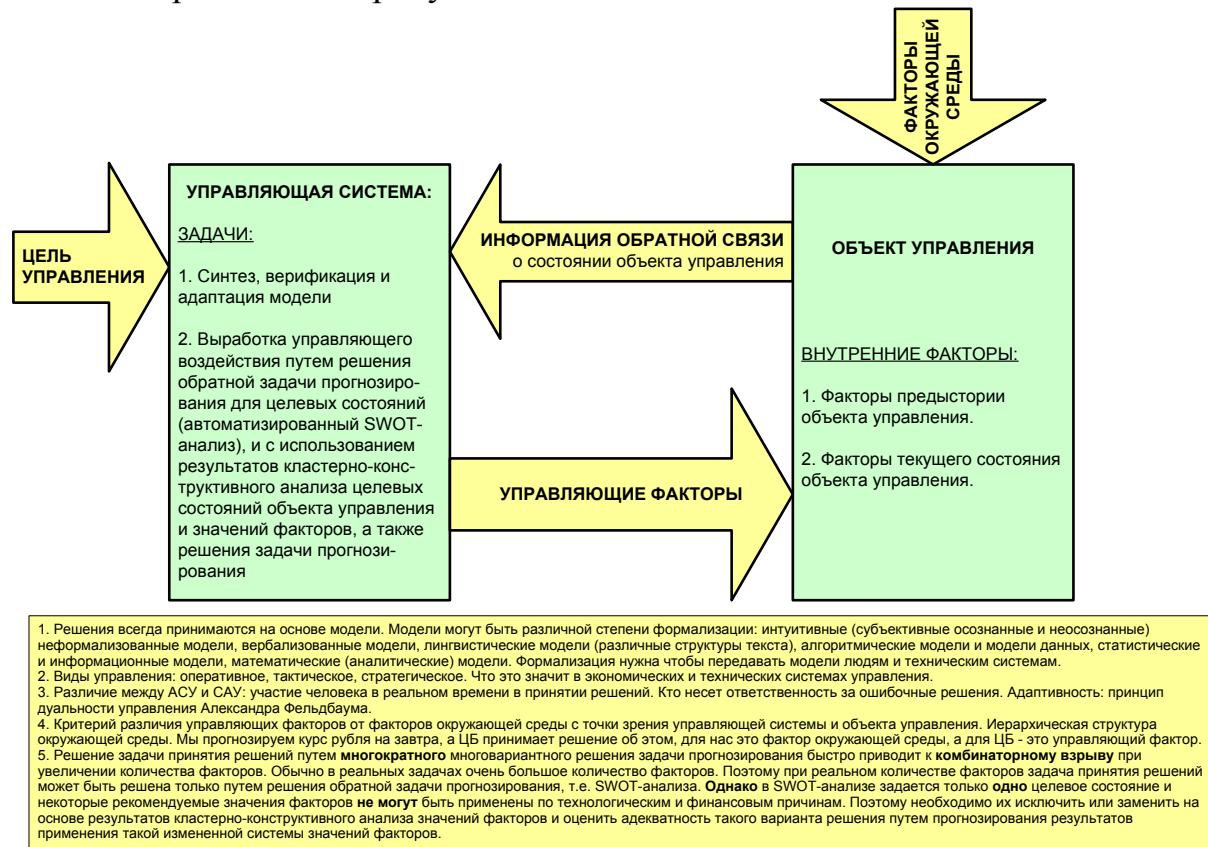
**Шаг 12-й.** **Заменяем** рекомендованные на шаге 6, но удаленные на шаге 9 значения факторов другими, сходными по влиянию на объект управления, но которые есть возможность использовать (4.1.1). Эти значения факторов для замены выбираются с использованием

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 20. 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:

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

Код	Наименование значения фактора
1	NAME-Alabama
2	NAME-Alaska
3	NAME-Arizona
4	NAME-Arkansas
5	NAME-California
6	NAME-Carolina

**SWOT-анализ значения фактора: 1 "NAME-Alabama" в модели: 6 "INF3"**

**СПОСОБСТВУЕТ:**

Код	Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ	Сила влиян...
1	MEDIAN AGE-1/3-{23.000000, 25.533333}	0.663

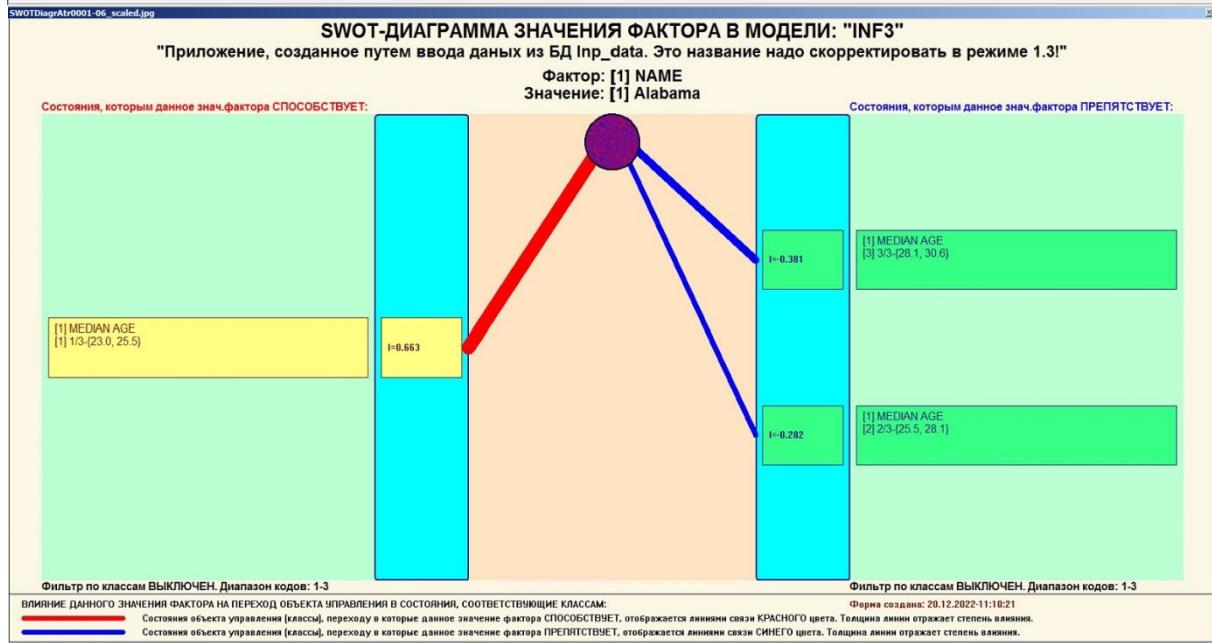
**ПРЕПЯТСТВУЕТ:**

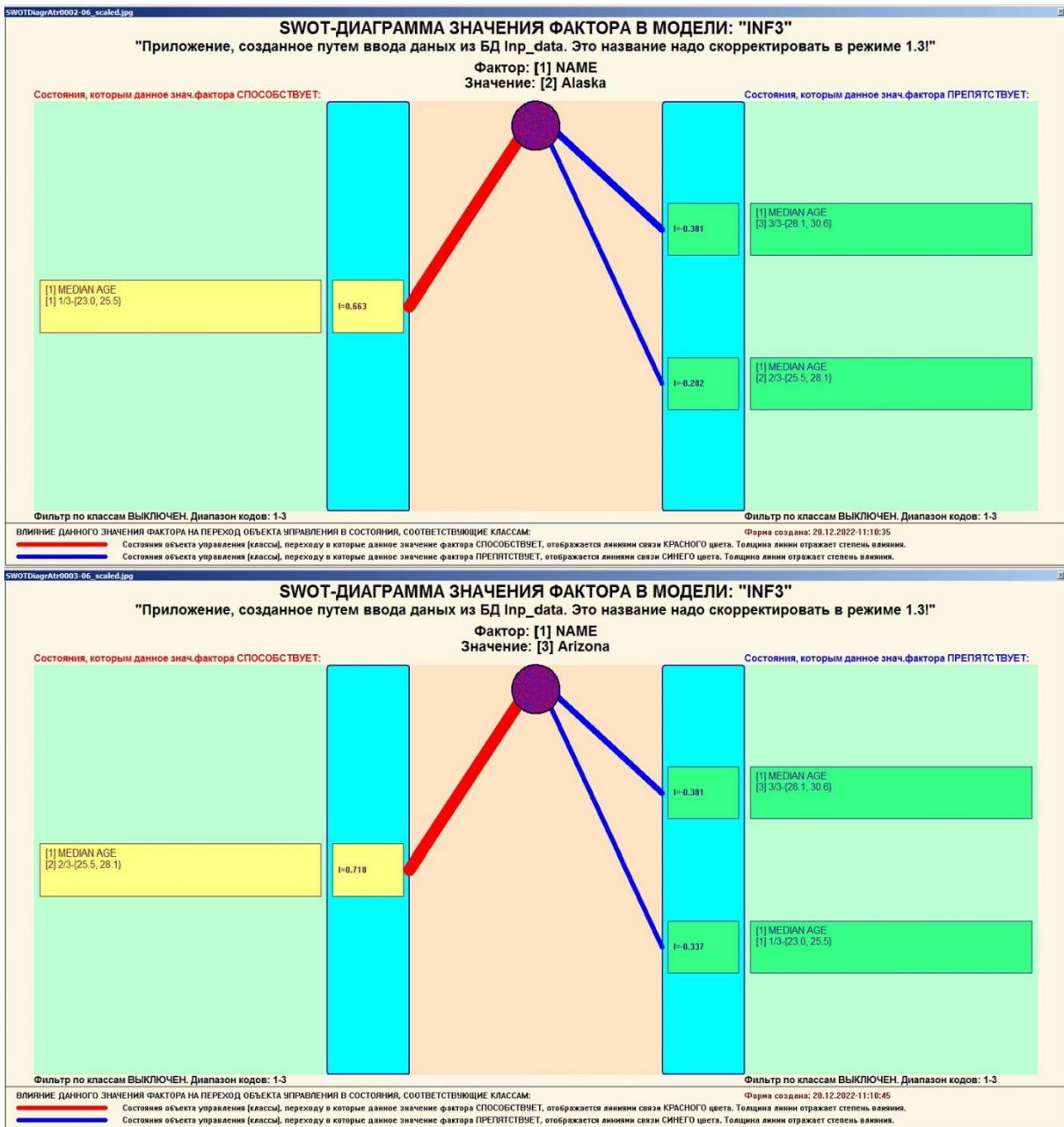
Код	Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ	Сила влияния
3	MEDIAN AGE-3/3-{28.066667, 30.600000}	-0.381
2	MEDIAN AGE-2/3-{25.533333, 28.066667}	-0.282

ВКЛЮЧИТЬ фильтр по кл.шкале      ВЫКЛЮЧИТЬ фильтр по кл.шкале

ВКЛЮЧИТЬ фильтр по кл.шкале      ВЫКЛЮЧИТЬ фильтр по кл.шкале

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





**Рисунок 21. 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.

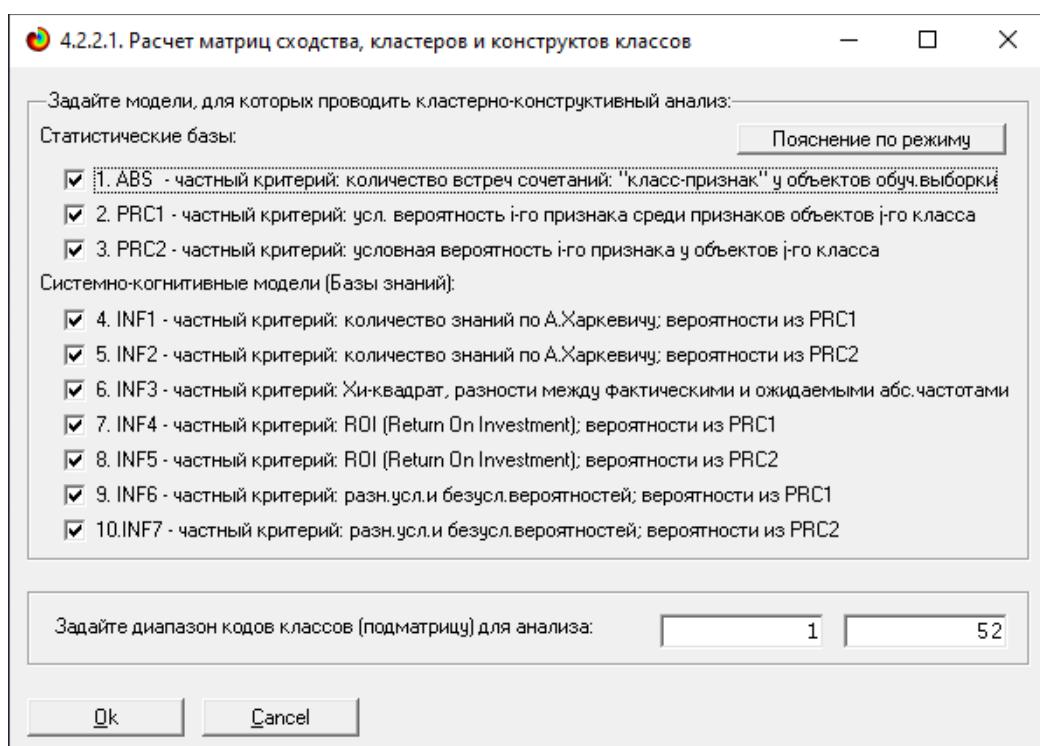
### 3.8.2. Кластерно-конструктивный анализ классов

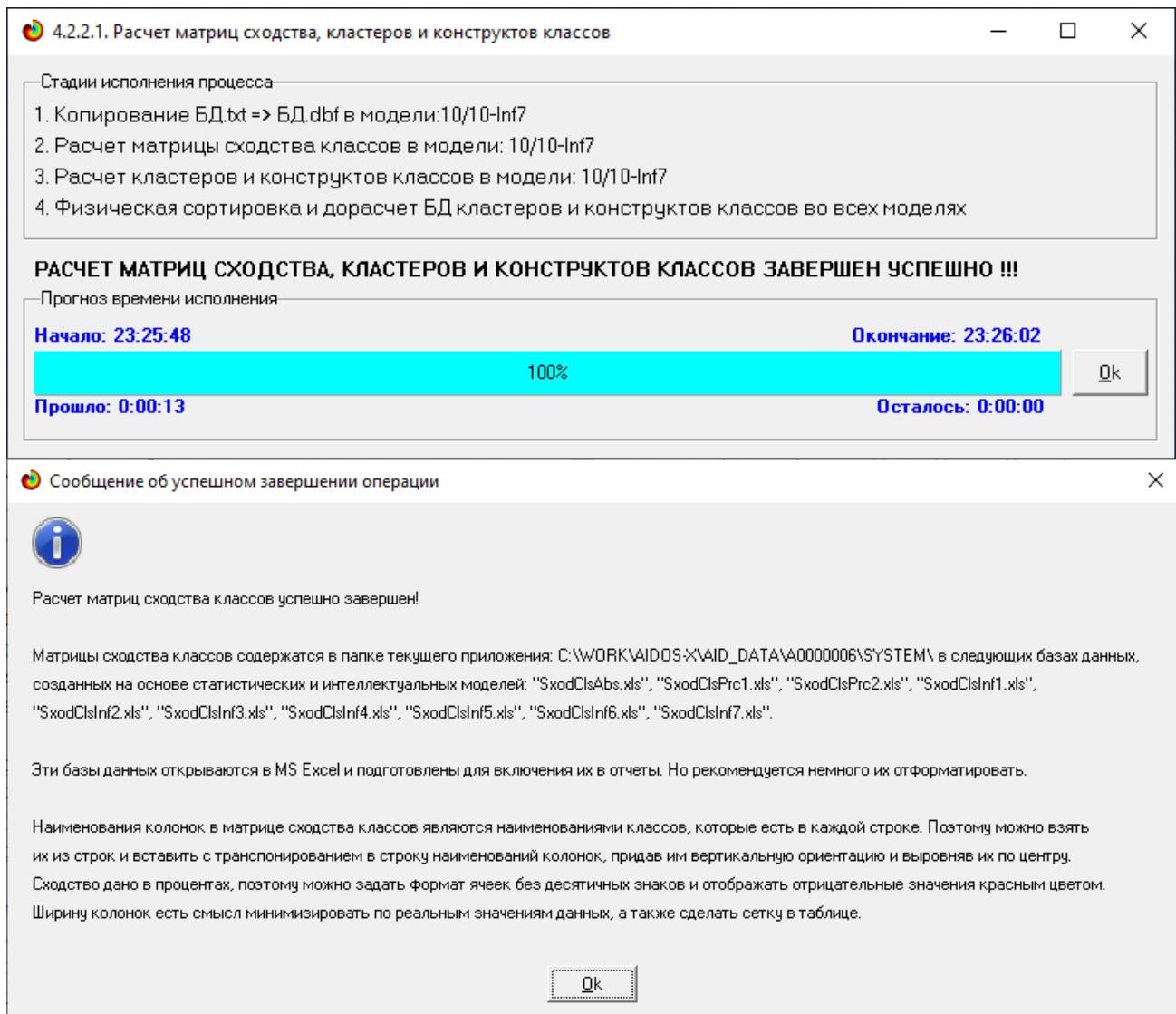
В системе «Эйдос» (в режиме 4.2.2.1, рисунок 25) рассчитывается матрица сходства классов (таблица 14) по системе их детерминации и на основе этой матрицы рассчитывается и выводится три основных формы:

- 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 сходства (таблица 14) используются 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 22. Экранные формы режима 4.2.2.1, обеспечивающего расчет матриц сходства классов**

**Таблица 5 – Матрица сходства классов в СК-модели INF3 (полностью)**

	A	B	C	D	E	F
1	KOD_CLS	KOD_CLSC	NAME_CLS	N1	N2	N3
2	1	1	MEDIAN AGE-1/3-[23.0000000, 25.5333333]	100,0000000	-20,1472624	-91,7422218
3	2	1	MEDIAN AGE-2/3-[25.5333333, 28.0666667]	-20,1472624	100,0000000	-20,4920107
4	3	1	MEDIAN AGE-3/3-[28.0666667, 30.6000000]	-91,7422218	-20,4920107	100,0000000

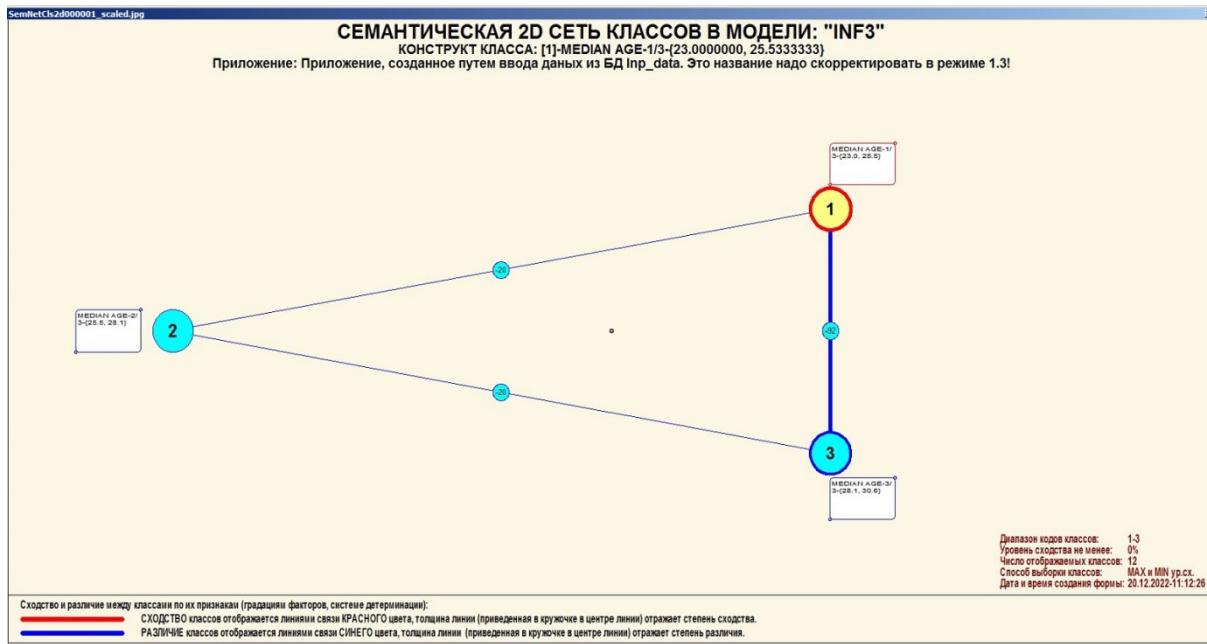


Рисунок 23. Круговая 2d-когнитивная диаграмма классов (режим 4.2.2.2)

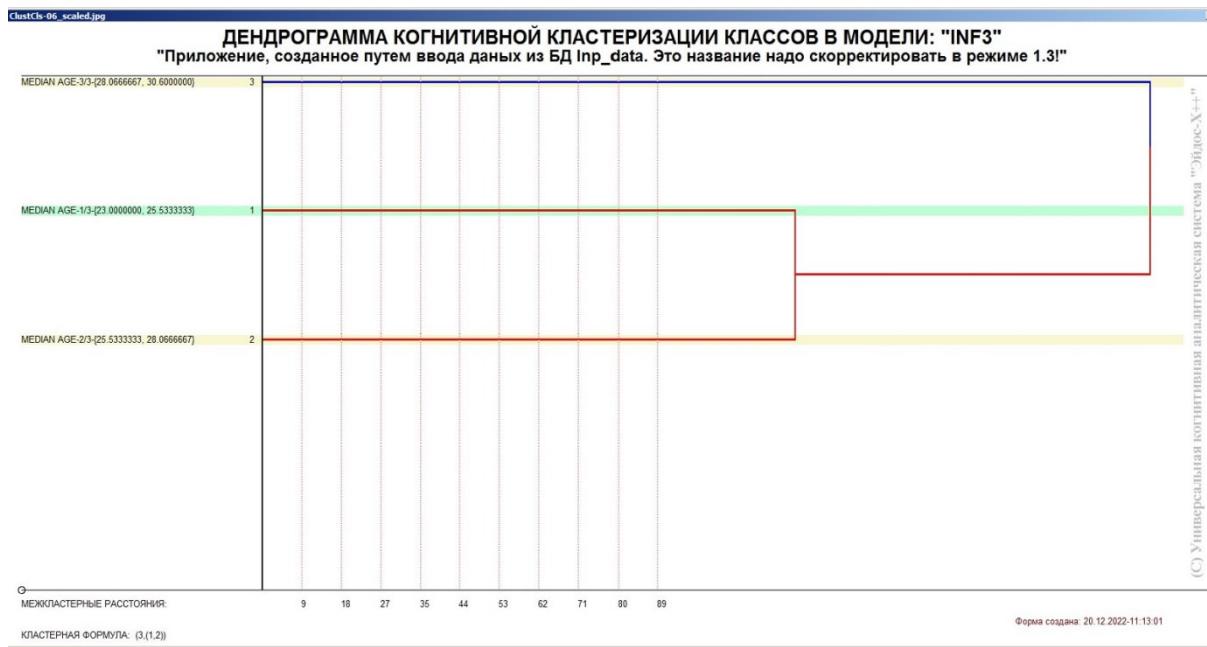
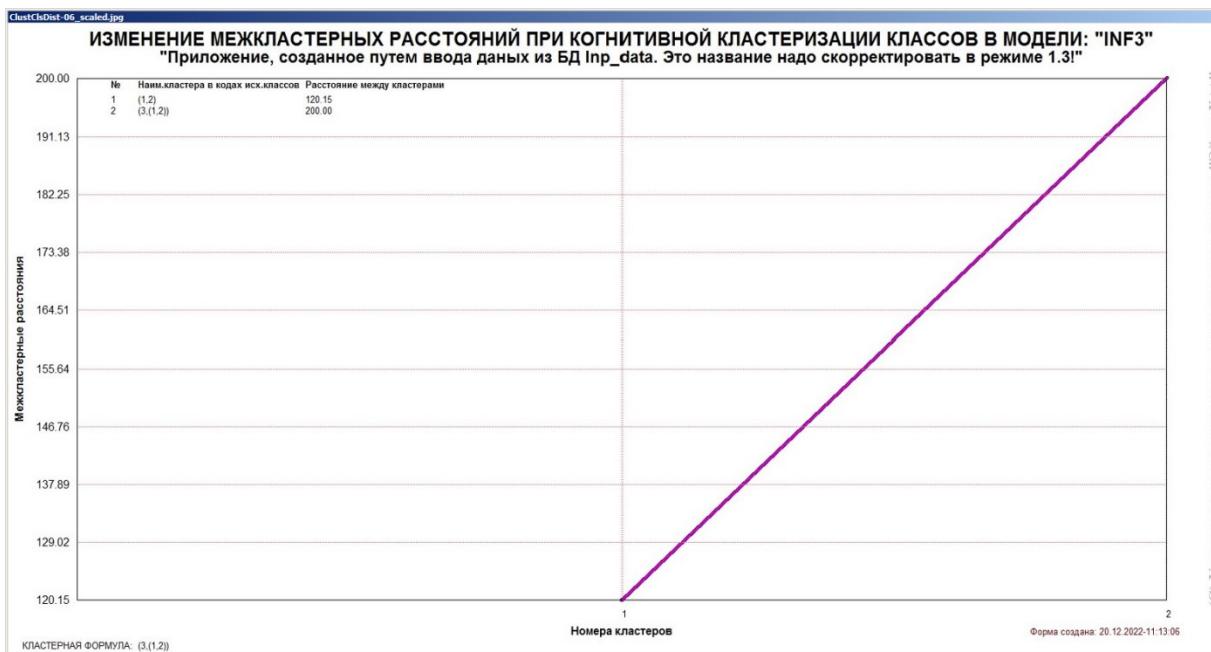


Рисунок 24. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации классов (режим 4.2.2.3)



**Рисунок 25. 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) рассчитывается матрица сходства признаков (таблица 15) по их смыслу и на основе этой матрицы рассчитывается и выводится четыре основных формы:

- круговая 2d-когнитивная диаграмма признаков (режим 4.3.2.2) (рисунок 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 межклusterных расстояний (режим 4.3.2.3) (рисунок 31).

Эта матрица сходства (таблица 15) 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: Пояснение по режиму

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 87

**Ok** **Cancel**

**4.3.2.1. Расчет матриц сходства, кластеров и конструктов признаков**

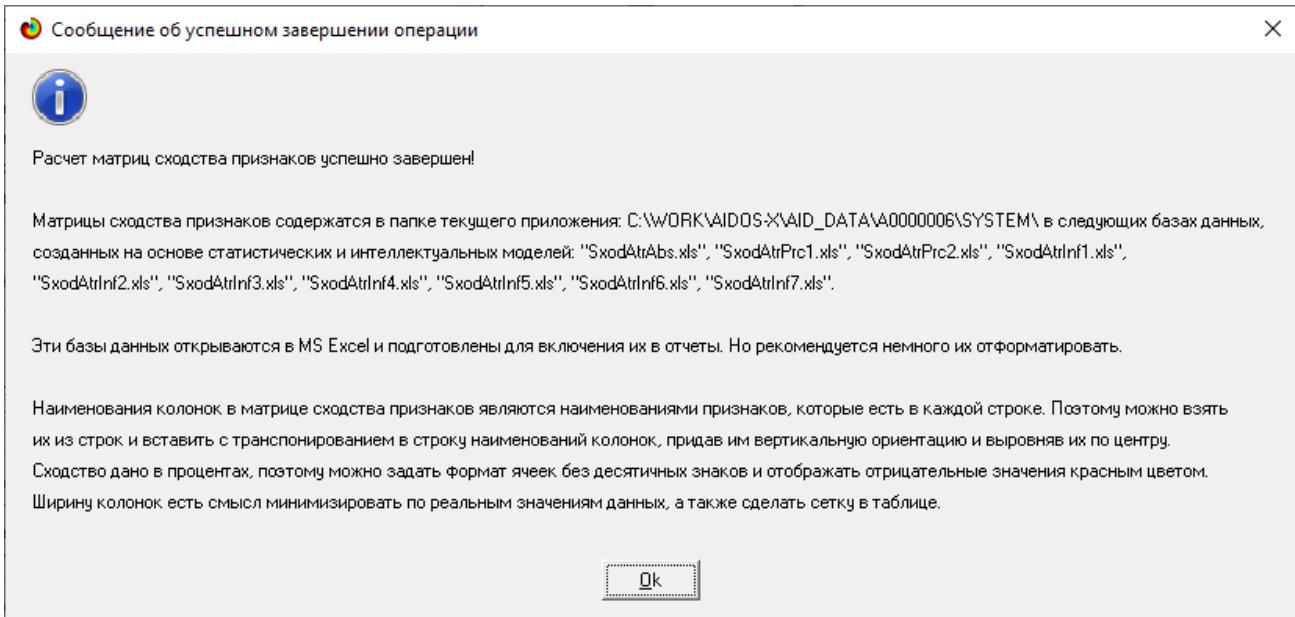
Стадии выполнения процесса:

1. Копирование БД.txt => БД.dbf в модели:10/10-Inf7
2. Расчет матрицы сходства признаков в модели: 10/10-Inf7
3. Расчет кластеров и конструктов признаков в модели: 10/10-Inf7
4. Физическая сортировка и дорасчет БД кластеров и конструктов признаков во всех моделях

**РАСЧЕТ МАТРИЦ СХОДСТВА, КЛАСТЕРОВ И КОНСТРУКТОВ ПРИЗНАКОВ ЗАВЕРШЕН УСПЕШНО !!!**

Прогноз времени исполнения

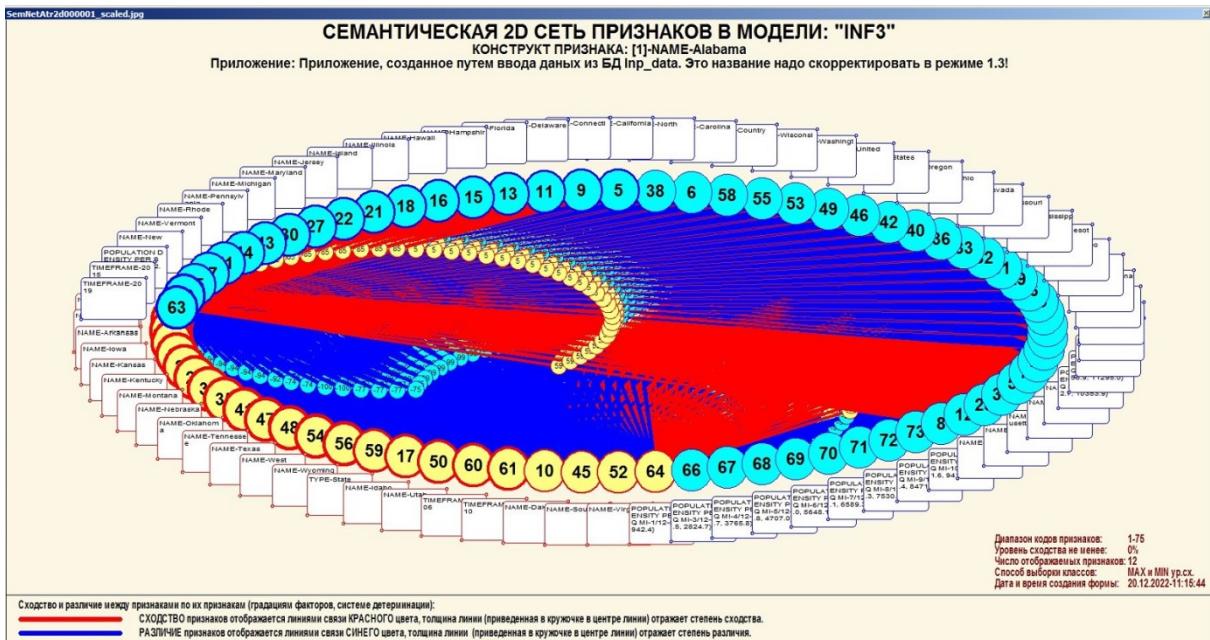
<b>Начало: 23:26:59</b>	<b>Окончание: 23:27:25</b>	<b>Ok</b>
<b>Прошло: 0:00:25</b>		<b>Осталось: 0:00:00</b>



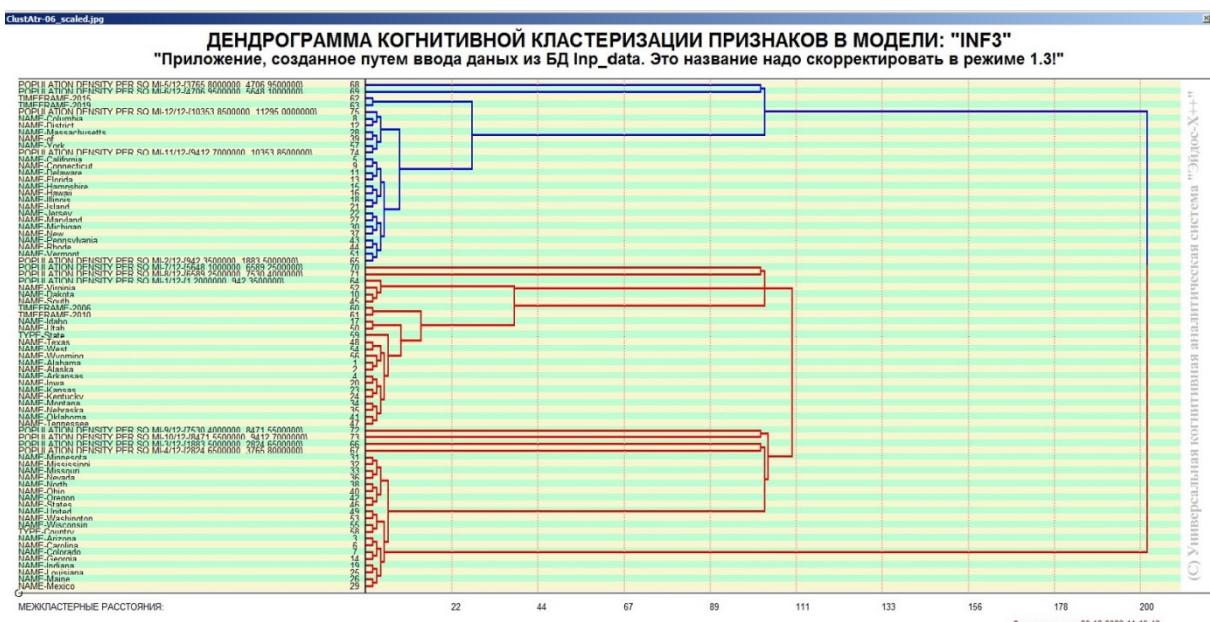
**Рисунок 26. Экранные формы режима 4.3.2.1, обеспечивающего расчет матриц сходства значений факторов**

**Таблица 6 – Матрица сходства значений факторов в СК-модели INF1 (полностью)**

A	B	C	D	E	F	G	H	I
4	3	1 NAME-Arizona	-39,1584878	-39,1584878	100,0000000	-39,1584878	-48,6197319	100,0000000
5	4	1 NAME-Arkansas	100,0000000	100,0000000	-39,1584878	100,0000000	-61,3677689	-39,1584916
6	5	1 NAME-California	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
7	6	1 NAME-Carolina	-39,1584916	-39,1584916	100,0000000	-39,1584916	-48,6197283	100,0000000
8	7	1 NAME-Colorado	-39,1584878	-39,1584878	100,0000000	-39,1584878	-48,6197319	100,0000000
9	8	1 NAME-Columbia	-28,0638328	-28,0638328	-77,3271063	-28,0638328	93,0048678	-77,3271037
10	9	1 NAME-Connecticut	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
11	10	1 NAME-Dakota	50,2450650	50,2450650	59,8808108	50,2450650	-99,0998078	59,8808075
12	11	1 NAME-Delaware	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
13	12	1 NAME-District	-28,0638328	-28,0638328	-77,3271063	-28,0638328	93,0048678	-77,3271037
14	13	1 NAME-Florida	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
15	14	1 NAME-Georgia	-39,1584878	-39,1584878	100,0000000	-39,1584878	-48,6197319	100,0000000
16	15	1 NAME-Hampshire	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
17	16	1 NAME-Hawaii	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
18	17	1 NAME-Idaho	92,5249238	92,5249238	-71,1378479	92,5249238	-26,8278774	-71,1378508
19	18	1 NAME-Illinois	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
20	19	1 NAME-Indiana	-39,1584878	-39,1584878	100,0000000	-39,1584878	-48,6197319	100,0000000
21	20	1 NAME-Iowa	100,0000000	100,0000000	-39,1584878	100,0000000	-61,3677689	-39,1584916
22	21	1 NAME-Island	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
23	22	1 NAME-Jersey	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283
24	23	1 NAME-Kansas	100,0000000	100,0000000	-39,1584878	100,0000000	-61,3677689	-39,1584916
25	24	1 NAME-Kentucky	100,0000000	100,0000000	-39,1584878	100,0000000	-61,3677689	-39,1584916
26	25	1 NAME-Louisiana	-39,1584878	-39,1584878	100,0000000	-39,1584878	-48,6197319	100,0000000
27	26	1 NAME-Maine	-39,1584878	-39,1584878	100,0000000	-39,1584878	-48,6197319	100,0000000
28	27	1 NAME-Maryland	-61,3677689	-61,3677689	-48,6197319	-61,3677689	100,0000000	-48,6197283



**Рисунок 27. Круговая 2d-когнитивная диаграмма значений факторов СК-модели INF1 (режим 4.3.2.2)**



**Рисунок 28. Агломеративная дендрограмма, полученная в результате когнитивной (истинной) кластеризации признаков (режим 4.3.2.3)**

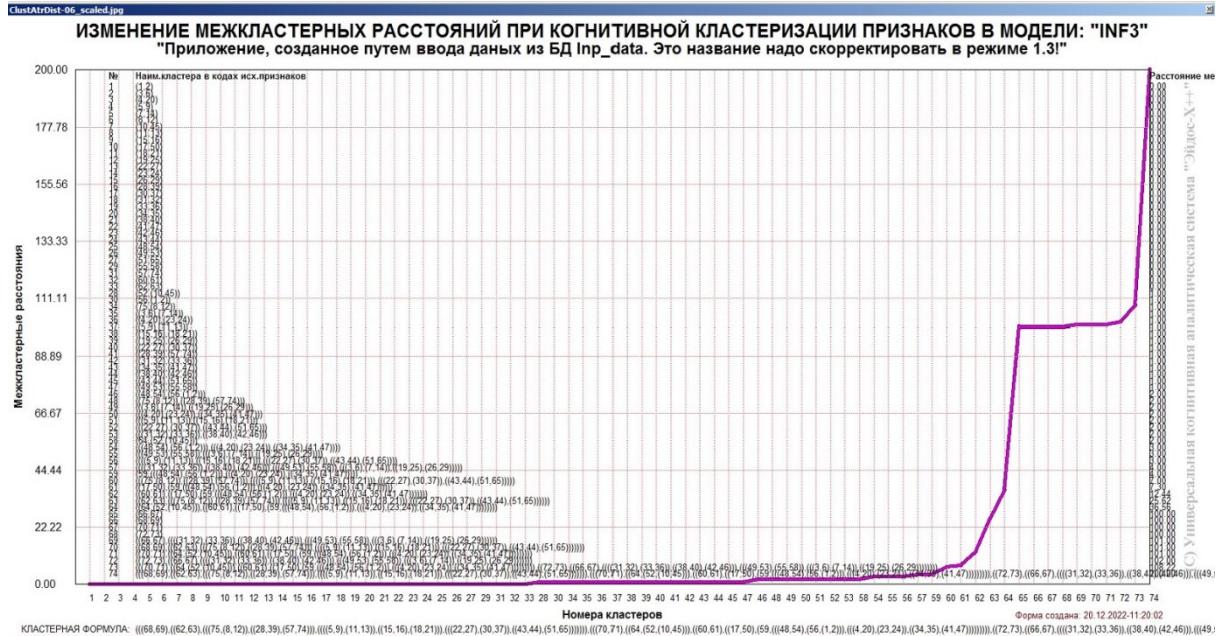


Рисунок 29. График изменения межкластерных расстояний (режим 4.3.2.3)

### 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.

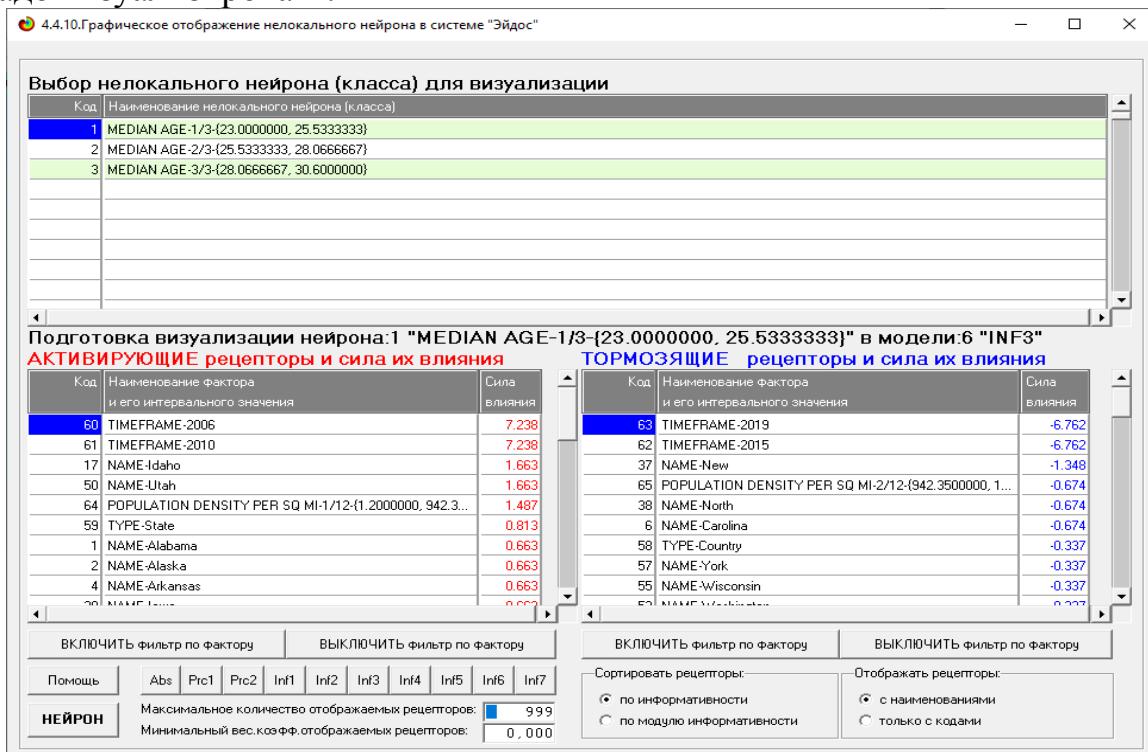
От нейросетевой модели представления знаний модель системы «Эйдос» отличается тем, что [20]:

- 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.



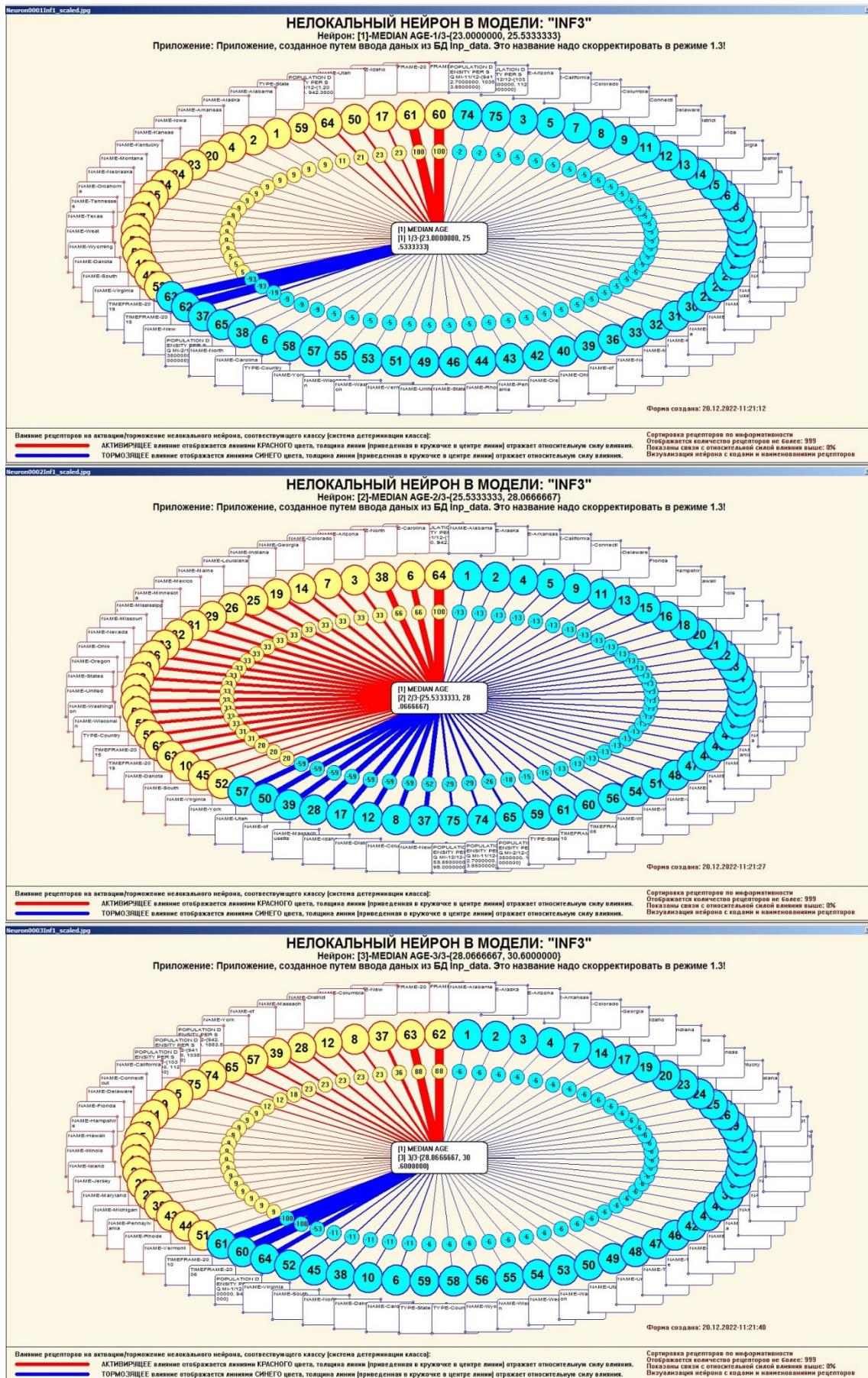


Рисунок 30. Примеры нелокальных нейронов, соответствующие классам

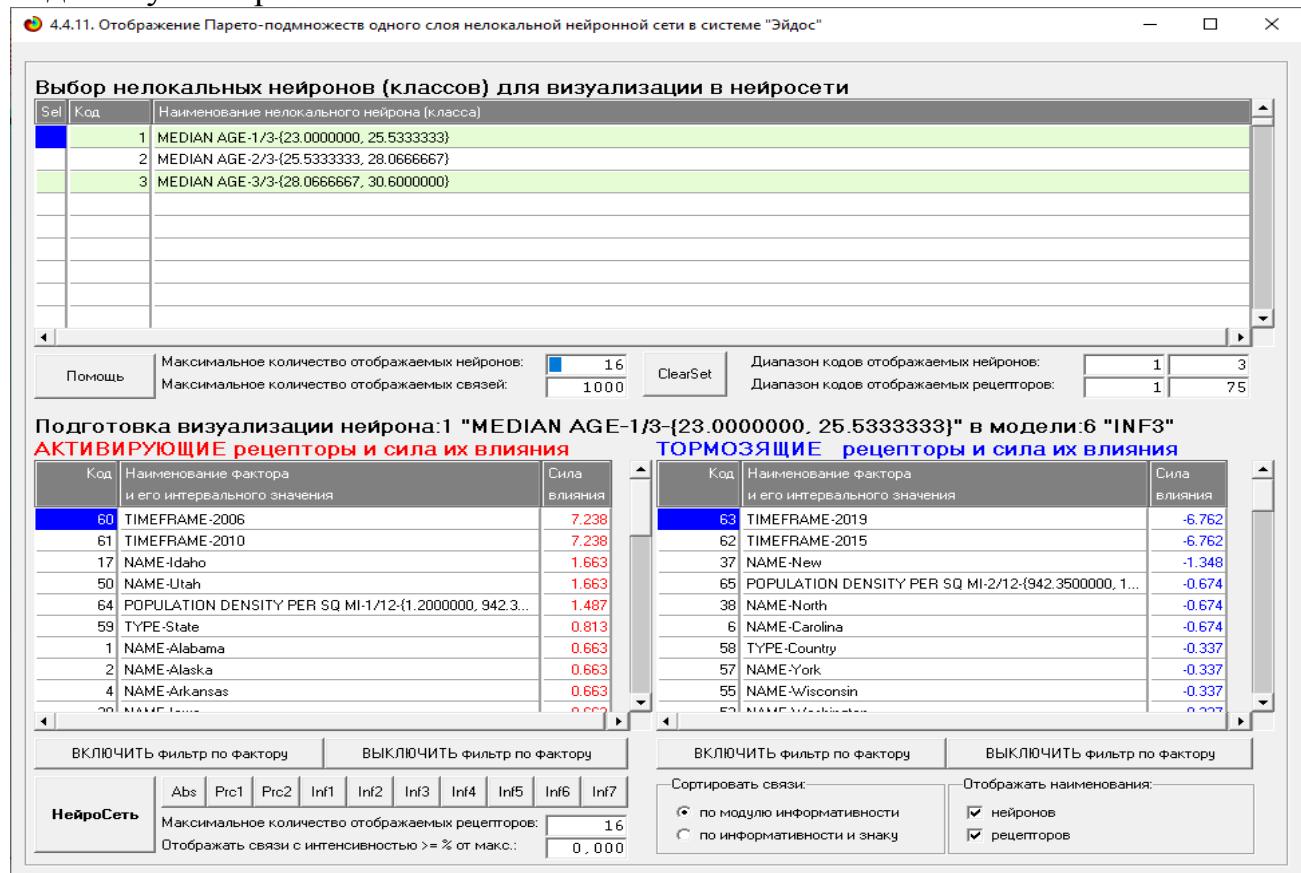
### 3.8.5. Нелокальная нейронная сеть

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.



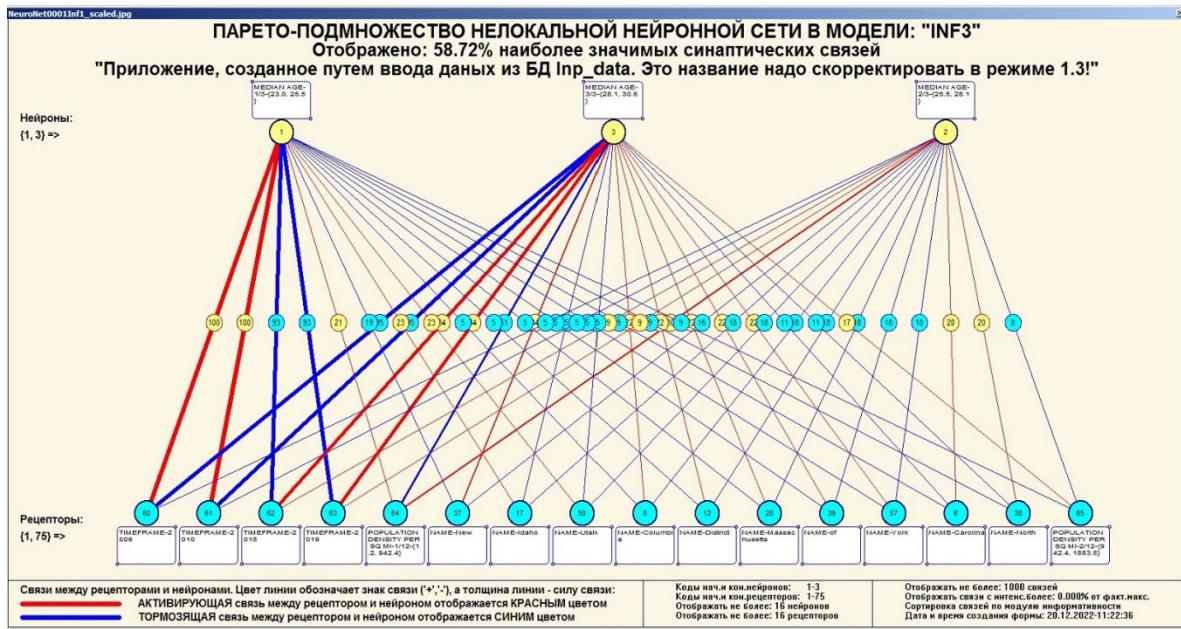
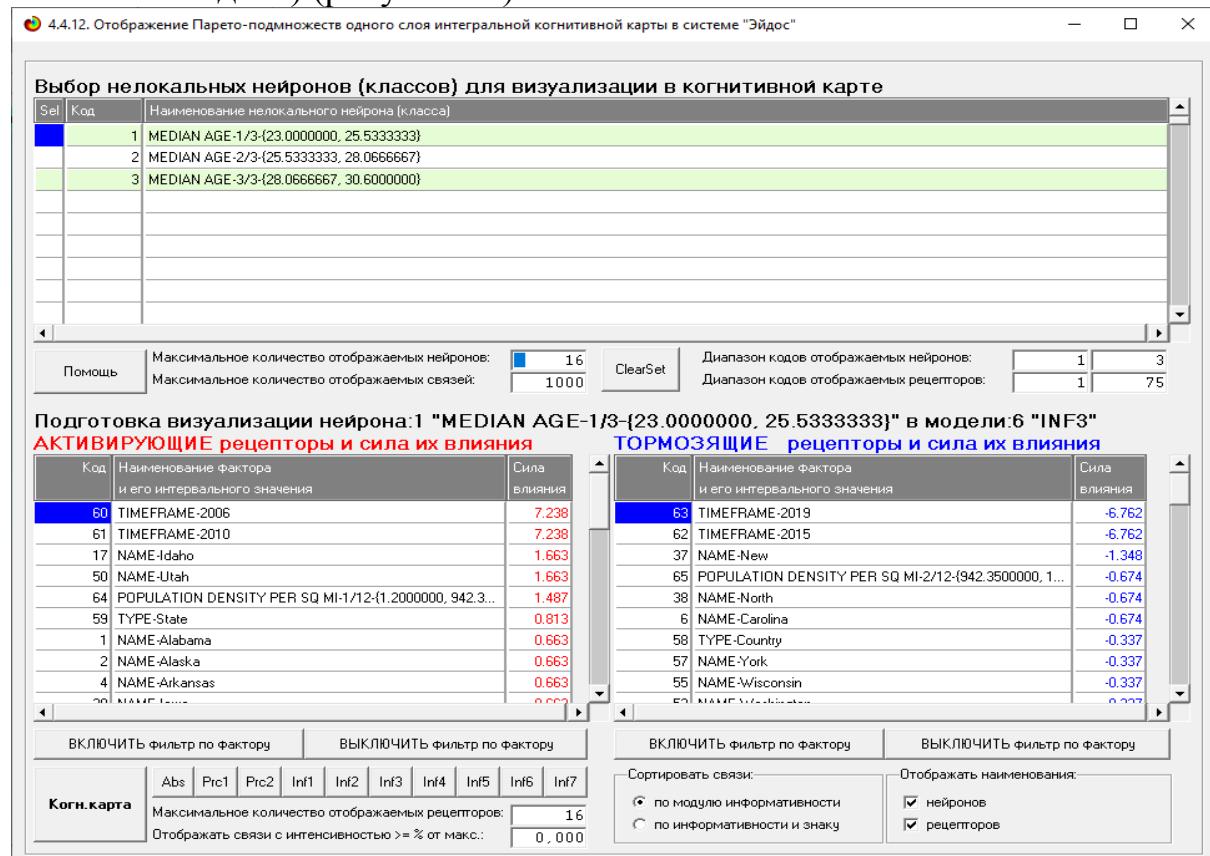


Рисунок 31. Нейронные сети в СК-модели INF3

### 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):



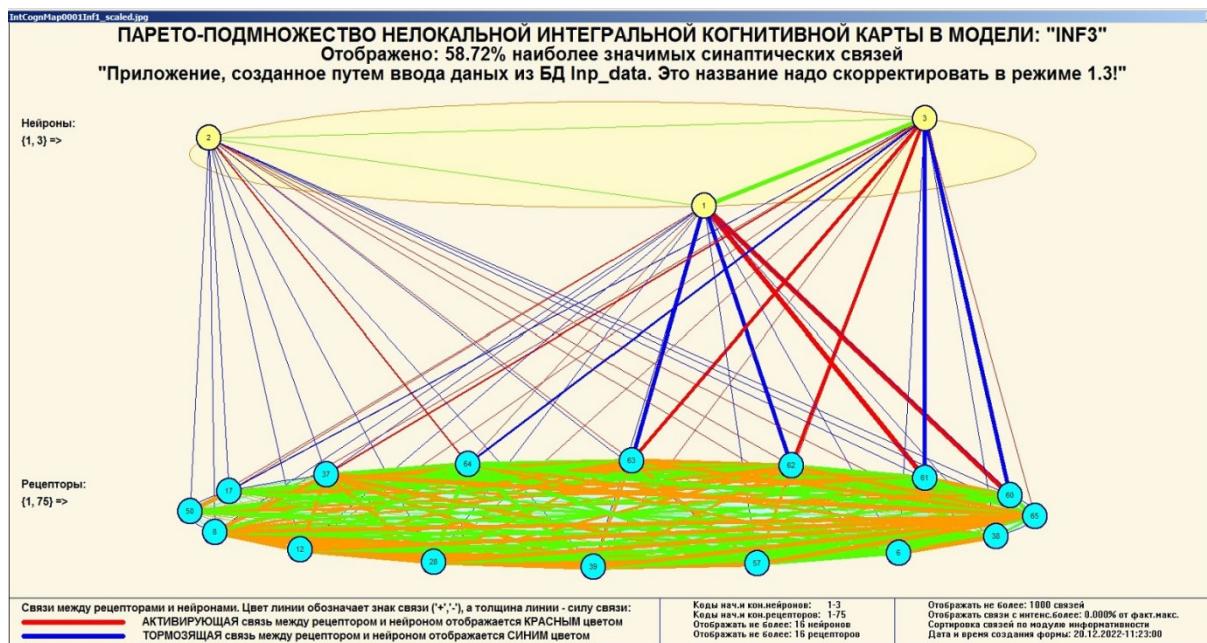


Рисунок 32. 3d-когнитивная диаграмма классов и признаков (режим 4.4.12)

### 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 521<sup>10</sup>. Позже об этом писалось в работе [3]<sup>11</sup> и ряде других работ автора, поэтому здесь подробнее рассматривать этот вопрос нецелесообразно.

**Пример опосредованных правдоподобных рассуждений.**

<sup>10</sup> [https://www.elibrary.ru/download/elibrary\\_18632909\\_64818704.pdf](https://www.elibrary.ru/download/elibrary_18632909_64818704.pdf), Таблица 7. 17, стр. 521

<sup>11</sup> <http://ej.kubagro.ru/2013/07/pdf/15.pdf>, стр.44.

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](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.

4.2.3. Когнитивные диаграммы классов. Задание параметров генерации выходных форм

Выбор классов для когнитивной диаграммы

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

Код	Наименование класса
0	ВСЕ КЛАССЫ
1	MEDIAN AGE-1/3-{23.0000000, 25.5333333}
2	MEDIAN AGE-2/3-{25.5333333, 28.0666667}
3	MEDIAN AGE-3/3-{28.0666667, 30.6000000}

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 ОПИСАТЕЛЬНЫЕ ШКАЛЫ	1	75
1	NAME	1	57
2	TYPE	58	59
3	TIMEFRAME	60	63
4	POPULATION DENSITY PER SQ MI	64	75

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 так количество отображаемых связей:

Помощь

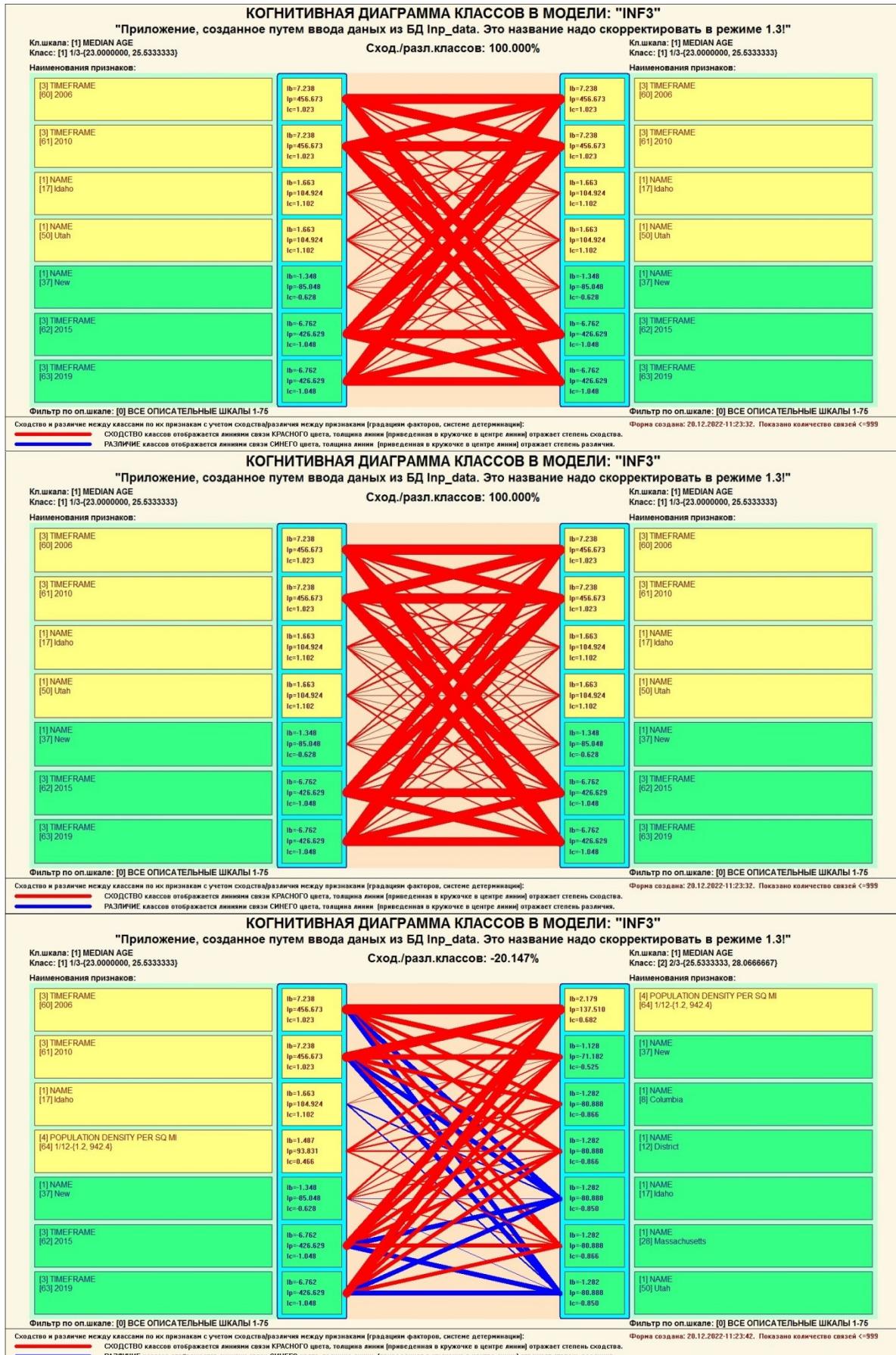
В диалоге заданы следующие параметры расчета когнитивных диаграмм:

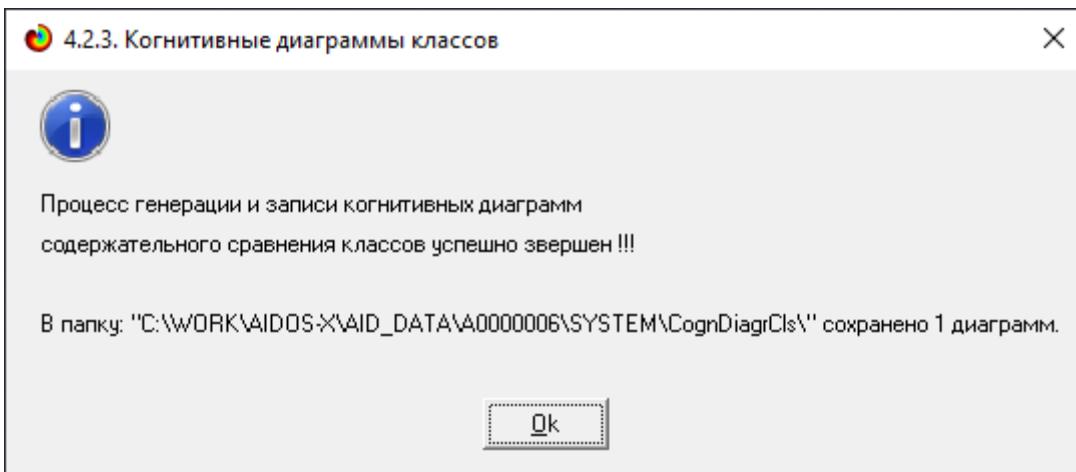
**Класс для левого инф.портрета:** [0] ВСЕ КЛАССЫ  
**Класс для правого инф.портрета:** [0] ВСЕ КЛАССЫ  
**Описат.шкала для левого инф.портрета:** [0] ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛЫ  
**Описат.шкала для правого инф.портрета:** [0] ВСЕ ОПИСАТЕЛЬНЫЕ ШКАЛЫ  
**Модели, заданные для расчета:** Abs, Prc1, Prc2, Inf1, Inf2, Inf3, Inf4, Inf5, Inf6, Inf7

Задайте режим вывода когнитивных диаграмм:

Показать все диаграммы с остановкой  
 Записать все диаграммы без показа

Ok Cancel





**Рисунок 33. 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. Задание параметров генерации выходных форм**

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

Код	Наименование признака
0	ВСЕ ПРИЗНАКИ
1	NAME-Alabama
2	NAME-Alaska
3	NAME-Arizona
4	NAME-Arkansas
5	NAME-California
6	NAME-Carolina

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

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

Код	Наименование классификационной шкалы	Минимальный код градации	Максимальный код градации
0	ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ	1	3
1	MEDIAN AGE	1	3

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

Задайте модели, в которых проводить расчеты когнитивных диаграмм:  
 Abs    Prc1    Prc2    Inf1    Inf2    Inf3    Inf4    Inf5    Inf6    Inf7

Задайте max количество отображаемых связей: 99999

Задайте режим вывода когнитивных диаграмм:  
 Показать все диаграммы с остановкой  
 Записать все диаграммы без показа

Ok   Cancel

**КОГНИТИВНАЯ ДИАГРАММА ПРИЗНАКОВ В МОДЕЛИ: "INF3"**

"Приложение, созданное путем ввода данных из БД Inp\_data. Это название надо скорректировать в режиме 1.3!"

Оп.шкала: [1] NAME  
Признак: [1] Alabama  
Наименования классов:

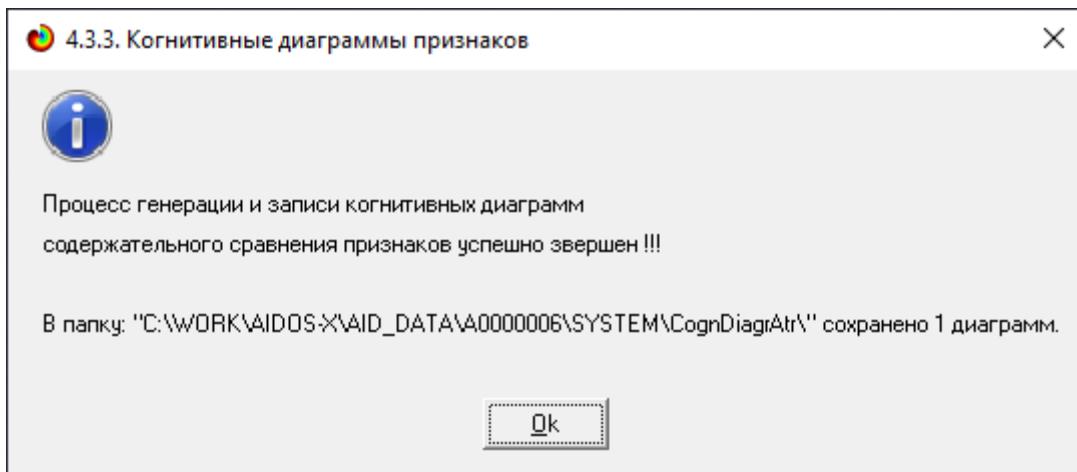
[1] MEDIAN AGE [1] 1/3-(23.0, 25.5)	[1] MEDIAN AGE [2] 2/3-(25.5, 28.1)	[1] MEDIAN AGE [3] 3/3-(28.1, 30.6)	[1] MEDIAN AGE [1] 1/3-(23.0, 25.5)	[1] MEDIAN AGE [2] 2/3-(25.5, 28.1)	[1] MEDIAN AGE [3] 3/3-(28.1, 30.6)
lb=0.663 lp=41.831 lc=0.386	lb=0.282 lp=17.795 lc=0.404	lb=-0.381 lp=-24.035 lc=-0.222	lb=0.663 lp=41.831 lc=0.386	lb=0.292 lp=17.795 lc=0.404	lb=0.381 lp=24.035 lc=0.222

Сход./разл.признаков: 100.000%      Оп.шкала: [1] NAME  
Признак: [1] Alabama  
Наименования классов:

Фильтр по кл.шкале: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ 1-3      Фильтр по кл.шкале: [0] ВСЕ КЛАССИФИКАЦИОННЫЕ ШКАЛЫ 1-3

Сходство и различие между признаками по их классам с учетом сходства/различия между классами по системе детерминирующих их признаков:  
  
 СХОДСТВО признаков отображается линии связи КРАСНОГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень сходства.  
 РАЗЛИЧИЕ признаков отображается линии связи СИНЕГО цвета, толщина линии [приведенная в кружочке в центре линии] отражает степень различия.

Форма создана: 20.12.2022-11:26:28. Показано количество связей <99999



**Рисунок 34. 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 121 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 логики 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].



## 4.5. Визуализация когнитивных функций



Что такое когнитивная функция:

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

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

<http://ej.kubagro.ru/2011/03/pdf/18.pdf>

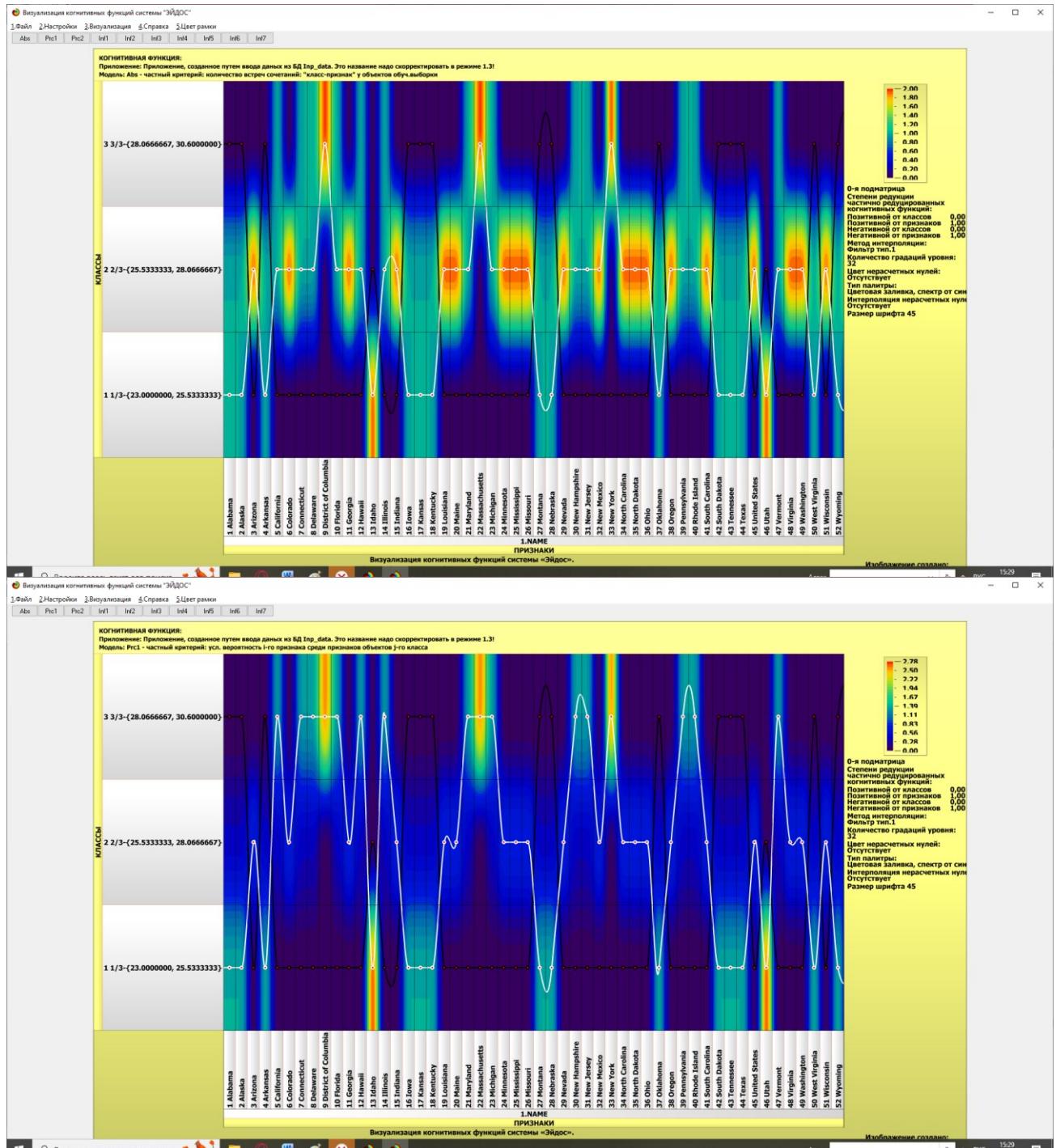
Задайте нужный режим:

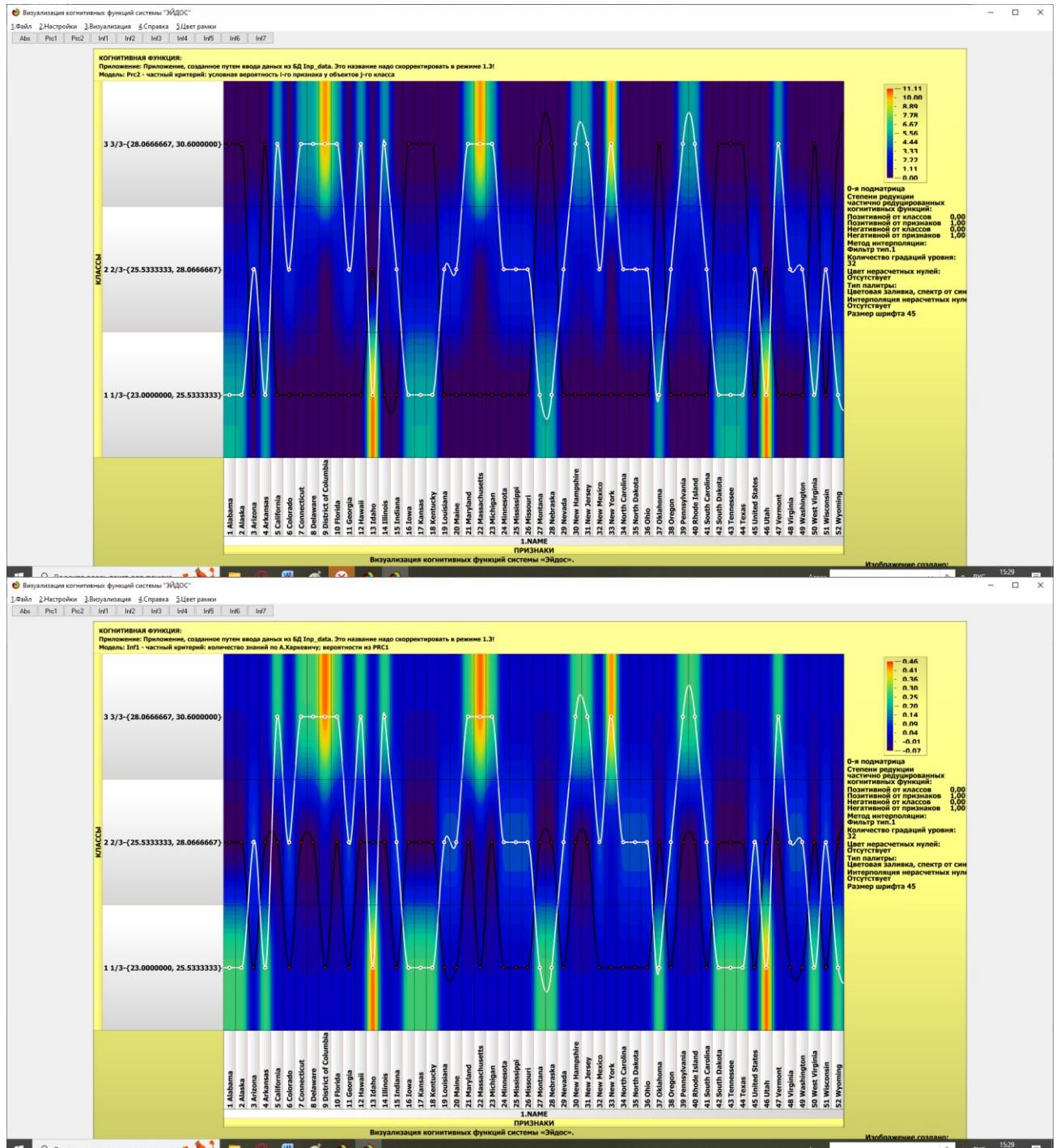
**Визуализации когнитивных функций**

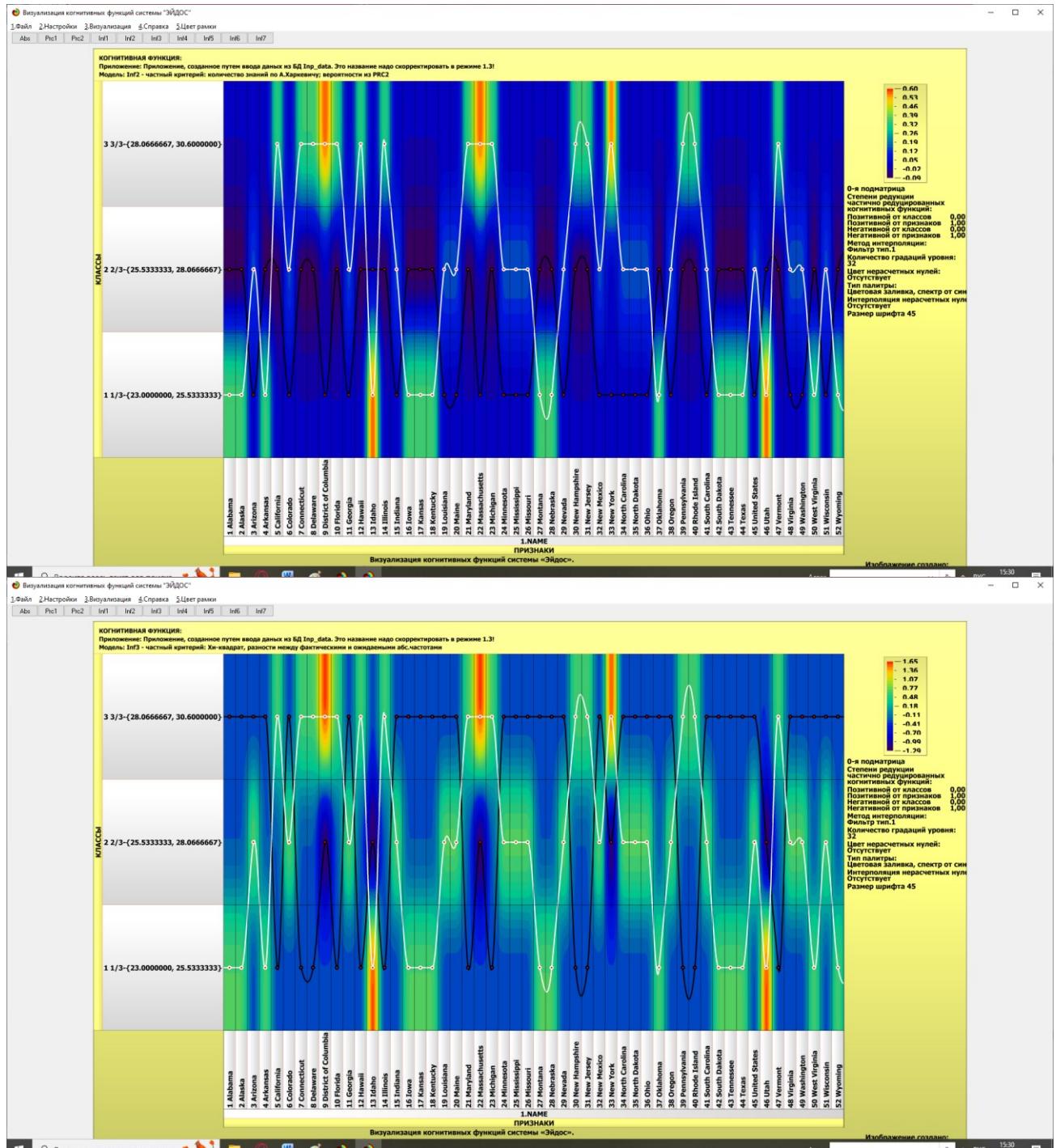
Литератур.ссылки на работы по когнитивным функциям

Литератур.ссылки на работы по когнитивным функциям

Литератур.ссылки на работы по управлению знаниями







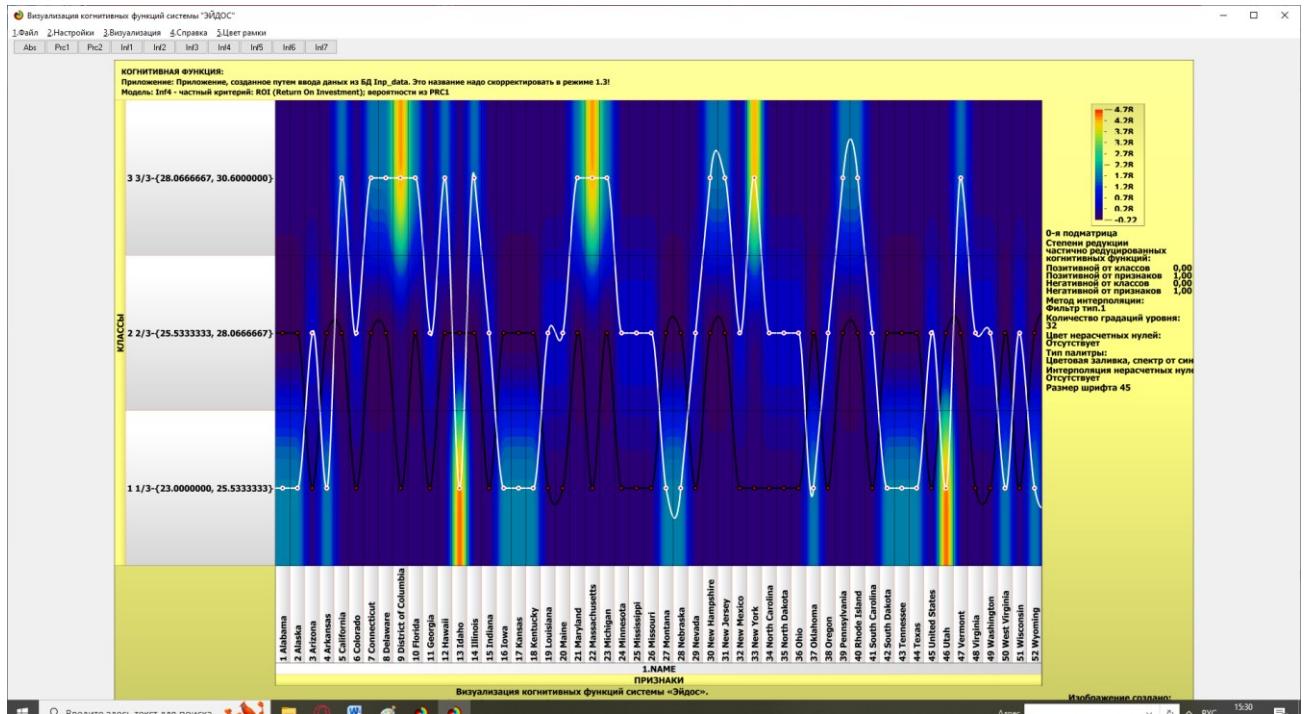


Рисунок 35. Примеры когнитивных функций в СК-модели INF1-INF7

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

### 3.8.9. Значимость описательных шкал и их градаций

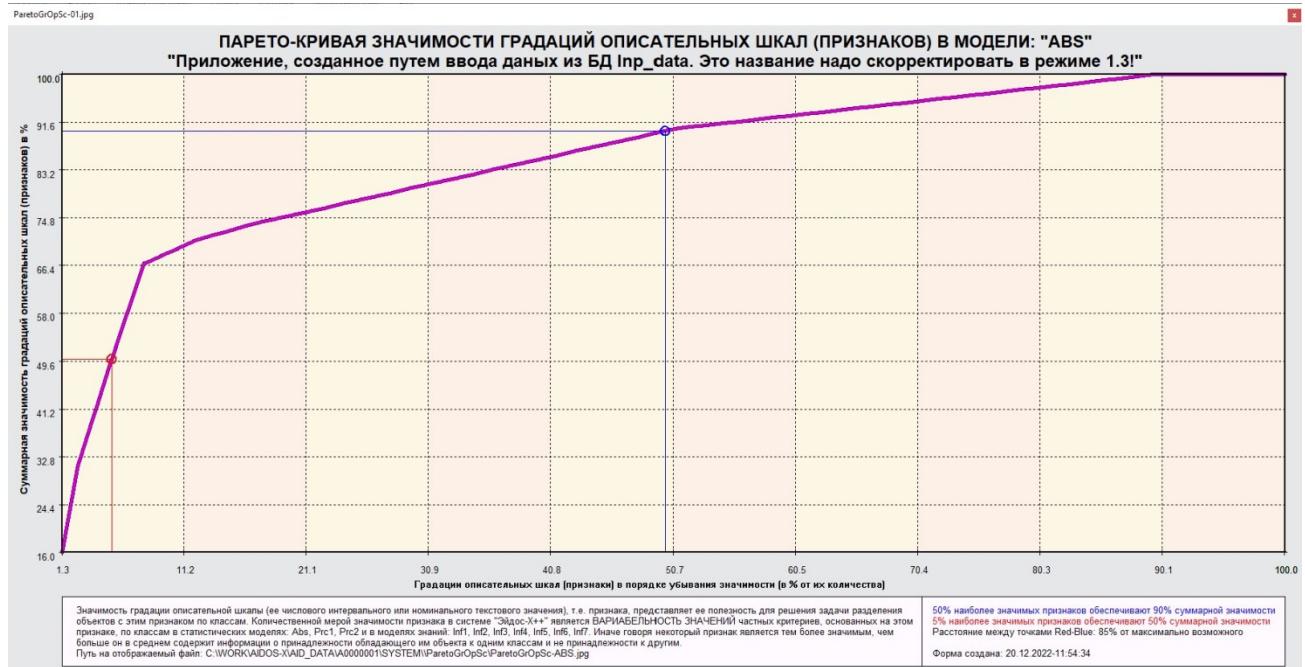
В АСК-анализе все факторы рассматриваются с одной единственной точки зрения: сколько информации содержится в их значениях о переходе объекта моделирования и управления, на который они действуют, в определенное будущее состояние, описываемое классом (градация классификационной шкалы), и при этом сила 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:



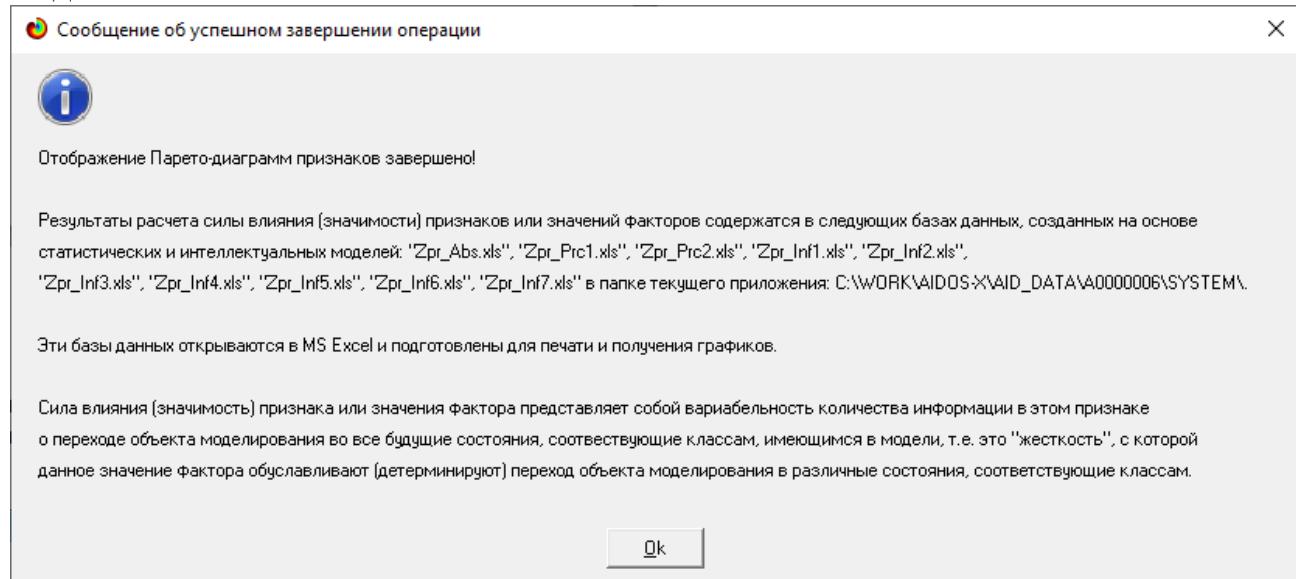
**Рисунок 36. Парето-кривая силы влияния значений факторов на поведение объекта моделирования в СК-модели INF3**

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.

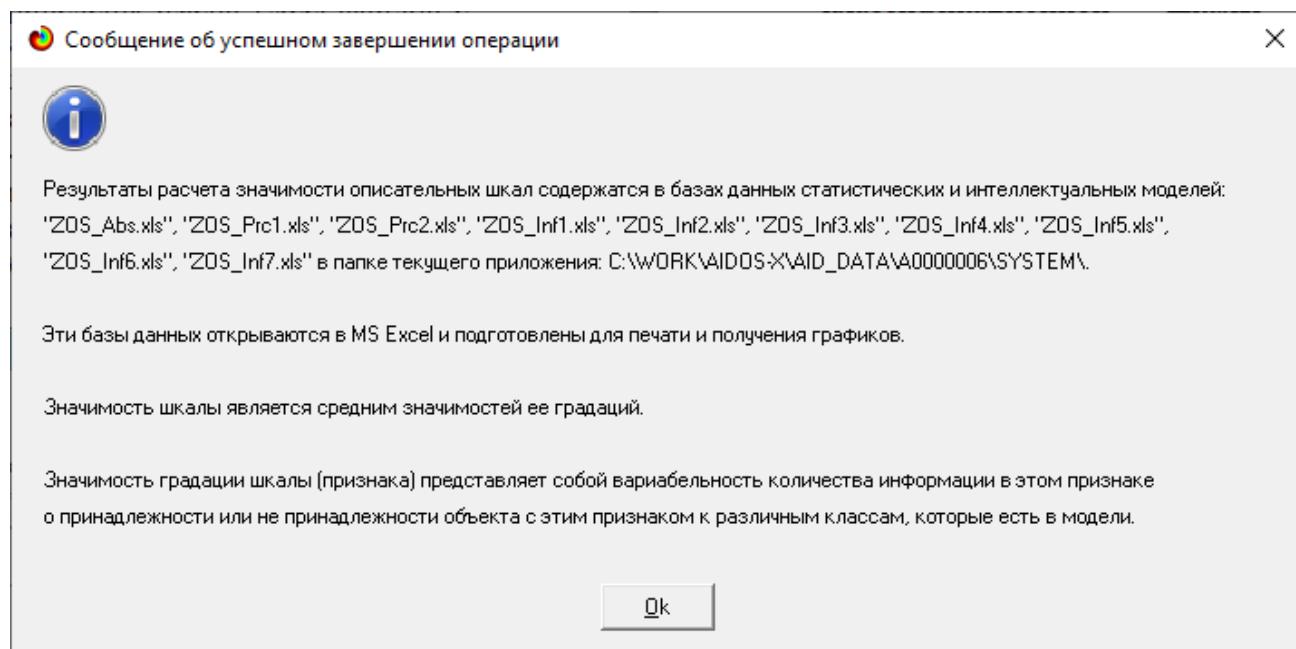
**Таблица 7 – Сила влияния значений факторов на поведение объекта моделирования в СК-модели INF3**

	A	B	C	D	E	F	G	H	I	J
1	NUM	NUM_PRC	KOD_ATR	NAME_ATR	KOD_OPSC	ZNACH_ATR	ZN_ATRNIT	ZNACH_PRC	ZN_PRCNIT	DELETE
2	1	1,3333333	60	TIMEFRAME-2006	3	7,0773184	7,0773184	9,2200609	9,2200609	
3	2	2,6666667	61	TIMEFRAME-2010	3	7,0773184	14,1546368	9,2200609	18,4401219	
4	3	4,0000000	62	TIMEFRAME-2015	3	6,4544453	20,6090821	8,4086056	26,8487275	
5	4	5,3333333	63	TIMEFRAME-2019	3	6,4544453	27,0635274	8,4086056	35,2573330	
6	5	6,6666667	64	POPULATION DENSITY PER SQ MI-1/12-[1.2000000, 942.3500000]	4	3,1942379	30,2577653	4,1613315	39,4186646	
7	6	8,0000000	37	NAME-New	1	2,1472576	32,4050229	2,7973654	42,2160300	
8	7	9,3333333	17	NAME-Idaho	1	1,5090331	33,9140560	1,9659109	44,1819409	
9	8	10,6666667	50	NAME-Utah	1	1,5090331	35,4230891	1,9659109	46,1478517	
10	9	12,0000000	8	NAME-Columbia	1	1,4796177	36,9027068	1,9275896	48,0754413	
11	10	13,3333333	12	NAME-District	1	1,4796177	38,3823245	1,9275896	50,0030309	
12	11	14,6666667	28	NAME-Massachusetts	1	1,4796177	39,8619422	1,9275896	51,9306233	
13	12	16,0000000	39	NAME-of	1	1,4796177	41,3415599	1,9275896	53,8582101	
14	13	17,3333333	57	NAME-York	1	1,4796177	42,8211776	1,9275896	57,0579977	
15	14	18,6666667	6	NAME-Carolina	1	1,2443003	44,0654779	1,6210271	57,4068268	
16	15	20,0000000	38	NAME-North	1	1,2443003	45,3097782	1,6210271	59,0278539	
17	16	21,3333333	65	POPULATION DENSITY PER SQ MI-2/12-[942.3500000, 1883.500000]	4	1,0736288	46,3834070	1,39686827	60,4265366	
18	17	22,6666667	74	POPULATION DENSITY PER SQ MI-11/12-[9412.7000000, 10353.8500000]	4	0,7398088	47,1232158	0,9637947	61,3903313	
19	18	24,0000000	75	POPULATION DENSITY PER SQ MI-12/12-[10353.8500000, 11295.0000000]	4	0,7398088	47,8630246	0,9637947	62,3541261	
20	19	25,3333333	59	TYPE-State	2	0,7045833	48,5676079	0,9179043	63,2720303	
21	20	26,6666667	10	NAME-Dakota	1	0,6621126	49,2297205	0,8625751	64,1346054	
22	21	28,0000000	45	NAME-South	1	0,6621126	49,8918331	0,8625751	64,9971805	
23	22	29,3333333	52	NAME-Virginia	1	0,6621126	50,5594547	0,8625751	65,8597556	
24	23	30,6666667	3	NAME-Arizona	1	0,6221501	51,1760598	0,8105135	66,6702691	
25	24	32,0000000	7	NAME-Colorado	1	0,6221501	51,7982459	0,8105135	67,4807825	
26	25	33,3333333	14	NAME-Georgia	1	0,6221501	52,4203960	0,8105135	68,2912960	
27	26	34,6666667	19	NAME-Indiana	1	0,6221501	53,0425461	0,8105135	69,1018094	
28	27	36,0000000	25	NAME-Louisiana	1	0,6221501	53,6646962	0,8105135	69,9123229	
29	28	37,3333333	26	NAME-Maine	1	0,6221501	54,2868463	0,8105135	70,7228364	
30	29	38,6666667	29	NAME-Mexico	1	0,6221501	54,9089964	0,8105135	71,5333498	
31	30	40,0000000	31	NAME-Minnesota	1	0,6221501	55,5311465	0,8105135	72,3438633	
32	31	41,3333333	32	NAME-Mississippi	1	0,6221501	56,1532966	0,8105135	73,1543768	
33	32	42,6666667	33	NAME-Missouri	1	0,6221501	56,7754467	0,8105135	73,9648902	
34	33	44,0000000	36	NAME-Nevada	1	0,6221501	57,3975968	0,8105135	74,7754037	
35	34	45,3333333	40	NAME-Ohio	1	0,6221501	58,0197469	0,8105135	75,5859171	
36	35	46,6666667	42	NAME-Oregon	1	0,6221501	58,6418970	0,8105135	76,3964306	
37	36	48,0000000	46	NAME-States	1	0,6221501	59,2640471	0,8105135	77,2069441	

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



**Рисунок 37. Имена Excel-файлов с информацией о силе влияния значений факторов в разных моделях**



**Рисунок 38. имена Excel-файлов с информацией о силе влияния факторов в разных моделях**

**Таблица 8 – Сила влияния факторов на поведение объекта моделирования в системно-когнитивной модели INF3**

A	B	C	D	E	F	G	H	I	J	K	L
NUM	NUM_PRC	KOD_OPSC	NAME_OPSC	N_GROPSC	KODGR_MIN	KODGR_MAX	ZNACH_OS	ZN_OSNIT	ZNACH_PRC	ZN_PRCNIT	DELETE
1	1	25,000000	3 TIMEFRAME		4	60	63	6,7658819	6,7658819	78,1643915	78,1643915
2	2	50,000000	1 NAME		57	1	57	0,7477588	7,5136407	8,6386538	86,8030453
3	3	75,000000	2 TYPE		2	58	59	0,6633667	8,1770074	7,6636949	94,4667402
4	4	100,000000	4 POPULATION DENSITY PER SQ MI		12	64	75	0,4789570	8,6559644	5,5332598	100,0000000

### 3.8.10. 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:



Сообщение об успешном завершении операции X

**i**

Результаты расчета значимости классификационных шкал содержатся в базах данных статистических и интеллектуальных моделей: "ZCS\_Abs.xls", "ZCS\_Prc1.xls", "ZCS\_Prc2.xls", "ZCS\_Inf1.xls", "ZCS\_Inf2.xls", "ZCS\_Inf3.xls", "ZCS\_Inf4.xls", "ZCS\_Inf5.xls", "ZCS\_Inf6.xls", "ZCS\_Inf7.xls" в папке текущего приложения: C:\WORK\VAIDOS\VAID\_DATA\A0000006\SYSTEM\.

Эти базы данных открываются в MS Excel и подготовлены для печати и получения графиков.

Значимость классификационной шкалы является средним значимостей ее градаций, т.е. классов.

Значимость градации классификационной шкалы, т.е. класса, представляет собой вариабельность количества информации в во всех признаках модели о принадлежности или не принадлежности объекта с этим признаками к данному классу.

Значимость градации классификационной шкалы (класса) - это степень детерминированности этого класса (см.режим 3.7.3).

Сообщение об успешном завершении операции X

**i**

Отображение Парето-диаграмм классов завершено!

Результаты расчета степени детерминированности (значимости) классов содержатся в следующих базах данных, созданных на основе статистических и интеллектуальных моделей: "Zkl\_Abs.xls", "Zkl\_Prc1.xls", "Zkl\_Prc2.xls", "Zkl\_Inf1.xls", "Zkl\_Inf2.xls", "Zkl\_Inf3.xls", "Zkl\_Inf4.xls", "Zkl\_Inf5.xls", "Zkl\_Inf6.xls", "Zkl\_Inf7.xls" в папке текущего приложения: C:\WORK\VAIDOS\VAID\_DATA\A0000006\SYSTEM\.

Эти базы данных открываются в MS Excel и подготовлены для печати и получения графиков.

Степень детермированности класса представляет собой вариабельность количества информации в всех признаках модели о принадлежности или не принадлежности объекта с этими признаком к данному классу, т.е. это "жесткость", с которой значения факторов обуславливают (детерминируют) переход объекта моделирования в состояние, соответствующее классу.

**Рисунок 39. Экранные формы режимов 3.7.2 и 3.7.3 системы «Эйдос»**

**Таблица 9 – Степень детерминированности классов в СК-модели ABS**

	A	B	C	D	E	F	G	H	I	J	K	L
1	NUM	NUM_PRC	KOD_CLSC	NAME_CLSC	N_GRLSC	KODGR_MIN	KODGR_MAX	ZNACH_CS	ZN_CSNT	ZNACH_PRC	ZN_PRCNT	DELETE
2	1	100,000000		1 MEDIAN AG		3	1	3	1,3775867	1,3775867	100,000000	100,000000

**Таблица 10 – Степень детерминированности классификационных шкал в системно-когнитивной модели INF3**

	A	B	C	D	E	F	G	H	I	J
1	NUM	NUM_PRC	KOD_CLS	NAME_CLS	KOD_CLSC	ZNACH_CLS	ZN_CLSNIT	ZNACH_PRC	ZN_PRCNIT	DELETE
2		1	33,333333	3 MEDIAN AGE-3/3-[28.0666667, 30.6000000]		1	1,7180354	1,7180354	41,5711369	41,5711369
3		2	66,6666667	1 MEDIAN AGE-1/3-[23.0000000, 25.5333333]		1	1,7167805	3,4348159	41,5407722	83,1119091
4		3	100,000000	2 MEDIAN AGE-2/3-[25.5333333, 28.0666667]		1	0,6979443	4,1327602	16,8880909	100,0000000

#### **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.

Рекомендуется ввести классификационные шкалы, отражающие влияние исследуемых факторов на объект моделирования не только в натуральном выражении, но и в стоимостном выражении. Перспективность и ценность результатов подобных исследований и

разработок для теории и практики не вызывает особых сомнений, что подтверждается работами автора в этой области.

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](http://lc.kubagro.ru/aidos/_Aidos-X.htm), а затем в диспетчере приложений (режим 1.3) установить интеллектуальное облачное Эйдос-приложение №342. 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](http://lc.kubagro.ru/aidos/How_to_make_your_own_cloud_Eidos-application.pdf).

## **5. CONCLUSIONS (ВЫВОДЫ)**

В работе решена задача выявления зависимости возраста при первом браке от индекса проживания, района проживания и плотности населения. На основе знания этих зависимостей решены задачи прогнозирования, принятия решений и исследования моделируемой предметной области путем исследования ее системно-когнитивной модели.

Спецификой данной задачи является то, что все независимые переменные являются лингвистическими (категориальными) переменными. Поэтому для решения данной задачи применяется лингвистический АСК-анализ, т.е. когнитивная математическая лингвистика. При этом сами показатели качества при выявлении возраста измеряется 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 [6].

Это достигается путем вычисления количества информации, содержащегося в градациях номинальных шкал и получении той или иной урожайности.

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

## REFERENCES (ЛИТЕРАТУРА)

1. Работы проф.Е.В.Луценко & С° по тематике, связанной с АПК, частности с когнитивной агрономией: [http://lc.kubagro.ru/aidos/Work\\_with\\_agricultural.htm](http://lc.kubagro.ru/aidos/Work_with_agricultural.htm)
2. Луценко, Е. В. Автоматизированный системно-когнитивный анализ в управлении активными объектами : (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем) / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2002. – 605 с. – ISBN 5-94672-020-1. – EDN OCZFHC.
3. Орлов, А. И. Системная нечеткая интервальная математика / А. И. Орлов, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет, 2014. – 600 с. – ISBN 978-5-94672-757-0. – EDN RZXZ.
4. 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](http://lc.kubagro.ru/aidos/Works_on_ASK-analysis_of_texts.htm)
5. 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](http://lc.kubagro.ru/aidos/Work_on_emergence.htm)
6. Луценко, Е. В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2013. – № 92. – С. 61-71. – EDN RNEGHR.
7. Сайт Е.В.Луценко: <http://lc.kubagro.ru/>.
8. Страница Е.В.Луценко в РесечГейт <https://www.researchgate.net/profile/Eugene-Lutsenko>
9. Страница Е.В.Луценко в РИНЦ: [https://elibrary.ru/author\\_profile.asp?id=123162](https://elibrary.ru/author_profile.asp?id=123162).
10. Кравченко Р.В. Влияние основной обработки почвы на ее агро-физические показатели в посевах сои / Р.В. Кравченко, Г.А. Дубовой // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2022. – №05(179). С. 320 – 331. – IDA [article ID]: 1792205021. – Режим доступа: <http://ej.kubagro.ru/2022/05/pdf/21.pdf>, 0,75 у.п.л.
11. Горпинченко, К. Н. Прогнозирование и принятие решений по выбору агротехнологий в зерновом производстве с применением методов искусственного интеллекта (на примере СК-анализа) / К. Н. Горпинченко, Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2013. – 168 с. – ISBN 978-5-94672-644-3. – EDN RAIMQL.
12. Луценко, Е. В. Инвариантное относительно объемов данных нечеткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергена в АСК-анализе и системе "Эйдос" / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 126. – С. 1-32. – DOI 10.21515/1990-4665-126-001. – EDN XXXBDV.
13. Луценко, Е. В. Сценарный и спектральный автоматизированный системно-когнитивный анализ / Е. В. Луценко. – Краснодар : Кубанский государственный аграрный университет им. И.Т. Трубилина, 2021. – 288 с. – DOI 10.13140/RG.2.2.22981.37608. – EDN ZQLITW.
14. 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.

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

16. Луценко, Е. В. Автоматизация функционально-стоимостного анализа и метода "Директ-костинг" на основе ACK-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) / Е. В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2017. – № 131. – С. 1-18. – DOI 10.21515/1990-4665-131-001. – EDN ZRXVFN.

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

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

19. Луценко, Е. В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе "Эйдос") / Е. В. Луценко, В. Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2011. – № 71. – С. 27-74. – EDN OIGYBB.

20. 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.

21. 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.

22. 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](http://lc.kubagro.ru/aidos/Works_on_identification_presentation_and_use_of_knowledge.htm)

23. Пойа Дерьдь. 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/rassuzhdlenija.htm>

24. 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.

25. Работы проф.Е.В.Луценко & C° по когнитивным функциям:  
[http://lc.kubagro.ru/aidos/Works\\_on\\_cognitive\\_functions.htm](http://lc.kubagro.ru/aidos/Works_on_cognitive_functions.htm)

26. 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.