Математические методы исследования

УДК 519.24

МАТЕМАТИЧЕСКИЙ МЕТОД АСК-АНАЛИЗА — СИСТЕМНАЯ ТЕОРИЯ ИНФОРМАЦИИ

© Е. В. Луценко[1]

Статья поступила 25 июля 2014 г.

Рассмотрен математический метод автоматизированного системно-когнитивного анализа (АСК-анализа), реализованный в его программном инструментарии — универсальной когнитивной аналитической системе «Эйдос-X++». Математический метод АСК-анализа основан на системной теории информации (СТИ), которая создана в рамках реализации программной идеи обобщения всех понятий математики, в частности теории информации, базирующихся на теории множеств, путем тотальной замены понятия множества на более общее понятие системы и тщательного отслеживания всех последствий этой замены. Математический метод, положенный в основу АСК-анализа, является непараметрическим и позволяет сопоставимо обрабатывать десятки и сотни тысяч градаций факторов и будущих состояний объекта управления (классов) при неполных (фрагментированных), зашумленных данных числовой и нечисловой природы, измеряемых в различных единицах измерения.

Ключевые слова: математический метод; автоматизированный системно-когнитивный анализ; АСК-анализ; системная теория информации; интеллектуальная система «Эйдос»; информационно-измерительная система.

Данная работа продолжает серию статей, начатую с публикации [1], посвященных:

1) синтезу адаптивных интеллектуальных измерительных систем в автоматизированном системно-когнитивном анализе (АСК-анализ) и его программном инструментарии — универсальной когнитивной аналитической системе «Эйдос-X++»;

2) применению интеллектуальных измерительных систем для системных измерений, т.е. для идентификации состояний сложных многофакторных нелинейных динамических систем (так называемая системная идентификация).

Статья [1] посвящена концептуальным основам построения интеллектуальных измерительных систем в АСК-анализе, данная работа (вторая) — математическому методу АСК-анализа, третью планируется посвятить рассмотрению численного примера синтеза интеллектуальной измерительной системы в системе «Эйдос-X++» и ее применения для системной идентификации состояний сложных систем.

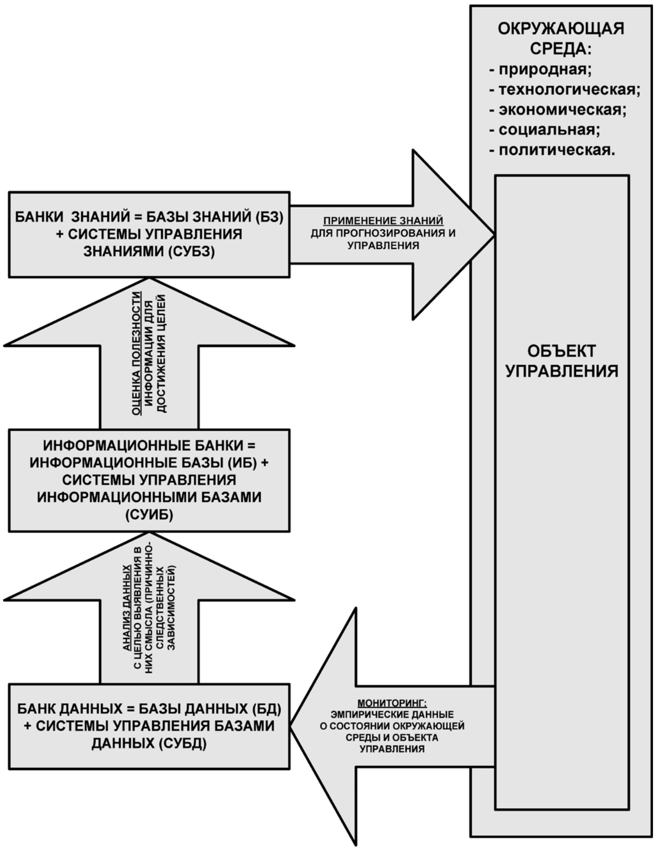
На основе работы [1] предлагаются следующие три принципа построения интеллектуальных измерительных систем в АСК-анализе.

Первый принцип заключается в осознании того обстоятельства, что когда мы получаем результаты измерения, то, по сути, получаем некоторое количество информации о том, в каком состоянии находится измеряемый объект. Однако традиционно результаты измерения выражаются в определенных единицах измерения (в частности, единицах измерения физических величин), а не в единицах измерения информации и этим в определенной степени маскируется или скрывается смысл самого измерения, выраженный в первом принципе.

Второй принцип связан с первым и состоит в понимании того, что когда мы получаем результаты измерения, то нас интересует не собственно сам результат, а количество информации, которое содержится в результате измерения о состоянии объекта измерения, т.е. о том, что нас собственно интересует. Например, когда врач измеряет температуру пациенту, то его интересует не температура сама по себе, как некоторые почему-то думают, а возможность на ее основе сделать выводы о состояния пациента, т.е. о том, болен он или нет, и, если болен, то насколько серьезно и какой у него диагноз.

Третий принцип состоит в том, что при построении измерительной системы на эмпирических примерах проводится градуировка или метризация измерительных шкал, т.е. нанесение на них делений, соответствующих различным степеням выраженности измеряемых свойств у объектов измерения. Затем, когда измерительная система применяется, т.е. при измерении по ранее полученным шкалам получаются неко-

[1] Кубанский государственный аграрный университет, г. Краснодар, Россия; e-mail: prof.lutsenko@gmail.com



Цикл преобразования эмпирических данных в информацию и знания и их применение для прогнозирования и принятия управленческих решений в АСК-анализе

торые значения, то на их основании делается вывод о том, что состояние измеряемого объекта близко к состоянию тех примеров, которые давали аналогичный результат измерений при построении шкал. По сути третий принцип состоит в том, что этап построения измерительной системы функционально сходен с этапом обучения системы распознавания образов, а этап ее применения сходен с использованием системы распознавания для идентификации состояния объекта измерения.

Для того чтобы реализовать сформулированные принципы в реальной интеллектуальной измерительной системе, необходим математический метод, обеспечивающий преобразование данных, полученных в результате измерений, в информацию о состоянии измеряемого объекта. Такой метод существует — это математический метод АСК-анализа, основанный на системной нечеткой интервальной математике (СНИМ) [2] и представляющий собой реализацию идей СНИМ в теории информации.

В этой связи необходимо определить соотношение содержания терминов: «данные», «информация» и «знание» (см. рисунок).

Данные рассматриваются как информация, записанная на носителях или находящаяся в каналах связи и представленная в определенной системе кодирования или на определенном языке и рассматриваемая безотносительно к ее смысловому содержанию.

Смысл данных согласно концепции Шенка – Абельсона известен и понятен тогда, когда известны причины и следствия меду событиями, которые описываются этими данными. Информация представляет собой осмысленные данные, т.е. данные, описывающие события, между которыми выявлены причинно-следственные связи. Знания — это информация, полезная для достижения целей, т.е. для управления.

В этой связи возникает вопрос о математической количественной мере причинно-следственных связей, которая бы адекватно отражала их силу и направление. Из вышесказанного следует, что естественной мерой причинно-следственных связей являются количественные меры информации и в качестве единицы измерения силы и направления причинно-следственных связей могут быть использованы единицы измерения информации. В связи с этой идеей необходимо отметить работу [3], суть которой состоит в применении теории информации для проверки статистических гипотез. (В этой связи нельзя не упомянуть о лемме Неймана – Пирсона, согласно которой более вероятна реализация той статистической гипотезы, в пользу которой имеется больше информации.) В предисловии к работе [3] А. Н. Колмогоров высоко оценивал это направление, но оно не получило в СССР должного распространения. По-видимому, АСК-анализ [4] в определенной мере можно рассматривать как развитие этого направления прикладной математической статистики, может быть, не столько в чисто математическом теоретическом плане, сколько в прагматически прикладном.

Однако известно довольно много различных количественных мер информации. Поэтому возникает вопрос о том, какая из них является наиболее подходящей в нашем случае. По мнению автора, это семантическая мера целесообразности информации А. Харкевича [5]. Основным свойством этой меры, предопределяющим ее выбор, является то обстоятельство, что в ее определение органично входит понятие цели. Кроме того, для вычисления меры А. Харкевича достаточно знать изменение вероятности достижения цели в условиях действия некоторого значения фактора и при его отсутствии, т.е. она вполне может быть рассчитана непосредственно на основе эмпирических данных, что очень важно для практических применений (поэтому и говорят, что эта мера прагматическая).

Операция преобразования данных в информацию называется «анализ» и предполагает выполнение следующих этапов:

1) разработка справочников, содержащих формальное кодированное описание, с одной стороны, будущих состояний объекта управления, а с другой — факторов их значений, влияющих на этот объект (классификационных и описательных шкал и градаций в терминологии АСК-анализа);

2) поиск в исходных данных событий, связанных с переходами объекта из одного состояния в другое, и значений факторов, под действием которых эти переходы происходят; при этом в качестве значений факторов могут выступать и переходы объекта из одного состояния в другое в прошлом;

3) преобразование базы исходных данных в базу событий, т.е. кодирование исходных данных с использованием справочников классов и факторов;

4) поиск причинно-следственных связей между прошлыми и будущими событиями в базе событий и формальное представление этих причинно-следственных связей в виде базы информативностей.

Таким образом, если исходные базы данных представляют собой временные ряды, то информационная база включает в себя дополнительно:

базы классификационных и описательных шкал и градаций;

базу событий, представляющую собой закодированную с помощью классификационных и описательных шкал и градаций базу исходных данных;

базу информативностей, содержащую информацию о силе и направлении влияния значений факторов на переход объекта управления в состояния, соответствующие классам.

Основываясь на работах [1, 2, 4, 6], рассмотрим математический метод АСК-анализа, обеспечивающий решение поставленных задач. Краткое описание автоматизированного системно-когнитивного анализа (АСК-анализа) приведено в работе [4].

В работе [6] (и ряде других) развита и обоснована идея системного обобщения математики. Эта идея актуальна по ряду причин разного рода.

Во-первых, в мире нет ничего, кроме систем, а понятие множества является абстракцией от понятия системы: множество — это система без внутренней структуры. Поэтому математика, основанная на понятии системы, имеет некоторые шансы быть более адекватной, чем классическая математика, в большой степени основанная на понятии множества.

Во-вторых, идея системного обобщения математики частично реализована в теории информации, в результате получены некоторые результаты в области системной теории информации (СТИ), в частности, — вариант выражения для семантической меры целесообразности информации А. Харкевича, удовлетворяющий принципу соответствия с формулой Р. Хартли для равновероятного детерминистского случая. Этим преодолена искусственная пропасть между «Теорией передачи данных по каналам связи», как совершенно справедливо называл свою теорию К. Шеннон, понимавший различие между данными и информацией, и семантической теорией информации А. Харкевича.

В-третьих, в созданной системной теории информации получены разнообразные формы различных коэффициентов эмерджентности — Хартли, Харкевича, Шеннона — для классических систем, подчиняющихся статистике Л. Больцмана [7] и квантовых систем, подчиняющихся статистикам Ферми – Дирака и Бозе – Эйнштейна [7]. Смысл этих коэффициентов раскрыт в работах [2, 6, 7] и др. Если резюмировать, то можно сказать, что эти подходы, по-видимому, открывают новые подходы математического моделирования процессов эволюции систем различного рода и масштаба от микро- до макро- и мегауровней [8, 9] и др.

Математический метод АСК-анализа основан на системной теории информации (СТИ), которая создана в рамках реализации программной идеи обобщения всех понятий математики, в частности теории информации, базирующихся на теории множеств, путем тотальной замены понятия множества на более общее понятие системы и тщательного отслеживания всех последствий этой замены [2, 6]. Благодаря математическому методу, положенному в основу АСК-анализа, этот метод является непараметрическим и позволяет в реализующей его системе «Эйдос-X++» сопоставимо обрабатывать десятки и сотни тысяч градаций факторов и будущих состояний объекта управления (классов) при неполных (фрагментированных), зашумленных данных числовой и нечисловой природы, измеряемых в различных единицах измерения [10].

Итак, будем считать, что информация содержится не только в самих базовых элементах системы, но и в ее подсистемах различной сложности, т.е. состоящих из 2, 3, ..., m, ..., M базовых элементов.

Классическая формула Хартли имеет вид [11]

$$I = \log_2 w. \tag{1}$$

Будем искать ее системное обобщение в виде

$$I = \log_2 W^{\varphi}, \tag{2}$$

где W — количество элементов в множестве; I — количество информации, которое содержится в факте извлечения одного элемента из множества; φ — коэффициент эмерджентности, названный автором коэффициентом эмерджентности Хартли[2].

Суммарное количество таких подсистем для систем, подчиняющихся статистике Ферми – Дирака [7], можно принять равным числу сочетаний. Поэтому примем, что системное обобщение формулы Хартли имеет вид

$$I = \log_2 \sum_{m=1}^{M} C_W^m, \tag{3}$$

где C_W^m — количество подсистем из m элементов; m — сложность подсистем; M — максимальная сложность подсистем (максимальное число элементов подсистемы).

[2] *Вяткин В.* Групповой плагиат: от студента до министра. — [Электронный ресурс]. URL: http://trv-science.ru/2011/11/08/gruppovojj-plagiat-ot-studenta-do-ministra

Поскольку $C_W^1 = W$, то при $M = 1$ система переходит в множество и выражение (3) приобретает вид (1), т.е. для него выполняется принцип соответствия, являющийся обязательным для более общей теории.

Учитывая, что при $M = W$

$$\sum_{m=1}^{M} C_W^m = 2^W - 1, \tag{4}$$

получаем

$$I = \log_2(2^W - 1). \tag{5}$$

Выражение (5) дает оценку максимального количества информации в элементе системы. Из выражения (5) видно, что при увеличении числа элементов W количество информации I быстро стремится к W и уже при $W > 4$ погрешность выражения (5) не превышает 1 %:

$$\lim_{W \to \infty} \frac{I}{W} = 1. \tag{6}$$

Приравняв правые части выражений (2) и (3), т.е.

$$I = \log_2 W^{\varphi} = \log_2 \sum_{m=1}^{M} C_W^m, \tag{7}$$

запишем выражение для коэффициента эмерджентности Хартли:

$$\varphi = \frac{\log_2 \sum_{m=1}^{M} C_W^m}{\log_2 W}. \tag{8}$$

Смысл этого коэффициента раскрыт в работах [2, 6, 7] и др. Здесь отметим лишь, что при $M \to 1$, когда система асимптотически переходит в множество, имеем $\varphi \to 1$ и (2) $\to$ (1), как и должно быть согласно принципу соответствия, предложенному Нильсом Бором в 1913 г.

С учетом (8) выражение (2) примет вид

$$I(W, M) = \log_2 W^{\frac{\log_2 \sum_{m=1}^{M} C_W^m}{\log_2 W}} \tag{9}$$

или при $M = W$ и больших W с учетом формул (4) и (5)

$$I(W, M) = \log_2 W^{W/\log_2 W} = W. \tag{10}$$

Выражение (9) и представляет собой искомое системное обобщение классической формулы Хартли, а выражение (10) — его достаточно хорошее приближение при большом количестве элементов в системе W.

Классическая формула А. Харкевича имеет вид

$$I_{ij}(W, M) = \log_2 \frac{P_{ij}}{P_{\Sigma j}}, \tag{11}$$

где P_{ij} — условная вероятность перехода объекта в j-е состояние при условии действия на него i-го значения фактора; $P_{\Sigma j}$ — безусловная вероятность перехода объекта в j-е состояние (вероятность самопроизвольного перехода или вероятность перехода, рассчитанная по всей выборке, т.е. при действии любого значения фактора).

Придадим выражению (11) следующий эквивалентный вид, который и будем использовать ниже:

$$I_{ij}(W, M) = \log_2 \frac{P_{ij}}{P_{i\Sigma}}, \tag{12}$$

где индекс i обозначает признак (значение фактора), $1 \le i \le M$; индекс j — состояние объекта или класс $1 \le j \le W$; P_{ij} — условная вероятность наблюдения i-го значения фактора у объектов j-го класса; $P_{i\Sigma}$ — безусловная вероятность наблюдения i-го значения фактора по всей выборке.

Из (12) видно, что формула Харкевича для семантической меры информации по сути является логарифмом от формулы Байеса для апостериорной вероятности (отношение условной вероятности к безусловной).

Известно, что классическая формула Шеннона для количества информации для неравновероятных событий преобразуется в формулу Хартли при условии, что события равновероятны, т.е. удовлетворяет фундаментальному принципу соответствия. Поэтому теория информации Шеннона справедливо считается обобщением теории Хартли для неравновероятных событий. Однако выражения (11) и (12) при подстановке в них реальных численных значений вероятностей P_{ij}, $P_{i\Sigma}$ и $P_{\Sigma j}$ не дают количества информации в битах, т.е. для этих выражений не выполняется принцип соответствия, обязательный для более общих теорий. Возможно, в этом состоит причина довольно сдержанного, а иногда и скептического отношения специалистов по теории информации Шеннона к семантической теории информации Харкевича.

Причину этого мы видим в том, что в выражениях (11) и (12) отсутствуют глобальные параметры *конкретной* модели W и M, т.е. в том, что А. Харкевич в своем выражении для количества информации не ввел зависимости от мощности пространства будущих состояний объекта W и количества значений факторов M, обуславливающих переход объекта в эти состояния.

Поставим задачу получить такое обобщение формулы Харкевича, которое бы удовлетворяло тому же самому принципу соответствия, что и формула Шеннона, т.е. преобразовывалось в формулу Хартли в предельном детерминистском равновероятном случае, когда каждому классу (состоянию объекта) соответствует один признак (значение фактора) и каждому признаку — один класс и эти классы (а, значит, и признаки) равновероятны, при этом каждый фактор однозначно, т.е. детерминистским образом, определяет переход объекта в определенное состояние, соответствующее классу.

В детерминистском случае вероятность P_{ij} наблюдения объекта j-го класса при обнаружении у него i-го признака составляет

$$P_{ij} = \begin{cases} 1 \text{ при } i = j \\ 0 \text{ при } i \neq j \end{cases}.$$

Будем искать обобщение (12) в виде

$$I_{ij}(W, M) = \log_2 \left(\frac{P_{ij}}{P_{i\Sigma}} \right)^{\Psi}. \quad (13)$$

Найдем такое выражение для коэффициента Ψ, названного автором «коэффициентом эмерджентности Харкевича»[2], которое обеспечивает выполнение для выражения (13) принципа соответствия с классической формулой Хартли (1) и ее системным обобщением (2) и (3) в равновероятном детерминистском случае.

Для этого потребуется выразить вероятности P_{ij}, P_j и P_i через частоты наблюдения признаков по классам (табл. 1, где рамкой обведена область значений, переменные определены ранее).

Алгоритм формирования матрицы абсолютных частот

Объекты обучающей выборки описываются векторами (массивами) $\mathbf{L} = \{L_i\}$ имеющихся у них признаков: $\mathbf{L} = \{L_i\} = \mathbf{n}$, если у объекта i-й признак встречается n раз.

Первоначально в матрице абсолютных частот все значения равны нулю. Затем организуется цикл по объектам обучающей выборки. Если у предъявленного объекта, относящегося к j-му классу, есть i-й признак, то

$$N_{ij} = N_{ij} + 1; \;\; N_{i\Sigma} = N_{i\Sigma} + 1;$$

$$N_{\Sigma j} = N_{\Sigma j} + 1; \;\; N_{\Sigma\Sigma} = N_{\Sigma\Sigma} + 1.$$

Здесь можно провести очень интересную и важную аналогию между способом формирования матрицы абсолютных частот и работой многоканальной системы выделения полезного сигнала из шума. Представим себе, что все объекты, предъявляемые для формирования обобщенного образа некоторого класса, в действительности являются различными реализациями одного объекта — «Эйдоса» (в смысле Платона), по-разному зашумленного различными случайными обстоятельствами. И наша задача состоит в том, чтобы подавить этот шум и выделить из него то общее и существенное, что отличает объекты данного класса от объектов других классов. Учитывая, что шум чаще всего является «белым» и имеет свойство при суммировании с самим собой стремиться к нулю, а сигнал при этом, наоборот, возрастает пропорционально количеству слагаемых, то увеличение объема обучающей выборки приводит ко все лучшему отношению сигнал/шум в матрице абсолютных частот, т.е. к выделению полезной информации из шума. Примерно так мы начинаем постепенно понимать смысл фразы, которую сразу не расслышали по телефону и несколько раз переспрашивали. При этом в повторах шум не позволяет понять то одну, то другую часть фразы, но в конце концов за счет использования памяти и интеллектуальной обработки информации мы понимаем ее всю. Так и объекты, описанные признаками, можно рассматривать как зашумленные фразы, несущие нам информацию об обобщенных образах классов — «Эйдосах» [6], к которым они относятся. И эту информацию мы выделяем из шума при синтезе модели.

Для выражения (11) получим

$$P_{ij} = N_{ij}/N_{i\Sigma}, \quad (14)$$

для (12) и (13)

$$P_{ij} = N_{ij}/N_{\Sigma j}, \quad (15)$$

для (11), (12) и (13) —

$$P_i = N_{i\Sigma}/N_{\Sigma\Sigma}, \;\; P_j = N_{\Sigma j}/N_{\Sigma\Sigma},$$

$$N_{i\Sigma} = \sum_{j=1}^{W} N_{ij}, \;\; N_{\Sigma j} = \sum_{i=1}^{M} N_{ij},$$

$$N_{\Sigma\Sigma} = \sum_{i=1}^{M} N_{i\Sigma} = \sum_{j=1}^{W} N_{\Sigma j} = \sum_{i=1}^{W} \sum_{j=1}^{M} N_{ij}. \quad (16)$$

В выражениях (16) использованы следующие обозначения: N_{ij} — суммарное количество наблюдений в исследуемой выборке факта: «действовало i-е значение фактора и объект перешел в j-е состояние»; $N_{\Sigma j}$ — суммарное по всей выборке количество встреч различных факторов у объектов, перешедших в j-е состояние; $N_{i\Sigma}$ — суммарное количество встреч i-го фактора у всех объектов исследуемой выборки; $N_{\Sigma\Sigma}$ — суммарное количество встреч различных значений факторов у всех объектов исследуемой выборки.

Формирование матрицы условных и безусловных процентных распределений. На основе анализа матрицы частот (см. табл. 1) классы можно сравнивать по наблюдаемым частотам признаков только в том случае, если количество объектов по всем классам одинаково, как и суммарное количество признаков по классам. Если же они отличаются, то корректно сравнивать классы можно только по условным и безусловным относительным частотам (оценкам вероятностей) наблюдений признаков, посчитанных на основе матрицы частот (см. табл. 1) в соответствии с выражениями (14) и (15), в результате чего получается матрица условных и безусловных процентных распределений (табл. 2).

При расчете матрицы оценок условных и безусловных вероятностей N_j (см. табл. 1) могут браться либо из предпоследней, либо из последней строки. В первом случае N_j представляет собой «суммарное количество признаков у всех объектов, использованных для формирования обобщенного образа j-го класса», а во втором случае — это «суммарное количество объектов обучающей выборки, использованных для формирования обобщенного образа j-го класса». Соответственно получаем различные, хотя и очень сходные семантические информационные модели, которые называем СИМ-1 и СИМ-2. Оба этих вида моделей поддерживаются системой «Эйдос».

Эквивалентность выражений (11) и (12) устанавливается, если подставить в них выражения относительных частот как оценок вероятностей P_{ij}, $P_{\Sigma j}$ и $P_{i\Sigma}$ через абсолютные частоты наблюдения признаков по классам из (14), (15) и (16). В обоих случаях из выражений (11) и (12) получается одно и то же выражение

$$I_{ij} = \log_2 \frac{N_{ij} N_{\Sigma\Sigma}}{N_{i\Sigma} N_{\Sigma j}}, \tag{17}$$

Таблица 1. Матрица абсолютных частот

		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				$N_{\Sigma j} = \sum_{i=1}^{M} N_{ij}$			$N_{\Sigma\Sigma} = \sum_{i=1}^{W} \sum_{j=1}^{M} N_{ij}$

Таблица 2. Матрица условных и безусловных процентных распределений

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

а из (13) — выражение

$$I_{ij} = \log_2 \left(\frac{N_{ij} N_{\Sigma\Sigma}}{N_{i\Sigma} N_{\Sigma j}} \right)^{\Psi}, \tag{18}$$

с которым и будем далее работать.

При взаимно-однозначном соответствии классов и признаков в равновероятном детерминистском случае имеем матрицу частот в виде табл. 3.

В этом случае к каждому классу относится один объект, имеющий единственный признак. Отсюда для всех i и j получаем

$$\forall ij{:}\;\; N_{ij} = N_{i\Sigma} = N_{\Sigma j} = 1. \tag{19}$$

В результате обобщенная формула А. Харкевича (18) с учетом (19) приобретает вид

$$I_{ij} = \log_2 N_{\Sigma\Sigma}^{\Psi} = \log_2 W^{\varphi}, \tag{20}$$

откуда

$$\Psi = \frac{\log_2 W^{\varphi}}{\log_2 N_{\Sigma\Sigma}}. \tag{23}$$

С учетом выражения для коэффициента эмерджентности Хартли (8)

$$\Psi = \frac{\log_2 W^{\frac{\log_2 \sum_{m=1}^{M} C_W^m}{\log_2 W}}}{\log_2 N_{\Sigma\Sigma}}. \tag{22}$$

Подставив коэффициент эмерджентности А. Харкевича (21) в выражение (18), получим

$$I_{ij} = \log_2 \left(\frac{N_{ij} N_{\Sigma}}{N_{i\Sigma} N_{\Sigma j}} \right)^{\Psi} = \log_2 \left(\frac{N_{ij} N_{\Sigma\Sigma}}{N_{i\Sigma} N_{\Sigma j}} \right)^{\frac{\log_2 W^{\varphi}}{\log_2 N}} =$$

$$= \frac{\log_2 W^{\varphi}}{\log_2 N_{\Sigma\Sigma}} \left[\log_2 \left(\frac{N_{ij}}{N_{i\Sigma} N_{\Sigma j}} \right) + \log_2 N_{\Sigma\Sigma} \right] =$$

$$= \log_2 \left(\frac{N_{ij}}{N_{i\Sigma} N_{\Sigma j}} \right)^{\frac{\log_2 W^{\varphi}}{\log_2 N_{\Sigma\Sigma}}} + \log_2 W^{\varphi}.$$

Таблица 3. Матрица частот в равновероятном детерминистском случае

		Классы					Сумма
		1	...	j	...	W	
Значения факторов	1	1					1
	...		1				1
	i			1			1
	...				1		1
	M					1	1
Сумма		1	1	1	1	1	$N_{\Sigma\Sigma}$

Запишем в окончательном варианте:

$$I_{ij} = \log_2 \left(\frac{N_{ij}}{N_{i\Sigma} N_{\Sigma j}} \right)^{\frac{\log_2 W^{\varphi}}{\log_2 N_{\Sigma\Sigma}}} + \log_2 W^{\varphi}. \qquad (23)$$

Отметим, что первая задача получения системного обобщения формул Хартли и Харкевича и вторая задача получения такого обобщения формулы Харкевича, которая удовлетворяет принципу соответствия с формулой Хартли, — это две разные задачи. Первая является более общей и при ее решении, которое приведено выше, автоматически решается и вторая задача, которая является, таким образом, частным случаем первой.

Однако представляет самостоятельный интерес и частный случай, в результате которого получается формула Харкевича, удовлетворяющая в равновероятном детерминистском случае принципу соответствия с классической формулой Хартли (1), а не с ее системным обобщением (2) и (3). Ясно, что эта формула получается из (23) при $\varphi = 1$:

$$I_{ij} = \log_2 \left(\frac{N_{ij}}{N_{i\Sigma} N_{\Sigma j}} \right)^{\frac{\log_2 W}{\log_2 N_{\Sigma\Sigma}}} + \log_2 W. \qquad (24)$$

Из выражений (21) и (22) видно, что в частном случае, когда система эквивалентна множеству ($M = 1$), коэффициент эмерджентности А. Харкевича приобретает вид

$$\Psi = \frac{\log_2 W}{\log_2 N_{\Sigma\Sigma}}. \qquad (25)$$

На практике для численных расчетов удобнее пользоваться не выражениями (23) или (24), а формулой, которая получается непосредственно из (18) после подстановки в него выражения (25):

$$I_{ij} = \frac{\log_2 W}{\log_2 N_{\Sigma\Sigma}} \log_2 \frac{N_{ij} N_{\Sigma\Sigma}}{N_{i\Sigma} N_{\Sigma j}}. \qquad (26)$$

Используя выражение (26) и данные табл. 1, непосредственно прямым счетом получаем матрицу знаний (табл. 4).

Когда количество информации $I_{ij} > 0$, i-й фактор способствует переходу объекта управления в j-е состояние, когда $I_{ij} < 0$, он препятствует этому переходу, а когда $I_{ij} = 0$, то никак не влияет на это. В векторе i-го фактора (строка матрицы информативностей) отображается, какое количество информации о переходе объекта управления в каждое из будущих состояний содержится в том факте, что данный фактор действует. В векторе j-го состояния класса (столбец матрицы информативностей) отображается, какое количество информации о переходе объекта управления в соответствующее состояние содержится в каждом из факторов.

Таким образом, матрица знаний (информативностей), приведенная в табл. 4, является обобщенной таблицей решений, в которой входы (факторы) и выходы (будущие состояния объекта управления) связаны друг с другом не с помощью классических (Аристотелевых) импликаций, принимающих только значения «истина» и «ложь», а различными значениями истинности, выраженными в битах и принимающими значения от положительного теоретически максимально возможного («максимальная степень истинности») до теоретически неограниченного отрицательного («степень ложности»). Это позволяет автоматически формулировать прямые и опосредованные правдоподобные высказывания с расчетной степенью истинности.

Таблица 4. Матрица знаний (информативностей)

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

Здесь $\bar{I}_i = \frac{1}{W}\sum_{j=1}^{W} I_{ij}$ — среднее количество знаний в i-м значении фактора.

Фактически предложенная модель дает возможность осуществить синтез обобщенных таблиц решений для различных предметных областей непосредственно на основе эмпирических исходных данных и продуцировать прямые и обратные правдоподобные (нечеткие) логические рассуждения по неклассическим схемам с различными расчетными значениями истинности, являющимися обобщением классических импликаций.

Данная модель позволяет рассчитать, какое количество информации содержится в любом факте о наступлении любого события в любой предметной области, причем для этого не требуется повторности этих фактов и событий. Если такие повторности осуществляются и при этом наблюдается некоторая вариабельность значений факторов, обуславливающих наступление тех или иных событий, то модель обеспечивает многопараметрическую типизацию, т.е. синтез обобщенных образов классов или категорий наступающих событий с количественной оценкой степени и знака влияния на их наступление различных значений факторов. Причем эти значения факторов могут быть как количественными, так и качественными и измеряться в любых единицах измерения. В любом случае в модели оценивается количество информации, которое в них содержится о наступлении событий, переходе объекта управления в определенные состояния или просто о его принадлежности к тем или иным классам. Другие способы метризации приведены в работе [10]. Все они реализованы в системно-когнитивном анализе и интеллектуальной системе «Эйдос» и обеспечивают сопоставление с градациями всех видов шкал числовых значений, имеющих смысл количества информации о принадлежности объекта к классу. Поэтому корректно применение интегральных критериев, включающих операции умножения и суммирования, для обработки числовых значений, соответствующих градациям шкал. Это позволяет единообразно и сопоставимо обрабатывать эмпирические данные, полученные с помощью любых типов шкал, применяя при этом все математические операции.

Информационный портрет класса — это список значений факторов, ранжированных в порядке убывания силы их влияния на переход объекта управления в состояние, соответствующее данному классу. Информационный портрет класса отражает систему его детерминации. Генерация информационного портрета класса представляет собой решение обратной задачи прогнозирования, так как при прогнозировании по системе факторов определяется спектр наиболее вероятных будущих состояний объекта управления, в которые он может перейти под влиянием данной системы факторов, а в информационном портрете, наоборот, по заданному будущему состоянию объекта управления определяется система факторов, детерминирующих это состояние, т.е. вызывающих переход объекта управления в это состояние. В начале информационного портрета класса идут факторы, оказывающие положительное влияние на переход объекта управления в заданное состояние, затем — факторы, не оказывающие на это существенного влияния, и далее — факторы, препятствующие переходу объекта управления в это состояние (в порядке возрастания силы препятствования). Информационные портреты классов могут быть отфильтрованы по диапазону факторов, т.е. мы можем отобразить влияние на переход объекта управления в данное состояние не всех отраженных в модели факторов, а только тех, коды которых попадают в определенный диапазон, например, относящиеся к определенным описательным шкалам.

Информационный (семантический) портрет фактора — это список классов, ранжированный в порядке убывания силы влияния данного фактора на переход объекта управления в состояния, соответствующие данным классам. Информационный портрет фактора называется также его семантическим портретом, так как в соответствии с концепцией смысла системно-когнитивного анализа, являющейся обобщением концепции смысла Шенка – Абельсона [12], смысл фактора состоит в том, какие будущие состояния объекта управления он детерминирует или обуславливает. Сначала в этом списке идут состояния объекта управления, на переход в которые данный фактор оказывает наибольшее влияние, затем состояния, на которые данный фактор не оказывает существенного влияния, и далее состояния, переходу в которые данный фактор препятствует. Информационные портреты факторов могут быть отфильтрованы по диапазону классов, т.е. мы можем отобразить влияние данного фактора на переход объекта управления не во все возможные будущие состояния, а только в состояния, коды которых попадают в определенный диапазон, например, относящиеся к определенным классификационным шкалам.

Прямые и обратные, непосредственные и опосредованные правдоподобные логические рассуждения с расчетной степенью истинности в системной теории информации. Одним из первых ученых, поднявших и широко обсуждавшим в своих работах проблематику правдоподобных рассуждений, был известный венгерский, швейцарский и американский математик Дьердь Пойа [13], книги которого подарил мне школьный учитель математики Михаил Ильич Перевалов, за что ему очень благодарен. Разве мог он тогда предположить, что через много лет в работе [6] мной будет предложена логическая форма представления правдоподобных логических рассуждений с расчетной степенью истинности, которая определяется в соответствии с системной теорией информацией непосредственно на основе эмпирических данных.

В качестве количественной меры влияния факторов предложено использовать обобщенную формулу А. Харкевича, полученную на основе предложенной эмерджентной теории информации. При этом непосредственно из матрицы абсолютных частот рассчитывается база знаний (см. табл. 4), которая и представляет собой основу содержательной информационной модели предметной области. Весовые коэффициенты табл. 4 непосредственно определяют, какое количество информации I_{ij} система управления получает о наступлении события: «активный объект управления перейдет в j-е состояние» из сообщения «на активный объект управления действует i-й фактор».

Принципиально важно, что эти весовые коэффициенты не определяются экспертами неформализуемым способом на основе интуиции и профессиональной компетенции (фактически «на глазок»), а рассчитываются непосредственно на основе эмпирических данных и удовлетворяют всем ранее обоснованным в работе [1] требованиям, т.е. являются сопоставимыми, содержательно интерпретируемыми, отражают понятия «достижение цели управления» и «мощность множества будущих состояний объекта управления» и т.д.

В работе [6] обосновано, что предложенная информационная мера обеспечивает сопоставимость индивидуальных количеств информации, содержащейся в факторах о классах, а также сопоставимость интегральных критериев, рассчитанных для одного объекта и разных классов, для разных объектов и разных классов.

Когда количество информации $I_{ij} > 0$, i-й фактор способствует переходу объекта управления в j-е состояние, когда $I_{ij} < 0$, он препятствует этому переходу, когда же $I_{ij} = 0$, никак не влияет на это. В векторе i-го фактора (строка матрицы информативностей) отображается, какое количество информации о переходе объекта управления в каждое из будущих состояний содержится в том факте, что данный фактор действует. В векторе j-го состояния класса (столбец матрицы информативностей) отображается, какое количество информации о переходе объекта управления в соответствующее состояние содержится в каждом из факторов.

Таким образом, матрица информативностей (см. табл. 4) является обобщенной таблицей решений, в которой входы (факторы) и выходы (будущие состояния активного объекта управления (АОУ) связаны друг с другом не с помощью классических (Аристотелевских) импликаций, принимающих только значения «истина» и «ложь», а различными значениями истинности, выраженными в битах и принимающими значения от положительного теоретически максимально возможного («максимальная степень истинности») до теоретически неограниченного отрицательного («степень ложности»).

Фактически предложенная модель позволяет осуществить синтез обобщенных таблиц решений для различных предметных областей непосредственно на основе эмпирических исходных данных и продуцировать на их основе прямые и обратные правдоподобные (нечеткие) логические рассуждения по неклассическим схемам с различными расчетными значениями истинности, являющимся обобщением классических импликаций (табл. 5).

Приведем пример более сложного высказывания, которое может быть рассчитано непосредственно на основе матрицы информативностей — обобщенной таблицы решений (см. табл. 4). Если A со степенью истинности $\alpha(A, B)$ детерминирует B и если C со степенью истинности $\alpha(C, D)$ детерминирует D и A совпадает по смыслу с C со степенью истинности $\alpha(A, C)$, то это вносит вклад в совпадение B с D, равный степени истинности $\alpha(B, D)$.

При этом в прямых рассуждениях как предпосылки рассматриваются факторы, а как заключение — будущие состояния АОУ, а в обратных, наоборот, как предпосылки — будущие состояния АОУ, а как заключение — факторы. Степень истинности i-й предпосылки — это просто количество информации I_{ij}, содержащейся в ней о наступлении j-го будущего состояния АОУ. Если предпосылок несколько, то степень истинности наступления j-го состояния АОУ равна суммарному количеству информации, содержащемуся в них. Количество информации в i-м факторе о наступлении j-го состояния АОУ рассчитывается в соответствии с выражениями системной теории информации (СТИ).

Прямые правдоподобные логические рассуждения позволяют прогнозировать степень достоверности наступления события по действующим факторам, а об-

Таблица 5. Прямые и обратные правдоподобные логические высказывания с расчетной в соответствии с системной теорией информации (СТИ) степенью истинности импликаций

	Прямые высказывания:	Обратные высказывания
1	Если A, то B (если действует фактор A, то мы предполагаем со степенью истинности I_{AB}, что АОУ перейдет в состояние B)	Если B, то A (если АОУ перешел в состояние B, то мы предполагаем со степенью истинности I_{AB}, что действовал фактор A)
2	Если A_1 и A_2 ..., и A_M, то B (прогноз влияния системы факторов на поведение АОУ; степень истинности обобщающей — итоговой импликации равна алгебраической сумме истинностей составляющих ее элементарных импликаций вида: «если A то B»)	Если B, то A_1 и A_2 ..., и A_M (информационный портрет класса B, т.е. влияние различных факторов A_i на переход АОУ в будущее состояние B, решение обратной задачи прогнозирования, т.е. выработка управления)
3	Если A, то B_1 или B_2 ..., или B_W (семантический портрет фактора A, т.е. его влияние на переход АОУ в различные состояния)	
4	Если A_1 и A_2 ..., и A_M, то B_1 или B_2 ..., или B_W (прогноз влияния системы факторов на переход АОУ в различные состояния)	

ратные — по заданному состоянию восстановить степень необходимости и степень нежелательности каждого фактора для наступления этого состояния, т.е. принимать решение по выбору управляющих воздействий на АОУ, оптимальных для перевода его в заданное целевое состояние.

Приведем простой пример, когда безупречная классическая бинарная логика Аристотеля дает сбой. Рассмотрим высказывания:

А) если студент хорошо сдал экзамен по информационным системам, значит, он умеет хорошо программировать;

Б) если студент умеет хорошо программировать, то он может стать хорошим специалистом в области прикладной информатики.

Отсюда средствами логики предикатов получаем вывод:

В) если студент хорошо сдал экзамен по информационным системам, то он может стать хорошим специалистом в области прикладной информатики.

Если при рассмотрении каждого высказывания «А» и «Б» по отдельности у нас не возникает особых возражений, хотя мы сразу чувствуем, что это не совсем так или не всегда так и легко можем привести вполне реальные примеры, когда эти высказывания могут быть и ложными, то высказывание «В» уже само по себе выглядит очень сомнительным, т.е., проще говоря, ложным, тогда как в логике предикатов оно является истинным. Интуитивно мы хорошо понимаем, почему так получается. Дело в том, что в этих высказываниях не отражен *контент*, т.е. та огромная слабо формализованная и вообще неформализованная информация об объекте моделирования, которой располагает человек, но не располагает логическая система. Например, в этих двух логических высказываниях не отражена информация, которой располагает каждый преподаватель и студент о том, каким образом иногда сдаются экзамены, когда оценка вообще никак не зависит от знаний. Иначе говоря, чтобы эти высказывания были истинны, необходимо, чтобы оценка определялась только знаниями. Но и этого мало. Предполагается, что факт получения хорошей оценки по дисциплине означает полное ее освоение, хотя все понимают, что для этого достаточно освоения только тех вопросов, которые были в билете и были заданы преподавателем (кроме того, существуют и весьма распространены и другие варианты).

При решении этой же задачи средами АСК-анализа мы формулируем эти высказывания в форме правдоподобных рассуждений:

А) если студент хорошо сдал экзамен по информационным системам, то в этом факте содержится $I(\text{А})$ информации о том, что он умеет хорошо программировать;

Б) если студент умеет хорошо программировать, то в этом факте содержится $I(\text{Б})$ информации о том, что он может стать хорошим специалистом в области прикладной информатики.

Откуда средствами АСК-анализа получаем результирующее высказывание:

В) если студент хорошо сдал экзамен по информационным системам, то в этом факте содержится $I(\text{В})$ информации о том, что он может стать хорошим специалистом в области прикладной информатики.

Это высказывание не выглядит как истинное или ложное и может быть и истинным, и ложным, причем в различной степени, в зависимости от знака и модуля его расчетной степени истинности $I(\text{В})$. Для расчета этой величины нужны конкретные эмпирические данные, являющиеся репрезентативными для отражения определенной предметной области (генеральной совокупности), в которой этот вывод и будет иметь эти значения знака и величины степени истинности.

Описанная математическая модель реализована в программном инструментарии АСК-анализа — универсальной когнитивной аналитической системе «Эйдос-Х++». Как следует из самого названия системы, это сделано в универсальной постановке, не зависящей от предметной области. Поэтому система «Эйдос-Х++» может быть применена и фактически была применена в самых различных предметных областях для построения интеллектуальных измерительных систем и интеллектуальных систем управления, а также для решения задач идентификации, прогнозирования и приятия решений [4].

ЛИТЕРАТУРА

1. **Луценко Е. В.** Системно-когнитивный анализ и система «Эйдос» и их применение для построения интеллектуальных измерительных систем / Заводская лаборатория. Диагностика материалов. 2014. Т. 80. № 5. С. 64 – 74.
2. **Орлов А. И., Луценко Е. В.** Системная нечеткая интервальная математика. Монография (научное издание). — Краснодар: КубГАУ, 2014. — 600 с.
3. **Кульбак С.** Теория информации и статистика. — М.: Наука, 1967. — 408 с.
4. **Луценко Е. В., Коржаков В. Е.** Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2013. № 04(088). С. 340 – 359. — IDA [article ID]: 0881304022. — Режим доступа: http://ej.kubagro.ru/2013/04/pdf/22.pdf
5. **Харкевич А. А.** О ценности информации. — М.: Физматгиз, 1960. Вып. 4. С. 53 – 57.
6. **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). — Краснодар: КубГАУ, 2002. — 605 с.[3]
7. **Луценко Е. В., Трунев А. П.** Коэффициент эмерджентности классических и квантовых статистических систем / По-

[3] Для удобства читателей некоторые источники из списка литературы размещены на сайте: http://lc.kubagro.ru

литематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2013. № 06(090). С. 214 – 235. — IDA [article ID]: 0901306014. — Режим доступа: http://ej.kubagro.ru/2013/06/pdf/14.pdf

8. **Луценко Е. В.** Универсальный информационный вариационный принцип развития систем / Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2008. № 07(041). С. 117 – 193. — Шифр Информрегистра: 0420800012\0091, IDA [article ID]: 0410807010. — Режим доступа: http://ej.kubagro.ru/2008/07/pdf/10.pdf
9. **Луценко Е. В.** Количественные меры возрастания эмерджентности в процессе эволюции систем (в рамках системной теории информации) / Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2006. № 05(021). С. 355 – 374. — Шифр Информрегистра: 0420600012\0089, IDA [article ID]: 0210605031. — Режим доступа: http://ej.kubagro.ru/2006/05/pdf/31.pdf
10. **Луценко Е. В.** Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2013. № 08(092). С. 859 – 883. — IDA [article ID]: 0921308058. — Режим доступа: http://ej.kubagro.ru/2013/08/pdf/58.pdf
11. **Хартли Р.** Передача информации / Теория информации и ее приложения / Пер. с англ.; под ред. А. А. Харкевича. — М.: Физматгиз, 1959. С. 5 – 35.
12. **Луценко Е. В.** Системно-когнитивный анализ как развитие концепции смысла Шенка – Абельсона / Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. — Краснодар: КубГАУ, 2004. № 03(005). С. 65 – 86. — IDA [article ID]: 0050403004. — Режим доступа: http://ej.kubagro.ru/2004/03/pdf/04.pdf
13. **Пойа Д.** Математика и правдоподобные рассуждения. 2-е изд. — М.: Наука, 1975. — 464 с. http://ilib.mccme.ru/djvu/polya/rassuzhdenija.htm