

УДК 004.45

Лаборатория информационных систем. Основные научные результаты, полученные в 2021 году

Марчук А.Г. (Институт систем информатики СО РАН),

Андреева Т.А. (Институт систем информатики СО РАН),

Городняя Л.В. (Институт систем информатики СО РАН),

Демин А.В. (Институт систем информатики СО РАН),

Крайнева И.А. (Институт систем информатики СО РАН),

Пономарев Д.К. (Институт систем информатики СО РАН),

Тихонова Т.И. (Институт систем информатики СО РАН),

Филиппова М.Я. (Институт систем информатики СО РАН)

В статье изложены основные научные результаты, полученные лабораторией информационных систем ИСИ СО РАН в 2021, их связь с мировыми исследованиями и работами предыдущих лет.

Ключевые слова: *Лаборатория информационных систем ИСИ СО РАН, семантические методы, исследования по истории науки, методики обучения программированию, технологии обработки данных, информационные ресурсы.*

1. Введение

Обязательный раздел, в конце которого можно поместить информацию о поддержке исследования грантами.

Выполняемый в лаборатории проект направлен на: создание фундаментальных подходов и технологий для автоматизации предметных областей, решения задач в сфере анализа данных, прогнозирования, принятия решений; формирование принципов, методологии и технологии для историко-ориентированных информационных систем, в том числе электронных архивов, реализующих фактографический подход; описание, формализация и

развитие моделей, методик и технологий обучения программированию, востребованных в современной ИТ-индустрии; изучение важных классов задач, связанных с обработкой больших данных, предложение и техническая реализация эффективных алгоритмов.

Более конкретно, в фокусе интереса исследований были: нахождение и обоснование новых способов комбинирования логических, алгебраических и вероятностных методов описания предметных областей, интеграция символьных и субсимвольных моделей анализа данных.

Изучались особенности создания электронных архивов, в частности архивов СО РАН, на основе междисциплинарного взаимодействия гуманитаристики и точных наук (математики, информатики), акцентируя внимание на источниковедческом аспекте. Формулировались технологические и организационные проблемы, выявленные при создании источниково-ориентированных ИС (инженерный и научный подходы).

Продолжалось создание новых методик обучения программированию, методики опробовываются на Летних школах юных программистов (ЛШЮП) и в спецкурсах НГУ.

Производилась классификация задач и схем структуризации, характерных для различных предметных областей. Это касается классических задач делопроизводства, социальных сетей, обработки физической, биологической, астрофизической информации. Анализ существующих реализаций различных схем структуризации и алгоритмов обработки. Выделение модельных задач, формирование бенчмарков.

2. Семантические методы интеллектуального анализа, инженерии знаний и управления

Цель исследования заключается в разработке фундаментальных подходов и технологий для автоматизации предметных областей и решения задач в сфере анализа данных, прогнозирования, принятия решений.

Актуальность проблемы, предлагаемой к решению, заключается в общественной потребности в создании методов интеллектуального анализа данных и знаний в построении систем поддержки принятия решений. Конкурентное преимущество обеспечивают технологии, которые позволяют производить аналитику и прогнозирование на данных большого объема, легко интерпретируемы и универсальны в настройке на предметную область. Это обеспечивается интеграцией семантических методов моделирования предметных областей и вероятностных моделей обработки данных, развитием фундаментальной и инженерной базы.

2.1 Обнаружение причин и причинно-следственных связей в информационных моделях событий

Большинство направлений в информатике связаны с математическими моделями, которые дают ответ на вопрос «что» (например, ответ на запрос из базы данных, результат автоматизированного логического вывода и пр.) и лишь немногие позволяют отвечать на вопросы «почему не...», «что если...». При взаимодействии пользователя с системой такие вопросы возникают часто. Способность отвечать на них зачастую продуктивнее, чем возможность системы предоставить объяснение в виде последовательности сделанных своих шагов. В информационных моделях, которые представляют сложные взаимосвязи событий, подобная функциональность является ключевой, а задача установления причинно-следственных связей – весьма нетривиальной.

Цель исследования: создание технологии интеллектуальных систем-ассистентов, позволяющих рассуждать о причинно-следственных связях событий, обнаруживать причины, производить контрфактический анализ событий.

Само понятие причины трудно поддается осмыслению с общих позиций и, тем более, формализации. Примеры из юриспруденции наглядно показывают, что установление причины является комплексным процессом: причиной может быть как действие, так и бездействие одного или нескольких акторов. Более ста лет в литературе идет дискуссия по поводу того, как определить понятие причины. Многие подходы и модели за это время потерпели поражение от найденных интуитивных контрпримеров, которые опровергают ту или иную теорию причин. В последнее время были развиты пропозициональные модели событий, в которых делается попытка построения алгоритмов для выявления причин. Например, J. Pearl с соавторами в [1] развил подход на основе баесовских сетей. В [2] J. Halpern реализовал концепции каузальности Дэвида Юма в рамках систем присваиваний, формулирующих зависимости одних переменных от других. К сожалению, и эти подходы неустойчивы к ряду наглядных контрпримеров. По-видимому, пропозициональные подходы не способны ухватить широкую специфику каузальности, поскольку их язык недостаточно выразителен.

Задача исследования: построение математической модели событий, не уступающей в функциональности пропозициональным моделям и превосходящей их в спектре охватываемых предметных областей и приложений. Построение алгоритмов обнаружения причин, контрфактического анализа событий в рамках модели.

Гипотеза исследования: для анализа причинно-следственных связей необходима модель,

включающая такие ингредиенты, как: действие, понятия шкалы времени, выраженное через последовательность действий, параллельные и последовательные действия, пред/постусловия действий, способность рассуждать о будущем и прошлом (forward/backward reasoning). Все они, кроме параллельных действий, присутствуют в формализме Исчисления Ситуаций [3], основанном на логике первого порядка. В работе изучается применимость этого подхода, как математической основы для причинно-следственного анализа. В сотрудничестве с факультетом информатики Ryerson University, Toronto разрабатывается расширение Исчисления Ситуаций параллельными действиями, развиваются новые методы причинно-следственного анализа, разрабатываются алгоритмы обнаружения причин в информационных моделях событий. Ранее в сотрудничестве ИСИ СО РАН-Ryerson University были разработаны новые методы компонентного представления и анализа теорий действий в Исчислении Ситуаций [4-5], которые сформировали основу для настоящих исследований.

Результаты проведенных за 2021 год исследований опубликованы в [6-8].

2.2 Интерпретируемые логико-вероятностные модели обучения с подкреплением

Направление исследований "обучение с подкреплением" (reinforcement learning, RL) [9] занимается проблемой обучения агента путем его взаимодействия с окружающей средой методом "проб и ошибок". В последнее время доминирующим в RL стал подход "глубокое обучение с подкреплением" ("deep reinforcement learning", DRL), объединяющий классический RL и глубокие нейронные сети. К примеру, алгоритмы DRL позволили машине автоматически обучиться играть в игры Atari, получая на вход лишь сырые пиксельные данные [10], а также побить чемпионов мира в игру го [11]. Однако, при более детальном рассмотрении можно заметить, что своим успехом многие современные эффективные DRL приложения в основном обязаны использованию моделей глубоких нейронных сетей для анализа входной сенсорной информации. В частности, Atari DRL эффективно использует глубокие сети для преобразования сырых пиксельных данных в сжатое представление, пригодное для использования RL методами. При этом по-прежнему остались нерешенными ряд традиционных проблем обучения с подкреплением, в частности: 1) проблема разбиения задачи на подзадачи (обнаружение подцелей) и 2) проблема интерпретируемости модели.

Решению первой проблемы занимается направление RL, называемое "иерархическое обучение с подкреплением" (hierarchical reinforcement learning, HRL) [12, 13], которое объединяет различные подходы к группировке элементарных действий для более

эффективного обучения агента и решения им задач. Однако одна из основных проблем большинства этих подходов заключается в необходимости заранее задавать подцели. Таким образом, задача автоматического обнаружения подцелей по прежнему остается крайне актуальной.

Вторая проблема, интерпретируемость модели, также является очень актуальной, поскольку доминирующие в RL нейросетевые модели работают по принципу "черного ящика". В настоящее время предлагается два основных решения: 1) обучение интерпретируемых моделей (деревья решений, набор правил и т.д.) по выходам уже обученной нейросетевой модели (к примеру, метод PIRL) [14]; 2) использование генетических методов [15]. Оба подхода нельзя напрямую отнести к RL методам, поскольку в первом случае необходимо сначала обучить нейросеть, а во втором требуется наличие популяции агентов. Таким образом, можно утверждать, что в настоящее время не существует RL подходов, дающих интерпретируемую модель сразу в процессе обучения агента.

Целью данного исследования является разработка интерпретируемых методов обучения с подкреплением с автоматическим формированием иерархии целей. В исследовании предлагается альтернативный подход, который, с одной стороны, использует идеи организации управления из нейрофизиологической Теории функциональных систем, а с другой стороны, логико-вероятностный методы (ЛВ-методы) машинного обучения для формализации модели и тренировки агента. Одно из преимуществ использования ЛВ-методов состоит в том, что обнаруженные в результате обучения закономерности имеют явную форму, т.е. представлены в виде логических формул. Это, во-первых, сразу дает интерпретируемую модель, во-вторых, позволяет анализировать и использовать полученные закономерности для построения мета-алгоритмов (извлечение иерархии подцелей, построение плана, автоматическое обнаружение категорий, не сформулированных человеком).

В предыдущих работах [16-19] были предложены логико-вероятностные RL-модели, основанных на обучении прогнозированию оценок результатов своих действий в определенных ситуациях (model-free модели). Данные модели были способны обнаруживать явные подцели и успешно решали двухэтапную задачу фуражирования. Однако недостатками предложенных ранее моделей являлось отсутствие поддержки обучения модели среды и способность обнаруживать только явные подцели (подцели, информация о достижении которых постоянно присутствует в сенсорном поле агента).

Были проведены следующие теоретические исследования:

- 1) Разработка логико-вероятностных RL-моделей с иерархией подцелей: model-free

(основанные на обучении отображению состояний и действий в прогноз награды, соответствуют большинству классических RL-моделей), model-based (основанные на обучении модели среды и формирующие прогноз награды на основе полученной модели).

2) Разработка логико-вероятностных алгоритмов обучения для разных типов RL-моделей с учетом иерархии целей.

3) Разработка методов автоматического построения иерархии целей, включая неявные подцели, на основе анализа закономерностей, полученных ЛВ-методами обучения, и истории работы агента.

4) Разработка методов автоматического построения концептов на основе анализа закономерностей (для формирования концептуального представления о среде и рефлексии агента с целью ускорения обучения, сокращения пространства перебора действий при исследовании сред, переноса опыта на новые среды).

Теоретические исследования подкреплялись практическими исследованиями и экспериментальными системами:

1) Программная реализация моделей и алгоритмов.

2) Разработка сред для экспериментального тестирования моделей.

3) Проведение экспериментальных исследований с целью апробации моделей и реализаций. Постановка задач для корректировки и доработки моделей и алгоритмов.

4) Сравнение предложенных моделей с другими RL-подходами на общедоступном бенчмарке (из репозитория <https://gym.openai.com/>).

Были получены следующие научные результаты:

1) Предложена логико-вероятностная RL-модель, основанная на обучении модели среды, и новый метод обнаружения глубоких неявных подцелей.

2) Проведены экспериментальные исследования предложенной модели на примере многоэтапной задачи фуражирования с выделением подцелей различной глубины. По результатам исследований были выполнены доработки модели и метода извлечения подцелей. Были выделены мета-параметры системы и исследовано их влияние на работу модели.

Результаты исследований изложены в [20].

В рамках научного сотрудничества с Mike Soutchansk (Department of Computer Science, Ryerson University, Toronto, Canada) проведены совместные исследования по математической формализации причин и причинно-следственных связей, в том числе для юридических кейсов. Начата работа над совместной статьей.

3 Системы и технологии поддержки исследований по истории науки, техники и образования

Цель исследования заключается в том, чтобы оценить современные потребности науки и образования в области создания и применения источник-ориентированных информационных систем, обобщить теоретический и практический опыт в области понятийного аппарата, подходов к их проектированию, методов и технологий создания, принципов использования в исследовательской и образовательной практике.

Сформировать принципы, методологии и технологии для историко-ориентированных информационных систем, в том числе электронных архивов, реализующих фактографический подход.

Актуальность данной работы базируется на успешной практической реализации в конкретных исследованиях инструментального подхода к публикации источников по истории науки, техники и образования. Высокие требования к публикационной активности гуманитариев, требования новизны и актуальности исследований с новой силой ставят вопрос о большей и оперативной доступности исследователю архивного, библиотечного и прочего контента наследия. В этом случае особую актуальность приобретает доступность, открытость архивов. Опыт ИСИ СО РАН, который в числе первых научных коллективов осуществил несколько проектов по созданию, научной интерпретации, методической и организационной работе в области академических цифровых архивов может быть актуализирован как технологичный, научно обоснованный и успешно апробированный.

3.1 Сохранение научного наследия на базе ИТ и исследования исторического характера

Большая часть научного наследия Сибирского отделения АН СССР/РАН с момента его образования в 1957 г., является объектом хранения Научного архива СО РАН (НАСО). С 2014 г. после сокращения штатов в результате реформы РАН, НАСО прекратил формирование научных коллекций. Перед коллективом Лаборатории информационных систем ИСИ СО РАН стоит задача мессианского характера: сохранить наследие Отделения насколько это возможно, разместив документы в электронных архивах. Нами выявляются как объекты наследия первостепенной значимости, так и случайные коллекции, которые нам предоставляют фондообразователи или их наследники (фондохранители). Таким образом наполняются электронные архивы СО РАН: Открытый архив и Фоторахив.

В настоящее время сформировано несколько направлений, по которым идет наполнение контента электронных архивов: персональные фонды, институциональные фонды, фонды общественных организаций, фонд устной истории. Наиболее активно формируются персональные фонды. В 2021 г. Открытый архив СО РАН пополнился несколькими значимыми коллекциями:

- Фонд академика Н.Л. Добрецова (1936-2020), заместителя председателя СО РАН (1989-1997), председателя СО РАН (1997-2008). (3697 сканов документов)
- Фонд д.ф.-м.н. Л.А. Боярского (1933-2020) – ведущего научного сотрудника Института неорганической химии СО РАН, профессора, зав. кафедрой физики низких температур НГУ, руководителя клуба любителей кино «Сигма» в ДУ СО РАН (1965-2020). (515 ск.)
- Фонд чл.-корр. А.А. Ляпунова (1911-1973) пополнился ценной перепиской с учениками, присланной нам их Чехословакии. (196 ск.)
- Фонд академика Г.И. Марчука (1925-2013) – положено начало формирования данного фонда, который планируется создать к 100-летию со дня рождения ученого. (224 ск.)
- История Дома ученых СО РАН пополнилась документами от нашего корреспондента из США. (62 ск.)

Всего в 2021 г. на платформе открытых архивов размещено 6674 скана документов (см. Фотоархив СО РАН – 430 сканов фото.

Открытый архив СО РАН – 6244 сканов документов.

Наши исследования лежат в русле мировых трендов. Информатизация научно-исследовательской деятельности – актуальное направление в области развития коммуникативных процессов, которое осуществляется через сетевую организацию вычислительной техники, и поддерживает доступность контента наследия, размещенного в Сети. Фокус-группа проектов – научное сообщество. Глобально и институционально это направление исследований и практических шагов связано с возрастанием потока информации. В то же время, мировое научное сообщество озабочено обеспечением качественной информацией, поэтому идея большого виртуального архива науки для историко-научных и прочих исследований гуманитарного направления вполне своевременна.

Идея архивации наследия оказалась настолько плодотворной, что вызвала к жизни несколько международных проектов по унификации подходов его оформлению. Одним из них стал общеевропейский проект The Open Archives Initiative (OIE, 1999) [21]. OIE «разрабатывает и продвигает стандарты функциональной совместимости, которые

направлены на содействие эффективному распространению контента. ОИЕ стремится улучшить доступ к электронным архивам как к средству повышения доступности научного общения. Однако фундаментальные технологические рамки и стандарты, которые разрабатываются для поддержки этой работы, не зависят ни от типа предлагаемого контента, ни от экономических механизмов, которые окружают этот контент, и обещают иметь гораздо более широкое значение для открытия доступа к целому ряду цифровых источников. По мере того, как ОИЕ получает больше знаний о масштабах применимости разрабатываемых базовых технологий и стандартов, понимает структуру и культуру различных сообществ последователей, предполагается внесение постоянные эволюционных изменений как в миссию, так и в организацию ОИЕ [22]. Группа ОИЕ разработала некоторые спецификации, такие как Протокол сбора метаданных, Руководство по внедрению Протокола ОИЕ для сбора метаданных и т. д. Ассоциация предложила кодирование XML в качестве механизма упаковки для собранных метаданных.

Практически одновременно был запущен проект euroCRIS – The International Organization of Research Information. EuroCRIS предусматривает, что Общий европейский формат исследовательской информации (CERIF) является всеобъемлющей информационной моделью в области научных исследований. Он предназначен для поддержки обмена исследовательской информацией между платформами CRISs. На этом основаны Руководящие принципы OpenAIRE для CRIS-менеджеров [23]. Технический комитет по взаимодействию и стандартам (TCIS) был создан в сентябре. 2020, параллельно с CERIF TG, и направлен на процесс принятия решений для развития CERIF и связанных с ним технических продуктов. TCIS призван установить дорожную карту и стратегию для широкого внедрения CERIF и его согласования с другими соответствующими технологиями, моделями и стандартами.

Среди последних отечественных разработок в области электронного хранения исторических артефактов назовем проект Европейского университета в Санкт-Петербурге «Прожито», запущенный в 2015 г. Здесь аккумулируются такие эго-документы, как дневники. Корпус включает датированные тексты на русском и украинском языках. Общий объем корпуса – более полумиллиона записей XVIII–XX вв. Загружено около 500 [4].

3.2 Исследования исторического характера

Пополнение электронных архивов позволяет решать исследовательские задачи. На базе имеющихся и вновь полученных документов проводятся исследования по истории науки в Сибири. Фокус исследований сосредоточен как на технических и технологических, так и на

гуманитарных вопросах. В частности, акцентируется внимание на персональных биографиях женщин-ученых, истории научных династий, на отдельных периодах истории науки (Великая отечественная война, начальный период становления СО АН СССР). Нами подготовлено и опубликовано несколько статей для журналов, сделаны научные доклады на профильных конференциях.

Так, на основе документов из коллекции академика Татьяны Ивановны Заславской (1927–2013) и доктора филологических наук Майи Ивановны Черемисиной (1924–2013), в девичестве сестер Карповых, проведен краткий анализ данного персонального фонда. В России известно несколько научных династий, которые ведут свою историю с начала XIX века: Ляпуновы-Анри, Капицы-Милевские, Лаврентьевы, Вернадские, Шмальгаузен, Ворожцовы и др. К подобным династиям принадлежали сестры Карповы, достойные потомки первой в семье женщины-ученой Ольги Карловны Крафт, которая получила степень доктора медицины в Париже в 1884 г. Дед по материнской линии Г.Г. Де-Метц (1861–1947) – физик, профессор, был деканом физико-математического факультета, затем ректором Киевского Императорского университета Св. Владимира, одним из организаторов Киевского политехнического университета и Кубанского государственного университета. Отец Иван Васильевич Карпов (1897–1965) – кандидат педагогических наук, историк, в качестве вольноопределяющегося воевал в Первой мировой. Исследуя историю семьи Де-Метц–Крафтов–Карповых, мы ставили задачу «прочитать» тексты персональных историй в контексте культуры и социума, выявить глубинное влияние семейной истории на личную. Кроме того мы хотели привлечь внимание коллег к данной коллекции [25].

В качестве еще одного примера использования материалов Открытого архива СО РАН с привлечением материалов Государственного архива РФ является реконструкция персональной истории одного из представителей культурной московской семьи Румеров – Исидора Борисовича Румера (1884–1938). Биография прослежена в актуальном контексте взаимоотношения российской интеллигенции и власти. Внимание концентрируется на особенностях репрессивной политики в отношении интеллигенции в довоенный период с акцентом на 1930-е гг. В это время происходили необратимые перемены в экономической, политической и культурной политике Советского государства. Они состояли в насаждении единой идеологической доктрины, командных методов управления учреждениями науки и культуры, в фактическом распоряжении властью всей собственностью. Несогласие сопровождалось массовыми репрессиями в отношении явных и мнимых противников. Среди них – представители отечественной интеллигенции, деятели культуры и науки и техники, в своей массе не являвшиеся непримиримыми противниками Советской власти. Их

критический настрой, как правило, носил латентный характер, являлся проявлением ментальности интеллигентов-гуманистов. Они озвучивали свое мнение в свободном общении в кругу единомышленников. Исследование проведено с привлечением архивных материалов, публикаций по теме, осмысления феномена отечественной интеллигенции [26].

Поскольку Открытые архивы СО РАН продолжают пополняться, они предоставляют неисчерпаемые возможности для исследований в области истории науки и техники, биографики, локальной истории науки. Результаты исследований изложены также в [27-30].

4 Развитие новых моделей, методик и технологий обучения программированию

Целью исследования является описание, формализация и развитие моделей, методик и технологий обучения программированию, накопленных за историю развития вычислительной техники и активно востребованных в современной ИТ-индустрии.

Проблема образования в информатике и программировании упоминается в лекциях многих лауреатов премии Тьюринга, тем не менее, её полного решения пока достичь не удалось, не все достижения этого направления сохранены в наши дни. Прецедент успешного образования программистов под Ершовским лозунгом «Программирование — вторая грамотность» формировался шире, чем чисто университетское или школьное обучение на уровне первичных навыков, дополнением было и самообразование, и дистанционное общение, и факультативные формы типа школ юных программистов (ШЮП).

Актуальность проблемы давно видна флагманам компьютерной индустрии, испытывающим серьёзные трудности в привлечении специалистов, имеющих навыки приаппаратного программирования и обладающих изобретательскими способностями для повторного программирования системных библиотек.

4.1 Развитие парадигмального анализа языков и систем программирования

При исследовании и сравнении парадигм программирования получено два результата:

- 1) Описана мультипарадигмальность параллельных вычислений. Получение этого результата опирается на методику парадигмального анализа языков и систем программирования, применённую к проблемам организации параллельных вычислений и многопоточных программ для многопроцессорных комплексов и распределённых систем. Такая методика позволила сделать вывод о неявной мультипарадигмальности параллельных

вычислений. На основе этого вывода сформулированы требования к учебно-производственному мультипарадигмальному языку параллельного программирования [31].

2) Представлены принципы функционального программирования. При получении этого результата методика парадигмального анализа языков и систем программирования применена к описанию особенностей функционального программирования как парадигмы, выполняющей функции проектно-конструкторского бюро для производственного и параллельного программирования. В результате сформулированы принципы функционального программирования, позволяющие отличать его от других парадигм программирования. Описаны следствия из этих принципов, конкретизированы принципы и следствия при переходе к проблемам параллельного программирования, а также при выполнении работ по повышению эффективности и производительности программ [31, 32].

В мире подобные задачи рассматривались в [33-36]. В настоящее время происходит расширение двух взаимодействующих, формально почти противоположных, направлений. Первое связано с экстенсивным порождением новых проблемно-ориентированных языков программирования (DSL) — их за последнее десятилетие насчитывается десятки тысяч. Второе отражает рост актуальности проблем производительности ПО, особенно на уровне энергопотребления, использования памяти и немонотонности скоростей исполнения программ, что было не очень заметно ещё пять лет назад.

Первое направление провоцирует избыточный объём производства ПО без принципиального роста производительности программ, связанного с техникой синтаксического конструирования новых ЯиСП преимущественно на базе Clang-LLVM над языками семейства C/C++, что ограничивает реализационное пространство возможностями прежней базы, сложившейся на однопроцессорных конфигурациях ещё в 1970-80-е годы, без перехода к рациональному освоению преимуществ новой элементной базы, многопроцессорных конфигураций и компьютерных сетей. Легко создать новый удобный компьютерный язык, трудно обеспечить заметный рост производительности программ, создаваемых с его помощью.

Второе направление приторможено риском новой реализации приаппаратных решений, для которых даже самые мощные производители аппаратуры испытывают дефицит необходимой высокой квалификации разработчиков. По этой причине по 20-30 лет эксплуатируются неизменные ядра компиляторов и ведутся массовые, не имеющие перспектив успеха, эксперименты по решению приаппаратных проблем изобретением методов высокоуровневого декаривования программ. Таких подходов уже опубликовано более ста, идёт поиск новых при отсутствии предложений по пересмотру языковых моделей

работы с аппаратурой. Легко измерять суммарную производительность программы, сращенной с аппаратурой, операционной системой и системой программирования. Трудно выделить из полученных результатов измерения вклад программируемых решений в производительность программ.

Таким образом, оба направления вышли на уровень трудно решаемых проблем. Обычно решение трудных проблем лежит через поиск альтернативного подхода или синтез с комплексом других направлений, уже имеющих технику решения сложных задач. Прецеденты работоспособной техники решения таких сложных задач имеются в современных оптимизирующих компиляторах и накоплены при исследовании методов верификации и моделирования параллельных процессов. Альтернативный подход может дать пересмотр методов реализации компьютерных языков, что требует дальнейшего исследования, особенно при их синтезе с методами верификации, логического вывода и функционально-информационного моделирования. В рассмотренных материалах доступных международных конференций таких обобщённых попыток обнаружить не удалось.

Анализ современного состояния исследований в области ЯиСП показывает существование трудно преодолимой дистанции между потенциалом современных ИТ, достижениями теоретических исследований и исторически сложившейся практикой создания информационных систем, что тормозит прогресс, особенно заметный в области параллельных вычислений, обретающих практический интерес на базе массового доступа к многопроцессорным архитектурам. Заметен яркий дисбаланс между мощностью средств интеграции программных комплексов и дефицитом инструментальных возможностей корректного выделения программных компонент без потери их функциональности. Это отмечено в материалах ICS International Conference on Supercomputing как ведущего интернационального форума по представлению результатов высокопроизводительных вычислений, отражающего все аспекты развития, исследования и применения суперскалярных экспериментальных и коммерческих программных систем поиска научных ответов на вызовы современных ИТ.

Это по существу создаёт почву для внедрения в практику результатов теоретических работ в области верификации и системного анализа, практическую основу которых может составить функциональный подход, поддержка жизненного цикла разработки безопасного ПО, инкрементального анализа систем (подразумевает быструю повторную проверку недавно измененного кода) и автоматическую генерацию тестов для сложных программных систем. В этом плане интересен Svace – необходимый инструмент жизненного цикла разработки безопасного ПО, обнаруживающий более 50 классов критических ошибок в

исходном коде. Имеется поддержка навигации по коду — разделение срабатываний на истинные и ложные;— миграция результатов между запусками и сокрытие ложных срабатываний. Привлекает внимание и ИСП Crusher – программный комплекс, комбинирующий несколько методов динамического анализа.

Ранее, в Лаборатории информационных систем ИСИ СО РАН по этому направлению были получены результаты [31, 32], показывающие основу для проведения дальнейших исследований в таком направлении.

Попытки описания принципов функционального программирования и методик обучения в мире регулярно рассматриваются на международных конференциях, посвященных перспективам функционального программирования, поддерживаемых АСМ и другими авторитетными организациями. [33-38].

Опыт реализации языков функционального программирования и современная тенденция к мультипарадигмальности новых языков программирования дали достаточные основания для перехода к ЯиСП, одновременно содержащим интерпретатор, компилятор, мемоизатор и декомпозитор программ, а также для создания новых средств более высокого уровня, сравнимого с механизмами языков сверх высокого уровня. Появились попытки выделения типовых семантических систем из определений разных ЯиСП. Само название «The next 700 semantics: a research challenge» [41] можно рассматривать как симптом избыточного разнообразия выделяемых компонент. Избыточность препятствует определению чётких критериев систематизации языков и парадигм программирования.

Исследование подходов к систематизации средств и методов разработки ПО активизировано в сентябре 2009 года Иваром Якобсоном, Бертраном Мейером и Ричардом Соули, выступившим с инициативой SEMAT, основы которой они изложили в своей книге «The Essence of Software Engineering: Applying the SEMAT Kernel». Их идею поддержали многие гуру программирования и включили разные корпорации. В настоящее время Бертран Мейер, автор языка Eiffel и книг по объектно-ориентированному проектированию, с другими коллегами разворачивает инициативу «PEGS - Project, Environment, Goals and System» по исследованию проблем определения инженерных требований к ПО. Учитывая, что профессиональная карьера Б.Мейера началось с работы по достаточно интересному языку спецификаций Z0, можно предполагать хорошие результаты [36, 37].

Общий вывод из содержания рассмотренных материалов по отношению ЯиСП и проблемам параллельного программирования и производительности программ сводится к констатации ряда не решённых проблем. Значение работ в области ЯиСП освещено на ряде конференций [33-38].

Нет средств удобного представления взаимодействующих потоков и их эффективного перевода в процессы, не происходит преодоление страха перед разработкой и отладкой параллельных алгоритмов, особенно в случае лёгкого доступа к готовым последовательным алгоритмам, не созданы методики решения образовательных проблем, включая осознание реальных возможностей аппаратуры, требующих проявления интуитивной грамматики параллельных процессов. Ещё одна проблема связана с отсутствием общепризнанной методики представления результатов измерения вклада программируемых решений в производительность программ. Хотя специалисты, исследующие эту проблему, отмечают низкий уровень заинтересованности программистского корпуса в измерениях, они потратили 20 лет на измерение производительности программных приложений.

Ранее, в Лаборатории информационных систем ИСИ СО РАН были получены результаты [31, 39-43], позволяющие наглядно представлять результаты анализа определений языков программирования и измерения эксплуатационных характеристик программируемых решений, влияющих на производительность программ.

4.2 Создание автоматизированной системы подготовки материалов для олимпиад и тестов

Целью исследований является разработка информационной системы, предназначенной для полной или частичной автоматизации подготовки материалов (задачных и тестовых наборов) для проведения олимпиад и проверочных работ в различных областях знаний (включая гуманитарные науки), позволяющей снизить трудоемкость и сложность подготовительных процессов и расширить за счет этого целевую аудиторию пользователей. Актуальность темы значительно возросла на фоне массового перехода образования в дистанционный формат. Однако исследования по данной теме практически не проводятся: за последние три года в научной электронной библиотеке e-Library зарегистрировано всего шесть работ по близкой тематике (см. [44-49]). Все они посвящены частным вопросам автоматизации проверок заданий, преимущественно в программировании и близких к нему науках. Попыток обобщить и распространить опыт автоматизации подготовки тестовых заданий на другие области школьного и вузовского образования не производилось.

Изначально целью исследований было создание комплексной информационной системы, предназначенной для подготовки материалов (задачных и тестовых наборов) для проведения олимпиад и тестов, позволяющей снизить трудоемкость и сложность подготовительных процессов и расширить за счет этого целевую аудиторию пользователей.

Ранее система подготовки тестовых комплектов была расширена до системы подготовки задачных комплектов, включающих условие задачи, спецификацию входных и выходных данных, тестовый набор, авторское решение, принцип вынесения вердикта о правильности предоставленного участником решения.

В текущем периоде была продолжена разработка системы, а) формализующей при помощи разработанного математического аппарата спецификации входных и выходных данных, заданных на естественном языке; б) генерирующей тестовые наборы исходя из формализованных спецификаций; в) контролирующей совместимость набора спецификаций и форматов данных; и д) проверяющей непротиворечивость, полноту и избыточность тестовых наборов.

Внесены дополнения и уточнения в математическую модель вынесения вердиктов о правильности решений на основе различных способов тестирования решений. Поскольку правильность вердикта существенно зависит от качества тестовых наборов, основное внимание было уделено разбору частных случаев некорректных и псевдо-корректных процедур оценивания.

Особое внимание было уделено возможности коллективного использования системы, поэтому возникла необходимость включить в систему дополнительные операции с тестовыми наборами: проверка соответствия тестов формализованным спецификациям, сведение воедино нескольких тестовых наборов, проверка непротиворечивости, полноты и избыточности тестовых наборов в соответствии с формализованными спецификациями.

В рамках исследования параллелизма процессов, возникающих при автоматизации коллективной разработки задачных наборов, были сделаны важные замечания, развитие которых планируется осуществить в дальнейшем.

4.3 Разработка методов автоматической верификации тестовых наборов

Многочисленные примеры тестовых систем, вошедшие в школьное и вузовское обучение на волне роста дистанционного образования (Moodle, SkySmart, ЯКласс, др.) выявили недостаток внимания к проблеме верификации тестов, создаваемых пользователями этих систем. Проверка правильности создаваемых тестов в этих системах совершенно не автоматизирована. Кроме того, не учитываются различия в типах тестовых заданий, которые могли бы сократить количество ошибочных заданий.

В дополнение к методам автоматической проверки генерируемых тестовых наборов, проводится разработка математически обоснованных методов автоматической и

полуавтоматической верификации полноты и непротиворечивости тестовых заданий, создаваемых пользователями вручную.

Поскольку правильность вердикта существенно зависит от качества тестовых наборов, основное внимание было уделено разбору частных случаев некорректных и псевдокорректных процедур оценивания.

Для наиболее полного охвата различных типов возможных ошибок проводится сбор примеров ошибочных заданий, созданных пользователями дистанционных образовательных систем (Moodle, SkySmart, ЯКласс, др.)

4.4 Практическая проверка

В условиях дистанционной работы со студентами НГУ промежуточные результаты исследований были подвергнуты практической проверке на примере системы MOODLE (НГУ), предоставляющей недостаточный инструментарий для создания тестов, а также на примере системы автоматической проверки решений, используемой на 1 курсе ФИТ НГУ (getlab.ccf.it.nsu.ru, автор Петров Е.С.).

По сформированным в ИСИ СО РАН методикам в 2021 году была проведена очередная Летняя школа юных программистов (ЛШЮП-2021) [50]. В связи с эпидемиологической обстановкой, школа проводилась дистанционно, в on-line режиме.

5 Технологии поддержки обработки больших структурированных данных

Целью исследований является изучение важных классов задач, связанных с обработкой больших объемов данных, предложение и техническая реализация эффективных алгоритмов обработки [51]. Разработка заказных и библиотечных решений, развитие и поддержка системы Polar [52].

Актуальность проблемы, предлагаемой к решению заключается в том, что происходящий процесс информатизации множества предметов и процессов деятельности, увеличивающиеся объемы вовлекаемой в обработку информации, интеграция данных и знаний, определяют актуальность разработки эффективных средств структурирования и обработки структурированных данных. Появилось множество подходов, объединенных общим понятием NoSQL, развиваются распределенные системы хранения/обработки, предлагаются и стандартизируются новые подходы с структуризации данных. Вместе с тем, объемы доступных для данных опережают технические возможности компьютерных систем, сетевые

средства ограничивают производительность и доступность большой обработки, что требует нахождения и реализации новых подходов [53].

Продолжались исследования задачи построения и обработки графа де Брёйна (de Bruijn) как одного из представителей актуальных видов данных большого и сверхбольшого размера. Данная модель используется в задаче восстановления цепочек нуклеотидов по результатам секвенирования [54]. Были рассмотрены варианты работы с графом, причем упор делался на экономное программирование как для одномашинной реализации, так и для многомашинной кластерной конфигурации. Полученные программы позволяют решать задач построения частного варианта графа Де Брёйна и решать на этом графе задачу формирования (отслеживания) цепочек узлов. Созданные программы частично сравнивались с решением с использованием инфраструктуры Hadoop в варианте Spark, полученные результаты свидетельствуют в пользу предложенного подхода [55].

Продолжились исследования по созданию средств обработки структурированных данных. Работы 2021 года концентрировались вокруг проблем обработки структурированных данных предельно большого размера. Такие данные появляются в результате масштабных физических экспериментов, больших систем сквозного измерения, накопления данных отдельных областей, биологических и медицинских массивов данных и т.д. Была построена модель основной части структурирования больших данных – последовательности однотипных элементов [56]. В частности, было сформировано представление об универсальном индексном построении. Модель позволяет в едином стиле рассматривать сосредоточенные и распределенные базы данных. Часть модели реализована в библиотеки PolarDB.

Было проведено исследование подходов к модернизации электронных архивов, создаваемых в ИСИ СО РАН. Некоторые из них эксплуатируются более 20 лет. Перевод на современные технологические платформы может сопровождаться решением дополнительных задач, таких как интеграция таких архивов. Были проведены эксперименты с Электронным архивом академика А.П.Ершова, Открытым архивом СО РАН и Фотоархивом СО РАН [57].

6 Создание и поддержание информационных ресурсов, ориентированных на поддержку науки и образования

Основными направлениями в 2021 г. были следующие:

Разработка новой концепции и нового графического дизайна сайта ИСИ СО РАН.

Дальнейшее развитие электронного Архива имени академика А.П. Ершова.

Разработка и поддержка сайтов проектов ИСИ СО РАН.

В 2021 году начата подготовка к обновлению сайта ИСИ СО РАН. Была выработана концепция обновленного сайта, запланированы изменения структуры, подготовлены макеты графического дизайна. Конкретно были выполнены следующие работы:

- 1) Миграция сайта в новое программное окружение – на платформу Drupal 9 и PHP 7;
- 2) адаптация сайта под мобильные устройства различных разрешений.
- 3) разработка новой структуры сайта с учетом добавления новых типов материалов, а также изменившихся требований к сайту со стороны вышестоящих организаций;
- 4) создание современного дизайна взамен устаревшего, который не изменялся с 2010 года.

1) Разработаны несколько вариантов макетов главной страницы, а также и нескольких типов второстепенных страниц для 4 основных разрешений (xs - мобильные телефоны, sm – планшеты, md – компьютеры, lg - компьютеры с широкоформатным монитором).

2) разработаны макеты для иллюстрации оформления отдельных видов материалов и структурных блоков – видео- и фотогалерей, списков, таблиц и т.д.

4) переработана структура расположения материалов и структурных элементов - например, визуализация меню;

5) разработка дизайна велась с учетом того, что основная масса сотрудников института - люди старшей возрастной группы, не очень любят перемены и должны иметь возможность легко ориентироваться на обновленном сайте.

В 2021 г. продолжались работы по дальнейшему Электронного архива академика А.П. Ершова, мигрированного на свободно распространяемое ПО с открытым кодом на платформе Drupal. Ведутся работы по увеличению производительности приложения. Были внедрены изменения в соответствии с запросами пользователей в части бэкенда и фронтенда.

В течение всего отчетного периода продолжались работы по поддержке и обеспечению хостинга сайтов, разработанных в ИСИ СО РАН. Эти работы включают как административные функции

- обеспечение резервного копирования,
- своевременное обновление модулей третьих сторон, что особенно важно при использовании Open Source разработок,
- быстрая реакция на непредвиденные обстоятельства – отключение электроэнергии, сбой аппаратуры и т.д..

так и работы по поддержке – коммуникации с владельцами, обновление материалов по просьбе владельцев, работа с письмами пользователей. В настоящее время на поддержке находятся следующие сайты:

- 1) Сайт ИСИ СО РАН;
- 2) Мемориальная библиотека имени А.П. Ершова;
- 3) Архив академика А.П. Ершова;
- 4) Сайт Бюллетеня Новосибирского Вычислительного Центра;
- 5) Сайт серии конференций PSI;
- 6) Сайт серии конференций PSI до 2017 года;
- 7) Сайт конференций SORUCOM-2014;
- 8) Сайт проекта «Информатика и программная инженерия»;
- 9) Сайт Интеграционного проекта СО РАН № 21 «Веб-пространство»;
- 10) Сайт «Хроника Сибирского отделения»;
- 11) Портал MathTree;
- 12) Коллекция старинных математических книг;
- 13) Сайт ВНТК "СТАРТ";
- 14) Сайт «50 лет Отделу программирования».
- 15) Сайт проекта Кронос;
- 16) Сайт Музея СО РАН;
- 17) Исторический портал ММФ НГУ Global MMF;
- 18) Сайт кафедры программирования ММФ НГУ;
- 19) Мемориальный сайт Г.И. Марчука;
- 20) Мемориальный сайт А.Ф. Пара;
- 21) Мемориальный сайт А.А. Берса;
- 22) Юбилейный сайт В.Е. Котова;
- 23) Юбилейный сайт А.Г. Марчука;
- 24) Исторический сайт «Аллея памяти».

7. Заключение

Все поставленные на 2021 год задачи были выполнены. Были получены существенные результаты по заявленным направлениям:

В исследованиях по причинно-следственному анализу разрабатывается расширение Исчисления Ситуация параллельными действиями, развивается метод причинно-следственного анализа, разрабатываются алгоритмы для цели настоящих исследований.

Была предложена логико-вероятностная RL-модель, основанная на обучении модели среды и новый метод обнаружения глубоких неявных подцелей. Были проведены экспериментальные исследования предложенной модели на примере многоэтапной задачи фуражирования с выделением подцелей различной глубины. По результатам исследований были выполнены доработки модели и метода извлечения подцелей. Были выделены мета-параметры системы и исследовано их влияние на работу модели.

Были выявлены объекты исторического наследия первостепенной значимости и коллекции, которые предоставляют фондообразователи. Открытый архив СО РАН пополнился несколькими значимыми коллекциями.

Выполнен парадигмальный анализ более десяти парадигм и ряда языков параллельного программирования. Создано описание принципов функционального программирования, рассматриваемого как мета-парадигма для организации параллельных вычислений.

С помощью созданного в ИСИ СО РАН программного обеспечения было создано значительное число промышленных и экспериментальных информационных систем. Основные действующие информационные системы в разное время прошли модернизацию и адаптацию к меняющимся технологиям базовых платформ. В 2021 году исследовались принципы сосуществования разных информационных систем, вопросы их интеграции и дезинтеграции. Были проведены содержательные эксперименты по частичному или полному включению ресурсов одного электронного архива в состав другого. Показано, что такое включение может осуществляться без разрушения целостности систем, представляющих авторскую композицию.

Был выполнен большой объем работ по поддержанию и модернизации информационных ресурсов, ориентированных на поддержку науки и образования. Ресурсы созданы в ИСИ СО РАН.

Проведенные исследования выполнены с применением научного метода и, в частности, удовлетворяют требованиям к полноте и опубликованности полученных результатов. В конкретных случаях полученные результаты используются в прикладных исследованиях и разработках или смогут стать основой создания программных продуктов, баз данных и технологий. Технично-экономическая эффективность внедрения результатов не оценивалась. Полученные результаты выполнения НИР находятся на лучшем отечественном и мировом уровне.

Список литературы

1. Judea Pearl, Dana Mackenzie. The Book of Why: The New Science of Cause and Effect // Basic Books, Inc., US, 2018.
2. Joseph Halpern. Actual Causality // MIT Press, 2019.
3. Raymond Reiter. Knowledge in Action // MIT Press, 2001.
4. D. Ponomaryov, M. Soutchanski. Progression of Decomposed Situation Calculus Theories. // Proc. 27th conference on Artificial Intelligence, AAAI'13, July 14-18, Bellevue WA, USA.
5. D. Ponomaryov, M. Soutchanski. Progression of Decomposed Local-Effect Action Theories. // ACM Transactions on Computational Logic (TOCL) 18(2):1-41, 2017.
6. Denis Ponomaryov On the Relationship Between the Complexity of Decidability and Decomposability of First-Order Theories ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2021, Vol. 42, No.

- 12, pp. 2905–2912. DOI: 10.1134/S199508022112026X Web-ссылка на статью должна появиться в конце ноября-начале декабря. WoS, Scopus
7. P. Emelyanov, M. Krishna, V. Kulkarni, S.K. Nandy, D. Ponomaryov, and S. Raha Factorization of Boolean polynomials: Parallel algorithms and experimental evaluation. *Programming and Computer Software* 47: 108-118, 2021. <https://doi.org/10.1134/S0361768821020043> <https://link.springer.com/article/10.1134/S0361768821020043> WoS, Scopus
 8. Пономарев, Д. К. Декомпозиция логических теорий: вычислительные проблемы и приложения // Всероссийская научная конференция «Математические основы информатики и информационно-коммуникационных систем». Сборник трудов. — Тверь : ТвГУ, 2021. — С. 57–60. <https://doi.org/10.26456/mfcsics-21-7>, РИНЦ
 9. Sutton R.S., Barto A.G. Reinforcement Learning. London // MIT Press. – 2012. – 320 p.
 10. Mnih V., Kavukcuoglu K., Silver D. et al. Human-level control through deep reinforcement learning // *Nature* 518. – 2015. – pp. 529-533.
 11. Silver D., Huang A., Maddison C. et al. Mastering the game of Go with deep neural networks and tree search // *Nature* 529. – 2016. – pp. 484-489.
 12. Al-Emran Mostafa. Hierarchical Reinforcement Learning: A Survey // *IJCDS Journal* 4(2) . – 2015. – pp. 137-142.
 13. Dietterich T.G. Hierarchical reinforcement learning with the MAXQ value function decomposition // *Journal of Artificial Intelligence Research*, 13. – 2000. – pp. 227–303.
 14. Verma, A.; Murali, V.; Singh, R.; Kohli, P.; and Chaudhuri, S. 2018. Programmatically interpretable reinforcement learning // In 35th International Conference on Machine Learning, ICML 2018, volume 11.
 15. Juang, C. F.; Lin, J. Y.; and Lin, C. T. 2000. Genetic reinforcement learning through symbiotic evolution for fuzzy controller design // *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 30(2).
 16. Vityaev E.E., Demin A.V., Kolonin Y.A. Logical probabilistic biologically inspired cognitive architecture // *Artificial General Intelligence - 13th International Conference, AGI 2020, Proceedings. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. – Springer Gabler, 2020. – V. 12177 LNAI. – p. 337-346.
 17. Vityaev E.E., Demin A.V. Cognitive architecture based on the functional systems theory // *Procedia Computer Science*. – Elsevier, 2018. – V. 145. – p. 623-628.
 18. Vityaev E.E., Demin A.V. Recursive subgoals discovery based on the Functional Systems Theory // *Biologically Inspired Cognitive Architectures 2011*, IOS Press, 2011. – p. 425-430.
 19. Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления // *Нейроинформатика*. – 2008. – Т. 3. – № 1. – С. 79-107.

20. Демин А.В. // Глубокое обучение адаптивных систем управления на основе логико-вероятностного подхода. – Известия Иркутского государственного университета. Серия «Математика» – Иркутск, 2021. – Т. 38. – С. 65-83.
21. Carl Lagoze and Herbert Van der Sompel, The Open Archives Initiative: Building a low-barrier interoperability framework // <https://www.openarchives.org/documents/jcdl2001-oai.pdf>
22. Open Archives Initiative Organization <https://www.openarchives.org/organization/>
23. Parinov S., International Professional Association of Research Information System Specialists euroCRIS and its Main Product CERIF // http://ceur-ws.org/Vol-1297/6-9_paper-2.pdf
24. Прожито <https://prozhito.org/> (дата обращения 04.04.2022).
25. Савелова О.А., Крайнева И.А. Электронный архив сестер Карповых: история научной династии // Исторический курьер. 2021. № 2 (16). С. 25–35. ВАК.
26. Крайнева И.А. Исидор Борисович Румер: страницы биографии // Диалог со временем. 2021. № 74. С. 140-155. Scopus/WoS
27. Крайнева И.А. Академик Будкер: «...будоражить умы и кресла» // История науки и техники. 2021. №3. С. 20-31. ВАК.
28. Крайнева И.А. Ржановы под Ленинградом: выжить, защищая город // Сборник 18-х международных чтений «Право на имя: Биографика XX века» памяти Вениамина Иофе. Санкт-Петербург, 2021. С. 43-54.
29. Krajevna I., Savelova O. The female face of programming (mid 1950s – early 21st century) / HISTELCON-21. Moscow, Nov. 10-12, 2021.
30. Крайнева И.А., Сэнборн К., Меристе М. Энн Тыгуу: история эстонского программиста / "Право на имя: Биографика 20 века" 19-е чтения памяти Вениамина Иофе. Санкт-Петербург, 20-22 апреля 2021.
31. Городняя Л.В. О неявной мультипарадигмальности параллельного программирования // Научный сервис в сети Интернет: труды XXIII Всероссийской научной конференции (20-23 сентября 2021 г.). — М.: ИПМ им. М.В.Келдыша, 2021. — С. 104-116.
32. [Gorodnyaya L.V. The Role of Functional Programming in the Organization of Parallel Computing](https://dl.acm.org/conference/ics) [CEUR Workshop Proceedings ICS '19: Proceedings of the ACM International Conference on Supercomputing](https://dl.acm.org/conference/ics) <https://dl.acm.org/conference/ics>
33. https://online.isprasopen.ru/docs/ISPRAS_Svace.pdf
34. <https://cs.brown.edu/~sk/Publications/Papers/Published/kle-next-700-semantics/paper.pdf>
35. 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. <http://plc.sfedu.ru/files/PLC-2017-proceedings.pdf>
36. <http://keldysh.ru/abrau/2020/> - Доклады, представленные на XIX Всероссийскую научную конференцию «Научный сервис в сети Интернет», 18-23 сентября 2020 года.

37. <https://www.sorusom.org/> - 5-я международная конференция «Развитие вычислительной техники в России, странах бывшего СССР и СЭВ (SORUCOM 2020)», 6–8 октября 2020 г. в НИУ ВШЭ, Москва.
38. <https://russianscdays.org/> - Суперкомпьютерные дни в России, 21–22 сентября 2020 г., Москва, <http://2020.nscf.ru/> - Национальный Суперкомпьютерный Форум (НСКФ-2020). 24-27 ноября 2020, Переславль-Залесский, ИПС имени А.К. Айламазяна РАН
39. Городня Л.В. Парадигмальный подход к факторизации определений языков и систем программирования // Системная информатика (System Informatics), No. 12 (2018), с. 1- 26. https://system-informatics.ru/files/issue_12_full.pdf (Рукопись поступила в редакцию 10.08.2018)
40. Городня Л.В. Методы декомпозиции программ // Препринт 182. Новосибирск. 2018. 27 с.3.2
41. Городня, Л. В. Парадигма программирования : учебное пособие для вузов / Л. В. Городня. - 2-е изд., стер. - Санкт-Петербург : Лань, 2021. - 232 с. - ISBN 978-5-8114-6680-1 : УДК 004.43 ББК 32.973-18я73
42. М.М. Лаврентьев, Л.В. Городня, М.А. Держо, Н.А. Иванчева, Д.В. Иртегов, Д.С. Мигинский, Б.Н. Пищик Вопросы мотивации обучения сложным профессиям // Вестник НГУ. Серия: Информационные технологии. 2021 Т.19, №1. С. 80–92. DOI: DOI 10.25205/1818-7900-2021-19-1-80-92
43. Л. В. Городня Визуализация результатов анализа языков программирования для их поверхностного сравнения // Вестник НГУ. Серия: Информационные технологии. 2021 Т.19, №2. С. 29–52. DOI: 10.25205/1818-7900-2021-19-2-29-52
44. Бешапошников Н.О., Дьяченко М.С., Леонов А.Г., Матюшин М.А., Орловский А.Е. Использование машинного обучения и нейронных сетей для автоматической верификации заданий в текстовом и графическом представлении и помощи преподавателю. // Успехи кибернетики. 2020. Т. 1. № 2. С. 39-45.
45. Димитриенко Ю.И., Губарева Е.А., Зубарев К.М., Алесин А.В., Иванова Т.Л. Автоматизация проверки математических заданий по курсу «Аналитическая геометрия» в системе NOMOTEX. // В сборнике: Цифровые технологии в инженерном образовании: новые тренды и опыт внедрения. Сборник трудов Международного форума. 2020. С. 206-208.
46. Кацман В.И., Козлов И.А., Новиков Ф.А. Игрофикация процесса решения типовых учебных задач на основе выбора правил преобразования. // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2020. № 9. С. 63-68.
47. Новиков Ф.А., Кацман В.И. Автоматическая проверка решений учебных задач на основе комбинации методов перебора логических правил и тестирования. // В сборнике: Цифровые технологии в инженерном образовании: новые тренды и опыт внедрения. Сборник трудов Международного форума. 2020. С. 266-273.
48. Тузов А.А. Автоматизированный практикум по решению вычислительных задач в среде «Кумир». // В сборнике: Проблемы и перспективы технологического образования в России и за

- рубежом. Сборник материалов III Международной научно-практической конференции. Отв. редактор Л.В. Козуб. Ишим, 2021. С. 50-53.
49. Тузов А.А. Практикум с автоматической проверкой решения задач для исполнителя робот системы кумир (расширяем круг задач). // В сборнике: Современный учитель дисциплин естественнонаучного цикла. Сборник материалов Международной научно-практической конференции. Ответственный редактор Т.С. Мамонтова. 2019. С. 184-186.
 50. Тихонова Т.И. Организация и проведение Летней школы юных программистов // «Сибирский учитель». Новосибирск, 2021. – № 1 (134). С. 83-88.
 51. Виктор Майер-Шенбергер, Кеннет Кукьер. Большие данные. Революция, которая изменит то, как мы живём, работаем и мыслим = Big Data. A Revolution That Will Transform How We Live, Work, and Think / пер. с англ. Инны Гайдюк. — М.: Манн, Иванов, Фербер, 2014. — 240 с. — ISBN 987-5-91657-936-9.
 52. Марчук А.Г. Архитектура и основные особенности библиотеки PolarDB работы со структурированными данными // Системная информатика, № 13, 2018. Стр. 25-34
 53. Min Chen, Shiwen Mao, Yin Zhang, Victor C.M. Leung. Big Data. Related Technologies, Challenges, and Future Prospects. — Springer, 2014. — 100 p. — ISBN 978-3-319-06244-0. — doi:10.1007/978-3-319-06245-7.
 54. Bankevich A., Nurk S. SPAdes: A New Genome Assembly Algorithm and Its Applications to SingleCell Sequencing // Journal of computational biology : a journal of computational molecular cell biology, □19. 2012. P. 455-477
 55. Марчук А.Г., Трошков С.Н. Некоторые эксперименты по построению и анализу графа Де Брёйна // Системная информатика, No. 16 (2020), стр. 47- 56
 56. Марчук А.Г. Последовательность как абстракция структурированного построения баз данных // Системная информатика, № 18 (2021), стр. 35-52, <https://system-informatics.ru/ru/article/284>
 57. Марчук, А.Г., Трошков С.Н, Крайнева И.А. К вопросу о модернизации и интеграции электронных архивов длительного срока жизни // Сб. конференции Научный сервис в сети Интернет. 2021. № 23. С. 214-227.

