УДК 004.415.2, 004.588

МЯГКИЕ ВЫЧИСЛЕНИЯ В СИСТЕМАХ ОЦЕНКИ И МОНИТОРИНГА ДИНАМИЧЕСКИХ СИСТЕМ

Рыжов Александр Павлович

Доктор технических наук, доцент;
Московский Государственный Университет им. М.В.Ломоносова, Механико-математический факультет,
Кафедра Математической Теории Интеллектуальных Систем;
119991, Москва, Ленинские горы, МГУ, д. 1, Главное здание;
e-mail: alexander.ryjov@gmail.com.

В работе описывается технология информационного мониторинга сложных систем. Приводится содержательная постановка проблемы информационного мониторинга, описываются технологические и научные аспекты разработки систем информационного мониторинга.

Ключевые слова: Нечеткая математика, иерархические системы, оценка и мониторинг нечетких и дискретных динамических систем.

SOFT COMPUTING IN SYSTEMS FOR EVALUATION AND MONITORING OF DYNAMIC SYSTEMS

Ryjov Alexander

Doctor of Engineering Science, associate professor;
Lomonosov Moscow State University, Department of Mathematics and Mechanics,
The Chair of Mathematical Theory of Intelligent Systems;
119991, Moscow, Leninskie Gory,1, Main building;
e-mail: alexander.ryjov@gmail.com.

Information monitoring technology for complex systems has been described in the article. Problem definition, technological and scientific basics are presented and discussed.

Keywords: Fuzzy sets, hierarchical systems, evaluation and monitoring of fuzzy discrete dynamic systems.

Введение

Термин «информационный мониторинг» был введен вначале как некоторый сленг для обозначения специализированных человеко-машинных интеллектуальных информационных систем, предназначенных для оценки состояния некоторого процесса или системы на базе фрагментарной, ненадежной, возможно недостоверной информации о ней и моделирования возможных вариантов развития проблемы в 1991 – 1992 годах и впервые использован в научной печати в 1994 г. В отличие от широко применяемых систем мониторинга (например, экологического, медицинского, технического и т.п.), базирующихся на анализе большого количества результатов показаний различных датчиков, приставка «информационный» подчеркивает, что для интересующих нас систем «входом» является не показание прибора, а некоторая информация (в частности, и показания измерительных устройств тоже, но не только они). Говоря совсем популярно, мы разрешаем системам мониторинга использовать результаты измерений не только технических устройств, но и человека как измерительного прибора.

1. Содержательная постановка задачи

Многие процессы в бизнесе, экономике, политике и других областях, называемых слабо (или плохо) формализуемыми, не возможно представить в виде набора уравнений, автоматов и других математических средств представления и анализа динамических систем, однако специалисты как-то решают задачи оценки состояния процесса и управления им. В общем виде задача заключается в оценке текущего состояния процесса на основе всей доступной информации, построении прогнозов его развития и выработке рекомендаций по управлению исходя из целей, стоящих перед специалистом. Можно привести следующие примеры проблем из различных областей:

- моделирование поведения клиента (маркетинг);
- сегментирование рынков (маркетинг);
- продвижение кандидата на выборах (политология);
- диагностика (медицина);
- оценка кредитного риска (финансовый анализ);
- оценка страховых рисков (страховое дело);
- развитие технологий/научных исследований (наукометрия).

В качестве примера процессов, не являющихся таковыми, можно привести взаимодействие двух тел или распространение колебаний в однородной среде. Имеются математические модели таких процессов, информация измерима и доступна, результат можно вычислить для любого момента времени.

Будем называть задачу оценки текущего состояния системы (процесса) и построении прогнозов ее развития задачей информационного мониторинга, а человеко-компьютерные системы, обеспечивающие информационную поддержку подобного рода информационных задач, системами информационного мониторинга.

2. Понятие систем информационного мониторинга

Основными элементами систем информационного мониторинга являются информационное пространство и аналитик. Существенными являются следующие их свойства.

Информационное пространство представляет собой совокупность различных информационных элементов, которые можно охарактеризовать следующим образом:

- разнородность носителей информации, т.е. фиксация информации в виде статей, газетных заметок, компьютерном виде, аудио - и видеоинформация и т.п.;
- фрагментарность: информация чаще всего относится к какому-либо фрагменту проблемы, причем разные фрагменты могут быть по-разному «покрыты» информацией;
- разноуровневость: информация может относиться ко всей проблеме в целом, к некоторой ее части, к конкретному элементу проблемы;
- различная степень надежности: информация может содержать конкретные данные различной степени надежности, косвенные данные, результаты выводов на основе надежной информации или косвенные выводы;
- возможная противоречивость: информация из различных источников может совпадать, слегка различаться или вообще противоречить друг другу;
- изменяемость во времени: проблема/процесс развивается во времени, поэтому и информация в разные моменты времени об одном и том же элементе проблемы может и должна различаться;
- возможная тенденциозность: информация отражает определенные интересы источника информации, поэтому может носить тенденциозный характер. В частном случае она может являться наме-

ренной дезинформацией (например, для политических проблем или для проблем, связанных с конкуренцией).

Аналитики являются активным элементом системы мониторинга и, наблюдая и изучая элементы информационного пространства, делают выводы о состоянии проблемы и перспективах ее развития с учетом перечисленных выше свойств информационного пространства. Обычно аналитики образуют некоторую структуру (министерство, агентство, консультационную службу, отдел маркетинга фирмы и т.п.). В этом случае каждый аналитик «нижнего уровня» имеет дело с некоторой частью проблемы и работает с элементами информационного пространства, аналитики «более высокого уровня» имеют дело с более крупными фрагментами проблемы или проблемой в целом, и работают уже с выводами предыдущих аналитиков. При этом они могут ознакомиться с выводами более низкого уровня вплоть до элементов информационного пространства.

3. Технологические основы разработки систем информационного мониторинга

Учитывая особенности информации и конкретных способов ее обработки, основные черты технологии информационного мониторинга можно изложить следующим образом.

Эта технология базируется на использовании ряда приемов, позволяющих обрабатывать подобного рода информацию. В частности:

- для реализации возможности обработки информации из разнородных источников, в базе данных системы хранятся как сами документы, так и ссылки на них с оценкой содержащейся в них информации, данной экспертом;
- для возможности обработки фрагментарной информации используется модель проблемы в виде дерева. Ясно, что для сложных проблем такое представление модели является несколько упрощенным, однако при этом достигается хорошая наглядность и простота работы с моделью проблемы;
- обработка разноуровневой информации достигается за счет предоставления пользователю возможности отнести оценку конкретного информационного материала к разным вершинам дерева-модели;
- обработка информации различной степени надежности и обладающей возможной противоречивостью или тенденциозностью достигается за счет использования лингвистических оценок экспертами данной информации;
- изменяемость во времени учитывается фиксацией даты поступления информации при оценке конкретного материала, т.е. время является одним из элементов описания объектов системы.

Таким образом, системы, построенные на базе этой технологии, позволяют иметь развивающуюся во времени модель проблемы на основе оценок аналитиков, подкрепленную ссылками на все информационные материалы, выбранные ими, с общими и частными оценками состояния проблемы или ее аспектов. Использование времени как параметра системы позволяет проводить ретроспективный анализ и строить прогнозы развития проблемы (отвечать на вопросы типа «Что будет, если ... ?»). В последнем случае возникает возможность выделения «критических путей», т.е. таких элементов модели, малое изменение которых может вызвать значительные изменения в состоянии всей проблемы. Знание таких элементов имеет большое практическое значение и позволяет выявить «слабые места» в проблеме на текущий момент времени, разработать мероприятия по блокированию нежелательных ситуаций или провоцированию желательных, т.е. в некоторой степени управлять развитием проблемы в интересах организации, ее отслеживающей.

Итак, системы информационного мониторинга позволяют:

- единообразно обрабатывать разнородную, разноуровневую, фрагментарную, ненадежную, меняющуюся во времени информацию;
- получать оценки состояния проблемы, отдельных ее аспектов;
- моделировать различные ситуации в предметной области мониторинга;

- выявлять «критические пути» развития проблемы, то есть выявлять те элементы проблемы, малое изменение состояния которых может качественно изменить состояние проблемы в целом.

Технологию информационного мониторинга можно назвать технологией информационной поддержки конкретного пользователя, которая, по нашему мнению, является естественным продолжением развития средств информационной поддержки.

4. Математические основы систем информационного мониторинга

Системы информационного мониторинга можно отнести к классу иерархических нечетких дискретных динамических систем. Теоретическую основу такого класса систем составляют методы анализа иерархий, теория нечетких множеств, теория измерений, теория автоматов, дискретная математика, которые были разработаны в работах Саати (Saaty) [12], Месаровича (Messarovich) [14], Заде (Zadeh) [1], Пфанцгаля (Pfanzgal) [5], Кудрявцева В.Б. [3] и других авторов.

Для эффективного практического применения предложенных технологических решений (раздел 3) необходима проработка ряда теоретических проблем, результаты котрой приводятся ниже.

4.1. Формулировка проблем

Предполагается, что эксперт описывает степень противоречивости полученной информации и имеющейся или возможность реализации некотрых процессов в предметной области в виде лингвистических значений. Субъективная степень удобства такого описания зависит от набора и состава таких лингвистических значений. Поясним сказанное на модельном примере.

Пример 1

Пусть нас просят описать рост человека. Рассмотрим две крайние ситуации.

Ситуация 1. Разрешено использовать только два значения: «низкий» и «высокий».

Ситуация 2. Разрашено использовать много значений: «очень низкий», «не очень высокий», … , «не высокий и не низкий», … , «очень-очень высокий».

Ситуация 1 является неудобной. Действительно, для многих людей оба разрешенных значения могут не подходить, и, описывая их, мы выбираем между двумя «плохими» значениями.

Ситуация 2 также является неудобной. Действительно, при описании конкретного человека могут подходить несколько из разрешенных значений. Мы опять испытываем дискомфорт, но теперь от того, что мы вынуждены выбирать между двумя или более «хорошими» значениями. Может ли быть множество лингвистических значений в этом смысле оптимальным?

Предполагается, что система отслеживает развитие проблемы, то есть изменение ее во времени. Предполагается также, что она интегрирует оценки различных экспертов. Это значит, что один объект может описываться разными экспертами. Поэтому желательно иметь гарантии того, что разные эксперты описывают один и тот же объект наиболее «единообразно».

Исходя из сделанных замечаний, мы можем сформулировать первую проблему следующим образом.

Проблема 1. Можно ли, учитывая некоторые особенности восприятия человеком объектов реального мира и их описания, сформулировать правило выбора оптимального множества значений признаков, по которым описываются эти объекты? Возможны два критерия оптимальности:

Критерий 1. Под оптимальными понимаются такие множества значений, используя которые человек испытывает минимальную неопределенность при описании объектов.

Критерий 2. Если объект описывается некоторым количеством экспертов, то под оптимальными понимаются такие множества значений, которые обеспечивают минимальную степень рассогласования описаний.

Технология информационного мониторинга предполагает хранение информационнх материалов (или ссылок на них) и их лингвистических оценок в базе данных системы. В связи с этим возникает следующая проблема.

Проблема 2. Можно ли определить показатели качества поиска информации в нечетких (лингвистических) базах данных и сформулировать правило выбора такого множества лингвистических значений, использование которого обеспечивало бы максимальные показатели качества поиска информации?

Решение сформулированных проблем приведено в разделах 4.2 и 4.3 соответственно.

Показано, что мы можем сформулировать методику выбора оптимального множества значений качественных признаков. Более того, показано, что такая методика является устойчивой, то есть возможные при построении функций принадлежности естественные маленькие ошибки не оказывают существенного влияния на выбор оптимального множества значений. Множества, оптимальные по критериям 1 и 2 совпадают. Данные результаты описаны в разделе 4.2.

Показано, что можно ввести показатели качества поиска информации в нечетких (лингвистических) базах данных и формализовать их. Показано, что возможно сформулировать методику выбора оптимального множества значений качественных признаков, которое обеспечивает максимальные показатели качества поиска информации. Более того, показано, что такая методика является устойчивой, то есть возможные при построении функций принадлежности естественные маленькие ошибки не оказывают существенного влияния на выбор оптимального множества значений. Данные результаты изложены в разделе 4.3.

Важной проблемой при разработке систем информационного мониторинга является задача выбора адекватных операторов агрегирования информации в модели проблемы. Эта задача возникает в силу того, что системы информационного мониторинга ориентированы на обработку разноуровневой фрагментарной информации. Это означает, что при вводе информации в систему пользователь может осуществлять привязку информационных материалов к узлам различных уровней иерархии в модели проблемы (и, соответственно подтверждать/изменять их оценки). Указанное допущение позволяет использовать системы информационного мониторинга при решении значительно более широкого класса практических задач, однако платой за это является необходимость разработки соответствующей теории и создания инструментария выбора адекватных операторов агрегирования информации. Поясним суть проблемы на следующем примере.

Пример 2

Пусть текущее состояние модели имеет вид, изображенный на рис. 1.

Будем для простоты считать, что оценки пользователя представляются в виде числа из отрезка [0,1]. Пусть, далее, в качестве операторов агрегирования информации используется операция взятия максимума. Как не трудно видеть, в этом случае модель является согласованной. Рассмотрим две ситуации.

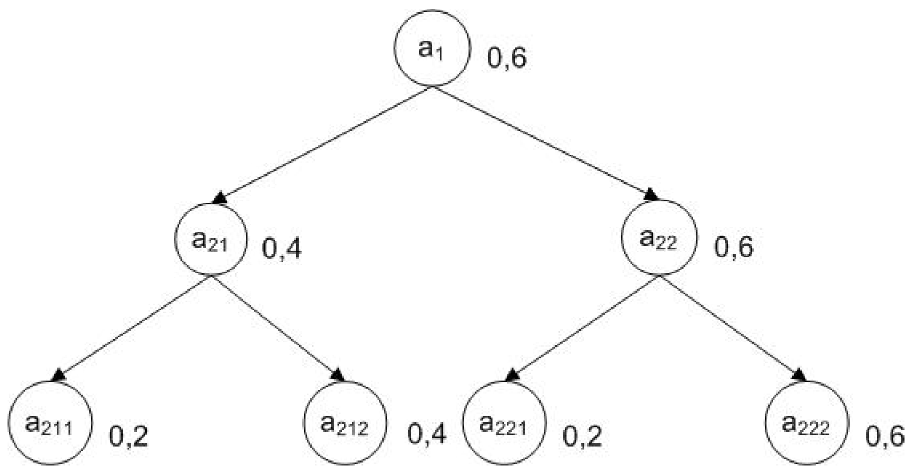

Рис. 1. Пример состояния модели

Ситуация 1. Пользователь изменил значение в узле a_{211} с текущего на 0.3. В этом случае согласованность остается.

Ситуация 2. Пользователь изменил значение в узле a_{22} с текущего на 0.5. В этом случае введенная оценка входит в противоречие как с оценками нижнего уровня, так и с оценкой верхнего уровня. Как быть в этом случае? Если оценки узлов a_1 и a_{212} являются не результатом вычислений в системе, а введены пользователем на основе предыдущей информации, мы должны для сохранения согласованности модели либо «заставить» пользователя изменить их, либо выбрать другие операторы агрегирования информации. Первый путь, несмотря на насильственный характер, иногда оказывается применимым при условии развитой подсистемы работы с архивом (например, в ситуации, когда оценки, которые необходимо менять, являются достаточно «старыми» и/или поставлены пользователем на основе недостаточно надежной информации). Однако, если пользователь одинаково уверен в правильности всех оценок, необходимо менять оператор агрегирования информации.

Проблема 3. Можно ли предложить процедуры выбора операторов агрегирования информации в нечетких иерархических динамических системах, минимизирующих противоречивость модели проблемы/процесса в системах информационного мониторинга?

Можно выделить следующие подходы к решению этой проблемы, базирующиеся на различных интерпретациях операторов агрегирования информации: геометрический, логический и подход на основе обучения, включающий вы себя обучение на основе генетических алгоритмов и обучение на основе нейронных сетей. Описанию этих подходов, формулировке и решению возникающих в их рамках задач посвящен раздел 4.4 настоящей работы.

4.2. Модель описания человеком объектов

Формализация

Моделью множества значений признака может служить полное ортогональное семантическое пространство [9, 10]. Тогда модель описания человеком объектов есть процедура выбора элемента этого пространства при описании объекта.

Напомним основные понятия, необходимые для формулировки результатов.

Частным случаем лингвистической переменной [1], имеющей широкий спектр практических приложений, может быть набор нечетких переменных [1], описывающих некоторое понятие (то есть лингвистическая переменная с фиксированным терм-множеством). Такие структуры можно интерпретировать, в частности, как набор лингвистических значений, описывающих степень противоречивости полученной информации и имеющейся или возможность реализации некоторых процессов. Для таких структур можно потребовать выполнение ряда естественных условий, в рамках которых возможно решение сформулированных выше проблем 1 и 2 (раздел 4.1).

Рассмотрим t нечетких переменных с именами $a_1, a_2, .. , a_t$, заданных на одном универсальном множестве. Будем называть такую совокупность *семантическим пространством* s_t .

Введем систему ограничений для функций принадлежности нечетких переменных, составляющих s_t. Будем для простоты обозначать функцию принадлежности a_j через μ_j. Будем считать, что:

1. $\forall \mu_j \ (1 \le j \le t) \ \exists \ U^1{}_j \ne \varnothing$, где $U^1{}_j = \{u \in U : \mu_j(u) = 1\}$, $U^1{}_j$ есть отрезок или точка;
2. $\forall \ j \ (1 \le j \le t) \ \mu_j$ не убывает слева от $U^1{}_j$ и не возрастает справа от $U^1{}_j$ (так как, согласно 1, $U^1{}_j$ является отрезком или точкой, понятия «слева» и «справа» определяются однозначно).

Требования 1 и 2 довольно естественны для функций принадлежности понятий, образующих семантическое пространство. Действительно, первое означает, что для любого используемого понятия в универсальном множестве существует хотя бы один объект, который является эталонным для данного понятия. Если таких эталонов много, то они расположены рядом, а не «раскиданы» по универсуму. Второе требование означает, что если объекты «близки» в смысле метрики в универсальном множестве, то они также «близки» в смысле принадлежности к некоторому понятию.

Далее нам понадобится наряду с функциями принадлежности использовать и характеристические функции, поэтому потребум выполнения следующего технического условия:

3. $\forall\, j\ (1 \le j \le t)\ \mu_j$ имеет не более двух точек разрыва первого рода.

Будем для простоты обозначать требования 1 – 3 через L.

Введем также систему ограничений для совокупностей функций принадлежности нечетких переменных, образующих s_t. А именно, будем считать, что:

4. $\forall\, u \in U\ \exists\, j\ (1 \le j \le t)\colon \mu_j(u) > 0$;

5. $\forall\, u \in U \sum_{j=1}^{t} \mu_j(u) = 1$.

Требования 4 и 5 также имеют довольно естественную интерпретацию. Требование 4, называемое полнотой, означает, что для любого объекта из универсального множества существует хотя бы одно понятие, к которому он может принадлежать. Это означает, что в нашем семантическом пространстве нет «дырок». Требование 5, называемое ортогональностью, означает, что мы не допускаем использования семантически близких понятий или синонимов, требуем достаточной различимости используемых понятий. Отметим также, что частота выполнения этого требования зависит от используемого метода построения функций принадлежности понятий, образующих семантическое пространство. Так, например, если у нас есть некоторое количество экспертов, мы им предъявляем объект $u \in U$ и допускаем только ответы «Да, $u \in a_j$» и «Нет, $u \notin a_j$» (ответа «Не знаю» не допускается), а в качестве значения функции принадлежности $\mu_j(u)$ берем отношение числа экспертов, ответивших положительно, к общему числу экспертов, то данное требование выполняется автоматически. Заметим также, что все приводимые ниже результаты справедливы при некотором ослаблении требования ортогональности [8], но для его описания необходимо введение ряда дополнительных понятий. Поэтому мы остановимся на этом требовании.

Будем для простоты обозначать требования 4, 5 через G.

Будем называть сематническое пространство, состоящее из нечетких переменных, функции принадлежности которых удовлетворяют требованиям 1 – 3, а их совокупности – требованиям 4 и 5, полным ортогональным семантическим пространством и обозначать его $G(L)$.

Основные результаты

Как видно из примера 1, разные семантические пространства имеют различную степень внутренней неопределенности. Можно ли измерять эту степень неопределенности? Для полных ортогональных семантических пространств ответ на этот вопрос положительный.

Для доказательства этого факта и вывода соответсвующей формулы, нам необходимо ввести ряд дополнительных понятий.

Пусть у нас есть некоторая совокупность из t функций принадлежности $s_t \in G(L)$. Пусть $s_t = \{\mu_1, \mu_2, \dots, \mu_t\}$.

Будем называть совокупность из t характеристических функций $\hat{s}_t = \{h_1, h_2, \dots h_t\}$ *ближайшей совокупностью характеристических функций*, если

$$h_j(u) = \begin{cases} 1, & \text{если } \mu_j(u) = \max\limits_{1 \le i \le t} \mu_i(u) \\ 0, & \text{иначе} \end{cases} \qquad (1 \le j \le t)$$

Не трудно заметить, что, если полное ортогональное семантическое пространство состоит не из функций принадлежности, а из характеристических функций, то при описании объектов в нем не возникает никакой неопределенности. Эксперт однозначно выбирает термин a_j, если объект находится в соответствующей области универсального множества. Несколько экспертов описывают один и тот же объект одним и тем же термином. Эту ситуацию можно проиллюстрировать следующим образом. Допустим, у нас весы достаточной точности и мы имеем возможность взвешивать некоторый материал. Кроме этого, мы договорились, что если вес материала попадает в определенный интервал, то он принадлежит одной из категорий. Тогда мы в точности будем иметь описанную ситуацию.

Проблема заключается в том, что для нашей задачи нет таких весов, а также возможности «взвешивать» на них интересующие нас объекты.

Однако, мы можем предполагать, что из двух семантических пространств имеет меньшую неопределенность то пространство, которое более «похоже» на пространство из совокупностей характеристических функций.

В математике степенью сходства может быть расстояние. Можно ли ввести расстояние между семантическими пространствами? Для полных ортогональных семантических пространств это возможно. Так как мы рассматриваем полные ортогональные семантические пространства, функции принадлежности используемых понятий удовлетворяют условиям L. Это подмножество известного пространства функций, интегрируемых на некотором отрезке U. Известно, что в нем можно ввести расстояние, например, следующим образом:

$$d(f,g)=\int_U |f(u)-g(u)|du\cdot$$

Расстояние в $G(L)$ можно вести с помощью леммы 1.

Лемма 1. Пусть $s_t, s_t' \in G(L)$, $s_t=\{\mu_1, \mu_{2,\dots,}\mu_t\}$, $s_t'=\{\mu_1', \mu_{2,\dots,}'\mu_t'\}$, $d(\mu,\nu)$ – метрика в L. Тогда

$$d(s_t,s_t')=\sum_{j=1}^{t} d\left(\mu_j,\mu_j'\right) \text{ – есть метрика в } G(L).$$

С доказательством леммы можно ознакомиться в [10].

Пусть $s_t \in G(L)$. Под мерой неопределенности s_t будем понимать значение функционала $\xi(s_t)$, определенного на элементах $G(L)$ и принимающего значения в [0,1] (то есть $\xi: G(L)\rightarrow [0,1]$), удовлетворяющего следующим условиям (аксиомам):

A1. $\xi(s_t)=0$, если s_t представляет собой совокупность характеристических функций;

A2. Пусть $s_t, s'_{t'} \in G(L)$, t и t' могут быть равны или не равны друг другу. Тогда $\xi(s_t)\le\xi(s'_{t'})$, если $d(s_t,\hat{s}_t)\le d(s'_{t'},\hat{s}'_{t'})$.

(Напомним, что $\hat{s}_t$ – ближайшая к s_t совокупность характеристических функций.)

Существуют ли такие функционалы? Ответ на этот вопрос дает следующая теорема.

Теорема 1 (существования). Пусть $s_t \in G(L)$. Тогда функционал:

$$\xi(s_t)=\frac{1}{|U|}\int_U f(\mu_{i_1^*}(u)-\mu_{i_2^*}(u))du, \tag{1}$$

где $\mu_{i_1^*}(u)=\max\limits_{1\le j\le t}\mu_{i_2^*}(u)$, $\mu_{i_2^*}(u)=\max\limits_{1\le j\le t, j\ne i}\mu_{j^*}(u)$, (2)

f удовлетворяет следующим условиям:

F1. $f(0)=1, f(1)=0$;

F2. f не убывает, является мерой неопределенности s_t, то есть удовлетворяет аксиомам A1 и A2.

Доказательство теоремы существования приведено в [10].

Существует много функционалов, удовлетворяющих условиям Теоремы 3.1. Они детально описаны в [8]. Простейшим таким функционалом является функционал, для которого функция f есть линейная функция. Не трудно видеть, что условиям F1 и F2 удовлетворяет единственная линейная функция $f(x) = 1-x$. Подставляя ее в (1), мы получаем следующую простейшую меру неопределенность для полного ортогонального семантического пространства:

$$\xi(s_t)=\frac{1}{|U|}\int_U (1-(\mu_{i_1^*}(u)-\mu_{i_2^*}(u)))du, \tag{3}$$

где $\mu_{i_1^*}(u)$, $\mu_{i_2^*}(u)$ определяются соотношениями (2).

Обозначим подынтегральную функцию в (3) через $\eta(s_t,u)$:

$$\eta(s_t, u) = 1 - (\mu_{i_1^*}(u) - \mu_{i_2^*}(u) . \qquad (4)$$

Тогда мы можем привести следующую интерпретацию меры неопределенности (3).

Интерпретация. Рассмотрим процесс описания объектов в рамках семантического пространства $s_3 \in G(L)$ (рис. 2).

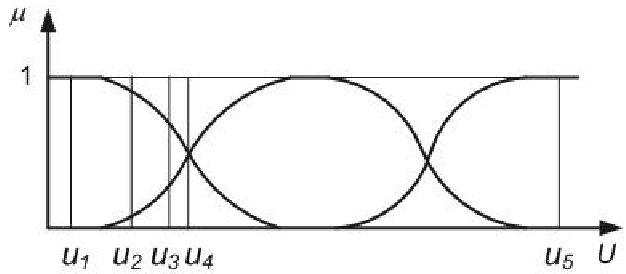

Рис. 2. Пример семантического пространства s_3

Для объектов u_1 и u_5, человек без всяких колебаний выбирает один из термов (a_1 и a_3 соответственно). При описании объекта u_2 человек начинает выбирать между термами a_1 и a_2. Такого рода колебания возрастают и достигают своего пика при описании объекта u_4: для этой точки термы a_1 и a_2 неразличимы. Если вспомнить процедуру построения функций принадлежности, описанную при анализе свойства ортогональности, мы можем также утверждать, что все эксперты будут единодушны при описании объектов u_1 и u_5, иметь некоторую рассогласованность при описании u_2 которая будет возрастать и достигнет своего пика для объекта u_4.

Теперь рассмотрим формулу (4). Не трудно видеть, что

$$0 = \eta(s_t, u_5) = \eta(s_t, u_1) < \eta(s_t, u_2) < \eta(s_t, u_3) < \eta(s_t, u_4) = 1.$$

Таким образом, $\eta(s_t,u)$ действительно отражает степень неопределенности, которую испытывает человек при описании объектов в рамках семантического пространства или степень рассогласования мнения экспертов при таком описании.

Тогда степень неопределенности $\xi(s_t)$ (3) есть усредненная степень неопределенности процесса описания всех объектов универсального множества.

Для изучения свойств степени нечеткости (3) определим следующие подмножества множества функций L:

$\overline{L}$ – множество кусочно-линейных функций из L, которые являются линейными на

$\overline{U} = \{u \in U : \forall j (1 \le j \le t)\ \ 0 < \mu_j(u) < 1\}$.

$\hat{L}$ – множество кусочно-линейных на U (включая $\overline{U}$) функций из L.

Справедливы следующие теоремы [10].

Теорема 2. Пусть $s_t \in G(\overline{L})$. Тогда $\xi(s_t) = \dfrac{d}{2|U|}$, где $d = |\overline{U}|$.

Теорема 3. Пусть $s_t \in G(\hat{L})$. Тогда $\xi(s_t) = \dfrac{d}{2|U|}$, где $c < 1$, $c = \text{Const}$.

Так как любая $s_t \in G(L)$ может быть (со сколь угодно большой точностью) аппроксимирована системой множеств $s_t \in G(\hat{L})$, то данное соотношение является справедливым для любой $s_t \in G(L)$.

Пусть g – некоторая взаимно-однозначная функция, определенная на U. Эта функция индуцирует преобразование некоторого полного ортогонального семантического пространства $s_t \in G(L)$, определенное на универсуме U в полное ортогональное семантическое пространство $g(s_t)$, определенное на универсуме U', где $U' = g(U) = \{u' = g(u), u \in U\}$.

Данное преобразование может быть определено следующим образом: $g(s_t)$ – есть множество функций принадлежности $\mu_j'(u)$, где $\mu_j'(u') = \mu_j'(g(u)) = \mu_j(g^{-1}(u')) = \mu_j(u)$, $(\mu_j(u) \in s_t$, $1 \le j \le t)$.

Это определение можно проиллюстрировать следующим примером.

Пример 3. Пусть $s_t \in G(L)$, U – универсум s_t и g – есть растяжение (сжатие) универсума U. Тогда $g(s_t)$ есть множество функций, получаемых из s_t таким же растяжением (сжатием).

Теорема 4. Пусть $s_t \in G(L)$, U – универсум s_t, g – некоторая линейная взаимно-однозначная функция, определенная на U и $\xi(s_t) \ne 0$. Тогда $\xi(s_t) = \xi(g(s_t))$.

Доказательство теоремы 4 приведено в [10].

Данное свойство означает, что человек описывает разнородные объекты, используя некоторое множество лингвистических значений, с одинаковыми трудностями, если физические параметры объектов одного типа могут быть получены из параметров объектов другого типа некоторым линейным преобразованием. Например, используя термины {*высокий, средний, низкий*} мы описываем людей, деревья, здания и т.п. с одинаковыми трудностями; используя множество термов {*очень близко, близко, не близко, очень далеко*} мы описываем расстояние между молекулами, расстояние между улицами в городе, расстояние между городами на карте с одинаковыми трудностями.

Важным аспектом практического использования любой модели является ее устойчивость. Достаточно очевидно, что при практическом применении любой модели и идентификации ее параметров (в нашем случае – при построении функций принадлежности) возникают (маленькие) ошибки измерения. Если модель чувствительна к такого рода естественным ошибкам, вопрос ее практического применения является достаточно проблематичным.

Рассмотрим ситуацию, когда функции принадлежности в семантическом пространстве заданы не абсолютно точно, а с некоторой «точностью» δ (рис. 3). Будем называть эту ситуацию δ-моделью и обозначать $G^{\delta}(L)$.

В этой ситуации мы можем вычислить верхние и нижние оценки степени неопределенности. Совокупности нечетких множеств из $G^{\delta}(L)$, имеющие минимальную и максимальную степень неопределенности представлены на рис. 4 (a) и 4 (b) соответственно.

Теорема 5. Пусть $s_2 \in G^{\delta}(\overline{\boldsymbol{L}})$. Тогда $\underline{\xi}(s_2) = \dfrac{d(1-\sigma_2)^2}{2|U|}$, $\overline{\xi}(s_2) = \dfrac{d(1+2\sigma_2)}{2|U|}$.

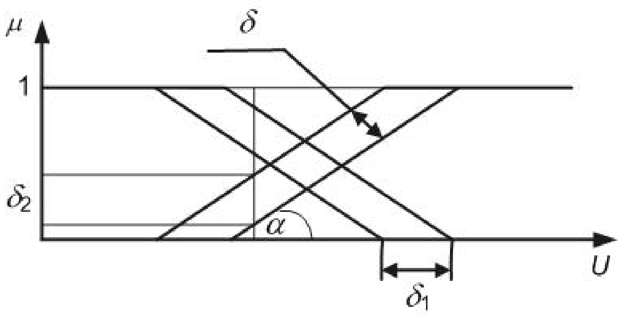

Рис. 3. δ-модель

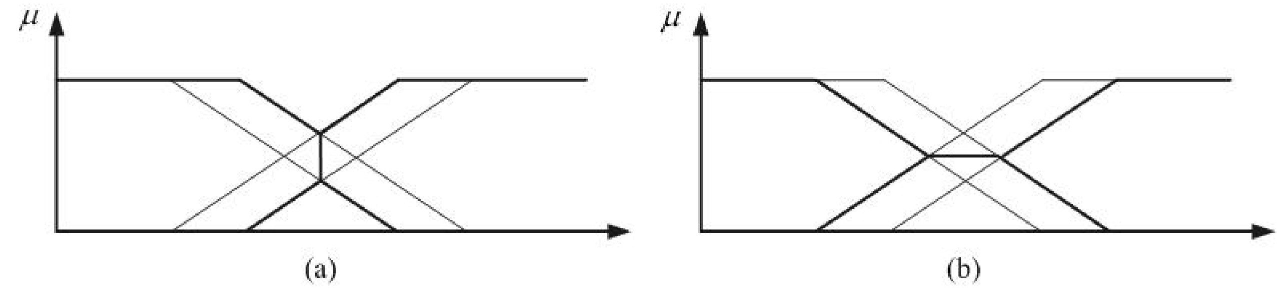

Рис. 4. Минимальная и максимальная степень нечеткости в δ-модели

Теорема 6. Пусть $s_t \in G^{\delta}(\overline{\boldsymbol{L}})$. Тогда $\underline{\xi}(s_t) = \dfrac{D(1-\sigma_2)^2}{2|U|}$, $\overline{\xi}(s_t) = \dfrac{D(1+2\sigma_2)}{2|U|}$, где

$$D = \sum_{j=1}^{t-1} d_{j,j+1}, \; d_{j,j+1} = \overline{u}_{jR} - \overline{u}_{j+1,L}.$$

Сравнивая теорему 2 и теоремы 5 и 6, мы видим, что для малых значений σ основные закономерности нашей модели сохраняются. Таким образом, мы можем использовать степень неопределенности $\xi(s_t)$ при решении практических задач, поскольку она обладает свойством стабильности.

Базируясь на этих результатах, мы можем предложить следующую методику выбора оптимального множества значений качественных признаков.

1. Генерируются все «разумные» множества значений лингвистической переменной.
2. Каждое из таких множеств представляется в форме полного ортогонального семантического пространства.
3. Для каждого из них вычисляется мера неопределенности (3).
4. В качестве оптимального множества значений как с точки зрения минимизации неопределенности описания объектов, так и с точки зрения минимизации степени рассогласования мнений экспертов, выбирается то множество, мера неопределенности которого минимальна.

Следуя этой методике, мы можем описывать объекты с минимально возможной неопределенностью, то есть гарантировать оптимальность свойств систем информационного мониторинга с этой точки зрения.

4.3. Модель поиска информации в нечетких базах данных

Формализация

Так же, как и в разделе 4.2, будем считать, что множество лингвистических значений задано в виде $G(L)$.

Можно ввести понятия потерь информации ($\Pi_X(U)$) и информационных шумов ($H_X(U)$), возникающих при поиске информации в лингвистических базах данных. Смысл этих понятий следующий. При общении с системой пользователь формулирует запрос, содержащий определенные значения лингвистических признаков, и получает ответ на запрос. Если бы он мог знать физические (не лингвистические) значения признаков, он, возможно, не принял бы некоторые записи из ответа на запрос (такие записи составляют информационный шум); если бы он имел возможность при этом «видеть» всю базу данных, он, возможно, дополнил бы некоторыми записями ответ на свой запрос (такие записи составляют потери информации). Такого рода потери информации и шумы порождаются нечеткостью лингвистических описаний объектов.

Эти величины описываются следующей формулой [7]:

$$\Pi_X(U) = \mathrm{H}_X(U) = \frac{1}{|U|}\sum_{j=1}^{t-1}(p_j + p_{j+1}) \times \int_U \mu_j(u)\mu_{j+1}(u)N(u)du, \tag{5}$$

где $X = \{a_1, \dots, a_t\}$ – набор значений признака, p_i ($i = 1, 2, \dots, t$) – вероятность запроса по i-значению признака.

Основные результаты

Справедливы следующие теоремы [7].

Теорема 7. Пусть $s_2 \in G(\overline{L})$, $N(u) = N = Const$. Тогда $\Pi_X(U) = \mathrm{H}_X(U) = \dfrac{Nd}{6|U|}$, где $d = |U_{12}|$.

Следствие 1. Пусть выполняются условия теоремы 7. Тогда $\Pi_X(U) = \mathrm{H}_X(U) = \dfrac{N}{3}\xi(s_2)$.

Теорема 8. Пусть $s_t \in G(\overline{L})$, $N(u) = N = Const$ и $p_j = \dfrac{1}{t}$ (j=1, …, t). Тогда $\Pi_X(U) = \mathrm{H}_X(U) = \dfrac{ND}{3t|U|}$, где $D = \sum_{j=1}^{t-1} d_{j,j+1}$, $d_{j,j+1} = |U_{j,j+1}|$.

Следствие 2. Пусть выполняются условия теоремы 8. Тогда $\Pi_X(U) = \mathrm{H}_X(U) = \dfrac{Nd}{3t}\xi(s_t)$.

Теорема 9. Пусть $s_2 \in G(L)$, $N(u) = N = Const$. Тогда $\Pi_X(U) = \mathrm{H}_X(U) = c\xi(s_2)$, где c – константа, зависящая только от N.

Теорема 10. Пусть $s_t \in G(L)$, $N(u) = N = Const$ и $p_j = \dfrac{1}{t}$ ($j = 1, \dots, t$). Тогда $\Pi_X(U) = \mathrm{H}_X(U) = \dfrac{c}{t}\xi(s_t)$, где c –константа, зависящая только от N.

Рассмотрим, как и в разделе 4.2, δ-модель. Справедливы следующие теоремы [15].

Теорема 11. Пусть $X = \{a_1, a_2\}$, $s_2 \in G^\delta(\overline{L})$, $N(u) = N = Const$. Тогда:

$$\underline{\Pi}_X(U) = \underline{\mathrm{H}}_X(U) = \frac{Nd(1-\sigma_2)^3}{6|U|}.$$

Следствие 3. Пусть выполняются условия теоремы 11. Тогда $\underline{\Pi}_X(U) = \underline{\mathrm{H}}_X(U) = \dfrac{N(1-\sigma_2)}{3}\underline{\xi}(s_2)$.

Теорема 12. Пусть $X = \{a_1, a_2\}$, $s_2 \in G^\delta(\overline{L})$, $N(u) = N = Const$. Тогда

$$\overline{\Pi}_X(U) = \overline{\Pi}_X(U) = \frac{Nd(1-\sigma_2)^3}{6|U|} + \frac{Nd\sigma_2}{|U|}.$$

Следствие 4. Пусть выполняются условия теоремы 12. Тогда

$$\overline{\Pi}_X(U) = \overline{\mathrm{H}}_X(U) = \frac{N\left(6\sigma_2 + (1-\sigma_2)^3\right)}{3(1+2\sigma_2)}\overline{\xi}(s_2).$$

Теорема 13. Пусть $X = \{a_1, a_2, \dots, a_t\}$, $s_t \in G^\delta(\overline{L})$, $N(u) = N = Const$ и $p_j = \dfrac{1}{t}$ (*j = 1, ..., t*). Тогда

$$\underline{\Pi}_X(U) = \underline{\mathrm{H}}_X(U) = \frac{ND(1-\sigma_2)^3}{3t|U|}, \text{ где } D = \sum_{j=1}^{t-1} d_{j,j+1},\quad d_{j,j+1} = |U_{j,j+1}|.$$

Следствие 5. Пусть выполняются условия теоремы 13. Тогда $\underline{\Pi}_X(U) = \underline{H}_X(U) = \frac{2N}{3t}\left(1-\sigma_2\right)\underline{\xi}\left(s_t\right)$.

Теорема 14. Пусть $X = \{a_1, a_2, \ldots, a_t\}$, $s_t \in G^\delta(\overline{L})$, $N(u) = N = Const$ и $p_j = \frac{1}{t}$ ($j=1, \ldots, t$). Тогда

$$\overline{\Pi}_X(U) = \overline{\Pi}_X(U) = \frac{ND\left(1-\sigma_2\right)^3}{3t|U|} + \frac{2ND\sigma_2}{t|U|}, \text{ где } D = \sum_{j=1}^{t-1} d_{j,j+1}, \quad d_{j,j+1} = |U_{j,j+1}|.$$

Следствие 6. Пусть выполняются условия теоремы 14. Тогда

$$\overline{\Pi}_X(U) = \overline{\Pi}_X(U) = \frac{2N}{t\left(1+2\sigma_2\right)}\left[\frac{\left(1-\sigma_2\right)^3}{3} + 2\sigma_2\right]\overline{\xi}\left(s_t\right).$$

Из приведенных теорем не трудно видеть, что при малых значениях σ основные закономерности модели информационного поиска в нечеткой среде сохраняются. Это означает, мы можем использовать как оценку степени неопределенности, так и модель информационного поиска в нечетких (лингвистических) базах данных при решении практических задач.

Таким образом, мы можем утверждать, что объемы потерь информации и информационных шумов, возникающие при поиске информации в нечетких (лингвистических) базах данных согласуются со степенью неопределенности описания объектов. Кроме этого, объемы потерь информации и информационных шумов и их связь с мерой неопределенности описания объектов являются устойчивыми. Это означает, что мы можем использовать предложенную модель при решении практических задач.

4.4. Модели агрегирования информации в иерархических системах

Формализация

Рассмотрим некоторую не концевую вершину d_{j_0} с подчиненными ей (в смысле рассматриваемого дерева *D*) вершинами $d_{j_1}, d_{j_2}, \ldots, d_{j_{N_0}}$. Тогда оператор агрегирования информации (ОАИ) O_{j_0} есть функция, определенная на множестве всех возможных значений подчиненных вершин и принимающая значения в множестве X_{j_0}: $O_{j_0} : X_{j_1} \times X_{j_2} \times \cdots \times X_{j_{N_0}} \to X_{j_0}$.

Множества X_j ($j = 0, \ldots, N_D$) представляет собой набор лингвистических значений $a_j^1, a_j^2, \ldots, a_j^{Nj}$.

Обозначим множество операторов агрегирования информации для вершины d_{j_0} через $M\left\lfloor O_{j_0}\right\rfloor$ Ясно, что для конкретного элемента модели проблемы d_{j_0} число возможных ОАИ является большим: $\left|M\left[O_{j_0}\right]\right| = \left|X_{j_0}\right|^{\left|X_{j_1}\right| \times \left|X_{j_2}\right| \times \cdots \times \left|X_{j_{N_0}}\right|}$.

Нашей задачей является выбор конкретного оператора $o_j \in M\left\lfloor O_j\right\rfloor$ для всех не концевых вершин d_j дерева-модели *D*. Этот выбор базируется на некоторой информации I_j об «идеальном» ОАИ $\hat{o}_j \in M\left\lfloor O_j\right\rfloor$. Эта информация представляет собой два множества: $I_j = I_j^{(1)} \cup I_j^{(2)}$, где $I_j^{(1)}$ – множество высказываний экспертов о «правильном поведении» $\hat{o}_j$; $I_j^{(2)}$ – множество результатов работы выбранного ОАИ.

Примерами элементов $I_j^{(1)}$ могут служить высказывания типа «При $d_{j_1} = a_{j_1}{}^1$ и $d_{j_2} = a_{j_2}{}^1$ и ... и $d_{j_{N_0}} = a_{j_{N_0}}{}^1$, значение $d_{j_0} = a_{j_0}{}^1$», «При сильном возрастании d_{j_1} значение d_{j_0} убывает», «Значение d_{j_0} монотонно изменяется по всем аргументам» и т.п.

Множество $I_j^{(2)}$ представляет собой таблицу вида табл. 1.

Таблица 1

$(d_{j_1}, d_{j_2}, \ldots, d_{j_{N_0}})$	d_{j_0}
$a_{j_1}{}^{1},\ a_{j_2}{}^{1},\ \ldots\ a_{j_{N_0}}{}^{1}$	$\tilde{o}(a_{j_1}{}^{1},\ a_{j_2}{}^{1},\ \ldots\ a_{j_{N_0}}{}^{1})$
$a_{j_1}{}^{1},\ a_{j_2}{}^{1},\ \ldots\ a_{j_{N_0}}{}^{2}$	$\tilde{o}(a_{j_1}{}^{1},\ a_{j_2}{}^{1},\ \ldots\ a_{j_{N_0}}{}^{2})$
. . .	. . .
$a_{j_1}{}^{1},\ a_{j_2}{}^{1},\ \ldots\ a_{j_{N_0}}{}^{N_{N_0}}$	$\tilde{o}(a_{j_1}{}^{1},\ a_{j_2}{}^{1},\ \ldots\ a_{j_{N_0}}{}^{N_{N_0}})$
. . .	. . .
$a_{j_1}{}^{N_{j1}},\ a_{j_2}{}^{N_{j2}},\ \ldots\ a_{j_{N_0}}{}^{N_{N_0}}$	$\tilde{o}(a_{j_1}{}^{N_{j1}},\ a_{j_2}{}^{N_{j2}},\ \ldots\ a_{j_{N_0}}{}^{N_{N_0}})$

В левом столбце таблицы расположены все попарно различные значения $d_{j_1}, d_{j_2}, \ldots, d_{j_{N_0}}$, в правом – значения d_{j_0}, полученные на основе информации $I_j^{(1)}$ (то есть для этих строк $\tilde{o}_j = \hat{o}_j$), или значения d_{j_0}, полученные в ходе работы с системой, или пустые значения.

В начале работы с системой таблица содержит только значения d_{j_0} первого типа (полученные на основе информации $I_j^{(1)}$). По мере получения и оценки пользователем информации таблица заполняется на основе вычислений для выбранного оператора $\hat{o}_j$ до момента, когда пользователь не согласен с «теоретическим» значением d_{j_0}. Если такое противоречие не возникает, значит оператор $\hat{o}_j$ выбран удачно и является адекватным ОАИ для данного узла дерева-модели. Если противоречие возникает – необходимо повторить процедуру выбора адекватного ОАИ, но на основе дополненной и, быть может, уточненной с экспертом информации $I_j'^{(1)}$ и $I_j^{(2)}$. Этот процесс повторяется до тех пор, пока вся таблица не будет заполнена. Полученная заполненная таблица и есть адекватный ОАИ для данной вершины d_{j_0}.

Основные результаты

Описанная укрупненная схема выбора ОАИ допускает несколько реализаций в зависимости от интерпретаций множества X_j. Мы можем интерпретировать данное множество как набор дискретных или как набор нечетких значений элемента d_j. Эта интерпретация зависит от свойств предметной области (проблемы). Так, если при оценке информации набора значений $a_j^1, a_j^2, \ldots, a_j^{Nj}$ достаточно (ситуации, когда оценка находится «между» соседними значениями, либо отсутствуют, либо их мало), то мы можем говорить о дискретной модели выбора ОАИ. Если же часто при оценке поступающей в систему информации возникают ситуации, когда оценка находится «между» соседними значениями a_j^i, a_j^{i+1}, причем пользователь может говорить, что она более близка, например, к a_j^i, чем к a_j^{i+1}, то мы должны использовать нечеткую модель выбора ОАИ. Заметим, что дискретная модель является частным случаем нечеткой и может быть получена при замене соответствующих нечетких множеств на множества уровня 0,5 для полных ортогональных семантических пространств. Мы ее выделяем, однако, потому, что в этом случае возможна разработка специальных алгоритмов выбора ОАИ. Заметим также, что, как показывает практика, дискретная модель является допустимой при степени нечеткости $\xi(d_j) \le 0{,}2$ [15].

В зависимости от доступной информации $I_j^{(1)}$ в рамках дискретной модели можно выделить два подхода к выбору ОАИ: геометрический и логический. Первый применим тогда, когда эксперт может только определить значение $\hat{o}_j$ на некоторых наборах значений $(d_{j_1}, d_{j_2}, \ldots, d_{j_{N_0}})$. Второй – когда кроме этого возможна формулировка некоторых условий на «поведение» $\hat{o}_j$. Эти подходы описаны [1, 6, 7].

В рамках нечеткого подхода также предполагается, что известны значения $\hat{o}_j$ на некоторых наборах значений $(d_{j_1}, d_{j_2}, \ldots, d_{j_{N_0}})$. Эта информация представляется в виде логических высказываний вида «Если $d_{j_1} = a_{j_1}^{k_1}$ и $d_{j_2} = a_{j_2}^{k_2}$ и ... и $d_{j_{N_0}} = a_{j_{N_0}}^{k_{N_0}}$, то $d_{j_0} = \hat{o}_{j_0}(a_{j_1}^{k_1}, a_{j_2}^{k_2}, \ldots a_{j_{N_0}}^{k_{N_0}})$» для всех известных наборов $(d_{j_1}, d_{j_2}, \ldots, d_{j_{N_0}})$. Такое представление $I_j^{(1)}$ позволяет использовать обучение либо на основе генетических алгоритмов (при выполнении достаточно естественных условий монотонности, коммутативности и ассоциативности ОАИ), либо на основе нейронных сетей (при не выполнении данных условий). Эти подходы описаны в [7, 11].

Таким образом, в зависимости от свойств предметной области системы информационного мониторинга и доступности экспертных заключений о свойствах операторов агрегирования информации, возможно применение различных процедур выбора таких операторов, обеспечивающих адекватность системы процессам, происходящим в предметной области.

Список литературы

1. Ананич И.С., Беленький А.Г., Пронин Л.Б., Рыжов А.П. Агрегирование информации в системах информационного мониторинга // Труды Международного семинара «Мягкие вычисления – 96» . – Казань, 1996. – С. 22-26.

2. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приблизительных решений. – М.: Мир, 1976. – С. 165.

3. Кудрявцев В.Б., Алешин С.В., Подколзин А.С. Введение в теорию автоматов. – М.: Наука, 1985. – С. 320.

4. Пфанцгаль И. Теория измерений / пер. с англ. – М.: Мир, 1976. – С. 263.

5. Рогожин С.В., Рыжов А.П. О нечетко заданных классах функций k-значной логики // V Всероссийская конференция «Нейрокомпьютеры и их применение»: сборник докладов. – Москва, 1999. – С. 460-463.

6. Рыжов А.П. Модели поиска информации в нечеткой среде. – М.: Издательство Центра прикладных исследований при механико-математическом факультете МГУ, 2004. – С. 96.

7. Рыжов А.П. Об агрегировании информации в нечетких иерархических системах // Интеллектуальные системы. – М.: МНЦ КИТ, 2001. – Т. 6. – Вып. 1-4. – С. 341.

8. Рыжов А.П. О степени нечеткости размытых характеристик // Математическая кибернетика и ее приложения в биологии / под ред. Л.В.Крушинского, С.В.Яблонского, О.Б.Лупанова. – М.: Изд-во МГУ, 1987. – С. 60-77.

9. Рыжов А.П. Оценка степени нечеткости и ее применение в системах искусственного интеллекта. // Интеллектуальные системы. – М.: МНЦ КИТ, 1996. – Т.1. – Вып.1-4. – С. 95-102.

10. Рыжов А.П. Элементы теории нечетких множеств и измерения нечеткости. – М.: Диалог-МГУ, 1998. – С. 116.

11. Рыжов А.П., Федорова М.С. Генетические алгоритмы в задаче выбора операторов агрегирования информации в системах информационного мониторинга // V Всероссийская конференция «Нейрокомпьютеры и их применение»: сборник докладов. – Москва, 1999 г. – С. 267-270.

12. Саати Т. Анализ иерархических процессов. – М.: Радио и связь, 1993. – С. 315.

13. Fuzzy Logic and Intelligent Technologies in Russia // Ed. by Da Ruan, Gert de Cooman, Alexander Ryjov. – Belgium, SCK*CEN BLG-699, 1995. – P. 285.

14. Messarovich M.D., Macko D., Takahara Y. Theory of hierarchical multilevel systems. – Academic Press, N.Y. – London, 1970. – P. 344.

15. Ryjov, A., Belenki, A., Hooper, R., Pouchkarev, V., Fattah, A. and Zadeh, L.A. Development of an Intelligent System for Monitoring and Evaluation of Peaceful Nuclear Activities (DISNA). – IAEA, STR-310, Vienna, 1998. – P. 122.