

# **ИНФОРМАТИКА И ЕЁ ПРИМЕНЕНИЯ**

**Научный журнал Отделения информационных технологий  
и вычислительных систем Российской академии наук**

Издается с 2007 года  
Журнал выходит ежеквартально

**Учредители:**  
**Российская академия наук**  
**Институт проблем информатики Российской академии наук**

## **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

академик С. В. Емельянов (главный редактор, член Редакционного совета)  
академик Ю. И. Журавлев (председатель Редакционного совета)  
академик С. К. Коровин  
академик Г. И. Савин  
академик А. Л. Стемпковский  
академик Ю. И. Шокин (член Редакционного совета)  
член-корреспондент РАН В. Л. Арлазаров  
член-корреспондент РАН А. Б. Жижченко  
член-корреспондент РАН И. А. Каляев  
член-корреспондент РАН Ю. С. Попков  
член-корреспондент РАН К. В. Рудаков  
член-корреспондент РАН И. А. Соколов (зам. главного редактора,  
член Редакционного совета)  
член-корреспондент РАН Ю. А. Флеров  
член-корреспондент РАН Б. Н. Четверушкин  
член-корреспондент РАН Р. М. Юсупов  
профессор, д.т.н. В. И. Будзко  
профессор, д.т.н. А. А. Зацаринный  
профессор, д.ф.-м.н. А. В. Печинкин  
профессор, д.т.н. И. Н. Сеницын  
профессор, д.ф.-м.н. С. Я. Шоргин (ответственный секретарь)

## **Редакция**

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков  
к.ф.-м.н. Е. Н. Арутюнов  
О. В. Ломакина

© Институт проблем информатики Российской академии наук, 2008

## **Адрес редакции:**

Москва 119333, ул. Вавилова 44, корп. 2, ИПИ РАН,  
редакция журнала «Информатика и её применения»  
Тел. 8(499)135-86-92, e-mail [rust@ipiran.ru](mailto:rust@ipiran.ru)

Подписной индекс журнала в каталоге «Пресса России» 88018 (годовая подписка)

# Информатика и её применения

Том 2 Выпуск 1 Год 2008

## СОДЕРЖАНИЕ

Архитектура промежуточного слоя предметных посредников для решения задач над множеством интегрируемых неоднородных распределенных информационных ресурсов в гибридной грид-инфраструктуре виртуальных обсерваторий <i>Д. О. Брюхов, А. Е. Вовченко, В. Н. Захаров, О. П. Желенкова, Л. А. Калининченко, Д. О. Мартынов, Н. А. Скворцов, С. А. Ступников</i>	2
Квазилинейные методы построения информационных моделей флуктуаций неравномерности вращения Земли <i>И. Н. Сеницын</i>	35
Порталы в системах органов государственной власти <i>А. В. Босов</i>	44
Система <i>Geo/G/1/∞</i> с одной «нестандартной» дисциплиной обслуживания <i>А. В. Печинкин, С. Я. Шоргин</i>	55
Размещения систем массового обслуживания, минимизирующие среднюю длину очереди <i>Т. В. Захарова</i>	63
Новый этап информатизации общества и актуальные проблемы образования <i>И. А. Соколов, К. К. Колин</i>	67
Предпосылки и факторы конвергенции информационной и компьютерной наук <i>И. М. Зацман, О. С. Кожунова</i>	77
Abstracts	98

---

Выпускающий редактор	<i>Т. Торжкова</i>
Технический редактор	<i>Л. Кокушкина</i>
Художественный редактор	<i>М. Седакова</i>

Сдано в набор 31.01.08. Подписано в печать 17.03.08. Формат 60 x 84 / 8  
Бумага офсетная. Печать офсетная. Усл.-печ. л. 12,5. Уч.-изд. л. 11,5. Тираж 200 экз.

Заказ №

Издательство «ТОРУС ПРЕСС», Москва 121614, ул. Крылатская, 29-1-43  
[torus@torus-press.ru](mailto:torus@torus-press.ru); <http://www.torus-press.ru>

Отпечатано в ППП «Типография «Наука» с готовых диапозитивов  
Москва 121099, Шубинский пер., д. 6.

# АРХИТЕКТУРА ПРОМЕЖУТОЧНОГО СЛОЯ ПРЕДМЕТНЫХ ПОСРЕДНИКОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ НАД МНОЖЕСТВОМ ИНТЕГРИРУЕМЫХ НЕОДНОРОДНЫХ РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ В ГИБРИДНОЙ ГРИД-ИНФРАСТРУКТУРЕ ВИРТУАЛЬНЫХ ОБСЕРВАТОРИЙ\*

Д. О. Брюхов<sup>1</sup>, А. Е. Вовченко<sup>2</sup>, В. Н. Захаров<sup>3</sup>, О. П. Желенкова<sup>4</sup>, Л. А. Калиниченко<sup>5</sup>, Д. О. Мартынов<sup>6</sup>, Н. А. Скворцов<sup>7</sup>, С. А. Ступников<sup>8</sup>

**Аннотация:** Рассматривается архитектура промежуточного слоя предметных посредников для решения научных задач над множеством интегрируемых неоднородных распределенных информационных ресурсов в гибридной грид-инфраструктуре виртуальной обсерватории (ВО). Архитектура реализована как объединение системы поддержки ВО АстроГрид, разработанной в Великобритании, и средств поддержки предметных посредников, созданных в ИПИ РАН. Реализован подход, при котором для класса приложений формируется спецификация предметной области посредников независимо от существующих информационных ресурсов. Создание прототипа гибридной архитектуры требует сопряжения исполнительных механизмов двух инфраструктур (АстроГрида и предметных посредников), поэтому основное внимание уделено проблемам переписывания запросов к посредникам в планы их реализации над конкретными информационными ресурсами. Приведено краткое описание объединенной архитектуры исполнительных механизмов АстроГрида и средств поддержки предметных посредников. Дан пример реализации предметного посредника для решения задач поиска далеких галактик в гибридной архитектуре. Показаны отличия реализованного подхода от известных прототипов интеграции баз данных в ВО. Разработанные средства планируется использовать при решении задач Российской ВО (РВО).

**Ключевые слова:** предметный посредник; каноническая информационная модель; виртуальная обсерватория; унификатор информационных моделей; уточнение; переписывание формул; семантическая интеграция неоднородных информационных ресурсов; регистрация ресурсов в посреднике; онтологическая модель; релевантные посреднику ресурсы; промежуточный слой; описание предметной области задачи, движимое приложениями

## 1 Введение

В различных областях науки наблюдается экспоненциальный рост объема получаемых экспериментальных (наблюдательных) данных. Например, в астрономии текущий и ожидаемый темп роста данных от наземных и космических телескопов удваивается в течение периода от шести месяцев до одного года. Сложность использования таких данных увеличивается еще и вследствие

их естественной разнородности. Число организаций, получающих данные наблюдений в отдельных областях науки в мире, велико. Разнообразие (информационная несогласованность) получаемой информации вызывается, в частности, не только большим числом организаций, производящих наблюдения, и их независимостью, но и разнообразием объектов наблюдения, непрерывным и быстрым совершенствованием техники наблюдений, вызывающим адекватные изменения структуры и со-

\* Работа выполнена при частичной финансовой поддержке РФФИ (гранты 06-07-08072-офи-а, 06-07-89188-а, 08-07-00157-а) и программы ОИТВС РАН «Фундаментальные основы информационных технологий и систем» (проект 1-10).

<sup>1</sup>Институт проблем информатики Российской академии наук, brd@ipi.ac.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, itsnein@gmail.com

<sup>3</sup>Институт проблем информатики Российской академии наук, vzakharov@ipiran.ru

<sup>4</sup>Специальная астрофизическая обсерватория Российской академии наук, zhe@sao.ru

<sup>5</sup>Институт проблем информатики Российской академии наук, leonidk@synth.ipi.ac.ru

<sup>6</sup>Институт проблем информатики Российской академии наук, domartynov@gmail.com

<sup>7</sup>Институт проблем информатики Российской академии наук, nskv@ipi.ac.ru

<sup>8</sup>Институт проблем информатики Российской академии наук, ssa@ipi.ac.ru

держания накапливаемой информации. Это приводит к необходимости использования неоднородной, распределенной информации, накопленной в течение значительного периода наблюдений технологически различными инструментами.

Чрезвычайно быстро развивается также программный инструментарий, включающий многообразные сервисы для поддержки различных видов обработки информации при решении научных задач и проведении исследований. Такие сервисы производятся различными научными организациями, их описания неоднородны и неполны.

Увеличивающийся разрыв между исследователями и источниками данных и сервисов приводит к необходимости поиска новых путей создания информационных систем, в которых особое внимание было бы сосредоточено на специальных средствах организации решения задач над множеством распределенных информационных ресурсов (данных и программ), накапливаемых в разнообразных научных центрах. Разработан (разрабатывается) ряд инфраструктур, которые технически способствуют организации решения задач в такой среде. Среди них Веб-сервисы, Гриды данных, Семантический Веб, инфраструктуры онтологического моделирования, интеграции информационных ресурсов, интероперабельные инфраструктуры промежуточного слоя и др. Вместе с тем сложны и все еще далеки от решения проблемы семантики, возникающие при:

- (1) определении понятий предметной области;
- (2) отображении и интеграции контекстов предметных областей информационных ресурсов в контекст предметной области задачи;
- (3) идентификации релевантных задач информационных ресурсов и формировании их композиций;
- (4) доказательно правильном отображении информационных моделей ресурсов в общую информационную модель задачи;
- (5) интеграции схем ресурсов в схеме предметной области задачи и устранении разнообразных конфликтов;
- (6) выявлении семантически подобных компонентов ресурсов в процессе интеграции схем;
- (7) адекватном преобразовании формул (запросов) программы решения задачи, выраженных в терминах схемы предметной области задачи, в формулы, выраженные в схемах релевантных ресурсов, и пр.

Следует заметить, что аналогичные проблемы создания и развития информационных систем возникают не только в научных применениях, но и в

корпоративных информационных системах, создаваемых для решения задач в разнообразных областях государственного управления, производства, бизнеса. Достаточно назвать системы, конструируемые в рамках программ «Электронная Москва», «Электронная Россия», системы для решения задач правоохранительных органов, для инспекции безопасности дорожного движения, банковские системы, системы управления транспортом и др.

Настоящая статья в качестве применений ограничивается рассмотрением инфраструктур информационных систем для науки, которые в последнее время приобретают вид международных ВО. Виртуальные обсерватории, в свою очередь, могут быть отнесены к классу корпоративных информационных систем, поэтому результаты настоящей работы применимы и к последним. Анализ инфраструктуры РВО для астрономии дан в ее аванпроекте [1]. Значительное внимание в нем занимает анализ состояния и решений наиболее продвинутых в мире проектов ВО: проекта Национальной ВО (NVO) США, проекта АстроГрид Великобритании, проекта Европейской ВО Euro-VO. Координация проекта РВО с международными проектами — основная стратегическая линия развития РВО. Основной координирующей организацией создания проектов ВО в различных странах мира является Альянс Международной ВО (IVOA) [2], в который входит и Россия. Альянс Международной ВО разрабатывает совокупность стандартов, которые должны обеспечить совместимость ВО, создаваемых в различных странах мира. Технически инфраструктура, продвигаемая IVOA, является сервис-ориентированной. В стандартах различаются сервисы доступа к астрономическим данным, вычислительные сервисы обработки данных, сервисы регистрации разнообразных сервисов, позволяющие публиковать и обнаруживать их. Подробный анализ разрабатываемых стандартов IVOA дан в [2, 3]. Слабым местом стандартов IVOA является недостаточность средств интеграции неоднородных астрономических ресурсов для решения над ними разнообразных научных задач. Поэтому в инфраструктуре РВО дополнительно к инфраструктуре IVOA вводится концепция предметных посредников, позволяющих задавать определение прикладных областей для формулирования и решения классов научных задач в терминах понятий этих областей, структур информационных объектов, декларативно объявляемых сервисов и процессов. Посредники располагаются между исследователями, формулирующими задачи в терминах посредников, и разнообразными распределенными информационными ресурсами (данными, сер-

висами, процессами), необходимыми для решения задачи.

В итоге инфраструктура РВО определена многоуровневой [1] и включает (при движении снизу вверх) уровень информационных ресурсов, промежуточный слой, обеспечивающий интероперабельность ресурсов благодаря технической унификации их интерфейсов и введению дистанционных механизмов обращения к ресурсам (адаптеры, осуществляющие преобразование запросов, выраженных в канонической информационной модели посредников, в их представление в информационной модели ресурса, включаются в состав промежуточного слоя), уровень *предметных посредников*, каждый из которых создает спецификацию предметной области для решения некоторого класса задач, используя каноническую информационную модель («эсперанто») для представления семантики предметной области и унифицированного отображения разнообразных видов информационных моделей ресурсов (моделей данных, сервисных моделей, онтологических моделей, процессных моделей), уровень задач (приложений), формулируемых в терминах одного или нескольких посредников. Следует заметить, что посредники, в свою очередь, могут рассматриваться как информационные ресурсы для использования в посредниках более высокого уровня. В такой инфраструктуре предметные посредники играют ключевую роль для решения семантических проблем, перечисленных выше. В частности, при интеграции неоднородных ресурсов в посреднике нужно уметь семантически отождествлять объекты, представленные в различных информационных моделях, и семантически правильно отображать схемы интегрируемых ресурсов в схему посредника. Поскольку в общем случае ресурсы неоднородны (представлены в различных моделях), при интеграции неоднородных ресурсов для однородного представления их семантики требуется приведение различных информационных моделей к унифицированному виду в рамках некоторой унифицирующей информационной модели, которая называется *канонической*.

Для унификации разнородных спецификаций прежде всего требуется умение сопоставлять спецификации различных ресурсов друг с другом так, чтобы можно было отвечать на вопрос, можно ли при реализации посредника использовать спецификацию существующего ресурса вместо фрагмента спецификации посредника. Для этого достаточно доказать, что рассматриваемые спецификации находятся в отношении уточнения. Говорят, что спецификация  $A$  уточняет спецификацию  $D$ , если  $A$  можно использовать вместо  $D$  так, что пользователь  $D$  не будет замечать этой замены. Средств

ва доказательства факта уточнения спецификации некоторого компонента спецификацией другого компонента (реализуемые на основе теоретико-модельных нотаций и соответствующего инструментария [4–6]) составляют фундамент применяемых методов конструирования унифицирующих (канонических) моделей представления информации в посредниках. Каноническая информационная модель служит в качестве общего языка, эсперанто, для адекватного выражения семантики разнородных моделей представления информации, используемых в разнообразных информационных ресурсах, релевантных посреднику. Методы отображения информационных моделей и синтеза расширяемых канонических информационных моделей для слоя предметных посредников подробно рассмотрены в [7]. Там же предложена архитектура Унификатора информационных моделей.

Задачей Унификатора моделей является унификация множества информационных моделей (называемых *исходными*), использующихся в различных информационных ресурсах некоторой информационной системы. Унификация исходной модели  $R$  есть приведение ее к канонической информационной модели  $C$ , т.е. создание такого расширения  $E$  канонической модели (которое может быть и пустым) и такого отображения  $M$  исходной модели в расширенную каноническую, что исходная модель *уточняет* расширенную каноническую модель. Уточнение моделей означает, что для любой допустимой спецификации  $r$  модели  $R$  ее образ  $M(r)$  при отображении  $M$  уточняется спецификацией  $r$ . Синтез канонической модели заключается в построении указанных расширений ядра канонической модели для всех исходных моделей и объединении таких расширений. В качестве ядра канонической информационной модели применяется язык СИНТЕЗ [8], который ориентирован на спецификацию предметных посредников и синтез канонических моделей. В настоящей работе Унификатор моделей предполагается входящим в качестве неотъемлемой части слоя предметных посредников. Его архитектура и подробности конструирования определены в [7].

Другой конкретной совокупностью семантических проблем создания посредников для ВО являются проблемы (1), (2), (5) и (6) из приведенного выше списка. Эти проблемы должны решаться при определении предметного посредника и регистрации в нем релевантных посреднику информационных ресурсов.

Спецификация предметного посредника для класса задач включает определения понятий предметной области, выражаемых соответствующими онтологическими спецификациями, специфика-

ции классов объектов предметной области, спецификации типов экземпляров названных классов и их методов, определяющих их поведение, спецификации процессов решения задач данного класса как совмещенных во времени последовательностей действий, реализуемых методами классов, сервисами и другими процессами. Предполагается, что такие спецификации преобразуются в спецификации канонической модели, имеющей формальную семантику. Результат определения предметного посредника, выполняемого заинтересованным научным сообществом, составляет спецификацию посредника, образуемую в результате достижения консенсуса в таком сообществе, а сама деятельность по спецификации посредника называется периодом его *консолидации*.

Регистрация релевантных посреднику ресурсов рассматривается как задача композиционного проектирования систем [9, 10]. Регистрация ресурсов есть процесс целенаправленной трансформации спецификаций, включающий декомпозицию спецификаций посредника на непротиворечивые фрагменты, поиск среди спецификаций релевантных ресурсов подходящих типов данных — кандидатов для уточнения ими спецификаций типов посредника, построение выражений, определяющих классы ресурсов в виде композиции классов посредника. Для подобного манипулирования спецификациями разработано специальное *исчисление спецификаций* [11]. В нем предложен принцип декомпозиции спецификаций типов в набор редуков спецификаций типов, служащих основными единицами повторного использования и композиции. Введена операция определения *наибольшего общего редукта* спецификаций типов ресурсов и посредника. На основе частично упорядоченного множества спецификаций типов определены также решетка и алгебра типов. Эти структуры послужили теоретической базой для разработки репозитория метаинформации, хранящего спецификации посредников, ресурсов, а также промежуточные спецификации, возникающие в процессе отображения таких определений.

Принципиальным моментом в этой схеме является реализация доказательства уточнения фрагментов спецификаций посредника спецификациями ресурсов в процессе построения отображений таких спецификаций. Идентификация релевантных ресурсов (предшествующая регистрации) основана на использовании трех моделей — *модели метаданных*, характеризующих свойства ресурсов, собранных в некотором реестре, *онтологической модели*, позволяющей формально определять понятия

предметной области посредника или ресурса, и *канонической модели*, позволяющей формально определять структуру и поведение объектов предметной области посредника и информационных ресурсов. Рассуждения в канонической и в онтологической модели основаны на семантике канонической модели и средствах доказательства уточнения. При этом в онтологической модели необходимо достичь согласования понятийной семантики спецификаций посредника и регистрируемых в нем ресурсов. Рассуждения в модели метаданных являются эвристическими на основе нефункциональных требований к нужным в классе задач ресурсам (к таковым относятся, в частности, показатели качества данных в ресурсах). Необходимые модели метаданных и алгоритмы поиска составляют часть метода. В целях проектирования спецификации посредника и ресурсов задаются в однородном их представлении в канонической модели, хотя для этого может потребоваться преобразование в такую модель из некоторого другого языка спецификаций, например из UML<sup>1</sup>.

Сложной проблемой регистрации ресурсов в посреднике является согласование прикладных контекстов посредника и конкретных ресурсов. В разработанном инструментарии регистрации [10] такое согласование осуществляется на основе онтологического подхода. Онтологические спецификации играют роль «клея» фрагментов спецификаций посредника и ресурсов для их семантического отображения и композиции. Онтологические определения аннотируют элементы спецификаций посредника и спецификаций ресурсов (заданных в форме типов, классов, процессов). В действующем подходе используется вербальная (задаваемая подобно определениям терминов в толковом словаре) форма определения понятий. Вербальное представление онтологии дополняется более формальными спецификациями на основе абстрактных типов данных канонической модели и техники доказательства уточнения.

Перечисленные подходы положены в основу разработанного прототипа средств идентификации и регистрации информационных ресурсов в посреднике [10]. При этом онтологические спецификации используются для идентификации классов посредника, семантически релевантных классам ресурса. Максимальное подмножество информации класса ресурса, релевантное классу посредника, устанавливается на основании максимального общего фрагмента спецификаций соответствующих типов экземпляров этих классов. Конкретизирующие типы, устраняющие возникающие конфликты (зна-

<sup>1</sup>UML — Unified Modeling Language — унифицированный язык моделирования.

чений, структур данных и поведения) в названных типах экземпляров, определяются так, чтобы тип экземпляра класса посредника уточнялся бы типом экземпляра класса ресурса. Основным результатом регистрации является выражение, определяющее, как класс ресурса выражается через посредство классов посредника. Архитектура и принципы построения инструмента регистрации ресурсов в посреднике, который является неотъемлемой частью слоя посредников, определены в [10] и рассматриваются в деталях в настоящей статье не будут.

Существенным для данной работы является то обстоятельство, что в Великобритании развивается система АстроГрид [12] как полная инфраструктура для создания систем решения научных задач в астрономических ВО. Инфраструктура АстроГрида позволяет конструировать гибкие распределенные структуры ВО, компонентами которых могут быть разнообразные средства хранения данных и доступа к ним, средства размещения и временного хранения файлов в процессе решения задач группами пользователей, реестры метаданных, в которых регистрируются ресурсы ВО, средства программирования приложений и организации их интероперабельного исполнения, средства создания и исполнения потоков работ, которые позволяют решать сложные задачи в распределенной системе. Разнообразные подмножества такой совокупности компонентов могут быть установлены на различных машинах. Взаимодействие между компонентами основано на использовании технологии Веб-сервисов, а сами компоненты являются Веб-приложениями. Компоненты удовлетворяют стандартам IVOA там, где такие стандарты уже существуют. В случае, если стандартов еще нет, АстроГрид предлагает проекты стандартов для рассмотрения IVOA.

По поддерживаемым функциям сервисы АстроГрида, в основном, соответствуют требуемым для реализации инфраструктуры РВО. АстроГрид установлен в Москве для использования астрономами в качестве Центра коллективного пользования [13]. Вместе с тем, АстроГрид не предусматривает развитых средств решения задач над множествами неоднородных информационных ресурсов, подобных предметным посредникам, поэтому в 2006 г. при поддержке РФФИ был инициирован проект создания гибридной архитектуры, объединяющей возможности АстроГрида с архитектурой предметных посредников. Целью настоящей статьи является обзор достигнутых в этом проекте результатов развития такой гибридной архитектуры, акцентируя новые методы и средства определения предметных областей, информационных ресурсов и решения задач в гибридной инфраструктуре посредников и АстроГрида.

Поскольку создание гибридной инфраструктуры касается, главным образом, сопряжения исполнительных механизмов двух инфраструктур, в статье им и уделено основное внимание. Дальнейшее изложение материала в статье структурировано следующим образом. В разд. 2 рассматриваются основные виды предметных посредников и методология интеграции неоднородных ресурсов в них. В разд. 3 определена архитектура исполнительных механизмов предметных посредников. В разд. 4 особое внимание уделено переписыванию запросов к посредникам в планы их реализации над конкретными информационными ресурсами. В разд. 5 рассматриваются особенности инфраструктуры системы АстроГрид. В разд. 6 приведено краткое описание объединенной архитектуры АстроГрида и средств поддержки исполнительного слоя предметных посредников. Пример реализации простого предметного посредника для решения задач поиска далеких галактик в объединенной архитектуре дан в разд. 7. В разд. 8 дана краткая характеристика развития методов интеграции неоднородных источников информации в инфраструктурах ВО. Наконец, заключение статьи подводит итог обсуждению и намечает планы дальнейшего развития работы.

## 2 Виды предметных посредников и методология интеграции неоднородных ресурсов в посредниках

Одним из широко распространенных взглядов на ВО является рассмотрение их как инфраструктур, предназначенных для интеграции данных и сервисов в различных исследовательских центрах, с целью предоставления таким образом всем ученым доступа к информации, необходимой для решения задач. Виртуальная обсерватория может рассматриваться как специальная модель межорганизационного взаимодействия. Основным аспектом моделирования такой распределенной деятельности в e-science, предусматривающей совместную работу организаций, является многообразие типов ресурсов для использования в процессе исследований и многообразие контекстов, в которых такие ресурсы могут рассматриваться. Часто такие ресурсы реализуются как унаследованные приложения. Перспективные подходы моделирования в e-science основаны на фундаментальной стратегии интеграции, при которой тот или иной фрагмент информации извлекается из некоторого исследовательского центра и помещается в иной контекст — контекст конкрет-

ной задачи. Формирование и реализация нужного контекста задачи и моделирование интегрируемой информации в этом контексте являются основными проблемами определения посредников.

Следующие соображения будут затронуты в ходе этого обсуждения. Различные приложения должны поддерживаться ВО, каждое из них работает в собственном, конкретном контексте (возможно, пересекающемся с контекстами других приложений), который должен быть определен семантически. Неоднородные информационные ресурсы различных видов (ресурсы данных, сервисные ресурсы, процессные ресурсы, онтологические ресурсы), релевантные ВО, должны использоваться в контексте конкретного приложения. Многие такие ресурсы автономны и эволюционируют во времени. Множество ресурсов, релевантных конкретному классу задач, может изменяться весьма быстро. Технологии, применяемые к релевантным ресурсам, также быстро эволюционируют. Поэтому доказательно правильная идентификация ресурсов, релевантных задаче, достижение семантической интеграции различных видов ресурсов в контексте конкретного приложения, достижение стабильности спецификации проблемной области в быстро развивающемся мире разнообразных ресурсов являются трудно разрешимыми задачами. Требуется создание новых методов и средств разработки приложений ВО над множествами распределенных неоднородных ресурсов.

Различаются два принципиально различных подхода к проблеме интегрированного представления описания предметной области задачи по отношению к множеству релевантных задаче информационных ресурсов: (1) двигаясь от ресурсов к задачам (схема посредника образуется как интегрированная схема множества ресурсов независимо от приложения) и (2) двигаясь от приложения к ресурсам (описание предметной области приложения образуется независимо от ресурсов (в терминах понятий, структур данных, функций, процессов), а затем релевантные приложению ресурсы отображаются в это описание). Первый подход, *движимый информационными ресурсами*, является немасштабируемым по отношению к числу ресурсов, не дает возможности достижения семантической интеграции ресурсов в контексте конкретного приложения, не ведет к доказательной идентификации релевантных приложению ресурсов, не способствует повышению стабильности спецификации посредника в процессе эволюции ресурсов, релевантных приложению. Эти недостатки являются характерными для подхода, при котором глобальная схема является взглядом (Global as View — GAV) [14, 15]. Схема GAV может служить в качестве базовой тех-

ники подхода, движимого информационными ресурсами.

Другой подход (*движимый приложениями*) предполагает создание предметного посредника, который поддерживает взаимодействие между приложением и ресурсом на основе определения прикладной области (определения посредника). Второй подход имеет очевидные преимущества по отношению к подходу, движимому информационными ресурсами. Процесс регистрации неоднородных информационных ресурсов в предметном посреднике в подходе, движимом приложениями, основан на технике GLAV, комбинирующей два подхода: LAV (Local as View) — когда локальная схема ресурса является взглядом над схемой посредника — и GAV. Согласно LAV [15], схемы регистрируемых ресурсов рассматриваются как материализованные взгляды над виртуальными классами посредника. В этом случае GAV взгляды служат для разрешения различных конфликтов между спецификациями ресурсов и посредника и обеспечивают правила трансформации результатов запроса от ресурса в представление в посреднике. Подобная техника регистрации обеспечивает стабильность спецификации приложения ВО при изменении конкретных информационных ресурсов и их фактического присутствия (удаление ресурса, добавление новых ресурсов и пр.), а также масштабируемость посредников по отношению к числу ресурсов, регистрируемых в них. В версии LAV, именуемой GLAV [16], голова определения правила взгляда LAV может содержать произвольный запрос над схемой ресурса и тем самым выразить случай, когда схема ресурса используется для определения составных частей глобальной схемы (GAV).

Основными принципами разработки посредников ВО, движимых приложениями, над множеством неоднородных информационных ресурсов, являются следующие:

- (1) независимость спецификации приложения (посредника) от существующих информационных ресурсов;
- (2) определение предметного посредника как результата консолидации усилий соответствующего сообщества;
- (3) применение семантических канонических определений в спецификации посредника;
- (4) независимость интерфейсов пользователя от ресурсов, регистрируемых в посреднике; пользователи должны только знать определение предметной области (определение посредника);
- (5) независимость публикации вновь разработанных информационных ресурсов от посредни-

ков (как результат такой публикации могут быть инициированы действия по регистрации ресурсов в соответствующем посреднике);

- (6) трехстадийная идентификация информационных ресурсов, релевантных посреднику, обеспечивающая (а) релевантность ресурсов нефункциональным требованиям, выражаемых метаданными; (б) онтологическую релевантность ресурсов; (в) структурную и поведенческую релевантность ресурсов;
- (7) семантическая интеграция релевантных неоднородных информационных ресурсов в канонической спецификации посредника;
- (8) интегрированный доступ к информационным ресурсам, зарегистрированным в посреднике, на основе канонической модели и системы переписывания запросов;
- (9) рекурсивная структура посредника: каждый посредник может быть зарегистрирован как новый информационный ресурс. Такая способность посредника обеспечивает возможность интеграции различных приложений, что особенно важно для разработки виртуальных организаций.

Принципы (1)–(8) определяют преимущества разработки ВО на основе идеи посредников, применения подход, движимый приложениями, по сравнению с подходом, движимым информационными ресурсами.

Настоящая статья основана, главным образом, на подходе, движимом приложениями.

### 3 Архитектура исполнительных механизмов слоя предметных посредников

Основные решения реализации исполнительных средств слоя посредников заключаются в следующем:

- предметный посредник представляет собой Веб-сервис;
- адаптеры неоднородных ресурсов, зарегистрированных в посреднике, представляют собой Веб-сервисы;
- посредники и адаптеры регистрируются в реестрах;

- один адаптер может работать с несколькими посредниками и с несколькими ресурсами;
- адаптеры могут содержать собственную систему управления базами данных (СУБД) для выполнения сложных планов запросов;
- связь между посредником и адаптерами обеспечивается посредством SOAP<sup>1</sup> протокола;
- база метаинформации посредников и вычисление остаточных запросов реализуются на СУБД Oracle;
- для задания запросов к посреднику пользователь может использовать либо Веб-портал, либо специальный клиент.

На рис. 1 показана общая архитектура исполнительных средств слоя посредников. Она содержит следующие компоненты.

**Application Client.** Клиентская программа, предоставляющая пользователю графический интерфейс для работы с посредником, включающий:

- графические дружественные пользователю средства, обеспечивающие поддержку применения канонического языка запросов на уровне посредников;
- поддержку запросов к посреднику на языке Syfs<sup>2</sup> [8];
- услуги по удобному представлению выводимой информации, включая сегментацию информационных объектов, их агрегирование и слияние.

**Portal.** Предоставляет доступ к посреднику через Интернет. По набору функций эквивалентен обычному клиенту.

**Run-time environment.** Исполнительная среда слоя посредников обеспечивает:

- обработку программ (запросов) пользователя;
- переписывание программ в запросы над зарегистрированными в посреднике ресурсами, отправку переписанных запросов на выполнение соответствующим адаптерам;
- получение результатов запросов от адаптеров;
- выполнение остаточного запроса и отправку результата запросов клиенту.

Исполнительная среда включает следующие компоненты:

<sup>1</sup>SOAP — Simple Object Access Protocol — протокол обмена структурированными сообщениями в распределенной вычислительной среде.

<sup>2</sup>Syfs — подмножество языка формул в языке СИНТЕЗ.

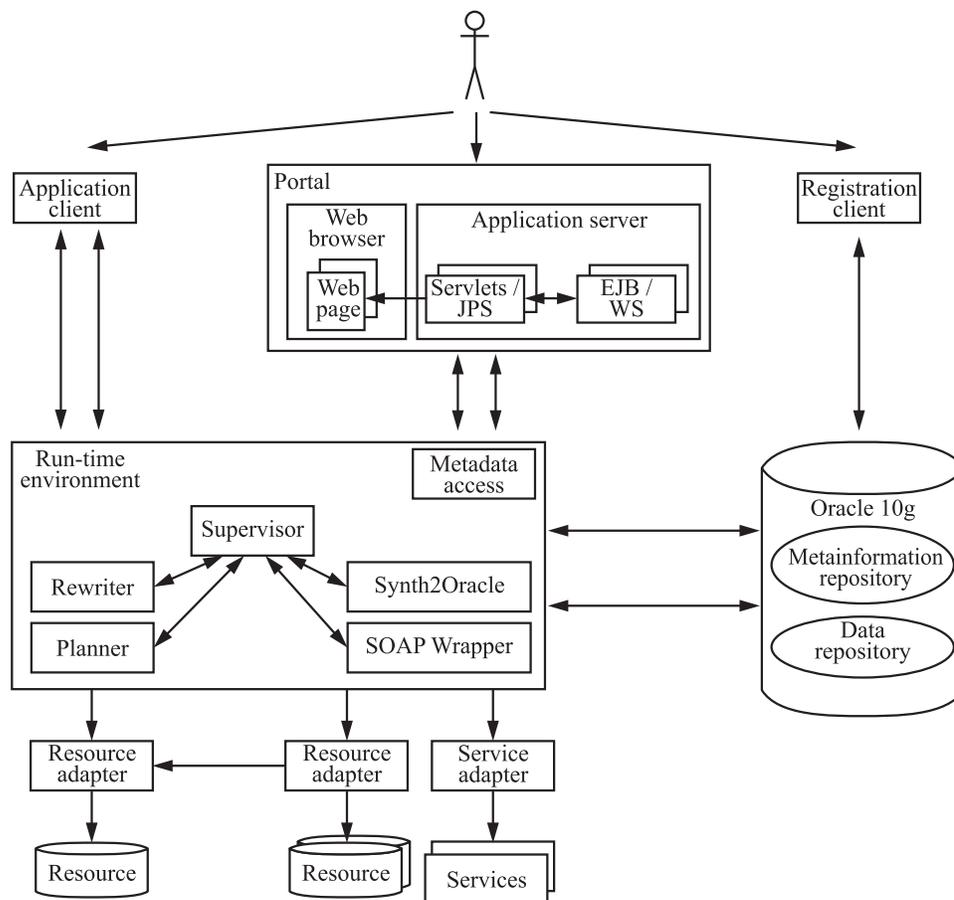


Рис. 1 Архитектура исполнительных средств слоя посредников

**Supervisor** является связующим компонентом посредника и определяет его интерфейс. Supervisor обрабатывает запросы пользователя и для их выполнения взаимодействует с другими компонентами посредника. Фактически код данного компонента реализует взаимодействие всех компонентов в посреднике.

**Rewriter** реализует функцию переписывания запроса пользователя на языке Syfs над схемой посредника в запрос над зарегистрированными ресурсами, который выражается на языке Asyfs<sup>1</sup> и передается компоненту Planner для дальнейшей обработки.

**Planner** реализует функцию планирования запроса над ресурсами. В результате планирования строится распределенный запрос, состоящий из удаленных запросов к отдельным ресурсам. Оптимизированная последовательность выполнения удаленных запросов определяет план, имеющий вид «дерева выполнения».

**SoapWrapper** инкапсулирует детали физической реализации сервисов, которые используют компонент Supervisor и адаптеры. Таким образом, адаптеры могут быть реализованы как Веб-сервисы либо как классы локальных программ.

**Synth2Oracle** реализует поддержку канонической модели данных посредника над базой Oracle в период исполнения. Этот компонент позволяет загружать и выгружать данные, которые в канонической модели представляются как набор классов, преобразованный в набор таблиц базы Oracle. Данные загружаются и выгружаются в формате VOClass<sup>2</sup>. Компонент Synth2Oracle реализует также поддержку выполнения остаточных запросов на языке Asyfs, транслируя их в язык SQL.

**Adapter.** Адаптеры связывают посредник с конкретными ресурсами. В адаптерах происходит преобразование запросов на языке посредника (Asyfs) в язык конкретного ресурса и, наоборот, преобразо-

<sup>1</sup>Asyfs — алгебраический вариант языка Syfs.

<sup>2</sup>VOClass — аналог представления таблиц в стандарте IVOA (в соответствии с XML Schema) для классов языка СИНТЕЗ.

вание результата запроса, полученного в формате, выдаваемом ресурсом, в формат посредника. Дополнительно, если это позволяет ресурс, адаптер реализует возможность загрузки в базу ресурса промежуточных классов периода выполнения запросов, которые используются в выполняемом адаптером запросе.

**Metainformation and Data repository.** Репозиторий состоит из двух частей: репозиторий метаданных и репозиторий данных. Репозиторий метаданных используется для хранения метаинформации посредника (такой как схема посредника, схемы зарегистрированных в посреднике ресурсов, правила, задающие соответствия между классами ресурсов и классами схемы посредника посредством GLAV), используемой при преобразовании запросов. Репозиторий данных используется для хранения промежуточных результатов запросов, возвращаемых ресурсами, и для выполнения остаточных запросов над этими данными.

## 4 Методы переписывания запросов и планирования в предметных посредниках

### 4.1 Основные характеристики ядра канонической модели

В настоящем проекте в качестве ядра канонической модели посредника выбран язык СИНТЕЗ [8], который является гибридной слабоструктурированной/объектной моделью [17]. В этом разделе приводятся основные особенности языка, необходимые для понимания примеров, демонстрирующих идеи переписывания запросов.

Рассматриваемое композиционное исчисление спецификаций [11] не зависит от какой-нибудь определенной информационной модели. Каноническая модель [8] предоставляет поддержку широкого диапазона данных — от слабоструктурированных до строго типизированных данных. Типизированные данные являются экземплярами *абстрактных типов данных* (АТД). Спецификация АТД включает имена и типы параметров операций, определяющих поведение типа. Операции определяются также предикативными спецификациями, описывающими смешанные пред- и постусловия. Абстрактные типы данных могут быть объектными и неobjектными. Объектный тип является подтипом неobjектного АТД с дополнительной операцией *self* на своем интерфейсе, предоставляющей уникальные идентификаторы объектов. В данной статье

используются только типизированные возможности языка СИНТЕЗ. Обсуждение переписывания запросов со слабоструктурированными (фреймовыми) данными планируется в будущих работах. *Множества* в языке (наряду с мультимножествами, последовательностями) рассматриваются в качестве общего механизма группирования значений АТД. Класс рассматривается как подтип типа множества. Поэтому эти, в общем, различные конструкции могут использоваться единым образом: класс может использоваться везде, где может использоваться множество. Например, для формул языка запросов исходные и результирующие данные представляются множествами значений АТД (т. е. коллекциями) или множествами объектов (т. е. классами).

#### 4.1.1 Операции над спецификациями типов

Семантика операций над классами в канонической модели объясняется в терминах композиционного исчисления спецификаций [11]. Действия в исчислении включают декомпозицию спецификаций типов в непротиворечивые фрагменты (редукты), выделение общих фрагментов, композицию выделенных фрагментов в более сложные спецификации соответствующих результирующих типов SPJ (select, project, join) операций. Исчисление использует следующие концепции и операции.

Сигнатура  $\Sigma_T$  спецификации типа  $T = \langle V_T, O_T, I_T \rangle$  включает множество операций  $O_T$ , задающих типы аргументов и результатов операций, и множество предикатов инвариантов типа  $I_T$ . Конъюнкция всех инвариантов из  $I_T$  формирует инвариант типа. Экстенционал  $V_T$  типа  $T$  (носитель типа) есть множество допустимых экземпляров типа.

**Определение 1.** *Редукт типа.* Сигнатура редукта  $R_T$  типа  $T$  определяется как подсигнатура  $\Sigma'_T$  сигнатуры типа  $\Sigma_T$ , которая включает носитель  $V_T$ , множество символов операций  $O'_T \subseteq O_T$ , множество символов инвариантов  $I'_T \subseteq I_T$ .

Это определение легко можно расширить с уровня сигнатур до уровня спецификаций таким образом, что тип редукта  $R_T$  может рассматриваться как подспецификация (с сигнатурой  $\Sigma_T$ ) спецификации типа  $T$ . Спецификация  $R_T$  должна формироваться так, что  $R_T$  становится супертипом типа  $T$ . Предполагается, что только состояния, допустимые для типа, остаются допустимыми для редукта этого типа (никакие другие состояния не являются допустимыми). Таким образом, носитель редукта предполагается равным носителю типа.

**Определение 2.** *Наиболее общий редукт  $R(T_1, T_2)$*  для типов  $T_1$  и  $T_2$  — это редукт  $R_{T_1}$  типа  $T_1$  такой,

что существует редукт  $R_{T_2}$  типа  $T_2$  такой, что  $R_{T_2}$  уточняет  $R_{T_1}$ , и не существует другого редукта  $R'_{T_1}$  такого, что  $R(T_1, T_2)$  является редуктом  $R'_{T_1}$ ,  $R'_{T_1}$  не равен  $R(T_1, T_2)$ , и существует редукт  $R'_{T_2}$  типа  $T_2$ , который уточняет  $R'_{T_1}$ .

Тип  $C$  является *уточнением* типа  $R$  тогда и только тогда, когда:

- существует биективное отображение  
 $Ops: O_R \rightarrow O_C$ ;
- существует всюду определенная функция абстракции  $Abs: V_C \rightarrow V_R$ ;
- для всех  $\nu \in V_C$ :  $Inv_C(\nu)$  влечет  $Inv_R(Abs(\nu))$ ;
- для всех  $o \in O_R$ ,  $o$  уточняется  $Ops(o)$ .

Операция  $o_1$  уточняет  $o_2$  тогда и только тогда, когда предусловие операции  $pre(o_2)$  влечет предусловие  $pre(o_1)$  и постусловие  $post(o_1)$  влечет постусловие  $post(o_2)$ . Для сигнатур операций уточнение означает, что сигнатуры совпадают с точностью до следующего отношения типов параметров: типы входных параметров операции  $o_1$  являются супертипами типов соответствующих входных параметров операции  $o_2$  и типы выходных параметров операции  $o_1$  являются подтипами типов соответствующих выходных параметров операции  $o_2$ .

Основываясь на понятиях редукта и уточнения типов, можно установить меру общей информации между типами в решетке типов  $T$ . Отношение *подтипа* определяется подобно уточнению, только  $Ops$  становится инъективным отображением.

**Определение 3.** *Операция пересечения типов.* В общем случае тип  $T$  — результат операции  $T_1 \& T_2$  — включает атрибуты, общие для типов  $T_1$  и  $T_2$ . Общие элементы типов определяются с помощью наиболее общих редуктов  $R(T_1, T_2)$  и  $R(T_2, T_1)$ . Формально  $O_{T_1 \& T_2}$  определяется следующим образом:

$$O_{T_1 \& T_2} = O_{R(T_1, T_2)} \cup O_{R(T_2, T_1)}.$$

Инвариант типа  $T$  определяется как дизъюнкция инвариантов операндов  $Inv_{T_1} \vee Inv_{T_2}$ .

Можно заметить, что если  $T_2(T_1)$  является подтипом типа  $T_1(T_2)$ , то  $T_1(T_2)$  является результатом пересечения типов. Тип  $T$  помещается в решетку типов как супертип типов операндов этой операции.

**Определение 4.** *Операция соединения типов.* В общем случае тип  $T$  — результат операции  $T_1 | T_2$  — включает объединение спецификаций типов  $T_1$  и  $T_2$ .

Общие элементы спецификаций типов  $T_1$  и  $T_2$  включаются в объединение (результатирующий тип)

только один раз. Общие элементы типов определяются с помощью наиболее общих редуктов  $R(T_1, T_2)$  и  $R(T_2, T_1)$ . Формально  $O_{T_1 | T_2}$  определяется как

$$O_{T_1 | T_2} = (O_{T_1} \setminus O_{R(T_1, T_2)}) \cup (O_{T_2} \setminus O_{R(T_2, T_1)}).$$

Инвариант типа  $T$  определяется как конъюнкция инвариантов операндов  $Inv_{T_1} \wedge Inv_{T_2}$ .

Можно заметить, что если  $T_2(T_1)$  является подтипом типа  $T_1(T_2)$ , то  $T_2(T_1)$  является результатом операции соединения типов. Тип  $T$  помещается в решетку типов как подтип типов операндов операции.

Операции композиционного исчисления формируют решетку типов [11] на основе отношения подтипа (как отношения частичного порядка). В языке СИНТЕЗ операции композиции типов используются для формирования типовых выражений, которые в языке используются для определения типа. Типовые выражения включают следующие конструкции:

- указатель типа ( $T$ );
- конструкция редукта ( $R[e_1, \dots, e_n]$ , где  $R$  — указатель типа,  $e_i$  — элементы редукта);
- выражения ( $T$ ),  $T|S$  и  $T \& S$ , где  $T$  и  $S$  — указатели типов.

Конструкция редукта имеет следующий вид:  $R[e_1, \dots, e_n]$ , где  $R$  — исходный АД,  $e_i$  — элементы редукта, каждый из которых определяет атрибут образуемого АД. Редукт определяет также функцию преобразования экземпляров исходного АД в экземпляры редукта. В общем случае элемент редукта имеет следующий вид:

$$a/T : t,$$

где  $a$  — идентификатор атрибута,  $T$  — тип,  $t$  — атрибутный путь (вида  $a_1.a_2. \dots .a_k$ ). Доступны следующие сокращенные формы элементов редукта:

- $a : t$  — сокращение для  $a/T : t$ , где  $T$  — тип термина  $t$  (т.е. для пути  $a_1. \dots .a_k$  его типом является тип атрибута  $a_k$ );
- $a/T$  и  $a$  — сокращения для  $a/T : a$  и  $a : a$  соответственно.

Редукт  $R[a_1/T_1 : t_1, \dots, a_n/T_n : t_n]$  образует АД  $T_R$ , который содержит атрибуты  $a_1, \dots, a_n$ , типы которых соответственно:  $T_1, \dots, T_n$ . Функция преобразования экземпляров АД  $R$  в экземпляры редукта типа определяется следующим образом:

$$P(x) = \{y \in T_R | y.a_1 = x.t_1, \dots, y.a_n = x.t_n\}.$$

#### 4.1.2 Подмножество языка запросов СИНТЕЗ

В статье рассматривается ограниченное подмножество языка запросов СИНТЕЗ, ориентированное на представление объединения конъюнктивных запросов с SPJ семантикой.

Для определения формулы запроса используется разновидность типизированного (многоуровневого) языка логики предикатов первого порядка. Предикаты в формуле соответствуют коллекциям (таким как множества и мультимножества необъектных экземпляров), классам, которые рассматриваются как множества объектных значений, и функциям. Предикат класса (или предикат коллекции) всегда унарный предикат. Предикат функции  $F$  синтаксически имеет вид  $n$ -арного предиката  $F(X, Y)$ , где  $X$  — последовательность термов, соответствующих входным параметрам функции,  $Y$  — последовательность типизированных результирующих атрибутов (конструкция вида  $y/Y$ , где  $Y$  определяет тип атрибута  $y$ ), соответствующих выходным параметрам функции (результатирующие атрибуты принимают значения соответствующих выходных параметров). В качестве термов могут выступать: переменные, константы, атрибуты, функции (включая арифметические функции), конструкция взятия атрибута  $t.a$ , где  $t$  — терм и  $a$  — атрибут из АД, типизирующего  $t$ . Частным случаем функции является операция (метод), где первым входным параметром передается экземпляр АД, в котором определен метод.

**Определение 5.** Конъюнктивным запросом СИНТЕЗа (*Synthesis Conjunctive Query, SCQ*) является правило следующего вида

$$q(\nu/T_\nu) : -C_1(\nu_1/T_{\nu 1}) \& \dots \& C_n(\nu_n/T_{\nu n}) \&$$

$$F_1(X_1, Y_1) \& \dots \& F_m(X_m, Y_m) \& B,$$

где  $q(\nu/T_\nu)$ ,  $C_1(\nu_1/T_{\nu 1}), \dots, C_n(\nu_n/T_{\nu n})$  — предикаты коллекций (классов);  $F_1(X_1, Y_1), \dots, F_m(X_m, Y_m)$  — предикаты функций,  $B$  — ограниченное, представляющее собой конъюнкцию арифметических предикатов. Каждый предикат  $C_i(\nu_i/T_{\nu i})$  или  $F_j(X_j, Y_j)$  называется *подцелью*. Подцели и  $B$  называются *телом правила*,  $q(\nu/T_\nu)$  называется *головой правила*. Запрос SCQ также называется правилом. Семантика SPJ тела SCQ будет приведена далее.

В языке Syfs допускаются правила, представляющие собой объединение правил с SPJ семантикой. Несколько правил на языке Syfs составляют программу.

Общая схема вычисления результирующей коллекции тела SCQ имеет следующий смысл:

$$C_1(\nu_1/T_{\nu 1}) \& \dots \& C_n(\nu_n/T_{\nu n}) \&$$

$$F_1(X_1, Y_1) \& \dots \& F_m(X_m, Y_m) \& B.$$

Сначала выполняется соединение коллекций, которые определяются предикатами классов  $C_i(\nu_i/T_{\nu i})$ . Предикат класса  $C(\nu/T)$  формирует результирующую коллекцию, которая строится на основе проекции класса  $C$ , где проекция определяется типом  $T$ . В общем случае  $T$  является редуктом  $R[a_1/T_1 : t_1, \dots, a_n/T_n : t_n]$ , и к экземплярам класса  $C$  применяется определяемая редуктом функция преобразования. Типом результирующей коллекции является АД  $R_{C(\nu/T)}$  с атрибутами  $\nu, a_1, \dots, a_n$ , типы которых есть, соответственно:  $T, T_1, \dots, T_n$ .

Множество экземпляров результирующей коллекции является следующим:

$$\{y \in R_{C(\nu/T)} \mid \exists x \in C :$$

$$y.\nu = x, y.a_1 = x.t_1, \dots, y.a_n = x.t_n\}.$$

Результатом соединения двух коллекций  $C_1$  и  $C_2$  с типами экземпляров  $T_1$  и  $T_2$ , соответственно, является коллекция  $C_1 \& C_2$ , тип экземпляров которой равен  $T_1|T_2$ , и множество экземпляров определяется следующей формулой:

$$C_1 \& C_2 = \{x \in T_1|T_2 \mid \exists y_1 \in C_1, y_2 \in C_2 :$$

$$x = y_1, x = y_2\}.$$

Затем по очереди вычисляются функции  $F_j$ , и к результирующей коллекции присоединяются (are appended) результирующие атрибуты.

Вычисление функции выполняется для каждого экземпляра исходной коллекции. Значения атрибутов текущего экземпляра могут использоваться при вычислении значений термов, аргументов для входных параметров функции, и текущий экземпляр расширяется результирующими атрибутами предиката функции.

После этого выбираются только те экземпляры результирующей коллекции, для которых выполняется ограничение  $B$ . Результирующий класс  $q$  формируется как проекция результирующей коллекции тела, которую описывает редукт  $T_\nu$  в голове правила.

Два запроса SCQs с головами  $q_1(\nu_1/T_1)$  и  $q_2(\nu_2/T_2)$  называются *сравнимыми*, если тип  $T_1$  является подтипом типа  $T_2$ . Для двух сравнимых запросов  $q_1$  и  $q_2$  запрос  $q_1$  *содержится* в запросе  $q_2$  (обозначается  $q_1 \subseteq q_2$ ), если для любого экземпляра базы данных все экземпляры результирующего класса  $q_1$  являются экземплярами результирующего класса  $q_2$ , т. е. результирующий класс  $q_1$  является подклассом результирующего класса  $q_2$ .

## 4.2 Регистрация ресурсов в посреднике и инверсные правила

В предметном посреднике процесс регистрации неоднородных информационных ресурсов основан на подходе LAV/GLAV. Подход GLAV позволяет голове взгляда LAV содержать элемент, выраженный запросом над схемой ресурса, и таким образом выражать случай, когда схемы ресурса используются для определения конструкций глобальной схемы (GAV). Подход GAV позволяет задавать функции разрешения различных видов конфликтов между спецификациями ресурса и посредника. Он также позволяет задавать правила преобразования результатов запросов от ресурса в посредник.

Таким образом, процесс регистрации заключается в конструировании множества взглядов GLAV, которые имеют следующий вид:

- правило LAV

$$V(\nu/T_\nu) :- Cm_1(\nu_1/T_{\nu 1}) \& \dots \& Cm_n(\nu_n/T_{\nu n}) \& Fm_1(X_1, Y_1) \& \dots \& Fm_m(X_m, Y_m) \& B_m ;$$

- правило GAV

$$V(\nu/T_\nu) :- Cs_1(\nu_1/T_{\nu 1}) \& \dots \& Cs_n(\nu_n/T_{\nu n}) \& Fs_1(X_1, Y_1) \& \dots \& Fs_m(X_m, Y_m) \& B_s ,$$

где  $Cm_i$  — предикат класса посредника;  $Fm_i$  — предикат функции посредника;  $Cs_i$  — предикат класса ресурса;  $Fs_i$  — предикат функции ресурса или функция разрешения конфликтов;  $B_m$  и  $B_s$  — ограничения.

Формально переписывание запросов основано на понятии инверсного правила (ИП). Рассматриваются два вида инверсных правил: инверсное правило класс в класс и инверсное правило функция в функцию. Инверсное правило класс в класс получается из взгляда LAV как пара головы взгляда и подцели взгляда. Инверсное правило класс в класс имеет следующий вид:

$$C(x/T) \leftarrow V(-/R) ,$$

где предикат  $C(x/T)$  — голова инверсного правила и  $V(-/R)$  — тело инверсного правила (символ подчеркивания обозначает анонимную переменную). Для взгляда

$$V(-/R) :- C_1(x_1/T_1) \& \dots \& C_n(x_n/T_n) \& F_1(t_1, y_1) \& \dots \& F_m(t_m, y_m) \& B$$

конструируются следующие инверсные правила класс в класс:

$$C_1(x_1/T_1) \leftarrow V(-/R), \dots, C_n(x_n/T_n) \leftarrow V(-/R) .$$

Инверсные правила определяют, как отдельные подцели класса, выраженные в терминах схемы посредника, могут быть переписаны в подцели, выраженные в терминах взглядов GAV и схем ресурсов. Например, для того чтобы получить переписывание правила, которое содержит вхождения  $C(x/T)$ , можно заменить вхождения  $C(x/T)$  вхождениями  $V(-/R)$ , если при этом существует инверсное правило  $C(x/T) \leftarrow V(-/R)$  и редукты  $T$  и  $R$  задают типы, связанные так, что  $R$  уточняет  $T$ .

В общем случае редукт  $R$  не включает все атрибуты, которые определяются в теле взгляда, и поэтому редукт  $R$  должен быть расширен в процессе сколемизации. Так как в нашей информационной модели атрибуты определяются как соответствующие *get*-функции, мы рассматриваем *get*-функции атрибутов, которые формируют расширение, как сколемовские функции (помечая их символом #). Например, во взгляде

$$\nu(-/R[a, b]) :- p(x/T_p[a, b]) \& q(y/T_q[b, c]) \& f(x, z)$$

редукт в голове не включает атрибуты  $x$ ,  $y$ ,  $c$  и  $z$ , которые определяются в теле взгляда, поэтому редукт расширяется сколемовскими атрибутами  $\#x$ ,  $\#y$ ,  $\#c$  и  $\#z$ :

$$R[a, b, x : \#x, y : \#y, c : \#c, z : \#z] .$$

Затем на основе сколемизированного взгляда строятся два инверсных правила:

$$p(x/T_p[a, b]) \leftarrow \nu(-/R[a, b, x : \#x])$$

и

$$q(y/T_q[b, c]) \leftarrow \nu(-/R[b, y : \#y, c : \#c]) .$$

Инверсные правила функция в функцию задаются непосредственно в процессе регистрации. Инверсное правило функция в функцию является формулой следующего вида:

$$all \ x_1/T_1, \dots, x_n/T_n (F(t, y) \leftarrow G(s, y)) ,$$

где  $x_1, \dots, x_n$  — переменные типов  $T_1, \dots, T_n$ ,  $t$  и  $s$  — термы с вхождениями этих переменных, а  $y$  — результирующий атрибут. Инверсное правило функция в функцию выражает отношение уточнения между двумя функциями таким образом, что тело инверсного правила  $G(s, y)$  уточняет голову инверсного правила  $F(t, y)$  как функций над переменными  $x_1, \dots, x_n, y$ . Инверсные правила функция в функцию используются для переписывания подцелей функции аналогично тому, как инверсные правила класс в класс используются для переписывания подцелей класса.

### 4.3 Переписывание запросов

Для поддержки подхода GLAV был разработан алгоритм переписывания запросов [18], состоящий из трех фаз: формальное переписывание и семантический анализ, реализующие подход LAV и третья фаза — разворачивание взглядов GAV. Для нерекурсивных программ запроса, состоящих из нескольких правил, переписывание применяется к каждому правилу в отдельности. Реализация подхода LAV основана на разновидности алгоритма инверсных правил [19], расширенного для типизированного языка запросов [20]. К результатам фазы формального переписывания применяется семантический анализ, который выполняет проверку на выполнимость правил и устраняет вхождения сколемовских функций.

Формальное переписывание применяется к телу правила. Для каждой подцели, которая не является встроенным предикатом, выбирается такое инверсное правило, голова которого унифицируется с этой подцелью. Унификация головы инверсного правила с подцелью также влечет следующее преобразование тела инверсного правила.

Инверсное правило класс в класс

$$C(x/T[x_1/T_1 : t_1, \dots, x_n/T_n : t_n]) \leftarrow \\ \leftarrow V(-/R[x/T : y, x_1/T_1 : y_1, \dots, x_n/T_n : y_n]),$$

где  $t_1, \dots, t_n$  — пути атрибутов, можно унифицировать с подцелью

$$C'(z/T'[z_1/S_1 : t_1.s_1, \dots, z_k/S_k : t_k.s_k]),$$

где  $k \leq n$ ,  $C'$  — подкласс класса  $C'$ ,  $S_i$  — подтип типа  $T_i$ , а  $s_1, \dots, s_2$  — пути атрибутов.

Преобразованное тело инверсного правила строится как следующий предикат класса:

$$V(-/R[z/T : y, z_1/S_1 : y_1.s_1, \dots, z_k/S_k : y_k.s_k]).$$

Инверсное правило функция в функцию

$$\text{all } x_1/T_1, \dots, x_n/T_n (F(t, y) \leftarrow G(s, y))$$

можно унифицировать с подцелью  $F(t', z)$ , если унификация термов  $t$  и  $t'$  может быть установлена со следующим наиболее общим унификатором как подстановкой

$$u = \{t'_1|x_1, \dots, t'_n|x_n, \dots, t_1|z_1, \dots, t_m|z_m\},$$

где переменные  $z_i$  имеют вхождения в терме  $t'$ ,  $t'_i$  — подтермы терма  $t'$ , а  $t_i$  — подтермы терма  $t$ . Преобразованное тело инверсного правила строится как следующая конъюнкция:

$$G(s', z) \& E,$$

где  $E = z_1 = t_1 \& \dots \& z_m = t_m, s'$  получается из  $s$  посредством применения  $u$ , т.е. вхождения переменных  $x_1, \dots, x_n$  заменяются на термы  $t'_1, \dots, t'_n$  соответственно, и вхождения переменных  $z_1, \dots, z_m$  заменяются на термы  $t_1, \dots, t_m$  соответственно. Например, инверсное правило функция в функцию  $\text{all } x/\text{real} (f(45, x, y) \leftarrow g(x, y))$  можно унифицировать с предикатом функции  $f(a, b, c)$  и с преобразованным телом инверсного правила:  $g(b, c) \& a = 45$ .

Для заданного правила

$$Q(-/R) : -C_1(x_1/T_1) \& \dots \& C_n(x_n/T_n) \& \\ F_1(t_1, y_1) \& \dots \& F_m(t_m, y_m) \& B$$

и набора инверсных правил такого, что унификация инверсного правила класс в класс с подцелью  $C_i(x_i/T_i)$  влечет преобразование тела правила  $V_i(-/R_i)$ , а унификация инверсного правила функция в функцию с подцелью  $F_j(t_j, y_j)$  влечет преобразование тела инверсного правила  $G_j(s_j, y_j) \& E_j$ , формируется следующее переписывание:

$$Q(-/R) : -V_1(-/R_1) \& \dots \& V_n(-/R_n) \& G_1(s_1, y_1) \& \dots \\ \dots \& G_m(s_m, y_m) \& E_1 \& \dots \& E_m \& B.$$

Построенное переписывание правила включается в исходное правило и таким образом формирует подзапрос.

### 4.4 Семантический анализ

Полученные в результате формального переписывания правила могут содержать сколемовские функции и невыполнимые ограничения, поэтому необходимо применить дополнительный анализ для того, чтобы образовать правила, не содержащие сколемовские функции, и отбросить невыполнимые правила. В ходе данной фазы отслеживаются атомарные значения всех термов из правила, которые привязываются к отдельным переменным.

Для терма  $t$ , который имеет не скалярный тип, такой как АД с атрибутами  $a_1, \dots, a_n$ , выполняется рекурсивно привязка к переменным термов  $t.a_1, \dots, t.a_n$ . Проекция и соединения транслируются во множество равенств над этими переменными. Предикаты функций транслируются в равенства вида  $f(\nu_1, \dots, \nu_n) = y$ .

Например, для правила

$$r(-/T[a, b]) : -p(-/T[a, b, c]) \& f(c, y)$$

атрибуты предиката головы привязываются к переменным  $\nu_1$  и  $\nu_2$ , атрибуты в предикате класса —

к переменным  $\nu_3, \nu_4, \nu_5$ , атрибут  $y$  — к переменной  $\nu_6$ . Правило транслируется в множество равенств:  $\nu_1 = \nu_3, \nu_2 = \nu_4, f(\nu_5) = \nu_6$ . Для построения конгруэнтных классов термов к множеству равенств применяется разновидность алгоритма конгруэнтного замыкания [21]. Множество конгруэнтных классов предоставляет удобную форму для представления множества равенств с учетом свойства транзитивности равенств.

Например, из множества равенств  $\nu_1 = \nu_3, \nu_2 = \nu_4, f(\nu_5) = \nu_6$  конструируется два конгруэнтных класса  $\nu_1, \nu_3$  и  $\nu_2, \nu_4$ . Анализ правила и взглядов LAV выполняется в терминах этого представления. Конгруэнтные классы, получаемые из взглядов LAV, на которые имеются ссылки в правиле (посредством предикатов классов), представляют равенства, имплицитные взглядами, и поэтому они безопасно могут быть исключены из множества равенств правила. Затем по множеству равенств восстанавливаются правила, которые уже не содержат сколемовские функции. Таким образом, реализуется устранение сколемовских функций в фазе семантического анализа.

Наряду с устранением сколемовских функций выполняется проверка выполнимости правил. Дополнительно собираются все арифметические предикаты, обнаруженные в ограничении правила и в ограничениях, использующихся в правиле LAV взглядов. Арифметические предикаты обрабатываются в терминах привязки к переменным, и следовательно, рассматриваются только переменные, константы и арифметические предикаты. Рассматриваются арифметические предикаты  $<, \leq, \geq, >$ , и для проверки выполнимости конструируется граф неравенств [22]. Также выделяется граф неравенств, который строится из арифметических предикатов, порождаемых взглядами, и используется для устранения арифметических предикатов, которые связаны со сколемовскими функциями.

## 4.5 Разворачивание взглядов GAV

Первые две фазы производят переписанные правила, которые могут ссылаться на взгляды GAV. Для того чтобы получить переписывания правил, выраженные в терминах локальных схем, выполняется разворачивание взглядов GAV. В нашем подходе проблема разворачивания взглядов GAV рассматривается в контексте ограниченного случая, так как в результате процесса регистрации образуются взгляды GAV, которые не имеют общих локальных функций и классов. Поэтому можно выполнять разворачивание каждого взгляда GAV

независимо. Разворачивание взгляда GAV в подцели правила выполняется как унификация головы взгляда GAV с этой подцелью. Унификация влечет преобразование тела взгляда GAV. Дополнительно, в качестве оптимизации, из преобразованного тела взгляда GAV удаляются все предикаты функции, результирующие атрибуты которых не используются. Полученным телом взгляда GAV заменяется вхождение данной подцели. Для того чтобы при подстановке тела взгляда замкнутые переменные и атрибуты (не встречающиеся в голове взгляда GAV) не вступали в конфликт с другими замкнутыми переменными правила, к идентификаторам замкнутых атрибутов и переменных приписывается уникальный суффикс.

Например, рассмотрим следующие два взгляда GAV:

$$r_1(-/[a, b, c]) : -p_1(x/[a, b, c])$$

$$r_2(-/[a, b]) : -p_2(x/[a, c]) \& p_3(y/[b, c])$$

и правило, для которого выполняется разворачивание этих взглядов:

$$r(-/[a]) : -r_1(/[a, c]) \& r_2(/[a, b]).$$

Результат разворачивания взглядов GAV будет следующим:

$$r(-/[a]) : -p_1(x_1/[a, b]) \& p_2(x_2/[a, c_1]) \& p_3(y_1/[b, c_1]).$$

## 4.6 Планирование распределенного запроса

В посреднике запрос переписывается в последовательность запросов, выраженных в терминах ресурсов, зарегистрированных в посреднике. Так как в реальности посредник имеет дело с удаленными неоднородными ресурсами, доступ к которым осуществляется посредством адаптеров к этим ресурсам, возникает необходимость в решении задачи планирования распределенного запроса. В процессе планирования выделяются подзапросы к отдельным ресурсам, а также остаточный запрос, который будет выполняться в посреднике над результатами, полученными от этих локальных ресурсов. В архитектуре посредника поддерживаются ресурсы, которые могут принимать внешние коллекции (т.е. предоставляют интерфейс для загрузки данных во временное хранилище ресурса), и позволяют использовать загруженные коллекции в запросах к этим ресурсам (в примере, который рассматривается в данной статье, в качестве ресурса с такими возможностями выступает ресурс SDSS<sup>1</sup>). Поэтому

<sup>1</sup>SDSS — Sloan Digital Sky Survey — Цифровой обзор неба Слоан.

в процессе планирования необходимо решить, куда передаются данные: от одного ресурса к другому или в посредник. Подробнее этап планирования запроса будет рассмотрен для примера задачи поиска далеких галактик.

## 5 АстроГрид как основа инфраструктуры Российской виртуальной обсерватории

Цель системы АстроГрид — поддержка инфраструктуры для решения научных задач в астрономических ВО. Для решения научных задач АстроГрид предоставляет средства доступа к астрономическим каталогам, цифровым обзорам и архивам изображений, а также к реестрам метаданных, в которых регистрируются ресурсы ВО. Система обеспечивает размещение файлов и хранение промежуточных результатов в процессе решения задач группами пользователей, средства доступа к реестрам метаданных, в которых регистрируются ресурсы ВО, средства создания и исполнения потоков работ, средства взаимодействия между различными внешними приложениями в распределенной среде ВО.

Система АстроГрид построена на основе архитектуры Веб-сервисов. Каждый компонент системы представляет собой Веб-сервис, который реализует определенную функцию. Веб-сервисы могут располагаться как на одном сервере, так и на разных серверах, что позволяет добиться максимальной гибкости и производительности. Далее рассматриваются основные компоненты АстроГрида.

**Registry** (реестр) представляет собой коллекцию метаданных — XML-документов, описывающих ресурсы, которые могут использоваться при решении задач с помощью ВО. Registry реализован на основе стандарта OAI PMH<sup>1</sup> [23], специализированного IVOA для нужд ВО. Поиск в Registry осуществляется по ключевым словам на этапе планирования и выполнения приложений. В качестве ресурсов могут выступать приложения, сервисы, информационные ресурсы, компоненты АстроГрида, а соответствующие им метаданные в Registry содержат сведения о том, где располагаются ресурсы, как они могут быть использованы, кем разработаны, для чего предназначены и т. д. Схемы для метаданных, хранящихся в Registry, а также интерфейсы сервисов определены спецификациями IVOA. Registry поддерживает функцию сбора информации о ресурсах, зарегистрированных в других реестрах ВО (harvesting).

**Community** обеспечивает регистрацию и персональную аутентификацию пользователей. Для управления учетными записями пользователей предоставляется Веб-интерфейс. Учетная запись пользователя позволяет использовать компонент АстроГрида MySpace, который может потребоваться для решения научных задач.

**MySpace** представляет собой виртуальное хранилище данных, к которым могут иметь доступ все сервисы системы АстроГрид. Компонент MySpace используется для хранения и передачи файлов между сервисами в рамках одной задачи и между задачами, решаемыми при помощи ВО.

**Common Execution Architecture** (CEA — Общая исполнительная архитектура) определяет способ оформления приложения в сервис АстроГрида. Приложения могут быть трех видов: Веб-сервисы, приложения командной строки unix, приложения Java. Приложение, оформленное как CEA, обладает рядом преимуществ:

- вызывается как из потоков работ, так и как настольное приложение, при этом не требуется изменения кода клиентских программ;
- регистрируется и обнаруживается в реестрах АстроГрида;
- приложение может считывать и записывать данные в MySpace;
- может выполняться в асинхронном режиме, что важно в случае продолжительных задач;
- приложение может быть защищено средствами безопасности системы АстроГрид, если требуется ограничить доступ к нему;
- приложение может использоваться без написания дополнительного кода: CEA предоставляет Веб-сервис, в который встраивается приложение.

**DataSet Access** (DSA) реализует подключение базы данных к системе АстроГрид. Компонент DSA предоставляет два интерфейса доступа: CEA-приложение, которое выполняет запросы к базе данных; Cone Search, который позволяет производить поиск в некоторой заданной области неба. В настоящее время для использования баз данных в DSA поддерживаются следующие СУБД: PostgreSQL, MySQL, MS SQL Server, HSQLDB.

**Astro Runtime** обеспечивает доступ к сервисам и ресурсам ВО посредством простого программного интерфейса, функции которого могут быть вызваны из любых приложений. Библиотеки

<sup>1</sup>OAI-PMH — Open Archives Initiative Protocol for Metadata Harvesting — архивный открытый протокол сбора метаданных.

для работы с Astro Runtime разработаны для ряда языков. Таким образом, обеспечивается возможность создания новых приложений, взаимодействующих с ВО.

**PLASTIC** (*PLatform for Astronomical Tool Inter Co-operation*) — это протокол, обеспечивающий взаимодействие между средствами АстроГрида и внешними программными инструментами в среде ВО. Посредством PLASTIC приложения могут обмениваться данными или управлять друг другом, например одно приложение может попросить другое загрузить изображение некоторой области неба и работать с полученным результатом запроса. Хотя подобные операции достаточно просты, но они открывают дополнительные возможности для взаимодействия между приложениями в среде ВО.

**Workbench** — это клиент системы АстроГрид, который позволяет работать с астрономическими ресурсами. Workbench включает:

- Task Launcher — позволяет запускать удаленные приложения (СЕА), реализующие запросы к астрономическим каталогам данных, и/или их обработку, например CrossMatch;
- Workflow Builder — позволяет конструировать потоки работ, состоящие из шагов, которые могут выполняться последовательно, параллельно или в цикле;
- MySpace Browser — позволяет манипулировать пользовательскими данными, находящимися в MySpace;
- LookOut — позволяет следить за выполняющимися задачами (или потоками работ), предоставляет возможности отмены, паузы или досрочного завершения задачи;
- AstroScore — позволяет находить все ресурсы, связанные с некоторым астрономическим объектом;
- HelioScore — обеспечивает доступ к данным по Солнцу.

Согласно аванпроекту информационной инфраструктуры РВО (РВОИИ) [1], основными принципами построения инфраструктуры являются следующие:

- архитектура реализуется как иерархия взаимодействующих между собой Веб-сервисов (воплощаемых в грид-архитектуре по мере того, как такие архитектуры будут созревать и будут окончательно стандартизированы мировым сообществом);
- приближение обработки к данным, мотивацией для чего является большой объем накопленных

данных и ресурсно-интенсивный характер прикладных задач ВО. Во многих случаях данные не существуют как удаленные, их еще нет во время формулирования задач: они могут быть извлечены из баз данных по запросам, формулируемым при реализации задач;

- создание программ для большинства астрономических проектов является основной проблемой при разработке и эволюции ВО, поэтому модульность архитектуры, способствующей повторному использованию и композиции программ является основополагающим принципом построения инфраструктуры РВО;
- учитывая существование тысяч астрономических архивов и каталогов в мире, формирование глобальной схемы базы данных для ВО является не масштабируемым подходом. Поэтому основным принципом инфраструктуры ВО является следование идее предметных посредников, обеспечивающих взаимодействие между исследователями и релевантными ресурсами информации на основе спецификации предметной области для соответствующего класса задач.

Анализ показывает, что использование АстроГрида в качестве основы РВОИИ позволяет реализовать перечисленные принципы, а именно: поддержку грид-интероперабельных сервисов, модульность архитектуры, возможность повторного использования и композиции сервисов, создание многослойной архитектуры. Компоненты АстроГрида являются непосредственно применимыми в качестве ядра архитектуры РВОИИ:

- Registry — для поиска ресурсов в реестрах на основе метаданных;
- MySpace — для управления виртуальными областями хранения данных совместного пользования;
- Workbench — для интерфейса пользователя ВО и решения задач на основе потоков работ;
- Community — для администрирования и управления доступом к данным ВО;
- СЕА — в качестве средства поддержки сети интероперабельных сервисов;
- DSA — в качестве средств хранения данных, размещаемых на требуемом уровне реализации системы (задачи).

## 6 Гибридная архитектура АстроГрида и исполнительных средств поддержки слоя предметных посредников

При разработке гибридной архитектуры во главу угла ставилась задача обеспечения интероперабельности компонентов слоя предметных посредников и АстроГрида. В качестве необходимых были выделены следующие возможности гибридной архитектуры:

- регистрация и поиск посредников в реестрах АстроГрида;
- просмотр пользователем схем посредников и другой информации о посредниках;
- задание запросов к посредникам как на языке Syfs, так и на языке ADQL<sup>1</sup> [24];
- использование программ посредников как шагов потоков работ в системе АстроГрид;
- просмотр результатов программ посредников в среде АстроГрида.

В результате анализа архитектур средств поддержки слоя посредников и АстроГрида выработаны следующие основные решения гибридной архитектуры:

- посредник реализуется как СЕА-приложение и регистрируется в реестре АстроГрида при помощи сервера-адаптера приложений. Метаданные, которыми характеризуется регистрируемый посредник, расширяются схемой посредника;
- на интерфейсе посредника имеются методы задания запросов на Syfs и на ADQL;
- результаты запросов к посреднику представляются в формате VOClass, для которого формат VOTable [2] является подмножеством, и сохраняются в MySpace;
- для подключения к посредникам DSA ресурсов, уже зарегистрированных в системе АстроГрид, используется адаптер к DSA ресурсам;
- для осуществления поиска по метаданным ресурсов или приложений в реестрах системы АстроГрид используется адаптер к реестрам АстроГрида, реестр регистрируется в посреднике как обычный ресурс;
- в качестве клиентского интерфейса и приложения используется Workbench;
- для поддержки астрономических онтологий, принятых в системе АстроГрид, разрабатываются средства онтологической интеграции

на основе UCD (Unified Content Description) [2] — унифицированных дескрипторов контента, широко используемых в астрономии.

В данном разделе рассматриваются, в основном, те компоненты гибридной архитектуры, которыми дополняется система АстроГрид, а также исполнительные средства слоя посредников (рис. 2). Компоненты самой системы АстроГрид и компоненты исполнительных средств слоя посредников описаны ранее.

Посредник может регистрироваться в реестрах АстроГрида в двух вариантах: как СЕА-приложение или как DSA-компонент. Стоит отметить, что один и тот же посредник может быть зарегистрирован одновременно как СЕА и как DSA.

Посредник, зарегистрированный как СЕА-приложение, выполняет запрос на языке Syfs над объектной схемой посредника и возвращает результат в формате VOClass. Для реализации этой возможности разработан компонент *M-CEA*, который инкапсулирует конкретный посредник и может быть зарегистрирован в реестре АстроГрида как СЕА приложение.

Посредник, зарегистрированный как DSA, выполняет запрос на языке ADQL над уплощенной схемой посредника и возвращает результат в формате VOClass. Поскольку схема посредника объектная, а ADQL — язык запросов к реляционной схеме, необходим специальный компонент посредника — *преобразователь схем*, который по объектной схеме посредника строит ее уплощение (реляционное представление). Запрос на ADQL формулируется в терминах плоской схемы, а затем компонентом *ADQL2Syfs* транслируется в запрос на языке Syfs к исходной объектной схеме посредника. *Преобразователь схем* и *ADQL2Syfs* используются разработанным компонентом *M-DSA*, который инкапсулирует конкретный посредник и может быть зарегистрирован в реестре АстроГрида как DSA-приложение. Оформление посредника как DSA — важная особенность, так как предоставляет пользователям АстроГрида возможность видеть посредник в привычном виде как DSA-источник.

В любом случае посредник оформляется в качестве СЕА-приложения, которое может быть использовано как шаг потока работ или как самостоятельное приложение. Таким образом, обеспечивается возможность использования посредника в системе АстроГрид, а также возможность удобного просмотра схемы посредника прямо из клиента Workbench системы АстроГрид. Метаданные для регистрации посредника генерируются на основании его схемы, определенной на языке СИНТЕЗ.

<sup>1</sup>ADQL — Astronomical Data Query Language — астрономический язык запросов, поддерживаемый IVOA. Основан на SQL 92.

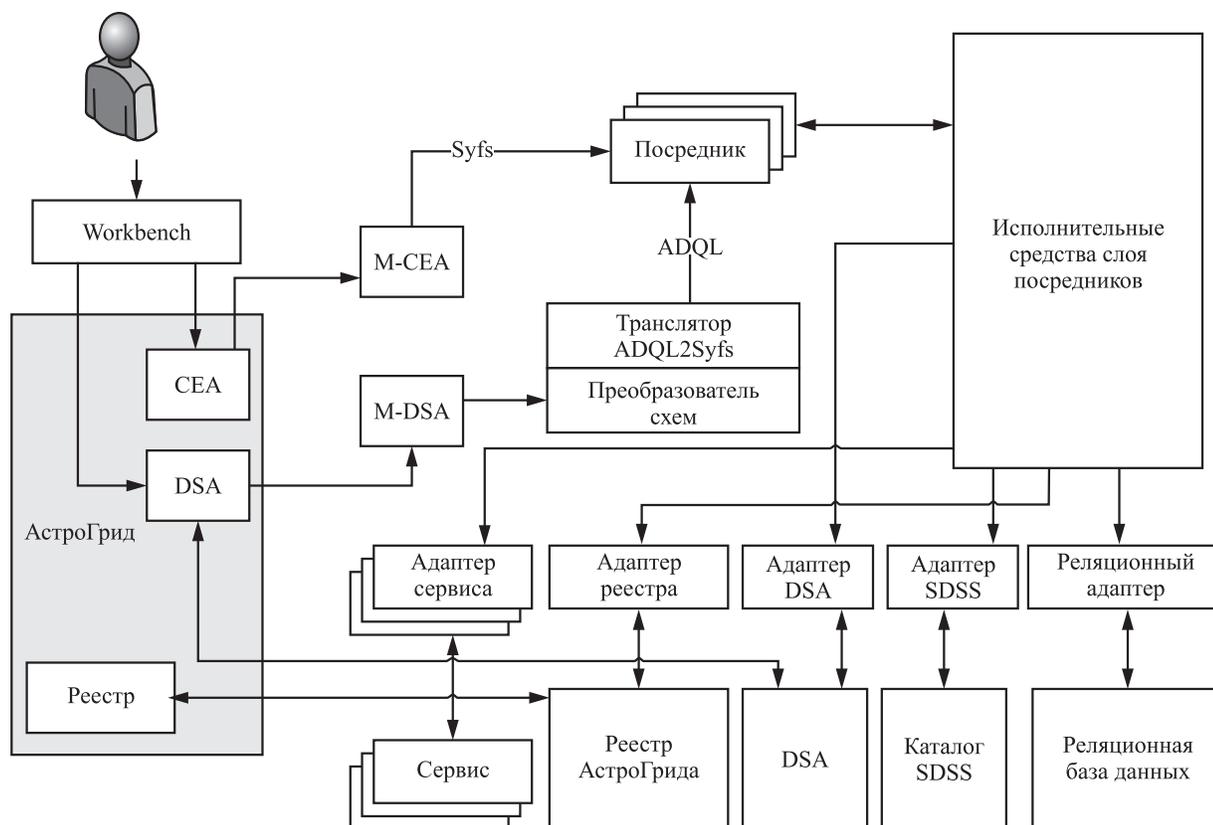


Рис. 2 Схема гибридной архитектуры АстроГрида и исполнительных средств слоя посредников

Адаптеры обеспечивают преобразование запросов на языке планов посредника (Asyfs) в запросы на языке ресурса, получение результата запроса от другого ресурса, а также преобразование объектов результирующего множества в объекты схемы посредника.

В действующем прототипе разработаны адаптер реляционных баз данных (*Relational Wrapper*), адаптер к астрономическому каталогу данных SDSS (*SDSS Wrapper*) с поддержкой возможности выполнения XMatch на сервере SDSS, адаптер DSA ресурсов (*DSA Wrapper*) и адаптер реестров АстроГрида (*Registry Wrapper*).

Для обеспечения возможности использования уже существующих ресурсов, зарегистрированных в реестрах АстроГрида как DSA ресурсы, был разработан адаптер к DSA ресурсам. Этот адаптер играет одну из ключевых ролей по реализации гибридной архитектуры, поскольку дополнительно обеспечивается возможность конструирования посредников над многочисленными ресурсами системы АстроГрид.

Реестр регистрируется в посреднике как обычный ресурс, к которому могут быть сформулированы запросы, что позволит идентифицировать ре-

сурсы, уже зарегистрированные в системе АстроГрид, и использовать их для конструирования новых посредников.

## 7 Пример: реализация средств поддержки задачи поиска далеких галактик в гибридной архитектуре

### 7.1 Задача поиска далеких радиогалактик

В качестве примера класса задач для решения в гибридной архитектуре слоя посредников и АстроГрида рассматривается поиск далеких галактик.

Радиогалактики являются самыми массивными звездными системами, наблюдаемыми на космологических расстояниях [25], но это достаточно редкие объекты по сравнению с обычными (нормальными) галактиками. Далекая галактика, порождающая мощное радиоизлучение, часто оказывается слабым объектом в оптическом диапазоне, что существенно затрудняет ее спектрофотометрические исследования и, следовательно, определе-

ние красного смещения. Усилия, которые были предприняты для поиска радиогалактик на больших красных смещениях, привели к обнаружению  $\sim 300$  радиогалактик с  $Z > 2$  (по данным из базы данных NED (<http://nedwww.ipac.caltech.edu>) для красных смещений, определенных по спектрам), но такого количества объектов недостаточно для статистических исследований свойств далеких галактик. Известны разные техники селекции радиоисточников для поиска далеких галактик. К ним относятся глубокая спектроскопия пустых полей, прямые снимки в узкополосных фильтрах, спектроскопия объектов вокруг радиогалактики с большим красным смещением [26, 27], использование показателей цвета для выбора кандидатов (color-dropouts) [28].

Один из часто применяемых методов селекции кандидатов в далекие радиогалактики основан на корреляции наблюдаемой крутизны радиоспектра с красным смещением [29]. Именно этот подход использовался в программе «Большое Трио» [30], которая в течение ряда лет проводилась в Специальной астрофизической обсерватории (САО) РАН под руководством академика Ю. Н. Парийского. В начале 1990-х гг. были начаты работы по исследованию SS (Steep Spectrum) выборки каталога RC. Наблюдения по этой программе выполнялись на 3 крупных телескопах: ПАТАН-600, БТА и VLA (NRAO, США). Каталог RC, полученный на основе материалов глубокого обзора «Холод» (ПАТАН-600) и радиокаталог UTRAO [31] использовались для отбора объектов, определения спектров; радиотелескоп VLA — для получения точных координат источников, радиоизображений, морфологии; 6-метровый оптический телескоп БТА — для отождествления и глубокой спектроскопии радиоисточников.

При составлении SS выборки каталога RC так же, как и в других группах, занимающихся поиском далеких радиогалактик, в качестве основного критерия отбора использовалась крутизна радиоспектра и еще дополнительные условия [32]:

- плотность потока (спектральная плотность потока излучения или плотность потока — количество электромагнитной энергии в единичном интервале частот, протекающей через единичную площадку за единицу времени. Внесистемная единица измерения, применяемая в радиоастрономии, — Янский ( $1 \text{ Ян} = 10^{-26} \text{ Вт}/(\text{м}^2 \text{ Гц})$ );
- морфология объектов. Были отобраны двойные радиоисточники;
- угловые размеры. На больших красных смещениях не зарегистрировано объектов очень

больших угловых размеров, поэтому учитывалась феноменологическая зависимость размеров двойных радиогалактик от красного смещения.

С появлением в конце XX — начале XXI в. новых больших по площади покрытия глубоких обзоров неба в радиодиапазоне методика селекции по крутизне радиоспектра не теряет своей значимости. Чтобы повысить эффективность изучения выборки радиоисточников с крутыми и ультракрутыми спектрами, часто используют дополнительные ограничения по показателю цвета [28] для отождествленных объектов. Трудоемкость работы с разнородными астрономическими каталогами и обзорами требует нового подхода при решении разных астрономических задач, в частности это относится и к задаче составления поисковых списков для дальнейших более детальных исследований.

Задача поиска далеких галактик была сформулирована в контексте подхода, развиваемого на основе программных средств ВО, а именно: как последовательность типичных действий (выборка данных, сравнение, визуализация) с астрономическими данными, доступными в Интернете по протоколам IVOA. В качестве ресурсов, используемых в задаче, были выбраны: каталог RC [30], полученный на частоте 3,9 ГГц из наблюдательного материала глубокого обзора 20' полосы неба, centered на склонение источника SS433, проведенного на радиотелескопе ПАТАН-600; каталог и обзор неба на частоте 1,4 ГГц FIRST [33]; обзор в оптическом диапазоне SDSS [34]. Отметим, что оптическое отождествление радиоисточников простыми алгоритмами кросс-мэтчинга, такими как ConeSearch или XMatch, возможно, если каталог имеет высокую точность координат и угловое разрешение (порядка  $1''$ ), для точечных источников. Каталог RC имеет координатную точность  $\sim 15''$  по прямому восхождению RA и  $\sim 45''$  по склонению DEC и не имеет данных по морфологии источников, поэтому необходимы дополнительные ограничения для оптических объектов, попадающих в область поиска, а именно, ограничения по показателю цвета. Высокое угловое разрешение и точность координат лучше уловой секунды при предельной чувствительности по плотности потока  $\sim 1 \text{ мЯн}$  обзора FIRST позволяют сделать выбор из кандидатов на оптическое отождествление.

Поток работ для такой задачи может выглядеть следующим образом:

- найти радиоисточники, у которых значение спектрального индекса (спектральный индекс — показатель степени  $\alpha$  в выражении  $S(\nu) \sim \nu^{-\alpha}$ , характеризующем изменение спек-

тральной плотности потока излучения  $S(\nu)$  у ряда космических объектов с изменением частоты  $\nu$  находится в заданном диапазоне  $[\alpha_1, \alpha_2]$ . Если этот параметр неизвестен, то для его определения необходимо

- выбрать радиоисточники, для которых есть данные по плотности потока, по крайней мере, на двух заданных частотах  $\nu_1$  и  $\nu_2$ ;
  - вычислить спектральный индекс  $\alpha$  для выбранных радиоисточников;
  - выдать список радиоисточников, у которых спектральный индекс  $\alpha$  находится в диапазоне  $[\alpha_1, \alpha_2]$ ;
- найти радиоисточники, у которых линейные размеры меньше заданного значения  $d$  и величина потока  $s$  для заданной частоты лежит в диапазоне  $[s_1, s_2]$ ;
  - найти оптические объекты, совпадающие по координатам (с учетом ошибок координат и размеров радиоисточников) с выбранными радиоисточниками;
  - выдать список радиоисточников, для которых найдены соответствующие оптические объекты;
  - по запросу пользователя для выбранной пары радио- и оптического источников показать изображение источника в оптическом диапазоне с нанесенным контуром радиоизображения;
  - выдать запрос пользователю на подтверждение отождествления радио- и оптического изображений;
  - найти значения потоков (звездных величин) для указанных фильтров для оптических источников из сформированного списка;
  - вычислить показатели цвета для оптических источников;
  - выбрать источники, у которых показатели цвета  $c_1$  и  $c_2$  лежат в интервалах  $[a_1, b_1]$  и  $[a_2, b_2]$  соответственно;
  - выдать список источников — кандидатов в далекие галактики.

Выполнение описанной выше последовательности действий для составления поискового списка радиоисточников осложняется тем, что не всегда требуемые данные имеются в выбранных каталогах, а также точностью определения параметров в отдельно взятом ресурсе.

Каталог RC имеет недостаточную для оптического отождествления координатную точность, и

размер области кросс-мэтчинга с оптическим каталогом будет  $45'' \times 45''$  (если мы используем утреннюю ошибку координат по склонению). По нашей оценке при средней плотности объектов SDSS ( $\sim 7-8$  на кв. угл. мин в полосе обзора «Холод») в область поиска для источников каталога RC (размеры ее варьируются от  $45''$  до  $10'$ ) попадает до тысячи объектов. Поскольку точность определения координат радиоисточников не позволяет выполнить оптическое отождествление по позиционному совпадению, то нужна дополнительная селекция по свойствам объектов в оптическом диапазоне. Красное смещение приводит к сдвигу спектрального распределения энергии (Spectral Energy Distribution или SED) далеких галактик в более длинноволновую область спектра по сравнению с близкими галактиками. В оптическом диапазоне скачок в SED для  $\lambda = 912 \text{ \AA}$  для галактик с большим  $Z$  смещается из голубой в красную область спектра, т. е. далекие объекты — более красные и более яркие в этой области спектра по сравнению с близкими галактиками. Аналогично работает избыток излучения в линии водорода (свойство, присущее радиогалактикам)  $Ly$  ( $\lambda = 1216 \text{ \AA}$ ). Используя звездные величины родительских галактик радиоисточников, можно определить показатель цвета или вторые разности показателей цвета и использовать их для выбора наиболее вероятного кандидата из числа объектов, попавших в область поиска. В качестве дополнительного селекционного правила использовались следующие ограничения:

$$\frac{u+r}{2} - g > 1; \quad \frac{g+i}{2} - r > 1; \quad \frac{r+z}{2} - i > 1,$$

является наиболее вероятным кандидатом для отождествления.

Для подтверждения отождествления отобранных источников из обзора FIRST извлекаются изображения, по которым строятся контурные карты, накладываемые на оптические изображения, и отмечаются каталожные положения. По такому рисунку исследователь принимает решение.

Таким образом, поток работ для задачи окончательно определяется выбранными ресурсами и для набора: каталог RC, радиообзор FIRST и оптический обзор SDSS — выглядит следующим образом:

- извлечение из каталога RC источников, попадающих в область пересечения с обзорами FIRST и SDSS;
- извлечение из обзора SDSS объектов, попадающих в эту же область;
- проведение кросс-идентификации двух выборок;

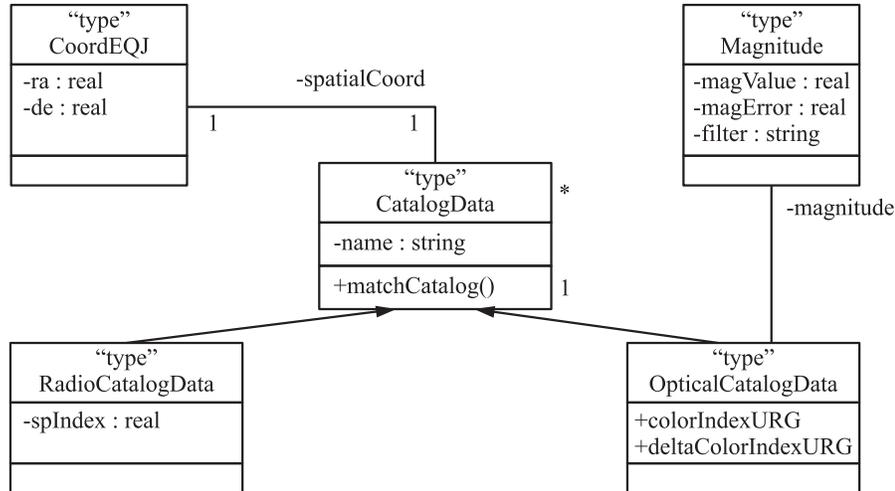


Рис. 3 Спецификация типов фрагмента схемы посредника

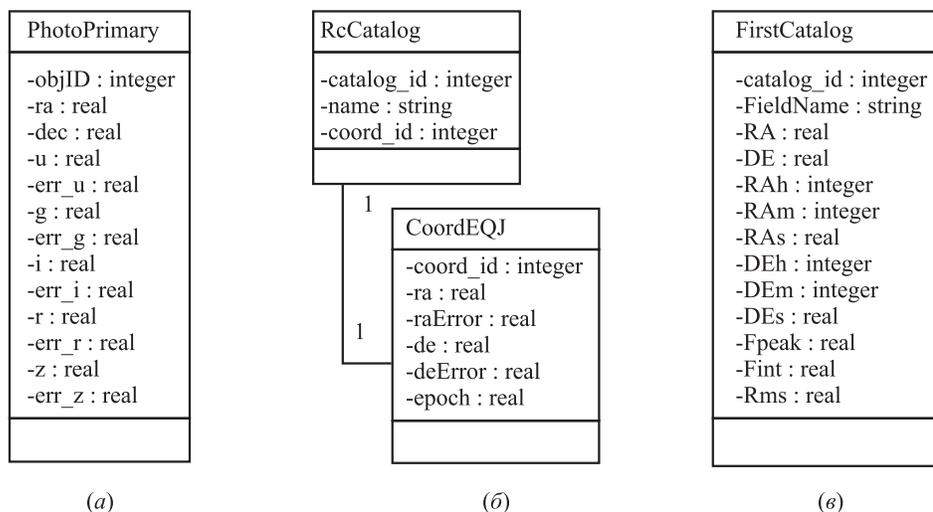
- вычисление для оптических объектов, попавших в область поиска, центрированную на положение источника каталога RC, показателей цвета по звездным величинам в фильтрах  $u$ ,  $g$ ,  $r$ ,  $i$  и  $z$  с учетом ошибок определения звездных величин;
- отбор из оптических объектов, удовлетворяющих одному из условий:  $(u + r)/2 - g > 1$ ,  $(g + i)/2 - r > 1$ ,  $(r + z)/2 - i > 1$ ;
- показ по запросу пользователя для выбранной пары радио- и оптического источников изображения источника в оптическом диапазоне с нанесенным контуром радиоизображения из обзора FIRST;
- занесение координаты оптического объекта в список при подтверждении отождествления.

## 7.2 Схема посредника

На рис. 3 приведен пример фрагмента схемы посредника для задачи поиска далеких галактик. В языке СИНТЕЗ этому фрагменту соответствует следующая спецификация (спецификации элементов схемы посредника объединены в модуль *Mediator*).

```
{Mediator; in: module;
  type:
    {CoordEQJ; in: type;
      ra: real;
      de: real;
    },
    {CatalogData; in: type;
      name: string;
      spatialCoord: CoordEQJ;
      matchCatalog: {in: function;
```

```
  params:
    {+o/CatalogData,
      +rad1/real,
      +rad2/real,
      -returns/boolean};
    };
  },
  {Magnitude; in: type;
    magValue: real;
    magError: real;
    filter: string;
  },
  {RadioCatalogData; in: type;
    supertype: CatalogData;
    spIndex: real;
  },
  {OpticalCatalogData; in: type;
    supertype: CatalogData;
    colorIndexURG: real;
    deltaColorIndexURG: real;
    magnitude: {set;
      type_of_element: Magnitude;};
  };
class_specification:
  {scienceData; in: class;
    instance_section: ScienceData;
  },
  {catalogData; in: class;
    superclass: scienceData;
    instance_section: CatalogData;
  },
  {radioCatalogData; in: class;
    superclass: catalogData;
    instance_section:
      RadioCatalogData;
  },
```



**Рис. 4** Спецификация типов экземпляров классов фрагментов схемы SDSS (а), RC (б) и FIRST (е)

```
{opticalCatalogData; in: class;
  superclass: catalogData;
  instance_section:
    OpticalCatalogData;
};
}
```

В ней определены типы (снабжены стереотипом “type”) и классы, соответствующие астрономическим наблюдениям, накапливаемым в различных радио- и оптических каталогах. Суперкласс *CatalogData* содержит объекты из обоих каталогов. Этот класс содержит атрибут *spatialCoord*, который определяет координаты астрономического объекта, и метод (предикат) *matchCatalog()*, который возвращает истинное значение, считая два астрономических объекта тождественными, если они «близки» друг к другу в пространстве.

В текущем примере в посреднике были зарегистрированы следующие каталоги и обзоры неба:

- RC catalog — каталог радиоисточников, полученный из наблюдательного материала обзора неба «Холод» на частоте 3,9 ГГц на радиотелескопе ПАТАН 600 [30];
- FIRST — обзор неба на частоте 1,4 ГГц, полученный на радиотелескопе VLA; включает коллекцию цифровых изображений и каталог радиоисточников [33];
- SDSS — глубокий фотометрический (в фильтрах *u*, *g*, *r*, *i* и *z*) и спектральный обзор неба в оптическом диапазоне [34].

В процессе регистрации RC и SDSS рассматриваемый фрагмент схемы ресурса RC включает классы *rcCatalog* и *coordEQJ*; рассматриваемый

фрагмент схемы ресурса SDSS включает класс *photoPrimary* и функцию модуля *xmatch*. Функция *xmatch* представляет хранимую процедуру, доступную для выполнения в SDSS, которая выполняет координатную кросс-идентификацию объектов каталога SDSS и объектов внешней коллекции. Фрагменты схем ресурсов RC, SDSS и FIRST приведены на рис. 4.

В качестве примера дан следующий модуль спецификации фрагмента схемы ресурса SDSS:

```
{SDSS; in: module, local;
  type:
    {PhotoPrimary; in: type;
      objID: integer;
      ra: real;
      dec: real;
      u: real;
      err_u: real;
      g: real;
      err_g: real;
      i: real;
      err_i: real;
      r: real;
      err_r: real;
      z: real;
      err_z: real;
    };
  function:
    {xmatch; in: function;
      params: {
        +ra1/real, +del/real,
        +ra2/real, +de2/real,
        +rad/real, -returns/boolean
      };
    };
};
```

```
class_specification:
  {photoPrimary; in: class;
   instance_section: PhotoPrimary;
  };
}
```

В результате сопоставления элементов схемы посредника и схем ресурсов RC и SDSS в процессе регистрации были построены следующие взгляды. Взгляд для FIRST строится аналогично RC.

Взгляды для SDSS:

– LAV<sub>1</sub>:

```
v_SDSS_OData(x/[name,ra,de,
  colorIndexURG,deltaColorIndexURG])
:- opticalCatalogData(x/[name,
  ra: spatialCoord.ra,
  de: spatialCoord.de,
  colorIndexURG,deltaColorIndexURG])
```

– GAV<sub>1</sub>:

```
v_SDSS_OData(x/[name,ra,de,
  colorIndexURG,deltaColorIndexURG])
:- SDSS.photoPrimary(x/[ra,de:dec,
  objID,u,r,g,err_u,err_r,err_g])
& integerToString(objID, name)
& id(((u + r)/2.0-g),colorIndexURG)
& I_SDSS.deltaColorIndexURG(err_u,
  err_r,err_g, deltaColorIndexURG)
```

Во взгляде GAV<sub>1</sub> используется функция разрешения конфликтов, которая вычисляет значение атрибута *deltaColorIndexURG* на основе значений атрибутов *err\_u*, *err\_r*, *err\_g* из класса *photoPrimary*. Для функций разрешения конфликтов в посреднике определена их реализация посредством PL/SQL.

Взгляды для RC:

– LAV<sub>2</sub>:

```
v_RC_RData(_/[name, ra, de])
:- radioCatalogData(x/[name,
  ra: spatialCoord.ra,
  de: spatialCoord.de])
```

– GAV<sub>2</sub>:

```
v_RC_RData(x/[name, ra, de])
:- RC.rcCatalog(x/[name, catalog_id,
  coord_id])
& RC.coordEQJ(y/[ra, de, coord_id])
```

На основе взглядов LAV строятся следующие инверсные правила отображения класса в класс:

– IR<sub>1</sub>:

```
radioCatalogData(x/[name,
  ra: spatialCoord.ra,
  de: spatialCoord.de]) <-
v_RC_RData(_/[x:#x, name, ra, de])
```

– IR<sub>2</sub>:

```
opticalCatalogData(x/[name,
  ra: spatialCoord.ra,
  de: spatialCoord.de, colorIndexURG,
  deltaColorIndexURG]) <-
v_SDSS_OData(_/[x:#x, name, ra, de,
  colorIndexURG, deltaColorIndexURG])
```

Дополнительно определяется следующее инверсное правило отображения функции в функцию:

– IR<sub>3</sub>:

```
all x/CatalogData, y/CatalogData,
  rad/real
matchCatalog(x, y, rad, rad, b) <-
xmatch(x.spatialCoord.ra,
  x.spatialCoord.de,
  y.spatialCoord.ra,
  y.spatialCoord.de, rad, b)
```

Инверсное правило IR<sub>3</sub> связывает метод *matchCatalog* в схеме посредника с функцией кросс-мэтчинга *xmatch* в схеме ресурса SDSS.

### 7.3 Реализация примера поддержки решения задачи поиска далеких галактик в гибридной архитектуре

Выше был описан пример фрагмента посредника для поддержки решения задачи поиска далеких галактик. Этот посредник был оформлен как СЕА-приложение, к которому реализуются запросы на языке Syfs. В системе АстроГрид был написан поток работ с участием данного посредника (рис. 5) для решения задачи поиска далеких галактик. Этот поток работ состоит из 3-х шагов: нахождение галактик-кандидатов, преобразование результата, получение изображений.

#### 7.3.1 Первый шаг — поиск кандидатов в далекие галактики

На первом шаге происходит нахождение объектов — кандидатов в далекие галактики. Этот шаг реализуется посредником как СЕА-приложение. Для нахождения кандидатов в далекие галактики задается запрос к посреднику, который на основании заданных условий и кросс-мэтчинга объектов из двух радиокаталогов (RC и FIRST) с объектами оптического каталога (SDSS) отбирает объекты — потенциальных кандидатов в далекие галактики. На рис. 5 изображен первый шаг потока работ, на котором видно, что используется СЕА-приложение *ipi.ac.ru/execute/SyfsQuery*, реализующее описанный выше посредник. В качестве входного параметра данному приложению передается содержимое фай-

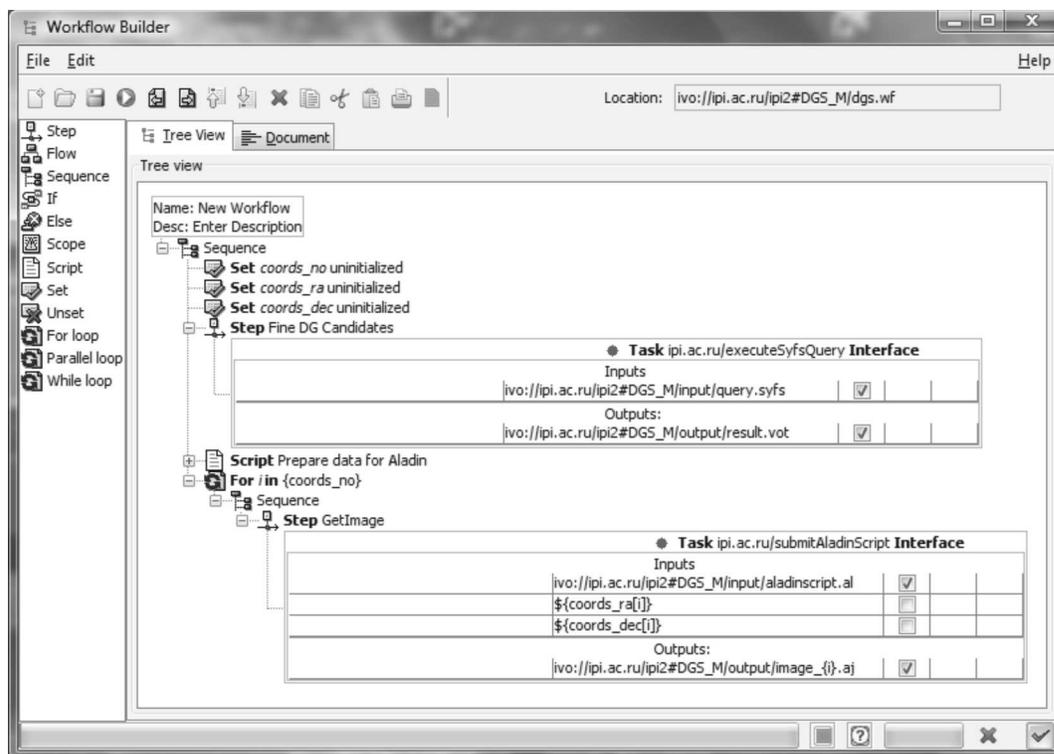


Рис. 5 Поток работ для задачи поиска далеких галактик

ла *query.syfs*, находящегося в виртуальном хранилище пользователя MySpace. В файле содержится следующий запрос на языке Syfs:

```
{ {r(x/[ra, de, name, name1, ra1, de1])
:-radioCatalogData(y/[name,
  ra: spatialCoord.ra,
  de: spatialCoord.de])
& opticalCatalogData(x/[name1: name,
  ra1: spatialCoord.ra,
  de1: spatialCoord.de,
  colorIndexURG,deltaColorIndexURG])
& matchCatalog(y, x, 45, 45, b)
& b = true
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
& ra1 >= 120.0 & ra1 <= 255.0
& de1 >= 4.39 & de1 <= 5.61
& colorIndexURG > deltaColorIndexURG}}
```

Запрос состоит из одного правила и является частью решения задачи поиска далеких галактик. По запросу должны возвращаться астрономические объекты, найденные в радио и оптических каталогах с применением функции кросс-мэтчинга, которые имеют координаты в заданном диапазоне и удовлетворяют специальному условию, ограничивающему значения атрибутов *colorIndexURG* и *deltaColorIndexURG*.

Таблица 1 описывает переписывание, которое применяется к телу правила. Подцели правила перечислены в первом столбце. Инверсные правила, подобранные для унификации с соответствующей подцелью, перечислены во втором столбце. Преобразованные в результате унификации тела инверсных правил перечислены в третьем столбце.

После того как унификация и преобразование тел выбранных инверсных правил выполнены, строится следующее переписывание правила:

```
r(_/[ra, de, name, name1, ra1, de1])
:- v_RC_RData(_/[y:#x, name, ra, de])
& v_SDSS_OData(_/[x:#x, name1: name,
  ra1: ra, de1: de, colorIndexURG,
  deltaColorIndexURG])
& xmatch(y.spatialCoord.ra,
  y.spatialCoord.de,x.spatialCoord.ra,
  x.spatialCoord.de, 45, b)
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
& ra1 >= 120.0 & ra1 <= 255.0
& de1 >= 4.39 & de1 <= 5.61
& colorIndexURG > deltaColorIndexURG
```

Полученное правило содержит сколемовские атрибуты *#x* во взгляде *v\_SDSS\_OData* и *#x* во взгляде *v\_RC\_RData*. Но в процессе семантического анализа устанавливается, что взгляд *v\_SDSS\_OData*

**Таблица 1** Переписывание подцелей правила

Подцель	ИП	Преобразованное тело ИП
radioCatalogData(y/[name, ra: spatialCoord.ra, de: spatialCoord.de])	IR <sub>1</sub>	v_RC_RData(_/[y:#x, name, ra, de])
opticalCatalogData(x/[name1: name, ra1: spatialCoord.ra, de1: spatialCoord.de, colorIndexURG, deltaColorIndexURG])	IR <sub>2</sub>	v_SDSS_OData(_/[x:#x, name1: name, ra1:ra, de1:de, colorIndexURG, deltaColorIndexURG])
matchCatalog(y, x, 45, 45, b)	IR <sub>3</sub>	xmatch(y.spatialCoord.ra, y.spatialCoord.de, x.spatialCoord.ra, x.spatialCoord.de, 45, b)

включает равенство переменной, привязанной к терму  $\#x.spatialCoord.ra$ , и переменной, привязанной к терму  $ra$ . Следовательно, с учетом того, что в правиле в проекции в подцели  $v\_SDSS\_OData$  приравниваются атрибуты  $x$  и  $\#x$ , а также атрибуты  $ra1$  и  $ra$ , из правила можно удалить равенства  $x.spatialCoord.ra = ra1$ . Это фактически означает, что в правиле вхождения термина  $x.spatialCoord.ra$  заменяются термом  $ra1$ . Аналогично семантический анализ выполняется для  $x.spatialCoord.de$  и  $de1$ , для  $y.spatialCoord.ra$  и  $ra$ , для  $y.spatialCoord.de$  и  $de$ . В результате образуется правило без вхождений сколемовских функций:

```
r(_/[ra, de, name, name1, ra1, de1])
:- v_RC_RData(_/[name, ra, de])
& v_SDSS_OData(_/[name1: name,
  ra1: ra, de1: de, colorIndexURG,
  deltaColorIndexURG])
& xmatch(ra, de, ra1, de1, 45, b)
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
& ra1 >= 120.0 & ra1 <= 255.0
& de1 >= 4.39 & de1 <= 5.61
& colorIndexURG > deltaColorIndexURG
```

Дальше выполняется разворачивание взглядов GAV (взгляда  $v\_RC\_RData$  и взгляда  $v\_SDSS\_OData$ ).

Голова взгляда GAV<sub>1</sub>

```
v_SDSS_OData(x/[name,ra,de,
  colorIndexURG,deltaColorIndexURG])
:- SDSS.photoPrimary(x/[ra,de:dec,
  objID,u,r,g,err_u,err_r,err_g])
& integerToString(objID, name)
& id(((u + r)/2.0 - g),colorIndexURG)
& I_SDSS.deltaColorIndexURG(err_u,
  err_r,err_g, deltaColorIndexURG)
```

унифицируется с подцелью

```
v_SDSS_OData(_/[name1: name, ra1: ra,
  de1: de, colorIndexURG,
  deltaColorIndexURG]),
```

которая заменяется на преобразованное тело взгляда

```
SDSS.photoPrimary(x1/[ra1:ra,de1:dec,
  objID,u,r,g,err_u,err_r,err_g])
& integerToString(objID, name1)
& id(((u+r)/2.0-g), colorIndexURG)
& I_SDSS.deltaColorIndexURG(err_u,
  err_r,err_g, deltaColorIndexURG)
```

Аналогично для взгляда  $v\_RC\_RData$  соответствующая подцель заменяется преобразованным телом взгляда GAV<sub>2</sub>:

```
RC.rcCatalog(x2/[name, catalog_id,
  coord_id])
& RC.coorEQJ(y2/[ra, de, coord_id])
```

Результатом разворачивания взглядов GAV является следующее правило, которое переписывает исходное правило, выражая его в терминах схем ресурсов SDSS и RC:

```
r(_/[ra, de, name, name1, ra1, de1])
:-RC.rcCatalog(x2/[name, catalog_id,
  coord_id])
& RC.coorEQJ(y2/[ra, de, coord_id])
& SDSS.photoPrimary(x1/[ra1:ra,
  de1:dec,objID,u,r,g,
  err_u,err_r,err_g])
& integerToString(objID, name1)
& id(((u + r) / 2.0 - g),
  colorIndexURG)
& I_SDSS.deltaColorIndexURG(err_u,
  err_r,err_g, deltaColorIndexURG)
& xmatch(ra, de, ra1, de1, 45, b)
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
& ra1 >= 120.0 & ra1 <= 255.0
& de1 >= 4.39 & de1 <= 5.61
& colorIndexURG > deltaColorIndexURG
```

Аналогично переписывание с использованием взглядов для ресурса FIRST приводит ко второму переписыванию исходного правила запроса:

```
r(_/[ra, de, name, name1, ral, del])
:-FIRST.firstCatalog (x2/[name:
  FieldName, ra: RA, de: DE])
& SDSS.photoPrimary(x1/[ral:ra,
  del:dec,objID,u,r,g,
  err_u,err_r,err_g])
& integerToString(objID, name1)
& id(((u + r) / 2.0 - g),
  colorIndexURG)
& I_SDSS.deltaColorIndexURG(err_u,
  err_r,err_g, deltaColorIndexURG)
& xmatch(ra, de, ral, del, 45, b)
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
& ral >= 120.0 & ral <= 255.0
& del >= 4.39 & del <= 5.61
& colorIndexURG > deltaColorIndexURG
```

Полученный переписанный запрос состоит из двух правил, и результирующий класс *r* образуется как объединение результирующих множеств этих двух правил. Правила запросов обращаются к классам трех ресурсов (RC, FIRST и SDSS), поэтому дальше необходим этап планирования, в результате которого получается распределенный план запроса.

Алгоритм планирования выделяет подзапросы, которые адресованы каждому ресурсу, и остаточный запрос, который будет выполняться в СУБД посредника над результатами, полученными от ресурсов. Также алгоритм выбирает, как данные с результатами подзапросов передаются от одного ресурса к другому ресурсу или посреднику. В этом примере учитывается способность ресурса SDSS принимать внешнюю коллекцию, которая затем участвует в выполнении запроса к SDSS, и то, что число объектов, удовлетворяющих условиям запроса в RC во много раз меньше, чем число объектов в SDSS (число объектов в SDSS исчисляется сотнями миллионов, что делает невозможным выполнение прямого подзапроса к SDSS с последующей передачей результатов в посредник). Поэтому строится следующий распределенный план (для простоты составляющие плана даны в логическом языке Syfs, хотя в реальной реализации план запроса представляется на алгебраическом языке Asyfs).

Подзапрос к ресурсу RC:

```
rc0(_/[ra, de, name])
:-RC.rcCatalog(x2/[name, catalog_id,
  coord_id])
& RC.coordeQJ(y2/[ra, de, coord_id])
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
```

Результаты выполнения этого подзапроса передаются адаптером ресурса RC в адаптер ресурса SDSS, который загружает их во временную коллек-

цию на ресурсе SDSS. Далее выполняется подзапрос к ресурсу SDSS:

```
sdss1(_/[ra, de, name, ral, del,
  objID, colorIndexURG])
:-rc1(_/[ra, de, name])
& SDSS.photoPrimary(x1/[ral:ra,
  del:dec,objID,u,r,g,
  err_u,err_r,err_g])
& id(((u + r) / 2.0 - g),
  colorIndexURG)
& xmatch(ra, de, ral, del, 45, b)
& ral >= 120.0 & ral <= 255.0
& del >= 4.39 & del <= 5.61
```

Результат выполнения этого подзапроса передается адаптером SDSS в посредник. Подзапрос к ресурсу FIRST:

```
first0(_/[ra, de, name])
:-FIRST.firstCatalog (x2/[name:
  FieldName, ra: RA, de: DE])
& ra >= 120.0 & ra <= 255.0
& de >= 4.39 & de <= 5.61
```

Результаты выполнения этого подзапроса передаются адаптером ресурса FIRST в адаптер ресурса SDSS, который загружает их во временную коллекцию на ресурсе SDSS. Далее выполняется подзапрос к ресурсу SDSS:

```
sdss2(_/[ra, de, name, ral, del,
  objID, colorIndexURG])
:-first0(_/[ra, de, name])
& SDSS.photoPrimary(x1/[ral:ra,
  del:dec,objID,u,r,g,
  err_u,err_r,err_g])
& id(((u + r) / 2.0 - g),
  colorIndexURG)
& xmatch(ra, de, ral, del, 45, b)
& ral >= 120.0 & ral <= 255.0
& del >= 4.39 & del <= 5.61
```

Результат выполнения второго подзапроса передается адаптером SDSS в посредник. В посреднике над полученными данными выполняется остаточный запрос, вычисляющий конечный результат рассматриваемого примера:

```
r(_/[ra, de, name, name1, ral, del])
:- sdss1(_/[ra, de, name, name1,
  ral, del, objID, colorIndexURG])
& integerToString(objID, name1)
& I_SDSS.deltaColorIndexURG(err_u,
  err_r,err_g, deltaColorIndexURG)
& colorIndexURG > deltaColorIndexURG
```

```
r(_/[ra, de, name, name1, ral, del])
:- sdss2(_/[ra, de, name, name1,
```

```

    ral, del, objID, colorIndexURG])
& integerToString(objID, name1)
& I_SDSS.deltaColorIndexURG(err_u,
    err_r, err_g, deltaColorIndexURG)
& colorIndexURG > deltaColorIndexURG

```

В качестве выходного параметра указывается адрес файла *result.vot* в виртуальном хранилище пользователя MySpace, в который по завершению работы посредника помещается результат в формате VOTable. Результат работы первого шага (посредника) представляется таблицей, каждая строчка которой — кандидат в далекие галактики. Фрагмент таблицы в формате VOTable выглядит следующим образом:

```

<?xml version=-1.0- encoding=-UTF-8-?>
<votable xmlns="http://
www.ivoa.net/xml/VOTable/v1.1">
<resource>
<table name="r">
<field name="de" datatype="double" />
<field name="del" datatype="double" />
<field name="name" datatype="char"
    arraysize="*" />
<field name="name1" datatype="char"
    arraysize="*" />
<field name="ra" datatype="double" />
<field name="ral" datatype="double" />
<data><tabledata>
<tr>
    <td>5.0155554</td>
    <td>5.02104130192953</td>
    <td>RC 0819+0517</td>
    <td>587734948577674422</td>
    <td>120.183464</td>
    <td>120.179933980257</td>
</tr>
<tr>
    <td>5.0155554</td>
    <td>5.02523423172066</td>
    <td>RC 0819+0517</td>
    <td>587734948577673600</td>
    <td>120.183464</td>
    <td>120.179594574991</td>
</tr>
...
</tabledata></data>
</table>
</resource>
</votable>

```

Формат VOTable — это XML документ. Тэгами *field* обозначаются столбцы таблицы, определяя ее структуру. Данные для одной строки заключаются в тэг *tr*, а внутри строки каждое значение заключается в тэг *td*.

Для оценки производительности работы посредника в гибридной архитектуре были проведены 10 тестовых запусков первого шага. Среднее время выполнения первого шага в АстроГриде составило 26 с, минимальное время — 19 с, максимальное время — 42 с. Подобная разница времени выполнения связана с географией расположения ресурсов, зарегистрированных в посреднике. Радиокаталог RC представляет собой DSA источник системы АстроГрид, физически располагающийся на установке АстроГрида в ИПИ РАН. Радиокаталог FIRST представляет собой базу данных, физически располагающуюся на сервере в ИПИ РАН. Каталог SDSS располагается в США. В соответствии с планом, данные из радиоисточников перемещаются в SDSS, где выполняется кросс-мэтчинг, и далее результат возвращается в Москву, в MySpace АстроГрида в ИПИ РАН.

### 7.3.2 Второй шаг

На втором шаге происходит обработка результата, чтобы на третьем шаге данные можно было передать приложению, получающему изображения в цикле. Данный шаг состоит не из приложения SEA, а из скрипта на языке Groovy [35]. В этом скрипте происходит загрузка в память таблицы, полученной на первом шаге. Затем по этой таблице строятся три массива: массив координат га (*coords\_ra*), массив координат dec (*coords\_dec*) и массив индексов (*coords\_no*). При построении массивов координат происходит их преобразование из вещественного представления в формат H:M:S (час:мин:с).

### 7.3.3 Третий шаг

На третьем шаге используется астрономическое приложение Aladin. Aladin — это атлас неба, обеспечивающий визуализацию цифровых астрономических изображений; суперпозицию записей из астрономических каталогов и баз данных; доступ в интерактивном режиме к астрономическим базам данных Simbad, Vizier и др. Приложение обладает как графическим пользовательским интерфейсом, так и командной строкой. В командной строке приложения команды задаются на скриптовом языке программы Aladin.

На третьем шаге в цикле вызывается разработанное SEA-приложение, суть которого сводится к выполнению в Aladin'е программы, написанной на его скриптовом языке. Результатом работы приложения является стек Aladin'а, содержащий результаты запросов, выполняемых в программе. В стеке сохраняются данные из каталогов, изображения и контуры изображений и т. д. Данные стека визуализируются программой Aladin.

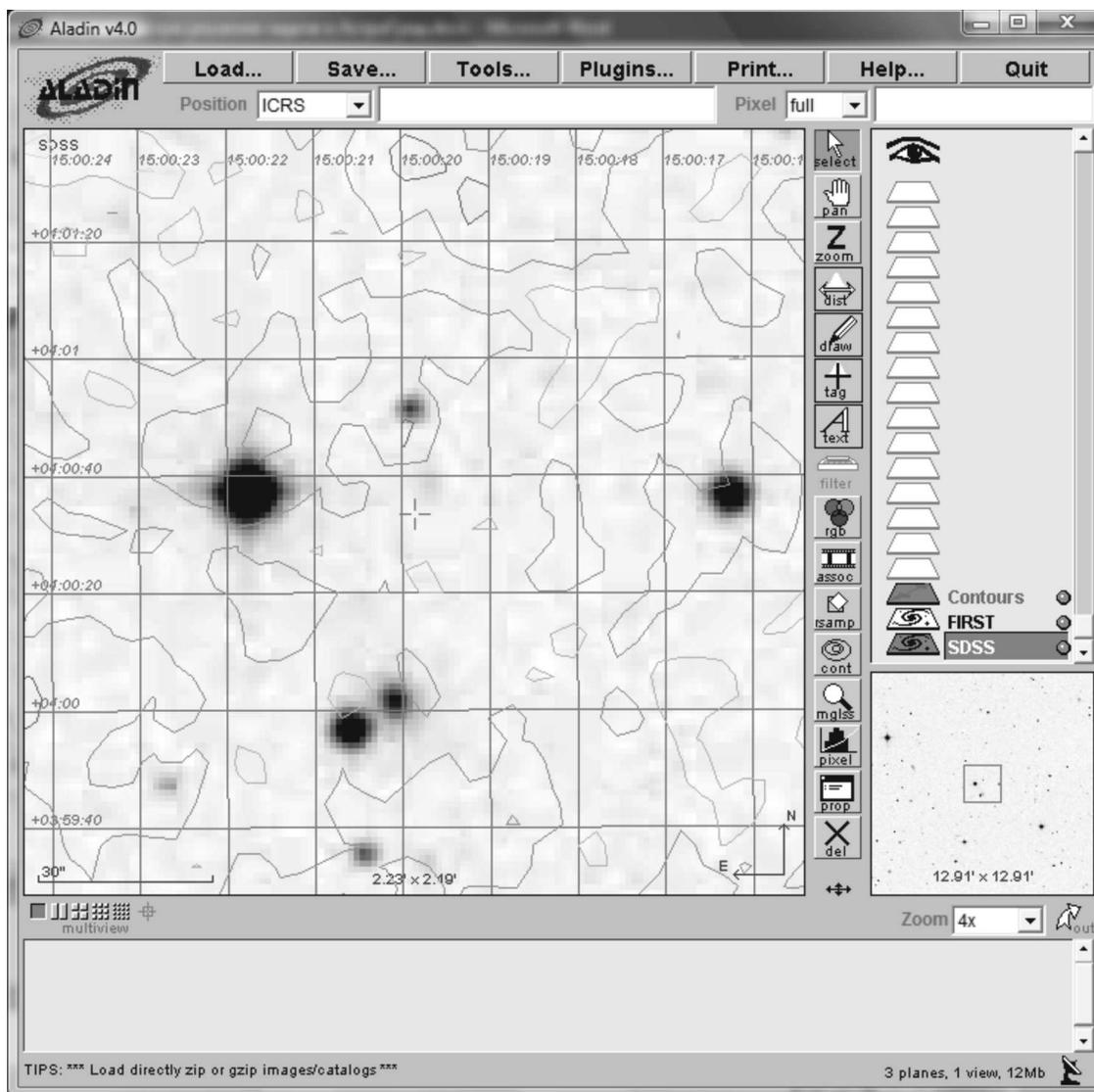


Рис. 6 Пример результата

На рис. 5 изображен вызов этого СЕА-приложения ([ipi.ac.ru/submitAladin.Script](http://ipi.ac.ru/submitAladin.Script)). В качестве второго и третьего входного параметра передаются координаты га и dec:  $i$ -е значение массивов  $coords\_ra[i]$  и  $coords\_dec[i]$  соответственно. В качестве первого входного параметра передается ссылка на файл *aladinscript.al*, находящийся в MySpace. Скрипт является параметризованным двумя переменными:  $\#Ra$  и  $\#Dec$ . В качестве значения этих переменных подставляются те значения, которые передаются СЕА-приложению вторым и третьим параметрами, соответственно. Таким образом, можно выполнять один и тот же скрипт для разных координат.

Скрипт загружает изображения из оптического обзора SDSS и радиообзора FIRST. На оптическое изображение накладываются контуры, построен-

ные по радиоизображению. Результирующее изображение радиооптика (рис. 6) предназначено для анализа специалистом.

## 8 Состояние проблемы интеграции неоднородных информационных ресурсов в посредниках

**Цели введения предметных посредников.** Отличительной особенностью введения посредников в настоящей работе по сравнению с подходами, рассматриваемыми в литературе [15, 36, 37], является их ориентация на решение задач над неоднородны-

ми распределенными информационными ресурсами [38–41] (интеграция информационных ресурсов является лишь частью этой проблемы). Одним из следствий такой цели является то, что схема посредника, наряду со спецификациями структур данных, включает спецификации функций (методов) и процессов. Уточняющая реализация функций посредника устанавливается в существующих ресурсах (сервисах) при их регистрации.

**Отображение схем и определение взглядов.** Отображение схем является одним из критических моментов интеграции информационных ресурсов. Проект Clío [42] интеграции данных и обмена данными поддерживает генерацию отображения схем. Для пары схем средства установления соответствия генерируют множество потенциальных соответствий схем (чисто синтаксических, основанных на отображении термов, имен полей и пр.). Отображение рассматривается как множество запросов из исходной к целевой схеме, осуществляющих преобразование данных ресурса в формат целевой схемы. Эти задачи являются интерактивными (полуавтоматическими). Clío поддерживает отображение реляционных и иерархических XML схем [42].

В другом методе, AutoMed [43], предпринята попытка поддержки трансформации схем для интеграции данных, применяя подходы LAV, GAV и GLAV. При этом используются низкоуровневая гиперграфовая модель данных и набор примитивных преобразований схем в этой модели. Модели данных более высокого уровня и преобразования в них выражаются в терминах общей низкоуровневой модели. Допускаются композиции примитивных трансформаций. Этот синтаксический метод является полуавтоматическим, его можно приспособить для разных структурированных моделей данных.

Настоящая статья представляет совершенно иной, *семантический* подход к трансформации информационных моделей. Во-первых, для информационной интеграции вводится расширяемая каноническая модель (в частности, ядро такой модели составляет гибридная объектно-ориентированная/фреймовая информационная модель [8]). Для любой модели ресурса (определяемой синтаксисом и семантикой двух языков — языка определения информации и языка манипулирования информацией) определяется отображение в расширение канонической модели. Такое отображение обычно конструируется реверсивно [44, 45] и поддерживается соответствующим транслятором (который играет также полезную роль в процессе генерации адаптеров). Подход, основанный на переписывании термов, обеспечиваемый средствами Meta Environment [46], применяется в настоящем

проекте для генерации компиляторов. Соответствие элементов схемы ресурса схеме посредника устанавливается онтологически [9, 10]. Релевантная посреднику схема ресурса, выраженная в некоторой исходной модели, компилируется в каноническую схему. Схема ресурса уточняет схему в канонической модели [45]. Уточнение, задаваемое отображением схем, может быть формально проверено [47]. Отображение схем порождается Унификатором моделей данных, подробно описанным в [7]. После этого, имея схемы, выраженные в одной и той же нотации (каноническая модель), применяется регистрация схемы ресурса в схеме посредника согласно подходу LAV/GLAV, полагаясь на онтологические связи между элементами схем (типами, атрибутами, функциями, параметрами, классами). В результате, применяя исчисление типов [11], формируются правила отображения для подхода LAV/GLAV, согласно описанному в разд. 4.

**Переписывание запросов.** Обзор исследований в области ответа на запрос посредством взглядов, которые простираются от теоретических оснований до алгоритмов соответствующих систем и их реализации, дан в [15]. Кроме того, оценки алгоритмов переписывания запросов даны в работах [19, 36]. Алгоритмы инверсных правил выделяются благодаря их концептуальной простоте, модульности и способности продуцировать максимально включенные переписывания за время, полиномиальное относительно длины запроса и взглядов. Переписывание объединений конъюнктивных запросов при использовании взглядов [19] на основе инверсных правил имеет преимущества по сравнению с известными алгоритмами, в частности этот алгоритм обобщает алгоритмы MiniCon [48] и Ujoin [49] и является более эффективным, чем алгоритм Bucket. Нахождение включаемого переписывания для запросов с объединением конъюнктивных запросов (когда как запрос, так и ограничения взгляда могут содержать встроенные предикаты) является важным свойством этого алгоритма [19]. В то же время исследования семантики объектных запросов в типизированной среде не известны. Одним из вкладов настоящей работы является расширение подхода к переписыванию запросов на основе взглядов в типизированной среде предметных посредников. В противоположность [19], в [20] было показано, как расширить конъюнктивные запросы объектной SPJ семантикой, основанной на отношении уточнения типов и типовом исчислении. Применяемый при этом подход LAV [20] показывает, что уточнение типов экземпляров классов ресурса является основным отношением, необходимым для достижения включения запросов в объектной сре-

де. В настоящей статье, при сохранении типовой среды, показано, как подход LAV/GLAV применяется для реализации функций разрешения конфликтов, возникающих при отображении схем ресурс/посредник. Головы правил определения LAV взглядов могут содержать произвольный запрос в терминах схемы ресурса, разрешающий конфликт при помощи конкретизирующего взгляда, выражаемого как GAV правило над соответствующим ресурсом. Для завершения переписывания взгляд над ресурсом должен быть разрешен в соответствии с подходом GAV (см. разд. 4).

**Прототипы интеграции баз данных для ВО.** Система OpenSkyQuery [50] в NVO [51] обеспечивает интерфейс для выполнения распределенных запросов над федерацией зарегистрированных астрономических архивов. Доступ к каждой базе данных обеспечивается посредством Веб-сервиса SkyNode. Это интерфейс SOAP, воспринимающий запрос на ADQL и возвращающий данные. Узлы в системе могут быть вида Basic или Full. Только узлы Full могут участвовать в операции XMatch. Возможности узла SkyNode Full включают: интерфейс QueryCost(), который выдает простой ADQL запрос для оценки плотности объектов в квадратном градусе в зависимости от условий запроса; возможность использования сложных структур в запросах на ADQL, выполняющих кросс-мэтчинг между содержащимся в базе обзором неба и таблицей в формате VOTable; реализацию метода ExecutePlan, который принимает ExecutePlan и передает релевантную часть плана следующему узлу. Все узлы должны быть зарегистрированы в Реестре ВО. Данные возвращаются обратно от каждого узла и над ними выполняется кросс-мэтчинг. Наконец, результат возвращается клиенту. Система OpenSkyNode использует GAV для интеграции схем баз данных.

Другой прототип распределенной обработки запросов был разработан в Манчестерском университете для гридовой архитектуры OGSA (Open Grid Services Architecture). OGSA-DQP<sup>1</sup> [52] является распределенным процессором запросов, выполненным по отношению к пользователю как OGSA-совместимый Грид сервис. Этот сервис поддерживает компиляцию и вычисление запросов над множеством сервисов Грида. DQP применяет посредник на основе GAV подхода к интеграции данных (объединение схем баз данных составляет глобальную схему). OGSA-DQP поддерживает только такие Веб-сервисы, которые воспринимают примитивные типы в качестве параметров и возвращают либо примитивные типы, либо массивы примитивного типа.

<sup>1</sup>DQP — Distributed Query Processing.

Предлагаемая в настоящей статье гибридная архитектура предметных посредников и АстроГрида отличается от известных прототипов применением расширяемой канонической объектной модели данных, применением движимого приложениями метода формирования предметных посредников на основании LAV/GLAV подхода, использованием развитых методов переписывания запросов к посредникам в типизированной (объектной) среде.

## 9 Заключение

В настоящей статье рассматриваются первые результаты создания промежуточного слоя предметных посредников в гибридной грид-инфраструктуре виртуальной обсерватории для решения научных задач над множеством интегрируемых посредниками неоднородных распределенных информационных ресурсов (таких как базы данных, программные сервисы, онтологические спецификации). Введение такого промежуточного слоя призвано решить ряд семантических проблем взаимодействия ученого, решающего задачу в некоторой предметной области, и разнообразных релевантных задачи результатов наблюдений и средств их обработки. Гибридная архитектура ВО реализована как объединение инфраструктуры системы поддержки ВО АстроГрид, разработанной в Великобритании, и средств поддержки исполнительного слоя предметных посредников, созданных в ИПИ РАН. В исследованной архитектуре предметных посредников реализован подход, движимый приложениями, при котором для класса приложений формируется спецификация предметной области независимо от существующих информационных ресурсов. Далее происходит идентификация ресурсов, релевантных задаче, и их регистрация в посреднике на основе техники GLAV. В различных разделах статьи показано, как, благодаря посредникам, могут быть решены проблемы семантики, возникающие при доказательно правильном отображении информационных моделей ресурсов в каноническую информационную модель посредников, при отображении и интеграции контекстов предметных областей информационных ресурсов в контекст предметной области задачи, при идентификации релевантных задаче информационных ресурсов и формировании их композиций, при интеграции схем ресурсов в схеме посредника и устранении разнообразных конфликтов, при адекватном преобразовании формул (запросов) программы решения задачи, выраженных в терминах схемы предметной области

задачи, в формулы, выраженные в схемах релевантных ресурсов. Показана важная роль уточнения как математически точного соотношения спецификаций при разрешении названных семантических проблем.

Создание прототипа гибридной инфраструктуры касается главным образом сопряжения исполнительных механизмов двух инфраструктур (АстроГрида и средств поддержки исполнительного слоя предметных посредников). Поэтому в статье основное внимание уделено проблемам переписывания запросов к посредникам в планы их реализации над конкретными информационными ресурсами, приведено краткое описание объединенной архитектуры исполнительных механизмов АстроГрида и средств поддержки предметных посредников, дан пример реализации простого предметного посредника для решения задач поиска далеких галактик в гибридной архитектуре.

Полученные результаты свидетельствуют о перспективности исследованного подхода. Планируется их использование при решении разнообразных задач РВО. Вместе с тем планируется существенное развитие предложенного подхода в ряде направлений, включая увеличение гибкости идеи посредников введением комбинированного подхода к их определению, движимому как приложениями, так и ресурсами; расширение средств языка формул, эффективно поддерживаемых совершенствуемыми алгоритмами переписывания запросов; дальнейшее развитие применения онтологического моделирования в совокупности с развитием средств семантической идентификации релевантных ресурсов и их регистрации в посредниках; развитие эффективности архитектуры посредников введением средств оптимизации планов реализации запросов в распределенной среде, совершенствованием адаптеров для повышения эффективности реализации запросов, более тесным сопряжением с Грид архитектурами; расширение применения предметных посредников в других научных областях и развитие методологии решения задач на основе идеи посредников.

## Литература

1. Briukhov D. O., Kalinichenko L. A., Zakharov V. N., et al. Information infrastructure of the Russian Virtual Observatory (RVO). 2nd ed. — IPI RAN, 2005. 173 p.
2. Virtual observatory architecture overview. Version 1.0, IVOA Note 15 June 2004.
3. Briukhov D., Kalinichenko L., Zakharov V. Diversity of domain descriptions in natural science: Virtual observatory as a case study // 7th Russian Conference on Digital Libraries RCDL2005 Proceedings. Yaroslavl, October 2005. P. 23–30.
4. Abrial J.-R. B-technology: Technical overview. — B-Core (UK) Ltd., 1993.
5. Abrial J.-R. The B-book: Assigning programs to meanings. — Cambridge: Cambridge University Press, 1996.
6. The B-toolkit online documentation. <http://www.b-core.com/ONLINEDOC/BToolkit.html>.
7. Захаров В. Н., Калиниченко Л. А., Соколов И. А., Ступников С. А. Конструирование канонических информационных моделей для интегрированных информационных систем // Информатика и её применения, 2007. Т. 1. Вып. 2. С. 15–38.
8. Kalinichenko L. A., Stupnikov S. A., Martynov D. O. SYNTHESIS: A language for canonical information modeling and mediator definition for problem solving in heterogeneous information resource environments. — М.: IPI RAS, 2007. 171 p.
9. Briukhov D. O., Kalinichenko L. A. Component-based information systems development tool supporting the SYNTHESIS design method // 2nd East-European Conference “Advances in Databases and Information Systems” Proceedings. — Berlin-Heidelberg: Springer-Verlag, 1998. P. 305–327.
10. Briukhov D. O., Kalinichenko L. A., Skvortsov N. A. Information sources registration at a subject mediator as compositional development // East-European Conference “Advances in Databases and Information Systems” Proceedings. Lithuania, Vilnius, Springer, LNCS No. 2151, 2001.
11. Kalinichenko L. A. Compositional specification calculus for information systems development // Conference “Advances in Databases and Information Systems (ADBS’99)” Proceedings. Maribor, Slovenia, September 1999, Springer Verlag, LNCS.
12. AstroGrid. <http://www.astrogrid.org>.
13. RVO public utility center. <http://synthesis.ipi.ac.ru/synthesis/projects/RVOAG/>.
14. Ullman J. D. Information integration using logical views // 6th International Conference on Database Theory (ICDT’97) Proceedings, 1997.
15. Alon Y. Halevy. Answering queries using views: A survey // VLDB J. 2001. Vol. 10. No. 4.
16. Friedman M., Levy A., Millstein T. Navigational plans for data integration // National Conference on Artificial Intelligence (AAAI) Proceedings, 1999.
17. Kalinichenko L. A. Integration of heterogeneous semistructured data models in the canonical one // 1st Russian Conference on Digital Libraries RCDL’99 Proceedings. — St.-Petersburg: St.-Petersburg University, 1999.
18. Briukhov D. O., Kalinichenko L. A., Martynov D. O. Source registration and query rewriting applying LAV/GAV techniques in a typed subject mediator // Труды Девятой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» RCDL, 2007. —

- Переславль-Залесский: Изд-во «Университет города Переславля», 2007. С. 253–262.
19. Wang J., Maher M., Topor R. Rewriting unions of general conjunctive queries using views // 8th International Conference on Extending Database Technology, EDBT'02 Proceedings. Prague, Czech Republic, March 2002.
  20. Kalinichenko L. A., Martynov D. O., Stupnikov S. A. Query rewriting using views in a typed mediator environment // East-European Conference “Advances in Databases and Information Systems” (ADBIS'04) Proceedings. Lecture notes in Computer Science. — Hungary, Budapest: Springer, 2004. Vol. 3255.
  21. Nieuwenhuis R., Oliveras A. Congruence closure with integer offsets // 10th International Conference Logic for Programming Proceedings. Artif. Intell. and Reasoning (LPAR). LNAI. 2003. Vol. 2850. P. 78–90.
  22. Afrati F., Li C., Mitra P. Rewriting queries using views in the presence of arithmetic comparisons // Theor. Comput. Sci. 2006. Vol. 368. No. 1–2. P. 88–123.
  23. The open archives initiative protocol for metadata harvesting. Protocol Version 2.0 of 2002-06-14, Document Version 2004/10/12T15:31:00Z. <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>.
  24. IVOA Astronomical Data Query Language Version 1.01. <http://www.ivoa.net/Documents/WD/ADQL/ADQL-20050624.pdf>.
  25. Spinrad H., Dey A., Stern D., Dunlop J., Peacock J., Jimenez R., Windhorst R. LBDS 53W091: An Old, Red Galaxy at  $z = 1.552$ . *ApJ*. 1997. Vol. 484. P. 581.
  26. Venemans B. P., Rottgering H. J. A., Overzier R. A., Miley G. K., De Breuck C., Kurk J. D., van Breugel W., Carilli C. L., Ford H., Heckman T., McCarthy P., Penttericci L. Discovery of six Ly $\alpha$  emitters near a radio galaxy at  $z \sim 5.2$  // *A&A*. 2004. Vol. 424. P. L17–L20.
  27. Overzier R. A., Miley G. K., Bouwens R. J. et al. Clustering of star-forming galaxies near a radio galaxy at  $z = 5.2$  // *AJ*. 2006. Vol. 637. P. 58–73.
  28. Madau P. Radiative transfer in a clumpy universe: The colors of high-redshift galaxies // *ApJ*. 1995. Vol. 441. P. 18–27.
  29. Rottgering H. J. A., Lacy M., Miley G. K., Chambers K. C., Saunders R. Samples of ultra-steep spectrum radio sources // *A&AS*. 1994. Vol. 108. P. 79–141.
  30. Parijskij Yu., Goss W., Kopylov A., et al. Investigation of RATAN-600 RC radio sources // *Bulletin of Special Astrophys. Obs.* 1996. Vol. 40. P. 5–124.
  31. Douglas J. N., Bash F. N., Torrence G. W., Wolfe C. The Texas survey — Preliminary + 18DEG Strip // *Publ. in Astronomy*. University of Texas. 1980. Vol. 17. No. 1.
  32. Копылов А. И., Госс В. М., Парийский Ю. Н., Соболева Н. С., Желенкова О. П., Темирова А. В., Витковский В. В., Наугольная М. Н., Верходанов О. В. Оптическое отождествление подвыборки радиисточников RC-каталога с крутыми спектрами с помощью 6-метрового телескопа САО РАН. Требования к координатной точности и глубине изображения, наблюдения на VLA и способы оценки фотометрического красного смещения // *АЖ*, 1995. Т. 72. № 4. С. 437–446.
  33. Becker R. H., Helfand D. J., White R. L., et al. The VLA FIRST survey: Faint images of the radio sky at twenty-centimeters. <http://sundog.stsci.edu/top.html>.
  34. Sloan Digital Sky Survey, Data Release 6. <http://www.sdss.org/dr6/>.
  35. Groovy: An agile dynamic language for the Java Platform. <http://groovy.codehaus.org/>.
  36. Grant J., Minker J. A logic-based approach to data integration // *Theory and Practice of Logic Programming*, 2002. Vol. 2(3). P. 293–321.
  37. Lenzerini M. Data integration: A theoretical perspective // *ACM Symposium on Principles of Database Systems (PODS) Proceedings*. 2002.
  38. Калиниченко Л. А., Ступников С. А., Земцов Н. А. Синтез канонических моделей для интеграции неоднородных источников информации. — М.: ИПИ РАН, 2005.
  39. Kalinichenko L. A., Stupnikov S. A., Vovchenko A. E., et al. Russian Virtual Observatory Community Centre for scientific problems solving over multiple distributed information sources // 18th Russian Conference on Digital Libraries RCDL2006 Proceedings. — Suzdal, Russia, 2006. P. 120–129.
  40. Kalinichenko L. Subject mediation approach for scientific problem solving in Virtual Observatories // XXVI General Assembly of the International Astronomical Union. Prague, August 14–25, 2006. P. 454–455.
  41. Kalinichenko L. A., Briukhov D. O., Martynov D. O., Skvortsov N. A., Stupnikov S. A. Mediation framework for enterprise information system infrastructures: Application-driven approach // 9th International Conference on Enterprise Information Systems ICEIS 2007 Proceedings. V. Databases and Information Systems Integration. 2007. P. 246–251.
  42. Haas L. M., Hernández M. A., Ho H., Popa M., Roth M. Clio grows up: From research prototype to industrial tool // *ACM SIGMOD Conference Proceedings*. — Baltimore, Maryland, USA. June 14–16, 2005. P. 805–810.
  43. Jasper E., Tong N., Mc.Brien P., Poulovassilis A. Generating and optimising views from both as view data integration rules // 6th Baltic Conference on Database and Information Systems (DBIS'04) Proceedings. Riga, June 2004.
  44. Kalinichenko L. A., Skvortsov N. A. Extensible ontological modeling framework for subject mediation // 4th All-Russian Conference on Digital Libraries, RCDL'2002 Proceedings. Dubna, October 15–17, 2002.
  45. Kalinichenko L. A. Canonical model development techniques aimed at semantic interoperability in the heterogeneous world of information modeling // Open INTEROP Workshop “Enterprise modeling and ontologies for interoperability” Proceedings at the 16th Conference on “Advanced Information Systems Engineering (CAiSE).” Riga, Latvia, 7–11 June, 2004. P. 101–116.
  46. Van Den Brand M. G., Heering J., Klint P., Oliver P. A. Compiling language definitions: The ASF + SDF compiler // *ACM TOPLAS*. 2002. Vol. 24. No. 4.

47. *Ступников С. А.* Отображение спецификаций, выраженных средствами ядра канонической модели, в язык AMN // Системы и средства информатики: Спец. вып. Формальные методы и модели в композиционных инфраструктурах распределенных информационных систем / Под ред. И. А. Соколова. — М.: ИПИ РАН, 2005. С. 69–95.
48. *Pottinger R., Levy A.* A scalable algorithm for answering queries using views // International Conference on Very Large Data Bases (VLDB) Proceedings, Cairo, Egypt, 2000.
49. *Qian X.* Query folding // International Conference on Data Engineering (ICDE) Proceedings. New Orleans, LA, 1996. P. 48–55.
50. SkyQuery. <http://www.skyquery.net>.
51. *Hanisch R., Quinn P.* The international virtual observatory. <http://www.ivoa.net/pub/info/>.
52. *Alpdemir M. N., Mukherjee A., Gounaris A., Paton N. W., Watson P., Fernandes A. A. A., Fitzgerald D. J.* OGSA-DQP: A service for distributed querying on the grid // Proceedings of EDBT 2004, Springer, LNCS 2992, March 2004.

# КВАЗИЛИНЕЙНЫЕ МЕТОДЫ ПОСТРОЕНИЯ ИНФОРМАЦИОННЫХ МОДЕЛЕЙ ФЛУКТУАЦИЙ НЕРАВНОМЕРНОСТИ ВРАЩЕНИЯ ЗЕМЛИ\*

И. Н. Сеницын<sup>1</sup>

**Аннотация:** В основе многих современных стохастических информационных технологий научных исследований лежат методы построения информационных моделей по априорным и апостериорным данным. В статье рассматривается применение корреляционных методов построения информационных моделей флуктуаций неравномерности вращения Земли на внутригодовых интервалах времени, основанных на эквивалентной линеаризации нелинейных стохастических уравнений. Приводятся примеры приложения методов и программного обеспечения из состава информационных ресурсов «Статистическая динамика вращения Земли».

**Ключевые слова:** флуктуации неравномерности вращения Земли; информационная модель; корреляционные характеристики; корреляционные методы; эквивалентная линеаризация; стохастические дифференциальные уравнения

## 1 Введение

В [1–10] на основе априорных данных о динамической гравитационно-приливной структуре Земли построены корреляционные и кинетические стохастические модели флуктуаций угловой скорости собственного вращения в составе совокупной модели вращательного движения Земли.

В [11] предложена линейная модель возмущенных регулярных собственных вращательных движений Земли на внутригодовых интервалах времени.

Влияние аддитивных и линейных мультипликативных (параметрических) гармонических и широкополосных гауссовских случайных возмущений на одно- и двумерные корреляционные характеристики флуктуаций приливной неравномерности вращения Земли изучено в [12].

Рассмотрим методы построения квазилинейных информационных моделей флуктуаций приливной неравномерности вращения Земли в случае нелинейного механизма диссипации на основе как априорной, так и апостериорной информации о неравномерности вращения Земли. При этом учтем также аддитивные и параметрические гармонические и гауссовские широкополосные случайные гравитационно-приливные диссипативные возмущения.

## 2 Стохастические дифференциальные уравнения флуктуаций неравномерности вращения Земли

Обобщая математические модели [1–12] флуктуаций угловой скорости собственного вращения Земли на случай воздействия как аддитивных, так и параметрических гармонических и широкополосных случайных гравитационно-приливных и флуктуационно-диссипативных возмущений на внутригодовых интервалах времени, представим дифференциальное уравнение изменения угла собственного вращения Земли  $\delta\varphi$  в следующем виде:

$$\begin{aligned} \delta\ddot{\varphi} = & M_{10}^S \cos(2\pi f_{\Gamma}t + \chi_1^S) + M_{20}^S \cos(4\pi f_{\Gamma}t + \chi_2^S) + \\ & + M_{m0}^L \cos(2\pi\nu_m t + \chi_m^L) + M_{f0}^L \cos(2\pi\nu_f t + \chi_f^L) + \\ & + X^S(t) + X^L(t) - \mu_1 \left[ 1 + \pi_{11}^{\mu S} \cos(2\pi f_{\Gamma}t + \chi_{11}^{\mu S}) + \right. \\ & + \pi_{12}^{\mu S} \cos(4\pi f_{\Gamma}t + \chi_{12}^{\mu S}) + \pi_{1m}^{\mu L} \cos(2\pi\nu_m t + \chi_{1m}^{\mu L}) + \\ & + \pi_{1f}^{\mu L} \cos(2\pi\nu_f t + \chi_{1f}^{\mu L}) + X_1^{\mu S}(t) + X_1^{\mu L}(t) \left. \right] \delta\dot{\varphi} + \\ & + \mu_n \left[ 1 + \pi_{n1}^{\mu S} \cos(2\pi f_{\Gamma}t + \chi_{n1}^{\mu S}) + \right. \\ & \left. + \pi_{n2}^{\mu S} \cos(4\pi f_{\Gamma}t + \chi_{n2}^{\mu S}) + \right. \end{aligned}$$

\* Работа выполнена при финансовой поддержке РФФИ (проект № 07-07-00031) и программы ОИТВС РАН «Фундаментальные основы информационных технологий и систем» (проект 1.5).

<sup>1</sup> Институт проблем информатики Российской академии наук, sinitsin@dol.ru

$$+ \pi_{nm}^{\mu L} \cos(2\pi\nu_m t + \chi_{nm}^{\mu L}) + \pi_{nf}^{\mu L} \cos(2\pi\nu_f t + \chi_{nf}^{\mu L}) + X_n^{\mu S}(t) + X_n^{\mu L}(t) \Big] F_n(\delta\varphi, \dot{\varphi}). \quad (1)$$

Здесь введены следующие обозначения:  $\delta\varphi = \varphi - r_*\omega_*^{-1}t$  — изменение угла собственного вращения Земли;  $\varphi$  — угол собственного вращения Земли, значения которого определены на дату  $t$  (он является параметром, характеризующим вращение земной системы координат по отношению к небесной);  $r_* = 7,292115 \cdot 10^{-5}$  рад/с — постоянная средняя составляющая угловой скорости собственного вращения Земли;  $\omega_*$  — угловая скорость обращения Земли по орбите;  $t$  — безразмерное время, измеряемое стандартными годами;  $f_\Gamma = 1, 2f_\Gamma$ ,  $\nu_m = 13,28$ ,  $\nu_f = 26,28$ ,  $M_{10}^S$ ,  $M_{20}^S$ ,  $M_{m0}^L$ ,  $M_{f0}^L$  и  $\chi_1^S$ ,  $\chi_2^S$ ,  $\chi_m^L$ ,  $\chi_f^L$  — частоты, амплитуды и начальные фазы аддитивных гармонических возмущений от Солнца ( $S$ ) и Луны ( $L$ ), соответствующие годовому, полугодовому, месячному и двухнедельному циклам;  $\mu_h$  ( $h = 1, n$ ) — коэффициенты диссипативных моментов сил, обусловленных разнообразием геофизических процессов (приливное трение океанических и земных приливов, атмосферные воздействия, океанические течения, перераспределение водных масс и т.п.);  $\pi_{h1}^{\mu S}$ ,  $\pi_{h2}^{\mu S}$ ,  $\pi_{hm}^{\mu L}$ ,  $\pi_{hf}^{\mu L}$  и  $\chi_{h1}^{\mu S}$ ,  $\chi_{h2}^{\mu S}$ ,  $\chi_{hm}^{\mu L}$ ,  $\chi_{hf}^{\mu L}$  ( $h = 1, n$ ) — амплитуды и начальные фазы параметрических гармонических диссипативных моментов сил на частотах  $f_\Gamma$ ,  $2f_\Gamma$ ,  $\nu_m$ ,  $\nu_f$ ;  $F_n = F_n(\delta\varphi, \dot{\varphi})$  — нелинейная составляющая диссипации (в случае рэлеевского механизма диссипации  $F_n(\delta\dot{\varphi}) = \delta\dot{\varphi}^3$ );  $X^S(t)$ ,  $X^L(t)$  и  $X_h^{\mu S}(t)$ ,  $X_h^{\mu L}(t)$  — нормальные (гауссовские) широкополосные аддитивные и параметрические случайные возмущения с известными математическими ожиданиями и ковариационными характеристиками.

### 3 Квазилинейные методы построения информационных корреляционных моделей по априорным данным

Составим приближенные уравнения для математических ожиданий, дисперсий, ковариаций и ковариационных функций переменных  $\delta\varphi = X_1$ ,  $\delta\dot{\varphi} = X_2$  при следующих трех основных допущениях.

<sup>10</sup> Аддитивный возмущающий момент допускает представление  $X^S(t) + X^L(t) = m_0^{SL} + V_1$ , где  $V_1$  является скалярным нормальным белым шумом интенсивности  $\nu_1 = \nu_1(t)$ .

<sup>20</sup> Возмущающие моменты  $X_3 = X_1^{\mu S}(t) + X_1^{\mu L}(t)$  и  $X_4 = X_n^{\mu S}(t) + X_n^{\mu L}(t)$  происходят от одного источника нормального белого шума  $V_2$  единичной интенсивности ( $\nu_2 = 1$ ). Они имеют конечные математические ожидания  $m_3$  и  $m_4$  и дисперсии  $\sigma_3^2$  и  $\sigma_4^2$  и удовлетворяют следующим скалярным уравнениям формирующего фильтра [13]:

$$\begin{aligned} \dot{X}_l^0 &= -\alpha_l X_l^0 + \sigma_l \sqrt{2\alpha_l} V_2; \\ X_l^0 &= X_l - m_l^*, \quad l = 3, 4, \quad \alpha_l > 0, \end{aligned} \quad (2)$$

где  $m_l^*$  — постоянные значения  $m_l$ .

<sup>30</sup> Нелинейные функции  $X_2 X_3$ ,  $F_n = F_n(X_1, X_2)$ ,  $F'_n = F'_n(X_1, X_2, X_4) = F_n(X_1, X_2) X_4$ , входящие в (1), допускают статистическую линеаризацию нелинейностей по формулам [13]:

$$\begin{aligned} X_2 X_3 &\approx m_2 m_3 + k_{23} + m_2 X_3^0 + m_3 X_2^0; \\ F_n &= F_n(X_1, X_2) \approx F_{n0} + F_{n1} X_1^0 + F_{n2} X_2^0; \\ F'_2 &= F_1(X_1, X_2) X_4 \approx \\ &\approx F'_{n0} + F'_{n1} X_1^0 + F'_{n2} X_2^0 + F'_{n4} X_4^0. \end{aligned}$$

Здесь  $m_i$  и  $k_{ij}$  — математические ожидания и ковариационные моменты переменных  $X_i$ ,  $X_i^0 = X_i - m_i$ ,

$$\begin{aligned} F_{n0} &= F_{n0}(m_1, m_2, k_{11}, k_{12}, k_{22}) = \\ &= \mathbf{M}_N^{(1,2)} [F_n(X_1, X_2)]; \end{aligned}$$

$$F_{n1} = \frac{\partial F_{n0}}{\partial m_1}; \quad F_{n2} = \frac{\partial F_{n0}}{\partial m_2};$$

$$\begin{aligned} F'_{n0} &= \\ &= F'_{n0}(m_1, m_2, m_4, k_{11}, k_{12}, k_{22}, k_{14}, k_{24}, k_{44}) = \\ &= \mathbf{M}_N^{(1,2,4)} [F'_n(X_1, X_2, X_4)]; \end{aligned}$$

$$F'_{n1} = \frac{\partial F'_{n0}}{\partial m_1}; \quad F'_{n2} = \frac{\partial F'_{n0}}{\partial m_2}; \quad F'_{n4} = \frac{\partial F'_{n0}}{\partial m_4},$$

где  $\mathbf{M}_N^{(1,2)}[\cdot]$  и  $\mathbf{M}_N^{(1,2,4)}[\cdot]$  — символы вероятностного осреднения для двух- и трехмерного нормального распределения переменных  $X_1, X_2$  и  $X_1, X_2, X_4$ .

В результате стохастическое нелинейное дифференциальное уравнение второго порядка (1) будет эквивалентно нелинейной системе для математических ожиданий  $m_i = \mathbf{M}X_i$  ( $i = \overline{1, 4}$ ):

$$\begin{aligned} \dot{m}_1 &= m_2; \\ \dot{m}_2 &= M_0^\ominus; \\ \dot{m}_l &= -\alpha_l (m_l - m_l^*) \quad (l = 3, 4), \end{aligned} \quad (3)$$

где введены обозначения:

$$\begin{aligned} M_0^\ominus &= \tilde{M}_{0t} + m_0^{SL} - (\mu_1 + \tilde{\mu}_{1t}) m_2 - \\ &- \mu_1 (m_2 m_3 + k_{23}) + (\mu_n + \tilde{\mu}_{nt}) F_{n0} + \mu_n F'_{n0}; \\ \tilde{M}_{0t} &= M_{10}^S \cos(2\pi f_\Gamma t + \chi_1^S) + \\ &+ M_{20}^S \cos(4\pi f_\Gamma t + \chi_2^S) + M_{m0}^L \cos(2\pi \nu_m t + \chi_m^L) + \\ &+ M_{f0}^L \cos(2\pi \nu_f t + \chi_f^L); \\ \tilde{\mu}_{lt} &= \mu_l \left[ \pi_{l1}^{\mu S} \cos(2\pi f_\Gamma t + \chi_{l1}^{\mu S}) + \right. \\ &+ \pi_{l2}^{\mu S} \cos(4\pi f_\Gamma t + \chi_{l2}^{\mu S}) + \pi_{lm}^{\mu L} \cos(2\pi \nu_m t + \chi_{lm}^{\mu L}) + \\ &\left. + \pi_{lf}^{\mu L} \cos(2\pi \nu_f t + \chi_{lf}^{\mu L}) \right] \quad (l = 1, 3), \quad (4) \end{aligned}$$

и линейной системе для центрированных составляющих:

$$\begin{aligned} \dot{X}_1^0 &= X_2^0; \\ \dot{X}_2^0 &= V_1 + [(\mu_n + \tilde{\mu}_{nt}) F_{n1} + \mu_n F'_{n1}] X_1^0 - \\ &- [(\mu_1 + \tilde{\mu}_{1t}) + \mu_1 m_3 - (\mu_n + \tilde{\mu}_{nt}) F_{n2} - \\ &- \mu_n F'_{n2}] X_2^0 - \mu_1 m_2 X_3^0 + \mu_n F'_{n4} X_4^0; \\ \dot{X}_l^0 &= \alpha_l X_l^0 + \sigma_l \sqrt{2\alpha_l} V_2. \end{aligned}$$

Обозначая через  $\varepsilon$  и  $\beta$  матрицы

$$\begin{aligned} \varepsilon = \varepsilon(t) &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} & \varepsilon_{24} \\ 0 & 0 & -\alpha_3 & 0 \\ 0 & 0 & 0 & -\alpha_4 \end{bmatrix}; \\ \beta &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \sigma_3 \sqrt{2\alpha_3} \\ 0 & \sigma_4 \sqrt{2\alpha_4} \end{bmatrix}, \end{aligned} \quad (5)$$

где

$$\begin{aligned} \varepsilon_{21} &= (\mu_n + \tilde{\mu}_{nt}) F_{n1} + \mu_n F'_{n1}; \\ \varepsilon_{22} &= -(\mu_1 + \tilde{\mu}_{1t}) - \mu_1 m_3 + \\ &+ (\mu_n + \tilde{\mu}_{nt}) F_{n2} + \mu_n F'_{n2}; \quad (6) \\ \varepsilon_{23} &= -\mu_1 m_2; \\ \varepsilon_{24} &= \mu_n F'_{n1}, \end{aligned}$$

представим уравнения для ковариационной матрицы  $K(t) = [k_{ij}(t)]$ ,  $k_{ij}(t) = M[X_i^0(t)X_j^0(t)]$  и матрицы для ковариационных функций  $K(t_1, t_2) = [K_{ij}(t_1, t_2)]$ ,  $K_{ij}(t_1, t_2) = M[X_i^0(t_1)X_j^0(t_2)]$  ( $i, j = \overline{1, 4}$ ) в следующем виде [13]:

$$\dot{K}(t) = \varepsilon(t)K(t) + K(t)\varepsilon^T(t) + \beta\nu\beta^T; \quad (7)$$

$$\begin{aligned} k_{ij}(0) &= k_{ij0}; \\ \frac{\partial K(t_1, t_2)}{\partial t_2} &= K(t_1, t_2)\varepsilon^T(t_2); \\ K_{ij}(t, t') &= k_{ij}(t), \end{aligned} \quad (8)$$

где  $\nu = [\nu_{ij}]$  ( $\nu_{11} = \nu_1(t)$ ,  $\nu_{12} = \nu_{12}(t)$ ,  $\nu_{22} = 1$ ) — матрица интенсивностей белых шумов  $V_1$  и  $V_2$ .

Уравнения (7) и (8) допускают следующую развернутую запись:

$$\begin{aligned} \dot{k}_{11} &= 2k_{12}; \\ \dot{k}_{12} &= \varepsilon_{21}k_{11} + \varepsilon_{22}k_{12} + \varepsilon_{23}k_{13} + \varepsilon_{24}k_{14} + k_{22}; \\ \dot{k}_{22} &= 2(\varepsilon_{21}k_{12} + \varepsilon_{22}k_{22} + \varepsilon_{23}k_{23} + \varepsilon_{24}k_{24}) + \nu_1; \\ \dot{k}_{13} &= k_{23} - \alpha_3 k_{13}; \\ \dot{k}_{23} &= \varepsilon_{21}k_{13} + \varepsilon_{22}k_{23} + \varepsilon_{23}k_{33} + \varepsilon_{24}k_{34} - \\ &- \alpha_3 k_{23} + \nu_{12}\sigma_3\sqrt{2\alpha_3}; \quad (9) \\ \dot{k}_{33} &= -2\alpha_3(k_{33} - \sigma_3^2); \\ \dot{k}_{14} &= k_{24} - \alpha_4 k_{14}; \\ \dot{k}_{24} &= \varepsilon_{21}k_{14} + \varepsilon_{22}k_{24} + \varepsilon_{23}k_{34} + \varepsilon_{24}k_{44} - \\ &- \alpha_4 k_{24} + \nu_{12}\sigma_4\sqrt{2\alpha_4}; \\ \dot{k}_{34} &= -(\alpha_3 + \alpha_4)k_{34} + 2\sigma_3\sigma_4\sqrt{\alpha_3\alpha_4} \\ \dot{k}_{44} &= -2\alpha_4(k_{44} - \sigma_4^2) \end{aligned}$$

при начальных условиях  $k_{ij}(0) = k_{ij0}$  и

$$\begin{aligned} \frac{\partial K_{11}(t_1, t_2)}{\partial t_2} &= K_{12}(t_1, t_2); \\ \frac{\partial K_{21}(t_1, t_2)}{\partial t_2} &= K_{22}(t_1, t_2); \\ \frac{\partial K_{31}(t_1, t_2)}{\partial t_2} &= K_{32}(t_1, t_2); \\ \frac{\partial K_{41}(t_1, t_2)}{\partial t_2} &= K_{42}(t_1, t_2); \\ \frac{\partial K_{12}(t_1, t_2)}{\partial t_2} &= \varepsilon_{21,t_2}K_{11}(t_1, t_2) + \varepsilon_{22,t_2}K_{12}(t_1, t_2) + \\ &+ \varepsilon_{23,t_2}K_{13}(t_1, t_2) + \varepsilon_{24,t_2}K_{14}(t_1, t_2); \\ \frac{\partial K_{22}(t_1, t_2)}{\partial t_2} &= \varepsilon_{21,t_2}K_{21}(t_1, t_2) + \varepsilon_{22,t_2}K_{22}(t_1, t_2) + \\ &+ \varepsilon_{23,t_2}K_{23}(t_1, t_2) + \varepsilon_{24,t_2}K_{24}(t_1, t_2); \\ \frac{\partial K_{32}(t_1, t_2)}{\partial t_2} &= \varepsilon_{21,t_2}K_{31}(t_1, t_2) + \varepsilon_{22,t_2}K_{32}(t_1, t_2) + \\ &+ \varepsilon_{23,t_2}K_{33}(t_1, t_2) + \varepsilon_{24,t_2}K_{34}(t_1, t_2); \quad (10) \\ \frac{\partial K_{42}(t_1, t_2)}{\partial t_2} &= \varepsilon_{21,t_2}K_{41}(t_1, t_2) + \varepsilon_{22,t_2}K_{42}(t_1, t_2) + \\ &+ \varepsilon_{23,t_2}K_{43}(t_1, t_2) + \varepsilon_{24,t_2}K_{44}(t_1, t_2); \\ \frac{\partial K_{13}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{13}(t_1, t_2); \\ \frac{\partial K_{23}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{23}(t_1, t_2); \\ \frac{\partial K_{33}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{33}(t_1, t_2); \end{aligned}$$

$$\begin{aligned}\frac{\partial K_{43}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{33}(t_1, t_2); \\ \frac{\partial K_{14}(t_1, t_2)}{\partial t_2} &= -\alpha_4 K_{14}(t_1, t_2); \\ \frac{\partial K_{24}(t_1, t_2)}{\partial t_2} &= -\alpha_4 K_{24}(t_1, t_2); \\ \frac{\partial K_{34}(t_1, t_2)}{\partial t_2} &= -\alpha_4 K_{34}(t_1, t_2); \\ \frac{\partial K_{44}(t_1, t_2)}{\partial t_2} &= -\alpha_4 K_{44}(t_1, t_2)\end{aligned}$$

при начальных условиях  $K_{ij}(t, t) = k_{ij}(t)$  ( $i, j = \overline{1, 4}$ ).

Совокупность уравнений (3), (9) и (10) определяет аналитическую нелинейную корреляционную модель флуктуаций приливной внутрigoдовой неравномерности вращения Земли.

**Замечание 1.** Усредняя по периодам основных гармоник, отвечающих частотам  $f_\Gamma, 2f_\Gamma, \nu_m$  и  $\nu_f$  для стационарных регулярных и нерегулярных колебаний и постоянных  $m_h = m_h^*$  ( $h = 2, 3, 4$ ), из уравнений (3), (9) и (10) находим соответствующие уравнения для математических ожиданий, ковариационной матрицы, матрицы ковариационных функций и матрицы спектральных плотностей.

**Замечание 2.** Уравнения (3), (9) и (10) применимы для нелинейных функций  $F_n = F_n(X_1, X_2)$ , допускающих угловые точки и даже разрывы. В случае гладких функций  $F_n$  рассмотренный квазилинейный метод переходит в метод непосредственной линеаризации в окрестностях математических ожиданий  $m_1$  и  $m_2$ .

**Замечание 3.** Если в исходном уравнении (1) возмущения  $X^S(t), X^L(t), X_1^{\mu S}(t), X_1^{\mu L}(t), X_n^{\mu S}(t)$  и  $X_n^{\mu L}(t)$  заданы совместными каноническими представлениями, то целесообразно использовать метод нормализации посредством канонических представлений [13].

**Замечание 4.** Для негауссовских возмущений в (1), как показано в [13–16], можно воспользоваться методом эквивалентной линеаризации, взяв в качестве осредняющего вероятностного распределения отрезков параметризованного (моментами, квазимоментами, семиинвариантами и др.) разложения одно- и двумерных плотностей.

## 4 Линейная корреляционная модель флуктуаций неравномерности вращения Земли (тестовые примеры 1 и 2)

*Пример 1.* Уравнения (3), (9) и (10) при  $\mu_1 = \mu$  и  $\mu_n = 0$  (отсутствуют параметрические и нелинейные возмущения) имеют следующий вид:

$$\begin{aligned}\dot{m}_1 &= m_2; \\ \dot{m}_2 &= \tilde{M}_{0t} - (\mu + \tilde{\mu}_t) m_2 - \\ &\quad - \mu(m_2 m_3 + k_{23}) + m_0^{SL};\end{aligned}\quad (11)$$

$$\begin{aligned}\dot{m}_3 &= -\alpha_3(m_3 - m_3^*); \\ \dot{k}_{11} &= 2k_{12}; \\ \dot{k}_{12} &= k_{22} - [\mu(1 + m_3) + \tilde{\mu}_t] k_{12} - \mu m_2 k_{13}; \\ \dot{k}_{22} &= -2[\mu(1 + m_3) + \tilde{\mu}_t] k_{22} - 2\mu m_2 k_{23} + \nu_{11}; \\ \dot{k}_{13} &= -\alpha_3 k_{13} + k_{23};\end{aligned}\quad (12)$$

$$\begin{aligned}\dot{k}_{23} &= -[\alpha_3 + \mu(1 + m_3) + \tilde{\mu}_t] k_{23} - \\ &\quad - \mu m_2 k_{33} + \nu_{12} \sigma_3 \sqrt{2\alpha_3}; \\ \dot{k}_{33} &= -2\alpha_3(k_{33} - \sigma_3^2)\end{aligned}$$

при начальных условиях  $K_{ij}(0) = k_{ij,0}$  ( $i, j = 1, 2, 3$ ) и

$$\frac{\partial K_{11}(t_1, t_2)}{\partial t_2} = K_{12}(t_1, t_2);$$

$$\frac{\partial K_{21}(t_1, t_2)}{\partial t_2} = K_{22}(t_1, t_2);$$

$$\frac{\partial K_{31}(t_1, t_2)}{\partial t_2} = K_{32}(t_1, t_2);$$

$$\begin{aligned}\frac{\partial K_{12}(t_1, t_2)}{\partial t_2} &= -[\mu(1 + m_{3t_2}) + \tilde{\mu}_{t_2}] K_{12}(t_1, t_2) - \\ &\quad - \mu m_{2t_2} K_{13}(t_1, t_2); \\ \frac{\partial K_{22}(t_1, t_2)}{\partial t_2} &= -[\mu(1 + m_{3t_2}) + \tilde{\mu}_{t_2}] K_{22}(t_1, t_2) - \\ &\quad - \mu m_{2t_2} K_{23}(t_1, t_2);\end{aligned}\quad (13)$$

$$\begin{aligned}\frac{\partial K_{32}(t_1, t_2)}{\partial t_2} &= -[\mu(1 + m_{3t_2}) + \tilde{\mu}_{t_2}] K_{32}(t_1, t_2) - \\ &\quad - \mu m_{2t_2} K_{33}(t_1, t_2); \\ \frac{\partial K_{13}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{13}(t_1, t_2); \\ \frac{\partial K_{23}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{23}(t_1, t_2); \\ \frac{\partial K_{33}(t_1, t_2)}{\partial t_2} &= -\alpha_3 K_{33}(t_1, t_2)\end{aligned}$$

при начальных условиях  $K_{ij}(t, t) = k_{ij}(t)$  ( $i, j = 1, 2, 3$ ). Здесь  $\tilde{\mu}_t = \tilde{\mu}_{lt}$  определены в (11), а  $\nu_1 = \nu_{11}$  — интенсивность белого шума  $V_1$ ;  $\nu_{12}$  — взаимная интенсивность шумов  $V_1$  и  $V_2$ .

Рассмотрим основные свойства модели (11)–(13). В условиях аддитивных гармонических и случайных моментов ( $\tilde{M}_{0t} \neq 0, X^S(t) \neq 0, X^L(t) \neq 0$ ) при отсутствии диссипативных моментов ( $\mu = 0$ ) имеют место следующие явные выражения для моментов первого и второго порядков:

$$m_1 = m_{10} + m_{20}t + m_0^{SL} \frac{t^2}{2} + \int_0^t \int_0^{t_1} \tilde{M}_{0t_1} dt_1 dt; \quad (14)$$

$$m_2 = m_{20} + m_0^{SL}t + \int_0^t \tilde{M}_{0t_1} dt_1,$$

$$k_{11} = \frac{1}{3}\nu_{11}t^3 + k_{22,0}t^2 + 2k_{12,0}t + k_{11,0};$$

$$k_{12} = \frac{\nu_{11}t^2}{2} + k_{22,0}t + k_{12,0}; \quad (15)$$

$$k_{22} = \nu_{11}t + k_{22,0};$$

$$K_{11}(t_1, t_2) = k_{11}(t_1) + k_{12}(t_1)(t_2 - t_1);$$

$$K_{12}(t_1, t_2) = k_{12}(t_1);$$

$$K_{22}(t_1, t_2) = k_{22}(t_1).$$

Из формул (14) и (15) следует, что при нулевых начальных условиях, во-первых, постоянный момент  $m_0^{SL}$  вызывает систематические временные дрейфы по переменным  $\delta\varphi$  и  $\delta\dot{\varphi}$ ; во-вторых, гармонический момент  $\tilde{M}_{0t}$  может приводить к накопленной интегральной ошибке по переменной  $\delta\varphi$ ; в-третьих, аддитивный белый шум  $V_1$  вызывает флуктуационный временной дрейф по переменным  $\delta\varphi$  и  $\delta\dot{\varphi}$ . Ненулевые начальные условия по переменной  $\delta\dot{\varphi}$  вызывают флуктуационный дрейф по переменной  $\delta\varphi$ .

*Пример 2.* При линейной диссипации, когда  $\mu \neq 0$  для  $t \gg \mu^{-1}$  и начальных условиях  $m_{10} = 0$ ,  $m_{20} = 0$ ,  $m_{30} = m_3^*$  и  $m_0^{SL} = 0$  усредненные на интервалах времени  $2\pi/f_\Gamma$ ,  $2\pi/\nu_m$  и  $2\pi/\nu_f$  математические ожидания, дисперсии и ковариационные моменты имеют вид:

$$\langle m_1 \rangle = -\frac{\lambda t^2}{2};$$

$$\langle m_2 \rangle = -\lambda t;$$

$$\langle m_3 \rangle = m_3^*,$$

$$\text{где } \lambda = \langle \tilde{\mu}_t \int_0^t \tilde{M}_{0t_1} dt_1 \rangle;$$

$$\langle k_{11} \rangle = \frac{\nu_{11}}{\mu^2} t;$$

$$\langle k_{12} \rangle = \frac{\nu_{11}}{2\mu^2};$$

$$\langle k_{22} \rangle = \frac{\nu_{11}}{2\mu};$$

$$\langle k_{13} \rangle = \sigma_3 \sqrt{2\alpha_3} \frac{\nu_{12}}{\alpha_3 (\mu + \alpha_3)};$$

$$\langle k_{23} \rangle = \sigma_3 \sqrt{2\alpha_3} \frac{\nu_{12}}{\mu + \alpha_3}.$$

Здесь  $\langle \dots \rangle$  — символ усреднения. Отсюда видно, что имеют место усредненные систематические дрейфы  $\langle m_1 \rangle$  и  $\langle m_2 \rangle$  и флуктуационный дрейф  $\langle k_{11} \rangle$ . При этом  $\langle k_{ij} \rangle$  ( $i, j = 2, 3$ ) постоянны и зависят от интенсивности  $\nu_{11}$  аддитивного шума  $V_1$  и взаимной интенсивности аддитивного и параметрического шумов  $V_1$  и  $V_2$ . Для  $t \ll \mu^{-1}$  соответствующие результаты следуют из (14) и (15).

## 5 Квазилинейная корреляционная модель флуктуаций при аддитивных возмущениях и нелинейной рэлеевской диссипации (тестовый пример 3)

*Пример 3.* В случае рэлеевского механизма диссипации, когда  $F_n = \delta\dot{\varphi}^3$ , имеем следующие выражения для коэффициентов статистической линеаризации функций  $F_n$  и  $F'_n$ :

$$F_n = X_2^3 \approx F_{0n} + F_{n2} X_2^0 = m_2^3 + 3k_{22}m_2 + 3(m_2^2 + k_{22}) X_2^0; \quad (16)$$

$$F'_n = X_2^3 X_4 \approx F'_{0n} + F'_{n2} X_2^0 + F'_{n4} X_4^0 = m_2^3 m_4 + 3m_2 m_4 k_{22} + 3m_2^2 k_{24} + 3k_{22} k_{24} + 3(m_2^2 m_4 + m_4 k_{22} + 2m_2 k_{24}) X_2^0 + m_2 (m_2^2 + 3k_{22}) X_4^0. \quad (17)$$

Тогда при аддитивных гармонических и случайных возмущениях уравнения (3) и (9) для  $\delta\varphi = X_1$  и  $\delta\dot{\varphi} = X_2$  имеют следующий вид:

$$\dot{m}_1 = m_2; \quad (18)$$

$$\dot{m}_2 = \tilde{M}_{0t} + m_0^{SL} - \mu_1 m_2 + \mu_3 (m_2^3 + 3m_2 k_{22});$$

$$\dot{k}_{11} = 2k_{12};$$

$$\dot{k}_{12} = k_{22} - \mu^\ominus k_{12}; \quad (19)$$

$$\dot{k}_{22} = -2\mu^\ominus k_{22} + 2\mu_3 m_2 (m_2^2 + k_{22}) + \nu_{11},$$

где  $\mu^\ominus = [\mu_1 - 3\mu_3 (m_2^2 + k_{22})]$ . Отсюда при  $m_0^{SL} = 0$  для  $t \ll 1/\mu^\ominus$  после усреднения стационарные решения  $m_2^*$  и  $k_{22}^*$  будут определяться уравнениями:

$$-\mu_1 m_2^* + \mu_3 (m_2^{*2} + 3k_{22}^*) m_2^* = 0; \quad (20)$$

$$-2 [\mu_1 - 3\mu_3 (m_2^{*2} + k_{22}^*)] k_{22}^* + 2\mu_3 m_2^* (m_2^{*2} + k_{22}^*) + \nu_{11} = 0. \quad (21)$$

При условии

$$\mu^\ominus = \frac{1}{2} \left( \mu_1 \pm \sqrt{\mu_1^2 - 6\mu_3\nu_{11}} \right) > 0,$$

во-первых, имеют место стационарные решения:

$$m_2^* = 0, \quad k_{22}^* = \frac{\mu_1 \pm \sqrt{\mu_1^2 - 6\mu_3\nu_{11}}}{6\mu_3},$$

отвечающие двум режимам нерегулярных колебаний, соответственно, при малых и больших  $\mu_3$ . Во-вторых, уравнения (20) и (21) допускают стационарные решения  $m_2^*$  и  $k_{22}^*$ , определяемые из уравнения:

$$-\mu_1 + \mu_3 (m_2^{*2} + 3k_{22}^*) = 0.$$

Они отвечают смещенным на величину  $m_2^*$  нерегулярным колебаниям.

**Замечание 5.** По переменной  $\delta\varphi = X_1$  стационарные регулярные и нерегулярные колебания сопровождаются систематическим и флуктуационным дрейфами, определяемыми в силу (9) уравнениями:

$$\begin{aligned} \dot{m}_1 &= m_2^*; \\ \dot{k}_{11} &= 2k_{12}; \\ \dot{k}_{12} &= k_{22} - \langle \mu^\ominus \rangle k_{12} - \mu_1 m_2^* k_{13} + \mu_3 m_2^* (m_2^{*2} + k_{22}^*) k_{14}; \\ \dot{k}_{13} &= k_{23}^* - \alpha_3 k_{13}; \\ \dot{k}_{14} &= k_{24}^* - \alpha_4 k_{14} \end{aligned}$$

при начальных условиях:  $m_1 = m_{10}$ ,  $k_{11} = k_{110}$ ,  $k_{12} = k_{120}$ ,  $k_{13} = k_{130}$ ,  $k_{14} = k_{140}$ .

**Замечание 6.** В силу известного линейного соотношения между вариациями скорости вращения Земли  $\delta\dot{\varphi}(t)$  и длительностью суток (*l.o. d(t)*) [14]

$$l.o. d(t) = -\frac{86400}{r_*} \delta\dot{\varphi}(t)$$

уравнения (3), (9), (10), а также (18)–(21) позволяют получить соответствующие уравнения для корреляционных характеристик *l.o. d(t)*.

## 6 Квазилинейные информационные модели флуктуаций неравномерности вращения Земли по апостериорным данным

Примем за информационные переменные  $Z_1$  и  $Z_2$ , допускающие измерения  $X_1$  и  $X_2$ . Положим

$$Z_1 = X_1 + V_3; \quad Z_2 = X_2 + V_4, \quad (22)$$

где  $X_1 = \delta\varphi$ ,  $X_2 = \delta\dot{\varphi}$ ;  $V_3$  и  $V_4$  — независимые нормальные белые шумы с интенсивностями  $\nu_3$  и  $\nu_4$  соответственно. Тогда совокупность уравнений (1), (2) и (22) будет представлять собой исходную систему уравнений для синтеза квазилинейного фильтра для обработки информации о флуктуациях неравномерности вращения Земли по апостериорным данным, т. е. по результатам измерения  $Z_1$  и  $Z_2$ .

Для построения квазилинейного нормального фильтра согласно [15, 16] перепишем уравнения (1), (2) и (22) в следующем стандартном виде:

$$\dot{X} = a(X, t) + b(t)\bar{V}_1; \quad Z = a_1(X, t) + \bar{V}_2. \quad (23)$$

Здесь

$$\begin{aligned} X &= [X_1 X_2 X_3 X_4]^T; \\ Z &= [Z_1 Z_2]^T; \\ \bar{V}_1 &= [V_1 V_2]^T; \\ \bar{V}_2 &= [V_3 V_4]^T; \\ a &= a(X, t) = \begin{bmatrix} X_2 \\ \tilde{M}_{0t} m_0^{SL} - [(\mu_1 + \tilde{\mu}_{1t}) + \mu_1 X_3] X_2 + \\ + [(\mu_n + \tilde{\mu}_{nt}) + \mu_n X_4] F_n(X_1, X_2) \\ -\alpha_3(X_3 - m_3^*) \\ -\alpha_4(X_4 - m_4^*) \end{bmatrix}; \end{aligned}$$

$$a_1 = a_1(X, t) = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix};$$

$$b = b(t) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \sigma_3 \sqrt{2\alpha_3} \\ 0 & \sigma_4 \sqrt{2\alpha_4} \end{bmatrix};$$

$$\bar{\nu}_1 = \begin{bmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{bmatrix};$$

$$\bar{\nu}_2 = \begin{bmatrix} \nu_3 & 0 \\ 0 & \nu_4 \end{bmatrix}.$$

Заменим (23) статистически линеаризованной системой уравнений, нелинейной относительно математических ожиданий  $m^x = [m_1^x m_2^x m_3^x m_4^x]^T$ ,  $m^z = [m_1^z m_2^z]$  и линейной относительно центрированных составляющих  $X^0 = X - m^x$ :

$$\dot{m}^x = a_{00}, \quad m^z = a_{10}, \quad (24)$$

$$\dot{X}^0 = a_{01} X^0 + \psi(t)\bar{V}_1, \quad Z^0 = a_{11} X^0 + V_2. \quad (25)$$

Здесь  $a_{ij} = a_{ij}(m^x, K^x, t)$  ( $i, j = 0, 1$ ) — коэффициенты статистической линеаризации функций  $a = a(X, t)$  и  $a_1 = a_1(X, t)$ . При этом ковариационная матрица  $K_x$  определяется уравнением вида (7):

$$\dot{K}^x = a_{01} K^x + K^x a_{01}^T + b\bar{\nu}_1 b^T. \quad (26)$$

Применяя к (24) и (25) уравнения фильтра Калмана–Бьюси [15, 16], получим

$$\dot{\hat{X}} = a_{00} - a_{01}m^x + a_{01}\hat{X} + Ra_{11}\bar{\nu}_2^{-1} \left( Z - a_{11}\hat{X} - a_{10} + a_{11}m^x \right); \quad (27)$$

$$\hat{X}_0 = MX(t_0);$$

$$\dot{R} = a_{01}R + Ra_{01}^T - Ra_{11}^T\bar{\nu}_2^{-1}a_{11}R + b\bar{\nu}_1b^T; \quad (28)$$

$$R_0 = M \left[ \left( X_0 - \hat{X}_0 \right) \left( X_0 - \hat{X}_0 \right)^T \right].$$

Совокупность фильтрационных уравнений (27) и (28) при условиях (24) и (26) определяет искомым квазилинейный нормальный фильтр для обработки информации о флуктуациях неравномерности вращения Земли по апостериорным данным, в том числе в темпе получения результатов наблюдения. Коэффициенты статистической линеаризации  $a_{00}$  и  $a_{01}$  приведены в разд. 2,  $a_{10} = 0$ ,  $a_{11} = I_2$  (в силу линейности второго уравнения (23)).

Фильтрационные уравнения (27) и (28) для произвольного линейного наблюдения, когда  $a_1(X, t) = b_1(t)X + b_0$ ,  $a_{11} = b_1(t)$ , принимают следующий вид:

$$\dot{\hat{X}} = a_{00} - a_{01}m^x + a_{01}\hat{X} + \beta \left[ Z - b_1(t)\hat{X} - b_0(t) + b_1(t)m^x \right]; \quad (29)$$

$$\hat{X}_0 = MX_0;$$

$$\dot{R} = a_{01}R + Ra_{01}^T - \beta b_1(t)R + b(t)\bar{\nu}_1(t)b(t)^T; \quad (30)$$

$$R_0 = M \left[ \left( X_0 - \hat{X}_0 \right) \left( X_0 - \hat{X}_0 \right)^T \right].$$

если через  $\beta$  обозначить коэффициент усиления фильтра  $\beta = Rb_1(t)^T\bar{\nu}_2^{-1}$ .

Коэффициенты статистической линеаризации  $a_{00}$ ,  $a_{01}$  и вспомогательная инструментальная матрица ошибки фильтрации  $R$  не содержат результатов наблюдений и могут быть определены отдельно (до получения результатов наблюдений). Таким образом, возможна априорная оценка точности квазилинейного фильтра.

**Замечание 7.** В случае, когда известны канонические представления возмущений в (1), используются соответствующие версии квазилинейного метода [15, 16].

**Замечание 8.** Для негауссовских возмущений в (1) можно воспользоваться методом эквивалентной линеаризации, взяв в качестве осредняющего распределения отрезок параметризованного распределения, например плотности, а затем использовать уравнения фильтра Калмана–Бьюси или Пугачева [15, 16].

## 7 Линейные фильтры для обработки информации о неравномерности скорости вращения Земли (тестовые примеры 4–6)

*Пример 4.* Пусть отсутствуют диссипативные силы ( $\mu_1 = 0$ ). Тогда уравнения (23), (27) и (28) можно представить в виде:

$$\begin{aligned} \dot{X}_2 &= \tilde{M}_{0t} + m_0^{SL} + V_1, \quad Z_2 = X_2 + V_2; \\ \dot{\hat{X}}_2 &= \tilde{M}_{0t} + m_0^{SL} + R\nu_2^{-1} \left( Z_2 - \hat{X}_2 \right), \quad \hat{X}_{2,t_0} = 0; \\ \dot{R} &= -R^2\nu_2^{-1} + \nu_1; \quad R_{t_0} = R_0. \end{aligned} \quad (31)$$

Решение (31) допускает аналитическую запись [16]:

$$R = \sqrt{\nu_1\nu_2} \frac{1 + \gamma \exp(-2\gamma_0 t)}{1 - \gamma \exp(-2\gamma_0 t)},$$

где

$$\gamma = \frac{R_0 - R_\infty}{R_0 + R_\infty}; \quad \gamma_0^2 = \nu_1\nu_2^{-1}, \quad R_\infty = \sqrt{\nu_1\nu_2}.$$

При  $t \rightarrow \infty$

$$R \rightarrow R_\infty \text{ и } \dot{\hat{X}}_2 = \tilde{M}_{0t} + m_0^{SL} + \sqrt{\frac{\nu_1}{\nu_2}} \left( Z_2 - \hat{X}_2 \right).$$

*Пример 5.* В случае чисто линейных диссипативных сил ( $\mu_1 \neq 0$ ,  $\tilde{\mu}_{1t} \equiv 0$ ,  $\mu_n = 0$ ,  $\mu_{nt} = 0$ ) уравнения (23), (29) и (30) можно представить в виде:

$$\begin{aligned} \dot{X}_2 &= -\mu_1 X_2 + \tilde{M}_{0t} + m_0^{SL} + V_1; \quad Z_2 = X_2 + V_2; \\ \dot{\hat{X}}_2 &= -\mu_1 \hat{X}_2 + \tilde{M}_{0t} + m_0^{SL} + \beta \left( Z_2 - \hat{X}_2 \right); \end{aligned} \quad (32)$$

$$\hat{X}_{2,t_0} = m_0;$$

$$\beta = R\nu_2^{-1}; \quad (33)$$

$$\dot{R} = -2\mu_1 R - R^2\nu_2^{-1} + \nu_1; \quad R_{t_0} = R_0. \quad (34)$$

Решение (34) имеет вид [16]:

$$R = R_\infty + \rho; \quad R_\infty = R_\infty^+; \quad (35)$$

$$R_\infty^\pm = \nu_2 \left( \sqrt{\mu_1^2 + \nu_1\nu_2^{-1}} \pm \mu_1 \right),$$

где

$$\rho = \frac{R_\infty^+ + R_\infty^-}{(R_0 + R_\infty^-) (R_0 + R_\infty^+)^{-1} e^{2\sqrt{\mu_1^2 + \nu_1\nu_2^{-1}}t} - 1},$$

причем  $R \rightarrow R_\infty$  при  $t \rightarrow \infty$ .

*Пример 6.* В условиях примера 5, если нормальные шумы  $V_1$  и  $V_2$  коррелированы ( $\nu_{12} \neq 0$ ), уравнения (32)–(35) имеют вид:

$$\begin{aligned}\dot{X}_2 &= -\mu_1 \hat{X}_2 + \tilde{M}_{0t} + m_0^{SL} + \\ &\quad + (R + \nu_{12}) \nu_2^{-1} (Z_2 - \hat{X}_2); \\ \dot{R} &= -2(\mu_1 + \nu_{12} \nu_2^{-1}) R + R^2 + \nu_2^{-1} + \\ &\quad + \nu_1 - \nu_{12}^2 \nu_2^{-1}; \\ R &= R_\infty + \rho;\end{aligned}$$

где

$$\begin{aligned}\rho &= \frac{(R_\infty^+ + R_\infty^-)(R_0 + R_\infty^+)}{(R_0 + R_\infty^-) \exp(\gamma t) - (R_0 + R_\infty^+)}; \\ \gamma^2 &= (\mu_1 + \nu_{12} \nu_2^{-1}) + \nu_1 \nu_2^{-1}; \\ R_\infty^\pm &= \nu_2 (\gamma \pm \mu_1 - \nu_{12} \nu_2^{-1}).\end{aligned}$$

## 8 Линейные фильтры для обработки информации о неравномерности угла поворота Земли (тестовые примеры 7 и 8)

*Пример 7.* При наличии чисто линейной диссипации ( $\mu_3 = 0$ ) уравнения (23), (27) и (28) имеют следующий вид:

$$\begin{aligned}\dot{X}_1 &= X_2; \quad \dot{X}_2 = -2\mu_1 X_1 + \tilde{M}_{0t} + m_0^{SL} + V_1; \\ Z_1 &= X_1 + V_2; \quad Z_2 = X_2 + V_3,\end{aligned}$$

$$\begin{aligned}\dot{\hat{X}}_1 &= \hat{X}_2 + R_{11} \nu_2^{-1} (Z_1 - \hat{X}_1) + \\ &\quad + R_{12} \nu_3^{-1} (Z_2 - \hat{X}_2); \\ \dot{\hat{X}}_2 &= -2\mu_1 \hat{X}_2 + \tilde{M}_{0t} + m_0^{SL} + \\ &\quad + R_{12} \nu_2^{-1} (Z_1 - \hat{X}_1) + R_{22} \nu_3^{-1} (Z_2 - \hat{X}_2),\end{aligned}$$

где

$$\begin{aligned}\dot{R}_{11} &= 2R_{12} + R_{11}^2 \nu_2^{-1} + R_{12}^2 \nu_3^{-1}, \\ \dot{R}_{12} &= R_{22} - 2\mu_1 R_{12} + R_{11} R_{12} \nu_2^{-1} + R_{12} R_{22} \nu_3^{-1}, \\ \dot{R}_{22} &= \nu_1 - 2\mu_1 R_{22} + R_{12}^2 \nu_2^{-1} + R_{22}^2 \nu_3^{-1}.\end{aligned}$$

*Пример 8.* В условиях примера 7, если измеряются только флуктуации угла  $\delta\varphi$ , следует пренебречь членами, содержащими  $\nu_3^{-1}$ , а в случае измерения только флуктуаций скорости  $\delta\dot{\varphi}$  — пренебречь членами, содержащими  $\nu_2^{-1}$ . В [16] содержится ряд других примеров для случая автокоррелированных шумов в уравнениях (23).

## 9 Квазилинейные фильтры для обработки информации о неравномерности вращения Земли в случае нелинейного рэлеевского механизма диссипации (тестовые примеры 9 и 10)

*Пример 9.* С учетом формул (16) и (17) в условиях примера 2, когда  $X_3 = 0$ ,  $X_4 = 0$  одновременно измеряются  $\delta\varphi$  и  $\delta\dot{\varphi}$ , в фильтрационных уравнениях (27) и (28) или (29) и (30) следует взять коэффициенты статистической линеаризации  $a_{00}$  и  $a_{01}$  в виде:

$$\begin{aligned}a_{00} &= a_{00}(m^x, K^x, t) = \\ &= \left[ -\mu_1 m_2 + \mu_3 (m_2^3 + 3m_2 K_{22}) \right]; \\ a_{01} &= a_{01}(m^x, K^x, t) = \frac{\partial a_{00}}{\partial m^x} = \\ &= \begin{bmatrix} 0 & 1 \\ 0 & -\mu_1 + 3\mu_3 m_2^2 (m_2^2 + K_{22}^2) \end{bmatrix}.\end{aligned}$$

*Пример 10.* В условиях примера 9, если  $X_3$  и  $X_4$  удовлетворяют уравнениям (2), то коэффициенты  $a_{00}$  и  $a_{01}$  определяются согласно (5), (6), (16) и (17). В [16] приведен ряд примеров, которые можно использовать в качестве тестирующих для автокоррелированных шумов в уравнениях состояния и наблюдения.

## 10 Заключение

Разработанные квазилинейные методы построения информационных моделей флуктуаций неравномерности вращения Земли по априорным и апостериорным данным реализованы в виде экспериментального программного обеспечения в среде MATLAB. Проведено тестирование программного обеспечения на примерах 1–10.

Квазилинейные методы, как показали вычислительные эксперименты и сравнение с результатами статистического моделирования, обеспечивают высокую точность фильтрации скорости  $\delta\dot{\varphi}$  (порядка 2%–3% для априорной стационарной информации и 0,5%–1% при апостериорной информации). Из-за отсутствия возвращающей силы по  $\delta\varphi$ , как видно

из уравнения (1), появляются дрейфы и накапливающиеся ошибки. Поэтому так необходимы точные измерения  $\delta\varphi$ .

Методы, алгоритмы, программное обеспечение и тестовые примеры включены в состав информационных ресурсов по фундаментальной проблеме РАН «Статистическая динамика вращения Земли».

Среди направлений дальнейших исследований следует выделить следующие:

- (1) учет влияния автокоррелированности различных возмущений в уравнениях (1);
- (2) оценка негауссовости распределений возмущений для оценок больших уклонений по  $\delta\varphi$ ;
- (3) оценивание потенциальной точности. Экстраполяция и интерполяция апостериорных данных измерений;
- (4) оценивание и распознавание возмущений в уравнениях (1) на основе апостериорной информации от нескольких нелинейных измерительных систем различной точности;
- (5) разработка комплексных статистических моделей вращения Земли, учитывающих флуктуации полюса и неравномерности вращения Земли, обобщающие результаты [1–10, 16, 17].

Автор благодарен Н. Н. Семендяеву за выполненные вычислительные эксперименты.

## Литература

1. Марков Ю. Г., Сеницын И. Н. Стохастическая модель движения полюса деформируемой Земли // ДАН, 2002. Т. 385. № 2. С. 189–192.
2. Марков Ю. Г., Сеницын И. Н. Флуктуационно-диссипативная модель движения полюса деформируемой Земли // ДАН, 2002. Т. 387. № 4. С. 482–486.
3. Марков Ю. Г., Сеницын И. Н. Нелинейные стохастические корреляционные модели движения полюса деформируемой Земли // Астрон. журн., 2003. Т. 80. № 2. С. 186–192.
4. Марков Ю. Г., Сеницын И. Н. Влияние параметрических флуктуационно-диссипативных сил на движение полюса Земли // ДАН, 2003. Т. 390. № 3. С. 343–346.
5. Марков Ю. Г., Сеницын И. Н. Многомерные распределения флуктуаций полюса Земли // ДАН, 2003. Т. 391. № 2. С. 194–198.
6. Марков Ю. Г., Сеницын И. Н. Спектрально-корреляционные модели флуктуаций вращательного движения Земли // ДАН, 2003. Т. 393. № 5. С. 618–623.
7. Марков Ю. Г., Сеницын И. Н. Спектрально-корреляционные и кинетические модели движения Земли // Астрон. журн., 2004. Т. 81. № 2. С. 184–192.
8. Марков Ю. Г., Сеницын И. Н. Стохастические корреляционные модели движения полюса деформируемой Земли // Космические исследования, 2004. Т. 42. № 1. С. 76–82.
9. Сеницын И. Н. Стохастические модели флуктуаций движения Земли в условиях пуассоновских возмущений // Системы и средства информатики. Спец. вып. Геоинформационные технологии. — М.: ИПИ РАН, 2004. С. 39–55.
10. Марков Ю. Г., Дасаев Р. Р., Перепелкин В. В., Сеницын И. Н., Сеницын В. И. Стохастические модели вращения Земли с учетом влияния Луны и планет // Космические исследования, 2005. Т. 43. № 1. С. 54–66.
11. Акуленко Л. Д., Марков Ю. Г., Перепелкин В. В. Неравномерности вращения Земли // ДАН, 2007. Т. 417. № 4. С. 483–488.
12. Марков Ю. Г., Сеницын И. Н. Корреляционная модель приливной неравномерности вращения Земли // ДАН (в печати).
13. Пугачев В. С., Сеницын И. Н. Теория стохастических систем. 2-е изд. — М.: Логос, 2004.
14. Одуан К., Гино Б. Измерение времени. Основы GPS. — М.: Техносфера, 2002.
15. Сеницын И. Н. Развитие теории фильтров Пугачева для оперативной обработки информации в стохастических системах // Информатика и её применения, 2007. Т. 1. Вып. 1. С. 3–13.
16. Сеницын И. Н. Фильтры Калмана и Пугачева. 2-е изд. — М.: Логос, 2007.
17. Сеницын И. Н. Корреляционные методы построения аналитических информационных моделей флуктуаций полюса Земли по априорным данным // Информатика и её применения, 2007. Т. 1. Вып. 2. С. 2–14.

## ПОРТАЛЫ В СИСТЕМАХ ОРГАНОВ ГОСУДАРСТВЕННОЙ ВЛАСТИ

А. В. Босов<sup>1</sup>

**Аннотация:** Технологии информационных порталов применяют многие организации, учреждения и сообщества для достижения самых разных целей. Наверное, наиболее полно порталы нашли свое применение в таких областях деятельности, как поддержка взаимодействия в коммерческой (e-Business) и в научной (e-Science) среде. Но и в других информационных сферах можно обнаружить множество задач, которые успешно решаются или могут решаться с помощью порталов. Предметом обсуждения данной работы является исследование состоявшихся результатов и перспектив применения технологий web-порталов в органах государственной власти (e-Government). Результативность и эффективность использования порталов в интересах органов государственной власти естественно всего просматривается с позиций задач, стоящих перед популярной и поддерживаемой в большинстве развитых стран концепции «электронного правительства».

**Ключевые слова:** «электронное правительство»; Интернет-технологии; порталы; стандарты Интернета

### 1 Портал — инструмент государственного управления

#### 1.1 Информационные технологии для государственных учреждений

На данный момент уже окончательно состоявшимся можно считать факт проникновения информационных технологий (ИТ) на все уровни государственного управления. В любом государственном учреждении от нижнего уровня конечных исполнителей до руководства любого уровня сегодня обязательно найдутся компьютеры и, что более важно, специфические для данного учреждения программные средства и информационные ресурсы, обеспечивающие функции назначения именно этого учреждения. К очевидным достоинствам применения ИТ в государственном учреждении можно отнести:

- сокращение числа разного рода ошибок;
- возможность получения доступа к различным инструментам автоматизации делопроизводительной, аналитической и коллективной деятельности, повышающим оперативность и качество принимаемых решений;
- улучшенное качество подготовки документов, в частности по причине большей информированности при принятии решений;
- укрепление исполнительской дисциплины, в частности в делопроизводстве организации;
- повышение уровня защищенности конфиденциальной информации граждан, которой располагает учреждение;

- быстрый обмен информацией как между сотрудниками учреждения, так и с другими учреждениями, в том числе возможность использования интегрированных межведомственных ресурсов;

- возможность более эффективного предоставления ориентированных на граждан услуг за счет использования инструментов автоматизации обработки клиентских запросов.

Кроме того, к не менее очевидным преимуществам можно отнести и повышение «коэффициента интеллекта» и авторитетности учреждения [1], сотрудниками которого активно используются компьютерные технологии, как в глазах общества, так и с точки зрения международной позиции государства в целом.

Ясно, что такая перспектива однозначно свидетельствует о стратегической перспективности информатизации любого госучреждения. Однако не все просто на пути достижения этой перспективы. Есть и не менее очевидные, чем достоинства, недостатки. Обобщая опыт многих проектов внедрения ИТ-решений в органах государственной власти, можно утверждать, что основные проблемы, которые приходится решать в рамках обсуждаемых проектов, носят организационный характер. К их числу относятся, например, такие:

- низкий уровень компьютерной грамотности служащих и, как следствие, необходимость осуществления в рамках проекта мероприятий по его повышению;

<sup>1</sup>Институт проблем информатики Российской академии наук, AVBosov@ipiran.ru

- реализация проектов осуществляется «без отрыва от производства»: учреждение должно выполнять возложенные на него функции вне зависимости от этапов выполнения ИТ-проекта; как следствие — возникновение сложной задачи комплексного планирования с участием всех уровней служащих учреждения, в том числе планирование развертывания и поддержки первых этапов эксплуатации с условием непрерывности работы учреждения;
- все выполняемые работы должны учитывать вопросы обеспечения безопасности, причем не только в целях обеспечения защиты от несанкционированного доступа извне, но и в целях блокирования возможностей некорректного использования ИТ-решения в личных целях служащими самой организации; к этому же кругу проблем можно отнести вопросы соблюдения авторских прав и прав собственности на интеллектуальный продукт.

Этот перечень позволяет сформулировать и оценить еще одну проблему, решение которой, к сожалению, целиком ложится на, как правило, малокомпетентное в данном вопросе руководство учреждения — выбор исполнителя проекта информатизации.

## 1.2 Концепция «электронного правительства»

Сравнительно недавно уровень развития ИТ не позволял сформировать сколько-нибудь обобщенные решения для широкого круга госучреждений. Да, компьютеры и отдельные ведомственные системы существуют давно. Взрывное развитие Интернета привело, в частности, к тому, что использование электронной почты стало необходимостью даже для самых небольших госучреждений. Однако даже постановки более существенных задач, например интеграция электронной почты с ведомственными базами данных и средствами управления контентом, представлялись нереальными задачами. Принципиально на изменение ситуации повлияло не столько развитие Интернета, сколько результаты деятельности Интернет-сообщества в области стандартизации форматов, протоколов взаимодействия и архитектур информационных систем для world-wide web. Результаты этой деятельности, о которых подробнее будет сказано далее, сделали возможным сформировать понимание того, какими методами можно достичь сформулированной стратегической перспективы информатизации органов государственного управления, выразившееся

в формировании концепции «электронного правительства».

Концепция «электронного правительства» состоит в определении необходимости прохождения госучреждением ряда этапов с конечной целью интегрировать весь комплекс услуг учреждения и обеспечить доступ к ним потребителей — граждан и других учреждений.

Этап первый предполагает использование Интернета для создания сайтов отдельных госучреждений с целью информирования граждан о работе учреждения, правилах и сроках подачи документов, получения справок и пр. Как и всем «статическим» web-ресурсам Интернета, таким сайтам свойственна *односторонняя связь*.

Следующий очевидный этап — наделение ведомственных сайтов интерактивными возможностями, превращающими их в инструмент *двустороннего общения* граждан и сотрудников учреждения. Этот этап начинается с появления на сайтах госучреждений электронных форм, которые предлагается заполнить пользователю с целью получения какой-либо услуги. Обратим внимание, что для реализации даже простейшего двустороннего обмена требуется, чтобы web-ресурс предоставлял ряд сервисов: аутентификации и авторизации пользователей, персонализации представлений, обработки пользовательских запросов на стороне сервера web-системы (портлеты). Строго говоря, наличие таких сервисов является основанием для признания такой системы web-порталом.

Большинство государств в настоящее время находится либо на первом, либо на втором этапе «электронного правительства». Следующий качественный уровень предполагает возможность использования web-ресурсов государства с целью *осуществления «финансовых» операций*, подразумевая под этим широкий спектр услуг, результат которых носит официальный статус. В качестве примеров можно было бы привести оплату всевозможных пошлин, сборов и штрафов, получение и/или продление различных лицензий, подачу заявлений для устройства на работу, подачу конкурсных материалов, получение субсидий и пр. Примеров в бизнесе структур, поддерживающих такой уровень информатизации, множество. Так, большинство банков, включенных в систему электронных платежей, имеют упомянутые web-ресурсы.

Последний, наиболее сложный этап состоит в превращении web-ресурса госучреждения в полноценный *web-портал*. При этом именно в данной концепции возможности порталных технологий востребованы в наиболее полном объеме. От портала органа госвласти потребуется:

- интегрированное представление всего комплекса услуг госучреждения;
  - возможность доступа к услугам граждан, исходя из их потребностей, а не из структуры подразделений учреждения;
  - непротиворечивое представление в рамках единого пользовательского интерфейса множества разнородных данных и услуг, поддерживаемых/поставляемых разнородными внутриведомственными системами автоматизации.
- необходимость консолидированного проявления политической воли со стороны высших государственных органов, в том числе для решения вопросов финансирования;
  - организационная сложность контроля качества выполнения проектов, в частности трудности определения критериев оценки их успешности;
  - необходимость вовлечения в проекты широкого круга граждан;
  - технические вопросы.

Интегрированный характер последнего этапа концепции «электронного правительства», который многие исследователи называют также этапом «информационной экономики» [2], конечно, не исчерпывается объединением внутренних ресурсов одного учреждения, что было бы не таким и сложным. Более значимым ожиданием на этом этапе является не только полная информатизация операций внутри учреждения, но и возможность выполнения любых межведомственных операций, требующихся гражданину, в электронной форме, т. е. интеграция не внутриведомственная, а уже межведомственная.

В одном из выступлений Билл Гейтс [2] заявил: «Пока ни в одной стране мира не удалось собрать все кусочки этой головоломки («электронного правительства») вместе. Тем не менее системы, развертываемые в настоящее время практически повсеместно, отлично работают, предоставляя гражданам дополнительные возможности. Они расширяют доступ к информации, что необходимо для совершенствования демократии, повышают эффективность государственного управления, избавляют от лишнего ожидания и обеспечивают информационную прозрачность». Пока ситуация в целом сохраняется: законченных решений такого уровня в мире нет, но на данный момент можно констатировать, что достижение этого этапа концепции уже не является фантастическим, и, по крайней мере, весь необходимый инструментарий и все нужные для этого технологии ИТ-сообществом разработаны и апробированы.

### 1.3 Проблемы реализации концепции

Как и любая инновация, реализация «электронного правительства» сталкивается с рядом проблем. К главным из них можно отнести:

- непонимание целей и задач концепции со стороны государственных чиновников, вплоть до сопротивления реализации конкретных проектов;

Ответить на вопрос, для чего нужно создавать «электронное правительство», не так просто. Для большинства чиновников решение этого вопроса представляется во внедрении компьютеров и ИТ в деятельность госучреждения, в частности с целью экономии затрат. Место же этой деятельности в общем контексте государственных реформ не рассматривается. На самом деле главный акцент следует делать на максимальную полноту информации в государственных структурах и максимально полное удовлетворение потребностей граждан ИТ-средствами. Иными словами, более важным в концепции является тезис о доступности информации и услуг, а уж за ним — электронные формы взаимодействия. В целом в мире сформировалось понимание того, что «электронное правительство» — это, прежде всего, применение ИТ как инструмента для изменения характера деятельности государственных служб [3], и признается, что ключ к успеху «электронного правительства» — ориентированный на граждан подход. Государство должно рассматривать граждан и частные компании как своих клиентов и предоставлять им услуги в соответствии с их запросами, а не на основе задач, стоящих перед самим государством. Одним из главных факторов, определяющих категорию, к которой может отнести ту или иную страну по показателю развития «электронного правительства», является степень использования государством данного подхода.

В этой связи высока вероятность противления ИТ-инновациям со стороны чиновников, поскольку такие преобразования, во-первых, потребуют новых методов работы от самих чиновников, во-вторых, сделают работу госучреждений более прозрачной. И то, и другое никогда не приветствовалось.

Для преодоления подобных препон в первую очередь следует прибегать к разного рода разъяснительным методам, но добиться полного успеха можно только за счет формирования консолидированной политической воли высших органов управления. Мировая практика показывает, что за большинством успешных проектов в этой области стоит

сильный государственный лидер, готовый связать с ним свою репутацию и тратить на него свое время, авторитет и ресурсы [2]. Любопытно отметить, что понимание этой ситуации давно сложилось и в России [4, 5], и в настоящее время действует серия государственных программ, поддерживающих концепцию «электронного правительства» на самом высоком уровне.

Но даже при наличии необходимой политической и финансовой поддержки реализация конкретных проектов испытывает значительные трудности. Прежде всего, речь идет о том, что очень сложно точно сформулировать критерии успешности проектов в этой области: непосредственных и быстрых экономических выгод нет, оценки степени прозрачности и числа вовлеченных пользователей весьма туманны. Таким образом, команде, реализующей проект на самых начальных стадиях, необходимо озаботиться и сформулировать по возможности более четкие критерии оценки реализации проекта. Здесь не последнее внимание необходимо уделять целевой аудитории, ради которой, собственно, и выполняется проект. Привлечь максимальное число заинтересованных граждан, информировать их о новых возможностях, организовать обратную связь — вот задачи, обязательные к постановке и решению в любом проекте из рассматриваемой области.

На ранних стадиях формирования концепции «электронного правительства» к важнейшим относились и технические вопросы, включая как организацию доступа к новым ресурсам, т. е. развитие и обеспечение доступности средств коммуникаций, так и возможности технологий, позволяющих создавать нужные решения. На данный момент можно констатировать, что острота этих вопросов в целом снята. Технологические проблемы, как и упоминалось выше, за последние годы практически решены — весь необходимый инструментариум имеется. И главное — имеются необходимые стандарты. Проблему доступа в Интернет тоже нельзя считать острой, по крайней мере, в нашей стране, чему немало способствуют и действующие национальные программы в области обучения.

Преодолев главные проблемы, госучреждение приступает к реализации концепции, выполняя первый, довольно простой шаг — выведение организации в Интернет. Решение этой задачи, т. е. создание собственного компетентного представительства в Интернете, действительно, довольно простая задача, но при ее решении с учетом перспективы постановки и последующих задач возникает еще одна проблема, хотя и не отнесенная к главным, но, несомненно, очень значительная. Это проблема выбора команды для реализации проек-

та информатизации. Грамотный подход к решению этой проблемы существенно влияет на общий успех проекта информатизации учреждения в целом, ведь начиная с малого — создания сайта — выбранной команде, реализующей концепцию «электронного правительства», предстоит практически сразу перейти к следующему этапу — налаживанию онлайн-новых отношений сотрудников учреждения и граждан. В целом проблема выбора команды обширна и выходит за рамки данной статьи. Перечислим только те задачи, которые должны решать разработчики, организуя электронную службу в Интернете для госучреждения:

- обеспечение компетентного представления учреждения в сети Интернет в кратчайшие сроки;
- обеспечение информирования граждан о наличии и возможностях Интернет-представительства учреждения;
- разработка и внедрение web-форм и системы обработки обращений граждан в среду созданного Интернет-представительства (скорее всего, с использованием порталных технологий), включив при этом в состав решения и средства персонализации;
- обеспечение высокого уровня конфиденциальности и безопасности при интеграции необходимых web-служб в состав созданного Интернет-представительства;
- интеграция сервисов, предоставляемых порталом учреждения, с имеющимися в учреждении средствами автоматизации;
- разработка предложений по совершенствованию бизнес-процессов, протекающих в учреждении, на основе новых сервисов, поддерживаемых созданным порталом, и предложений по снижению расходов на сопровождение ИТ и рисков, связанных с технической поддержкой, за счет использования современных архитектур и открытых стандартов;
- предвидение и прогнозирование будущих потребностей сотрудников госучреждения и граждан и учет их в текущих решениях.

Таким образом, проблема определения исполнителя, реализующего проект в области «электронного правительства», требует значительного внимания и с полным основанием может быть отнесена к главным проблемам реализации концепции.

## 2 Порталы и «электронное правительство»

### 2.1 Задачи, решаемые порталами

Основной задачей при реализации концепции «электронного правительства» является обеспечение открытости деятельности органов госвласти, т. е. возможность публичного доступа к информации государственного учреждения. В любой стране именно государство является крупнейшим поставщиком услуг населению и предприятиям. Государство также является и крупнейшим потребителем информации и информационных услуг предприятий и граждан.

На данный момент общепризнано, что оптимальной формой организации публичного доступа к информации являются web-порталы. В большинстве развитых стран, поддерживающих концепцию «электронного правительства», доступ в Интернет имеет подавляющее большинство граждан. В РФ доступность необходимых средств коммуникаций по разным причинам не столь масштабна, однако более 10% населения эту возможность имеет, причем это — наиболее активная и деятельная часть населения.

Первая задача портала органа государственного управления — обеспечивать возможность доступа к информации заинтересованных лиц. Эта задача решается путем публикации информации на сайте портала, а необходимые для этого инструменты предоставляются любым порталным решением. Вторая задача, также легко решаемая стандартными средствами портала, — это активное информирование пользователей портала (рассылки). Соответствующие инструменты обеспечивают возможность рассылки новостей, нормативных актов, сообщений и прочих видов информационного содержания, публикуемого на портале, пользователям, подписавшимся на рассылку.

Продвижение по этапам реализации концепции «электронного правительства» ставит следующую задачу — обеспечение возможности работы с обращениями граждан и организаций, т. е. речь идет об организации интерактивного взаимодействия. Портальные инструменты, применяемые для решения этой задачи, должны обеспечивать возможность обращения через сайт портала к сотрудникам госучреждения с целью решения своих вопросов. Кроме того, данный инструментарий должен давать возможность отслеживания хода выполнения запросов, формирования поручений и отчетов, т. е. реализовывать элементы делопроизводства.

Обеспечивая решение этой задачи, портал позволяет радикально изменить практику работы со-

трудников государственного учреждения, повысить эффективность и уменьшить затраты, в том числе и затраты времени граждан, которое в противном случае тратится на личное посещение учреждения.

Далее естественным образом возникает очередная задача — централизация и интеграция услуг, предоставляемых как порталом отдельного учреждения, так и множеством порталов различных ведомств. Практика реализации концепции «электронного правительства» показывает, что большинством стран-лидеров в этой области данная задача формулируется как задача создания единых центров доступа [6]. Такие центры доступа, будучи, очевидно, очередным шагом по реализации концепции, во-первых, также оформляются в виде порталов (т. е. инструментом для решения данной задачи продолжает оставаться web-портал), во-вторых, являются наиболее ощутимым результатом реализации концепции. Кроме того, центры доступа оказываются мощным инструментом интеграции государственных услуг в целом, существенно способствуя преодолению межведомственных барьеров, а также — толчком к ускорению развития сетевых услуг в стране. Из сказанного следует, что интеграция услуг государственных учреждений с помощью порталов означает непременно и реформирование деятельности государства, т. е. высказанный выше тезис о необходимости «политической воли» сильно проявляется именно на этой стадии реализации концепции «электронного правительства».

На текущий момент общемировую практику реализации решения задачи создания порталных центров доступа (так называемых правительственных порталов) характеризуют следующие факторы [6]:

- большинство правительственных порталов ограничивается переадресацией пользователя на другие web-узлы;
- отмечается общая тенденция движения в сторону предоставления информации и услуг в виде, не требующем от пользователя знания деталей устройства бюрократического аппарата;
- успешными оказываются порталные проекты, которые, во-первых, ориентированы на клиента, во-вторых, решают задачи интеграции информации, систем, процессов и услуг, предоставляемых государственной инфраструктурой в целом.

Решение задачи организации правительственного портала — центра доступа, или правительственного шлюза, — наиболее полно задействует возможности современных порталных технологий, так как создание «единой точки входа», обеспечивающей функции «единого окна» при доступе

к различным данным и услугам, и есть задача web-портала.

Создание правительственного центра доступа обеспечивает решение и ряда сопутствующих задач, определенных концепцией «электронного правительства», а именно:

- полное и объективное информирование граждан о всей совокупности услуг, предоставляемых государственным аппаратом в целом;
- вовлечение граждан и организаций, обеспечение социальной поддержки инициатив в области «электронного правительства»;
- обеспечение постоянства и оперативности взаимодействия государственных учреждений с гражданами и хозяйствующими субъектами, в том числе достижение прозрачности и повышение эффективности этого взаимодействия и оптимизация взаимодействия государственных структур.

## 2.2 О видах порталов

Термин «портал», первоначально появившийся в ИТ-области безотносительно к сети и технологиям Интернета, означает единые ворота или единую точку доступа ко множеству разнообразных данных и услуг. В принципе, для реализации портала вовсе нет необходимости использовать механизмы world-wide web. Однако именно лавинообразное развитие Интернета и связанных с глобальной сетью технологий стало реальным толчком к созданию действующих порталных решений. Самая очевидная задача, решаемая как первыми web-порталами, так и наиболее известными порталами, действующими в настоящее время, — это создание «единой точки входа» в Интернет для всех пользователей. Такие общедоступные порталы, называемые еще по признаку охвата пользовательской аудитории «горизонтальными», непременно имеют в своем составе глобальную поисковую машину и простейшие средства сбора, систематизации и представления в удобном для пользователя виде как можно большего числа полезных информационных ресурсов. Как правило, такие порталы — уникальные штучные решения, поддерживаемые большими ИТ-компаниями ([www.google.com](http://www.google.com), [www.msn.com](http://www.msn.com), [www.rambler.ru](http://www.rambler.ru), [www.yandex.ru](http://www.yandex.ru) и пр.).

В рамках концепции «электронного правительства» речь идет, конечно, не о таких порталах. Уместнее говорить о «вертикальных» порталах, ограниченных по сравнению с общедоступными порталами как целевой аудиторией, так и специальным набором поддерживаемых ресурсов и сервисов. Если общедоступные порталы, подчеркивая их

«горизонтальный» статус, называют еще Интернет-порталами, то для «вертикальных» порталов уместно использовать название интранет-порталов, подчеркивая тем самым, что у такого портала есть существенная часть, предназначенная не для общего доступа, а для поддержки деятельности ограниченного круга сотрудников, выполняющих обязанности в рамках некоторого предприятия.

Внутренняя, корпоративная часть интранет-портала обеспечивает создание персонализированного рабочего места сотрудника учреждения, позволяющего ему получать доступ ко всем необходимым внутренним и внешним ресурсам и сервисам в рамках своих полномочий. Внешняя часть интранет-портала, по сути аналогичная Интернет-порталу, позволяет широкому кругу внешних пользователей получать доступ (естественно, ограниченный) к той части корпоративной информации и услуг, что предназначена для этого.

На текущий момент практически все крупные производители программного обеспечения имеют «коробочные» продукты, позволяющие учреждению реализовывать собственное порталное решение. Конечно, порталы от IBM, Oracle, Microsoft, Hummingbird, Plumtree, Lotus, Vignette и пр. отличаются по своим функциональным возможностям, так что выдача каких-либо рекомендаций по выбору порталной платформы без детального анализа функций назначения и информационно-сервисного обеспечения портала органа госвласти не представляется возможной. Отметим в связи с многообразием возможных решений следует только одно: все они не являются законченными web-порталами предприятия, которые достаточно развернуть и сразу начать использовать, а представляют собой только инструментарий, который надо изучить и грамотно применить в конкретном проекте.

## 3 Подходы к разработке портала «электронного правительства»

### 3.1 Состав порталного решения и требования к компонентам

Определение детальных требований к portalу органа госвласти должно основываться на определениях: состава целевой аудитории; состава информации, размещаемой на страницах сайта портала; состава информационных услуг, предоставляемых службами портала; возможностей и уровня подготовки сотрудников, обеспечивающих актуализацию и поддержку порталного контента (ресурсов и

служб). Однако можно выделить обязательный минимальный набор компонентов портала, необходимый для любого решения такого уровня. В состав таких компонентов непременно должны входить:

- сайт портала или компонент представления (обеспечивает формирование конечного пользовательского web-интерфейса);
- собственное хранилище портала (обеспечивает размещение собственного информационного содержания портала);
- система управления содержанием или блок управления контентом (обеспечивает решение задачи формирования информационного содержания портала как в собственном хранилище, так и посредством выполнения запросов в ресурсах портала, например корпоративных хранилищах);
- компонент поиска (обеспечивает формирование и исполнение разнообразных поисковых запросов пользователей в собственном хранилище портала и во «внешних» ресурсах);
- система безопасности и компонент персонализации (решают задачи аутентификации и авторизации пользователей (возможно также и аудит) и формирования собственных персональных настроек).

В дополнение к данному базовому набору, как правило, присутствуют и еще несколько стандартных компонентов, таких как компоненты обмена сообщениями (форумы, чаты, электронная почта), компоненты коллективной работы (делопроизводство), а для порталов, готовых к работе с большим числом разнообразных «внешних» источников, — компонент интеграции, обеспечивающий единое представление разнородных ресурсов в унифицированном порталном интерфейсе. Наличие интеграционного сервиса непременно означает и наличие в портале ряда адаптеров — компонентов, обеспечивающих преобразование запросов и форматов данных между ядром портала и «внешними» ресурсами.

Общим для всех порталных компонентов требованием является возможность управления ими, т.е. использования всех поддерживаемых ими функций, с помощью «тонкого клиента» — посредством программы web-обозревателя (браузера). Из других важных требований остановимся на тех, что предъявляются к компоненту представления и системе управления содержанием.

Для интегрированной архитектуры портала «электронного правительства», предполагающей представление на его сайте большого многообразия ресурсов и услуг, непременным условием вос-

требованности является обеспечение максимального удобства представления информации, исходя исключительно из интересов пользователей. Компонент представления должен, во-первых, обеспечивать гибкие возможности по размещению информационного содержания на сайте, как минимум позволяя группировать информацию по тематическим областям. И во-вторых, компонент представления должен поддерживать многообразие индивидуальных параметров, позволяющих авторизованному пользователю настраивать интерфейсные элементы «под себя», в частности иметь возможность формировать состав представляемых ресурсов и услуг согласно собственным потребностям и вкусам.

Служба управления содержанием (Content Management System, CMS) является, видимо, самой значимой и сложной частью портала, хотя конечному пользователю, как правило, она не видна вовсе. Требования, предъявляемые к порталной CMS:

- возможность использования шаблонов представления, обеспечивающих разделение данных и кода, формирующего конечное оформление информации;
- возможность обновления информационного содержания, состава информационных услуг портала и состава «внешних» ресурсов в реальном масштабе времени;
- возможность динамической сборки содержания web-страниц из разных ресурсов, подключенных к portalу.

Еще несколько слов следует сказать по поводу дополнительных сервисов, получающих распространение в настоящее время в связи с развитием коммуникационных средств. Именно широкому кругу потенциальных пользователей государственных порталов сегодня становятся доступны средства коммуникаций, позволяющие активно использовать видеоизображения и телевидение (WebTV), подключаться к portalу с использованием мобильного телефона или карманного PDA-компьютера. Таким образом, граждане имеют возможность сами определять, как, когда, каким образом и какие услуги они хотят получать, причем эти потребности уже не исчерпываются возможностями домашнего компьютера и Dialup-соединения. Задача правительственного портала — обеспечивать удовлетворение этих потребностей, т.е. поддерживать богатый мультимедийный контент, средства Интернет-вещания и специальные версии сайта, предназначенные для просмотра с мобильных устройств.

Не столь очевидными, как перечисленные, но не менее важными являются и требования, которые следует предъявлять к архитектуре порталного решения. На текущем этапе реализации концепции «электронного правительства» в разных странах выполняются серии проектов по созданию порталов для отдельных органов федеральной, региональной и местной власти. В некоторых странах выполняются, в других — предполагается, что будут выполняться, проекты по созданию центральных правительственных порталов, которые должны объединить часть услуг порталов отдельных ведомств. Таким образом, формируется существенное требование к реализации инфраструктурной составляющей портала — она должна позволять разработчикам использовать механизмы интеграции услуг различных провайдеров с учетом существующей иерархии ответственности. Как удовлетворить этому требованию, обсуждается дальше, а здесь отметим, что прежде всего это требование влияет на центральный правительственный портал, который должен рассматриваться не только как информационная витрина правительства, но и как инфраструктурный компонент, отвечающий за обеспечение единых интерфейсов для разработчиков государственных систем.

### 3.2 Технологии и стандарты Интернета

Базой для реализации любого порталного решения, естественно, являются web-сайты — сама основа Интернета, т.е. распределенная IP-среда, в которой присутствуют web-серверы, поддерживающие стандартный протокол HTTP, и web-клиенты — программы-обозреватели (web-браузеры), обеспечивающие представление полученных по протоколу HTTP страниц, размеченных на языке HTML (сейчас актуальнее говорить о разметке XHTML). Однако считать, что именно возможности этих технологий являются необходимым минимумом для создания портала, было бы серьезной ошибкой.

Первая проблема, с которой сталкивались разработчики порталов, — отсутствие общепринятого стандарта для обмена информацией между разными системами, а ведь в связи с решением именно этой задачи — обеспечить «единую точку входа» к данным разных источников — и появилась терминология порталов. Стереть ограничения и барьеры непонимания между множеством обязанных взаимодействовать систем, научить их «понимать» данные друг друга мог только общепринятый язык, оформленный как открытый стандарт, достаточно простой и в то же время выразительный, чтобы выступать в качестве базового стандарта для ин-

теграции любых данных. Таким языком «лингва франка» для Интернета стал XML (eXtensible Markup Language, расширяемый язык разметки) [7], утвержденный World-Wide Web Consortium (W3C) и предоставляющий механизм для обеспечения информационной совместимости, признаваемый на данный момент всеми без исключения членами ИТ-сообщества.

Следующая проблема — обеспечение функциональной совместимости. Попытки ее решения осуществлялись давно и безуспешно (можно сослаться на такие технологии, как Corba и DCOM), но единого общепринятого стандарта не образовывалось вплоть до утверждения W3C стандарта протокола доступа к простым объектам SOAP (Simple Object Access Protocol) [8]. Протокол SOAP описывает способ использования языка XML и протокола HTTP для создания механизмов доставки информации и вызова удаленных процедур. Протокол SOAP является основой технологии web-сервисов, сделавшей реальным применение в промышленных масштабах архитектур SOA (Service-Oriented Architecture, ориентированная на сервисы архитектура). Идея SOA [9] создавать системы, формируя их из наборов отдельных сервисов, предоставляющих возможности доступа к простым услугам посредством стандартных интерфейсов, поначалу находила применение только в системах распределенных вычислений, а с недавнего времени — в сетях GRID. Но настоящую путевку в жизнь этой архитектуре дали именно web-сервисы [10], основой которых является протокол SOAP.

Другие важные компоненты технологии web-сервисов оформлены в стандартах UDDI и WSDL. Спецификации UDDI (universal description, discovery and integration specification, спецификация универсального описания, обнаружения и интеграции web-сервисов) [11] задают способ публикации и поиска информации о web-сервисах путем описания конкретных функций, предоставляемых этим сервисом. Описание функций, в свою очередь, оформляется на языке WSDL (web-service definition language, язык описания web-сервисов) [12].

Эта тройка, собственно, и сделала возможным создание реальных информационных порталов, обеспечивающих возможности доступа посредством унифицированного web-интерфейса к множеству разнородных данных и услуг, формируемых в федеративной среде портала множеством «внешних» ресурсов. В результате применения этих технологий в порталных решениях получается открытая система, позволяющая другим информационным системам находить и использовать конкретную услугу посредством стандартного механизма web-сервисов.

Наконец, в связи с web-сервисами обязательно следует отметить, а разработчикам — использовать, еще ряд сопутствующих стандартов, разработанных с целью дальнейшего повышения интероперабельности систем, основанных на web. Эти стандарты подготовлены сообществом Web Services Interoperability Organization (WS-I, [www.ws-i.org](http://www.ws-i.org)) и касаются вопросов:

- поддержки транзакционности (WS-Transaction) [13];
- безопасности передачи сообщений (WS-Security) [14];
- гарантированной доставки сообщений (WS-Reliable-Messaging) [15].

## 4 Некоторые решения

На данный момент трудно сколько-нибудь полно описать все проекты, выполненные в разных странах в рамках реализации концепции «электронного правительства». Здесь перечислены только некоторые решения, интересные либо с исторической точки зрения (первые решения), либо с точки зрения информационного и сервисного содержания.

По-видимому, первыми Интернет-ресурсами, получившими общегосударственный статус, можно считать MAXI (<http://www.maxi.com.au/>), созданный в австралийском штате Виктория, и eCitizen Centre (<http://www.ecitizen.gov.sg/>), созданный в Сингапуре. Австралийцы, видимо, наиболее активно продвигают идеи «электронного правительства» в коммерческую среду (см., например, австралийский государственный информационный бизнес-портал ВЕР (business entry point) <http://www.business.gov.au/>).

Значительное число успешных проектов реализовано в США. Вот только несколько примеров: [www.cio.gov](http://www.cio.gov) — Совет директоров по информационным технологиям США; [www.gao.gov](http://www.gao.gov) — ресурс Административно-бюджетного управления; [www.nascio.org](http://www.nascio.org) — Национальная ассоциация ИТ-директоров.

Представляется, что наиболее комплексно и системно к реализации проектов «электронного правительства» подходят в Великобритании. Координирующей группой UK GovTalk выполняется проект с широким числом участников, цель которого — разработка согласованных стандартов (схем и структур данных), объединяющих многочисленные разрозненные системы государственного сектора на основе XML. На сайте <http://www.govtalk.gov.uk/> содержатся проекты схем данных и согласованные схемы, соответствующие требованиям и ра-

бочим процедурам государственных органов Великобритании, рекомендации по их практическому использованию и инструментальные средства. Практическая реализация этой деятельности представлена, например, правительственным шлюзом <http://www.gateway.gov.uk/>, где в режиме онлайн могут быть получены услуги государственных департаментов сельского хозяйства и занятости (DARDNI и DEFRA).

В Европе рядом аналогичных проектов по разработке XML-схем, ориентированных на обслуживание граждан, руководит Европейская комиссия (программа eEurope, стандарты опубликованы на <http://www.e-eurostandards.org>). В качестве действующего и востребованного примера можно привести портал <http://www.eu-careers.com/>, предоставляющий услуги при поиске и подготовке к трудоустройству в европейских странах.

Имеются и успешно действующие межгосударственные проекты. Любопытный пример в этой области оформился в результате деятельности правоохранительных сообществ США: открытый ресурс <http://it.ojp.gov>, созданный по инициативе американского Департамента юстиции активно используется и, главное, пополняется специалистами правоохранительной сферы всего мира.

Аналогичных ресурсов в Интернете на данный момент огромное множество. К странам-лидерам в области реализации концепции следует отнести Австралию, Германию, Великобританию, Испанию, США, Швецию, Тайвань, Сингапур и др. [16].

Позиции Российской Федерации в области «электронного правительства» хотя и нельзя называть лидирующими, но определенные небезуспешные результаты у нас есть. Отметим, прежде всего, что основная деятельность в России в этой сфере достаточно централизованно координируется в рамках государственной программы «Электронная Россия». Так, на данный момент можно считать реализованным этап программы, в рамках которого предполагалось создание группы сайтов, публикующих информацию органов власти — федеральных министерств и ведомств. В типовой набор информации, представленной на таком сайте, входят:

- основные ведомственные нормативные документы;
- справочная информация (расположение, расписание работы, персоналии);
- актуальные материалы информационно-новостного характера;
- проекты ведомственных документов;
- рекомендации физическим или юридическим лицам по решению их типичных проблем.

О завершении этого первого этапа объявлялось еще в 2000 г. [4, 5], хотя и на данный момент далеко не все даже федеральные органы имеют свои представительства в Интернете. Наиболее удачные и информационно полные ресурсы: официальный сайт президента РФ [www.kremlin.ru](http://www.kremlin.ru), сайт правительства РФ [www.government.ru](http://www.government.ru), сайты Совета Федерации [www.council.gov.ru](http://www.council.gov.ru), Государственной думы [www.duma.gov.ru](http://www.duma.gov.ru), Совета безопасности [www.scrf.gov.ru](http://www.scrf.gov.ru), Счетной палаты [www.ach.gov.ru](http://www.ach.gov.ru) и др. Важным результатом «Электронной России» является собственный сайт программы — [www.e-rus.ru](http://www.e-rus.ru), выполняющий начальные функции шлюза «электронного правительства». На данный момент на этом сайте можно найти ссылки на сайты практически всех федеральных министерств и ведомств.

Можно также отметить, что имеются и успешные примеры реализации второго этапа концепции «электронного правительства», а именно предоставление населению некоторых услуг. Два хорошо известных россиянам примера этого представляют сайт Федеральной налоговой службы ([www.nalog.ru](http://www.nalog.ru)) и официальный сайт президента РФ. Так, использование Интернета налоговой службой существенно облегчило для граждан прохождение процедуры подачи налоговых деклараций, а услуги президентского сайта позволяют участвовать в ежегодном интерактивном общении с Президентом РФ. В настоящее время осуществляются и другие попытки создания обратной связи с гражданами вплоть до создания правительственного портала [17]. При этом постепенно просматривается тенденция перехода на более высокий уровень сервиса в сравнении с информационно-справочным. К характерным примерам, иллюстрирующим уровень интерактивности Интернет-ресурсов государственных органов в РФ, следует отнести: сайт [www.government.ru](http://www.government.ru) (позволяет пользователю обратиться в правительство РФ), сайт [www.ach.gov.ru](http://www.ach.gov.ru) (обращения в Счетную палату РФ), сайт [www.mvdinform.ru](http://www.mvdinform.ru) (обращения в Министерство внутренних дел). При этом предоставляемый сервис ограничивается возможностью отправки сообщения, т. е. о полноценной обратной связи говорить пока рано.

Важнейшим примером тенденции смены этапа «электронного правительства» в РФ является сайт государственных закупок [www.zakupki.gov.ru](http://www.zakupki.gov.ru), заработавший в полном объеме в прошлом году в рамках исполнения Федерального закона от 21 июля 2005 г. № 94-ФЗ «О размещении заказов на поставки товаров, выполнение работ, оказание услуг для государственных и муниципальных нужд». Этот сайт реализует пользовательский интерфейс уже полнофункционального порталного решения, предоставляя весь необходимый спектр услуг, обеспе-

чивающий выполнение Закона о государственных закупках. Его услугами пользуются все имеющие бюджетное финансирование организации, а само его наличие является существенным инструментом обеспечения открытости информации об использовании бюджетных средств.

Примерами успешной реализации ведомственных порталов являются Информационный веб-портал Российской академии наук [www.ras.ru](http://www.ras.ru), ресурсы Российского фонда фундаментальных исследований [www.rfbr.ru](http://www.rfbr.ru), портал «Почта России» [www.russianpost.ru](http://www.russianpost.ru).

В целом, однако, говоря о всей совокупности Интернет-проектов органов государственной власти в РФ, приходится констатировать отставание от общемировых тенденций: перехода от «информационного» этапа концепции «электронного правительства» к этапу «услуг» в государственном масштабе пока не происходит. Любопытно отметить, что «коммерческий» сектор российского Интернета уже ни в чем не уступает ведущим мировым лидерам. В этой связи надо сказать, что рассмотренные выше проблемы реализации концепции в РФ все-таки решаются: имеется необходимый уровень понимания со стороны государственных чиновников, достаточно высок уровень финансирования, созданы центры государственной компетенции. Проблемными остаются два направления:

- из-за все еще ограниченных возможностей доступа широкого круга граждан к средствам коммуникаций эффективность реализуемых государством Интернет-проектов остается невысокой;
- затруднения возникают в технической составляющей — сложным остается формирование квалифицированной команды технических специалистов высокого уровня, нехватка которых имеется в стране в целом.

## Заключение

Успешная реализация поддерживаемой всем мировым сообществом концепции «электронного правительства» является необходимым условием создания информационного общества или, как часто говорят, информационной экономики. Текущее состояние этой реализации в целом соответствует этапности, определенной для концепции ИТ-обществом в конце прошлого века, подтверждают жизнеспособность «электронного правительства», соответствие получаемых результатов высказываемым ранее ожиданиям, хотя и с некоторым экстенсивным оттенком. Сформулированные выше

задачи «электронного правительства», проблемы реализации проектов, требования к уровням и способам достижения поставленных целей и технологический базис позволяют эффективно решать практически все задачи, определенные концепцией. Движущей силой для технической составляющей реализации концепции являются стандарты и технологии world-wide web, сформировавшиеся к настоящему времени в законченном виде, и использующие их порталные технологии. Именно порталы являются ключевым звеном большинства проектов, реализуемых в рамках «электронного правительства», и от успешности дальнейшего проникновения порталных решений в проекты «электронного правительства» в значительной степени зависит успех всего направления в целом.

## Литература

1. Государство в XXI веке. Электронный документооборот и делопроизводство // Информационный бюллетень Microsoft. Специальный выпуск, 2003.
2. Государство в XXI веке. Материалы конференции для государственных деятелей // Информационный бюллетень Microsoft. Выпуск 18. Декабрь 2002 г.
3. Realizing the vision: 2002 Global accenture study on eGovernment // Бюллетень Microsoft Insight — Government. 2002. [http://www.microsoft.com/europe/insight/Government/Analyst\\_insights/item117.htm](http://www.microsoft.com/europe/insight/Government/Analyst_insights/item117.htm).
4. Правительство—население: диалог в виртуальном пространстве // Всероссийская объединенная конференция «Интернет и современное общество», ноябрь 2000. <http://ims2000.nw.ru/src/ТЕХТ50.HTMЛ>.
5. Кабинет Касьянова обещает облегчить к себе доступ: Интервью с руководителем департамента правительственной информации правительства РФ А. Коротковым // «Новая газета», август, 2000. <http://2000.novayagazeta.ru/nomer/2000/36n/n36n-s21.shtml>.
6. Государство в XXI веке. Интеграция государственных информационных систем и организация межведомственного взаимодействия // Информационный бюллетень Microsoft. Выпуск 21. Сентябрь, 2003.
7. Extensible Markup Language (XML) 1.0 (4th ed.). <http://www.w3.org/TR/2006/REC-xml-20060816>.
8. SOAP Version 1.2 Part 1: Messaging Framework (2nd ed.). <http://www.w3.org/TR/soap12-part1>.
9. Web Services Glossary. <http://www.w3.org/TR/2004/NOTE-ws-gloss-20040211>.
10. Web Services Architecture. <http://www.w3.org/TR/ws-arch>.
11. Universal Description, Discovery and Integration specification. [www.uddi.org](http://www.uddi.org).
12. Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language. <http://www.w3.org/TR/wsdl20>.
13. Web Services Transactions specifications. <http://www-128.ibm.com/developerworks/library/specification/ws-tx>.
14. Web Services Security. <http://www-128.ibm.com/developerworks/library/specification/ws-secure>.
15. Web Services Reliable Messaging. <http://www-128.ibm.com/developerworks/library/specification/ws-rm>.
16. Государство в XXI веке. Реализация проектов электронного правительства // Информационный бюллетень Microsoft. Выпуск 19. Март, 2003.
17. Шелян И. Ворота во власть // Computerworld, 2003. No. 1. Изд-во «Открытые Системы». [http://www.osp.ru/cw/2003/01/025\\_2.htm](http://www.osp.ru/cw/2003/01/025_2.htm).

# СИСТЕМА $Geo/G/1/\infty$ С ОДНОЙ «НЕСТАНДАРТНОЙ» ДИСЦИПЛИНОЙ ОБСЛУЖИВАНИЯ\*

А. В. Печинкин<sup>1</sup>, С. Я. Шоргин<sup>2</sup>

**Аннотация:** Рассматривается система массового обслуживания (СМО)  $Geo/G/1/\infty$  с дисциплиной обслуживания, при которой в момент поступления в систему новой заявки ее длина сравнивается с (остаточной) длиной заявки на приборе и та из них, длина которой меньше, занимает прибор, а другая становится первой в очереди, сдвигая остальную очередь на единицу. Для этой системы найдены основные стационарные характеристики функционирования. В частности, показано, что, в отличие от непрерывного времени, в дискретном времени стационарное распределение числа заявок в системе не является инвариантным относительно загрузки.

**Ключевые слова:** система массового обслуживания; дискретное время; «нестандартная» дисциплина обслуживания

## 1 Введение

Хорошо известно, что применение различных дисциплин обслуживания может существенно улучшить пользовательские характеристики СМО.

Для СМО с ожиданием в классе консервативных дисциплин (дисциплина называется консервативной, если процесс обслуживания не зависит от процесса поступления заявок и в любой момент времени, когда в системе имеются заявки, суммарная скорость обслуживания равна единице) абсолютным чемпионом в смысле минимальной длины очереди является дисциплина преимущественного обслуживания заявки минимальной остаточной длины, или SRPT (Shortest Remaining Processor Time) [1–4].

Отметим, что математические соотношения для расчета показателей функционирования (в частности, стационарного распределения числа заявок) системы  $M/G/1/\infty$  с дисциплиной SRPT [4–9] весьма сложны, демонстрируют, скорее всего, возможности современных математических конструкций и вряд ли могут быть использованы для практических расчетов. Поэтому важно было найти границы изменения показателей функционирования СМО  $M/G/1/\infty$  с дисциплиной SRPT при фиксированных значениях параметров обслуживания, например загрузке системы [10, 11].

Интересная гипотеза была высказана А. Д. Соловьевым: при заданной загрузке в стационарном режиме функционирования число заявок в СМО  $M/G/1/\infty$  с дисциплиной SRPT максимально в

смысле упорядочения « $\prec$ » функций распределения (т. е. такого упорядочения, при котором  $\xi \prec \eta$ , если  $F_\xi(x) \geq F_\eta(x)$  для всех  $x$ ) при постоянной длине заявок. Гипотезу Соловьева удалось доказать в [12] с помощью введения специальной дисциплины обслуживания, при которой так же, как и при дисциплине SRPT, поступающая новая заявка ставится на прибор, если ее длина меньше остаточной длины обслуживаемой заявки, вытесняя последнюю на первое место в очереди. Однако если ее длина больше остаточной длины обслуживаемой заявки, то новая заявка сама становится на первое место в очереди (вне зависимости от того, есть ли в очереди заявки более короткой длины). Иными словами, место на приборе и первое место в очереди разыгрывают между собой не все заявки, а только новая заявка и заявка на приборе. Доказательство гипотезы Соловьева основывалось на том, что при введенной дисциплине стационарное распределение числа заявок в СМО  $M/G/1/\infty$  зависит только от загрузки системы.

В дальнейшем появилось много работ, обобщающих данную дисциплину в разных направлениях [13–28]. Так, в [13] было показано, что постоянная длина заявок является не только худшей при дисциплине SRPT, но и лучшей при обычной дисциплине обслуживания заявок в порядке поступления (FCFI — First-Come-First-In).

Обратимся теперь к нескольким иным задачам современной теории массового обслуживания. Применение в инфотелекоммуникационных сетях со-

\* Работа выполнена при поддержке РФФИ, гранты 06-07-89056 и 08-07-00152.

<sup>1</sup> Институт проблем информатики Российской академии наук, apetchinkin@ipiran.ru

<sup>2</sup> Институт проблем информатики Российской академии наук, sshorgin@ipiran.ru

временных технологий привело к новому всплеску исследований СМО, функционирующих в дискретном времени (см., например, [29–35]). Заметим, что при кажущемся упрощении методов исследования СМО в дискретном времени по сравнению со СМО в непрерывном в некоторых случаях результаты анализа значительно усложняются. Это связано, в первую очередь, с тем, что, в отличие от непрерывного времени, в дискретном времени может одновременно происходить несколько изменений состояний (в частности, окончание обслуживания заявки на приборе и поступление в систему новой заявки).

В настоящей работе для системы  $Geo/G/1/\infty$  с введенной в [12] дисциплиной, являющейся дискретным аналогом СМО  $M/G/1/\infty$  с той же самой дисциплиной, найдены основные стационарные характеристики функционирования. В частности, показано, что в дискретном времени стационарное распределение числа заявок в системе зависит не только от загрузки, т. е. не является инвариантным относительно загрузки.

## 2 Описание системы

Рассмотрим СМО в дискретном времени  $Geo/G/1/\infty$ , в которую поступает геометрический поток заявок с вероятностью  $a$  поступления заявки на одном такте (далее будем называть тактом как интервал времени между соседними изменениями состояния системы, так и сами моменты, в которые происходят эти изменения). Распределение времени обслуживания заявки является произвольным дискретным с вероятностью  $b_i$ ,  $i \geq 0$ , того, что обслуживание заявки продлится  $i$  тактов (предполагается, что  $b_0 = 0$ ).

Далее будем использовать следующие обозначения:

$\bar{a} = 1 - a$  — вероятность непоступления заявки на такте;

$B_i = \sum_{j=i}^{\infty} b_j$ ,  $i \geq 0$ , — вероятность того, что обслуживание заявки продлится не менее  $i$  тактов;

$\bar{B}_i = \sum_{j=0}^{i-1} b_j = 1 - B_i$ ,  $i \geq 1$ , — вероятность того, что обслуживание заявки продлится менее  $i$  тактов;

$\bar{b} = \sum_{i=0}^{\infty} i b_i = \sum_{i=1}^{\infty} B_i$  — среднее время обслуживания заявки.

Будем предполагать, что загрузка системы  $\rho = a\bar{b}$  меньше единицы. Это условие является необходимым и достаточным для существования стационарного режима функционирования системы.

Дисциплина обслуживания заключается в следующем. Предполагается, что в любой момент времени известна остаточная длина (далее будем

говорить просто длина) каждой заявки в системе, т. е. число тактов, необходимое для окончания обслуживания данной заявки. В момент поступления в систему новой заявки ее длина сравнивается с (остаточной) длиной заявки на приборе, и та из них, длина которой меньше, занимает прибор, а другая заявка становится на первое место в очереди, сдвигая остальную очередь на единицу. Для определенности будем считать, что при равенстве длин обеих заявок заявка на приборе продолжит обслуживаться. Если же в некоторый момент одновременно оканчивается обслуживание одной заявки и в систему поступает другая, то независимо от длин всех находящихся в очереди заявок на прибор становится вновь поступившая заявка. Заявки с прерванным обслуживанием дообслуживаются.

## 3 Стационарное распределение очереди

Найдем стационарное распределение числа заявок в системе.

Введем обозначения:

$p_0$  — стационарная вероятность того, что непосредственно после очередного такта система будет пуста;

$p_n(i)$ ,  $i \geq 1$ , — стационарная вероятность того, что непосредственно после очередного такта в системе будет  $n$  заявок и до окончания обслуживания заявки на приборе останется  $i$  тактов.

Положим

$$\bar{P}_n(j) = \sum_{i=j+1}^{\infty} p_n(i), \quad n \geq 1, \quad j \geq 0;$$

$$P_n(j) = \sum_{i=1}^j p_n(i) = \bar{P}_n(0) - \bar{P}_n(j), \quad n \geq 1, \quad j \geq 1;$$

$$p_n = \sum_{i=1}^{\infty} p_n(i) = \bar{P}_n(0) \quad n \geq 1.$$

Выпишем систему уравнений равновесия (СУР), которой удовлетворяют введенные функции. Для этого удобно поступить следующим образом [12]. Введем новую СМО  $Geo/G/1/n$  с конечным числом  $n$  мест ожидания, отличающуюся от исходной только тем, что если в очереди находится  $n$  заявок и поступает новая заявка, то после сравнения длин этой заявки и заявки на приборе в системе (на приборе) остается заявка большей длины, а вторая заявка покидает систему. Используя прием, введенный в [12] и подробно изложенный в [4], нетрудно показать, что стационарные вероятности состояний в исходной и новой СМО отличаются

лишь на постоянный множитель. Это дает возможность записать следующую СУР:

$$p_0 = p_0\bar{a} + p_1(1)\bar{a}; \quad (1)$$

$$p_1(i) = p_0ab_i + p_1(i+1)\bar{a} + P_1(i+1)ab_i + p_1(i+1)a\bar{B}_i, \quad i \geq 1; \quad (2)$$

$$p_n(i) = p_{n-1}(i+1)aB_i + \bar{P}_{n-1}(i+1)ab_i + p_n(i+1)\bar{a} + P_n(i+1)ab_i + p_n(i+1)a\bar{B}_i, \quad n \geq 2, \quad i \geq 1, \quad (3)$$

к которой необходимо добавить условие нормировки

$$p_0 + \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} p_n(i) = p_0 + \sum_{n=1}^{\infty} \bar{P}_n(0) = p_0 + \sum_{n=1}^{\infty} p_n = 1. \quad (4)$$

Для того чтобы решить СУР (1)–(4), просуммируем сначала уравнения (2) и (3) по  $i$  от  $j$  до  $\infty$ . Тогда

$$\bar{P}_1(j-1) = [p_0 + p_1]aB_j + \bar{P}_1(j)(1 - aB_j), \quad j \geq 1; \quad (5)$$

$$\bar{P}_n(j-1) = [p_n + \bar{P}_{n-1}(j)]aB_j + \bar{P}_n(j)(1 - aB_j), \quad n \geq 2, \quad j \geq 1. \quad (6)$$

Полагая  $j = 1$  в соотношении (6), имеем

$$p_n(1)\bar{a} = [p_{n-1} - p_{n-1}(1)]a, \quad n \geq 2. \quad (7)$$

В терминах производящей функции (ПФ)

$$\bar{P}(z, j) = \sum_{n=1}^{\infty} z^n \bar{P}_n(j), \quad j \geq 0,$$

из (5) и (6) получаем уравнение

$$\bar{P}(z, j-1) = [zp_0 + P(z)]aB_j + \bar{P}(z, j)[1 - a(1-z)B_j], \quad j \geq 1, \quad (8)$$

где

$$P(z) = \sum_{n=1}^{\infty} z^n p_n = \bar{P}(z, 0). \quad (9)$$

Решение уравнения (8) имеет вид

$$\bar{P}(z, j) = \frac{zp_0 + P(z)}{1-z} \times \left( 1 - \prod_{i=j+1}^{\infty} (1 - a(1-z)B_i) \right), \quad j \geq 0, \quad (10)$$

где  $P(z)$  определяется из соотношений (10) при  $j = 0$  и (9), которые приводят к формуле

$$P(z) = zp_0 \left( 1 - \prod_{i=1}^{\infty} (1 - a(1-z)B_i) \right) \times \left( \prod_{i=1}^{\infty} (1 - a(1-z)B_i) - z \right)^{-1}. \quad (11)$$

Наконец, вероятность  $p_0$  получается с помощью правила Лопиталья из условия нормировки (4) и имеет обычный вид

$$p_0 = 1 - \rho.$$

Дифференцируя формулу (11) в точке  $z = 1$  соответствующее число раз, можно найти моменты любого порядка стационарного распределения числа заявок в системе. Так, математическое ожидание определяется выражением

$$\bar{m} = P'(1) = \rho + \frac{1}{1-\rho} \Sigma^{(2)}, \quad (12)$$

где

$$\Sigma^{(2)} = \sum_{\substack{i,k=1 \\ k>i}}^{\infty} a^2 B_i B_k. \quad (13)$$

Последняя формула показывает, в частности, что, в отличие от непрерывного времени, в дискретном времени для СМО  $Geo/G/1/\infty$  с рассматриваемой дисциплиной обслуживания отсутствует свойство инвариантности стационарного распределения очереди от распределения длины заявки при фиксированной загрузке  $\rho$ .

Решения уравнений (5) и (6) можно записать без использования ПФ. Например,

$$\bar{P}_1(j) = P_1(1 - \Pi_j), \quad j \geq 0; \quad (14)$$

$$\bar{P}_2(j) = P_2(1 - \Pi_j) - P_1 \Pi_j \tilde{\sigma}_j^{(1)}, \quad j \geq 0; \quad (15)$$

$$\bar{P}_3(j) = P_3(1 - \Pi_j) - P_2 \Pi_j \tilde{\sigma}_j^{(1)} - P_1 \Pi_j \tilde{\sigma}_j^{(2)}, \quad j \geq 0, \quad (16)$$

где

$$P_n = \sum_{i=0}^n p_i, \quad n \geq 0;$$

$$\Pi_j = \prod_{i=j+1}^{\infty} (1 - aB_i), \quad j \geq 0;$$

$$\tilde{\sigma}_j^{(1)} = \sum_{i=j+1}^{\infty} \frac{aB_i}{1 - aB_i}, \quad j \geq 0;$$

$$\tilde{\sigma}_j^{(2)} = \sum_{\substack{i,k=j+1 \\ k>i}}^{\infty} \frac{a^2 B_i B_k}{(1 - aB_i)(1 - aB_k)}, \quad j \geq 0;$$

$$\begin{aligned} p_1 &= \bar{P}_1(0) = p_0(1 - \Pi_0)\Pi_0^{-1}; \\ p_2 &= \bar{P}_2(0) = P_1 \left(1 - \Pi_0 + \Pi_0\tilde{\sigma}_j^{(1)}\right) \Pi_0^{-1}; \\ p_3 &= \bar{P}_3(0) = \left[ P_2 \left(1 - \Pi_0 + \Pi_0\tilde{\sigma}_j^{(1)}\right) + \right. \\ &\quad \left. + P_1\Pi_0\tilde{\sigma}_j^{(2)} \right] \Pi_0^{-1}. \end{aligned}$$

Впрочем, формулы (14)–(16) можно получить и из ПФ (10) и (11), дифференцируя их в точке  $z = 0$ .

Отметим, что для численных расчетов удобно воспользоваться исходными соотношениями (1)–(3), которые вкпе с равенством (7) приводят к следующей рекуррентной по  $n$  и по  $j$  процедуре:

$$\begin{aligned} p_1(1) &= \frac{a}{\bar{a}} p_0, \\ p_n(i+1) &= \frac{1}{1 - aB_{i+1}} \{p_n(i) - [p_0 + P_1(i)] ab_i\}, \quad i \geq 1, \\ p_n(1) &= \frac{a}{\bar{a}} [p_{n-1} - p_{n-1}(1)], \quad n \geq 2, \\ p_n(i+1) &= \frac{1}{1 - aB_{i+1}} \{p_n(i) - p_{n-1}(i+1)aB_i - [\bar{P}_{n-1}(i+1) + P_n(i)] ab_i\}, \quad n \geq 2, \quad i \geq 1. \end{aligned}$$

Программная реализация приведенной процедуры не вызывает никаких затруднений.

#### 4 Стационарное распределение времени пребывания заявки в системе

Определим теперь стационарное распределение времени пребывания заявки в рассматриваемой системе.

Заметим сначала, что период занятости системы (ПЗ), открываемый заявкой длины  $i$ , имеет ПФ

$$\gamma_i(z) = (z[a\gamma(z) + \bar{a}])^i, \quad i \geq 1, \quad (17)$$

где  $\gamma(z)$  — ПФ ПЗ, открываемого произвольной заявкой, удовлетворяющая уравнению

$$\gamma(z) = \beta(z[a\gamma(z) + \bar{a}]). \quad (18)$$

Здесь  $\beta(z) = \sum_{i=1}^{\infty} z^i b_i$  — ПФ времени обслуживания заявки.

Стационарное распределение времени ожидания начала обслуживания заявки длины  $i$  имеет ПФ

$$\begin{aligned} \psi_i(z) &= p_0 + \sum_{n=1}^{\infty} \left( p_n(1) + \sum_{j=2}^{i+1} p_n(j)\gamma_{j-1}(z) + \bar{P}_n(i+1) \right), \quad i \geq 1. \quad (19) \end{aligned}$$

Распределение полного времени обслуживания заявки длины  $i$  (т. е. времени от момента первого поступления заявки на прибор и до момента ухода ее из системы) имеет ПФ

$$\begin{aligned} \delta_1(z) &= z; \\ \delta_i(z) &= z \left( a \left[ \sum_{j=1}^{i-2} \gamma_j(z) b_j + B_{i-1} \right] + \bar{a} \right) \delta_{i-1}(z), \quad i \geq 2. \quad (20) \end{aligned}$$

Наконец, стационарное распределение времени пребывания в системе заявки длины  $i$  имеет ПФ

$$\varphi_i(z) = \psi_i(z)\delta_i(z). \quad (21)$$

Соответствующие безусловные характеристики определяются усреднением по длине заявки. В частности, стационарное распределение времени пребывания в системе произвольной заявки имеет ПФ

$$\varphi(z) = \sum_{i=1}^{\infty} \varphi_i(z) b_i. \quad (22)$$

Дифференцируя формулы (17)–(22), можно получить выражения для моментов соответствующих стационарных характеристик. В частности, средняя длина ПЗ и средняя длина ПЗ, открываемого заявкой длины  $i$ , определяются формулами

$$\bar{g} = \frac{\bar{b}}{1 - \rho}, \quad \bar{g}_i = \frac{i}{1 - \rho};$$

математическое ожидание стационарного распределения времени ожидания начала обслуживания заявки длины  $i$  и математическое ожидание полного времени обслуживания заявки длины  $i$  — формулами

$$\bar{v}_i = \frac{1}{1 - \rho} \sum_{n=1}^{\infty} \sum_{j=2}^{i+1} (j-1) p_n(j), \quad i \geq 1;$$

$$\bar{d}_1 = 1;$$

$$\bar{d}_i = 1 + \frac{a}{1 - \rho} \sum_{j=1}^{i-2} j b_j + \bar{d}_{i-1}, \quad i \geq 2;$$

математические ожидания стационарных распределений времен пребывания в системе заявки длины  $i$  и произвольной заявки — формулами

$$\bar{w}_i = \bar{v}_i + \bar{d}_i, \quad i \geq 1;$$

$$\bar{w} = \sum_{i=1}^{\infty} \bar{w}_i b_i.$$

## 5 Неравенства для числа заявок в системе при некоторых дисциплинах обслуживания

В этом разделе будут приведены границы, в которых изменяется среднее число заявок в рассматриваемой системе при фиксированной загрузке, а также неравенства для длины очереди в системе  $Geo/G/1/\infty$  при некоторых дисциплинах обслуживания.

Можно показать, что для любых функций  $B_i = \sum_{j=i}^{\infty} b_j$ ,  $i \geq 0$ , определяемых распределением  $\{b_i, i \geq 0\}$  положительной целочисленной случайной величины, справедливы неравенства

$$1 < \sum_{i=1}^{\infty} B_i^2 \leq \bar{b} - \alpha(1 - \alpha),$$

где через  $\alpha$  обозначена дробная часть  $\bar{b}$ . Тогда из (12) и (13) имеем

$$\begin{aligned} \rho + \frac{a^2}{2(1-\rho)} \left[ \bar{b}^2 - \bar{b} + \alpha(1-\alpha) \right] &\leq \\ &\leq \bar{m} < \rho + \frac{a^2}{2(1-\rho)} \left[ \bar{b}^2 - 1 \right]. \end{aligned} \quad (23)$$

Левая часть неравенства (23) превращается в равенство при времени обслуживания заявки, принимающем всего 2 значения:  $[\beta]$  с вероятностью  $(1 - \alpha)$  (здесь  $[\beta]$  — целая часть числа  $\bar{b}$ ) и  $[\beta] + 1 - \alpha$  с дополнительной вероятностью  $\alpha$ .

Правая часть неравенства (23) превращается в равенство в пределе на последовательности «расплывающихся» распределений времен обслуживания заявок, например на последовательности распределений длин заявок, принимающих такие 2 значения: 1 с вероятностью  $(1 - 1/n)$  и  $n$  с вероятностью  $(b - 1 + 1/n)/n$ .

Если увеличивать среднюю длину  $\bar{b}$  заявки и одновременно уменьшать вероятность  $a$  поступления заявки таким образом, что загрузка  $\rho = a\bar{b}$  остается постоянной, то в пределе нижняя и верхняя границы изменения среднего числа заявок в системе совпадут и будут равны  $\rho + \rho^2 / [2(1 - \rho)]$ . Это соответствует тому факту, что в пределе получается система  $M/G/1/\infty$  с рассматриваемой дисциплиной обслуживания, стационарные вероятности числа заявок

в которой при фиксированной загрузке инвариантны относительно распределения длины заявки, о чем уже говорилось выше.

Введем бесконечномерный процесс  $\nu(t) = (\nu_1(t), \nu_2(t), \dots)$ , где  $\nu_n(t)$  — суммарная (остаточная) длина всех заявок в системе после  $t$ -го такта, за исключением заявок, находящихся на последних  $(n - 1)$  местах в очереди (заявка на приборе считается находящейся на первом месте в очереди). В частности,  $\nu_1(t)$  — суммарная длина всех заявок после  $t$ -го такта. Если в системе в момент  $t$  меньше  $n$  заявок, то  $\nu_n(t) = 0$ . Естественно, при заданных параметрах системы значение процесса  $\nu(t)$  зависит от дисциплины обслуживания.

Сравним значения процессов  $\nu(t)$  при следующих дисциплинах:

- (1) при рассматриваемой дисциплине;
- (2) при инверсионном порядке обслуживания без прерывания обслуживания (с точки зрения распределения числа заявок в системе эта дисциплина совпадает со стандартной дисциплиной обслуживания заявок в порядке поступления);
- (3) при дисциплине SRPT — наилучшей в смысле минимизации числа заявок в системе.

Значение процесса  $\nu(t)$  при  $i$ -й дисциплине снабдим верхним индексом ( $i$ ).

Справедливы неравенства:

$$\nu^{(3)}(t) \leq \nu^{(1)}(t) \leq \nu^{(2)}(t). \quad (24)$$

Неравенства (24) доказываются математической индукцией по времени  $t$  на основе покоординатного сравнения процессов.

Поскольку число  $\eta(t)$  заявок в системе в момент  $t$  определяется выражением

$$\eta(t) = \max\{n : \nu_n(t) > 0\},$$

то из неравенств (24) получаем неравенства

$$F_{\eta^{(3)}(t)}(x) \prec F_{\eta^{(1)}(t)}(x) \prec F_{\eta^{(2)}(t)}(x),$$

связывающие распределения чисел заявок в системах с дисциплинами 1–3 в смысле упорядочения « $\prec$ ». В частности, такое упорядочение справедливо и для стационарных распределений числа заявок в этих же системах.

## 6 Вариант дисциплины

Обратимся теперь к варианту исследуемой дисциплины, отличающемуся от рассмотренного ранее только лишь правилом постановки заявки на прибор в том случае, когда одновременно оканчивается обслуживание одной заявки и в систему поступает

новая заявка. В этой ситуации сравниваются длины новой заявки и первой заявки из очереди, кратчайшая из них становится на прибор, а оставшаяся занимает первое место в очереди. Здесь также для определенности будем считать, что при равенстве длин заявок на прибор становится заявка из очереди.

Введем обозначения:

$p_0$  — стационарная вероятность того, что непосредственно после окончания очередного такта система будет пуста;

$p_n(i, j), i, j \geq 1, n \geq 2$ , — стационарная вероятность того, что непосредственно после окончания очередного такта в системе будет  $n$  заявок и до окончания обслуживания заявки на приборе осталось  $j$  тактов, а первой заявки из очереди —  $i$  тактов;

$p_n(j), j \geq 1, n \geq 1$ , — стационарная вероятность того, что непосредственно после окончания очередного такта в системе будет  $n$  заявок и до окончания обслуживания заявки на приборе осталось  $j$  тактов, в частности  $p_n(j) = \sum_{i=1}^{\infty} p_n(i, j), j \geq 1, n \geq 2$ .

Положим

$$P_n(i, j) = \sum_{k=1}^j p_n(i, k), \quad n \geq 2, j \geq 1;$$

$$\bar{P}_n(j) = \sum_{i=j+1}^{\infty} p_n(i), \quad n \geq 1, j \geq 0;$$

$$P_n(j) = \sum_{i=1}^j p_n(i) = \bar{P}_n(0) - \bar{P}_n(j), \quad n \geq 1, j \geq 1;$$

$$p_n = \sum_{i=1}^{\infty} p_n(i) = \bar{P}_n(0), \quad n \geq 1.$$

Введенные функции удовлетворяют СУР

$$p_0 = p_0\bar{a} + p_1(1)\bar{a}; \quad (25)$$

$$p_1(j) = p_0ab_j + p_1(j+1)\bar{a} + P_1(j+1)ab_j + p_1(j+1)a\bar{B}_j; \quad j \geq 1, \quad (26)$$

$$p_n(i, j) = p_n(i, j+1)\bar{a} + [P_n(i, j+1) - p_n(i, 1)]ab_j + p_n(i, j+1)a\bar{B}_j; \quad n \geq 2, j > i \geq 1; \quad (27)$$

$$p_n(j, j) = p_{n-1}(j+1)ab_j + p_n(j, j+1)\bar{a} + [P_n(j, j+1) - p_n(j, 1)]ab_j + p_n(j, j+1)a\bar{B}_j + p_n(j, 1)ab_j; \quad n \geq 2, j \geq 1; \quad (28)$$

$$p_n(i, j) = p_{n-1}(i+1)ab_j + p_{n-1}(j+1)ab_i + p_n(i, j+1)\bar{a} + [P_n(i, j+1) -$$

$$- p_n(i, 1)]ab_j + p_n(i, j+1)a\bar{B}_j + p_n(i, 1)ab_j + p_n(j, 1)ab_i, \quad n \geq 2, i > j \geq 1, \quad (29)$$

с условием нормировки

$$p_0 + \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} p_n(i) = p_0 + \sum_{n=1}^{\infty} \bar{P}_n(0) = p_0 + \sum_{n=1}^{\infty} p_n = 1. \quad (30)$$

Произведя элементарные преобразования соотношений (25)–(30), получим следующие формулы:

$$p_0 = 1 - \rho; \quad (31)$$

$$p_1(1) = \frac{a}{a} p_0; \quad (32)$$

$$p_1(j+1) = \frac{1}{1 - aB_{j+1}} \{p_1(j) - [p_0 + P_1(j)]ab_j\}, \quad j \geq 1, \quad (33)$$

$$p_n(1, 1) = \frac{ab_1}{1 - ab_1} p_{n-1}(2), \quad n \geq 2; \quad (34)$$

$$p_n(i, 1) = \frac{1}{1 - a\bar{B}_{i+1}} \{[p_n(1, 1) + \dots + p_n(i-1, 1)]ab_i + p_{n-1}(i+1)a\bar{B}_{i+1} + [P_{n-1}(i) - p_{n-1}(1)]ab_i\}, \quad n \geq 2, i \geq 2; \quad (35)$$

$$p_n(1) = \sum_{i=1}^{\infty} p_n(i, 1), \quad n \geq 2; \quad (36)$$

$$p_n(j+1) = \frac{1}{1 - aB_{j+1}} \{p_n(j) - P_n(j)ab_j - [p_{n-1} - p_{n-1}(1)]ab_j - p_{n-1}(j+1)a + p_n(j, 1)\bar{a}\}, \quad n \geq 2, j \geq 1. \quad (37)$$

Формулы (31)–(37) позволяют рекуррентно вычислять  $p_0, p_n(i, 1), n \geq 2, i \geq 2$ , и  $p_n(j), n \geq 1, j \geq 1$ . При этом сначала по формулам (31)–(33) находят  $p_0$  и  $p_1(j), j \geq 1$ , а затем по формулам (34)–(37) последовательно по  $n$ , начиная с  $n = 2$ , определяются  $p_n(i, 1), i \geq 1, p_n(1), p_n(j+1), j \geq 1$ .

Формулы для вероятностей  $p_n(i, j), n \geq 2, i \geq 1, j \geq 2$ , здесь не приводятся, поскольку эти вероятности не используются при вычислении основных стационарных показателей функционирования системы.

Однако следует отметить, что, несмотря на простоту приведенных выше формул, не удалось, в отличие от первоначального варианта, получить аналитические выражения даже для стационарного среднего числа заявок в системе.

Вычисление стационарных распределений, связанных с временем пребывания заявки в системе,

мало отличается от вычисления их при дисциплине, введенной в начале статьи. Приведем здесь только те формулы, в которых имеются отличия от формул разд. 4.

Формула для ПФ стационарного распределения времени ожидания начала обслуживания заявки длины  $i$  (аналог формулы (19)) принимает вид

$$\psi_i(z) = p_0 + \sum_{n=1}^{\infty} \left( \sum_{k=1}^i p_n(k, 1) \gamma_k(z) + \sum_{k=i+1}^{\infty} p_n(k, 1) + \sum_{j=2}^{i+1} p_n(j) \gamma_{j-1}(z) + \bar{P}_n(i+1) \right), \quad i \geq 1.$$

Для того чтобы записать аналог формулы (20), введем «укороченный» ПЗ, открываемый заявкой длины  $i$ , под которым будем понимать интервал времени от момента поступления в свободную систему заявки длины  $i$  и до того первого момента (такта), перед которым в системе окажется единственная заявка, обслуживаемая на последнем такте, т. е. до того такта, когда система впервые освободится от заявок, если не считать той заявки, которая, возможно, поступит на этом такте. Производящая функция «укороченного» ПЗ, открываемого заявкой длины  $i$ , имеет вид

$$\tilde{\gamma}_i(z) = z^i (a\gamma(z) + \bar{a})^{i-1}, \quad i \geq 1.$$

Тогда ПФ времени, на которое прервется обслуживание заявки, имеющей в начале такта длину  $i$ , определяется выражением

$$\hat{\gamma}_i(z) = a \left( \sum_{j=1}^{i-1} \tilde{\gamma}_j(z) b_j \hat{\gamma}_i(z) + B_i \right) + \bar{a}, \quad i \geq 1,$$

откуда находим:

$$\hat{\gamma}_i(z) = (aB_i + \bar{a}) \left( 1 - a \sum_{j=1}^{i-1} \tilde{\gamma}_j(z) b_j \right)^{-1}, \quad i \geq 1.$$

Наконец, распределение полного времени обслуживания заявки длины  $i$  имеет ПФ

$$\begin{aligned} \delta_1(z) &= z; \\ \delta_i(z) &= z \hat{\gamma}_{i-1}(z) \delta_{i-1}(z), \quad i \geq 2. \end{aligned}$$

Обозначая через  $\nu^{(4)}(t)$  введенный в предыдущем разделе процесс при рассматриваемой в этом разделе дисциплине, нетрудно получить неравенства

$$\nu^{(3)}(t) \leq \nu^{(4)}(t) \leq \nu^{(1)}(t) \leq \nu^{(2)}(t).$$

## 7 Заключение

Итак, в настоящей работе рассмотрены два варианта функционирующей в дискретном времени СМО  $Geo/G/1/\infty$  со следующей дисциплиной обслуживания. Предполагается, что в любой момент времени известна остаточная длина каждой заявки в системе, т. е. число тактов, необходимое для окончания обслуживания данной заявки. В момент поступления в систему новой заявки ее длина сравнивается с (остаточной) длиной заявки на приборе, и та из них, длина которой меньше, занимает прибор, а другая заявка становится первой в очереди, сдвигая остальную очередь на единицу.

Для этой системы найдены основные стационарные характеристики функционирования: распределение числа заявок в системе и распределения времени ожидания начала обслуживания и времени пребывания заявки в системе. В частности, показано, что, в отличие от непрерывного времени, в дискретном времени стационарное распределение числа заявок в системе зависит не только от загрузки, т. е. не является инвариантным относительно загрузки.

Приведены границы, в которых изменяется среднее число заявок в рассматриваемой системе при фиксированной загрузке, а также неравенства для длины очереди в системе  $Geo/G/1/\infty$  при некоторых дисциплинах обслуживания.

## Литература

1. *Schrage L., Miller L.* The queue  $M/G/1$  with the shortest remaining processing time discipline // *Oper. Res.*, 1966. Vol. 14. P. 670–684.
2. *Schrage L.* A proof of the optimality of the shortest remaining processing time discipline // *Oper. Res.*, 1968. Vol. 16. P. 687–690.
3. *Козлов В. В.* Оптимальная дисциплина обслуживания для систем массового обслуживания // *Научные труды Кубанского университета*, 1977. Т. 247. С. 33–37.
4. *Бочаров П. П., Печинкин А. В.* Теория массового обслуживания. — М.: Изд-во РУДН, 1995.
5. *Печинкин А. В., Соловьев А. Д., Яшков С. Ф.* О системе с дисциплиной обслуживания первым требованием с минимальной оставшейся длиной // *Изв. АН СССР. Технич. кибернет.*, 1979. № 5. С. 51–58.
6. *Schassberger R.* The steady-state appearance of the  $M/G/1$  queue under the discipline of shortest remaining processing time // *Adv. Appl. Probab.*, 1990. Vol. 22. P. 456–479.
7. *Grishechkin S. A.* On a relationship between processor-sharing queues and Cramp–Mode–Jagers branching processes // *Adv. Appl. Probab.*, 1992. Vol. 24. P. 653–698.

8. Печинкин А. В. Нестационарные характеристики системы обслуживания с дисциплиной SRPT // Вестник РУДН. Сер. Прикл. матем. и информ., 1996. № 1. С. 77–85.
9. Печинкин А. В. Система  $MAP/G/1/\infty$  с дисциплиной SRPT // Теория вероятностей и ее прим., 2000. Т. 45. Вып. 3. С. 589–595.
10. Печинкин А. В. О верхней и нижней оценках средней очереди в системе с дисциплиной Шраге // Техника средств связи. Сер. Системы связи, 1980. Вып. 3. С. 24–28.
11. Печинкин А. В. Границы изменения стационарной очереди в системах обслуживания с различными дисциплинами обслуживания // Проблемы устойчивости стохастических моделей. Труды семинара. — М.: ВНИИСИ, 1985. С. 118–121.
12. Печинкин А. В. Об одной инвариантной системе массового обслуживания // Math. Operationsforsch. und Statist. Ser. Optimization, 1983. Vol. 14. No. 3. P. 433–444.
13. Нагоненко В. А. О характеристиках одной нестандартной системы массового обслуживания. I // Изв. АН СССР. Технич. кибернет., 1981. № 1. С. 187–195.
14. Нагоненко В. А. О характеристиках одной нестандартной системы массового обслуживания. II // Изв. АН СССР. Технич. кибернет., 1981. № 3. С. 91–99.
15. Нагоненко В. А., Печинкин А. В. О большой нагрузке в системе с инверсионным обслуживанием и вероятностным приоритетом // Изв. АН СССР. Технич. кибернет., 1982. № 1. С. 86–94.
16. Нагоненко В. А., Печинкин А. В. О малой нагрузке в системе с инверсионным порядком обслуживания и вероятностным приоритетом // Изв. АН СССР. Технич. кибернет., 1984. № 6.
17. Печинкин А. В. Инверсионный порядок обслуживания с вероятностным приоритетом в системе обслуживания с неординарным входящим потоком / Случайные процессы и их приложения. Математические исследования. — Кишинев: Штиинца, 1989. Вып. 109.
18. Печинкин А. В. Нестационарные характеристики СМО с инверсионным порядком обслуживания и вероятностным приоритетом / Вероятностные задачи дискретной математики: Межвузовский сборник. — М.: МИЭМ, 1990.
19. Таташев А. Г. Многоканальная система массового обслуживания с потерей кратчайших требований // Автоматика и телемеханика, 1991. № 7. С. 187–189.
20. Таташев А. Г. Одна система массового обслуживания с инвариантной дисциплиной // Автоматика и телемеханика, 1992. № 7. С. 92–96.
21. Таташев А. Г. Одна система массового обслуживания с инвариантными вероятностями состояний // Кибернетика и системный анализ, 1993. № 5. С. 183–186.
22. Таташев А. Г. Система массового обслуживания с групповым поступлением и инверсионной дисциплиной // Кибернетика и системный анализ, 1995. № 6. С. 163–165.
23. Таташев А. Г. Одна инверсионная дисциплина обслуживания в системе с групповым поступлением // Автоматика и вычисл. техника, 1995. № 1. С. 53–59.
24. Таташев А. Г. Одноканальная система массового обслуживания с потерей заявки наибольшей длины // Кибернетика и системный анализ, 1997. № 3. С. 187–188.
25. Таташев А. Г. Одна инверсионная дисциплина обслуживания в одноканальной системе с разнотипными заявками // Автоматика и телемеханика, 1999. № 7. С. 177–181.
26. Таташев А. Г. Одноканальная система с инверсионной дисциплиной обслуживания и разнотипными заявками // Кибернетика и системный анализ, 2000. № 3. С. 170–174.
27. Печинкин А. В., Свищева Т. А. Система  $MAP/G/1/\infty$  с инверсионным порядком обслуживания и вероятностным приоритетом // Вестник Российского университета дружбы народов. Сер. Прикладная математика и информатика, 2003. № 1. С. 109–118.
28. Pechinkin A., Svishcheva T. The stationary state probability in the  $BMAP/G/1/r$  queueing system with inverse discipline and probabilistic priority // Transactions of XXIV International Seminar on Stability Problems for Stochastic Models. Jurmala, Latvia. September 10–17, 2004. P. 141–147.
29. Chaudhry M. Invited talk: Queue-length and waiting-time distributions of discrete-time  $GI(X)/Geom/1$  queueing systems with early and late arrivals // Queueing systems: Theory and applications, 1997. Vol. 25. No. 1–4. P. 307–324.
30. Desert B., Daduna H. Discrete time tandem networks of queues effects of different regulation schemes for simultaneous events // Performance Evaluation, 2002. Vol. 47. No. 2. P. 73–104.
31. He J., Sohraby Kh. An extended combinatorial analysis framework for discrete-time queueing systems with general sources // IEEE/ACM Transactions on Networking (TON), 2003. Vol. 11. No. 1. P. 95–110.
32. Chaudhry M., Gupta U. C., Goswami V. On discrete-time multiserver queues with finite buffer:  $GI/Geom/m/N$  // Computers and Operations Research, 2004. Vol. 31. No. 13. P. 2137–2150.
33. Fiems D., Steyaert B., Bruneel H. Discrete-time queues with generally distributed service times and renewal-type server interruptions // Performance Evaluation, 2004. Vol. 55. No. 3–4. P. 277–298.
34. Atencia I., Moreno P. A discrete-time  $Geo/G/1$  retrial queue with general retrial times // Queueing Systems, 2004. No. 48. P. 5–21.
35. Akar N. A matrix analytical method for the discrete time Lindley equation using the generalized Schur decomposition // ACM International Conference Proceeding Series, 2006. Vol. 201. No. 12.

# РАЗМЕЩЕНИЯ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ, МИНИМИЗИРУЮЩИЕ СРЕДНЮЮ ДЛИНУ ОЧЕРЕДИ

Т. В. Захарова<sup>1</sup>

**Аннотация:** В статье рассмотрен класс систем массового обслуживания с вызовами, возникающими на плоскости. Решается задача оптимизации расположения станций обслуживания по критерию средней суммарной длины очереди. Проводится сравнение оптимальных расположений по критерию средней суммарной длины очереди и критерию среднего суммарного времени ожидания.

**Ключевые слова:** система массового обслуживания; вызовы на плоскости; длина очереди; время ожидания; оптимизация расположения

## 1 Введение

В статье рассмотрен класс систем массового обслуживания с вызовами, возникающими на плоскости. Спецификой изучаемого класса систем является необходимость использования информации о положении обслуживающих приборов, положении поступающих вызовов, их плотности распределения. Такие модели систем массового обслуживания служат для изучения реальных систем, где обслуживание производится территориально расположенными объектами.

Как наилучшим образом расположить эти объекты?

В данной статье рассматривается проблема оптимизации расположения станций обслуживания по критерию средней суммарной длины очереди. Проводится сравнение оптимальных расположений по критерию средней суммарной длины очереди и критерию среднего суммарного времени ожидания.

Станции функционируют как независимые системы массового обслуживания типа  $M|G|1$ . Описываются свойства оптимальных размещений и приводятся алгоритмы построения асимптотически оптимальных размещений.

## 2 Постановка задачи

На плоскости  $\mathbb{R}^2$  возникают требования в точках, являющихся реализациями некоторого случайного вектора  $\xi$ , распределенного с плотностью распределения  $p$ . Для обслуживания этих требований имеется  $n$  станций. Моменты поступления требований образуют пуассоновский поток с параметром  $\lambda$ .

Интенсивность входящего потока  $\lambda$  изменяется с ростом числа станций  $n$ . В случае, когда нужно подчеркнуть эту зависимость, параметр входящего потока будем обозначать  $\lambda(n)$ .

**Определение 1.** Размещением  $n$  станций обслуживания на плоскости  $\mathbb{R}^2$  назовем множество точек плоскости  $\{x_1, \dots, x_n\}$ , в которых они расположены.

Обозначать размещение станций будем символом  $x$ , то есть  $x = \{x_1, \dots, x_n\}$ . Станцию обслуживания и точку плоскости, где она расположена, будем обозначать одним и тем же символом.

**Определение 2.** Зоной влияния станции  $x_i$  назовем множество  $C_i$  тех точек плоскости, для которых эта станция является ближайшей, т. е.

$$C_i = \{v \in \mathbb{R}^2 : |v - x_i| \leq |v - x_j|, \quad j = 1, 2, \dots, n\}.$$

Расстояние  $|u - v|$  между точками  $u$  и  $v$  плоскости  $\mathbb{R}^2$  задается евклидовой нормой.

Станции обслуживают заявки только из своих зон влияния. Обслуживание осуществляется прибором,двигающимся только по прямой и с постоянной скоростью. При поступлении заявки прибор со станции перемещается в точку вызова, заявка обслуживается прибором некоторое случайное время  $\eta$ , затем прибор возвращается обратно на станцию. Дисциплина обслуживания вызовов следующая: если в момент поступления вызова прибор занят, то поступающий вызов ставится в очередь. После освобождения прибора первая заявка из очереди поступает на обслуживание.

Обозначим через  $\lambda_i$  интенсивность потока вызовов, поступающих на станцию  $x_i$ ;  $\beta_{i1}$  и  $\beta_{i2}$  —

<sup>1</sup>Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, кафедра математической статистики, lsa@cs.msu.su

соответственно первый и второй моменты времени обслуживания на  $x_i$ .

Предполагается, что станции функционируют как независимые системы массового обслуживания типа  $M|G|1$ . Тогда их средняя суммарная длина очереди  $L(x)$  при размещении  $x$  определяется по формуле

$$L(x) = \sum_{i=1}^n \frac{\lambda_i^2}{2} \frac{\beta_{i2}}{1 - \lambda_i \beta_{i1}}$$

при условии, что нагрузка каждой станции меньше единицы, т. е.  $\max_{1 \leq i \leq n} \lambda_i \beta_{i1} < 1$ .

Задача заключается в нахождении размещений, минимизирующих введенный критерий оптимальности  $L(x)$ .

В работе [1] рассматривалась аналогичная постановка задачи, но для систем с малой нагрузкой и с более жесткими условиями на плотность распределения  $p$ .

**Определение 3.** Размещение  $x^*$  назовем оптимальным, если  $L(x^*) \leq L(x)$  для любого размещения  $x$  такого, что  $|x| = |x^*|$ . Через  $|x|$  здесь обозначено число элементов размещения  $x$ .

**Определение 4.** Правильной  $n$ -решеткой назовем покрытие плоскости конгруэнтными правильными  $n$ -угольниками, пересекающимися лишь по границе. Фундаментальным регионом — правильный  $n$ -угольник.

Введем следующие обозначения:

$$f(k) = \int_{M_k} |u| du;$$

$$g(k) = \int_{M_k} |u|^2 du,$$

где  $M_k$  — правильный  $k$ -угольник единичной площади с центром в нуле;

$$|p|_m = \left( \int p^m(u) du \right)^{1/m};$$

$$E\eta = \beta_1; \quad E\eta^2 = \beta_2;$$

$$x_A^* = x^* \cap A;$$

$\{x\}$  — последовательность размещений.

### 3 Основные результаты

В следующей теореме описываются асимптотические свойства оптимальных размещений для исходной модели.

**Теорема 1.** Если плотность  $p$  ограничена и интегрируема по Лебегу,  $E|\xi|^2 < \infty$ , а интенсивность входящего потока  $\lambda(n)$  изменяется так, что

$$\lim_{n \rightarrow \infty} \frac{\lambda(n)}{n} = \frac{\rho}{\beta_1}, \quad \rho \in [0, 1),$$

то для всякой последовательности оптимальных размещений  $\{x^*\}$  справедливы равенства:

$$(1) \quad \lim_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x^*) = 0,5\beta_2(1 - \rho)^{-1};$$

$$(2) \quad \lim_{n \rightarrow \infty} \frac{|x_A^*|}{n} = \int_A p(u) du,$$

где  $A$  — произвольное измеримое по Лебегу множество.

Рассмотрим модель, в которой обслуживание состоит только лишь в перемещении прибора со станции до точки, где возникло требование, и обратно. В этом случае оптимальные размещения обладают другими свойствами.

**Теорема 2.** Если плотность  $p$  ограничена, интегрируема по Лебегу,  $E|\xi|^2 < \infty$  и интенсивность входящего потока требований изменяется так, что

$$\lim_{n \rightarrow \infty} \frac{\lambda(n)}{n^{3/2}} = \frac{\rho}{2f(6)|p|_{2/3}}, \quad \rho \in [0, 1),$$

тогда для всякой последовательности оптимальных размещений  $\{x^*\}$

$$(1) \quad \lim_{n \rightarrow \infty} \frac{n^2}{\lambda^2(n)} L(x^*) = 2g(6)|p|_{2/3}^2(1 - \rho)^{-1};$$

$$(2) \quad \lim_{n \rightarrow \infty} \frac{|x_A^*|}{n} = |p|_{2/3}^{-2/3} \int_A p^{2/3}(u) du$$

для любого измеримого по Лебегу множества  $A$ .

В работе [2] исследовались свойства оптимальных размещений по критерию среднего суммарного времени ожидания

$$W(x) = \frac{1}{2n} \sum_{i=1}^n \frac{\lambda_i \beta_{i2}}{1 - \lambda_i \beta_{i1}}$$

в аналогичной постановке задачи.

В данной статье также рассматривается проблема оптимизации расположения станций обслуживания по критерию  $W(x)$ , но с более мягкими условиями на плотность распределения  $p$ .

**Теорема 3.** Пусть плотность  $p$  ограничена,  $p^{2/3}$  интегрируема по Лебегу,  $E|\xi|^2 < \infty$  и  $\lambda_n = o(n^{1/2})$ , тогда для всякой последовательности оптимальных размещений  $\{x^*\}$

$$(1) \lim_{n \rightarrow \infty} \sqrt{n} \left( \frac{n}{\lambda_n} W(x^*) - 0,5\beta_2 \right) = 2f(6) \beta_1 |p|_{2/3};$$

$$(2) \lim_{n \rightarrow \infty} \frac{|x_A^*|}{n} = |p|_{2/3}^{-2/3} \int_A p^{2/3}(u) du,$$

где  $A$  — произвольное измеримое по Лебегу множество.

В случае, когда  $E\eta = 0$ , т.е. обслуживание заключается лишь в перемещении прибора до вызова и обратно, справедлива следующая теорема.

**Теорема 4.** Если плотность  $p$  ограничена,  $p^{1/2}$  интегрируема по Лебегу,  $E|\xi|^2 < \infty$  и  $\lambda_n = o(n^{1,5})$ , тогда для любой последовательности оптимальных размещений  $\{x^*\}$ :

$$(1) \lim_{n \rightarrow \infty} \frac{n^2}{\lambda_n} W(x^*) = 2g(6)|p|_{1/2};$$

$$(2) \lim_{n \rightarrow \infty} \frac{|x_A^*|}{n} = |p|_{1/2}^{-1/2} \int_A p^{1/2}(u) du$$

для любого измеримого по Лебегу множества  $A$ .

**Замечание.** Размещение, являющееся оптимальным по критерию  $W$  (стационарное среднее суммарное время ожидания начала обслуживания), не минимизирует критерий  $L$  (стационарная средняя суммарная длина очереди всех станций).

## 4 Доказательство основных результатов

Приведем только доказательство теоремы 1, доказательства остальных теорем проводятся аналогичными методами.

Для любой станции  $x_i$  с зоной влияния  $C_i$  некоторого размещения  $x$  первые два момента времени обслуживания оцениваются как

$$\beta_{i1} = E \left( \frac{2|\xi - x_i| + \eta}{\xi} \in C_i \right) \geq E\eta = \beta_1;$$

$$\beta_{i2} = E \left( \frac{(2|\xi - x_i| + \eta)^2}{\xi} \in C_i \right) \geq E\eta^2 = \beta_2,$$

а интенсивность потока поступающих на нее требований есть  $\lambda_i = \lambda(n)P(C_i)$ .

С учетом этого, а также используя выпуклость функции  $f(u, v) = u^2(1 - v)^{-1}$ , оценим снизу критерий  $L(x)$ :

$$L(x) = \sum_{i=1}^n \frac{\lambda_i^2}{2} \frac{\beta_{i2}}{1 - \lambda_i \beta_{i1}} \geq$$

$$\geq 0,5\lambda^2\beta_2 \sum_{i=1}^n \frac{P^2(C_i)}{1 - \lambda\beta_1 P(C_i)} \geq$$

$$\geq 0,5\lambda^2\beta_2 \frac{1}{n} \left( \sum_{i=1}^n P(C_i) \right)^2 \times \left( 1 - \lambda\beta_1 \frac{1}{n} \sum_{i=1}^n P(C_i) \right)^{-1},$$

т.е.

$$\frac{n}{\lambda^2(n)} L(x^*) \geq 0,5\beta_2 \left( 1 - \frac{\lambda\beta_1}{n} \right)^{-1}.$$

Устремляя  $n$  к бесконечности, получим

$$\liminf_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x^*) \geq 0,5\beta_2(1 - \rho)^{-1}.$$

Предположим сначала, что плотность  $p$  — простая функция, т.е.  $p = \sum_{j=1}^r p_j \mathbf{1}_{K_j}$ , где  $K_j$  — измеримые по Лебегу непересекающиеся множества на плоскости.

Для такой плотности построим асимптотически оптимальное размещение  $x$ , т.е. такое, что при  $|x| = |x^*|$  выполняется равенство

$$\lim_{n \rightarrow \infty} \frac{L(x)}{L(x^*)} = 1.$$

Для каждого множества  $K_j$  выберем элементарное множество  $L_j$  так, что  $\mu(K_j \Delta L_j) < \varepsilon$ , где  $\varepsilon$  — некоторое произвольное положительное число, затем  $L_j$  покрываем правильной 6-решеткой с площадью фундаментального региона  $\sigma_j = K_j/m_j$ , где  $m_j = m(1 - \delta)p_j K_j$ ,  $m$  — некоторое натуральное число и  $0 < \delta < 1$ .

Если  $\mu(\sigma_j \cap L_j) > 0$ , то в центры таких  $\sigma_j$  помещаем станцию обслуживания. Число таким образом размещенных станций обозначим через  $n_j$ . Построим последовательность вложенных расширяющихся квадратов  $K$  с центром в нуле, которые удовлетворяют условию  $E|\xi|^2 \mathbf{1}_{K^c} = o(m^{-1})$ .  $[m\delta] + 1$  станцию разместим равномерно на множествах  $(K_i \setminus L_i) \cap K$ . Общее число размещенных станций обозначим через  $n$ . Тем самым получено некоторое размещение  $x = \{x_1, \dots, x_n\}$ , для которого

$$\begin{aligned} L(x) &\leq \\ &\leq \sum_{j=1}^r n_j \frac{\lambda^2 p_j^2 \sigma_j^2}{2} \frac{\beta_2 + 4\beta_1 f(6) \sigma_j^{1/2} + 4g(6) \sigma_j}{1 - \lambda \beta_1 p_j \sigma_j - 2\lambda f(6) p_j \sigma_j^{3/2}} = \\ &= \frac{\lambda^2}{2} \sum_{j=1}^r \frac{\beta_2 p_j K_j (n(1 - \delta))^{-1} + o(n^{-1})}{1 - \lambda \beta_1 (n(1 - \delta))^{-1} + o(n^{-1})} + \\ &\quad + O \left( \frac{\varepsilon}{n \delta} \right). \end{aligned}$$

Устремим  $\varepsilon$ , а затем и  $\delta$  к нулю, тогда

$$\limsup_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x) \leq 0,5\beta_2(1 - \rho)^{-1}.$$

Так как всегда  $L(x^*) \leq L(x)$ , то с учетом нижней оценки для  $L(x^*)$  получаем, что  $x$  — асимптотически оптимальное размещение и

$$\lim_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x^*) = 0,5\beta_2(1 - \rho)^{-1}.$$

Пусть  $p$  — произвольная функция, удовлетворяющая условиям доказываемой теоремы. Введем простые функции  $\bar{p}_k(u)$  по правилу:

$$\bar{p}_k(u) = \frac{m + 1}{k},$$

если  $m/k < p(u) \leq (m + 4)/k$ , для  $k \in \mathcal{N}$ ,  $m = 0, 1, \dots$

Очевидно, что  $p(u) \leq \bar{p}_k(u)$ , а для простых функций выше была получена предельная оценка сверху, поэтому справедливо следующее неравенство

$$\limsup_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x^*) \leq 0,5\beta_2 |\bar{p}_k|_1 (1 - \rho)^{-1}.$$

При  $k \rightarrow \infty |\bar{p}_k|_0 \rightarrow |p|_1 = 1$ . И с учетом оценки снизу для  $L(x^*)$  получаем, что

$$\lim_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x^*) = 0,5\beta_2(1 - \rho)^{-1}.$$

Докажем второй пункт теоремы 1.

Рассмотрим размещение  $x_A = x \cap A$ . Пусть  $k = |x_A|$  — число станций, попадающих в множество  $A$  при размещении  $x$ . Определим  $L(x_A)$  как  $L(x)$  для  $p \mathbf{1}_A$ :

$$\begin{aligned} L(x_A) &= \sum_{i=1}^k \frac{\lambda_i^2}{2} \frac{\beta_{i2}}{1 - \lambda_i \beta_{i1}} \geq \\ &\geq 0,5\lambda^2\beta_2 \sum_{i=1}^k \frac{P^2(C_i \cap A)}{1 - \lambda\beta_1 P(C_i \cap A)} \geq \\ &\geq 0,5\lambda^2\beta_2 \frac{1}{k} \left( \sum_{i=1}^k P(C_i \cap A) \right)^2 \times \\ &\times \left( 1 - \lambda\beta_1 \frac{1}{k} \sum_{i=1}^k P(C_i \cap A) \right)^{-1}, \end{aligned}$$

значит,

$$\frac{k}{\lambda^2(n)} L(x_A) \geq 0,5\beta_2 P^2(A) \left( 1 - \frac{\lambda\beta_1 P(A)}{k} \right)^{-1}.$$

Следовательно,

$$\liminf_{n \rightarrow \infty} \frac{k}{\lambda^2(n)} L(x_A) \geq 0,5\beta_2 P^2(A) \left( 1 - P(A) \frac{\rho}{\gamma_1} \right)^{-1},$$

где  $\gamma_1$  — предельная точка последовательности  $\{|x_A|n^{-1}\}$ .

Пусть теперь  $x$  — асимптотически оптимальное размещение для критерия  $L$ . Тогда

$$\limsup_{n \rightarrow \infty} \frac{n}{\lambda^2(n)} L(x_A) \leq 0,5\beta_2 P(A)(1 - \rho)^{-1}.$$

Из двух последних неравенств следует, что

$$\gamma_1 \geq \frac{P^2(A)}{1 - P(A)\rho\gamma_1^{-1}} \frac{1 - \rho}{P(A)} = \frac{\gamma_1 P(A)(1 - \rho)}{\gamma_1 - P(A)\rho},$$

т. е.

$$\gamma_1 - P(A)\rho \geq P(A)(1 - \rho) \Rightarrow \gamma_1 \geq P(A).$$

Пусть  $\gamma_2$  — предельная точка последовательности  $\{|x_B|n^{-1}\}$ , где  $B = A^c$ .

Для  $\gamma_2$  аналогично доказывается соответствующее неравенство

$$\gamma_2 \geq P(B).$$

Так как

$$1 = \gamma_1 + \gamma_2 \geq P(A) + P(B) = 1,$$

то в неравенстве достигнуто равенство. Это возможно, только если

$$\gamma_1 = P(A), \quad \gamma_2 = P(B).$$

Тем самым для любого асимптотически оптимального размещения

$$\lim_{n \rightarrow \infty} \frac{|x_A|}{n} = \int_A p(u) du,$$

а значит, это равенство верно и для оптимального размещения. Отсюда следует утверждение второго пункта теоремы 1.

## Литература

1. Назаров Л. В., Смирнов С. Н. Обслуживание вызовов, распределенных в пространстве // Изв. АН СССР. Сер. Техническая кибернетика, 1982. № 1. С. 95–99.
2. Захарова Т. В. Оптимизация расположения станций обслуживания на плоскости // Изв. АН СССР. Сер. Техническая кибернетика, 1987. № 6. С. 83–91.

# НОВЫЙ ЭТАП ИНФОРМАТИЗАЦИИ ОБЩЕСТВА И АКТУАЛЬНЫЕ ПРОБЛЕМЫ ОБРАЗОВАНИЯ

И. А. Соколов<sup>1</sup>, К. К. Колин<sup>2</sup>

**Аннотация:** Анализируются основные черты и особенности современного этапа процесса информатизации общества и связанные с этим актуальные проблемы модернизации образования. Показано, что новые условия существования человека в глобальном информационном обществе создают для него принципиально новые возможности и проблемы, которые еще недостаточно учитываются в системе образования и поэтому требуют ее существенной модернизации. Рассматривается состояние информационного общества в России и актуальные проблемы российского образования, связанные с переходом страны к инновационной стратегии развития.

**Ключевые слова:** глобальная информатизация общества; информационно-коммуникационные технологии; информационное общество в России; новая парадигма образования; опережающее образование; стратегия инновационного развития

## 1 Введение

Одним из основных направлений развития цивилизации в XXI в. является глобальная информатизация общества на основе создания новых средств информатики и информационно-коммуникационных технологий (ИКТ) и их все более широкого использования во всех сферах социальной практики. Овладение информацией и знаниями, а также обеспечение доступности социально значимой информации для широких слоев населения — это две важнейшие задачи развития общества на ближайшие десятилетия. Их решение является необходимым условием для развития и эффективного использования человеческого потенциала, развития экономики, науки, образования и культуры, а также для решения многих глобальных проблем современности, которые требуют объединенных усилий мирового сообщества. К числу таких проблем можно отнести снижение остроты существующего в настоящее время социального неравенства, решение многих проблем занятости населения, повышение общего уровня образованности и культуры людей, а также качества их жизни.

Анализ показывает, что новые условия существования человека в глобальном информационном обществе создают для него принципиально новые возможности и новые проблемы, которые еще недостаточно учитываются в современной системе образования и требуют ее существенной модернизации. Именно поэтому проблемы образования в информационном обществе являются сегодня ис-

ключительно актуальными и достаточно широко обсуждаются на страницах печати как в России, так и за рубежом. В статье рассматривается состояние и ближайшие перспективы развития информационного общества в России и связанные с этим актуальные проблемы модернизации российского образования.

## 2 Отличительные особенности развития процесса информатизации общества в XXI веке

События последнего десятилетия свидетельствуют о том, что процесс информатизации общества в XXI в. перешел на качественно новый этап своего развития. Он стал приобретать системный, скоординированный характер и направлен уже не только на распространение средств ИКТ в передовых странах и отдельных регионах мира, но также и на формирование глобального информационного общества, в котором создаются принципиально новые условия для жизни и деятельности человека, развития экономики, образования и науки. В развитии этого процесса усиливается роль международных организаций, таких как ООН, ЮНЕСКО, ЮНЕП, ИФЛА и др. Начиная с 2000 г., проблемы формирования информационного общества регулярно обсуждаются на авторитетных междунаро-

<sup>1</sup>Институт проблем информатики Российской академии наук, isokolov@ipiran.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, kolinkk@mail.ru

ных форумах с высоким уровнем представительства. В их работе принимают участие не только ученые и общественно-политические деятели, но также и руководители государств, международных организаций, члены парламентов и правительств различных стран мира. Активно участвует в этих форумах и Россия.

Обсуждение наиболее актуальных и стратегически значимых проблем становления информационного общества показало, что такими проблемами на ближайшие десятилетия являются следующие:

- проблема эффективного использования тех новых возможностей, которые открывают достижения научно-технического прогресса в сфере ИКТ, в интересах социально-экономического развития общества и эффективного использования человеческого потенциала;
- проблема информационного неравенства между людьми, странами и отдельными регионами в активно формирующейся новой информационной среде общества, которая порождает новые формы социального неравенства и во многих случаях усиливает остроту этой проблемы;
- проблема обеспечения информационной безопасности человека, общества и государства в условиях информационного общества;
- проблема формирования новой информационной культуры общества и новой парадигмы образования, которые были бы адекватными новым условиям жизни и деятельности людей в информационном обществе.

Общие подходы к решению этих проблем международным научно-образовательным сообществом в основном уже выработаны, но каждая страна решает эти проблемы по-разному, с учетом своей специфики, национальных традиций, состояния экономики, науки, образования и культуры. Россия здесь не является исключением. Поэтому процесс развития информационного общества в нашей стране также имеет свои особенности, которые в своих основных чертах будут рассмотрены ниже.

### 3 Современное состояние процесса формирования информационного общества в России

**Общая оценка.** По экспертным оценкам отечественных и зарубежных специалистов, информационное развитие России идет достаточно быстрыми

темпами. Активно развивается отечественный рынок информационных и коммуникационных технологий, продуктов и услуг. Продолжается компьютеризация многих отраслей хозяйства, а также сфер государственного и регионального управления, создается современная правовая и нормативная база в информационной сфере. В рамках Федерального целевого проекта «Образование» осуществляется массовая компьютеризация российских школ, в 2007 г. завершено их подключение к сети Интернет. Функционирует созданный в 2006 г. Совет главных конструкторов информатизации регионов России, который осуществляет координацию работ в этой области, в том числе ряда проектов по созданию информационно-технологической инфраструктуры страны.

В 2007 г. общий объем отечественного рынка ИКТ вырос на 25,4% и, по предварительным оценкам, составил около 1,5 трлн руб. При этом объем услуг электросвязи оценивается значением 983 млрд руб. (рост на 25,8%), услуг почтовой связи — 67 млрд руб. (рост на 24,8%), а объем рынка информационных технологий увеличился на 24,5% и приблизился к значению 450 млрд руб. Наиболее быстрыми темпами (40,4%) растет рынок программных средств, объем которого в 2007 г. составил около 80 млрд руб. Рынок аппаратных средств составил 252 млрд руб. (рост на 18,5%), а рынок информационных услуг населению страны достиг значения 117,9 млрд руб. (рост на 28,4%).

По оценкам специалистов Мининформсвязи, приведенные выше показатели развития отрасли ИКТ во многом обусловлены привлечением в нее отечественных инвестиций, которые в 2007 г. составили 206 млрд руб., что на 20,8% больше по сравнению с инвестициями 2006 г. Тем не менее, в традиционном ежегодном рейтинге готовности стран мирового сообщества к «электронному развитию», который в апреле 2007 г. опубликован международной исследовательской компанией Economist Intelligence Unit (EIU), входящей в состав Economist Group, Россия сегодня занимает только 57 место из 69 стран, обследованных с точки зрения их готовности внедрять новые информационные технологии. Год назад Россия занимала в этом списке 52 место, а сегодня она уже уступает Индии, Филиппинам и Китаю, которые занимают в данном рейтинге соответственно 54, 55 и 56 места. Лидируют же в данном рейтинге Дания, США и Швеция, а аутсайдерами являются Азербайджан и Иран [1].

В качестве главных причин отставания нашей страны в данной области зарубежные эксперты называют происходящую в России модернизацию системы образования и административную реформу правительства, которая была проведена в 2004 г.

С этой оценкой трудно не согласиться. Ведь проводимая в настоящее время в стране государственная политика в сфере образования уже не рассматривает ее информатизацию в качестве приоритетного направления, как это имело место ранее. В особенности это касается содержательных аспектов образования, переподготовки педагогических кадров и формирования новой информационной культуры общества, адекватной современным требованиям формирующегося информационного общества.

**Новый этап развития информационного общества в России.** В 2007 г. руководством нашей страны был принят ряд важных организационных решений, направленных на формирование в России основ информационного общества. Президентом РФ одобрена Стратегия развития информационного общества в России на период до 2015 г. [2]. Ее реализация должна обеспечить повышение конкурентной способности страны, эффективности сферы государственного управления, повышение уровня и качества жизни граждан России, укрепление их конституционных прав, а также создание равных возможностей для доступа к необходимой информации и ИКТ. По своему содержанию Стратегия представляет собой политический документ, который определяет цели, принципы и основные направления государственной политики России в области развития ИКТ, науки, образования и культуры для продвижения страны на пути к информационному обществу. Кроме того, она содержит контрольные значения показателей развития информационного общества в России на период до 2015 г.

По предложению Президента России В. В. Путина создана Комиссия по вопросам развития информационного общества, которая будет функционировать при Совете при Президенте РФ по науке, технологиям и образованию. В ее состав вошли ученые, губернаторы, руководители ИКТ-компаний, депутаты Государственной Думы России. Председателем Комиссии назначен Министр информационных технологий и связи РФ Леонид Рейман. Основной задачей Комиссии является объединение усилий государства, бизнеса и гражданского общества по построению информационного общества в России. Она будет разрабатывать предложения по реализации Стратегии и системе прогнозных показателей, характеризующих уровень развития информационного общества в России, проводить по поручению президента страны экспертизу проектов федеральных законов и нормативных актов, а также осуществлять подготовку ежегодного Национального доклада о развитии информационного общества.

В Общественной палате РФ создана Рабочая группа по развитию информационного общества в России, ближайшей целью которой является общественная экспертиза процессов информатизации и организация диалога между представителями общественности, государственных органов власти и бизнеса по проблемам формирования информационного общества.

В 2007 г. Правительством России была также одобрена Концепция формирования «Электронного правительства» на период до 2010 г. Это новая форма деятельности органов государственной власти, основанная на широком применении ИКТ для получения гражданами и организациями государственных услуг.

Перечисленные выше политические и организационные решения, которые были приняты в 2007 г. Президентом и Правительством РФ, а также Советом безопасности и Общественной палатой России, свидетельствуют о том, что в стране начался новый этап процесса формирования информационного общества. Можно ожидать, что этот этап будет отличаться большей целенаправленностью, а также более высоким уровнем координации между основными участниками этого весьма сложного по своей реализации социально-технологического, научно-технического и культурологического процесса дальнейшего развития нашей страны и обеспечения ее национальной безопасности.

**Состояние основных направлений формирования информационного общества в России.** В последние годы были предприняты значительные усилия по развитию отрасли ИКТ, в результате которых Россия сегодня является мировым лидером по темпам развития этой отрасли. Общий объем отрасли в экономике России достиг значения в 1,5 трлн руб., а ее доля в структуре ВВП увеличилась с 3,2% в 2000 г. до 4,8% в 2007 г. Это означает, что данная отрасль становится локомотивом экономического роста страны. Согласно программе социально-экономического развития Российской Федерации, вклад ИКТ в общий рост экономики страны должен составить 0,5%, что будет сопоставимо с вкладом от добычи нефти (0,6%). Основной рост здесь ожидается в области разработки программного обеспечения и системной интеграции. В этих сегментах рынка доля отечественных компаний весьма значительна, и уже сегодня они успешно конкурируют на внутреннем рынке с зарубежными фирмами. Предполагается, что в результате реализации основных мероприятий Стратегии развития информационного общества в России доля отечественных товаров и услуг на внутреннем рынке ИКТ к 2015 г. превысит 50% [2].

Анализ показывает, что по мере перемещения производства комплектующих изделий в страны Азии и снижения стоимости оборудования основная конкуренция на мировом рынке ИКТ сосредоточивается в области производства программных продуктов. Ожидается, что к 2010 г. объем мирового рынка услуг в области разработки программного обеспечения достигнет 140 млрд долларов. Имеющиеся в России интеллектуальные ресурсы, а также опыт отечественных разработок дают нашей стране реальные шансы занять на этом рынке достойное место. Однако для этого необходима дальнейшая государственная поддержка отрасли, ее обеспечение необходимыми кадрами и создание благоприятной среды предпринимательства.

*Макроэкономическая среда предпринимательства.* Стратегия развития информационного общества в России ставит задачу вхождения нашей страны к 2015 г. в число 20 передовых стран мира в области развития информационного общества. При этом предполагается, что основные направления этой Стратегии должны быть реализованы силами преимущественно российских компаний. Поэтому важнейшей задачей сегодня является формирование инфраструктуры их деятельности и повышения конкурентной способности. С этой целью в 2006 г. постановлением Правительства РФ создан и с 2008 г. начнет полноценно функционировать Российский инвестиционный фонд информационно-коммуникационных технологий. Его уставной капитал составляет 1,45 млрд руб., причем 100% акций этого фонда находится в федеральной собственности.

Состоялось и учредительное собрание Национальной ассоциации инноваций и развития информационных технологий (НАИРИТ). Ее основной задачей является реализация проектов, призванных поднять уровень отечественных ИКТ, а также привлечение в данную сферу зарубежных инвесторов. В состав этой ассоциации выразили желание войти более 200 инновационных компаний, связанных с ИКТ. Создана Ассоциация предприятий отечественных производителей телекоммуникационного оборудования «Совет главных конструкторов». В настоящее время в России формируется также и система венчурного финансирования, которая должна обеспечить доступ ИКТ-предприятий к финансовым ресурсам.

*Первые отечественные технопарки в сфере ИКТ.* Государственная программа создания технопарков в сфере ИКТ получила в последние годы свое новое развитие. Такие технопарки создаются сегодня в семи регионах России: Московской области,

Санкт-Петербурге, Новосибирске, Нижнем Новгороде, Казани, Тюмени и Калуге. В дальнейшем планируется придать им статус особых экономических зон, что даст дополнительные экономические стимулы для привлечения инвестиций в их развитие со стороны бизнес-структур. Основная функция технопарков заключается в формировании инфраструктуры для развития российских предприятий. Кроме того, по прогнозам Мининформсвязи, развитие технопарков позволит создать порядка 75 тыс. новых рабочих мест, а годовая стоимость выпускаемой ими продукции и оказываемых услуг может превысить 100 млрд руб.

Можно ожидать, что создание технопарков позволит также уменьшить отток из нашей страны интеллектуального капитала, и в особенности талантливых программистов. Однако реализация этого проекта требует решения ряда проблем, связанных с подготовкой кадров. В настоящее время система образования России еще не обеспечивает подготовки необходимого числа квалифицированных специалистов для новой информационной отрасли, особенно с учетом ее планируемого расширения. Сегодня в нашей стране имеется острый дефицит квалифицированных программистов, специалистов среднего звена, менеджеров и руководителей проектов.

*Система связи и информационных коммуникаций.* Инфраструктура связи в России за последние годы продолжает улучшаться. В результате модернизации и реконструкции электронных междугородных АТС, а также ввода в эксплуатацию новых станций, кабельных и радиорелейных линий, плотность фиксированной телефонной связи в стране увеличена до 32 аппаратов на 100 человек. При этом очередь на установку телефона сократилась в полтора раза — с 1,5 млн до 1 млн заявок. В 2007 г. было установлено более 73 тыс. таксофонов и организовано 16 тыс. пунктов коллективного доступа в Интернет. В 2008 г. планируется обеспечить телефонизацию всех населенных пунктов страны, среди которых около 8% еще не имеют фиксированной телефонной связи, и будет завершён первый этап внедрения универсальных услуг связи в масштабах всей страны. К 2010 г. плотность телефонной сети будет повышена в полтора раза по сравнению с 2005 г. Появится реальная конкуренция на рынках междугородной и мобильной связи.

В 2007 г. в России начата модернизация всей системы почтовой связи с переводом ее на современную технологическую базу. Восстановлена единая структура почты страны, создано Федеральное государственное унитарное предприятие «Почта России», которое приняло на свой баланс все наличное

имущество почты, в состав которой сегодня входит 42 тыс. почтовых отделений. Это исключительно важный и социально значимый федеральный проект, реализация которого позволит не только обеспечить современный уровень услуг почтовой связи, но также и предоставить населению новые виды услуг, в том числе возможности доступа к сети Интернет и использования электронной почты. В 2007 г. в почтовых отделениях создано более 3 тыс. пунктов коллективного доступа в Интернет, а их общее количество достигло 23 тыс.

Число пользователей сети Интернет в России в 2007 г. приблизилось к значению 35 млн чел., что почти на 40% превышает показатели 2006 г. При этом существенно увеличивается количество линий широкополосного доступа пользователей к сети Интернет. В соответствии с контрольными значениями показателей Стратегии развития информационного общества в России, число таких линий на 100 чел. населения должно составить к 2010 г. 15 линий, а к 2015 г. — 35 линий. Это уже весьма неплохие показатели по сравнению с мировым уровнем.

Что же касается использования персональных компьютеров, то, по оценкам Мининформсвязи, их число в России в настоящее время составляет 31,2 млн шт., что превышает аналогичный показатель 2006 г. почти на 36%.

*Глобальная навигационная спутниковая система.* В настоящее время в России проводятся работы по восстановлению и повышению работоспособности отечественной глобальной навигационной спутниковой системы ГЛОНАСС, в состав которой входит 24 высокоорбитальных спутника Земли. К концу 2007 г. их число было увеличено до 18, что уже обеспечивает возможность непрерывной навигации на всей территории России с точностью определения местоположения подвижных объектов порядка 10 м.

В 2008 г. планируется изготовить 400–500 тыс. абонентских навигационных комплектов этой системы, а в 2009 г. — обеспечить ее работу в полном составе. При этом точность навигационных определений будет сопоставима с точностью аналогичной американской системы GPS. В настоящее время в России создана ассоциация разработчиков, производителей и потребителей оборудования и приложений на основе глобальных навигационно-спутниковых систем «ГЛОНАСС/ГНСС-Форум». Эта ассоциация создана в рамках осуществления Комплексного плана мероприятий по форсированному развитию системы ГЛОНАСС в интересах национальной безопасности, а также для социально-экономических, научных и коммерческих целей.

Необходимо отметить, что ГЛОНАСС и GPS — уникальные навигационные системы глобального масштаба, которые сегодня не имеют конкурентов и играют важную роль в решении проблем обеспечения национальной безопасности многих развитых стран, и в первую очередь США и России.

*Готовность потребителей и бизнеса.* Основными потребителями ИКТ в России продолжают оставаться государство и крупные российские компании. Около 40% спроса приходится на предприятия финансовой и нефтегазовой сферы, торговли и связи. Металлургия, машиностроение, транспорт и другие отрасли еще отстают в использовании ИКТ. По оценкам специалистов, все еще недостаточно обеспечены современными средствами ИКТ силовые министерства и ведомства России, машиностроительные отрасли, строительная промышленность, сельское хозяйство, а также образование, медицина и сфера культуры.

Необходимыми мерами, с помощью которых государство может увеличить спрос на ИКТ со стороны предприятий всех отраслей экономики, являются: увеличение в стране числа специалистов по ИКТ, повышение компьютерной грамотности и информационной культуры занятого населения, а также повышение доступности оборудования за счет снижения импортных пошлин и сокращения сроков амортизации компьютерной техники.

*Наука и образование.* Социологические исследования показывают, что научные работники России, хотя и отстают еще от своих европейских коллег по уровню использования ИКТ, но в целом лидируют в этой области по сравнению с работниками других сфер деятельности. Интернет и электронную почту сегодня используют более 80% российских научных учреждений и вузов, но веб-сайты имеет лишь треть научных организаций. Около 40% ученых используют Интернет для выполнения совместных проектов с зарубежными партнерами. Однако еще не все работники российской науки имеют навыки работы с компьютерами. Так, например, компьютерной грамотностью обладают лишь 87% столичных гуманитариев, в то время как для специалистов в области естественных наук этот показатель равен 100%.

Достаточно быстро развиваются в России научно-образовательные сети (RUNNet/RBNet и др.), однако значительная часть используемого в них оборудования морально устарела. Сохраняются высокие цены на аренду цифровых каналов. Инфраструктура российских научно-образовательных сетей имеет значительно меньшую пропускную способность по сравнению с европейской, и, кроме того, она существенно неоднородна. Так, например, если скорость передачи данных между Моск-

вой и Санкт-Петербургом составляет 2,4 Гбит/с (среднеевропейский уровень), то между Новосибирском и Хабаровском — всего 4 Мбит/с.

Тем не менее в 2007 г. в России начата реализация программы «СКИФ Университеты», которая имеет целью объединение в суперкомпьютерную сеть вузов России и Белоруссии. Планируется оснастить эти вузы суперкомпьютерами СКИФ, а в дальнейшем объединить их в единую вычислительную сеть «СКИФ Полигон». Предполагается, что в состав этой сети войдут суперкомпьютеры, установленные в МГУ им. М. В. Ломоносова, Томском, Южно-Уральском, Владимирском, Белгородском и Санкт-Петербургском государственных университетах, а также в Объединенном институте проблем информатики НАН Белоруссии. Ожидается, что уже к середине 2008 г. суммарная вычислительная мощность этой сети составит более 100 терафлопсов. Обучение и переподготовку необходимых специалистов будет обеспечивать НИВЦ МГУ.

Важным этапом в развитии информационного общества в России является завершение в 2007 г. процесса подключения к сети Интернет всех российских школ, общее число которых составляет 62 тыс. В рамках приоритетного национального проекта «Образование» в течение 2006—2007 гг. к сети Интернет было подключено более 53 тыс. школ в 89 областях РФ. Более 60% из них находится в сельской местности, а 7,5 тыс. — в труднодоступных районах страны, где пришлось использовать спутниковые технологии. Уже к началу 2007/2008 учебного года выход в Интернет по выделенным каналам имели более 51 тыс. школ (98%). При этом минимальная скорость передачи данных составляла 128 кбит/с. Необходимо отметить, что к сети Интернет сегодня подключены не только средние образовательные учреждения, но также вечерние и кадетские школы, школы-интернаты, коррекционные и специальные школы, а также школы, находящиеся в составе исправительных учреждений.

В результате реализации данного проекта учащиеся и преподаватели как городских, так и сельских школ смогут пользоваться электронными библиотеками и электронными образовательными ресурсами, что создает новые возможности для существенного повышения доступности качественного образования. Однако для этого необходимо будет осуществить массовую переподготовку преподавателей школ и работников школьных библиотек, которые сегодня в своем большинстве еще не обладают необходимым уровнем информационной компетентности.

*Здравоохранение и медицина.* Информатизация сферы здравоохранения является сегодня в России

исключительно актуальной и важной социальной проблемой. В приоритетном национальном проекте «Здоровье» значительное внимание уделяется внедрению методов дистанционного обслуживания пациентов с использованием ИКТ, которые получили название «телемедицины». Для России с ее огромной территорией значение этих методов трудно переоценить. В последнее время в решении данной проблемы наметились определенные сдвиги. Так, например, в Ханты-Мансийском автономном округе уже функционирует 6 телемедицинских центров и 52 телемедицинских пункта, которые обеспечивают возможность оперативного дистанционного консультирования пациентов. Однако это лишь одно из направлений информатизации системы здравоохранения. Не менее актуальным является перевод в электронный формат всей системы обслуживания пациентов, информатизация рабочих мест врачей и других работников поликлиник, внедрение «электронных медицинских карт», создание электронных баз данных пациентов и т. п. Все это позволило бы резко ускорить оформление медицинских документов и высвободить время врачей для их непосредственной работы с больными. Сегодня же это направление в России пока еще только начинает развиваться.

*Социальная и культурная среда общества.* Со стороны населения России спрос на информационную технику и технологии продолжает расти, хотя все еще остается недостаточным, особенно в регионах Сибири, Севера и Дальнего Востока. По экспертным оценкам, более 30% российских семей имеют дома персональный компьютер. Число абонентов сотовой связи в 2007 г. достигло 180 млн, что существенно превышает общую численность населения страны, а в крупных городах — Москве и Санкт-Петербурге — численность их населения. Мобильная связь является важным фактором информатизации общества, так как она обеспечивает гражданам России колоссальную экономию общественно полезного времени. Важный социальный аспект реализации приоритетного национального проекта «Образование» состоит в том, что он стимулирует развитие телекоммуникационной инфраструктуры в регионах нашей страны, в результате чего во многих населенных пунктах России впервые появилась возможность доступа в Интернет по приемлемым для населения ценам.

Исключительно важной и актуальной задачей формирования информационного общества в России является обеспечение свободного доступа населения к социально значимым информационным ресурсам. Решение этой задачи осуществляется сегодня по трем основным направлениям: обеспе-

чение доступа населения к информации органов власти и государственным услугам («Электронное правительство»), создание центров доступа к правовой информации и создание общедоступных публичных информационных ресурсов (электронные библиотеки, архивы, музеи и т. п.).

Реализацию Концепции «Электронного правительства» в России планируется осуществить в период до 2010 г. в два этапа. На первом этапе (2008 г.) разрабатываются нормативные и правовые документы, завершается проектирование и создание опытных образцов межведомственных систем «Электронного правительства». На втором этапе (2009–2010 гг.) должно быть обеспечено внедрение типовых технологических и организационных решений. Реализация данной концепции будет происходить в рамках ФЦП «Электронная Россия». При этом ожидается, что трудозатраты органов государственной власти на обмен информацией между ведомствами снизятся примерно на 50%, уменьшится административная нагрузка на граждан и организации, что позволит стране ежегодно экономить до 10 млрд руб. на административных расходах. Необходимо отметить, что некоторые регионы РФ уже сегодня имеют и реализуют собственные концепции «Электронного правительства». Так, например, Концепция «Электронного правительства» Нижегородской области была утверждена еще в середине 2004 г.

Важной составной частью будущего «Электронного правительства» России должна стать единая информационная система в сфере здравоохранения и социального развития страны. По оценкам специалистов, имеющаяся в настоящее время информация для населения о наборе предоставляемых государством услуг является явно недостаточной. Необходима комплексная информатизация деятельности всей социальной сферы. Одной из задач здесь является создание автоматизированной информационной системы «Электронный социальный регистр населения», которая внедряется в Саратовской области, а с 2008 г. будет распространена и на все другие регионы России.

В области же создания системы общедоступных публичных информационных ресурсов имеется ряд серьезных проблем, прежде всего правового характера. По мнению специалистов, входящих в состав Российской ассоциации электронных библиотек, сегодня необходима разработка Стратегии развития публичных информационных ресурсов России, а также принятие ряда новых и уточнение действующих законодательных и нормативных актов.

*Правовое обеспечение электронного развития.* По оценкам специалистов, существующая в России

нормативная правовая база в области информационной сферы общества еще не соответствует современным требованиям и нуждается в существенном дальнейшем развитии. В частности, еще недостаточно разработаны механизмы защиты прав на интеллектуальную собственность и обеспечения патентного права, что приводит к потерям доходов российскими экспортерами и мешает привлечению в Россию крупных международных компаний.

Институтом государства и права РАН еще в 2006 г. разработана Концепция развития информационного законодательства в Российской Федерации, которая предусматривает системное развитие информационного законодательства в нашей стране на основе принятия ряда новых федеральных законов. В Концепции правового регулирования в сфере информационных технологий, разработанной Мининформсвязи, также предусмотрено внесение изменений в действующие законы и принятие ряда новых законов («Об электронном документообороте», «Об информации персонального характера», «Об электронной торговле», «О праве на информацию», «Об участии в международном информационном обмене» и др.). Российской ассоциацией электронных библиотек подготовлены проекты Федеральных законов «Об обязательном электронном экземпляре документа» и «О праве граждан на доступ к информации о деятельности органов государственной власти и местного самоуправления». Однако многие из этих проектов еще находятся на разных стадиях своего прохождения в правительстве и Государственной Думе.

Принятие Стратегии развития информационного общества в России, создание Комиссии по вопросам развития информационного общества при Совете при Президенте РФ по науке, технологиям и образованию, а также Рабочей группы по информационному обществу в Общественной палате РФ должны дать новый импульс развитию отечественного информационного законодательства. Ведь отставание России от мирового уровня в области развития информационного общества все еще остается весьма существенным и представляет реальную угрозу для ее национальной безопасности.

#### 4 Актуальные проблемы дальнейшего развития информационного общества в России

Активизация процессов развития информационного общества в России обостряет многие про-

блемы информационной безопасности государства, человека и общества и требует адекватных мер противодействия не только на федеральном, но и на региональном уровне, а также на уровне корпораций, предприятий и общественных организаций практически во всех сферах социальной активности общества [3]. С этой целью необходима организация соответствующей просветительской работы в обществе, а также решение задач кадрового обеспечения информационного развития страны.

*Развитие информационной инфраструктуры.* Анализ показывает, что современный уровень развития российской информационной инфраструктуры, использования информационно-коммуникационных технологий в общественном производстве и государственном управлении еще не соответствует задачам инновационного развития страны и повышения ее конкурентоспособности, диверсификации экономики, повышения благосостояния и качества жизни граждан, укрепления обороноспособности и безопасности. По основным показателям он существенно уступает развитым странам мира. Поэтому в ближайшие годы необходимо сосредоточить усилия на устранении инфраструктурных диспропорций в области связи и телекоммуникаций, что позволит снизить остроту проблемы «цифрового неравенства» между отдельными регионами России.

В нашей стране сегодня практически отсутствует производство собственной конкурентоспособной продукции микроэлектронной промышленности, телекоммуникационного оборудования и средств вычислительной техники. В результате этого зависимость развития российской информационной инфраструктуры от поставок зарубежных ИКТ, по оценкам специалистов, в настоящее время превышает критический уровень и представляет собой одну из актуальных проблем.

*В сфере науки и образования.* Система образования России сегодня еще не обеспечивает в необходимом объеме качественное воспроизводство трудовых ресурсов, требуемое для повышения конкурентоспособности страны в условиях постиндустриального развития и становления информационного общества, основанного на знаниях. Уровень компьютерной грамотности и информационной компетентности преподавателей образовательных учреждений, и в особенности общеобразовательных школ в сельской местности, является сегодня чрезвычайно низким и не обеспечивает эффективного использования современных ИКТ и электронных образовательных информационных ресурсов общества. Школьные библиотеки еще не стали современными центрами информационной

поддержки образовательного процесса, а их работники не имеют педагогического статуса и соответствующего уровня оплаты труда.

Несмотря на принимаемые государством меры по повышению уровня оплаты труда научных работников и преподавателей, проблема преемственности поколений в науке и образовании остается достаточно острой. По оценкам ряда специалистов, уровень общей образованности российского общества снижается. Качество государственных образовательных стандартов и учебников во многом еще не отвечает современным требованиям, особенно в области базовых дисциплин (математики, физики, филологии, истории, информатики). Содержание образования в целом еще не ориентировано на новые условия существования человека в информационном обществе, а социальные аспекты информатизации общества, в том числе проблемы информационной безопасности, в системе образования практически не изучаются. Не развернуты также и работы по формированию новой информационной культуры личности и общества, адекватной современным требованиям.

*В социальной сфере общества.* В современной России наблюдается значительное неравенство в доступе к информации и современным ИКТ различных групп населения и регионов, что является важным фактором, тормозящим социально-экономическое развитие общества и повышение качества жизни населения. Кроме того, «цифровое неравенство» усиливает социальное расслоение российского общества и поэтому является одной из угроз для национальной безопасности страны.

*В сфере культуры.* Работы по использованию ИКТ для сохранения культурного наследия России, формированию электронных ресурсов библиотек и использования их в режиме удаленного доступа сегодня еще недостаточно развернуты. Одной из причин здесь является нерешенная ситуация в области авторского права, так как электронные копии документов должны распространяться бесплатно. Другая причина — все еще низкий уровень информатизации учреждений культуры, который не позволяет обеспечивать потребности населения в социально значимой информации. Так, например, только 15% российских библиотек используют современные ИКТ и лишь 9% из них подключено сегодня к сети Интернет. Материально-техническая база большинства библиотек безнадежно устарела и нуждается в качественном обновлении. В стране практически не строятся новые здания и специализированные помещения для публичных библиотек.

**Первоочередные задачи.** Формирование информационного общества в России должно стать од-

ним из важнейших приоритетов ее государственной политики на ближайшие годы. Необходимо также организовать проведение широкомасштабной разъяснительной кампании о значении, целях и задачах формирования в России информационного общества.

В числе первоочередных задач нужно обеспечить безусловное выполнение планов Мининформсвязи РФ по подключению в 2008 г. всех населенных пунктов страны к системе стационарной телефонной связи, а также по созданию точек коллективного доступа в Интернет во всех населенных пунктах страны. При этом к июлю 2008 г. минимальная скорость передачи данных должна достигнуть значения 256 кбит/с. Кроме того, количество таксофонов должно возрасти с 73 до 140 тыс.

*В сфере науки и образования.* Необходимо сформировать программу массовой переподготовки преподавателей школ и работников школьных библиотек с целью обеспечения необходимого уровня их компьютерной грамотности и информационной компетентности для эффективного использования ИКТ и электронных информационных образовательных ресурсов общества и приступить к практической реализации этой программы. Нужно также внести необходимые коррективы в Программу модернизации российского образования с учетом основных положений Концепции формирования информационного общества в России. Информатизация сферы образования вновь должна стать приоритетным направлением этой программы.

*В области правового обеспечения информационного развития страны.* Существующая законодательная и нормативная правовая база в информационной сфере России еще не соответствует современным требованиям, а работы по ее формированию ведутся очень медленно. В первую очередь здесь необходимо рассмотреть и принять уже подготовленные проекты законов. Кроме того, необходимо ввести в текст Федерального закона «Об обязательном экземпляре документов» ряд поправок, регламентирующих поставку обязательных экземпляров документов в электронной форме, с целью правового обеспечения работ по формированию национального электронного информационно-библиотечного фонда социально значимой информации [4].

## 5 Инновационное развитие России и новая парадигма образования

Выступая на расширенном заседании Государственного Совета 8 февраля 2008 г., Президент Рос-

сии В. В. Путин поставил стратегическую задачу радикального изменения курса дальнейшего развития страны и реализации крупных преобразований практически во всех основных сферах нашего общества уже в период до 2020 г. Понятно, что для решения задач такого масштаба и значимости потребуются принципиально новый подход к проблеме развития и использования человеческого потенциала, подготовки кадров для инновационной России. Какой же должна стать отечественная система образования для того, чтобы обеспечить решение этих задач? Анализ процессов развития глобальных проблем современности и тех основных качеств, которыми должны обладать люди для своей успешной социальной адаптации в новых условиях стремительно формирующегося информационного общества, показывает, что перспективная система образования должна обладать рядом принципиально новых качеств [5, 6].

Эти качества настолько существенны, что их совокупность может рассматриваться как *новая образовательная парадигма*, ориентированная на принципиально новые условия существования человека и общества в XXI в. В числе этих качеств наиболее важными представляются следующие:

1. *Опережающий характер* образования, его направленность на решение проблем формирования информационной цивилизации, развитие творческих способностей человека, его умения самостоятельно принимать ответственные решения в условиях все более динамично развивающегося общества.
2. Существенное расширение и качественное развитие *высшей школы*, которая должна обеспечить необходимое для условий XXI в. число специалистов с высшим образованием.
3. *Фундаментализация образования* за счет его все большей ориентации на изучение новейших достижений науки в области познания фундаментальных закономерностей развития природы, человека и общества, а также роли информации в реализации этих закономерностей.
4. Существенно большая *доступность системы образования* для широких масс населения, которая необходима не только для повышения общего уровня интеллектуального и духовного развития общества, но и для достижения большей социальной стабильности в обществе, уменьшения социального неравенства.
5. *Непрерывность образования*, переход к практической реализации концепции ЮНЕСКО «Образование через всю жизнь». Развитие системы открытого образования и повышения квалификации дипломированных специалистов. Создание системы образования и самообразования для

взрослого населения, а также для инвалидов и людей с ограниченной мобильностью.

Таким образом, речь идет о необходимости перехода к новой философской концепции образования, целью которого должна быть признана, прежде всего, высокая образованность человека, а не подготовка специалиста узкого профиля, как это в большинстве случаев имеет место сегодня. Суть этой концепции заключается в следующем: перестроить содержание и методологию учебного процесса во всех звеньях системы образования таким образом, чтобы она оказалась способной своевременно готовить людей к новым условиям существования, давать им такие знания, умения и навыки, которые позволили бы им успешно адаптироваться, жить и работать в новой социальной и информационной среде общества.

## 6 Заключение

Процесс информатизации общества в XXI в. перешел на качественно новый этап своего развития — он стал системным и направлен на формирование глобального информационного общества, в котором создаются принципиально новые условия для жизни и деятельности человека. Развитие средств и методов информатики, их использование практически во всех сферах социальной практики осуществляется стремительно и в последние годы принимает масштабы одного из важнейших направлений развития не только научно-технического прогресса, но и всей современной цивилизации. Однако эти процессы еще мало учитываются в системе образования, которая сегодня не обеспечивает подготовки людей к новым условиям существования в информационном обществе.

Проблема формирования информационного общества в России ставит перед научно-образовательным сообществом, и прежде всего перед системой образования, новые стратегические задачи, которые должны в значительной степени изменить существующую образовательную политику, в первую очередь в области развития высшей школы. Суть этих изменений заключается в том, чтобы привести структуру и содержание образования в соответствие с теми новыми требованиями, которые выдвигает на повестку дня информационное общество. При этом недостаточно будет обеспечить лишь подготовку необходимого числа специалистов в области ИКТ, информационной экономики и бизнеса, хотя, конечно же, это нужно сделать в первую очередь. Принятая в 2007 г. Стратегия развития информационного общества в России на период до 2015 г. требует соответствующей модернизации всей системы образования. При этом прежде всего

необходимо обеспечить переподготовку педагогических кадров средней и высшей школы, а также существенно переориентировать направленность и содержание диссертационных исследований, выполняемых в аспирантуре и магистратуре.

Стратегически важной и существенно более сложной задачей является формирование новой информационной культуры российского общества. Этой культурой должна будет, прежде всего, овладеть элитарная часть общества: ученые, преподаватели и аспиранты вузов, а также молодые специалисты, уже получившие высшее образование. Большую работу предстоит провести и в области переподготовки государственных служащих, причем не только чиновников, но также и медицинских работников, специалистов социальной сферы, правоохранительных органов, библиотечного сообщества, сферы культуры.

Для эффективного использования возможностей, которые создает развитие информационного общества, в стране должны быть соответствующим образом подготовлены люди. Миллионы российских граждан должны непрерывно получать новые знания, умения и навыки, необходимые им для жизни и деятельности в быстро изменяющихся условиях новой среды обитания, характерной для информационного общества. В значительной степени должна будет измениться и философия образования, в которой процессы глобальной информатизации общества должны будут получить свое необходимое отражение. Информатизация общества в России, имеющей огромную территорию, может и должна стать приоритетным направлением ее инновационного развития, так как она открывает новые возможности для использования высокого интеллектуального потенциала нашей страны.

## Литература

1. Рейтинг «электронной готовности» стран. <http://e-commerce.ru/cgi/print.asp>.
2. Стратегия развития информационного общества в России // Открытое образование, 2006. № 4(63). С. 4.
3. Доктрина информационной безопасности Российской Федерации // Научные и методологические проблемы информационной безопасности (сборник статей) / Под ред. В. П. Шерстюка. — М.: МЦНМО, 2004. С. 149–197.
4. Антапольский А. Б. О стратегии развития электронных публичных информационных ресурсов России // Открытое образование, 2007. № 6(65). С. 58–63.
5. Кинелев В. Г. Образование для информационного общества // Открытое образование, 2007. № 5(64). С. 46.
6. Колин К. К. Человек в информационном обществе: новые задачи образования, науки и культуры // Открытое образование, 2007. № 5(64). С. 40–46.

## ПРЕДПОСЫЛКИ И ФАКТОРЫ КОНВЕРГЕНЦИИ ИНФОРМАЦИОННОЙ И КОМПЬЮТЕРНОЙ НАУК\*

И. М. Зацман<sup>1</sup>, О. С. Кожунова<sup>2</sup>

**Аннотация:** Аналитический обзор посвящен проблеме конвергенции информационной науки (information science) и компьютерной науки (computer science), а также взаимосвязям этих наук с информационно-коммуникационными технологиями (ИКТ). Интерес к проблеме конвергенции возник более сорока лет назад, и одна из формулировок этой проблемы — The computer and information sciences: A new basic discipline — стала названием работы С. Горна, опубликованной в 1963 г. [1]. В наше время актуальность проблемы конвергенции существенно возросла, о чем свидетельствуют рассматриваемые в обзоре отдельные приоритетные направления исследований и разработок по ИКТ 7-й Рамочной программы Европейского Союза, принятой на период 2007–2013 гг. [2]. Приоритетные направления ИКТ позиционируются в обзоре как внешние факторы конвергенции. Кроме внешних факторов рассматриваются исторические предпосылки конвергенции информационной и компьютерной наук. Факторы и предпосылки конвергенции рассматриваются в контексте разработки научных основ создания новых поколений ИКТ.

**Ключевые слова:** информационная наука; компьютерная наука; информационно-компьютерная наука; информационно-коммуникационные технологии; взаимосвязи направлений информационно-коммуникационных технологий и компьютерной науки.

### 1 Введение

В соответствии с концепцией А. Соломоника научная парадигма любой «зрелой» науки состоит из следующих четырех составляющих, которые могут разрабатываться отдельно, но объединяются в единую и цельную конструкцию: философские основания, аксиоматика, классификация объектов предметной области исследования и система терминов [3]. При этом сам термин «научная парадигма» трактуется А. Соломоником в соответствии с теорией Т. Куна [4].

В рамках концепции Соломоника суть проблемы конвергенции можно выразить следующей фразой: описание парадигмы единой области знаний, охватывающей предметные области компьютерной и информационной наук, с общими философскими основаниями, включая позиционирование этой области знаний в системе современного научного мировоззрения, с общей аксиоматикой, единой классификацией объектов, процессов и явлений этой области знаний, общей и конвенциональной системой терминов.

Что касается последней составляющей научной парадигмы (системы терминов), то здесь необходимо учитывать объективно существующие трудности перевода с одного естественного языка на

другой. Например, в английском языке существует как минимум четыре устойчивых словосочетания — information science, computer science, computer and information science, computational science — которые нередко переводятся на русский язык одним словом «информатика». Приведенные англоязычные словосочетания обозначают разные научные дисциплины и их приложения. Однако при переводе на русский язык единым словом «информатика» содержательные отличия этих наук, различие их предметных областей и приложений не отражаются лексически.

Кроме того, в английском и русском языках разные по смыслу базовые понятия информационной науки (information science), компьютерной науки (computer science), информационно-компьютерной науки (computer and information science) и вычислительной науки (computational science) также часто лексически выражены одним словом, что затрудняет сравнительное описание предметных областей и приложений этих дисциплин. Например, слово “information” — «информация» — при его использовании в перечисленных предметных областях может трактоваться по-разному.

Различия в трактовках слова «информация», существенные для этого обзора, рассматриваются далее.

\* Работа выполнена при частичной поддержке РФФИ, грант № 06-07-07001ано.

<sup>1</sup> Институт проблем информатики Российской академии наук, im@a170.ipi.ac.ru

<sup>2</sup> Институт проблем информатики Российской академии наук, okozhunova@ipiran.ru

Для описания парадигмы единой области знаний, охватывающей предметные области компьютерной и информационной наук, необходимо явно определить значения используемых терминов. Ключевой характер явного определения системы терминов и смысла базовых понятий в любой области знаний отметил В. Ф. Турчин в своей книге «Феномен науки». Он писал: «Собственно говоря, ввести основные понятия — это и значит уже определить данную науку, ибо остается только добавить: описание мира с помощью этой вот системы понятий и есть данная конкретная наука» [5, с. 17].

Применяя это положение к предметной области обзора, можно сказать, что явно зафиксировать базовые понятия и определить систему терминов для описания парадигмы единой области знаний, охватывающей предметные области компьютерной и информационной наук, — это и значит уже во многом определить ее научную парадигму. Это и является причиной того, что в обзоре акцентируется внимание на терминологии.

Здесь возникает закономерный вопрос: а являются ли в настоящее время сами компьютерная и информационная науки по отдельности «зрелыми» науками? Этот вопрос выходит за пределы настоящего обзора и заслуживает самостоятельного изучения. Однако отметим, что в 1980 г. один из классиков информационной науки писал: «Едва ли теоретическая информационная наука уже существует. Я различаю рассеянные обрывки теории, некоторые из которых имеют более или менее стройный вид, но все же они не складываются в единую область знаний. Поэтому общих допущений (явных или неявных), которые могли бы претендовать на теоретические основания, попросту не существует. Информационная наука рассматривает большое число приложений, которые все больше требуют участия компьютера. Ни о каких основаниях она не может заявлять, если опирается только на здравый смысл в области языковых исследований, коммуникаций, отношений знаний и информации, на приложения компьютерных и телекоммуникационных технологий. Однако состояние компьютерной науки не намного лучше. В философском отношении информационная наука пребывает в подвешенном состоянии, так как у нее нет теоретических оснований» [6, с. 125].

Что касается проблемы конвергенции компьютерной и информационной наук, то в обзоре принята попытка показать, с одной стороны, что научная парадигма единой области знаний еще не сформировалась, с другой стороны, что по отдель-

ным позициям уже обозначился ряд предпосылок и факторов, стимулирующих процессы конвергенции. Здесь важно подчеркнуть, что кроме внутренних предпосылок существует важный внешний фактор, существенно влияющий на процессы конвергенции — достаточно четкие и явно эксплицированные потребности в разработке научных основ создания новых поколений ИКТ.

Одним из примеров явно обозначенных потребностей может служить 7-я Рамочная программа Европейского Союза, принятая на период 2007–2013 гг. В документах этой программы сформулировано восемь приоритетных направлений исследований и разработок, включая в качестве отдельного направления «Перспективные ИКТ» [2, 7].

Цели проектов, финансируемых в рамках приоритетного направления «Перспективные ИКТ», в программе на 2007–2008 гг. сформулированы следующим образом: «Своевременная идентификация и обоснование новых тематических направлений исследований и разработок, которые имеют большой научно-технический потенциал и могут стать основой для разработки ИКТ следующих поколений<sup>1</sup>. Эти проекты должны включать междисциплинарные исследования новых и альтернативных подходов к разработке ИКТ будущего и быть нацеленными на фундаментальное переосмысление системы теоретических, прикладных, методологических и технологических принципов, подходов и понятий, используемых сегодня в сфере ИКТ» [8, с. 54].

Важно отметить, что в цитируемом документе можно найти идентификацию и обоснование целого ряда конкурсных тем в рамках приоритетного направления «Перспективные ИКТ», включая тему «ИКТ долговременного применения» [8, с. 57–63].

Сфера применения результатов исследований и разработок по конкурсной теме «ИКТ долговременного применения», ее цели и задачи сформулированы следующим образом: «Широкое распространение и применение ИКТ, информационно-компьютерных и других цифровых систем в социально-значимых сферах деятельности кардинально увеличивает требования к их надежности, безопасности и долговечности. Это требует новых решений для того, чтобы обеспечить доверие пользователей в процессе их использования, защитить от несанкционированного к ним доступа и сохранить их функциональность в течение долгого периода времени в условиях крайне децентрализованной и быстрой смены поколений ИКТ, информацион-

<sup>1</sup>В цитируемом документе подчеркивается, что для разработки ИКТ следующих поколений особенно важны инновационные исследования и разработки, а не поэтапная модернизация существующих и используемых сегодня теоретических, прикладных, методологических и технологических принципов, подходов и понятий в сфере ИКТ.

но-компьютерных и других цифровых систем». Далее формулируются актуальные исследовательские проблемы в рамках темы «ИКТ долговременно-го применения». Приведем формулировки только двух проблем [8, с. 62–63]:

- (1) *Разработать теоретические и прикладные основы создания долговечных систем*, обеспечивающих их эволюцию при минимизации затрат на их развитие в условиях многократной смены поколений программно-аппаратных и сетевых средств и/или форматов данных. Другими словами, долговечные системы должны быть способны к сохранению своей первоначальной социально-значимой функциональности в течение долгого периода времени и изменять ее в случае необходимости. Методы сохранения и изменения функциональных возможностей должны быть машинно-независимыми и должны обеспечивать устойчивую эволюцию долговечных систем.
- (2) *Разработать новые подходы к представлению и сохранению знаний*, ориентированные на долговременный и безотказный к ним доступ в условиях локальной генерации отдельных «квантов» знаний, их интеграции, а также глобального использования систем представления и сохранения знаний с учетом контекста и временной эволюции систем. При этом должна быть обеспечена долговременная устойчивость систем представления и сохранения знаний в условиях многообразия их использования и эволюции семантики во времени.

Таким образом, в документах 7-й Рамочной программы Европейского Союза четко обозначена потребность в разработке научных основ создания ИКТ следующих поколений. Потребность в конкретных теоретических основаниях иногда формулируется в косвенной форме и нередко является импликацией (следствием) приведенных формулировок актуальных исследовательских проблем.

Например, фраза «локальная генерация отдельных «квантов» знаний. . .» во второй проблеме подразумевает возможность членения знаний на «кванты» некоторым способом, их представления в цифровой среде<sup>1</sup> и обеспечение доступа к сохраненным представлениям «квантов» знаний и отношениям между ними. Эти вопросы рассматриваются далее в обзоре, так как они являются ключевыми для определения предметной области, относящейся одновременно к компьютерной и информационной наукам. Здесь отметим только три вопроса, на которые необходимо ответить при описании любого

варианта парадигмы единой области знаний, охватывающей предметные области компьютерной и информационной наук, если ставится цель предложить вариант парадигмы как теоретическую основу создания новых поколений ИКТ.

Во-первых, допускает ли предлагаемый вариант парадигмы возможность членения системы знаний на «кванты», и если допускает, то является ли положение о возможности членения аксиомой или следствием других аксиом?

Во-вторых, допускается только единственный способ членения системы знаний или предполагается существование множества разных способов?

В-третьих, учитывается ли эволюция системы знаний человека во времени, и если учитывается, то как этапы (стадии) эволюции отражаются в способе (способах) членения, или аксиоматически предполагается, что в предлагаемом варианте парадигмы знания человека в разные моменты времени являются самоидентифицируемыми?

Отметим, что фраза «эволюция семантики во времени» в формулировке второй проблемы имеет непосредственное отношение к третьему вопросу. Все три перечисленных вопроса являются ключевыми (но далеко не единственными). На них необходимо дать ответы при описании любого варианта парадигмы, если она предлагается в качестве теоретической основы создания новых поколений ИКТ в трактовке новизны ИКТ в рамках 7-й Рамочной программы Европейского Союза.

Ответы на поставленные вопросы затрагивают все четыре составляющих научной парадигмы. Например, к философским основаниям парадигмы относится вопрос: «Являются ли знания человека в разные моменты времени самоидентифицируемыми?» [9]. Вопрос о возможности членения знаний на «кванты» скорее всего относится к аксиоматике этой области знаний. В классификации объектов, процессов и явлений единой области знаний необходимо описать отношения между процессами генерации отдельных «квантов» знаний, их интеграции и использования. В системе терминов необходимо дать определение «кванта» знаний и назвать этот «квант», например, концептом, понятием, значением или другим словом, которое и использовать далее как термин только в этом значении в рамках предлагаемого варианта парадигмы. Отметим, что в обзоре «квант» системы знаний человека, понятие и концепт рассматриваются как синонимы.

Далее будет показано, что сформулированные проблемы рассматриваются и в компьютерной, и в

<sup>1</sup>Цифровая среда — сочетание элементов цифровой вычислительной техники, средств телекоммуникации, информационно-компьютерных систем, иных цифровых средств ввода, хранения, поиска, передачи и других процессов обработки данных.

информационной науках. Таким образом, в настоящее время существуют внешние факторы конвергенции в виде приоритетных направлений ИКТ и актуальных теоретических проблем, возникающих в процессе создания новых поколений ИКТ и относящихся к предметным областям обеих наук. Как видно из приведенных примеров, иногда в самих формулировках проблем содержатся явно эксплицированные или косвенно сформулированные потребности в конкретных теоретических основаниях. Кроме внешних факторов в обзоре далее рассматриваются исторические предпосылки конвергенции информационной и компьютерной наук.

Прежде чем завершить введение к обзору, остановимся на еще одном документе, появление которого привлекло внимание лиц, принимающих решения в сфере научно-технической политики, к необходимости создания новых поколений ИКТ в интересах обеспечения конкурентоспособности национальной экономики, в том числе к разработке теоретических основ создания новых поколений ИКТ как ключевой составляющей общества, основанного на знаниях (knowledge-based society). Речь идет об аналитическом отчете по вопросам обеспечения конкурентоспособности США в XXI в., подготовленным Консультативным комитетом по информационным технологиям при Президенте США [10]. Информацию об этом отчете можно найти в работе [11].

В аналитическом отчете информационные технологии позиционируются как важная составляющая триады «научная теория – научный эксперимент – информационные технологии, обеспечивающие проведение эксперимента», являющаяся основой процессов научного познания практически во всех областях знаний. Чтобы выразить в явной форме сферу применения, роль и функции информационных технологий в триаде научного познания, авторы отчета определяют междисциплинарную область исследований и разработок, которую называют «computational science», что в этом обзоре переводится буквально как «вычислительная наука».

Определение этой области исследований и разработок, предлагаемое в отчете, имеет следующий вид [10, с. 10]: «Вычислительная наука — это

быстро растущая мультидисциплинарная предметная область, в которой используются возможности передового компьютеринга (advanced computing) для понимания и решения сложных проблем. Вычислительная наука интегрирует три компонента:

- (1) *алгоритмы (численные и нечисленные), программное обеспечение* для моделирования и имитирования, разработанные для решения проблем естественных, гуманитарных и инженерных наук;
- (2) *информационно-компьютерная наука*, которая разрабатывает и оптимизирует современные аппаратные, программные и сетевые средства, а также компоненты управления данными, которые необходимы для решения вычислительно сложных проблем;
- (3) *вычислительная инфраструктура*, которая поддерживает решение научных и инженерных проблем, а также развитие информационно-компьютерной науки»<sup>1</sup>.

В этом определении используется словосочетание *информационно-компьютерная наука*. Это словосочетание одним из первых использовал американский ученый С. Горн в 1963 г., с той лишь разницей, что тогда оно употреблялось им во множественном числе (computer and information sciences) [1]. С 1983 г. Горн начал использовать этот термин в единственном числе [12].

Таким образом, идея конвергенции была отражена С. Горном в самом названии научной дисциплины — *информационно-компьютерная наука* (computer and information science) — которая в аналитическом отчете по вопросам обеспечения конкурентоспособности США в XXI в. позиционируется как одна из трех составляющих вычислительной науки. Отметим, что трактовки информационно-компьютерной науки в этом отчете и в работах Горна отличаются. Описание и анализ трактовки Горна является одной из задач обзора.

Структура предлагаемого обзора имеет следующий вид. Разделы 2 и 3 посвящены отдельным вопросам становления информационной и компьютерной наук соответственно, а также формированию предпосылок их конвергенции. Краткий

<sup>1</sup>Computational science is a rapidly growing multidisciplinary field that uses advanced computing capabilities to understand and solve complex problems. Computational science fuses three distinct elements:

- (1) Algorithms (numerical and non-numerical) and modeling and simulation software developed to solve science (e.g., biological, physical, and social), engineering, and humanities problems;
- (2) Computer and information science that develops and optimizes the advanced system hardware, software, networking, and data management components needed to solve computationally demanding problems;
- (3) The computing infrastructure that supports both the science and engineering problem solving and the developmental computer and information science [10, с. 10].

обзор работ Горна и Шрейдера включен в разд. 4, за которым следует заключение.

## 2 Информационная наука

### 2.1 Становление информационной науки

По мнению скандинавского ученого, специалиста в области информационной науки Петера Ингверсена, изучаемая им наука — дисциплина молодая. В своей работе “Information and information science” он подчеркивает, что самое раннее использование термина «информационная наука» в научных кругах пришлось на 1958 г., когда был сформирован Institute of Information Scientists (IIS) в Великобритании. По планам его основателя, Джейсона Фаррадейна, предполагалось, что «использование термина «информационный ученый» поможет различать ученых, занимающихся информационной наукой, и ученых-естествоиспытателей, поскольку сотрудники института имели дело в основном со сбором, хранением и обработкой научно-технической информации» [13, с. 137].

Сотрудники вышеупомянутого института специализировались в разных областях знаний, зачастую очень сильно отличающихся друг от друга. В круг их основных обязанностей входили организация информационного обслуживания и предоставление научной информации исследователям из других институтов и промышленных лабораторий. Пионерами информационной науки были Б. К. Брукс, С. Клевердон, Р. Фейтхорн, Е. Гарфильд, М. Кочен, И. Ланкастер, Дж. Солтон, Д. де Солла Прайс и Б. Викери [13, с. 137].

В своей работе Ингверсен подробно объясняет, почему сотрудники IIS называли себя информационными учеными: «Называя себя информационными учеными, они, очевидно, хотели подчеркнуть важность научного подхода к изучению информации и процессов научных коммуникаций. Их работа являлась продолжением предыдущих теоретических и эмпирических попыток исследовать проблемы организации, роста и распространения информации, которая была накоплена перед второй мировой войной» [13, с. 137].

Чтобы проследить процесс становления информационной науки, Ингверсен обращается к ее истории и к предпосылкам, которые привели к формированию предметной области информационной науки. По его мнению, традиционно профессионалы, занимающиеся хранением документальных и издательских форм представления научных результатов, были известны как документалисты (поз-

же — информационные ученые) и библиотекари: «Первые из упомянутых обычным образом обучались какой-либо научной дисциплине и занимались прикладными аспектами передачи научно-технической информации применительно к своей дисциплине. Библиотекари обучались библиотечному делу (т. е. работе в библиотеках). Для них передача информации в таких учреждениях, как библиотеки, часто символизирует и социальные, и значимые культурные аспекты. Несмотря на идентичные способы обработки документов и информации и похожее использование информационных технологий, разделение на две группы продолжало существовать и в послевоенное время во многих странах, например в Скандинавии и Соединенных Штатах. Результат этого разделения можно увидеть в другом названии этой области знания: библиотечно-информационная наука. В дополнение к этому, для библиотечного сообщества сама по себе библиотечная наука иногда рассматривалась в качестве научной альтернативы информационной науке. Однако социокультурные коммуникации с помощью библиотек не могут существовать без процессов передачи информации» [13, с. 138].

Поэтому Петер Ингверсен отдельно оговаривает, что «библиотечная наука — это информационная наука и исследовательские методы, примененные к конкретному учреждению под названием «библиотека». Разделение на библиотечную и информационную науки является непродуктивным и носит искусственный характер. Наука не может быть целиком посвящена некоторому учреждению; например, медицина не может быть представлена на концептуальном уровне как больничная наука» [13, с. 138].

С одной стороны, такое разделение внесло свой вклад в кризис идентификации этой науки и фрагментировало стройное здание ее теории, построение которого потребовало нескольких десятилетий. С другой стороны, настаивая на большем числе социальных и гуманитарных аспектов, ассоциированных с передачей информации, библиотечное сообщество в 90-е гг. прошлого века способствовало эволюции в направлении консенсуса в информационной науке. В конечном счете, фокусирование только лишь на научных коммуникациях и передаче информации является слишком ограниченной основой для научной дисциплины. Влияние сообщества информационных ученых расширило концепцию предметной области за счет информационных процессов в тех сферах человеческой деятельности, в которых *знания и информация играли жизненно важную роль*, таких как коммерческая деятельность и социокультурные коммуникации» [13, с. 138; 14, 15].

В своем исследовании информационной науки и ее предмета Ингверсен обращается к истокам и основаниям этой науки. Он задается вопросами: что же было движущей силой попыток основания науки, занимающейся, главным образом, обработкой документов, и почему было так интересно и даже необходимо серьезно изучать вопрос обеспечения эффективной передачи желаемой и доступной информации от человека-генератора человеку-пользователю? Один из возможных ответов заключается в том, что информационная наука как область знания возникла вследствие осознания проблем как физического, так и интеллектуального доступа к чрезвычайно быстро растущему объему научных знаний (послевоенный информационный взрыв) [13, с. 138].

Ингверсен полагает, что подобного рода ответы получили широкое распространение, но при этом они являются частично поверхностными. В приведенном объяснении предполагается, что информацию можно приравнивать к документам, таким как публикации и другие физические сущности, содержащие какие-либо сообщения. Однако информация в информационной науке не обозначает физические сущности вроде документов. Если что и произошло в течение XX столетия, так, в первую очередь, это впечатляющий «документальный взрыв» как в науке, так и в обществе, что повлекло за собой усложнение доступа — и физического к документам, и интеллектуального — для получения адекватной информации» [13, с. 139].

Главный вывод, который делает Ингверсен о том, каким должно быть понятие информации в информационной науке, заключается в следующем. В пределах предметной области информационной науки понятие информации должно удовлетворять двум требованиям. С одной стороны, информация является результатом преобразования в *знаковую форму когнитивных структур человека-генератора*. При этом учитывается модель представления знаний, имеющихся у получателя этих знаковых форм. С другой стороны, это нечто такое, при восприятии и осознании чего имеющиеся знания получателя информации *подвергаются влиянию и трансформируются*. В результате, словосочетание «информационное общество» также подразумевает, что общество зависит от того, как оно использует информацию, а не только от того, как оно ее производит [13, с. 139].

Здесь необходимо отметить, что в обзоре, в целях различения и сопоставления точек зрения разных ученых с использованием лексически отличающихся-

ся выражений разных концептов, информацию как результат преобразования в знаковую форму когнитивных структур человека будем называть *знаковой информацией*.

Далее Ингверсен пишет: «... по существу, никому не требуется наука для обеспечения доступа к документам. Если что-то и необходимо, то это улучшенные методы, позволяющие людям поспевать за ростом документов. Эта практическая работа уже почти пять тысяч лет выполняется архивариусами (хранителями архивов), библиотекарями и документалистами. Они постоянно извлекали пользу из информационных технологий, доступных им в каждый исторический период времени, начиная с глиняных табличек и заканчивая пергаментом, бумагой и компьютерными методами» [13, с. 139]. Отметим, что первое использование компьютерных технологий для поиска научных документов отнесится еще к 1960-м гг. прошлого века.

Главной движущей силой использования технологических инноваций была потребность в быстром получении документов, релевантных какой-либо цели или проблеме. Неудивительно, что методы решения проблем доступа к документам и, что более существенно, к потенциально значимой информации *все больше определялись используемыми технологиями* [13, с. 139]. Здесь важно отметить идею зависимости методов решения проблем доступа к информации от степени развития используемых технологий. Эта идея, имеющая прямое отношение к проблеме конвергенции, более подробно будет рассмотрена далее.

С момента создания в 1958 г. IIS неоднократно предпринимались попытки установить основные направления исследований в информационной науке и определить ее границы с другими областями знания. Основная проблема заключалась в недостатке *базовых философских подходов к описанию информационных процессов*, кроме подразумеваемых рационалистических взглядов, унаследованных от физических наук. Причиной этому была неопределенность ее положения в системе научного познания<sup>1</sup>. Ряд специалистов настаивали на том, что информационную науку необходимо рассматривать как составляющую естествознания. Поэтому они стремились сформулировать и действительно сформулировали фундаментальные «законы» информации, которые вследствие особенностей человеческого подхода к использованию информации для познавательных целей можно рассматривать лишь как индикаторы наличия информационных процессов. Однако важно отметить и эти попыт-

<sup>1</sup>С точки зрения концепции А. Соломоника здесь речь идет о первой составляющей научной парадигмы информационной науки — философских основаниях и позиционировании этой области знаний в системе современного научного мировоззрения.

ки «форсированной научной эволюции» (основанной на желании совершенствоваться), поскольку без них данная область знания была бы поглощена близкими когнитивными дисциплинами еще в 1960-х гг. [13, с. 141].

В своем исследовании информационной науки Ингверсен, говоря об истории развития этой области, резюмирует, что за период ее существования предпринимались попытки слияния с другими областями знаний, с тем чтобы утвердить более весомую научную позицию в системе научного познания. Прослеживаются два основных направления: (1) движение в направлении теории коммуникаций и (2) попытка слияния с компьютерной наукой. В то же время ряд ученых приложили значительные усилия для сохранения независимости информационной науки с ее собственной индивидуальностью [13, с. 141; 16–18].

## 2.2 Фаррадейн: предметная область информационной науки

Одним из главных сторонников формирования и развития информационной науки как независимой области знаний являлся Дж. Фаррадейн, основатель ИИС.

Предметную область этой науки он описал следующим образом: «информационная наука по большей части когнитивная наука, т. е. имеет дело с мыслительными процессами, одной из самых сложных областей исследований. Конечно, она является частью более обширной области коммуникаций, преподавания и обучения. Но даже такие практические аспекты этой науки, как хранение и поиск информации, постоянно остаются в тени понятия релевантности, ментальной оценки, индивидуальной для каждого отдельного получателя информации и зависящей от его первоначального уровня знаний» [16, с. 75].

Такой подход к вопросу определения информационной науки приводит Фаррадейна к следующему выводу: «Чем больше мы изучаем то, что принадлежит когнитивным границам в информационной науке, т. е. ментальные процессы, которые генерируют информацию, и когнитивные процессы, происходящие при получении информации, тем больше у нас возможностей улучшать и контролировать процессы хранения и поиска информации для получения желаемых результатов» [16, с. 75].

В описании предметной области информационной науки Фаррадейн сопоставляет знания и информацию, в том числе ее трансформацию в знания, как центральные понятия и процессы обсуждаемой науки. Картина событий при передаче ин-

формации, которую описывает ученый, не содержит какого-либо лингвистического анализа, кроме описания ряда ограничений, накладываемых естественным языком. Зато она содержит описание процессов мышления человека-генератора информации и получателя, ее преобразования в знания. Фаррадейн определяет «знания» как отпечаток процессов понимания и осознания, происходящих в памяти человека, как нечто, доступное лишь в пределах памяти человека. При этом он отмечает, что сами процессы понимания и осознания в настоящее время остаются невыясненными.

Информация определяется им как сущность, заменяющая знания, например язык, и используемая для коммуникаций. Важно отметить, что определение информации по Фаррадейну, которую будем называть «языковой информацией», во многом совпадает со «знаковой информацией» по Ингверсену, но принципиально отличается от «ментальной информации» Брукса, которая будет рассмотрена далее.

О свойствах информации и ее связях со знаниями Фаррадейн говорит следующее: «... она [информация] нейтральна в том смысле, что она не обязательно должна быть новой для воспринимающего ее субъекта. Нам известно много практических примеров обработки информации в системах поиска и хранения, но ее отношение к знаниям не менее важно для развития информационной науки как науки об информации» [16, с. 77].

Ученый придает большое значение моделированию поисковых процессов и связанных с ними понятий: «Поисковые процессы (в частности, в памяти человека), если их можно было бы обнаружить и моделировать, могли бы предоставить лучшие методики поиска, чем методы поиска по образцу, которые используются в поисковых системах в настоящее время. Как правило, человек испытывает более серьезные затруднения при переводе своих потребностей в новых знаниях в лингвистическую форму запроса, чем при представлении уже имеющихся у него знаний в форме информации. В этом нет сомнений, поскольку пробел, породивший потребность в знаниях, действительно пуст, без каких-либо очевидных связей (отношений) с существующими у человека когнитивными структурами его знания.

Таким образом, подразумевается, что структура неполных отношений какой-либо области знания и порождает информационные потребности пользователя. Если бы существующие знания проблемной области могли бы быть полностью отображены в системе ментальных, связанных между собой, понятий, то, возможно, получилось бы выразить эту «потребность» в более точной форме, чем в той,

в которой был сформулирован вопрос пользователя. В комбинации с информацией, хранимой в структурированном виде, мог бы быть произведен более точный и полный поиск средствами автоматизированной системы» [16, с. 79]. Отметим, что, по мнению Фаррадейна, информационный поиск относится к предметной области информационной науки.

### 2.3 Брукс: основания информационной науки

Бертрам Брукс (его главная работа в данной области относится к 1981 г.) также был сотрудником IIS в Великобритании. Он активно изучал вопросы оснований информационной науки, поскольку в ней к 1980-м гг. накопились нерешенные теоретические вопросы, касающиеся ее оснований и системы терминов.

Одним из центральных вопросов при рассмотрении философских оснований информационной науки Брукс считает вопрос соотношения субъективного и объективного знания в научной деятельности. По его мнению, понятие информации, безусловно, является главным в системе терминов информационной науки, но оно же предполагает определенные сложности для ученого-теоретика. Даже на уровне здравого смысла, т. е. того, как мы ее себе представляем, информация является сущностью, которая проникает во все сферы человеческой деятельности. Поэтому особенно сложно отслеживать информационные явления в изоляции, обособленности, которой обычно требует научный вопрос. Даже процесс описания результатов наблюдений каких-то явлений сам по себе уже является информационной деятельностью, в рамках которой не так просто отделить объективное знание от субъективного.

Все социальные науки сталкиваются с подобной проблемой, но ни одна из них, согласно Бруксу, не предназначена для ее решения. Информационная наука наиболее близко связана со взаимодействиями ментальных и физических процессов, субъективных и объективных способов мышления. Поэтому, с точки зрения Брукса, на информационную науку возлагается особая ответственность. Она должна, насколько возможно, прояснить эти моменты, что является одной из основных ее задач.

#### 2.3.1 Миры Карла Поппера: мир 3 и информационная наука

В попытке сформулировать основания информационной науки Брукс исследовал философские труды. Наиболее значимой работой, связанной с

проблемами и основаниями информационной науки, стала для него книга Карла Поппера «Объективное знание». По мнению Брукса, в чем действительно нуждается информационная наука как в основе, так это в объективной, а не в субъективной теории знания [6, с. 127].

По мере углубления в вопрос соотношения объективного и субъективного знания в науке Поппер формирует следующие онтологии, так называемые миры Поппера:

- мир 1 — физический мир (мир материальных сущностей);
- мир 2 — мир субъективных человеческих знаний;
- мир 3 — мир объективных знаний, нашедших выражение в текстах на естественных языках, искусствах, науках, технологиях.

По мнению Брукса, именно мир 3 Поппера должен быть рекомендован библиотечным и информационным ученым, поскольку он впервые предлагает логическое обоснование их профессиональной деятельности, которое выражается в терминах, отличных от терминов практических приложений. А практическая деятельность библиотечных и информационных ученых складывается из собрания и хранения записей мира 3, в то время как теоретической проблемой является изучение взаимодействий между мирами 2 и 3, их описанием и объяснением способности или неспособности систематизировать знания (скорее их, чем документы) для более эффективного использования. События мира 2 — наших индивидуальных ментальностей — происходят в наших индивидуальных частных пространствах и потому являются субъективными. Для того чтобы объективизировать наши индивидуальные мысли, нам необходимо каким-то образом выразить их и занести эти записи в мир 3. Там они становятся доступными другим людям [6, с. 129–130].

Но такое разъяснение объективного знания и субъективного предполагает проблемы, которые, по мнению Брукса, Поппер не учел: «Не слишком верно полагать, что любое выражение мысли (или чувства), сохраненное в мире 3, является непосредственно доступным всякому, кто его ищет, как некая объективность, объективная реальность» [6, с. 130].

После этого заключения Брукс переходит к рассмотрению двух вопросов, которые являются ключевыми при описании теоретических оснований информационной и компьютерной наук:

- допускается ли возможность членения знаний на «кванты»;

- допускается ли эволюция системы знаний человека во времени или знания человека в разные моменты времени являются самоидентифицируемыми?

### 2.3.2 Фундаментальное уравнение информационной науки

Основополагающими понятиями информационной науки Брукс, как и остальные ученые, занимающиеся информационной наукой, считает информацию и знания. Именно поэтому вопрос их соотношения для него является центральным при обсуждении оснований этой науки. Брукс рассматривает знания как структуру взаимосвязанных понятий, а информацию — как небольшую часть этой структуры, т.е. ученый относит информацию к ментальной сфере [6, с. 131].

Он выразил это соотношение знаний и информации в виде следующего выражения, названного им «фундаментальным уравнением»:

$$K[S] + \Delta I = K[S + \Delta S].$$

Это уравнение выражает в абстрактной символической форме идею, что структура знаний  $K[S]$  под влиянием информации  $\Delta I$  меняется на новую модифицированную структуру  $K[S + \Delta S]$ , где  $\Delta S$  обозначает эффект модификации структуры знаний.

Брукс акцентирует внимание на следующем свойстве своего уравнения: «Фундаментальное уравнение информационной науки также подчеркивает, что определенная таким образом информация не является идентичной с философскими данными-ощущениями. Конечно, информация может зависеть от сенсорно воспринимаемых результатов наблюдений, но данные-ощущения, полученные таким образом, должны быть субъективно интерпретированы структурой знаний, чтобы стать информацией. (...) Уравнение также имеет своей целью вывести факт, что рост знаний не просто увеличивается. Поглощение информации структурой знаний может повлечь за собой не просто добавление, а даже такое уточнение структуры, как изменение системы отношений, связывающих понятия. В науках прирост информации иногда приводил к серьезной перестройке структуры знаний» [6, с. 131].

Отсюда следует, что Брукс допускает возможность членения знаний на «кванты», которые он называет «информацией». Кроме того, он допускает эволюцию знаний во времени и не считает знания человека в разные моменты времени самоидентифицируемыми. Однако у Брукса остается открытым вопрос о том, как этапы эволюции системы

знаний связаны со способом (способами) членения системы знаний.

Ответ на вопрос о способе (способах) членения системы знаний играет существенную роль при разработке ИКТ. Покажем это на примере формулировки второй проблемы, упомянутой в разд. 1 обзора, а именно: «разработка новых подходов к представлению и сохранению знаний, ориентированных на долговременный и безотказный к ним доступ в условиях *локальной генерации отдельных «квантов» знаний, их интеграции, а также глобального использования систем представления и сохранения знаний с учетом контекста и временной эволюции систем при условии обеспечения долговременной устойчивости систем представления и сохранения знаний в ситуации многообразия их использования и эволюции семантики во времени.*»

Если парадигма единой области знаний допускает только один способ членения знаний, например с помощью одного естественного языка, то на этапе интеграции результатов любых локальных процессов генерации «квантов» знаний разработчик ИКТ будет иметь дело, в терминах лингвистики и семиотики, только с одним планом выражения и одним, соответствующим ему, планом содержания. Однако такая парадигма может быть использована при разработке только моноязычных ИКТ.

Если допускается использование нескольких способов членения знаний, например с помощью разных естественных языков, то на этапе интеграции результатов любых локальных процессов генерации «квантов» знаний разработчик ИКТ будет иметь дело одновременно с несколькими планами выражения и содержания. Важно отметить, что разные способы членения знаний объединяет только сама их возможность устанавливать соответствие между множеством «квантов» знаний и множеством информационных объектов ИКТ, например слов естественного языка.

В остальном каждый способ «определяет» свои правила членения знаний и соответствия между этими двумя множествами. Поэтому «объемы значений» практически любых соотносимых в двуязычных словарях пар слов не совпадают [19]. При этом правила членения и соответствия могут изменяться во времени, и со временем могут изменяться «объемы значений» слов в каждом естественном языке.

Одновременный учет нескольких способов членения знаний необходим при разработке мультиязычных ИКТ и интеллектуальных систем, так как это позволяет соотносить «объемы значений» слов разных естественных языков, включая описание тех случаев, когда в одних языках отсутствуют слова для отдельных «квантов» знаний, выражаемых одним

- 0 — Общий отдел. Наука и знание. Информация. Документация. Библиотечное дело. Организации. Публикации в целом.
- 00 — Общие вопросы науки и культуры. Пропедевтика.
- 004 — Информационные технологии. Вычислительная техника. Теория. Технология и применение вычислительных машин и систем.
- 004.8 — Искусственный интеллект.
- 004.82 — Представление знаний.
- 004.822 — Сети знаний. Семантические сети.
- 004.823 — Фреймовые системы. Фреймы. Схемы. Сценарии.
- 004.824 — Множественные миры.
- 004.825 — Порождающие системы. Системы правил вывода.
- 004.826 — Модель черной доски.
- 004.827 — Представление неоднозначности. Неопределенность. Пробелы знаний.

**Рис. 1** Отношения концепта «Представление знаний» с другими концептами УДК

словом в других языках. Например, в английском языке есть пять основных значений существительного *office* [20]:

- (1) место выполнения профессиональных обязанностей;
- (2) функция или должностные обязанности (*it is my office to open the mail* — в мои обязанности входит вскрывать почту);
- (3) должность в организации или позиция в иерархии институциональной системы;
- (4) ведомство, министерство, управление, организация или административная единица (*Foreign Office* — Министерство иностранных дел);
- (5) сотрудники организации или ее подразделения (*the whole office was late the morning of the blizzard* — из-за снежной бури утром целый офис опоздал на работу).

В этом примере английскому слову *office* соответствуют пять «квантов» знаний как основных значений этого слова в английском языке. Однако в других языках для представления некоторых «квантов» знаний необходимые слова могут отсутствовать. Например, как отмечается в материалах проекта EuroWordNet [20], в испанском и итальянском языках нет слов для выражения пятого значения в приведенном примере для слова *office*.

Таким образом, вопрос использования одного или нескольких способов членения знаний является ключевым при описании любого варианта парадигмы единой области знаний, охватывающей предметные области компьютерной и информационной наук.

В примерах, иллюстрирующих актуальность вопроса использования одного или нескольких способов членения знаний, в качестве конкретных

способов упоминались только естественные языки. Говоря словами Р. Барта: «язык есть область артикуляции, а смысл в первую очередь есть результат членения» [21, с. 277]. Однако при разработке ИКТ и интеллектуальных систем могут использоваться и другие способы членения знаний, основанные на той или иной системе классификации, а также такие искусственные способы членения знаний, как семантические словари, тезаурусы и онтологии электронных библиотек.

Для иллюстрации членения знаний на основе системы классификации рассмотрим фрагмент Универсальной Десятичной Классификации (УДК) на рис. 1 и отношения концепта «Представление знаний» с другими концептами этого фрагмента. В УДК концепт «Представление знаний» располагается на пятом уровне десятичной иерархии, имеет номер 004.82 и включает в себя 6 концептов шестого уровня с номерами от 004.822 до 004.827. Отметим, что все эти 6 концептов в УДК являются терминальными, т. е. дальнейшее их членение в УДК отсутствует.

Концепт «Представление знаний» включен в концепт «Искусственный интеллект» четвертого уровня, который, в свою очередь, включен в концепт «Информационные технологии. Вычислительная техника. Теория. Технология и применение вычислительных машин и систем» третьего уровня и т. д. до концепта «Общий отдел. Наука и знание. Информация. Документация. Библиотечное дело. Организации. Публикации в целом» самого верхнего уровня.

В заключение отметим важный терминологический аспект. Далее определение информации по Бруксу будем называть «ментальной информацией». При употреблении этого словосочетания будем считать, что в информационной науке допускается членение знаний на «кванты» и учитывается воз-

возможность эволюции системы знаний человека во времени.

## 2.4 Хьорланд: информационные процессы

Для того чтобы очертить предметную область информационной науки, Биргер Хьорланд приводит определение, которое сформулировало Американское общество информационной науки (American Society for Information Science — ASIS): «Предметом исследований в информационной науке являются процессы генерации, накопления, организации, интерпретации, хранения, поиска, распространения, преобразования и использования информации, а также применение современных технологий в перечисленных процессах. В качестве дисциплины она стремится создать и структурировать научные и технологические знания, имеющие отношение к передаче информации. Информационная наука включает теоретические направления исследований безотносительно к конкретным приложениям и прикладным исследованиям, ориентированным на создание продуктов и оказание услуг» [22, с. 509].

Анализируя это определение, Хьорланд отмечает явление, которое он называет «понятийным хаосом» в информационной науке, проявляющийся, в частности, в том, что в научной литературе насчитывается несколько сотен определений этой науки. Что касается определения, данного ASIS, то, по мнению Хьорланда, в этом определении в явном виде не выражена *соотнесенность информации со знаниями*. Он отмечает, что в информационной науке гораздо более распространенным является исследование информации посредством рассмотрения процесса модификации знаний ее получателя (ментальная информация Брукса) или в качестве снятия неопределенности (статистическая информация, которая рассматривается в следующем разделе обзора).

С точки зрения проблемы конвергенции в этом определении важно отметить изолированность информационных процессов от применяемых технологий. Десять перечисленных процессов — от генерации до передачи информации — являются предметом исследований в информационной науке, а по отношению к технологиям, в том числе компьютерным, предметом исследований является только их применение в этих процессах.

Это определение отражает достаточно длительный исторический этап использования тех методов обработки информации, которые не зависели от цифровой среды по причине ее полного отсутствия.

С появлением цифровой среды стал актуальным вопрос применения ИКТ в перечисленных процессах. Со временем ИКТ стали оказывать влияние на создание новых методов обработки информации, возможности которых стали во многом определяться характеристиками и свойствами цифровой среды и, наоборот, на ее развитие оказывают влияние новые методы обработки информации.

В наше время подобное взаимное влияние становится *одной из ключевых предпосылок конвергенции информационной и компьютерной наук*. Например, современные методы представления знаний в медицине и процессы обработки медицинских данных, генерируемых компьютерными томографами, достаточно сильно зависят от возможностей ИКТ и часто неразрывно интегрированы с ними [23].

## 3 Компьютерная наука

В начале обзора уже говорилось, что в английском языке существует как минимум четыре устойчивых словосочетания — *information science*, *computer science*, *computer and information science*, *computational science*, — которые нередко переводятся на русский язык одним словом «информатика». Кроме того, необходимо учитывать, что в английском языке смысл рассматриваемого в этом разделе термина “*computer science*” — «компьютерная наука» — за последние 40 лет существенно изменился.

Например, проект *Computing Curricula*, в рамках которого были подготовлены «Рекомендации по преподаванию программной инженерии и компьютерной науки в университетах», ведет свой отсчет с 1968 г., когда была опубликована первая версия рекомендаций. С тех пор эти рекомендации обновлялись примерно раз в десять лет совместным комитетом по образованию под эгидой профессиональных ассоциаций *Association for Computing Machinery (ACM)* и *IEEE Computer Society*. В конце 1990-х гг. стало ясно, что область знаний, связанная с ИКТ, очень сильно разрослась и ее трудно, если вообще возможно, полностью осветить в рамках одного университетского курса. В связи с этим было принято решение о его разделении на четыре основные специальности — *computer science* (компьютерная наука), *software engineering* (программная инженерия), *hardware engineering* (проектирование аппаратных платформ) и *information systems* (информационные системы) [24, с. 5].

После этого деления предметная область компьютерной науки в проекте *Computing Curricula* стала включать следующие 14 разделов: дискретные структуры, основы программирования, алгоритмы

и теория сложности, архитектура и организация ЭВМ, операционные системы, распределенные вычисления, языки программирования, взаимодействие человека и машины, графика и визуализация, интеллектуальные системы, управление информацией, социальные и профессиональные вопросы программирования, программная инженерия, методы вычислений [24, с. 193].

Естественно, что эти четыре специальности тематически частично пересекаются. Например, тему «Дискретные структуры» изучают в рамках специальности «Программная инженерия», а ряд тем «Программной инженерии» преподается для студентов специальности «Компьютерная наука».

Значительное число основополагающих работ, составляющих теоретические основы компьютерной науки, относится к первой половине XX в. Среди ученых, которые глубоко исследовали теоретические вопросы, связанные с алгоритмами и их возможностями, были Клини, Черч, Тьюринг и Пост. Неформальное понятие алгоритма для решения некоторого класса задач подразумевает некоторый набор правил, с помощью которых решение любой указанной задачи этого класса может быть найдено в случае выполнения этого набора правил. Так подходит к определению алгоритма Г. Эббингаус в своей статье «Машины Тьюринга и вычислимые функции I. Уточнение понятия алгоритма» [25, с. 9–11].

Вплоть до 30-х гг. прошлого столетия понятие алгоритма оставалось интуитивно понятным, имевшим скорее методологическое описание, нежели математическое определение. В истории науки известно много ярких примеров алгоритмов. Среди них алгоритм Евклида нахождения наибольшего общего делителя двух натуральных чисел или двух целочисленных многочленов, алгоритм Гаусса решения системы линейных уравнений, алгоритм разложения многочлена одной переменной на неприводимые множители. Перечисленные алгоритмы позволяли решать задачи путем указания и выполнения конкретных процедур. Для решения подобных задач было достаточно интуитивного понимания алгоритма.

Однако в начале XX в. был сформулирован ряд алгоритмических проблем, решение которых потребовало разработки и применения новых логических средств. Это связано с тем, что доказательство существования и разработку разрешающего алгоритма можно осуществить и с помощью интуитивного понимания алгоритма. Если же требуется доказать, что для решения задачи не существует

алгоритма, то в этом случае необходимо точное определение того, что такое алгоритм.

Определение алгоритма было предложено в первой половине XX в. в двух формах: на основе понятия рекурсивной функции и на основе описания процесса, осуществимого на абстрактной машине. Был сформулирован тезис («тезис Тьюринга»), утверждающий, что любой алгоритм может быть реализован на соответствующей машине Тьюринга. Оба подхода, а также другие подходы (Маркова и Поста) привели к одному и тому же классу алгоритмически вычислимых функций и подтвердили целесообразность использования тезиса Тьюринга для решения алгоритмических проблем.

В настоящее время теория алгоритмов является краеугольным камнем фундамента компьютерной науки. С ее помощью были уточнены такие понятия, как доказуемость, эффективность, разрешимость, перечислимость и другие.

В этом разделе обзора сначала будут рассмотрены отдельные вопросы становления компьютерной науки, а затем в ее предметной области будут обозначены те тематические направления, которые являются ключевыми для конвергенции компьютерной и информационной наук.

### 3.1 Машина Тьюринга и универсальные вычислительные машины

С момента создания первой универсальной вычислительной машины ENIAC (Electronical Numerical Integrator and Computer), которую в момент ее создания называли математическим роботом<sup>1</sup>, компьютерная наука была тесно связана с проектами разработки компьютеров, которые в первую очередь были ориентированы на решение вычислительных задач.

Вычислительную машину фон Нейман определял как «устройство, которое может выполнять команды для вычислений значительной сложности». Центральным моментом в работе фон Неймана была формулировка требований к структуре вычислительной машины. В них фон Нейман фактически описал структурную схему аналитической машины Бэббиджа, состоящей (в современной терминологии) из арифметического устройства (АУ), устройства управления (УУ), оперативного запоминающего устройства (ОЗУ), внешнего запоминающего устройства (ВЗУ), устройств для ввода и вывода информации (УВВ). Однако в архитектуре фон Неймана были существенные отличия от схемы аналитической машины Бэббиджа. В частности, фон

<sup>1</sup>Словосочетание «математический робот» использовалось в первом пресс-релизе о компьютере ENIAC от 16 февраля 1946 г. (см. <http://www.americanhistory.si.edu/csr/comphist/pr1.pdf>). Разработка этого компьютера была начата в июле 1943 г. и завершена осенью 1945 г.

Нейман рекомендовал использовать в устройстве не механическую, а электронную элементную базу и не десятичную, а двоичную систему исчисления. Бэббидж искал аналогии между блоками вычислительной машины и структурными производственными единицами (мельница, склад), в то время как фон Нейман находил аналогии в живом организме (точнее, в нейронных сетях) [26].

В процитированной работе [26] проводится сравнение универсальной вычислительной машины с «фон-неймановской архитектурой» и «универсальной машины Тьюринга» (Universal Turing Machine — UTM). Английский математик Алан Тьюринг предложил использовать UTM для исследования «проблемы разрешимости» (the Hilbert Entscheidungsproblem), которую сформулировал в 1900 г. немецкий математик Гильберт [27].

В 1936 г. Тьюринг доказал, что эта проблема не имеет решения, и опубликовал полученные результаты в статье «О вычислимых числах применительно к проблеме разрешимости» [28]. Важно то, что для этого доказательства Тьюринг использовал предложенную им гипотетическую машину UTM.

Краткое описание одного из вариантов UTM приводит Эббинхаус в своей статье: «Машина Тьюринга Т состоит из операционного исполнительного устройства, которое может находиться в одном из дискретных состояний  $q_0, \dots, q_s$ , принадлежащих некоторой конечной совокупности, комбинированной читающей и пишущей головки, счетной ленты  $\langle \dots \rangle$  и лентопротяжного механизма. При этом  $q_0$  называется начальным состоянием [машины] Т. Ячейки ленты пронумерованы, начиная с крайней левой, числами 0, 1, 2, . . . Читающая и пишущая головка находится в каждый данный момент времени над некоторой ячейкой ленты — текущей рабочей ячейкой. С помощью лентопротяжного механизма одна из ячеек, соседняя с рабочей ячейкой, может быть помещена под читающей и пишущей головкой; в таком случае мы будем говорить, что рабочая ячейка сдвинулась на одну ячейку вправо или влево. Читающая и пишущая головка может читать буквы алфавита  $A = \{a_1, \dots, a_t\}$  и букву  $a_0$  (у Тьюринга этот символ означает «пусто», то есть отсутствие информации в ячейке), стирать их и печатать;  $A$  называется рабочим алфавитом [машины] Т. . . » [25, с. 24–25].

Тьюринг показал, что его машина «за данный большой, но конечный промежуток времени способна справиться с любым вычислением, которое сможет выполнить всякая сколь угодно мощная вычислительная машина» [26]. Далее в цитируемой работе сравнивается концепция «фон-неймановской архитектуры» со свойствами UTM:

- Тьюринг фактически впервые выдвинул концепцию вычислительной машины с хранимыми в памяти командами (программой);
- поскольку операции UTM на каждом такте зависят, в частности, от результата последнего действия, можно говорить, что машина выполняет команду условного перехода;
- Тьюринг показал, что результатом работы машины может быть группа символов, которые, будучи введены в другую UTM, заставят ее действовать так же, как первую; иными словами, машина может «генерировать» или видоизменять программу, и Тьюринг понимал, что это ее свойство должно быть перенесено на реальную ЭВМ;
- любую универсальную вычислительную машину можно запрограммировать так, что она будет моделировать работу некоей специализированной машины (создатели ENIAC делали то же самое, когда настраивали свою универсальную машину на решение конкретной задачи) [26].

В описании UTM содержится существенное ограничение: все вычисления выполняются на одномерной ленте, то есть допускается только линейная конкатенация символов алфавита UTM при построении символьных выражений. При этом в работе Тьюринга есть замечание о том, что в математике используются двумерные символьные выражения, которые всегда могут быть преобразованы в одномерные, что позволяет обрабатывать их с помощью UTM. Говоря о размещении символов алфавита UTM в ячейках одномерной ленты, Тьюринг отмечает, что в одной ячейке может располагаться линейная последовательность символов, трактуемая как единый сложный символ, и проводит аналогию между сложными символами и словами европейских языков. Другими словами, при обработке с помощью UTM эти слова предлагается трактовать как сложные символы, размещаемые в ячейках одномерной ленты [28].

### 3.2 Статистическая информация

Параллельно с развитием понятия алгоритма и созданием универсальной вычислительной машины формировалось еще одно направление исследований, связанное с именами Норберта Винера и Клода Шеннона. В построении теоретического фундамента этого направления использовались статистические и вероятностные методы.

Норберт Винер в гл. 3 своей книги [29, с. 119] о предметной области статистической науки говорит следующим образом: «Существует широкий класс явлений, в которых объектом наблюдения

служит какая-либо числовая величина или последовательность числовых величин, распределенных во времени. Температура, непрерывно записываемая самопишущим термометром; курс акций на бирже в конце каждого дня; сводка метеорологических данных, ежедневно публикуемая бюро погоды, — все это временные ряды, непрерывные или дискретные, одномерные или многомерные. Эти временные ряды меняются сравнительно медленно, и их вполне можно обрабатывать посредством вычислений вручную или при помощи обыкновенных вычислительных приборов, таких как счетные линейки и арифмометры. Их изучение относится к обычным разделам статистической науки».

О статистическом характере исследуемых объектов и явлений Винер пишет далее: «Все эти временные ряды и все устройства, работающие с ними, будь то в вычислительном бюро или в телефонной схеме, связаны с записью, хранением, передачей и использованием информации. Что же представляет собой эта информация, и как она измеряется? Одной из простейших, наиболее элементарных форм информации является запись выбора между двумя равновероятными, простыми альтернативами, например между орлом и решкой при бросании монеты. Мы будем называть решением однократный выбор такого рода. Чтобы оценить теперь количество информации, получаемое при совершенно точном измерении величины, которая заключена между известными пределами  $A$  и  $B$  и может находиться с равномерной априорной вероятностью где угодно в этом интервале, положим  $A = 0$ ,  $B = 1$  и представим нашу величину в двоичной системе бесконечной двоичной дробью  $0, a_1 a_2 a_3 \dots a_n \dots$ , где каждое  $a_1, a_2, a_3, \dots, a_n$  имеет значение 0 или 1» [29, с. 120].

Для Винера центральным было понятие информации как термина статистической науки. В статистической науке, которая является смежной с компьютерной наукой, слово «информация» включено в единую систему терминов со словами «шум», «помеха», «вероятность», «энтропия» и т. д., но не со словом «знания», как это сделано в информационной науке. Поэтому далее информацию статистической науки будем называть статистической информацией, которая понимается совсем не в том смысле, который вкладывается в словосочетания «языковая информация» по Фаррадейну, «знаковая информация» по Ингверсену и «ментальная информация» по Бруксу в разд. 2 данного обзора.

В качестве меры статистической информации в совокупности сообщений Шеннон предлагает использовать логарифмическую функцию: «Если ряд сообщений из множества сообщений конечен, то этот ряд или его произвольную монотонную функ-

цию можно рассматривать как меру информации, произведенной на тот момент, когда выбирается одно сообщение из множества сообщений, причем все выборы равнозначны. Как заметил Хартли, наиболее естественным выбором является логарифмическая функция. Хотя это определение должно быть подвергнуто значительному обобщению, когда мы рассматриваем влияние статистики сообщения и когда мы располагаем постоянным диапазоном сообщений, в любом случае мы будем использовать, по сути, логарифмическую меру» [30].

Выбор основания логарифма соответствует выбору единицы измерения информации. Если используется основание 2, то результирующие единицы можно назвать двоичными числами или более кратко битами, термином, который предложил Дж. Теки. Устройство с двумя устойчивыми состояниями, такое как реле или триггер, может хранить один бит информации;  $N$  таких устройств могут хранить  $n$  бит, суммарное число возможных состояний равно  $2^n$ , а  $\log_2 2^n = n$  [30].

Шеннон подчеркивает, что, говоря о статистической информации, ее передаче и соответствующих системах передачи, он имеет в виду именно физическую модель передачи сообщений, а содержательная сторона передаваемого сообщения при этом не рассматривается. Иначе говоря, по смыслу «информация» статистической науки принципиально отличается от «информации» информационной науки. Таким образом, наблюдается только *совпадение последовательности 10 литер: «информация», но отличаются «кванты» знаний, соответствующие этим литерам в статистической и информационной науках.*

Отметим, что теоретические исследования Шеннона и Винера были необходимы для практических приложений и результаты этих исследований широко использовались. Например, Шеннон исследовал природу передаваемой по каналу связи статистической информации и вопросы оптимизации этого процесса. Теория связи была востребована во время второй мировой войны, а затем получила развитие в многочисленных приложениях в мирное время. Винер, в свою очередь, полагал, что методы обработки статистической информации могут коренным образом изменить взгляд на использование вычислительных машин. Он считал, что «при применении этих машин становится все более очевидным, что они требуют специальных математических методов, совершенно отличных от тех, к которым прибегали в ручных расчетах или на малых машинах» [29, с. 206].

Отметим, что дальнейшее развитие методов статистической науки, востребованное в кибернетике, компьютерной науке и теории автоматическо-

го управления, нашло свое отражение в работах Р. Е. Калмана, В. С. Пугачева и ряда других исследователей [31–33]. Разработка специальных математических методов, ориентированных на применение ЭВМ и совершенно отличных от тех, к которым прибегали в ручных расчетах, является примером того, что ИКТ оказывают влияние на создание новых методов в широком спектре областей знаний.

### 3.3 Представление знаний и управление информацией в компьютерной науке

Рассмотренные ранее в этом разделе вопросы, относящиеся к предметной области компьютерной науки, характеризуют исторические аспекты ее становления, но не дают представление о тематическом содержании компьютерной науки в ее современном понимании и о влиянии этой науки на создание новых поколений ИКТ.

Значительно более полное представление о ее предметной области дает перечень из 14 позиций, приведенный в начале этого раздела, от дискретных структур (1-я позиция) до методов вычислений (14-я позиция). Каждая из 14 позиций достаточно подробно описана в Рекомендациях по преподаванию компьютерной науки в университетах [24]. При их описании используются два уровня детализации для каждой позиции.

Выделим в детализированных описаниях формулировки вопросов представления знаний и управления информацией, которые являются ключевыми одновременно для компьютерной и информационной наук. Начнем с позиции «Языки программирования», которая на первом уровне детализации включает 11 пунктов, в том числе семантику языков программирования [24, с. 197]. Для этого пункта рассмотрим второй уровень детализации, который включает 5 следующих тем: неформальная семантика, обзор формальной семантики, денотационная семантика, аксиоматическая семантика и операционная семантика [24, с. 329].

Таким образом, в программу изучения компьютерной науки включена тема «Неформальная семантика», относящаяся и к информационной науке. При этом важно отметить, что это не единственный пример подобного тематического пересечения. Отметим еще две позиции в исходном списке, в которых наблюдается аналогичное явление тематического пересечения — «Интеллектуальные системы» и «Управление информацией».

Позиция «Интеллектуальные системы» на первом уровне детализации включает 10 пунктов: основные вопросы (связанные с интеллектуальными

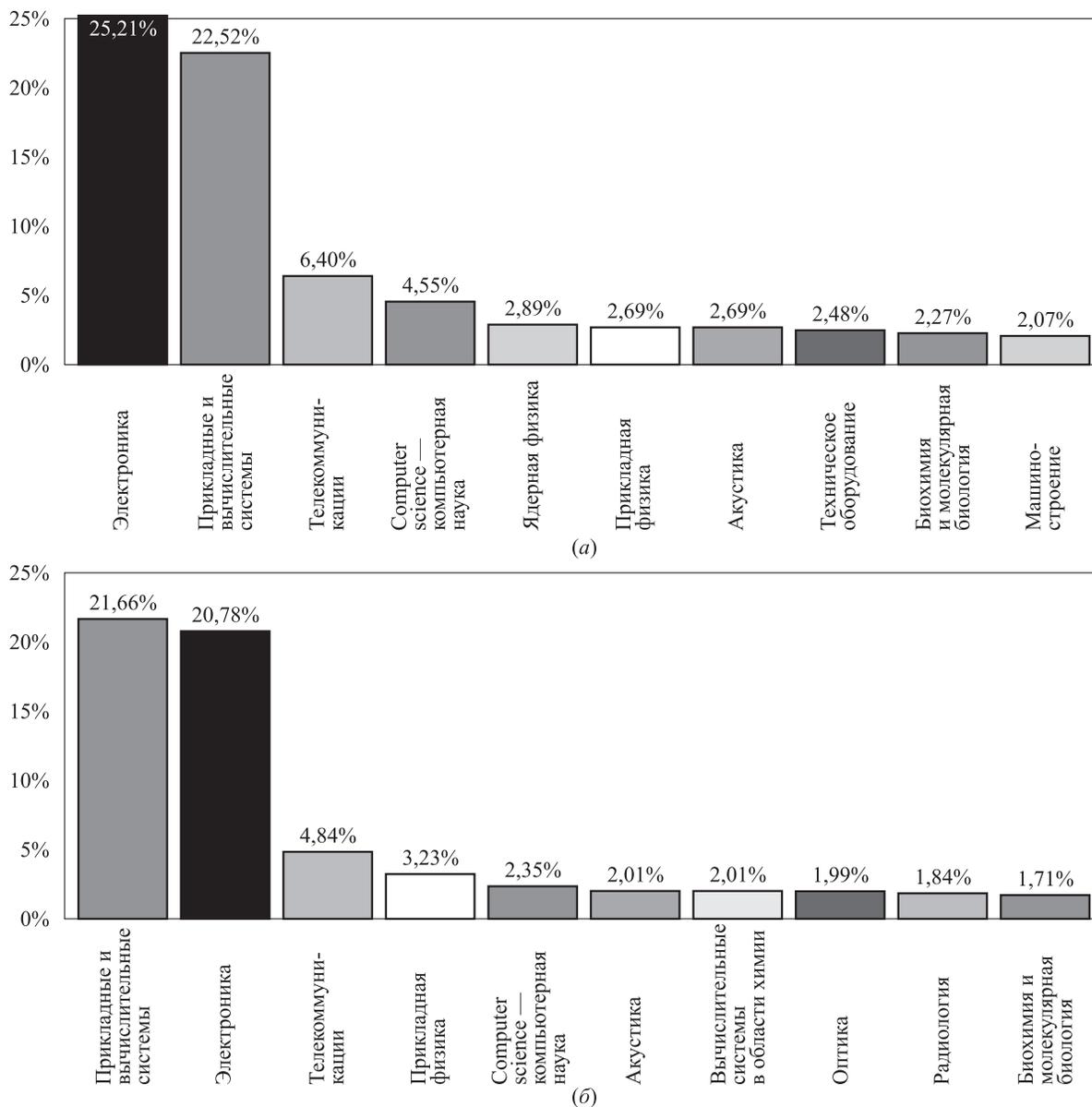
системами), поиск решений, методы представления знаний, углубленное изучение методов поиска, углубленное изучение методов представления знаний, агенты, обработка текстов на естественном языке, методы обучения компьютеров и нейронные сети, системы искусственного интеллекта с планируемым поведением, робототехника [24, с. 342–347].

С точки зрения процесса конвергенции привлекает внимание пункт с методами представления знаний, включающий на втором уровне детализации тему «*Структурное представление знаний*», а также пункт обработки текстов на естественном языке, включающий следующие две темы: «*Методы, основанные на совокупности текстов*» и «*Информационный поиск*» [24, с. 342–347], которые также являются примерами тематического пересечения информационной и компьютерной наук.

Позиция «Управление информацией» на первом уровне детализации включает 14 пунктов: информационные модели и системы, системы баз данных, моделирование данных, реляционные базы данных, языки запросов к базам данных, проектирование реляционных баз данных, обработка транзакций, распределенные базы данных, проектирование физической структуры базы данных, извлечение информации, хранение и поиск информации, гипертекст и гипермедиа, мультимедийная информация и системы мультимедиа, цифровые библиотеки. С точки зрения процесса конвергенции привлекают внимание вопросы хранения, извлечения и поиска информации, а также цифровые библиотеки. Наиболее явно тематическое пересечение информационной и компьютерной наук проявляется на втором уровне детализации пункта «Хранение и поиск информации», который включает следующие темы [24, с. 342–347]:

- *информационные потребности пользователя*, релевантность и оценка эффективности поиска;
- тезаурус, онтология, классификация и категоризация;
- *библиографическая информация и библиометрия*;
- *резюмирование* и визуализация информации;
- *интеграция ключевых слов и схем классификации*.

Следовательно, в тематических разделах, характеризующих предметную область компьютерной науки в ее современном понимании, наблюдается пересечение с информационной наукой в первую очередь по вопросам представления знаний и управления информацией, которые являются ключевыми одновременно для компьютерной и информационной наук.



**Рис. 2** Количественные взаимосвязи направлений научных исследований с ИТ для европейских патентов (а) и для патентов США (б)

С одной стороны, наблюдаемое тематическое пересечение, естественное для любой области знаний, в процентном отношении от общего числа тем второго уровня детализации не превышает нескольких процентов. С другой стороны, если необходимо анализировать наблюдаемое тематическое пересечение с позиций потребностей в разработке научных основ создания новых поколений ИКТ, то нельзя ограничиваться только процентным отношением от общего числа тем.

Для подобного анализа необходимо использовать количественные меры (индикаторы) связей областей технологического развития и тех направ-

лений научных исследований, результаты которых используются в процессе создания новых технологий [34]. Знание значений этих индикаторов для направлений компьютерной науки позволило бы оценить их влияние на развитие ИКТ.

Оценка влияния отдельных направлений компьютерной науки требует учета как минимум первого уровня детализации для каждой из 14 позиций в тематическом делении этой науки. Однако в настоящее время отсутствуют индикаторные оценки влияния отдельных направлений компьютерной науки как для первого уровня детализации ее предметной области, так и для каждой из 14 позиций в целом.

Имеются сведения о взаимосвязях информационных технологий (ИТ) только со всей компьютерной наукой в целом и рядом других областей знаний. На рис. 2 приведены диаграммы, которые иллюстрируют индикаторы связей ИТ с теоретическими и прикладными дисциплинами, включая компьютерную науку. Первая диаграмма (рис. 2, а) построена на основе данных Европейского патентного ведомства, вторая диаграмма (рис. 2, б) — на основе данных Патентного ведомства США<sup>1</sup>.

На каждой из диаграмм приведены процентные отношения для 10 теоретических и прикладных дисциплин за период 1992–1996 гг. На рис. 2, а для компьютерной науки указано процентное отношение 4,55%. Это число было определено следующим образом: из европейских патентов за период 1992–1996 гг., относящихся к сфере ИТ, были извлечены все ссылки на научные публикации, которые затем были распределены по направлениям теоретических и прикладных исследований. Число 4,55% — это доля научных публикаций по компьютерной науке от общего числа публикаций по всем теоретическим и прикладным дисциплинам, на которые есть ссылки в европейских патентах по ИТ. На обеих диаграммах указаны только первые 10 дисциплин с наибольшим числом публикаций. Отметим, что информационная наука не попала в первую десятку научных дисциплин на этих диаграммах.

Приведенные диаграммы иллюстрируют существенные отличия в цитировании научных публикаций для разных теоретических и прикладных дисциплин. Существуют отличия и в региональном разрезе, например доля научных публикаций по компьютерной науке, на которые есть ссылки в патентах США по ИТ, равна 2,35%, что на 2,2% меньше, чем доля научных публикаций по компьютерной науке, на которые есть ссылки в европейских патентах по ИТ (см. рис. 2).

Существуют аналогичные диаграммы и для телекоммуникационных технологий. Однако данные имеющихся диаграмм не позволяют количественно оценить взаимосвязи ИКТ с тематическим пересечением предметных областей информационной и компьютерной наук с позиций потребностей в разработке научных основ создания ИКТ, так как в этих данных компьютерная наука рассматривается как единое целое, без деления на составляющие ее направления.

В настоящее время можно говорить только о *существенных отличиях в доле цитирования в патентах научных публикаций* из разных теоретических и прикладных дисциплин. Поэтому, скорее

всего, нельзя ограничиваться только процентным отношением тематического пересечения информационной и компьютерной наук к общему числу тем последней в задачах оценки роли этого пересечения для создания новых поколений ИКТ.

## 4 Информационно-компьютерная наука

Материал разд. 2 и 3 позволяет предположить, что одновременно с формированием и институционализацией информационной и компьютерной наук как самостоятельных научных дисциплин и областей применения их результатов наблюдалось развитие отдельных предпосылок их конвергенции. Отметим, что многоаспектное исследование проблемы конвергенции началось более 40 лет назад.

Термин «информационно-компьютерная наука», который вынесен в название этого раздела, одним из первых использовал американский ученый С. Горн в 1963 г., с той разницей, что тогда этот термин употреблялся во множественном числе. Единственное число использовалось ученым начиная с 1983 г. [12]. Однако уже в 1963 г. потенциальный результат конвергенции информационной и компьютерной наук позиционировался С. Горном как новая фундаментальная область знаний, что нашло отражение в самом названии его работы — “a new basic discipline” [1].

В течение двадцати лет им была опубликована серия статей о предметной области и методологии информационно-компьютерной науки [1, 12, 35–37]. В качестве смежных дисциплин С. Горн называет библиотекведение, теорию информационного поиска, информационную науку, кибернетику, когнитивную психологию, искусственный интеллект, семиотику, лингвистику и математику. Среди сфер применения результатов этой науки он выделяет разработку компьютеров, менеджмент и сферу образования [12, с. 121].

В 1963 г. Горн предпринял попытку перечислить вопросы, изучаемые информационно-компьютерной наукой: «Примерами основных вопросов исследования в этой области могут быть системы программирования, проектирование компьютерных систем, искусственный интеллект, информационный поиск и т. д. Вероятностная информационная теория Шеннона определенно принадлежит к этой области знания, но помимо нее существует еще теория информации искусственных языков и

<sup>1</sup>Эти диаграммы взяты из научно-технического отчета ‘Linking Science to Technology — Bibliographic References in Patents’, подготовленного на основе результатов проектов, выполненных по контракту ERBHPV2-СТ-1993-03 в рамках 5-й и 6-й Рамочных программ ЕС.

ее обработки, которую также необходимо включить в предметную область этой науки. Одним из центральных вопросов этой новой дисциплины, скорее всего, станет синтез и анализ искусственных языков и их процессоров» [1, с. 150].

В отличие от авторов аналитического доклада [10], Горн в явном виде включает в перечень направлений, изучаемых информационно-компьютерной наукой, искусственный интеллект, информационный поиск, синтез и анализ искусственных языков. После перечисления этих вопросов, информационно-компьютерная наука далее рассматривается им уже как учебная дисциплина и говорится о необходимости описать способы различения новой области знаний от соседних с ней областей в учебном процессе. К примеру, каким образом абитуриент может узнать, относится ли сфера его интересов именно к этой новой области знаний, а не к одной из уже устоявшихся дисциплин? Какое ему необходимо образование для того, чтобы углубиться в эту новую область? И в чем результат его обучения существенным образом будет отличаться от того образования, которое потребовалось бы ему в другой области?

В 60-х гг. прошлого века ощущалась потребность в оценке перспектив развития этой новой дисциплины, в ее позиционировании среди существовавших уже тогда областей знаний и учебных дисциплин. Горн рассматривает эти вопросы, отталкиваясь в своих рассуждениях от профессиональных интересов ученых в этой области еще на стадии получения ими образования:

«Информационно-компьютерная наука рассматривает прагматические аспекты использования символов их пользователями и интерпретаторами в качестве еще одного центрального вопроса таким же образом, как эти аспекты должны исследоваться специалистами в области лингвистики, психологии, философии и инженерных наук.

Таким образом, студент, изучающий численный анализ, в процессе разработки или анализа какого-либо алгоритма мыслит себя как математик, если его единственный интерес заключается в доказательстве существования алгоритма или определения его точности. Но он является специалистом в области информационно-компьютерной науки, если рассматривает этот алгоритм прагматически, например с точки зрения его реализации (обработки процессором), и интересуется эффективностью его работы, временными затратами, распределением памяти и т. д.

Аналогично студент, изучающий процедуру адаптивного управления, описывающую поведение животного в некоторой ситуации, позиционирует себя как психолог, если его главной задачей

является выяснение того, обладает ли он хорошей моделью поведения этого животного. Если его интересует проблема искусственного интеллекта как одного из направлений информационно-компьютерной науки, то он интересуется применимостью этой процедуры независимо от того, является ли она моделью поведения животного или не является.

Студент, занимающийся порождающей грамматикой, мыслит себя как лингвист, если его больше всего интересует, действительно ли естественный язык работает так, а не иначе. Однако он думает как ученый в области информационно-компьютерной науки, если его занимает вопрос, каким образом можно использовать эту грамматику в информационной системе. Лингвист может рассматривать механизм стековой памяти, но с глубиной не более семи из-за ограниченных возможностей локальной памяти человека, но для решения информационно-компьютерных задач такой глубины явно недостаточно» [1, с. 154]. Рассмотрев в статье 1963 г. эти примеры, Горн предлагает перечень тех дисциплин, которые должны преподаваться студентам, изучающим информационно-компьютерную науку, включая математику, физику, философию, лингвистику, психологию, вычислительную технику и компьютерное программирование.

Предложенный подход к изучению информационно-компьютерной науки уже тогда начал реализовываться в Пенсильванском университете. Через двадцать лет, когда уже накопился большой опыт ее преподавания, Горн пишет, что его понимание концепции информационно-компьютерной науки заключается в том, что *эта область знаний не является ветвью математики, так как она должна соотноситься с прагматическими вопросами, от которых математика не должна зависеть* [12, с. 137].

Следует отметить, что процитированная статья начинается со следующей фразы: «Позвольте мне, прежде всего, выбрать более короткое название, чем *информационно-компьютерная наука*. Я выбираю термин *«информатика»*, созвучный французскому *Informatique* и немецкому *Informatik*. Он несет в себе идею информации, а оканчивается так же, как и математика, подразумевая формализованную теорию. Плохо то, что при использовании слова *«информатика»* теряется компьютерная составляющая в названии и, кроме того, оно не вызывает ассоциаций с какой-либо экспериментальной основой» [12, с. 121].

Следовательно, Горн, используя в 1983 г. термин *«информатика»*, подразумевает под ним именно информационно-компьютерную науку. Ученый обращается к истокам этой дисциплины, чтобы дать четкое определение информатике: «Все, что я до сих пор говорил о вычислениях, ориентиро-

вано на практическую деятельность и связано с компьютером. Но сама теория вычислений уже сформировалась и существовала к тому времени, когда появились цифровые компьютеры. (...) Специалисты в области символической логики уже исследовали логические пределы вычислений; была описана универсальная машина Тьюринга и доказана неразрешимость проблемы остановки; Гедель продемонстрировал пределы формализма при помощи своих теорем о неразрешимости; Черч, Клини и Карри проанализировали вычисления в теории рекурсивных функций и комбинаторной логике; Туэ и Пост, а в более позднее время Марков, рассмотрели вычисления с синтаксической точки зрения. (...) Поэтому, когда появились компьютеры, обсуждение лингвистики естественных языков Ноамом Хомским происходило в ракурсе вычислений (...). В результате этих новых разработок появились лингвистические описания процессов программирования, математическая теория автоматов и формальные языки. Эти результаты, в свою очередь, повлияли на разработки языков программирования и программируемых вычислительных машин. (...) Теперь под информатикой мы понимаем нечто, связанное с синтезом и анализом символических выражений, а также синтез и анализ процессоров, которые интерпретируют, транслируют и обрабатывают такие выражения. Если говорить более прозаично, то информатика занимается изучением, проектированием и использованием структур данных и их обработкой. . . » [12, с. 131].

Главный вывод Горна о составе и статусе новой области знания, которым он завершает статью, состоит в следующем: «(...) не следует отделять компьютерную науку от информационной науки, а следует пытаться отстаивать единую область знаний — информатику. Любая попытка поощрить такое разделение (...) повлечет за собой отделение практической деятельности от знаний, как это произошло с математикой Пифагора, риторикой софистов, метафизикой и органом Аристотеля, грамматикой стоиков, логикой и грамматикой логических позитивистов. Такое разделение будет причиной прекращения деятельного кипения, которое поддерживается сплавом знаний и практической деятельности» [12, с. 139–140].

Проиллюстрировать последствия подобного отделения практической деятельности от знаний можно было бы количественно с помощью диаграмм, подобных рис. 2, на которых доля компьютерной науки равна 4,55% для европейских патентов и 2,35% для патентов США за период 1992–1996 гг. Однако для полноты картины явно не хватает исходных данных для выявления тренда изменений этих долей во времени.

Вернемся к проблеме конвергенции. Важным этапом в развитии идеи Горна, позиционирующей информатику как единую область знаний и охватывающую предметные области компьютерной и информационной наук, были работы Ю. А. Шрейдера [38, 39]. В статье «Информация и знание» говорится, что не существует двух информатик (информационной науки и компьютерной науки), а есть два облика информатики. Первый из них (информационная наука) дополнительно нагружен представлениями о традиционном информационном обслуживании специалистов-ученых и инженеров в области их профессиональных интересов. Второй облик (компьютерная наука) неправомерно искажен чисто программистскими проблемами, не специфичными для информатики. Специфические же проблемы информатики оказываются там, где возникают задачи *информационного представления знаний в форме, удобной для обработки, передачи и творческого реконструирования знаний в результате усилий пользователя* [39, с. 51].

В этой же работе Ю. А. Шрейдер формулирует ряд положений научной парадигмы информатики: «информация есть общественное достояние, она в принципе социальна, в то время как знание, вообще говоря, соотносено с конкретной личностью, с тем, кто им владеет и непосредственно пользуется. (...) Информация должна пройти через «когнитивный экран» тех, для кого она представляет ценность. Так возникает необходимость считаться не только с существованием мира объективированного социализированного знания, т.е. информации как превращенной формы знания, но и с феноменом личностного знания. (...) Тождественность информации и знания при этом исключается, но информация как превращенная форма знания сохраняет следы своего происхождения. (...) наиболее принципиальные вопросы информатики всегда возникали на стыке информации и знания, там, где речь шла о превращении одного в другое». Далее Шрейдер пишет о пропасти, разделяющей *информацию и знания как сущности разной природы* [39, с. 50–51].

## 5 Заключение

Обзор многолетней истории проблемы конвергенции информационной и компьютерной наук позволяет сделать следующие выводы.

Во-первых, приведенные положения из работ Горна и Шрейдера являются ключевыми для описания научной парадигмы информатики как информационно-компьютерной науки, но не включают всех необходимых ее составляющих, в том числе *аксиоматику, классификацию объектов,*

процессов и явлений этой области знаний, а также систему терминов.

Во-вторых, процессы понимания, осознания и экспликации знаний в настоящее время по-прежнему остаются во многом невыясненными. В информационной науке они исследуются как когнитивные и креативные процессы, с которыми неразрывно связаны процессы генерации информации, социальных коммуникаций и понимания информации. В этой науке знания человека (в том числе ментальная информация Брукса) и информация (знаковая информация Ингверсена и языковая информация Фэррадейна) соотносятся между собой как *сущности разной природы* [6, 13, 16].

В-третьих, в компьютерной науке в качестве базовых понятий используются, как правило, «символы абстрактного алфавита», в явном виде не соотношенные со знаниями человека и ментальной информацией Брукса, а также со знаковой информацией Ингверсена и языковой информацией Фэррадейна. Например, в классической работе Тьюринга слова «знания» и «информация» не используются, а рассматриваются лишь линейные символьные выражения. Однако в этой работе отмечается, что в одной ячейке может располагаться линейная последовательность символов, трактуемая как единый сложный символ, и проводится аналогия между сложными символами и словами европейских языков [28].

Таким образом, имеется *непустое пересечение множества символьных выражений «языка» компьютерной науки и множества слов естественных языков информационной науки*, являющихся знаковой информацией, т. е. знаковыми формами представления знаний и главной сущностью социокультурных коммуникаций. Это объектное пересечение (т. е. пересечение объектов *исследования в компьютерной и информационной науках*) относится одновременно к предметным областям обеих наук. Однако эти объекты в компьютерной науке трактуются и обрабатываются как абстрактное множество символьных выражений, а в информационной науке эти же объекты трактуются и обрабатываются как множества конкретных слов естественных языков с их собственными планами выражения и содержания. Для интеграции двух подходов к трактовке и обработке этого объектного пересечения необходима новая научная парадигма информационно-компьютерной науки.

Заключительная фраза статьи С. Горна говорит о том, что не следует отделять компьютерную науку от информационной науки, а следует пытаться отстаивать единую область знаний. В настоящее время остаются открытыми главные вопросы проблемы конвергенции:

- ◇ На каких теоретических основаниях должна строиться информационно-компьютерная наука как единая область знаний?
- ◇ В какой системе аксиом и с использованием каких терминов следует строить информационно-компьютерную науку как единую область знаний?

Научная парадигма информационно-компьютерной науки, которая будет предлагать ответы на поставленные вопросы, должна включать описание системы аксиом и теоретических оснований этой науки. При этом с единых концептуальных позиций должны быть описаны и классифицированы множество абстрактных символов компьютерной науки и множество конкретных знаковых систем (в том числе естественные языки) информационной науки.

Должны быть описаны отношения между этими множествами символов и знаков на стыке абстрактного и конкретного, включая процессы превращения одного в другое, а также *процессы локальной генерации отдельных «квантов» знаний, их интеграции в цифровой среде и глобального использования в среде социальных коммуникаций на символьно-знаковой основе*.

В завершение необходимо отметить, что настоящий обзор был подготовлен при поддержке РФФИ и в соответствии с правилами Фонда должен включать также описание результатов проектов по тематике обзора, финансируемых по грантам РФФИ. Однако авторам не удалось найти публикаций с изложением результатов проектов РФФИ, посвященных проблеме конвергенции информационной и компьютерной наук, а также вопросам определения количественных индикаторов взаимных связей информационной и компьютерной наук с ИКТ.

## Литература

1. Gorn S. The computer and information sciences: A new basic discipline // SIAM Review, April, 1963. Vol. 5. No. 2. P. 150–155.
2. Decision No 1982/2006/EC of the European Parliament and of the Council of 18 December 2006 concerning the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007–2013) // Official J. of the European Union L412 30.12.2006. P. 1–41.
3. Соломоник А. Парадигма семиотики. — Минск: МЕТ, 2006.
4. Кун Т. Структура научных революций. — М.: АСТ, 2001.

5. Турчин В. Ф. Феномен науки: кибернетический подход к эволюции. — М.: Наука, 1993.
6. Brookes B. C. The foundations of information science. Part I. Philosophical aspects // J. Information Science, 1980. No. 2. P. 125–133.
7. CORDIS ICT Programme Home — [http://cordis.europa.eu/fp7/ict/programme/home\\_en.html](http://cordis.europa.eu/fp7/ict/programme/home_en.html) (состояние страницы на 27.07.2007).
8. ICT FP7 Work Programme — [ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2007-08\\_en.pdf](ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/ict-wp-2007-08_en.pdf) (состояние файла на 27.07.2007).
9. Мамардашвили М. Классический и неклассический идеалы рациональности. — М.: Логос, 2004. 240 с.
10. Computational science: Ensuring America's competitiveness. Report to the President. Arlington, VA: National Coordination Office for Information Technology Research and Development, 2005.
11. Колин К. К. Новая стратегическая компьютерная инициатива США и задачи России в области развития фундаментальной информатики // Информационные технологии, 2006. № 7. С. 2–5.
12. Gorn S. Informatics (computer and information science): Its ideology, methodology, and sociology // The studies of information: Interdisciplinary messages / Ed. by F. Machlup and U. Mansfield. — New York: Wiley, 1983. P. 121–140.
13. Ingwersen P. Information and information science // Encyclopaedia of library and information science / Ed. by A. Kent. — New York: Marcel Dekker Inc., 1995. Vol. 56, sup. 19. P. 137–174.
14. Nonaka I., Takeuchi H. The knowledge-creating company. — N.Y.: Oxford University Press, 1995. (Перевод: Нонака И., Такеучи Х. Компания — создатель знания. — М.: ЗАО «Олимп-Бизнес», 2003. 384 с.)
15. Колин К. К. Фундаментальные основы информатики: социальная информатика. — М.: Академический проект, 2000. 350 с.
16. Farradane J. Knowledge, information, and information science // J. Information Science, 1980, No. 2. P. 75–80.
17. Арский Ю. М., Гиляревский Р. С., Туров И. С., Черный А. И. Инфосфера: информационные структуры, системы и процессы в науке и обществе. — М.: ВИНТИ, 1996. 489 с.
18. Информатика как наука об информации: информационный, документальный, технологический, экономический, социальный и организационный аспекты / Под ред. Р. С. Гиляревского. — М.: ФАИР-ПРЕСС, 2006.
19. Кибрик А. Е. Язык // Языкознание: Большой энциклопедический словарь. — М.: Большая Российская энциклопедия, 1998. С. 604–606.
20. Vossen P., ed. EuroWordNet General Document (Version 3) (URL: <http://www.illc.uva.nl/EuroWordNet/docs/GeneralDoc>).
21. Барт Р. Основы семиологии // Французская семиотика: От структурализма к постструктурализму. — М.: Прогресс, 2000. С. 247–310.
22. Hjørland B. Library and information science: Practice, theory, and philosophical basis // Information Processing and Management, 2000. No. 36. P. 501–531.
23. Digital imaging and communications in medicine (DICOM). Part 16: Content mapping resource. — Rosslyn, Virginia: National Electrical Manufacturers Association, 2004.
24. Рекомендации по преподаванию программной инженерии и computer science в университетах (Software Engineering 2004: Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering; Computing Curricula 2001: Computer Science) / Пер. с англ. — М.: Интернет-Университет Информационных Технологий, 2007. 462 с.
25. Эббинхаус Г. Д. Машины Тьюринга и вычислимые функции I. Уточнение понятия алгоритма // Машины Тьюринга и рекурсивные функции. — М.: Мир, 1972.
26. Полунов Ю. Автора!!! // PC Week/RE, 2006. № 20–21.
27. Hilbert D., Ackermann W. Grundzüge der Theoretischen Logik. — Berlin, 1931.
28. Turing A. M. On computable numbers, with an application to the Entscheidungsproblem // November, No. 12, 1936 (<http://www.abelard.org/turpap2/tp2-ie.asp>).
29. Винер Н. Кибернетика, или управление и связь в животном мире. — М.: Наука, 1983.
30. Shannon C. E. A mathematical theory of communication // The Bell System Technical J., July, October, 1948. Vol. 27. P. 379–423, 623–656.
31. Калман Р. Е. Об общей теории систем управления // Труды ИФАК, 1961. Т. 2. — М.: Изд-во АН СССР. С. 521–547.
32. Пугачев В. С. Основы автоматического управления. — М.: Наука, 1974.
33. Пугачев В. С. Теория вероятности и математическая статистика. — М.: Наука, 2002.
34. Зацман И. М., Шубников С. К. Принципы обработки информационных ресурсов для оценки инновационного потенциала направлений научных исследований // Тр. 9-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» — RCDL'2007 (Переславль, 15–18 октября 2007 г.). — Переславль: Изд-во «Университет города Переславля», 2007. С. 35–44.
35. Gorn S. The individual and political life of information systems // Proc. Symposium on Education for Information Science. — New York: Spartan Books, 1965. P. 33–40.
36. Gorn S. Computer and information sciences and the community of disciplines // Behavioral Science, November, 1967. Vol. 12. No. 6. P. 433–452.
37. Gorn S. The identification of the computer and information sciences: Their fundamental semiotic concepts and relationships // Foundations of Language, November, 1968. Vol. 4. No. 4. P. 339–372.
38. Шрейдер Ю. А. ЭВМ как средство представления знаний // Природа, 1986. № 10. С. 14–22.
39. Шрейдер Ю. А. Информация и знание // Системная концепция информационных процессов. — М.: ВНИИСИ, 1988. С. 47–52.

### THE MIDDLEWARE ARCHITECTURE OF THE SUBJECT MEDIATORS FOR PROBLEM SOLVING OVER A SET OF INTEGRATED HETEROGENEOUS DISTRIBUTED INFORMATION RESOURCES IN THE HYBRID GRID-INFRASTRUCTURE OF VIRTUAL OBSERVATORIES

D. O. Briukhov<sup>1</sup>, A. E. Vovchenko<sup>2</sup>, V. N. Zakharov<sup>3</sup>, O. P. Zhelenkova<sup>4</sup>, L. A. Kalinichenko<sup>5</sup>, D. O. Martynov<sup>6</sup>, N. A. Skvortsov<sup>7</sup>, and S. A. Stupnikov<sup>8</sup>

<sup>1</sup>IPI RAN, brd@ipi.ac.ru

<sup>2</sup>IPI RAN, itsnein@gmail.com

<sup>3</sup>IPI RAN, vzakharov@ipiran.ru

<sup>4</sup>Special Astrophysical Observatory RAN, zhe@sao.ru

<sup>5</sup>IPI RAN, leonidk@synth.ipi.ac.ru

<sup>6</sup>IPI RAN, domartynov@gmail.com

<sup>7</sup>IPI RAN, nskv@ipi.ac.ru

<sup>8</sup>IPI RAN, ssa@ipi.ac.ru

The middleware architecture of subject mediators for scientific problem solving over a set of heterogeneous distributed information resources in the hybrid grid-infrastructure of virtual observatory (VO) is considered. The VO hybrid architecture is implemented as a binding of the AstroGrid VO system developed in the U.K. and of the facilities supporting subject mediators developed at the Institute of Informatics Problems of the Russian Academy of Sciences. An approach is implemented according to which for a class of applications, a specification of subject domain is formed independently of preexisting information resources. The hybrid architecture is implemented as the binding of execution engines of two infrastructures (AstroGrid and subject mediators). That is why, in the paper, the main attention is drawn to the problems of rewriting the mediator queries into the plans of their implementation over specific information resources, to the brief description of the hybrid architecture of execution engines of the AstroGrid and subject mediators. An example of implementation in the hybrid architecture of a subject mediator for solving distant galaxies discovery problem is revealed. The distinguishing features of the presented approach comparing to the well-known prototypes of database integration in VOs are overviewed. The middleware architecture of subject mediators is planned to be used for solving the Russian VO problems.

**Keywords:** subject mediator; canonical information model; virtual observatory; unifier of the information models; refinement, formulae rewriting; semantic integration of heterogeneous information resources; resource registration at the mediator; ontological model; resources relevant to a mediator; middleware; specification of the problem's subject domain driven by an application

### QUASI-LINEAR METHODS FOR THE INFORMATION MODEL BUILDING FOR THE EARTH TIDAL IRREGULAR ROTATION

I. N. Sinitsyn

IPI RAN, sinitsin@dol.ru

Modern stochastic information technologies for scientific research (using *a priori* and *a posteriori* data) are based on the stochastic correlational model building methods. Off-line and on-line quasi-linear methods based on equivalent statistical linearization of nonlinear stochastic differential equations of the Earth tidal irregular rotation (on half year intervals) are considered. Ten testing examples for MATLAB software from informational resources on RAS fundamental problem "Statistical Dynamics of the Earth Motion" are presented.

**Keywords:** Earth tidal irregular rotation; informational model; correlational characteristics; correlational methods; equivalent linearization; stochastic differential equations

## PORTALS FOR e-GOVERNMENT SYSTEMS

A. V. Bosov

IPI RAN, AVBosov@ipiran.ru

Nowadays, different companies, organizations, and associations use informational portal technologies for various purposes. Most likely portal greatest utility are in such practice as cooperation and coordination of business and science. But in other information spheres one can find a lot of problems for successful decision with portal technologies. The subject of the paper is an investigation of current results and perspectives of web-portal technologies applications for government institutions (e-Government). The potency and effectiveness of web-portals in e-Government activity, potentially decided problems are discussed and a lot of examples are briefly considered.

**Keywords:** e-Government; Internet technologies; portal; Internet standards

 $Geo/G/1/\infty$ -QUEUE WITH ONE “NONSTANDARD” DISCIPLINE OF SERVICEA. V. Pechinkin<sup>1</sup> and S. Ya. Shorgin<sup>2</sup><sup>1</sup>IPI RAN, apechinkin@ipiran.ru<sup>2</sup>IPI RAN, sshorgin@ipiran.ru

The object of consideration is a queueing system  $Geo/G/1/\infty$  with the service discipline according to which upon the arrival of a new customer, its length is compared to the length of (remaining) length of the customer on the server. The customer with the minimum length occupies the server whereas the other becomes the first in the queue thus shifting the remaining queue for one place. For this system, main nonstationary characteristics are found. In particular, it is demonstrated that, unlike the continuous time case, for the discrete time case, the stationary distribution of the number of customers in the system is not invariant with respect to the loading.

**Keywords:** queueing system; discrete time; “nonstandard” discipline of the service

## QUEUEING SYSTEMS ALLOCATIONS MINIMIZING EXPECTED QUEUE LENGTH

T. V. Zakharova

Moscow State University, Faculty of Computational Mathematics and Cybernetics, lsa@cs.msu.su

A class of queueing systems is considered with claims emerging on the plane. The problem of optimal allocation of servers is solved with respect to the criterion of expected total queue length. The optimal allocations are compared according to the expected total queue length and expected total waiting time.

**Keywords:** queueing system; claims on a plane; queue length; waiting time; optimal allocation

## NEW STAGE OF THE SOCIETY INFORMATIZATION AND ACTUAL PROBLEMS OF EDUCATION

I. A. Sokolov<sup>1</sup> and K. K. Kolin<sup>2</sup>

<sup>1</sup>IPI RAN, isokolov@ipiran.ru

<sup>2</sup>IPI RAN, kolinkk@mail.ru

The basic features of the present stage of a society informatization process and corresponding actual problems of the education modernization are analyzed. It is shown that the new conditions for a person in a global information society create essentially new opportunities and problems, which are still insufficiently considered in an education system and consequently demand its essential modernization. The current stage of the information society in Russia and the actual problems of the education in Russia connected with the transition to innovative strategy of development are considered.

**Keywords:** global informatization of a society; information and communication technologies; an information society in Russia; a new paradigm of education; advancing education; strategy of innovative development

## PREREQUISITES AND FACTORS OF THE INFORMATION AND COMPUTER SCIENCES CONVERGENCE

I. M. Zatsman<sup>1</sup> and O. S. Kozhunova<sup>2</sup>

<sup>1</sup>IPI RAN, im@a170.ipi.ac.ru

<sup>2</sup>IPI RAN, okozhunova@ipiran.ru

Analytical review is devoted to a problem of convergence of information and computer sciences, and also to interrelations of these sciences with information and communication technologies (ICT). Interest in the convergence problem arouse more than forty years ago. One of its formulations — computer and information sciences: a new basic discipline — became the title of the work by S. Gorn published in 1963. At present, the urgency of a convergence problem has essentially increased. Considered in the review priority guidelines of research and development on ICT within 7th Frame Program of the European Union accepted for the period of 2007–2013 are the evidence of it. In the review, priority ICT guidelines are positioned as external factors of convergence. Except for external factors, historical preconditions of convergence of information and computer sciences are viewed. Factors and preconditions of convergence are considered in a context of development of scientific basis of new ICT generations' creation.

**Keywords:** information science; computer science; information and computer science; information and communication technologies (ICT); correlation between ICT and computer science

# Правила подготовки рукописей статей для публикации в журнале «Информатика и её применения»

Журнал «Информатика и её применения» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информатики и ее приложений. Журнал издается на русском языке. Тематика журнала охватывает следующие направления:

- теоретические основы информатики;
- математические методы исследования сложных систем и процессов;
- информационные системы и сети;
- информационные технологии;
- архитектура и программное обеспечение вычислительных комплексов и сетей.

1. В журнале печатаются результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях. Публикация не должна нарушать закон об авторских правах. Направляя свою рукопись в редакцию, авторы автоматически передают учредителям и редколлегии неисключительные права на издание данной статьи на русском языке и на ее распространение в России и за рубежом. При этом за авторами сохраняются все права как собственников данной рукописи. В связи с этим авторами должно быть представлено в редакцию письмо в следующей форме: Соглашение о передаче права на публикацию:

*«Мы, нижеподписавшиеся, авторы рукописи « \_\_\_\_\_ », передаем учредителям и редколлегии журнала «Информатика и её применения» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторского права других лиц или организаций. Подписи авторов: (ф. и. о., дата, адрес)».*

Редколлегия вправе запросить у авторов экспертное заключение о возможности опубликования представленной статьи в открытой печати.

2. Статья подписывается всеми авторами. На отдельном листе представляются данные автора (или всех авторов): фамилия, полное имя и отчество, телефон, факс, e-mail, почтовый адрес. Если работа выполнена несколькими авторами, указывается фамилия одного из них, ответственного за переписку с редакцией.

3. Редакция журнала осуществляет самостоятельную экспертизу присланных статей. Возвращение рукописи на доработку не означает, что статья уже принята к печати. Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редакционной коллегии о принятии статьи к печати или ее отклонении сообщается авторам. Редколлегия не обязуется направлять рецензию авторам отклоненной статьи.

5. Корректурa статей высылается авторам для просмотра. Редакция просит авторов присылать свои замечания в кратчайшие сроки.

6. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки. Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху — 2, снизу — 2, от края до нижнего колонтитула — 1,3. Основной текст: стиль — «Обычный»; шрифт Times New Roman, размер 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине. Рекомендуемый объем рукописи — не свыше 25 страниц указанного формата. Ознакомиться с шаблонами, содержащими примеры оформления, можно по адресу в Интернете: <http://www.ipiran.ru/journal/template.doc>.

7. К рукописи, предоставляемой в 2-х экземплярах, обязательно прилагается электронная версия статьи (как правило, в форматах MS WORD (.doc) или LaTeX (.tex), а также — дополнительно — в формате .pdf) на дискете, лазерном диске или по электронной почте. Сокращения слов, кроме стандартных, не применяются. Все страницы рукописи должны быть пронумерованы.

8. Статья должна содержать следующую информацию на русском и английском языках: название, Ф.И.О. авторов, места работы авторов и их электронные адреса, аннотация (не более 100 слов), ключевые слова. Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в порядке их первого упоминания. Все фамилии авторов, заглавия статей, названия книг, конференций и т. п. даются на языке оригинала, если этот язык использует кириллический или латинский алфавит.

9. Присланные в редакцию материалы авторам не возвращаются.

10. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- использовать attach (присоединение);
- в случае больших объемов информации возможно использование общеизвестных архиваторов (ZIP, RAR);
- в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.

11. Журнал «Информатика и её применения» является некоммерческим изданием, и гонорар авторам не выплачивается.

**Адрес редакции:** Москва 119333, ул. Вавилова, д. 44, корп. 2, ИПИ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05 E-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru)