
ТОЧКА
ЗРЕНИЯ

КАК РЕОРГАНИЗОВАТЬ ВЫЧИСЛИТЕЛЬНЫЕ НАУКИ И ТЕХНОЛОГИИ?

© 2019 г. В.П. Ильин

Институт вычислительной математики и математической геофизики СО РАН, Новосибирск, Россия
Новосибирский государственный университет, Новосибирск, Россия
Новосибирский государственный технический университет, Новосибирск, Россия
E-mail: ilin@sscc.ru

Поступила в редакцию 22.10.2018 г.

Поступила после доработки 22.10.2018 г.

Принята к публикации 22.10.2018 г.

Основы современного прогресса зиждутся на беспредельных темпах роста вычислительно-информационных мощностей постпетафлопсных масштабов, предоставляющих немислимые ещё вчера возможности получения новых фундаментальных знаний и внедрения инноваций в производственные, природопользовательские, экономические и социальные сферы. Ключевую роль здесь обретает предсказательное математическое моделирование процессов и явлений, включающее теоретические и прикладные исследования, супервычисления и крупномасштабные компьютерные эксперименты, операции со сверхбольшими данными и искусственный интеллект. Решение возникающих сверхзадач требует огромных объёмов интеллектуального программного обеспечения нового поколения, невозможных без формирования концепций и парадигм деятельности, архитектурных решений и конструктивных технологий. Мировые тренды ведут к созданию интегрированных вычислительно-информационных окружений, образующих инструментальную среду для автоматизации построения моделей и алгоритмов, их отображения на платформы суперЭВМ и создания комфортных интерфейсов для пользователей с разной профессиональной подготовкой. Такая экосистема рассчитана на длительный жизненный цикл с непрерывным развитием и согласованным участием различных групп разработчиков. Фактически она приводит к образованию новой отрасли с массовыми профессиями высокой суперкомпьютерной грамотности. Обозначенные научно-технологические вызовы требуют компетентных организационных и инфраструктурных решений, включающих межведомственную координацию академических, образовательных и производственных коллективов.

Ключевые слова: супервычисления, междисциплинарные задачи, математическое моделирование, вычислительные методы и технологии, программное обеспечение, искусственный интеллект, интегрированное инструментальное окружение, суперкомпьютерный эксперимент.

DOI: <https://doi.org/10.31857/S0869-5873893232-242>

Китайская мудрость гласит: не дай вам Бог жить в эпоху перемен. Однако исторический опыт свидетельствует, что такая сложная динамическая система, как мировая цивилизация,

не ограничивается только эволюцией, но обязательно проходит через периоды революционных преобразований. Подобная амбивалентность иллюстрируется русской поговоркой "Под лежащий камень вода не течёт" и библейским изречением "Время разбрасывать камни, и время собирать камни". Революция, как правило, связана с разрушением старого уклада, а созидание нового — это уже путь эволюционного развития, не терпящий суеты. Хотя в каждом правиле есть исключения.

В нашей стране сегодня фактически официально провозглашено наступление судьбоносного для страны этапа. За шесть лет надо построить цифровую экономику, качественно поднять производительность труда и ВВП, кардинально укрепить национальную безопасность [1]. В этом же



ИЛЬИН Валерий Павлович — доктор физико-математических наук, главный научный сотрудник ИВМиМГ СО РАН, профессор НГУ и НГТУ.

направлении стремительно движутся и остальные передовые государства, а футурологи предрекают человечеству эру "технологической сингулярности" — эру неограниченных энергетических ресурсов, новых материалов с трудновообразимыми сегодня фантастическими свойствами и биомедицинских технологий, дарующих вечную молодость и бессмертие [2]. Наблюдая невиданные темпы технологического прогресса, являющегося следствием новых научных достижений, уместно вспомнить слова великого Леонардо да Винчи, полагавшего, что тот, кто любит практику без теории, подобен моряку, который садится на корабль без руля и компаса, а потому никогда не знает, куда его может занести.

При всей гармоничности развития различных фундаментальных направлений в XXI в. очевидным оказывается значение математики как движущей силы этого развития. Её роль кардинально возросла после массового появления персональных компьютеров, а затем в связи с рождением многопроцессорных супер-ЭВМ терафлопсного и петафлопсного уровня (10^{12} – 10^{15} операций в секунду). Скорость увеличения компьютерных мощностей более 50 лет подчиняется закону Г. Мура, одного из основателей компании "Интел", — экспоненциальный рост в 1000 раз за 11 лет! Хотя по физическим законам уже должно было наступить насыщение, аномальная динамика сохраняется. В настоящее время рекордный суперкомпьютер имеет около 200 петафлопс, а появление эксафлопсников (10^{18} арифметических действий в секунду) ожидается около 2020 г. Такая многопроцессорная вычислительная система (МВС) будет иметь сотни миллионов расчётных устройств, и поколение, вступающее сейчас в трудовую жизнь, получит доступ к супервычислениям принципиально новых масштабов. Возможность хранения и обработки огромных объёмов числовых данных приведёт к потрясениям, сравнимым по резонансу с последствиями двух предыдущих информационных взрывов — появлением сначала сети Интернет, а затем мобильного телефона, кардинально изменивших профессиональную деятельность и быт человека. Массовый удалённый доступ к суперкомпьютерным центрам с "облачными" и сетевыми технологиями в самой ближайшей перспективе сделает обыденным фактом высокопроизводительное моделирование сложнейших процессов и явлений во всех сферах — от фундаментальных исследований до производства и социальных услуг. Отметим, что мы говорим только о традиционных для последних двух десятилетий кластерных архитектурах, однако на подходе принципиально новые компьютеры на программируемых логических интегральных схемах (ПЛИС), на квантовых и других физиче-

ских эффектах, обсуждение которых мы оставляем на будущее.

С появлением суперкомпьютеров возникают новые суперзадачи: наше время отличается фантастическим прогрессом всех наук и технологий и работой с физико-математическими задачами такой сложности, что ещё десятилетие назад их изучение было возможным только при внесении в модели значительных упрощающих предположений. В определённом смысле теперь физику-теоретика или прикладному математику требуется преодолеть психологический барьер, чтобы понять, какого уровня проблемы становятся подвластны современным вычислительно-информационным методологиям.

Надо отдавать себе отчёт в том, что развитие вычислительных и информационных технологий — процесс необратимый, как приход цунами на берег, если оно уже зародилось даже за тысячи миль в глубинах океана. К счастью, волна компьютеризации — созидаящая, хотя какие-то атавистические производственные и общественные отношения неизбежно нарушатся. Главное — быть профессионально и этически готовыми к предстоящим изменениям, чтобы обратить их во благо человека и общества. Исторические уроки демонстрируют (но не всегда учат!), что технические открытия — обоюдоострое оружие, и неслучайно кибертерроризм стал одной из самых обсуждаемых тем на политических саммитах.

Передовые страны уже вступили на путь четвёртой промышленной революции, характеризующейся широким внедрением цифровой экономики, аддитивных технологий, 3D-принтеров, виртуальной реальности и виртуальных двойников, а также другими атрибутами постиндустриального общества. Задача стремящегося к современному развитию государства — вскочить на подножку уходящего экспресса, чтобы не отстать навсегда. Далее мы постараемся показать, какие задачи в связи с этим встают перед отечественной наукой, прежде всего математикой.

СТРУКТУРА МАТЕМАТИЧЕСКОГО ЗНАНИЯ И МЕСТО В НЕЙ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Для начала необходимо определиться с таким ёмким понятием, как *математика*, поскольку его содержание вызывает нешуточные дискуссии. Один из основоположников математики XXI в. академик В. И. Арнольд отождествлял теоретическую математику и теоретическую физику, более того, горячо боролся за их единство, воспринимая нарастающий методологический и образовательный разрыв этих наук как историческую трагедию. При этом он фактически не признавал

вычислительной математики и информатики, называя их в своих полемических работах "ремесленничеством" [3]. Принципиальным заочным оппонентом В.И. Арнольда был академик-программист А.П. Ершов, безусловный мировой авторитет на заре компьютеризации в 1960–1980-е годы, создатель уникальной Сибирской школы информатики, автор таких ставших азбучными терминов, как "компьютерная грамотность" и "школьная информатика", который с гордостью говорил о себе: "Я – математик". Его заслуги в области математики были по достоинству отмечены присуждением премии им. П.Л. Чебышева АН СССР за открытие и разработку смешанных вычислений. В подтверждение неразрывности теоретической и прикладной математики можно вспомнить великих И. Ньютона, Л. Эйлера и К. Гаусса, чьи имена носят популярные и сейчас численные алгоритмы.

Мы будем понимать под математикой *многоликую структуру*, включающую теоретическую и прикладную математику, вычислительную математику и математическое моделирование, теоретическое, системное и прикладное программирование, а также математические основы искусственного интеллекта, работы с данными, технологии знаний (big data, deep knowledge) и компьютерных архитектур. Деление на перечисленные составляющие полностью соответствует самому общему определению математики как науки о математических объектах. Конечно, полное право на существование имеют понятия "прикладная математика", "вычислительная информатика", "вычислительная геометрия и топология" и распространённое в англоязычной литературе словосочетание "computer science", дословный перевод которого "вычислительные науки" плохо приживается в отечественной профессиональной терминологии. Что касается таких направлений, как вычислительная физика, вычислительная химия, вычислительная биология и др., то их следует классифицировать в качестве применения математических методов и технологий в соответствующих предметных областях.

Среди обозначенного разнообразия дисциплин на первое место в последние десятилетия, безусловно, выходит *математическое моделирование*, которое можно в общем виде определить как *изучение процессов и явлений математическими методами*. Каждая из используемых здесь категорий требует своего содержательного раскрытия. Процессы и явления подразумеваются в первую очередь реальные — технические, естественные или социальные, например, производство металлов и новых материалов, природные или техногенные катастрофы, демографическая динамика. Однако актуальным является также изучение аб-

страктных моделей, описывающих чисто математические объекты. Более того, в действительности так и происходит: физическая (химическая, биологическая и т.д.) модель представляется какой-то математической моделью, которая практически всегда содержит пренебрежение малозначительными эффектами. И в её грамотном построении заключается задача моделлера (англ. "modeller") — специалиста, занимающегося моделированием в конкретной прикладной области.

Принципиально, что все рассматриваемые нами дисциплинарные направления составляют единую математику: теоретические исследования — база для вычислительных алгоритмов и технологий, реализуемых в программном обеспечении суперкомпьютеров, на которых проводятся расчётные эксперименты по решению фундаментальных и прикладных задач. Важно, кроме того, что вся эта картина очень быстро и взаимосвязанно изменяется, причём симбиоз фундаментальных исследований и практики представляет собой дорогу с двусторонним движением. Фундаментальные результаты становятся источником прикладных изысканий, но одновременно информационные технологии, включая искусственный интеллект, выступают эффективным средством для формульных преобразований, изучения функциональных свойств и качественного анализа абстрактных объектов с их наглядным отображением, что совершенно меняет характер работы исследователя-теоретика.

Ещё более кардинально меняется работа специалистов из прикладных областей, будь то учёный-химик, геофизик или инженер, занимающийся разработкой или эксплуатацией какого-либо устройства. Удобное средство для оперативного моделирования ситуации может также играть роль навигатора для автомобилиста или автопилота для лётчика. Конечно, в сложных случаях можно сделать заказ на анализ и прогнозирование процесса или явления (так называемый аутсорсинг), но это сути дела не меняет — всегда необходим эксперт, профессионально владеющий методологией и практикой моделирования.

Миссия моделирования — быть главным проводником фундаментальных идей и технологий в естественные и технические науки, в производство, природопользование и гуманитарные сферы. В определённом смысле математические методы исследования играют в науке ту же роль, что лимфатическая, кровеносная и нервная системы, питающие и поддерживающие все жизненно важные органы человека. Всевозможные состояния вещества описываются конечным набором типов дифференциальных и/или интегральных соотношений. Напряжённо-деформированные состояния твёрдых тел, гидро-газодинамические

течения, электромагнитные поля, в том числе с учётом фазовых переходов, химические взаимодействия, релятивистские и квантовые эффекты — все они подчиняются условиям сохранения массы, импульса, энергии и других субстанций, которые дают нам ключи сначала к познанию, а затем, если повезёт, и к управлению процессами и явлениями.

В 1930-е годы был популярен лозунг “Мы не должны ждать милостей от природы, взять их у неё — наша задача!”. Хотя ещё математик и философ Р. Декарт в XVIII в. пытался найти абсолютный метод постижения всех закономерностей, дело обстоит не так просто. Проникая в тайны Вселенной, мы не знаем даже строения ядра своей планеты. Дьявол кроется в деталях, а именно — у нас далеко не всегда имеется необходимая информация. Если для самой сложной, так называемой прямой, многомерной задачи известны начальные и граничные условия, обеспечивающие существование единственного решения, то мы, по-видимому, сможем найти его с требуемой точностью, прибегнув к современным вычислительным методам и технологиям. Гораздо сложнее решаются обратные задачи, у которых часть исходных данных или представлена параметрически, или совсем отсутствует. При этом решение надо определить по некоторым дополнительным косвенным измерениям, как правило, доступным с некоторой ошибкой. Говоря математическим языком, на искомом решении формулируется целевой функционал, который требуется минимизировать в пространстве параметров. Именно такие обратные задачи представляют наибольший интерес, так как они связаны с актуальными проблемами либо идентификации параметров модели, либо оптимизации производственного процесса, либо проектирования какого-то устройства с заданными характеристиками. Актуальной обратной задачей является, например, поиск полезных ископаемых по данным геологоразведки.

Частым “отягчающим” фактором оказывается междисциплинарность решаемых задач, предполагающая изучение взаимодействия процессов различной природы, описываемых системами уравнений с большим числом неизвестных функций, что значительно увеличивает ресурсоёмкость проблемы. Необходимо учитывать, что современные требования вынуждают ещё и обеспечивать высокую точность (или разрешение) моделирования. Всё это порождает экстремальные междисциплинарные суперзадачи, которые были не под силу самым мощным ЭВМ прошлого столетия. Яркий пример такой супер-задачи — моделирование ядерного взрыва, когда вещество претерпевает всевозможные фазовые переходы, а рассчитывать необходимо и электро-магнитные поля, и гидрогазодинами-

ку, и напряжённо-деформированное состояние твёрдого тела, и радиационный теплоперенос, и многие другие процессы.

Сложность насущных суперзадач — только одна причина стремительного наращивания мощностей суперкомпьютеров. Другая причина — это невообразимая многоликость всевозможных проблем моделирования. Здесь на помощь приходит системный подход, позволяющий посмотреть на ситуацию с высоты современных классификационных методологий. Разбиение множества задач моделирования можно провести по различным типам спецификаций, то есть в разных координатных направлениях. По отраслевым применениям, например, различаются энергетика, машиностроение, металлургия, транспорт, химические производства, сельское хозяйство, медицина, экономика и т.д., причём каждый из сегментов может быть разделён на достаточно автономные части. Приложений можно назвать очень много, но на абстрактном уровне все перечисленные проблемы описываются ограниченным набором типов уравнений: система Максвелла — для электромагнетизма, Навье—Стокса — для гидродинамики, Ламе — для упругопластичности, Больцмана — для кинетики. Конечно, этот список не исчерпывает всех случаев, но и другие уравнения могут быть классифицированы по дифференциально-интегральным характеристикам, а также с помощью подходов теории групп.

Если подразумевать под моделированием решение задач до логического конца — до понятных пользователю результатов, то в квалификацию методов моделирования необходимо добавить многообразие вычислительных методов и компьютерных технологий. Но об этом будет сказано несколько позже. Сейчас же можно констатировать, что методология моделирования является инструментом исследования практически всех прикладных задач. Разумеется, для каких-то постановок готовых решений пока не найдено, но это предмет продолжающихся фундаментальных исследований. Более важно другое: имеется огромное количество реальных приложений, для которых все главные научные вопросы закрыты, но не доведены до внедрения, то есть до инноваций с большим потенциальным экономическим эффектом. А чтобы поставить точку над *i*, необходимо развить математические методологии до уровня рабочей технологии моделирования. Например, достаточно представить, какого огромного экономического эффекта можно было бы достичь в строительстве за счёт компьютерной оптимизации теплосетей и конструкции зданий с учётом разнообразных региональных климатических условий России.

ТРИ ИСТОЧНИКА И ТРИ СОСТАВЛЯЮЩИЕ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Чтобы представить функциональность и структуру математического моделирования в целом, надо сначала охарактеризовать вкратце обе стороны медали — математическую и компьютерную. Концепция крупномасштабного *вычислительного эксперимента* была заложена в научных школах Г.И. Марчука, А.А. Самарского и Н.Н. Яненко. За последующие десятилетия она, безусловно, претерпела изменения, особенно в связи с появлением многопроцессорных вычислительных систем кластерного типа с распределённой и иерархической общей памятью, архитектура которых сейчас в главных своих чертах устоялась. Хотя в целом структура ЭВМ была и будет состоять из трёх основных компонентов — арифметических процессоров, устройств памяти и средств коммуникаций, их непрерывная эволюция привела к переходу количества в качество, и сейчас компьютерное сообщество стоит на пороге смены вычислительных информационных методов и технологий. Огромное число расчётных узлов и ядер — основание парадигмы масштабируемого параллелизма. Одновременно с ростом производительности увеличивается и объём оперативной памяти (исторически соблюдается баланс — машина с быстродействием терафлопс или петафлопс имеет память размером примерно терабайт или петабайт), которая к тому же имеет много уровней с различными временами доступа. К этому нужно добавить гетерогенность суперкомпьютеров, имеющих графические ускорители со сверхбыстрой арифметикой, но с более ограниченной логикой.

Из представленного наброска архитектуры супер-ЭВМ наглядно видны две главные составляющие математического моделирования. Первая — высокопроизводительные вычисления, по скорости которых изначально характеризуется класс машины. Распараллеливание алгоритмов, являющееся здесь основой основ, — это и наука, и технологии, и искусство (а также терпение) вычислителя-программиста. Успех измеряется реальным быстродействием исполняемого кода в сравнении с пиковой, или теоретической, производительностью. Средним достижением обычно считается цифра около 20%, её увеличение будет означать осуществление голубой мечты менеджеров сделать науку более экономной.

Вторая вычислительная часть моделирования заключается в хранении, обработке и передаче данных. Ресурсоёмкость больших задач, как правило, сбалансированно характеризуется не только

количеством арифметических операций, но и размером используемых числовых массивов. Такой класс проблем, к которым относятся и наукоёмкие задачи моделирования, в англоязычной литературе получил название "data intensive computing" (большие данные, большие вычисления). Отметим, что наш случай — это не специфические информационные проблемы типа "big data", которые возникают в сейсмической геологоразведке, космическом мониторинге или поиске в мировой сети Интернет. Однако в многомерных задачах моделирования стремление к повышению производительности — это борьба с коммуникационными потерями, которые не только резко замедляют расчёты, но обладают ещё и таким отрицательным свойством, как огромная энергозатратность. А потребление электроэнергии суперкомпьютером уже становится существенным фактором его эксплуатационных расходов.

Третья составляющая часть моделирования, не столь очевидная, поскольку актуализируется на достаточно высоком уровне вычислительно-информационных технологий, — это искусственный интеллект. На математическом языке данная проблема была сформулирована Г.И. Марчуком в 1980-е годы: автоматизация построения алгоритмов и их отображение на архитектуру ЭВМ. При решении математических задач неизбежно возникает необходимость выполнения аналитических формульных преобразований, теоретико-множественных и других операций, не связанных с арифметическими действиями. Первые шаги в компьютеризации такой деятельности делались лет 50 назад, а затем сформировалось самостоятельное направление программирования на основе когнитивных принципов и онтологий. Успехи искусственного интеллекта налицо в машинном переводе текстов, компьютерных играх, распознавании объектов и других областях.

Если говорить об эффективной системе моделирования для широкого класса задач, то она должна иметь базу знаний по математическим постановкам и алгоритмам. Требуются языки естественного типа для формулировок, средства их анализа, каталогизированный архив типовых примеров, инструменты для выбора соответствующих методов и т.д. Фактически в данном случае речь идёт о построении дружественного интерфейса нового поколения с пользователем-математиком.

На рисунке 1 представлена схема, отражающая в общих чертах структуру и логику системы моделирования. Её верхний уровень — "пользовательские приложения" — фактически объединяет множество интерфейсов для различных категорий моделлеров. Это и есть интегрированное окружение, или оболочка, системы моделирова-

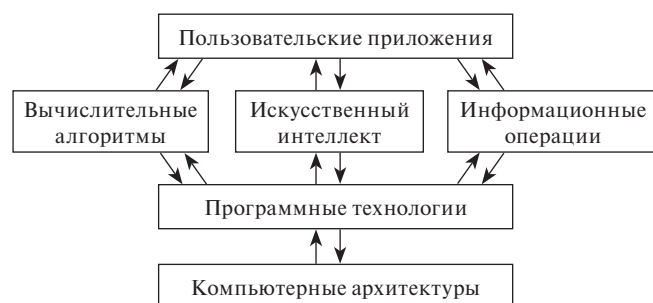


Рис. 1. Структура системы моделирования

ния для внешнего мира. Нижний уровень — это компьютерное оборудование (hardware) вместе со своей системной "платформой", которое может быть реализовано в виде вычислительного центра коллективного пользования (ВЦКП). Такой вычислительный центр включает "облачные" и сетевые технологии, определяющие расчётные мощности всего конгломерата. Сегмент "программные технологии" (software) — ключевой по количеству человеко-лет, которые необходимо затратить на реализацию всего проекта. Принципиальный момент заключается в том, что этот продукт должен разрабатываться профессиональными коллективами программистов, взаимодействующими по согласованным техническим заданиям. Кроме того, важно, что данное программное обеспечение должно иметь производственное сопровождение и поддержку, а также средства для постоянного развития и тестирования кодов на различных технических устройствах. Последний из сегментов является главным и состоит из трёх источников (и одновременно составных частей) математического моделирования: вычислительные алгоритмы, средства работы с данными, инструменты искусственного интеллекта. В терминологии А.А. Дородницына, долгие годы возглавлявшего делегации СССР по вычислительной математике и информатике на международных форумах, данный уровень можно назвать brainware, поскольку он фактически представляет фундаментальное ядро всей системы моделирования.

КОНЦЕПЦИЯ ИНТЕГРИРОВАННОГО ВЫЧИСЛИТЕЛЬНОГО ОКРУЖЕНИЯ

Разработка наукоёмкого программного обеспечения для решения задач математического моделирования имеет многолетнюю историю, и на мировом рынке накопилось огромное количество продуктов коммерческого и академического характера, представляющих неоценимый интеллектуальный потенциал. Однако вступление передовых стран в эпоху постиндустриализации, связанное с массовым появлением многопроцес-

сорных суперкомпьютеров и умножением количества актуальных сверхбольших задач, привело к всеобщему кризису программирования. Ситуация обостряется тем, что рост производительности труда программиста катастрофически отстаёт от стремительного увеличения производительности ЭВМ. А глобализация моделирования с приходом новых пользователей фактически вынуждает переходить на новую парадигму программирования.

Прикладное программное обеспечение (ППО) с организационной точки зрения делится на общедоступное, или открытое (open source), и на закрытое, или коммерческое. Традиционная классификация по функционально-методологическим признакам определяет три основных типа ППО.

- *Пакеты прикладных программ (ППП)* — это проблемно-ориентированное обеспечение конкретного класса (или классов) практических задач, включающее инструментальную поддержку основных стадий моделирования и относительно слабозависящее от внешних программных продуктов. Примеры высокоразвитых и широко распространённых ППП — NASTRAN, ANSYS, FreeFEM, NGSolve, FlowVision. Другое используемое название — прикладные программные комплексы (ППК), примерами которых являются ЛОГОС и НИМФА, разработанные во ВНИИЭФ (г. Саров).

- *Библиотеки алгоритмов и программ* — методо-ориентированное обеспечение, содержащее достаточно представительный набор методов для решения определённого математического класса задач. К такому обеспечению относятся "всеобъемлющая" библиотека NETLIB, библиотеки алгебраических методов MKL INTEL, HYPRE, PETSc, EIGEN, SPARSE KIT и т. д.

- *Технологические инструментари* по своей структуре могут также быть библиотечного типа, но они больше связаны с информационным обеспечением соответствующих вычислительных процессов. В качестве примеров назовём генераторы сеток NETGEN, GMESH, TETGEN, графические пакеты PARAVIEW, VISUAL STUDIO, инструменты для преобразования графов METIS, PARMETIS, SCOTCH и т. д.

В классификацию ППО не включается системное программное обеспечение (СПО) общего назначения, а также такие самодостаточные продукты, как САПРовские разработки (CAD, CAM, CAE, PLM), имеющие свой традиционный рынок, но проявляющие тенденции интеграции со средствами математического моделирования. Примерами отечественных (или полуотечественных, то есть разрабатываемых с использованием "чужих" компонент) систем такого рода являются TFLEX, COMPAS и ГЕРБАРИЙ.

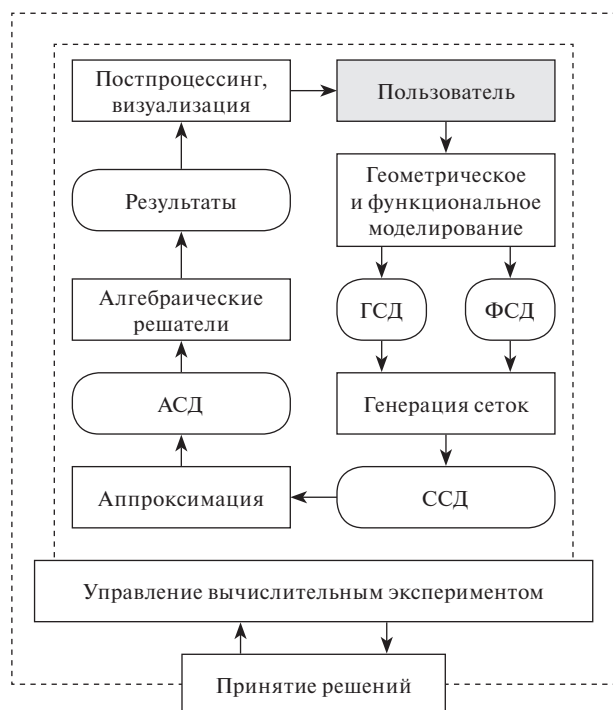


Рис. 2. Структура функциональных и информационных компонентов ядра базовой системы моделирования

ГСД, ФСД, ССД, и АСД обозначают геометрическую, функциональную, сеточную и алгебраическую структуры данных соответственно

Тенденция последних десятилетий — переход к концепции интегрированных вычислительных окружений (ИВО), то есть к созданию открытой инструментальной среды, ориентированной на автоматизированное построение вычислительных моделей, алгоритмов и технологий, классифицированных по всем основным стадиям моделирования широкого круга процессов и явлений. К таким проектам относятся OPEN FOAM, DUNE (Distributed Unified Numerical Environment), MATLAB (коммерческий продукт университетской направленности, имеющий свою открытую версию OCTAVE), INMOST (разработка Института вычислительной математики им. Г.И. Марчука РАН), а также создаваемая в Институте вычислительной математики и математической геофизики СО РАН базовая система моделирования (БСМ), рассчитанная на комплексную поддержку всех технологических этапов моделирования.

БСМ предназначена как для разработчиков ППО, так и для конечных пользователей с разнообразной профессиональной подготовкой. Концепция ИВО двухцелевая: с одной стороны, это комплекс средств для качественного повышения эффективности труда разработчиков математического и программного обеспечения, с другой — инструментальный ящик для оперативной

сборки и эксплуатации высокопроизводительных конфигураций ППО из готовых блоков функционального ядра БСМ для конкретных приложений по аналогии с интеллектуальным конструктором LEGO.

Архитектурно БСМ представляет собой совокупность достаточно автономных методоориентированных подсистем, каждая из которых отвечает за свою функциональную стадию и связана с остальными блоками только посредством согласованных структур данных. Такая схема воплощает формулу авторитетного специалиста в области информатики и программирования Н. Вирта: "программа = алгоритмы + структуры данных".

Все технологические стадии наукоёмкого машинного эксперимента состоят из следующего ограниченного набора: геометрическое и функциональное моделирование (описание и модификация расчётной области, решаемых уравнений и дополнительных условий, идентификация и анализ свойств задачи), построение сеток (дискретизация задачи, поддержка параллельных алгоритмов декомпозиции областей и много-сеточных подходов), аппроксимация исходных уравнений (методы конечных объёмов и конечных элементов различных порядков и т. д.), решение алгебраических задач (наиболее трудоёмкая часть, так как современные системы насчитывают до 10^{10} и более неизвестных), постобработка и визуализация численных результатов, методы оптимизации для решения обратных задач, управление вычислительным процессом, система принятия решения по итогам расчётов.

На рисунке 2 представлена структура функциональных и информационных компонентов (прямоугольники и овалы соответственно) ядра БСМ. Схема фактически описывает алгоритм решения прямых задач, а стадия решения обратных задач методами оптимизации (путём многократного решения прямых) для краткости включена в этап "управление вычислительным экспериментом". Каждый из восьми блоков — высокопроизводительная реализация отдельного содержательного направления вычислительной математики, и ни один из компонентов не может быть изъят из неразрывной технологической цепочки математического моделирования.

Главный архитектурный принцип построения БСМ заключается в том, что каждая из её подсистем — не просто некоторая библиотека программ, а интегрированное инструментальное окружение для соответствующих вычислительно-информационных технологий. Следовательно, и сама БСМ — это тоже интегрированное окружение.

Принципами построения ИВО (техническими требованиями) являются:

- гибкая расширяемость состава моделей решаемых междисциплинарных проблем, а также методов решения прямых и обратных задач, без программных ограничений на число степеней свободы и на количество используемых вычислительных процессоров и/или ядер;

- адаптируемость к эволюции компьютерных архитектур, предполагающая компонентные технологии обеспечения согласованности внутренних и внешних межмодульных интерфейсов;

- универсальность и конвертируемость структур данных, согласованных с имеющимися пространственными форматами, и поддерживающие возможности эффективного переиспользования внешних программных продуктов;

- многоязыковость и кроссплатформенность программного функционального наполнения, открытость к согласованному участию в проекте различных групп разработчиков;

- высокая производительность программного кода с масштабируемым параллелизмом на основе средств гибридного программирования на гетерогенных МВС с распределённой и иерархической общей памятью;

- наличие интеллектуальных пользовательских и внутренних интерфейсов с ориентацией на широкое применение в различных производственных сферах специалистами с разной профессиональной подготовкой.

Сформулированные требования призваны обеспечить длительный жизненный цикл и срок эксплуатации проекта, а также такие противоречащие друг другу качества, как эффективность, универсальность, высокая производительность и технологичность эксплуатации создаваемого наукоёмкого программного обеспечения. Данные требования предъявляются на всех этапах разработки, сопровождения и использования продуктов, что позволяет существенно увеличить эффективность труда разработчиков и массовую востребованность их результатов конечными пользователями.

Богатство закладываемых функциональных возможностей системы моделирования и высокие требования к производительности и качеству программирования требуют развитой системной поддержки разработки ядра БСМ, среди компонентов и методологий которой можно указать следующие:

- средства автоматизации для верификации, валидации и тестирования различных типов алгоритмов, в том числе в разных внешних окружениях;

- конфигурационное управление многоверсионных реализаций программных модулей и формирование ППП или их фрагментов для пользователей; важная частная проблема — создание учебных версий для подготовки кадров;

- управление вычислительным процессом и организация расчётных сеансов, в том числе в "облачных" технологиях;

- автоматизация построения и распараллеливания алгоритмов, их отображение на архитектуру ЭВМ, требующие детального знания и методов, и особенностей оборудования, а также умения эффективно использовать высокопроизводительные наукоёмкие инструментари — Sparse Blas, MAPLE, METIS и др.;

- конкретизация и поддержка стилей программирования — модульного, сборочного, фрагментарного, объектно-ориентированного, функционального и т.д.;

- компонентные технологии для поддержки многоязыковости и кроссплатформенности проекта — COM/DCOM, CCA (common component architecture); фабрики языков (SIDL — Scientific Interface Definition Language, DSL — Domain Specific Languages).

Немаловажное значение имеют и способы организации командной работы при создании больших программных проектов. Примерами служат методики SCRUM и AGILE. Самостоятельный вопрос при корпоративных разработках — удалённое взаимодействие различных групп, в том числе по принципу аутсорсинга — менеджмент с раздачей заданий различным исполнителям.

Цель перечисленных инструментальных подходов — кардинальное повышение производительности программирования, а также создание интегрированного математического и прикладного программного обеспечения, которое по уровню эксплуатационной технологичности было бы сопоставимо со штатной операционной системой или компилятором. Выполнение проекта в определённом смысле означает производство средств производства массовых наукоёмких программных продуктов. Говоря языком строительной индустрии, речь идёт о переходе от индивидуальной технологии построения зданий к воздвижению завода (или сети заводов) по созданию железобетонных изделий и других архитектурных конструкций. Важно подчеркнуть, что такая стратегия для вычислительно-информационных компонентов имеет большое преимущество в сравнении с материальным производством. Например, если в стоимость конкретного забиваемого гвоздя входит цена металла, из которого он изготовлен, и его транспортировки, то аналогичные расходы на копирование и пересылку программного модуля практически отсутствуют. Продолжая эту аналогию, можно выделить три производственных уровня, связанных с гвоздём: его непосредственное использование, производство гвоздя на станке и создание самого станка на соответствующем заводе. Так вот

создание ИВО относится к высшему уровню, то есть к строительству станкостроительного завода. Характеризуя интегрирующую роль данного проекта, следует отметить, что он сам базируется на системообразующем фундаменте, включая трёх научно-технологических “китов”: высокопроизводительные вычисления, структуры больших данных и искусственный интеллект.

ЧТО ДЕЛАТЬ?

Несколько десятилетий назад в нашей стране часто задавался сакраментальный вопрос: почему Индия может зарабатывать миллиарды долларов на продаже своего программного обеспечения, а Россия вынуждена тратить огромные суммы на его покупку или пользоваться “ворованными” продуктами? Чтобы найти решение этой и близких к ней проблем, нужно перейти от чисто научных задач к обсуждению организационных аспектов и даже коснуться проблематики научной дипломатии. Прежде всего сформулируем ряд положений, следующих из рассмотренного выше.

Во-первых, научно-технический прогресс в ближайшем будущем приведёт к значительному усилению разницы в экономическом уровне передовых и слаборазвитых стран. Сегодня решается, в какую группу попадёт наша страна.

Во-вторых, мощным катализатором четвёртой индустриальной революции выступает математическое моделирование на суперкомпьютерах.

В-третьих, растиражированные в зарубежных и отечественных средствах массовой информации названия “цифровая экономика”, “цифровые двойники”, “умные дома” и “Интернет вещей” могут ввести в заблуждение, будто бы инициированная суперкомпьютерами революция заключается исключительно в информационных технологиях, не подкреплённых фундаментальной наукой и математическим моделированием. Серьёзным предостережением (и уроком) можно назвать развернувшуюся в 1970-е годы по всему СССР пропагандистскую кампанию АСУнизации всей страны. Автоматизация и управление производством, дело огромной важности, пало жертвой чиновничьей некомпетентности и профанации, когда директор предприятия просто покупал компьютер и рапортовал в обкоме партии, что у него заработала АСУ.

В-четвёртых, предстоящие капитальные преобразования хозяйственной и общественной жизни невозможны без разработки огромных объёмов нового поколения математического программного обеспечения ЭВМ, стоимость создания которого не меньше расходов на техническое оборудование (сейчас цена петафлопсного

компьютера порядка 1 млрд руб.). Надо забыть об образе кабинетного учёного типа Г.Я. Перельмана, которому для работы ничего не нужно, кроме ручки и листа бумаги. Супервычисления и моделирование — серьёзная индустрия, требующая значительных вложений, но они, безусловно, многократно окупятся.

В-пятых, создание отечественного наукоёмкого программного обеспечения — приоритетная проблема национальной безопасности, особенно в связи с современной политической обстановкой в мире. Достаточно только представить гипотетическую ситуацию, когда Госдеп США в качестве очередных санкций запретит компаниям Intel, Microsoft и др. продавать России свою продукцию. Важен и другой момент: далеко не каждая страна способна создать полнокровную индустрию суперкомпьютерного математического и программного обеспечения. Но именно благодаря уровню российских научных школ наша страна это сделать может и должна.

Отправляясь от сформулированных пяти пунктов, отражающих текущую ситуацию в области научно-технического прогресса и неразрывно связанного с ним положения на мировой арене отдельных государств, можно обсудить возможные планы реорганизации управления развитием вычислительных наук и технологий в нашей стране. При этом очень полезно критически осмыслить опыт передовых стран, которые опережают нас по уровню развития данной отрасли.

Среди государственных инициатив в этом направлении отметим сформированную в 2013 г. в Германии Национальную программу SPPEXA по супервычислениям (Strategic Priority Program 1648 “Software for Exascale Computing” [4]), базирующуюся на консорциуме из более чем 40 институтов и трансформирующуюся сейчас в двух- и трёхсторонние междисциплинарные мегапроекты с партнёрами из Франции и Японии. Аналогичные стратегические инициативы имеются в США, Китае и ряде других стран.

Среди неправительственных инициатив особого внимания заслуживают IESP — международная программа по экзамасштабному программированию, SIAM — общество индустриальной и прикладной математики, выпускающее более 10 ведущих мировых инструментов по всем направлениям computer science, организующее крупнейшие конференции и другие мероприятия.

В нашей стране имеется достаточно большое количество государственных органов и иных организаций, де-юре причастных к управлению развитием науки и технологий. Помимо Российской академии наук это Министерство науки и высшего образования РФ, Министерство

цифрового развития, связи и массовых коммуникаций РФ, Совет по науке и образованию при Президенте РФ, Фонд перспективных исследований, Агентство стратегических исследований, Фонд развития информационных технологий РФ, Фонд "Сколково", Российская венчурная компания и др. Однако цельной национальной программы и единой инфраструктуры по программному обеспечению суперкомпьютерного моделирования в стране нет. Очевидно, что здесь нужна вертикаль, объединяющая и академическую науку, и образовательную систему специальной подготовки кадров, и организации-разработчики производственного программного продукта. К последним следует отнести отечественных производителей суперкомпьютерных систем (Т-платформы, РСК, ЭЛЬБРУС, АНГАРА и др.), заинтересованных в продвижении своих сервисов на рынок. Важной научно-организационной проблемой, наряду с созданием интегрированного вычислительного окружения, оказывается формирование инфраструктуры для сопровождения, эксплуатации и поддержки длительного жизненного цикла ИВО, включающего огромный объем программного обеспечения.

Международная практика показывает большую пользу от создания ассоциаций пользователей программного обеспечения. В них активное участие будут принимать представители государственных и частных корпораций (Ростех, Росатом, Роскосмос, Роснано, Роснефть, Газпром, Русал и др.), которые должны быть главными заказчиками и заинтересованными сторонами новых научно-технических инициатив.

Для организации работ представляется целесообразным создать Российский фонд алгоритмов и программ как усовершенствованный на современных принципах вариант существовавшего в советское время Государственного фонда алгоритмов и программ. В задачи такой структуры будет входить не только межведомственная координация разработок математического и программного обеспечения, но и широкое внедрение компьютерного моделирования в отраслевых сферах и организация обязательного международного сотрудничества.

Анализ назревших проблем фактически свидетельствует о необходимости формирования новой государственной научно-технологической политики, предполагающей, в частности, разработку и реализацию стратегической Программы фундаментальных исследований РАН, которая будет осуществляться в академических институтах, а координировать эти работы будут тематические отделения РАН. Данная структура российской науки исторически себя оправдала, а попытка раз-

рушить её в ходе продолжающегося реформирования научного сектора нарушает универсальный принцип Гиппократ "Не навреди!".

Принципиальный момент заключается в том, что бюджетное финансирование долгосрочных фундаментальных исследований РАН должно быть не меньше грантовых программ Российского научного фонда и Российского фонда фундаментальных исследований. При этом бюджет должен включать оплату внутренних командировок и участия в зарубежных конференциях, обеспечение иностранной научной литературой. Перефразируя известное изречение, можно утверждать, что государство, которое не может прокормить свою науку, будет кормить чужую науку. Наша страна тратит огромные средства на подготовку квалифицированных кадров, но лучшие из них затем успешно работают в зарубежных университетах и промышленных компаниях.

Что касается подготовки молодых кадров по вычислительным наукам и суперкомпьютерного образования в целом, то ситуация здесь не безнадёжная, но обновлённое Минобрнауки России должно озаботиться проблемой отсутствия квалифицированных лекторов и специалистов по ряду актуальных научных направлений. Без активного поиска новых форм привлечения зарубежных учёных тут, конечно, не обойтись.

Главное условие конечного успеха — инвестиции отраслевых структур и широкая востребованность наукоёмких разработок. На этом этапе научно-организационная деятельность сталкивается с жёсткими законами бизнеса, но последнее слово должно быть за органами государственного управления. Достаточно вспомнить, что в советское время директор каждого предприятия обязан был расходовать не менее 5% средств по статье "новая техника", и в результате Академия наук СССР более 50% финансирования получала из внебюджетных источников. В современном мире мощный стимул научно-технических инноваций — налоговые льготы, которые у нас пока что находятся в статусе перспективных, то есть весьма отдалённых, рычагов стимулирования прогресса.

Предстоящие месяцы и годы поистине должны стать судьбоносными для российской науки в целом и для возрастания роли математизации и вычислительно-информационных технологий в частности. Выполнение задуманных масштабных проектов — это проблема, требующая мобилизации значительных сил страны, но она очень своевременна и фактически неизбежна. Становящиеся хроническими пертурбации в мировой политике убеждают, что создание "информационно-технологического щита" — это проблема национальной и глобальной безопас-

ности, сродни атомному и ракетному проектам, реализация которых обеспечила миру паритет и достаточно устойчивое равновесие.

ЛИТЕРАТУРА

1. Владимир Путин в Новосибирске посетил международный форум "Технопром-2018". https://www.1tv.ru/news/2018-08-28/351287-vladimir_putin_v_novosibirske_posetil_mezhdunarodnyy_forum_tehnoprom_2018
2. Курцвейл Р. Эволюция разума. М.: ЛитРес, 2016.
3. Арнольд В.И. Что такое математика? М.: МЦНМО, 2004.
4. Software for Exascale Computing – SPP EXA-2013–2015 // Ed. by H.J. Bungartz, P. Neuman, W. Nagel. Springer International Publishing, 2016.

HOW TO REORGANIZE COMPUTER SCIENCE AND TECHNOLOGIES?

© 2019 V.P. Ilin

The Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia

Novosibirsk State University, Novosibirsk, Russia

Novosibirsk State Technological University, Novosibirsk, Russia

E-mail: ilin@sscc.ru

Received: 22.10.2018

Revised version received: 22.10.2018

Accepted: 22.10.2018

The fundamentals of modern progress are based on the infinite growth rates of the computational and informational capacities of a post-petaflop scale, which provide unthinkable opportunities for acquiring new knowledge and introducing innovation into production, environmental management, and the economic and social spheres. The key method here is based upon the predictive mathematical modeling of processes and phenomena, including theoretical and applied research, super-computations and large-scale computer experiments, operations with mega-data sets, and artificial intelligence. The solution to emerging supertasks requires huge amounts of new generation intelligent software, which is impossible without the formation of new concepts and paradigms of activity, architectural solutions, and constructive technologies. The world is embracing the creation of integrated computing and information environments, forming an instrumental context for automating the construction of models and algorithms, their mapping onto supercomputer platforms, and the creation of comfortable interfaces for users with different professional backgrounds. Such an ecosystem is designed for a long life cycle with continuous development and the coordinated participation of various groups of developers, as well as widespread demand. In fact, this is leading to the formation of a new industry with mass professions requiring high supercomputer literacy. The identified scientific and technological challenges require competent organizational and infrastructural solutions, including the interdepartmental coordination of academic, educational, and production teams.

Keywords: supercomputing, interdisciplinary tasks, mathematical modeling, computational methods and technologies, software, artificial intelligence, integrated instrumental environment, supercomputer experiment.