

РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
ИНСТИТУТ АРХЕОЛОГИИ И ЭТНОГРАФИИ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ГУМАНИТАРНЫХ ИССЛЕДОВАНИЯХ

Выпуск 4

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКИЕ МЕТОДЫ В АРХЕОЛОГИИ

Материалы научного отчета
по интеграционной программе СО РАН
за 2000-2002 гг. (проект № 82)

Ответственный редактор
Член-корр. РАН, доктор исторических наук Ю. П. Холюшкин

Новосибирск
2002

ББК 60
И 74

Издание осуществлено при финансовой поддержке интеграционной программы № 82
Сибирского Отделения РАН

И 74 Информационные технологии в гуманитарных исследованиях:.

Выпуск 4. Информационные технологии и математические методы в археологии. Материалы научного отчета по интеграционной программе СО РАН за 2000-2002 гг.(проект № 82). Новосибирск: Новосибирский госуниверситет, 2002. 95 с.

Настоящий выпуск представляет собой коллективную монографию с материалами научного отчета по итогам трехлетнего совместного исследования Института археологии и этнографии СО РАН, Института систем информатики СО РАН, Института искусственного интеллекта Минсвязи РФ и кафедры систем информатики НГУ. В монографии излагаются подходы к подготовке, созданию, обработке и представлению информации в археологии и этнографии. Выпуск рассчитан на археологов, историков, этнографов и на широкий круг исследователей, интересующихся информационными технологиями в гуманитарных исследованиях и образовании.

ББК 60

© Институт археологии и этнографии СО РАН, 2002

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1. Создание информационных ресурсов по истории, археологии и этнографии Северной Азии	5
2. Онтологическая методология создания информационных ресурсов в археологии.....	18
3. Разработка новых математических и статистических методов и технологий обработки археологических данных для палеолитических объектов	30
Публикации 2000-2002 гг. по теме проекта.....	65

ВВЕДЕНИЕ

В настоящей монографии представлены материалы научного отчета о выполнении интеграционной междисциплинарной программы СО РАН "Новые информационные технологии и математические методы в археологии, культурной и социальной антропологии" в 2000-2002 гг. В материалы включены исследования и разработки, представленные разделом II научного отчета "Создание центра по разработке информационных ресурсов по истории, археологии и этнографии, предназначенные для накопления, систематизации, сохранения и организации широкого доступа к информации о культуре, традициях и искусстве Северной Азии". Исследования и разработки выполнены совместно коллективами научных сотрудников Института археологии и этнографии СО РАН, Института систем информатики СО РАН, Института искусственного интеллекта Минсвязи РФ и кафедры систем информатики НГУ.

Соруководители проекта по разделу II:

д.и.н. Холюшкин Ю.П.

д.ф.-м.н. Марчук А.Г.

к.т.н. Загоруйко Ю.А.

Материалы отчета написаны и подготовлены к публикации авторским коллективом в составе:

Институт археологии и этнографии СО РАН:

Холюшкин Ю.П.

Воронин В.Т.

Ростовцев П.С.

Костин В.С.

Федоров С.А.

Бердников Е.В.

Дунаев Д.С.

Жилицкая Г.Ю.

Илларионов В.А.

Институт систем информатики СО РАН:

Марчук А.Г.

Загоруйко Ю.А.

Костов Ю.В.

Ибрагимов И.И.

Институт искусственного интеллекта Минсвязи РФ:

Нариньяни А.С.

Боровикова О.И.

Загоруйко Г.Б.

I. СОЗДАНИЕ ИНФОРМАЦИОННЫХ РЕСУРСОВ ПО ИСТОРИИ, АРХЕОЛОГИИ И ЭТНОГРАФИИ СЕВЕРНОЙ АЗИИ

В ходе реализации проекта осуществлена комплексная программа исследований по созданию и развитию проблемно-ориентированной среды по гуманитарным наукам и разработке на этой основе гуманитарных информационных ресурсов.

Исходя из этого принято и реализовано решение о создании информационной сети, включающей информационный узел и специализированные сайты. Информационный узел содержит ссылки на сайты сервера САТИ с краткой информацией об их содержании и назначении. В число специализированных, решающих тематические задачи, входят четыре сайта («Информационный центр института археологии и этнографии СО РАН», Виртуальный музей «Древности Сибири», «Сибирика», «Геоинформационные системы» и информационный портал «Гуманитарная паутина»).

Сайт *"Информационный центр института археологии и этнографии СО РАН"* (URL: <http://www.sati.archaeology.nsc.ru:8101/infoc/>) представлен данными о деятельности Института археологии и этнографии СО РАН, его подразделений, базе данных сотрудников, рядом электронных публикаций сотрудников института.

В создании этой части сайта принимали участие Ю.П.Холушкин (концепция и руководство), С.В.Елагина (web-мастер, база данных по сотрудникам, наполнение сайта), Ю.Д.Луговская, Я.В.Балаева (разработка web-страниц и наполнение). Сайт создан в 2001 году. Сотрудниками ИСИ СО РАН разработали и реализовали на этом сайте информационно-справочную систему, обеспечивающую доступ через Интернет к информационным ресурсам Института археологии и этнографии СО РАН.

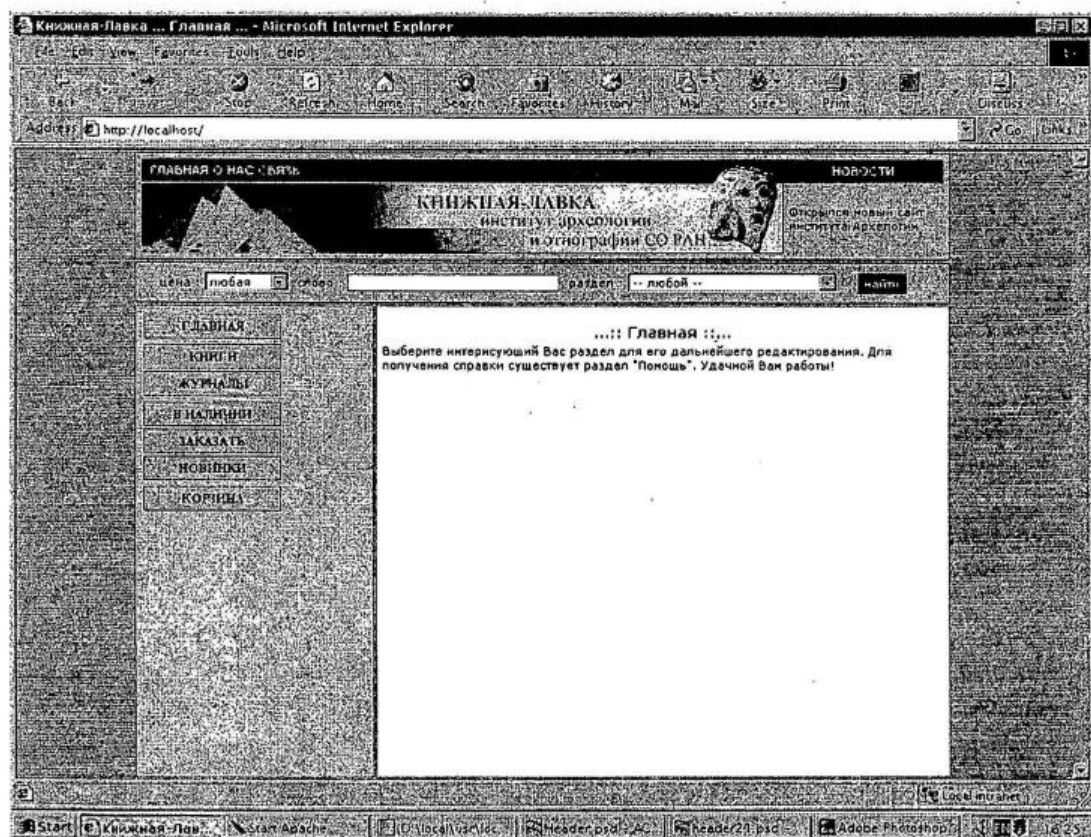


Рис 1. Каталог научных изданий Института археологии и этнографии СО РАН.

Эта система, реализованная в технологии клиент-сервер, позволяет любому пользователю сети Интернет ознакомиться с издаваемыми в ИАЭТ книгами и журналами (рис. 1), а также заказать и купить интересующее его издание (книгу или копию статьи журнала). Для данной информационно-справочной системы кроме

клиентской части было также разработано приложение (администраторская часть), с помощью которого можно вводить в базу данных системы информацию (в том числе, графическую) о новых книгах и журналах.

Для расширения доступа к накопленным информационным ресурсам, сосредоточенным в фондах научной библиотеки, в Информационном центре ИАЭТ СО РАН, организованном в секторе археологической теории и информатики (САТИ), был выполнен комплекс проектов и программ по созданию электронной библиотеки, включающей не только электронные библиографические базы данных научной библиотеки ИАЭТ СО РАН, но и библиотечные ресурсы, разработанные сотрудниками САТИ: Web-энциклопедии и информационно-справочные системы пилотного типа, электронные издания. Эти электронные материалы составляют основное содержание наполнения электронной библиотеки ИАЭТ СО РАН и служат предметом каталогизации ее ресурсов в форме электронного каталога.

Первоначальная тематика охватывает область археологии и этнографии и предусматривает создание и пополнение электронного каталога по археологии и этнографии Сибири за период с 1990 по 2000 год. Эта разработка получила также финансовую поддержку РГНФ (проект № 01-01-12013в), рассчитанную на трехлетний период (2001-2003 гг.).

Электронный ресурс содержит три части:

1) каталог изданий и рукописей, хранящихся в научной библиотеке Института археологии и этнографии СО РАН;

2) каталог электронных публикаций, размещенных на сайте информационного центра института археологии и этнографии СО РАН и других web-серверах в России и за рубежом.

Ресурс включает в себя предметный, систематический и именной каталоги. Доступ к ресурсам каталога имеет два режима:

свободный по сети Интернет – для пользователей (научных сотрудников, преподавателей, студентов и учащихся) – к базам данных электронного каталога,

авторизованный – для программистов (разработка и встраивание прикладных программ для облегчения наполнения и редактирования базы данных) и библиотечных работников (наполнение и редактирование данных каталога).

Создание электронного каталога библиотеки Института археологии и этнографии СО РАН с обеспечением доступа к его данным в сети Интернет предназначено не только для предоставления доступа к имеющимся фондам библиотеки Института, но и для накопления, систематизации, сохранения новых поступлений в эти фонды в будущем.

Источниками информации являются фонды научной библиотеки ИАЭТ СО РАН, библиографические справочники и реферативные издания. На отчетный период объем электронного каталога составляет около 6000 изданий (книг, журналов). В дальнейшем ожидается его пополнение в пределах 1000-2000 единиц ежегодно.

Наполнение и доступ к информации электронного каталога осуществляется под управлением PostgreSQL в операционной системе Linux. Обслуживание пользовательских запросов и генерация HTML-страниц производится с помощью программ на языке PHP/PI и JavaScript.

В 2001-2002 годах осуществлялась отработка технологии хранения данных и структур в базе данных электронного каталога; доработка существующих и создание новых прикладных программ, реализующих функции каталогизации, а также программ, автоматизирующих наполнение БД; доработка HTML-форм для взаимодействия пользователя с прикладными программами; разработка стандартных запросов к базе данных. Итогом этих усилий разработчиков и исследователей стал опытный образец (релиз) электронного каталога, обладающий основными функциональными возможностями (просмотр содержимого, навигация и поиск), сбор и подготовка данных, проведение опытной эксплуатации релиза системы (рис. 2).

В дальнейшем предстоит пополнение базы данных, завершение наполнения электронного каталога и реализация всех запланированных функций системы

(многоаспектный поиск, удобный пользовательский интерфейс для пользователей и администраторов системы электронных каталогов), оптимизация программного кода и структур данных электронного каталога, исследование вариантов использования возможностей электронного каталога, как системы для решения задач различного класса в археологии и этнографии.

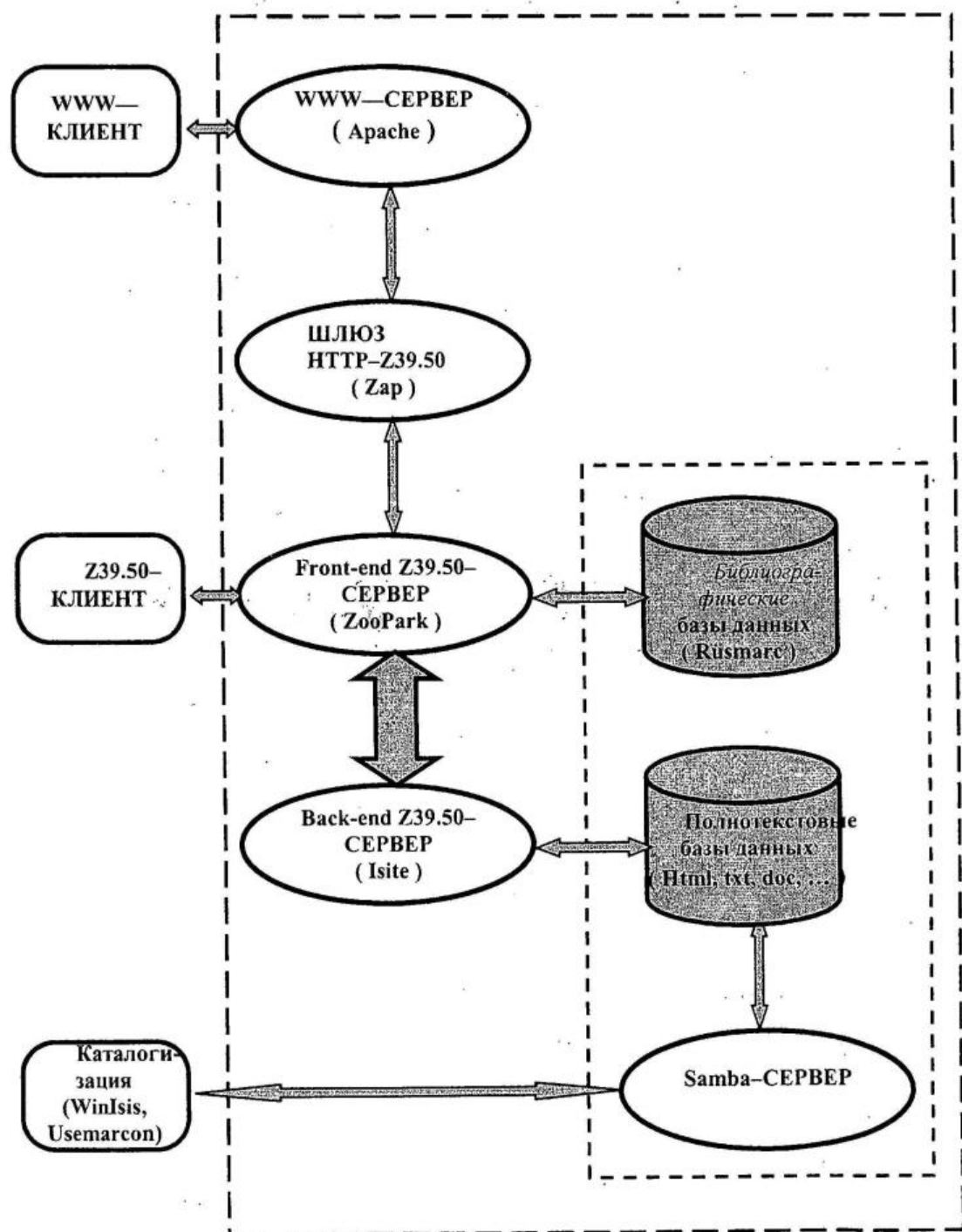


Рис. 2. Структура программного комплекса.

Современное состояние в области электронных каталогов характеризуется стремительным ростом числа электронных библиотек в России и за рубежом. Однако электронные каталоги гуманитарного направления по предметным областям

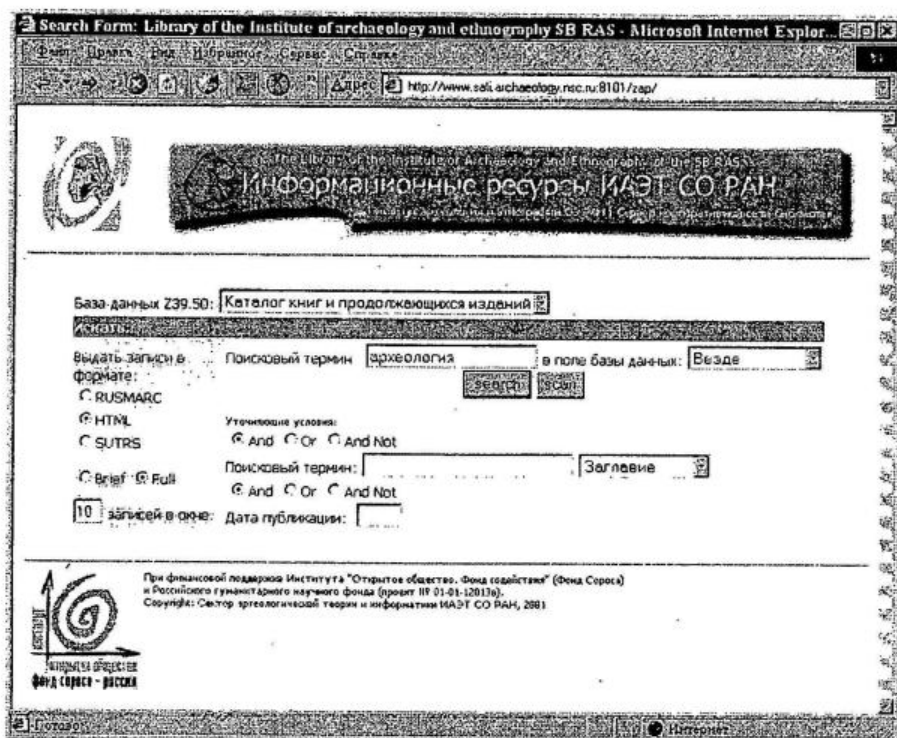


Рис. 3. Web-страница WWW-Z39.50 шлюза к базам данных Z-сервера сектора археологической теории и информатики ИАЭТ СО РАН.

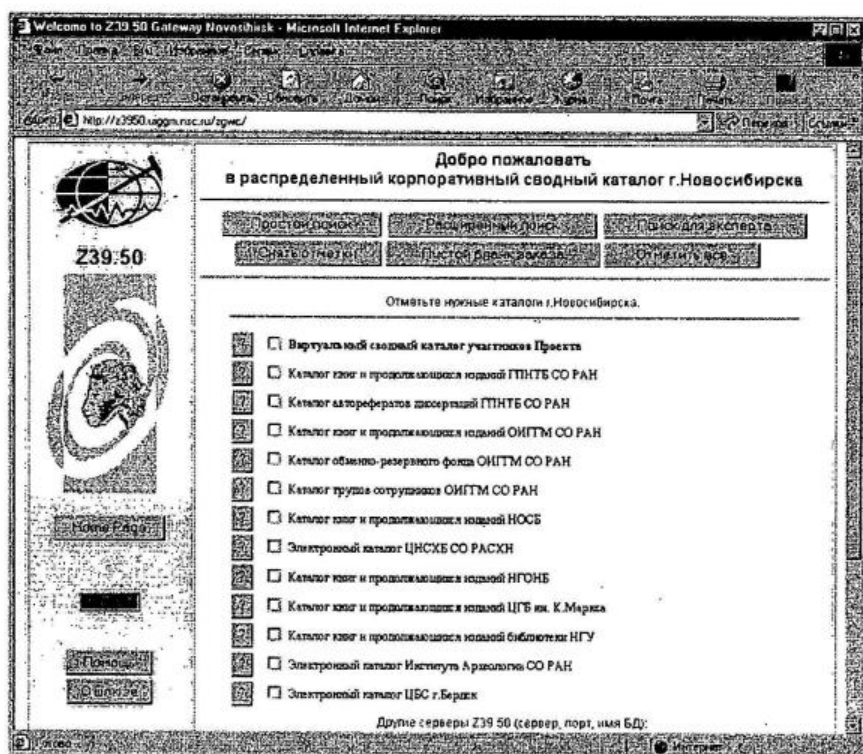


Рис. 4. Web-страница сводного каталога Новосибирской региональной распределенной корпоративной библиотечной системы.

представлены мало (от 300 до 700 наименований изданий в каждом). Библиографические материалы по авторам и названиям статей в периодических изданиях и сборниках научных трудов в создаваемых электронных библиотеках не

имеются. Описываемые в данной статье исследования и разработки ориентированы преимущественно на создание электронного каталога по материалам периодических изданий и сборников научных трудов. В связи с этим особое внимание в проекте уделяется включению в электронные каталоги материалов электронных изданий.

Электронный каталог размещен на одном из сайтов Информационного центра САТИ (URL: <http://www.sati.archaeology.nsc.ru:8101/zap/>, см. рис. 3). В соответствии с целями и задачами проекта создания библиотечной системы по гуманитарным наукам в Сибири ресурсы электронного каталога через Web-шлюз и Z39.50-шлюз включены в сводный каталог региональной распределенной корпоративной библиотечной системы Новосибирской области (URL: <http://z3950.uiggm.nsc.ru/zgwc/>, см. рис. 4).

Разрабатываемая библиографическая система представляет собой программно-информационный комплекс, обеспечивающий доступ пользователю сети Internet к информационным ресурсам научной библиотеки и сектора археологической теории и информатики (САТИ) Института археологии и этнографии (ИАЭТ) СО РАН.

В проекте были использованы идеи и решения по созданию открытых программных систем, ориентированных на поддержку протоколов Z39.50 и Web:

серверная и клиентская платформа баз данных ZooPARK (разработка ОИГГМ СО РАН),

серверная платформа Isite (разработка CNIDR – the Clearinghouse for Networked Information Discovery and Retrieval);

информационная система CDS/ISIS for Windows 1.4 (разработка ЮНЕСКО),

результаты анализа проектов информационных систем, используемых в зарубежных и отечественных библиотеках (ГПНТБ СО РАН, ОИГГМ СО РАН, ОРАС-97, библиотека Конгресса США, Bibliotheks-Verbund Bayern (BVB), ГПНТБ России, БЕН, библиотека Агропрома и др.).

Целью проекта создания библиотечной системы по гуманитарным наукам в Сибири является включение информационных ресурсов научной библиотеки Института археологии и этнографии (ИАЭТ) СО РАН в объединенные информационные ресурсы Новосибирской региональной распределенной корпоративной библиотечной системы и обеспечение свободного доступа к ним библиотек и открытых информационных центров региона и РФ.

Все работы по проекту являются естественным продолжением проводимых в Институте археологии и этнографии СО РАН работ по созданию электронной библиотеки, включающей электронные библиографические базы данных, Web-энциклопедии и информационно-справочные системы пилотного типа, и отработке технологии удаленного открытого доступа к этим информационным ресурсам. Проект является переходом от создания информационных ресурсов к организации широкого и свободного доступа к ним на региональном, российском и международном уровне.

В качестве базы для программного-технического обеспечения проекта на протяжении 1998-2000 гг. в ИАЭТ СО РАН создан, наращивается и активно используется программно-аппаратный комплекс, обеспечивающий разработку и размещение на Web-сервере САТИ (<http://www.sati.archaeology.nsc.ru>) электронных информационных ресурсов, наполнением которых являются полнотекстовые электронные публикации, электронные коллекции, электронные каталоги, перечисленные выше. Для доступа к данным применяется развитый сервис (каталоги, средства навигации, указатели, комментарии, авторизованный доступ для пополнения и редактирования ресурсов, открытый доступ для их просмотра).

Проект создания электронного каталога является новым этапом проводимых работ и его естественным продолжением, переходом от создания информационных ресурсов к организации широкого и свободного доступа к ним на региональном, российском и международном уровне.

Для этих целей на сервере САТИ дополнительно при любезной помощи и поддержке ОИГГМ СО РАН установлены программно-технические средства поддержки протоколов Web и z39.50 (разработчик О.Л.Жижимов) и развернута информационная система Isite (CNIDR). Интеграция этих программно-технических

средств (см. рис. 1) позволила организовать в пилотном режиме открытый и свободный доступ к базам данных САТИ ИАЭТ и научной библиотеки СО РАН на основе международных стандартов. Таким образом, выполнение проекта обеспечило ввод информационных ресурсов электронной библиотеки в более широкий научный и культурно-образовательный оборот.

При выполнении проекта в прешествующем, 2001 году исполнителями проекта предварительно была осуществлена адаптация электронных информационных ресурсов научной библиотеки и САТИ ИАЭТ СО РАН для их представления в Корпоративной информационной системе в стандартных обменных форматах RUSMARC и HTML по стандартным протоколам взаимодействия (Z39.50, HTTP). Для этих целей были разработаны абстрактные схемы и модели данных для баз данных электронной библиотеки САТИ ИАЭТ СО РАН в соответствии с протоколом Z39.50, а также произведено отображение информационных ресурсов электронной библиотеки на абстрактные схемы данных, приобретено, установлено, протестировано и введено в действие необходимое для проекта оборудование и осуществлена адаптация программного обеспечения доступа к базам данных в соответствии со спецификациями протокола Z39.50.

Одновременно с этим выполнена ретроконверсия электронного каталога научной библиотеки ИАЭТ СО РАН за 1998-2000 г., в рамках которой осуществлен перевод в новую структуру ретроспективной информации из электронного каталога за 1998-2000 годы.

Наиболее важной работой является организация публичного доступа к информационным ресурсам научной библиотеки и САТИ ИАЭТ СО РАН из Интернет по стандартным протоколам Z39.50, HTTP, включающей:

- настройку сервера Zoopark,

- адаптацию программного обеспечения информационно-поисковой системы Isite;

- развертывание программного обеспечения WWW-Z39-50 шлюза ZAP на web-сервере САТИ;

- представление информационных ресурсов библиотечными электронными базами данных (каталог книг и продолжающихся изданий, каталог электронных публикаций научных сотрудников Сибирского отделения РАН по гуманитарной тематике, сосредоточенных на сервере САТИ ИАЭТ СО РАН).

Существенное значение в программе проекта имеет участие в программе обмена библиографическими записями в рамках Корпоративной системы Новосибирского региона, в программе по организации электронной доставки документов и в разработке нормативных и методических документов для работы в Корпоративной системе с каталогами и документами виртуальных коллекций и экспозиций.

На сайте *"Виртуальный музей"* (URL: http://sati.archaeology.nsc.ru/gen-ivirtual_e.htm) размещены коллекции самого крупного к востоку от Урала виртуального музея по древней истории, древней культуре и древнему искусству Северной Азии. Такого масштаба Археолого-этнографический виртуальный музей предлагается впервые в СНГ. Его экспозиции доступны учебным заведениям, научным учреждениям, общественным организациям и частным лицам, интересующихся культурой, историей и древним искусством северного региона.

Сайт представлен в Интернет в двух версиях: на русском и частично на английском языках.

Целью проекта создания и поддержки виртуального музея является не только сохранение в электронном виде культурного наследия, но и поддержка развития мировой культуры за счет включения музея в международные информационные сети.

На сайте размещено 12 экспозиций и галерей, в том числе восемь – на русском языке и четыре – на английском, из них три – на русском и английском, пять – на русском, одна – на английском языках.

Базовым материалом для создания галерей послужили коллекции артефактов и их иллюстрации с подрисовочными подписями, а также набор текстов комментария. Авторами этих материалов являются известные исследователи Сибири и Дальнего

Востока – археологи, этнографы и антропологи – научные сотрудники Института археологии и этнографии СО РАН.

Содержимое экспозиций виртуального музея представлено в виде разделов и подразделов. Для доступа к материалам создана "древовидная" система просмотра. Научная функция музея состоит в обеспечении пользователей научной продукцией через создание, поддержку и публикацию монографий. Такая информация может быть полезна для исследователей различных специальностей (историков, археологов, этнографов, краеведов), а также преподавателей, искусствоведов и т.д. Образовательная функция ресурса состоит в обеспечении пользователей данными по вопросам религиозных представлений коренных народов Сибири, знакомством с материальной культурой и древним искусством.

Этим способом материалы могут привлекаться для образовательных целей в контексте ряда гуманитарных дисциплин, например, в учебном процессе на гуманитарных факультетах Российских и зарубежных университетов и в средней школе при изучении археологии, краеведения, истории, этнографии, искусствоведения. Потенциальные пользователи такой информации школьники, студенты и преподаватели ряда специальностей. Культурно-просветительская функция системы состоит в обеспечении пользователей общекультурной и научно-популярной информацией. Общекультурное значение данных материалов несомненно. Учитывая тот факт, что речь идет о множестве разнообразных в этническом и культурном отношении народах, обитавших в течении многих тысячелетий на громадных пространствах между древними цивилизациями Европы и Азии, представленный материал выходит за рамки узко-образовательного или регионально-значимого контекста и приобретает яркое общекультурное и общеисторическое значение.

В галерее *"Древнее искусство Приамурья"* представлены (на русском и английском языках) уникальные археологические и этнографические находки Приамурья и Приморья, которые были сделаны сибирскими археологами, возглавляемыми академиком А.П.Окладниковым в 60-70-х годах. Эта галерея включает богатую коллекцию археологических материалов 4-3 тысячелетий до н.э., а так же коллекции этнографических материалов из музеев Новосибирска, Санкт-Петербурга и Хабаровска, которые отражают близкие связи с традициями более древнего неолитического искусства. Вместе они подтверждают идею, что Амур – колыбель древней цивилизации, и что современная культура народов Дальнего Востока сложилась на основе древних традиционных обществ.

Экспозиция *"Литые кресты и иконы"* демонстрирует (на русском языке) коллекцию предметов старообрядческого культа в частных собраниях Новосибирской области (меднолитые иконы и кресты). Как памятники, осевшие в народных, частных музеях, крестьянских домах, домах-моленных, экспонаты экспозиции имеют огромное историческое и художественное значение. До сих пор этот колоссальный никем не востребованный этнографический материал отражает традиционное религиозное сознание значительной части русских старожилов Западной Сибири – старообрядцев.

Галерея *"Жертвенные покрывала обских угров"* (на русском языке) впервые знакомит с неопубликованными материалами и комплексом вопросов, связанных с феноменом жертвенных покрывал в религиозно-обрядовой практике северных групп манси и хантов. Вводит в курс проблем генезиса и эволюции жертвенных покрывтий, семантики имеющихся на них изображений, использования данного типа культовой атрибутики во время религиозных церемоний.

Виртуальная галерея *"Искусство хантов Нижней Оби"* (на русском и английском языках) представляет собой электронный альбом, на материалах которого описывается длительный процесс этногенеза и этнической истории хантов. Их широкое расселение при изолированности отдельных групп, контакты с соседями определили сложный характер культуры народа. Как этническая общность ханты сложились в 1 тыс. до н. э. – 1 тыс. н. э. в результате слияния продвинувшихся с юга угорских племен с племенами охотников и рыболовов таежного Зауралья. В литературе неоднократно отмечались

особенности языка и культуры разных групп хантов. Исследователи выделяют три большие группы: северную, южную и восточную.

В альбоме представлены одежда, обувь, бытовые вещи и предметы культа части северных хантов, проживающих по рекам Сыня, Куноват и на участках, прилегающих в этом регионе к Оби (Шурышкарский район Ямала-Ненецкого автономного округа).

Исключительно интересный историко-этнографический материал представлен в галерее "*Угорское наследие*" (на русском языке). Ученым и путешественникам, писавшим в конце XIX – начале XX вв. о Западной Сибири, вся ее территория, особенно таежная зона, виделась огромным болотисто-лесным захолустьем, где редкое население влачило унылую, неизменную во времени жизнь. Выработанный стереотип оказался живучим. В 70-х годах нашего столетия про газовиков и нефтяников, осваивающих север Тюменской области, говорили, что они "разбудили" его. Представление это автоматически переносилось и на древность: многие до сих пор убеждены в том, что таежное население Западной Сибири всегда значительно отставало от своих более развитых соседей на западе и юге. Сознывая всю ошибочность подобного убеждения, проистекающего от незнания, авторы собраний материалов в галерее попытались показать, чем многим мы обязаны наследию культуры древних обских угров, и представить эти собрания в виде некоего "концентрата" вещей и идей, отказавшись от обычного ретроспективного подбора по эпохам и сведя весь материал в разделы по следующим темам:

1. Прикладное искусство (изделия из глины, бересты, дерева, тканей, бронзовая художественная пластика).
2. Древние производства (обработка камня, металлургия и металлообработка).
3. Вооружение и военное дело.
4. Культурные связи.

Большая часть представленных экспонатов получена в результате раскопок, произведенных Уральской археологической экспедицией на территории Сургутского Приобья в урочищах Барсова Гора и Сайгатино.

Галерея "*Селькупская мифология*" (на русском языке) представляет собой электронную версию книги – результат многолетних экспедиционных работ под руководством выдающегося ученого-этнографа Галины Ивановны Пелих. Экспедиции в отдаленные таежные уголки Сибири позволили накопить и сохранить для потомков огромный материал по духовной и материальной культуре коренных малочисленных народов Сибири, в том числе уникального народа, живущего в трудных условиях северных просторов. Значение этого материала трудно переоценить. Она полезна ученым гуманитарных направлений, практикам-управленцам и политикам. В силу специфики работы на данной территории они должны знать и использовать традиции материальной и духовной культуры народов, ее населяющих. Это важно и для решения национального вопроса. В экспозиции представлены календарные системы, словарь шаманского языка, словник, галерея и записи песен селькупов.

Галерея "*Народы Саяно-Алтая*" (на русском и английском языках) знакомит с бережно хранимым народами Южной Сибири наследием предков – знаниями и навыками, связанными с традиционными отраслями хозяйства.

Экспозиция галереи "*Святыни манси верховьев Северной Сосьвы*" (на русском языке) посвящены вопросам религиозно-обрядовой практики народа манси. Основное место уделено ранее не описанным в научной литературе домашним, поселковым и региональным святыням. Показаны механизмы формирования и прослежена эволюция святынь с учетом влияния на этот процесс русской колонизации и православия, внутренних миграций, межэтнических взаимодействий.

Галерея-выставка "*Секреты Сибири*" (на английском языке). Представлены материалы буклета выставки экспозиций музейных материалов Института археологии и этнографии СО РАН в Японии и Австралии.

Информационные ресурсы, размещенные на страницах созданного в 2002 году сайте "*Сибирика*" (URL: <http://www.sati.archaeology.nsc.ru/sibirica/news/index.html>),

представлены разделами новостей, энциклопедиями, публикациями, и ссылками на другие создаваемые участниками проекта ресурсы.

Новости на сервере отражают текущие события научной и культурно-образовательной жизни внутри института (новости сайта,) и за его пределами (новости науки, культуры, образования, конференции, научные юбилеи, премии и награды, памятные даты, научные программы и фонды и пр.) (около 3000 документов). В основу функционирования раздела положена концепция регулярной, практически ежедневной обновляемости ресурса.

Другой раздел сайта представлен энциклопедиями. Структура их имеет иерархическую модель с возможностью создания вложенности любого порядка и не обязательно древовидного вида. Система редактирования в указанных энциклопедиях основана на принципах коллективного заполнения материалов, что позволяет организовать участие в наполнении системы информацией ряда лиц и коллективов, разбросанных территориально.

Основные преимущества созданной системы заключается в следующем:

1. Работа через Internet, а следовательно из любого места где есть доступ к WWW.
2. Полная интерактивность редактирования базы – обновление реально используемых данных.

3. Простой, доступный для обычного пользователя, интерфейс.

При создании системы были задействованы следующие средства:

- JavaScript,
- формы html,
- CGI.

Для работы используется браузер Internet Explorer 4.0.

В целом редактирование разбито на отдельные составляющие:

- редактирование разделов (названия и структура), позволяет добавлять, удалять и редактировать разделы и подразделы любой степени вложенности.
- включение терминов в раздел, позволяет добавлять в раздел термины из списка имеющихся.
- удаление из раздела, позволяет удалить термин из раздела.
- добавление и редактирование термина позволяет создавать и переименовывать термин.

– редактирование содержимого термина (основная статья, этимология, словоформа, дополнительные тексты, иллюстрации), позволяет изменить содержимое статьи, добавить, изменить или удалить дополнительные тексты, иллюстрации, изменить содержимое полей.

Менеджер разделов. Предоставляет возможность навигации по дереву разделов в стандартном для Windows режиме спуска/подъема по дереву.

Визуализатор терминов. Отображает выбранный в настоящее время термин.

Для поиска используются два стандартных средства: навигация по разделам и поиск по названию. Навигация по разделам сделана с помощью гипертекстовых ссылок, поиск по названию – через систему запроса.

Технические характеристики:

Серверная платформа:	Redhat Linux 5.2 + Russian Apache
Клиентская платформа:	Web клиент (4-ой версии)
Язык программирования:	PHP3
СУБД:	PstgreSQL.

На сайте в настоящее время размещены следующие энциклопедии:

Археологическая Web-энциклопедия. Содержит 2765 терминов по мировой археологии. В основу энциклопедии был положен археологический словарь Брея и Трампа (1500 терминов), но значительно расширен авторами проекта.

Web энциклопедия «Археология и этнография Приобья». Содержит около 800 терминов по: основным понятиям этнографии; этнографическим направлениям и школам; этнографическим теориям, направлениям и школам; этнографическим субдисциплинам и смежным наукам; социально-экономическим отношениям и

соционормативной культуре; мифологии манси; общие понятия археологии из словаря Л.С.Клейна.

Материалы к словарю для доисториков (М.Брезийона). Русско-французский описательный словарь из 400 терминов по палеолиту, включающий в себя термины по описательной морфологии, технике расщепления камня, типологии и т.д.

Плейстоценовые гоминиды. Иллюстрированный справочник на английском языке по находкам древнейших гоминидов, содержащий 123 находки и 185 иллюстраций.

Титулованные роды Российской империи (около 800 фамилий). Представлена информация об истории возникновения рода, а иногда и полная родословная большинства титулованных родов Российской империи.

Дворянские роды Российской империи (более 700 фамилий)

Бронза Сибири. В этой ИПС были размещены: коллекция бронзовых предметов из Овгортского музея. Подобные находки (бронзовые изображения всадников, бобров, птиц, пряжки с головами медведей) встречались этнографам в домашних и на поселковых святилищах обских угров и ранее (материалы А.Ф. Теплоухова, С.И. Руденко, В.Н. Чернецова, И.Н. Гемуева и А.М. Сагалаева, и др.), общее число их невелико, как и число работ, специально, посвященных данному вопросу [Теплоухов, 1947; Гемуев, Молодин, Сагалаев, 1984; Зыков и др., 1994]. Представленная коллекция замечательна как по числу экспонатов, так и по разнообразию затронутых в них сюжетов. Дополнительную значимость находкам придает тот факт, что они были обнаружены на современных хантыйских святилищах и сопровождалась комментариями местных жителей. Кроме того были введены в научный оборот целый ряд средневековых бронзовых артефактов, обнаруженных в ходе этнографических экспедиций последних лет в Нижнем Приобье (Березовский, Белоярский и Шурышкарский районы Тюменской области).

Серебро Сибири. В этой ИПС большой интерес представляет размещенная на сервере небольшая коллекция серебряных предметов, найденная в святилищах. Следует напомнить, что восточное серебро (блюда, чаши, кувшины, подносы, и др.) активно поступало на Север уже в VII - VIII вв.: среднеазиатские купцы обменивали его на пушнину, моржовый клык и даже ловчих птиц. Ввиду особой ценности и "священности" белого металла, серебряные изделия чаще всего попадали на сибирские языческие святилища, где продолжали свою жизнь в качестве ритуальных атрибутов. С исчезновением или разрушением культовых мест серебро уходило под землю, откуда спустя сотни лет вторично появлялось на свет в составе т.н. "кладов". Тем не менее, уникальные средневековые изделия иной раз сохраняются в домашних и на поселковых святилищах хантов и манси.

Ритуальные колчаны обских угров Другой коллекцией размещенной на сервере представлены ритуальные колчаны. Колчаны на сегодняшний день обнаружены только в пос. Тугияны (р. Обь) Белоярского района Тюменской области. Они однотипны по форме: сшиты из прямоугольных кусков сукна. У некоторых изделий входное отверстие обшито лисьим мехом. Подобный способ оформления известен из сюжетов средневековых миниатюр киданей, которые окаймляли горловину колчана хвостом пушного зверя. На лицевой стороне колчана техникой вшивания кусков сукна двух цветов выполнены семь фигур медведей, ящериц и выдр; шесть фигур размещены парами, седьмая - отдельно, восьмая фигура пришивается с лентами к углам. Рядом с фигурами или на них нашиты монеты, их кладут и внутрь изделия (как при изготовлении колчана, так и в последующие годы его хранения). В комплекте с колчанами в "святых" сундуках хранятся votивные стрелы или их наконечники

Атрибутика медвежьего праздника у обских угров. Представлена в созданной ИПС коллекцией масок и медвежьих шкур. Известно, что специфической особенностью атрибутики медвежьего праздника у сынских хантов являлись маски, выполненные не из бересты, а из дерева. Их практически нет в музеях России; крупнейшая коллекция хранится в музее пос. Овгорт Шурышкарского района Ямало-Ненецкого АО. В экспозиции виртуального музея использованы также материалы полевых исследований А.В. Бауло у сынских хантов в 2000 - 2002 гг. Вторая часть

экспозиции представлены материалами экспедиций А.В. Бауло 1998 – 2002 гг. к березовским, казымским и сынским хантам.

Каменные изваяния Алтая. Большой интерес представляет размещенная на сервере коллекция 116 средневековых изваяний Центральной Азии, в основном, изображающих сидящих мужчин монголоидного облика с широкоскулым лицом, раскосыми глазами, очень часто с усами и бородой. На отдельных изваяниях изображены вдетые в уши серьги, на шее гладкие гривны и ожерелья. На многих изваяниях показана одежда: разнотипные головные уборы, халаты и кафтаны с отворотами на груди, узкими или широкими рукавами с манжетами. На узком поясе, набранном различными по форме бляшками, подвешены кинжал и сабля в ножнах, а также сумочка, точило и другие предметы. Почти все изваяния держат в правой руке сосуд, левая, как правило, изображена лежащей на поясе или оружии.

Искусство чжурчженей. Об искусстве чжурчжэньского населения Дальнего Востока, об уровне его миропонимания и эстетических представлениях можно частично узнать из небольшой коллекции предметов, касающихся как элементов материальной культуры (одежда, украшения); так и духовного мира-верования (шаманизм, буддизм) и искусства.

Культурное наследие народа Саха. В размещенной якутской коллекции насчитывается несколько десятков уникальных предметов, теперь редко встречающихся даже в среде самих якутов.

Энциклопедия мифов обских угров. В энциклопедии размещены, по согласованию с авторами, словарные статьи изданной в 2001 году книги «Мифология хантов».

Ряд созданных информационных систем: *Религии народов мира*; *Народы мира*, *Иллюстрированный атлас первобытного человека* и ряд других доступны лишь для внутреннего пользования. Это связано с проблемой авторских прав.

Большая часть размещенных на страницах секции электронных энциклопедий богато иллюстрирована и содержит более 15000 статей и около 4000 иллюстраций. Кроме того, на сайте размещена энциклопедия наших коллег *Мифология Коми*.

В будущем предполагается разместить на сайте энциклопедии мифов других угрофинских народов. Там же будет размещена разработанная в 2002 информационная система для интернет *Системная классификация археологической науки*.

Электронные издания. Участниками проекта разработана система автоматической верстки электронных изданий, включающая в себя полноценные электронные публикации (журналов, бюллетеней, сборников научных статей). Известно множество проектов создания полнотекстовых баз данных, доступных в среде Интернет, однако говорить о том, что в мире существует какая-либо общепризнанная технология создания электронных публикаций нельзя. Каждый из реализуемых проектов использует, как правило, свою технологию электронных версий журналов и бюллетеней, основываясь лишь на унификации средств доступа к данным. Российские гуманитарии используют не самые технологичные, с точки зрения их простоты, стандартные средства создания HTML- документов. Кроме того технология создания и публикация электронных изданий требует и специальных знаний о web-дизайне. Поэтому программа автоматизации этого процесса преследовала цель предоставить возможность делать это обычным пользователям ПК знакомых лишь с Word из MS Office'95, или поздними версиями.

Созданная система электронной верстки основывается на следующей информационной структуре:

Заголовочная информация. Поскольку в своей основе она стандартна для всех изданий, то было легко выделить основные структурные компоненты для создания формы ее заполнения. Она включала в себя: название раздела, гриф организации, название издания, сведения об авторах, город, год издания, аннотацию и др.

Содержание, представляющее собой информацию о иерархии расположения текстов в издании и которая может быть разбита на отдельные компоненты. Система построена так, что из этих компонент можно собрать содержание любой конфигурации, соответствующее создаваемому изданию.

Тексты отражают конкретное содержание публикаций. Для каждого издания они разбиваются на отдельные блоки – файлы. Файлы могут быть подготовлены в MS Word и предоставлены в публикацию как в HTML-формате, так и в DOC-формате.

В рубрике выставлены полнотекстовые версии целого ряда изданий, которые могут быть распределены на следующие группы:

- монографии (11 изданий);
- периодические издания (5 изданий);
- электронные журналы и бюллетени (14 изданий);
- сборники научных трудов (1 издание);
- авторефераты диссертаций (10 изданий);
- учебно-методические пособия (3 издания).

Именно эта часть является наиболее востребуемой в научно-образовательном плане.

В настоящее время электронная библиотека содержит труды по археологии, этнографии, философии, истории, социальной биологии, этнографии, науковедению, социальной психологии и др. дисциплинам.

В разделе ссылки представлены следующие разработки участников проекта:

Картография и ГИС (http://sati.archaeology.nsc.ru/lnk/gis_links.html).

В состав этих баз входят следующие геоинформационные системы:

ГИС "Палеолит Северной Азии". Завершено заполнение баз данных по палеолиту Забайкалья.

ГИС "Религиозно-мифологические представления народов Западной Сибири" <http://www.sati.archaeology.nsc.ru:8101/ethno/>;

Создан опытный Web-сайт, на котором размещены основные инструменты авторизованного наполнения и редактирования баз данных ГИС, а также средства доступа, поиска и извлечения информации из них. Доступ к данным, поиск и извлечение организован в двух режимах: иерархическом (через кнопки, клавиши меню) и контекстный (через поисковую строку). Кроме клавишей доступа к тематическим разделам, на экране пользователю предлагаются клавиши доступа к традиционным сервисным разделам «Гостевая книга», «Контактная информация», «О проекте», «Ссылки» (на другие сайты со смежной этнографической тематикой).

Базы данных включают следующие тематические разделы:

Материальная культура (подразделы: Олений транспорт, Упряжное собаководство, Лыжи, Лодки, Жилища, Верхняя одежда, Головные уборы);

Духовная культура (подразделы: Шаманские бубны, Атрибутика медвежьего праздника Обских Угров, Культовая атрибутика, Восточное серебро в культовых комплексах Обских Угров, Изделия из свинца в обрядовой практике);

Культура жизнеобеспечения (подразделы: Охота, Рыболовство, Собирательство, Виды и способы приготовления пищи);

Соционормативная культура (подразделы: Типы семей, Отношения собственности, Порядок заключения браков, Обычное право);

Народы (29 ссылок на данные о сибирских этносах и этнических общностях).

К настоящему моменту наполнение разделов "Материальная культура" и "Народы" завершено полностью. Наполнение данными разделов "Духовная культура" выполнено в существенной части. Наполнение разделов "Культура жизнеобеспечения" и "Соционормативная культура" развернуто. Можно уверенно сказать, что накопленные данные в БД ГИС уже в текущем состоянии обладают критической массой для того, чтобы быть востребованными в Интернет.

ГИС "Корпус радиоуглеродных дат археологических памятников северных территорий Китая".

Справочники:

1. «Международный справочник исследователей Арктики по социальным наукам» – 1000 персон.
2. «Археологи и этнографы России» – 450 персон.
3. «Мировая Антропология, этнография и археология в лицах» - около 300 персон.

Электронные издания представлены: электронным вариантом издания научного междисциплинарного альманаха "Немецкий этнос в Сибири", выпуск 2, 2000 и Библиотекой Япониста.

Наиболее значимые информационные ресурсы гуманитарного направления сосредоточены на сайте *Информационного портала "Гуманитарная паутина"*» (URL: http://www.sati.archaeology.nsc.ru:8101/hum_web/) в виде ссылок на более 7370 мировых и российских гуманитарных ресурсов, из них относящихся к сфере науки на более 6000. К каждой ссылке прилагается аннотация их содержания. Все ссылки объединены в древовидный каталог. В корневой части каталога на главной странице портала выделены разделы (см. таблицу 1).

Таблица 1. Информационные ресурсы информационного портала "Гуманитарная паутина."

Разделы	Кол-во сайтов в разделе	Тематическая ориентация сайтов	Информационные ресурсы
Официальная Россия	5	официальные серверы органов государственной власти	информация о деятельности высшей законодательной, исполнительной и судебной власти
Региональная власть	26	официальные сайты субъектов Федерации, межрегиональных объединений, Федеральных округов	информация о деятельности законодательной и исполнительной власти субъектов Федерации
Партийная жизнь	45	сайты политических партий, объединений, движений, политических деятелей	политические программы, уставы, органы управления, фонды, интересы, направления деятельности
Четвертая власть	168	сайты информационных агентств, газет, журналов, телеканалов и др. СМИ	периодические издания, программы телепередач, репортажи, интервью, информационные материалы, новости
Правовое поле	12	правовые документы РФ	Конституция РФ, законопроекты, законы и постановления Государственной Думы, Федерального собрания, указы Президента РФ
Сибирь и Интернет	112	адреса сибирских интернет-провайдеров и наиболее интересных и содержательных сайтов в городах Сибири	информация о провайдерских услугах, сервисе в городах Сибири, о транспортных узлах Сибири и Дальнего Востока, об общественных организациях
Образование	686	сайты высших и средних учебных заведений, органов управления образовательным процессом	информация о проблемах и перспективах образования и подготовки кадров РФ
Культура	179	сайтов музеев, библиотек, художественных выставок, театров, музыки, кино, архитектуры	культура, искусство и литература в Интернет (виртуальные музеи, электронные библиотеки, художественные выставки, театральная жизнь, журналы, музыка, кино, архитектура)
Религия	45	сайты конфессий, религиозных, культовых организаций и обществ	информационные материалы по религиоведению, языческим культам древней Руси, шаманизму, сатанизму, зороастризму
Наука	6029	сайты научной и научно-информационной направленности	информация о научных сообществах, научных фондах, журналах и газетах, разделах гуманитарных наук, а также сайты с развлекательной околонуучной, квазинаучной и прочей информацией
Архивы	3	государственные, научные и региональные архивы	государственная, научная и региональная архивная служба, международный совет архивов
Поиск в Интернет	44	информационно-поисковые сайты	основные справочно-поисковые системы по информационным ресурсам в Интернет.

В основе концепции Информационного портала лежит понимание того, что российская наука, образование и культура испытывают в наши дни потребность в концентрации и обобщении накопленной информации по гуманитарным наукам и эффективного ее использования. Удовлетворение этой потребности затрудняется тем, что в силу многоплановости и многоаспектности информационные ресурсы гуманитарного направления в пределах российского подпространства Интернет рассредоточены на удаленных страницах множества сайтов, основная тематическая направленность которых относится скорее к естественнонаучной, технической и технологической сфере, нежели к сфере гуманитарной.

Кроме того, потребители этой информации в своем большинстве не являются профессионалами в области информационных технологий, особенно в части создания распределенных ресурсов или обеспечении удобного доступа к ним. Именно эти задачи (сведение гуманитарных ресурсов в единое адресное пространство, возможность открытого удобного доступа к ним и поддержка целостности этих ресурсов) легли в основу концепции информационного портала "Гуманитарная паутина".

Разработанная система представляет из себя структурированный каталог гиперссылок, построенный на реализации идеи файлоподобной системы:

- структура каталога подобна файловой системе, где разделы и подразделы соответственно подобны каталогам и подкаталогам файловой структуры, а гиперссылка – файлу;

- у каждой гиперссылки есть признак активности (активна – неактивна), неактивные гиперссылки может видеть только администратор системы;

- для каждой гиперссылки есть название и краткое описание;

- сделана функция контекстного поиска по названию и описанию гиперссылки;

- любой пользователь системы после регистрации может добавить новую гиперссылку в любой из уже существующих каталогов; добавленная гиперссылка получает статус "неактивна" для обеспечения контроля на вводимые данные;

- администратор системы может свободно редактировать структуру каталога (добавлять, редактировать и удалять разделы и подразделы и подразделы), удалять, редактировать и добавлять ссылки, делать их активными и неактивными, видеть информацию об авторе добавленной ссылки.

2. ОНТОЛОГИЧЕСКАЯ МЕТОДОЛОГИЯ СОЗДАНИЯ ИНФОРМАЦИОННЫХ РЕСУРСОВ В АРХЕОЛОГИИ

Работа над порталом «Гуманитарная паутина» показала, что представленные в сети Интернет громадные объемы информации становятся практически недоступными из-за неэффективной работы поисковых систем, не учитывающих семантику запросов. Проблема усугубляется еще и тем, что в качестве основной задачи поисковых систем в настоящее время рассматривается индексация ресурсов сети при полном отсутствии средств анализа смысла представленной в них информации.

Поэтому для решения задачи повышения эффективности поиска в сети Интернет предлагается строить *порталы знаний*, каждый из которых предоставляет доступ к ресурсам сети Интернет определенной тематики. Основу таких порталов знаний составляют онтологии, содержащие описание структуры и типологии соответствующих сетевых ресурсов.

Понятие *онтология*, заимствованное из философии, сейчас активно применяется в информатике и искусственном интеллекте. Напомним, что в философии онтология – это учение о бытии, о сущем, о его формах и фундаментальных принципах, о наиболее общих определениях и категориях бытия. В таком понимании онтология, являясь философской дисциплиной, изучает объекты, свойства которых являются общими для всего сущего.

Когда же природа объектов различна, то в зависимости от контекста и целей использования онтологии, существуют следующие интерпретации этого понятия. Одни исследователи рассматривают онтологию как концептуальную "семантическую" сущность. Тогда онтология является концептуальной системой, которая предлагается в качестве базиса определенной базы знаний. Другие исследователи трактуют онтологию как специальный семантический объект. В данном случае онтология – это представление концептуальной схемы логической теории, ее словарь или спецификация.

Для систем искусственного интеллекта (ИИ), которые, как правило, являются замкнутыми системами, существует только то, что уже в них представлено или может быть представлено, поэтому в области ИИ самым распространенным определением онтологии является определение, данное в работе. Согласно этому определению, онтология является точной спецификацией концептуализации. Точный смысл этой интерпретации зависит от понимания терминов "спецификация" и "концептуализация". Под концептуализацией может пониматься некоторая абстракция, т.е. упрощенное представление мира, построенное для определенной цели. Концептуализация включает объекты, понятия и другие сущности, которые предполагаются существующими в рассматриваемой области, а также отношения между ними. С этой точки зрения каждая база знаний, система ИИ или интеллектуальный агент явно или неявно фиксируются некоторой концептуализацией.

В контексте ИИ основу онтологии составляет множество (словарь) представленных в ней терминов, организованных в таксономию. В такой онтологии определения связывают имена сущностей предметной области, (например, классы, отношения, функции и другие объекты) с текстом на естественном языке, описывающим, что означают эти имена, и формальными аксиомами, ограничивающими интерпретацию и корректное использование терминов. Заметим, что при таком подходе понятие онтологии сильно пересекается с уже принятым в информатике и лингвистике понятием тезауруса.

В настоящее время ведется большое число исследований в области онтологий, в том числе и в работах по интеллектуализации информационного поиска, в первую очередь в среде Интернет. Общей целью таких проектов является разработка новых подходов к построению пространств знаний и средств работы с ними.

Особенностью онтологий порталов знаний, ориентированных на поиск информации в Интернет, является наличие в них описания сетевых ресурсов наряду с традиционным описанием предметной области. Такая онтология на основе предварительного индексирования группирует ссылки на Интернет-ресурсы в категории и соотносит их с понятиями, которые в ней описаны. На основе свойств сетевых ресурсов онтология отслеживает взаимосвязи между информационными источниками для соотнесения их с определенной тематикой, группой, интересами и другими параметрами.

Онтологии порталов знаний могут использоваться не только для локализации поиска в проиндексированных данным порталом ресурсах, но и для уточнения формулировки запроса при выполнении поиска во всей сети.

Заметим, что для решения задачи поиска информации также должны быть созданы специальные средства пополнения онтологии и ее интеграции с другими предметными онтологиями.

Задачи между участниками проекта были распределены следующим образом:

- разработка системного тезауруса: ИАЭт СО РАН (сектор археологической теории и информатики);
- концептуальная модель портала: ИСИ СО РАН, НИИ ИИ Минсвязи РФ

2.1. Системная классификация археологической науки и археологических понятий

Начало современному этапу исследований проблем систематизации в гуманитарных науках было положено возникновением фонологии в лингвистике. Именно фонология обнаружила конкретные фонологические системы в языке, выявила их структуру и проявила стремление к открытию общих законов, либо найденных индуктивным путем, либо выведенных логически. Таким образом в одной из гуманитарных наук удалось выявить отношения, имеющие, по словам крупнейшего представителя фонологии Н. Трубецкого, всеобъемлющий и абсолютный характер. Событие такого масштаба должно было заставить представителей смежных гуманитарных дисциплин проверить вытекающие из этого последствия и возможности распространения на факты иного порядка. При этом конечно же возникла опасность пойти по ложному пути, который заключался в формальном употреблении и механическом переносе терминов из лингвистики (фонем, морфем) в смежные науки. Примером такого переноса терминов из лингвистики в археологию могут служить работы Д.Хаймса и Д.Дитца. Теоретическое обоснование такого переноса Д.Хаймс выводил из общности основных посылок обеих дисциплин, а Д. Дитц считал язык и производственную деятельность человека явлениями одной природы, сводимой к моторной активности человека. Эти методические посылки, направленные на признание различных форм социальной жизни в качестве систем поведения и являющихся некой проекцией на плоскость сознательного и обобществленного мышления, однако требуют проведения глубоких исследований, поддающихся экспериментальной проверке.

Отсюда требование создания некоего всеобщего кода, способного выразить общие свойства, присущие каждой из специфических структур, соответствующих отдельным областям. При достижении указанной цели исследователи, по мнению К. Леви-Строса, окажутся в состоянии выяснить, удалось ли наиболее полно постичь природу этих структур, а также определить, состоят ли они из реалий одного типа. Как считал К. Леви-Строс, многие семиотические проблемы могли быть решены современными вычислительными машинами. "С их помощью можно было бы получить нечто, вроде периодической системы элементов, которой современная химия обязана Менделееву. Тогда нам осталось бы только разместить исследованные языки, непосредственное исследование которых еще недостаточно для того, чтобы познать их теоретически, и даже найти место для языков исчезнувших, будущих и просто предполагаемых".

Таким образом, в указанной цитате, К. Леви-Строс близко подошел к пониманию основных свойств, служащих критериями системных классификаций, разработанных в 1985 году Е.Д. Гражданниковым.

Первый критерий — упорядоченность (ранжированность) всех объектов по определенному критерию (критерий упорядочения). В периодической системе элементов Д. И. Менделеева критерий упорядочения — атомная масса; в системе общественно-экономических формаций — время, соответствующее уровню развития производительных сил; у К. Леви-Строса — некий всеобщий код, способный выразить общие свойства, присущие каждой из специфических структур.

Второй критерий — периодичность классификации, т.е. тот научный закон, который лежит в основе системной классификации. Периодическая система элементов Д. И. Менделеева основана на периодическом законе; марксистская периодизация — на законе смены общественно-экономических формаций; у К. Леви-Строса — проекции универсальных законов, регулирующих бессознательную структуру разума, на уровень сознания и социализованной мысли.

Третий критерий — структурированность (критерий структурированности). В периодической системе — это таблица элементов, расположенных по возрастанию атомного веса; у К. Леви-Строса — структура разума, состоящая из бинарных оппозиций. Выявив эту бессознательную структуру человеческого разума путем раскрытия структуры мифов, он, по его мнению, раскрыл и структуру мира, который "за тысячи, миллионы, миллиарды лет не делал ничего другого, кроме того, что

соответствует обширной мифологической системе". В стремлении описывать структуру этнографических явлений в терминах диалектических противопоставлений можно увидеть некоторое влияние К. Маркса и Гегеля на К. Леви-Строса.

Четвертый критерий — теоретическая обоснованность построений. У К. Леви-Строса она видится в окончательном оформлении структурной антропологии. Заслуга К. Леви-Строса заключается и в том, что он одним из первых пришел к выводу о необходимости применения к анализу структуры мифа многомерных классификаций. При этом он подчеркивал, что "продолжая сравнение вариантов (мифов), мы должны будем пользоваться столь многомерными схемами, что интуитивное представление о них станет невозможным и потребуются новые методические и инструментальные вычислительные средства. А в настоящее время вся путаница и отсутствие содержательных выводов в изучении мифологии происходит оттого, что исследователи не умеют пользоваться многомерными системами отсчета".

И, наконец, следует упомянуть еще одно условие для построения системной классификации — это уровень ее универсальности, т.е. возможности ее применения к достаточно широкому классу понятий. С этим требованием смыкается одно из сформулированных К. Леви-Стросом условий построения модели структуры, а именно, такая модель должна быть построена таким образом, чтобы ее применение охватывало все наблюдаемые явления. Об универсальности бинарных (двоичных) противопоставлений в практике этнографических и культурно-исторических работ свидетельствует их довольно частая встречаемость, например, в упоминаемых Вяч. Вс. Ивановым, описаниях сибирских шаманских бубнов; древнекитайской картине мира и др. Утверждению представлений об универсальности дихотомических классификаций, основанных на четверичном и восьмиричном принципах способствовали также работы Лича, Нидэма, Бейдельмана, Эванса-Притчарда, посвященные исследованиям систем религии и родства, символики латеральности и оппозиции правого и левого.

В ходе этих и других исследований было установлено, что бинарная оппозиция — это способ установления двух символических средств, чьи явные противоположные качества или количества предполагают, в понятиях ассоциативных правил культуры семантическую оппозицию. Так, В. Тэрнером было выделено ряд типов бинарных оппозиций. Среди них для нас представляют интерес следующие:

а) бинарная оппозиция иногда может возникнуть между комплексами символических средств, каждый из которых содержит систему доминантных и второстепенных символов; один из них может быть активным, а другой — пассивным (критерий определяющего влияния и мощности объема понятий);

б) один из членов диады может мыслиться как производный от другого (критерии порядка следования и родовой);

в) они могут быть похожими или непохожими, но равными по ценности (случай нестандартного классификационного фрагмента, когда элементы диады могут не различаться по критериям аналитичности — синтетичности; первичности — вторичности; частности — общности).

Конечно же дихотомические классификации не могли исчерпать всех видов первобытных форм классификаций. Да и сам К. Леви-Строс обнаружил, что у многих племен наблюдается кажущееся противоречие в описании структуры поселения. В случае с племенем бороро в Бразилии, члены одной половины племени описывали ее как радиальную, разделенную пополам между половинами. Члены другой половины описывали ее как концентрическую. По их словам, жилища одной половины вписаны в круг жилищ другой половины. Это дало К. Леви-Стросу основания для построения четкой формальной модели. Оказалось, что одна из половин, в свою очередь, делится на две половины. Поэтому вся система может описываться и как двоичная — радиальная, и как троичная — концентрическая. Два описания не противоречат друг другу, а являются дополнительными и даже переходящими, согласно их толкованию, друг в друга.

Кроме того, В. Тэрнеру в ходе изысканий в области ритуальной символики ндембу также пришлось убедиться в том, что практически любую форму дуализма следует

рассматривать как часть более широкой, трехчленной классификации. Рассматривая цветовую триаду "белое-красное-черное" в качестве архетипа человека в процессе переживания наслаждения и боли, В. Тэрнер приходит к выводу, что "восприятие этих цветов и осознание триадных и диадных отношений в космосе и обществе, непосредственное или метафорическое, является производным этого изначального психофизиологического опыта... Для зачатия требуется двое, и двое участвуют также в акте кормления, в борьбе и убийстве (Каин и Авель), а в формировании семьи уже участвуют трое. Множества накладывающихся друг на друга классификаций, образующих системы идеологии, которые контролируют социальные отношения, суть уже производные от этих изначальных двоих и троих, очищенные от их первоначального эмоционального фона". Таким образом триада, по В.Тэрнеру, "это сокращенное или концентрированное обозначения больших областей психофизиологического опыта, затрагивающих как разум, так и все органы чувств, и связанных с первичными групповыми отношениями. Лишь в результате последующего абстрагирования от этих конфигураций возникают другие виды используемой человеком социальной классификации". Так Тэрнером было установлено, что в некоторых "типах ритуальных комплексов могут применяться и другие типы классификации... Не существует единой иерархии классификаций, которую можно было бы рассматривать как охватывающую все типы ситуаций. Скорее существуют различные уровни классификаций, которые пересекаются друг с другом и в которых составные бинарные пары (или триадные рубрики) вступают лишь во временные связи...". Из этого следует заключение, что указанные простые структуры при всей своей универсальности не исчерпывают всех структур, присущих первобытным формам классификации. Таким образом К. Леви-Строс и В. Тэрнер вплотную подошли к теоретическому решению задачи построения периодической системной классификации понятий. Они смогли уловить ряд системообразующих правил их построения и нарисовать достаточно убедительную картину взаимосвязей и взаимопереходов диадных и триадных структур, служащую действенным методологическим орудием познания. Однако создание всеобщей системы понятий и обеспечение единства разнообразных структур в рамках этой системы оказалось для них непосильной задачей.

В российской науке гипотезу о возможности существования всеобщего периодического закона в природе высказывали в первую очередь философы и социологи. Одним из первых эту гипотезу высказал И.Ф. Зубков (1985). Свой вклад в формирование некоторых структур всеобщих системных классификаций внесло так же открытие новосибирским философом Итэсем триады триад, из существования которой предсказывалось наличие пентадных классификаций. В 1985 году, одновременно с И. Ф. Зубковым, гипотезу о существовании всеобщего фрагментного периодического закона высказал и сформулировал Е. Д. Гражданников. В ходе реализации этой гипотезы были построены периодическая система философских и социологических категорий, системная классификация исторических наук, а в Институте археологии и этнографии СО РАН системная классификация археологии.

В основе метода системного классификационного анализа понятий лежит гипотеза существования всеобщего фрагментного периодического закона, которая была выдвинута Е.Д. Гражданниковым в 1985 г. Формулировка этого закона была следующей: "...понятия, упорядоченные по критерию первичности-вторичности, образуют периодическую систему, роль периода в которой играет классификационный фрагмент, состоящий из опорного понятия, двойной и тройной групп, шести диадно-триадных понятий, альтернативно-тождественного понятия и пятиэлементной группы понятий.

Приведенная формулировка соответствует стандартной схеме классификационного фрагмента. Однако может быть и нестандартная схема расположения понятий, но при этом сохраняется определенное соответствие стандартной и нестандартной схем, так как существует некоторое пространство, в котором можно объективно фиксировать положение всех понятий.

С учетом данного обстоятельства получаем обобщенную формулировку всеобщего фрагментного периодического закона: понятия, упорядоченные по универсальным критериям, образуют периодическую систему, роль периода в которой играет классификационный фрагмент, в стандартном случае состоящий из опорного понятия, двойной и тройной групп, шести диадно-триадных понятий, альтернативно-тождественного понятия и пятиэлементной группы понятий, а в нестандартном случае занимающий пространство, задаваемое стандартной моделью.

В этой формулировке остаются неопределенными понятия "группа", "универсальные критерии", "диадно-триадные понятия" и "пространство, задаваемое стандартной моделью". Определения этих понятий носят характер законов, которые можно назвать внутрифрагментными законами, так как они действуют внутри классификационного фрагмента.

Можно ввести пять внутрифрагментных законов: 1) закон классификационных групп; 2) закон универсальной критериальной упорядоченности; 3) закон перекрестного варьирования диадно-триадных групп; 4) закон равнопервичности позиционно-групповых понятий и 5) закон внутрифрагментного смыслового соответствия (закон понятийной когерентности). Исходя из сказанного можно сформулировать Закон классификационных групп: каждое понятие входит в набор понятий, который обладает смысловым единством, определенностью числа и состава и упорядоченностью по определенному критерию. По типу упорядоченности группы могут быть *ранговыми, позиционными и межъярусными*.

Выявление классификационных групп - один из основных приемов диалектического мышления. Чтобы овладеть им необходимо научиться выявлять смысловые связи между понятиями. Методы их выяснения могут быть различными, но чаще всего может быть использована формула диалектической логики: *тезис - антитезис - синтез*.

Всеобщий фрагментный периодический закон служит путеводной нитью при построении системной (т.е. упорядоченной, периодической и иерархической) классификации.

Основными процедурами, используемыми при построении системной классификации, являются следующие:

1. Составление классификационных групп.
2. Упорядочение по критериям первичности-вторичности, антиэнтропийности-энтропийности и общности-частности.
3. Проверка возможности перекрестного варьирования диадно-триадных групп.
4. Проверка равно первичности позиционно-групповых понятий
5. Установление внутри фрагментного смыслового соответствия.
6. Установление принадлежности к определенному ярусу.
7. Построение межфрагментных рядов первичности-вторичности.
8. Выявление межфрагментных аналогов и т.д.

Возможны два основных типа моделей классификационного фрагмента - знаковые и геометрические модели. В данном случае знаковая модель - это набор букв и цифр (индексная модель) или только цифр (шифровая модель); геометрическая модель - прямоугольник, разделенный на прямоугольные площадки (чертежная модель) (рис. 1), или набор координат этих площадок (координатная модель). Чертежную модель, на которой приведены названия понятий или их символические обозначения, будем называть семантической картой.

Возможны две основные схемы классификационного фрагмента - стандартная и нестандартная. Символически стандартная форма классификационного фрагмента может быть представлена в следующем виде: О, А, Б, 1, 2, 3, 1А, 1Б, 2А, 2Б, 3А, 3Б, П, П1, П2, П3, П4, П5, где О - опорное понятие; А - первое диадное понятие; Б - второе диадное понятие; 1 - первое триадное понятие; 2 - второе диадное понятие; 3 - третье триадное понятие; 1А - первое диадно-триадное понятие; 1Б - второе диадно-триадное понятие; 2А - третье диадно-триадное понятие; 2Б - четвертое диадно-триадное понятие; 3А - пятое диадно-триадное понятие; 3Б - шестое диадно-триадное понятие; П

- альтернативно-тождественное понятие; П1 - первое пентадное понятие; П2 - второе пентадное понятие; ПЗ - третье пентадное понятие; П4 - четвертое пентадное понятие; П5 - пятое пентадное понятие.

О					
А			Б		
1		2		3	
1А	1Б	2А	2Б	3А	3Б
П					
П1	П2	ПЗ	П4	П5	

Рис. 5. Семантическая карта стандартного варианта классификационного фрагмента

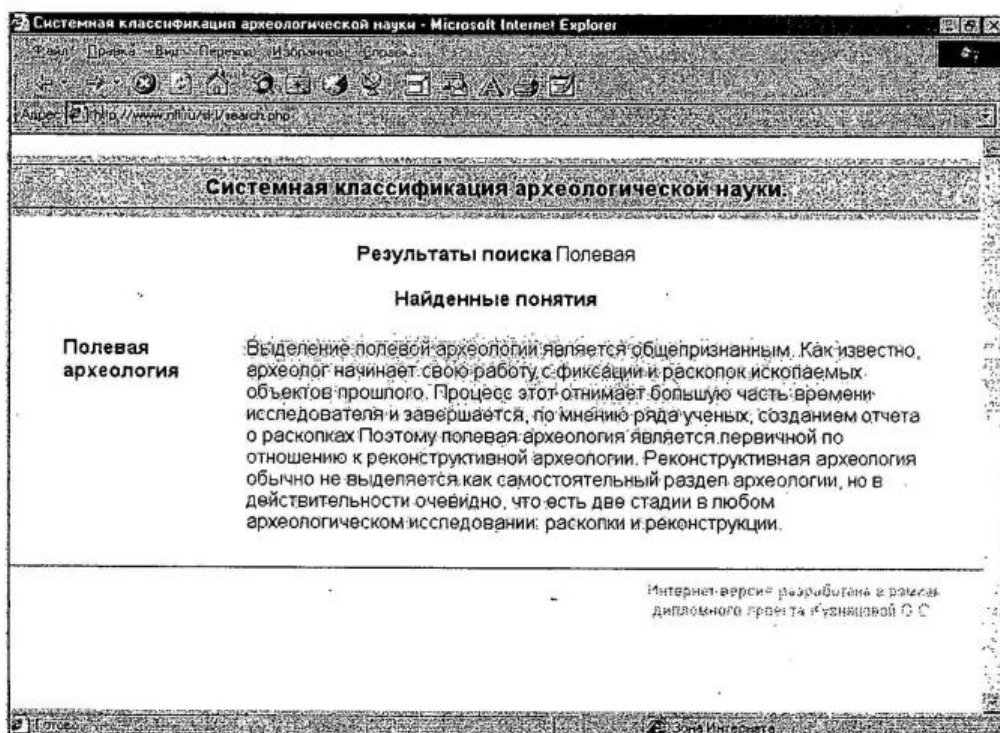


Рис. 6. Результаты поиска археологических понятий.

В ходе разработки системной классификации археологии была разработана система из 54 классификационных фрагментов, включающих в себя 972 понятия. Была так же поставлена задача представить опытный вариант программы с представлением в Интернет.

В ходе решений этой задачи, были проведены следующие работы:

Были осуществлены операции по связыванию классификационных фрагментов.

Была разработана пользовательская система, наглядно отображающая иерархию и связи между классификационными фрагментами.

Для удобства пользователя была создана система поиска понятий (Рис. 6).

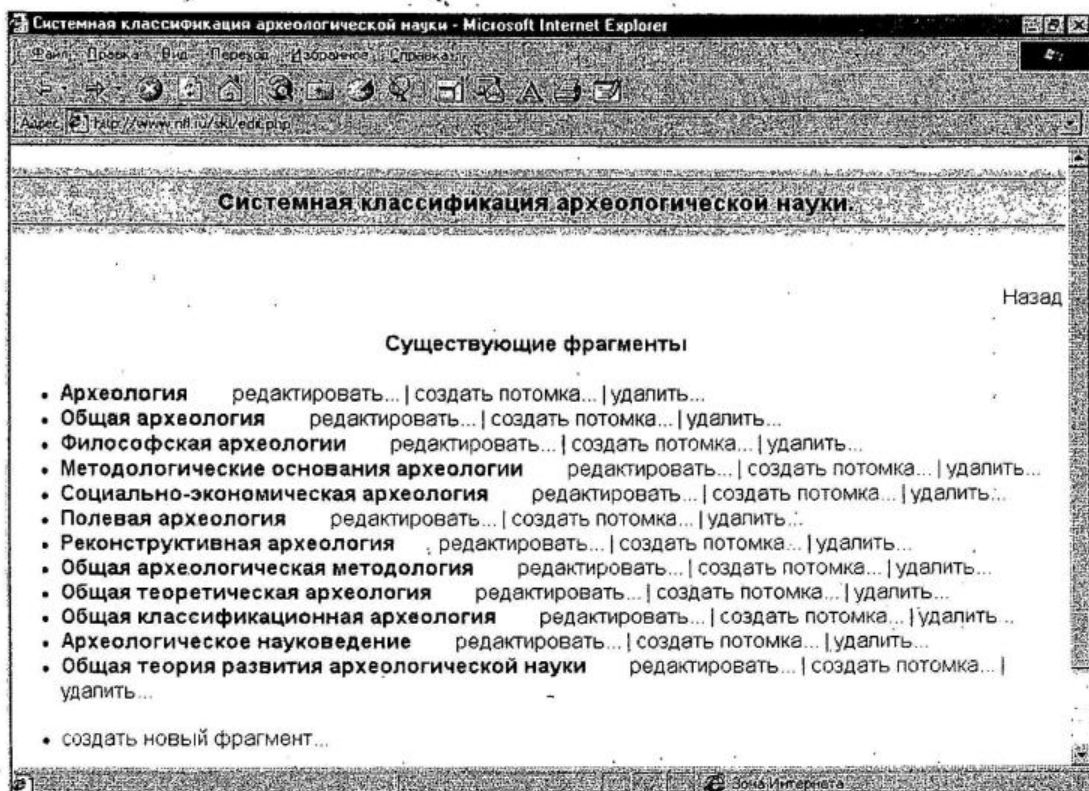


Рис.7. Система редактирования понятий

Построенное дерево классификационных фрагментов не является законченным и для этого была создана система редактирования, позволяющая создавать новые фрагменты, из понятий имеющих своих потомков (Рис. 7).

В результате был разработан опытный вариант программы, позволяющий создавать, иерархически построенную системную классификацию понятий, не только археологической науки, но и пригодный для классификации в других областях знаний.

2.2. Структура портала знаний

В ходе работ была разработана структура портала знаний, который мы предлагаем в качестве средства обеспечения эффективного доступа к Интернет-ресурсам определенной тематики.

Каждый портал знаний может иметь иерархическую или сетевую структуру, т.е. состоять из более специализированных порталов, связанных некоторыми отношениями, например, отношением "общее-частное". Такая архитектура делает портал знаний гибким и легко расширяемым. В основе порталов знаний лежат онтологии, содержащие описание устройства и типологии соответствующих сетевых ресурсов.

На Рис.8. представлена структура предлагаемого портала знаний. Его основными компонентами являются: онтология, коллекционер онтологической информации о ресурсах, конструктор запросов и формирователь ответов. Рассмотрим каждый из них подробнее.

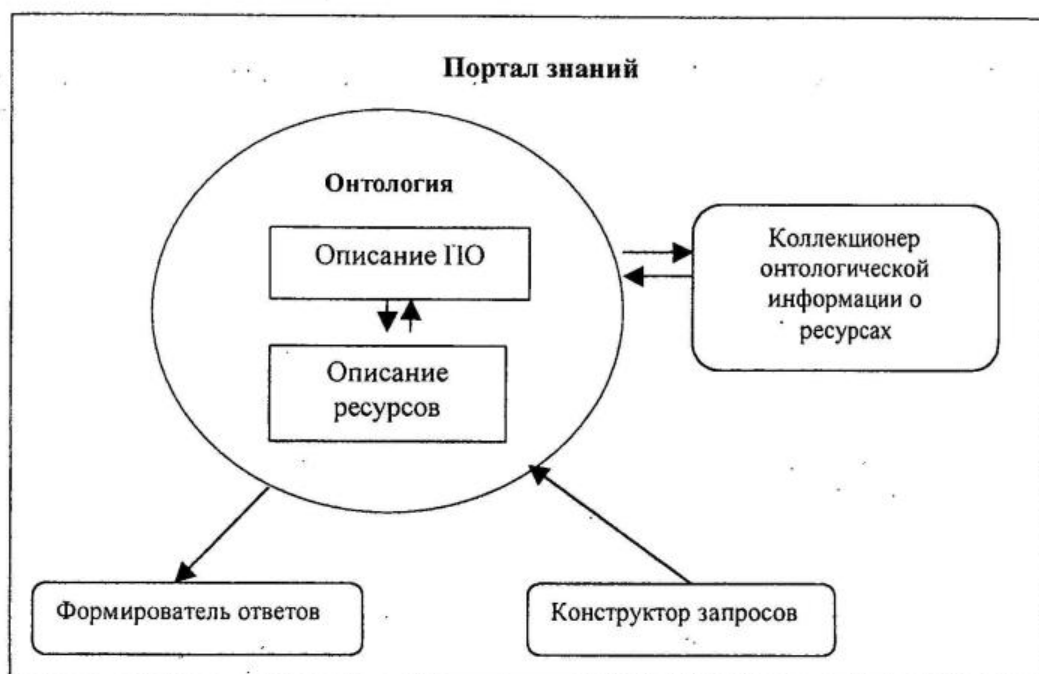


Рис.8. Структура портала

Онтология. Онтологии порталов знаний включают как описание предметной области (ПО), так и описание релевантных ей ресурсов. Описание предметной области включает совокупность терминов и отношений, семантически значимых для данной предметной области, а также правил, согласно которым можно строить утверждения об элементах ПО. Описание ресурса сети включает список ссылок на страницы и сайты, а также описание связей между ними.

Коллекционер онтологической информации о ресурсах, используя стандартные механизмы поиска, последовательно обрабатывает ссылки на документы, страницы, сайты, определяя их тематику, структуру, степень релевантности между тематикой и информацией, которая в них содержится. Полученные сведения накапливаются в онтологии. Коллекционер должен анализировать новые источники и проверять старые.

Формирователь ответов. Обрабатывая источники, формирователь ответов отслеживает степень их релевантности понятиям и выдает список наиболее подходящих для данного понятия ресурсов.

Конструктор запросов. Портал знаний имеет язык запросов, близкий к естественному языку. Это требует наличия в нем соответствующего лингвистического обеспечения, одним из основных компонентов которого является словарь терминов, органически связанный с онтологией.

На основе такого лингвистического обеспечения должны решаться следующие важные задачи:

- задание лингвистической информации о данной предметной области на ЕЯ. Для этого необходимо построение словаря, организованного с учетом проблемно-ориентированной лексики, морфологии, словообразования, синонимии;
- описание системы знаний о предметной области в виде комплекса понятий, связанных между собой отношениями;
- для описания входного языка, близкого к естественному, используется специализированный словарь-тезаурус разработанной в РосНИИ ИИ системы Alex. В этом словаре отражены в виде иерархических лексических шаблонов все возможные вхождения терминов (слов и словосочетаний) конкретной предметной области, что позволяет задавать в запросе термины не только на русском, но и других языках. В

дальнейшем эта иерархия понятий может быть использована для интеграции словаря-тезауруса и онтологии понятий в единую сеть знаний;

– благодаря интеграции словаря-тезауруса и онтологии портал знаний становится способным “понимать” разноязычные ресурсы и воспринимать запросы на разных языках. При этом онтология становится независимой от конкретных языков, так как описание языковых (лингвистических) свойств понятий отражаются в тезаурусе, в то время как в онтологии представлена информация только об их семантике.

2.3. Описание ресурса

Как уже было сказано выше, онтология портала знаний включает как описание предметной области, так и описание релевантных ей ресурсов.



Рис. 9. Фрагмент иерархии понятий предметных областей

Часть онтологии, описывающая конкретную предметную область, включает в себя совокупность терминов и отношений, семантически значимых для данной предметной области, а также правил, согласно которым можно строить утверждения об элементах ПО.

Процесс построения онтологии распадается на серию подпроцессов. Обычно сначала составляется глоссарий терминов, который в дальнейшем используется для исследования свойств и характеристик представленных в нем терминов, затем на естественном языке создается список точных определений. После этого на основе таксономических отношений строятся деревья классификации понятий (рис. 9).

Для фиксации значимых отношений между терминами выделяются основные связи между ними, которые можно графически отобразить с помощью диаграмм бинарных отношений. Такие связи в дальнейшем могут послужить основой для интеграции различных онтологий.

На примере, изображенном на рис. 10, показано объединение онтологий, соответствующих таким предметным областям, как *представление знаний*, *организация*, *человек*, *проект* и *ресурс*, на основе отношений, существующих между представленными в них понятиями.

Описание ресурса сети включает ссылки на ресурсы, описание страниц, сайтов и связей между ними. Для этого в онтологию заносятся такие характеристики ресурса,

как название, адрес, автор, структура, тематика, язык, тип информации (текст, графика, мультимедиа, гиперссылки, дата обновления, число посетителей).

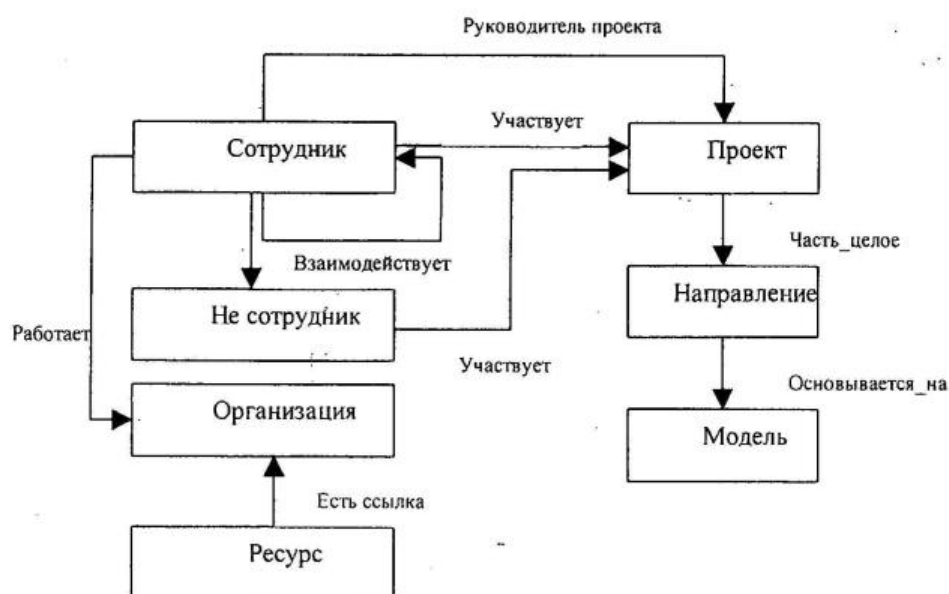


Рис.10. Фрагмент диаграммы бинарных

Важной характеристикой ресурса является его *степень релевантности* данному понятию ПО. Она может быть представлена некоторой числовой величиной от 0 до 1, определяющей на диаграмме вес дуги, связывающей данное понятие с ресурсом.

Таким образом, в онтологии становится возможным описывать такие свойства понятий, как "ссылка на релевантный ресурс в сети Интернет" с указанием значения атрибута "степень релевантности". С каждым ресурсом могут быть связаны и другие свойства, полезные при поиске информации, например, "ссылка другого ресурса на ресурс" и др.

Указанные свойства и атрибуты приписываются понятиям и связанным с ними ресурсам в соответствии с результатами предварительного индексирования сети Интернет и последующей оценки обработанных ресурсов на релевантность.

2.4. Схема организации поиска

Построенный по предложенной выше схеме портал знаний может поддерживать несколько видов поиска:

- "локальный" (когда поиск информации производится непосредственно среди ссылок на ресурсы, которые содержатся в онтологии),
- "глобальный" (когда запрос перед передачей той или иной поисковой системе предварительно уточняется с помощью онтологии и переводится в соответствующий формат с использованием ключевых слов (терминов) и логических связей).

Выбор вида поиска осуществляется в пользовательском интерфейсе портала знаний, который обеспечивает также возможность составления поискового запроса на естественном языке и просмотра результатов поиска.

На первом этапе поисковый запрос, представленный на естественном языке, обрабатывается системой Alex, которая осуществляет лексический анализ запроса посредством сопоставления текста запроса с множеством лексических образцов-шаблонов, заданных в словаре-тезаурусе.

Затем, полученный в результате лексического анализа набор терминов (понятий) предметной области подвергается семантическому анализу, который выполняется на основе онтологии, составляющей со словарем-тезаурусом единую сеть знаний. При этом рассматриваются как свойства понятий предметной области и связи между ними, так и свойства представленных в онтологии ресурсов, сопоставленных указанным в запросе понятиям предметной области. Результирующий список ссылок на ресурсы формируется с учетом степени релевантности каждого ресурса полученной совокупности понятий.

При формировании данного списка мы различаем относительную r_o и абсолютную r_a релевантность ресурса запросу.

В простейшем случае, когда не учитывается взаимосвязь понятий и их степень важности в запросе, относительная релевантность r_o ресурса R_k запросу Q , включающему n понятий P_i данной предметной области, может быть вычислена по формуле

$$r_o = \frac{\sum_{i=1}^n r_i}{n},$$

где r_i – релевантность ресурса R_k понятию P_i .

Абсолютная релевантность ресурса запросу вычисляется по формуле

$$r_a = \prod_{i=1}^n r_i.$$

Из формулы видно, что абсолютная релевантность ресурса R_k запросу Q становится нулевой, если этот ресурс нерелевантен хотя бы одному понятию из запроса. Поэтому использование абсолютной релевантности ресурса предполагает более жесткий критерий отбора ресурсов, чем использование относительной релевантности.

В то же время, очевидно, что описанные выше способы вычисления релевантности как в отдельности, так и в комбинации, дают довольно грубую оценку релевантности ресурса запросу. В связи с этим нами разрабатывается метод вычисления релевантности, учитывающий взаимосвязь понятий, а также их степень важности в запросе.

Предложенный подход, позволяет повысить эффективность поиска в сети Интернет путем организации порталов знаний, каждый из которых предоставляет пользователю доступ к ресурсам определенной тематики. В основе таких порталов знаний лежат онтологии, содержащие наряду с традиционным описанием предметной области соотнесенное с ним описание структуры и типологии соответствующих сетевых ресурсов. Главным преимуществом данного подхода является то, что порталы знаний позволяют значительно сократить время обработки запроса пользователя и количество выдаваемых ресурсов за счет более точного определения степени их релевантности и хранения ссылок на них непосредственно на портале знаний.

Пока нет достаточного опыта построения и статистики использования специализированных порталов знаний, так как они только стали появляться, но мы считаем, что порталы знаний, основанные на принципах, представленных в данной разработке, должны обеспечить более быстрые и релевантные ответы на запросы, чем существующие на данный момент поисковые системы.

3. РАЗРАБОТКА НОВЫХ МАТЕМАТИЧЕСКИХ И СТАТИСТИЧЕСКИХ МЕТОДОВ И ТЕХНОЛОГИЙ ОБРАБОТКИ АРХЕОЛОГИЧЕСКИХ ДАННЫХ ДЛЯ ПАЛЕОЛИТИЧЕСКИХ ОБЪЕКТОВ

3.1. ЗНАЧИМОСТЬ И УСТОЙЧИВОСТЬ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ В ЗАДАЧЕ ПОИСКА ОПТИМАЛЬНОГО РАЗБИЕНИЯ

Введение

Задача проведения оптимальной (по количеству классов) автоматической классификации поставлена достаточно давно, но предложенные решения носят скорее эвристический, чем статистический характер, что оставляет вопрос об оптимальной классификации открытым.

Автоматическая классификация, синонимами которой можно считать термины "разбиение", "таксономия", "кластерный анализ", применяется в целях получения гипотез о логической структуре изучаемой статистической совокупности объектов. Слово автоматическая подчеркивает тот факт, что разделение проводится без предварительного обучения с помощью учителя или обучающей выборки, на которой все объекты разнесены по классам. Результатом классификации является разбиение исходной совокупности объектов на некоторое число классов (таксонов, кластеров). Содержательный смысл деления на классы состоит в выделении качественно различных состояний объектов, характеризующихся своими особыми закономерностями. Дальнейшее исследование может как подтвердить, так и опровергнуть гипотезу о существовании обнаруженной кластерной структуры. Подтверждением правильности классификации является ее объяснительная и предсказательная сила при выходе за пределы исходной выборки как по признакам, так и по объектам.

Классификация применяется на предварительном этапе исследования, когда до проведения основной статистической обработки (построения регрессионных моделей, оценки параметров генеральной совокупности и т.д.) добиваются расслоения этого множества на однородные (в смысле проводимого затем статистического анализа) порции данных.

Современная технология автоматической классификации характеризуется набором различных алгоритмов распределения объектов по кластерам, из-за чего проверка статистической значимости в действительности мало неприменима.

В данной работе описаны несколько первых шагов на пути превращения автоматической классификации из набора алгоритмов в статистический метод, который позволит не только получать практически полезный результат, но и оценивать статистическую значимость нулевой гипотезы (об отсутствии кластерной структуры, или, что эквивалентно, - об однокластерности статистической совокупности объектов).

Описываемый далее подход является развитием идей, развиваемых участниками проекта и их коллегами на протяжении многих лет. Ранее была описана общая постановка задачи автоматического определения оптимальной сложности разбиения для самых разных методов. В случае кластерного анализа сложность разбиения интерпретируется как количество кластеров. В качестве критерия оптимальности предлагалось использовать такие показатели качества разбиения, как значимость гипотезы однокластерности и устойчивость полученной кластерной структуры, измеряемая процентом оставшихся в своих кластерах объектов при генерации повторных выборок методом Boot-Strap.

В целом продолжая предложенные подходы, мы все же несколько сместили акценты. Если результаты, полученные в монографии, можно использовать для создания статистических таблиц значимости при кластеризации по стандартизованным

данным, то здесь мы напрямую (в статистических экспериментах) измеряем значимость разбиений для определения оптимального количества кластеров. Такое смещение акцента связано с существенным ростом за последние 5 лет доступных рядовому пользователю вычислительных мощностей персональных компьютеров, что делает возможным проведение полномасштабных вычислительных экспериментов в реальном времени.

Освободив показатель устойчивости классификации от его основных обязанностей – служить критерием оптимальности разбиения, авторы нашли для него новое применение – составление "структурного портрета" полученного разбиения, который выявляет взаимопереходы кластеров, их взаимодействия через обмен объектами в экспериментах с имитацией повторного сбора данных. Такие "структурные портреты" дают дополнительную информацию о кластерах, позволяя преодолеть недостатки выбранного метода кластеризации (к-средних), который является оптимальным по скорости, но накладывает существенные ограничения на форму выделяемых кластеров в виде сферических сгущений точек (объектов). Большая частота перетекания объектов между двумя соседними кластерами (межкластерная неустойчивость) дает основания выдвинуть гипотезу о единстве этих кластеров и в дальнейшем интерпретировать их как один кластер несферической формы.

Перед тем, как приступить к описанию способов определения значимости, рассмотрим кратко, что представляет из себя критерий качества разбиения и из каких шагов складывается алгоритм кластеризации. Из соображений простоты реализации и скорости выполнения нами были выбраны Евклидова метрика пространства признаков и простейший алгоритм кластеризации (к-средних), хотя способ измерения значимости не накладывает особых ограничений ни на метрику, ни на сам алгоритм. Но все же необходимо отметить, что к поведению алгоритма предъявляются повышенные требования, поэтому мы были вынуждены внести в него некоторые изменения, которые позволили улучшить качество разбиения и, как побочный эффект, сократить время вычислений (за счет уменьшения количества итераций).

Критерий качества разбиения

Результатом кластерного анализа является разбиение имеющихся в исходной выборке объектов на некоторое количество кластеров. При этом каждый кластер включает в свой состав один или более объектов. Мы будем описывать кластер не просто как множество объектов, а как некий идеальный объект, имеющий свои координаты в пространстве признаков. Принимая, далее, что эти координаты являются арифметическим средним координат реальных объектов, входящих в состав кластера, то есть координаты кластера совпадают с координатами его центра.

Теперь зададимся вопросом, что будет, если мы, вместо реальных координат объекта, будем использовать координаты кластера, к которому он отнесен? Очевидно, при этом мы теряем часть информации об объекте, поскольку внутри кластеров объекты не идентичны. Но, чем лучше кластерная структура описывает реальные скопления объектов в признаковом пространстве, тем меньшая часть информации будет потеряна и тем большую ее часть будет нести в себе кластер, как представитель всех принадлежащих ему объектов.

Исходя из этих соображений, логично выбрать критерием качества кластеризации какую-либо характеристику потерь информации при описании объектов кластерами. Наиболее простой оценкой таких потерь является остаточная дисперсия, представляющая собой сумму квадратов отклонений объектов от центров их кластеров. Если же мы хотим иметь дело с безразмерной величиной, то наиболее удобным будет нормировать эту величину на остаточную дисперсию однокластерной структуры. Полученная таким образом *доля остаточной дисперсии* принимает значения в диапазоне от нуля до единицы. Чем лучше качество разбиения, тем ближе она к нулю, и наилучшему разбиению на кластеры будет соответствовать *минимум* доли остаточной

дисперсии. А сама эта величина напрямую показывает, какую часть информации мы потеряем при замене объектов на кластеры.

Алгоритм кластеризации

Поскольку способ измерения значимости гипотезы отсутствия кластерной структуры требует проведения кластеризации не только на исходных данных, но и на десятках, сотнях случайно сгенерированных выборок, требования к скорости алгоритма по сравнению с традиционным кластерным анализом существенно ужесточаются. С учетом этого обстоятельства выбор метода k -средних становится почти вынужденным.

Но измерение значимости требует от алгоритма не только скорости, но и "хорошего поведения", поскольку значимость определяется при сравнении остаточной дисперсии на исходных данных с усредненной остаточной дисперсией в статистических экспериментах (малая разность двух больших величин намного более чувствительна к ошибкам, чем сами эти величины). Мало того, нас интересует не просто значимость сама по себе, а минимум значимости при изменении количества кластеров от двух до некоторого предельного значения. Мы можем позволить себе ошибиться в величине значимости при некотором k , но не должны допускать ошибки при сравнении значимостей на разных количествах кластеров. Таким образом, от алгоритма требуется по меньшей мере монотонность критерия при увеличении количества кластеров, а также как можно меньший дефект алгоритма (в данном случае - это "недожатая", по сравнению с глобальным минимумом, доля остаточной дисперсии, которая вносит дополнительный шум в оценку значимости).

Исходный метод k -средних можно разложить на следующие 4 шага:

- 1) случайно выбираются k объектов - центров кластеров;
- 2) каждый объект относится к ближайшему кластеру (по расстоянию до центра);
- 3) вычисляются новые центры кластеров - средние арифметические по объектам;
- 4) если произошло уменьшение критерия качества кластеризации - осуществляется переход к п. 2).

Самый слабый пункт в этом алгоритме (и единственный, подлежащий модификации без качественного изменения идеи алгоритма) - это первый шаг. Поскольку метод дает не глобальный, а локально-оптимальный результат, то и качество разбиения, и скорость зависят в решающей степени от удачного начального разбиения.

В описанный алгоритм было внесено 2 изменения, хотя резервы улучшения этим далеко не исчерпаны. Во-первых, учитывая то, что поиск оптимального количества кластеров требует последовательного проведения кластеризации для диапазона значений $k > 1$, мы решили не отбрасывать результат предыдущего разбиения на k кластеров, а использовать полученные центры в качестве начального приближения на $k+1$ шаге. Эта модификация позволила исключить случаи немонотонного уменьшения остаточной дисперсии с ростом количества кластеров. Во-вторых, для определения $k+1$ кластера мы берем не пустое множество, как предлагалось в [4: 88], а поступаем следующим образом. Среди k кластеров находим такой, который при разбиении на два дает наибольший эффект уменьшения критерия. Эти две части кластера и дают координаты k -го и $k+1$ -го кластеров.

Оценка значимости

Критерий качества разбиения, описанный выше, позволяет определить оптимальное разбиение при заданном количестве кластеров, но не может служить основанием для сравнения внутри ряда оптимальных разбиений. Это связано с тем, что значение этого критерия при возрастании количества кластеров ведет себя как монотонно убывающая функция (пока не достигнет нуля), но мы не располагаем информацией, насколько быстро она *должна* убывать. Для получения такой информации мы ввели в качестве эталона для сравнения случайно сгенерированную, в соответствии с нулевой гипотезой, выборку данных того же объема, что и исходная. Нулевая гипотеза заключается в том, что распределение объектов в пространстве

признаков соответствует многомерной нормальности. При этом длины осей эллипсоида рассеяния для случайной выборки совпадают с таковыми в исходной. В дальнейшем тексте нулевую гипотезу будем называть гипотезой однокластерной структуры или гипотезой отсутствия кластерной структуры.

Для определения оптимального количества кластеров мы использовали оценку значимости гипотезы однокластерной структуры. Минимум значимости достигается на числе кластеров, которое дает наиболее отклоняющуюся (в сторону меньших значений) долю остаточной дисперсии на исходных данных по сравнению со средним значением доли остаточной дисперсии в статистических экспериментах.

При разработке алгоритма оценки значимости необходимо принять два решения:

- как генерировать случайную выборку с однокластерной структурой (по построению), которая была бы по остальным параметрам сравнима с исходной выборкой?;

- как измерять значимость гипотезы однокластерности, располагая значениями критерия качества разделения (при k кластерах) на исходных данных и для всего множества случайных выборок?

Рассмотрим эти проблемы по порядку.

В монографии предлагалось имитировать однокластерную структуру выборкой, нормально распределенной в сферически симметричной области, а сами исходные данные преобразовывать к такому виду с помощью ориентации в многомерном пространстве вдоль главных компонент (найденных факторным анализом). Поскольку при этом мы получаем облако рассеяния в виде многомерного эллипсоида, то следующим шагом растягиваем его оси так, чтобы их длина стала одинаковой, то есть превращаем эллипсоид в сферу. Такой подход страдает несколькими недостатками. Во-первых, при переходе к главным компонентам размерность пространства может уменьшиться за счет линейной зависимости выбранных переменных, а во-вторых, после растягивания осей мы меняем метрику пространства и в результате ищем уже совсем не ту кластерную структуру, которую обнаружили бы в исходном пространстве. Таким образом, решая задачу оценки значимости отсутствия кластерной структуры, нам пришлось бы накладывать слишком жесткие ограничения на саму процедуру кластеризации.

Учитывая указанные недостатки, мы пришли к альтернативному решению — преобразовывать не исходные данные к сферическому распределению, а генерировать случайные выборки в эллипсоиде, длины осей которого совпадают с главными компонентами. При этом оси нулевой длины не приносят неприятностей (умножать на ноль всегда проще, чем делить), а на процедуру кластеризации не накладываются дополнительных ограничений.

Теперь перейдем к вопросу измерения значимости гипотезы однокластерности. Поскольку критерием качества кластеризации была выбрана доля остаточной дисперсии, которая принимает значения от нуля до единицы, то естественно будет аппроксимировать распределение значений этого критерия в серии случайных экспериментов Бэта-распределением, которое работает для случайных величин, имеющих ограниченную с обеих сторон область допустимых значений:

$$f_{a,b}(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \quad (1)$$

Параметры a и b Бэта-распределения легко определить по выборочному среднему M и дисперсии D :

$$a(k) = \frac{M(k)(M(k)(1-M(k)) - D(k))}{D(k)} \quad (2)$$

$$b(k) = \frac{a(k)(1 - M(k))}{M(k)} \quad (3)$$

Получив оценки параметров Бэта-распределения, мы можем вычислить значимость гипотезы однокластерности при любом значении $x(k)$ – доли остаточной дисперсии на k кластерах. Значимость вычисляется как интеграл Бэта-распределения в пределах от нуля до остаточной дисперсии $d(k)$:

$$p(k) = \int_0^{d(k)} \frac{x^{a(k)-1} (1-x)^{b(k)-1}}{B(a(k), b(k))} dx \quad (4)$$

Учет микро-кластеров

Итак, мы имеем способ генерации случайной выборки и способ измерения значимости гипотезы однокластерности. Казалось бы, этого достаточно, чтобы считать метод определения оптимального количества кластеров готовым к применению. Однако пробные расчеты на реальных данных показали, что это далеко не так. Оказалось, что не существует данных, на которых этот метод в чистом виде дал бы ожидаемый результат – оптимальное количество кластеров. Авторы столкнулись с эффектом систематического уменьшения значимости (см. рис. 1) при увеличении количества кластеров вплоть до $n-1$ (n – количество объектов в выборке). Получалось, что в области практически интересного количества кластеров минимум значимости не локализуется.

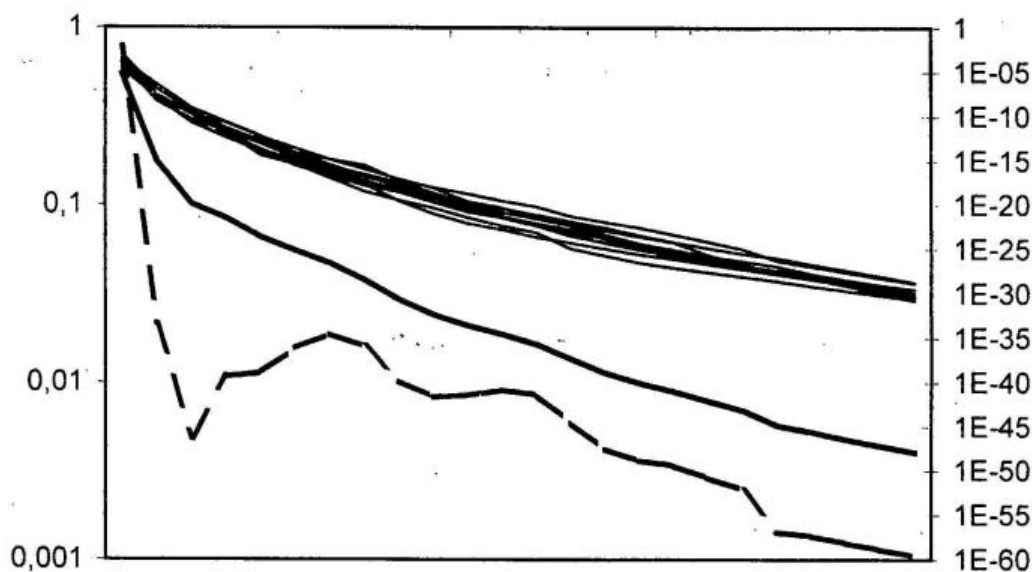


Рис. 1. "Хвост значимости" не позволяет определить оптимальное количество кластеров.

Какие способы выхода из этой ситуации можно предложить? Их по крайней мере три:

1) скорректировать случайные данные так, чтобы на них распределение расстояний между ближайшими соседями не отличалось от того же распределения на исходных данных;

2) скорректировать исходные данные так, чтобы на них распределение расстояний между ближайшими соседями не отличалось от того же распределения на случайных данных, то есть искусственно рандомизировать исходные данные, разрушив микро-кластеры;

3) генерировать случайные данные с учетом реальных одномерных распределений и ковариационной матрицы. При этом все "квантовые эффекты" будут гарантированно скопированы с исходных данных.

Каждый из этих способов обладает своими достоинствами и недостатками, в частности, третий способ требует дополнительной проработки и пока остается непроверенным.



Рис. 2. Распределение межобъектных расстояний.

В программе мы выбрали последний способ, то есть после генерации каждой случайной точки ее координаты округляются до ближайшего значения в реальном одномерном распределении. Как видно из сравнения рис. 3 с рис. 1, такой способ корректировки генерируемых точек позволяет эффективно бороться с хвостом значимости.

Поиск причины появления "хвоста значимости" немедленно привел к исследованию распределения межобъектных расстояний (см. рис. 2). Построение графика распределения расстояний между ближайшими соседями выявило резкое отличие этого распределения на реальных данных от аналогичного распределения на случайных. А именно — в реальных выборках распределение систематически смещено в область нулевых расстояний, то есть наблюдается повышенное содержание "слипшихся точек", которые образуют множество "микро-кластеров", состоящих из 2-3 объектов.

Что является причиной повышенной концентрации "слипшихся точек" в реальных данных, остается открытым вопросом, однако одна гипотеза все же есть. Согласно этой гипотезе, причина кроется в ограниченной точности измерения данных (в отличие от практически неограниченной точности представления случайных данных), которая наблюдается в одномерных распределениях значений в виде многократно повторяющихся значений (дублей). Таким образом, здесь мы, возможно, сталкиваемся со статистическим "квантовым порогом", за пределами которого предположения о непрерывности распределения становятся неприменимыми, что и приводит к катастрофе классического подхода.

Дефекты алгоритма

Алгоритм определения значимости нулевой гипотезы (отсутствия кластерной структуры) в том виде, как он здесь описан, обладает существенным недостатком: он вносит большой шум в измеряемую им значимость, и величина этого шума уменьшается только с ростом числа статистических экспериментов. В то же время, из-за высокой вычислительной сложности проводимых экспериментов, мы не можем себе позволить добиться приемлемой точности измерения методом грубой силы.

Рассмотрим несколько подробнее основные источники шума. Для этого необходимо проанализировать способ измерения значимости. Из уравнения 4 видно, что значимость нулевой гипотезы является преобразованным значением остаточной дисперсии, и определяющую роль в этом преобразовании играют параметры $a(k)$ и $b(k)$, которые, в свою очередь (уравнения 2, 3), полностью определяются статистическими параметрами распределения $M(k)$ и $D(k)$ значений остаточной дисперсии по всем проведенным экспериментам. Итак, мы пришли к пониманию того, что шум измерения имеет своим источником статистический разброс параметров $M(k)$ и $D(k)$.

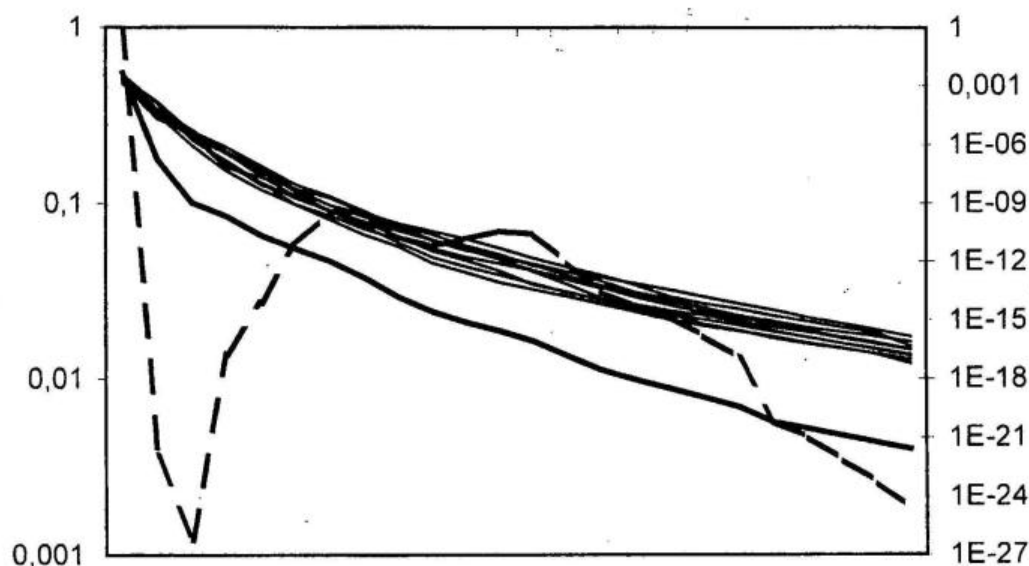


Рис. 3. Значимость после корректировки.

Теперь, зная источник шума, попробуем наметить (на будущее) подходы к его уменьшению:

1. Увеличение количества проводимых экспериментов (метод грубой силы).
2. Сглаживание рядов $M(k)$ и $D(k)$ при помощи цифровых фильтров.
3. Экспериментальное исследование зависимостей M и D от k с определением класса функций для их аппроксимации и аппроксимация рядов этими функциями путем подбора параметров.

Самым перспективным нам представляется последний подход, который, однако, требует дополнительного времени на исследование.

Следует отметить, что и само значение остаточной дисперсии, как предел интегрирования в уравнении 4, также является источником шума, причина которого кроется в дефекте алгоритма кластеризации, направленного на поиск локального, а не глобального минимума остаточной дисперсии, за счет чего остаточная дисперсия "не выжимается на все 100%", и эта "недожатая" дисперсия порождает дополнительный разброс величины критерия как в исходных данных, так и в статистических экспериментах.

Неустойчивость классификации как проявление межкластерных связей

Если раньше интегральный показатель устойчивости классификации предлагалось использовать как критерий выбора оптимального количества кластеров, то в настоящей работе эта ниша занята показателем значимости гипотезы однокластерности. Но это не означает, что исследование устойчивости утратило необходимость. Напротив,

определилась задача, в решении которой результаты измерений устойчивости (а правильное было бы сказать - неустойчивости) оказываются применимы напрямую.

Эта задача связана с преодолением недостатка выбранного в данной работе метода кластеризации (к-средних). Дело в том, что метод к-средних накладывает определенные ограничения на форму выделяемых кластеров, источником которых (ограничений) является критерий оптимальности искомой структуры в виде минимума внутрикластерного разброса. Поскольку этот разброс вычисляется по расстояниям от центра кластера, до всех объектов, входящих в этот кластер, то в результате форма выделяемых кластеров получается близкой к сферической. Но в тех случаях, когда прообразы кластеров в признаковом пространстве оказываются вытянутыми или тороидальными (для циклически протекающих процессов), метод к-средних, в лучшем случае, выдаст цепочку следующих друг за другом шарообразных кластеров.

Восстановление исходной формы (и количества) кластеров, искаженных описанным недостатком алгоритма, и является той задачей, которая в данной работе возложена на описательную статистику неустойчивости кластерной структуры. Показатель неустойчивости определяется как доля объектов, перебегающих (в среднем) между каждой парой кластеров, и представляется графически в виде отрезка, соединяющего центры кластеров. Ширина отрезка пропорциональна показателю неустойчивости. Кластеры, соединенные самыми широкими полосами неустойчивости, являются первыми кандидатами на объединение в один класс. Правомерность такого объединения требует подтверждения независимыми методами, а потому является только гипотетическим.

3.2. СТРУКТУРНАЯ ИЗОГЕОМЕТРИЯ ЛОКАЛЬНЫХ СПЛАЙНОВ С ДОПОЛНИТЕЛЬНЫМИ УЗЛАМИ

3.2.1. ЗАДАЧИ СТРУКТУРНО-ИЗОГЕОМЕТРИЧЕСКОГО АНАЛИЗА ДАННЫХ В ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЯХ

Понятие изогеометрии в информационных технологиях

В решении информационных проблем в гуманитарных исследованиях, разработках и образовательном процессе важная роль отводится математическим методам. Математические средства позволяют преодолеть те трудности в решении возникающих задач, с которыми с помощью иных, не математических методов, справиться не удастся. Наиболее часто эти затруднения возникают в обработке данных, представленных в таблично-числовой форме. Наряду с текстами и изображениями таблично-числовая форма представления информации в гуманитарных исследованиях относится к числу наиболее наглядных и распространенных. Среди задач по обработке подобной информации следует выделить те, которые непосредственно связаны с проблемами полноты и сопоставимости, а также структурного и графического представления табличных данных.

Решение этих проблем средствами математики можно выделить в отдельный класс так называемых изогеометрических задач, содержание которых излагается ниже. Всякую таблицу или ее часть (подтаблицу) соблазнительно представить в виде некоторой поверхности или линии, чтобы получить возможность изучения иных аспектов проблематики исследований или обучения и использования новых исследовательских и познавательных инструментов и средств. Наиболее интересным направлением может служить исследование изогеометрических свойств.

Под изогеометрическими свойствами математических, информационных или иных объектов понимаются такие инварианты, которые в наиболее наглядной и адекватной форме могут быть представлены с помощью изобразительно-геометрических (графических) средств. В качестве форм проявления подобных свойств можно

привести, к примеру, видимые (или воображаемые) на линии или на поверхности зоны сгущения или разреженности, вершины, впадины, участки подъема, спуска, перевалы (седловые точки) и т.п.

В математическом контексте такими свойствами могут быть перепад градиентов через крутизну или пологость участков поверхности или кривых, а также выпуклость или вогнутость, монотонность убывания или возрастания, перемена знака или знакопостоянство и т. д. Непрерывность или разрывы тоже можно считать изогеометрическими свойствами, ибо их также можно демонстрировать с помощью графики. К числу наиболее важных изогеометрических свойств в первую очередь следует отнести свойство монотонности, выпуклости (вогнутости) и знакопостоянства, с которым непосредственно связано множество других изогеометрических и структурных качеств исследуемых объектов.

Традиционные методы математической обработки данных в информационных технологиях, используемых не только в гуманитарных исследованиях, но и в естественнонаучных изысканиях и в технических приложениях, пока еще слабо проработаны применительно к исследованию изогеометрических свойств.

В исследовательской практике археолога, этнографа, антрополога, историка или социолога часто возникают задачи, для решения которых приходится обрабатывать множество таблиц с выборочными распределениями из разных генеральных совокупностей. Когда интервалы, на которые разбиваются значения варьирующего признака, для всех выборок одинаковы, такого рода задачи могут быть в какой-то мере решены на основе традиционных статистических методов.

Но условие совпадения границ интервалов зачастую может быть нарушено вследствие их неодинаковой длины, в частности, при неравномерном разбиении значений варьирующего признака. Из-за этого границы интервалов для сопоставляемых выборок могут быть не одни те же, и как следствие результаты выборочных обследований могут оказаться несопоставимыми. В обработке данных эта ситуация часто обостряется тем, что повторное обследование оказывается невозможным (что почти всегда характерно для большинства гуманитарных исследований) или слишком дорогим. Учитывая то, что каждый исследователь или обучающийся имеет свои вкусы и приемы сбора и обработки информации, результаты исследований или измерений часто оказываются статистически несопоставимы.

Трудности, обусловленные несопоставимостью данных, тесно смыкаются с не менее важной проблемой восполнения данных. Она возникает тогда, когда исходная информация для решения многих задач, а также промежуточные или конечные результаты расчетов заданы на неравномерной сетке или не во всех точках, по которым такие решения предусматриваются. Во многих случаях такого рода неполнота информации делает ее несопоставимой. Поэтому возникает двуединая задача: обеспечить используемым данным одновременно сопоставимость и полноту.

Математическое решение проблемы восполнения данных сводится к тому, что на основе исходной информации вычисляются по определенным алгоритмам, требования к которым будут изложены ниже, промежуточные данные, сохраняющие исходную изогеометрическую структуру. Так, например, если исходные данные, заданные в табличной форме как значения некоторой функции в узлах исходной сетки, представляют собой или содержат в себе последовательность неотрицательных монотонно возрастающих или монотонно убывающих чисел, то рассчитанные по указанным алгоритмам недостающие данные, будучи вставленными в соответствующие места этой последовательности, не должны нарушать этой изогеометрической закономерности.

Обеспечение сопоставимости данных достигается аналогично, если исходные данные представляют собой выборочные распределения. В последнем случае могут быть использованы две схемы.

При первой схеме сначала обычными средствами (с помощью накопленных частот) определяются значения функции распределения выборочных данных на границах интервалов. Затем с помощью алгоритмов интерполяции, требования к которым тоже

будут приведены ниже, строится неотрицательная монотонная (неубывающая) функция (желательно гладкая), значения которой на границах интервалов совпадают со значениями функции распределения выборочных данных. В этом случае, очевидно, сохранение знака осуществляется автоматически.

Во второй схеме первый шаг (построение функции распределения выборочных данных) тот же. На втором шаге вычисляются значения производной этой функции на границах интервалов. Наконец, на третьем шаге осуществляется гладкая изогеометрическая интерполяция этой производной, сохраняющая не только монотонность, но и знак. Сохранение знака требуется из-за неотрицательности распределений.

Одной из наиболее важных задач исследований, разработок или обучения является выявление структурных особенностей в строении изучаемых объектов. В частности, такая проблема часто возникает при обработке числовых таблиц. Как правило, значительная часть структурных особенностей объектов может быть представлена через изогеометрические свойства. Для этих целей нужна соответственно такая аппроксимация табличных данных, которая всегда сохраняла бы изогеометрические свойства объектов в той мере, в какой они зафиксированы в таблице.

Особо следует сказать о роли графического представления данных в исследовательском процессе. Графика позволяет придавать первичным данным, а также промежуточным или конечным результатам наглядный и обобщенный вид, который позволяет исследователю оценивать эти данные в количественной и качественной форме. Наиболее наглядным графическим представлением данных, как правило, является такое, при котором в минимальной мере используются "негладкие" формы. Например, обычный график гладкой кривой намного привычнее и нагляднее для исследователя, нежели аналогичная ломаная кривая. С другой стороны, графика наиболее информативна, когда с ее помощью фиксируются структурные характеристики представленного объекта. Но гладкие кривые, создаваемые по традиционным технологиям и не сохраняющие изогеометрических свойств объектов, пока еще значительно уступают в этом отношении кусочно-линейным графикам. Недаром в наиболее известных и мощных статистических пакетах для персональных компьютеров, таких, например, как Microsoft Excel, Quattro Pro и других, графическое оформление результатов производится преимущественно с помощью кусочно-линейных кривых и поверхностей, а не "гладких" технологий.

Обработанные графически данные, однако, непросто исследовать без хорошо обоснованных и отработанных методов классификации и анализа строения объектов и структур. Это относится в первую очередь к парадигме представления данных, подлежащих последующему анализу. Об этом свидетельствует практика оформления данных в форме разнообразных таблиц.

Изогеометрические структуры: формы и проявления

Как правило, таблицы с данными отражают только изменения нулевого уровня признаков. На этом уровне фиксируются только такие изменения в структуре объектов или процессов, которые можно отразить лишь с помощью знака и абсолютной величины значения изменяющегося признака. Для объекта подобные изменения фиксируются в некотором пространстве (физическом, организационном, виртуальном и т.д.), для процесса – во времени. Здесь важнейшей характеристикой является уровень отсчета, от (или вокруг) которого отклоняются (колеблются) значения признаков. Структура на нулевом уровне определяется соотношением двух соседних по таблице значений признаков. Графически же такое изменение можно представить с помощью некоторого графика, где табличные значения располагаются вдоль некоторой кривой, в некотором отдалении друг от друга. Разумеется, подобное представление структурных изменений является наиболее простым и экономным. Однако такая простота и лаконичность дает, к сожалению, минимально возможное количество информации для

первичного структурно-изометрического анализа данных, который не предполагает какой-либо предварительной обработки данных.

Если же подвергнуть некоторой обработке данные нулевого уровня, то можно получить скрытую информацию, отражающую более глубокие и устойчивые структурные характеристики, позволяющие более полно представлять детали строения объектов или процессов. Традиционной формой первичной изометрической обработки данных является вычисления средних значений или разделенных разностей. Средние значения отражают соответственно изменения первого уровня объектов или процессов. Они могут быть представлены непосредственно таблицей, и тогда первичная обработка предполагает восстановление данных нулевого уровня.

С помощью разделенных разностей появляется возможность фиксировать детали структуры объектов или процессов. Такая возможность реализуется с помощью сравнения двух соседних разностей, через которое представляется интенсивность (плотность) изменения признака. Естественно, для такого представления необходимы три значения данных в таблице, если мы имеем дело только с исходной таблицей данных нулевого уровня.

Известно, что для тех случаев, когда значения признака изменяются непрерывно на всем промежутке (интервале) или области, где производится измерение, учет или фиксация значений признака, среднее значение представляет некоторую скорость течения изменений во внутренней точке промежутка (области). Но реально обосновать попытку приписать значение этой скорости какой-либо конкретной точке промежутка (области) мы не можем. Зато мы можем сравнивать соседние разделенные разности как средние скорости перемен и из этого сравнения с определенной точностью делать заключение о возможных скоростях на границах интервалов (областей). Кроме того, разделенные разности позволяют судить нам о характере изменений в целом на каждом интервале. Конкретные же структурные изменения внутри каждого промежутка на первом уровне нам, однако, недоступны.

При исходных данных, заданных в таблице нулевого уровня, можно, очевидно, выходить на более высокий уровень с помощью вычисления разделенных разностей более высоких степеней (вторых, третьих и т.д.). Но и как первые разделенные разности, разделенные разности более высоких степеней, давая представление о некоторых деталях изометрической структуры каждого интервала (области) в целом, не дают основания судить о деталях их внутреннего строения.

Для этих целей требуется дополнительно к данным нулевого уровня принципиально другая информация, которая, как и информация нулевого уровня, характеризовала бы изменения не в среднем, как разделенные разности, а в отдельных точках (целесообразно в тех же, в которых задаются исходные данные нулевого уровня). Для общности будем считать, что ранг уровня этой информации будет равен рангу разделенной разности.

Допустим теперь, нам заданы таблицей не только изменения какого-либо признака (информация нулевого уровня), но и скорость такого изменения в тех же точках ("мгновенные скорости"). Имея значения производной (скорости) на концах интервала и значения разделенной разности для этого интервала как среднего значения производной во внутренней точке, мы можем уже иметь представление о том, есть ли у нас условия для каких-либо структурных особенностей строения объекта или процесса (монотонности, выпуклости и т.д.) или нет.

Еще больше информации дают дополнительные данные о "мгновенных скоростях" более высокого порядка. На их основе мы можем судить о "сгущениях" или "разреженностях" качеств объекта или сменах интенсивности процессов и о возможных точках таких особенностей внутри интервалов. Это еще более "содержательная" информация, так как она позволяет более полно, в сравнении с информацией более низкого ранга, фиксировать детали структуры.

Однако для получения подобной комплексной информации требуется хорошо отработанная технология. Потребность в подобной технологии сейчас весьма велика, ибо многие не только гуманитарные, но и физические, технические и социально-

экономические задачи, целью которых является выявление всякого рода тонких структур, пока еще не имеют необходимых для этого инструментов.

В настоящем отчете приводится один класс методов изогеометрической обработки данных, позволяющих интерполировать, сохраняя изогеометрические свойства, не только гладкие функции, но и их первые и вторые производные. Этот класс включает методы локальных параболических и кубических сплайнов с дополнительными узлами.

3.2.2. ПАРАБОЛИЧЕСКИЕ СПЛАЙНЫ

Сохранение знака

Пусть на исходной сетке:

$$\Delta: x_0 = a < x_1 < \dots < x_n = b \quad (1)$$

заданы значения интерполируемой функции $f(x) \in C^1[a, b]$ и ее производной $f'(x)$:

$$f_i = f(x_i), f'_i = f'(x_i), x_i \in \Delta. \quad (2)$$

Предлагается строить подобный сплайн с помощью одного дополнительного узла в середине каждого интервала в предположении, что:

1) сплайн гарантированно сохраняет знак на каждом интервале в том случае, когда значения интерполируемой функции на концах этого интервала не имеют разных знаков:

$$f_i f_{i+1} \geq 0, i = 0, \dots, n-1. \quad (3)$$

2) на исходной сетке заданы только значения функции $f(x)$, а производную в узлах этой сетки требуется предварительно аппроксимировать с необходимой точностью;

3) во всех интервалах дополнительные узлы строятся только в их середине, что существенно упрощает алгоритм, сокращает объемы вычислений, повышает их скорость и точность.

Аппроксимацию производной сплайна в узлах сетки Δ будем осуществлять традиционно, а затем корректировать вычисленные значения производной в том случае, когда они не обеспечивают возможность сохранения знака функции во внутренних точках интервалов при ее интерполяции с помощью локального параболического сплайна с дополнительными узлами в середине интервалов.

Традиционный алгоритм аппроксимации производной во внутренних узлах сетки Δ таков:

$$S'_i = S'(x_i) = \frac{f[x_{i-1}, x_i]h_i + f[x_i, x_{i+1}]h_{i-1}}{h_{i-1} + h_i}, i = 1, \dots, n-2, \quad (4)$$

где

$$h_i = x_{i+1} - x_i$$

$$f[x_i, x_{i+1}] = \frac{f_{i+1} - f_i}{h_i}, i = 0, \dots, n-1. \quad (5)$$

Выбор производных S'_0, S'_n на концах сетки Δ будем делать, исходя из эвристических соображений и смысла решаемой задачи.

При внесении поправок в значения производных во внутренних узлах сплайна с целью сохранения знака во всех интервалах, в которых выполняются условия (3), будем руководствоваться следующими правилами, которые опираются на

Утверждение 1. Пусть в условиях (3) вместо (4)

$$S'_i = -4 \frac{f_i}{h_i}, i = 1, \dots, n-2, \quad (6)$$

если $f_i f_{i+1} > 0, f_i (4f_i + h_i S'_i) < 0, S'_i S'_{i+1} < 0$;
или

$$S'_i = 4 \frac{f_i}{h_{i-1}}, i = 1, \dots, n-2, \quad (7)$$

если $f_{i-1}f_i > 0$, $f_i(4f_i - h_{i-1}S'_i) < 0$, $S'_{i-1}S'_i < 0$.

Тогда локальный параболический сплайн с дополнительными узлами в середине интервалов

$$S_j(x) = S_{ij}(x) = a_{ij}(x - x_{ij})^2 + b_{ij}(x - x_{ij}) + c_{ij}, \\ x \in [x_{ij}, x_{i,j+1}], j = 0, 1, \quad (8)$$

$$x_{i0} = x_i, x_{i1} = x_i + \frac{1}{2}h_i, x_{i2} = x_{i+1}, i = 0, \dots, n-1,$$

такой, что

$$S_{i0}(x_{i0}) = f_i, S'_{i0}(x_{i0}) = S'_i,$$

$$S_{i0}(x_{i1}) = S_{i1}(x_{i1}), S'_{i0}(x_{i1}) = S'_{i1}(x_{i1}),$$

$$S_{i1}(x_{i2}) = f_{i+1}, S'_{i1}(x_{i2}) = S'_{i+1},$$

сохраняет знак в тех интервалах $[x_i, x_{i+1}]$, $i = 0, \dots, n-1$, где соблюдаются необходимые условия (3).

Сохранение монотонности

Для сохранения монотонности предлагается наглядная формулировка, которая будет использована далее при исследовании проблемы сохранения кусочной выпуклости локального кубического сплайна с дополнительными узлами.

Будем рассматривать локальный параболический сплайн с одним или двумя дополнительными узлами

$$S(x_j) = S(x_{ij}) = a_{ij}(x - x_{ij})^2 + b_{ij}(x - x_{ij}) + c_{ij}, \quad (9)$$

$$x \in [x_{ij}, x_{i,j+1}], j = 0, 1, 2,$$

$$x_{i0} = x_i < x_{i1} = x_i + k_{i1}h_i \leq x_{i2} = x_{i+1},$$

$$0 < k_{ij} < 1, j = 1, 2,$$

$$0 < k_{i1} + k_{i2} \leq 1.$$

Здесь при $k_{i1} + k_{i2} = 1$ строится один дополнительный узел, иначе два.

Очевидно, если

$$f(x) \in C^1[a, b] \text{ монотонна на интервале } [x_i, x_{i+1}], i = 0, \dots, n-1,$$

то

$$f'(x_j) f[x_i, x_{i+1}] \geq 0, f'(x_{i+1}) f[x_i, x_{i+1}] \geq 0, \text{ если } f''[x_i, x_{i+1}] \neq 0, \quad (10)$$

$$f'(x_j) = f'(x_{i+1}) = f[x_i, x_{i+1}] = 0, \quad \text{иначе.} \quad (11)$$

Утверждение 2. Для того чтобы в условиях (10) сплайн (9) сохранял монотонность на интервале $[x_i, x_{i+1}]$, $i = 0, \dots, n-1$, необходимо и достаточно, чтобы

$$\frac{k_{i1}f'_i + k_{i2}f'_{i+1}}{f[x_i, x_{i+1}]} \leq 2. \quad (12)$$

Справедливость утверждения 2 становится очевидной, если принять во внимание свойства разделенной разности $f[x_i, x_{i+1}]$ как среднего значения на интервале.

3.2.3. КУБИЧЕСКИЕ СПЛАЙНЫ: сохранение монотонности

Локальная монотонность и выпуклость (вогнутость) кубических полиномов

Поскольку кубический сплайн представляет собой цепочку, построенную из кубических полиномов, то предварительно целесообразно исследовать локальную монотонность и локальную выпуклость (вогнутость) этих звеньев.

Отметим, что если некоторая гладкая функция $f(x)$ (нестрого) монотонна на отрезке $[x_1, x_2]$, $x_1 \neq x_2$, то

$$\frac{f'(x_1)}{f[x_1, x_2]} \geq 0, \quad \frac{f'(x_2)}{f[x_1, x_2]} \geq 0, \text{ если } f[x_1, x_2] \neq 0, \quad (1)$$

$$f'(x_1) = f'(x_2) = 0, \text{ иначе.} \quad (1')$$

где $f[x_1, x_2]$ - разделенная разность: $\frac{f(x_2) - f(x_1)}{h}$,

$h = x_2 - x_1$ - длина промежутка $[x_1, x_2]$.

Причем, если функция $f(x) \in C^2$ (нестрого) монотонна на отрезке $[x_1, x_2]$, то при $f[x_1, x_2] = 0$ помимо (1')

$$f''(x_1) = f''(x_2) = 0.$$

Если некоторая гладкая функция $f(x)$ (нестрого) выпукла (вогнута) на отрезке $[x_1, x_2]$, $x_1 \neq x_2$, то

$$f''(x_1) f''(x_2) \geq 0 \text{ или} \quad (2)$$

$$f'(x_1) \leq f[x_1, x_2] \leq f'(x_2) \quad (f'(x_1) \geq f[x_1, x_2] \geq f'(x_2)). \quad (2')$$

Пусть в точках x_1, x_2 ($x_1 < x_2$) будут известны значения $f(x_1)$ и $f(x_2)$ полинома третьей степени

$$f(x) = a(x - x_1)^3 + b(x - x_1)^2 + c(x - x_1) + d \quad (3)$$

и значения его первой производной $f'(x_1)$ и $f'(x_2)$.

Тогда через эти значения можно однозначно выразить его коэффициенты a, b, c, d и его вторые производные на концах интервала $f''(x_1)$ и $f''(x_2)$.

Имеем:

$$a = -\frac{2f[x_1, x_2] - f'(x_1) - f'(x_2)}{h^2}, \quad b = \frac{3f[x_1, x_2] - 2f'(x_1) - f'(x_2)}{h},$$

$$c = f'(x_1), \quad d = f(x_1), \quad f''(x_1) = 2 \frac{3f[x_1, x_2] - 2f'(x_1) - f'(x_2)}{h},$$

$$f''(x_2) = -2 \frac{3f[x_1, x_2] - f'(x_1) - 2f'(x_2)}{h}.$$

О поведении полинома (3) на промежутке $[x_1, x_2]$ будем судить по поведению его производной

$$f'(x) = 3a(x - x_1)^2 + 2b(x - x_1) + c.$$

Покажем, что наличие или отсутствие монотонности полностью определяется соотношением между $f[x_1, x_2]$, $f'(x_1)$ и $f'(x_2)$.

Утверждение. Пусть для полинома (3) в точках x_1, x_2 , $x_1 < x_2$, выполнены условия (1), (1'). Для того, чтобы полином сохранял монотонность на всем промежутке $[x_1, x_2]$, необходимо и достаточно, чтобы

$$f[x_1, x_2] \geq \frac{1}{3}(f'(x_1) + f'(x_2) - \sqrt{f'(x_1)f'(x_2)}), \text{ если } f[x_1, x_2] \geq 0, \quad (4)$$

$$f[x_1, x_2] \leq \frac{1}{3}(f'(x_1) + f'(x_2) + \sqrt{f'(x_1)f'(x_2)}), \text{ иначе.} \quad (4')$$

Для доказательства утверждения отметим, что в условиях (1)-(1') полином $f(x)$ будет монотонным на промежутке $[x_1, x_2]$ в трех случаях (см. рис. 1):

1) $a \cdot f[x_1, x_2] \leq 0$ (на рисунке эта область ограничена снизу прямой, соединяющей точки $(x_1, f'(x_1))$, $(x_2, f'(x_2))$). Здесь при $a \cdot f[x_1, x_2] \neq 0$ вершина параболы $f'(x)$ лежит выше $f[x_1, x_2] > 0$ или ниже $f[x_1, x_2] < 0$. При $f[x_1, x_2] = 0$ парабола вырождается в прямую;

2) $f''(x_1) f''(x_2) \geq 0$ (на рисунке эта область ограничена снизу и сверху тонкими дугами парабол). Здесь производная $f'(x)$ монотонна на промежутке $[x_1, x_2]$. Соответственно, полином (3) в этом промежутке является выпуклым (вогнутым);

3) $b^2 - 3ac \leq 0$ (на рисунке эта область ограничена снизу и сверху жирными дугами парабол). Это критическая область, в которой производная $f'(x)$ не имеет некрратных действительных корней. Причем, кратный корень для положительной разделенной разности имеет только нижняя граница. Зеркальная картина наблюдается для отрицательной разделенной разности.

Очевидно, неравенству 1) эквивалентны неравенства:

$$f[x_1, x_2] \geq \frac{1}{3}(f'(x_1) + f'(x_2)), \text{ если } f[x_1, x_2] \geq 0,$$

$$f[x_1, x_2] \leq \frac{1}{3}(f'(x_1) + f'(x_2)), \text{ иначе.}$$

Соответственно случай 2) характеризуют неравенства

$$(f[x_1, x_2] - \frac{1}{2}(f'(x_1) + f'(x_2)))^2 \leq (\frac{1}{6}(f'(x_2) - f'(x_1)))^2,$$

если вместо $f''(x_1)$, $f''(x_2)$ подставить их выражения через $f[x_1, x_2]$, $f'(x_1)$, $f'(x_2)$, h , или

$$\frac{2}{3}f'_{\min} + \frac{1}{3}f'_{\max} \leq f[x_1, x_2] \leq \frac{1}{3}f'_{\min} + \frac{2}{3}f'_{\max},$$

где

$$f'_{\min} = \min(f'(x_1), f'(x_2)),$$

$$f'_{\max} = \max(f'(x_1), f'(x_2)).$$

Наконец, нетрудно убедиться, что для случая 3) справедливо неравенство

$$(3 f[x_1, x_2] - f'(x_1) - f'(x_2))^2 \leq f'(x_1) f'(x_2),$$

если вместо коэффициентов a , b , c подставить их выражения через $f[x_1, x_2]$, $f'(x_1)$, $f'(x_2)$, h , или

$$\frac{1}{3}(f'(x_1) + f'(x_2) - \sqrt{f'(x_1)f'(x_2)}) \leq f[x_1, x_2] \leq \frac{1}{3}(f'(x_1) + f'(x_2) + \sqrt{f'(x_1)f'(x_2)}).$$

Очевидно, случаи 1) и 2) частично перекрывают друг друга. Несложно показать, что и случаи 2) и 3) тоже перекрываются, но уже с другой стороны. Объединяя результаты, мы приходим к заключению о справедливости критерия (4)-(4').

Замечание 1. В условиях (1), (1') нижняя (для $f[x_1, x_2] \geq 0$) или верхняя (для $f[x_1, x_2] \leq 0$) граница монотонности полинома колеблется в пределах

$$\frac{1}{4} f'_{\max} \leq \frac{1}{3} (f'(x_1) + f'(x_2) - \sqrt{f'(x_1)f'(x_2)}) \leq \frac{1}{3} f'_{\max}, \text{ если } f[x_1, x_2] \geq 0,$$

$$\frac{1}{4} f'_{\min} \geq \frac{1}{3} (f'(x_1) + f'(x_2) + \sqrt{f'(x_1)f'(x_2)}) \geq \frac{1}{3} f'_{\min}, \text{ иначе,}$$

причем, при равенстве

$$f'_{\max} = 3 f[x_1, x_2] - \frac{1}{2} f'_{\min} + \sqrt{3 f'_{\min} (f[x_1, x_2] - \frac{1}{4} f'_{\min})}, \text{ если } f[x_1, x_2] \geq 0,$$

$$f'_{\min} = 3 f[x_1, x_2] - \frac{1}{2} f'_{\max} - \sqrt{3 f'_{\max} (f[x_1, x_2] - \frac{1}{4} f'_{\max})}, \text{ иначе.}$$

Замечание 2. Из условий (4)-(4') вытекает, что для того чтобы в условиях (1), (1') полином (3) сохранял монотонность на всем промежутке $[x_1, x_2]$, необходимо, чтобы

$$f[x_1, x_2] \geq \frac{1}{6} (f'(x_1) + f'(x_2)), \text{ если } f[x_1, x_2] \geq 0, \quad (5)$$

$$f[x_1, x_2] \leq \frac{1}{6} (f'(x_1) + f'(x_2)), \text{ иначе.} \quad (5')$$

Эрмитовы сплайны

При построении сплайнов обычно возникают два класса задач:

1) задачи выявления структуры данных, фиксируемых и описываемых методами сплайн-интерполяции (задачи анализа);

2) задачи поиска возможностей сохранения с помощью сплайн-интерполяции структуры данных (задачи реконструкции).

Сформулированные нами точные и приближенные критерии локальной выпуклости и монотонности кубических полиномов являются мощным инструментом для решения задач обоих классов. Для задач первого класса эти критерии являются способом, с помощью которого выясняется вопрос о существовании тех или иных структурных свойств и качеств, присущих анализируемому данным. Для задач второго класса этот инструмент служит способом обоснования адекватности методов реконструкции действительной сущности восстанавливаемых объектов.

Имея это в виду, рассмотрим последовательно методологические схемы изометрической сплайн-аппроксимации, начиная от простых схем и переходя к более сложным и интересным.

Наиболее простой формой сплайн-аппроксимации является сплайн-интерполяция эрмитовыми сплайнами непрерывных таблично заданных функций, принадлежащих к классу C^1 .

Пусть в узлах сетки Δ : $a = x_0 < x_1 < \dots < x_n = b$ заданы значения функции $f(x) \in C^1[a, b]$ и ее (первая) производная:

$$f_i = f(x_i), f'_i = f'(x_i), x_i \in \Delta.$$

Обозначим шаг сетки и разделенные разности соответственно через:

$$h_i = x_{i+1} - x_i,$$

$$f[x_i, x_{i+1}] = \frac{f_{i+1} - f_i}{h_i}.$$

Для интерполяции $f(x)$ в каждом интервале $[x_i, x_{i+1}]$, $i = 1, \dots, n-1$, строим кубический полином подобно (3):

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i,$$

$$x \in [x_i, x_{i+1}], i = 0, \dots, n-1. \quad (6)$$

Очевидно, монотонность сплайна в каждом интервале $[x_i, x_{i+1}]$, $i = 0, \dots, n-1$, полностью определяется с помощью критерия (4)-(4'). Выяснение этого является наиболее интересным вариантом задач первого класса.

Переходя к задачам второго класса, допустим, что нам заданы только значения интерполируемой функции в узлах сетки Δ . Задача интерполяции состоит в том, чтобы найти такие значения производных эрмитова сплайна, которые обеспечивали бы сохранение монотонности интерполируемой функции в тех интервалах, где имеются необходимые условия (1)-(1').

Расчет производных в узлах сетки Δ целесообразно делать в два шага.

На первом шаге производным присваиваются значения, обеспечивающие сплайну высокую сходимость и точность. Для этих целей обычно используется формула:

$$S'_i = S'_i(x_i) = \frac{f[x_{i-1}, x_i] \cdot h_i + f[x_i, x_{i+1}] \cdot h_{i-1}}{h_{i-1} + h_i}, i = 1, \dots, n-1. \quad (7)$$

Выбор значений S'_0, S'_n будем делать, исходя из эвристических соображений и смысла решаемой задачи.

На следующем шаге будем использовать рассчитанные по правилам (7) производные в качестве начального приближения, делая затем необходимые поправки, обеспечивающие сплайну (6) монотонность в тех интервалах, где имеются условия (1)-(1').

При внесении поправок в значения S'_i , $i = 1, \dots, n-1$, наиболее желательным является использование критерия (4)-(4') с учетом замечания 1. Очевидно, для расчета этих поправок нужен, по крайней мере, двухпроходный алгоритм, поскольку приходится раздельно рассматривать ограничения на величину производных в смежных узлах. Для построения более быстрых алгоритмических схем можно применять упрощенный критерий.

Исходя из этого, положим:

$$S'_i = \begin{cases} 3f[x_{i-1}, x_i], & \text{если } \frac{S'_i}{f[x_{i-1}, x_i]} > 3, \\ 3f[x_i, x_{i+1}], & \text{если } \frac{S'_i}{f[x_i, x_{i+1}]} > 3, \end{cases} \quad i = 1, \dots, n-1, \quad (8)$$

$$S'_0 = 3f[x_0, x_1], \text{ если } \frac{S'_0}{f[x_0, x_1]} > 3, \quad (8')$$

$$S'_n = 3 f[x_{n-1}, x_n], \text{ если } \frac{S'_n}{f[x_{n-1}, x_n]} > 3. \quad (8'')$$

Очевидно, подобная мера гарантирует сохранение монотонности для эрмитова сплайна $S_i(x)$ там, где необходимые условия (1)-(1') имеются.

Теперь введем еще одно ограничение на производные S'_i :

$$\text{если } f[x_{i-1}, x_i] f[x_i, x_{i+1}] \leq 0, \text{ то } S'_i = 0. \quad (9)$$

Кроме того, будем считать, что

$$\text{если } S'_0 f[x_0, x_1] \leq 0, \text{ то } S'_0 = 0, \quad (9')$$

$$\text{если } S'_n f[x_{n-2}, x_{n-1}] \leq 0, \text{ то } S'_n = 0. \quad (9'')$$

Тогда в каждом интервале может быть построен (нестрого) монотонный сплайн.

Будем называть сплайн, монотонный на каждом интервале, но не обязательно монотонный на смежных интервалах, кусочно монотонным. Таким образом нами доказано

Утверждение. (Существование кусочно-монотонного эрмитова сплайна). Для любой таблично заданной функции можно построить кусочно монотонный эрмитов сплайн (6).

Замечание. Следует отметить высокую степень сходства описанного сплайна с кусочно-линейной интерполяцией таблично заданной функции. Как и при аппроксимации подобных функций ломаными (для функций от одного аргумента) и плоскими прямоугольными участками (для функций от нескольких аргументов) все локальные экстремумы находятся в узлах исходной сетки, а не внутри интервалов. При аппроксимации же обычными (основными) сплайнами значения функций в узлах исходной сетки не всегда являются локальными экстремумами, а часто оказываются на "излете" кривой или поверхности сплайна. Это свойство монотонного эрмитова сплайна придает ему высокую графическую значимость, так как подобное гладкое представление таблично заданной функции является более наглядным, нежели ломаная.

Очевидно, если выполняются условия (2') на некотором интервале $[x_i, x_{i+1}]$, $i = 1, \dots, n-1$, то сплайн (6) на этом интервале будет выпуклым или вогнутым.

Локальные сплайны с дополнительными узлами

Разумеется, с помощью локальных эрмитовых сплайнов можно строить только функции, принадлежащие к классу C^1 . Для выхода на построение сплайнов класса C^2 , гарантирующих сохранение монотонности, требуется приравнять вторые производные в узлах сетки слева и справа, ликвидируя возможный разрыв между ними. Пока еще нет доказательств того, что это всегда достижимо для нестрогой монотонности.

Есть однако другой путь решения этой проблемы: использовать локальные сплайны с дополнительными узлами.

Разработаны две схемы вычислений. В первой схеме используются ограничения на первые и вторые производные в узлах исходной сетки. Во второй схеме эти ограничения снимаются.

Вначале рассматривается первая схема. Отметим, что введение дополнительных узлов внутри сетки Δ дает возможность разделить процессы вычисления первых (и вторых) производных и расчеты коэффициентов сплайна. Это разделение позволяет фиксировать изогеометрическую ситуацию в каждом интервале и строить алгоритмы по ее сохранению в той мере, в какой она зафиксирована в узлах исходной сетки. Кроме того, дополнительные узлы позволяют задавать легко интерпретируемое аналитическое представление интерполируемой функции.

Определение.

Пусть в узлах сетки Δ : $a = x_0 < x_1 < \dots < x_n = b$ дополнительно к значениям функции $f(x) \in C^2[a, b]$ и ее первой производной $f'_i = f'(x_i)$, $f''_i = f''(x_i)$, $x_i \in \Delta$, заданы значения второй производной $f''_i = f''(x_i)$, $x_i \in \Delta$.

Пусть в каждом интервале сетки Δ будут теперь задано по два дополнительных узла x_{i1}, x_{i2} , $[x_i < x_{i1} < x_{i2} < x_{i+1}]$, $i = 0, \dots, n-1$.

Обозначим $x_{i0} = x_i$, $x_{i3} = x_{i+1}$, $i = 0, \dots, n-1$.

Будем называть локальным кубическим сплайном с дополнительными узлами x_{i1}, x_{i2} , $i = 0, \dots, n-1$, сплайн

$$S_i(x) = S_{ij}(x) = a_{ij}(x - x_{ij})^3 + b_{ij}(x - x_{ij})^2 + c_{ij}(x - x_{ij}) + d_{ij}, \quad (10)$$

$$x \in [x_{ij}, x_{i,j+1}], j = 0, 1, 2,$$

такой, что

$$S_{i0}(x_{i0}) = f_i, S'_{i0}(x_{i0}) = f'_i, S''_{i0}(x_{i0}) = f''_i,$$

$$S_{ij}(x_{i,j+1}) = S_{i,j+1}(x_{i,j+1}), S'_{ij}(x_{i,j+1}) = S'_{i,j+1}(x_{i,j+1}),$$

$$S''_{ij}(x_{i,j+1}) = S''_{i,j+1}(x_{i,j+1}), j = 0, 1,$$

$$S_{i3}(x_{i3}) = f_{i+1}, S'_{i3}(x_{i3}) = f'_{i+1}, S''_{i3}(x_{i3}) = f''_{i+1},$$

где $S'_{ij}(x)$, $S''_{ij}(x)$ – производные сплайна (10) в точке $x \in [x_{ij}, x_{i,j+1}]$, $j = 0, 1$,

2.

Обозначим:

$$S''_i = S''_i(x_i), i = 0, 1, \dots, n.$$

Изогеометрические свойства интерполируемой функции будем поддерживать таким образом, чтобы в каждом подинтервале выполнялись ограничения типа (4)-(4') или (8)-(8''), или же соответствующими смещениями дополнительных узлов относительно концов интервалов.

Для сохранения монотонности достаточно использовать наиболее удобные в вычислительном отношении равноотстоящие узлы, т.е.

$$x_{i1} = x_i + \frac{1}{3} h_i, x_{i2} = x_{i+1} + \frac{2}{3} h_i.$$

Чтобы выяснить условия, при которых обеспечивается монотонность сплайна (10) в соответствующих интервалах, целесообразно сначала вычислить значения первой и второй производной сплайна в дополнительных узлах через $f[x_i, x_{i+1}]$, S'_i , S'_{i+1} , S''_i , S''_{i+1} , h_i , используя последние в качестве параметров. Через значения производных в узлах исходной сетки и в дополнительных узлах легко вычислить и коэффициенты сплайна для подинтервалов.

Теперь допустим, что на сетке Δ заданы только значения функции $f(x) \in C^2[a, b]$. Аппроксимацию ее первой и второй производной сплайна (10) в узлах сетки Δ будем осуществлять таким образом, чтобы, не теряя точности, сохранять монотонность сплайна на тех интервалах, где это возможно.

Устанавливая ограничения на S'_i и S''_i по аналогии с эрмитовым сплайном, будем вначале придерживаться того, чтобы

$$f[x_{i-1}, x_i] \leq S'_i \leq f[x_i, x_{i+1}], S''_i \geq 0, \quad (11)$$

если $f[x_{i-1}, x_i] < f[x_i, x_{i+1}]$, или

$$f[x_{i-1}, x_i] \geq S'_i \geq f[x_i, x_{i+1}], S''_i \leq 0, \quad (12)$$

если $f[x_{i-1}, x_i] > f[x_i, x_{i+1}]$.

Как и для рассмотренных ранее эрмитовых сплайнов, при аппроксимации производных сплайна (10) в узлах исходной сетки Δ будем использовать рассчитанные из соображений (11)-(12) производные S'_i , S''_i , $i = 0, 1, \dots, n$, в качестве начального

приближения. Если окажется, что вычисленные таким образом производные сплайна (10) не будут удовлетворять условиям, достаточным, чтобы он был монотонным, то будем искать вносить необходимые поправки.

Для этих целей будем исследовать свойства сплайна (10) в каждом интервале $[x_i, x_{i+1}]$, $i = 1, \dots, n-1$, исходя из поведения его первой производной, представленной набором попарно сопряженных парабол.

Целью анализа является выявление условий сохранения знака каждой из этих парабол и точек их склейки. Для анализа будем применять критерий 8)-(8''), поскольку более мощный критерий (4)-(4') использовать для подинтервалов весьма затруднительно.

Обозначим:

$$P_{i1} = S'_i + \frac{1}{6} h_i S''_i,$$

$$P_{i2} = S'_{i+1} - \frac{1}{6} h_i S''_{i+1},$$

$$Q_i = f[x_i, x_{i+1}] - \frac{1}{54} (S''_{i+1} - S''_i) h_i$$

Утверждение 3. (Достаточные условия монотонности сплайна с дополнительными узлами). Для того, чтобы сплайн с дополнительными узлами (10) был монотонен на интервале $[x_i, x_{i+1}]$ в условиях (1)-(1'), достаточно, чтобы одновременно

$$f[x_i, x_{i+1}] P_{i1} \geq 0, \quad (13)$$

$$f[x_i, x_{i+1}] P_{i2} \geq 0, \quad (14)$$

а также

$$f[x_i, x_{i+1}] (Q_i - \frac{1}{4} (P_{i1} + P_{i2})) \geq 0, \quad (15)$$

если $\frac{1}{3} \leq \frac{P_{i1}}{P_{i2}} \leq 3$, или

$$f[x_i, x_{i+1}] (Q_i - \frac{1}{3} P_{i1}) \geq 0, \quad (16)$$

$$f[x_i, x_{i+1}] (Q_i - \frac{1}{3} P_{i2}) \geq 0, \text{ иначе.} \quad (17)$$

Для доказательства утверждения отметим, что в условиях (1)-(1') критерий (13)-(17) гарантирует первым производным в дополнительных узлах $S'_i(x_{i1})$, $S'_i(x_{i2})$ и разделенным разностям для каждого подинтервала $[x_{ij}, x_{i,j+1}]$, $j = 0, 1, 2$, знак разделенной разности $f[x_i, x_{i+1}]$ и, кроме того, обеспечивает для каждого подинтервала $[x_{ij}, x_{i,j+1}]$, $j = 0, 1, 2$, выполнение критерия (8)-(8'').

Отметим с удовлетворением, что ограничения для первых производных в узлах исходной сетки для сплайна (10), налагаемые критерием (13)-(17), существенно слабее, нежели соответствующие ограничения для эрмита сплайна (8)-(8'') и даже (4)-(4'). Однако при этом, очевидно, условия (13)-(17) определяют совместные ограничения на значения первых и вторых производных в узлах исходной сетки.

Следует заметить, что условия (15)-(17) можно разделить так, чтобы одна часть из них относилась к соотношениям между разделенной разностью и производными правого узла, а вторая часть - к соотношениям между разделенной разностью и производными левого узла. Тогда при разработке алгоритма интерполяции можно рассматривать каждый i -й узел сетки Δ отдельно, независимо от остальных, учитывая

только соседние разделенные разности $f[x_{i-1}, x_i]$ и $f[x_i, x_{i+1}]$, и строить на их основе ограничения для каждого узла слева и справа независимо.

Для этой цели разделим на две части:

$$Q_{i1} = \frac{1}{2} (f[x_i, x_{i+1}] + \frac{1}{54} h_i S''_i),$$

$$Q_{i2} = \frac{1}{2} (f[x_i, x_{i+1}] - \frac{1}{54} h_i S''_{i+1}),$$

$$Q_i = Q_{i1} + Q_{i2}.$$

Тогда условия (13)-(17) в утверждении 3 можно заменить на следующие:

$$f[x_i, x_{i+1}] P_{i1} \geq 0, \quad (18)$$

$$f[x_i, x_{i+1}] Q_{i1} \geq 0, \quad (19)$$

$$f[x_i, x_{i+1}] (Q_{i1} - \frac{1}{3} P_{i1}) \geq 0, \quad (20)$$

$$f[x_i, x_{i+1}] P_{i2} \geq 0, \quad (21)$$

$$f[x_i, x_{i+1}] Q_{i2} \geq 0, \quad (22)$$

$$f[x_i, x_{i+1}] (Q_{i2} - \frac{1}{3} P_{i2}) \geq 0. \quad (23)$$

Нетрудно убедиться, что тогда в условиях (8)-(8'') всегда можно подобрать такие производные в узлах исходной сетки для соответствующих интервалов, которые будут удовлетворять перечисленным ограничениям.

Теперь допустим, что для сплайна (10) выполняются ограничения (9)-(9'').

Тогда, очевидно, в каждом интервале может быть построен (нестрого) монотонный сплайн. Таким образом нами доказано

Утверждение 4. (Существование монотонного кубического сплайна с дополнительными узлами). Для любой таблично заданной функции можно построить кусочно монотонный кубический сплайн с дополнительными узлами (10).

Как и эрмитов сплайн (6) с ограничениями (9')-(9''), описанный сплайн напоминает кусочно-линейную интерполяцию таблично заданной функции. Здесь справедливо замечание 3, относящееся к эрмитову сплайну.

Замечание 4. Утверждения 3, 4 позволяют строить алгоритмы монотонной интерполяции таблично заданной функции и аппроксимации ее первых и вторых производных. На первый взгляд, они должны быть громоздкими из-за необходимости построения в каждом интервале двух дополнительных узлов. Однако нетрудно убедиться, что легко построить схему вычислений, по сложности сравнимой с алгоритмами локальных В-сплайнов. Как и для последних, для расчета производных в узлах исходной сетки и в дополнительных узлах, а также соответствующих им коэффициентов сплайна для каждого внутреннего интервала нужны значения функции только для 4 соседних исходных узлов. Вследствие этого построенная на подобной основе вычислительная схема может быть не менее эффективной и экономной, нежели В-сплайны. Однако по сравнению с последними она, кроме подобия кусочно-линейной интерполяции, обладает рядом других не менее замечательных свойств, которые делают ее во многих случаях незаменимой. Наиболее интересными из них являются слабая реакция на неравномерность шага исходной сетки и слишком малые или большие перепады градиентов, а также слабая зависимость внутренних узлов от краевых условий.

Замечание 5. Использование дополнительных узлов позволяет строить локальные монотонные кубические сплайны не только в тех случаях, когда заданы табличные значения интерполируемых функций. Могут быть построены схемы интерполяции не только самих функций, но и их производных первого ($f'(x_i)$) и второго порядка $f''(x_i)$, i

$= 0, 1, \dots, n-1$. Однако здесь могут возникать критические ситуации, когда условия монотонности (13)-(17) могут не соблюдаться.

Отметим, что подобные критические ситуации для сплайнов, построенных на равноотстоящих дополнительных узлах, можно разделить на два класса.

Ситуации первого класса (не выполняются условия (13) или (14)) имеют место тогда, когда значения первых и вторых производных (с учетом шага) функции $f(x_i)$ в узлах исходной сетки сильно различаются по абсолютной величине. Для их преодоления достаточно сдвигать должным образом дополнительные узлы вправо или влево.

Ситуации второго класса (когда не выполняются условия (15), (16) или (14)) связаны с тем, разделенные разности оказываются существенно меньше по абсолютной величине, нежели первые производные в узлах исходной сетки (с поправкой, вносимой значениями вторых производных с учетом текущего шага исходной сетки). В этом случае обеспечить монотонность за счет сдвига дополнительных узлов в наиболее критических ситуациях вправо или влево не удастся. Например, если для некоторого интервала $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$, в условиях (1)-(1")

$$f[x_i, x_{i+1}] (f[x_i, x_{i+1}] - \frac{1}{6} (f'(x_i) + f'(x_{i+1})) - (f''(x_{i+1}) - f''(x_i)) h_i)) \leq 0,$$

$$f[x_i, x_{i+1}] (f[x_i, x_{i+1}] - \frac{1}{3} (f'(x_i) + f'(x_{i+1})) - \frac{1}{2} (f''(x_{i+1}) - f''(x_i)) h_i)) \leq 0,$$

$$f''(x_i) f[x_i, x_{i+1}] \leq 0, f''(x_{i+1}) f[x_i, x_{i+1}] \geq 0,$$

когда значения производных $f'(x_i)$ и $f'(x_{i+1})$ являются достаточно близкими, то для внутреннего подинтервала необходимые условия (5)-(5') не выполняются даже при любых сдвигах дополнительных узлов. Для того чтобы убедиться в этом, достаточно сделать не очень сложные вычисления.

Можно показать, что и при трех дополнительных узлах в этом случае проблема монотонности сплайна тоже не решается. И лишь при четырех дополнительного узлах в особо критических ситуациях можно строить монотонные сплайны. Здесь для компенсации относительно малых значений разделенной разности как среднего значения функции на интервале требуется вставлять в середину кривой сплайна кусок постоянной прямой (ступень), которому соответствует на уровне производной сплайна некоторый отрезок на оси абсцисс.

Выбор дополнительных узлов на интервалах, где имеются подобные критические ситуации и где условия (13), (14) выполняются как строгие неравенства, но не выполняются условия (15) или (16), (17), т.е.

$$f[x_i, x_{i+1}] P_{11} > 0, f[x_i, x_{i+1}] P_{12} > 0, f[x_i, x_{i+1}] \neq 0,$$

определяется соотношением знаков первой и второй производной функции в узлах сетки Δ . При этом левые два узла определяется соотношением знаков первой и второй производной функции на левом конце интервала, а правые два узла - соотношением знаков первой и второй производной функции на правом конце интервала соответственно.

Для каждой пары могут быть три типа ситуаций:

1) знаки совпадают; 2) знаки противоположны; 3) вторая производная равна нулю.

Нетрудно показать, что в ситуациях 1) и 2)

$$\text{при } \frac{2f[x_i, x_{i+1}]}{f'(x_i) + f'(x_{i+1}))} < \frac{1}{3}$$

существуют такие узлы

$x_{ij}, j = 1, 2, 3, 4, x_i < x_{i1} < x_{i2} < x_{i3} < x_{i4} < x_{i+1}, x_{i2} < x_i + \frac{1}{2} h_i; x_{i3} > x_i + \frac{1}{2} h_i$, при которых сплайн на интервале $[x_i, x_{i+1}]$ будет монотонным. В ситуациях типа 3) для этих целей можно обойтись и тремя дополнительными узлами.

3.2.4. КУБИЧЕСКИЕ СПЛАЙНЫ: сохранение выпуклости

Для выяснения условий, при которых сплайн (10) сохраняет выпуклость (вогнутость) интерполируемой функции на интервале $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$, отметим, что если $f(x)$ выпукла (вогнута) на этом интервале, то

$$S'_i < f[x_i, x_{i+1}] < S'_{i+1} \quad (S'_i > f[x_i, x_{i+1}] > S'_{i+1}), \quad (24)$$

$$S''_i f[x_i, x_{i+1}] \geq 0, \quad S''_{i+1} f[x_i, x_{i+1}] \geq 0, \quad (25)$$

если $f''[x_i, x_{i+1}] \neq 0$, или

$$S'_i = f[x_i, x_{i+1}] = S'_{i+1}, \text{ иначе.} \quad (26)$$

Очевидно, в условиях (26) сплайн (10) на интервале $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$ представляет собой отрезок прямой, соединяющий точки (x_i, f_i) , (x_{i+1}, f_{i+1}) , и не требует дополнительных узлов. Поэтому сосредоточим свое внимание на условиях (24)-(25). Здесь сохранение выпуклости (вогнутости) сплайном (10) означает сохранение знака его второй производной, в том числе и в точках, где она принимает среднее значение $f''[x_i, x_{i+1}]$.

Отметим, что для сохранения выпуклости сплайном на интервале $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$, необходимо выполнение условия, аналогичного (12) для параболического сплайна:

$$\frac{k_{i1} f''_i + k_{i2} f''_{i+1}}{f''[x_i, x_{i+1}]} \leq 2, \quad (27)$$

Равноотстоящие дополнительные узлы

Сначала рассмотрим схему сплайна (10) с равноотстоящими дополнительными узлами:

$$x_{i1} = x_i + \frac{1}{3} h_i, x_{i2} = x_i + \frac{2}{3} h_i. \text{ (здесь } k_{i1} = k_{i2} = 1/3).$$

Обозначим:

$$p_{i1} = \frac{S''_{i1}(x_{i1})}{f''[x_i, x_{i+1}]},$$

$$p_{i2} = \frac{S''_{i2}(x_{i2})}{f''[x_i, x_{i+1}]},$$

$$q_{i1} = \frac{1}{27} (5 p_{i1} + 2 p_{i2}),$$

$$q_{i2} = \frac{1}{27} (2 p_{i1} + 5 p_{i2}).$$

$$r_{i1} = \frac{f[x_i, x_{i+1}] - S'_i}{S'_{i+1} - S'_i},$$

$$r_{i2} = \frac{S'_{i+1} - f[x_i, x_{i+1}]}{S'_{i+1} - S'_i}.$$

$$t_{i1} = r_{i1} - \frac{1}{9} - q_{i1},$$

$$t_{i2} = r_{i2} - \frac{1}{9} - q_{i2}.$$

Утверждение 5. Для того чтобы в условиях (24), (25), (27) сплайн (10) с равноотстоящими узлами был выпуклым (вогнутым) на интервале $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$, необходимо и достаточно, чтобы одновременно

$$t_{i1} \geq 0, t_{i2} \geq 0, \quad (28)$$

$$q_{i1} \geq 0, q_{i2} \geq 0, \quad (29)$$

причем, если $q_{i1} q_{i2} \neq 0$,

$$\text{то } \frac{2}{3} \leq \frac{q_{i1}}{q_{i2}} \leq \frac{5}{2}, \quad (30)$$

$$\text{иначе } q_{i1} = q_{i2} = 0. \quad (31)$$

Для доказательства утверждения покажем, что условия (28)-(31) обеспечивают одновременно сохранение знака вторым производным сплайна во внутренних (дополнительных) узлах и на концах интервала.

Сначала отметим, что условия (28), (30), (31) обеспечивают сохранение знака вторым производным в дополнительных узлах. Чтобы убедиться в этом, заметим, проделав несложные вычисления, что

$$p_{i1} = \frac{9}{7} (5q_{i1} - 2q_{i2}),$$

$$p_{i2} = -\frac{9}{7} (2q_{i1} - 5q_{i2}).$$

Кроме того, условия (29) обеспечивают сохранение знака второй производной сплайна и на концах интервала. Для доказательства вычислим значения второй производной сплайна на концах интервала, используя уравнения связи (11). Имеем:

$$S''_i = \frac{54}{7} f''[x_i, x_{i+1}] t_{i1},$$

$$S''_{i+1} = \frac{54}{7} f''[x_i, x_{i+1}] t_{i2}.$$

Нетрудно убедиться, что при равноотстоящих дополнительных узлах возможны критические ситуации, в которых условия (28)-(31) обеспечивать не удастся. Кроме того, эти условия явно указывают точные границы для $f[x_i, x_{i+1}]$, в пределах которых может возникнуть возможность сохранения выпуклости сплайном (10). Эти пределы фиксирует

Утверждение 6. Для того, чтобы в условиях (24)-(25) сплайн (10) был выпуклым (вогнутым) на интервале $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$, необходимо, чтобы

$$\frac{8}{9} S'_i + \frac{1}{9} S'_{i+1} \leq f[x_i, x_{i+1}] \leq \frac{1}{9} S'_i + \frac{8}{9} S'_{i+1}, \text{ если } S'_i < S'_{i+1};$$

$$\frac{8}{9} S'_i + \frac{1}{9} S'_{i+1} \geq f[x_i, x_{i+1}] \geq \frac{1}{9} S'_i + \frac{8}{9} S'_{i+1}, \text{ иначе.}$$

Следует отметить, что условия (30)-(31) при $r_{i1} r_{i2} \neq 0$ накладывают несимметричные ограничения на вторые производные сплайна на концах интервала. Эти различия усиливаются по мере различий между r_{i1} и r_{i2} . Пусть, например, $r_{i1} < r_{i2}$. Для того чтобы выполнялись условия (28)-(31), необходимо

$$\frac{S''_{i+1}}{f''[x_i, x_{i+1}]} \geq \frac{S''_i}{f''[x_i, x_{i+1}]} + 27(\frac{1}{3} - p_{i1}),$$

что не всегда достижимо при равноотстоящих узлах даже в условиях утверждения 6.

Очевидно, если для каких-либо интервалов условия утверждений 5 не выполняются, то требуется корректировать схему сплайна (10). Здесь, вероятно, могут быть три пути для интервалов, где имеются критические ситуации:

- 1) использовать сдвиг дополнительных узлов;
- 2) расширить сетку Δ , вводя в нее промежуточные узлы;
- 3) строить более двух дополнительных узлов.

Покажем, что в некоторых случаях достаточно сдвига дополнительных узлов, хотя расширение сетки Δ и большее число дополнительных узлов в критических ситуациях может решить проблему полностью.

Подвижные дополнительные узлы

Пусть теперь дополнительные узлы в интервалах, где имеется критическая ситуация, будут подвижными:

$$x_{i1} = x_i + k_{i1}h_i, \quad x_{i2} = x_{i+1} - k_{i2}h_i \quad (32)$$

$$(0 < k_{i1} < 1, 0 < k_{i2} < 1, 0 < k_{i1} + k_{i2} < 1).$$

Сформулируем аналоги утверждений 4 и 5 применительно к подвижным узлам. С этой целью переобозначим параметры $p_{ij}, q_{ij}, j=1, 2$:

$$p_{i1} = (1 - k_{i1}) \frac{S''_{i1}(x_{i1})}{f'[x_i, x_{i+1}]},$$

$$p_{i2} = (1 - k_{i2}) \frac{S''_{i2}(x_{i2})}{f'[x_i, x_{i+1}]},$$

$$q_{i1} = \frac{1}{6}((2 - k_{i1})p_{i1} + (1 - k_{i1})p_{i2}),$$

$$q_{i2} = \frac{1}{6}((1 - k_{i2})p_{i1} + (2 - k_{i2})p_{i2})$$

$$t_{i1} = r_{i1} - \frac{1}{3}k_{i2} - q_{i1},$$

$$t_{i2} = r_{i2} - \frac{1}{3}k_{i1} - q_{i2},$$

Утверждение 7. Пусть в условиях (24)-(25), (27) для интервала $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n-1$, выполняются условия

$$t_{i1} \geq 0, t_{i2} \geq 0, \quad (33)$$

$$q_{i1} \geq 0, q_{i2} \geq 0, \quad (34)$$

причем, если $q_{i1} q_{i2} \neq 0$,

$$\text{то } \frac{1 - k_{i1}}{2 - k_{i2}} \leq \frac{q_{i1}}{q_{i2}} \leq \frac{2 - k_{i1}}{1 - k_{i2}}, \quad (35)$$

$$\text{иначе } q_{i1} = q_{i2} = 0. \quad (36)$$

Тогда существуют такие дополнительные узлы (32), при которых сплайн (10) на интервале будет выпуклым (вогнутым).

Для доказательства утверждения покажем, что условия (33)-(36) обеспечивают одновременно сохранение знака вторым производным сплайна во внутренних (дополнительных) узлах и на концах интервала.

Сначала отметим, что условия (33)-(36) обеспечивают сохранение знака вторым производным в дополнительных узлах. Чтобы убедиться в этом, заметим, проделав несложные вычисления, что

$$p_{i1} = 6 \frac{((2 - k_{i2})q_{i1} - (1 - k_{i1})q_{i2})}{3 - k_{i1} - k_{i2}},$$

$$p_{i2} = -6 \frac{((1 - k_{i2})q_{i1} - (2 - k_{i1})q_{i2})}{3 - k_{i1} - k_{i2}}.$$

Кроме того, условия (33) обеспечивают сохранение знака второй производной сплайна и на концах интервала. Для доказательства вычислим значения второй производной сплайна на концах интервала, используя уравнения связи в (10). Имеем:

$$S_i'' = \frac{6}{k_{i1}(3 - k_{i1} - k_{i2})} f''[x_i, x_{i+1}] t_{i1},$$

$$S_{i+1}'' = \frac{6}{k_{i2}(3 - k_{i1} - k_{i2})} f''[x_i, x_{i+1}] t_{i2}.$$

Следует отметить, что условия (35)-(36) (как и при равноотстоящих дополнительных узлах, когда $r_{i1} r_{i2} \neq 0$) накладывают (еще более) несимметричные ограничения на вторые производные сплайна на концах интервала. Однако эти ограничения могут быть "нейтрализованы" сдвигом дополнительных узлов. Все же и в этом случае данные условия исключают возможность построения сплайнов на тех интервалах, где вторые производные сплайна (10) на их концах оказываются одинаково малы или одинаково велики.

Вернемся к нейтрализации критических ситуаций за счет сдвига дополнительных узлов. Пусть, например, в условиях (24), (25), (27) будет $r_{i1} < r_{i2}$. Тогда, очевидно, для того чтобы выполнялись условия (33)-(36), необходимо согласовать между собой четыре величины — значения вторых производных сплайна на концах интервала и параметры k_{i1} и k_{i2} . Чтобы выполнялось левое условие (33), требуется, чтобы были достаточно малы по абсолютной величине параметр k_{i2} и вторая производная сплайна на левом конце интервала. С другой стороны, чтобы выполнялось правое условие (33), требуется, чтобы в пределах ограничений (27) и (32) были достаточно велики по абсолютной величине параметр k_{i1} и вторая производная сплайна на правом конце интервала. Зеркальная картина будет при $r_{i1} > r_{i2}$.

Разумеется, подобные ограничения на значения вторых производных в узлах исходной сетки Δ могут быть конфликтными для смежных интервалов со сходными значениями r_{i1} и r_{i2} . Можно показать, что для преодоления этих трудностей требуется больше число узлов в соответствующих интервалах.

3.3. НЕДООПРЕДЕЛЕННЫЕ ВЫЧИСЛЕНИЯ

3.3.1. Недоопределенная переменная

Существуют по крайней мере два основных, качественно различных понятия переменной. *Классическая переменная* — базовое понятие математики — представляет некоторую неизвестную величину, связанную условиями задачи с другими известными и неизвестными величинами. При достаточной полноте условий задачи сопоставленная данной переменной величина, т.е. ее значение, может быть получена точно. Таким образом, значение классической переменной отражает некоторую конкретную, заданную условиями задачи сущность, или *денотат*, представляемую в задаче именем данной переменной. В рамках одной задачи денотат-значение переменной не может меняться — он может быть либо известен, либо неизвестен.

Алгоритмическая переменная, связанная с использованием алгоритмов и появлением языков программирования, является, по сути дела, ячейкой абстрактной памяти, в которую могут помещаться различные значения, меняющиеся по ходу исполнения соответствующей процедуры.

И в математике, и в традиционном программировании с каждой переменной можно связать только одно значение из множества допустимых значений (или *универсума*).

Далее понятие переменной будет использоваться в расширенном классическом смысле: в Н-моделях переменной сопоставляется *недоопределенное значение* (или *Н-значение*), являющееся оценкой реального значения-денотата на основе доступной нам в данный момент информации. Н-значение является промежуточным между полной определенностью (точным значением) и полной неопределенностью (всем универсумом) и может уточняться по мере получения более точных данных. Например, не зная точного возраста Петрова, мы можем оценить его как «между 35 и 40 годами».

Таким образом, Н-значение есть непустое подмножество универсума, содержащее внутри себя значение-денотат, которое остается пока неизвестным (вернее, известным с точностью до данного недоопределенного значения) ввиду недостатка информации. В процессе уточнения, т.е. при поступлении более точных данных, Н-значение становится все более определенным и в пределе может стать точным, а именно равным денотату данной недоопределенной переменной (*Н-переменной*).

Это означает, что для Н-переменной, вне зависимости от ее типа, следует различать два значения:

- реальное неизвестное нам значение-денотат, которое она представляет, и ее
- текущее Н-значение, являющееся доступной оценкой этого реального значения.

Недоопределенность может характеризовать не только значения параметров существующих объектов или процессов, но и виртуальных объектов, находящихся в процессе создания. В этом случае Н-значение выступает в качестве ограничения на вычисляемое значение.

Примеры:

- здесь нужен провод диаметром от 0.25 до 0.32 мм;
- в этом редукторе придется использовать коническую или цилиндрическую передачу.

В процессе вычислений Н-значение может становиться только более точным, гарантируя тем самым монотонность вывода. Для завершаемости вычислений существенно, чтобы число различных Н-значений одного объекта было конечным.

Иллюстративный пример

Рассмотрим теперь пример, поясняющий алгоритм вывода, используемый в недоопределенных моделях.

Пусть требуется найти решение системы двух линейных уравнений с двумя неизвестными:

$$y = x - 1; \quad (F_1)$$

$$2 * y = 3 * (2 - x); \quad (F_2)$$

Каждое уравнение можно рассматривать как неявную функцию (F_1 и F_2) двух переменных

(x и y). Графики этих функций изображены на рис. 1.1.

Предположим, что известна начальная оценка значения переменной x : где-то между -1 и 4 (такую оценку можно рассматривать как интервал $[-1, 4]$).

Идея недоопределенных вычислений состоит в том, что по текущей оценке поочередно вычисляются проекции функций F_1 и F_2 на x и y . Например, проекция F_1 на y для x , равного $[-1, 4]$, есть интервал $[-2, 3]$. Результат такой проекции изображен на рис. 1.2.

Теперь, если по y вычислять проекцию F_2 на x , то получим новое значение x , равное интервалу $[0, 10/3]$ (рис. 1.3).

Продолжая этот процесс, мы постепенно приближаемся к искомому решению.

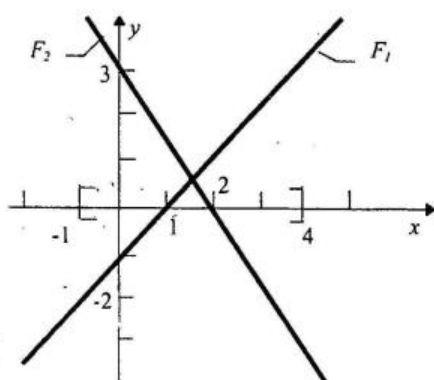


Рис. 0.1

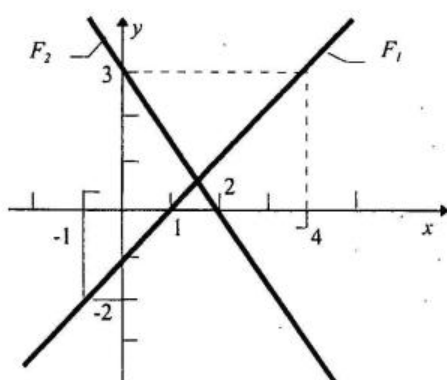


Рис. 0.2

На рис. 1. 4 изображены две спирали, которые показывают, каким образом мы приближаемся к решению — снизу и сверху соответственно.

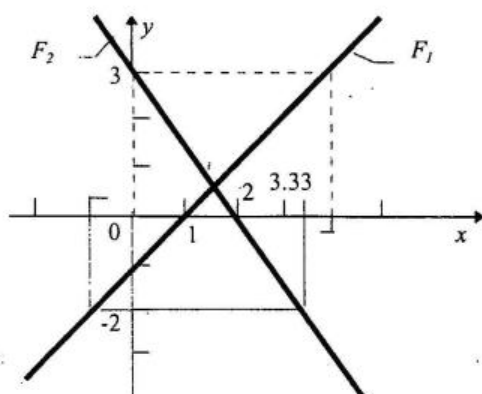


Рис. 0.3

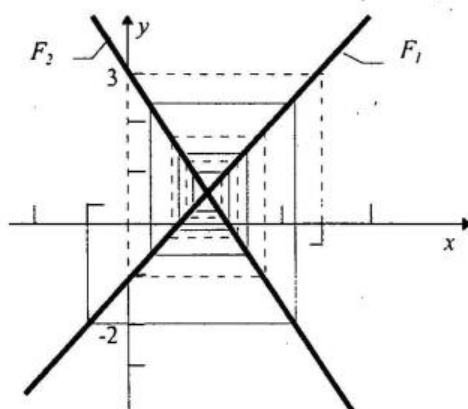


Рис. 0.4

Отметим, что параметры реальных задач всегда имеют начальные оценки своих значений, поскольку даже в тех случаях, когда решающий задачу затрудняется в определении исходных ограничений на область значений того или иного числового параметра, оценка его значения от минус бесконечности до плюс бесконечности будет представлена конкретными числами.

3.3.2. Недоопределенные расширения

При работе с недоопределенными значениями необходимо выбрать для каждого значения некоторый способ его представления. При этом необходимо представлять и его полностью неопределенное H -значение (т.е. универсум), и противоречивую оценку (т.е. пустое множество). Это приводит к следующему определению.

Исходные определения

Для того чтобы для данной традиционной формальной системы построить ее аналог, способный оперировать H -значениями, необходимо сформировать область значений для H -переменных, представляющих переменные исходной системы.

Недоопределенным расширением (H -расширением) произвольного универсального множества X является любая конечная система его подмножеств, замкнутая относительно операции пересечения и содержащая весь универсум и пустое множество. В случае бесконечного множества X в качестве универсума рассматривается некоторое его конечное подмножество $X_0 \subset X$. Например, если X — множество вещественных чисел, то X_0 может быть множеством чисел, представимых в

памяти компьютера, таких, что любые два числа отличаются не более чем на некоторый $\varepsilon > 0$.

Ниже будут использоваться обозначения *H-функция* и *H-отношение* вместо *недоопределенная функция* и *недоопределенное отношение*.

Для одного и того же универсума существуют различные возможные H-расширения. Далее будем обозначать через $*X$ произвольное H-расширение универсума X .

Независимо от вида выбранного H-расширения, приведенное выше определение гарантирует *однозначное представление* любого множества X_0 из универсума X в его H-расширении $*X$. Такое представление, обозначаемое $*[X_0]$, рассматривается как наименьший (в смысле включения) элемент H-расширения, содержащий данное подмножество.

Рассмотрим некоторые виды H-расширений, которые используются в программных технологиях, базирующихся на аппарате H-моделей.

1) Наиболее простым является H-расширение, в котором каждый его элемент представлен точным значением ($*X = X \text{ Single}$):

$$X^{\text{Single}} = \{ \{x\} \mid x \in X \} \cup \{\emptyset\} \cup \{X\}.$$

Это H-расширение добавляет в обычный универсум два специальных значения: не определено (\emptyset) и противоречие (X).

2) Перечислимое H-расширение представляется множеством всех подмножеств (которое обозначим 2^X):

$$X^{\text{Enum}} = 2^X.$$

Данное H-расширение можно применять лишь к конечным универсумам.

В случае, когда X является решеткой (множеством с определенными на нем ассоциативными и коммутативными операциями, подчиняющимися законам поглощения и идемпотентности), можно задать такие виды H-расширений X , как интервалы и мультиинтервалы [11 – 16].

3) Интервальное H-расширение:

$$X^{\text{Interval}} = \{ [x^{Lo}, x^{Up}] \mid x^{Lo}, x^{Up} \in X \}.$$

Здесь x^{Lo} обозначает нижнюю, а x^{Up} – верхнюю границу интервала. Пустое множество (\emptyset) представляется любым интервалом $[x^{Lo}, x^{Up}]$, где $x^{Lo} > x^{Up}$.

4) Мультиинтервальное H-расширение:

$$X^{\text{MultiInterval}} = \{ x \mid x = \bigcup x_k, x_k \in X^{\text{Interval}}, k = 1, 2, \dots \}.$$

Пример. Пусть универсум переменной v – это множество целочисленных значений, а ее текущее значение равно множеству $\{3, -2, 7, 8, 9, 4\}$. Рассмотрим его недоопределенное представление в различных H-расширениях множества целых чисел:

H-расширение	H-значение v
<i>Single</i>	(полная неопределенность)
<i>Enum</i>	$\{3, -2, 7, 8, 9, 4\}$
<i>Interval</i>	$[-2, 9]$
<i>MultiInterval</i>	$\{[-2, -2], [3, 4], [7, 9]\}$

5) Рассмотрим теперь универсум, заданный в виде декартова произведения множеств: $X = X_1 \times \dots \times X_n$. Как и к любому множеству, к нему применимы Н-расширения X^{Single} и X^{Enum} . Более того, если каждый из X_i является решеткой, то X тоже является решеткой. В этом случае к X также применимы Н-расширения $X^{Interval}$ и $X^{MultiInterval}$. Но к универсуму X применимо еще одно Н-расширение – структурное:

$$*X = *X_1 \times \dots \times *X_n.$$

3.3.3. Недоопределенные операции и отношения

Рассмотрим, каким образом над приведенными выше Н-расширениями можно определить операции и отношения.

Пусть в предметной области задана операция над обычными универсумами

$$f: X_1 \times \dots \times X_n \rightarrow X_{n+1}$$

и некоторые Н-расширения $*X_i$ ($i = 1, \dots, n+1$). Тогда Н-расширение операции f есть операция

$$*f: *X_1 \times \dots \times *X_n \rightarrow *X_{n+1},$$

результатом которой является представление образа функции f на множестве значений $X_1 \times \dots \times X_n$ в недоопределенном расширении $*X_{n+1}$:

$$*f(x_1, \dots, x_n) = *[\{f(a_1, \dots, a_n) \mid a_i \in x_i, i = 1, \dots, n\}].$$

Под образом функции f на множестве значений X понимается множество значений $\{f(x)\}$ для всех $x \in X$. Например, в случае перечислений (*Enum*) вычисление Н-расширенной операции может заключаться в переборе возможных значений ее аргументов, а в случае интервалов (*Interval*) и мультиинтервалов (*Multiinterval*) можно воспользоваться широко известными правилами интервальной арифметики.

Пример. Пусть имеется операция извлечения корня из вещественного числа, которая вычисляет не только арифметическое значение корня, но и его отрицательный аналог. Областью определения этой функции является множество неотрицательных чисел. Посмотрим на некоторые виды ее Н-расширений на примере извлечения корня из недоопределенного вещественного числа, заданного в виде интервала.

Н-расширение	Значение корня
<i>Single</i>	(полная неопределенность)
<i>Enum</i>	$\{-5, -5 + \epsilon, -5 + 2\epsilon, \dots, -2 - \epsilon, -2, 2, 2 + \epsilon, 2 + 2\epsilon, \dots, 5 - \epsilon, 5\}$
<i>Interval</i>	$[-5, 5]$
<i>MultiInterval</i>	$\{[-5, -2], [2, 5]\}$

Пусть теперь в предметной области задано отношение $r(x_1, \dots, x_n)$, определяющее некоторое подмножество декартова произведения $X_1 \times \dots \times X_n$. Проблема организации автоматического решения для моделей традиционно связывается с функциональной интерпретацией всякого отношения $r(X')$, которое является набором функций, вычисляющих значение каждой переменной из X' на основании заданных значений других переменных. Таким образом, функция f_i ($i = 1, \dots, n$) называется функцией интерпретации отношения r , если

$$x_i = f_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

на множестве $r(x_1, \dots, x_n)$.

Например, уравнение

$$x + y = z ;$$

интерпретируется тремя функциями:

$$z = x + y ;$$

$$x = z - y ;$$

$$y = z - x ;$$

Такие функционально интерпретируемые отношения в дальнейшем будем называть ограничениями.

Задача удовлетворения ограничений на Н-моделях описывается как множество объектов предметной области, каждому из которых сопоставлено недоопределенное значение, и множество ограничений на этих объектах. При этом каждому объекту приписывается универсум и вид Н-расширения.

Все функции интерпретации ограничений, если их рассматривать как операторы, действующие на множестве Н-значений объектов модели, являются монотонными по включению и сужающимися. Двух этих свойств и конечности множества всех Н-значений достаточно для существования наибольшей (по включению) неподвижной точки произвольной системы таких операторов и гарантированного достижения этой точки за конечное число шагов методом последовательных приближений. Его разновидностью является используемый в Н-моделях потоковый алгоритм вычислений, который подробнее будет рассмотрен в следующей главе.

Свойства недоопределенных расширений

Заметим, что некоторые свойства обычных операций существенно меняются в соответствующих недоопределенных расширениях.

Рассмотрим, например, хорошо известное свойство аддитивных операций над числами (целыми или вещественными), — сумма любого числа и обратного ему равна единственному нулевому элементу

$$a + (-a) = 0. \quad (1)$$

Пусть в качестве Н-расширения множества чисел рассматривается Н-расширение Interval, а соответствующие Н-расширения операций минус (*-) и плюс (*+) определены согласно известным правилам интервальной арифметики. Таким образом, Н-расширение унарного минуса имеет вид:

$$*-: \quad [a^{Lo}, a^{Up}] = [-a^{Up}, -a^{Lo}],$$

а Н-расширение сложения ($a = b *+ c$) вычисляется следующим образом:

$$*+: \quad [a^{Lo}, a^{Up}] = [b^{Lo} + c^{Lo}, b^{Up} + c^{Up}].$$

Тогда Н-расширение выражения (1) имеет следующий вид:

$$\begin{aligned} [a^{Lo}, a^{Up}] *+ (*-[a^{Lo}, a^{Up}]) &= [a^{Lo}, a^{Up}] *+ [-a^{Up}, -a^{Lo}] = \\ &= [a^{Lo} + (-a^{Up}), a^{Up} + (-a^{Lo})] = [a^{Lo} - a^{Up}, a^{Up} - a^{Lo}]. \end{aligned}$$

Достаточно очевидно, что когда нижняя и верхняя границы Н-числа не совпадают (т.е. число недоопределено), интервал $[a^{Lo} - a^{Up}, a^{Up} - a^{Lo}]$ всего лишь содержит нулевой элемент, но не равен в точности ему.

Такое изменение свойства унарного минуса приводит к тому, что закон дистрибутивности для обычных аддитивных и мультипликативных операций переходит в так называемый закон субдистрибутивности:

$$*a * \times (*b *+ *c) \subseteq *a * \times *b *+ *a * \times *c.$$

Из сказанного выше можно сделать вывод, что в Н-моделях существенно то, в каком виде представлены условия задачи: какие Н-расширения выбраны для переменных и каким образом представлены ограничения (уравнения и неравенства).

3.3.4. Недоопределенные модели

Обобщенные вычислительные модели

Недоопределенные модели являются частным случаем *обобщенных вычислительных моделей* (ОВМ), которые имеют более широкую область применения, чем решение задач удовлетворения ограничений. Ниже мы даем определение ОВМ и алгоритма вычислений на них, указывая при необходимости отличия Н-моделей от ОВМ.

Обобщенная вычислительная модель $M = (V, W, C, R)$ состоит из следующих четырех множеств:

V — множество объектов из заданной предметной области,

R — множество ограничений на значения объектов из V ,

W — множество функций присваивания и

C — множество функций проверки корректности.

Каждому объекту $v \in V$ сопоставлены:

универсум X_v ,

- начальное значение из универсума (точное, недоопределенное или полная неопределенность),

- функция присваивания W_v ,

- функция проверки корректности C_v .

Функция присваивания — это двухместная функция, работающая при каждой попытке присваивания очередного значения объекту $v \in V$ и определяющая новое значение объекта как функцию от текущего и присваиваемого значений.

Функция проверки корректности — это унарный предикат, который выполняется в случае, если значение объекта x изменилось, и проверяет правильность этого нового значения.

Ограничения из R должны быть функционально интерпретируемыми.

На уровне интерпретации ОВМ представляется двудольным ориентированным графом (ОВМ-сеть), в котором выделены два типа вершин: объекты и функции. Дуги связывают функциональные и объектные вершины. Входящие в вершину-функцию дуги соотносят с ней объекты, значения которых выступают в качестве входных аргументов для функции, исходящие — указывают на объекты, в которые должна производиться запись вырабатываемых функцией результатов.

Каждой объектной вершине сопоставляются тип и значение, а также с ней связываются функции присваивания и проверки корректности.

С каждой функциональной вершиной соотнесены целое число, играющее роль приоритета, и разметка входящих и исходящих дуг.

Алгоритм удовлетворения ограничений на ОВМ

Процесс вычислений на ОВМ имеет потоковый характер — изменение объектных вершин сети активирует (вызывает к исполнению) функциональные вершины, для которых эти объектные вершины являются входными аргументами, а исполнение функциональных вершин, в свою очередь, может вызывать изменение результирующих объектных вершин. Вычисления заканчиваются тогда, когда либо не останется активных функциональных вершин (УСПЕХ), либо функция проверки корректности вырабатывает значение «ложь» (НЕУДАЧА).

Допустим, что вместо обычных универсумов X_v рассматриваются некоторые их недоопределенные расширения $*X_v$. Пусть все функции присваивания в ОВМ производят пересечение Н-значений:

$$w_i(\xi^{old}, \xi^{new}) = \xi^{old} \cap \xi^{new},$$

а функции проверки корректности — проверку Н-значения на непустоту:

$corr_i(\xi) = \text{if } \xi \neq \emptyset \text{ then true else false fi.}$

Именно этот класс обобщенных вычислительных моделей называется недоопределенными моделями, или Н-моделями.

Используются известные доказанные утверждения:

- (i) Процесс удовлетворения ограничений в Н-моделях завершается за конечное число шагов.
- (ii) Достижение процессом НЕУДАЧИ или УСПЕХА предопределено входными данными (начальными Н-значениями переменных и ограничениями) и не зависит от конкретной стратегии выбора очередного ограничения для интерпретации.
- (iii) В случае УСПЕХА процесса, при одних и тех же входных Н-значениях переменных, их выходные Н-значения не зависят от конкретной стратегии выбора очередного ограничения для интерпретации.
- (iv) В случае УСПЕХА процесса решение задачи (если оно существует) лежит внутри полученного результата (декартова произведения Н-значений).

Пример работы алгоритма

Поясним работу алгоритма удовлетворения ограничений для Н-моделей на примере достаточно простой системы двух линейных уравнений с двумя целочисленными неизвестными:

$$\begin{cases} x + y = 12; \\ 2 * x = y; \end{cases} \quad (2)$$

Предположим, что значения x и y ограничены следующими неравенствами:

$$0 \leq x \leq 100; 0 \leq y \leq 100. \quad (3)$$

Множество объектов X данной модели содержит две целочисленные переменные (x, y), множество ограничений R — два уравнения и четыре неравенства.

Для определения соответствующей Н-модели требуется выбрать некоторый вид Н-расширения целого числа. Пусть недоопределенное целое число (Н-число) будет представлено Н-расширением *Interval*. Функция присваивания (обозначим ее $PRint$) реализует следующее правило изменения значений Н-числа: нижняя граница может только увеличиваться, а верхняя — только уменьшаться. Функция проверки корректности (обозначим ее $PRDint$) проверяет, чтобы нижняя граница Н-числа не превосходила верхнюю. Операции для таких Н-чисел определяются в соответствии с правилами интервальной математики.

Так как множество X содержит две переменные типа Н-число, множество функций присваивания (W) и множество функций проверки корректности (C) содержат по два элемента:

$$W = \{ PRint(x), PRint(y) \}; \quad C = \{ PRDint(x), PRDint(y) \};$$

Множество функций интерпретации уравнений (2) состоит из четырех элементов ($f_1 - f_4$):

$$\begin{aligned} f_1: y &\leftarrow 12 - x; & f_2: x &\leftarrow 12 - y; \\ f_3: y &\leftarrow 2 * x; & f_4: x &\leftarrow y / 2; \end{aligned}$$

Согласно правилам интервальной арифметики, семантика функций интерпретации ($f_1 - f_4$) представляется следующим образом:

$$\begin{aligned} f_1: [y^{Lo}, y^{Up}] &\leftarrow [12 - x^{Up}, 12 - x^{Lo}]; \\ f_2: [x^{Lo}, x^{Up}] &\leftarrow [12 - y^{Up}, 12 - y^{Lo}]; \\ f_3: [y^{Lo}, y^{Up}] &\leftarrow [\min \{ 2 * x^{Lo}, 2 * x^{Up} \}, \max \{ 2 * x^{Lo}, 2 * x^{Up} \}]; \\ f_4: [x^{Lo}, x^{Up}] &\leftarrow [\min \{ x^{Lo} / 2, x^{Up} / 2 \}, \max \{ x^{Lo} / 2, x^{Up} / 2 \}]; \end{aligned}$$

Четыре ограничения (3) можно проинтерпретировать на этапе генерации Н-сети, и тогда нижние и верхние границы x и y станут равными 0 и 100 соответственно. Таким образом, на первом шаге процесса потоковых вычислений имеем:

$$x = [0, 100]; \quad y = [0, 100];$$

В табл. 1 приведена последовательность шагов исполнения потокового процесса для данной задачи.

Таблица 1. Протокол исполнения Н-модели

N	Активные функции	Н-значения текущее новое	Флаг	Добавить функции
1	$f_1 f_2 - f_4$	$y = [0, 100] [0, 12]$	да	f_2, f_4
2	$f_2 f_3, f_4$	$x = [0, 100] [0, 12]$	да	f_1, f_3
3	$f_3 f_4, f_1$	$y = [0, 12] [0, 12]$	нет	
4	$f_4 f_1$	$x = [0, 12] [0, 6]$	да	f_1, f_3
5	$f_1 f_3$	$y = [0, 12] [6, 12]$	да	f_2, f_4
6	$f_3 f_2, f_4$	$y = [6, 12] [6, 12]$	нет	
7	$f_2 f_4$	$x = [0, 6] [0, 6]$	нет	
8	$f_4 $	$x = [0, 6] [3, 6]$	да	f_1, f_3
9	$f_1 f_3$	$y = [6, 12] [6, 9]$	да	f_2, f_4
10	$f_3 f_2, f_4$	$y = [6, 9] [6, 9]$	нет	
11	$f_2 f_4$	$x = [3, 6] [3, 6]$	нет	
12	$f_4 $	$x = [3, 6] [3, 4]$	да	f_1, f_3
13	$f_1 f_3$	$y = [6, 9] [8, 9]$	да	f_2, f_4
14	$f_3 f_2, f_4$	$y = [8, 9] [8, 8]$	да	f_2, f_4
15	$f_2 f_4$	$x = [3, 4] [4, 4]$	да	f_1, f_3
16	$f_4 f_1, f_3$	$x = [4, 4] [4, 4]$	нет	
17	$f_1 f_3$	$y = [8, 8] [8, 8]$	нет	
18	$f_3 $	$y = [8, 8] [8, 8]$	нет	

В таблице предполагается, что на каждом шаге исполняется первая функция из множества активных функций (она отделена от остальных символом '|'). На первом шаге итерации исполняется функция f_1 . В результате работы этой функции получаем значение $y = [-8, 12]$. После этого вызывается функция присваивания $PRint(y)$. Текущее значение y равно $[0, 100]$, новое — $[-88, 12]$. Функция присваивания вырабатывает значение $[0, 12]$ и присваивает его y . Как видим, у переменной y изменилась только верхняя граница (т.к. новая нижняя граница (-88) меньше текущей (0)). Изменение значения y в целом все-таки произошло (Флаг = да), поэтому вызывается функция проверки корректности $PRDint(y)$. Условие корректности для y ($0 \leq 12$) не нарушается, и процесс исполнения Н-модели продолжается, т.е. происходит активация функций интерпретации f_2 и f_4 , для которых y является входным аргументом. Ввиду того что эти функции уже входят в множество активных функций, их повторная активация не происходит. Далее исполняется следующая функция (f_2) из очереди.

Вычисления заканчиваются тогда, когда нижняя и верхняя границы как x , так и y становятся равными друг другу. Это произойдет при значениях x^{Lo} и x^{Up} , равных 8, а

y^{Lo} и y^{Up} — равных 4. При таких значениях исполнение любой функции $f_1 - f_4$ не изменяет значения своего результата, и множество активных функций становится пустым.

3.3.5. ТЕХНОЛОГИЯ ИНТЕРВАЛЬНЫХ ТАБЛИЦ INSRESH

Технология интервальных электронных таблиц (InSpreSh), сохраняя эффективность и простоту использования, свойственные обычным электронным таблицам, предоставляет уникальные вычислительные возможности для решения широкого класса задач в условиях неполноты и неточности исходных данных в различных областях.

Новые качества технологии

Технология InSpreSh базируется на использовании современных методов искусственного интеллекта и интервальной математики. По сравнению с обычными электронными таблицами (например, Excel) InSpreSh предоставляет следующие принципиально новые возможности:

1. Позволяет выполнять вычисления не только с точной, но и с частично известной информацией, заданной в виде интервалов допустимых значений соответствующих параметров. При решении задачи эти интервалы автоматически сужаются (либо превращаются в точные значения) как в результате вычислений, так и при поступлении дополнительной информации извне. Такая возможность особенно важна для задач финансового планирования. Например, лишь приблизительно можно оценить величину будущей инфляции, сезонные изменения цен на продукцию или колебания курса обмена валют.

2. Ячейки электронной таблицы могут быть связаны формулами произвольного вида и сложности (линейными и нелинейными уравнениями, неравенствами, логическими выражениями), причем количество связей для одной ячейки не ограничено. Это позволяет решать реальные задачи повышенной сложности, недоступные другим системам.

3. Все формулы, выражающие связи между ячейками таблицы, допускают как прямые, так и обратные вычисления, что позволяет определять не только последствия принимаемых решений, но и формировать сами эти решения, исходя из желаемого результата в будущем.

4. Современный пользовательский интерфейс InSpreSh позволяет структурировать таблицы произвольным образом, определяя иерархическую структуру строк и столбцов.

Вычислительные возможности технологии InSpreSh основаны на методе недоопределенных моделей — оригинальном математическом аппарате. Сегодня этот аппарат превосходит аналогичные западные работы, объединенные общим термином — constraint programming (программирование в ограничениях). Сейчас на программном рынке известно несколько продуктов, в которых электронные таблицы интегрируются с методами программирования в ограничениях. Большинство из них обеспечивают решение только простых задач и в отличие от технологии InElTab не способны решать сложные неоднородные системы ограничений и обрабатывать интервальные значения.

Технология InSpreSh реализована в виде прикладной программной системы, предназначенной для широкого круга пользователей. Ядром этой системы является универсальный решатель Unicalc.

Опыт применения системы InSpreSh показал, что она позволяет достаточно эффективно решать задачи, которые включают сотни показателей, связанные тысячами ограничений. При этом время, затрачиваемое на поиск решения, прямо зависит от степени неопределенности начальных данных. Чем большее число показателей задано интервальными значениями и чем шире эти интервалы, тем больше требуется времени для нахождения сбалансированного решения. При работе с точными значениями эффективность нашей системы сравнима с обычными электронными таблицами.

Монографии

1. Деревянко А.П., Холюшкин Ю.П., Ростовцев П.С., Воронин В.Т.. Неандертальская проблема как задача статистического анализа. –Новосибирск, Изд. НГУ., 2001, 71. с.71 (7,5 п.л.).
2. Derevianko A.P., Kholiouchkine Y.P., Rostovtsev P.S., Voronine V.T. L 'Analyse statistique desensembles de Paleolithique moyen du proche –et Moyen-Orient . – Изд: НГУ., Новосибирск, 2001, 48 с. (3.0 п.л.).
3. Деревянко А.П., Холюшкин Ю.П., Ростовцев П.С., Воронин В.Т. Статистический анализ мустьерских индустрий Кавказа. Ч.1. Технологические индексы. – Новосибирск, Изд. НГУ, 2002, 53 с. (6.75 п.л.).
4. Деревянко А.П., Холюшкин Ю.П., Ростовцев П.С., Воронин В.Т. Статистический анализ мустьерских индустрий Кавказа. Ч.2. Типология. – Новосибирск, Изд. НГУ, 2002, 45 с. (5.5 п.л.).

Статьи и тезисы докладов

1. Ершов Ю.Л., Деревянко А.П., Холюшкин Ю.П. и др. Развитие проблемно-ориентированной информационной сети на базе Сибирского центра информационной поддержки гуманитарных наук, культуры и образования.// Новые информационные технологии в университетском образовании. – Новосибирск, 2001: 89-90.
2. Топчий А.Т., Гемуев И.Н., Холюшкин Ю.П., Воронин В.Т., Бауло А.В., Соловьев А.И., Воробьев В.В., Топчий Р.А., Творогова И.Ю. О концепции создания электронных библиотек по истории, археологии и этнографии //Историческая наука на рубеже веков. – Томск, 2001:246 – 258.
3. Холюшкин Ю.П. Археологические ресурсы в сети Интернет //Информационные технологии в гуманитарных исследованиях. Вып.2. – Новосибирск, 2000: 87–93.Холюшкин .П., Воронин .Т.. Сектор археологической теории и информатики: итоги пятилетия //Информационные технологии в гуманитарных исследованиях. Вып.2. – Новосибирск, 2000: 33-46.
4. Холюшкин Ю.П., Воронин В.Т., Воробьев В.В. Информационный портал «Гуманитарная паутина» как форма представления исторического знания и интеллектуальная культура в Интернет. // Историческое знание и интеллектуальная культура. Часть2., М., 2001, с. 47-51. (0,3 п.л.).
5. Деревянко А.П., Холюшкин Ю.П., Воронин В.Т., Воробьев В.В.. Технология WWW и программа создания виртуальных музеев. // Историческое знание и интеллектуальная культура. Часть2., М., 2001, с. 62-64. (0,2 п.л.).
6. Холюшкин Ю.П., Воронин В.Т., Федоров С.А., Бердников Е.В. Технология доступа к полнотекстовым базам данных ИАЭТ СО РАН из Интернет // Технологии информационного общества. Санкт-Петербург, 2001, с. 228-230 (0,2 п.л.).
7. А.П.Деревянко, Ю.П.Холюшкин, В.Т.Воронин, В.В.Воробьев, С.А.Федоров, Е.В.Бердников, С.В.Елагина. Создание информационного центра сектора археологической теории и информатики ИАЭТ СО РАН 9 проблемы и решения) //Информационные технологии в гуманитарных исследованиях. – Новосибирск, Изд. НГУ, 2002, с. 6-14 (0.7 п.л.).
8. Холюшкин Ю.П., Воронин В.Т., Федоров С.А., Бердников Е.В., Жилицкая Г.Ю. Электронный каталог научной библиотеки ИАЭт СО РАН //Информационные технологии в гуманитарных исследованиях. – Новосибирск, Изд. НГУ, 2002, с. 15-20 (0.45 п.л.).
9. Жданов А.С., Костин В.С. Значимость и устойчивость автоматической классификации в задаче поиска оптимального разбиения //Информационные технологии в гуманитарных исследованиях. – Новосибирск, Изд. НГУ, 2002, с. 36-42 (0.5 п.л.).

10. Деревянко А.П., Холюшкин Ю.П., Ростовцев П.С. Воронин В.Т. Статистический анализ технологических индексов мустьерских индустрий Кавказа // Информационные технологии в гуманитарных исследованиях. – Новосибирск, Изд. НГУ, 2002, с. 43-65 (1,7 п.л.).
11. Деревянко А.П., Холюшкин Ю.П., Ростовцев П.С., Воронин В.Т. Статистический анализ индексов мустьерских памятников Средней Азии // Проблемы каменного века Средней и Центральной Азии. – Новосибирск, Изд. ИАЭТ СО РАН, 2002, с. 92-101 (1,0 п.л.).
12. Бердников Е.В., Дунаев Д.С., Марнуев П.Е.. Опыт создания сетевого информационного ресурса по палеолиту Северной Азии на основе ГИС-технологий. // Культурология и история древних и современных обществ Сибири и Дальнего Востока (материалы XLII РАЭСК), Омск, 2002, 586-587 стр. (0,1 п.л.).
13. Жигалов В.А., Загорюлько Ю.А., Нариньяни А.С., Россеева О.И. Предел однородности поиска в Интернет. // Системная информатика. Выпуск 8: Теория и методология программирования – Новосибирск: Наука. Сибирская издательская фирма РАН, 2002. – с.29-71.
14. Боровикова О.И., Загорюлько Ю.А. Организация порталов знаний на основе онтологий // Труды международного семинара Диалог'2002 по компьютерной лингвистике и ее приложениям. -Протвино, 2002. -Т.2. -с.76-82.
15. Загорюлько Ю.А., Кононенко И.С., Костов Ю.В., Сидорова Е.В. Представление знаний в интеллектуальной системе документооборота // Труды 8-й национальной конференции по искусственному интеллекту - КИИ'2002. –Москва: Физматлит, 2002. -Т.2. -С.867-875.
16. Загорюлько Ю.А., Кононенко И.С., Костов Ю.В., Сидорова Е.В. Подход к разработке интеллектуальной системы документооборота инвестиционной компании // Труды IV-й международной конференции "Проблемы управления и моделирования в сложных системах". - Самара: Самарский Научный Центр РАН, 2002. -С. 366-372.
17. Кононенко И.С., Сидорова Е.В. Обработка делового письма в системе документооборота // Труды международного семинара Диалог'2002 по компьютерной лингвистике и ее приложениям. -Протвино, 2002. -Т.2. -с.299-310.
18. Загорюлько Ю.А., Кошечев В.О., Мамонтов П.Г., Парамзин Д.Ю. Новый подход к разработке решателя UniCalc // Труды 8-й национальной конференции по искусственному интеллекту - КИИ'2002. –Москва: Физматлит, 2002. -Т.2. -С.720-728.
19. Яхно Т.М. Алгоритмы муравьиных колоний: еще одна альтернатива для задач оптимизации? // Там же. -Т.1. -С.372-380.

Научное издание

**Информационные технологии в гуманитарных исследованиях:
Выпуск 4. Информационные технологии и математические
методы в археологии.**

Материалы научного отчета по интеграционной программе СО РАН
за 2000–2002 гг. (проект № 82).

Подписано в печать 04.12.2002 Формат 60х84 1/8. Офсетная печать.

Уч.-изд. л. 8,25. Тираж 200 экз. Заказ № 591

Лицензия ЛР № 021285 от 6 мая 1998 г.

Редакционно-издательский центр НГУ 630090, Новосибирск-90, ул. Пирогова, 2.

