

**Российская академия наук  
Сибирское отделение**

**Федеральное государственное бюджетное учреждение науки**

**Институт систем информатики  
имени А.П. Ершова СО РАН**

**Отчет о деятельности  
в 2017 году**

**Новосибирск  
2017**

**Институт систем информатики имени А.П.Ершова СО РАН**

**630090, г. Новосибирск, пр. Лаврентьева, 6**

**e-mail:** [iis@iis.nsk.su](mailto:iis@iis.nsk.su)

**http:** [www.iis.nsk.su](http://www.iis.nsk.su)

**тел:** (383) 330-86-52

**факс:** (383) 332-34-94

**Директор**

**д.ф.-м.н.**

**Марчук Александр Гурьевич**

**e-mail:** [mag@iis.nsk.su](mailto:mag@iis.nsk.su)

**http:** [www.iis.nsk.su](http://www.iis.nsk.su)

**тел:** (383) 330-86-52

**Заместитель директора по научной работе**

**к.ф.-м.н.**

**Мурзин Федор Александрович**

**e-mail:** [murzin@iis.nsk.su](mailto:murzin@iis.nsk.su)

**http:** [www.iis.nsk.su](http://www.iis.nsk.su)

**тел:** (383) 330-70-68

**Заместитель директора по экономическим вопросам**

**Филиппов Владимир Эдуардович**

**e-mail:** [fil@iis.nsk.su](mailto:fil@iis.nsk.su)

**http:** [www.iis.nsk.su](http://www.iis.nsk.su)

**тел:** (383) 332-96-58

**Ученый секретарь**

**к.ф.-м.н.**

**Промский Алексей Владимирович**

**e-mail:** [promsky@iis.nsk.su](mailto:promsky@iis.nsk.su)

**http:** [www.iis.nsk.su](http://www.iis.nsk.su)

**тел:** (383) 330-70-68

## Введение

Институт систем информатики имени А.П.Ершова Сибирского отделения РАН (ИСИ СО РАН) создан в апреле 1990 г. Постановлением Президиума Сибирского отделения РАН № 268 от 20.08.1997 г. определены основные научные направления института – теоретические и методологические основы создания систем информатики, в том числе:

- теоретические основания информатики;
- методы и инструменты построения программ повышенной надежности и эффективности;
- методы и системы искусственного интеллекта;
- системное и прикладное программное обеспечение перспективных вычислительных машин, систем, сетей и комплексов.

Среднесписочная численность сотрудников института в 2017 г. составила 119 человек, из них 65 научных сотрудников, в том числе 5 докторов наук и 33 кандидата наук.

В 2017 г. в институте проводились исследования в области теоретических и методологических основ информатики, включая все перечисленные выше направления. Все задания 2017 г. выполнены.

Сотрудниками института в 2017 г. опубликовано: 1 монография, 51 статья в рецензируемых отечественных журналах, 19 статей — в зарубежных рейтинговых журналах, 48 докладов в трудах международных конференций, получено 2 свидетельства о государственной регистрации программы.

В 2017 г. для участия в работе международных конференций, чтения лекций и проведения совместных научных исследований за рубеж выезжали 6 сотрудников института.

## Структура Института.

### Краткая характеристика подразделений

На 01.12.2017 г. в структуре Института имелось 8 лабораторий и 1 научно-исследовательская группа.

<b>Лаборатория теоретического программирования</b>	<b>Лаборатория автоматизации проектирования и архитектуры СБИС</b>	<b>Лаборатория искусственного интеллекта</b>
<b>Лаборатория системного программирования</b>	<b>Лаборатория конструирования и оптимизации программ.</b>	<b>Лаборатория смешанных вычислений</b>
<b>Лаборатория моделирования сложных систем</b>	<b>Лаборатория теории параллельных процессов</b>	<b>НИГ переносимых систем программирования</b>

### Лаборатория теоретического программирования

*Заведующий лабораторией к.ф.-м.н. Валерий Александрович Непомнящий.*

**Кадровый состав:** всего сотрудников — 22, из них научных сотрудников — 14 (в том числе 1 доктор и 7 кандидатов наук).

### **Основные направления исследований:**

- исследование формальных моделей и методов описания семантики, спецификации и верификации программ и систем.

### **Лаборатория автоматизации проектирования и архитектуры СБИС**

*Заведующий лабораторией д.ф.-м.н. Александр Гурьевич Марчук.*

**Кадровый состав:** всего сотрудников — 19, из них научных сотрудников — 9 (в том числе 1 доктор и 5 кандидатов наук).

### **Основные направления исследований:**

- разработка систем автоматизации проектирования и программирования;
- создание информационных и телекоммуникационных систем и сетей.

### **Лаборатория искусственного интеллекта**

*Заведующий лабораторией к.т.н. Юрий Алексеевич Загорюлько.*

**Кадровый состав:** всего сотрудников — 9, из них научных сотрудников — 7 (в том числе 2 кандидата наук).

### **Основные направления исследований:**

- методы и системы искусственного интеллекта.

### **Лаборатория системного программирования**

*Заведующий лабораторией к.т.н. Владимир Иванович Шелехов.*

**Кадровый состав:** всего сотрудников — 5, из них научных сотрудников — 4 (в том числе 1 кандидат наук).

### **Основные направления исследований:**

- создание методов и экспериментальных инструментов конструирования и спецификаций программ в окружениях надежного программирования.

### **Лаборатория конструирования и оптимизации программ**

*Заведующий лабораторией д.ф.-м.н., проф., член-корр. РАН Виктор Николаевич Касьянов.*

**Кадровый состав:** всего сотрудников — 10, из них научных сотрудников — 7 (в том числе 2 доктора и 2 кандидата наук).

### **Основные направления исследований:**

- развитие теории трансформационного программирования и разработка методов и средств конструирования эффективных и надежных программ;
- разработка программно-методических средств поддержки преподавания фундаментальных основ информатики и программирования;
- создание инструментально-информационной системы по оптимизирующим и реструктурирующим преобразованиям программ для ЭВМ параллельных архитектур;
- подготовка «Энциклопедии по алгоритмам и методам теории графов для программистов».

## Лаборатория смешанных вычислений

*Заведующий лабораторией к.ф.-м.н. Михаил Алексеевич Бульонков.*

**Кадровый состав:** всего сотрудников — 7, из них научных сотрудников — 6 (в том числе 4 кандидата наук).

**Основные направления исследований:**

- теория и практика смешанных вычислений.

## Лаборатория моделирования сложных систем

*Заведующий лабораторией к.ф.-м.н. Мурзин Федор Александрович.*

**Кадровый состав:** всего сотрудников — 11, из них научных сотрудников — 10 (в том числе 8 кандидатов наук).

**Основные направления исследований:**

- разработка сложных алгоритмов и программных систем для применения в различных областях: обработка изображений и сигналов, биоинформатика, поиск нефти, обработка текстов на естественном языке.

## Лаборатория теории параллельных процессов

*Заведующий лабораторией д.ф.-м.н. Вирбицкайте Ирина Бонавентуровна.*

**Кадровый состав:** всего сотрудников — 7, из них научных сотрудников — 7 (в том числе 1 доктор и 4 кандидата наук).

**Основные направления исследований:**

- теоретико-категорное исследование взаимосвязей параллельных моделей с реальным временем и их эквивалентностей;
- изучение свойств достижимости, безопасности, управления моделей различных классов динамических и гибридных систем;
- разработка дискретно-временных стохастических расширений алгебр параллельных процессов, построение стохастических алгебраических и поведенческих эквивалентностей и исследование их взаимосвязей;
- проектирование алгоритмов параметрической верификации различных классов временных сетей Петри.

## Научно-исследовательская группа переносимых систем программирования

*Руководитель группы Андрей Дмитриевич Хапугин.*

**Кадровый состав:** всего сотрудников — 4, из них научных сотрудников — 1.

**Основные направления исследований:**

- теоретические основы и инструментальные программные системы, поддерживающие разработку переносимых программных систем на базе объектно-ориентированного подхода.

Научная и научно-организационная деятельность научных подразделений координируется Ученым советом.

## Основные научные результаты, полученные в 2017 году

### 1. Тема № 39.1.1. Исследования фундаментальных основ структуризации данных, управления информационными ресурсами, создание информационно-вычислительных систем и сред для науки и образования.

*Руководитель темы: Марчук А.Г.*

Предложен способ организации процесса извлечения знаний логико-вероятностными методами, позволяющий эффективнее анализировать данные, имеющие большой размер пространства входных признаков.

Исследована алгоритмическая сложность семантического комбинирования онтологий, основанного на механизме согласования их моделей, и обобщающего операцию простого объединения онтологий, как логических теорий.

Реализован метод вычисления явных определений понятий относительно онтологий в дескрипционной логике EL на основе исчисления направленного вывода. Разработан алгоритм вычисления явных определений понятий, основанный на трассировке доказательств в логике EL, который позволяет проверить, существует ли определение интересующего понятия с использованием выбранных пользователем терминов и отношений, и построить определение в явном виде. Алгоритм реализован в виде прототипа плагина (расширения) для широко используемого редактора онтологий Protégé и имеет практическую ценность для анализа и декомпозиции объемных онтологий, т.к., в частности, позволяет установить, возможно ли устранить те или иные термины и отношения из определений понятий в онтологии.

Для широкого класса дескрипционных логик получена алгоритмическая классификация проблемы логического следования из семантически связанных онтологий. В рамках рассмотренного класса логик показано, что ациклическое семантическое комбинирование приводит к повышению сложности проблемы логического следования на одну экспоненту по сравнению со сложностью логики, в которой сформулированы онтологии. Результаты исследований дают основу для реализации новых практических методов комбинирования онтологий, в том числе, в целях спецификации терминологий естественнонаучных предметных областей.

Предложен метод декомпозиции таблиц реляционных баз данных, основанный на ранее полученном алгоритме факторизации булевских полиномов, имеющем полиномиальную временную сложность.

Разработан инструмент в виде .Net библиотеки для конвертирования изображений в Deep Zoom Image и стандартные изображения требуемого размера, используя технологию Polar.BD. Инструмент может применяться для отображения изображений высокого разрешения в браузере в формате одного изображения выбранного размера, при этом оно динамически составляется из частей нужного размера, и с помощью технологии Deep Zoom от Microsoft, Seadragon и Open Seadragon.

Выполнен содержательный анализ специфики основных и фундаментальных парадигм программирования, сформулированы ведущие критерии их упорядочения и классификации языков программирования, поддерживающих разные парадигмы.

### 2. Тема № 39.1.2. Методы и технологии конструирования эффективного и надежного программного обеспечения для суперкомпьютеров и компьютерных сетей.

*Руководитель темы: Касьянов В.Н.*

Разработаны методы и создана экспериментальная версия среды параллельного и функционального программирования CSS для поддержки облачных супервычислений, ориентированная на работу с Cloud Sisal-программами и супервычислители Сибирского суперкомпьютерного центра СО РАН (ССКЦ СО РАН) и пригодная для учебных применений.

Предложены критерии оценки полноты покрытия тестами для проверки оптимизирующих компиляторов. Разработаны новые методы и алгоритмы автоматической генерации тестов для проверки распараллеливающих и векторизирующих преобразований циклов в компиляторе, ориентированные на полноту покрытия оптимизаций циклов тестами. Созданы программные средства проверки распараллеливающих и векторизирующих преобразований циклов в компиляторе для языка Cloud-Sisal.

Для генерации тестов разработан метаязык, который напоминает БНФ-нотацию и функциональный язык. Грамматика задается в одном файле и является набором строк с поддержкой комментариев. Правило – это идентификатор, для которого задано, каким образом он может быть переписан. Правило может быть как генерирующим, так и ограничивающим. Любой идентификатор – это последовательность символов, не начинающаяся с цифры. Имя параметра шаблона контекста в левой части правила может быть любым именем распознанной группы символов регулярного выражения языка Python, стоящей в предыдущем шаблоне контекста в левой части правила. Правая часть правила может задаваться либо альтернативами, либо многострочным правилом. Для более тонкой работы с правилами можно использовать специальные распознаватели контекста, предоставляющие возможности по извлечению объектов из контекста и по их переименованию, а также по изменению самого контекста для всех идентификаторов в правой части правила. Помимо этого для описания и распознавания контекста можно использовать так называемые условия – специального вида булевские выражения. Также в языке грамматики предусмотрено несколько встроенных функций для удобной работы со списками, условиями и числами. Задаваемая при этом параметрическая КС-грамматика по наглядности сопоставима с обычной КС-грамматикой, но позволяет описывать более широкий класс КЗ-языков.

Создана система ATG (Automatic Test Generator), которая позволяет в автоматическом режиме тестировать компилятор при помощи порождаемых генератором тестов следующим образом. Каждый сгенерированный тест проходит две проверки – при компиляции с оптимизациями и без. Если тест успешно компилируется при помощи компилятора с отключенными оптимизациями и включенной проверкой указателей и запускается, то он является статически и динамически детерминированной программой, и результат ее исполнения объявляется эталонным. Здесь стоит сделать уточнение насчет успешного запуска: успешный запуск – это отсутствие сообщений о проблемах с указателями и нулевой код завершения, а также отсутствие завершения программы по установленному тайм-ауту. Это первая ступень проверки. Следующая ступень, это проверка компиляции с оптимизациями. Если тест вызвал на данном этапе проблемы, то он уже полезен для тестирования и сохраняется для дальнейшей обработки. Если компиляция прошла успешно, то проверяется корректность исполнения программы с ее эталоном. И если программа проваливает эту проверку, то она также сохраняется. В случае, когда программа проходит проверку с эталоном, она просто удаляется и запускается процесс генерации нового теста. Сохраненные тесты проходят дальнейшую обработку при помощи нашей системы упрощения компиляторных тестов Reduce. Также производится попытка найти компоненту компилятора, виновную в проблеме, при помощи функционала компилятора, который позволяет отключать применяемые оптимизации. Затем тесты сохраняются в базу системы с пометкой, на какой системе, когда и с какой проблемой они были получены. С целью оптимизации места, занимаемого хранилищем тестов, сохраняются 10 минимальных по размеру тестов на каждую проблемную оптимизацию. Но если возможно найти дату проблемной сборки, то это позволяет сохранять по 10 тестов на каждую дату.

Осуществлено расширение словаря по теории графов WikiGRAPP и энциклопедии теоретико-графовых алгоритмов WEGA новыми базовыми терминами и алгоритмами, в том

числе связанными с кодированием и обработкой деревьев, а также произведено пополнение существующих статей словаря и энциклопедии дополнительной информацией, в том числе статическими и динамическими иллюстрациями.

Разработан пакет программ, расширяющий MediaWiki средствами подготовки динамических иллюстраций работы теоретико-графовых алгоритмов обработки деревьев.

Разработаны новые методы и алгоритмы визуализации информации на основе атрибутированных иерархических графовых моделей.

Создана и прошла государственную регистрацию программная система Visual Graph, предназначенная для визуализации сложных больших данных на основе иерархических графовых моделей большого размера.

### **3. Тема № 39.1.3. Методы и средства повышения надежности программных систем, базирующиеся на формальной спецификации и верификации.**

*Руководители темы: Непомнящий В.А., Вирбицкайте И.Б.*

Получены новые результаты о топологических и теоретико-множественных аспектах теории вычислений на несчетных структурах, значимые для дескриптивной теории множеств, общей топологии и теории вычислений.

Разработаны новые подходы к оценке дескриптивной и арифметической сложности вычислений над непрерывными данными и на их основе проведен анализ сложности непрерывных констрейнтов. Ряд теорем эффективной ДТМ перенесены на некоторый класс эффективных пространств, строго содержащий классы вычислимых польских и вычислимых квази-польских пространств. Доказана знаменитая теорема Райса-Шапиро для вычислимых элементов модульных T-0-пространств.

В iRRAM-пакете разработаны и реализованы методы представления непрерывных констрейнтов, реализован алгоритм проверки истинности линейных констрейнтов. Разработаны новые подходы к вычислениям с обобщенными моделями Тейлора.

Определена и исследована эквивалентность флюидных стохастических сетей Петри (ФССП) на основе комкуемости (lumpability) для цепей Маркова и дифференциальных уравнений. Предложены подходы к определению флюидных бисимуляционных эквивалентностей мест ПФССП поднятием бисимуляции непрерывных мест на гибридные маркировки и применением дуг детерминированного скачка флюида из расширенных ФССП. Для флюидных следовой и бисимуляционной эквивалентностей помеченных флюидных стохастических сетей Петри (ПФССП) построена и исследована логическая характеристика оригинальными флюидными модальными логиками.

В контексте модели реляционных структур, обобщающих различные событийные модели параллельных процессов, установлены взаимосвязи «аксиом параллелизма», позволяющие изучать отношения между событиями параллельных/распределенных систем.

Предложены методы построения и сопоставления изоморфных абстракций в виде систем переходов, представляющих изменения структуры и поведения при функционировании событийно-ориентированных моделей различных классов.

Для случая произвольных конечных систем переходов над двоичным алфавитом построена характеристика генерируемых сетями Петри пространств состояний, позволяющая осуществить минимальную аппроксимацию сверху конечных языков языками сетей Петри.

Предложены новое более общее определение концептуальных систем перехода (КСП) и двухэтапный метод применения КСП к разработке формальной онтологической операционной семантики языков программирования, базирующийся на новом определении КСП. Разработана формальная онтологическая операционная семантика для модельных процедурных языков программирования из семейства MPL и фрагмента языка C# как типовом представителе объектно-ориентированных языков.

Разработанная ранее система верификации программ на языке C-light расширена с целью применения символического метода элиминации инвариантов циклов для финитных итераций над массивами.

Язык спецификаций распределенных систем MSC расширен посредством временных конструкций. Разработана и реализована программная система анализа и верификации MSC-спецификаций, которые расширены этими конструкциями.

Разработаны и реализованы новые версии программных средств анализа и верификации распределенных систем, представленных на языках спецификаций SDL, UCM и MSC.

Разработан мультиагентный подход к разрешению кореференции в процессе извлечения информации из текстов для пополнения онтологии.

#### **4. Тема № 39.1.4. Методы и технология создания и сопровождения интеллектуальных информационных систем и систем поддержки принятия решений.**

*Руководитель темы: Загоруйко Ю.А.*

В рамках разработки технологии создания и сопровождения интеллектуальных информационных систем (ИИС) поддержки научной, производственной и образовательной деятельности разработаны новые методы интеграции таких ИИС, построенных с использованием технологий Semantic Web, с облаком Linked Open Data (LOD). Разработаны новые методы пополнения контента ИИС данными из облака LOD и публичных веб-сайтов. Исследованы вопросы использования паттернов онтологического проектирования для разработки онтологий научных предметных областей (ПО). Разработаны наборы структурных и содержательных паттернов, с помощью которых могут единообразно задаваться основные сущности онтологий различных научных ПО.

В рамках развития технологии анализа текста разработаны новые семантические модели и методы, используемые для извлечения информации из текстов. Предложена модель представления предметной лексики, позволяющая описывать сложную семантику терминов, и осуществлена ее интеграция в модель представления факта. Разработаны новые методы жанрового анализа текстов. Разработан ряд лингвистических ресурсов, предназначенных для решения задач классификации и извлечения информации из текстов технической документации и медицинских интернет-источников. Предложен подход к синхронизации аннотированных текстов с их переводами на другие языки, а также с сопутствующими мультимедийными материалами.

В рамках создания технологии построения интеллектуальных систем поддержки принятия решений разработана методология информационно-аналитической поддержки (ИАП) разработчиков СППР, которая представляет собой систему принципов, подходов и способов организации теоретической и практической деятельности по обеспечению ИАП и использованию её средств. Разработан сервис, реализующий метод недоопределенных вычислений и метод рассуждений на основе экспертных правил с использованием клиент-серверной архитектуры.

#### **5. Тема № 39.1.5. Алгоритмы и программные средства для моделирования сложных систем.**

*Руководители темы: Мурзин Ф.А., Бульонков М.А.*

Проведен комплекс исследований по компьютерной лингвистике и реализовано разнообразное программное обеспечение для обработки текстов на естественных языках: русском, английском, казахском и турецком. Рассмотрен ряд задач: разработка системы

семантико-синтаксических связей для тюркских языков, определение релевантности текста поисковому запросу, определение тем текстов, применение риторических структур для анализа текстов научно-технической направленности, разрешение анафорических связей, анализ данных из социальных сетей.

Получены результаты по кодированию источников со счётным алфавитом. Получены совпадающие по порядку убывания оценки избыточности универсального кодирования источников с бесконечным входным алфавитом для множеств источников с ограниченными моментами.

Разработан новый многокритериальный алгоритм установления авторов научных публикаций на основе сопоставления русскоязычной базы данных elibrary и англоязычного сайта SpringerLink.com. Благодаря сопоставлению русскоязычного и англоязычного источников данных удалось повысить полноту обнаруживаемых данных, а за счет реализации нового алгоритма кластеризации авторов научных публикаций удалось повысить точность идентификации авторов.

**В 2017 г. Институт проводил исследования по следующим программам:**

**Интеграционные проекты РАН и СО РАН:**

**Интеграционный проект РАН 15/10** «Математические и методологические аспекты интеллектуальных информационных систем».

*Руководитель Марчук А.Г.*

**Гранты РФФИ:**

**Проект РФФИ 15-01-05974** «Онтологический подход к формальной семантике языков программирования»

*Руководитель: Ануреев И.С.*

*Сроки: 2015-2017 гг.*

**Проект РФФИ № 16-07-00569а** «Методы и средства комплексной поддержки процесса разработки интеллектуальных СППР в слабоформализованных предметных областях на основе сервис-ориентированного подхода и технологий Semantic Web»

*Руководитель: к.т.н., заведующий лабораторией Ю.А. Загоруйко*

**Проект РФФИ 15-07-04144а** «Методы и технологии создания предметно-ориентированных систем извлечения информации из документов на основе мультиагентного подхода»

*Руководитель: Сидорова Е.А.*

**Проект РФФИ № 16-01-00498а**

Технология предикатного и автоматного программирования

*Руководитель – к.т.н., зав.лаб. В.И. Шелехов*

**Проект РФФИ 15-07-020029** «Методы и средства функционального программирования для поддержки облачных супервычислений».

*Руководитель: д.ф.-м.н., профессор В.Н. Касьянов*

**Проект РФФИ 16-01-00499** «Применение методов теории графов для анализа структурной информации»

*Руководитель — д.ф.-м.н., профессор В.А. Евстигнеев*

**Проект РФФИ N 14-07- 00386А** «Исследование и разработка технологий реализации и анализа больших графовых данных научного профиля»

*Руководитель: д.ф.-м.н. Марчук А.Г.*

**Проект РФФИ 17-07-20538** Проект организации 4-й Международной научной конференции «Развитие вычислительной техники в России и странах бывшего СССР: история и перспективы» SoRuCom-2017.

*Руководитель д.ф.-м.н. А.Г. Марчук (член ПК).*

**Проект РФФИ и фонда ДСТ (Индия) 17-51-45125** «Новые методы анализа и оптимизации реконфигурируемых компьютерных архитектур»

*Руководитель – к.ф.-м.н Пономарев Д.К.*

**Проект РФФИ N 17-11-00042 Д** «Издание научного труда "Парадигмы программирования"».

*Руководитель – к.ф.-м.н. Л.В. Городняя.*

**Проект РФФИ N 15-07-06345** “Становление и развитие научных школ программирования в ведущих научных центрах СССР.

*Руководитель – к.ф.-м.н. С.П. Прохоров.*

*Участник от ИСИ СО РАН Городняя Л.В.*

**Проект РФФИ N 15-07-06345** «Научные школы программирования в Академии наук СССР и ведущих советских вузах».

*Руководитель – к.ф.-м.н. С.П. Прохоров, Институт истории естествознания и техники РАН, г. Москва.*

*Участник от ИСИ СО РАН Крайнева И.А.*

#### **Гранты РНФ:**

**Грант РНФ** «Логико-вероятностный инструментарий искусственного интеллекта нового поколения»

*Руководитель академик С.С. Гончаров ИМ СО РАН*

*Участник от ИСИ СО РАН Пономарев Д.К.*

#### **Гранты РГНФ:**

**Проект РГНФ 13-01-12003в**

*Руководитель – д.ф.-м.н Панина Н.Л.*

**Проект РГНФ № 16-02-00221** «Моделирование процесса освоения северных территорий и акваторий России: игровой подход на основе геоинформационных технологий»

*Руководитель: Малов В.Ю.*

*Исполнители: Бульонков М.А., Филаткина Н.Н., Тарасова О.В., Мелентьев Б.В.*

**Грант благотворительного фонда В.Потанина** «Российский образовательный проект по курсу «Формальные Методы в Программной Инженерии»

*Руководитель – к.т.н., зав.лаб. В.И. Шелехов*

**Грант Благотворительного фонда В. Потанина** «Разработка магистерской программы «Логические методы в информатике».

*Руководитель Д.К. Пономарев*

#### **Международные проекты:**

**Международный проект « Сравнительный анализ и верификация для параллельных систем с повышенными требованиями к корректности — Проект DFG (CAVER, грант No BE 1267/14-1) и РФФИ (грант No 14-01-91334)**

*Руководители: А. Бест (Ольденбургский университет, Германия), П. Бухгольц (Дортмундский университет, Германия), Л. Попова-Цейгманн (Берлинский университет, Германия)*

*Участник: И.Б. Вирбицкайте*

*Сроки: 2014 – 2017 гг.*

**Международный проект «Computable analysis – theoretical and applied aspects», EU—грант № PIRSES-GA-2011-294962**

*Руководители: Дитер Шприн (Зиген, Германия), Виктор Селиванов (ИСИ СО РАН)*

*Участник: Коровина М.В.*

*Сроки: 2014 – 2017 гг.*

**Международный проект Испанского правительства "Formal Analysis and Applications of Web Services and Electronic Contracts (DArDOS)", грант TIN2015-65845-C03-02.**

*Руководитель: Maria Emilia Cambroner Piqueras, Gregorio Diaz Descalzo*

*Участник: Тарасюк И.В.*

*Сроки: 2016 – 2018 гг.*

**Проект Немецкого исследовательского сообщества “Технологии экспертной поддержки пользователей в рамках когнитивных технических систем”**

*Иностранные партнеры: Институт искусственного интеллекта при факультете информатики университета г. Ульм, Германия*

*Координаторы проекта: Сюзанна Биундо-Штефан (Ульм, Германия), Андреас Вендемут (Магдебург, Германия)*

*Участник: Пономарев Д.К.*

*Сроки: 2013 – 2017 гг.*

**Седьмая европейская рамочная программа, № контракта (гранта) 305280, раздел HEALTH.2012.2.1.1-3. MIMOmics – Methods for Integrated analysis of Multiple Omics datasets (методы интегрированного анализа множества омов**

*Совместно с компанией geneXplain, ответственный исполнитель от ИСИ СО РАН – к.ф.-м.н. Т.Ф. Валеев*

*Валеев Т.Ф.*

*Сроки: 2012 – 2017 гг.*

**Грант Министерства образования и науки РК (аналог ФЦП) – Разработка информационно-поискового тезауруса (с учетом морфологии казахского языка) в полнотекстовых базах данных по ИТ-технологиям, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана.**

*Сроки: 2015-2017 гг.*

*Научные руководители: д.ф.-м.н. Тусупов Д.А. (Казахстан), член-корр. Федотов А.М. (НГУ), ответственный исполнитель от ИСИ СО РАН – к.ф.-м.н. Т.В. Батура*

*Сроки: 2015 – 2017гг.*

## **Общая характеристика исследований лаборатории теоретического программирования**

*Зав лабораторией к.ф.-м.н. Непомнящий В.А.*

### **Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе**

**Приоритетное направление IV.39.** Архитектура, системные решения, программное обеспечение и информационная безопасность информационно-вычислительных комплексов и сетей новых поколений. Системное программирование.

**Программа IV.39.1.** Теоретические и прикладные проблемы создания эффективных надежных программных систем и информационных технологий

**Проект IV.39.1.3.** «Методы и средства повышения надежности программных систем, базирующиеся на формальной спецификации и верификации»

Объектом исследования являются методы теории вычислений на несчетных структурах, методы исследования систем дискретного и непрерывного времени, проблемы надежности сложных параллельных программных систем реального времени, методы описания формальной семантики языков программирования, методы создания средств дедуктивной верификации программ, методы и средства моделирования и верификации распределенных систем, мультиагентный подход к проблеме извлечения информации из текстов для пополнения онтологии.

Цель работы: разработка методов исследования топологических пространств, используемых для вычислимого анализа; методов сравнительного анализа и доказательства корректности параллельных и динамических моделей с временными, стохастическими и гибридными характеристиками; методов описания семантики языков программирования и технологии создания средств дедуктивной верификации программ; методов и средств моделирования и верификации распределенных систем; разработка методов пополнения систем организаций знаний.

Методы: логические, автоматные и сложностные методы исследования систем дискретного и непрерывного времени; методы теории вычислимости и теории моделей; методы теорий параллелизма, алгебраических решеток, категорий, стохастических процессов, флюидных потоковых систем, интервального и вычислимого анализа, теории моделей, теории нумераций; методы и средства моделирования и верификации последовательных программ, распределенных и мультиагентных систем.

Результаты: Получены новые результаты о топологических и теоретико-множественных аспектах теории вычислений на несчетных структурах, значимые для дескриптивной теории множеств, общей топологии и теории вычислений.

Разработаны новые подходы к оценке дескриптивной и арифметической сложности вычислений над непрерывными данными и на их основе проведен анализ сложности непрерывных констрейнтов. Ряд теорем эффективной ДТМ перенесены на некоторый класс эффективных пространств, строго содержащий классы вычислимых польских и вычислимых квази-польских пространств. Доказана знаменитая теорема Райса-Шапиро для вычислимых элементов модульных T-0-пространств.

В iRRAM-пакете разработаны и реализованы методы представления непрерывных констрейнтов, реализован алгоритм проверки истинности линейных констрейнтов. Разработаны новые подходы к вычислениям с обобщенными моделями Тейлора.

Определена и исследована эквивалентность флюидных стохастических сетей Петри (ФССП) на основе комкуемости (lumpability) для цепей Маркова и дифференциальных

уравнений. Предложены подходы к определению флюидных бисимуляционных эквивалентностей мест ПФССП поднятием бисимуляции непрерывных мест на гибридные маркировки и применением дуг детерминированного скачка флюида из расширенных ФССП. Для флюидных следовой и бисимуляционной эквивалентностей помеченных флюидных стохастических сетей Петри (ПФССП) построена и исследована логическая характеристика оригинальными флюидными модальными логиками.

В контексте модели реляционных структур, обобщающих различные событийные модели параллельных процессов, установлены взаимосвязи «аксиом параллелизма», позволяющие изучать отношения между событиями параллельных/распределенных систем.

Предложены методы построения и сопоставления изоморфных абстракций в виде систем переходов, представляющих изменения структуры и поведения при функционировании событийно-ориентированных моделей различных классов.

Для случая произвольных конечных систем переходов над двоичным алфавитом построена характеристика генерируемых сетями Петри пространств состояний, позволяющая осуществить минимальную аппроксимацию сверху конечных языков языками сетей Петри.

Предложены новое более общее определение концептуальных систем перехода (КСП) и двухэтапный метод применения КСП к разработке формальной онтологической операционной семантики языков программирования, базирующийся на новом определении КСП. Разработана формальная онтологическая операционная семантика для модельных процедурных языков программирования из семейства MPL и фрагмента языка C# как типовом представителе объектно-ориентированных языков.

Разработанная ранее система верификации программ на языке C-light расширена с целью применения символического метода элиминации инвариантов циклов для финитных итераций над массивами.

Язык спецификаций распределенных систем MSC расширен посредством временных конструкций. Разработана и реализована программная система анализа и верификации MSC-спецификаций, которые расширены этими конструкциями.

Разработаны и реализованы новые версии программных средств анализа и верификации распределенных систем, представленных на языках спецификаций SDL, UCM и MSC.

Разработан мультиагентный подход к разрешению кореференции в процессе извлечения информации из текстов для пополнения онтологии.

### **Краткое описание проведенных научных исследований**

Изучение классов топологических пространств, допускающих «хорошую» дескриптивную теорию множеств и представляющих интерес для вычислимого анализа, является актуальной проблемой. Классическая дескриптивная теория множеств развита для случая Польских пространств. Ранее В.Л. Селивановым была развита существенная часть дескриптивной теории множеств для счетно-базируемых областей Ершова-Скотта. Недавно эта теория была обобщена М. де Брехтом на случай так называемых квази-Польских пространств.

Для анализа вычислений над непрерывными данными требуется развитие различных подходов к оценке сложности вычислений. В настоящее время результаты теории вычислимости над непрерывными структурами активно используются для создания верификационных алгоритмов для анализа безопасности комплексных систем, таких как гибридные системы. На сегодняшний момент нет устоявшейся теории оценки сложности вычислений с непрерывными данными и она требует дальнейшего развития и обоснования.

Известно, что в теории параллелизма достаточно много наработок в области анализа и синтеза структуры и поведения параллельных/распределенных систем. Однако есть необходимость в дальнейшем исследовании взаимосвязей отношений (причинной зависимости,

параллелизма, конфликта и их ослаблений/усилений) между событиями таких систем, в разработке различных изоморфных представлений изменения структуры и поведения при функционировании событийно-ориентированных моделей и в сопоставлении классических языков и языков, порожаемых параллельными моделями на основе синтеза последних. Кроме того, в контексте новой модели сетей Петри с параметрическими, временными и флюидными стохастическими характеристиками сетей Петри для моделирования сложных физических систем реального времени и с текучими веществами необходимо разработать поведенческие эквивалентности, которые могли бы упростить анализ таких систем за счет уменьшения количества их состояний, объединяемых в классы эквивалентности.

Разработка новых подходов к описанию формальной операционной семантики языков программирования представляет значительный интерес. Для достижения этой цели полезен онтологический подход. Поэтому возникает задача развития этого подхода и его апробация для известных языков программирования.

Дедуктивная верификация С-программ является важной и открытой проблемой. Известные системы верификации не имеют средств построения инвариантов циклов, что является трудной задачей для пользователей. Поэтому применение символического метода элиминации инвариантов циклов для финитных итераций над массивами в С-программах является актуальной задачей.

Проблема анализа и верификации распределенных систем, представленных на языке спецификаций MSC, является актуальной для современного программирования. Временные характеристики играют важную роль в современных программных системах. Поэтому возникает задача временного расширения MSC-диаграмм и разработки метода верификации этих диаграмм с временными конструкциями.

Представляет интерес разработка и реализация новых версий программных средств анализа и верификации распределенных систем, представленных на языках спецификаций SDL и UCM. Эти версии программных средств будут включать модули локализации ошибок.

Проблема анализа требований к параллельным, распределенным и мультиагентным программным системам является актуальной для поиска и анализа информации. Естественно возникают задачи разработки онтологии шаблонов требований к таким системам, а также пополнения онтологии предметной области.

Целями проекта в целом являются следующие:

Исследования теории вычислений на дискретных и непрерывных структурах, значимые для дескриптивной теории множеств, общей топологии и теории вычислений.

Разработка новых формальных методов для оценки арифметической и дескриптивной сложности непрерывных констрейнтов и построения решений полиномиальных динамических систем с использованием обобщенных моделей Тейлора.

С целью выявления и установления взаимосвязей между базовыми отношениями ((слабой/сильной) причинной зависимости, синхронного/асинхронного параллелизма, симметричного/асимметричного конфликта и т.д.) между событиями параллельных/распределенных систем предполагается определить и исследовать «аксиомы параллелизма» для различных событийных и реляционных моделей. С целью отслеживания корректности функционирования моделируемых систем будут построены различные изоморфные абстракции структуры и поведения событийно-ориентированных моделей в виде систем переходов для таких классов, как расслоенные, дуальные, с самоконфликтными событиями, стабильные, обобщенные, с динамически меняющимися причинной зависимостью и конфликтом, с приоритетами и др. Характеризация пространств состояний, синтезируемых сетями Петри в общем случае и специальными классами сетей Петри позволит разработать алгоритмы их синтеза. В контексте параметрических, временных, стохастических, флюидных стохастических сетей Петри планируется разработать и исследовать ряд отношений

поведенческой эквивалентности с целью редукции моделей с сохранением функциональности и производительности и верификации на основе логической характеристики эквивалентностей.

Разработка методов описания семантики языков программирования и технологии создания средств дедуктивной верификации программ; методов и средств моделирования и верификации распределенных систем; разработка методов пополнения систем организаций знаний.

Целью работы в 2017г. является:

Разработка методов исследования топологических пространств, используемых для вычислимого анализа; проведение теоретических исследований, связанных с разработкой методов анализа, верификации и синтеза параллельных и динамических моделей и моделей с временными, стохастическими и гибридными характеристиками; разработка технологии создания средств дедуктивной верификации программ; разработка методов и средств моделирования и верификации распределенных систем; разработка методов пополнения систем организаций знаний.

Для достижения цели были поставлены задачи:

- Изучение классов топологических пространств, допускающих «хорошую» дескриптивную теорию множеств и представляющих интерес для вычислимого анализа. Классическая дескриптивная теория множеств развита для случая Польских пространств. Ранее В.Л. Селивановым была развита существенная часть дескриптивной теории множеств для счетно-базируемых областей Ершова-Скотта. Недавно эта теория была обобщена М. де Брехтом на случай так называемых квази-Польских пространств. Планируется расширение этих результатов на более широкий класс пространств, включая пространства непрерывных функционалов Клини-Крайзеля.
- Исследование структур степеней по различным сводимостям на подмножествах топологических пространствах, а также их расширений на функции, в частности на  $k$ -разбиения топологических пространств. Предполагается исследование структуры степеней Вэджа борелевских  $k$ -разбиений Бэрвского и Канторовского пространства. Наряду с классической сводимостью Вэджа, в последнее время активно исследуются и применяются некоторые ее ослабленные варианты, введенные Вайраухом. Предполагается изучить стандартные вопросы определимости в начальных сегментах таких структур, в том числе вопросы разрешимости теории первого порядка и строения группы автоморфизмов. Планируется также поиск новых сводимостей, полезных для характеристики топологической сложности задач на топологических структурах.
- Исследование арифметической и дескриптивной сложности проблемы равенства мажорантно-вычислимых функций, заданных над эффективно перечислимыми топологическими пространствами. Реализация в iRRAM-пакете алгоритмов построения решений полиномиальных динамических систем с использованием

обобщенных моделей Тейлора. Разработка подходов к эффективному представлению непрерывных констрейнтов в iRRAM-пакете.

- Определение и исследование в различных семантиках и на основе различных поведенческих структур эквивалентностей флюидных и стохастических сетей Петри, а также построение логической характеристики эквивалентностей оригинальными флюидными модальными логиками.
- Определение и установление взаимосвязей «аксиом параллелизма» в контексте модели реляционных структур, обобщающей различные событийные модели параллельных процессов, что позволит лучше понять отношения между событиями и природу процессов, протекающих в параллельных/распределенных системах.
- Построение и сопоставление изоморфных абстракций изменения структуры и поведения в процессе функционирования широкого ряда моделей структур событий.
- Характеризация генерируемых сетями Петри пространств состояний, позволяющая осуществить минимальную аппроксимацию сверху конечных языков языками сетей Петри.
- Развитие онтологического подхода к формальной операционной семантике языков программирования и его апробация на семействе модельных процедурных языков программирования MPL и фрагменте языка C#.
- Применение символического метода элиминации инвариантов циклов для финитных итераций над массивами для верификации программ на языке C-light.
- Расширение языка спецификаций распределенных систем MSC временными конструкциями. Разработка системы анализа и верификации расширенных MSC-спецификаций.
- Разработка мультиагентного подхода к разрешению кореференции в процессе извлечения информации из текстов для пополнения онтологии.

Все запланированные работы были выполнены, обзор результатов 2017 года приведен в тексте данного отчета.

## **1. Логические и сложностные методы исследования систем**

Получены новые результаты о топологических и теоретико-множественных аспектах теории вычислений на несчетных структурах, значимые для дескриптивной теории множеств, общей топологии и теории вычислений.

Установлена тесная связь между конструктивными числовыми полями и упорядоченным полем действительных чисел и найдены приложения этих результатов к доказательству вычислимости решений важного класса дифференциальных уравнений в частных производных.

При этом установлена вычислимость важных задач линейной алгебры (таких как вычисление ортонормированного базиса собственных векторов вычислимых симметрических матриц и матричных пучков), играющих принципиальную роль в построении устойчивых разностных схем для изучаемых уравнений в частных производных (таких как схема Годунова). Установлена неразрешимость элементарных теорий многих решеток эффективно открытых множеств вычислимых топологических пространств и областей Ершова-Скотта. Для ряда конкретных важных примеров таких пространств установлена точная оценка сложности элементарных теорий решеток открытых множеств. Охарактеризована структура степеней Вэджа  $k$ -разбиений бэровского пространства для третьего уровня иерархии Бореля. Эта структура играет важную роль в характеристике топологической сложности таких задач как вычисление корней многочленов.

## **2. Развитие онтологического подхода к формальной операционной семантике языков программирования**

В рамках развития онтологического подхода к формальной семантике языков программирования предложены новое более общее определение концептуальных систем перехода (КСП) и двухэтапный метод применения КСП к разработке формальной онтологической операционной семантики языков программирования, базирующийся на новом определении КСП. Разработана формальная онтологическая операционная семантика для модельных процедурных языках программирования из семейства MPL и фрагмента языка C# как типовом представителе объектно-ориентированных языков.

КСП в новом определении являются специальным видом систем перехода, базирующемся на понятии концептуальных структур. Концептуальные структуры строятся из атомов с помощью операций последовательной композиции, абсолютной типизации и относительной типизации. Операция последовательной композиции строит из последовательности концептуальных структур составную структуру. Операция абсолютной типизации сопоставляет концептуальной структуре конечный набор типов, представляемых концептуальными структурами и называемых абсолютными типами. Эта операция категоризирует концептуальные структуры, разбивая их по категориям, представляемых абсолютными типами. Она также моделирует конструкторы экземпляров для этих категорий. Абсолютные типы соответствуют понятиям онтологии. Операция относительной типизации сопоставляет концептуальной структуре, входящей в составную структуру, конечный набор типов, представляемых концептуальными структурами и называемых относительными типами. Эта операция категоризирует концептуальные структуры в контексте составной структуры, элементами которой они являются, разбивая их по категориям, представляемых относительными типами. Она также моделирует доступ к подструктурам через их относительные типы.

КСП определяется как тройка, включающая множество атомов, из которых строятся концептуальные структуры, отношение перехода (бинарное отношение на множестве состояний), и множество начальных состояний. Элементы, состояния и значения в КСП представляются унифицированным образом с помощью концептуальных структур. Составные элементы, состояния и значения являются составными концептуальными структурами.

В соответствие с результатом доступа относительные типы делятся на атрибуты и мульти-атрибуты. Относительный тип является атрибутом КСП, если в любом состоянии результат доступа содержит не более одной подструктуры. В противном случае, относительный тип является мульти-атрибутом КСП. Атрибуты КСП соответствуют атрибутам онтологии.

Определены специальные виды КСП: КСП с возвращаемыми значениями, КСП с программами и КСП с явной остановкой. КСП с возвращаемыми значениями имеют атрибут value. Переходы (пары, состоящие из входного и выходного состояний) в таких системах возвращают значения. Значение, возвращаемого переходом, определяется значением атрибута

value в выходном состоянии. КСП с возвращаемыми значениями могут порождать исключительные ситуации (исключения). Исключение, порождаемое переходом, является возвращаемым значением этого перехода, имеющим абсолютный тип exception. Переход выполняется нормально, если он не порождает исключений. Элементы, которые инициируют переходы, называются выполняемыми элементами. КСП с программами имеют атрибут program. Значением этого атрибута являются составная структура, называемая программой. Программа и ее элементы являются выполняемыми элементами. Переходы инициируются последовательным выполнением элементов программы. Программы в КСП соответствуют программам в языках программирования и абстрактных машинах для этих языков. Состояние называется заключительным, если не существует перехода из этого состояния. КСП с явной остановкой имеют атрибут stop такой, что если значение атрибута stop в некотором состоянии определено, то это состояние является заключительным. Значение этого атрибута специфицирует причины остановки выполнения КСП.

В соответствие с новым определением КСП изменены синтаксис и семантика языка концептуальных систем перехода CTSL.

Двухэтапный метод применения КСП к разработке формальной онтологической операционной семантики языков программирования состоит в построении концептуальной модели абстрактной машины (АМ) для языка программирования (ЯП) и спецификации операционной семантики этой концептуальной модели на языке CTSL. В силу того, что операционная семантика строится для концептуальной (онтологической) модели АМ, операционная семантика ЯП, выражаемая в этой двухэтапной схеме, называется онтологической операционной семантикой.

Концептуальная модель АМ – это четверка, включающая множество элементов, представляющих типы ЯП, множество элементов, представляющих значения АМ (в частности, значения типов ЯП), множество состояний, представляющих состояния АМ, и множество выполнимых элементов, представляющих выполнимые конструкции АМ (в частности, элементы программ ЯП). Таким образом, этап построения концептуальной модели АМ состоит из 4 подэтапов, на которых разрабатываются представления типов, значений, состояний и выполнимых конструкций АМ на языке CTSL, соответственно.

Так как концептуальные структуры имеют конструктивное определение, КСП общего вида позволяют конструктивно определять концептуальные модели АМ. Однако, чтобы разработать операционную семантику концептуальных моделей АМ, требуется также уточнить структуру отношения перехода. Предложен специальный вид КСП – операционные КСП – которые уточняют структуру отношений перехода. Операционные КСП выделяют следующие специализированные отношения перехода: программные отношения перехода, отношения перехода, определяемые правилами перехода, и атомарные отношения перехода. Первые два вида отношений являются параметрическими, базируются на сопоставлении с образцом, и определяют переходы, инициируемые программами. Параметрами программного отношения перехода является образец и функция, которая сопоставляет экземплярам образца отношение перехода. Это отношение перехода выполняется в том случае, когда первый элемент программы сопоставляется с образцом. Параметрами правила перехода (и, соответственно, порождаемого им отношения перехода) являются образец и последовательность элементов, называемых телом правила. В случае сопоставления с образцом сопоставляемый элемент программы заменяется на соответствующий экземпляр тела правила. Предложены специальные виды правил переходов: с вычисляемыми и кватированными (невчисляемыми) переменными образца, условные (с условием на переменные образца и их значения), с просачиванием исключительных (исключения) и абнормальных (исключения или неопределенное значение) ситуаций по атрибуту value и переменным образца. Атомарные отношения перехода определены для перехода, в которых входные состояния не содержат программу или она пуста. Язык CTSL расширен для операционных КСП. Операционная семантика концептуальной модели АМ включает набор предопределенных программных и атомарных отношений перехода

и множество правил перехода, определяющих выполнимые элементы концептуальной модели (и, соответственно, выполнимые конструкции АМ, представлениями которых они являются).

В рамках решения задачи разработки онтологической операционной семантики модельных процедурных языков из семейства MPL предложены методология построения концептуальной модели АМ, включающей типизированные глобальные и локальные переменные, элементарные типы (int, nat, bool) и составные типы (массивы, структуры и указатели), и методология построения операционной семантики для стандартных операторов процедурных ЯП таких, как присваивание, блок, условный оператор, операторы цикла, операторы передачи управления (exit, return, continue, break), операций над массивами, структурами и указателями (включая new и delete), деклараций переменных, функций и структур.

В рамках решения задачи разработки онтологической операционной семантики для фрагмента языка С# предложена методология построения концептуальной модели для классов и их членов (полей, свойств, методов, индексаторов и т. д.) и методология построения операционной семантики для ряда операторов и выражений языка С# и нескольких алгоритмов АМ (просачивание исключений, разрешение перегрузки методов, определение динамического типа).

### **3. Система дедуктивной верификации С-программ**

Разработана и реализована новая версия системы верификации программ на языке C-light. Символический метод элиминации инвариантов циклов для финитных итераций над массивами применяется в этой системе. Метод содержит правило вывода для итерации без инвариантов, которое использует специальную функцию, выражающую действие тела цикла. Данное правило было реализовано в генераторе условий корректности, являющемся частью системы верификации C-light программ. Для доказательства порождённых условий корректности применяется метод математической индукции, вызывающий сложности у SMT-решателей. В нашей системе верификации на стадии доказательства используется SMT-решатель CVC4. Для преодоления упомянутой трудности применяется стратегия переписывания условий корректности. Для доказательства условий корректности предложен метод, основанный на расширении теории новыми теоремами.

### **4. Методы и средства моделирования и верификации распределенных систем**

Проблема анализа и верификации распределенных систем является актуальной для современного программирования. Для представления таких систем часто используются языки спецификаций MSC и UCM. Естественно возникает задача верификации спецификаций распределенных систем, представленных на этих языках. Временные характеристики играют важную роль в современных программных системах. Поэтому при проектировании систем полезно описывать временные требования с помощью расширения языка спецификаций MSC посредством временных конструкций. Предложено такое расширение, в котором временная задержка может быть задана для любого события или пары событий в MSC-диаграмме. Она представляет собой либо целочисленное значение, либо целочисленный интервал. Временные задержки могут быть абсолютными и относительными (относительно текущего времени глобальных часов). Разработана и реализована программная система анализа и верификации MSC-спецификаций, которые расширены временными конструкциями. Эта система включает

транслятор временных MSC-спецификаций во временные раскрашенные сети Петри и верификатор этих сетей Петри. Разработанная система была применена для анализа и верификации временных MSC-спецификаций некоторых коммуникационных протоколов.

Разработана и реализована новая версия программного средства анализа и верификации распределенных систем, представленных на языке спецификаций UCM.

## **5. Мультиагентные алгоритмы извлечения информации из текстов для пополнения онтологии**

Предложена онтология шаблонов требований к параллельным и распределенным программным системам. Эта онтология комбинирует известные виды шаблонов с новыми шаблонами, встречающимися при разработке и верификации программных систем. Онтология описывает требования, использующие булевские комбинации шаблонов следующих видов: качественные, линейного, ветвящегося и реального времени, содержащих составные события, количественные характеристики событий и простые утверждения о данных. Шаблоны требований имеют строгую семантику, выраженную формулами соответствующих темпоральных и модальных логик.

Разработан оригинальный мультиагентный подход к разрешению кореференции в процессе извлечения информации из входных данных для пополнения онтологии. Определены специальные агенты-классы, соответствующие классам онтологии. Разработаны мультиагентные алгоритмы означивания атрибутов экземпляров, обнаружения дубликатов и эквивалентов экземпляров, фиксирования кореференциальных отношений и определения веса информационных отношений. При разрешении кореференции агенты-классы учитывают многофакторную меру сходства извлеченных объектов. В процессе разрешения агенты-классы взаимодействуют друг с другом согласно иерархии классов онтологии, что позволяет снизить объем вычислений. Разработан широкий спектр характеристик и мер сравнения информационных объектов, основанных на их как онтологических, так и контекстных, свойствах. Эти меры улучшают качество работы агента-класса, позволяя устанавливать более точные референциальные связи между агентами-экземплярами.

Получены новые результаты о топологических и теоретико-множественных аспектах теории вычислений на несчетных структурах, значимые для дескриптивной теории множеств, общей топологии и теории вычислений.

Разработаны новые подходы к оценке дескриптивной и арифметической сложности вычислений над непрерывными данными и на их основе проведен анализ сложности непрерывных констрейнтов. Развита высшая эффективная дескриптивная теория множеств, т.е. ряд теорем эффективной ДТМ перенесены на некоторый класс эффективных пространств, строго содержащий классы вычислимых польских и вычислимых квази-польских пространств. Доказана знаменитая теорема Райса-Шапиро для вычислимых элементов модульных T-0-пространств.

В iRRAM-пакете разработаны и реализованы методы представления непрерывных констрейнтов, реализован алгоритм проверки истинности линейных констрейнтов. Разработаны новые подходы к вычислениям с обобщенными моделями Тейлора.

В контексте моделей флюидных стохастических сетей Петри (ФССП) определена эквивалентность на базе эквивалентностей стохастических сетей Петри, извлекаемых из комкуемости (lumpability) для цепей Маркова и некоторых классов дифференциальных уравнений. Показано, что комкуемые отношения эквивалентности можно получить как из дискретизированного пространства состояний стохастического процесса, так и непосредственно из ФССП. Предложены подходы к определению флюидных бисимуляционных

эквивалентностей мест ПФССП поднятием (lifting) бисимуляции непрерывных мест на гибридные маркировки и применением дуг детерминированного скачка флюида из расширенных ФССП. Кроме того, для флюидных следовой и бисимуляционной эквивалентности ФССП даны логические характеристики новыми флюидными модальными логиками. Построены прикладные примеры системы подготовки документов и производственной линии, демонстрирующие поведенческий анализ на основе категоризации по флюидной бисимуляционной эквивалентности, а также верификацию и сравнение линейно-временного и ветвисто-временного поведения с помощью формул упомянутых флюидных логик.

В рамках дискретно-временного стохастического исчисления боксов Петри с мгновенными мультидействиями (dtsiPBC) разработаны и сравнены структурная, построенная на системах переходов, и шаговая стохастическая бисимуляционная, используемая для сокращения соответствующих алгебраическим процессным выражениям поведенческих структур, таких как помеченные вероятностные системы переходов, полумарковские цепи и дискретно-временные цепи Маркова. С помощью интерпретации вероятностей стохастических мультидействий как параметров алгебраических спецификаций выяснен эффект изменения этих вероятностей на стационарную функцию масс вероятностей категоризованной полумарковской цепи и на соответствующие индексы производительности обобщенной в прикладном примере системы с разделяемой памятью.

В контексте реляционных структур, позволяющих описывать отношения слабой/сильной причинной зависимости, синхронного/асинхронного параллелизма, симметричного/асимметричного конфликта и т.д. между событиями параллельных/распределенных систем, введено понятие K-плотности, обобщающее различные модификации «аксиом параллельности», и дана его альтернативная характеристика.

Разработаны и исследованы методы построения изоморфных абстракций в виде систем переходов, представляющих изменения структуры и поведения при функционировании различных классов (расширенных первичных, расслоенных, потоковых, стабильных, дуальных, с разрешимым конфликтом, конфигурационных) событийно-ориентированных моделей.

Для произвольных конечных систем переходов над двоичным алфавитом построена характеристика в виде минимальных выпуклых оболочек множеств точек геометрического пространства состояний сетей Петри и нереклексивных сетей. На основе характеристики представлен алгоритм минимальной аппроксимации сверху конечных языков языками сетей Петри.

Предложены новое более общее определение концептуальных систем перехода (КСП) и двухэтапный метод применения КСП к разработке формальной онтологической операционной семантики языков программирования, базирующийся на новом определении КСП. Разработана формальная онтологическая операционная семантика для модельных процедурных языков программирования из семейства MPL и фрагмента языка C# как типовом представителе объектно-ориентированных языков.

Разработанная ранее система верификации программ на языке C-light расширена с целью применения символического метода элиминации инвариантов циклов для финитных итераций над массивами.

Язык спецификаций распределенных систем MSC расширен посредством временных конструкций. Разработана и реализована программная система анализа и верификации MSC-спецификаций, которые расширены этими конструкциями.

Разработаны и реализованы новые версии программных средств анализа и верификации распределенных систем, представленных на языках спецификаций SDL, UCM и MSC.

Разработан мультиагентный подход к разрешению кореференции в процессе извлечения информации из текстов для пополнения онтологии.

## Проекты РФФИ

**Проект РФФИ 15-01-05974** «Онтологический подход к формальной семантике языков программирования»

*Руководитель: Ануреев И.С.*

*Сроки: 2015-2017гг.*

## Список публикаций лаборатории

### Центральные издания

1. Anureev I.S., Promsky A.V. Conceptual transition systems and their application to development of conceptual models of programming languages // System Informatics, 2017, № 9, 133–154.
2. Anureev I.S. Operational conceptual transition systems and their application to development of conceptual operational semantics of programming languages // System Informatics, 2017, № 9, 155–200.
3. Гаранина Н.О., Зюбин В.Е., Лях Т.В. Онтологический подход к организации шаблонов требований в рамках системы поддержки формальной верификации распределенных программных систем // Системная информатика, 2017, № 9, 111-132.
4. Maryasov I.V., Nepomniaschy V.A., Kondratyev D.A. Invariant Elimination of Definite Iterations over Arrays in C Programs Verification // Моделирование и анализ информационных систем, Т.24, №6, 2017, 743–754.
5. I.V. Maryasov, V.A. Nepomniaschy, D.A. Kondratyev Verification of Definite Iteration over Arrays with a Loop Exit in C Programs // System Informatics, 2017, №10, 57–66.
6. Селиванов В.Л., Селиванова С.В. О конструктивных числовых полях и вычислимости решений дифференциальных уравнений в частных производных // Доклады РАН, Серия «Математика», 2017, 477, № 3, 278-281 (WoS, Scopus).

### Зарубежные издания

1. Natalia Garanina, Elena Sidorova, Irina Kononenko, Sergei Gorlatch. Using Multiple Semantic Measures For Coreference Resolution In Ontology Population // International Journal of Computing, 2017, V.16, № 3, 166-176 (Scopus)
2. N. O. Garanina and E. A. Sidorova. An Approach to Verification of a Family of Multiagent Systems for Conflict Resolution // Automatic Control and Computer Sciences, Allerton Press, 2017, V. 51, № 7, 498–506. (WoS, Scopus)
3. N. V. Vizovitin, V. A. Nepomniaschy, A. A. Stenko Application of Colored Petri Nets for Verification of Scenario Control Structures in UCM Notation // Automatic Control and Computer Sciences, Allerton Press, 2017, V. 51, № 7, 489–497 (WoS, Scopus).

4. Selivanov V.L., Kudinov O. First order theories of some lattices of open sets // Logical Methods in Computer Science, 13, № 3, 2017, 1-18 (Scopus).
5. Selivanov V.L. Towards a descriptive theory of cb0-spaces // Mathematical structures in computer science, 2017, V. 27, № 8, 1553-1580 (WoS, Scopus).
6. Selivanov V.L., Selivanova S.V. Computing solution operators of boundary-value problems for symmetric hyperbolic systems of PDEs // Logical Methods in Computer Science, V. 13, № 4, 2017, 1–31 (Scopus).

### **Материалы международных конференций**

1. Natalia Garanina. An Ontological Approach to a System of Requirement Patterns // Тезисы VIII Международного научно-исследовательского семинара «Программные семантики, спецификации и верификация» (PSSV-2017). Москва: МАКС пресс, 2017, 13-23.
2. Garanina N., Sidorova E., Kononenko I., Gorlatch S. Using Multiple Semantic Measures in a Framework for Coreference Resolution in the Process of Ontology Population // Proc. of the 26th International Workshop on CS&P. Warsaw, Poland, 2017 (Scopus).
3. Selivanov V.L. Extending Wadge theory to k-partitions // Proc. Int. conf. CiE 2017, LNCS 10307, 387-399, 2017 (Scopus).

### **Материалы российских конференций**

1. Кондратьев Д.А. Расширение системы C-light символическим методом верификации финитных итераций. – Труды XVII Всероссийской конференции молодых ученых по математическому моделированию и информационным технологиям // Вычислительные технологии, 2017, Т. 22, 44-59.
2. Сидорова Е.А., Гаранина Н.О., Кононенко И.С. Подход к разрешению референциальной неоднозначности текста при пополнении онтологии // Труды XXII Байкальской Всероссийской конференции «Информационные и математические технологии в науке и управлении». – 2017, № 3, 154-167.
3. Гаранина Н.О., Серый А.С., Сидорова Е.А. Мультиагентный подход к разрешению кореференции при извлечении информации из текстов на основе онтологии. // Материалы Всероссийской конференции с международным участием «Знания-Онтологии-Теории» (ЗОНТ-2017). Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017, Т.1, 88-98.
4. Серый А.С., Сидорова Е.А., Гаранина Н.О. Многодокументный анализ референции в процессе извлечения информации на основе онтологии предметной области // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы.

Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции, Новосибирск: ИВТ СО РАН, 167-173, 2017.

5. Черненко С.А., Непомнящий В.А. Верификация моделей распределенных систем, представленных временными MSC-диаграммами // Материалы 10-й Всероссийской мультikonференции по проблемам управления (МКПУ-2017). Ростов-на-Дону, Южный федеральный университет, Т. 3, 2017, 119-122.
6. Кондратьев Д. А. Верификация программ инженерной математики в системе C-light // Материалы XVIII Всероссийской конференции молодых ученых по математическому моделированию, ИВТ СО РАН, 2017, с. 78 .

### **Международное сотрудничество**

#### **Членство в редакциях международных журналов, другие формы сотрудничества Командировки**

##### **Членство в редколлегиях научных изданий**

Непомнящий В.А. – член редколлегии журналов “Системная информатика” и “Bulletin of the Novosibirsk Computing Center. Series: Computer Science”.

Ануреев И.С. – заместитель главного редактора журнала “Системная информатика”.

Селиванов В.Л. – член редколлегии журнала “Системная информатика”.

## **Общая характеристика исследований лаборатории конструирования и оптимизация программ**

*Зав лабораторией д.ф.-м.н., профессор Касьянов В.Н.*

### **Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе**

Цель исследований, ведущихся в лаборатории, - разработка методов и средств повышения качества матобеспечения вычислительных систем и сетей, главным образом его эффективности и надежности. Лаборатория ведет фундаментальные исследования, направленные на достижение данной цели, а также осуществляет экспериментальные и прикладные проекты, базирующиеся на разрабатываемых теоретических концепциях и методах.

#### **Результат, включенный в список основных результатов института за 2017 год**

**«Методы и система для визуализации сложных больших данных на основе атрибутированных иерархических графовых моделей»**

**Авторы:** д.ф.-м.н., профессор Касьянов В.Н., Золотухин Т. А., к.ф.-м.н., доцент Касьянова Е.В.

Разработаны методы и эффективные алгоритмы визуализации сложных больших данных на основе атрибутированных иерархических графовых моделей. Создана и прошла государственную регистрацию программная система Visual Graph, позволяющая с использованием атрибутированных иерархических графовых моделей исследовать сложные данные через их визуальные представления. Система поддерживает стандартные языки описания графов, включая язык GraphML, и обеспечивает плавность выполнения основных операций над графами, содержащими до 100000 элементов.

#### **Публикации по результату:**

1. Касьянов В.Н., Золотухин Т.А. Программная система для визуализации сложных больших данных на основе графовых моделей (Visual Graph). Свидетельство о государственной регистрации программы № 2017612824 от 03.03.2017.
2. Kasyanov V., Zolotuhin T. A system for structural information visualization based on attributed hierarchical graphs, WSEAS Transactions on Computers, 2017, Vol. 16, pp. 193-201.
3. Касьянов В.Н., Касьянова Е.В. Теоретико-графовые методы и системы программирования, Проблемы информатики, 2016, № 1, С. 26-38.

#### **Краткое описание проведенных научных исследований**

**Приоритетное направление IV.39.** Архитектура, системные решения, программное обеспечение и информационная безопасность информационно-вычислительных комплексов и сетей новых поколений. Системное программирование.

## **Программа IV.39.1.** Теоретические и прикладные проблемы создания эффективных надежных программных систем и информационных технологий

### **Проект IV.39.1.2.** Методы и технологии конструирования эффективного и надежного программного обеспечения для суперкомпьютеров и компьютерных сетей.

*Руководитель — д.ф.-м.н., профессор В.Н.Касьянов*

В процессе выполнения первого этапа работы были достигнуты все поставленные промежуточные цели проекта и получены следующие результаты.

1. Разработаны методы и создана экспериментальная версия среды параллельного и функционального программирования CSS для поддержки облачных супервычислений, ориентированная на работу с Cloud Sisal-программами и супервычислители Сибирского суперкомпьютерного центра СО РАН (ССКЦ СО РАН) и пригодная для учебных применений.

Цель работ – дать возможность широкому кругу лиц, находящихся в удаленных населенных пунктах или в местах с недостаточными вычислительными средствами, но имеющих выход в Интернет, дистанционно и без установки дополнительного программного обеспечения на своих недорогих вычислительных устройствах в визуальном стиле создавать и отлаживать переносимые параллельные программы на языке Cloud Sisal, а также затем дистанционно (в облаке) осуществлять эффективное решение своих задач, исполняя на некоторых супервычислителях, доступным им по сети, созданные и отлаженные переносимые Cloud-Sisal-программы, предварительно адаптировав их под используемые супервычислители с помощью облачного оптимизирующего кросс-компилятора, предоставляемого средой.

Среда CSS содержит интерпретатор, визуальный отладчик и облачный оптимизирующий кросс-компилятор. Она использует внутреннее теоретико-графовое представление функциональных и параллельных программ, ориентированное на их визуальную обработку и основанное на атрибутированных иерархических графах.

Разработанный входной язык системы (язык Cloud Sisal) продолжает традицию предыдущих версий языка Sisal], оставаясь функциональным потоковым языком, ориентированным на написание больших научных программ, содержащим циклы и массивы, и расширяет их возможности средствами визуальной поддержки облачных супервычислений.

Язык Cloud Sisal реализует модель всюду завершаемых частичных вычислений, что означает наличие специального «ошибочного» значения, просачиваемого по графу потока данных. Например, при вычислении матрицы часть элементов в результате может иметь ошибочное значение, что означает ошибку. Таким образом, семантика языка является изначально удобной для параллельных вычислений, и сложного контроля со стороны системы времени исполнения не требуется.

Язык Cloud Sisal допускает использование утверждений (прагм), описывающих известные программисту свойства программы в виде формализованных комментариев.

2. Предложены критерии оценки полноты покрытия тестами для проверки оптимизирующих компиляторов. Разработаны новые методы и алгоритмы автоматической генерации тестов для проверки распараллеливающих и векторизирующих преобразований циклов в компиляторе, ориентированные на полноту покрытия оптимизаций циклов тестами. Созданы программные средства проверки распараллеливающих и векторизирующих преобразований циклов в компиляторе для языка Cloud-Sisal.

Для генерации тестов разработан метаязык, который напоминает БНФ-нотацию и функциональный язык. Грамматика задается в одном файле и является набором строк с поддержкой комментариев. Правило – это идентификатор, для которого задано, каким образом он может быть переписан. Правило может быть как генерирующим, так и ограничивающим. Любой идентификатор – это последовательность символов, не начинающаяся с цифры. Имя параметра шаблона контекста в левой части правила может быть любым именем распознанной группы символов регулярного выражения языка Python, стоящей в предыдущем шаблоне контекста в левой части правила. Правая часть правила может задаваться либо альтернативами,

либо многострочным правилом. Для более тонкой работы с правилами можно использовать специальные распознаватели контекста, предоставляющие возможности по извлечению объектов из контекста и по их переименованию, а также по изменению самого контекста для всех идентификаторов в правой части правила. Помимо этого для описания и распознавания контекста можно использовать так называемые условия – специального вида булевские выражения. Также в языке грамматики предусмотрено несколько встроенных функций для удобной работы со списками, условиями и числами. Задаваемая при этом параметрическая КС-грамматика по наглядности сопоставима с обычной КС-грамматикой, но позволяет описывать более широкий класс КЗ-языков.

Создана система ATG (Automatic Test Generator), которая позволяет в автоматическом режиме тестировать компилятор при помощи порождаемых генератором тестов следующим образом. Каждый сгенерированный тест проходит две проверки – при компиляции с оптимизациями и без. Если тест успешно компилируется при помощи компилятора с отключенными оптимизациями и включенной проверкой указателей и запускается, то он является статически и динамически детерминированной программой, и результат ее исполнения объявляется эталонным. Здесь стоит сделать уточнение насчет успешного запуска: успешный запуск – это отсутствие сообщений о проблемах с указателями и нулевой код завершения, а также отсутствие завершения программы по установленному тайм-ауту. Это первая ступень проверки. Следующая ступень, это проверка компиляции с оптимизациями. Если тест вызвал на данном этапе проблемы, то он уже полезен для тестирования и сохраняется для дальнейшей обработки. Если компиляция прошла успешно, то проверяется корректность исполнения программы с ее эталоном. И если программа проваливает эту проверку, то она также сохраняется. В случае, когда программа проходит проверку с эталоном, она просто удаляется и запускается процесс генерации нового теста. Сохраненные тесты проходят дальнейшую обработку при помощи нашей системы упрощения компиляторных тестов Reduce. Также производится попытка найти компоненту компилятора, виновную в проблеме, при помощи функционала компилятора, который позволяет отключать применяемые оптимизации. Затем тесты сохраняются в базу системы с пометкой, на какой системе, когда и с какой проблемой они были получены. С целью оптимизации места, занимаемого хранилищем тестов, сохраняются 10 минимальных по размеру тестов на каждую проблемную оптимизацию. Но если возможно найти дату проблемной сборки, то это позволяет сохранять по 10 тестов на каждую дату.

3. Осуществлено расширение словаря по теории графов WikiGRAPP и энциклопедии теоретико-графовых алгоритмов WEGA новыми базовыми терминами и алгоритмами, в том числе связанными с кодированием и обработкой деревьев, а также произведено пополнение существующих статей словаря и энциклопедии дополнительной информацией, в том числе статическими и динамическими иллюстрациями.

Среди новых аналитических статей, расширивших энциклопедию, выделим следующие: «В-деревья без явного задания параметров кэша», «Алгоритм DC-дерева для k серверов на деревьях», «Взвешенное связное доминирующее множество», «Географическая маршрутизация», «Компоновка схемы», «Критический диапазон для беспроводных сетей», «Лес Штейнера», «Миграция данных», «Модель без явного задания параметров кэша», «Модель ввода-вывода», «Обмен ближайшими соседями и относительные расстояния», «Обобщенная двухсерверная задача», «Параллельное программирование, взаимное исключение», «Построение тонкослойной филогенетической сети», «Преобразование Барроуза-Уилера», «Причинно-следственное упорядочение, логические часы, репликация конечного автомата», «Распределение каналов и маршрутизация в беспроводных ячеистых мультирадиосетях», «Ресинхронизация схемы: инкрементный подход», «Связное доминирующее множество», «Сжатие и индексация дерева», «Сортировка без явного задания параметров кэша».

Разработан пакет программ, расширяющий MediaWiki средствами подготовки динамических иллюстраций работы теоретико-графовых алгоритмов обработки деревьев.

4. Разработаны новые методы и алгоритмы визуализации информации на основе атрибутированных иерархических графовых моделей.

Создана и прошла государственную регистрацию программная система Visual Graph, предназначенная для визуализации сложных больших данных на основе иерархических графовых моделей большого размера.

Система Visual Graph работает под управлением ОС Windows, Linux и MacOS, поддерживает обработку произвольных атрибутированных иерархических графов (в том числе составных и кластерных графов), ориентирована на визуализацию структур данных, возникающих в компиляторах, и позволяет одновременно работать с ними как в графовой, так и в текстовой форме. Система предоставляет богатые возможности для навигации по графовой модели, работы с атрибутами ее элементов, а также настройки системы на нужды конкретного пользователя. Система поддерживает стандартные языки описания графов, включая язык GraphML, и обеспечивает плавность выполнения основных операций над графами, содержащими до 100000 элементов.

Проведено исследование существующих визуальных техник описания поведения алгоритмов на графах. Разработаны эффективные методы и созданы программные средства поддержки динамической визуализации процессов исполнения теоретико-графовых алгоритмов, позволяющие совмещать визуализацию шагов алгоритма с их визуальными эффектами.

### **Результаты работы по грантам**

**Грант РФФИ 15-07-020029** «Методы и средства функционального программирования для поддержки облачных супервычислений»,

*Руководитель — д.ф.-м.н., профессор В.Н. Касьянов*

**Грант РФФИ 16-01-00499** «Применение методов теории графов для анализа структурной информации»

*Руководитель — д.ф.-м.н., профессор В.А. Евстигнеев*

### **Список публикаций лаборатории**

#### **Журналы и патенты**

1. Касьянов В.Н., Золотухин Т.А. Программная система для визуализации сложных больших данных на основе графовых моделей (Visual Graph). Свидетельство о государственной регистрации программы № 2017612824 от 03.03.2017.

#### **Статьи в рецензируемых журналах**

1. Панкратов С.Б. Автоматическая генерация тестов для GFX-offload компилятора Intel // Проблемы информатики. – 2017. – N 2– С. 13-23.
2. Kasyanov V.N., Kasyanova E.V. Methods and Tools of Parallel Programming // CEUR Workshop Proceedings. - 2017. - Vol. 1839. - pp. 141-154.

3. Kasyanov V., Zolotuhin T. A system for structural information visualization based on attributed hierarchical graphs // WSEAS Transactions on Computers. – 2017. – Vol. 16. – pp. 193-201.
4. Kasyanov V., Kasyanova E. A visual cloud system for parallel and functional programming teaching and learning // WSEAS Transactions on Computers. – 2017. – Vol. 16. – pp. 208-215.
5. Kasyanov V.N., Zolotuhin T.A. A system for visualization of big attributed hierarchical graphs // Information Technology in Industry. – 2018. – 6 с. (в печати)
6. Гордеев Д.С. Обзор техник визуализации алгоритмов на графах // Научная визуализация. – 2018. – 33 с. (в печати)
7. Kasyanov V.N., Kasyanova E.V. A system of functional programming for supporting cloud supercomputing // WSEAS Transactions on Information Science and Applications. – 2018. – 10 с. (в печати)
8. Kasyanov V.N., Zolotuhin T.A. A system for big attributed hierarchical graph visualization // WSEAS Transactions on Computers. – 2018. – 5 с. (в печати)

### **Материалы международных конференций**

1. Касьянов В.Н. Теоретико-графовые методы и системы программирования // Языки программирования-2017. Труды конференции. - Ростов-на-Дону: Изд-во Южного федерального университета, 2017. - С.129-133.
2. Касьянов В.Н., Касьянова Е.В. Методы и средства дистанционного обучения программированию // Материалы XVII Международной конференции "Информатика: проблемы, методология, технология". – Воронеж: Вэлборн, 2017. – Том 5. –С. 343-348.
3. Касьянов В.Н., Панкратов С.Б. Автоматическая генерация тестов для оптимизирующего компилятора // Материалы XVII Международной конференции "Информатика: проблемы, методология, технология". – Воронеж: Вэлборн, 2017. – Том 1. –С. 65-69.
4. Касьянова Е.В., Касьянова С.Н. Использование кластеров при решении задач томографии // Материалы XVII Международной конференции "Информатика: проблемы, методология, технология". – Воронеж: Вэлборн, 2017. – Том 4. –С. 195-200.
5. Касьянова Е.В. Методы и средства дистанционного обучения программированию // Языки программирования-2017. Труды конференции. - Ростов-на-Дону: Изд-во Южного федерального университета, 2017. - С.134-136.
6. Касьянов В. Облачные средства поддержки параллельного программирования // Международная конференция «Актуальные проблемы чистой и прикладной математики, посвященная 100-летию со дня рождения академика Тайманова Асана Дабсовича. - Алматы: ИМММ, 2017. - С. 133-135.
7. Касьянов В.Н. Российская информатика в лицах: академик Андрей Петрович Ершов // Труды Всероссийской научно-практической конференции (с международным участием) «Гуманитарное образование и наука в техническом вузе». –Ижевск: Изд-во ИжГТУ, 2017. – С. 124-128. - ISBN 978-5-7526-0781-3

### **Международное сотрудничество**

#### **Командировки (в том числе инициативные, не оплачиваемые Институтом)**

1. Касьянов В.Н. (24.09.17-01.10.17) - участие с докладом в Международной конференции APSAC-2017, г. Дубровик, Хорватия;
2. Касьянова Е.В.(24.09.17-01.10.17) - участие с докладом в Международной конференции APSAC-2017, г. Дубровик, Хорватия.

#### **Членство в международных научных организациях**

Касьянов В.Н. – член Американского математического общества.

**Участие в международных программах сотрудничества,  
зарубежные гранты, членство в редакциях международных журналов,  
другие формы сотрудничества**

Касьянов В.Н. — член редколлегии международного журнала «Проблемы программирования», г. Киев.

## **Общая характеристика исследований лаборатории искусственного интеллекта**

*Зав лабораторией к.т.н. Загорулько Ю.А.*

### **Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе**

**Приоритетное направление IV.39.** Архитектура, системные решения, программное обеспечение и информационная безопасность информационно-вычислительных комплексов и сетей новых поколений. Системное программирование

**Программа IV.39.1.** Теоретические и прикладные проблемы создания эффективных надежных программных систем и информационных технологий

**Проект IV.39.1.4.** Методы и технологии создания интеллектуальных информационных систем и систем поддержки принятия решений

Номер проекта в ИСГЗ ФАНО 0317-2014-0004

#### **Ответственные исполнители блоков проекта:**

**Блок 1:** к.т.н., заведующий лабораторией Загорулько Ю.А.

**Блок 2:** к.ф.-м.н., с.н.с. Сидорова Е.А.

**Блок 3:** к.т.н., заведующий лабораторией Загорулько Ю.А., н.с. Загорулько Г.Б.

Объектами исследования в данной работе являются интеллектуальные информационные системы поддержки научной и производственной деятельности; онтологии предметных областей; интернет-ресурсы; тексты на естественном языке; семантические словари; интеллектуальные системы поддержки принятия решений.

Цель работы: разработка методов, программных средств и технологий создания и сопровождения интеллектуальных информационных систем; разработка методов и программных средств построения и пополнения онтологий; разработка методов и программных средств анализа текста и извлечения информации; разработка методологии и программных средств построения интеллектуальных систем поддержки принятия решений.

Методами проведения работы являются: онтологический подход к проектированию и разработке информационных систем; методы морфологического, семантического и жанрового анализа текста; машинного обучения; методы онтологического инжиниринга.

В процессе выполнения проекта были получены следующие результаты:

- В рамках разработки технологии создания и сопровождения интеллектуальных информационных систем (ИИС) поддержки научной, производственной и образовательной деятельности разработаны новые методы интеграции таких ИИС, построенных с использованием технологий Semantic Web, с облаком Linked Open Data (LOD). Разработаны новые методы пополнения контента ИИС данными из облака LOD и публичных веб-сайтов. Исследованы вопросы использования паттернов онтологического проектирования для разработки онтологий научных предметных областей (ПО). Разработаны наборы структурных и содержательных паттернов, с помощью которых могут единообразно задаваться основные сущности онтологий различных научных ПО.
- В рамках развития технологии анализа текста разработаны новые семантические модели и методы, используемые для извлечения информации из текстов. Предложена модель

представления предметной лексики, позволяющая описывать сложную семантику терминов, и осуществлена ее интеграция в модель представления факта. Разработаны новые методы жанрового анализа текстов. Разработан ряд лингвистических ресурсов, предназначенных для решения задач классификации и извлечения информации из текстов технической документации и медицинских интернет-источников. Предложен подход к синхронизации аннотированных текстов с их переводами на другие языки, а также с сопутствующими мультимедийными материалами.

- В рамках создания технологии построения интеллектуальных систем поддержки принятия решений разработана методология информационно-аналитической поддержки (ИАП) разработчиков СППР, которая представляет собой систему принципов, подходов и способов организации теоретической и практической деятельности по обеспечению ИАП и использованию её средств. Разработан сервис, реализующий метод недоопределенных вычислений и метод рассуждений на основе экспертных правил с использованием клиент-серверной архитектуры.

### Краткое описание проведенных научных исследований

#### Блок 1. Разработка методов, программных средств и технологии создания и сопровождения интеллектуальных информационных систем поддержки научной, производственной и образовательной деятельности.

В рамках развития технологии создания интеллектуальных информационных систем (ИИС) поддержки научной деятельности, построенных с использованием технологий Semantic Web, исследованы существующие и разработаны новые методы интеграции такого типа ИИС с облаком данных Linked Open Data (LOD).

Разработанные методы интеграции предполагают конструирование правил отображения сущностей онтологии ИИС на сущности выбранной онтологии из ресурса LOD и последующее связывание данных ИИС с соответствующими данными из LOD с помощью стандартной конструкции языка OWL *owl:sameAs*.

В качестве основной онтологии LOD для интеграции данных ИИС выступала DBpedia. Таким образом, правила отображения устанавливали соответствие между сущностями этих двух онтологий.

В простейшем случае правила устанавливают соответствия между конкретными классами онтологии ИИС и DBpedia. Например, класс *Проект* онтологии ИИС соответствует классу *dbo:ResearchProject* в DBpedia.

Для более сложных случаев задаются правила соответствия классов, атрибутов и отношений. В частности, с помощью правил можно ограничить возможные экземпляры класса теми, у которых значение определенного атрибута равно конкретному значению. Например, на Рис. 1 представлено правило отображения множества экземпляров класса *Событие*, у которых значение атрибута *Тип события* равно “конференция”, на класс *dbo:AcademicConference*.



Рис. 1. Правило отображения класса *Событие*

С помощью правил можно также устанавливать функции преобразования значений атрибутов. Например, атрибут *foaf:name* ресурса DBpedia является конкатенацией атрибутов *Фамилия*, *Имя* и *Отчество* онтологии ИИС (см. Рис. 2).



Рис. 2. Правило преобразования атрибутов

Сконструированные таким образом правила соответствия классов используются при установлении соответствий между данными, представленными экземплярами классов (объектами).

Для того чтобы установить соответствие данных, применяется модификация разработанного ранее метода пополнения онтологии, опирающегося на поиск и сравнение экземпляров. Данный метод применяется для автоматического обновления контента ИИС, когда он пополняется объектами, извлекаемыми из внешних источников. При этом извлеченные данные могут уже содержаться в хранилище данных ИИС, дополнять информацию об имеющихся экземплярах онтологии или породить новые экземпляры.

Чтобы связать данные ИИС с данными LOD, необходимо для каждого экземпляра онтологии ИИС выполнить проверку на наличие в LOD его аналогов. Проверка проводится в несколько этапов.

**1. Поиск потенциальных аналогов объекта в онтологии DBpedia.** Поисквые запросы формируются по правилам соответствия классов (Рис. 3).

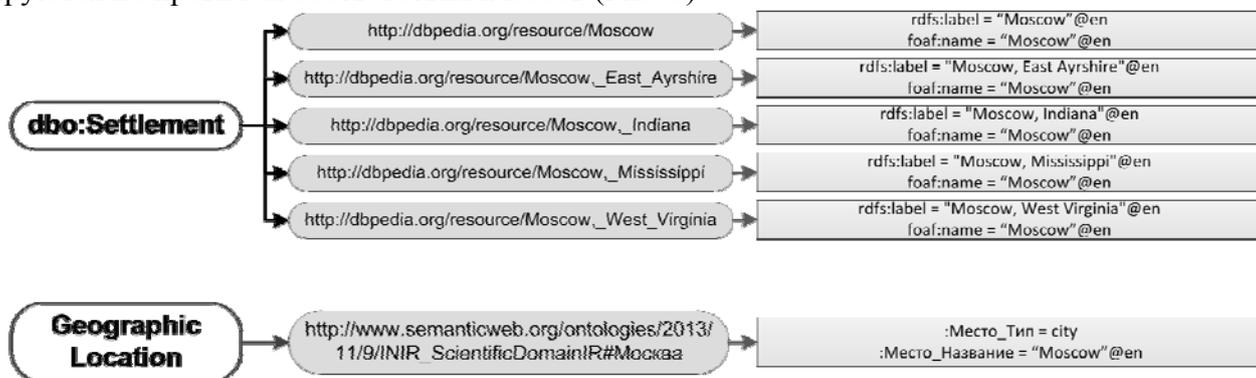


Рис. 3. Поиск потенциальных аналогов

**2. Сравнение объекта с потенциальными аналогами, поиск наиболее подходящего.** На этом этапе применяются формулы расчета семантического сходства объектов.

**3. Установление отношения *owl:sameAs* между объектом и его аналогом в том случае, если он существует (см. Рис. 4).** В противном случае объект будет считаться новым и в будущем может быть интегрирован в LOD как самостоятельная сущность.

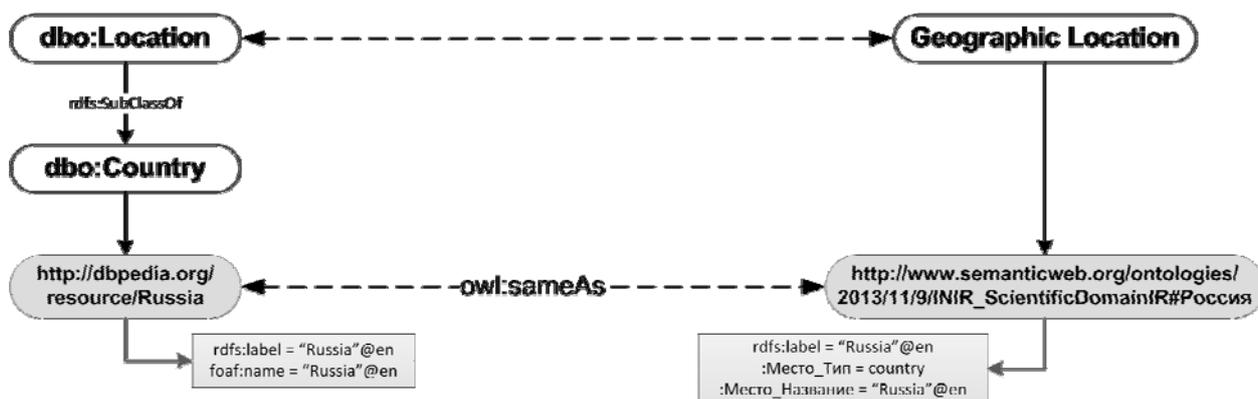


Рис. 4. Соответствие между экземплярами

Установление соответствия между данными является последовательным итеративным процессом (Рис. 5). Если на текущем шаге невозможно принять решение о соответствии объекта онтологии ИИС объекту из DBpedia, процесс вернется к нему на следующем шаге. Остановка происходит тогда, когда на текущем шаге не было установлено ни одного соответствия.

Данный метод применяется также для автоматического обновления контента ИИС, когда он дополняется объектами, извлекаемыми из внешних источников.

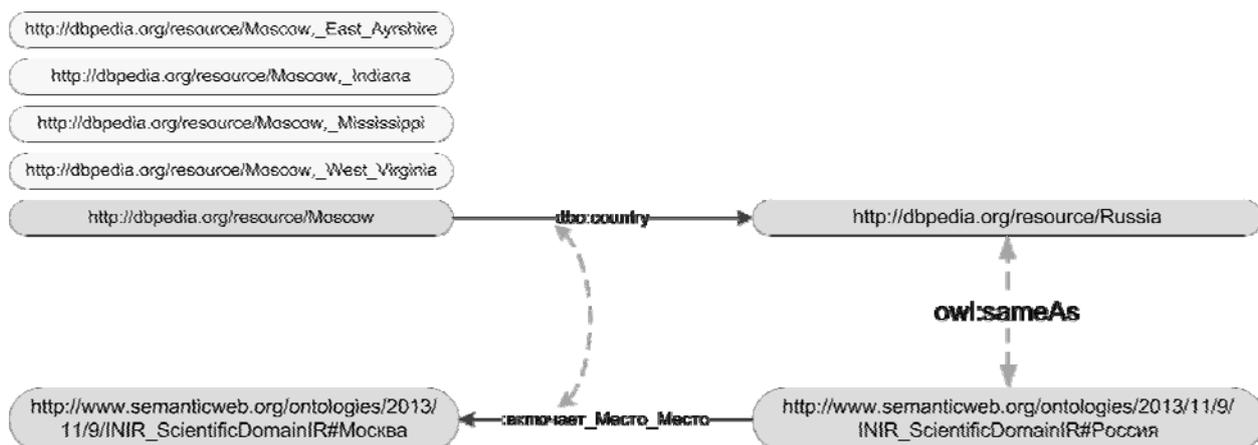


Рис. 5. Разрешение неоднозначностей соответствия

Разработаны новые методы пополнения контента ИИС поддержки научной деятельности данными из LOD и других публичных информационных ресурсов (веб-сайтов). В частности, построена онтология структуры веб-страниц; разработаны и реализованы методы сегментации веб-страниц, позволяющие выделять наиболее важные сегменты веб-страниц: Основной контент, Меню, Текущий пункт меню и т.п.; разработан и реализован метод сегментации веб-страниц с помощью лексических маркеров.

Исследованы вопросы использования паттернов онтологического проектирования (Ontology Design Patterns или ODP) для разработки онтологий научных предметных областей (ПО). Паттерны онтологического проектирования представляют собой документально зафиксированные описания проверенных на практике решений проблем онтологического моделирования. ODP позволяют описывать как типовые, так и специфические проблемы, возникающие при разработке онтологий, а также предлагаемые разработчиками рекомендации и соглашения для их решения.

Основной каталог паттернов онтологического проектирования представлен на портале Ассоциации ODP (Association for Ontology Design & Patterns), созданного в рамках проекта NeOn. В зависимости от проблем, для решения которых предназначены паттерны

онтологического проектирования, различают структурные паттерны (Structural ODPs), паттерны соответствия (Correspondence ODPs), паттерны содержания (Content ODPs), паттерны логического вывода (Reasoning ODPs), паттерны представления (Presentation ODPs) и лексико-синтаксические паттерны (Lexico-Syntactic ODPs).

В рамках выполнения данного проекта разработаны наборы структурных и содержательных паттернов, с помощью которых могут единообразно задаваться основные сущности онтологий различных научных ПО. Показано, что использование таких паттернов не только обеспечивает согласованное представление всех сущностей онтологии, но и позволяет экономить ресурсы и избегать ошибок при разработке онтологий.

## Блок 2. Разработка методов и программных средств извлечения информации из текстов на основе лингвистических моделей и ресурсов.

В рамках развития технологии фактографического анализа текстов были разработаны новые **семантические модели и методы**, используемые для извлечения информации.

Была расширена концепция построения семантических словарей и предложена новая модель представления предметной лексики (Рис. 6), позволяющая описывать структуру семантики терминов, которая позволяла бы сопоставлять текстовым единицам их смысловые эквиваленты. Модель включает грамматическую, тезаурусную и семантическую информацию о термине, а также необходимые данные для описания валентностной структуры предикатных слов.

The screenshot displays a software application window titled "Словарь предметной лексики - Анализ сценария\_v1.1.vc". The interface is divided into several sections:

- Table of Terms:** A table with columns "Лексема" (Lexeme) and "Часть речи" (Part of speech). The term "останавливаться" is highlighted, with its part of speech listed as "Глагол" (Verb).
- Term Details:** A section for the selected term "останавливаться", showing its normal form, morphological features (e.g., "Вид: нс, Переходность: нп"), and a paradigm of 41 forms.
- Semantic Tree:** A hierarchical tree structure for the word "Признак". The tree includes categories like "Значимые" (Significant), "Процесс" (Process), "Действие" (Action), and "Фаза" (Phase). The "Управление" (Control) category is highlighted, with sub-categories like "Управляющее действие" (Controlling action) and "Управляющее действие" (Controlling action).
- Alternative Groups:** A section showing alternative groups of terms, such as "{ Управляющее действие, sem. действие, sem. состояние, sem. отр-состояние }".
- Status Bar:** At the bottom, it shows statistics: "Терминов: 176 (176)", "СК: 26 (26)", "Стоп-терминов: 60 (60)", and "Стоп-СК: ...".

Рис. 6. Пример описания семантики термина.

Для кодирования семантической информации о слове предусмотрены следующие возможности:

а) Семантический класс. Термин может быть отнесен к определенному семантическому классу. Иерархия классов также позволяет отнести термин к определенному уровню иерархии, более общему или конкретному с наследованием свойств общего класса.

б) Семантический атрибут. Для представления лексического значения термина используются семантические атрибуты. Совокупность значений атрибутов, приписанных слову, в определенной мере моделирует компонентную семантическую структуру слова. Основные компоненты семантической структуры термина могут рассматриваться как тезаурусные дескрипторы.

в) Группировка семантических классов и атрибутов для описания многозначного слова. Если термин имеет более одного формально различимого контекстом значения, формируется соответствующее число семантических статей слова, объединяющих в себе с помощью механизма группировки семантический класс и совокупность семантических атрибутов с их значениями.

Разработан подход и осуществлена интеграция лексической модели термина со сложной семантикой в модель представления факта. С этой целью был реализован конвертор, обеспечивающий представление многозначного термина в виде лексического объекта с заданным набором формальных атрибутов, часть из которых фиксирована и опирается на внутреннюю модель описания термина в словаре, а часть формируется в соответствии с параметрами, заданными экспертом при описании семантики словаря.

Лексический объект наравне с информационным объектом может выступать в качестве аргумента при построении факта. В рамках модели могут задаваться грамматические и семантические ограничения на атрибуты объектов.

Структуру лексического объекта можно представить следующим образом:

### **LexObject**

```
Name: string;           // название термина
Descriptor: string;     // имя дескриптора
Semantic: set of LexClass; // множество лексико-семантических классов
Neg: bool;              // наличие отрицания
Attributes: set of Semantic_attribute; // множество атрибутов
Value: string;          // значение или форма, в которой встретился термин в тексте
Position: int;          // позиция термина в тексте
Grammatics: set of gram_parameters; // грамматические характеристики термина
```

Здесь параметр *Name* задает нормализованную форму термина, полученное либо в соответствии с правилами морфологической нормализации (в случае однословного термина), либо в соответствии с правилами согласования слов в словосочетании (для многословных терминов), либо как имя лексического шаблона (для символно-числовых конструкций). Параметр *Descriptor* определяет имя понятия, к которому относится термин. В частности, дескриптор служит для формирования синсетов или групп синонимов, когда все термины с одинаковым дескриптором являются синонимами или квазисинонимами в данной предметной области. Поле *Semantic* определяет лексико-семантический класс(ы) термина. Наличие нескольких значений в поле *Semantic* лексического объекта соответствует составному значению термина (например, сложносоставное слово *пятиметровый* имеет лексическое значение *Числа* и *Единицы измерения*) в отличие от неоднозначности, когда формируется несколько лексических объектов *LexObject* (например, по термину *останавливаться* будут созданы объекты с семантикой *Действия* и *Процесса*). Отметим, что морфологическая омонимия не

порождает отдельных лексических объектов, а разрешается, если необходимо, в рамках одного объекта. Поле *Neg* задает наличие отрицания при значении термина, которое в тексте может выражаться, например, приставкой *не*. Поле *Attributes* задает множество семантических атрибутов для представления структуры семантики термина. Поля *Grammatics*, *Position* и *Value* определяются словарными и текстовыми характеристиками термина.

Таким образом, методы фактографического анализа текста обогатились возможностью представления и обработки семантической структуры терминов, включая составные и многозначные термины с их сложной семантикой, что позволяет более полно использовать лексикографические ресурсы, разрабатываемые профессиональными лингвистами, и осуществлять более тонкий семантический анализ текста.

Разработаны новые модели и методы **тематического и жанрового анализа** текстов, взятых из интернет-источников.

Предлагаемые решения по анализу, классификации и извлечению информации из текстов опираются на понятие жанра как совокупности содержательных и формально-лингвистических (логико-композиционных и лексико-грамматических) аспектов. В рамках проведенных исследований по данной теме была разработана формальная жанровая модель текста, которая представляется системой вида:

$GM_T = \langle g, V_G, V_M, P_G, S \rangle$ , где

$g$  – жанровый тип документа,

$V_G$  – словарь жанровой лексики,

$V_M$  – словарь жанровых маркеров, которые задаются с помощью терминов из  $V_G$ ,

$P_G: V_M \rightarrow F_S$  – множество жанровых шаблонов, которые связывают маркеры  $V_M$  и структурные блоки текста  $F_S$ , где каждый блок  $fr \in F_S$  имеет тип и определяется начальной и конечной текстовой позицией,

$S = \langle F_S, R_S \rangle$  – логико-композиционная структура текста, определяемая набором структурных блоков  $F_S$  и их взаимосвязями  $R_S$ , задающими взаиморасположение в тексте.

Таким образом, жанровые особенности текста передаются его разбиением на содержательные блоки, которые: а) включают определенную жанровую лексику, б) имеют определенную структурную организацию, в) реализуются в рамках определенных формальных сегментов.

Логико-композиционная структура текста выявляется с помощью лексикона жанровых маркеров и шаблонов, выделяющих содержательные блоки. Маркеры извлекаются из текстов, и прежде всего – из заголовков подразделов, вводных предложений и списков, включенных в состав документа. Простые маркеры сопоставляются терминам предметного словаря или группе терминов (синонимы). Более сложные формируются на основе простых: поддерживаются альтернативы, совместная встречаемость, а также вложенное использование маркеров. Особенностью анализа интернет-текстов является опора на маркеры, представленные в тегах, и соответствующее выделение фрагмента.

Рассмотрим схему жанровой модели текста на примере стандартной xml-разметки протоколов клинических испытаний, представленных в интернет-источниках (см. Рис. 7). Данные тексты включают жанровые сегменты, маркированные жанровыми тегами, на основе которых могут извлекаться фрагменты текста для поиска той или иной информации.

<p><b>TextGenre</b>  <b>Block</b> genre_segment              <b>Block</b> genre_segment              ...              <b>Block</b> genre_segment              <b>Block</b> genre_segment              ...</p>	<pre> &lt;study_design_info&gt;   &lt;allocation&gt;Randomized&lt;/allocation&gt;   &lt;intervention_model&gt;Parallel Assignment&lt;/intervention_model&gt;   &lt;primary_purpose&gt;Prevention&lt;/primary_purpose&gt;   &lt;masking&gt;Double Blind (Participant, Care Provider, Investigator, Outcomes Assessor)&lt;/masking&gt; &lt;/study_design_info&gt; </pre>
---	--

Рис. 7. Фрагмент жанровой модели текста, описывающего дизайн клинического исследования.

В рамках развития технологии семантического аннотирования текстов были предложены модели и методы синхронизации аннотированных текстов с их переводами, а также аудио и видео ресурсами на основе онтологии предметной области.

Предложен подход к синхронизации аннотированных текстов с их переводами на другие языки, а также с сопутствующими мультимедийными материалами. Синхронизация осуществляется на основе разметки ресурсов понятиями, отношениями и свойствами онтологии предметной области. Единица аннотирования представляет собой связь фрагмента текстового ресурса и элемента онтологии, к которым в нашем подходе относятся: концепт (*Class*), экземпляр класса, тип атрибута (*Data Property*), значение атрибута (значение *Data Property*), тип отношения (*Object Property*) и экземпляр отношения (значение *Object Property*). Помимо аннотирования выделяется еще один метод разметки — синхронизация ресурсов. В случае синхронизирующей разметки, ресурс разбивается на фрагменты (связь между фрагментом и элементом онтологии при этом может отсутствовать), которые связываются с фрагментами разметок или аннотаций других ресурсов. Мультимедийные ресурсы размечаются аналогично текстовым, отличие заключается лишь в способе выделения фрагмента: границы текстового интервала указывают начальный и конечный символы в тексте, границы мультимедийного — начальную и конечную секунду воспроизведения или границу выделенного графического элемента в случае изображения.

Таким образом, разработанная ранее онтология аннотирования и разметки текстовых ресурсов была расширена отношениями синхронизации между фрагментами ресурсов или множествами размеченных фрагментов — разметками. Один ресурс может быть размечен более чем одним способом, т.е. иметь несколько независимых разметок. Синхронизация разметок и/или аннотаций различных ресурсов и связывание фрагментов с элементами онтологии дает возможность быстрого доступа и анализа информации об интересующем объекте предметной области, распределенной по нескольким разнородным ресурсам или представленной на разных языках.

Данный подход апробирован на электронном портале «Фольклор народов Сибири». Оригинальные тексты фольклорных произведений на национальных языках построчно синхронизированы с переводами на русский язык (Рис. 8).

The screenshot shows the 'Электронный портал "ФОЛЬКЛОР НАРОДОВ СИБИРИ"' (Electronic portal "FOLKLORE OF PEOPLES OF SIBERIA"). The main content is a list of folk songs with their original titles in Cyrillic, line numbers, and Russian translations. Below the list is a media player with several video thumbnails. A metadata box on the right provides information about the researcher and the resource.

Original Title	Line Number	Translation
Урянай-ханай н-уриянай	1	В давнюю давность
Урин сагаан сагта	2	В благодатное светлое время*
Тохонойн ганса Төлэн	3	Тохоной Ганса Төлэн
Байга бэту юумайдаа.	4	Жил в ту пору.
Ганса хадаа хүн	5	«От одного человека
Хүүн болохо ёно үбээ,	6	Род не пойдёт,
Ганса хадаа сусал	7	От одной головёшки
Гал болхо ёно үбээ,	8	Огонь не запыляет».*-
Гэжэн-ханай хэлэжэ	9	Так, говорят, сказав,
Түхөөрбэ гэшүүдаа.	10	Собираться он стал.
Түхөөрэлэн биеэрээ:	11	Сборы его такими были:
Баруун-ханай гараараа	12	Правой своей рукой
Дайни тоного түхөөрбэ,	13	Снаряжение для войны собирал,
Зүүн-ханай гараараа	14	Левой своей рукой -
Морин байин тоног	15	Снаряжение для коня.
Түхөөрэлэн биеэрээ,	16	Сборы его такими были:
Далан бугайн н-эбэрээр	17	Из рогов семидесяти изюбрей
Номо-ханай хэбүүдаа,	18	Лук он сделал,
Далан буган шандааһаар	19	Из жил семидесяти изюбрей

Metadata box content:  
 Исследователь: Гомбожапов А.Г.  
 Название ресурса: Иван Лубсанцренович Дабаев  
 Дополнительная информация: современный сказитель-улигершин, род. в 1939 г., живёт в с. Хурашша Иволгинского района Республики Бурятия

Рис. 8. Синхронизация ресурсов на портале "Фольклор народов Сибири".

Аннотированные фрагменты текстовых и мультимедийных ресурсов, таких как фотографии, аудио- и видеозаписи, связаны посредством двух предметных онтологий: «Героический эпос бурят» и «Песни медвежьего праздника хантов». Информация о связанных фрагментах доступна для пользователей портала (<http://infra.iis.nsk.su/folklore>).

При апробации разрабатываемых ЕЯ-технологий были разработан ряд **лингвистических ресурсов**, предназначенных для решения задач классификации и извлечения информации из текстов технической документации.

Разработан базовый семантический словарь для извлечения информации из технической документации. Универсальная часть словаря содержит классы, соответствующие универсальным категориям: *Объект, Время, Пространство*. В классе *Объект* далее выделяются подклассы *Предмет, Вещество, Персона, Организация* и т.п. Более конкретная таксономия определяется требованиями конкретной предметной области. Еще одна ветвь иерархии соответствует классам предикатной лексики, соответствующей обозначениям различных типов ситуаций (класс *Ситуация*), участниками которых являются объектные сущности: *Состояния, Процессы, Действия*. Дальнейшая классификация определяется тематическими классами предикатов: *Физическое воздействие, Перемещение, Ментальные, Восприятие, Речевое воздействие, Количественное изменение* и т.п.

В рамках данного словаря предложена универсальная модель для извлечения параметрической информации, представляющей оценки значений количественных параметров объектов (сущностей и событий). Эта информация характерна для многих предметных областей и представлена в текстах различных жанров – деловых и технических текстах, научных статьях, сообщениях отраслевых СМИ, медицинской документации. Известная трехкомпонентная модель *<Объект, Параметр, Значение>*, лежащая в основе базовой ситуации параметрической оценки, расширена для описания всего спектра ситуаций сопоставления величин. Так, *Абсолютная оценка*, которая соотносит значение параметра с числовой шкалой или с некоторым эталонным значением, требует уточнения компонента *Значение*: *<Оператор, Константа>* с соответствующим доопределением типа *Константы*: *<Число, Мера>* или *<Эталон>*. Включение в рассмотрение *Сравнительных оценок* потребовало трехчастной фреймовой структуры с дублированием базовой схемы и вводом дополнительного компонента *<Функция, Разность>*, соответствующего собственно сравнению. Вариантом трехчастной схемы являются ситуации *Динамической оценки*, в которой в описания двух сравниваемых по *Параметру Объектов* вводится темпоральный компонент, а соответствующие *Значения* представляются как *Начальное* и *Конечное*.

Создан аннотированный корпус текстов технической документации, на основе которого разработан тематический словарь, предназначенный для определения поджанра технической документации.

### **Блок 3. Разработка методов и технологии построения интеллектуальных систем поддержки принятия решений.**

#### **3.1. Разработка методологии информационно-аналитической поддержки разработчиков СППР.**

Была разработана методология информационно-аналитической поддержки разработчиков СППР, которая описывает подход к организации такой поддержки, её методы и средства, а также порядок их использования. Согласно методологии, средством такой поддержки является информационно-аналитический ресурс (ИАИР) по поддержке принятия решений (ППР). На Рис. 9 представлены основные действия, которые могут осуществляться на ресурсе его создателем и пользователем.

Ресурс был создан с использованием технологии разработки тематических интеллектуальных интернет-ресурсов, ранее предложенной в ИСИ СО РАН. При его создании на сервере устанавливается оболочка будущего ресурса. Затем, с помощью блока настроек

задаются элементы пользовательского интерфейса – название ресурса, его цветовая гамма, «шапка». Дальнейшие шаги по созданию ресурса сводятся к разработке его онтологии и наполнению контента.

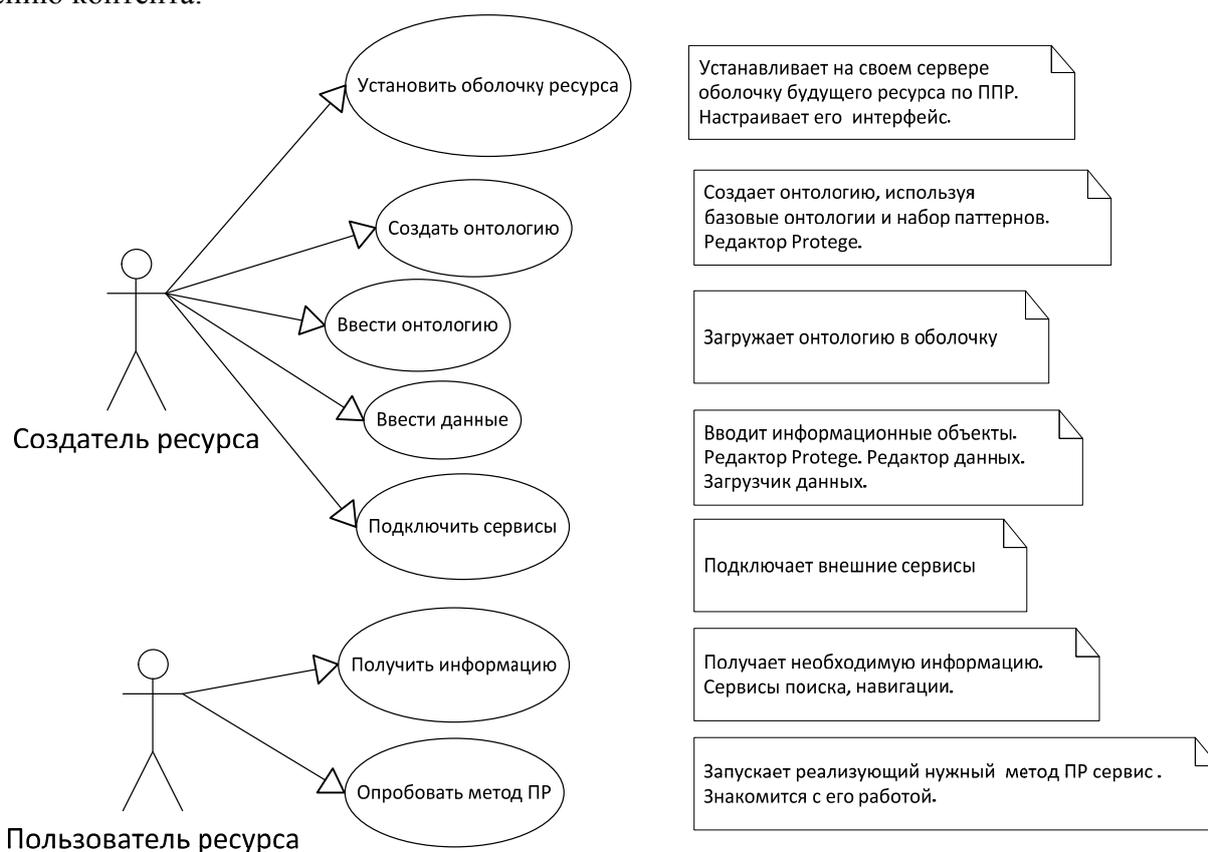


Рис. 9. Варианты работы с ресурсом.

**Разработка онтологии.** Онтология является концептуальной основой любого ИАИР, созданного на основе вышеупомянутой технологии. Она служит для формализации и систематизации различных видов знаний, данных, информационных ресурсов, средств обработки и анализа информации, интегрируемых в информационное пространство ресурса, а также для организации удобного содержательного доступа к ним.

Разработка таких онтологий осуществляется на основе небольшого, но представительного набора базовых онтологий, включающих только самые общие сущности, не зависящие от конкретных областей знаний. В этот набор входят: (1) онтология научного знания, (2) онтология научной деятельности, (3) базовая онтология задач и методов, (4) базовая онтология научных информационных ресурсов. Данные онтологии содержат понятия и отношения, описывающие, соответственно область знаний ИНИР, осуществляемую в ее рамках деятельность, решаемые в ней задачи и методы их решения, относящиеся к ней информационные ресурсы. Разработка конкретной онтологии заключается в конкретизации и дополнении содержащихся в базовых онтологиях понятий.

Для упрощения и упорядочения этого процесса, а также снижения возможности внесения ошибок были использованы три вида паттернов онтологического проектирования: структурные логические, паттерны содержания, паттерны представления.

**Структурные логические паттерны** фиксируют способы решения проблем представления элементов онтологии, вызванные ограничениями выразительных возможностей (экспрессивности) языков описания онтологий. Базовые онтологии, и онтологии, создаваемые на их основе, разработаны на языке OWL. При построении онтологии ИАИР ППР возникла необходимость в использовании структурных логических паттернов для представления таких сложных конструкций, как области допустимых значений, атрибутированные и многоместные

отношения, отношения между объектом и классом. На Рис. 10 показан паттерн бинарного атрибутивного отношения и пример его использования.

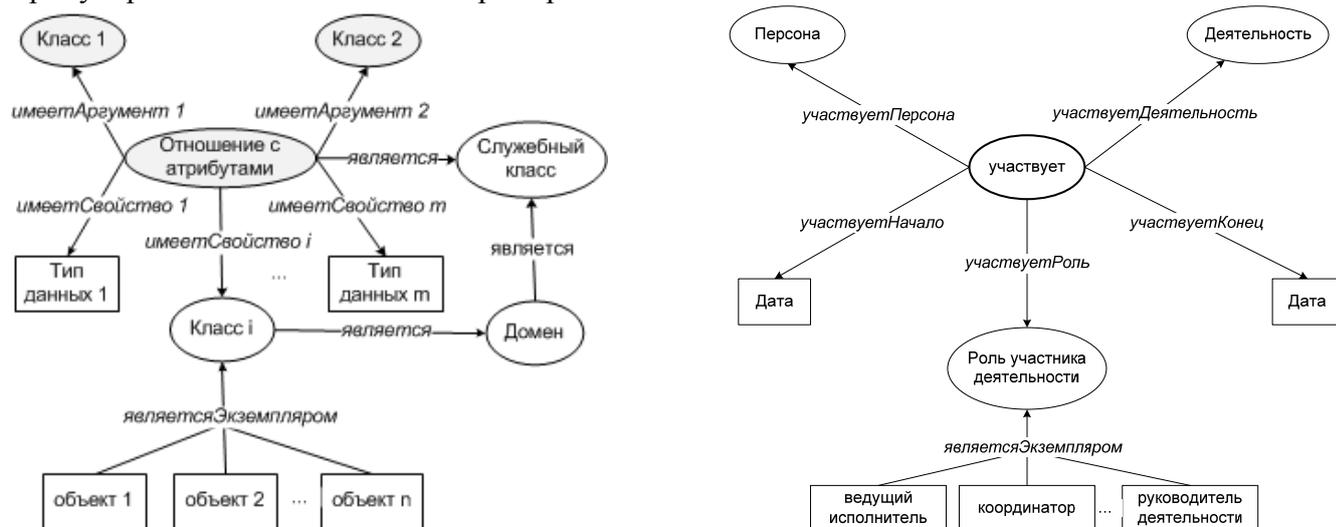


Рис. 10. Паттерн отношения с атрибутами и пример его использования.

При создании онтологии области знаний ИАИР ППР важно единообразно и согласованно представлять используемые в базовых онтологиях научные понятия и их свойства. Для решения этой проблемы требуются **паттерны содержания**, задающие способы представления типовых фрагментов онтологий, на основе которых могут строиться онтологии целого класса предметных областей. При разработке онтологии области знаний ППР используются паттерны содержания, сложившиеся ранее в ходе выполнения проектов по созданию научных интернет-ресурсов для ряда научных предметных областей. На основе таких паттернов строятся базовые онтологии научной деятельности и научного знания. Эти паттерны, служащие для представления научных результатов, методов, объектов и предметов исследования, выполняемой деятельности и др., являются общими для многих научных предметных областей. Такие паттерны, в свою очередь, могут быть частью более сложных (составных) паттернов содержания.

Паттерны содержания могут строиться из структурных паттернов, которые, по сути, являются предметно-независимыми. Так, например, содержательные паттерны для представления участия персоны или организации в какой-то деятельности или каком-то событии задаются с помощью структурного паттерна, служащего для представления бинарного атрибутивного отношения (Рис. 10).

В паттерны содержания включаются не только представляемые с помощью них понятия, но и понятия из «смежных» паттернов, а также паттернов других базовых онтологий, что позволяет задавать связанное описание моделируемой области.

Фактически, паттерны содержания являются фрагментами базовых онтологий, которые после конкретизации (специализации) содержащихся в них понятий и дополнения новыми понятиями становятся составными частями онтологии области знаний ППР.

На Рис. 11 представлен фрагмент онтологии области знаний ППР, являющийся специализацией паттернов для представления элементов базовых онтологий и иллюстрирующий взаимосвязи между паттернами, задающими понятия *Метод ППР* и *Задача принятия решений*.

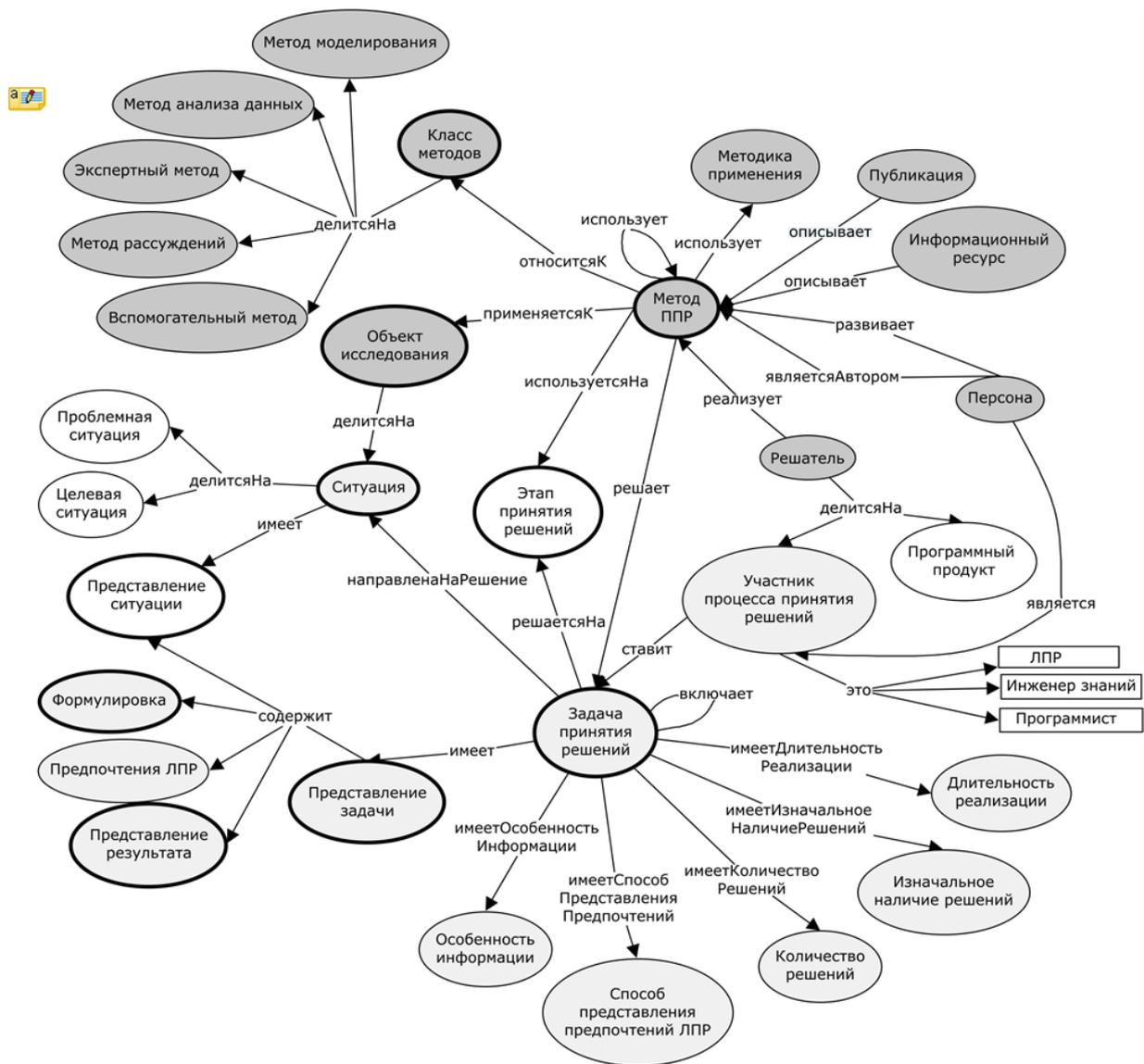


Рис. 11. Фрагмент онтологии области знаний ППР.

Очень важно обеспечить удобное и понятное для пользователей представление онтологии. Для этих целей задаются **паттерны представления**, определяющие соглашения по именованию, аннотированию и визуализации элементов онтологии для конечного пользователя.

Для реализации таких паттернов в онтологию были введены специальные свойства (Рис. 12) *order*, *link* и *shortlink*, служащие для задания порядка (*order*) показа атрибутов при визуализации и формирования кратких (*shortlink*) и полных (*link*) ссылок на объекты. (Ссылка на объект – это его имя, составленное из значений его свойств и свойств связанных с ним объектов). Эти свойства декларированы в онтологии как *Annotation Properties*, что позволяет хранить информацию о правилах визуализации контента непосредственно в онтологии.

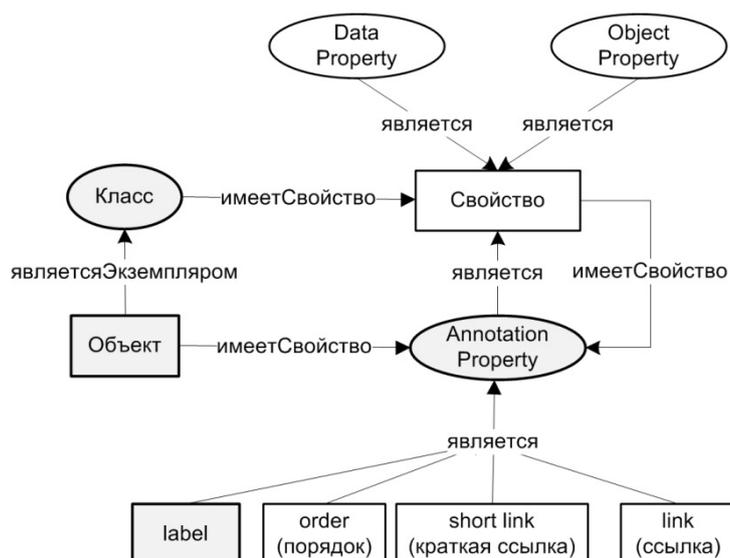


Рис.12. Паттерн представления

На Рис. 13 показано использование паттерна представления при задании онтологии и при отображении информации об объекте.

last name	Zagorulko
name	Galina
middle name	Borisovna
initials	Galina B.
<b>participate_Person_in_Activity</b>	
Activity	
<a href="#">Development of intelligent scientific Internet resources</a>	
role	project executor
<a href="#">Integrated support of development of intelligent decision support systems</a>	
role	project executor
start date	2016
<b>work_Person_in_Organization</b>	
Organization	
<a href="#">A.P. Ershov Institute of Informatics Systems, Siberian Branch of the Russian Academy of Sciences</a>	
hold the position of	researcher
Start date	1990
<b>is_Author_of</b>	
Publication	
<a href="#">Zagorulko Galina B. Development of ontology for intelligent scientific Internet resources on decision-making support in weakly formalized domains 2016</a>	
<a href="#">Zagorulko Galina B., Zagorulko Yury A. Ontology-Based Technology for Development of Intelligent Scientific Internet Resources 2015</a>	
<a href="#">Zagorulko Galina B., Zagorulko Yury A. Architecture of Extensible Tools for Development of Intelligent Decision Support Systems 2011</a>	

Рис.13. Использование паттерна представления

**Ввод онтологии и данных.** Согласно предлагаемой методике (технологии разработки ИНИР), онтология ППР разрабатывается на основе базовых онтологий и паттернов средствами редактора Protege (Рис. 13) и затем загружается в хранилище, с которым работает ИАИР ППР.

Данные в контент ресурса можно также вводить через редактор Protege. Однако более удобные средства предоставляет редактор данных, входящий в оболочку ИАИР. На Рис. 14 представлен интерфейс редактора данных.



снабжен пользовательским интерфейсом, он позволяет запустить метод, задав необходимые для его работы данные и посмотреть результат. Познакомившись с работой метода и оценив, насколько он подходит для решения поставленных задач, пользователь может встроить подходящий метод в разрабатываемую им СППР, используя описание программного интерфейса соответствующего сервиса.

### 3.2. Разработка интеллектуальных методов поддержки принятия решений

Разработана новая версия сервиса, реализующего метод недоопределенных вычислений с использованием клиент-серверной архитектуры. При обращении к серверу можно настроить протоколы передачи данных и указать порты, задать пути к папкам на локальном компьютере, в которые будут записываться служебная информация о работе сервера (Рис. 15).

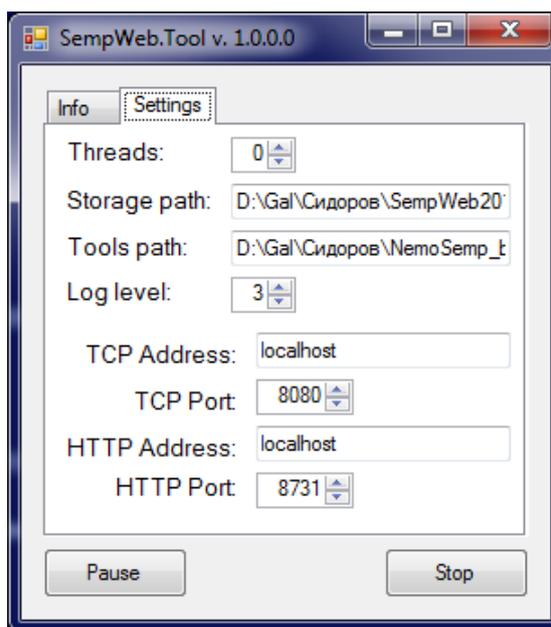


Рис. 15. Серверная часть сервиса.

При запуске клиентской части сервиса загружаются файлы с описанием решаемой задачи на языке NeMo+. Клиент позволяет в окне редактирования внести изменения в описание задачи, запустить задачу на счет, посмотреть результаты или сообщения об ошибках. При желании результаты можно сохранить в файл (Рис. 16).

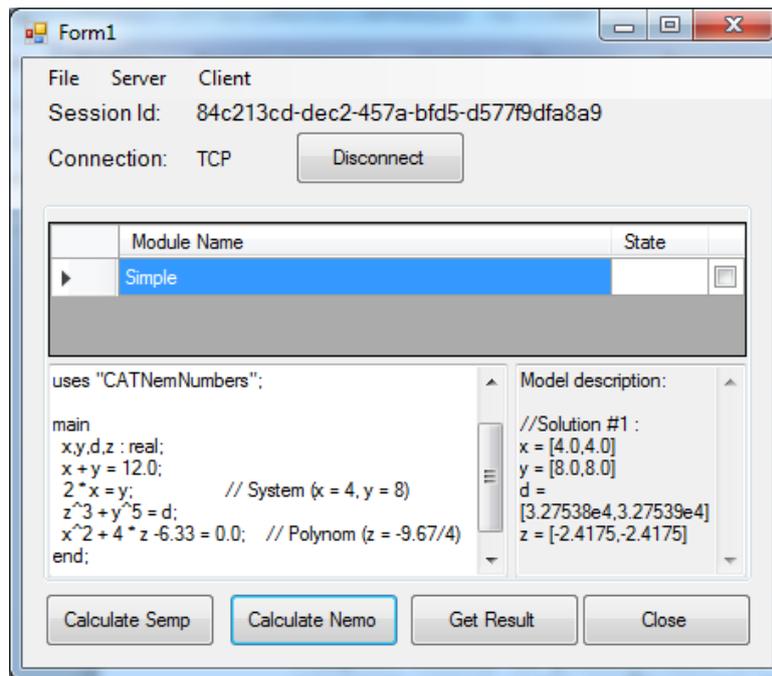


Рис. 16. Клиентская часть сервиса

## Заключение

Все поставленные в проекте задачи в целом были выполнены. Получены результаты по всем заявленным направлениям.

В рамках создания методов, программных средств и технологии создания и сопровождения интеллектуальных информационных систем (ИИС) поддержки научной, производственной и образовательной деятельности исследованы вопросы использования паттернов онтологического проектирования для разработки онтологий научных предметных областей (ПО). Разработаны наборы структурных и содержательных паттернов, с помощью которых могут единообразно задаваться основные сущности онтологий различных научных ПО. Показано, что использование таких паттернов не только обеспечивает согласованное представление всех сущностей онтологии, но и позволяет экономить ресурсы и избегать ошибок при разработке онтологий.

Исследованы существующие и разработаны новые методы интеграции интеллектуальных информационных систем (ИИС) поддержки научной деятельности, построенных с использованием технологий Semantic Web, с облаком Linked Open Data (LOD).

Разработаны новые методы пополнения контента ИИС поддержки научной деятельности данными из облака LOD и публичных веб-сайтов. В частности, построена онтология структуры веб-страниц; разработаны и реализованы методы сегментации веб-страниц, позволяющие выделять наиболее важные сегменты веб-страниц; разработан и реализован метод сегментации веб-страниц с помощью лексических маркеров.

В рамках развития технологии анализа текста разработаны новые семантические модели и методы, используемые для извлечения информации из текстов. Предложена модель представления предметной лексики, позволяющая описывать сложную семантику терминов, и осуществлена ее интеграция в модель представления факта.

Разработаны новые методы жанрового анализа текстов. Формально описана жанровая модель текста и расширено представление жанровых маркеров, учитывающее особенности интернет-представления текста.

Разработан ряд лингвистических ресурсов, предназначенных для решения задач классификации и извлечения информации из текстов технической документации и медицинских интернет-источников.

В рамках создания технологии построения интеллектуальных систем поддержки принятия решений разработана методология информационно-аналитической поддержки (ИАП) разработчиков СППР, которая представляет собой систему принципов, подходов и способов организации теоретической и практической деятельности по обеспечению ИАП и использованию её средств.

Разработан сервис, реализующий метод недоопределенных вычислений и метод рассуждений на основе экспертных правил с использованием клиент-серверной архитектуры. Сервис снабжен пользовательским интерфейсом, позволяющим описать задачу или выбрать готовое описание на соответствующем языке, запустить соответствующий метод и посмотреть результаты его работы. Кроме того сервис предоставляет программный интерфейс, позволяющий встраивать реализованные в нем методы в другие программы.

Важнейшим результатом выполнения данного этапа проекта, обладающим существенной научной новизной, можно отнести следующий:

- Разработка подхода к синхронизации аннотированных текстов с их переводами на другие языки, а также с сопутствующими мультимедийными материалами. В основе этого подхода лежит разработанная ранее онтология аннотирования и разметки текстовых ресурсов, расширенная отношениями синхронизации между фрагментами ресурсов или множествами размеченных фрагментов. В рамках подхода разработана технология создания семантически-аннотированных интернет-ресурсов (корпусов текстов, мультимедийных коллекций), обеспечивающая удаленную разметку ресурсов и последующий интеллектуальный поиск на основе предложенной онтологии.

#### **Проект РФФИ № 15-07-04144а «Методы и технологии создания предметно-ориентированных систем извлечения информации из документов на основе мультиагентного подхода»**

*Руководитель проекта – к.ф.-м.н., с.н.с. Е.А. Сидорова*

Целью проекта является на решение фундаментальной проблемы - автоматическая обработка и анализ разнородной информации, представленной на естественном языке. Конкретной фундаментальной задачей, решаемой в рамках данного проекта, является разработка эффективных методов извлечения информации из текстов ограниченной тематики и жанра на основе мультиагентного подхода.

В 2017 году были выполнены следующие работы.

Разработаны модели и методы представления знаний, ориентированные на задачи автоматической обработки текста и извлечения информации.

Предложены модель информационного покрытия текста, позволяющая описывать процесс извлечения информации, модель описания предметной лексики, обеспечивающая представление тезаурусной и семантической информации о терминах, в том числе семантику элементов параметрических конструкций, и структурно-жанровая модель текста, позволяющая выявлять фрагменты текста для поиска необходимой информации.

Разработана оригинальная мультиагентная модель анализа текста, включающая множество лексических, информационных, управляющих и понятийных агентов, которые обеспечивают извлечение информации, снятие неоднозначности и разрешение кореференции на основе онтологии предметной области.

Разработаны алгоритмы извлечения информации из текста на основе мультиагентного подхода и создана уникальная технологическая среда для разработки систем извлечения информации, которая включает словарный компонент, подсистему жанрового анализа текста и специализированную мультиагентную платформу, обеспечивающую реализацию предложенных мультиагентных алгоритмов.

**Проект РФФИ № 16-07-00569а «Методы и средства комплексной поддержки процесса разработки интеллектуальных СППР в слабоформализованных предметных областях на основе сервис-ориентированного подхода и технологий Semantic Web»**

*Руководитель проекта – к.т.н., заведующий лабораторией Ю.А. Загоруйко*

Целью данного проекта является разработка методологии и средств, направленных на обеспечение комплексной поддержки процесса разработки интеллектуальных систем поддержки принятия решений для слабоформализованных предметных областей.

В 2017 году были выполнены следующие работы.

Реализованы и оптимизированы программные компоненты, необходимые для полноценного функционирования информационно-аналитического интернет-ресурса по поддержке принятия решений (ИАИР ППР), являющегося важным компонентом комплексной поддержки процесса разработки интеллектуальных СППР, обеспечивающим содержательный эффективный доступ к систематизированным знаниям и информационным ресурсам области знаний «Поддержка принятия решений» (ППР).

Для поддержки поисковых и навигационных функций пользовательского интерфейса ИАИР ППР была разработана подсистема доступа к данным в терминах онтологии предметной области. Разработаны и оптимизированы средства визуализации знаний и данных, представленных в контенте ИАИР ППР. Разработана рабочая версия редактора данных, ориентированная на экспертов, позволяющая специалистам в области знаний ППР самим пополнять онтологии, не обращаясь за помощью к инженерам знаний или разработчикам ресурса.

Разработаны и реализованы методы автоматизированного наполнения контента ИАИР ППР. Предложенные методы осуществляют поиск и извлечение информации о научной деятельности в области знаний ППР из сети Интернет, а также ее интеграцию в информационное пространство ИАИР ППР. При поиске информационных ресурсов онтология используется как для генерации поисковых запросов к глобальным поисковым системам, так и для оценки релевантности найденной информации тематике ИАИР ППР.

Для интеграции объектов, извлекаемых из информационных ресурсов, в онтологию ИАИР ППР разработан и реализован оригинальный алгоритм идентификации, который выполняет их сравнение с информационными объектами, содержащимися в онтологии ИАИР ППР, и принимает решение относительно каждого такого объекта, исходя из того является ли он для онтологии новым или уточняет уже существующий в ней объект.

Разработано информационное наполнение (контент) информационно-аналитического интернет-ресурса по поддержке принятия решений. При этом онтология ИАИР ППР была расширена новыми понятиями, позволяющими более детально описать его область знаний. Для единообразного и согласованного представления используемых в онтологии научных понятий, а также для сокращения времени, требуемого на пополнение онтологии конкретными сущностями, были разработаны специальные паттерны онтологического проектирования (паттерны содержания), задающие способы представления типовых фрагментов онтологий. С помощью таких паттернов были описаны и внесены в контент ИАИР ППР известные методы ППР, объекты и предметы исследования ППР, полученные в области ППР научные результаты, созданные программные продукты и др.

Разработана библиотека методов, реализованных в виде сервисов. Включение сервиса в библиотеку означает создание в контенте ИАИР ППР информационного объекта, описывающего данный сервис, его связи с решаемыми задачами, реализуемыми методами, авторами-разработчиками, типом используемой информации, методикой его применения.

Исследователями из ИСЭМ СО РАН и ИСИ СО РАН создана научная среда, обеспечивающая поддержку процесса разработки ИСППР в области интеллектуальных энергетических систем (ИЭС). Данная среда состоит из двух взаимодействующих интернет-ресурсов и имеет, соответственно, две точки входа. Первый ресурс, построенный на основе

ИАИР ППР, обеспечивает содержательный доступ к систематизированной информации об области ИЭС, методах и средствах поддержки принятия коллективных (согласованных) решений. Его основное назначение – удовлетворение информационных потребностей пользователей. Второй ресурс предназначен для поддержки коллективной деятельности экспертов и исследователей. Данный ресурс предоставляет доступ к средствам разработки ИСППР ИЭС.

### **Список публикаций лаборатории**

#### **Российские журналы**

1. Боровикова О.И., Загорулько Г.Б., Загорулько Ю.А., Шестаков В.К. Использование паттернов для разработки онтологии информационно-аналитического интернет-ресурса «поддержка принятия решений» // Информационные и математические технологии в науке и управлении. —2017.— № 3 (7). — С. 144–153.
2. Сидорова Е.А., Гаранина Н.О., Кононенко И.С. Подход к разрешению референциальной неоднозначности текста при пополнении онтологии // Информационные и математические технологии в науке и управлении. – 2017. – № 3(7). – С. 154–167.
3. Ахмадеева И.Р., Серый А.С., Шестаков В.К. Некоторые особенности реализации платформы для построения информационно-аналитических интернет-ресурсов // Информационные и математические технологии в науке и управлении – 2017. № 3 (7). – С. 168–175.
4. Боровикова О.И., Кононенко И.С., Сидорова Е.А. Подход к извлечению информации из протоколов клинических испытаний на основе медицинской онтологии // Системная информатика. – 2017. – № 9. – С. 93–110.

#### **Зарубежные журналы**

1. Natalia Garanina, Elena Sidorova, Irina Kononenko, Sergei Gorlatch. Using multiple semantic measures for coreference resolution in ontology population // International Journal of Computing. –2017. – Vol. 16, Iss. 3. – P.166–176. (CSP- Scopus Elsevier)  
<http://www.computingonline.net/computing/article/view/900>
2. Yuliya Rubtsova. Reducing the Degradation of Sentiment Analysis for Text Collections Spread over a Period of Time // Knowledge Engineering and Semantic Web / Przemysław Różewski, Christoph Lange (Eds.) / 8th International Conference, KESW 2017 Szczecin, Poland, November 8–10, 2017, Proceedings. Communications in Computer and Information Science book series. – Cham: Springer, 2017. – Vol. 786. – P. 3–18.

#### **Материалы международных конференций**

1. Yury A. Zagorulko, Olesya I. Borovikova, Galina B. Zagorulko. Methodology for the development of ontologies for thematic intelligent scientific Internet resources // Proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russky Island, Russia, 25-29 September, 2017. – IEEE Xplore digital library. – P. 194–198. – URL: <http://ieeexplore.ieee.org/document/8168097> (дата обращения: 10.12.2017) ISBN: 978-1-5386-1206-4. DOI: 10.1109/RPC.2017.8168097
2. Ludmila Braginskaya, Valery Kovalevsky, Andrey Grigoryuk, Galina Zagorulko. Ontological approach to information support of investigations in active seismology // Proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russky Island, Russia, 25-29 September, 2017. – IEEE Xplore digital library. – P. 27–29. – URL: <http://ieeexplore.ieee.org/document/8168060> (дата обращения: 10.12.2017) ISBN: 978-1-5386-1206-4. DOI: 10.1109/RPC.2017.8168060
3. Zagorulko, Yu.A., Borovikova, O.I., Zagorulko, G.B. Application of ontology design patterns in the development of the ontologies of scientific subject domains. 2017. Selected Papers of the XIX International Conference on Data Analytics and Management in Data Intensive Domains

- (DAMDID/RCDL 2017). – CEUR Workshop Proceedings. – Vol. 2022. – P. 258-265. – URL: <http://ceur-ws.org/Vol-2022> (дата обращения: 10.12.2017)
4. Sidorova, E.A., Kononenko, I.S., Zagorulko, Yu.A. An approach to filtering prohibited content on the web. 2017. Selected Papers of the XIX International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2017). – CEUR Workshop Proceedings. – Vol. 2022. – P. 64-71. – URL: <http://ceur-ws.org/Vol-2022> (дата обращения: 10.12.2017)
  5. Загорулько Ю.А., Загорулько Г.Б., Шестаков В.К. Подход к разработке информационно-аналитического интернет-ресурса по поддержке принятия решений // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Междунар. науч. конф., Банное, Россия, 1-5 марта 2017 г. (ИТиС - 2017) : науч. электрон. изд. (1 файл 28,0 Мб) / отв. ред. Ю. С. Попков, А. В. Мельников. – Челябинск : Изд-во Челяб. гос. ун-та, 2017. – С. 113– 116.
  6. Загорулько Г.Б., Боровикова О.И. Проблемы разработки онтологии для информационно-аналитического интернет-ресурса по поддержке принятия решений // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Междунар. науч. конф., Банное, Россия, 1-5 марта 2017 г. (ИТиС - 2017) : науч. электрон. изд. (1 файл 28,0 Мб) / отв. ред. Ю. С. Попков, А. В. Мельников. – Челябинск: Изд-во Челяб. гос. ун-та, 2017. – С. 108– 112.
  7. Сидорова Е. А., Боровикова О. И. Подход к жанровой классификации текстовых ресурсов // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Междунар. науч. конф., Банное, Россия, 1-5 марта 2017 г. (ИТиС - 2017) : науч. электрон. изд. (1 файл 28,0 Мб) / отв. ред. Ю. С. Попков, А. В. Мельников. – Челябинск: Изд-во Челяб. гос. ун-та, 2017. – С. 264–269.

#### **Материалы российских конференций**

1. Серый А.С., Сидорова Е.А., Гаранина Н.О. Многодокументный анализ референции в процессе извлечения информации на основе онтологии предметной области // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.167–173. – ISBN: 978-5-905569-10-4.
2. Вишнев К.Е., Загорулько Г.Б., Молородов Ю.И. Средства управления распределёнными табличными данными в информационно-аналитических интернет-ресурсах, основанных на онтологиях // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.135–141. – ISBN: 978-5-905569-10-4.
3. Ахмадеева И.Р. Методы поиска научных ресурсов в сети интернет на основе онтологии // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.174–179. – ISBN: 978-5-905569-10-4.
4. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. Особенности разработки информационно-аналитического интернет-ресурса по поддержке принятия решений // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.157–166. –ISBN: 978-5-905569-10-4.
5. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. Использование паттернов онтологического проектирования для разработки онтологий предметных областей // Материалы Всероссийской конф. с международным участием «Знания – Онтологии –

- Теории» (ЗОНТ-2017). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017. – Т.1. – С. 139–148.
6. Брагинская Л.П., Григорюк А.П., Ковалевский В.В., Загорулько Г.Б. Информационная поддержка вибросейсмических исследований // Материалы Всероссийской конференции с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2017). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017. – Т1. – С. 72–79.
  7. Гаранина Н.О., Серый А.С., Сидорова Е.А. Мультиагентный подход к разрешению кореференции при извлечении информации из текстов на основе онтологии. // Материалы Всероссийской конференции с международным участием «Знания-Онтологии-Теории» (ЗОНТ-2017). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017. – Т1. – С. 88–98.

#### **Участие в конференциях**

1. 8th International Conference: Knowledge Engineering and Semantic Web, KESW 2017. Szczecin, Poland, November 8–10, 2017.
2. XXII Байкальская Всероссийская конференция с международным участием «Информационные и математические технологии в науке и управлении». Иркутск – Байкал, 29 июня-8 июля 2017 г.
3. Вторая Российско-Тихоокеанская конференция по компьютерным технологиям и приложениям (RPC 2017), Владивосток, остров Русский, Россия. 25-29 сентября 2017 г.
4. XIX международная конференция “Аналитика и управление данными в областях с интенсивным использованием данных” DAMDID/RCDL'2017, МГУ, Москва, Россия, 10-13 октября 2017 г.
5. VI Всероссийская конференция «ЗНАНИЯ – ОНТОЛОГИИ – ТЕОРИИ» с международным участием (ЗОНТ-2017), 2–6 октября 2017 г., Новосибирск.
6. XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017), Новосибирск, ИВТ СО РАН, 4-7 декабря 2017 г.
7. Международная конференция «Вычислительная и прикладная математика 2017» (ВПМ 2017), НГУ, Новосибирск, 25-30 июня 2017 г.
8. 6-я научная конференция с международным участием «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ и СИСТЕМЫ», санаторий “Юбилейный”, Республика Башкортостан, Россия, 1-5 марта 2017 г.
9. Межрегиональный форум «Инфосибирь», 17-18 мая 2017, Новосибирск.

#### **Участие в программных комитетах конференций**

1. Загорулько Ю.А. – член программного комитета VII Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2017). Республика Беларусь, Минск, 16–18 февраля 2017.
2. Загорулько Ю.А. – член программного комитета The 16th International Conference on Intelligent Software Methodologies, Tools and Techniques (SoMet 2017), September 26-28, 2017. Kitakyushu, JAPAN.
3. Загорулько Ю.А. – член программного комитета Международной Крымской конференции «СВЧ-техника и телекоммуникационные технологии» (КрыМиКо-2017).

4. Загорулько Ю.А. – член программного комитета XIX Международной конференции “Аналитика и управление данными в областях с интенсивным использованием данных” DAMDID / RCDL’2017. Москва, Россия, 10-13 октября 2017 г.
5. Загорулько Ю.А. – член программного комитета VI Всероссийской конференции «ЗНАНИЯ – ОНТОЛОГИИ – ТЕОРИИ» с международным участием (ЗОНТ-2017), 2–6 октября 2017 г., Новосибирск.
6. Загорулько Ю.А. – председатель подсекции секции «Информационные технологии» 55-й Международной студенческой конференция "Студент и научно-технический прогресс", Новосибирск, апрель 2017 г.

#### **Членство в национальных научных организациях**

1. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б., Сидорова Е.А. – члены Российской ассоциации искусственного интеллекта.

#### **Членство в редколлегиях научных изданий**

- Журнал «Проблемы информатики» — Ю.А. Загорулько – член редколлегии.
- Журнал «Онтология проектирования» — Ю.А. Загорулько – член редколлегии.
- Научный электронный журнал «Системная информатика» — Ю.А. Загорулько – член редколлегии.

#### **Международное сотрудничество**

#### **Членство в международных научных организациях**

1. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б., Сидорова Е.А. – члены Европейской ассоциации искусственного интеллекта.

## Общая характеристика исследований лаборатории системного программирования

*Зав. лабораторией к.т.н. Шелехов В.В.*

### Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе

Разработан метод автоматического синтеза предикатной программы по спецификации и набору теорий предметной области. Для синтеза используется та же система правил доказательства корректности, что и при дедуктивной верификации. Применение правил дает систему формул корректности, которую необходимо разрешить относительно неизвестных функций и предикатов. Решением является подстановка успешной унификации формул корректности с леммами теорий. Используется также перебор термов методом контрпримеров с проверкой через SMT-решатели. Синтез применением SMT-решателей невозможен без соответствующих лемм. Однако определить требуемый набор лемм априори невозможно. Помочь может специализированный интерактивный решатель с удобной визуализацией генерируемых формул корректности и набором преобразований. Применение решателя позволит также вдвое сократить затраты на дедуктивную верификацию предикатных программ.

На примере программы управления лифтом описывается технология автоматного программирования, представленная правилами интеграции автоматного, предикатного, объектно-ориентированного и модели-ориентированного программирования, а также баланса информационных и управляющих связей автоматной программы. Модель автоматной программы конструируется в виде набора *слоев* (аспектов), реализующих конкретную функциональность. Модель записывается на формальном языке требований и представляет собой шаблон, в котором фиксирована лишь управляющая часть программы, а информационная часть схематично представлена набором произвольных переменных и операторов, конкретизируемых при построении автоматной программы по модели. Для класса систем управления определены и описаны *базисные слою*: простейшая циклическая модель, слой отказоустойчивости, слой мониторинга.

Разработаны методы и алгоритмы оптимизации предикатных программ с трансляцией на язык C++, названной оптимизацией среднего уровня, существенно отличающейся от классической оптимизации императивных программ. Результатом трансформаций является эффективная императивная программа. Применяется набор оптимизирующих трансформаций: склеивание переменных, замена хвостовой рекурсии циклом, подстановка определения предиката на место вызова, упрощения и оформления. В целях оптимизации проводится потоковый анализ предикатной программы, в частности, определяются области жизни переменных.

Описываются оптимизирующие трансформации для операций над списками и деревьями в системе предикатного программирования. Кодирование операций представлено набором правил, определяющих замену исходной операции на ее образ в императивном языке. Построен набор правил эффективной трансляции операций алгебраических типов с использованием потокового анализа программы. Для АВЛ-деревьев разработан метод эффективной трансформации операций с деревьями.

Синтаксически ориентированный редактор предикатных программ реализован в рамках известной системы Eclipse и предоставляет среду, в которую интегрированы подсистемы дедуктивной верификации и программного синтеза.

Разработана модель типов языка предикатного программирования  $P$  в целях реализации семантического анализа программ. Система типов включает подтипы, параметрические (зависимые) типы, алгебраические типы и предикатные типы. Проверка принадлежности подтипу осуществляется при помощи SMT-решателей. В трансляторе с языка  $P$  реализовано представительное подмножество системы типов, в частности, параметрические типы.

В целях реализации дуального (текстового и графического) редактора автоматных программ разработана модель текстового образа программы графического языка Дракон.

## Описание проведенных научных исследований

### 1. Предикатное программирование.

*Предикатная программа*  $H(x: y)$  с аргументами  $x$  и результатами  $y$  является *вычислимым предикатом*. Язык предикатного программирования  $P$  построен последовательным расширением ядра  $P_0$ . В языке  $P_0$  пять операторов: *вызов предиката*  $V(x: y)$ , *оператор суперпозиции*  $V(x: z); C(z: y)$ , *параллельный оператор*  $V(x: y) || C(x: z)$ , *условный оператор* **if** ( $e$ )  $V(x: y)$  **else**  $C(x: y)$  и *оператор каррирования* (аналог лямбда-функции)  $D(y: z) \{V(x, y: z)\}$ . Каждый оператор имеет эквивалентное определение в виде предиката – логической формулы.

Для *предусловия*  $P(x)$  и *постусловия*  $Q(x, y)$  программа  $H(x: y)$  должна удовлетворять условиям *частичной корректности*  $P(x) \& H(x: y) \Rightarrow Q(x, y)$  и *тотальности*  $P(x) \Rightarrow \exists y. H(x: y)$ . Для дедуктивной верификации и программной синтеза разработана универсальная система правил, существенно упрощающая доказательство корректности разных операторов.

**Дедуктивная верификация** программы  $H(x: y)$  относительно ее спецификации  $[P(x), Q(x, y)]$  реализуется применением правил корректности операторов. Получается набор формул корректности программы, доказываемых применением SMT-решателей CVC3, CVC4 и Z3, либо интерактивно в системе PVS. Доказательство формул корректности гарантирует правильность программы относительно спецификации.

Реализована экспериментальная система предикатного программирования. Подсистема дедуктивной верификации применялась в рамках курса «Формальные методы в программной инженерии» в 2013-2017гг. для генерации формул корректности, которые далее доказывались студентами в системе PVS.

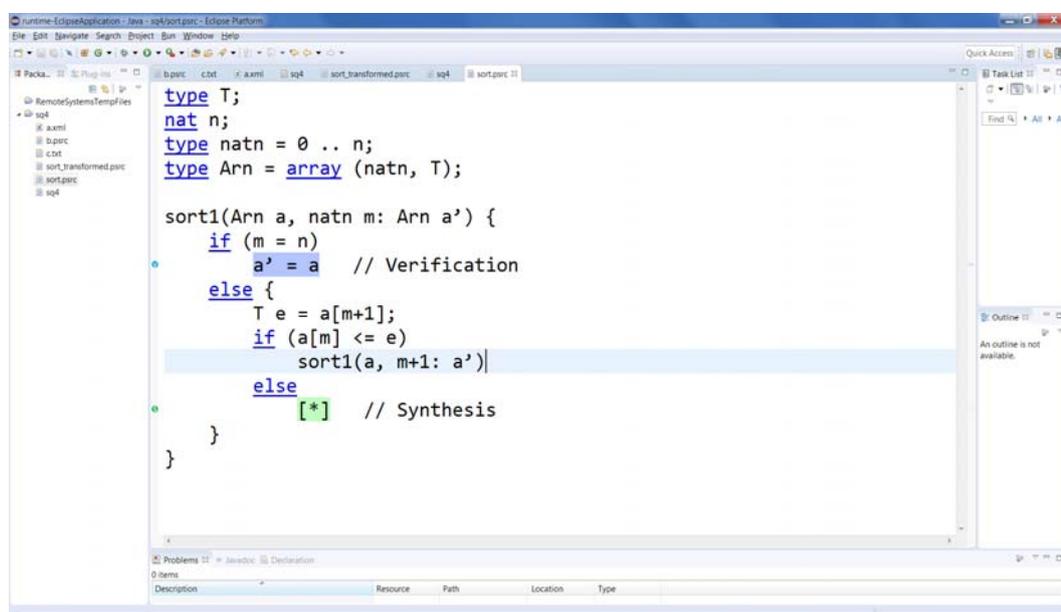
**Программный синтез** предикатной программы  $H(x: y)$  по ее спецификации  $[P(x), Q(x, y)]$  заключается в автоматическом построении тотальной программы  $H(x: y)$ , удовлетворяющей формуле корректности  $P(x) \& H(x: y) \Rightarrow Q(x, y)$ . Спецификация дополняется теориями предметной области с наборами лемм, используемых в процессе синтеза. Рассматривается также задача **синтеза фрагментов**, операторов и выражений, отмеченных в тексте программы  $H(x: y)$  комбинацией символов:  $[*]$ . Вместо комбинации  $[*]$  синтезатор вставляет вызов неизвестной функции или неизвестного предиката с набором параметров – переменных, определяемых в соответствии с текущим контекстом. Применение правил доказательства корректности к преобразованной программе дает систему формул корректности, которую необходимо разрешить относительно неизвестных функций и предикатов. Далее синтезатор просматривает леммы, проводя унификацию формулы каждой леммы с формулами корректности. Успешная подстановка унификации определяет значения неизвестных функций и предикатов.

Значения неизвестных функций определяются также перебором всевозможных термов в порядке возрастания сложности. Очередной терм подставляется на место неизвестной функции. Истинность полученной формулы корректности проверяется применением SMT-решателей. Здесь также используется метод синтеза через контрпримеры. Для спецификации

$P(x)$  &  $Q(x, y)$  с помощью SMT-решателей генерируется набор примеров, на которых спецификация истинна. При этом истинность формулы корректности для очередного терма сначала проверяется на сгенерированном наборе примеров.

При наличии соответствующих лемм синтез фрагментов программ средней сложности с применением SMT-решателей является вполне реалистичной задачей. Однако написать заранее нужный набор лемм практически невозможно. Определить требуемые леммы возможно лишь посредством анализа формул корректности. Необходим собственный специализированный решатель, работающий интерактивно в контексте синтезируемой программы и набора теорий. Решатель проводит преобразования (перебор и унификацию термов, подстановки) и удобную визуализацию генерируемых формул корректности. Применение решателя позволит вдвое сократить затраты на дедуктивную верификацию предикатных программ.

**Редактор предикатных программ** построен в виде набора плагинов в рамках известной системы Eclipse и предоставляет среду, в которой традиционный стиль построения предикатной программы интегрирован с подсистемами дедуктивной верификации и программного синтеза. Редактор синтаксически ориентирован. Выделяются служебные слова. Редактор также интегрирован с инструментарием системы предикатного программирования.



```
type T;
nat n;
type natn = 0 .. n;
type Arn = array (natn, T);

sort1(Arn a, natn m: Arn a') {
  if (m = n)
    a' = a // Verification
  else {
    T e = a[m+1];
    if (a[m] <= e)
      sort1(a, m+1: a')
    else
      [*] // Synthesis
  }
}
```

**Оптимизирующие трансформации.** Разработаны методы и алгоритмы оптимизирующей трансформации предикатных программ с трансляцией на язык C++. Это оптимизация *среднего уровня*, существенно отличающаяся от классической оптимизации императивных программ. Результатом трансформаций является эффективная императивная программа. В языке P нет циклов и указателей. Вместо указателей в язык P включены *алгебраические типы*: списки и деревья. Набор базисных оптимизирующих трансформаций: склеивание переменных, замена хвостовой рекурсии циклом, открытая подстановка, кодирование алгебраических типов через массивы и указатели. В целях оптимизации проводится потоковый анализ предикатной программы. Определяются аргументы и результаты операторов программы и области жизни переменных.

Построен набор правил эффективной трансляции операций алгебраических типов с использованием потокового анализа программы. Для AVL-деревьев разработан метод эффективной трансформации операций с деревьями. Язык P расширен средствами модификации поддеревя, доступного по некоторому произвольному пути в дереве. Для реализации списков и строк используются «плавающие» буфера. Введены средства сканирования, аналогичные итераторам в императивных языках.

**Анализ типов языка предикатного программирования.** Система типов языка P включает подтипы, параметрические (зависимые) типы, алгебраические типы и предикатные типы. Проверка принадлежности подтипу осуществляется при помощи SMT-решателей. Правила семантики определяют три вида отношений между соседними конструкциями: совместимость, тождество и согласованность (существование общей мажоранты). Отношение вложенности для типов определяется в виде системы правил, описывающих предпорядок «<:» на типовых термах. Рекурсивные алгебраические типы определены через аппарат наименьшей неподвижной точки. Для конструкций с неявной типизацией сформулированы правила восстановления типов переменных из контекста. Обобщенные типы (параметры других типов или предикатных программ) анализируются на базе *концептов* – наборов ограничений для типов. Разработаны алгоритмы построения дерева вызовов с использованием достаточно точной аппроксимации множеств значений переменных предикатных типов.

В трансляторе с языка P реализовано представительное подмножество системы типов, в частности, параметрические типы. Это позволило снять чувствительные ограничения на задачи для студентов в рамках курса «Формальные методы в программной инженерии».

## 2. Автоматное программирование

ориентировано на класс программ – *реактивных систем*, реализующих взаимодействие с внешним окружением программы и реагирующих на определенный набор событий (сообщений) в окружении программы. Язык автоматного программирования – компактный язык, который строится как расширение *базисного языка* предикатного или императивного программирования. В качестве базисного язык P предоставляет развитую систему типов, удобную для описания объектов окружения и состояния автоматной программы, с гарантией ее полного семантического контроля, поддержанного SMT-решателями.

Спецификация реактивной системы определяется набором требований. Определение требований должно проводиться методами *инженерии требований* в соответствии с последним стандартом ISO/IEC/IEEE 29148. *Требование* – утверждение, определяющее потребность и связанные с ней измеримые условия и ограничения. *Функциональные требования* определяют поведение программы.

Автоматная программа строится переписыванием функциональных требований на формальный *язык требований*, являющийся языком автоматного программирования. Технология автоматного программирования представлена в виде свода правил. Это правила интеграции автоматного, предикатного, объектно-ориентированного и модели-ориентированного программирования, а также баланса информационных и управляющих связей автоматной программы.

**Моделе-ориентированное программирование** (model-driven engineering) принципиально упрощает построение сложной реактивной системы. Модель автоматной программы конструируется в виде набора *слоев* (аспектов), реализующих конкретную функциональность. Слой может быть реализован независимой утилитой с предоставлением определенного интерфейса. Для включения в модель такого слоя достаточно лишь поставить вызовы его утилит в нужных местах. Однако в общем случае включение в модель нового слоя требует существенной модификации исходной модели.

Модель записывается на формальном языке требований и представляет собой шаблон, в котором фиксирована лишь управляющая часть программы, а информационная часть схематично представлена набором произвольных переменных и операторов, конкретизируемых при построении автоматной программы по модели.

**Верификация автоматных программ.** Верификация автоматной программы ограничена полным семантическим контролем (соответствия типов), а также проверкой истинности *инвариантов управляющих состояний*. Проведение формальной верификации не гарантирует правильности программы. Во-первых, инварианты обычно неполны или

отсутствуют. Во-вторых, ошибочными могут быть исходные содержательно сформулированные требования. Формализация требований способствует устранению ошибок, но не всегда.

Эффективным инструментом верификации является **моделирование** автоматных программ. Для этого необходимо использовать программные модели, имитирующие датчики произвольной природы. Моделирование предполагает визуализацию функционирования автоматной программы в числовом и графическом виде, в том числе и в реальном масштабе времени через *панель управления* в виде графического интерфейса, моделирующего работу разнообразных портов ввода и вывода. Предполагается реализация различных видов анализа исполняемой программы с возможностью пошагового исполнения. Планируется также обеспечить возможность воспроизведения повторного исполнения автоматной программы посредством запоминания истории вводимых наборов данных.

**Редактор автоматных программ** на базе дуального (текстового и графического) представления с использованием графического языка Дракон. Разработана модель текстового образа программы языка Дракон. Рассматриваются различные архитектуры дуального редактора с выходом на любой из пяти существующих редакторов языка Дракон.

### 3. Классификация программ для разработки адекватной технологии программирования.

Методы программной инженерии, доказавшие свою эффективность, не всегда успешно применимы для всех программ. Причина здесь в различиях архитектур программ. Это ставит задачу *классификации программ*, т.е. построения системы классов программ и разработки адекватной технологии программирования для каждого класса программ. Теория программ каждого класса должна определять методы спецификации, верификации (в широком смысле), моделирования и эффективной реализации программ.

Генеральная классификация определяет два класса программ: не взаимодействующие программы (или *программы-функции*) и реактивные системы (или *программы-процессы*). Данные два класса составляют более 90% всех программ. Имеются другие, более сложные классы, например, языковые процессоры и операционные системы.

**Класс программ-процессов.** *Программа-процесс* является *реактивной системой*, реагирующей на определенный набор событий (сообщений) во внешнем окружении программы. Программа-процесс является либо автоматной программой, либо она определяется в виде композиции нескольких автоматных программ, исполняемых параллельно и взаимодействующих между собой через прием / посылку сообщений и разделяемые переменные.

*Автоматная программа* состоит из одного или нескольких сегментов. *Сегмент* имеет один *вход*, помеченный меткой – *управляющим состоянием*. Сегмент имеет один или несколько выходов. Автоматная программа определяет конечный автомат в виде *гиперграфа* с набором управляющих состояний в качестве вершин и набором сегментов в качестве ориентированных гипердуг. *Состояние* автоматной программы определяется значениями набора переменных, модифицируемых в программе.

Частью задачи классификации программ является определение структуры класса реактивных систем. Определена иерархия подклассов в классе программ-процессов (реактивных систем), включающая гибридные системы, автоматы (временные, недетерминированные, вероятностные), системы управления, информационные системы и т.д.

Подклассом реактивных систем являются *гибридные системы*, соединяющие дискретное и непрерывное поведение. Часть переменных состояния гибридной системы соответствует непрерывным параметрам (типа **real**), изменение которых реализуется независимо от программы гибридной системы (вне ее) по определенным законам, обычно формулируемым в виде дифференциальных уравнений. Важнейшими подклассами гибридных систем являются контроллеры систем управления и временные автоматы.

*Система управления* реализует взаимодействие с *объектом управления* для поддержания его функционирования в соответствии с поставленной целью. Системы управления

используются в аэрокосмической отрасли, энергетике, медицине, робототехнике, массовом транспорте и др. отраслях. На каждом шаге вычислительного цикла *контроллер системы управления* получает входную информацию из окружения и обрабатывает ее. Результаты вычисления используются для передачи управляющего сигнала для воздействия на объект управления. Большинство систем управления являются встроенными системами.

**Модель реактивной системы в виде набора слоев.** Возможна иная классификация реактивных систем в виде *дерева наборов слоев*. Вершины идентифицируются набором слоев. Включение в исходную модель нового слоя определяет дугу в дереве от набора слоев новой модели к набору слоев исходной модели. Задача классификации заключается в описании структуры этого дерева для класса реактивных систем.

Для класса систем управления определены и описаны *базисные слои*: простейшая циклическая модель, слой отказоустойчивости, слой мониторинга. Зафиксированы другие слои, в частности, защиты от несанкционированного доступа и интеграции автоматического и ручного управления.

## Список публикаций лаборатории

### Российские журналы

1. Тумуров Э.Г., Шелехов В.И. Технология автоматного программирования на примере программы управления лифтом // «Программная инженерия», Том 8, № 3, 2017. – С.99-111. <http://persons.iis.nsk.su/files/persons/pages/lift1.pdf>
2. Шелехов В.И. Предикатная программа вставки в AVL-дерево // Системная информатика, № 9. — Новосибирск, 2017. — С. 23-42. [http://persons.iis.nsk.su/files/persons/pages/avl\\_insert.pdf](http://persons.iis.nsk.su/files/persons/pages/avl_insert.pdf)
3. Булгаков К.В., Каблуков И.В., Тумуров Э.Г., Шелехов В.И. Оптимизирующие трансформации списков и деревьев в системе предикатного программирования // Системная информатика, № 9. — Новосибирск, 2017. — С. 63-92. <http://www.system-informatics.ru/files/article/105.pdf>
4. Зубарев А.Ю. Анализ типов в трансляторе с языка предикатного программирования // Системная информатика, № 9. — Новосибирск, 2017. — С. 1-22. <http://persons.iis.nsk.su/files/persons/pages/zubarev.pdf>
5. Каблуков И.В., Шелехов В.И. Реализация оптимизирующих трансформаций в системе предикатного программирования // Системная информатика. — Новосибирск, 2018. — 18с. <http://persons.iis.nsk.su/files/persons/pages/opttransform4.pdf> На рецензии

### Материалы международных конференций

1. Шелехов В.И. Синтез операторов предикатной программы // Труды конф. «Языки программирования и компиляторы '2017», Ростов-на-Дону — 2017 — С.258-262. <http://persons.iis.nsk.su/files/persons/pages/sintr.pdf>
2. Чушкин М.С. Завершение эскизов предикатных программ методом синтеза через контрпримеры // Конф. «Компьютерная безопасность и криптография» SIBECRYPT'17 / Прикладная дискретная математика. Приложение. — 2017, №.10, С.151-153. <http://www.mathnet.ru/links/3bc251f8a0f2d95801e4e7efc9501433/pdma319.pdf>
3. Шелехов В.И. Технология модельно-ориентированного автоматного программирования // Научный сервис в сети Интернет 2017. Новороссийск. — 2017. — С.468-476. <http://persons.iis.nsk.su/files/persons/pages/mdea.pdf>
4. Зубарев А.Ю. Анализ типов в языке предикатного программирования // Знания-Онтологии-Теории (ЗОНТ-2017), Том 1. — Институт Математики СО РАН, Новосибирск, 2017. — С. 149-157. [http://persons.iis.nsk.su/files/persons/pages/zont\\_11.pdf](http://persons.iis.nsk.su/files/persons/pages/zont_11.pdf)
5. Шелехов В.И. Доказательное построение, верификация и синтез предикатных программ //

Знания-Онтологии-Теории (ЗОНТ-2017), Том 2. — Институт Математики СО РАН, Новосибирск, 2018. — С. 156-165. <http://persons.iis.nsk.su/files/persons/pages/lbase.pdf>

#### **Учебные материалы**

1. Шелехов В.И., Чушкин М.С., Пономарев Д.К. Формальные методы в программной инженерии. Видеолекции. — ИСИ СО РАН, Новосибирск, 2017. <http://wasp.iis.nsk.su/>

#### **Участие в конференциях**

1. Языки программирования и компиляторы '2017. Всероссийская научная конференция памяти А.Л. Фуксмана. 3–5 апреля 2017г., Ростов-на-Дону. — Докладчик: Шелехов В.И.
2. 16-ая всероссийская конференция Сибирская научная школа-семинар с международным участием «Компьютерная безопасность и криптография» SIBECRYPT'17. — Сибирский государственный университет науки и технологий им. М. Ф. Решетнева, г. Красноярск, 4–9 сентября 2017 г. — Докладчик: Чушкин М.С.
3. XIX Всероссийская конференция Научный сервис в сети Интернет, 18 – 23 сентября 2017 г. Новороссийск. — Докладчик: Шелехов В.И.
4. VI Всероссийская конференция «ЗНАНИЯ – ОНТОЛОГИИ – ТЕОРИИ» с международным участием 2 – 6 октября 2017 г. Новосибирск. — Докладчики: Зубарев А.Ю. , Шелехов В.И.

#### **Работы по грантам**

##### **Проект РФФИ № 16-01-00498а**

Технология предикатного и автоматного программирования

*Руководитель – к.т.н., зав.лаб. В.И. Шелехов*

##### **Грант благотворительного фонда В.Потанина «Российский образовательный проект по курсу «Формальные Методы в Программной Инженерии»**

Подготовлена большая часть видеолекций по первой половине курса "Формальные методы в программной инженерии": <http://wasp.iis.nsk.su/>

*Руководитель – к.т.н., зав.лаб. В.И. Шелехов*

## **Общая характеристика исследований лаборатории смешанных вычислений**

*Зав. лабораторией к.ф.-м.н. Бульонков М.А.*

### **Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе**

**Приоритетное направление IV.39.** Архитектура, системные решения, программное обеспечение, стандартизация и информационная безопасность информационно-вычислительных комплексов и сетей новых поколений. Системное программирование

**Программа IV.39.1.** Теоретические и прикладные проблемы создания эффективных надежных программных систем и информационных технологий

**Проект:** Алгоритмы и программные средства для моделирования сложных систем

*Научные руководители:*

*к.ф.-м.н., доцент Ф.А. Мурзин, к.ф.-м.н., доцент М.А. Бульонков*

### **Краткое описание проведенных научных исследований**

#### **1. Система МИКС.**

В рамках системы МИКС (модульная информационно-картографическая система) получила развитие система автоматизации экономических исследований МИКС-ПРОСТОР, которая, по-существу, создавала окружение для решателя транспортной задачи. Решаемая на тот момент проблема состояла в том, что как описание входных данных, так и получаемый результат представляет собой весьма объёмные и сложно устроенные объекты. Система МИКС-ПРОСТОР, с одной стороны, предоставляет удобный пользовательский интерфейс для определения всех параметров транспортной сети, таких как топология, длина дорог, тарифы на провоз и обработку, объёмы производства и потребления конкретных продуктов в узлах и т.п, а с другой, - позволяют вывести результаты моделирования «прямо на карте», причём, по необходимости, выделить, отдельные аспекты, такие как загруженность дорог, перевозку отдельного продукта и т.п. Типичный снимок работы в систем МИКС-ПРОСТОР показан на рис. 1.

*Рис 1. Система МИКС-ПРОСТОР*

После решения этой первоочередной задачи, поскольку проведение одного эксперимента сократилось и появилась возможность проводить их в большом количестве, появилась необходимость систематизировать проведённые эксперименты: хранить и сортировать по различным параметрам, входные и выходные данные, сравнивать результаты моделирования и т.п. Теперь, обеспечив достаточную автоматизацию для рутинной и чисто технической работы, мы вышли на новый уровень проблематики: пониманию того, что же пытаются найти

экспериментатор с помощью предоставленного инструментария и как в этом ему помочь. Система позволяет эксперту выделить набор интересующих его входных параметров и для каждого из них задать диапазон возможных значений. После этого формируются тестовые наборы параметров равномерным или вероятностным методом, а далее для каждого тестового набора решается задача оптимизации, определяющая значения переменных. На основе собранной информации мы пытаемся дать ответы на содержательные вопросы эксперта. Так на рис.2 приводится диаграмма зависимости использования отрезка транспортного маршрута от сочетания варьируемых тарифов:

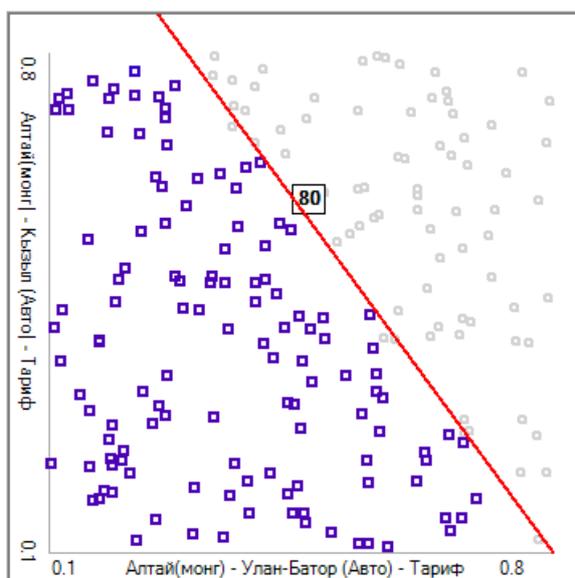


Рис 2. Двумерные проекции зависимости от параметров

## 2. Алгоритмические вопросы конъюнктивной декомпозиции булевых формул.

Разработаны алгоритмы декомпозиции таблиц с явным и (бинарные таблицы) неявным атрибутом, связывающим компоненты посредством равенства значений (в терминах SQL, обращение оператора JOIN T1, T2 ON T1.A1=T2.A2). Данные алгоритмы могут применяться для оптимизации представления реляционных баз данных (поиск функциональных зависимостей для целей нормализации и разбиения таблиц), для оптимизации и концепт-анализа данных в задачах D&KM, а также оптимизации представления булевых функций

Показано, что всякое представление логических формул, допускающее эффективную реализацию:

1. Означивание в 0 и 1.
2. Вычисление исключаящего ИЛИ.
3. Сравнение на равенство. С учетом пункта 2 достаточно сравнения с 0 (проверка тождественной ложности).

позволяет эффективно отыскать (если существует) разбиение множества переменных формулы так, чтобы она представима в виде конъюнкции формул с непересекающимися множествами переменных, соответствующих компонентам разбиения. Если также допускается эффективная реализация дизъюнкции и конъюнкции, то можно эффективно получить сами компоненты конъюнктивного разложения в том же самом представлении.

Реализован алгоритм конъюнктивной декомпозиции с использованием BDD для представления булевых функций, что является де-факто стандартом для разного рода систем оптимизации

булевых функций. Это позволяет выйти за рамки представления функций в виде Sum-of-Products (СДНФ и АНФ). Исследовались возможности использования ZDD для задачи конъюнктивной декомпозиции. Подготовлена инфраструктура (конверторы) для проведения масштабных экспериментов по декомпозиции булевых функций из признанных в мире бенчмарков по оптимизации логических схем.

### **3. Методы анализа и визуализации данных большого объема**

1) Продолжались исследования про кросс-языковой идентификации сущностей. Эксперименты на примере сотрудников ИСИ СО РАН показали, что ошибки идентификации авторов научных публикаций значительно снижают их рейтинг в таких базах данных как Scopus и Web of Science. Поэтому был разработан новый многокритериальный алгоритм установления авторов научных публикаций на основе сопоставления русскоязычной базы данных eLibrary и англоязычного сайта SpringerLink.com. Благодаря сопоставлению русскоязычного и англоязычного источников данных удалось повысить полноту обнаруживаемых данных, а за счет реализации нового алгоритма кластеризации авторов научных публикаций удалось повысить точность идентификации авторов. Также было показано, что сопоставление русскоязычных и англоязычных источников данных позволяет обнаруживать ошибки в источниках обоих типов.

2) В связи с работами по кросс-языковой идентификации сущностей были исследованы возможности использования визуального сравнения текстов и метаданных научных публикаций. Написан обзор, посвященный тому, каким образом методы анализа больших коллекций научных публикаций связаны с методами визуализации этих коллекций. Рассмотрены такие методы анализа коллекций научных публикаций как вероятностные тематические модели, включая латентное размещение Дирихле, а также неотрицательное матричное разложение (Non-negative Matrix Factorization, NMF) и предсказательные модели дистрибутивной семантики, такие как непрерывный Мешок Слов (Continuous Bag Of Words, CBOW) и Skip-gram. Показано, каким образом использование более эффективных методов анализа коллекций публикаций расширяет возможности визуализации, в частности обеспечивая интерактивные скорости работы с коллекциями большого объема, а также позволяя работать с несколькими взаимосвязанными представлениями коллекций на разных уровнях детализации.

### **Результаты работы по грантам**

**Проект №15/10** «Математические и методологические аспекты интеллектуальных информационных систем» Программы фундаментальных исследований президиума РАН №15 «информационные, управляющие и интеллектуальные технологии и системы»

*Руководитель – д.ф.-м.н. Марчук А.Г.*

**ГРАНТ РФФИ и фонда ДСТ (Индия) 17-51-45125** «Новые методы анализа и оптимизации реконфигурируемых компьютерных архитектур»

*Руководитель – к.ф.-м.н Пономарев Д.К.*

**Проект РГНФ № 16-02-00221** «Моделирование процесса освоения северных территорий и акваторий России: игровой подход на основе геоинформационных технологий»

*Руководитель: Малов В.Ю.*

*Исполнители: Бульонков М.А., Филаткина Н.Н., Тарасова О.В., Мелентьев Б.В.*

## Список публикаций лаборатории

### Российские издания

1. Апанович З.В. Кросс-языковая идентификация авторов публикаций/Открытые системы. СУБД. 2017. № 1. С. 43-45. ВАК
2. Апанович З.В. Методы визуализации коллекций научных публикаций //Материалы Всероссийской конференции с международным участием “Знания-Онтологии\_Теории” (ЗОНТ-2017) 2-6 октября 2017 г. Новосибирск С. 22-28.
3. Апанович З.В. Опыт преподавания принципов, методов и средств Semantic Web//Материалы Всероссийской конференции с международным участием “Знания-Онтологии\_Теории” (ЗОНТ-2017) 2-6 октября 2017 г. Новосибирск С. 29-37.
4. Апанович З.В. Современные тенденции визуализации коллекций научных публикаций // Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2017. — С. 3-8. — URL: <http://keldysh.ru/abrau/2017/36.pdf> doi:10.20948/abrau-2017-36
5. Апанович З.В. Преподавание методов Semantic Web разработчикам программного обеспечения // Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2017. — С. 9-20. — URL: <http://keldysh.ru/abrau/2017/37.pdf> doi:10.20948/abrau-2017-3

### Материалы международных конференций

1. Emelyanov P. and Ponomaryov D. Cartesian Decomposition in Data Analysis // *Proceedings of the Siberian Symposium on Data Science and Engineering (SSDSE 2017)*, 12-13 April 2017, Novosibirsk, Russia, pp. 55-60, 2017, DOI: 10.1109/SSDSE.2017.8071964.
2. В.Ю.Малов, М.А.Бульонков, Н.Н.Филаткина. Монгольский транзит: оценка конкурентноспособности транспортных маршрутов. // III Международный форум ассоциации экспертных центров Китая, Монголии и России "Экономический коридор Китай - Монголия - Россия: дорожная карта". - 19-20 сентября, Улан-Удэ, 2017.

### Материалы российских конференций

1. Апанович З.В. Кросс-языковая идентификация сущностей для сопоставления данных разноязычных ресурсов.//Труды пятнадцатой национальной конференции по искусственному интеллекту с международным участием, Смоленск, Универсум, 2016, С. 5-13.
2. Апанович З.В. Сопоставление данных разноязычных ресурсов и кросс-языковая идентификация авторов //Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19-24 сентября 2016 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2016. — С. 36-45.

### **Участие в конференциях**

1. Научный сервис в сети Интернет: XIX Всероссийская научная конференция (18-23 сентября 2017 г., г. Новороссийск)
2. “Знания-Онтологии\_Теории” (ЗОНТ-2017) Всероссийская конференция с международным участием (2-6 октября 2017 г. Новосибирск)
3. Siberian Symposium on Data Science and Engineering (SSDSE 2017), 12-13 April 2017, Novosibirsk, Russia.

### **Командировки**

**(в том числе инициативные, не оплачиваемые Институтом)**

1. Апанович З.В. (18.09.2016-23.09.2017) поездка в Новороссийск, участие в конференции «Научный сервис в сети Интернет -2017»
2. Емельянов П.Г. с 14 по 24 июня (Париж, командировка от НГУ). Посетил:
  - Ecole des Ponts (Марн-ля-Вале).
  - EDF Lab (R&D подразделение Electricité de France, Палезо).
  - Ecole Polytechnique (Палезо).
  - Télécom ParisTech.
3. Емельянов П.Г. с 09.10.2017 г. по 12.10.2017 (Москва, МГУ, командировка от НГУ) участвовал в совместном заседании ФУМО в сфере высшего образования по УГСН 01.00.00 Математика и механика и ФУМО в сфере высшего образования по УГСН 02.00.00 Компьютерные и информационные науки.

# Общая характеристика исследований лаборатории САПР и архитектуры СБИС

*Зав лабораторией д.ф.-м.н. Марчук А.Г.*

## Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе

### Краткое описание проведенных научных исследований

#### 1. Разработка многоуровневой иерархической сети распознавания образов на основе логико-вероятностных методов извлечения знаний

В настоящее время сверточные нейронные сети лидируют в области распознавания изображений, а также показывают хорошие результаты в распознавание аудиоданных и текстов. Однако существенным недостатком нейронных сетей является невозможность интерпретировать найденные ими закономерности в понятном для человека виде. В данной работе предлагается использовать идею многоуровневой иерархической обработки данных для построения многоуровневых сетей распознавания образов на основе логико-вероятностных методов извлечения знаний, результаты работы которых представимы в форме интерпретируемых логических правил.

Используя идеи построения сверточных нейронных сетей, в работе была предложена многоуровневая сеть, состоящая из следующих слоев: слой локальной перцепции, слой подвыборки и слой семантического предсказания. Слой локальной перцепции, по аналогии со слоем свертки у сверточных нейронных сетей, обрабатывает исходный образ по фрагментам, при котором «окно» фрагмента перемещается по всему образу и осуществляет классификацию, формируя карту признаков для обработки последующими слоями. Основное отличие от сверточных нейронных сетей состоит в том, что вместо операции свертки над фрагментом выполняется классификация без учителя. Слой подвыборки, по аналогии со слоем пулинга у сверточных сетей, выполняет уменьшение размерности карты признаков, выбирая один наиболее вероятный класс среди соседних фрагментов. Указанные слои могут чередоваться, уменьшая с каждым слоем размер карты признаков, но увеличивая при этом сложность и «глубину» признаков. Заключительным слоем сети является слой семантического предсказания, который осуществляет окончательное распознавание по последней карте признаков, используя семантический вероятностный вывод.

С предложенной сетью были проведены эксперименты по исследованию эффективности модели и сравнению с другими методами. В качестве теста использовалась задача распознавания рукописных цифр (150 примеров размером 14x14 пикселей). Тестирование проводилось методом скользящего контроля, при котором на каждой итерации из общей выборки исключалось по 50 тестовых примеров, а остальные 100 примеров использовались для обучения. Оценивалась точность распознавания на тестовой выборке. Результаты сравнения приведены в следующей таблице:

Метод	Точность
Предложенная модель	94.5 %
Классический СВВ	79.8 %

Neural Net	~70 %
Convolution Neural Net	97.1 %

В рассмотренном эксперименте предложенная модель была реализована с одним слоем локальной перцепции, одним слоем подвыборки и слоем предсказания.

Для сравнения использовались следующие методы:

- Классический метод семантического вероятностного вывода (СВВ).
- Классическая нейронная сеть с одним скрытым слоем и сигмоидальной функцией активации, обучающаяся методом RProp (Resilient Propagation).
- Глубокая сверточная нейронная сеть, реализованная при помощи библиотек Keras и Theano. Сеть имеет один сверточный слой, один MaxPooling слой и один полносвязный слой.

Несмотря на то, что в эксперименте предложенный метод немного уступил в точности сверточной нейронной сети, тот факт, что его результаты значительно превосходили классический семантический вероятностный вывод и классическую нейронную сеть, показывает, что используемые в архитектуре сверточных сетей идеи достаточно универсальны и могут быть с успехом использованы совместно с другими методами, в частности, с логико-вероятностными методами извлечения знаний, для улучшения качества распознавания.

## **2. Исследования по адаптации и оптимизации логико-вероятностных методов извлечения знаний для работы с данными больших размеров**

Существующие логико-вероятностные методы извлечения знаний из данных имеют ограничения при работе с данными больших размеров, как по количеству объектов, так и по количеству признаков, поскольку использует в своей работе элементы перебора. В данном исследовании ставится задача адаптации указанных методов для работы с данными, имеющими большие признаковые пространства.

В результате проведенных исследований предложен способ организации процесса извлечения знаний логико-вероятностными методами, позволяющий эффективнее анализировать данные, имеющие большой размер пространства входных признаков. Предложенный метод основан на идеях метода случайных подпространств (RSM) и стэкинга (stacking). Суть метода состоит в разбиении исходного признакового пространства на подвыборки меньшего размера и формирование иерархии слоев логических предикторов, где каждый предиктор обучается на своей подвыборке признаков, а выходы предикторов предыдущего слоя формируют пространство мета-признаков для предикторов последующего слоя. Каждый логический предиктор выполняет задачу прогнозирования целевого значения, основываясь на своем наборе признаков и используя семантический вероятностный вывод для обнаружения вероятностных закономерностей, предсказывающих целевой предикат. Поскольку выходными значениями предикторов являются прогнозируемые вероятности появления целевого предиката, то для формирования мета-признаков для последующих слоев, осуществляется кодирование выходов, к примеру, путем разбиения возможного диапазона значений на отрезки.

Преимуществом предложенного подхода является существенное уменьшение объема перебора за счет разбиения исходного большого пространства признаков на небольшие подпространства и дальнейшей композиции результатов работы методов извлечения знаний на этих подпространствах. При этом сохраняется возможность интерпретации используемых предикторами правил в логической форме, удобной для восприятия человеком. Кроме того, для предложенного метода легко организовать параллельные вычисления, поскольку все предикторы одного слоя могут осуществлять процесс извлечения знаний независимо друг от друга. Недостатком метода является потеря свойства гарантированной полноты извлечения всех вероятностных закономерностей по сравнению полным применением семантического

вероятностного вывода ко всему объему данных.

Были проведены экспериментальные исследования эффективности предложенного алгоритма на примере задачи распознавания рукописных символов. С использованием скользящего контроля оценивалась точность и время вычислений (без распараллеливания) предложенной модели с различным количеством предикторов в сравнении с классическим семантическим вероятностным выводом (СВВ). Результаты приведены в следующей таблице:

Метод	Кол-во признаков на один предиктор	Точность	Время работы
Классический СВВ	392	87.33 %	5320.55s
2 предиктора	196	84 %	1409.44s
3 предиктора	131	82.66 %	802.55s
4 предиктора	98	83.33 %	540.01s
6 предикторов	66	86.66 %	418.50s

Предложенный метод может быть использован для анализа данных с большим количеством признаков, когда применение стандартных логико-вероятностных методов извлечения знаний не возможно из-за большого объема вычислений, а также когда требуется провести анализ данных за ограниченное время.

### **3. Разработка методов адаптивного управления модульными механическими системами с большим числом степеней свободы**

В предыдущих исследованиях были разработаны и успешно протестированы частные модели системы управления для модульных гиперизбыточных систем, состоящих из симметричных по функциональности модулей. В соответствии с планом исследований следующим шагом является построение обобщенной модели системы управления и проведение экспериментального исследования по управлению некоторым ограниченным классом модульных роботов с произвольно заданной конструкцией.

В ходе проведенных за отчетный период исследований были выполнены работы по подготовке базы для экспериментального исследования. Целью исследования является проверка универсальности модели, т.е. что модель способна самостоятельно, без участия человека, настроиться и научиться управлять любой заданной конструкцией. Суть эксперимента заключается в том, что программная среда автоматически случайным образом генерирует конструкции роботов в рамках некоторого класса, а система управления должна научиться управлять движением данных конструкций.

Для планируемого эксперимента была подготовлена программная среда, моделирующая трехмерное виртуальное пространство с физическими законами. Был предложен класс модульных роботов, состоящих из одинаковых элементов, и разработана система для автоматической генерации случайных конструкций роботов для данного класса.

### **4. Математические свойства онтологических построений**

Исследована алгоритмическая сложность семантического комбинирования онтологий,

основанного на механизме согласования их моделей, и обобщающего операцию простого объединения онтологий, как логических теорий. Для широкого класса дескрипционных логик получена алгоритмическая классификация проблемы логического следования из семантически связанных онтологий. В рамках рассмотренного класса логик показано, что ациклическое семантическое комбинирование приводит к повышению сложности проблемы логического следования на одну экспоненту по сравнению со сложностью логики, в которой сформулированы онтологии. Результаты исследований дают основу для реализации новых практических методов комбинирования онтологий, в том числе, в целях спецификации терминологий естественнонаучных предметных областей.

Во многих предметных областях традиционно используются стандартизованные терминологические системы, представляющие определения ключевых понятий, используемых экспертами. В целях автоматизированной работы с такими системами, они представляются в компьютерных языках, основанных, к примеру, на дескрипционных логиках – фрагментах логики первого порядка. Стандартным примером такого языка является OWL (Web Ontology Language); сотни терминологических систем (онтологий), представленных в виде наборов утверждений (аксиом) в языке OWL, опубликованы в сети Интернет и используются в информационных системах, в том числе, для естественнонаучных предметных областей. Повсеместно возникает проблема построения более специфичных, узкоспециализированных онтологий на основе онтологий общепринятых терминов предметной области. Эта проблема подразумевает заимствование информации из существующих онтологий при разработке новых. Заимствование информации в целом можно характеризовать как статическое, либо динамическое. В первом случае из интересующей онтологии извлекается компонента, содержащая (либо аппроксимирующая в некотором смысле) информацию об интересующих понятиях. Эта компонента затем становится подмножеством разрабатываемой онтологии. В простейшем случае в качестве такой компоненты выбирается специальное подмножество аксиом онтологии. В более сложных подходах генерируется множество логических следствий онтологии, которые не обязательно являются ее аксиомами.

Динамический подход к заимствованию информации отличается от статического тем, что обмен информацией между связанными онтологиями происходит в виде обмена запросами каждый раз, когда необходимо произвести действия с какой-либо онтологией. Например, если требуется узнать синонимы некоторого термина онтологии  $O_1$ , которая использует онтологию  $O_2$ , то инициируется обмен запросами между этими онтологиями, в результате которого из  $O_2$  извлекается необходимая информация. Динамический подход является предпочтительным в случае версионности онтологий. Ранее не было известно методов динамического комбинирования онтологий, которые бы поддерживали полное заимствование семантики понятий из внешних онтологий. С целью решить эту проблему ранее нами был предложен новый механизм (семантического) импортирования онтологий, ориентированный на обмен сообщениями - логическими следствиями между онтологиями. Сформулирована семантика импортирования онтологий, обобщающая простое объединение онтологий. В результате работы получены результаты относительно предложенной семантики о сложности проблемы логического следования из онтологии с учетом связанных с ней онтологий. Для широкого класса дескрипционных логик от EL до SROIQ получены следующие сложностные результаты:

- для онтологий, сформулированных в логике EL, проблема логического следования с учетом импортов рекурсивно изоморфна проблеме останковки и, таким образом, неразрешима; при ациклическом комбинировании онтологий:

- для онтологий, сформулированных в логиках от EL до EL++, проблема логического следования с учетом импортов EXPTIME-полна;

- для онтологий в логиках от ALC до SHIQ, проблема логического следования с учетом импортов 2EXPTIME-полна;

- для онтологий в логиках от R до SRIQ указанная проблема следования 3EXPTIME-полна;

- для онтологий в логиках от ALCHOIF до SHOIQ указанная проблема соN2EXPTIME-полна;

- для онтологий в логиках от ROIF до SROIQ указанная проблема соN3EXPTIME-полна.

Полученные результаты показывают, что ациклическое семантическое импортирование приводит к повышению сложности проблемы логического следования на одну экспоненту по сравнению со сложностью логики, в которой сформулированы онтологии. Результаты проведенных исследований дают основу для реализации новых практических методов комбинирования онтологий.

Реализован метод вычисления явных определений понятий относительно онтологий в дескрипционной логике EL на основе исчисления направленного вывода. Разработан алгоритм вычисления явных определений понятий, основанный на трассировке доказательств в логике EL, который позволяет проверить, существует ли определение интересующего понятия с использованием выбранных пользователем терминов и отношений, и построить определение в явном виде. Алгоритм реализован в виде прототипа плагина (расширения) для широко используемого редактора онтологий Protégé и имеет практическую ценность для анализа и декомпозиции объемных онтологий, т.к., в частности, позволяет установить, возможно ли устранить те или иные термины и отношения из определений понятий в онтологии.

При анализе объемных онтологий типичным является вопрос, возможно ли определить то или иное понятие, используя только некоторый ограниченный набор терминов. Простейший пример, если определить понятие A через некоторые термины B и C как:  $A = B \text{ and } (C \text{ or not } C)$ , то из этого следует, что A можно определить, не используя термин C. Необходимость вычисления явных определений понятий (то есть тех определений, которые возможны, но не представлены в онтологии синтаксически) возникает в процессе работы с большими онтологиями, при извлечении из них компонент, касающихся выделенных терминов. Известно, что во многих дескрипционных логиках явные определения термина можно вычислить на основе анализа доказательств формул, логически следующих из онтологии. В частности, на основе анализа доказательств в табличных исчислениях.

Многие современные оптимизированные алгоритмы автоматического доказательства в дескрипционных логиках основаны на модификации известного метода резолюций и называются алгоритмами направленного логического вывода (consequence-based reasoning algorithms). В сравнении с другими известными методами доказательства, например, табличными методами, они показывают большее быстродействие и удобство реализации. Известно, что для проверки выводимости формулы в классической пропозициональной логике необходимо и достаточно вычислить (конечное) замыкание расширения заданного множества формул относительно правила резолюции и проверить, лежит ли формула - противоречие в построенном замыкании. Исчисления направленного вывода также ориентированы на проверку выводимости формул определенного класса и представляют собой правила, которые минимизируют применение правила резолюции в тех случаях, когда получаемые по нему формулы заведомо не требуются для проверки выводимости. Этим на практике достигается существенное сужение пространства поиска. К примеру, до недавнего времени задача классификации понятий (вычисление is-a иерархии понятий) онтологий в дескрипционной логике EL (сложность проверки выводимости в которой невысока ограничена квадратичным полиномом от размера онтологии), содержащих сотни тысяч терминологических определений, считалась на практике нерешаемой. Современные алгоритмы, использующие направленный логический вывод, решают эту задачу за секунды.

В результате работы для логики EL реализован метод вычисления явных определений понятия относительно онтологий на основе трассировки доказательств в исчислении направленного вывода. Разработан алгоритм вычисления явных определений понятий, который позволяет проверить, существует ли определение интересующего понятия с использованием выбранных пользователем терминов и отношений, и построить определение в явном виде. Алгоритм реализован в виде прототипа расширения (плагина) для широко используемого редактора онтологий Protégé и имеет практическую ценность для анализа и декомпозиции

объемных онтологий, т.к., в частности, позволяет установить, возможно ли устранить те или иные термины и отношения из определений понятий в онтологии.

## **5. Декомпозиция таблиц реляционных баз данных на основе факторизации булевских полиномов**

Предложен метод декомпозиции таблиц реляционных баз данных, основанный на ранее полученном алгоритме факторизации булевских полиномов, имеющем полиномиальную временную сложность. Результат имеет практическую значимость для оптимизации представления булевых функций, заданных в СДНФ, и в контексте самонастраивающихся баз данных для рефакторинга схем реляционных баз данных, поскольку позволяет обнаружить возможность более компактного представления реляционных таблиц на основе анализа их содержимого.

В ряде информационных задач возникает проблема компактного и экономного представления данных. Если данные представлены в виде конечного (атрибутированного) отношения, то указанная проблема означает нахождение его представления в виде комбинации нескольких отношений так, чтобы их суммарный размер был меньше, чем размер исходного отношения. Возможность для комбинации определяется заданным набором операций над отношениями. В таком виде проблема оптимизации возникает при анализе булевых функций, представленных в СДНФ (списком выполняющих наборов значений ее переменных), а также при рефакторинге (реструктуризации) схем реляционных баз данных. Одной из операций, позволяющей потенциально снизить размер представления данных, является декартова декомпозиция отношений, которая означает представление отношения в виде декартового произведения двух (или более) отношений с дизъюнктивными набором атрибутов. С точки зрения булевых функций такая декомпозиция соответствует представлению функции в виде конъюнкции функций от дизъюнктивных наборов переменных.

В результате работы показано, что проблема вычисления декомпозиции (атрибутированных) отношений эквивалентна проблеме факторизации булевских полиномов. На более ранних этапах работы было установлено, что последняя проблема полиномиально разрешима (от размера входного полинома, представленного своей естественной записью в виде слова). Был приведен соответствующий полиномиальный по времени алгоритм факторизации. В результате работы выполнена SQL-реализация данного алгоритма в применении к декартовой декомпозиции атрибутированных отношений. Реализация алгоритма за полиномиальное время от размера входного отношения позволяет установить, допускает ли оно декартову декомпозицию и найти минимальные компоненты, то есть те компоненты, которые декомпозицию уже не допускают.

## **6. Исследования по парадигмам программирования**

Изучение парадигмальных характеристик языков и систем программирования, включая взаимосвязи эксплуатационной и реализационной прагматик, а также их представление на уровне абстрактной машины в применении к проблеме декомпозиции сложных определений информационных систем, позволило сформулировать методику парадигмального анализа, нацеленного на выделение типовых автономно развиваемых компонентов, образующих основу для конструирования программ, правильных по построению. В результате исследований выполнен содержательный анализ специфики основных и фундаментальных парадигм программирования, сформулированы ведущие критерии их упорядочения и классификации языков программирования, поддерживающих разные парадигмы. Результаты анализа представлены в монографии «Парадигмы программирования: анализ и сравнение».

Результаты исследования представлены на международных и российских конференциях.

## **7. Исследования по определению языка начального ознакомления с миром параллелизма**

Работы велись по двум направлениям.

7.1. Разработка изобразительных средств для представления параллельных алгоритмов решения учебных задач, нацеленных на демонстрацию основных явлений и методов.

Предложен подход к представлению лексико-синтаксических средств, допускающих аспектно-ориентированное отделение процессов вычисления от последовательности представления фрагментов программ в терминах макрогенерации с контролем синтаксической правильности вида параметров. Результаты исследования представлены на международных конференциях.

7.2. Исследования по методам реализации учебных и экспериментальных языков программирования

Продолжена работа студенческого проекта по реализации ряда учебных роботов, приспособленных к решению задач на клетчатой доске с целью демонстрации основных проблем организации взаимодействия процессов.

## **8. Исследования по автоматизации генерирования тестовых наборов**

Системы автоматической проверки решений задач находят применение не только на олимпиадах различных уровней, но и в учебном процессе, особенно вузовского уровня. Для расширения области использования подобных систем необходимо иметь возможность автоматизировать не только собственно процесс проверки, но и процесс подготовки к этой проверке.

Наиболее важной частью задачного комплекта является тестовый набор, поскольку от его полноты существенно зависит правильность вердикта относительно проверяемого решения.

Целью исследований является создание системы, а) облегчающей формализацию спецификаций входных и выходных данных, заданных на естественном языке, б) предлагающей базовое разбиение пространства входных данных на классы эквивалентности, в) генерирующей тестовые наборы, исходя из выработанных спецификаций, и г) контролирующей совместимость спецификаций, указанных и полученных форматов данных, а также полноту и избыточность полученных тестовых наборов.

Продолжалась теоретическая разработка подходов к автоматизации генерирования задачных комплектов на основе спецификаций входных и выходных данных для систем автоматической проверки задач по программированию. Набирается статистика по использованию различных форматов спецификаций, представленных на естественном языке.

Результаты исследований представлены на всероссийских и международных конференциях.

## **9. Методы и технологии обработки структурированных данных**

Продолжались исследования вопросов, связанных со структурированием и обработкой данных, в том числе, больших данных. На этом этапе, работа была в основном связана с изучением применимости подхода и решений к типовым и нестандартным задачам. В частности, были продолжены исследования и разработки по созданию эффективных ORM решений. Была получена конкурентная стабильная версия ORM на базе PolarDB. Другая группа экспериментов касалась распределенных решений обработки баз данных. Был сформирован подход к распределенному хранению данных и асинхронной распределенной обработке таких данных. Подход заключается в расслоении используемых последовательностей в соответствии с синтаксическими свойствами ключа и распределении расслоенных данных по узлам хранения и обработки. Был проведен развитый эксперимент, подтвердивший потенциальную эффективность предложенного подхода. Была проведена переработка множества решений, входящих в библиотеку PolarDB. Продолжалось изучение способов ускорения работы механизма ORM по времени работы, а также оптимизация по использованию памяти.

Разработан инструмент в виде .Net библиотеки для конвертирования изображений в Deep Zoom Image и стандартные изображения требуемого размера, используя технологию Polar.BD. Инструмент может применяться для отображения изображений высокого разрешения в браузере в формате одного изображения выбранного размера, при этом оно динамически составляется из частей нужного размера, и с помощью технологии Deep Zoom от Microsoft, Seadragon и Open Seadragon.

Продолжаются исследования веб-публикации изображений высокого разрешения. Работа продолжается по направлениям: 1) отображение одного изображения с помощью технологии Open Seadragon; 2) отображение коллекции изображений с помощью технологии Open Seadragon; 3) создание формата файла, содержащего данные изображения.

Продолжаются разработки технологии создания полнотекстовых индексов с помощью Polar.DB и алгоритма n-грамм. Началось исследование системы сбора фактов из необработанных тестов на русском языке с целью их анализа в дальнейшем.

Продолжаются исследования в области Semantic Web: обработки больших RDF данных с технологией Polar.DB и на её основе индексов, языком запросов SPARQL версии 1.1. В частности, производились разработка таблицы имен, создание трансляторов SPARQL и Turtle с помощью генераторов по грамматикам, изучались вопросы оптимизаций для случаев ограниченности объёма оперативной памяти сверху и не ограниченности, когда её объём позволяет обрабатывать больше данных, чем поддерживаемые языком стандартные коллекции.

## **10. Миграция системы Библиотека «БИБЛИОТЕКА» на свободное ПО**

В 2017 г. были проведены работы по созданию веб-приложения Мемориальная библиотека имени А.П. Ершова.

Целью проекта было:

1) создание веб-приложения, предоставляющего доступ к фондам библиотеки пользователям Интернет; миграция каталогов библиотеки; миграция данных, относящихся к единицам хранения.

2) Сохранение мемориального ПО - системы «БИБЛИОТЕКА».

Библиотечная система «БИБЛИОТЕКА» была разработана Я.М. Курляндчиком в начале 1980-х гг. на платформе MS DOS с использованием FoxPro. Система «Библиотека» используется до настоящего времени в Мемориальной библиотеке им. А.П. Ершова для хранения, управления фондами, обработки и публикации новых поступлений в библиотеку.

В рамках данной работы создано веб-приложение с использованием открытой платформы Drupal. В приложении поддержана модель данных исходной системы «БИБЛИОТЕКА».

Приложение в настоящее время находится в стадии тестирования. Производятся итерации по миграции исходных данных системы «БИБЛИОТЕКА».

## **11. Создание, поддержание и развитие информационных ресурсов института**

Были созданы следующие информационные ресурсы: Сайт конференции PSI 2019, Ежегодная XII Ершовская лекция по программированию, раздел «Аспирантура» на сайте ИСИ. Были произведены модернизации и развитие следующих информационных ресурсов: основной сайт ИСИ, Электронный архив академика А.П.Ершова, Электронный журнал «Системная информатика», аллея памяти ИСИ, метрики сайтов ИСИ. Производились хостинг, администрирование и поддержка более, чем 20 сайтов и информационных ресурсов.

## **12. Работы в области учебной, в том числе школьной информатики**

Продолжается работа по анализу существующих и разработке новых систем для эффективного процесса обучения в области информатики образования. В качестве технических

средств поддержки конкурсной работы и преподавания информатики велась отладка и разработка новых форм проведения олимпиад, в том числе уделялось существенное внимание дистанционной и сетевой работе с учащимися. Велась работа по созданию и апробация систем учебной информатики. Практическое изучение со школьниками новых языков программирования и инструментов использовано в качестве базовой основы индивидуальной траектории развития.

Разрабатываются и апробируются методы и программы для профильной подготовки учащихся, программы для изучения основных и факультативных курсов информатики, программы для олимпиадной и предпрофессиональной подготовки школьников. Создаются новые формы работы по интенсификации конкурсной деятельности, предназначенные для эффективности образовательного процесса.

Проведена 42-я Летняя школа юных программистов (ЛШЮП). В работе школы участвовали школьники, студенты и преподаватели из Абакана, Бердска, Кольцово, Новокузнецка, Новосибирска, Милана (Италия), Миасса, Москвы, Оби, Омска, Санкт-Петербурга, Санта-Клары (Калифорния, США), Саратова.

Для научных институтов СО РАН существенно, что сотрудники, аспиранты и студенты принимают участие в работе ЛШЮП, выполняют несколько проектов по темам, разрабатываемым сотрудниками, развивается эксперимент по обучению в форме мастерских, идея которых сформулирована и внедрена сотрудниками ИСИ, выявлена заинтересованность молодежи в новых формах экспериментальной работы в области систем учебной информатики, а именно, предлагается эксперимент по организации Школы программирования (воскресной и вечерней, дистанционной) для наиболее подготовленных мастерских в течение учебного года, продолжает развиваться традиционный механизм, потеря которого может повредить интересам ИСИ в отношении плановой темы "Исследование основ информатики и методов преподавания информатики и программирования", привлечено внимание к методическим наработкам. После ЛШЮП в прошлом году было организовано 2 группы для заинтересованных школьников и студентов ВКИ. В этом году групп тоже 2, занятия будут проводиться в течение года на технике, купленной по программе повышения конкурентоспособности НГУ.

Продолжалось изучение подходов к образовательной деятельности, велась разработка курсов и программ изучения программирования в различных образовательных учреждениях учебного типа. («Школа раннего развития» в Детском Технопарке – 5, 6, 7, 8 классы). Велась деятельность по подготовке и проведению конкурсов, олимпиад и конференций учащихся, включающая в себя разработку форм и методов их проведения, а также тематические разработки задачного материала согласно особенностям школьного подхода к изучению алгоритмов (городская, областная, региональная научно-практическая конференция школьников «Эврика», секция «Информатика» – работа в жюри, подготовка задач, решений и тестов для школьного этапа Всероссийской олимпиады).

### **13. История отечественного программирования и отечественной информатики**

Завершено исследование по истории школ программирования в Академии наук СССР и ведущих советских вузах. Актуальность данной проблематики обоснована социально-культурными и научными следствиями деятельности программистов АН СССР позднесоветского времени. Исследование истории школ программирования в АН СССР в их взаимодействии актуально и в свете изучения научного наследия А.П. Ершова, поскольку как генезис, так и сплочение этого сообщества всегда было в фокусе его внимания, особенно в период его работы в межведомственной Комиссии по системному математическому обеспечению АН СССР в 1979–1988 гг. Архив этой комиссии с привлечением других источников позволил выявить основные направления развития школ отечественного программирования в Академии наук СССР и связанных с ней организаций. Наследие советских академических школ программирования имело значение не только в рамках институции, оно

формировалось исходя из различных народнохозяйственных нужд, подпитывало и поддерживало общий уровень отечественного программирования и науки в целом.

Разработана и обоснована периодизация становления научных школ программирования в СССР:

1. 1951– сер. 1950-х г. – стадия топоса<sup>1</sup>: освоение навыков программирования и формирование профессионального сообщества специалистов по МО из числа математиков, физиков и вычислителей. Формирование центров программирования в АН СССР (Киев, Москва, Ленинград).
2. Вторая половина 1950-х г.– сер. 1960-х гг. – стадия размежевания: рост и самоопределение программистского сообщества, появление специфических проектов, публикаций, сети цитирования, конференций, формирование парадигм и школ программирования, наиболее авторитетные из которых сформировались в Москве, Ленинграде, Новосибирске и Киеве.
3. Вторая пол. 1960-х – 1980-е г. – акматическая стадия (зрелость): углубление теоретического базиса программирования. Формирование кафедр и факультетов по прикладной математике. Информатизация общеобразовательной школы. Отстранение Академии наук СССР от экспертной политики в области вычислительной техники и программирования в условиях осуществления проекта Единой системы ЭВМ. Попытка консолидации академических и ведомственных программистов в рамках межведомственной Комиссии по системному математическому обеспечению (КОСМО, 1978).

Продолжалось наполнение электронных ресурсов СО РАН. Создан новый фонд академика Т.И. Заславской и доктора филологических наук М.И. Черемисиной (родные сестры). Документы переданы внучкой М.И. Черемисиной – Е.В. Шиплюк. Введены сканы трудов д.ф.-м.н. Ю.Б. Румера. Польский историк М. Сикора прислал сканы 7 документов из Национального архива США в Вашингтоне в архив А.П. Ершова, 2 фото прислала вдова К. Костера (Нидерланды). Продолжается описание документов из архива С.С. Лаврова. Проводится тестирование Электронного архива А.П. Ершова, который переносится на платформу Drupal.

Даты	2013	2014	2015	2016	2017
Листы					
Фотоархив СО РАН	429	279	132		136
Открытый архив СО РАН	9295	16365	10610	6679	4269
Архив АПЕ	2000	2000	2307		9

### Результаты работы по грантам

<sup>1</sup> Топос – повторяющийся мотив, формирующиеся склонности, предпочтения, знаковые слова какой-либо социальной группы (формирующееся, не общепринятое значение).

**Грант РФФ «Логико-вероятностный инструментарий искусственного интеллекта нового поколения»**

*Руководитель академик С.С. Гончаров ИМ СО РАН*

*Участник от ИСИ СО РАН Пономарев Д.К.*

Для широкого класса дескрипционных логик получена алгоритмическая классификация проблемы логического следования из семантически связанных онтологий. В рамках рассмотренного класса логик показано, что ациклическое семантическое комбинирование приводит к повышению сложности проблемы логического следования на одну экспоненту по сравнению со сложностью логики, в которой сформулированы онтологии. Результаты исследований дают основу для реализации новых практических методов комбинирования онтологий, в том числе, в целях спецификации терминологий естественнонаучных предметных областей. Реализован метод вычисления явных определений понятий относительно онтологий в дескрипционной логике EL на основе исчисления направленного вывода. Разработан алгоритм вычисления явных определений понятий, основанный на трассировке доказательств в логике EL, который позволяет проверить, существует ли определение интересующего понятия с использованием выбранных пользователем терминов и отношений, и построить определение в явном виде. Алгоритм реализован в виде прототипа плагина (расширения) для редактора онтологий Protégé и имеет практическую ценность для анализа и декомпозиции объемных онтологий, т.к., в частности, позволяет установить, возможно ли устранить те или иные термины и отношения из определений понятий в онтологии.

**Грант РФФИ и Департамента науки и технологии правительства Индии «Новые методы анализа и оптимизации реконфигурируемых компьютерных архитектур»**

*Руководитель Д.К. Пономарев*

Предложен метод декомпозиции таблиц реляционных баз данных, основанный на ранее полученном алгоритме факторизации булевских полиномов, имеющем полиномиальную временную сложность. Результат имеет практическую значимость для оптимизации представления булевых функций, заданных в СДНФ, и в контексте самонастраивающихся баз данных для рефакторинга схем реляционных баз данных, поскольку позволяет обнаружить возможность более компактного представления реляционных таблиц на основе анализа их содержимого.

**Грант Благотворительного фонда В. Потанина «Разработка магистерской программы «Логические методы в информатике».**

*Руководитель Д.К. Пономарев*

Разработана и реализована концепция новой магистерской программы для ММФ НГУ. Магистерская программа предлагает фундаментальные и практические знания методов формальной логики, которые являются ключевыми в информационных технологиях, применяемых в электронике, аэрокосмической отрасли, робототехнике, программной инженерии, обработке и хранении данных. Образовательный проект имеет своей целью подготовку специалистов, обладающих фундаментальными знаниями и навыками применения логических методов в работе со сложными информационными задачами. Сверхзадача образовательного проекта - переломить тенденцию подготовки специалистов - пользователей существующих зарубежных информационных технологий и сформировать предпосылки для воспроизводства специалистов-исследователей, способных к разработке новых технологий работы с информацией. Совместно с соисполнителями гранта разработано 5 новых курсов и модифицировано содержание 12 существующих курсов для целей разрабатываемой магистерской программы.

**Проект РФФИ N 17-11-00042 Д «Издание научного труда "Парадигмы программирования"».**

*Руководитель – к.ф.-м.н. Л.В. Городняя.*

Издана монография «Парадигмы программирования: анализ и сравнение» Из-во СО РАН, тиражом 300 экземпляров. Монография представляет результаты изучения и классификации ряда языков и систем программирования, позволяющие сформулировать методику парадигмального анализа сложных информационных систем и мультипарадигмального подхода к обучению программированию.

**Проект РФФИ N 15-07-06345** “Становление и развитие научных школ программирования в ведущих научных центрах СССР.

*Руководитель – к.ф.-м.н. С.П. Прохоров.*

*Участник от ИСИ СО РАН Городняя Л.В.*

Продолжено описание истории становления ряда отечественных научных школ программирования и краткая характеристика их основных достижений. Проанализировано большое число материалов по школам отечественного программирования и по аналогичным зарубежным работам с целью уяснения идейных приоритетов по отдельным вопросам технологии программирования. Результаты представлены в центральных изданиях и на международных конференциях.

**Проект РФФИ N 15-07-06345** «Научные школы программирования в Академии наук СССР и ведущих советских вузах».

*Руководитель – к.ф.-м.н. С.П. Прохоров, Институт истории естествознания и техники РАН, г. Москва.*

*Участник от ИСИ СО РАН Крайнева И.А.*

А. Исследована история программирования в странах советской Балтии, становление которого пришлось на конец 1950-х – начало 1960-х годов. Этот феномен формировался в рамках кибернетического направления, а в более широком контексте – на волне научно-технической модернизации советской экономики в период Холодной войны. В процессе реабилитации кибернетика была объявлена «наукой величайших возможностей» и «одним из основных средств созидания коммунистического общества». В АН СССР были созданы институты кибернетики в ряде советских республик, в том числе в 1960 г. в Таллинне, в 1976 г. при институте было организовано СКБ вычислительной техники, в 1977 организован Институт математики и кибернетики АН Литовской ССР (Вильнюс), ВЦ в Латвийском государственном университете (Рига) был создан в 1959 г. В Латвии был организован Институт электроники и вычислительной техники АН ЛатССР (1960).

Эти организации, появившиеся в системе Академии наук и в вузах Прибалтийских республик стали базовыми во многих аспектах программирования, сетевых технологий, профильного образования. Для появления данных учреждений в Прибалтике были и внутренние причины, обусловленные экономическими обстоятельствами. Так, создание Института кибернетики в Эстонии, например, было обусловлено бурным развитием химической и энергетической отраслей, автоматизация и средства контроля в которых играли значительную роль. Значительный импульс развитию стран Балтии придали реформы по децентрализации экономики 1957–1965 гг. Поскольку новые направления развития науки и техники требовали навыков, некоторое количество специалистов по вычислительной технике, имевших опыт программирования, для Прибалтийских республик были подготовлены в 1960-е годы в Ленинградском политехническом институте (ЛПИ) и Московском энергетическом институте (МЭИ), специалисты высшей квалификации – в Новосибирске, Москве, Киеве, Минске.

Б. История программирования в Беларуси в период хрущевской оттепели также связана с модернизационной стратегией на основе научно-технического прогресса, а именно кибернетики и ЭВМ. Во второй половине 1950-х годов в СССР развернулось создание новой отрасли машиностроения. В 1956 г. было принято постановление ЦК КПСС и Совета Министров СССР о мерах по расширению производства ЭВМ, предусматривалось строительство Минского и Казанского заводов ЭВМ, в конце 1957 г. – появилось постановление о строительстве Завода полупроводниковых приборов в Воронеже<sup>2</sup>. Постановление 1956 г. не предусматривало создание программного обеспечения (ПО) ЭВМ. В создавшихся условиях проблема ПО массовых ЭВМ вставала достаточно остро. На минском заводе к ее решению подошли творчески: на республиканском уровне в 1959 г. было принято решение о создании при СКБ завода лаборатории математического обеспечения по предложению Геннадия Константиновича Столярова, недавно окончившего Военмех в Ленинграде и заочно матмех ЛГУ. Столярову и его коллективу принадлежит роль создателей софтверной индустрии на Минском заводе вычислительных машин им. С. Орджоникидзе: коллектив лаборатории, поначалу состоявший из нескольких сотрудников, увеличился до 120 человек к концу 1960-х годов. Здесь трудились такие математики как Г.М. Генделев, В.И. Цагельский, М.С. Марголин, Е.В. Ковалевич, Г.Д. Смирнов, Л.М. Романовская и другие.

**Комплексная программа фундаментальных научных исследований СО РАН № П.2 П на 2017 год.** Современные тенденции в актуализации исторического опыта формирования идентичностей в Сибирском регионе. Номер проекта в ИСГЗ ФАНО 0326-2015-0019. Приоритетное направление 190. Изучение эволюции человека, обществ и цивилизаций, человек в истории и история повседневности, традиции и инновации в общественном развитии, анализ взаимоотношений власти и общества.

*Руководитель – д.и.н. О.Н. Шелегина, Институт истории СО РАН, г. Новосибирск.*

На основе концепции исторической идентичности ученых как акторов истории исследовалась конкретно-индивидуальная практика формирования *научного наследия*. Конкретизировались теоретические подходы к пониманию научного наследия. Под научным наследием ученых понимаем всю совокупность результатов научного труда ученых: опубликованные труды (литературное наследие), архив (черновые рукописи, неопубликованные работы, письма, отзывы, рецензии, заметки, фото), научные разработки (техника, лекарства, удобрения, новые породы животных и виды растений и т.д.), а также библиотеки и сопутствующие научные коллекции. Научное наследие индивида является частью мирового культурного наследия, даже если его влияние не выходит за рамки места развития.

Содержание категории *научного наследия* выходит за рамки научного вклада, который состоит, как правило, в новых положениях, решениях, теориях, признанных (или имеющих отсроченное признание) в научном сообществе, или реализованных в практическом плане. Оценка научного вклада может быть изменена со временем в ту и другую стороны, поскольку решения, предложенные при жизни ученого, могут опережать свое время, или потерять свою актуальность (историчность категории научное наследие). Научный мемориальный архив – часть наследия, может оказаться более значимым, чем научный вклад. Научный архив реализует свое значение, когда изучается, вводится в научный оборот, порождает новое знание в истории науки. В этом его непреходящее значение. Поэтому научный архив является, возможно, более важным компонентом научного наследия, чем научный вклад.

**Грант РФФИ 17-07-20538** Проект организации 4-й Международной научной конференции «Развитие вычислительной техники в России и странах бывшего СССР: история и перспективы» SoRuCom-2017.

---

<sup>2</sup> <http://www.computer-museum.ru/calendar/7.htm> (04.04.2017)

*Руководитель д.ф.-м.н. А.Г. Марчук (член ПК).*

4-я Международная конференция «Развитие вычислительной техники в России и странах бывшего СССР: история и перспективы» (SoRuCom-2017, <http://www.sorucum.org/>) прошла в Зеленограде 3–5 октября 2017 г. на базе Национального исследовательского университета МИЭТ. Данная конференция является уникальным в России мероприятием, где представляются и обсуждаются результаты исследований по проблематике советской вычислительной техники в исторической ретроспективе.

На конференцию SoRuCom-2017 был принят 61 доклад (81 докладчик, из них 1 академик, 24 доктора наук, 30 кандидатов наук, 26 научных сотрудников и инженеров). В составе докладчиков 40 представителей вузов, 20 из институтов РАН, 17 из проектно-конструкторских учреждений и компаний, 4 – от учреждений культуры. 61 доклад опубликован в трудах.

География: Россия (Зеленоград, Москва, Мытищи, Санкт-Петербург, Новосибирск, Самара, Казань, Екатеринбург, Нижний Новгород, Ростов-на-Дону), Беларусь, Польша, Прибалтика, Финляндия, США.

Формат конференции: пять секций (история вычислительной техники, история программирования, сетевые проекты, информатика и образование, сохранение историко-научного наследия), открывающая и завершающая пленарные сессии.

Необходимо отметить возросший уровень научного содержания докладов, представленных некоторыми инженерно-техническими специалистами, расширение тематики конференции в рамках ее устоявшейся организационной структуры. С момента проведения в июле 2006 года первой международной конференции «История вычислительной техники и ее программного обеспечения в России и странах бывшего СССР: история и перспективы» сформировалось сообщество SoRuCom, состоящее из активных исследователей, обеспокоенных проблемой сохранения наследия в области вычислительной техники и ее программного обеспечения, курирующих вопросы сбора и обработки уникальных научных материалов, публикации аналитических материалов, повышения качества обучения по истории информатики, развития информационных технологий в России.

Помимо организации и проведения конференций сообщество SoRuCom активно развивает сайт Виртуального компьютерного музея, публикует книги и статьи в ведущих рецензируемых журналах, в том числе на английском языке при поддержке Международной федерации по обработке информации IFIP и Международного Института инженеров электротехники и электроники IEEE, проводит локальные семинары и конференции (ИСИ СО РАН, ИНИОН РАН, Политехнический музей, Музей истории вычислительной техники Казанского завода ЭВМ (Казань), ИВМиМГ СО РАН, СПИИРАН, ИПИ РАН и др.), поддерживает контакты с зарубежными историками информатики Л. Грэхэмом, С. Геровичем, П. Джозефсоном, Б. Петерсом, К. Татарченко и другими. Происходит обмен идеями, литературой, Музей истории компьютеров из Маунтин Вью (Computer History Museum, CA, USA) и Российский фонд фундаментальных исследований (РФФИ) неизменно оказывают финансовую поддержку конференции. Членами сообщества создаются и поддерживаются электронные ресурсы по истории информатики, высоко оцененные международным сообществом: электронный архив академика А.П. Ершова (<http://ershov.iis.nsk.su>), сайт Новосибирского филиала Института точной механики и вычислительной техники АН СССР (<http://nfitmivt.ru>) и др.

Вместе с тем приходится констатировать, что конференция не охватывает всей проблематики народнохозяйственных применений ЭВМ (экономика, сфера обслуживания). Участие зарубежных специалистов на конференции не слишком велико (порядка 5-6 человек), хотя в состав программного комитета они охотно входят. Их непосредственное участие в данной конференции зависит от личного настроения, интереса к проблематике и стремления к живому общению.

В Зеленограде, отмечавшем 55-летие своего основания как центра микроэлектроники, конференция не вызвала того резонанса, на который могла бы рассчитывать. Принимающий вуз не проявил к ней должного интереса, не рекомендовал своим студентам включиться в ее

работу. Префектура Зеленограда также осталась к ней равнодушна. Экскурсия в музей "Ангстрема" – одного из отечественных флагманов микроэлектроники вызвала скорее сожаление, чем гордость, поскольку продемонстрировала упадок интереса к отечественному научно-техническому наследию со стороны администрации предприятия. Тем не менее, участники конференции искреннее благодарны принимающей стороне – Национальному исследовательскому университету МИЭТ в Зеленограде.

Решением участников конференции было высказано намерение провести следующий форум SoRuCom в 2020 г.

Идет подготовка к публикации 40 избранных докладов в электронных ресурсах IEEE CS, которая будет завершена в 2018 г. (ответственная к.и.н. И.А. Крайнева). В это число вошли 2 доклада сотрудников ИСИ СО РАН (соавторы Л.В. Городняя, И.А. Крайнева, А.Г. Марчук).

### Список публикаций лаборатории

#### Монографии

1. Городняя Л.В. Парадигмы программирования: анализ и сравнение. Из-во СО РАН, РФФИ 17-11-00042. 216 с.

#### Центральные издания

1. Городняя Л.В., Крайнева И.А., Марчук А.Г. Школа программирования Института кибернетики АН УССР (1962-1990)// История науки и техники, 2017. №1. С. 42–64. ВАК.

2. Городняя Л.В., Крайнева И.А., Марчук А.Г. Библиография и источники по истории школ программирования в Академии наук СССР и ведущих советских вузах (1950–1917)// История науки и техники, 2017. Т. 38. №3. С. 482–529. ВАК.

3. Городняя Л.В. Движущие силы информатики. Дистанционное и виртуальное обучение № 2 (116) 2017, с. 75-90. Издательство Современного гуманитарного университета (Москва), ISSN: 1561-2449 ВАК.

4. Сапрыкина Г.А., Тихонова Т.И. КОНФЕРЕНЦИИ ШКОЛЬНИКОВ КАК СПОСОБ ВЫЯВЛЕНИЯ СПЕЦИАЛЬНЫХ ОДАРЕННОСТЕЙ //Сибирский учитель. 2017. № 4 (113). С. 55-58. ВАК.

5. Куперштох Н.А., Крайнева И.А. История новосибирского Института радиофизики и электроники (1957–1967)// Гуманитарные науки в Сибири. 2017. Т.24. №2. С.109–113. ВАК.

6. Крайнева И.А. Генезис дисциплины в поле науки: вычислительное дело – программирование – информатика// Вестник ТГУ. 2017. N 421. С. 118–128. Web of Science

#### Зарубежные издания

1. Demin A.V. Vityaev E.E. Adaptive Control of Modular Robots // Biologically Inspired Cognitive Architectures (BICA) for Young Scientists. – Springer, 2018. – pp. 204-212. (Scopus).

2. D. Ponomaryov and M. Soutchanski Progression of decomposed local-effect action theories. ACM Transactions on Computational Logic (TOCL) 18(2), 2017. (Scopus, WoS).

<http://dl.acm.org/citation.cfm?id=3091119>. Журнал входит в 1-ый квартиль журналов по специализации «информатика» по версии Scopus.

3. Gregor Behnke, Florian Nielsen, Marvin Schiller, Denis Ponomaryov, Pascal Bercher, Birte Glimm, Wolfgang Minker, and Susanne Biundo. To Plan for the User Is to Plan with the User: Integrating User Interaction into the Planning Process. Companion Technology. A Paradigm Shift in Human-Technology Interaction. Susanne Biundo and Andreas Wendemuth, eds. Springer International

Publishing, 2017. ISBN: 978-3-319-43664-7 (Scopus). <https://www.springerprofessional.de/to-plan-for-the-user-is-to-plan-with-the-user-integratinguser-i/15277618> . Глава в зарубежной монографии.

4. Platonov Y., Artamonova E., "Creation of Information Portraits Using 'Polar' System" // Information Technology in Industry» 2018. Issues: 1-4 (в печати, индексируется в Web of Science).

### **Материалы международных конференций**

1. Y. Kazakov and D. Ponomaryov. On the complexity of semantic integration of OWL ontologies. Proc. DL '2017, The Description Logic Workshop, Montpellier, France, July 2017. <http://ceur-ws.org/Vol-1879/paper59.pdf> (Scopus).

2. Крайнева И.А., Городняя Л.В., Марчук А.Г. О работах по системному математическому обеспечению в странах Советской Балтии (1960–1990)/ В сборнике: Труды SORUCOM-2017. Материалы Четвертой Международной конференции. 2017. С. 135–144.

3. Крайнева И.А., Городняя Л.В. из истории программирования в Беларуси (1959–1990)/ В сборнике: Труды SORUCOM-2017. Материалы Четвертой Международной конференции. 2017. С. 145–155.

4. Городняя Л.В., Кирпотина И.А. О проблеме достоверности доступной в Интернете исторической фактографии // ТРУДЫ SORUCOM — 2017- с. 46-55.

5. Андреева Т.А. Автоматизация создания задачного комплекта // Электронные ресурсы в непрерывном образовании: труды VI Международного научно-методического симпозиума «ЭРНО-2017» (Адлер). – Ростов-на-Дону: Издательство Южного федерального университета, 2017. – 230 с. ISBN 978-5-9275-2080-0. – с. 145-148.

6. Y. Platonov ; E. Artamonova. Материалы 2017 Siberian Symposium on Data Science and Engineering (SSDSE), Managing big data using specialized system polar . - 2017. <http://ieeexplore.ieee.org/document/8071953/> (дата обращения 28.12.2017), г.Новосибирск, 2017. (Индексируется в SCOPUS).

7. Тихонова Т.И. Исполнители Звенигородского: эпоха от «Агатов» 1980-х до современности //ТРУДЫ SORUCOM-2017. Четвертая Международная конференция Развитие вычислительной техники в России и странах бывшего СССР: история и перспективы. 3–5 октября 2017 года. – Москва, Зеленоград. – с. 358-361.

8.Крайнева И.А., Пивоваров Н.Ю., Шилов В.В. Советский атомный проект и становление отечественной вычислительной техники/ В сборнике: Труды SORUCOM-2017. Материалы Четвертой Международной конференции. 2017. С. 127–134.

9. Andreyeva T.A. Automated preparation of problem complexes // Наука сегодня: теоретические и практические аспекты: материалы международной научно-практической конференции, г. Вологда, 27 декабря 2017 г.: в 2 частях. Часть 1. - Вологда: ООО <Маркер>, 2018. - 200 с. ISBN 978-5-906850-89-8 (Часть 1) – с. 25-26

### **Материалы российских конференций**

1. P. Emelyanov and D. Ponomaryov. Cartesian decomposition in data analysis. Proc. SSDSE, IEEE Symposium on Data Science and Engineering, Novosibirsk, Russia, April 2017. <http://ieeexplore.ieee.org/document/8071964/> .

2. Городняя Л.В. Резервы синтаксически ориентированного конструирования систем программирования. //Научный сервис в сети Интернет: труды XIX Всероссийской научной

конференции (18-23 сентября 2017 г., г. Новороссийск). -- М.: ИПМ им. М.В.Келдыша, 2017. -- 480 с., ISBN 978-5-98354-037-8 doi:10.20948/abrau-2017 URL: <http://keldysh.ru/abrau/2017/proc.pdf> .

3. Городняя Л.В. Учебный язык параллельного программирования СИНХРО – с. 92-97. Языки программирования и компиляторы — 2017 труды конференции / Южный федеральный университет; под ред. Д. В. Дуброва. — Ростов-на-Дону: Издательство Южного федерального университета, 2017 .— 282 с.ISBN 978-5-9275-2349-8 <http://plc.sfedu.ru/files/PLC-2017-proceedings.pdf> .

4. М. М. Лаврентьев, Т. С. Васючкова, Л. В. Городняя, В. С. Бартош, И. В. Белого, М. А. Держо, Н. А. Иванчева, О. А. Федотова. Использование Интернет для профориентации школьников в области IT на Факультете информационных технологий НГУ. Всероссийская объединённая конференция «Интернет и современное общество». <http://ojs.ifmo.ru/index.php/IMS/article/view/515> .

5. Андреева Т. А. Система подготовки тестовых наборов для автоматической проверки решений // Материалы IX Всероссийской (с международным участием) научно-практической конференции «Информационные технологии в образовании» «ИТО-Саратов-2017». – эл. публ. <http://saratov.ito.edu.ru/2017/section/235/99713/> .

6. Крайнева И.А. Междисциплинарный дискурс в научном наследии А.А. Ляпунова// Материалы III Всероссийской научно-практической конференции «Современные тенденции в развитии музеев и музееведения» (Новосибирск, 9–12 октября 2017 г.). Новосибирск: Институт истории Сибирского отделения Российской академии наук. 2017. С. 136–141.

### **Прочие публикации**

1. Платонов Ю.Г., Артамонова Е. В. Возможности технологии Object Relational Mapping для нереляционных баз данных // Современные тенденции развития науки и технологий. - 2017. - №3 (4) - стр. 108-113 URL: [http://issledo.ru/wp-content/uploads/2017/04/Sb\\_k-3-4.pdf](http://issledo.ru/wp-content/uploads/2017/04/Sb_k-3-4.pdf) (дата обращения: 09.04.2017). г. Белгород, 2017.

### **Статьи в сборниках**

1. Андреева Т. А. Возможность автоматизации процесса генерирования тестовых наборов // ж. Universum: технические науки. – №8(41). М., Изд. «МЦНО», 2017. – с.5-7.

2. Г. А. Сапрыкина, Т.И. Тихонова. КОНФЕРЕНЦИИ ШКОЛЬНИКОВ КАК СПОСОБ ВЫЯВЛЕНИЯ СПЕЦИАЛЬНЫХ ОДАРЕННОСТЕЙ.// «Грани творчества». – Новосибирск, 2017. – с. 106-120.

3. Крайнева И.А. Феномен исторической идентичности в поле науки// Роль архивов в информационном обеспечении исторической науки: сборник статей/ Авт.-сост. Е.А. Воронцова; / отв. ред. В.Ю. Афиани, Ю.А. Петров. М: Этерна, 2017. С. 374–386.

### **Участие в конференциях**

**Д. Пономарев** (3 доклада):

1. *Пленарный доклад*: Д. Пономарев. "Знания в действии: как рассуждать логически о событиях и их последствиях". Всероссийская конференция ЗОНТ (Знания-Онтологии- Теории), Новосибирск, октябрь, 2017 (доклад).

2. P. Emelyanov and D. Ponomaryov Cartesian decomposition in data analysis. SSDSE, IEEE Symposium on Data Science and Engineering, Novosibirsk, Russia, April 2017 (доклад). 3. Y.

Kazakov and D. Ponomaryov. On the complexity of semantic integration of OWL ontologies. DL '2017, The Description Logic Workshop, Montpellier, France, July 2017 (доклад).

**Городняя Л.В.** (6 докладов):

1. Международная конференция SORUCOM-2017 – 3 доклада.

2. Всероссийская научная конференция «Научный сервис в сети Интернет» (18-23 сентября 2017 г., г. Новороссийск) — 1 доклад.

3. Всероссийская научная конференция «Языки программирования и компиляторы» (2-5 апреля 2017 г., г. Ростов-на-Дону) — 1 доклад.

4. Всероссийская научная конференция «Суперкомпьютерный форум» Секция «Посткремниевые вычисления» — 1 доклад (через Скайп).

**Андреева Т.А.** (2 доклада):

1. VI Международный научно-методический симпозиум «ЭРНО-2017» (Адлер).

2. IX Всероссийская (с международным участием) научно-практическая конференция «Информационные технологии в образовании» «ИТО-Саратов-2017».

**Платонов Ю.Г.** (2 доклада):

1. 2017 Siberian Symposium on Data Science and Engineering (SSDSE) 12-13 Apr 2017 Technopark of Novosibirsk Akademgorodok, Russia. – 1 доклад (Платонов Ю.Г., Артамонова Е.В).

2. 2-я Российско-Тихоокеанская международная конференция по Компьютерным Технологиям и Приложениям (RPC 25-29 сентября 2017, г. Владивосток) – 1 доклад (Платонов Ю.Г., Артамонова Е.В.).

#### **Участие в оргкомитетах конференций, школ, олимпиад**

1. Пономарев Д.К. – член программного комитета конференции IJCAI'17 (<https://ijcai-17.org>) International Joint Conference on Artificial Intelligence (Melbourne, Australia, August, 2017).

2. Пономарев Д.К. – член программного комитета AAAI'18 (<https://aaai.org/Conferences/AAAI-18/>) The Thirty-Second AAAI Conference on Artificial Intelligence (New Orleans, USA, February 2018). *Конференции IJCAI и AAAI являются двумя ключевыми конференциями по тематике Искусственного Интеллекта.*

3. Пономарев Д.К. – член программного комитета KESW'17 (<http://2017.kesw.ru>) International Conference on Knowledge Engineering and Semantic Web (Szczecin, Poland, November 2017).

4. Пономарев Д.К. – приглашенный рецензент на конференции WG'17 (<http://www.win.tue.nl/wg2017/>) - 43rd International Workshop on Graph-Theoretic Concepts in Computer Science (Eindhoven, Netherlands, June, 2017).

5. Марчук А.Г. - научный руководитель 42-ой Летней школы юных программистов им. академика А.П. Ершова (ЛШЮП-17), ДОЛ «Солнечная поляна, д. Бурмистрово, Новосибирская область, 19 июля - 1 августа 2017г. и председатель жюри XVIII Открытой Всесибирской олимпиады по программированию им. И.В.Поттосина, Новосибирск, НГУ, ноябрь 2017 года.

6. Занина И.В. -- член оргкомитета 42-ой Летней школы юных программистов им. академика А.П. Ершова (ЛШЮП-17), ДОЛ «Солнечная поляна, д. Бурмистрово, Новосибирская область, 19 июля - 1 августа 2017г.

7. Тихонова Т.И. - координатор, член оргкомитета 42-ой Летней школы юных программистов им. академика А.П. Ершова (ЛШЮП-17). ДОЛ «Солнечная поляна, д. Бурмистрово, Новосибирская область, 19 июля - 1 августа 2017г.

#### **Членство в редколлегиях научных изданий**

1. Журнал “Bulletin of the Novosibirsk Computing Center. Series: Computer Science” — д.ф.-м.н., проф. А.Г. Марчук – член редколлегии.
2. Журнал «Проблемы информатики» ИВМ и МГ СО РАН — д.ф.-м.н., проф. А.Г. Марчук в редакционном совете.
3. Журнал «Вестник НГУ, серия: Математика, механика, информатика» — д.ф.-м.н., проф. А.Г. Марчук в редакционном совете.
4. Научный электронный журнал «Системная информатика» (сайт журнала <http://www.system-informatics.ru/>) -- проф. А.Г. Марчук — главный редактор.

### Участие в международных программах сотрудничества, зарубежные гранты, другие формы сотрудничества

**1. Проект Немецкого исследовательского сообщества** “Технологии экспертной поддержки пользователей в рамках когнитивных технических систем”

**Иностранные партнеры:** Институт искусственного интеллекта при факультете информатики университета г. Ульм, Германия (адрес: James-Franck-Ring, 89081 Ulm, Germany)

**Координаторы проекта:** Сюзанна Биундо-Штефан (Ульм, Германия), Андреас Вендемут (Магдебург, Германия).

**Участник от ИСИ Пономарев Д.К.**

**Сроки:** январь 2013 – декабрь 2017.

**Тема работы:** применение автоматизированного логического вывода в интеллектуальных системах, направленных на экспертную поддержку пользователей.

**Результаты, полученные в 2017 году:**

I. Опубликована глава Gregor Behnke, Florian Nielsen, Marvin Schiller, Denis Ponomaryov, Pascal Bercher, Birte Glimm, Wolfgang Minker, and Susanne Biundo. To Plan for the User Is to Plan with the User: Integrating User Interaction into the Planning Process.

<https://www.springerprofessional.de/to-plan-for-the-user-is-to-plan-with-the-user-integratinguser-i/15277618> в зарубежной монографии Companion Technology. A Paradigm Shift in Human-Technology Interaction. Susanne Biundo and Andreas Wendemuth, eds. Springer International Publishing, 2017. ISBN: 978-3-319-43664-7 <https://www.springerprofessional.de/companion-technology/15277572> .

II. Опубликована статья в трудах конференции:

Y. Kazakov, D. Ponomaryov. On the complexity of semantic integration of OWL ontologies. Proc. DL '2017, The Description Logic Workshop, Montpellier, France, July 2017.

<http://ceur-ws.org/Vol-1879/paper59.pdf> .

III. Начато исследование алгоритмической сложности проблемы логического следования из семантически связанных онтологий в дескрипционных логиках семейства DL-Lite.

**2. Проект РФФИ и Департамента науки и технологии правительства Индии** «Новые методы анализа и оптимизации реконфигурируемых компьютерных архитектур».

**Иностранные партнеры:** Лаборатория САПР, Суперкомпьютерный центр г. Бангалор, Индия (адрес: CEDT Building, SERC, Indian Institute of Science, Bangalore - 560 012, India).

**Координаторы проекта:** Soumitra Kumar Nandy (Бангалор, Индия), Д.К. Пономарев (Новосибирск).

**Сроки:** июнь 2017 – декабрь 2018.

**Тема работы:** Исследование новых методов конъюнктивной декомпозиции булевых функций и реализация практических алгоритмов оптимизации электронных схем на их основе.

**Результаты, полученные в 2017 году:**

Предложен метод декомпозиции таблиц реляционных баз данных, основанный на ранее полученном алгоритме факторизации булевских полиномов, имеющем полиномиальную

временную сложность. Результат имеет практическую значимость для оптимизации представления булевых функций, заданных в СДНФ, и в контексте самонастраивающихся баз данных для рефакторинга схем реляционных баз данных, поскольку позволяет обнаружить возможность более компактного представления реляционных таблиц на основе анализа их содержимого. На основе полученных результатов начата разработка нового метода AND/OR-декомпозиции булевых функций, заданных в СДНФ.

**3. Грант Благотворительного фонда В. Потанина** «Разработка магистерской программы Логические методы в информатике».

**Иностранные партнеры:** Биркбек, Университет г. Лондон, Англия (адрес: Malet Street, London WC1E 7HX, United Kingdom).

**Координаторы:** Stanislav Kikot (London, UK), Денис Пономарев (Новосибирск).

**Сроки:** июнь 2017 – июнь 2018.

**Тема работы:** Разработка учебного курса для магистерской программы по информатике.

**Результаты, полученные в 2017 году:**

Создан учебный курс уровня магистратуры “Семантические технологии работы с данными и знаниями”, включая список примерных задач и вопросов для контроля успеваемости.

**4. Совместная научная деятельность (без проекта)** с факультетом информатики университета Райерсона, г. Торонто, Канада. (адрес: 245 Church Street, Toronto, Canada).

**Координаторы:** Mikhail Soutchanski (Торонто, Канада), Денис Пономарев (Новосибирск).

**Сроки:** январь 2012 – июнь 2017.

**Тема работы:** исследование компонентных свойств теорий в Ситуационном Исчислении.

**Результаты, полученные в 2017 году:**

Опубликована журнальная статья: D. Ponomaryov and M. Soutchanski. Progression of decomposed local-effect action theories. ACM Transactions on Computational Logic (TOCL) 18(2), 2017. <http://dl.acm.org/citation.cfm?id=3091119>

Журнал входит в первый квартиль журналов по специализации «информатика» по версии Scopus.

#### **Членство в международных научных организациях**

И.А. Крайнева – член IEEE (Institute of Electrical and Electronics Engineers – Институт инженеров по электротехнике и электронике).

#### **Командировки (в том числе инициативные, не оплачиваемые Институтом)**

1. Всероссийская научная конференция «Научный сервис в сети Интернет» (18-23 сентября 2017 г., г. Новороссийск) - Городня Л.В.

## **Общая характеристика исследований Лаборатории моделирования сложных систем**

*зав. лабораторией к.ф.-м.н. Мурзин Ф.А.*

### **Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе**

**Приоритетное направление IV.39.** Архитектура, системные решения, программное обеспечение, стандартизация и информационная безопасность информационно-вычислительных комплексов и сетей новых поколений. Системное программирование.

**Программа IV.39.1.** Теоретические и прикладные проблемы создания эффективных надежных программных систем и информационных технологий

**Пункт программы ФНИ:** Алгоритмы и программные средства для моделирования сложных систем.

Цель проекта – разработка новых, совершенствование имеющихся алгоритмов и создание соответствующего наукоемкого программного обеспечения для моделирования сложных систем. Области применения: обработка сигналов и изображений, поиск нефти, обработка текстов на естественном языке и анализ информации из социальных сетей, анализ генетических последовательностей, обработка физиологических сигналов, создание систем анализа и модернизации старого программного обеспечения очень больших объемов, параллельная и распределенная обработка больших объемов данных, анализ экономических и социальных процессов, моделирование и анализ транспортных потоков и др.

Коды критических технологий: 1.2.1, 1.3.4, 1.6.4, 2.4.2

### **Исследования, вошедшие в список основных результатов Института:**

#### **Разработка системы разрешения анафоры с использованием методов машинного обучения**

Проведен комплекс исследований по компьютерной лингвистике и реализовано разнообразное программное обеспечение для обработки текстов на естественных языках: русском, английском, казахском и турецком. Рассмотрен ряд задач: разработка системы семантико-синтаксических связей для тюркских языков, определение релевантности текста поисковому запросу, определение тем текстов, применение риторических структур для анализа текстов научно-технической направленности, разрешение анафорических связей, анализ данных из социальных сетей.

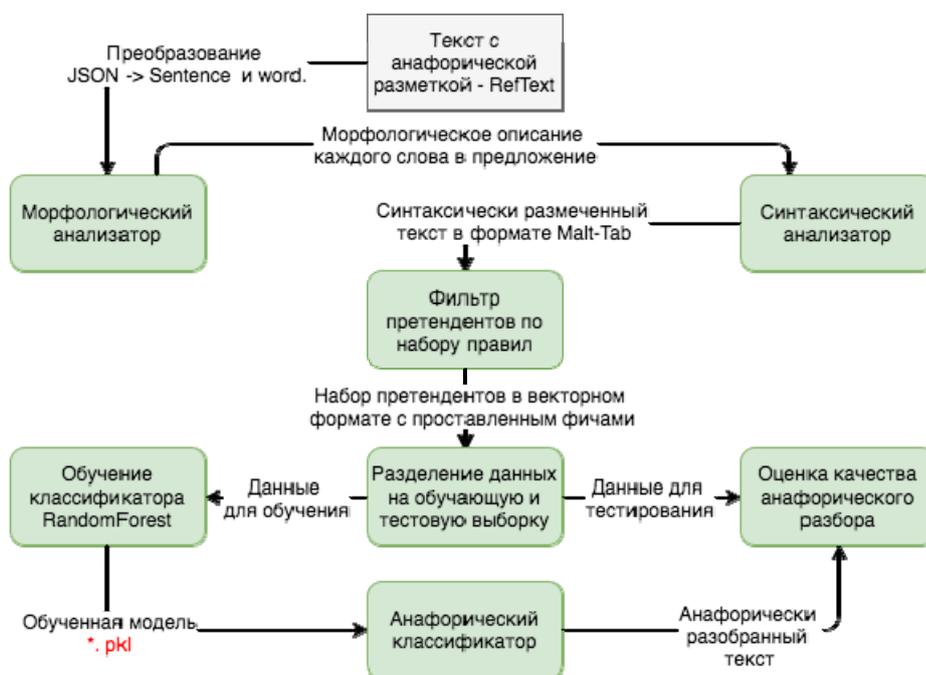


Рис. 1. Схема анафорического классификатора на основе машинного обучения

## Описание проведенных научных исследований

### 1. Исследования по теории информации

В 2017 году получены результаты по кодированию источников со счётным алфавитом. Получены совпадающие по порядку убывания оценки избыточности универсального кодирования источников с бесконечным входным алфавитом для множеств источников с ограниченными моментами. Верхние оценки были получены ранее в работах В.Ф. Бабкина, Б.Д. Кудряшова и А.В. Порова. Опубликована работа в журнале «Автометрия» и её перевод, индексируемый в Web of Science.

### 2. Исследования в области биоинформатики

#### 2.1. Алгоритмы и программные средства поиска сайтов связывания с транскрипционными факторами

Проведены доработки и дополнительные исследования алгоритма построения весовых матриц для поиска сайтов связывания, имеющих два ядра, разделенных между собой участком переменной длины (обычно от 0 до 4 нуклеотидов). Эти консервативные участки выделяются с помощью неполного перебора (называемого Gibbs Sampling).

Также проведена работа по исследованию вариаций в геноме человека с целью понимания, как эти вариации влияют на связывания сайтов различных транскрипционных факторов. Выявлены значимые отклонения в том, как ведут себя различные факторы (матрицы).

В некоторых факторах находятся позиции, в которых подверженность мутациям в 2 раза чаще, чем в случайных последовательностях. Это может быть обусловлено тем, что изменился транскрипционный фактор или по каким-то причинам природным механизмам необходимо уменьшить количество связывания с этим ТФ и поэтому в популяции больше мутаций направленных на уменьшение аффинности связывания.

## **2.2. Информационные платформы для исследований в биоинформатике**

В 2017 году проводились исследования: 1) в рамках совместной деятельности с лабораторией биоинформатики Конструкторско-технологического института вычислительной техники СО РАН по развитию программной платформы BioUML, предназначенной для моделирования биологических систем и анализа результатов экспериментов в биоинформатике; 2) в рамках сотрудничества с компанией geneXplain GmbH (Германия) по развитию программного продукта geneXplain; 3) в рамках международного проекта MIMOmics по интеграции программного инструментария для биологических исследований.

## **2.3. Анализ данных секвенирования защищённой рибосомами матричной РНК по методу RiboSeq**

Проведен первичный анализ данных секвенирования защищённой рибосомами матричной РНК по методу RiboSeq. В результате сопоставления данных от плюрипотентных стволовых и дифференцированных клеток крысы, а также сопоставления с данными секвенирования RNAseq удалось установить по-разному транскрируемые гены в процессе дифференцировки клеток. Разработанный ранее инструментарий применён для анализа уникальных данных прямого секвенирования РНК (DRS, direct sequencing) в мухе и для выявления функции длинных некодирующих генов VLINC в ответе на 30 противораковых препаратов в клеточной линии K562. Проведено сравнение точности определения уровня транскрипции генов на различных платформах секвенирования относительно метода ddPCR (цифровой капеллярной ПЦР). Усовершенствован метод выявления трисомии плода по данным секвенирования ДНК материнской крови при пренатальной диагностике генетических нарушений.

Усовершенствован программный комплекс построения оцифрованных цепей ДНК по стеку изображений флуоресцентных меток из высокопроизводительного секвенатора. Разработаны программные комплексы, реализующие методы машинного обучения для анализа спортивной статистики и звуковых дорожек. Разработан прототип алгоритма и программный комплекс для применения данных иммуносигнатур для малоинвазивной диагностики рака молочной железы. Проанализированы данные секвенирования ДНК нового поколения пациентов, страдающих кардиомиопатией. Найдены и предложены для дальнейшего анализа геномные полиморфизмы.

## **2.4. Исследование транскрипции микро РНК в плюрипотентных стволовых клетках крысы**

Опубликована статья Sherstyuk et al. в журнале Scientific Reports о полном профилировании транскрипции микро РНК в плюрипотентных стволовых клетках крысы. Найдены участвующие в процессе дифференцировки и поддержания стволовости: а) известные микро РНК крысы с известной функцией; б) известные микро РНК крысы с неизвестной функцией; в) ранее не детектируемые в крысе микро РНК, гомологичные известным микро РНК других организмов; г) полностью не известные микро РНК; д) неизвестные микро РНК гомологичные малым ядрышковым РНК. Для неизвестных микро РНК, которых найдено 374 кандидата, предсказаны гены-мишени. Проведен биоинформатический анализ биологических процессов с участием микро РНК и их генов-мишеней, предсказана их вовлечённость в процесс репрограммирования клеток.

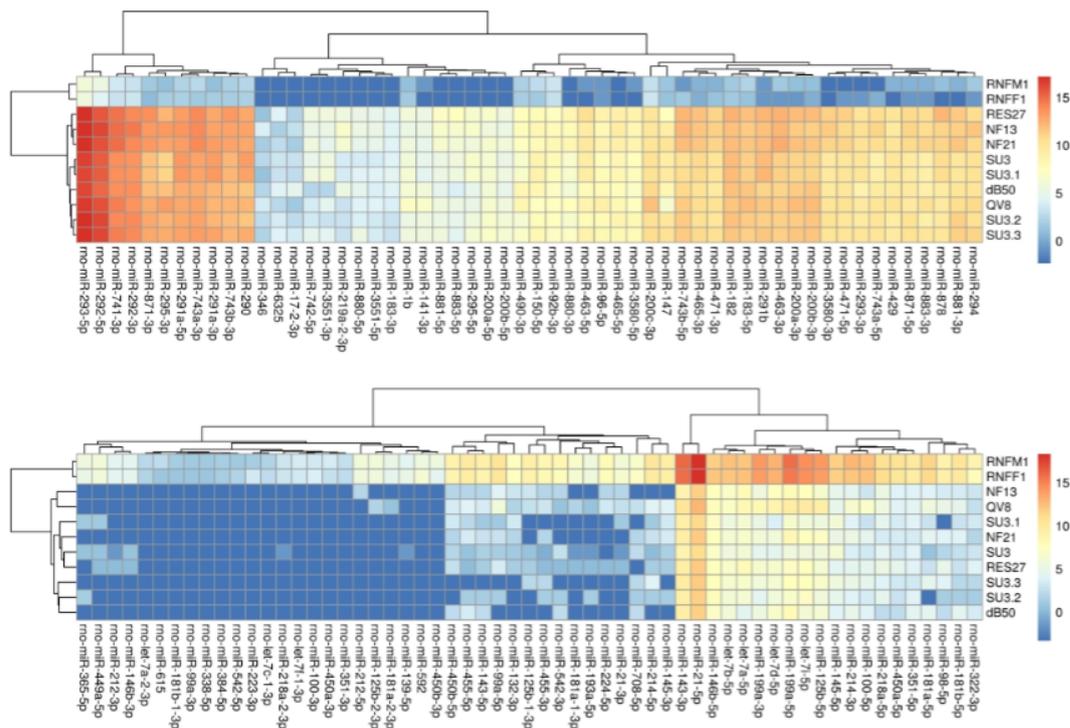
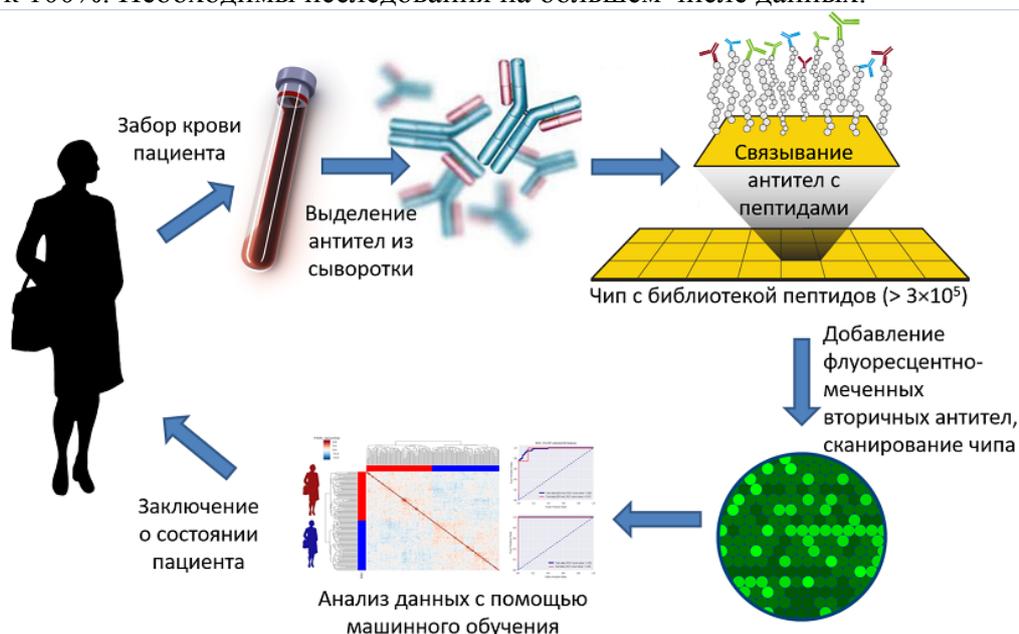


Рис. 2. Дифференциально экспрессированные микро РНК в эмбриональных стволовых клетках (RES27, dB50), индуцированных плюрипотентных стволовых клетках (NF13, QV8, SU3, NF21) и эмбриональных фибробластах (RNFM1, RNFF1). Уровень экспрессии обозначен цветом.

## 2.5. Разработка прототипа метода анализа данных иммуносигнатур

В сотрудничестве с Российско-американским противораковым центром при Алтайском государственном университете разработан прототип метода анализа данных иммуносигнатур. Метод позволяет выделять сигнал в количественных данных большого числа антител, представленных в крови и определять вероятность наличия заболевания раком молочной железы даже на ранних стадиях. На независимой тестовой выборке точность метода оказалась близка к 100%. Необходимы исследования на большем числе данных.



## **2.6. Обработка данных секвенирования нового поколения**

В настоящее время секвенирование нового поколения все чаще используется не только для фундаментальных биологических исследований, но и для решения задач трансляционной медицины. Стратегия секвенирования индивидуальных молекул, используемая компанией SeqLL (США), упрощает процесс пробоподготовки и позволяет избежать ошибок, обусловленных ПЦР-амплификацией. Ранее нами был разработан новый алгоритм анализа изображений, получаемых секвенатором SeqLL, и расшифровки нуклеотидных последовательностей секвенируемых молекул. Первоначальная обработка изображений: шумопонижение и усиление сигнала, регистрация изображений (относительно фона) осуществляется с помощью функций из библиотеки компьютерного зрения OpenCV. Для выделения областей интереса и последующего извлечения олигонуклеотидных последовательностей был разработан метод на основе коэффициентов сходства Жаккара. Алгоритм был реализован с помощью языка программирования Python (2.7.10) с использованием библиотеки компьютерного зрения cv2 (3.0.0) и пакет numpy для научных вычислений (1.9.2). Кроме того, был разработан алгоритм, восстанавливающий последовательности олигонуклеотидов на основе предсказаний качества сигналов. Предсказательная модель была создана с помощью алгоритма случайного леса (Random Forest), реализованного в библиотеке sklearn для языка программирования Python. В результате тестирования было показано, что разработанная программа демонстрирует последовательные и воспроизводимые результаты, которые хорошо согласуются с суммарной концентрацией олигонуклеотидов. По сравнению с оригинальным ПО для анализа изображений и определения последовательностей олигонуклеотидов новая программа обеспечивает большую чувствительность при аналогичном уровне специфичности – чувствительность примерно на 30% выше. В последовательностях олигонуклеотидов, определенных новой программой, содержится меньшее количество ошибочных инсерций и замен по сравнению с предыдущим алгоритмом, но при этом количество ошибочных делеций несколько выше. Среднее количество ошибок составляет 3% (3 ошибки на 100 нуклеотидов), обеспечивая высокую точность расшифровки первичных структур анализируемых нуклеиновых кислот. В настоящее время продолжается работа по оптимизации разработанного алгоритма. Ведется работа по увеличению специфичности/чувствительности нашего метода и работа над увеличением быстродействия разработанного программного комплекса. Кроме того, в настоящее время разрабатывается прототип нового алгоритма распознавания нуклеотидов, основанный на методах глубокого обучения.

## **2.7. Методы предсказания эпитопов**

Эпитоп (англ. epitope), или антигенная детерминанта – часть макромолекулы антигена, которая распознаётся иммунной системой (антителами, В-лимфоцитами, Т-лимфоцитами).

Антиген (англ. antigen от antibody-generator – «производитель антител») – любое вещество, которое организм рассматривает как чужеродное или потенциально опасное и против которого организм обычно начинает вырабатывать собственные антитела (иммунный ответ). Обычно в качестве антигенов выступают белки, однако простые вещества, даже металлы, также могут становиться антигенами в сочетании с собственными белками организма и их модификациями

В 2017 году продолжалась работа над совершенствованием методов предсказания Т-клеточных эпитопов на основе анализа аминокислотных последовательностей олигопептидов и молекул МНС. Начата работа по созданию моделей для предсказания Т-хелперных эпитопов, рестриктированных различными алломорфами молекул МНС II класса.

### **3. Алгоритмы и протоколы для распределенных децентрализованных систем хранения и обработки данных**

Продолжена разработка протоколов, обеспечивающих работу одноранговых распределенных хранилища данных высокой доступности на основе избыточно кодирования. Разработан и реализован в виде программного прототипа протокол Hydra Checksums. В отличие от предыдущего разработанного протокола Hydra Shards, данный протокол позволяет хранить объекты данных без разделения на фрагменты в рамках кластера, что позволяет их использовать напрямую, без ассемблирования. Как и в ранее разработанном, в Hydra Checksums используются коды Рида-Соломона и метод легковесной блокировка. Ранее заимствованные элементы протокола Dynamo (Amazon) были удалены, так как не применимы при хранении объектов без фрагментации.

Начата работа по формальной спецификации и верификации протоколов в парадигме eventual consistency на основе вероятностных автоматных моделей.

### **4. Алгоритмы и библиотека программных средств для анализа разупорядочения структуры клеточных стенок**

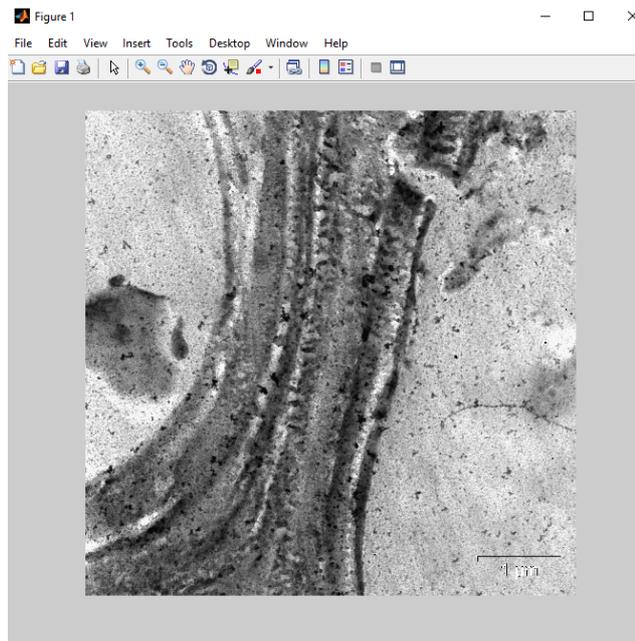
Объектами исследований и разработок в данной проекте являются: микрофотографии растительного сырья, размолотого на специальных мельницах. Работа выполнена для Института химии твердого тела и механохимии СО РАН. Исследуется вопрос о применимости наборов текстурных признаков для анализа экспериментальных данных с целью выявить на микрофотографиях характерные участки, которые в будущем можно будет увязать с пористостью, химической реактивностью и т.д.

В 2017 году были предложены алгоритмы и разработаны следующие программные средства.

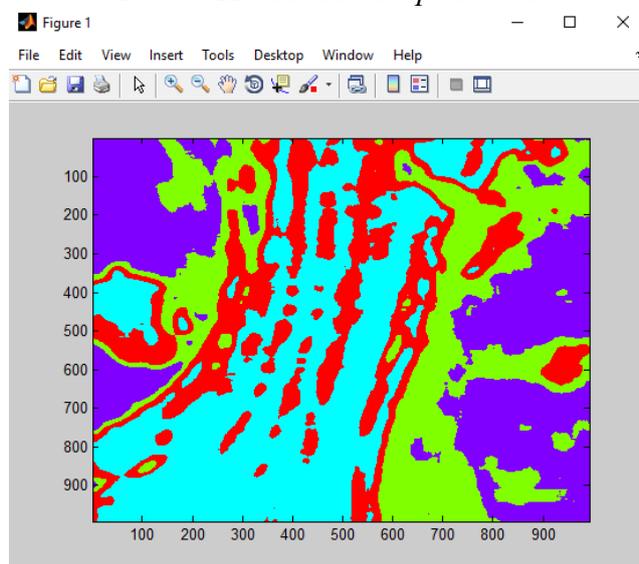
1. Усовершенствованная программа для вычисления различных текстурных признаков. Цель программы состоит в том, чтобы проводить компьютерные эксперименты и накапливать количественные данные (значения текстурных признаков) для различных участков изображений клеточных стенок, обработанных при различной температуре. Программа имеет развитые интерфейсы, позволяет выбирать прямоугольные области произвольного размера, представляющие интерес, для которых в отдельных всплывающих окнах в виде диаграмм отображаются значения текстурных признаков. Всего производится расчет 19 текстурных признаков. Результаты расчетов можно сохранять.

2. Реализованы программы для проведения кластерного анализа, в стандартном варианте и иерархического, и были проведены эксперименты

3. Реализована новая программа R/S – анализа изображений. Проведены новые эксперименты с применением R/S – анализа показавшие интересные результаты, которые могут быть полезными на практике, а именно, для поддержки работ в области химии и для диагностики рака молочной железы по изображениям, полученных с помощью тепловизора.



*Рис. 4. Исходное изображение*



*Рис. 5. Пример кластеризации параметров изображения*

## **5. Интерактивный 3d-симулятор нейромышечной системы нематоды *C. Elegans***

Продолжается работа в рамках проекта по созданию виртуального живого организма на основе нематоды *C. elegans*, а также над нашей разработкой, предназначенной для решения этой задачи - 3D симулятором *Sibernetic* ([sibernetic.org](http://sibernetic.org)), ориентированным на задачи моделирования в области биомеханики беспозвоночных. Добавлена возможность сохранения траекторий движения *C. elegans* для последующего анализа, возможность построения диаграмм кривизны тела от времени, исследована точность моделирования в зависимости от используемого шага интегрирования по времени. С помощью симулятора *Sibernetic* успешно произведена проверка возможности воспроизведения результатов эксперимента (Fouad et al., 2017), в котором передняя и задняя части тела *C. elegans* осуществляли осцилляторные движения на различных частотах.

## **6. Математическая лингвистика, обработка текстов на естественном языке и распознавание речи**

### **6.1. Создание методов обработки научно-технических текстов на основе теории риторических структур**

Исследована возможность применения теории риторических структур для анализа текстов научно-технической тематики. Формально описаны признаки некоторых отношений. Каждое отношение характеризуется набором из трех параметров: тип отношения, полярность и направление. Первый из параметров – тип отношения. Отношения могут быть трех типов:  $TYPE = \{C, T, A\}$ , C – каузальные (causal), T – темпоральные (temporal), A – аддитивные (additive). Каузальность содержит указание времени и причины, темпоральность – только времени, аддитивность не содержит каких-либо указаний. Полярность отношений подразумевает их деление на положительные (positive) и отрицательные (negative):  $POL = \{P, N\}$ . Положительность означает, что ситуация, которая представлена первой, развивается во второй, присоединенной к ней ситуации. Негативность предполагает, что ожидаемая связь ситуаций прекращается, наоборот, присутствует противопоставление. Направление может быть прямым (forward), обратным (backward) и двунаправленным (bi-directional) в зависимости от порядка упоминания событий в тексте:  $DIR = \{B, F, BD\}$ .

Сформулированы утверждения, связывающие выделенные признаки между собой. Утверждение 1. Аддитивность невозможна для обратного направления.

Пусть  $TYPE = \{C, T, A\}$  и  $DIR = \{B, F, BD\}$ .

Тогда, если  $R \in A$ , то  $R \in F$  или  $R \in BD$ , но  $R \notin B$ .

Утверждение 2. Не существует двунаправленной каузальности.

Пусть  $TYPE = \{C, T, A\}$  и  $DIR = \{B, F, BD\}$ .

Тогда, если  $R \in C$ , то  $R \in B$  или  $R \in F$ , но  $R \notin BD$ .

Утверждение 3. Не существует отрицательной двунаправленной темпоральности.

Пусть  $TYPE = \{C, T, A\}$ ,  $DIR = \{B, F, BD\}$  и  $POL = \{P, N\}$ .

Тогда, если  $R \in T$  и  $R \in BD$ , то  $R \in P$  и  $R \notin N$ .

Доказательство этих утверждений следует из определений множеств  $TYPE$ ,  $DIR$  и  $POL$ .

Автоматическое определение риторических отношений в тексте позволяет установить местоположение ядра и сателлита. Поиск ядра зависит от одного из параметров отношений – направления. Поскольку ядро содержит наиболее важную часть высказывания, то создаваемые методы предположительно будут использоваться в системах автореферирования и извлечения информации из текстов.

### **6.2. Разработка системы разрешения анафоры с использованием методов машинного обучения**

Предложен и реализован метод разрешения анафоры местоимений третьего лица в текстах на русском языке. Разработанная система включает в себя следующие компоненты:

- Морфологический и синтаксический анализ.
- Компонент создания множества кандидатов в антецеденты. Он основан на применении набора фильтров: дистанционного, морфологического, синтаксического.
- Классификатор выбора наиболее вероятного кандидата из множества предложенных, учитывающий 78 различных признаков.

В качестве алгоритма классификации был выбран алгоритм машинного обучения, основанный на деревьях принятия решений. Сам по себе процесс построения каждого дерева

можно описать следующим образом. Рассмотрим выборку  $X$ . Пусть  $x_j$  –  $j$ -й признак;  $t$  – порог значения признака;  $x_j \leq t$  – условие разбиения;  $Q(X, j, t)$  – критерий, характеризующий ошибку разбиения. Требуется найти наилучшие параметры  $j$  и  $t$ , при которых ошибка разбиения будет минимальна.

Качество работы системы оценивалось стандартными метриками: точность, полнота, аккуратность и F-мера. Получены значения точности – 53 %, полноты – 67 %, аккуратности – 64 % и F-меры – 59 %. По результатам тестирования было замечено, что качество анализатора улучшается, если не учитывать морфологический фактор падежа.

### **6.3. Методы машинного обучения и распознавание речи**

Методы глубокого обучения демонстрируют очень высокое качество в решении различных практических задач: от анализа изображений и распознавания речи, до постановки медицинских диагнозов и создания программных систем для самоуправляемых автомобилей. Начата работа по разработке ПО для распознавания собеседников в монофонических аудиозаписях. В рамках данной работы планируется использовать возможности методов глубокого обучения. Создание таких моделей поможет в сегментации записей конференций и т.п. и позволит в дальнейшем упростить аннотацию и облегчить поиск необходимой информации в таких записях. В настоящее время решение этой задачи является активной областью исследований, проводятся соревнования, направленные на поиск наилучших методов. Кроме того, отдельной важной задачей предлагаемой работы является выбор оптимальной схемы параметризации сигнала. В настоящее время разработка прототипа ПО ведется на языке программирования Python с использованием специализированных библиотек для анализа аудиосигналов (*librosa*) и для глубокого обучения (*tensorflow*, *keras*).

## **7. Исследования по геометрии и моделям пространства-времени**

### **7.1 Анализ волновых функций элементарных частиц на основе некомпактной группы $U(1,1)$**

Развиваются известные исследования, которые, в основном, были проведены математиками школы Сигала (I.E.Segal, 1918-1998, Boston, MIT, USA). Ими были проанализированы элементарные частицы на основе компактной параллелизующей группы  $U(2)$ . М. Кон, А. Levicehev ввели в рассмотрение коммутативную  $U(2)$ - $U(1,1)$  диаграмму. В итоге удалось преодолеть трудности, связанные с тем, что конформная группа  $GF = SU(2,2)$  действует на  $U(1,1)$  с сингулярностями. В 2017 году помимо исследования  $U(1,1)$ -случая, был обнаружен и исправлен серьезный пробел в доказательстве Сигала основной теоремы по  $U(2)$ -параллелизации.

### **7.2. Многоуровневая модель кварк-глюонной среды.**

Предложена модель кварк-глюонной среды, основанная на цепочке канонических (т.е., по главным минорам) вложений групп:  $U(2)$  в  $U(3)$ ,  $U(2)$  в  $U(4)$ , и т.д. Эти группы удобно называть уровнями материи:  $U(2)$  – нулевой (т.е., наш обычный),  $U(3)$  – первый,  $U(4)$  – второй и т.д. Уровни соответствуют поколениям (кварков), а «аромат» и «цвет» также вводятся строго математически. В модели получается, что кварки – это «притопленные» (на соответствующий уровень) протоны. Глюоны вводятся исходя из протонов и антипротонов, они (на каждом уровне) могут быть интерпретированы как «цветные» фотоны. Это первое описание данных объектов теоретической физики на таком уровне математической строгости. Модель совместима с обнаружением точечных компонент внутри протонов в глубоко неупругих электрон-протонных рассеяниях и с упругим рассеянием электронов на кварках. Не все из математических следствий модели совпадают со стандартными предположениями о кварках. В

частности, число цветов зависит от уровня. Модель предсказывает наличие *трёх* новых кварков четвёртого поколения. В то время как на основе Стандартной Модели сейчас ведутся поиски *двух* кварков четвёртого поколения.

### Участие в грантах

**1. MIMOmics** – Methods for Integrated analysis of Multiple Omics datasets (методы интегрированного анализа множества -омов) // Седьмая европейская рамочная программа, период контракта 2012-2017 гг., № контракта (гранта) 305280, раздел HEALTH.2012.2.1.1-3.

*Совместно с компанией geneXplain, ответственный исполнитель от ИСИ СО РАН – к.ф.-м.н. Т.Ф. Валеев*

**2. Грант Министерства образования и науки РК (аналог ФЦП)** – Разработка информационно-поискового тезауруса (с учетом морфологии казахского языка) в полнотекстовых базах данных по ИТ-технологиям, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, 2015-2017.

*Научные руководители: д.ф.-м.н. Тусупов Д.А. (Казахстан), член-корр. Федотов А.М. (НГУ), ответственный исполнитель от ИСИ СО РАН – к.ф.-м.н. Т.В. Батура*

### Участие в международных программах сотрудничества

**Совместная научная и коммерческая деятельность** ведется с партнерами из США, Германии, Китая, Казахстана и других стран. Тематика сотрудничества: анализ генетических последовательностей и микрочиповых данных; математическая лингвистика и обработка текстов на естественном языке; анализ данных из социальных сетей.

### Список публикаций лаборатории

1. Трофимов В.К., Храмова Т.В. Оптимальное равномерное по выходу кодирование для объединения различных множеств источников // Автометрия. 2017. Т.53 N1. – С 53-62. [Scopus, ВАК]
2. Trofimov V.K., Khramova T.V. Optimal output-uniform coding for a union of different sets of sources // Optoelectronics, Instrumentation and Data Processing.2017.Т.53. N1. – P. 43-50. [Scopus, ВАК] (Перевод статьи под номером 1)
3. Gureenkova O. A., Batura T. V., Kozlova A. A., Svischev A. N. Complex approach towards algorithm learning for anaphora resolution in Russian language // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue-2017”. V. 1. Iss. 16. 2017. pp. 89–97. ISSN 2221-7932 [Scopus, ВАК]
4. Barakhnin V.B., Fedotov A.M., Bakiyeva A.M., Bakiyev M.N., Tazhibayeva S.Zh., Batura T.V., Kozhemyakina O.Yu., Tussupov D.A., Sambetbaiyeva M.A., Lukpanova L.Kh. The software system for the study the morphology of the Kazakh language // The European Proceedings of Social & Behavioural Sciences. 2017. V. XXXIII. P. 18–27. URL: <http://www.futureacademy.org.uk/publication/EpSBS/ICPE2017InternationalConferenceonPsychologyandEducation> [Web of Science]
5. Мурзин Ф.А., Батура Т.В., Еримбетова А.С., Семич Д.Ф., Бакиева А.М., Ефимова Л.В. О системе поиска информации на основе грамматики связей. В сборнике: Распределенные информационно-вычислительные ресурсы. Наука-цифровой экономике (DICR-2017). Труды XVI Всероссийской конференции. Институт вычислительных технологий СО РАН. 2017. С. 110-114.
6. Бакиева А.М., Батура Т.В. Исследование применимости теории риторических структур для автоматической обработки научно-технических текстов // Cloud of Science. 2017. Т. 4. № 3. С. 450–464. [ВАК]

7. Батура Т.В. Методы автоматической классификации текстов // Программные продукты и системы. Тверь: ЗАО НИИ ЦПС, 2017. Т. 30. № 1. С. 85-99. DOI: 10.15827/0236-235X.030.1.085-099. [BAK]
8. Барахнин В. Б., Федотов А. М., Бакиева А. М., Бакиев М. Н., Тажибаева С. Ж., Батура Т. В., Кожемякина О. Ю., Тусупов Д. А., Самбетбаева М. А., Лукпанова Л. Х. Алгоритмы генерации и стемматизации словоформ казахского языка // Cloud of Science. 2017. Т. 4. № 3. С. 434–449. [BAK]
9. Соколов А.В., Батура Т.В. Разработка системы разрешения анафоры на основе методов машинного обучения // Программные продукты и системы. Тверь: ЗАО НИИ ЦПС, 2017. Т. 30. № 3. С. 461–468. DOI: 10.15827/0236-235X.030.3.461-468. [BAK] URL: <http://www.swsys.ru/files/2017-3/461-468.pdf>
10. Барахнин В.Б., Бакиева А.М., Бакиев М.Н., Тажибаева С.Ж., Батура Т.В., Лукпанова Л.Х. Стемматизация и генерация словоформ казахского языка для систем автоматической обработки текстов // Вычислительные технологии. Новосибирск: ИВТ. 2017. Т. 22, № 4. С. 11–21. [BAK, zbMath] URL: <http://www.ict.nsc.ru/jct/annotation/1811>
11. Kozlova A., Gureenkova O., Svishev A., Batura T. A hybrid approach for anaphora resolution in the Russian language // IEEE, Proceedings of 2017 Siberian Symposium on Data Science and Engineering (SSDSE). 2017. Novosibirsk, 12–13 April, 2017. P. 36–40. ISBN 978-1-5386-1592-8 [IEEE Xplore] URL: <http://ieeexplore.ieee.org/document/8071960/>
12. Barakhnin V., Bakiyeva A., Batura T. Stemming and word forms generation in automatic text processing systems in the Kazakh language // The 15th International Scientific Conference Information Technologies and Management (IT&M 2017), April 27–28, 2017, ISMA University, Riga, Latvia. pp. 85–86. ISSN 1691-2489. URL: [http://geoml.info/wp-content/uploads/2016/02/ITM2017\\_Theses.pdf](http://geoml.info/wp-content/uploads/2016/02/ITM2017_Theses.pdf)
13. Бакиева А.М., Батура Т.В. Методы автоматического реферирования и определения тем текстов // Материалы XVIII Всероссийской конференции молодых ученых по математическому моделированию и информационным технологиям. г. Иркутск, Россия, 21–25 августа 2017. Новосибирск: ИВТ СО РАН, 2017. С. 65. URL: [http://conf.nsc.ru/files/conferences/ym2017/419536/YM2017\\_Abstracts.pdf](http://conf.nsc.ru/files/conferences/ym2017/419536/YM2017_Abstracts.pdf)
14. Бакиева А.М., Батура Т.В. Система автоматического реферирования и определения тем научно-технических текстов // XVI Российская конференция «Распределенные информационно-вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4–7 декабря 2017 г.). 2017. Новосибирск: ИВТ СО РАН. С. 75–80. ISBN 978-5-905569-10-4.
15. Бакиева А.М., Батура Т.В. Применение теории риторических структур для автоматической обработки текстов // Марчуковские научные чтения - 2017 (MSR 2017). Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. С. 149.
16. Семенова Н. А., Батура Т. В. Автоматическое составление словаря оценочной лексики для конкретной предметной области // Марчуковские научные чтения - 2017 (MSR 2017). Тезисы. Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. С. 159.
17. Murzin F.A., Batura T.V., Yerimbetova A.S., Semich D.F., Bakiyeva A.M., Efimova L.V. About the search information system based on link grammar // Distributed Information-Computational Resources. Science - The Digital Economy (DICR-2017): Proceedings of the XVI All-Russian conference, Novosibirsk, Akademgorodok, 4 - 7 December 2017. Novosibirsk: ICT SB RAS, 2017. – P. 100–114, ISBN 978-5-905569-10-4.
18. Batura T.V., Murzin F.A., Semich D.F., Bakiyeva A.M., Yerimbetova A.S. Link Grammar Parser and estimation of the document relevance to the search query // Марчуковские научные чтения - 2017 (MSR 2017). Тезисы. Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. – С. 200.
19. G.B. Abdikerimiva, A.L. Bychkov, Wei Xin Yu, F.A.Murzin, N.E.Russkikh, E.I.Ryabchikova, S.S.Khairulin. Algorithms and software for analysis of textural features and the space spectral

- properties of images. // Марчуковские научные чтения - 2017 (MSR 2017). Тезисы. Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. – С. 200.
20. Мурзин Ф.А., Еримбетова А.С., Сагнаева С.К., Батура Т.В., Семич Д.Ф., Бакиева А.М. Алгоритмы и программные инструменты для определения релевантности текста поисковому запросу и определения тем текстов. // Международная конференция «Актуальные проблемы чистой и прикладной математики, посвященная 100-летию со дня рождения академика Тайманова Асана Дабсовича. Алматы 22-25 августа 2017 года. Тезисы докладов. С. 141-142.
21. Абдикеримова М., Бычков А., Синьют В., Мурзин Ф., Русских Н., Рябчикова Е., Хайруллин С. Алгоритмы и программный инструментарий для анализа текстур на микрофотографиях. // Международная конференция «Актуальные проблемы чистой и прикладной математики, посвященная 100-летию со дня рождения академика Тайманова Асана Дабсовича. Алматы 22-25 августа 2017 года. Тезисы докладов. С.124-126.
22. Мурзин Ф.А., Еримбетова А.С., Батура Т.В., Семич Д.Ф., Ефимова Л.В., Бакиева А.М. О новых инструментах поиска информации на основе грамматики связей // В сборнике: Интеллектуальный анализ сигналов, данных и знаний: методы и средства сборник статей Всероссийской научно- практической конференции с международным участием. Новосибирский государственный технический университет. 2017. – С. 161-166.
23. Абдикеримова Г.Б., Бычков А.Л., Вей С., Мурзин Ф.А., Русских Н.Е., Рябчикова Е.И., Хайруллин С.С. Библиотека программ для анализа текстур на микрофотографиях // В сборнике: Интеллектуальный анализ сигналов, данных и знаний: методы и средства сборник статей Всероссийской научно- практической конференции с международным участием. Новосибирский государственный технический университет. 2017. – С. 98-102.
24. Мурзин Ф.А., Сагнаева С.К., Еримбетова А.С., Дайырбаева Э.Н. Разработка системы связей для тюркских языков // Вестник Казахской академии транспорта и коммуникаций им. М. Тынышпаева. 2017. № 3 (102). – С. 102-107.
25. V.Sherstyuk, S.Medvedev, E.Elisaphenko, E.Vaskova, M.Ri, Y.Vyatkin, O.Saik, D.Shtokalo, E.Pokushalov, S.Zakian. Genome-wide profiling and differential expression of microRNA in rat pluripotent stem cells // Scientific reports. 2017 Jun 5;7(1):2787. doi: 10.1038/s41598-017-02632-0. (Scopus, Web of Science)
26. St. Laurent G III, Seilheimer B, Tackett M, Zhou J, Shtokalo D, Vyatkin Y, Ri M, Toma I, Jones D and McCaffrey TA. Deep Sequencing Transcriptome Analysis of Murine Wound Healing: Effects of a Multicomponent, Multitarget Natural Product Therapy-Tr14. // Front. Mol. Biosci. 4:57. 2017. doi: 10.3389/fmolb.2017.00057.
27. David Lung, Stephen Larson, Andrey Palyanov, Sergey Khayrulin, Pdraig Gleeson, Manuel Zimmer, Radu Grosu and Ramin M. Hasani. A Simplified Cell Network for the Simulation of C. elegans' Forward Crawling // Proc. NIPS 2017 Workshop on Worm's Neural Information Processing
28. Palyanov A.Y., Chekmarev S.F. Hydrodynamic description of protein folding: the decrease of the probability fluxes as an indicator of transition states in two-state folders. [Journal of Biomolecular Structure and Dynamics](#). 2017. T. 35. № 14. С. 3152-3160.
29. D.V.Antonets, N.E. Russkikh Using machine learning approach to improve base calling in next generation sequencing data // Moscow Conference on Computational Molecular Biology (MCCMB), abstracts, p. 182, Moscow, Russia, July 27-30.
30. Levichev A. V. --- On key properties of the intertwining operators ornament in the matrix multi-level model of the quark-gluon media // В Материалах Всероссийской конференции с международным участием "Знания-Онтологии-Теории" (ЗОНТ-2017), 6-8 октября 2017, ISSN 0568-661X Под ред. Д.Е.Пальчунова, Новосибирск: Изд-во Ин-та математики, 2017. Том 2. С. 41--47.
31. Левичев А.В. --- О сингулярных  $SU(n,n)$ -действиях в группах  $U(p,q)$ :  $n=p+q=2$  // Междунар. конф. Математика в современном мире: Тез. докл. Нск, 2017. 125.

32. Левичев А.В. --- О построении операторов переплетения в многоуровневой модели кварк-глюонной среды: свойства орнамента первого уровня// Междунар. конф. Математика в современном мире: Тез. докл. Нск, 2017. 578.
33. Левичев А.В. --- On  $U(2)$ -, and  $U(1,1)$ -parallelizations of chronometric bundles// Междунар. конф. Дни геометрии в Новосибирске — 2017: Тез. докл. Нск, 2017. 14.
34. Иртегов Д.В., Нестеренко Т.В., Чурина Т.Г. Разработка систем автоматизированной оценки знаний по программированию. Системная информатика. 2017. № 11. С. 91-116.
35. Карасюк П.К., Мигинский Д.С. Протокол для децентрализованной системы хранения с избыточным кодированием. Программные продукты и системы. 2017. № 4. С. 699-705.
36. Yevshin I.S., Sharipov R.N., Valeev T.F., Kel A.E., Kolpakov F.A. Trd: a database of transcription factor binding sites identified by chip-seq experiments В книге: Высокопроизводительное секвенирование в геномике II Всероссийская конференция с международным участием. Сер. "ActaNaturau" 2017. С. 25.
37. Yevshin I., Sharipov R., Valeev T., Kel A., Kolpakov F.G. Trd: a database of transcription factor binding sites identified by chip-seq experiments Nucleic Acids Research. 2017. Т. 45. № D1. С. D61-D67.

### **Свидетельства о государственной регистрации программ для ЭВМ**

1. Свидетельство о государственной регистрации программы для ЭВМ «Модуль анализа информации: классификатор текстов» №2017611829 от 09.02.2017г.

*Авторы: Павловский Е.Н., Масловский И.А., Батура Т.В. (ИСИ), Дюбанов В.В.*

*Правообладатель: НГУ.*

# Общая характеристика исследований лаборатории теории параллельных процессов

*Зав. лабораторией д.ф.-м.н. Вирбицкайте И.Б.*

## Основные результаты научных исследований за год, их практическое использование и применение в учебном процессе

**Кадровый состав:** всего сотрудников — 9, из них научных сотрудников — 8 (в том числе 1 доктор и 7 кандидатов наук).

### Основные направления исследований:

- теоретико-категорное исследование взаимосвязей параллельных моделей с реальным временем и их эквивалентностей;
- изучение свойств достижимости, безопасности, управления моделей различных классов динамических и гибридных систем;
- разработка дискретно-временных стохастических расширений алгебр параллельных процессов, построение стохастических алгебраических и поведенческих эквивалентностей и исследование их взаимосвязей;
- проектирование алгоритмов верификации различных классов гибридных и параллельных моделей.

## Краткое описание проведенных научных исследований в 2017 году

### 1. Анализ сложности вычислений с непрерывными данными

Развита теория оценки арифметической сложности проблем над непрерывными данными. Доказана  $Pi-0-2$  полнота проблем равенства и включения эффективно открытых подмножеств евклидовых пространств. Для них также установлена справедливость теоремы Райса и теоремы Райса-Шапира. До настоящего времени был открыт вопрос о семантическом описании вычислимости частичных вещественно-значных функций, определенных на эффективно перечислимых топологических пространствах. Предложен класс мажорантно-вычислимых вещественно-значных функций, определенных на эффективно перечислимых топологических пространствах. Для этого класса доказана теорема Райса, установлена  $Sigma-1-1$  полнота проблемы равенства функций,  $Pi-1-1$  полнота проверки всюду неопределенности. Также исследована дескриптивная сложность образов частично вычислимых функций над вычислимыми Польскими пространствами. Доказана теорема о  $Sigma-1-1$  полноте образов. Были предложены требования на эффективно перечислимые  $T_0$ -пространства, которые гарантируют выполнимость теоремы Райса-Шапира для вычислимых элементов этих пространств. Ослабление этих требований приводит классам, на которых теорема Райса-Шапира не выполняется. Предложены конструкции эффективно перечислимых топологических пространств по  $wp$ -семействам и вычислимым деревьям без вычислимых бесконечных путей.

Развита высшая эффективная дескриптивная теория множеств, т.е. ряд теорем эффективной ДТМ перенесены на некоторый класс эффективных пространств, строго содержащий классы вычислимых польских и вычислимых квази-польских пространств.

В *iRRAM*-пакете реализованы операции с обобщенными линейными моделями Тейлора. Проведен сравнительный анализ сложности вычислений с действительными числами на основе различных их представлений, в частности интервальных и обобщенными линейными моделями Тейлора. Разработаны и реализованы методы представления непрерывных констрейтов в *iRRAM*-пакете, реализован алгоритм проверки истинности линейных констрейтов.

## 2. Анализ корректности моделей параллельных процессов и процессов с временными и стохастическими характеристиками

Определена эквивалентность флюидных стохастических сетей Петри (ФССП) на базе эквивалентностей стохастических сетей Петри, извлекаемых из комкуемости (lumpability) для цепей Маркова и некоторых классов дифференциальных уравнений. Комкуемость для дифференциальных уравнений основана на конечной дискретизации и перестановках в непрерывной части ФССП. Агрегирование эквивалентных состояний, дающее сокращение их числа, применено для более эффективного анализа ФССП с использованием дискретизации. Показано, что комкуемые отношения эквивалентности можно получить как из дискретизированного пространства состояний стохастического процесса, так и непосредственно из ФССП.

Предложены новые подходы к определению флюидных бисимуляционных эквивалентностей мест ПФССП посредством поднятия (lifting) бисимуляции непрерывных мест на гибридные маркировки на основе действительных расширений мультимножеств, а также применением дуг детерминированного скачка флюида, описанным для расширенных ФССП.

Для флюидных следовой и бисимуляционной эквивалентности помеченных флюидных стохастических сетей Петри (ПФССП) даны логические характеристики новыми флюидными модальными логиками, разработанными на базе логики Хеннесси-Милнера. Флюидная бисимуляционная эквивалентность применена для упрощения качественного и количественного анализа ПФССП с использованием категоризации их поведенческих формализмов: дискретных графов достижимости, непрерывно-временных цепей Маркова и стохастических флюидных моделей. Построены прикладные примеры системы подготовки документов и производственной линии, демонстрирующие поведенческий анализ на основе категоризации по флюидной бисимуляционной эквивалентности, а также верификацию и сравнение линейно-временного и ветвисто-временного поведения с помощью формул упомянутых флюидных логик.

В рамках дискретно-временного стохастического исчисления боксов Петри с мгновенными мультидействиями (dtsiPBC) разработаны следующие поведенческие эквивалентности: структурная, относительно систем переходов и шаговая стохастическая бисимуляционная. Проведено сравнение различающей способности всех этих отношений. Последняя эквивалентность применена для сокращения соответствующих алгебраическим процессным выражениям поведенческих структур, таких как помеченные вероятностные системы переходов, полумарковские цепи и дискретно-временные цепи Маркова, с сохранением характеристик функциональности и производительности. С помощью интерпретации вероятностей стохастических мультидействий как параметров алгебраических спецификаций выяснен эффект изменения этих вероятностей на стационарную функцию масс вероятностей категоризованной полумарковской цепи и на соответствующие индексы производительности обобщенной в прикладном примере системы с разделяемой памятью, что дало возможность оптимизации ее количественного поведения и повышения производительности.

В контексте реляционных структур, являющихся расширением событийно-ориентированных моделей за счет возможности описания отношений слабой/сильной причинной зависимости, синхронного/асинхронного параллелизма, симметричного/асимметричного конфликта и т.д. между событиями параллельных/распределенных систем, введено понятие  $K$ -плотности, обобщающее различные модификации «аксиом параллельности», изученные в контексте моделей с классическими отношениями причинной зависимости, параллелизма и конфликта. Установлена взаимосвязь между различными вариантами свойства  $K$ -плотности при выполнении свойств треугольничности-свободы (отсутствия локальных тупиков). Также, исследованы различные подходы к

определению свойства N-плотности (локальной K-плотности) и установлено влияние выбора подхода на изменение характеристик K-плотности.

Разработаны два метода построения абстракций в виде систем переходов, представляющих изменения структуры и поведения при функционировании различных классов (расширенных первичных, расслоенных, потоковых, стабильных, дуальных, с разрешимым конфликтом, конфигурационных) событийно-ориентированных моделей. В первом методе в качестве состояний генерируемой системы переходов рассматриваются конфигурации (выполненные процессы) моделей, а во втором – «остаточные» структуры, т.е. что осталось от структуры исходной модели после выполнения ее конфигурации. Для формирования «остаточных» структур событий для каждого класса моделей был разработан и исследован оператор «обрезки», который не удаляет события, конфликтные с событиями в конфигурации, а преобразует их в невозможные для выполнения. Благодаря этому, удалось доказать изоморфизм двух генерируемых систем переходов в контексте шаговой семантики, что позволит перенести ряд известных из литературы результатов для «поведенческих» систем переходов на «структурные» и получить теоретико-категорную характеристику операторов построения таких систем переходов.

Для случая произвольных конечных систем переходов над двоичным алфавитом построена характеристика генерируемых сетями Петри пространств состояний для общего случая и случая нереклексивных сетей. Характеристика представляет генерируемые пространства состояний в виде минимальных выпуклых оболочек множеств точек геометрического пространства. На основе характеристики представлен алгоритм минимальной аппроксимации сверху конечных языков языками сетей Петри.

### **Список публикаций лаборатории**

#### **Центральные издания**

7. Tarasyuk I.V., Buchholz P. Equivalences for fluid stochastic Petri nets. Siberian Electronic Mathematical Reports 14, p. 317-366, S.L. Sobolev Institute of Mathematics, Novosibirsk, April 2017, <http://semr.math.nsc.ru/v14/p317-366.pdf> (Web of Science, Scopus, Zentralblatt Math indexed).
8. Korovina M., Kudinov O. On images of partial computable functions over computable Polish spaces. Siberian Electronic Mathematical Reports, 2017, v. 14, pp. 418-432. (Web of Science, Scopus)
9. Korovina M., Kudinov O. Spectrum of the field of computable real numbers. Algebra and Logic, 2017, Vol. 55, pp. 485-500 (Web of Science, Scopus).

#### **Зарубежные издания**

7. Korovina M., Kudinov O. Computable elements and functions in effectively enumerable topological spaces // Mathematical Structures in Computer Science, Vol. 27, Issue 8, pp. 1466-1494, 2017. (Web of Science, Scopus).
8. Ishihara H., Korovina M., Pauly A., Seisenberger M., Spreen D. (2017) Preface to the special issue: Continuity, computability, constructivity: from logic to algorithms 2013 // Mathematical Structures in Computer Science Vol. 27, Issue 8, pp. 1285-1286. (Web of Science, Scopus).
9. Korovina M., Kudinov O. On higher effective descriptive set theory // Lecture Notes in Computer Science, 2017, Vol. 10307, pp. 282-291. (Scopus).
10. Korovina M., Kudinov O. Outline of partial computability in computable topology // Lecture Notes in Computer Science, 2017, Vol. 10307, pp. 64-76. (Scopus).
11. Korovina M., Kudinov O. Highlights of Rice-Shapiro theorem in computable topology // Logical Methods in Computer Science, 2017 (В печати) (Scopus).
12. Bause F., Buchholz P., Tarasyuk I.V., Telek M. Equivalence and lumpability of FSPNs // Lecture Notes in Computer Science Vol. 10378, pp. 16-31, 2017. (Scopus).

13. Best E., Gribovskaya N., Virbitskaite I. Configuration- and Residual-Based Transition Systems for Event Structures with Asymmetric Conflict // Lecture Notes in Computer Science, 2017, Vol. 10139, pp. 132-146. (Scopus).

#### **Материалы международных конференций**

4. Bozhenkova, E., Virbitskaite, I.: Some Space-Time Characteristics of Relational Structures // Proceedings of Fourth Russian-Finnish Symposium (RuFiDim), May 2017, TUCS Lecture Notes – 2017 – No. 26 – pp. 27-38. (ISSN 1797-8831 (digital), ISSN 1797-8823 (print))
5. Erofeev, E., Wimmel, H.: Reachability Graphs of Two-Transition Petri Nets // Proceedings of the International Workshop on Algorithms & Theories for the Analysis of Event Data, 2017, CEUR Workshop Proceedings, Vol. 1847, P. 39-54.
6. Brauße F., Korovina M., Müller N. Th.: ERA: Applications, Analysis and Improvements // Proceedings of International Conference “Continuity, Computability, Constructivity - From Logic to Algorithms” (CCC 2017), France, 2017, pp. 27-30.

#### **Прочие публикации**

1. Tarasyuk I.V., Macia S.H., Valero R.V. Stochastic equivalence for performance analysis of concurrent systems in dtsiPBC. CoRR abs/1702.07478 (arXiv:1702.07478), 69 p., Computing Research Repository, Cornell University Library, Ithaca, NY, USA, February 2017.
2. Tarasyuk I.V., Buchholz P. Behavioural equivalences for fluid stochastic Petri nets. CoRR abs/01706.2641 (arXiv:1706.02641), 54 p., Computing Research Repository, Cornell University Library, Ithaca, NY, USA, June 2017, <http://itar.iis.nsk.su/files/itar/pages/fspneqarxiv.pdf>, <http://arxiv.org/pdf/1706.02641.pdf>. DBLP indexed.
3. Bozhenkova, E.: The notion of N-density for Relational Structures. Bulletin of the Novosibirsk Computing Center, Series Computer Science, IIS Special Issue 40, pages 17-26, NCC Publisher, Novosibirsk, Russia, 2017 (ISSN 1680-6972)

#### **Участие в международных программах сотрудничества, зарубежные гранты**

##### **1. Сравнительный анализ и верификация для параллельных систем с повышенными требованиями к корректности**

###### **Проект DFG (CAVER, грант No BE 1267/14-1)**

###### **Иностранные партнеры:**

- А. Бест (Университет Ольденбурга им. Карла фон Оссетского, 26111 Ольденбург, Германия)
- П. Бухгольц (Дортмундский университет, Отто-Хан-Штрассе, 16, 44227 Дортмунд, Германия)
- Н. Мюллер (Факультет Информатики, Трирский Университет, D-54286 Трир, Германия)
- Л. Попова-Цейгманн (Берлинский университет им. Гумбольдта, ул. Унтер ден Линден, 10099 Берлин, Германия)

**Координаторы:** проф. И.Б. Вирбицкайте (ИСИ СО РАН им. А.П. Ершова), проф. А. Бест (Ольденбургский университет, Германия)

**Сроки:** 06.2014-12.2018гг.

###### **Результаты, полученные в 2017 году:**

Построены и реализованы в iRRAM-пакете новые методы работы с обобщенными моделями Тейлора. На основе экспериментальных результатов проведена оценка сложности алгоритмов и сравнительный анализ с классическими интервальными методами.

Определена и исследована эквивалентность флюидных стохастических сетей Петри (ФССП) на основе комкуемости (lumpability) для цепей Маркова и дифференциальных

уравнений. Для флюидных следовой и бисимуляционной эквивалентностей помеченных флюидных стохастических сетей Петри (ПФССП) построена и исследована логическая характеристика оригинальными флюидными модальными логиками.

В контексте модели реляционных структур, обобщающих различные событийные модели параллельных процессов, установлены взаимосвязи «аксиом параллелизма», позволяющие изучать отношения между событиями параллельных/распределенных систем.

Предложены методы построения и сопоставления изоморфных абстракций в виде систем переходов, представляющих изменения структуры и поведения при функционировании событийно-ориентированных моделей различных классов.

Для случая произвольных конечных систем переходов над двоичным алфавитом построена характеристика генерируемых сетями Петри пространств состояний, позволяющая осуществить минимальную аппроксимацию сверху конечных языков языками сетей Петри.

## **2. Формальный анализ и приложения веб-сервисов и электронных контрактов**

**(подпроект: Формальная разработка и анализ сложных систем в распределенных контекстах: основания, инструментальные средства и приложения)**

Университет Кастильи - Ла Манчи, г. Альбасете

Проект Испанского правительства (Министерство науки и инноваций), Европейский фонд регионального развития (FEDER), грант TIN2015-65845-C03-02.

**Руководители проекта с испанской стороны:** Мария Эмилия Камбронеро Пикерас (Maria Emilia Cambronero Piqueras) и Грегорио Диас Дескальсо (Gregorio Diaz Descalzo); Высшая школа инженерии информатики, Университет Кастильи – Ла Манчи, Авенида де Эспанья, б/н, 02071, г. Альбасете, Испания.

**Участник проекта с российской стороны:** к.ф.-м.н. Тарасюк И.В.(ИСИ СО РАН им. А.П. Ершова)

**Сроки:** 01.01.2016 - 31.12.2018

**Результаты, полученные в 2017 году:**

В рамках дискретно-временного стохастического исчисления боксов Петри с мгновенными мультидействиями (dtsiPBC) разработаны поведенческие отношения изоморфизма, эквивалентности относительно систем переходов и шаговой стохастической бисимуляционной эквивалентности, а также выяснены их взаимосвязи. Последняя эквивалентность применена для сокращения соответствующих алгебраическим процессным выражениям поведенческих структур, таких как помеченные системы переходов, полумарковские цепи и дискретно-временные цепи Маркова, с сохранением характеристик функциональности и производительности. В прикладном примере обобщенной системы с разделяемой памятью описан метод моделирования, оценки производительности и сохраняющей поведение редукции параллельных систем на основе категоризации (по данной эквивалентности), которая позволяет упростить анализ таких систем. С помощью интерпретации вероятностей стохастических мультидействий в качестве параметров алгебраических спецификаций выяснен эффект изменения этих вероятностей на стационарную функцию масс вероятностей категоризованной полумарковской цепи и на соответствующие индексы производительности обобщенной системы с разделяемой памятью, что дало возможность оптимизации ее поведения и повышения производительности.

### **Командировки**

**(в том числе инициативные, не оплачиваемые Институтом)**

1. Боженкова Е.Н. (15.05.2017 - 19.05.2017) – участие с докладом в 4-ом Российско-Финском симпозиуме по дискретной математике, г. Турку, Финляндия.

2. Вирбицкайте И.Б. (01.01.2017 - 22.01.2017) – научная работа в рамках международного проекта CAVER DFG/РФФИ, Ольденбургский университет, г. Ольденбург, Германия.
3. Вирбицкайте И.Б. (15.05.2017 - 19.05.2017) – участие с докладом в 4-ом Российско-Финском симпозиуме по дискретной математике, г. Турку, Финляндия.
4. Коровина М.В. (13.06.2017-17.06.2017) - участие с докладом в конференции «Вычислимость в Европе», Университет г. Турку, Финляндия.
5. Коровина М.В. (16.11.2017 - 08.12.2017) – научная работа в рамках международного проекта CAVER DFG, Университет г. Трира, Германия.
6. Коровина М.В. (26.11.2017 - 01.12.2017) - участие с докладом в «Dagstuhl Seminar 17481 Reliable Computation and Complexity on the Reals», г. Вадерн, Германия.
7. Коровина М.В. (29.03.2017 - 08.05.2017) научная работа в рамках международного проекта CAVER DFG, Университет г. Трира, Германия.

### **Список принятых в институте иностранцев**

1. Норберт Мюллер (Norbert Mueller)  
Дата рождения: 25.01.1960  
Должность: Профессор, зам. директора по науке, Университет г.Трир, Германия  
Адрес: FB IV – Факультет Информатики, Университет Трира, D-54286 Трир, Германия  
Дата визита: 20.03.2017 – 29.03.2017
2. Франц Браузе (Franz Brausse)  
Дата рождения: 21.03.1987  
Должность: аспирант, Университет г.Трир, Германия  
Адрес: FB IV – Факультет Информатики, Университет Трира, D-54286 Трир, Германия  
Дата визита: 26.06.2017 – 23.07.2017

### **Членство в программных комитетах международных конференций**

д.ф.-м.н. Вирбицкайте И.Б.:

1. the 38th International Conference on Application and Theory of Petri Nets and Concurrency (25-30 June 2017, Zaragoza, Spain)
2. 11-ая международная конференция «Перспективы систем информатики» (27 – 29 июня 2017 г., г. Москва)

### **Членство в редколлегиях научных изданий**

Проф. И.Б. Вирбицкайте -- Journal “Programming and Computer Software” (Журнал «Программирование»)

## Сводные данные по институту

### Деятельность Ученого совета

За отчетный период проведено 7 заседаний Ученого совета, на которых обсуждались различные вопросы деятельности Института. Важнейшие из них: о финансовом положении Института; о планах редакционной подготовки на 2018 год; о планах проведения конференций; об итогах годового Общего собрания СО РАН и РАН; о реформе РАН; о подготовке основных заданий к плану НИР на 2018 год; о важнейших результатах Института по итогам научной деятельности в 2017 году; о работе аспирантуры Института. Кроме того, рассматривались различные кадровые вопросы.

### Издательская деятельность

В 2017 году продолжалась публикация статей в электронном журнале «Системная информатика», <http://www.system-informatics.ru/>. Институтом подготовлен выпуск BULLETIN of the Novosibirsk Computing Center. Series: Computer Sciences. В Мемориальной библиотеке А.П.Ершова ежемесячно проводились выставки новой литературы.

### Международные научные связи

В 2017 г. Институт систем информатики имени А.П.Ершова СО РАН осуществлял сотрудничество с зарубежными организациями по следующим грантам:

**Международный проект « Сравнительный анализ и верификация для параллельных систем с повышенными требованиями к корректности — Проект DFG (CAVER, грант No BE 1267/14-1) и РФФИ (грант No 14-01-91334)**

*Руководители: А. Бест (Ольденбургский университет, Германия), П. Бухгольц (Дортмундский университет, Германия), Л. Попова-Цейгманн (Берлинский университет, Германия)*

*Участник: И.Б. Вирбицкайте*

*Сроки: 2014 – 2017 гг.*

**Международный проект «Computable analysis – theoretical and applied aspects», EU—грант № PIRSES-GA-2011-294962**

*Руководители: Дитер Шприн (Зиген, Германия), Виктор Селиванов (ИСИ СОРАН)*

*Участник: Коровина М.В.*

*Сроки: 2014 – 2017 гг.*

**Международный проект Испанского правительства "Formal Analysis and Applications of Web Services and Electronic Contracts (DArDOS)", грант TIN2015-65845-C03-02.**

*Руководитель: Maria Emilia Cambronero Piqueras, Gregorio Diaz Descalzo*

*Участник: Тарасюк И.В.*

*Сроки: 2016 – 2018 гг.*

## **Проект Немецкого исследовательского сообщества “Технологии экспертной поддержки пользователей в рамках когнитивных технических систем”**

*Иностранные партнеры: Институт искусственного интеллекта при факультете информатики университета г. Ульм, Германия*

*Координаторы проекта: Сюзанна Биундо-Штефан (Ульм, Германия), Андреас Вендемут (Магдебург, Германия)*

*Участник: Пономарев Д.К.*

*Сроки: 2013 – 2017 гг.*

## **Седьмая европейская рамочная программа, № контракта (гранта) 305280, раздел HEALTH.2012.2.1.1-3. MIMOmics – Methods for Integrated analysis of Multiple Omics datasets (методы интегрированного анализа множества омов)**

*Совместно с компанией geneXplain, ответственный исполнитель от ИСИ СО РАН – к.ф.-м.н.*

*Т.Ф. Валеев*

*Сроки: 2012 – 2017*

## **Грант Министерства образования и науки РК (аналог ФЦП) – Разработка информационно-поискового тезауруса (с учетом морфологии казахского языка) в полнотекстовых базах данных по ИТ-технологиям, Евразийский национальный университет им. Л.Н. Гумилева, г. Астана, 2015-2017.**

*Научные руководители: д.ф.-м.н. Тусупов Д.А. (Казахстан), член-корр. Федотов А.М. (НГУ), ответственный исполнитель от ИСИ СО РАН – к.ф.-м.н. Т.В. Батура*

*Сроки: 2015 – 2017*

### **Календарь зарубежных командировок по странам**

- *Касьянов В.Н. (24.09.17-01.10.17) - участие с докладом в Международной конференции APSAC-2017, г. Дубровик, Хорватия;*
- *Касьянова Е.В.(24.09.17-01.10.17) - участие с докладом в Международной конференции APSAC-2017, г. Дубровик, Хорватия.*
- *Емельянов П.Г. ( 14.06.2017 – 24.06.2017) Париж, (командировка от НГУ)*
- *Боженкова Е.Н. (15.05.2017 - 19.05.2017) – участие с докладом в 4-ом Российско-Финском симпозиуме по дискретной математике, г. Турку, Финляндия.*
- *Вирбицкайте И.Б. (01.01.2017 - 22.01.2017) – научная работа в рамках международного проекта CAVER DFG/РФФИ, Ольденбургский университет, г. Ольденбург, Германия.*
- *Вирбицкайте И.Б. (15.05.2017 - 19.05.2017) – участие с докладом в 4-ом Российско-Финском симпозиуме по дискретной математике, г. Турку, Финляндия.*
- *Коровина М.В. (13.06.2017-17.06.2017) - участие с докладом в конференции «Вычислимость в Европе», Университет г. Турку, Финляндия.*
- *Коровина М.В. (16.11.2017 - 08.12.2017) – научная работа в рамках международного проекта CAVER DFG, Университет г. Трира, Германия.*
- *Коровина М.В. (26.11.2017 - 01.12.2017) - участие с докладом в «Dagstuhl Seminar 17481 Reliable Computation and Complexity on the Reals», г. Вадерн, Германия.*
- *Коровина М.В. (29.03.2017 - 08.05.2017) научная работа в рамках международного проекта CAVER DFG, Университет г. Трира, Германия.*

## Список принятых в институте иностранцев

### **Франц Браузе (Franz Brausse)**

Дата рождения: 21.03.1987

Должность: аспирант, Университет г.Трир, Германия

Адрес: FB IV – Факультет Информатики, Университет Трира, D-54286 Трир, Германия

Дата визита: 26.06.2017 – 23.07.2017

**Norbert Mueller**, профессор, зам. директора по науке, Университет г. Трир, Германия

FB IV – Факультет Информатики, Университет г.Трир, D-54286 Трир, Германия

Дата визита: с 20/03/ 2017 по 29/03/2017

Цель визита: проведение совместных исследований в рамках договора о совместном сотрудничестве между Университетом г. Трир и ИСИ СОРАН.

### **Членство в национальных и международных научных организациях**

- Европейская ассоциация искусственного интеллекта — *к.т.н. Ю.А. Загорулько, к.ф.- м.н. О.И. Боровикова, н.с. Г.Б. Загорулько.*
- Российская ассоциация искусственного интеллекта — *к.т.н. Ю.А. Загорулько, к.ф.- м.н. О.И. Боровикова, н.с. Г.Б. Загорулько.*
- Ассоциация по вычислительной технике (АСМ) — *к.ф.-м.н. М.А. Бульонков.*
- Институт инженеров по электронике и электротехнике (IEEE) — *к.ф.-м.н. М.А. Бульонков.*
- Российская академия естественных наук — *член-корр. В.Н. Касьянов.*
- Американское математическое общество (AMS) — *проф. В.Н. Касьянов, проф. В.Л. Селиванов, к.ф.-м.н. Ф.А. Мурзин.*
- Европейская ассоциация по теоретической информатике (EATCS) — *проф. В.Н. Касьянов, к.ф.-м.н. В.А. Непомнящий.*
- Общество по индустриальной и прикладной математике (SIAM) — *проф. В.Н. Касьянов.*
- Европейская ассоциация по компьютерной логике (EACSL) — *к.ф.-м.н. В.А. Непомнящий.*

### **Членство в редколлегиях научных изданий**

- Периодическое издание ИАЭТ «Информационные технологии в гуманитарных исследованиях» — *к.т.н. Ю.А. Загорулько.*
- Серия сборников статей «Системная информатика», изд - во «Наука» — *д.ф.-м.н. В.Н. Касьянов, к.ф.-м.н. В.А. Непомнящий.*
- Журнал «Проблемы информатики» ИВМ и МГ СО РАН — *проф. А.Г. Марчук (в редакционном совете).*
- Бюллетень ИВМ и МГ, Специальный выпуск ИСИ СО РАН (BULLETIN of the Novosibirsk Computing Center, Series: Computer Sciences) — *д.ф.-м.н. В.Н. Касьянов, д.ф.-м.н. А.Г. Марчук, к.ф.-м.н. В.А. Непомнящий.*
- Журнал «Вестник НГУ, серия: Математика, механика, информатика» — *проф. А.Г. Марчук.*
- Научный электронный журнал «Системная информатика» (сайт журнала <http://www.system-informatics.ru>) -- *проф. А.Г. Марчук — главный редактор.*
- Международный журнал «Проблемы программирования», г. Киев, — *проф. В.Н. Касьянов.*
- Журнал «Программирование» — *проф. И.Б. Вирбицкайте – член редколлегии*
- Международный журнал «Enterprise Information Systems» (Taylor & Francis Group) – *проф. В.Н. Касьянов – член редколлегии*

## Научно-педагогическая деятельность и популяризация науки

### 1. Крупные мероприятия

**1.1. XIV Открытая Всесибирская олимпиада по программированию им. И.В.Поттосина** организуется и проводится совместно с Новосибирским государственным университетом с 2000 года. Эта олимпиада является одним из наиболее эффективных инструментов выявления и подготовки одаренных молодых людей, вносящих затем существенный вклад в развитие отечественных современных компьютерных технологий. Основные цели олимпиады — повышение качества подготовки специалиста в области информационных технологий, развитие знаний и умений студентов вузов по ключевым направлениям профессиональной деятельности, повышение качества набора в вузы с привлечением к участию в олимпиаде одаренных школьников.

Открытая Всесибирская олимпиада по программированию им. И.В.Поттосина является командной, в ней принимают участие студенты не только российских вузов, но и стран ближнего зарубежья (Белоруссия, Украина, Казахстан, Киргизия, Грузия, Узбекистан, Армения). Олимпиада проходит в два-три тура (<http://olimpic.nsu.ru/>). Первые один-два тура проводятся с помощью Интернет, последний, очный — на базе НГУ.

Интернет-тур проводится по традиционным правилам международного студенческого чемпионата ACM (Association for Computing Machinery). Задачи, решения, тесты, программы, проверяющие правильность решений, также как и система автоматической проверки решений, разрабатываются жюри олимпиады.

Очный тур нацелен на искусство постановки задач и выбора методов решения. Здесь оценивается умение корректно поставить задачу на основании формулировки проблемы и ее контекста; умение проанализировать множество вариантов решений и, исходя из различных критериев эффективности, выбрать самый оптимальный. В рамках очного тура проводится две номинации.

В жюри и оргкомитете олимпиады принимают участие преподаватели НГУ и ведущих вузов России: Московского, Санкт-Петербургского, Саратовского госуниверситетов, Санкт-Петербургского государственного университета информационных технологий, механики и оптики.

Полная информация, в том числе задачи, тесты, решения жюри, рейтинг команд по проведенным олимпиадам выложена на сайте <http://olimpic.nsu.ru/>.

**1.2. XXXXII Летняя школа юных программистов (ЛШЮП) имени А.П. Ершова.** Летняя школа, созданная в 1976 году академиком Андреем Петровичем Ершовым, выполняла функции обкатки методики преподавания программирования в образовательных учреждениях, дала начало информатике как учебной дисциплине в школах, сформировала круг специалистов, являющихся лидерами в мировом сообществе программистов. Летняя школа и на сегодняшний день имеет важное значение как мероприятие, направленное на развитие творческой личности, которой дается в руки мощный инструмент для применения в любой области деятельности.

Практическая работа ЛШЮП организована в форме учебно-производственных проектов – мастерских – под руководством специалистов из числа студентов НГУ, сотрудников СО РАН и IT-компаний, преподавателей НГУ. Ежегодно спектр тематик мастерских разнообразен и широко охватывает область не только классического программирования, но и прикладные задачи других наук: математику, физику, биоинформатику.

Участники ЛШЮП по традиции – школьники 5-11 классов, студенты, преподаватели из числа научных сотрудников СО РАН, НГУ и других вузов и программистских фирм, в том числе и не только из Новосибирска.

Главными задачами ЛШЮП является отбор талантливых старшеклассников, заинтересованных в овладении профессиональным программированием, обучение учеников среднего звена навыкам коллективной работы с применением современных информационных

технологий и содействие развитию способностей к практическому программированию учащихся младших классов, а также поддержка педагогов, успешно преподающих информатику и программирование в общеобразовательной системе.

## **2. Олимпиады, конкурсы юных программистов и др.**

2.1 Заочная Олимпиада по программированию на языке Лого для школьников 3-7 классов. В олимпиаде принимают участие порядка 70 человек из различных регионов России (Новосибирск, Барнаул, Кемерово, Челябинск, Чебоксары, Москва, Санкт-Петербург и др.), а также из Казахстана.

2.2 Региональная командная олимпиада по программированию на языке Лого для 3-7 классов.

2.3 Городской конкурс «Триатлон» для учащихся 5-6 классов, включающий в себя Очную, Дистанционную формы обучения и конкурсной работы в средах Лого, Муравей и Скретч (совместно с Городским центром «Эгида»

2.4 Конкурс «Триатлоша» для учащихся начальных классов.

## **3. Чтение научно-популярных лекций**

3.1. В процессе работы Летней школы юных программистов сотрудниками ИСИ были прочитаны лекции по различным темам.

3.2. В «День науки» прочитаны научно-популярные лекции для школьников. Пальянов А., Хайрулин С.

3.3. Организация методических семинаров для школьных учителей. Тихонова Т.И.

## Научно-педагогическая деятельность

### Объединенный семинар ИСИ СО РАН и НГУ «Конструирование и оптимизация программ»

*Руководитель профессор В.Н. Касьянов*

### Научное руководство студентами и аспирантами

Аспирантов — 29

Студентов 3–5 курсов — 60

### Новосибирский государственный университет

#### Основные курсы (ММФ)

- Программирование (проф. В.Н. Касьянов, С.Н. Касьянова, Е.В. Касьянова, Р.И. Идрисов)
- Теория алгоритмов (проф. В.Н. Касьянов)
- Теория вычислений (проф. В.Н. Касьянов)
- Основы работы на ЭВМ (С.Н. Касьянова)
- Программирование-2 (Е.В. Касьянова)
- Практикум на ЭВМ (С.Н. Касьянова, Е.В. Касьянова, Р.И. Идрисов)
- Базы данных и экспертные системы (доцент Ю.А. Загорулько)
- Программирование-1 (ст. преподаватель Загорулько Г.Б.)
- Программирование-2 (ст. преподаватель Загорулько Г.Б.)
- Теория программирования (семинары, полугодовой) ассистент Ахмадеева И.Р.
- Программирование (доцент Городня Л.В.)
- Формальные методы в программной инженерии (Пономарев Д.К.)
- Программирование (Тихонова Т.И.)
- Теория программирования (доцент М.А.Бульонков)
- Программирование (доцент М.А.Бульонков)
- Программирование (Емельянов П.Г.)
- Информационные системы (Мурзин Ф.А)

- Программирование (семинары)  
(Нестеренко Т.В.)
- Программирование – 2(семинары)  
(Нестеренко Т.В.)

### **Спецкурсы (ММФ)**

- Методы верификации программ  
(доцент Непомнящий В.А.)
- Введение в параллельное программирование  
(профессор Вирбицкайте И.Б.)
- Теория параллельного программирования  
(профессор Вирбицкайте И.Б.)
- Методы и системы искусственного интеллекта  
(доцент Загоруйко Ю.А.)
- Объектно-ориентированное программирование  
(ассистент Соловьев В.В.)
- Стандарты XML  
(проф. Марчук А.Г.)
- Клиент - серверные технологии  
(проф. Марчук А.Г.)
- Функциональное программирование  
(доцент Городня Л.В.)
- Парадигмы программирования  
(доцент Городня Л.В.)
- Предикатное программирование  
(Шелехов В.И.)
- Система автоматизации доказательств PVS  
(Шелехов В.И.)
- Основы методов трансляции  
(Михеев В.В.)
- Методы оптимизирующей трансляции  
(Михеев В.В.)
- Документирование программных систем  
(Андреева Т.А.)
- Визуализация графов  
(Апанович З.В.)
- Методы обработки дискретной информации  
(Мурзин Ф.А.)
- Биоинформатика  
(Черемушкин Е.С.)
- Введение в обработку изображений и вычислительную геометрию  
(Мурзин Ф.А. совместно с Куликовым А.И., ИВМ и МГ СО РАН)

### **Спецкурсы (ММФ, ФИТ)**

- Язык Perl  
(П.А. Дортман)
- Графы в программировании  
(профессор В.Н. Касьянов)
- Язык программирования Zonnon  
(Е.В. Касьянова)

### **Спецкурсы (ФИТ)**

- Верификация и анализ программ» (лекции, полугодовой, магистратура 1 курс)  
(ассистент Марьясов И.В.)
- Задачи и методы параллельного программирования  
(профессор Вирбицкайте И.Б.)
- Модели и методы искусственного интеллекта  
(доцент Загорулько Ю.А.)
- Языки и системы искусственного интеллекта  
(доцент Загорулько Ю.А.)
- Визуализация информации при помощи графов  
(З.В.Апанович)
- Парадигмы программирования  
(доцент Городня Л.В.)
- Стандартизация программной документации  
(Андреева Т.А.)
- Проектирование программных систем  
(Никитин А.Г.)
- Теоретические основы обработки информации  
(Батура Т.В.)
- Теория вычислительных процессов  
(Мурзин Ф.А.)
- Геометрические методы в компьютерной графике  
(Мурзин Ф.А. совместно с Куликовым А.И., ИВМ и МГ СО РАН)
- Продвинутое программирование на C# (лекции, полугодовой, 4-й курс)  
(ст. преп. Быстров А.В.)
- Системное программирование в .NET» (лекции, полугодовой, 4-й курс)  
(ст. преп. Быстров А.В.)

### **Основные курсы (ФИТ)**

- Задачи и методы параллельного программирования  
(профессор Вирбицкайте И.Б.)
- Инженерия знаний  
(доцент Загорулько Ю.А.)
- Программирование на языке высокого уровня  
(ст. преподаватель Петров Е.С.)
- Современные информационные технологии.  
(доцент Городня совместно с А.М. Федотовым)
- Информатика  
(доцент Городня совместно с А.М. Федотовым)
- Формальные методы в описании языков и систем программирования  
(Шелехов В.И.)
- Программирование (практика, годовой, 1-й курс)  
(ст. преп. Алексеев Г.И., ст. преподаватель Мыльников С.П.)

### **Спецкурсы (ФФ)**

- Тьюториал по программированию  
(доцент Быстров А.В.)

- Представление знаний и искусственный интеллект (доцент Загорулько Ю.А.)
- Проектирование программных систем (Никитин А.Г.)
- Машинная графика (Валеев Т.Ф.)
- Динамическая 3Д-графика (Валеев Т.Ф.)

#### **Специальные семинары (ММФ, ФИТ)**

- Теоретическое и экспериментальное программирование (Непомнящий В.А. и Шилов Н.В.)
- Интеллектуальные системы (руководитель к.т.н., с.н.с. Загорулько Ю.А.)
- Системное программирование (проф. Марчук А.Г.)
- Системное программирование (к.ф.-м.н. М.А.Бульонков, Филаткина Н.Н.)

#### **Основные курсы (ГумФак)**

- Современные информационные технологии (Тихонова Т.И. - лекции, семинары.)

#### **Лекции (ВКИ НГУ)**

- Основы программирования (Нестеренко Т.В.)
- Теория алгоритмов (Нестеренко Т.В.)

## Список наиболее важных публикаций за 2017 год

### Монографии

1. Городняя Л.В. Парадигмы программирования: анализ и сравнение. Из-во СО РАН, РФФИ 17-11-00042. 216 с.

### Центральные издания

1. Anureev I.S., Promsky A.V. Conceptual transition systems and their application to development of conceptual models of programming languages // System Informatics, 2017, № 9, 133–154.
2. Anureev I.S. Operational conceptual transition systems and their application to development of conceptual operational semantics of programming languages // System Informatics, 2017, № 9, 155–200.
3. Гаранина Н.О., Зюбин В.Е., Лях Т.В. Онтологический подход к организации шаблонов требований в рамках системы поддержки формальной верификации распределенных программных систем // Системная информатика, 2017, № 9, 111-132.
4. Maryasov I.V., Nepomniaschy V.A., Kondratyev D.A. Invariant Elimination of Definite Iterations over Arrays in C Programs Verification // Моделирование и анализ информационных систем, Т.24, №6, 2017, 743–754.
5. I.V. Maryasov, V.A. Nepomniaschy, D.A. Kondratyev Verification of Definite Iteration over Arrays with a Loop Exit in C Programs // System Informatics, 2017, №10, 57–66.
6. Селиванов В.Л., Селиванова С.В. О конструктивных числовых полях и вычислимости решений дифференциальных уравнений в частных производных // Доклады РАН, Серия «Математика», 2017, 477, № 3, 278-281 (WoS, Scopus).
7. Панкратов С.Б. Автоматическая генерация тестов для GFX-offload компилятора Intel // Проблемы информатики. – 2017. – N 2– С. 13-23.
8. Kasyanov V.N., Kasyanova E.V. Methods and Tools of Parallel Programming // CEUR Workshop Proceedings. - 2017. - Vol. 1839. - pp. 141-154.
9. Kasyanov V., Zolotuhin T. A system for structural information visualization based on attributed hierarchical graphs // WSEAS Transactions on Computers. – 2017. – Vol. 16. – pp. 193-201.
10. Kasyanov V., Kasyanova E. A visual cloud system for parallel and functional programming teaching and learning // WSEAS Transactions on Computers. – 2017. – Vol. 16. – pp. 208-215.
11. Kasyanov V.N., Zolotuhin T.A. A system for visualization of big attributed hierarchical graphs // Information Technology in Industry. – 2018. – 6 с. (в печати)
12. Гордеев Д.С. Обзор техник визуализации алгоритмов на графах // Научная визуализация. – 2018. – 33 с. (в печати)
13. Kasyanov V.N., Kasyanova E.V. A system of functional programming for supporting cloud supercomputing // WSEAS Transactions on Information Science and Applications. – 2018. – 10 с. (в печати)
14. Kasyanov V.N., Zolotuhin T.A. A system for big attributed hierarchical graph visualization // WSEAS Transactions on Computers. – 2018. – 5 с. (в печати)
15. Боровикова О.И., Загорюлько Г.Б., Загорюлько Ю.А., Шестаков В.К. Использование паттернов для разработки онтологии информационно-аналитического интернет-ресурса

- «поддержка принятия решений» // Информационные и математические технологии в науке и управлении. —2017.— № 3 (7). — С. 144–153.
16. Сидорова Е.А., Гаранина Н.О., Кононенко И.С. Подход к разрешению референциальной неоднозначности текста при пополнении онтологии // Информационные и математические технологии в науке и управлении. – 2017. – № 3(7). – С. 154–167.
17. Ахмадеева И.Р., Серый А.С., Шестаков В.К. Некоторые особенности реализации платформы для построения информационно-аналитических интернет-ресурсов // Информационные и математические технологии в науке и управлении – 2017. № 3 (7). – С. 168–175.
18. Боровикова О.И., Кононенко И.С., Сидорова Е.А. Подход к извлечению информации из протоколов клинических испытаний на основе медицинской онтологии // Системная информатика. – 2017. – № 9. – С. 93–110.
19. Тумуров Э.Г., Шелехов В.И. Технология автоматного программирования на примере программы управления лифтом // «Программная инженерия», Том 8, № 3, 2017. – С.99-111. <http://persons.iis.nsk.su/files/persons/pages/lift1.pdf>
20. Шелехов В.И. Предикатная программа вставки в AVL-дерево // Системная информатика, № 9. — Новосибирск, 2017. — С. 23-42. [http://persons.iis.nsk.su/files/persons/pages/avl\\_insert.pdf](http://persons.iis.nsk.su/files/persons/pages/avl_insert.pdf)
21. Булгаков К.В., Каблуков И.В., Тумуров Э.Г., Шелехов В.И. Оптимизирующие трансформации списков и деревьев в системе предикатного программирования // Системная информатика, № 9. — Новосибирск, 2017. — С. 63-92. <http://www.system-informatics.ru/files/article/105.pdf>
22. Зубарев А.Ю. Анализ типов в трансляторе с языка предикатного программирования // Системная информатика, № 9. — Новосибирск, 2017. — С. 1-22. <http://persons.iis.nsk.su/files/persons/pages/zubarev.pdf>
23. Каблуков И.В., Шелехов В.И. Реализация оптимизирующих трансформаций в системе предикатного программирования // Системная информатика. — Новосибирск, 2018. — 18с. <http://persons.iis.nsk.su/files/persons/pages/opttransform4.pdf> На рецензии
24. Tarasyuk I.V., Buchholz P. Equivalences for fluid stochastic Petri nets. Siberian Electronic Mathematical Reports 14, p. 317-366, S.L. Sobolev Institute of Mathematics, Novosibirsk, April 2017, <http://semr.math.nsc.ru/v14/p317-366.pdf> (Web of Science, Scopus, Zentralblatt Math indexed).
25. Korovina M., Kudinov O. On images of partial computable functions over computable Polish spaces. Siberian Electronic Mathematical Reports, 2017, v. 14, pp. 418-432. (Web of Science, Scopus)
26. Korovina M., Kudinov O. Spectrum of the field of computable real numbers. Algebra and Logic, 2017, Vol. 55, pp. 485-500 (Web of Science, Scopus).
27. Апанович З.В. Кросс-языковая идентификация авторов публикаций //Открытые системы. СУБД. 2017. № 1. С. 43-45. ВАК
28. Трофимов В.К., Храмова Т.В. Оптимальное равномерное по выходу кодирование для объединения различных множеств источников // Автометрия. 2017. Т.53 N1. – С 53-62. [Scopus, ВАК]
29. Trofimov V.K., Khranova T.V. Optimal output-uniform coding for a union of different sets of sources // Optoelectronics, Instrumentation and Data Processing.2017.Т.53. N1. – P. 43-50. [Scopus, ВАК] (Перевод статьи под номером 1)
30. Gureenkova O. A., Batura T. V., Kozlova A. A., Svischev A. N. Complex approach towards algorithm learning for anaphora resolution in Russian language // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue-2017”. V. 1. Iss. 16. 2017. pp. 89–97. ISSN 2221-7932 [Scopus, ВАК]
31. Barakhnin V.B., Fedotov A.M., Bakiyeva A.M., Bakiyev M.N., Tazhibayeva S.Zh., Batura T.V., Kozhemyakina O.Yu., Tussupov D.A., Sambetbaiyeva M.A., Lukpanova L.Kh. The software system for the study the morphology of the Kazakh language // The European Proceedings of

- Social & Behavioural Sciences. 2017. V. XXXIII. P. 18–27. URL: <http://www.futureacademy.org.uk/publication/EpSBS/ICPE2017InternationalConferenceonPsychologyandEducation> [Web of Science]
32. Бакиева А.М., Батура Т.В. Исследование применимости теории риторических структур для автоматической обработки научно-технических текстов // Cloud of Science. 2017. Т. 4. № 3. С. 450–464. [BAK]
  33. Батура Т.В. Методы автоматической классификации текстов // Программные продукты и системы. Тверь: ЗАО НИИ ЦПС, 2017. Т. 30. № 1. С. 85-99. DOI: 10.15827/0236-235X.030.1.085-099. [BAK]
  34. Барахнин В. Б., Федотов А. М., Бакиева А. М., Бакиев М. Н., Тажибаева С. Ж., Батура Т. В., Кожемякина О. Ю., Тусупов Д. А., Самбетбаева М. А., Лукпанова Л. Х. Алгоритмы генерации и стемматизации словоформ казахского языка // Cloud of Science. 2017. Т. 4. № 3. С. 434–449. [BAK]
  35. Соколов А.В., Батура Т.В. Разработка системы разрешения анафоры на основе методов машинного обучения // Программные продукты и системы. Тверь: ЗАО НИИ ЦПС, 2017. Т. 30. № 3. С. 461–468. DOI: 10.15827/0236-235X.030.3.461-468. [BAK] URL: <http://www.swsys.ru/files/2017-3/461-468.pdf>
  36. Барахнин В.Б., Бакиева А.М., Бакиев М.Н., Тажибаева С.Ж., Батура Т.В., Лукпанова Л.Х. Стемматизация и генерация словоформ казахского языка для систем автоматической обработки текстов // Вычислительные технологии. Новосибирск: ИВТ. 2017. Т. 22, № 4. С. 11–21. [BAK, zbMath] URL: <http://www.ict.nsc.ru/jct/annotation/1811>
  37. Мурзин Ф.А., Сагнаева С.К., Еримбетова А.С., Дайырбаева Э.Н. Разработка системы связей для тюркских языков // Вестник Казахской академии транспорта и коммуникаций им. М. Тынышпаева. 2017. № 3 (102). – С. 102-107.
  38. V.Sherstyuk, S.Medvedev, E.Elisaphenko, E.Vaskova, M.Ri, Y.Vyatkin, O.Saik, D.Shtokalo, E.Pokushalov, S.Zakian. Genome-wide profiling and differential expression of microRNA in rat pluripotent stem cells // Scientific reports. 2017 Jun 5;7(1):2787. doi: 10.1038/s41598-017-02632-0. (Scopus, Web of Science)
  39. St. Laurent G III, Seilheimer B, Tackett M, Zhou J, Shtokalo D, Vyatkin Y, Ri M, Toma I, Jones D and McCaffrey TA. Deep Sequencing Transcriptome Analysis of Murine Wound Healing: Effects of a Multicomponent, Multitarget Natural Product Therapy-Tr14. // Front. Mol. Biosci. 4:57. 2017. doi: 10.3389/fmolb.2017.00057.
  40. David Lung, Stephen Larson, Andrey Palyanov, Sergey Khayrulin, Pdraig Gleeson, Manuel Zimmer, Radu Grosu and Ramin M. Hasani. A Simplified Cell Network for the Simulation of C. elegans' Forward Crawling // Proc. NIPS 2017 Workshop on Worm's Neural Information Processing
  41. Palyanov A.Y., Chekmarev S.F. Hydrodynamic description of protein folding: the decrease of the probability fluxes as an indicator of transition states in two-state folders. *Journal of Biomolecular Structure and Dynamics*. 2017. Т. 35. № 14. С. 3152-3160.
  42. Иртегов Д.В., Нестеренко Т.В., Чурина Т.Г. Разработка систем автоматизированной оценки знаний по программированию. Системная информатика. 2017. № 11. С. 91-116.
  43. Карасюк П.К., Мигинский Д.С. Протокол для децентрализованной системы хранения с избыточным кодированием. Программные продукты и системы. 2017. № 4. С. 699-705.
  44. Yevshin I.S., Sharipov R.N., Valeev T.F., Kel A.E., Kolpakov F.A. Trd: a database of transcription factor binding sites identified by chip-seq experiments В книге: Высокопроизводительное секвенирование в геномике II Всероссийская конференция с международным участием. Сер. "Acta Naturau" 2017. С. 25.
  45. Yevshin I., Sharipov R., Valeev T., Kel A., Kolpakov F.G. Trd: a database of transcription factor binding sites identified by chip-seq experiments *Nucleic Acids Research*. 2017. Т. 45. № D1. С. D61-D67.
  46. Городня Л.В., Крайнева И.А., Марчук А.Г. Школа программирования Института кибернетики АН УССР (1962-1990) // История науки и техники, 2017. №1. С. 42–64. ВАК.

47. Городня Л.В., Крайнева И.А., Марчук А.Г. Библиография и источники по истории школ программирования в Академии наук СССР и ведущих советских вузах (1950–1917)// История науки и техники, 2017. Т. 38. №3. С. 482–529. ВАК.
48. Городня Л.В. Движущие силы информатики. Дистанционное и виртуальное обучение
49. № 2 (116) 2017, с. 75-90. Издательство Современного гуманитарного университета (Москва), ISSN: 1561-2449 . ВАК.  
Сапрыкина Г.А., Тихонова Т.И. Конференции школьников как способ выявления специальных одаренностей // Сибирский учитель. 2017. № 4 (113). С. 55-58. ВАК.
50. Куперштох Н.А., Крайнева И.А. История новосибирского Института радиофизики и электроники (1957–1967)// Гуманитарные науки в Сибири. 2017. Т.24. №2. С.109–113. ВАК.
51. Крайнева И.А. Генезис дисциплины в поле науки: вычислительное дело – программирование – информатика// Вестник ТГУ. 2017. N 421. С. 118–128. Web of Science

### **Зарубежные издания**

1. N. O. Garanina and E. A. Sidorova. An Approach to Verification of a Family of Multiagent Systems for Conflict Resolution // Automatic Control and Computer Sciences, Allerton Press, 2017, V. 51, № 7, 498–506. (WoS, Scopus)
2. N. V. Vizovitin, V. A. Nepomniaschy, A. A. Stenenko Application of Colored Petri Nets for Verification of Scenario Control Structures in UCM Notation // Automatic Control and Computer Sciences, Allerton Press, 2017, V. 51, № 7, 489–497 (WoS, Scopus).
3. Selivanov V.L., Kudinov O. First order theories of some lattices of open sets // Logical Methods in Computer Science, 13, № 3, 2017, 1-18 (Scopus).
4. Selivanov V.L. Towards a descriptive theory of cb0-spaces // Mathematical structures in computer science, 2017, V. 27, № 8, 1553-1580 (WoS, Scopus).
5. Selivanov V.L., Selivanova S.V. Computing solution operators of boundary-value problems for symmetric hyperbolic systems of PDEs // Logical Methods in Computer Science, V. 13, № 4, 2017, 1–31 (Scopus).
6. Natalia Garanina, Elena Sidorova, Irina Kononenko, Sergei Gorlatch. Using multiple semantic measures for coreference resolution in ontology population // International Journal of Computing. –2017. – Vol. 16, Iss. 3. – P.166–176. (CSP- Scopus Elsevier)  
<http://www.computingonline.net/computing/article/view/900>
7. Yuliya Rubtsova. Reducing the Degradation of Sentiment Analysis for Text Collections Spread over a Period of Time // Knowledge Engineering and Semantic Web / Przemysław Różewski, Christoph Lange (Eds.) / 8th International Conference, KESW 2017 Szczecin, Poland, November 8–10, 2017, Proceedings. Communications in Computer and Information Science book series. – Cham: Springer, 2017. – Vol. 786. – P. 3–18.
8. Korovina M., Kudinov O. Computable elements and functions in effectively enumerable topological spaces // Mathematical Structures in Computer Science, Vol. 27, Issue 8, pp. 1466-1494, 2017. (Web of Science, Scopus).
9. Ishihara H., Korovina M., Pauly A., Seisenberger M., Spreen D. (2017) Preface to the special issue: Continuity, computability, constructivity: from logic to algorithms 2013 // Mathematical Structures in Computer Science Vol. 27, Issue 8, pp. 1285-1286. (Web of Science, Scopus ).
10. Korovina M., Kudinov O. On higher effective descriptive set theory // Lecture Notes in Computer Science, 2017, Vol. 10307, pp. 282-291. (Scopus).

11. Korovina M., Kudinov O. Outline of partial computability in computable topology // Lecture Notes in Computer Science, 2017, Vol. 10307, pp. 64-76. (Scopus).
12. Korovina M., Kudinov O. Highlights of Rice-Shapiro theorem in computable topology // Logical Methods in Computer Science, 2017 (В печати) (Scopus).
13. Bause F., Buchholz P., Tarasyuk I.V., Telek M. Equivalence and lumpability of FSPNs // Lecture Notes in Computer Science Vol. 10378, pp. 16-31, 2017. (Scopus).
14. Best E., Gribovskaya N., Virbitskaite I. Configuration- and Residual-Based Transition Systems for Event Structures with Asymmetric Conflict // Lecture Notes in Computer Science, 2017, Vol. 10139, pp. 132-146. (Scopus).
15. Demin A.V. Vityaev E.E. Adaptive Control of Modular Robots // Biologically Inspired Cognitive Architectures (BICA) for Young Scientists. – Springer, 2018. – pp. 204-212. (Scopus).
16. D. Ponomaryov and M. Soutchanski Progression of decomposed local-effect action theories. ACM Transactions on Computational Logic (TOCL) 18(2), 2017. (Scopus, WoS).
17. <http://dl.acm.org/citation.cfm?id=3091119>. Журнал входит в 1-ый квартиль журналов по специализации «информатика» по версии Scopus.
18. Gregor Behnke, Florian Nielsen, Marvin Schiller, Denis Ponomaryov, Pascal Bercher, Birte Glimm, Wolfgang Minker, and Susanne Biundo. To Plan for the User Is to Plan with the User: Integrating User Interaction into the Planning Process. Companion Technology. A Paradigm Shift in Human-Technology Interaction. Susanne Biundo and Andreas Wendemuth, eds. Springer International Publishing, 2017. ISBN: 978-3-319-43664-7 (Scopus).  
<https://www.springerprofessional.de/to-plan-for-the-user-is-to-plan-with-the-user-integratinguser-i/15277618> . Глава в зарубежной монографии.
19. Platonov Y., Artamonova E., "Creation of Information Portraits Using 'Polar' System "// Information Technology in Industry» 2018. Issues: 1-4 (в печати, индексируется в Web of Science).

#### **Материалы международных конференций**

1. Natalia Garanina. An Ontological Approach to a System of Requirement Patterns // Тезисы VIII Международного научно-исследовательского семинара «Программные семантики, спецификации и верификация» (PSSV-2017). Москва: МАКС пресс, 2017, 13-23.
2. Garanina N., Sidorova E., Kononenko I., Gorlatch S. Using Multiple Semantic Measures in a Framework for Coreference Resolution in the Process of Ontology Population // Proc. of the 26th International Workshop on CS&P. Warsaw, Poland, 2017 (Scopus).
3. Selivanov V.L. Extending Wadge theory to k-partitions // Proc. Int. conf. CiE 2017, LNCS 10307, 387-399, 2017 (Scopus).
4. Касьянов В.Н. Теоретико-графовые методы и системы программирования // Языки программирования-2017. Труды конференции. - Ростов-на-Дону: Изд-во Южного федерального университета, 2017. - С.129-133.
5. Касьянов В.Н., Касьянова Е.В. Методы и средства дистанционного обучения программированию // Материалы XVII Международной конференции "Информатика: проблемы, методология, технология". – Воронеж: Вэлборн, 2017. – Том 5. –С. 343-348.
6. Касьянов В.Н., Панкратов С.Б. Автоматическая генерация тестов для оптимизирующего компилятора // Материалы XVII Международной конференции "Информатика: проблемы, методология, технология". – Воронеж: Вэлборн, 2017. – Том 1. –С. 65-69.
7. Касьянова Е.В., Касьянова С.Н. Использование кластеров при решении задач томографии // Материалы XVII Международной конференции "Информатика: проблемы, методология, технология". – Воронеж: Вэлборн, 2017. – Том 4. –С. 195-200.

8. Касьянова Е.В. Методы и средства дистанционного обучения программированию // Языки программирования-2017. Труды конференции. - Ростов-на-Дону: Изд-во Южного федерального университета, 2017. - С.134-136.
9. Касьянов В. Облачные средства поддержки параллельного программирования // Международная конференция «Актуальные проблемы чистой и прикладной математики, посвященная 100-летию со дня рождения академика Тайманова Асана Дабсовича. - Алматы: ИМММ, 2017. - С. 133-135.
10. Касьянов В.Н. Российская информатика в лицах: академик Андрей Петрович Ершов // Труды Всероссийской научно-практической конференции (с международным участием) «Гуманитарное образование и наука в техническом вузе». –Ижевск: Изд-во ИЖГТУ, 2017. – С. 124-128. - ISBN 978-5-7526-0781-3
11. Yury A. Zagorulko, Olesya I. Borovikova, Galina B. Zagorulko. Methodology for the development of ontologies for thematic intelligent scientific Internet resources // Proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russky Island, Russia, 25-29 September, 2017. – IEEE Xplore digital library. – P. 194–198. – URL: <http://ieeexplore.ieee.org/document/8168097> (дата обращения: 10.12.2017) ISBN: 978-1-5386-1206-4. DOI: 10.1109/RPC.2017.8168097
12. Ludmila Braginskaya, Valery Kovalevsky, Andrey Grigoryuk, Galina Zagorulko. Ontological approach to information support of investigations in active seismology // Proceedings of the 2nd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, Russky Island, Russia, 25-29 September, 2017. – IEEE Xplore digital library. – P. 27–29. – URL: <http://ieeexplore.ieee.org/document/8168060> (дата обращения: 10.12.2017) ISBN: 978-1-5386-1206-4. DOI: 10.1109/RPC.2017.8168060
13. Zagorulko, Yu.A., Borovikova, O.I., Zagorulko, G.B. Application of ontology design patterns in the development of the ontologies of scientific subject domains. 2017. Selected Papers of the XIX International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2017). – CEUR Workshop Proceedings. – Vol. 2022. – P. 258-265. – URL: <http://ceur-ws.org/Vol-2022> (дата обращения: 10.12.2017)
14. Sidorova, E.A., Kononenko, I.S., Zagorulko, Yu.A. An approach to filtering prohibited content on the web. 2017. Selected Papers of the XIX International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2017). – CEUR Workshop Proceedings. – Vol. 2022. – P. 64-71. – URL: <http://ceur-ws.org/Vol-2022> (дата обращения: 10.12.2017)
15. Загорулько Ю.А., Загорулько Г.Б., Шестаков В.К. Подход к разработке информационно-аналитического интернет-ресурса по поддержке принятия решений // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Междунар. науч. конф., Банное, Россия, 1-5 марта 2017 г. (ИТиС - 2017) : науч. электрон. изд. (1 файл 28,0 Мб) / отв. ред. Ю. С. Попков, А. В. Мельников. – Челябинск : Изд-во Челяб. гос. ун-та, 2017. – С. 113– 116.
16. Загорулько Г.Б., Боровикова О.И. Проблемы разработки онтологии для информационно-аналитического интернет-ресурса по поддержке принятия решений // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Междунар. науч. конф., Банное, Россия, 1-5 марта 2017 г. (ИТиС - 2017) : науч. электрон. изд. (1 файл 28,0 Мб) / отв. ред. Ю. С. Попков, А. В. Мельников. – Челябинск: Изд-во Челяб. гос. ун-та, 2017. – С. 108– 112.
17. Сидорова Е. А., Боровикова О. И. Подход к жанровой классификации текстовых ресурсов // Информационные технологии и системы [Электронный ресурс]: тр. Шестой Междунар. науч. конф., Банное, Россия, 1-5 марта 2017 г. (ИТиС - 2017) : науч. электрон. изд. (1 файл 28,0 Мб) / отв. ред. Ю. С. Попков, А. В. Мельников. – Челябинск: Изд-во Челяб. гос. ун-та, 2017. – С. 264–269.
18. Шелехов В.И. Синтез операторов предикатной программы // Труды конф. «Языки программирования и компиляторы '2017», Ростов-на-Дону — 2017 — С.258-262. <http://persons.iis.nsk.su/files/persons/pages/sintr.pdf>
19. Чушкин М.С. Завершение эскизов предикатных программ методом синтеза через

- контрпримеры // Конф. «Компьютерная безопасность и криптография» SIBECRYPT'17 / Прикладная дискретная математика. Приложение. — 2017, №.10, С.151-153. <http://www.mathnet.ru/links/3bc251f8a0f2d95801e4e7efc9501433/pdma319.pdf>
20. Шелехов В.И. Технология модельно-ориентированного автоматного программирования // Научный сервис в сети Интернет 2017. Новороссийск. — 2017. — С.468-476. <http://persons.iis.nsk.su/files/persons/pages/mdea.pdf>
  21. Зубарев А.Ю. Анализ типов в языке предикатного программирования // Знания-Онтологии-Теории (ЗОНТ-2017), Том 1. — Институт Математики СО РАН, Новосибирск, 2017. — С. 149-157. [http://persons.iis.nsk.su/files/persons/pages/zont\\_11.pdf](http://persons.iis.nsk.su/files/persons/pages/zont_11.pdf)
  22. Шелехов В.И. Доказательное построение, верификация и синтез предикатных программ // Знания-Онтологии-Теории (ЗОНТ-2017), Том 2. — Институт Математики СО РАН, Новосибирск, 2018. — С. 156-165. <http://persons.iis.nsk.su/files/persons/pages/lbase.pdf>
  23. Шелехов В.И. Синтез операторов предикатной программы // Труды конф. «Языки программирования и компиляторы '2017», Ростов-на-Дону — 2017 — С.258-262. <http://persons.iis.nsk.su/files/persons/pages/sintr.pdf>
  24. Чушкин М.С. Завершение эскизов предикатных программ методом синтеза через контрпримеры // Конф. «Компьютерная безопасность и криптография» SIBECRYPT'17 / Прикладная дискретная математика. Приложение. — 2017, №.10, С.151-153. <http://www.mathnet.ru/links/3bc251f8a0f2d95801e4e7efc9501433/pdma319.pdf>
  25. Шелехов В.И. Технология модельно-ориентированного автоматного программирования // Научный сервис в сети Интернет 2017. Новороссийск. — 2017. — С.468-476. <http://persons.iis.nsk.su/files/persons/pages/mdea.pdf>
  26. Зубарев А.Ю. Анализ типов в языке предикатного программирования // Знания-Онтологии-Теории (ЗОНТ-2017), Том 1. — Институт Математики СО РАН, Новосибирск, 2017. — С. 149-157. [http://persons.iis.nsk.su/files/persons/pages/zont\\_11.pdf](http://persons.iis.nsk.su/files/persons/pages/zont_11.pdf)
  27. Шелехов В.И. Доказательное построение, верификация и синтез предикатных программ // Знания-Онтологии-Теории (ЗОНТ-2017), Том 2. — Институт Математики СО РАН, Новосибирск, 2018. — С. 156-165. <http://persons.iis.nsk.su/files/persons/pages/lbase.pdf>
  28. Bozhenkova, E., Virbitskaite, I.: Some Space-Time Characteristics of Relational Structures // Proceedings of Fourth Russian-Finnish Symposium (RuFiDim), May 2017, TUCS Lecture Notes – 2017 – No. 26 – pp. 27-38. (ISSN 1797-8831 (digital), ISSN 1797-8823 (print))
  29. Erofeev, E., Wimmel, H.: Reachability Graphs of Two-Transition Petri Nets // Proceedings of the International Workshop on Algorithms & Theories for the Analysis of Event Data, 2017, CEUR Workshop Proceedings, Vol. 1847, P. 39-54.
  30. Brauße F., Korovina M., Müller N. Th.: ERA: Applications, Analysis and Improvements // Proceedings of International Conference “Continuity, Computability, Constructivity - From Logic to Algorithms” (CCC 2017), France, 2017, pp. 27-30.
  31. Emelyanov P. and Ponomaryov D. Cartesian Decomposition in Data Analysis // *Proceedings of the Siberian Symposium on Data Science and Engineering (SSDSE 2017)*, 12-13 April 2017, Novosibirsk, Russia, pp. 55-60, 2017, DOI: 10.1109/SSDSE.2017.8071964.
  32. В.Ю.Малов, М.А.Бульонков, Н.Н.Филаткина. Монгольский транзит: оценка конкурентоспособности транспортных маршрутов. // III Международный форум ассоциации экспертных центров Китая, Монголии и России "Экономический коридор Китай - Монголия - Россия: дорожная карта". - 19-20 сентября, Улан-Удэ, 2017.
  33. Y. Kazakov and D. Ponomaryov. On the complexity of semantic integration of OWL ontologies. Proc. DL '2017, The Description Logic Workshop, Montpellier, France, July 2017. <http://ceur-ws.org/Vol-1879/paper59.pdf> (Scopus).
  34. Крайнева И.А., Городняя Л.В., Марчук А.Г. О работах по системному математическому обеспечению в странах Советской Балтии (1960–1990)/ В сборнике: Труды SORUCOM-2017. Материалы Четвертой Международной конференции. 2017. С. 135–144.

35. Крайнева И.А., Городняя Л.В. из истории программирования в Беларуси (1959–1990)/ В сборнике: Труды SORUCOM-2017. Материалы Четвертой Международной конференции. 2017. С. 145–155.
36. Городняя Л.В., Кирпотина И.А. О проблеме достоверности доступной в Интернете исторической фактографии // ТРУДЫ SORUCOM — 2017- с. 46-55.
37. Андреева Т.А. Автоматизация создания задачного комплекта // Электронные ресурсы в непрерывном образовании: труды VI Международного научно-методического симпозиума «ЭРНО-2017» (Адлер). – Ростов-на-Дону: Издательство Южного федерального университета, 2017. – 230 с. ISBN 978-5-9275-2080-0. – с. 145-148.
38. Y. Platonov ; E. Artamonova. Материалы 2017 Siberian Symposium on Data Science and Engineering (SSDSE), Managing big data using specialized system polar . - 2017. <http://ieeexplore.ieee.org/document/8071953/> (дата обращения 28.12.2017), г.Новосибирск, 2017. (Индексируется в SCOPUS).
39. Тихонова Т.И. Исполнители Звенигородского: эпоха от «Агатов» 1980-х до современности //ТРУДЫ SORUCOM-2017. Четвертая Международная конференция Развитие вычислительной техники в России и странах бывшего СССР: история и перспективы. 3–5 октября 2017 года. – Москва, Зеленоград. – с. 358-361.
40. Крайнева И.А., Пивоваров Н.Ю., Шилов В.В. Советский атомный проект и становление отечественной вычислительной техники/ В сборнике: Труды SORUCOM-2017. Материалы Четвертой Международной конференции. 2017. С. 127–134.
41. Andreyeva T.A. Automated preparation of problem complexes // Наука сегодня: теоретические и практические аспекты: материалы международной научно-практической конференции, г. Вологда, 27 декабря 2017 г.: в 2 частях. Часть 1. - Вологда: ООО <Маркер>, 2018. - 200 с. ISBN 978-5-906850-89-8 (Часть 1) – с. 25-26
42. Kozlova A., Gureenkova O., Svishev A., Batura T. A hybrid approach for anaphora resolution in the Russian language // IEEE, Proceedings of 2017 Siberian Symposium on Data Science and Engineering (SSDSE). 2017. Novosibirsk, 12–13 April, 2017. P. 36–40. ISBN 978-1-5386-1592-8 [IEEE Xplore] URL: <http://ieeexplore.ieee.org/document/8071960/>
43. Barakhnin V., Bakiyeva A., Batura T. Stemming and word forms generation in automatic text processing systems in the Kazakh language // The 15th International Scientific Conference Information Technologies and Management (IT&M 2017), April 27–28, 2017, ISMA University, Riga, Latvia. pp. 85–86. ISSN 1691-2489. URL: [http://geoml.info/wp-content/uploads/2016/02/ITM2017\\_Theses.pdf](http://geoml.info/wp-content/uploads/2016/02/ITM2017_Theses.pdf)
44. Мурзин Ф.А., Еримбетова А.С., Сагнаева С.К., Батура Т.В., Семич Д.Ф., Бакиева А.М. Алгоритмы и программные инструменты для определения релевантности текста поисковому запросу и определения тем текстов. // Международная конференция «Актуальные проблемы чистой и прикладной математики, посвященная 100-летию со дня рождения академика Тайманова Асана Дабсовича. Алматы 22-25 августа 2017 года. Тезисы докладов. С. 141-142.
45. Абдикеримова М., Бычков А., Синьют В., Мурзин Ф., Русских Н., Рябчикова Е., Хайруллин С. Алгоритмы и программный инструментарий для анализа текстур на микрофотографиях. // Международная конференция «Актуальные проблемы чистой и прикладной математики, посвященная 100-летию со дня рождения академика Тайманова Асана Дабсовича. Алматы 22-25 августа 2017 года. Тезисы докладов. С.124-126.
46. Левичев А.В. О сингулярных  $SU(n,n)$ -действиях в группах  $U(p,q)$ :  $n=p+q=2$ // Междунар. конф. Математика в современном мире: Тез. докл. Нск, 2017. 125.
47. Левичев А.В. --- О построении операторов переплетения в многоуровневой модели кварк-глюонной среды: свойства орнамента первого уровня// Междунар. конф. Математика в современном мире: Тез. докл. Нск, 2017. 578.
48. Левичев А.В. On  $U(2)$ -, and  $U(1,1)$ -parallelizations of chronometric bundles// Междунар. конф. Дни геометрии в Новосибирске — 2017: Тез. докл. Нск, 2017. 14.

## Материалы российских конференций

1. Кондратьев Д.А. Расширение системы C-light символическим методом верификации финитных итераций. – Труды XVII Всероссийской конференции молодых ученых по математическому моделированию и информационным технологиям // Вычислительные технологии, 2017, Т. 22, 44-59.
2. Сидорова Е.А., Гаранина Н.О., Кононенко И.С. Подход к разрешению референциальной неоднозначности текста при пополнении онтологии // Труды XXII Байкальской Всероссийской конференции «Информационные и математические технологии в науке и управлении». – 2017, № 3, 154-167.
3. Гаранина Н.О., Серый А.С., Сидорова Е.А. Мультиагентный подход к разрешению кореференции при извлечении информации из текстов на основе онтологии. // Материалы Всероссийской конференции с международным участием «Знания-Онтологии-Теории» (ЗОНТ-2017). Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017, Т.1, 88-98.
4. Черненко С.А., Непомнящий В.А. Верификация моделей распределенных систем, представленных временными MSC-диаграммами // Материалы 10-й Всероссийской мультиконференции по проблемам управления (МКПУ-2017). Ростов-на-Дону, Южный федеральный университет, Т. 3, 2017, 119-122.
5. Кондратьев Д. А. Верификация программ инженерной математики в системе C-light // Материалы XVIII Всероссийской конференции молодых ученых по математическому моделированию, ИВТ СО РАН, 2017, с. 78 .
6. Серый А.С., Сидорова Е.А., Гаранина Н.О. Многодокументный анализ референции в процессе извлечения информации на основе онтологии предметной области // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.167–173. – ISBN: 978-5-905569-10-4.
7. Вишнев К.Е., Загорюлько Г.Б., Молородов Ю.И. Средства управления распределёнными табличными данными в информационно-аналитических интернет-ресурсах, основанных на онтологиях // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.135–141. – ISBN: 978-5-905569-10-4.
8. Ахмадеева И.Р. Методы поиска научных ресурсов в сети интернет на основе онтологии // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.174–179. – ISBN: 978-5-905569-10-4.
9. Загорюлько Ю.А., Боровикова О.И., Загорюлько Г.Б. Особенности разработки информационно-аналитического интернет-ресурса по поддержке принятия решений // XVI Российская конференция «Распределенные информационно - вычислительные ресурсы.

- Наука – цифровой экономике» (DICR-2017): Труды XVI Всероссийской конференции (4-7 декабря 2017 г.). Новосибирск / Под ред. О.Л. Жижимова, А.М. Федотова. – Новосибирск: ИВТ СО РАН, 2017. – С.157–166. –ISBN: 978-5-905569-10-4.
10. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. Использование паттернов онтологического проектирования для разработки онтологий предметных областей // Материалы Всероссийской конф. с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2017). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017. – Т.1. – С. 139–148.
  11. Брагинская Л.П., Григорюк А.П., Ковалевский В.В., Загорулько Г.Б. Информационная поддержка вибросейсмических исследований // Материалы Всероссийской конференции с международным участием «Знания – Онтологии – Теории» (ЗОНТ-2017). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017. – Т1. – С. 72–79.
  12. Гаранина Н.О., Серый А.С., Сидорова Е.А. Мультиагентный подход к разрешению кореференции при извлечении информации из текстов на основе онтологии. // Материалы Всероссийской конференции с международным участием «Знания-Онтологии-Теории» (ЗОНТ-2017). – Новосибирск: Институт математики им. С.Л. Соболева СО РАН, 2017. – Т1. – С. 88–98.
  13. P. Emelyanov and D. Ponomaryov. Cartesian decomposition in data analysis. Proc. SSDSE, IEEE Symposium on Data Science and Engineering, Novosibirsk, Russia, April 2017. <http://ieeexplore.ieee.org/document/8071964/>.
  14. Городняя Л.В. Резервы синтаксически ориентированного конструирования систем программирования. //Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). -- М.: ИПМ им. М.В.Келдыша, 2017. -- 480 с., ISBN 978-5-98354-037-8 doi:10.20948/abrau-2017 URL: <http://keldysh.ru/abrau/2017/proc.pdf>.
  15. Городняя Л.В. Учебный язык параллельного программирования СИНХРО – с. 92-97. Языки программирования и компиляторы — 2017 труды конференции / Южный федеральный университет; под ред. Д. В. Дуброва. — Ростов-на-Дону: Издательство Южного федерального университета, 2017. — 282 с. ISBN 978-5-9275-2349-8 <http://plc.sfedu.ru/files/PLC-2017-proceedings.pdf>.
  16. М. М. Лаврентьев, Т. С. Васючкова, Л. В. Городняя, В. С. Бартош, И. В. Белаго, М. А. Держо, Н. А. Иванчева, О. А. Федотова. Использование Интернет для профориентации школьников в области IT на Факультете информационных технологий НГУ. Всероссийская объединённая конференция «Интернет и современное общество». <http://ojs.ifmo.ru/index.php/IMS/article/view/515>.
  17. Андреева Т. А. Система подготовки тестовых наборов для автоматической проверки решений // Материалы IX Всероссийской (с международным участием) научно-практической конференции «Информационные технологии в образовании» «ИТО-Саратов-2017». – эл. публ. <http://saratov.ito.edu.ru/2017/section/235/99713/>.
  18. Крайнева И.А. Междисциплинарный дискурс в научном наследии А.А. Ляпунова// Материалы III Всероссийской научно-практической конференции «Современные тенденции в развитии музеев и музееведения» (Новосибирск, 9–12 октября 2017 г.). Новосибирск: Институт истории Сибирского отделения Российской академии наук. 2017. С. 136–141.
  19. Бакиева А.М., Батура Т.В. Методы автоматического реферирования и определения тем текстов // Материалы XVIII Всероссийской конференции молодых ученых по математическому моделированию и информационным технологиям. г. Иркутск, Россия, 21-25 августа 2017. Новосибирск: ИВТ СО РАН, 2017. С. 65. URL: [http://conf.nsc.ru/files/conferences/ym2017/419536/YM2017\\_Abstracts.pdf](http://conf.nsc.ru/files/conferences/ym2017/419536/YM2017_Abstracts.pdf)
  20. Бакиева А.М., Батура Т.В. Система автоматического реферирования и определения тем научно-технических текстов // XVI Российская конференция «Распределенные информационно-вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017):

- Труды XVI Всероссийской конференции (4–7 декабря 2017 г.). 2017. Новосибирск: ИВТ СО РАН. С. 75–80. ISBN 978-5-905569-10-4.
21. Бакиева А.М., Батура Т.В. Применение теории риторических структур для автоматической обработки текстов // Марчуковские научные чтения - 2017 (MSR 2017). Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. С. 149.
  22. Семенова Н. А., Батура Т. В. Автоматическое составление словаря оценочной лексики для конкретной предметной области // Марчуковские научные чтения - 2017 (MSR 2017). Тезисы. Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. С. 159.
  23. Batura T.V., Murzin F.A., Semich D.F., Bakiyeva A.M., Yerimbetova A.S. Link Grammar Parser and estimation of the document relevance to the search query // Марчуковские научные чтения - 2017 (MSR 2017). Тезисы. Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. – С. 200.
  24. G.B. Abdikerimiva, A.L. Bychkov, Wei Xin Yu, F.A.Murzin, N.E.Russkikh, E.I.Ryabchikova, S.S.Khairulin. Algorithms and software for analysis of textural features and the space spectral properties of images. // Марчуковские научные чтения - 2017 (MSR 2017). Тезисы. Новосибирск: Омега Принт, 2017. Новосибирск, 25 июня–14 июля 2017 г. – С. 200.
  25. Мурзин Ф.А., Еримбетова А.С., Батура Т.В., Семич Д.Ф., Ефимова Л.В., Бакиева А.М. О новых инструментах поиска информации на основе грамматики связей // В сборнике: Интеллектуальный анализ сигналов, данных и знаний: методы и средства сборник статей Всероссийской научно- практической конференции с международным участием. Новосибирский государственный технический университет. 2017. – С. 161-166.
  26. Абдикеримова Г.Б., Бычков А.Л., Вей С., Мурзин Ф.А., Русских Н.Е., Рябчикова Е.И., Хайрулин С.С. Библиотека программ для анализа текстур на микрофотографиях // В сборнике: Интеллектуальный анализ сигналов, данных и знаний: методы и средства сборник статей Всероссийской научно- практической конференции с международным участием. Новосибирский государственный технический университет. 2017. – С. 98-102.
  27. Murzin F.A., Batura T.V., Yerimbetova A.S., Semich D.F., Bakiyeva A.M., Efimova L.V. About the search information system based on link grammar // Distributed Information-Computational Resources. Science - The Digital Economy (DICR-2017): Proceedings of the XVI All-Russian conference, Novosibirsk, Akademgorodok, 4 - 7 December 2017. Novosibirsk: ICT SB RAS, 2017. – P. 100–114, ISBN 978-5-905569-10-4.
  28. Мурзин Ф.А., Батура Т.В., Еримбетова А.С., Семич Д.Ф., Бакиева А.М., Ефимова Л.В. О системе поиска информации на основе грамматики связей. В сборнике: Распределенные информационно-вычислительные ресурсы. Наука-цифровой экономике (DICR-2017). Труды XVI Всероссийской конференции. Институт вычислительных технологий СО РАН. 2017. С. 110-114.
  29. Levichev A. V. On key properties of the intertwining operators ornament in the matrix multi-level model of the quark-gluon media// В Материалах Всероссийской конференции с международным участием "Знания-Онтологии-Теории" (ЗОНТ-2017), 6-8 октября 2017, ISSN 0568-661X Под ред. Д.Е.Пальчунова, Новосибирск: Изд-во Ин-та математики, 2017. Том 2. С. 41--47.
  30. Апанович З.В. Современные тенденции визуализации коллекций научных публикаций // Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2017. — С. 3-8. — URL: <http://keldysh.ru/abrau/2017/36.pdf> doi:10.20948/abrau-2017-36
  31. Апанович З.В. Преподавание методов Semantic Web разработчикам программного обеспечения // Научный сервис в сети Интернет: труды XIX Всероссийской научной конференции (18-23 сентября 2017 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2017. — С. 9-20. — URL: <http://keldysh.ru/abrau/2017/37.pdf> doi:10.20948/abrau-2017-3

32. Апанович З.В. Методы визуализации коллекций научных публикаций //Материалы Всероссийской конференции с международным участием “Знания-Онтологии\_Теории” (ЗОНТ-2017) 2-6 октября 2017 г. Новосибирск С. 22-28.
33. Апанович З.В. Опыт преподавания принципов, методов и средств Semantic Web//Материалы Всероссийской конференции с международным участием “Знания-Онтологии\_Теории” (ЗОНТ-2017) 2-6 октября 2017 г. Новосибирск С. 29-37.
34. D.V.Antonets, N.E. Russkikh Using machine learning approach to improve base calling in next generation sequencing data // Moscow Conference on Computational Molecular Biology (MCCMB), abstracts, p. 182, Moscow, Russia, July 27-30. 2017

#### **Учебные пособия**

1. Шелехов В.И., Чушкин М.С., Пономарев Д.К. Формальные методы в программной инженерии. Видеолекции. — ИСИ СО РАН, Новосибирск, 2017.  
<http://wasp.iis.nsk.su/>

#### **Патенты и свидетельства о государственной регистрации программ для ЭВМ**

**Свидетельство о государственной регистрации программы для ЭВМ № 2017612824 от 03.03.2017.** «Программная система для визуализации сложных больших данных на основе графовых моделей (Visual Graph).»

*Касьянов В.Н., Золотухин Т.А.*

*Правообладатель: ИСИ СО РАН.*

**Свидетельство о государственной регистрации программы для ЭВМ №2017611829 от 09.02.2017г** «Модуль анализа информации: классификатор текстов»

*Павловский Е.Н., Масловский И.А., Батура Т.В. (ИСИ), Дюбанов В.В.*

*Правообладатель: НГУ.*