

Нечеткие системы: обзор итогов и тенденций развития

Аннотация. В статье приводится результат анализа итогов и тенденций развития нечетких систем как динамично развивающегося научного направления. Основу для такого анализа составили международные конгрессы IFSA'05 (КНР, г. Пекин) и IFSA'07 (Мексика, г. Канкун).

Ключевые слова: Ключевые слова: нечеткие системы, нейронные сети, генетические алгоритмы, системы принятия решений.

Введение

Нечеткая логика как научное направление отметила в 2005 г. свое 40-летие. Ученые, занимающиеся нечеткой логикой, объединены в International Fuzzy Systems Association (IFSA), которая проводит один раз в два года мировые конгрессы. Материалы таких конгрессов естественным образом подытоживают работы за два прошедших года. Многие тенденции, разумеется, остаются значимы не в течение двух лет, а в течение больших периодов (5-7 лет). Ранее в статье [13] были проанализированы материалы конгрессов IFSA'97 и IFSA'03 и сделаны некоторые прогнозы развития. С момента выхода статьи прошло 4 года, в течение которых состоялись 2 конгресса IFSA: IFSA'05 (КНР, г. Пекин) [15] и IFSA'07 (Мексика, г. Канкун) [16]. Особым событием стала конференция BISCSE'05 (Berkeley Initiative Soft Computing Special Event) in Honor of Prof. Lotfi A. Zadeh [14], состоявшая в Беркли (США, Калифорния) и посвященная 40-летию нечеткой логики. Материалы перечисленных конференций представляют собой исходные данные для анализа основных результатов и тенденций развития нечеткой логики.

1. Два конгресса IFSA и конференция BISCSE'05

Конгресс IFSA'05 состоялся 28-31 июля 2005 в Пекине (Китай), а IFSA'07 – с 18 по 21

июня 2007 г. в г. Канкун (Мексика). Каждый из конгрессов выпустил серьезные многотомные материалы. Материалы IFSA'05 представлены 3 томами опубликованных статей, в том числе 5 статьями приглашенных пленарных докладчиков [15]: Л. Заде, Г. Клира, Д. Дюбуа, В. Редрича, И. Лиу и Д. Жанга. В Канкуне все материалы были опубликованы издательством Springer в 3-х томах - Lectures in Artificial Intelligence (1 том) и Advances in Soft Computing 41,42 (2 тома). На IFSA'07 были сделаны 7 пленарных докладов: Л. Заде, Р. Югера, К. Сироты, Я. Капржика, В. Редрича, Дж. Менделя, Г. Чена.

К сожалению, ранее проводимые школы в рамках мировых конгрессов ни в Пекине, ни в Канкуне не состоялись. Но очень интересной новой традицией конгрессов стали круглые столы, посвящаемые базовым задачам нечеткой логики.

Особый характер приобретали круглые столы конференции BISCSE'05, так как конференция была приурочена к 40-летию нечеткой логики, а точнее, пионерской статьи Л. Заде [17]. Именно на этой конференции участники круглых столов старались не только сформулировать основные результаты развития нечеткой логики, но и осветить их в широком научном контексте, в том числе учитывая 40-летнюю историю самой нечеткой науки.

Два конгресса, хоть и являются мероприятиями одной ассоциации, очень отличались друг от друга. Конгресс в Пекине был юбилей-

ным не только из-за 40-летия знаменитой статьи Л. Заде, но и в связи с 20 – летием IFSA. Конгресс имел очень широкую географию участников и обширный охват всех теоретических и прикладных областей исследований (участники из 41 страны, 340 зарегистрированных делегатов и 620 принятых докладов). На конгрессе были проведены специальные юбилейные заседания, которым организаторы придали некоторые черты торжественных «китайских» церемоний. Конгресс был отлично организован и прошел на базе современного Синхуа - университета.

Конгресс в Канкуне был ограничен более узким количеством участников (Заявлено 400 докладов, отобрано для участия в работе конгресса около 270 работ).

2. Сравнение содержания пленарных докладов

Особый интерес на конгрессах вызывают доклады пленарных докладчиков. Сам выбор участников обычно позволяет уловить самую популярную тему в развитии нечеткой логики в настоящий момент. Естественно, представляются доклады и по фундаментальным вопросам, разрабатываемым десятки лет. Но обычно присутствует некоторая модная тема, обещающая стать едва ли не новой парадигмой развития.

Пленарные доклады представлены в Табл.1 и Табл.2.

Табл.1. Пленарные доклады IFSA'05

Автор	Наименования докладов
Л. Заде	Toward a Computational Theory of Precisation of Meaning Based on Fuzzy Logic – The Concept of Cointensive Precisation. Перспективы вычислительной теории уточнения смысла основанной на нечеткой логике – концепция согласованной точности
И. Лиу, Д. Жанг	Order Structure, Topology and Fuzzy Sets. Упорядоченные структуры, топология и нечеткие множества
Дж. Клир	From Natural Language to Formalized Language and Back. От естественного языка к формализованному языку и обратно
Д. Дюбуа	On the Links between Probability and Possibility Theories. О связи между теориями вероятностей и возможностей.
В. Педрич	Knowledge-based Clustering for Human-Centric Systems. Кластеризация, основанная на знаниях, для гуманистических систем

Табл.2. Пленарные доклады IFSA'07

Автор	Наименование
Л. Заде	Fuzzy Logic as the Logic of Natural Languages. Нечеткая логика как логика естественных языков
Р. Ягер	Computing with Words and Granules . Вычисления со словами и гранулами
К. Хирота	Control/Robotics Applications based on Soft Computing Technology. Роботика и контроллеры, основанные на технологии мягких вычислений
Я. Капржык	Computing with words, usuality qualification and linguistic quantifiers: tools for human-centric computing. Вычисление со словами, квалификация полезности и лингвистические квантификаторы: инструменты для гуманистических вычислений
В. Педрич	Dynamic and Distributed (D^2) Fuzzy Modeling. Динамичное и распределенное нечеткое моделирование
Дж. Мендель	Novel Weighted Averages as a Computing With Words Engine. Метод взвешенных средних как машина вычисления со словами
Г. Чен	Uncertainty and Fuzziness in Knowledge Discovery from Large Databases. Неопределенность и нечеткость в интеллектуальном анализе данных в больших базах данных
Дж. Келлер (приглашенный лектор)	OWA Operators for Gene Product Similarity, Clustering, and Knowledge Discovery. OWA операторы для генетического поиска подобия, кластеризации и интеллектуального анализа данных

2.1. Основные положения докладов Л. Заде. Комментарии

Ключевыми моментами конгрессов, конечно, являются доклады Л. Заде, как отца-основателя нечеткой логики. Его доклады очень хорошо иллюстрируют известную мысль о том, что любое научное направление в начале своего развития богаче идеями, чем в развитом состоянии. Л. Заде на правах живого классика обращается в таких докладах к истокам развития нечеткой логики, вводит новые концепции в парадигму нечетких систем, причем неизменным источником таких идей является семантика естественного языка. В то время как большинство докладчиков ограничиваются узкими задачами теории и практики нечетких систем, не принимая во внимание контекст, связанный с развитием интеллектуальных систем в целом.

Общим свойством обоих докладов Л. Заде является его неизменный интерес к решению основных задач искусственного интеллекта, построению семантики естественного языка. Тео-

рия нечетких систем рассматривается сквозь призму моделирования интеллекта с помощью вычислений. Нечеткие множества, предложенные Л. Заде для выражения человеческих понятий, сохраняют для него, прежде всего это назначение. Это свойство нужно особенно подчеркнуть, так как для последних лет развития нечеткой логики характерен сдвиг в область математики. Исследуются действительно очень интересные и важные свойства нечетких структур в алгебре и топологии, а задачи интеллектуального моделирования привлекают все меньше исследователей.

Основная идея доклада Л. Заде на IFSA'05 заключается в развитии гранулярных вычислений. Докладчик начал с обоснованного утверждения о необходимости выбора уровня точности значений, согласованного с требованиями реальной задачи. Развивая это положение, Л. Заде разработал *Theory of Precisation of Meaning (TPM)*, теорию уточнения значений. Основные положения этой теории и были изложены как в докладе, так и в статье [18]. Интересно сопоставить содержание этой статьи со стартовой статьей [17], чтобы ясно представить себе путь, пройденный нечеткой логикой за 40 лет развития. Презентация Л. Заде была разделена на следующие разделы:

1. Концепция точности/неточности планов выражения и содержания понятий. Каждое понятие имеет содержание (*value*), которое может быть задано точно или не точно (*v-precise*, *v-imprecise*). Каждое понятие характеризуется и своей формой значения (*meaning*), которая также может быть выражена точно или не точно (*m-precise*, *m-imprecise*). Атрибут *m-precise* Л. Заде использует, как аналог термина «математически определенный». Например, если задана пропозиция

- $p: x \text{ is } X$,
- X гауссова случайная переменная с математическим ожиданием m и дисперсией σ , m и σ – точные действительные числа, то говорят, что p имеет атрибуты *v-imprecise* и *m-precise*. Данная концепция лучше всего выражается кратким лозунгом: **теория нечетких систем – это точная наука о неточности.**

2. Грануляция является необходимым следствием *v-imprecise*. Для представления неточного значения вместо единичного значения (синглтона) необходимо использовать: интервал,

распределение какой-либо функции множества, т.е. гранулу сложной структуры. В общем смысле можно говорить об экстенциональном и интенциональном (*attribute-based*) представлении значения. А возможность выполнять операции над гранулами приводит к гранулярным вычислениям.

3. Для определения гранулы используется **принцип обобщенных ограничений** (*generalized constraint*). Обобщенное ограничение задается как

$X \text{ is } R$, где X – ограниченная переменная, r – тип модальности, R – ограничивающее (нечеткое) отношение.

Типы ограниченных переменных:

- X – n -арная переменная, $X = (X_1, \dots, X_n)$
- X – это пропозиция,
- X – функция другой переменной: $X = f(Y)$
- X – обусловлена другой переменной X/Y
- X имеет структуру, например, $X = \text{Location}(\text{Residence}(\text{Carol}))$
- X обобщенное ограничение $X: Y \text{ is } R$

X групповая переменная $G[A]: (\text{Name}_1, \dots, \text{Name}_n)$, с каждым элементом группы Name_i , $i=1, \dots, n$, ассоциируется атрибут A_i .

Типы обобщенных ограничений $X \text{ is } R$:

- r : = ограничение эквивалентности: $X=R$ аббревиатура $X \text{ is } R$
- r : $\leq$ ограничения неэквивалентности: $X \leq R$
- r : $\subset$ ограничения вложенности: $X \subset R$
- R : *blank* возможностное ограничение; $X \text{ is } R$; R распределение возможности на X
- r : v истинностное ограничение; $X \text{ is } v R$; R – распределение истины на X
- r : p вероятностное ограничение; $X \text{ is } p R$; R распределение вероятностей на X
- r : *bm* бимодальное ограничение; X случайная переменная; $R \text{ is}$ бимодальное распределение
- r : *rest* ограничение случайных множеств; $X \text{ is } r R$; R многозначное распределение вероятностей на X
- r : *fg* ограничения нечеткого графа; $X \text{ is } fg R$; X функция и $R \text{ is}$ ее нечеткий граф
- r : *u* ограничения «традиции» (привычной практики usually); $X \text{ is } u R$ «обычно означает» ($X \text{ is } R$)
- r : *g* групповое ограничение; $X \text{ is } g R$ означает, что R ограничивает все значения атрибутов

4. Для символической записи гранулярных пропозиций предлагается Язык гранулярных

вычислений: *Generalized Constraint Language (GCL)*. Дедукцию можно рассматривать как распространение ограничений (*deduction = generalized constraint propagation*). Символическая запись дедуктивных рассуждений ведется с помощью языка протоформ *ProtoForm Language (PFL)*. **PF(p):** - абстрактная форма, глубинная структура **p**. На заданном уровне абстракции, объекты **p** и **q** **PF**-эквивалентны, если **PF(p)=PF(q)**. Например,

p: Большинство шведов высокие **Count (A) is Q**
q: Некоторые профессора богаты **Count (A) is Q**

5. Конструктивный характер *Theory of Precisation of Meaning (TPM)* придает развитый символичный аппарат дедукции протоформ и вычислительный аппарат гранулярных вычислений. Ключевым моментом доклада стала новая формулировка принципа расширения Заде.

6. В докладе все общие положения иллюстрированы конкретными гранулярными вычислениями (Табл. 3) «Высокие шведы» и «низкие итальянцы», богатые и не очень профессора, юная Мэри, Карол из Сан-Франциско и Роберт, возвращающийся домой около 6 часов, черный ящик с шарами, в большинстве своем тяжелыми, и иногда с маленькими удачно разясняли сложные абстрактные построения.

Доклад Л. Заде на IFSA'07 служил продолжением доклада IFSA'05, но включал в себя ряд новых положений. В частности, логическая схема доклада может быть представлена, как последовательность следующих рассуждений.

1. Неопределенность – основной атрибут информации. Со времени К. Шеннона изучают прежде всего статистическую природу неопределенности. Но теория обобщенных ограничений неопределенности (*Generalized Theory of Uncertainty, GTU*) отличается по существу.

2. Тезис о статистической природе неопределенности заменяется в GTU тезисом о том, что информация – это обобщенные ограничения, а статистическое представление информации – это только частный случай.

3. Бивалентность наличия свойства (в том числе истинности) заменяется степенью проявления свойства.

4. Главная цель GTU – способность описать информацию на естественном

языке (NL-capability). В символической форме можно записать: $I(X) = GC(X)$, где X – переменная, определенная на U , $I(X)$ – информация о X , GC – обобщенные ограничения.

NL-capability любой теории неопределенности Л. Заде считает критерием ее приемлемости в эпоху построения гуманистических систем, основанных на знаниях.

2.2. О докладе Д. Дюбуа на IFSA'05

Доклад Д. Дюбуа «О связи между теориями вероятностей и возможностей» был построен как классификация современных теорий неопределенности с указанием характерных особенностей, ограничений и областей применения. Рассмотрены и сопоставлены теория возможностей и вероятностей, в том числе теория возможностей и теория субъективных вероятностей; количественная и качественная теории возможностей. Области применения каждой теории рассматривались по их способности описывать гетерогенные данные и знания.

2.3. О докладах В. Педрича

Доклады В. Педрича на IFSA'05 и IFSA'07 представляют собой последовательное развитие идеи формирования гранул информации с помощью кластеризации.

Табл. 3. Базовые формулы гранулярных вычислений

Правило дедукции	Символьное (на языке протоформ) правило	Вычислительное гранулярное правило.
Композиционное правило вывода	$\frac{X \text{ is } A \quad (X, Y) \text{ is } B}{Y \text{ is } A \circ B}$	$\mu_B(v) = \sup(\mu_A(u) \wedge \mu_B(u, v))$
Принцип расширения Заде	$\frac{X \text{ is } A \quad Y = f(X)}{Y = f(A)}$	$\mu_y(v) = \sup_u(\mu_A(u))$ $v = f(u)$
Правило дедукции на основе принципа расширения Заде	$\frac{X \text{ is } A}{f(X) \text{ is } B}$	$\mu_B(v) = \sup_u(\mu_A(u))$ $v = f(u)$
Обобщение принципа расширения	$\frac{f(X) \text{ is } A}{g(X) \text{ is } B}$	$\mu_B(v) = \sup_u(\mu_A(f(u)))$ $v = g(u)$

Доклад на IFSA'05 назывался «Кластеризация, основанная на знаниях, в гуманистических системах» был основан на хорошо известном FCM-методе кластеризации, адаптированном для формирования гранул информации. Исходными данными являются временные ряды. Доклад IFSA'07 «Динамичное и распределенное нечеткое моделирование» был посвящен динамической кластеризации пространства данных. Динамическая кластеризация пространства рассматривалась как причина динамического развития гранул информации.

2.4. Остальные доклады и выводы

Доклад Дж. Клира «От естественного языка к формализованному языку и обратно», сделанный на IFSA'05, посвящен возможностям нечеткой логики в представлении утверждений естественного языка. Для характеристики такой возможности используются меры информативности нечетких утверждений.

Доклад Р. Ягера на IFSA'07 «Вычисления со словами и гранулами» рассматривал, что именно добавили гранулярные вычисления к теории приближенных рассуждений Л. Заде. Автор доклада обоснованно доказывал, что грануляция задает рамки как для построения вопросно-ответных интеллектуальных систем, так и для семантического Web.

Пленарный доклад Я. Капржыка на IFSA'07 «Вычисление со словами, квалификация полезности и лингвистические квантификаторы: инструменты для гуманистических вычислений» рассматривал возможности классификации на основе интуиционистских множеств, формирования запросов к базам данных на основе логических связей, лингвистического обобщения временных рядов на основе интегралов Шокэ. Довольно сложные рассуждения рассматривались на примере нечетких предпочтений избирателей на выборах.

В целом, обзор пленарных докладов двух конгрессов показывает стремление многих исследователей понять и обобщить результаты развития нечеткой логики, полученные в юбилейные годы. Вероятно, поиски места нечеткой логики в представлении семантики естественного языка и в интеллектуальных системах в целом связаны с определенным «возвращением к истокам».

3. Круглые столы конгресса IFSA'07 и конференции BISCSE

Темы круглых столов (Табл. 4) отражают «итоговый» характер событий, каждый круглый стол так или иначе пытался угадать основные тенденции развития науки о нечеткости.

Табл. 4. Темы круглых столов

BISCSE'05	IFSA'07
Fuzzy systems: Past, Present, Future (Global Issue). Нечеткие системы: прошлое, настоящее, будущее Модераторы: Я. Капржык, В. Нисканен	A GLIMPSE INTO THE FUTURE. Взгляд в будущее Модераторы: Б. Турксен и М. Никравеш
FLINT and Semantic Web. FLINT и семантический Web Модераторы: Е. Санчес и М. Никравеш	FUZZY SETS IN INFORMATION SYSTEMS. Нечеткие множества в информационных системах П. Боск и М. Никравеш
Fuzzy Sets in Information Systems. Нечеткие множества в информационных системах Модераторы: П. Боск	SOFT COMPUTING: PAST, PRESENT, FUTURE (GLOBAL ISSUE). Мягкие вычисления: прошлое, настоящее, будущее (глобальное эссе) Модераторы: Я. Капржык, В. Нисканен
A Glimpse into the Fuzzy Future. Взгляд в будущее нечетких систем Модераторы: Л. Заде и М. Никравеш	

Все проведенные круглые столы в той или иной мере стремились рассмотреть прошлое, настоящее и будущее нечеткой логики, причем, нечеткая логика трактовалась самым широким образом, как теория степени проявления свойства. Пожалуй, лучше всего удалось описать прошлое нечеткой логики. Богатые достижения прошлого даже позволило сделать историю нечеткой логики предметом историков науки. Недаром, на BISCSE'05 и IFSA'07 начали работу секции, посвященные истории нечеткой логики (председатель секций В. Нисканен).

Текущее состояние нечеткой логики также было ясно очерчено на круглых столах. Прогнозы, как всегда были даны довольно расплывчатые. Но важно, что наряду с ожиданием решения ряда математических проблем, от нечеткой логики ожидается существенный вклад в развитие методов искусственного интеллекта.

Многие участники круглых столов подчеркивали, что теория нечетких систем остается единственной теорией, способной на вычисления с объектами, описанными на естественном языке.

4. Сравнение тематик (topics) двух конгрессов

Материалы (Proceedings) конгрессов обширны и, разумеется, необозримы в полном объеме. Чтобы увидеть тенденции развития нечеткой логики рассмотрим материалы конгрессов с точки зрения группировки статей (наименований тем и секций), которую выполнили программные комитеты. Так как исходными данными научного обзора послужили материалы двух конгрессов IFSA, то возникает законное опасение, что подбор статей на конгрессы определили не тенденции развития нечеткой логики, а разный состав программных комитетов. Сравнение их составов показывает, что количество постоянных членов программных комитетов достаточно велико. Как отмечалось в [13],

25 человек из 52 членов программного комитета IFSA'97 работали в комитете IFSA'03. В программных комитетах обоих конгрессов 2005г. (52 члена) и 2007 г. (64 члена) работали 16 человек.

Причем, в большинстве своем это авторитетные ученые, поэтому можно рассчитывать, что выделенные тенденции и задачи будут определять развитие нечеткой логики в будущем.

Для сравнения тематик двух конгрессов воспользуемся следующим приемом. Если составить графы тем и секций двух конгрессов, то получим совершенные разные деревья, с одним общим свойством – эти деревья двухуровневые. Для сравнения будем считать, что существует «идеальное» тематическое дерево, с которым будем сравнивать тематику обоих конгрессов. Возьмем в качестве такого дерева структуру журнала Fuzzy Sets and Systems.

Обзор тематики статей IFSA'05 и IFSA'07 дополним данными ранних конгрессов IFSA'97 и IFSA'03, чтобы можно было видеть долгосрочные тенденции, и представим в Табл. 5.

Табл. 5. Тематика статей IFSA'97 и IFSA'03,05 и 07

Темы Fuzzy sets and systems	Кол-во статей IFSA'97/ Доля рубрики IFSA'97 (%)	Кол-во статей IFSA'03/ Доля рубрики IFSA'03 (%)	Кол-во статей IFSA'05/ Доля рубрики IFSA'05 (%)	Кол-во статей IFSA'07/ Доля рубрики IFSA'07 (%)	Тенденция
1. Основания	33/9	20/8	16/5	20/9	= ↓ ↑
1.1. Основы нечетких множеств и связи	25/7	10/4	11/4	14/7	= ↓ ↑
1.2. Нечеткие отношения и моделирование предпочтений	8/2	10/4	5/1	6/2	↑ ↓ =
2. Математика	67/18	44/19	52/17	41/19	===
2.1. Неклассические логики и теория нечетких множеств	7/2	5/2	11/4	4/2	= ↓ ↑
2.2. Нечеткие дифференциальные уравнения и регрессия	5/1	6/3	3/1	2/1	↑ ↓ =
2.3. Топология и теория категорий	20/5	1/-	9/3	3/1	↓ ↑ ↓
2.4. Обобщенные меры, теория возможностей и теоретические аспекты нечетких чисел	25/7	26/11	20/6	13/6	↑ ↓ =
2.5 Алгебра	10/3	6/3	9/3	19/9	= = ↑
3. Искусственный интеллект	36/10	15/6	25/8	20/9	↑ ↓ =
3.1. Приближенные рассуждения и представление знаний	25/7	9/4	23/7.5	13/6	↑ ↓ =
3.2. Обучение	11/3	6/2	2/0.5	7/3	= ↓ ↑
4. Обработка информации	50/13	43/18	71/24	48/22	↑ ↑ =
4.1. Нечеткие случайные переменные и статистика	8/2	6/3	21/7	0/0	= ↑ ↓
4.2. Классификация и обработка образов	23/6	15/6	28/9	16/7	= ↑ ↓
4.3. Нечеткие базы данных и информационные системы	19/5	22/9	22/8	32/15	↑ ↑ ↑

Темы Fuzzy sets and systems	Кол-во статей IFSA'97/ Доля рубрики IFSA'97 (%)	Кол-во статей IFSA'03/ Доля рубрики IFSA'03 (%)	Кол-во статей IFSA'05/ Доля рубрики IFSA'05 (%)	Кол-во статей IFSA'07/ Доля рубрики IFSA'07 (%)	Тенденция
5. Нечеткие системы	107/28	44/19	69/23	34/16	↓↑↓
5.1. Нечеткое моделирование и нейро-нечеткие системы	55/15	25/11	26/9	17/8	↓↓↓
5.2. Нечеткий контроль	52/14	19/8	43/14	17/8	↓↑↓
6. Принятие решений и оптимизация	48/13	26/11	31/10	31/14	↓ = ↑
6.1. Комбинаторная оптимизация и генетические алгоритмы	15/4	11/5	18/6	19/9	== ↑
6.2. Нечеткое математическое программирование	33/9	15/6	13/4	12/5	↓ ==
7. Приложения	38/10	24/10	32/11	24/11	===
7.1. Гранулярные вычисления и нечеткая инженерия	34/9	18/8	19/7	5/3	== ↓
7.2. Роботика и промышленные приложения	2/0.5	5/2	12/4	18/8	↓↑↑
7.3. нечеткое аппаратное обеспечение	2/0.5	1/0	1/0	1/0	===
			296	218	

Табл. 5 показывает, ряд изменений, а именно увеличилась доля работ по таким направлениям, как

- алгебра,
- нечеткие базы данных и информационные системы,
- роботика и промышленные приложения,
- комбинаторная оптимизация и генетические алгоритмы.

Сократилась доля работ по направлениям:

- топология и теория категорий (вероятно, многие исследования просто не были представлены),
- нечеткие системы (как по нечеткому моделированию и нейро-нечетким системам, так и по нечеткому контролю),
- гранулярные вычисления и нечеткая инженерия,
- нечеткое математическое программирование,
- нечеткие случайные переменные и статистика.

На каждом конгрессе появляются новые работы, которые обещают вырасти в полноценное направление исследований, например, исследования хаотических систем, фракталов (IFSA'97) или исследования в области биоинформатики (IFSA'03). Однако на следующих конгрессах работы в этих направлениях не представлены. Это связано с тем, что перечисленные научные направления сформировались в самостоятельные с собственным сообществом исследователей и календарем конференций, симпозиумов.

Разумеется, будущее нынешних новых направлений остается неясным. К таким направлениям, появившимся на IFSA'05 (Пекин), отнесем работы по эвентологии нечеткого события [6], а среди заявивших о себе на BISCSE'05 и IFSA'07 – работы по интервальному анализу [14].

5. Обзор специальных сессий двух конгрессов

На конгрессах IFSA организуются специальные сессии, причем их тематики и участники собираются по своей инициативе. Темы специальных сессий обычно посвящаются новым перспективным приложениям, которые концентрируют значительное количество работ или тем на теоретических направлениях, которые являются текущими точками роста. Некоторые инициативы специальных сессий не вызывают отклика у специалистов и в них обычно включают 2-3 доклада только инициаторов сессии, зато другие получают широкий отклик.

На IFSA'05 и BISCSE'05 специальные сессии не проводились. Зато на IFSA'07 такие события были широко представлены:

1. Соотношение между интервальными и нечеткими методами (Relation between interval and fuzzy techniques, 6 докладов).
2. Нечеткие деревья (Fuzzy Trees, 4 доклада).
3. Философские и гуманистические аспекты мягких вычислений (Philosophical and Human-scientific aspects of Soft Computing, 7 докладов).

4. Нечеткое мультикритериальное программирование/применения нечеткой логики (Fuzzy Multi-Objective Programming/ Fuzzy Logic Applications, 7 докладов).

5. Применение нечеткой логики и мягких вычислений для формирования гибких запросов (The application of fuzzy logic and soft computing in flexible quering, 30 докладов).

6. Мягкие вычисления в нефтяной промышленности (Soft Computing in Petroleum Applications, 9 докладов).

7. Интеллектуальный анализ данных и принятие решений, основанные на восприятиях (Perception Based Data Mining and Decision Making, 7 докладов).

8. Нечеткая /возможностная оптимизация Fuzzy /Possibilistic Optimization, 9 докладов).

9. Интуиционистские нечеткие множества и их применения (Intuitionistic fuzzy sets and their applications, 13 докладов).

10. Машины поиска и информационно-поисковые системы (Search Engine and Information Processing and Retrieval, 9 докладов).

11. Алгебраические основы мягких вычислений (Algebraic Foundations of Soft Computing, 12 докладов).

12. Объединенное обучение, основанное на модели и на данных: подход нечеткой логики (Joint model-based and data-based learning: The fuzzy logic approach, 5 докладов).

13. Мягкие вычисления в медицинских науках (Soft Computing in Medical Sciences, 3 докладов).

14. Нечеткая логика и мягкие вычисления в распределенных вычислениях (Fuzzy Logic and Soft Computing in Distributed Computing, 7 докладов).

Заявленные темы специальных сессий можно укрупнить, выделяя большое количество исследований в области баз данных (нечеткие базы данных, текстовые базы данных, Data Mining, семантический Web), в одну группу. Но с IFSA'05 можно выделить новое направление исследований - **ДМ баз данных временных рядов**. Такие темы специальных сессий, как история или философские и гуманистические аспекты мягких вычислений свидетельствуют о зрелости нечеткой логики, как научной теории и поисках ее места в общем научном мировоззрении.

6. Новое направление исследований – интеллектуальный анализ баз данных временных рядов (Data Mining for Time Series Data Base)

В трудах IFSA'05 в разных секциях рассыпаны работы, посвященные извлечению правил из баз данных временных рядов (ВР). Особенно интересной постановкой задачи является извлечение правил из нечетких временных рядов. Перечислим основные работы, посвященные данному направлению и опубликованные в томах IFSA'05:

1. M. Nasiri. *Fuzzy regression modeling of false twist texturing yarn*;
2. A. Shapiro. *Soft computing and statistical techniques from an insurance perspective*;
3. J.L. Aznarte, J.M. Benitez, J. L. Castro. *Equivalence relationships between fuzzy additive systems for time series analysis and smooth transition models*;
4. N. Yarushkina, I. Semoushin. *A Soft computing-based Integration Environment for Assessing the Performance of a Complex Enterprise*;
5. D. Marchek. *Determination of Fuzzy Relations for Economic Fuzzy time series models by SCL techniques*;
6. F. Yu, W. Pedrycz, J. Yuan. *Finding Fuzzy Rules from Granular Time series*.

Данные статьи, по нашему мнению отражают основные этапы развития теории нечетких ВР:

1. **Этап нечеткой регрессии.** Первыми были исследованы модели нечеткой регрессии (примером исследования является статья 1).
2. **Этап мягких вычислений.** Развитие методов мягких вычислений породило большое количество работ, исследующих эффективность мягких вычислений для анализа ВР (статьи 2,3,4).
3. **Этап нечетких временных рядов.** Следующим этапом был этап перехода к анализу нечетких временных рядов, а не использования нечетких методов для анализа четких ВР. (статья 5).
4. **Этап извлечения правил (Data Mining) из нечетких (гранулированных ВР).** Развитие методов нечетких баз данных и методов ДМ для реляционных баз

данных позволило перейти к данному этапу (статья 6, специальная сессия на IFSA'07).

6.1. Определение нечеткой регрессии

В статье 1 приводятся основные этапы развития методов нечеткой регрессии. В 1982 г. Х. Танка [9] рассмотрел модель линейной регрессии с нечетким коэффициентом и использовал методы линейного программирования. В 1987 г. А. Келминс и П. Диамонд ввели анализ нечеткой регрессии, основанной на методе наименьших квадратов [10,11]. Большинство работ, посвященных нечеткой регрессии были основаны на следующих базовых определениях.

Пусть дано множество наблюдений: $(y_j, x_{j1}, \dots, x_{jn}), j = 1, \dots, m$, необходимо найти нечеткую модель по следующей форме:

$$\tilde{Y} = A_0 + A_1 x_1 + \dots + A_n x_n$$

Где $A_i(a_i^c, s_i^L, s_i^R), i = 1, \dots, n$ - триангулярные нечеткие числа, a_i^c - среднее значение A_i , s_i^L, s_i^R показывают левый и правый разброс соответственно.

Используются два критерия, чтобы определить нечеткие коэффициенты модели.

1. Для всех наблюдений принадлежность значения y_j к его нечеткой оценке Y_j должна быть как минимум $Y_j(y_j) \geq h, j = 1, \dots, m$, где h - уровень доверия, выбранный лицом, принимающим решения.
2. Общая нечеткость предсказываемого значения зависимой переменной должна быть минимизирована. Это может быть достигнуто минимизацией суммы разбросов нечетких чисел для всех наборов данных. Итак, проблему настройки нечеткой модели с заданными данными $(y_j, x_{j1}, \dots, x_{jn}), j = 1, \dots, m$ можно решить как эквивалентную задачу линейного программирования:

Найти $a = (a_0^c, \dots, a_n^c), s_-^L = (s_0^L, \dots, s_n^L), s_-^R = (s_0^R, \dots, s_n^R)$, которые минимизируют

$$Z = m(s_0^L + s_0^R)s_0^L + \sum_{i=1}^n \left[(s_i^L + s_i^R) \sum_{j=1}^m |x_{ij}| \right].$$

Чтобы оценить качество настройки нечеткой регрессии используют метод наименьших квадратов (MSE). Для нечеткой регрессии MSE определяется следующим образом:

$$MSE = \frac{1}{m} \sum_{j=1}^n [y_j - def(Y_j)]^2, \text{ где } def(Y_j) -$$

дефазифицированное значение зависимой переменной.

Большинство работ, посвященные данной тематике на конгрессах IFSA, либо уточняли данную схему нечеткой регрессии, либо находили новые приложения для ее успешного применения.

6.2. Мягкие вычисления

Статья 2 рассматривает возможности и эффективность применения техники мягких вычислений для задач страхования. Традиционно в этой области использовалась техника статистических вычислений. Статья рассматривает данные методы в сравнении.

Статья 3 сопоставляет использование систем нечетких правил и традиционных моделей авторегрессии для анализа ВР. В частности выясняются отношения между нечеткими аддитивными системами (TSK - системами) и авторегрессионными моделями (Box-Jenkins).

Статья 4 предлагает новую архитектуру средств мягких вычислений для анализа поведения экономических показателей крупного промышленного предприятия.

6.3. Нечеткие ВР

В статье 5 рассмотрены основные понятия авторегрессии нечеткого ВР. В отличие от традиционного ВР значениями нечеткого ВР являются нечеткие множества, а не действительные числа наблюдений. Q. Song и B. Chisson [12] первыми дали определение моделей нечетких ВР.

Пусть $X_t, (t = 1, \dots) \subset R^1$ - универсум, на котором определены нечеткие множества $y_t^i, (i = 1, 2, \dots)$ и Y_t - коллекция $y_t^i, (i = 1, 2, \dots)$. Тогда $Y_t, (t = 1, 2, \dots)$ называется нечетким ВР.

На практике в большинстве ВР последовательные наблюдения зависимы:

$$R = \{(y_t, y_{t-1}), (y_{t-1}, y_{t-2}) \dots\} \subseteq Y_t \times Y_{t-1},$$

где Y_t, Y_{t-1} обозначает переменные, а y_t, y_{t-1} - наблюдаемые значения переменных. Наиболее

частой моделью зависимости является явная функция отображения:

$$f: Y_{t-1} \rightarrow Y_t$$

представленная линейной функцией (марковским процессом, модель AR):

$$y_t = f(y_{t-1}, \phi, \varepsilon) = \phi y_{t-1} + \varepsilon,$$

где ε - случайная ошибка, шум.

В случае нечеткого ВР в качестве модели авторегрессии используется нечеткое разностное уравнение:

$$y_t^j = y_{t-1}^i \circ R_{ij}(t, t-1),$$

$$y_t^i \in Y_t, y_{t-1}^i \in Y_{t-1}, i \in I, j \in J, \circ - \max \min$$

Следовательно,

$$Y_t \rightarrow Y_{t-1}, Y_t = Y_{t-1} \circ R(t, t-1),$$

где $R(t, t-1) = \bigcup_{i,j} R_{ij}(t, t-1)$. Уравнение R на-

зывают моделью нечеткого ВР первого порядка, данная модель – важный частный случай общей модели порядка p :

$$Y_t = (Y_{t-1} \times Y_{t-2} \times \dots \times Y_{t-p}) \circ R(t, t-p),$$

$$R(t, t-p) = \max_p \left\{ \min_{j, i_1, i_2, \dots, i_p} \{y_t^j, y_{t-1}^{i_1}, \dots, y_{t-p}^{i_p}\} \right\}.$$

6.4. Понятие гранулированного ВР

Развитие гранулярных вычислений привело к формированию понятия гранулированного ВР (статья 5). Традиционная сегментация (дискретизация) ВР выполняется методом скользящего окна заданной ширины k на X .

Пусть $W(x, k) = \{w_i | i = 1, 2, \dots, n - k + 1\}$ обозначает множество всех k -широких окон на X . Заддим меру, расстояние между двумя подпоследовательностями w_i . Если выполнить любым из известных способов кластеризацию таких подпоследовательностей, получим s кластеров: $C_i (i = 1, 2, \dots, s)$. Алфавит $\Omega = \{a_i | i = 1, 2, \dots, s\}$ представляет символы образцов ВР. Дискретную версию ВР $D(x) = \{a_{j_1}, a_{j_2}, \dots, a_{j_m}\}$ называют символьным ВР.

Переход к гранулярному ВР предполагает представление вышеописанных кластеров (информационных гранул) нечеткими множествами A_i . В статье 6 предложен кластерный подход к извлечению нечетких правил из символьных гранулярных ВР.

7. Анализ основных работ, по DM баз данных временных рядов

На IFSA'07 данное научное направление было представлено параллельной сессией «*Time series prediction*» и специальной сессией «*Perception based Data Mining and Decision Making*». На параллельной секции были сосредоточены прикладные доклады, а специальная сессия рассматривала теоретические основы нового направления, на ней состоялись следующие доклады.

1. L. Zhang, Zh. Pei, and H. Chen. *Extracting fuzzy linguistic summaries based on including degree theory and FCA*.
2. J. Kacprzyk, A. Wilbik, and S. Zadrozny. *Linguistic summarization of time series by using the Choquet integral*.
3. O. Hryniewicz. *Looking for dependencies in short time series using imprecise statistical data*.
4. I. Batyrshin and L. Sheremetov. *Perception Based Time Series Data Mining for Decision Making*.
5. M. Steinbrecher and R. f Kruse. *Visualization of Possibilistic Potentials*.
6. A. Velmisov, A. Stetsko, N. Yarushkina. *Data mining for fuzzy relational data servers*.
7. V. Sadovnychy, S. Sadovnychiy, V. Ponomaryov. *Computational Intelligence Models of the Distributed Technological Complexes*.

7.1. Обзор базовых понятий извлечения знаний из нечетких временных рядов на основе гранулярных вычислений (Computing with Words and Perceptions, CWP)

Статья 4 из вышеприведенного списка содержит обзор базовых понятий и методов складывающегося научного направления: извлечения знаний из нечетких временных рядов на основе гранулярных вычислений. Основные базовые идеи таких вычислений были рассмотрены ранее при изложении содержания пленарных докладов Л. Заде. Методология вычисления со словами и восприятиями (*computing with words and perceptions CWP*) определяет основную задачу анализа нечетких ВР: распознавание образцов – паттернов ВР (восприятий perception) и извлечения ассоциативных правил в

лингвистической форме. Форма правил определяется **принципом обобщенных ограничений** (generalized constraints). В состав правил входят переменные, принимающие **гранулированные** значения.

На основе новой методологии решаются традиционные задачи анализа ВР:

сегментация – разбиение ВР на значимые сегменты;

кластеризация – поиск группировок ВР или их паттернов;

классификация – назначение ВР или их паттернам одного из заранее определенных классов;

индексирование – построение индексов для эффективного выполнения запросов к базам данных ВР;

резюмирование (summarization) – формирование краткого описания ВР, содержащего существенные черты с точки зрения решаемой задачи;

обнаружение аномалий – поиск новых, не типичных паттернов;

частотный анализ – поиск часто проявляющихся паттернов;

прогнозирование – прогноз очередного значения ВР на основе истории ВР;

извлечение ассоциативных правил – поиск правил, относящихся к паттернам ВР.

В соответствии с методологией CWP основные направления работ сгруппированы в следующие классы:

- Уточнение (precisiation) паттернов, основанных на восприятии;
- Обработка ВР на основе **принципа обобщенных ограничений** (generalized constraints);
- Извлечение ассоциативных правил;
- Преобразование ассоциаций на основе **принципа обобщенных ограничений** (generalized constraints);
- Использование экспертных знаний в системах поддержки принятия решений.

При анализе ВР эксперт представляет свои суждения с помощью нечетких понятий, относящихся ко многим объектам:

- временные области: интервалы времени (*несколько дней*), абсолютная или относительная позиция на временной шкале (*ближкое будущее*), периодические или сезонные интервалы (*неделя до Рождества*);

- ранг значений ВР (*высокая цена, очень низкий уровень производства*);
- набор паттернов ВР (*быстро растущий, слегка выпуклый*);
- набор ВР, их атрибутов, как элементов системы (*фондовый индекс новой компании*);
- набор отношений между ВР, атрибутами или элементами (*тесно связанный*);
- множество значений возможности или вероятности (*непохоже, очень возможно*).

Традиционное выделение паттернов было связано с выделением участков с постоянным знаком первой и второй производной: *возрастающий и выпуклый, убывающий и гладкий* и т.д. Различные шкалы и методы гранулярных вычислений Заде использовались для описания паттернов линейных трендов: *рост, падение, резкий рост, медленное падение* и т.д. Параметрические методы выпукло-гладкой модификации линейных функций и нечеткая грануляция выпукло-гладких паттернов позволили получить лингвистическое описание для ВР, подобное следующему: *медленно убывающий и строго гладкий*. В результате многочисленных исследований создана онтология различных паттернов ВР, в том числе для колебаний и хаоса: *осцилляция, разрушение* и др. Для описания паттернов в одной из работ [5] был предложен и язык описания паттернов: *Shape Definition Language (SDL)*.

В основу такого языка можно положить **принцип обобщенных ограничений** (generalized constraints) для формирования информационных гранул, описывающих ВР и паттерны. Основными методами гранулирования служат точечные ограничения (*point-wise constraints*), ограничения скользящего окна заданной ширины (*window-wise constraints*) и их комбинации (*composite constraints*), например: *среди дней зимних месяцев было N дней с высоким уровнем производства нефти и M дней с низким уровнем производства*. В данной фразе композитные ограничения задают временные ограничения, которые могут включать интервальную логику Алена: *до, содержит, перекрытие интервалов* и т.д. Не менее часто композитные ограничения используются для задания пространственных ограничений.

Композитные ограничения могут служить основой для извлечения ассоциативных правил

из ВР. Такие правила извлекают на основе четырех методов:

1. ассоциативные правила на основе мер доверия;
2. ассоциативные правила на основе частотности;
3. ассоциативные правила на основе корреляции ВР;
4. ассоциации локальных трендов.

7.2. Основная схема извлечения ассоциативных правил из нечетких ВР

Основная схема извлечения ассоциативных правил следующая. Пусть даны m ВР и их значений: $Y = \{Y_i\}, i = 1, \dots, m; Y_i = (y_{i1}, \dots, y_{N_i})$ в моменты времени $T_i = (T_{i1}, \dots, T_{N_i})$. Пусть нечеткое *point-wise* ограничение A определено на интервалах времени $T_A = \bigcup_{i \in I_A} T_i, I_A \subseteq \{1, \dots, m\}$,

так что $\mu_A(t_{iq})$ указывает полноту ограничения A в момент времени $t_{iq} \in T_A$. Аналогично, пусть T_B будет временной областью для нечеткого ограничения B . Предположим $D \subseteq T_A \times T_B$ - множество пар моментов времени (t_{iq}, t_{jp}) , используемых для оценки ассоциативного правила и $|D|$ - количество пар. Для каждой пары (t_{iq}, t_{jp}) получим $\mu_A(t_{iq}, t_{jp}) = \mu_A(t_{iq}), \mu_B(t_{iq}, t_{jp}) = \mu_B(t_{jp})$, пределим меры s и c для ассоциативного правила: $R: A \Rightarrow B, \langle s, c \rangle$ с *point-wise* ограничениями A и B с помощью формул:

$$s = \frac{1}{|D|} \sum_{(t_{iq}, t_{jp}) \in D} T(\mu_A(t_{iq}), \mu_B(t_{jp}))$$

$$c = \frac{\sum_{(t_{iq}, t_{jp}) \in D} T(\mu_A(t_{iq}), \mu_B(t_{jp}))}{\sum_{(t_{iq}, t_{jp}) \in D} \mu_A(t_{ip})},$$

где T - t -норма, например, $\min$.

На основе оценок s и c можно преобразовать ассоциации в соответствие **принципом обобщенных ограничений** (*generalized constraints*):

$$A \rightarrow \text{BisrP}(s, c)$$

например:

$$A - \text{Increa sin } g \rightarrow B - \text{Increa sin } g : \text{isrP}(W).$$

7.3. Проблемы DM баз данных нечетких временных рядов

Открытыми проблемами данного нового научного направления являются следующие:

1. использование извлеченных ассоциативных правил в экспертных системах;
2. лингвистическая трансляция результатов кластеризации паттернов ВР в правила;
3. резюмирование совокупности нечетких ВР;
4. прогнозирование на основе распознанной тенденции нечеткого ВР: *If trend is F then next point is Y*;

5. использование результатов *Data Mining нечетких* временных рядов в системах поддержки принятия решений, в том числе теоретические проблемы использования интегралов Шоке (*Choquet*).

Заключение

Таким образом, итоги и перспективы основных направлений теоретических исследований по нечетким системам связаны с развитием гранулярных вычислений. Построение гранулы информации основано на математическом развитии теорий степени проявления свойств. Фундаментальный аспект исследования сейчас, как и 40 лет назад, вновь связывается с исследованием семантики естественного языка.

Уже достигнутые теоретические результаты позволяют рассматривать нечеткую логику как перспективный инструмент для анализа гетерогенной информации, в виде сложных качественных лингвистических описаний и количественных данных.

Основными активно исследуемыми прикладными направлениями становятся:

1. Исследование нечетких баз данных, в том числе баз данных нечетких временных рядов;
2. Интеллектуальный анализ данных и текстов (*Data and Text Mining*);
3. Построение интеллектуальных информационных систем;
4. Формирование семантического Web-пространства.

Литература

1. IFSA'97 Proceedings of Seventh International Fuzzy Systems Associations World Congress. Czech Republic, Prague: Academia, In 4 Vol., 1997.

2. Fuzzy structures. Current trends. Editors: R. Mesiar, B. Riecan, Tatra Mountains Mathematical Publications, Bratislava: Mathematical Institute Slovak Academy of Sciences, Vol. 13, 1997.
3. Fuzzy Sets and Systems – IFSA 2003. Eds. T. Bilgic, B. De Baets, O. Kaynak, Turkey, Istanbul, Proceedings of 10th International Fuzzy Systems Associations World Congress, Springer, 2003.
4. IFSA'03 Proceedings of 10th International Fuzzy Systems Associations World Congress. Turkey, Istanbul, 2003.
5. R. Agrawal, G. Psaila, E. Wimmers, M. Zait, Quering shapes of histories, in: Proc. VLDB'95, Zurich, Switzerland, 1995, pp. 502-514.
6. A. Vorob'ev Eventological means of fuzzy events, in Proc. IFSA'05, Beijing, China, vol. 1, pp. 330-333.
7. L.A. Zadeh Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and Systems, Vol. 90, 1997, p. 111-127.
8. И.З. Батыршин. Общий взгляд на основные черты и направления развития нечеткой логики Л. Заде. Новости искусственного интеллекта, № 2-3 (44-45), 2001, с. 25-28.
9. H. Tanaka, S. Uejima, K. Asai., Linear regression analysis with fuzzy model, IEEE Systems, Man & Cybernetics, vol. 12, pp. 903-907, 1982.
10. Celmins, Least squares model fitting to fuzzy vector data, Fuzzy Sets Syst., vol. 22, pp. 260-269, 1987.
11. P. Diamond, Least squares fitting of several fuzzy variables, Proc. Of Second IFSA Congress, Tokyo, pp. 20-25, 1987.
12. Q. Song, B. Chissov Fuzzy time series and its models, Fuzzy Sets and Systems, 54, pp. 269-277.
13. Ярушкина Н.Г. Нечеткие и гибридные системы: обзор итогов и тенденций развития // Новости искусственного интеллекта, № 6, 2003.
14. M. Ceberio, V. Kreinovich, A. Pownuk, B. Bede: From Interval Computations to Constraint-Related Set Computations: Towards Faster Estimation of Statistics and ODEs Under Interval, p-Box, and Fuzzy Uncertainty, Foundations of Fuzzy Logic and Soft Computing: P. Melin, O. Castillo, L. Aguilar, Ja. Kacprzyk, W. Pedrycz (Eds.), 12th International Fuzzy Systems Association World Congress, IFSA 2007, Cancun, Mexico, June 18-21, 2007, Proceedings. Lecture Notes in Computer Science 4529 Springer 2007, pp. 33-42.
15. Fuzzy Logic, Soft Computing & Computational Intelligence: Eleventh International Fuzzy Systems Association World Congress//Ed. Y. Liu, G. Chen, M. Ying, ISBN 7-302-11377-7/TP-7494, Springer, 2005.
16. Theoretical Advances and Applications of Fuzzy Logic and Soft Computing // Eds O. Castillo, P. Melin, O. Montiel Ross, R. Cruz, W. Pedrycz, J. Kacprzyk, Avances in Soft Computing 42, Springer, 2007.
17. Zadeh, Lotfi A. (1965). Fuzzy sets. Information and Control 8, 378-53.
18. Zadeh, Lotfi A. (2006) Generalized theory of uncertainty (GTU) – principal concepts and ideas, Computational statistic & Data analysis, 51, pp. 15-46, 2006.

Ярушкина Надежда Глебовна. Проректор по научной работе, зав. каф. «Информационные системы», ГОУ ВПО Ульяновский государственный технический университет, д.т.н., профессор. Почетный работник высшей школы. Количество научных работ: около 200, из них в рецензируемых журналах – 30, монографий – 5 (2 – без соавторов). Области интересов: интеллектуальные системы, теория и приложения нечетких систем, мягкие вычисления