

МАКСИМЕНКО ОЛЬГА ИВАНОВНА

**ФОРМАЛЬНЫЕ МЕТОДЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ
СИСТЕМ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ТЕКСТА**

Специальность 10.02.21 - прикладная и математическая лингвистика

**Автореферат
диссертации на соискание ученой степени
доктора филологических наук**

Москва 2003

Диссертация выполнена на кафедре теоретической и прикладной лингвистики Московского государственного областного университета.

Официальные оппоненты: доктор филологических наук, профессор

О.П. Крюкова

доктор филологических наук, доцент

А.А. Поликарпов

доктор филологических наук, профессор

А.И. Новиков

Ведущая организация: Московский государственный лингвистический университет

Защита состоится 3 октября 2003 года в 11.30 час. на заседании диссертационного совета Д 212.155.04 по защите диссертаций на соискание ученой степени доктора филологических наук при Московском государственном областном университете по адресу: 107042 г. Москва, Переведеновский переулок д. 5/7, ауд. 33.

С диссертацией можно ознакомиться в библиотеке Московского государственного областного университета по адресу: 105005, г. Москва, ул. Радио, 10А

Автореферат разослан 30 августа 2003 г.

Ученый секретарь диссертационного совета доктор филологических наук профессор



Г.Т. Хухуни

Реферируемая работа посвящена исследованию формальных методов оценки эффективности интеллектуальных систем обработки текста, результатом которого явилось создание действующих экспертных систем на базе нечеткой логики.

За последнее десятилетие существенно возросло количество и произошли качественные изменения систем автоматической обработки текста, опирающихся на глубинные структуры языка и оперирующих сложными, плохо формализуемыми элементами. (Апресян Ю.Д., Белоногов Г.Г., Берштейн Л.С., Гринбаум О.Н., Иомдин Л.Л., Кузнецов И.П., Нариньяни А.С., Пиотровский Р.Г., Arnold D., Carl M., Graven M., Streiter O. и мн. др.). Подобные системы требуют серьезной экспертной оценки. В наши дни наблюдается существенный разрыв между возможностями аппаратных средств компьютеров и применяемыми методами решения прикладных задач. Наиболее освоенные на сегодня методы базируются на хорошо формализованных алгоритмах, полученных в результате построения математических моделей предметных областей. Чаще всего это трудоемкие расчеты по известным формулам, либо сравнительно простые последовательности действий, приводящие после многократного повторения к желаемому результату (итерационные алгоритмы). В практической же деятельности многие актуальные задачи относятся к типу плохо формализуемых, особенно если дело касается естественного языка, для которого неизвестны аналитические зависимости или цепочки действий, приводящие к результату без интеллектуального вмешательства человека. До последнего времени для решения плохо формализуемых лингвистических задач просто не хватало ресурсов вычислительной техники, и поэтому было бессмысленно ставить саму проблему серьезного их решения. Сегодня создание алгоритмов и прикладных программ, наиболее полно использующих ресурсы современных компьютеров для решения нечетких задач, к которым относятся и лингвистические, является весьма актуальной проблемой.

Как правило, в плохо формализуемых задачах имеется некоторый набор параметров, описывающих объекты предметной области, причем, нельзя с уверенностью сказать, что этот набор параметров наверняка полон и адекватен, а сами измеренные значения параметров также в совокупности полны, непротиворечивы и не искажены. Все это не позволяет применять для решения плохо формализуемых задач традиционные статистические методы, методы оптимизации и аппроксимации.

Современные методы, основанные на нейросетевых алгоритмах, нечеткой логике, генетических алгоритмах, а также экспертные системы и ряд других направлений, составляют базу современных технологий прогнозирования на основе временных рядов данных, извлечения закономерностей из данных, создания систем поддержки принятия решений, что аккумулируется термином "интеллектуальный анализ данных" (Дударь З.В., Кемени Дж., Кузьмин В.Б., Моисеев Н.Н., Орлов А.И., Подиновский В.В., Потапова Р.К. и мн. др.). Нечеткая логика позволяет работать с параметрами, имеющими скорее качественную природу типа: "большой - маленький", а не точное значение, выражаемое определенной цифрой. Экспертные системы служат способом привлечения знаний экспертов для решения плохо формализованных задач. Все более широкое распространение получают экспертные системы на базе нечеткой логики (Вишневский Р.В., Масалович А.И., Феклистов А.И. др.).

Современное общество называют информационным - в последние десятилетия резко возросли потоки информации во всех сферах человеческой деятельности и появились мощные технические средства ее передачи и обработки, однако, для обмена информацией по-прежнему основной трудностью остается языковой барьер. Традиционные методы преодоления языковых барьеров - это изучение иностранных языков и переводческая деятельность. Но с ростом экономических, политических и культурных связей между странами, а также в связи с повышением интенсивности информационных потоков, эти методы становятся недостаточными. Поэтому, несмотря на все трудности, используя многие доступные формальные методы, продолжают развиваться интеллектуальные лингвистические системы, такие как системы машинного перевода, переводческие накопители, огромные по размеру и наполнению базы знаний поисковых систем, действующие как автономно, так и в среде сети Интернет - в этом серьезная роль и большая заслуга прикладной лингвистики, к задачам которой относится и оценка этих интеллектуальных систем.

Актуальность темы: в условиях информационного общества обилие разработанных и разрабатываемых интеллектуальных систем обработки текста, к которым относятся системы машинного перевода и информационно-поисковые системы в том числе, требует быстрой и адекватной оценки. Важно иметь представление о качестве систем как на стадии начальной разработки, так и на стадии готового продукта. Решение подобной задачи традиционными экспертными способами чрезвычайно трудоемко и, как правило, дает субъективный результат. Создание формальной методики решения этой проблемы - веление

времени, тем самым тема диссертационной работы актуальна, как в теоретическом, так и практическом смысле.

Научная новизна работы определяется следующими результатами:

1. Разработан новый подход к оценке эффективности систем автоматической обработки текста с применением современного математического аппарата нечеткой логики, по сути экспертная система оценки интеллектуальных систем на базе нечеткой математики.
2. Создан новый комплекс критериев оценки интеллектуальных систем.
3. Разработанный подход является универсальным: нечеткие экспертные системы позволяют оценивать качество интеллектуальных систем с единых позиций в зависимости от требуемой степени точности и сложности оценки.
4. Проведено тестирование созданных действующих систем оценки качества на трех разных типах интеллектуальных систем автоматической обработки текста - автоматизированном дистрибутивно-статистическом анализе (ДСА), системе машинного перевода (СМП), информационно-поисковой системе (ИПС), в разработке которых автор также принимал активное участие.
5. В работе предложена новая типологическая классификация электронных словарей и экспериментальная классификация серьезности ошибок при русско-английском МП.

Научные положения, изложенные в работе, опираются на обширные аналитические исследования имеющихся в мировой науке работ по близкой тематике. В итоге создано новое направление оценки плохо формализуемых интеллектуальных систем на базе нечеткой логики.

Объект исследования: существующие на настоящий момент формальные методы оценки различных по типу интеллектуальных систем автоматической обработки текста, таких как автоматизированный дистрибутивно-статистический анализ, системы машинного перевода и информационно-поисковые системы.

Цели и задачи:

- изучение и анализ существующих современных формальных методов автоматической обработки естественного языка;
- изучение и анализ формальных методов оценки интеллектуальных систем (ИС);
- обоснование необходимости формального метода оценки интеллектуальных лингвистических систем на базе нечеткой

логики и создание универсальной системы оценки на базе этого метода;

- тестирование разработанной системы на ИС разного типа: автоматизированном варианте ДСА, СМП и ИПС.

Теоретическая значимость: системы экспертных оценок представляют собой сложный аппарат, строящийся, как правило, на статистических методах. Такие методики не слишком хороши для систем автоматической обработки текста на естественном языке, поскольку язык относится к числу плохо формализуемых систем. Результаты исследования и оценок автоматических систем обработки текста способны внести существенный вклад как в прикладное, так и теоретическое языкознание. Научные результаты фактически открывают новое направление создания систем оценки интеллектуальных лингвистических систем.

Практическая ценность: впервые создана универсальная система оценки автоматических систем обработки текста на базе нечеткой логики, способная оценивать эффективность разнообразных интеллектуальных систем по комплексу изменяемых критериев. Экспертная система на базе нечеткой логики реально действует в интегрированной среде MATLAB.

Методы исследования: в первую очередь системный метод анализа существующих достижений в области создания автоматизированных систем и разнообразных систем оценки и определения научной целесообразности применения аппарата нечеткой логики для оценки лингвистических систем; метод моделирования при проектировании и разработке собственно экспертной системы на базе нечеткой логики; описательно-аналитический и типологический методы при сравнении полученных результатов оценки.

Материалом исследования послужили следующие автоматизированные системы обработки текста:

- 1) автоматизированный дистрибутивно-статический анализ, которым был обработан массив текстов объемом более 100 тыс. словоупотреблений, состоящий из текстов разных жанров (монография, научные статьи, рефераты из реферативного журнала, массив толкований терминологического словаря научно-технической терминологии по системам связи и управления);
- 2) система машинного перевода АСПЕРА, тестирование которой проводилось на текстах по 5 тематикам (бизнес, история философии, охрана окружающей среды, информатика, медицина) общим объемом 100 Кб;

3) информационно-поисковая система КАСКАД, содержащая около 1000 статей.

Объем исследования подтверждает достоверность полученных результатов.

Апробация: результаты исследования изложены в монографии "Формальные методы в современной прикладной лингвистике" (2002 г.) и других печатных работах автора, а также представлены в виде докладов и выступлений на конференциях:

- Третья конференция по Машинному фонду русского языка. Москва, ИРЯ АН СССР, 1989 г.;
- Международный семинар «ЭВМ и перевод», Тбилиси, 1989 г. ;
- Вторая Республиканская конференция "Компьютерные программы в обучении белорусскому и иностранным языкам". Минск, 1995 г. ;
- Научно-методический симпозиум «Язык в действии», МАПРЯЛ, Тбилисский государственный университет, 1997 г. ;
- Международная конференция "Актуальные проблемы обучения иностранных студентов", Харьковский государственный политехнический университет, 1998 г. ;
- Межвузовская научно-методическая конференция "Информационные технологии и фундаментализация высшего образования", Москва, РГУ им. Губкина, 1999 г. ;
- Межвузовская научно-практическая конференция "Традиционные и новые концепции, методы и приемы обучения иностранным языкам", Москва, МГИМО, 2000 г. ;
- Научно-практическая конференция "Проблемы обучения иностранных граждан на современном этапе" Волгоградский ГПУ, 2000 г. ;
- Научно-практическая конференция студентов, аспирантов и преподавателей МПУ, Москва, 2001 г.;
- Межвузовский научно-практический семинар «Теория и практика преподавания русского языка иностранным учащимся в вузе". Москва, МГИМО, 1997 г., 1998 г., 1999 г., 2000 г., 2002 г., 2003 г.

Результаты исследования использованы в теоретических курсах "Математические методы в лингвистике", "Формальные модели в лингвистике", "Автоматическая обработка естественного языка", которые читаются студентам кафедры теоретической и прикладной лингвистики МГОУ.

Структура работы: в соответствии с поставленными задачами диссертационная работа состоит из введения, четырех глав, заключения, списка использованной литературы и ряда приложений.

Диссертация прошла предварительную экспертизу на совместном заседании кафедр теоретической и прикладной лингвистики; переводоведения; теории языка и англистики МГОУ в июне 2003 года.

По теме диссертации опубликовано более 20 научных работ, общим объемом более 40 п.л., включая монографию (16 п.л.) и два учебных пособия.

В Главе 1 дан обзор формальных методов, использующихся в современной науке и, в частности, прикладной лингвистике. Развитие научных исследований невозможно без определенной формализации. Эта мысль подтверждается многочисленными наблюдениями за развитием гуманитарных наук в наши дни. Даже те области знания, где ранее формальные методы применялись редко или неэффективно, проявляют закономерный интерес к более строгим средствам описания и интерпретации интересующих их фактов и экспериментов (Бутенко И.А., Калеватых А.В., Карпова Н.А., Кононенко И., Раушенбах Г.В., Cushing B.E., Deboeck G. J., Fung P., и др.).

Формальные методы представляют собой искусственно созданные человеком приемы познания, поскольку они не содержатся в предметах или окружающей действительности наподобие закономерностей, правил, отношений. Однако методы не существуют в отрыве от этих закономерностей, правил и отношений. Правильно выбранный и примененный в исследовании метод помогает вскрывать и фиксировать закономерности и свойства, не лежащие на поверхности явлений. Выбор методов определяется спецификой предмета, накопленными знаниями о нем, целью исследования, а также зависит от того, какие стороны и свойства предмета предстоит анализировать. Учет всех этих факторов обеспечивает эффективность применения методов.

Исключительное внимание следует уделять интерпретации полученных результатов, чтобы теоретические выводы и заключения не навязывались материалу, а естественно вытекали из него в ходе использования того или иного метода. Метод часто представляет собой формализуемую процедуру анализа, но сам процесс его использования и интерпретация полученных при этом результатов - качественное явление, основанное на опыте, знаниях и интуиции специалиста. Наибольшего доверия заслуживает в науке тот метод, при использовании которого получают результаты, способные воспроизводиться в других экспериментах на новом материале, то есть проверяемые результаты.

Поскольку язык сам служит исследовательским инструментом, для его познания необходимо применить новый инструмент, или, как его называют в современной лингвистике, метаязык описания или исследования. На протяжении фактически всей истории языкознания специалисты использовали в качестве метаязыка сам язык, т. е. сущность языка раскрывалась и описывалась с помощью того же языка. Если метод - это прием исследования, то метаязык - инструмент исследования. Прием и инструмент, или метод и метаязык, представляют собой необходимые компоненты всякого научного анализа языковых фактов.

Развитие лингвистики в направлении формализации понятий и методов обусловлено как внутренними, так и внешними факторами. С одной стороны, внутреннее развитие лингвистики в XX веке привело к использованию более точных и объективных методов для изучения естественного языка (ЕЯ) и созданных на его основе текстов. С другой стороны, бурное развитие вычислительной техники, зарождение и становление кибернетики, достижения в смежных науках о человеке явились внешними факторами, активизировавшими использование точных методов в лингвистике в связи с применением ЕЯ в автоматизированных информационных системах, системах искусственного интеллекта, в организации диалога человека с компьютером и т.д.

Первой из математических наук в лингвистику пришла статистика, заложив основы количественной лингвистики (Алексеев П.М., Мартыненко Г., Орлов А.И., Поликарпов А.А., Скороходько Э.Ф., Сухотин Б.В., Фрумкина Р.М., Тулдава Ю., Abney S., Kaski S. и др.). Построение частотных словарей, словарей-индексов, конкордансов, создание машинных фондов лексики - все это невозможно без применения методов количественной лингвистики. Грамматический и семантический аспекты количественного исследования лексики привели к построению математической теории словообразования, определению лексико-семантических групп. Такие современные методы анализа текстов как кластерный анализ, контент-анализ, лексико-стилистический анализ также используют достижения количественной лингвистики. Метод дистрибутивно-статистического анализа, используемый для обработки текстов, представляет особый интерес (Поликарпов А.А., Шайкевич А.Я. и др.)

В последнее время появились так называемые компьютерные "интеллектуальные системы", способные достаточно адекватно работать с естественным языком. К ним относятся, в первую очередь, информационно-поисковые системы, системы машинного перевода и

экспертные системы. При создании таких систем применяются самые разнообразные формальные методы. Активное применение компьютерных технологий вынуждает исследователей пытаться формализовать те уровни языка и языковые явления, о возможности формализации которых ранее и речи не шло.

Лингвистическими проблемами искусственного интеллекта являются проблемы, связанные с исследованием, проектированием и реализацией лингвистических процессоров. В их число входят семантические языки и способы представления знаний в лингвистических процессорах, анализаторы текста как модели различных видов понимания, алгоритмы синтеза текста для вербализации заданного содержания; алгоритмы организации диалога между человеком и системой; алгоритмы перевода (Апресян Ю.Д., Белоногов Г.Г., Иомдин Л., Королев Э.И., Марчук Ю.Н., Нариньяни А.С., Пиотровский Р.Г., Потапова Р.Г., Узилевский Г.Я. и др.)

В течение нескольких последних десятилетий продолжались исследования в области автоматической обработки текста (АОТ), в результате которых были выработаны основные принципы анализа текста на естественном языке: семантически-ориентированный подход; эффективное использование знаний (как проблемных, так и общих) во время и после лингвистических этапов процесса анализа; организованное сообщество активных агентов, выступающее в качестве аппарата спецификации лингвистического обеспечения; распределенная, идущая "снизу-вверх" организация процесса обработки, базирующаяся на управлении по данным и/или событиям.

В качестве примера системы АОТ в работе рассматривается американская система Link Grammar Parser (LinkParser), доведенная до программной реализации. Она включает в себя все элементы анализа текста - от начального (графематического) до первичного семантического (условно) анализа английского языка.

Одно из наиболее интересных направлений формализации языкового мира представляют собой семантические сети, в частности, нейронные семантические сети. Самые простые сети, использующиеся в системах искусственного интеллекта, - это реляционные графы, состоящие из узлов, соединенных дугами. Для представления связного текста разработаны пропозиционные сети, узлы которых представляют собой целые предложения. В системах МП семантическая сеть помогает парсеру разрешить семантическую неоднозначность при порождении языка. Семантическая сеть может быть использована для автоматической классификации документов в задачах классификации и полнотекстового поиска. Семантические нейронные сети применяются

в качестве формального языка описания и обработки смысла текста на естественном языке. Последние версии семантических сетей стали более мощными и гибкими и составляют конкуренцию фреймовым системам, логическому программированию и другим языкам представления (Дударь З.В., Шуклин Д.Е., Феклистов А.И., Харламов А.А., Шингарева Е.А., Щетинин В.Г., Graven M., Toshio F. и др.)

Дистрибутивно-статистический метод анализа лексики (ДСА) (Поликарпов А.А., Прицкер А.И., Шайкевич А.Я. и др.), окончательно сформировавшийся в последние десятилетия XX века, дает один из наиболее адекватных результатов описания поведения лексики, как ядерной, так и терминологической, позволяя, в частности, различать жанры произведений в информационном потоке. Дистрибутивный вариант ДСА основан на сходстве распределений слов, т.е. всех возможных контекстов, в которых эти слова могут встречаться. Статистический вариант ДСА базируется на числовых характеристиках совместной встречаемости элементов текста в контекстах определенной длины.

Эксперимент по оценке алгоритма ДСА (Максименко 1989) заключался в анализе с помощью дистрибутивной и статистической методик текстов на естественном языке и построении графов дистрибутивных связей с различными порогами и графов статистической связности с разными уровнями доверительной вероятности. Также были определены оптимальные дистрибутивные пороги и уровни доверительной вероятности в задаче жанровой атрибуции текстов, для чего был разработан специальный алгоритм, описанный в Главе 2, где также дается оценка эффективности предложенного алгоритма. Анализ проводился на массиве текстов, общий объем которого превышал 100 тыс. словоупотреблений. В массив вошли тексты разных жанров (монография, главы из книги, статьи отдельных авторов, рефераты из реферативного журнала (РЖ) и толковательная часть терминологического словаря). Все тексты относились к одной тематической области - системы связи и управления. Единственно возможными параметрами сравнения оказались полнота и точность лексики графов. Было определено, что для разных жанров текста характерны разные типы оптимальных графов. Монография и статьи характеризуются компактностью дистрибутивных графов, в которых особенно насыщен второй уровень. Оба члена заглавной пары коррелируют со всеми членами второго уровня. Статистический граф имеет 5 - 7 уровней, основная масса терминологических гиперлексем (единиц анализа), включая родовые, находится на 3 - 5 уровнях. Выборка из РЖ характеризуется близостью

структуры дистрибутивного и статистического графов, в которых основная масса гиперлексем располагается на 3-ем и последующих уровнях. Разделы главы книги описываются дистрибутивным графом с одинаковым количеством членов на 2 - 4 уровнях и 5-уровневым статистическим графом с основной массой терминологических гиперлексем на 2-ом и 4-ом уровнях. Словарь отличается возможностью использования более жестких порогов (дистрибутивный - 0,4 - 0,5 и статистический - 0,99). Однако терминологические гиперлексеми в единый граф не объединяются, а расходятся по 6 - 9 отдельным графам, причем статистическая методика при анализе словарного текста склонна игнорировать родовые термины.

Нечеткая логика представляет собой, по всей видимости, одну из наиболее спорных, но мощных и своеобразных методик, используемых для формализации плохо формализуемых систем и работы с ними. Достижения нечеткой логики в последнее время стали использоваться в ряде гуманитарных исследований. (Аверкин А.Н., Вишневский Р.В., Масалович А.И., Мелихова О.А., Zadeh L., Zemankova-Leech M. и др.). На наш взгляд, теория нечеткой логики может быть использована для формализации задачи оценки эффективности, например, систем машинного перевода или информационно-поисковых систем. В последующих главах работы приводятся варианты нечетких экспертных систем оценки качества таких лингвистических ИС.

Прикладная лингвистика зачастую имеет дело со слабо формализованными или трудно формализуемыми задачами. Аппарат нечеткой логики как нельзя лучше подходит для описания условий и метода решения таких задач на языке, близком к естественному, для построения моделей приближенных рассуждений человека и использования их в компьютерных системах. Однако до настоящего времени в прикладной лингвистике аппарат нечеткой логики практически не использовался и никаких оценочных (экспертных) систем, а также систем принятия решений на базе нечеткой логики создано не было.

У систем, построенных на базе нечеткой логики, есть ряд отличительных особенностей: возможность оперирования входными данными, заданными нечетко, например, значениями, которые невозможно задать однозначно (мнения экспертов); возможность формализации нечетких критериев оценки и сравнения, таких как, "хороший", "плохой", "хуже", "лучше" и пр.; возможность проведения качественных оценок как входных данных, так и выходных результатов: исследователь оперирует не только самими значениями

данных, но и степенью их достоверности; возможность быстрого моделирования сложных систем и их сравнительного анализа с заданной степенью точности: оперируя принципами поведения системы, которые описаны методами нечеткой логики, исследователь может оценить разные варианты выходных значений, не составляя точных уравнений.

При построении нечетких экспертных систем оценки качества, предлагаемых в работе, использовались следующие основные понятия нечеткой логики (Круглов В.В.):

- нечеткое подмножество A универсального множества X (нечеткое множество) определяется как множество упорядоченных пар $A = \{\mu_A(x)/x\}$, где $\mu_A(x)$ - характеристическая функция принадлежности (ФП), принимающая значения в некотором множестве M (например, $M = [0,1]$), x - элемент универсального множества X . Функция принадлежности указывает степень принадлежности элемента x подмножеству A ;

- нечеткая переменная характеризуется тройкой параметров (α, X, A) , где α - наименование переменной, X - универсальное множество (область определения α), A - нечеткое множество на X , описывающее ограничения на значения нечеткой переменной α (т.е. ФП $\mu_A(x)$);

- лингвистическая переменная (ЛП) - набор из пяти параметров (β, T, X, G, M) , где β - наименование ЛП, T - множество ее значений (базовое терм-множество), представляющих собой наименования нечетких переменных, областью определения каждой из которых является множество X , G - синтаксическая процедура, позволяющая оперировать элементами терм-множества T , в частности, генерировать новые термы (значения), формируя $G(T)$ - множество сгенерированных термов. И наконец, M - семантическая процедура, позволяющая превратить каждое новое значение ЛП, образуемое процедурой G , в нечеткую переменную, т. е. сформировать соответствующее нечеткое множество;

- нечеткое число определяется как нечеткое множество A на множестве действительных чисел с ФП $\mu_A(x) \in [0, 1]$, где x - действительное число. Особым классом являются нечеткие числа (L-R)-типа, ФП которых относятся к типу так называемых (L-R)-функций, среди которых треугольная, трапецеидальная, прямоугольная, экспоненциальная функции, а также функция Гаусса и некоторые другие. При использовании функций (L-R)-типа нечеткое число A (нечеткое множество) задается следующим образом:

$$\mu_A(x) = \begin{cases} L(x, \alpha), & \text{при } x \leq a_1 \\ 1, & \text{при } a_1 < x \leq a_2 \\ R(x, \beta), & \text{при } x > a_2 \end{cases},$$

где $\alpha > 0$, $\beta > 0$ - левый и правый коэффициенты нечеткости, a_1 и a_2 - границы толерантности, определяющие промежуток $[a_1, a_2]$, в котором значение ФП равно 1. Таким образом, при выбранных функциях L и R нечеткое число задается четверкой параметров $A = (a_1, a_2, \alpha, \beta)$;

- нечеткое отношение (импликация) между множествами X и Y определяется как функция $R: (X \times Y) \rightarrow [0, 1]$, которая ставит в соответствие каждой паре элементов $(x, y) \in X \times Y$ величину $\mu_R(x, y) \in [0, 1]$. Нечеткое отношение кратко записывается в виде xRy ;

- композиция (свертка) двух нечетких отношений. Пусть R_1 - нечеткое отношение $R_1: (X \times Y) \rightarrow [0, 1]$ между множествами X и Y , а R_2 - нечеткое отношение $R_2: (Y \times Z) \rightarrow [0, 1]$ между множествами Y и Z . Нечеткое отношение между X и Z , обозначаемое $R_2 \circ R_1$ и определенное через R_1 и R_2 выражением

$$\mu_{R_1 \circ R_2}(x, z) = \max_y \left(\min \left(\mu_{R_1}(x, y), \mu_{R_2}(y, z) \right) \right),$$

называется (max - min)-композицией отношений R_1 и R_2 . Операция МИНИМУМ может быть заменена алгебраическим умножением, тогда говорят о (max - prod)-композиции;

- база знаний (правил вывода) - это формируемая специалистами предметной области совокупность нечетких предикатных правил вида: Π_1 : если x есть A_1 , то y есть B_1 ; Π_2 : если x есть A_2 , то y есть B_2 ; и т.д. Здесь x - входная ЛП, y - выходная ЛП (переменная вывода), A и B - ФП, определенные соответственно на множествах значений x и y .

- композиционное правило вывода заключается в следующем: знание эксперта, выраженное в базе правил вывода, отражает нечеткое отношение предпосылки A и заключения B , т.е. $R: A \rightarrow B$, где " \rightarrow " - операция нечеткой импликации. Процесс получения нечеткого результата вывода B' с использованием данного наблюдения A' и знания $A \rightarrow B$ представляется в виде формулы:

$$B' = A' \circ R = A' \circ (A \rightarrow B),$$

где " \circ " - операция композиции. Данное правило лежит в основе работы большинства нечетких экспертных систем.

Общий логический вывод в предлагаемых нечетких экспертных системах оценки качества осуществляется в четыре этапа.

1. Введение нечеткости (фазификация). Функции принадлежности, определенные на входных ЛП, применяются к их фактическим значениям для определения степени истинности каждой предпосылки каждого правила вывода.

2. Логический вывод (импликация). Вычисленное значение истинности для предпосылок каждого правила применяется к заключениям каждого правила. Это приводит к модифицированному нечеткому множеству, которое будет назначено каждой переменной вывода для каждого правила. В качестве правил логического вывода обычно используются только операции \min (МИНИМУМ) или prod (УМНОЖЕНИЕ).

3. Объединение (композиция). Все модифицированные нечеткие множества для каждой переменной вывода (во всех правилах), объединяются вместе, чтобы сформировать одно комбинированное нечеткое множество для каждой переменной вывода. При подобном объединении обычно используются операции \max (МАКСИМУМ) или sum (СУММА).

4. Приведение к четкости (дефазификация). Процесс преобразования выходного нечеткого множества в скалярное четкое значение. Чаще всего в качестве итогового значения используется "центр тяжести" ФП выходного нечеткого множества (центроидный метод дефазификации, см. ниже).

Наиболее распространенными являются два алгоритма нечеткого вывода - алгоритм Mamdani и алгоритм Larsen. Пусть, для простоты, базу знаний составляют два нечетких правила вывода:

П₁: если x есть A_1 и y есть B_1 , то z есть C_1 ;

П₂: если x есть A_2 и y есть B_2 , то z есть C_2 ,

где x и y - входные ЛП, z - переменная вывода, $A_1, A_2, B_1, B_2, C_1, C_2$ - заданные ФП. Если четкое значение z_0 необходимо определить на основе данной информации для четких значений x_0 и y_0 , то алгоритм вывода Mamdani записывается так:

$$\mu_z(z) = C(z) = \max(C_1'(z), C_2'(z)) = \max(\min(\alpha_1, C_1(z)), \min(\alpha_2, C_2(z))),$$

а алгоритм вывода Larsen выглядит следующим образом:

$$\mu_z(z) = C(z) = \max(C_1'(z), C_2'(z)) = \max(\alpha_1 C_1(z), \alpha_2 C_2(z)).$$

Здесь α_1 и α_2 - степени истинности предпосылок каждого из правил для данных значений x_0 и y_0 , $C(z)$ - комбинированное нечеткое множество выходной переменной, соответствующее логическому выводу. Приведение к четкости (для нахождения z_0) в обоих алгоритмах осуществляется с применением центроидного метода:

$$z_0 = \frac{\int_{\Omega} zC(z)dz}{\int_{\Omega} C(z)dz}.$$

Для нечетких экспертных и управляющих систем характерны определенные недостатки, например, исходный набор постулируемых нечетких правил формулируется экспертом-человеком и может оказаться неполным или противоречивым; вид и параметры ФП, описывающих входные и выходные переменные, выбираются субъективно и могут оказаться не вполне отражающими реальную действительность. Несмотря на это возможность эффективного использования аппарата нечеткой логики в системах принятия решений и экспертных системах доказана к настоящему времени целым рядом теорем.

Алгоритмы нечетких выводов реализованы в пакетах нечеткой логики. Один из наиболее развитых пакетов - Fuzzy Logic Toolbox, входящий в систему технических вычислений MATLAB, - представляет собой совокупность прикладных программ, относящихся к теории нечетких множеств и позволяющих конструировать нечеткие экспертные и/или управляющие системы. Предлагаемые в настоящей работе экспертные системы оценки качества лингвистических ИС реализованы в прикладном пакете Fuzzy Logic Toolbox.

В целом приведенный в Главе 1 обзор показывает широту формальных методов, отработанных за столетия и появившихся в последнее время. Однако становится ясно, что многие формальные методы, как правило, вероятностно-статистического направления, активно применяющиеся в лингвистике, оказываются либо излишне сложными, либо малоэффективными, поэтому для создания экспертных систем оценки интеллектуальных лингвистических систем был выбран аппарат нечеткой логики.

В Главе 2 рассмотрены современные методы оценки интеллектуальных систем (ИС), изложены основные положения предлагаемого в работе формального подхода к оценке лингвистических ИС, базирующегося на нечеткой логике, а также приведены основные результаты оценки автоматизированного ДСА с помощью предлагаемого метода.

Экспертные оценки эффективности необходимы в любых областях науки, в частности, такие оценки требуются интеллектуальным системам (ИС) (Бешелев С.Д., Винокуров А.А., Григорьев В.А., Емельянов С.В., Жигалов В.А., Ивакин Я.А.,

Кормышев В.М., Лахути Д.Г., Панкова Л.А. и др.). Современные методы оценки отличаются широким диапазоном возможностей, но принципиальные вопросы, определяющие надежность и компетентность оценок, по-прежнему не решены. Современные математические методы экспертных оценок - это в основном статистические методы, применяемые к объектам нечисловой природы. Экспертные оценки объединяют методы организации работы со специалистами-экспертами и методы обработки мнений экспертов, выраженных в количественной и/или качественной форме с целью подготовки информации для принятия решений лицами, принимающими решения (ЛПР). В настоящее время не существует научно обоснованной классификации методов экспертных оценок и, тем более, однозначных рекомендаций по их применению.

Эффективность информационных систем, содержащих элементы искусственного интеллекта, решающим образом определяется качеством баз знаний, имеющихся в системе. Однако до последнего времени отсутствовали надежные методы определения степени достоверности экспертных заключений при заполнении баз знаний. Аппарат теории нечетких множеств позволяет учесть нечеткий вид экспертных знаний и правил, а также такие трудно формализуемые факторы, как теоретические знания эксперта; его практический опыт; особенности памяти; логические способности; "инженерное чутье", интуицию и т.д. Недавно появились нечеткие алгоритмы определения компетентности эксперта.

Важную роль в методе экспертных оценок играют математические модели поведения экспертов, которые в основном базируются на предположении о возможности вероятностно-статистического описания поведения экспертов с помощью непрерывных функций распределения (Григорьев В.А.). Методы анализа экспертных оценок основаны на предположении, что эксперты оценивают интересующий ЛПР параметр с некоторыми ошибками, поэтому при анализе мнений экспертов применяются самые разнообразные статистические методы, а также методы статистики объектов нечисловой природы. Однако традиционная теория вероятностей и математическая статистика испытывают серьезные трудности при применении в методе экспертных оценок. В первую очередь проблема состоит в том, что распределение ответов экспертов не всегда подчиняется заранее известному закону, например, закону Гаусса (нормальное распределение). Выходом вновь является использование аппарата нечеткой логики, который оказывается

гораздо более эффективным при реализации метода экспертных оценок, чем традиционная теория вероятностей.

Задачи получения оценок, в частности оценок эффективности, решаются в различных областях человеческой деятельности. В задачах управления и контроля часто используется модель принятия решений на основе композиционного правила вывода. Нечеткие выводы широко используются в задачах классификации и диагностики. Отдельную проблему представляет собой оценка знаний. Под процессом оценивания знаний понимается взаимодействие эксперта и обучаемого, состоящее из двух фаз: фазы представления модели знаний обучаемого и фазы восприятия экспертом этой модели. Важной является проблема оценки эффективности информационного обеспечения (ИО) той или иной области человеческой деятельности. Доказано, что эффективность ИО определяется в первую очередь эффективностью имеющихся в информационной системе специфических методов обработки информации. Для оценки информационных технологий (ИТ) используется целая гамма специальных методик, призванных оценивать нематериальные преимущества, которые дают ИТ. Большинство подходов заимствовано из мира финансов и стратегии бизнеса.

В настоящее время не существует единого формального метода оценки лингвистических систем. Существуют разрозненные методы оценки качества СМП, ИПС и других систем обработки текстов. Например, метод оценки качества СМП, совмещающий оценку качества программного обеспечения и лингвистической характеристики, метод оценки соответствия смысла запроса на естественном языке (ЕЯ) предметной области ИПС, основанный на вычислении функции семантической корреляции, которая оценивает вероятность выдачи релевантного ответа на конкретный запрос ИПС с определенной предметной областью. Традиционно оценка качества систем обработки текстовой информации проводится по следующим параметрам: качество программного обеспечения; удобство интерфейса пользователя; лингвистическая характеристика; организационное значение; цена.

Можно с уверенностью говорить об отставании темпов продвижения методов оценки качества лингвистических систем от бурного роста и развития самих систем. Одной из причин этого является сложность формализации понятия "качество системы". Традиционные подходы к оценке эффективности лингвистических систем основаны в основном на методе экспертных оценок. Суть экспертного тестирования в случае лингвистической системы чаще

всего сводится к определению максимально достижимого уровня сложности запросов, эффективно обрабатываемых этой системой. Ранжирование запросов по уровню сложности позволяет получить представление о качестве функционирования системы независимо от мнений экспертов и экспериментально подтвердить адекватность оценки. К другим методам оценки относятся методы фреймов и нечетких множеств. Последние являются эффективным способом формализации некоторых свойств естественного интеллекта.

Базой для формирования единого подхода к оценке качества функционирования ИС, на наш взгляд, являются методы теории нечеткой логики. Новый формальный метод оценки качества интеллектуальных лингвистических систем (СМП, ИПС и т.д.) базируется на создании нечетких экспертных систем оценки качества на основе композиционного правила вывода. Предлагаемый подход претендует на универсальность, поскольку нечеткие экспертные системы позволяют оценивать качество работы ИС с единых позиций в зависимости от степени сложности требуемой оценки. Основные этапы предлагаемого подхода следующие.

Определяется критерий оценки качества ИС, который выбирается экспертом-филологом, тестирующим систему, в зависимости от конкретного типа ИС. Для формализации этого параметра вводится соответствующая ЛП "Выбранный критерий". В предлагаемых экспертных системах ее базовое терм-множество состоит из трех нечетких переменных $T = \{ "A_1", "A_2", "A_3" \}$, определенных в некотором диапазоне от X_1 до X_2 . Выбранный критерий оценки качества представляет из себя входную переменную нечеткой экспертной системы.

Формализация понятия "качество системы" осуществляется с помощью введения ЛП "Качество" (Q). Понятие "качество" определяется, в первую очередь, критерием оценки, взятым за основу при конструировании нечеткой экспертной системы. "Качество" является переменной вывода экспертной системы. Ее базовое терм-множество состоит из трех нечетких переменных $T_Q = \{ "Плохое", "Среднее", "Хорошее" \}$. Было решено оценивать качество функционирования системы в процентах, поэтому областью определения ЛП "Качество" является отрезок $Z = [0, 100]$.

Процедура задания на области определения ЛП соответствующих ФП носит в большинстве случаев субъективный характер. В нашем случае автор сам задавал для каждой нечеткой переменной значения ее ФП. От выбора ФП во многом зависит конкретный результат работы экспертной системы. Во всех

предлагаемых экспертных системах значениями ЛП являются нечеткие числа, что позволяет использовать для задания их ФП нечеткие числа (L-R) - типа. Вопрос о выборе оптимальных ФП пока остается открытым.

Нечеткие переменные C_1 = "Плохое", C_2 = "Среднее", C_3 = "Хорошее" были представлены нечеткими числами (L-R) - типа с гауссовскими функциями
$$L(z) = \exp\left(-\frac{(a_1 - z)^2}{\alpha^2}\right) \quad \text{и} \quad R(z) = \exp\left(-\frac{(z - a_2)^2}{\beta^2}\right).$$

Границы толерантности a_1 и a_2 , а также коэффициенты нечеткости α и β , для каждого нечеткого числа подбирались экспериментально.

Основой оценивающей экспертной системы является база нечетких правил логического вывода. Правила формулируются экспертом (автором), исходя из его компетентности и логических соображений. Система обрабатывает все правила вывода одновременно, определяя степень их влияния на оценку качества функционирования СМП, ИПС или системы обработки текста с помощью автоматизированного ДСА. В работе не ставилась цель выбора оптимального алгоритма нечеткого вывода, поэтому использовались два наиболее распространенных алгоритма Mamdani и Larsen. Результаты работы алгоритмов оказались очень близки. Таким образом, было доказано, что конкретная реализация алгоритма нечеткого вывода не играет существенной роли при создании экспертных систем оценки качества ИС.

Предлагаемые нечеткие экспертные системы оценки качества допускают неограниченное включение различных критериев (не обязательно количественных), определяющих итоговое понятие качества. Для этого достаточно формализовать новый критерий путем введения соответствующей ЛП и модифицировать базу правил вывода, чтобы, наряду с имеющимися, она включала новый критерий оценки качества лингвистической системы. После этого необходимо предоставить нечеткой экспертной системе конкретные значения входных переменных (критериев качества), определенные экспертом. В результате совместной обработки базы правил вывода будет получена итоговая оценка. Предлагаемая формальная методика позволяет количественно сравнивать эффективность работы различных лингвистических систем одного типа.

Продолжением эксперимента по использованию автоматизированного ДСА для обработки корпуса текстов различных жанров, принадлежащих к одной тематической области, описанного в Главе 1, явилась разработка модели полуавтоматической динамической

экспертной системы, построенной на базе нечетких правил вывода, которая может быть использована для оценки качества обработки текста с помощью автоматизированного ДСА в задаче жанровой атрибуции текстов. Предлагается использовать количество терминов, выделенных из текста с помощью ДСА, и наличие их в тезаурусе той предметной области, которой принадлежит текст, в качестве критериев оптимальности выбора значений дистрибутивного порога и уровня доверительной вероятности и, как следствие, качества построенных в результате ДСА графов дистрибутивной и статистической связности. Для определения оптимального порогового значения разработаны итерационные алгоритмы для обоих видов ДСА. Ниже приведено краткое описание и результаты работы алгоритма для дистрибутивного варианта ДСА.

Будем характеризовать систему обработки текста с помощью ДСА рассмотренным ранее понятием "качество". Качество считается наивысшим, если в результате анализа построен оптимальный (в соответствии с введенными критериями) граф дистрибутивной связности. Оценка качества осуществляется с помощью специально разработанной для этой цели нечеткой экспертной системы. Работа алгоритма начинается с задания некоторого (достаточно низкого) начального значения дистрибутивного порога (b). Затем проводится автоматизированный ДСА выбранного текста, и эксперт определяет количество ключевых слов в получающемся дистрибутивном графе и процент ключевых слов, содержащихся в тезаурусе. По полученным данным описанная ниже экспертная система определяет начальную оценку качества ДСА (q). Далее вычисляется новое значение дистрибутивного порога: $b = b + 0,1 * (1 - 0,01 * q)$. Коэффициент 0,1 определяет скорость сходимости алгоритма и подбирается эмпирически. Проводится ДСА с найденным порогом. Вновь определяются конкретные значения выбранных критериев оценки качества, которые вводятся в экспертную систему для получения нового значения оценки качества ДСА, которое сравнивается с предыдущим. Если новое значение выше предыдущего, то значение дистрибутивного порога увеличивается, в противном случае осуществляется уменьшение порога. Таким образом, дистрибутивный порог итерационно повышается, если "качество" ДСА возрастает, и понижается, если "качество" начинает падать. Процесс прекращается, когда "качество" перестает меняться. В результате описанной процедуры для выбранного текста определяется оптимальное значение дистрибутивного порога и строится граф связности, по

характеристикам которого эксперт осуществляет жанровую атрибуцию текста.

При работе с текстами различных жанров (монография, главы из книги, статьи из специализированных журналов, РЖ) удавалось достичь наивысшего значения оценки качества за сравнительно небольшое число итераций. Удовлетворительное значение полноты и точности лексики графов, построенных при оптимальных значениях порогов, позволило произвести жанровую атрибуцию исследованных текстов с высокой степенью достоверности. На рисунке 1 показана динамика поведения дистрибутивных порогов для нескольких типов текстов. Динамика оценки качества дистрибутивного варианта автоматизированного ДСА для соответствующих текстов показана на рисунке 2. Все оптимальные пороги соответствовали "хорошему" качеству системы. Для монографии и одной из статей оптимальный дистрибутивный порог равен 0,9, для РЖ и остальных статей - 0,8. Эксперимент всегда начинался с самого низкого значения дистрибутивного порога, равного 0,6.

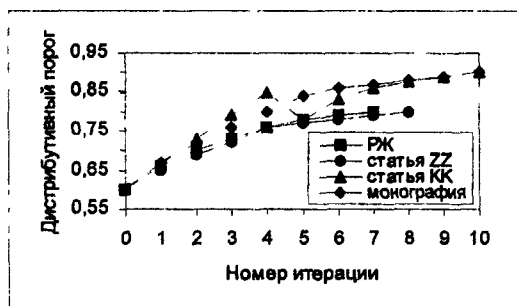


рис. 1

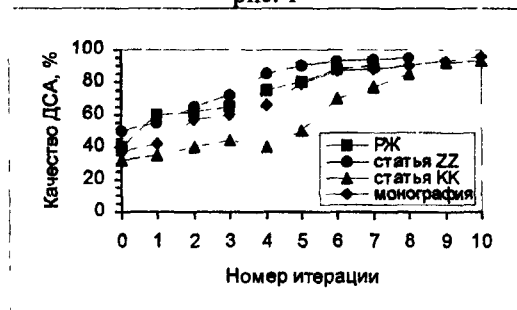


рис. 2

Общие принципы построения нечеткой экспертной системы были рассмотрены выше. Для формального описания критериев

оценки качества были введены две ЛП: "Количество терминов" (N) и "Процент в тезаурусе" (P), которые являются входными переменными нечеткой экспертной системы. Их базовые терм-множества состоят из трех нечетких переменных $T_N = \{A_1 = \text{"Малое"}, A_2 = \text{"Среднее"}, A_3 = \text{"Большое"}\}$, $T_P = \{B_1 = \text{"Небольшой"}, B_2 = \text{"Средний"}, B_3 = \text{"Большой"}\}$. Областью определения переменных A_i был взят отрезок $X = [0, 20]$. Выбор значения для максимального количества выделенных терминов определялся результатами проведенного автором эксперимента по ДСА. Областью определения переменных B_i является отрезок $Y = [0, 100]$. В рассматриваемой нечеткой экспертной системе нечеткие множества A_i и B_i были заданы следующими наборами параметров: $A_1 = (-3, 3, 3, 3)$; $A_2 = (8, 12, 3, 3)$; $A_3 = (17, 23, 3, 3)$; $B_1 = (-20, 20, 10, 10)$; $B_2 = (40, 60, 10, 10)$; $B_3 = (80, 120, 10, 10)$. Выходной переменной экспертной системы является ЛП "Качество" ДСА. Соответствующие ФП нечетких множеств C_i имеют параметры: $C_1 = (-10, 10, 20, 20)$; $C_2 = (40, 60, 20, 20)$; $C_3 = (90, 110, 20, 20)$.

База правил вывода была сформулирована на основе результатов эксперимента по использованию автоматизированного ДСА. База правил состоит из девяти утверждений:

Если количество терминов малое, и процент наличия их в тезаурусе небольшой, то качество системы плохое;

Если количество терминов малое, и процент их наличия в тезаурусе средний, то качество системы плохое;

Если количество терминов малое, и процент их наличия в тезаурусе большой, то качество системы хорошее;

Если количество терминов среднее, и процент их наличия в тезаурусе небольшой, то качество системы плохое;

Если количество терминов среднее, и процент их наличия в тезаурусе средний, то качество системы среднее;

Если количество терминов среднее, и процент их наличия в тезаурусе большой, то качество системы хорошее;

Если количество терминов большое, и процент их наличия в тезаурусе небольшой, то качество системы среднее;

Если количество терминов большое, и процент их наличия в тезаурусе средний, то качество системы хорошее;

Если количество терминов большое, и процент их наличия в тезаурусе большой, то качество системы хорошее.

Поскольку в левой части всех правил вывода используется связка "И", нахождение степени истинности для предпосылок каждого из правил осуществлялось с помощью операции МИНИМУМ. Полученные значения степени истинности использовались для

нахождения модифицированных ФП заключений правил либо в качестве уровней "отсечения" (алгоритм Mamdani), либо в качестве весовых коэффициентов (алгоритм Larsen). Композиция модифицированных нечетких подмножеств переменной вывода осуществлялась с помощью операции МАКСИМУМ. Для получения четкого значения оценки "качества" ДСА проводилась операция приведения к четкости с помощью центроидного метода.

Ниже приведены некоторые результаты моделирования работы системы в пакете нечеткой логики Fuzzy Logic Toolbox. На рисунке 3 (а, б) показаны ФП входных нечетких переменных экспертной системы. Левая и правая области нечеткости всех переменных определяются функциями Гаусса заданной ширины. Область толерантности подбиралась экспериментально. ФП нечетких переменных, являющихся значениями ЛП "Качество" ДСА, показаны на рисунке 4. Итоговая поверхность отклика исследуемой экспертной системы представлена на рисунке 5. Переменные X,Y соответствуют входным переменным, а переменная Z отражает "качество" работы автоматизированного ДСА.

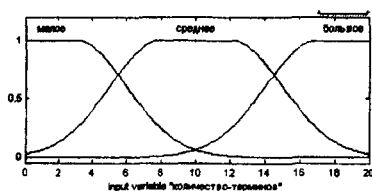


рис. 3а

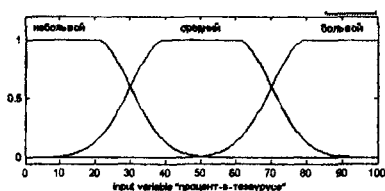


рис. 3б

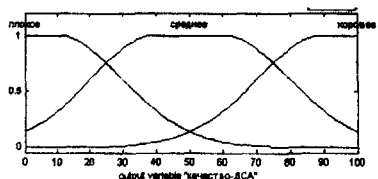


рис. 4

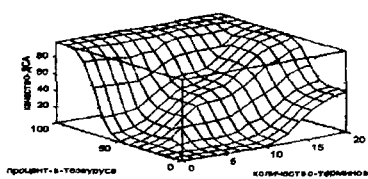


рис. 5

Проведенный эксперимент показал, что жанровую атрибуцию текстов можно осуществлять и с использованием графов статистической связности. Проблема заключается в оптимальном выборе уровня доверительной вероятности при построении графа. Для решения этой задачи была разработана нечеткая экспертная система оценки качества статистического варианта автоматизированного ДСА, с помощью которой можно получить оптимальное значение уровня

доверительной вероятности и построить наиболее полный и точный граф статистической связности. Данная система принципиально ничем не отличается от описанной выше. Для монографии и РЖ оптимальный уровень доверительной вероятности равен 0,9, для статей - либо 0,8, либо 0,7. Качество системы при этих уровнях оценивалось как "хорошее". Жанровая атрибуция с использованием оптимальных графов статистической связности была проведена с высокой степенью точности.

Глава 3 посвящена формальным решениям в системах машинного перевода (СМП). Создание СМП является одной из наиболее интересных проблем, связанных с моделированием языковой деятельности человека. Чтобы получить идеальный машинный перевод текста, необходимо воссоздать на компьютере, как минимум, три модели человеческой деятельности: модель мышления, модель речевой деятельности и модель мира, в котором живет человек.

За последние полвека активных работ предлагалось большое количество формальных решений по созданию СМП. В настоящее время перспектива появления более совершенных систем, обеспечивающих лучшее качество собственно машинного перевода, связана с разработкой моделей, более точно воспроизводящих действия человека-переводчика, использующих формализованные синтаксические и семантические методы анализа и синтеза, эквивалентно преобразующих смысловое содержание в тех пределах, которые допустимы для того, чтобы перевод мог бы еще называться переводом, а не переходил бы в реферирование. Прогресс в этих направлениях дается с большим трудом.

В данный момент существуют следующие типы компьютерного перевода:

1. Системы машинного перевода - программы, осуществляющие полностью автоматизированный перевод. Большинство современных программ машинного перевода выполняет перевод, опирающийся на набор заранее определенных правил.
2. Системы с функцией Translation Memory (переводческие накопители), предоставляющие средства для перевода, выполняемого человеком с помощью компьютера. Переводческий накопитель представляет собой специальным образом организованную индивидуальную базу данных.
3. Контролируемый язык и машинный перевод на основе базы знаний. Используется, чтобы упростить выражения исходного текста и повысить качество перевода.

4. Онлайн-переводчики, выполняющие перевод текста прямо в окне браузера сети Интернет, не требуя установки программы на компьютер пользователя.
5. Онлайн-словари, способные переводить слова и словосочетания, вводимые с клавиатуры, а также осуществлять поиск переводных эквивалентов.
6. Услуги по переводу в сети Интернет, предлагаемые компаниями-производителями СМП.

По современным представлениям процесс машинного перевода текстов с одного естественного языка на другой может быть разделен на три этапа: семантико-синтаксический анализ текста на входном языке; трансфер; семантико-синтаксический синтез текста на выходном языке. В различных СМП эти этапы не всегда бывают четко выделены, но, тем не менее, они всегда присутствуют.

Получить полноценный машинный перевод довольно трудно - существенно мешает омонимия и полисемия слов, несовпадение типов языка (аналитический, синтетический) и связанные с этим проблемы, и многое другое. Наиболее часто встречающимися ошибками при текстовом анализе являются неправильное распознавание антецедентов местоимений, анафорических отношений и эллиптических структур, что влечет за собой ошибки в синтезе. Кроме того в каждом языке и культуре существуют свои особенности, которые иногда весьма трудно предусмотреть при машинном переводе. Современные СМП можно использовать для перевода больших массивов однородных текстов с небольшим узкоспециализированным словарным запасом и упрощенной грамматикой.

Одним из решений перевода с помощью компьютера являются электронные словари. В работе предлагается их типологическая классификация. Все многообразие существующих словарей, созданных и использующихся при помощи компьютера, будем называть "электронными словарями" (родовое понятие), в то время как "машинными словарями" имеет смысл называть словари, созданные для работы в среде ИПС, "автоматическими словарями" - словари, работающие в СМП со встроенным в структуру машинного перевода языком-посредником, а "компьютерными словарями" - словари любого типа (переводные, учебные, терминологические, толковые и т.п.), предназначенные для использования на современных компьютерах, т.е. словари, обладающие свойствами развитой поисковой системы, дружелюбным интерфейсом (со структурой гиперссылок и пр.), активно использующие мультимедийные средства.

Особенностью лучших современных российских компьютерных переводных словарей являются их обратимость, гибкость и динамичность. Под обратимостью переводного словаря понимается возможность произвольно, по желанию пользователя, менять местами входной и выходной языки. Обратимый словарь должен иметь два разноязычных входа. Переводной эквивалент в словарной статье такого словаря фиксирует равенство абсолютных ценностей двух разноязычных слов. Гибкость компьютерного словаря - это комплекс лингвистических и программных приемов, которые упрощают обращение к словарю, расширяют возможности пользователя при работе с иноязычным текстом, снижают требования к уровню знания входного языка, а также позволяют использовать компьютерный словарь не только для хранения лексических массивов, но и для решения других задач АОТ и автоматизации лексикографических работ. Наиболее важным средством повышения гибкости словаря является упрощение формы запроса, т.е. обращения к словарю. Динамичность компьютерных словарей означает, что отбор лексики не прекращается после составления словаря, а продолжается в течение всего времени его функционирования. Переменность состава и способность воспринимать изменения в ходе функционирования являются принципиальными характеристиками, отличающими компьютерные словари от обычных.

Проблема обращения к компьютерным словарям в естественной языковой форме является одной из основных при их построении. Без решения этой проблемы в полном объеме компьютерный словарь не может претендовать на успешное решение своих основных задач - помощь человеку при работе с иноязычными текстами и автоматизацию лексикографических работ.

Что касается формальных решений с точки зрения теории машинного перевода, то наиболее интересными разработками, на наш взгляд, являются следующие: СМП ЭТАП-3, базирующаяся на компьютерной модели естественного языка на основе лингвистической теории "Смысл <--> Текст" И.А. Мельчука (ИППИ РАН); система фразеологического машинного перевода RETRANS, рассматривающая фразеологические сочетания, выражающие понятия, отношения между понятиями и типовые ситуации в качестве основных единиц смысла при переводе текста (ВИНИТИ); комплекс ЯРАП для лингвистических исследований по японско-русскому автоматическому переводу (Институт востоковедения РАН). Годы исследований показали, что ни один из подходов изолированно не решит все проблемы машинного перевода за обозримый промежуток времени. В той же степени

маловероятно, что в ближайшее время появится и будет до определенного уровня разработан новый "идеальный" подход. Существенный прогресс может быть достигнут только при сочетании преимуществ разных подходов.

В Главе 3 дается описание системы русско-английского машинного перевода АСПЕРА, которая создавалась преимущественно в ВЦП РАН под руководством д.ф.н. Королева Э.И. с участием автора диссертации. Это описание позволяет получить подтверждение адекватности проведенного эксперимента по оценке эффективности СМП.

Существует немало методик оценки СМП, чаще всего они опираются на статистические данные и мнение эксперта-переводчика, реже на мнение конечного пользователя. Основной проблемой при сравнении считается неравноценность программ перевода. По-прежнему важным является практический подход к оценке переводов, выполненных с помощью СМП, который до настоящего времени осуществляется в большей степени эмпирическим путем на основании профессионального опыта и знаний оценивающего. С точки зрения теории перевода адекватность переводов традиционно сводится к категориям семантической (смысловой) полноты и точности, дополняемым стилистической эквивалентностью, включающей, в частности, принцип соответствия текста перевода стилистическим нормам языка перевода (Комиссаров В.Н., Нелюбин Л.Л. и др.). Именно на основе этих параметров чаще всего и выводятся оценки качества перевода.

Многие разработчики и эксперты полагают, что основными критериями эффективности перевода являются понятность и адекватность. Существуют различные методики оценки понятности переведенного машинной текста, вводятся разнообразные шкалы понятности (Марчук Ю.Н., Nagaо M., Wall R.E. и пр.) . Критерий адекватности необходим для подтверждения правильной передачи смысла оригинала. Для обоих критериев оценки качества перевода формальных методик пока нет.

Необходимо учитывать также тот факт, что, собственно, оценивается - сама система или перевод, выполненный ею. Оценка самой системы, ее лингвистического обеспечения вызывает особые трудности. Едва ли не единственным методом тестирования является метод "черного ящика". Оценки качества машинного перевода, как правило, опираются на мнения экспертов, которые, в свою очередь, базируются на типах ошибок, встречающихся при переводе. Единого стандарта определения эффективности действующих систем не

создано. Однако весьма перспективной и отражающей реальную ситуацию может стать оценка эффективности СМП с помощью нечетких экспертных систем.

Предложенные в работе нечеткие экспертные системы на основе композиционного правила вывода могут использовать различные критерии оценки качества СМП. Наиболее простым критерием является количество ошибок, допущенных системой при переводе контрольного текста. Более сложным критерием, требующим разработки специальной шкалы оценок, является степень серьезности ошибок машинного перевода.

Для проведения экспертизы оцениваемой СМП предъявляется заранее подобранный контрольный текст определенной длины и заданной тематики. Впоследствии результат работы СМП обрабатывается экспертом-филологом, который определяет количество допущенных при переводе ошибок или степень их серьезности, после чего полученное значение критерия оценки качества вводится в экспертную систему, оценивающую качество работы тестируемой СМП. Определение единого понятия "качество" для текстов разной тематики позволяет сравнить работу различных СМП количественно в рамках единой шкалы оценок. Используя один и тот же контрольный текст, можно сравнивать СМП, находящиеся на одной стадии функционирования (разработка, отладка, коммерческий продукт).

Количество ошибок, допущенных СМП при переводе, представляет собой входную переменную нечеткой экспертной системы. Базовое терм-множество ЛП "Количество ошибок" состоит из трех нечетких переменных $T_N = \{A_1 = \text{"Малое"}, A_2 = \text{"Среднее"}, A_3 = \text{"Большое"}\}$, определенных в диапазоне от 0 до 40. В описываемом ниже эксперименте с СМП АСПЕРА использовались контрольные тексты средним объемом 2 Кб (одна машинописная страница). Выбор такого размера текста определялся рядом причин, в частности тем, что для фрагмента меньшего размера (предложения или абзаца) может получиться более качественный перевод, поскольку в них есть возможность более точно учитывать особенности связного текста, анафорические связи и пр. Если количество ошибок при переводе такого текста превышает 40, то, на наш взгляд, подобную СМП нельзя считать работоспособной. Входные нечеткие переменные были представлены нечеткими числами (L-R) - типа со следующими параметрами: $A_1 = (-6, 6, 6, 6)$; $A_2 = (16, 24, 4, 4)$; $A_3 = (34, 46, 6, 6)$. Соответствующие ФП приведены на рисунке 6. Выходной переменной нечеткой экспертной системы является ЛП "Качество" СМП. Нечеткие множества C_i , использующиеся в экспертных системах, оценивающих

качество СМП и ИПС, были заданы параметрами: $C_1 = (-10, 10, 20, 20)$; $C_2 = (40, 60, 10, 10)$; $C_3 = (90, 110, 20, 20)$.

Нечеткая экспертная система работает на основе следующей базы правил вывода:

Если количество ошибок малое, то качество системы хорошее;

Если количество ошибок среднее, то качество системы среднее;

Если количество ошибок большое, то качество системы плохое.

Для получения логического вывода используется алгоритм Larsen на основе композиционного правила. Результат работы системы представлен на рисунке 7 в виде графика зависимости оценки качества СМП от количества ошибок перевода. На этом же рисунке различными символами отмечены оценки качества СМП АСПЕРА, полученные с помощью описанной нечеткой экспертной системы для контрольных текстов различной тематики: информатика и вычислительная техника (●), медицина (●), история философии (▲), охрана природы (✕), бизнес (■). Выбор тематики был определен знанием автора состава словаря СМП АСПЕРА. Количественные результаты тестирования сведены в Таблице 1 (см. ниже).

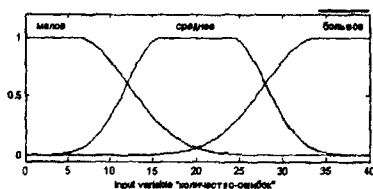


Рис. 6

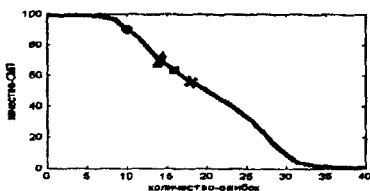


Рис. 7

Другим вариантом экспертной системы, оценивающей качество СМП, как уже упоминалось, является система, критерием оценки в которой является характер допущенных при переводе ошибок (степень их серьезности). Чтобы использовать степень серьезности ошибок в качестве критерия оценки эффективности СМП, необходимо создать "шкалу серьезности" ошибок, т.е. расположить (ранжировать) возможные ошибки, допускаемые при машинном переводе, например, в порядке возрастания их степени серьезности, присвоив каждой ошибке некоторый коэффициент. Это задача эксперта-филолога, которая решается путем анализа заданного корпуса текстов и их переводов различными СМП. Очевидно, что разные эксперты могут получить разные шкалы, и соответственно оценки качества одной и той же СМП будут различными. Вопрос разработки универсальной, единой оценочной шкалы ошибок весьма непрост. В работе предлагается

алгоритм построения нечеткой экспертной системы на основе разработанного варианта такой шкалы.

Пусть необходимо оценить качество СМП по степени серьезности допущенных при переводе ошибок. При наличии нескольких ошибок различной степени серьезности вычисляется итоговый коэффициент серьезности. Предложено несколько вариантов его расчета, оптимальным оказалось вычисление среднего значения. Средний коэффициент серьезности ошибок используется в качестве входной переменной нечеткой экспертной системы. Формально он описывается с помощью ЛП "Коэффициент серьезности ошибок" с базовым терм-множеством $T_G = \{B_1 = \text{"Небольшой"}, B_2 = \text{"Средний"}, B_3 = \text{"Большой"}\}$. Шкала серьезности ошибок может быть нормирована, поэтому областью определения ЛП "Коэффициент серьезности ошибок" является отрезок $Y = [0, 1]$. Входные нечеткие множества имеют параметры: $B_1 = (-0.01, 0.01, 0.2, 0.2)$; $B_2 = (0.45, 0.55, 0.15, 0.15)$; $B_3 = (0.99, 1.01, 0.2, 0.2)$. ФП нечетких чисел B_i показаны на рисунке 8. Выходной переменной по-прежнему является ЛП "Качество" СМП.

База правил вывода в данном случае имеет вид:

Если коэффициент серьезности ошибок небольшой, то качество системы хорошее;

Если коэффициент серьезности ошибок средний, то качество системы среднее;

Если коэффициент серьезности ошибок большой, то качество системы плохое.

Логический вывод производится на основе композиционного правила вывода с использованием алгоритма Mamdani. График зависимости оценки качества СМП от коэффициента серьезности ошибок перевода изображен на рисунке 9. Символами на графике отмечены результаты оценки качества СМП АСПЕРА.

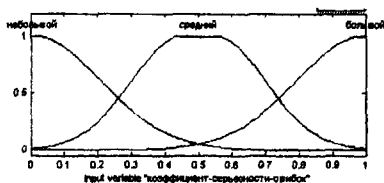


рис. 8

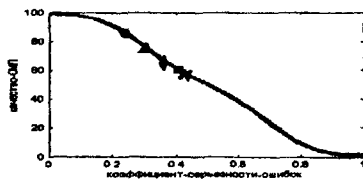


рис. 9

При переводе текста по информатике были допущены лишь незначительные ошибки. В текстах по медицине и истории философии встретились ошибки средней степени серьезности (от 0.4 до 0.7).

Грубые ошибки были зафиксированы в переводах текстов по бизнесу и охране природы. Количественные результаты тестирования представлены в Таблице 1 (см. ниже).

Исследование показало, что наиболее оптимальной является экспертная система оценки качества СМП, использующая оба приведенных выше критерия. Такая система определяет комбинированную оценку качества, в которой учитывается как общее количество ошибок, допущенных СМП при переводе, так и степень их серьезности. Для формального описания входных переменных используем введенные ранее ЛП "Количество ошибок" с областью определения $X = [0, 40]$ и "Коэффициент серьезности ошибок" с областью определения $Y = [0, 1]$. Выходной переменной экспертной системы является введенная ранее ЛП "Качество" СМП, определенная на отрезке $Z = [0, 100]$.

База правил вывода состоит из девяти утверждений:

Если количество ошибок малое, и коэффициент серьезности ошибок небольшой, то качество системы хорошее;

Если количество ошибок малое, и коэффициент серьезности ошибок средний, то качество системы хорошее;

Если количество ошибок малое, и коэффициент серьезности ошибок большой, то качество системы среднее;

Если количество ошибок среднее, и коэффициент серьезности ошибок небольшой, то качество системы хорошее;

Если количество ошибок среднее, и коэффициент серьезности ошибок средний, то качество системы среднее;

Если количество ошибок среднее, и коэффициент серьезности ошибок большой, то качество системы плохое;

Если количество ошибок большое, и коэффициент серьезности ошибок небольшой, то качество системы среднее;

Если количество ошибок большое, и коэффициент серьезности ошибок средний, то качество системы плохое;

Если количество ошибок большое, и коэффициент серьезности ошибок большой, то качество системы плохое.

Использовались оба рассмотренных в работе алгоритма нечеткого вывода. Итоговая поверхность отклика экспертной системы представлена на рисунке 10. В виде плоской "карты уровней" поверхность отклика показана на рисунке 11. В левом нижнем углу находится область "хорошего" качества СМП, соответствующая малому количеству ошибок и небольшим значениям коэффициента серьезности, правый верхний угол соответствует "плохому" качеству.

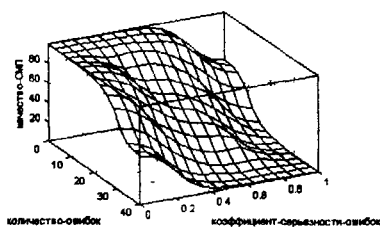


рис. 10

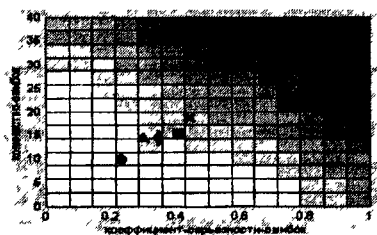


рис. 11

Результаты оценки качества СМП АСПЕРА на основе анализа перевода контрольных текстов отмечены символами на поверхности отклика. Значения итоговых оценок качества СМП АСПЕРА приведены в Таблице 1. Видно, что все три экспертные системы выставили близкие оценки качества на основе анализа одного и того же текста конкретной тематики.

Таблица 1

Тематика текста	Количество ошибок	Качество "1", %	Качество СМП АСПЕРА, %
	Козф-фициент серьезности	Качество "2", %	
Информатика	10 0.24	92 86	88
Бизнес	16 0.41	63 60	67
Медицина	14 0.371	71 65	74
История философии	14 0.314	71 74	77
Охрана природы	18 0.433	56 57	59

Эксперимент, проведенный с СМП АСПЕРА, позволяет сделать вывод об эффективности и перспективности предлагаемого формального подхода к оценке качества систем машинного перевода. Предлагаемая методика позволяет создавать нечеткие экспертные системы оценки качества СМП, использующие различные подходы к выбору критерия оценки. Например, при разделении ошибок перевода на грамматические, синтаксические и семантические можно проводить оценку качества системы перевода по типам ошибок. В этом случае

качество СМП будет характеризоваться несколькими параметрами, и способ оценки станет более гибким.

Глава 4 посвящена формальным решениям в информационно-поисковых системах (ИПС). В последнее время существенно возросла скорость появления и накопления новых данных в самых разных областях человеческой деятельности. В связи с этим появилась необходимость создания информационных систем (ИС), которые не только содержали бы в себе средства поиска фактических данных (информации), но и помогали бы находить нетривиальные способы решения разнообразных проблем. Несколько лет назад возникло новое направление - data mining (буквально – добыча данных). Результатом его реализации должны стать системы, которые самостоятельно "рассмотрев" данные, будут делать нетривиальные выводы. Особенно актуальна проблема "добычи данных" в сети Интернет. Извлекать даже простые знания из текстовых баз данных весьма непросто, и существующие в настоящее время поисковые системы сети Интернет еще далеки от совершенства.

Автоматизированная ИПС (АИПС) является формальной моделью таких сложных интеллектуальных процессов, как анализ, синтез, обобщение, логический вывод, экстракция сведений и т.д. Данные процессы моделируются в АИПС посредством двух типов формализованных процедур: составления поисковых образов документов (индексирования) и информационного поиска. АИПС на базе программных продуктов с открытым словарем реализуют также имитацию освоения предметной области путем накопления и интерпретации лексики.

Для оценки степени соответствия поисковых образов документа (ПОД) и запроса (ПОЗ) формулируется критерий смыслового соответствия, состоящий из меры их формальной релевантности (меры близости) и ее порогового значения, при превышении которого документ признается формально релевантным запросу. Однако на практике формальная релевантность вовсе не означает содержательного соответствия выданного документа запросу. Таким образом, множество формально релевантных (выданных) документов необходимо рассматривать совместно с множеством истинно релевантных документов. Применяются различные способы оценки качества, именуемые критериями оценки эффективности АИПС или способности системы выдавать релевантные и задерживать нерелевантные запросы документы.

В целях количественного описания уровня качества поиска АИПС исторически первыми были предложены частные критерии оценки: полнота и точность. Коэффициент полноты поиска определяется отношением количества выданных релевантных документов к общему числу релевантных документов в базе. Коэффициент точности поиска представляет собой отношение количества выданных релевантных документов к общему числу документов в выдаче (формально релевантных). Эти показатели измеряют на основе экспериментов в контролируемых условиях и, как правило, в процентах. Это требует наличия экспериментальной системы с фиксированным числом документов, стандартного набора запросов и множества документов, как релевантных, так и нерелевантных каждому из обрабатываемых в эксперименте запросов. Доказано, что одновременно достигнуть стопроцентной полноты и точности поиска принципиально невозможно. К тому же ясно, что отдельно взятое значение одного из частных критериев не отражает полностью качество поиска, поэтому для оценки поисковых систем обычно используют пару критериев, однозначно описывающую выдачу. Используя эту пару критериев, можно разработать нечеткую экспертную систему оценки качества ИПС.

В Главе 4 описана структура автоматизированной многоязычной базы данных с поисковыми элементами КАСКАД, в разработке которой автор принимал участие. Эта ИПС предназначена для обеспечения зарубежных пользователей информацией об отечественных программных продуктах.

На основе наиболее широко распространенных частных критериев оценки эффективности АИПС - коэффициентов полноты и точности поиска - была сконструирована универсальная нечеткая экспертная система оценки качества, входными переменными которой являются коэффициенты полноты и точности, определяемые экспертом, а переменной выхода служит оценка качества тестируемой АИПС. Входные переменные описываются ЛП "Коэффициент полноты" и "Коэффициент точности", выходная переменная описывается ЛП "Качество" ИПС. Базовые терм-множества входных ЛП одинаковы и состоят из трех нечетких переменных: $A_1 = B_1 =$ "Небольшой", $A_2 = B_2 =$ "Средний", $A_3 = B_3 =$ "Большой", определенных на отрезке $[0, 1]$. Для каждой введенной нечеткой переменной задана ее ФП (нечеткое множество на области определения). Поскольку рассматриваемые критерии оценки качества АИПС равнозначны и в равной степени должны влиять на переменную выхода, ФП нечетких множеств A_i и B_i одинаковы. Они показаны на рисунке 12. Нечеткие

множества A_i и B_i , представленные нечеткими числами (L-R) - типа с гауссовскими ФП, имеют параметры: $A_1 = B_1 = (-0.2, 0.2, 0.1, 0.1)$; $A_2 = B_2 = (0.4, 0.7, 0.1, 0.1)$; $A_3 = B_3 = (0.9, 1.1, 0.1, 0.1)$. Границы толерантности и левый и правый коэффициенты нечеткости были подобраны экспериментально. Нечеткие множества C_i базового термножества ЛП "Качество" ИПС были определены ранее.

База правил вывода экспертной системы имеет вид:

Если коэффициент полноты небольшой, и коэффициент точности небольшой, то качество системы плохое;

Если коэффициент полноты небольшой, и коэффициент точности средний, то качество системы плохое;

Если коэффициент полноты небольшой, и коэффициент точности большой, то качество системы среднее;

Если коэффициент полноты средний, и коэффициент точности небольшой, то качество системы плохое;

Если коэффициент полноты средний, и коэффициент точности средний, то качество системы среднее;

Если коэффициент полноты средний, и коэффициент точности большой, то качество системы хорошее;

Если коэффициент полноты большой, и коэффициент точности небольшой, то качество системы среднее;

Если коэффициент полноты большой, и коэффициент точности средний, то качество системы хорошее;

Если коэффициент полноты большой, и коэффициент точности большой, то качество системы хорошее.

Итоговая поверхность отклика в виде плоской "карты уровней" изображена на рисунке 13. Область "хорошей" оценки качества ИПС ($> 90\%$) занимает правый верхний угол рисунка. Эта область соответствует большим значениям коэффициентов полноты и точности ($> 0,9$). Малые значения коэффициентов ($< 0,2$) определяют область "плохой" оценки качества ($< 10\%$). Если один из коэффициентов "большой", а второй - "небольшой", то оценка качества ИПС составляет около 50%.

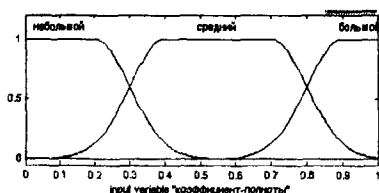


рис. 12

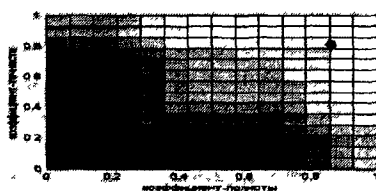


рис. 13

Нечеткая экспертная система была использована для получения оценки качества ИПС КАСКАД. Состав базы данных этой ИПС известен автору, поэтому тестирование проводилось на контрольном наборе документов. Оценка качества ИПС, полученная на основе анализа материала, выданного в ответ на первоначальный запрос, сформулированный на ЕЯ, составила $Q = 54\%$ (коэффициент точности $P = 0,675$, коэффициент полноты $R = 0,49$). Наличие "информационного шума" в выдаче объясняется тем, что одно лишь упоминание слов запроса в выданной статье еще не гарантирует ее истинной релевантности запросу. Путем применения методики модификации запроса удалось повысить коэффициент точности до значения $P = 0,76$. Оценка качества ИПС при этом возросла до $Q = 69\%$. Увеличить значение коэффициента полноты можно путем добавления в запрос синонимов ключевых слов. Если поисковая система не имеет словаря синонимов, или добавленный синоним не содержится в имеющемся словаре, то количество релевантных документов в выдаче увеличивается. При тестировании ИПС КАСКАД добавление синонимов существенно повысило значение коэффициента полноты ($R = 0,62$). Коэффициент точности стал равен $P = 0,8$. Оценка качества составила $Q = 82\%$.

В то же время добавление синонимов-словосочетаний в случае, если поисковая система не имеет словаря словосочетаний, приводит к выдаче как дополнительных релевантных, так и нерелевантных документов, поскольку словосочетание интерпретируется как набор отдельных слов. В нашем эксперименте добавление синонимов-словосочетаний повысило коэффициенты полноты и точности до значений $R = 0,86$ и $P = 0,81$, соответственно. При этом качество системы было оценено как "хорошее" ($Q = 92\%$). Итоговая оценка качества ИПС КАСКАД отмечена точкой на рисунке 13.

Таким образом, оценка качества ИПС определяется не только техническими характеристиками самой системы (наличием словаря синонимов, словаря словосочетаний, мощностью базы данных), но во многом зависит и от формулировки запроса, и может быть улучшена путем модификации запроса. Эксперимент показал, что предлагаемая нечеткая экспертная система может с успехом быть использована для формальной оценки качества поисковых систем.

В **Заключении** подводятся итоги исследования, определяется значимость полученных результатов и перспективы развития предложенного направления оценки интеллектуальных систем.

Основной теоретический вывод выполненной работы в том, что впервые предложено использовать теорию нечетких множеств для создания нового подхода к проведению оценки лингвистических интеллектуальных систем на базе нечеткой логики. Подход претендует на универсальность, т.к. с его помощью можно оценить эффективность ИС разного типа. Это подтвердил проведенный эксперимент, в результате которого были созданы действующие модели систем оценки качества СМП, ИПС и систем обработки текстов с помощью ДСА.

Основные положения предлагаемого подхода таковы: в ряде естественных и гуманитарных наук применяются нечеткие экспертные системы, использующие композиционное правило вывода для получения тех или иных заключений на основе определенного набора входных параметров. Принцип работы таких систем состоит в том, что существует определенный набор предпосылок и определенный набор заключений, связанных друг с другом с помощью так называемой базы правил вывода. Нечеткое правило вывода - это нечеткое отношение предпосылки и заключения. При конкретных значениях входных параметров степень влияния каждой предпосылки на каждое заключение различна, т.к. для математического описания предпосылок (и заключений) используются нечеткие множества. Степень истинности предпосылок модифицирует функции принадлежности заключений правил вывода, осуществляя тем самым операцию, называемую в теории нечеткой логики нечеткой импликацией. Существуют различные алгоритмы нечеткой импликации. Все модифицированные функции принадлежности одного и того же заключения объединяются, таким образом осуществляется второй этап алгоритма нечеткого вывода - композиция. Получаемое выходное нечеткое множество приводится к четкости тем или иным способом, что дает конкретное значение выходной переменной.

Для создания нечеткой экспертной системы оценки качества ИС необходимо: 1) формализовать понятие качества системы, 2) определить критерии качества, 3) формализовать эти критерии, 4) разработать базу правил вывода, связывающую предпосылки (формализованные критерии оценки) и заключение (оценку качества ИС). Формализация входных и выходного параметров осуществляется в предлагаемом подходе с помощью введения соответствующих лингвистических переменных, значениями которых служат нечеткие переменные. Их формальное описание осуществляется с помощью нечетких множеств. Критерии качества являются входными переменными экспертной системы. Оценка качества служит выходной

переменной. Ядром системы является база правил вывода. Разработка базы правил вывода - наиболее сложная часть создания экспертной системы, которая является основной задачей эксперта-лингвиста, создающего нечеткую оценивающую систему.

Преимуществами предлагаемого подхода являются:

- 1) универсальность, включающая в себя единое понятие качества для всех ИС, а также независимость от типа конкретной ИС, что позволяет производить сравнительную оценку эффективности однотипных ИС на одинаковой стадии разработки и функционирования;
- 2) открытость, подразумевающая возможность учета любых критериев (факторов), влияющих на оценку качества ИС, не обязательно количественных, а таких, как, например, удобство использования и др., которые могут быть формализованы с помощью введения нечетких множеств. Нечеткая логика позволяет количественно описать такие понятия как плохой, хороший, много, мало и т.д.;
- 3) простота, что выражается в возможности достаточно быстрой реализации предлагаемых экспертных систем в современных математических пакетах, таких как MATLAB, содержащих элементы нечеткой логики;
- 4) наглядность, т.е. формальное описание таких оценок как "качество системы хорошее, если она допускает мало ошибок" и т.д.

В диссертационной работе описана идеология созданных действующих моделей экспертных систем оценки качества интеллектуальных лингвистических систем разного типа, разработаны критерии оценки и базы правил вывода для всех систем. Тестирование предложенных систем оценки на базе нечеткой логики проводилось на трех типах интеллектуальных систем автоматизированной обработки текста (ДСА, СМП, ИПС).

Впервые в интегрированной среде MATLAB Fuzzy Logic Toolbox разработаны действующие модели нечетких экспертных систем оценки качества СМП, ИПС и ДСА на основе композиционного правила вывода. Результаты исследования и оценок автоматических систем способны внести заметный вклад как в прикладное, так и теоретическое языкознание.

Проведенные эксперименты подтвердили эффективность предлагаемого подхода, позволяющего создавать универсальные оценивающие нечеткие экспертные системы для любого типа ИС при любом количестве критериев оценки качества. Тем самым, идея

использовать нечеткие экспертные системы для оценки эффективности ИС открывает новое направление в компьютерной лингвистике: создание универсальных систем оценки качества интеллектуальных лингвистических систем.

Основное содержание диссертации отражено в следующих публикациях:

Монография и учебно-методические пособия:

1. Формальные методы в современной прикладной лингвистике. - "Народный учитель", МГОУ, М., 2002 г. - 16 п.л.
2. Краткий справочник по курсу «Новые информационные технологии» для студентов-филологов. Часть I, II, III Внешсигма, М., 1998-1999. - 6 п.л.
3. Словарь Computing and multimedia. Внешсигма, М., 1996.- 10 п.л.

Статьи и материалы конференций:

1. Лингвистические проблемы применения дистрибутивной и статистической методик в автоматизированном варианте для анализа научно-технических текстов. //Материалы Третьей конференции по Машинному фонду русского языка. М., ИРЯ АН СССР, 1989 - 0,1 п.л.
2. Сопоставительный анализ автоматизированных дистрибутивной и статистической методик // Материалы Международного семинара «ЭВМ и перевод». Тбилиси, 1989 - 0,1 п.л.
3. Система автоматической классификации текстовой информации. //Техника средств связи. М., № 3, 1989. - 0,5 п.л. (в соавт.)
4. Особенности применения автоматизированного ДСМ для анализа НТТ по системам связи и управления. МГУ им. Ломоносова, М., 1989 //Деп. ИНИОН АН СССР № 37622 - 0,5 п.л.
5. Лингвистические проблемы применения автоматизированной ДС методики для анализа НТТ по системам связи и управления. //Автореферат кандидатской диссертации, М., 1989. - 1 п.л.
6. Проблемы курса «Компьютерная лингвистика» // Материалы Второй Республиканской конференции "Компьютерные программы в обучении белорусскому и иностранным языкам". Минск, 1995 г. Минск: МГПУ ин. яз., 1995. - 0,1 п.л.
7. Проблемы снятия полисемии и омонимии терминов и ядерной лексики в системах машинного перевода. //Материалы научно-методического симпозиума «Язык в действии», МАПРЯЛ, ТГУ, 1997. - 0,1 п.л.
8. Возможности использования систем машинного перевода в обучении русскому языку как иностранному. //Материалы

межвузовского научно-практического семинара "Теория и практика преподавания русского языка иностранным учащимся в вузе", М., 1997. - 0,1

9. Преподавание русского языка как иностранного и автоматизированные обучающие системы. //Вестник Центра международного образования МГУ им. М.В. Ломоносова, М., 1998 г. - 0,5 п.л. (в соавт.)

10. Использование систем машинного перевода при разработке автоматизированных обучающих систем. // Компьютерная лингвистика и обучение языкам. Сборник научных статей. Минский ГЛУ, 2000. 0,5 п.л.

11. Использование электронных словарей при обучении русскому языку как иностранному. //Материалы ежегодного межвузовского научно-методического семинара «Теория и практика преподавания русского языка иностранным учащимся в вузе». Май 1999г., МГИМО. - 0,2 п.л. (в соавт.)

12. Некоторые формальные методы в лингвистике. // Филологические науки. Сборник научных трудов № 3 (18). МГИМО, 2000. - 0,4 п.л.

13. Новые перспективы автоматизации обучения языкам //Филологические науки. Сборник научных трудов № 5 (19).МГИМО, 2000. 0,4 п.л.

14. Формальные методы в современной лингвистике. // Материалы научно-практической конференции МПУ, апрель 2001. - 0,2 п.л.

15. Некоторые вопросы использования формальных методов в преподавании языков. // Материалы научно-практической конференции МГИМО, 2002 г. - 0,1 п.л.

16. Методы оценки эффективности интеллектуальных систем. //Филологические науки. Сборник научных трудов № 13 (28). МГИМО, 2003. -0,8 п.л.

17. Нечеткая логика в лингвистике. //Филологические науки. Сборник научных трудов № 12 (27). МГИМО, 2003. -0,6 п.л.

18. Электронные словари и машинный перевод. //Филологические науки. Сборник научных трудов № 14 (29). МГИМО, 2003. -0,8 п.л.

19. Экспертные оценки в гуманитарных науках. //Сборник научных трудов N1, МГОУ, 2003. -0,6 п.л.

20. Методы оценки систем машинного перевода. //Филологические науки. Сборник научных трудов № 15 (30). МГИМО, 2003. -0,6 п.л.

21. Типологии ошибок при машинном переводе. //Филологические науки. Сборник научных трудов № 16 (31). МГИМО, 2003. - 0,6 п.л.



Подписано в печать 10.07. 2003 г. Бумага офсетная. Гарнитура «Times New Roman».

Печать офсетная. Формат бумаги 60/84 ^{1/16} Усл. п.л. 2,6.

Тираж 100 экз. Заказ № 230.

Отпечатано в Издательстве МГОУ с готового оригинал-макета.

105005. г. Москва, ул. Радио, д. 10-а, тел.: 265-41-63, факс:265-41-62.

РНБ Русский фонд

2006-4

9909

25 сен 2003