РЕКОНСТРУКЦИЯ ПУТЕЙ СИГНАЛЬНОЙ ТРАНСДУКЦИИ, УЧАСТВУЮЩИХ В РАЗВИТИИ РАССЕЯННОГО СКЛЕРОЗА У ЧЕЛОВЕКА

Шарипов Р.Н.[1,2,*], Евшин И.С.[1,3], Филипенко М.Л.[4], Кондрахин Ю.В.[5,1], Боярских У.А.[4], Дымшиц Г.М.[2,3], Колпаков Ф.А.[5,1]

[1]ООО «Институт системной биологии», Новосибирск, Россия; *shrus79@gmail.com

[2]Институт цитологии и генетики СОРАН, Новосибирск, Россия

[3]Новосибирский государственный университет, Новосибирск, Россия

[4]Институт химической биологии и фундаментальной медицины СОРАН, Новосибирск, Россия

[5]Конструкторско-технологический институт вычислительной техники СОРАН, Новосибирск, Россия

Рассеянный склероз (РС) – заболевание человека, проявляющееся, в основном, в молодом возрасте, характеризующееся повреждением или утратой миелиновой оболочки нейронов центральной нервной системы и ведущее к инвалидности. Диагностика РС затруднена из-за разной степени и частоты проявления симптомов у больных, их периодичности с последующей ремиссией или постепенном прогрессировании, а также из-за того, что симптоматика может быть схожа с другими заболеваниями. Среди возможных причин развития РС рассматриваются аутоиммунная реакция, инфекция или генетическая предрасположенность [1].

С целью определения ключевых компонентов путей сигнальной трансдукции, играющих важную роль в формировании РС, проведен анализ восьми выборок микрочиповых данных из базы данных GEO(www.ncbi.nlm.nih.gov/geo/) с применением платформы BioUML (http://www.biouml.org). В результате мета-анализа были выявлены гены со значимо повышенной или пониженной по сравнению с контролем экспрессией, которые классифицировались по системе GeneOntology (биологические процессы), том числе, в категории «воспаление», «апоптоз», «иммунные реакции». Классификация по биологическим путям показала принадлежность выявленных генов, в том числе, к HIF1, Fas, TNFR1, p53 путям, регулирующим воспаление и апоптоз. В ходе промоторного анализа обнаружены транскрипционные факторы, регулирующие эти гены, а на следующем этапе - ключевые белки-ферменты, контролирующие активность выявленных факторов. Таким образом, проведенный анализ позволил реконструировать клеточные пути передачи сигнала, демонстрирующие значимые отличия в активности у больных РС по сравнению с нормой.

Работа поддержана в рамках междисциплинарного проекта №17 СО РАН и проекта №33 программы «Фундаментальные науки – медицине» Президиума РАН.

ЛИТЕРАТУРА:

1. Rosati G. The prevalence of multiple sclerosis in the world: an update. Neur. Sci. 2001. V.22, №2. P. 117–139.

О НАРУШЕНИИ ОТНОШЕНИЯ ТОЛЕРАНТНОСТИ В ПРОЦЕССЕ СРАВНЕНИЯ ГЕОМЕТРИЧЕСКИХ ОБРАЗОВ

Чуракова М. Л.

Тюменский государственный университет

e-mail: mind.86_reg@mail.ru

Семёнов Д.А.

e-mail: dasem@mail.ru

При формализации процесса экспертного сравнения образов, понятие сходства автоматически превращается в отношение толерантности: рефлексивное и симметричное, но не транзитивное. Симметричность воспринимается как естественное допущение. За этим стоит представление о том, что простая перестановка образов местами не меняет самого факта сходства. На самом деле симметричность вовсе не сводится к перестановке образов местами. Реальный же процесс сравнения протекает так, что сравнение может не быть симметричным.

Проведенная нами серия экспериментов была нацелена на изучения правомерности предположения о симметричности сходства простых геометрических образов.

Образы генерировались случайно: на первом этапе из данных, генерированных датчиком случайных чисел, извлекалась малая выборка; далее строилась гистограмма распределения, после чего сглаживалась; получившийся набор графиков функций одной переменной использовался для эксперимента.

Общая схема эксперимента проста: в поле зрения испытуемого есть два окна, в каждом окне демонстрируется график, от испытуемого требуется дать ответ на вопрос «Похожи ли картинки?» После каждого ответа картинка в правом окне заменяется следующей из списка. После демонстрации всего списка образов в правом окне, образ в левом окне также заменяется следующим в списке. Все ответы фиксируются, и в дальнейшем анализируется бинарная матрица ответов.

Испытуемыми во всех экспериментах были студенты-психологи Тюменского государственного университета. Первый эксперимент проводился с использованием 100 различных образов. Полное попарное сравнение в таком тесте занимало более двух часов. Разумеется, испытуемые очень уставали, что было видно и по результатам теста. По результатам данного эксперимента мы определились с критериями оценки деятельности эксперта и смогли выбросить из выборки наименее информативные (ни на что не похожие) графики. Основным же предварительным результатом оказалось крайне малое число симметричных результатов.

В дальнейшем критерии успешного прохождения теста сводились к тому, что испытуемый должен узнавать не менее 70% рефлексивных пар; считать похожими не только совпадающие графики; количество симметричных узнаваний должно быть значимо больше случайного совпадения. То есть испытуемый должен стараться находить

сходство между образами в классическом понимании сходства как отношения толерантности.

В двух последующих экспериментах варьировалось число образов, и соответственно время эксперимента. Если в одном эксперименте предлагался набор из 70 графиков, то в другом только из 26-ти. Хотя первый эксперимент продолжался более часа, но его результаты гораздо лучше удовлетворяли заранее выбранным критериям успешности теста. Именно эти данные легли в основу дальнейшего анализа.

Спустя год группе испытуемых, наиболее удачно прошедших эксперимент по попарному сравнению 70 графиков, было предложено повторить тестирование. О целях и ожидаемых результатах они не были осведомлены. Этот эксперимент был необходим для ответа на вопрос: «Является ли асимметричность узнавания продуктом невнимательности испытуемых? Исчерпывается ли это явление представлением о случайной ошибке?»

Продемонстрировано воспроизведение экспертных оценок, в том числе воспроизведение асимметричной части сравнения. Для этого достаточно посчитать число совпадений при наложении бинарных матриц сравнения—оно всегда больше, чем в случае, когда одна из матриц транспонируется.

В наших экспериментах продемонстрировано, что часть экспертных оценок всегда асимметрична. При этом показано, что эти асимметричные оценки воспроизводятся (в нашем эксперименте спустя год!) Таким образом, асимметрия сравнения не является следствием невнимательности эксперта—это источник систематической ошибки и ее необходимо учитывать при создании систем распознавания образов.

В дальнейшем существует как минимум три альтернативы для развития деятельности по формализации экспертной оценки: а) ставить задачу детального воспроизведения мнения эксперта, в том числе воспроизводить несимметричность экспертной оценки; б) ограничиться лишь симметричными оценками, а асимметричные результаты исключить, считая их, например, трудными для эксперта; в) акцентировать внимание именно на существовании асимметричных сравнений и исследовать причины появления таких сравнений.

Путь а) кажется наиболее бесперспективным, так как по сути это попытка обучить машину систематической ошибке, свойственной человеку.

Вариант б) может быть первым шагом на пути к постепенному построению непротиворечивой системы распознавания образов. Отношение толерантности удобно тем, что позволяет ввести расстояние между образами. Для формализации процесса сравнения на первом этапе можно воспользоваться симметричными экспертными оценками, не пытаясь воспроизводить оценки в трудном для эксперта случае. После введения расстояния между образами, можно вернуться к работе с асимметричной частью оценок. Можно, например, пытаться повторно опросить экспертов именно на материале образов, доставляющих асимметричные оценки. Предварительно можно вычислить расстояние между такими образами, чтобы оценить значимость дообучения экспертной

системы в каждом случае. Не исключено, что в результате дообучаться будет именно эксперт.

Вариант c) тоже допускает дальнейшее развитие работы, ведь нами продемонстрировано только, что воспроизводимые асимметричные сравнения существуют, но никак не рассмотрены особенности таких сравнений, а значит и не изучены причины. Асимметричные сравнения имеют самостоятельную ценность, если решается задача изучения именно человеческих алгоритмов узнавания образов. Именно там, где алгоритм дает сбой, можно многое узнать об алгоритме.

Асимметрия не является чем-то неожиданным. Рассмотрим простой пример: будем считать похожими девушек с совпадающим цветом волос и глаз. Мерой сходства будет просто вероятность такого совпадения. Тогда ответ на вопрос: «Похожа ли блондинка на голубоглазую?» не совпадет с ответом на вопрос: «Похожа ли голубоглазая девушка на блондинку?» Ответ можно преобразовать в более «контрастный» бинарный вид, введя пороги для превращения количественной оценки в «да» или «нет».

В рассмотренном примере было заранее выбрано два признака, исчерпывающе описывающих образы. При этом толерантность отсутствовала. Процедура формализации сравнения образов заставляет вводить некоторый набор признаков. При этом отношение толерантности гарантирует существование базиса признаков, но не наоборот [1].

Применительно к нашей задаче мы сконструировали простой пример: сравнивали между собой всего пять графиков. Заранее были введены формальные критерии, описывающие положение и высоту трех основных пиков графика. Ширина и высота пика задавали интервал, в который должны были попасть параметры сравниваемой картинки. Собственно, именно эта привязка параметров сравнения к характеристикам первой картинки и явилась причиной асимметричности сравнения. Это явление вполне естественно: один из сравниваемых образов присутствует в поле зрения гораздо дольше другого; человек параметризует этот образ, выделяет только существенные (по его мнению) признаки и ищет только их в остальных образах. Симметричная процедура теперь не сводится к простой перестановке образов, а связывается с изменением набора существенных признаков.

Пополнение существующих баз данных (пополнение каталогов спектров, например) может сопровождаться сравнением нового образа со всеми предыдущими. Обратное сравнение при этом считается не нужным, так как процедура узнавания считается симметричной. Это может быть источником систематических ошибок в самих экспертных оценках, что делает дальнейшую формализацию трудновыполнимой, противоречивой и малоэффективной.

Вполне возможно, что «дообучение» экспертов позволит уменьшить долю асимметричных сравнений. Также вполне вероятно, что у опытных или более успешных экспертов доля ассиметричных сравнений изначально меньше.

1. Шрейдер Ю.А. Равенство, сходство, порядок. М. «Наука» 1971.

МОДЕЛЬ ТЕПЛОВОГО РЕЖИМА ПОЧВЫ В ПРОСТРАНСТВЕННО-ДИФФЕРЕНЦИРОВАННЫХ ТЕХНОЛОГИЯХ ТОЧНОГО ЗЕМЛЕДЕЛИЯ[1]

Л.А. Хворова
Алтайский государственный университет
e-mail: hla@math.asu.ru

Аннотация

Статья посвящена проблеме построения многомерной модели теплового режима почвы. Рассматриваются алгоритм и численный метод решения двумерной задачи.

MODEL OF THE THERMAL REGIME OF SOILS IN THE SPACE-DIFFERENTIATION TECHNOLOGY OF PRECISION AGRICULTURE

Khvorova L.A.

Abstract

The article is devoted to the problem of constructing a multidimensional model of the thermal regime of soil. We consider an algorithm and a numerical method for solving two-dimensional problem.

Разработка математических моделей, корректно учитывающих процессы теплопереноса в почве, является сложной и актуальной задачей. Соседствующие почвенные массивы (выделенные единицы управления в рамках одного поля) характеризуются различными теплофизическими величинами (параметрами), которые, в свою очередь, зависят от соотношения твердой, жидкой и газообразной составляющих, текстурных и структурных особенностей грунтов, состояния влаги и температуры.

Точное (ориентированное) земледелие или локально специфическое – это эффективное, рациональное управление процессами роста растений в соответствии с их потребностями в питательных веществах и условиях произрастания с различной степенью дифференциации. Точное земледелие – это система хозяйствования на земле с использованием пространственно-дифференцированных технологий, опирающихся на использование новейших достижений в области информатики, моделирования и техники: компьютерных систем генерации агротехнологических решений, глобальных систем позиционирования (GPS), геоинформационных технологий (ГИС), новейших информационных технологий, современной сельскохозяйственной техники, управляемой бортовой ЭВМ, а также многофункционального программного обеспечения, позволяющего принимать оптимальные решения при управлении сельскохозяйственным предприятием [1].

Урожайность сельскохозяйственной культуры на различных участках одного и того же поля, как правило, различна. Это связано в первую очередь с тем, что процесс получения продукции растениеводства реализуется в пространстве и во времени на конкретной территории, которая не является однородной даже в пределах одного поля. В традиционном земледелии все характеристики поля, как правило, считаются одинаковыми для всех его участков. Точное земледелие предполагает динамическую оптимизацию при выполнении агротехнических операций для каждого однородного участка поля в зависимости от складывающихся агрохимических, агрофизических, фитосанитарных факторов.

Одним из факторов, способствующих развитию сельского хозяйства, является использование информационных технологий, ядром которых являются динамические модели, позволяющие учитывать пространственную неоднородность различных участков в границах одного поля.

[1] Работа выполнена при поддержке ведомственно-аналитической программы «Развитие научного потенциала Высшей школы 2010–2011» №2.2.2.4/4278.

В подавляющем большинстве современные модели, описывающие продукционный процесс сельскохозяйственных растений, рассматривают однородный фиктивный посев, а стратификация его характеристик производится в единственном вертикальном направлении. В подобных моделях расчет производится отдельно для каждой опорной точки поля с параметрами, характерными только для данного типа почвы. Все точки считаются независимыми друг от друга, но предполагается, что все окружение данной точки обладает теми же свойствами и, соответственно, никаких горизонтальных перетоков вещества и энергии не происходит.

Для целей точного земледелия горизонтальная неоднородность поля является важнейшим фактором, влияющим на выбор агротехники и определяющим результат хозяйствования. Учет взаимодействия динамики продукционного процесса на соседних участках неоднородного поля требует построения принципиально иной, более сложной многомерной модели.

Один из упрощенных вариантов решения данной проблемы предложен в [2]. Влияние разных участков посева друг на друга в [2] предполагается осуществлять в рамках системы поливариантных расчетов при условии поддержки ею механизма параллельных вычислений нескольких сценариев для одномерного случая.

Многомерная модель теплового режима почвы

Колебания температуры – важный компонент почвенного микроклимата. Температура почвы существенно влияет на многие протекающие в ней процессы. С тепловым режимом почв тесно связаны начало и конец вегетационного периода, пространственное размещение растений, характер распространения корневых систем, скорость поступления к корням питательных элементов.

В рамках концепции пространственно-дифференцированных технологий точного земледелия рассмотрим трехмерную модель теплового режима почвы.

Математические модели, связанные с описанием явления теплопереноса в пределах почвенного компартмента, основаны на нестационарных трехмерных уравнениях параболического типа. Теплота, поступающая на поверхность почвы, под действием создаваемого градиента температур перераспределяется в объеме почвенного компартмента Ω.

Пусть $P=P(x,y,z)$ – точка почвенного компартмента Ω, $P\in\Omega\subset R^3$, $T(P,t)$ – температура в точке P почвенного компартмента в момент времени t, $t\in[0,T]$. Тогда уравнение теплопереноса в почвенном компартменте можно Ω записать в виде:

$$\rho\, c\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(\chi\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\chi\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(\chi\frac{\partial T}{\partial z}\right)+f(x,y,z,t), \qquad (1)$$

где $\rho(x,y,z)$ – плотность почвы, $c(w(x,y,z))$ – теплоемкость, χ – коэффициент теплопроводности, зависящий от влажности почвы w: $\chi=\chi(w(x,y,z))$. Теплоперенос осуществляется вдоль координатных осей *Ox, Oy, Oz*; а $f(x,y,z,t)$ – функция источника тепла. Заметим, что влажность почвы w считается здесь заданной функцией.

Искомая функция $T(P,t)$ должна удовлетворять начальным условиям:

$$T(P,t)|_{t=0}=T(P,0) \text{ для } P\in\Omega \qquad (2)$$

и некоторым граничным условиям.

Нижняя граница помещается, как правило, на глубине, на которой температура либо полагается постоянной, либо зависящей от времени и точек границы известным образом. Следовательно, при $y=-H$ (на нижней плоскости почвенного компартмента Ω выполняется

$$T(-H,t)=\varphi_H(t). \qquad (3)$$

В качестве верхнего граничного условия следует записать соотношение, обеспечивающее «сшивание» решений задачи в почве и в приземном воздухе. Наиболее

корректным представляется условие теплового баланса на поверхности почвы вида (условие третьего рода):

$$\chi \frac{\partial T}{\partial n} + \beta(T - T_a) = q(x,t)\,. \qquad (4)$$

Здесь T – температура поверхности почвы, T_a – температура воздуха, $n = (n_1, n_2, n_3)$ – вектор внешней нормали к верхней границе деятельной поверхности почвы; $\frac{\partial T}{\partial n} = \frac{\partial T}{\partial x} n_1 + \frac{\partial T}{\partial y} n_2 + \frac{\partial T}{\partial z} n_3$, функция q задает тепловой поток в почву, затраты тепла на турбулентный перенос в атмосферу и на испарение и т.д., β – коэффициент теплообмена с внешней средой.

Общее условие теплового баланса (4) на границе двух сред, в частности, на поверхности почвы, записанное в виде [2]:

$$(1 - A_s)\cdot Q(0) + \Delta F(0) = \rho_a c_p D^T_{soil}(T_0 - T_a(NL)) +$$
$$+\rho_a \theta \cdot D^q_{soil}(q_0 - q_a(NL)) + \frac{2}{h_0 + h_1}\chi_{0,1}(T_0 - T_1)\,, \qquad (4^*)$$

представляет собой разностный аналог соотношения (4). В его левую часть входят приходные статьи теплового баланса: баланс коротковолновой ($(1 - A_s)\cdot Q(0)$) и длинноволновой ($\Delta F(0)$) радиации; A_s – альбедо почвы. Слагаемые в правой части (4*) характеризуют затраты тепла на турбулентный перенос в атмосферу и на испарение (первое и второе слагаемые), а также поток тепла в почву (третье слагаемое). Здесь ρ_a – плотность воздуха, c_p – удельная теплоемкость воздуха при постоянном давлении, D^T_{soil}, D^q_{soil} – коэффициенты проводимости тепла и паров воды, T_0, T_1 – температура верхних слоев почвы, $T_a(NL)$ – температура нижнего яруса посева, номер которого равен NL, θ – скрытая теплота парообразования, q_0 – концентрация паров воды в поровом пространстве почвы у поверхности, $q_a(NL)$ – удельная влажность воздуха, h_0, h_1 – высоты верхних слоев почвы, $\chi_{0,1}$ – теплопроводность верхнего слоя почвы. Заметим, что последнее слагаемое в (4*) представляет собой разностный аналог величины потока тепла на границе.

Численное решение задачи (1)–(4) можно найти при помощи метода стабилизирующей поправки [3, 4]. Разностная схема второго порядка аппроксимации для решения уравнения (1) может быть представлена в следующем общем виде:

$$\frac{T^{k+1/3} - T^k}{\Delta t} = \Lambda_1 T^{k+1/3} + \Lambda_2 T^k + \Lambda_3 T^k + F^k\,,$$
$$\frac{T^{k+2/3} - T^{k+1/3}}{\Delta t} = \Lambda_2 (T^{k+2/3} - T^k),$$
$$\frac{T^{k+1} - T^{k+2/3}}{\Delta t} = \Lambda_3 \left(T^{k+1} - T^k \right),$$

где Λ_i, $(i = 1,2,3)$ есть разностные аналоги соответствующих дифференциальных операторов ($\Lambda_i \sim \frac{\partial}{\partial x_i}\left(K \frac{\partial}{\partial x_i} \right)$, $x_i = x, y, z$ для $i = 1,2,3$, соответственно). Правая часть F сформирована здесь после введения коэффициента температуропроводности K и последующего приведения (1) к дивергентному виду. Не останавливаясь подробно на реализации метода стабилизирующей поправки для решения трехмерной задачи, перейдем к изучению двумерной аппроксимации задачи о распределении температуры в массивах почвы, имеющих вертикальную (относительно направления силы тяжести) границу раздела, связанную с неоднородностью структурных пластов почвы.

Двумерная задача

Рассмотрим следующий вариант задачи (1)–(4). Пусть неоднородный почвенный компартмент Ω состоит из двух участков, значительно отличающихся по влиянию характеристик поля на продукционный процесс посева и на движение почвенных растворов (в действительности свойства почвы меняются от точки к точке непрерывно и случайным образом). Целью «размежевания» поля на единицы управления является уменьшение теоретически бесконечной вариабельности условий произрастания к ограниченному набору вариантов. Тогда $\Omega = \Omega_1 \cup \Omega_2$, где $\Omega_1 = \{-x_1 \le x \le 0; -H \le y \le 0\}$, $\Omega_2 = \{0 \le x \le x_2; -H \le y \le 0\}$. Здесь полагается, что границы участков Ω_1 и Ω_2 являются известными и прямолинейными. В случае криволинейных границ областей Ω_1 и Ω_2 задача также может быть сформулирована и успешно решена.

Пусть система координат выбрана таким образом, что ось *0y* проходит по границе раздела областей Ω_1 и Ω_2. Функция T_1 определяет температуру почвы в области Ω_1, а T_2 – температуру почвы в области Ω_2. Тогда в силу почвенной однородности областей Ω_1 и Ω_2 можно записать условия:

$$\frac{\partial T_1}{\partial x} = 0 \text{ при } x = -x_1; \tag{5}$$

$$\frac{\partial T_2}{\partial x} = 0 \text{ при } x = x_2. \tag{6}$$

На границе раздела компартментов Ω_1 и Ω_2 $(x = 0)$ должны выполняться условия непрерывности температур и тепловых потоков:

$$T_1 = T_2 \text{ и } \chi_1 \frac{\partial T_1}{\partial x} = \chi_2 \frac{\partial T_2}{\partial x} \text{ при } x = 0. \tag{7}$$

Уравнение (1) в двумерном случае будет иметь вид:

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\chi \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y}\left(\chi \frac{\partial T}{\partial y} \right) + f(x, y, t) \tag{8}$$

Численное исследование задачи о распределении температуры в областях Ω_1 и Ω_2 будет производиться с использованием конечно-разностных методов.

Введем коэффициент температуропроводности K: $K = \frac{\chi}{\rho c}$, который также будет функцией пространственных координат x, y и перепишем уравнение (8) в следующем дивергентом виде:

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(K \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y}\left(K \frac{\partial T}{\partial y} \right) + \frac{K}{\rho c}\left[\frac{\partial(\rho c)}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial(\rho c)}{\partial y} \cdot \frac{\partial T}{\partial y} \right] + \frac{1}{\rho c} f(x, y, t) \tag{9}$$

Для численного решения уравнения (9), описывающего процесс теплопереноса, применяется численный метод, разработанный в [5], с использованием продольно-поперечной конечно-разностной схемы (метод переменных направлений), формально имеющей второй порядок аппроксимации. Схема расчета записывается в следующем общем виде:

$$\frac{T^{k+\frac{1}{2}} - T^k}{0.5 \cdot \Delta t} = \left[K T_x \right]_x^k + \left[K T_y \right]_y^{k+1/2} + F^k,$$

$$\frac{T^{k+1} - T^{k+\frac{1}{2}}}{0.5 \cdot \Delta t} = \left[K T_x \right]_x^{k+1} + \left[K T_y \right]_y^{k+1/2} + F^k. \tag{10}$$

Здесь $F=\frac{K}{\rho c}\left[\frac{\partial(\rho c)}{\partial x}\cdot\frac{\partial T}{\partial x}+\frac{\partial(\rho c)}{\partial y}\cdot\frac{\partial T}{\partial y}\right]+\frac{1}{\rho c}f(x,y,t)$, Δt – шаг по времени, $T^k=T(t_k,\cdot)$, $t_k=k\cdot\Delta t,\ k=1,2,\ldots$.

Для реализации представленной схемы вводится равномерная разностная сетка (x_n,y_m) для каждой области Ω_i, $i=1,2$, следующим образом:

для области Ω_1: (x_n,y_m), где $x_n=-x_1+(n-1)h_1,\ n=1,2,\ldots,N_1,\ h_1=\frac{x_1}{N_1}$;

$$y_m=-H+(m-1)h_y,\ m=1,2,\ldots,M,\ h_y=\frac{H}{M};$$

для области Ω_2: (x_n,y_m), где $x_n=(n-1)h_2,\ n=1,2,\ldots,N_2,\ h_2=\frac{x_2}{N_2}$;

$$y_m=-H+(m-1)h_y,\ m=1,2,\ldots,M,\ h_y=\frac{H}{M};$$

Значение сеточной функции $T(x,y,t)$ в узлах сетки обозначим $T^k_{n,m}=T(x_n,y_m,t^k)$. При этом используется следующая разностная аппроксимация для входящих в (10) слагаемых:

$$\left[KT_x\right]_x\approx\bar{K}_{n+1}\frac{T_{n+1,m}-T_{n,m}}{h_x^2}-\bar{K}_n\frac{T_{n,m}-T_{n-1,m}}{h_x^2}.$$

Здесь

$$\bar{K}_{n+1}=K_{n+1/2,m},\ K_{n+1/2,m}=K(x_{n+1/2},y_m),$$
$$x_{n+1/2}=x_n+0.5h_x,\ h_x=h_1 \text{ или } h_x=h_2.$$

В результате требуется решить системы линейных алгебраических уравнений

$$-a_{n,m}T^{k+1/2}_{n,m-1}+b_{n,m}T^{k+1/2}_{n,m}-c_{n,m}T^{k+1/2}_{n,m+1}=d_{n,m},$$
$$-a_{n,m}T^{k+1}_{n-1,m}+b_{n,m}T^{k+1}_{n,m}-c_{n,m}T^{k+1}_{n+1,m}=d_{n,m},$$

соответствующие (10). Данные системы решаются методом прогонки. При этом в направлении *y* используется обычный вариант данного метода [6].

Граничные значения температуры T_1 и T_2 следуют из (3), (4).

Для определения T_1 и T_2 на слое $(k+1)$ мы используем условия непрерывности температур и тепловых потоков на границе раздела (7) и представление решения (т.е. температуры в каждой из областей) в таком виде, когда $(T_1)_{n,m}$ и $(T_2)_{n,m}$ выражаются через неизвестные значения температуры $(T_1)_{N_1+1,m}=(T_2)_{1,m}$ на границе раздела $x=0$. Представления вида

$$T_{1n,m}=\beta^1_{n,m}+\gamma^1_{n,m}\cdot\bar{T}_m,$$
$$T_{2n,m}=\beta^2_{n,m}+\gamma^2_{n,m}\cdot\bar{T}_m,$$

где $\bar{T}_m$ – температура на границе раздела областей Ω_1 и Ω_2, позволяют организовать своеобразную прогонку с параметрами, коими являются граничные значения температуры $\bar{T}_m$, и найти сначала сами эти значения, а затем и распределение температуры в областях Ω_1 и Ω_2.

Первые производные, входящие в (7), внутри расчетной области аппроксимируются традиционно симметричными конечно-разностными аналогами со вторым порядком. Первые производные на границах расчетной области аппроксимируются несимметричными конечно-разностными аналогами также второго порядка.

Общая схема численного решения задачи состоит в осуществлении следующих этапов.

1. Переход на новый временной слой t^{k+1} начинается с расчета температуры $T_1^{k+\frac{1}{2}}$ и $T_2^{k+\frac{1}{2}}$ на промежуточном временном слое $t^{k+\frac{1}{2}}$. Расчет производится в каждой из областей Ω_1 и Ω_2

2. Затем, с помощью прогонки с параметрами, вычисляются значения температур $T_i^{k+1}, i=1,2$, на слое $(k+1)$ одновременно в обеих областях Ω_1 и Ω_2.

Отметим, что с использованием схемы (10) и введения итерационного параметра, может быть осуществлено вычисление распределения температуры в стационарном случае.

Дальнейшее математическое и численное моделирование двумерных задач следует проводить с учетом криволинейности границ раздела, а также подвижных границ раздела двух фаз и постановки на этих границах условий сопряжения и обобщенных условий Стефана.

ЛИТЕРАТУРА

[1]. Якушев В.П., Полуэктов Р.А., Смоляр Э.И. Оценка технологий точного земледелия: аналитический обзор. Агрохим. вестник. 2002. № 3. С. 36–40.

[2]. Полуэктов Р.А., Смоляр Э.И., Терлеев В.В., Топаж А.Г. Модели продукционного процесса сельскохозяйственных культур. СПб.: Изд-во С.-Петерб. ун-та, 2006. 396 с.

[3]. Марчук Г.И. Методы расщепления. М.: Наука, 1988. 264 с.

[4]. Гончарова О.Н. Метод расщепления по физическим процессам для расчета трехмерных задач конвекции. Известия АлтГУ. 2007. №1. С. 39–44.

[5]. Яненко Н.Н. Метод дробных шагов решения многомерных задач математической физики. Новосибирск: Наука, 1967. 196 с.

[6]. Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. М.: Наука, 1978. 590 с.

ДАЛЬНЕВОСТОЧНАЯ ШКОЛА МАТЕМАТИЧЕСКОЙ БИОЛОГИИ: О ФЛЮКТУАЦИЯХ ЧИСЛЕННОСТИ КАК РЕЗУЛЬТАТЕ ПРОЦЕССОВ ДИНАМИЧЕСКОЙ САМООРГАНИЗАЦИИ ПОПУЛЯЦИЙ И СООБЩЕСТВ*

*Исследования проведены при частичной финансовой поддержке РФФИ (проекты № 09-04-00146-а, № 11-01-98512-р_восток_а)

Е.Я. Фрисман
Учреждение Российской Академии наук
Институт комплексного анализа региональных проблем ДВО РАН
e-mail: frisman@mail.ru

Кратко обсуждается процесс развития Математической биологии во второй половине XX века в основных научных школах России. Особо подчеркивается роль Новосибирской школы. Подробно комментируется возникновение и развитие Дальневосточной школы и ее место в науке того времени. Отмечается, что именно в рамках этой школы сформировались представления о сценариях перехода к детерминированному хаосу в популяционных моделях. Далее в докладе обсуждаются некоторые эволюционные механизмы, обеспечивающие переход к нерегулярной популяционной динамике. Показано, что эволюционное изменение частот аллелей адаптивных генов, сопровождающееся ростом средней приспособленности популяции, может привести к циклическим и хаотическим режимам динамики ее численности. При этом на фоне колебаний численности, создаются условия для возникновения "конкуренции" между генотипами, чувствительными к уровню численности, т.е. условия для r- и K-отбора. Действие такого отбора приводит к возникновению сложных режимов динамики уже не только численности, но и генетической структуры популяции. Таким образом, в докладе показаны возможные механизмы появления очень сложной временной организации генетического биоразнообразия, весьма чувствительного к внешнему воздействию.

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ПОДДЕРЖКИ МЕДИКО-ГЕНЕТИЧЕСКИХ ИССЛЕДОВАНИЙ

И.В. Тяжев[1], Н.И. Толстых[1], Е.Н. Кудрявцева[2], Р.Н. Шарипов[1,3,*], У.А. Боярских[2], М.Л. Филипенко[2], Г.И. Лифшиц[4], Ф.А. Колпаков[5,1]

*[1]ООО «Институт системной биологии», Новосибирск, Россия; *shrus79@gmail.com*

[2]Институт химической биологии и фундаментальной медицины СОРАН, Новосибирск, Россия

[3]Институт цитологии и генетики СОРАН, Новосибирск, Россия

[4]Центр новых медицинских технологий СО РАН, Новосибирск, Россия

[5]Конструкторско-технологический институт вычислительной техники СОРАН, Новосибирск, Россия

Современные медицинские центры оперируют большими объемами данных, которые требуют соблюдения соответствующих мер безопасности и контроля доступа к ним, а также нуждаются в программной поддержке статистических вычислений и анализа полученных результатов. В рамках междисциплинарных проектов СО РАН нами был разработан и внедрен на базе ЦНМТ СО РАН специальный программный комплекс для поддержки медико-генетических исследований мультифакториальных заболеваний.

Комплекс построен по классической трехуровневой архитектуре: 1 уровень – базы данных (MySQL); 2 уровень – сервер приложений, используемый для реализации основной части бизнес-логики и анализа статистических данных; 3 уровень – пользовательский интерфейс. Применение портальных технологий предполагает использование в качестве пользовательского интерфейса обычного веб браузера («тонкий» клиент). Для генерации презентационной логики и обеспечения основной части бизнес-логики используется технология BeanExplorer EE (http://www.beanexplorer.com), которая представлена набором сервлетов на сервере приложений и использует протокол JDBC для работы с сервером баз данных. Комплекс обеспечивает решение целого ряда задач от ведения реестра пациентов до формирования выборок пациентов и статистического анализа данных средствами модуля BioUML EE (http://www.biouml.org), а также с привлечением внешних программ (например, R/Bioconductor). Для статистического анализа используются только обезличенные данные пациентов (категория 4), созданные средствами комплекса на основе персональных данных (категория 1), которые хранятся в отдельной защищенной базе данных.

Внедрение разработанного нами программного комплекса позволило значительно облегчить, унифицировать и ускорить ввод информации в базу данных, а также обработку и проведение статистического анализа данных пациентов.

Работа поддержана в рамках междисциплинарного проекта №17 СО РАН и проекта №33 программы «Фундаментальные науки – медицине» Президиума РАН.

А.А.УХТОМСКИЙ КАК ПРЕДШЕСТВЕННИК КИБЕРНЕТИКИ В РОССИИ.

Зуева Е.Ю., Ефимов Г.Б.
ИПМ им. М.В. Келдыша РАН, Москва, Россия
Ezueva@keldysh.ru , efimov@keldysh.ru

А.А.Ухтомский(1875-1942) – русский советский физиолог, автор учения о доминанте. При изучении процессов возбуждения и торможения в нервной системе ставил ряд основополагающих вопросов:

об активности и целенаправленности поведения организма,

о механизмах, обеспечивающих согласованную работу организма как целого,

о динамике развития организма и необратимости времени для живых существ,

о зависимости реакции организма на внешнее воздействие от текущего состояния организма (и в этой связи о влиянии решаемой задачи и предшествующего опыта на восприятие и поведение),

о физиологических аспектах ряда психологических понятий – мотивации, внимания, предметного мышления и др,

о физиологических аспектах нравственных понятий – о необходимости для человека собеседника, «другого лица», о возможности полноценного общения и препятствиях к нему.

В данном докладе мы попытались, используя не только научные статьи, но и опубликованные в последние годы письма, дневники и записные книжки Ухтомского [1-4], выделить основные философские и методологические положения его учения. Было проведено сравнение их с теми идеями, которые привели к возникновению кибернетики, и с теми, из которых позже выкристаллизовалось еще одно научное направление – синергетика [5],. Для этого использовались книги [6,7]. Были выявлены совпадения по важным пунктам:

необратимость времени в жизненном процессе,

наличие двух принципиально различных стадий процесса – детерминированной фазы и фазы неустойчивости,

интеграция в единую систему как объединение ритмических процессов, при этом в детерминированной фазе один из процессов преобладает и подавляет остальные (доминанта),

неравновесность процесса (наличие энергетического обмена со средой), и ряд других.

1.А.А. Ухтомский. Заслуженный собеседник. Рыбинское подворье. 1997.

2.А.А. Ухтомский. Интуиция совести. Спб. 1996

3.Алексей Ухтомский. Доминанта. «Питер», серия Психология-классика, Спб, 2002.

4.А.А.Ухтомский. Доминанта души. Рыбинск, 2001.

5.Е.Ю.Зуева, Г.Б.Ефимов. Принцип доминанты Ухтомского как подход к описанию живого. Препринт ИПМ им. Келдыша, №14 за 2010г. http://library.keldysh.ru/prep

6.Норберт Винер. Кибернетика, или управление и связь в животном и машине. «Советское радио», М, 1968

7.Е.Н.Князева, С.П.Курдюмов. Основания синергетики. Синергетическое мировидение. М., КомКнига, 2005. – 240с. (Синергетика: от прошлого к будущему).

КЛАСТЕРИЗАЦИЯ ОРГАНИЗМОВ ПО ХАРАКТЕРИСТИКАМ СТРОЯ ИХ ДНК

В работе рассмотрены подходы для анализа структуры знаковых последовательностей разной природы. Особо выделен анализ строя цепи, который позволяет исследовать структуру цепи безотносительно к её природе, непосредственно учитывая взаимное расположение элементов.

Даётся краткое описание строя цепи как нового математического объекта – особым образом организованного кортежа на основе данной знаковой последовательности. Определена декомпозиция строя на однородные знаковые цепи. Определены две числовые характеристики строя – средняя удалённость и регулярность элементов цепи.

Приведены вычисленные значения характеристик строя 29 нуклеотидных последовательностей живых организмов от простейших до позвоночных. Будучи упорядоченными по характеристике средней удалённости организмы разделились на три группы – между вышеупомянутыми выделилась группа беспозвоночных.

В дальнейших исследованиях производилась автоматическая кластеризация в двумерном λ-пространстве для следующих пар характеристик: <длина нуклеотидной цепи, средняя удалённость>, <регулярность, средняя удалённость>, <регулярность, нормированная удалённость>. Результаты автоматической кластеризации представлены в табличном и графическом виде и хорошо совпадают с экспертными оценками.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РАЗНОМАСШТАБНЫХ ГИДРОФИЗИЧЕСКИХ ПРОЦЕССОВ КАК ОСНОВНЫХ ФАКТОРОВ ТЕПЛО- И МАССОПЕРЕНОСА В ЭКОСИСТЕМЕ ОЗЕРА БАЙКАЛ*

Цветова Е.А.

Озеро Байкал – настоящий мировой центр биоразнообразия, значимость которого трудно переоценить. Уникальная экосистема озера существует благодаря специфическим гидродинамическим условиям. В свою очередь экосистема является активным элементом, влияющим на качество вод озера.

В докладе представлены результаты численного моделирования гидродинамических процессов различных масштабов, полученные с помощью трехмерной негидростатической модели озера со свободной поверхностью. В модели учтены реальная конфигурация берегов и рельефа дна, притоки основных рек и сток Ангары. Термическое взаимодействие озера с атмосферой описывается с помощью соотношений баланса тепла и влаги на поверхности. Динамическое взаимодействие с атмосферой параметризуется заданием переменных в пространстве и времени сценариев ветровых нагрузок.

Обсуждаются результаты моделирования термобара, естественной и вынужденной конвекции, возникновения и разрушения мезотермического максимума, процессов обновления глубинных вод. Все эти процессы и явления имеют специфические особенности проявления в озере.

* Работа поддержана Программами фундаментальных исследований № 4 и 20 Президиума РАН, №3 Отделения математических наук РАН, а также проектом РФФИ 11-01-00673.

НАПРАВЛЕНИЕ: ***ТЕОРЕТИЧЕСКОЕ И СИСТЕМНОЕ ПРОГРАММИРОВАНИЕ***

АЛГЕБРАИЧЕСКАЯ СТРУКТУРА И МОДЕЛЬ ВЫЧИСЛЕНИЙ ДЛЯ АРИФМЕТИКИ ОГРАНИЧЕННЫХ ЦЕЛЫХ НЕОТРИЦАТЕЛЬНЫХ ЧИСЕЛ

Ю. Г. Сметанин
Вычислительный центр им. А. А. Дородницына РАН
smetanin.iury2011@yandex.ru

М. В. Ульянов
Московский государственный университет печати; Научно-исследовательский университет Высшая школа экономики
muljanov@mail.ru

Аннотация

В докладе предлагается математический формализм компьютерной целочисленной арифметики в виде оригинальной алгебраической структуры с частичными операциями для арифметики целых неотрицательных чисел. Предложенный формализм ограниченность представления целых чисел в реальном компьютере и порождаемые этой ограниченностью особенности арифметических операций. В качестве конструктивной реализации введённой алгебраической структуры предложена специальная модель вычислений с предусловиями выполнения команд, представляющая собой формальную модель компьютера с ограниченной целочисленной арифметикой. Для частного случая представления арифметического выражения полиномом от его операндов, показана возможность применения предложенного аппарата для решения задачи проверки входов арифметических выражений на результатов в ограниченном носителе. Дополнительно рассматривается подход к построению эквивалентных арифметических выражений с целью обеспечения замкнутости в носителе задаваемых ими промежуточных результатов для конкретного входа.

Предложенный формализм и модель вычислений могут быть использованы при теоретическом анализе и исследовании компьютерных алгоритмов, а также для решения задач проверки допустимости входов арифметических выражений и выполнения их эквивалентных преобразований.

СУПЕРКОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ – ОСНОВА УСТОЙЧИВОГО СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ РОССИИ В XXI ВЕКЕ

В.Б. Бетелин

1. Создание в конце XVIII века технологий серийного производства паровых машин заложило основы как новых отраслей – энергетического, транспортного (железнодорожный и водный транспорт), а затем и общего машиностроения, так и **мирового рынка сложных технических систем**. Развитие этих отраслей, обусловленное, прежде всего стремлением к экономическому и военному превосходству в мире, в последующие 200 лет, собственно и **определило путь мирового социально-экономического развития**, итогом которого явилось создание и интенсивное развитие рукотворной технической среды обитания человека, основанной на принципе использования сложных технических систем, не имеющих аналогов в природной среде.

Одновременно с созданием промышленности машиностроения, в течение XIX века, на основе достижений фундаментальной науки, являющихся результатом изучения природной среды, сформировались новая отрасль знаний - технические науки и новая система профессиональной подготовки - инженерное образование, нацеленные на решение задачи создания технической среды обитания человека и ее базового элемента – сложной технической системы. Основной предмет технических (инженерных) наук – методы создания возможно более точных **информационных моделей сложных технических систем,** обеспечивающих, как возможность априорного предсказания наиболее важных параметров функционирования этих систем, так и их последующего изготовления и эксплуатации.

Результатом интенсивного развития технических наук и инженерного образования является создание в конце XIX века паровой турбины, бензинового и дизельного двигателей с существенно более высоким коэффициентом полезного действия, чем у паровой машины, следствием чего явилось развитие новых отраслей и соответствующих ниш мирового рынка (теплоэлектроэнергетика, автомобилестроение, авиастроение), а также создание **основ мирового рынка углеводородов.**

Созданием в XX веке эффективных систем теплоэнергетики на углеводородах, а также газотурбинных двигателей и атомных энергетических установок фактически завершился процесс эволюционного формирования современного облика отраслей нефтегазодобычи и переработки, а также энергетического и транспортного машиностроения и соответствующих им ниш мирового рынка. Лидерство страны на этих рынках – гарантия притока в национальную экономику стабильных финансовых потоков, **достаточных**

для обеспечения ее устойчивого социально-экономического развития и военной безопасности [1].

По данным Министерства энергетики США [6] **85% генерируемой в США тепловой энергии** вырабатывается за счет сжигания углеводородов (нефть, природный газ, уголь). Прогнозируется, что главенствующее положение энергетики углеводородов сохранится в США на ближайшие **30-50 лет**, вследствие огромной стоимости уже созданной инфраструктуры этой энергетики. Фактически это означает, что, по крайней мере, **до середины 21 века** нефтегазовая и машиностроительные отрасли по-прежнему **будут основой устойчивого социально-экономического развития** и военной безопасности этой страны.

В США развертывается двадцатилетняя программа создания и использования экзафлопных технологий (1018 оп/сек) прежде всего, с целью обеспечения конкурентоспособности страны на глобальном энергетическом рынке, за счет **«перехода промышленности от эмпирических методов проектирования и конструирования, опирающихся на натурные тесты, к научным методикам, опирающимся на предсказательное компьютерное моделирование».**

Обеспечение конкурентоспособности энергетики углеводородов на мировом энергетическом рынке невозможно без огромных долгосрочных (15-20 лет) финансовых вложений в модернизацию, на основе экзафлопных технологий, (создание которых является самостоятельной чрезвычайно сложной и дорогостоящей проблемой) существующей энергетической инфраструктуры. При этом **прогнозируемый эффект,** полученный в результате всех этих усилий, измеряется всего лишь **полутора-двумя десятками процентов**. Аналогичным образом обстоит дело и с атомной энергетикой (снижение на **20%** стоимости строительства АС в результате **15-летних усилий**) и с альтернативными и возобновляемыми источниками энергии (за счет биотоплива сократить потребление бензина на **20% за 10 лет, за двадцать лет довести долю электроветроэнергии до 20%)[2].** Одним из наиболее важных результатов этого двадцатилетнего «экзафлопного скачка» будет создание нескольких сверхкрупных компаний, которые монополизируют и поделят мировые нефтегазовые и энергомашиностроительные рынки, аналогично тому, как в настоящее время BOEING и AIRBUS уже фактически поделили мировой рынок магистральных авиалайнеров.

В России **88%** потребляемой энергии вырабатывается за счет сжигания углеводородов на теплоэнергетических установках, около **80%** которых построены в **1960-1985гг. [7]**, то есть основой устойчивого социально-экономического развития России, также как и США, на ближайшие **30-50 лет** будет энергомашиностроение и нефтегазодобыча и переработка, модернизированные на основе экзафлопных технологий. Очевидно, что необходимым условием обеспечения этого устойчивого развития является достижение паритета с США в области экзафлопных технологий.

2. Россия унаследовала от СССР высокоэффективный промышленно-научно-образовательный комплекс, обеспечивающий создание и промышленное производство современных сложных технических систем машиностроительного профиля. Поэтому к началу 1990 годов в России были все предпосылки для достижения лидерства национальной экономики на мировых машиностроительных рынках путем реформирования промышленной компоненты этого комплекса. Однако эта возможность не была реализована, прежде всего, потому, что в течение всех восемнадцати лет реформ российская экономика была, по сути, нацелена на вхождение и последующее достижение лидирующих позиций на мировых финансовых, а не машиностроительных рынках, что собственно и является основной причиной сложившегося бедственного положения науки, образования и машиностроительной отрасли России, включая ОПК[1] [1].

Действительно, уже в течение многих лет основным критерием успеха любого российского предприятия реального сектора экономики, является величина прибыли, полученная от вложенных в производство средств за минимальное время. Чем больше размер этой прибыли, тем больше и успех предприятия. Весьма существенно, что **получение этой прибыли не обусловливается необходимостью производства и реализации на внутреннем или мировом рынке какой-либо промышленной продукции, в том числе и машиностроительной,** то есть **натуральные показатели предприятия не являются критерием успеха его деятельности. К** критериям успеха предприятия идеологи реформ **не относят также ни уровень заработной платы, ни число рабочих мест.**

Другими словами **финансовая конкурентоспособность и прибыль** как главные цели российских хозяйственных субъектов, фактически освобождают отечественный бизнес (большой, средний, малый) от какой-либо социальной ответственности как в части занятости и доходов населения России, так и в части увеличения эффективности и масштабов производства, то есть увеличения контролируемой доли мирового рынка.

Об этом, например, свидетельствует статья в газете «Аргументы и факты» за 2004 год, озаглавленная «Нищета – не наша забота», в которой от имени крупной финансовой компании прямо заявляется, что ***борьба с***

---

[1] «В 90-е годы была реализована наивная идея всемогущей «невидимой руки рынка». Считалось, что нам достаточно естественных преимуществ в виде больших запасов нефти и газа. На самом деле, в Советской России существовали очень неплохие высокотехнологические отрасли: машино-, судо-, самолетостроение и др., которые при определенных усилиях государства могли бы успешно конкурировать на мировом рынке. Но в результате у нас распались крупные отраслевые объединения, в то время как во всем мире концентрация производства и капитала, наоборот, нарастала. Всем известный Airbus - самая натуральная госкорпорация. А с гигантами могут бороться только гиганты. Поэтому нужно было либо похоронить отечественную перерабатывающую промышленность окончательно, либо создавать своих «монстров"». *Руслан Гринберг, директор Института экономики РАН:*
Известия, 10 августа 2009; http://www.izvestia.ru/economic/article3131628/

бедностью не дело частного бизнеса, а дело государства, а также, что «***Наше дело – зарабатывать деньги для акционеров и клиентов*** *в рамках закона.* ***Других обязанностей у нас нет****»*[2]*, «Будут хорошие прибыли – мы останемся в этом бизнесе. А если предложат хорошую цену – продадим.* ***Наша цель – зарабатывание денег».***

Реализация этих принципов неизбежно приводит к тому, что интересы конкретного предприятия или ведомства вступают в противоречие с национальными интересами экономики страны и общества в целом. Результат этих противоречий – масштабные проблемы с поддержанием и развитием национальной технической и природной среды обитания, решить которые отечественный финансово-ориентированный бизнес очевидно не в состоянии. Об этом свидетельствуют и авария на Саяно-Шушенской ГЭС и масштабы лесных пожаров лета 2010 года.

3. Переход России к финансово-ориентированной экономике практически обесценил накопленный за 200 лет национальный потенциал в области создания и производства сложных технических систем. А это запустило цепную реакцию обесценивания естественных наук, высшего инженерного образования, общего обязательного естественно-научного среднего образования и тем самым предопределило неизбежность деградации всего промышленно-научно-образовательного комплекса, унаследованного Россией от СССР. При этом деградация национальных школ технических наук и инженерного образования влекут за собой наиболее тяжелые долговременные последствия. За этой деградацией с лагом 5-10 лет неизбежно следуют деградация среднего образования, и далее невозможность возрождения, даже в отдаленном будущем, сколь-нибудь конкурентоспособного на мировом рынке национального реального сектора экономики. Перефразируя слова президента США Б. Обамы[3] можно сказать, что *отстав в естественно-научном образовании, государство отстанет и во всех остальных областях глобальной мировой экономики.*

Из всего изложенного выше следует очевидный вывод: расплатой за попытку обретения Россией статуса финансовой державы оказывается необратимая утрата уже в среднесрочной перспективе сколь-нибудь значимых позиций на мировых рынках сложных технических систем, за которой последует утрата каких-либо национальных возможностей по формированию и сохранению национальной природной и технической среды

---

[2] Акционеры же могут своими решениями изымать на личное потребление непомерно большую долю прибыли предприятия, или даже всю прибыль плюс оборотные средства, разоряя тем самым предприятие и уничтожая рабочие места в России. Впечатляющие примеры подобного рода приведены в статье газеты **Ведомости** от 17.08.2009: ***Постпикалевская Россия****: Идеология временщиков* [2].

[3] «... мы знаем, что страна, которая опередит нас в образовании сегодня, завтра обгонит нас и в других областях» [3]

обитания. Подчеркнем, что будет утрачена и защитная компонента национальной среды – индустрия создания и эксплуатации систем вооружения и военной техники [1].

4. С кризисом 2008 года мировое сообщество стало приходить к осознанию того, что ***«триумф либерализма, обернулся пирровой победой, непомерно разросшаяся, неконтролируемая финансовая деятельность оборачивается беспорядком и общим кризисом».*** Ведущие мировые экономисты стали задаваться вопросами: ***«Можно ли доверить ответственность за принятие производственных решений операторам рынка, которые заботятся только о краткосрочной рентабельности и не имеют компетенций в индустриальной сфере***?» [4]; ***«Банковская и финансовая сферы росли колоссальными темпами за последние два десятилетия, и выросли до того, что они сейчас являются крупнейшими глобальными индустриями, если смотреть на доход, долгосрочную выгодность и долю ВВП. Странно, что индустрия, единственной функцией которой является превращение сбережений в реальные инвестиции, стала так доминировать. Мы должны выяснить, как получилось, что эта роль посредника так дорого обходится обществу?»*** [5].

В России подобные вопросы пока не ставятся, и мы пожинаем плоды "экономики денег", в основе которой лежит принцип «максимальная прибыль за минимальное время».

Литература

[1] Е.П. Велихов, В.Б. Бетелин, А.Г. Кушниренко «Промышленность, инновации, образование и наука в России», М. - НАУКА 2009 140 с.

[2] Газета Ведомости №152 (2422) от 17.08.2009, Никита Кричевский, Владислав Иноземцев. Первая часть статьи «Постпикалевская Россия: Идеология временщиков»
http://www.vedomosti.ru/newspaper/article.shtml?2009/08/17/209939

[3] Речь президента США Б. Обамы перед Национальной Академией Наук США 27 апреля 2009 г., перевод на русский, газета «Троицкий вариант» 26 мая 2009 г.
http://www.scientific.ru/trv/2009/029/obama.html

[4] К новому Бреттон-Вудсу, К. Матье, А. Стердиньяк – *Центр экономических исследований при Институте политических наук, Париж*

Оригинал: Mathieu C., Sterdyniak H. Vers un nouveau Bretton Woods // Lettre de l'OFCE. – P., 2008. – 5 dec.# 303 // Mode of access: www.ofce.sciences-po.fr/pdf/lettres/303.pdf
Перевод с французского специально для сайта «Перспективы» http://www.perspektivy.info/oykumena/ekdom/k_novomu_bretton_vudsu_2008-11-18-41-14.htm

[5] Глава британской службы финансового надзора Эдэйр Тернер. «Часть финансовых инноваций – совершенно бесполезны в социальном смысле», Новостной ресурс slon.ru, публикация от 29.08.09 http://slon.ru/articles/123500/

[6] The Role of EXASCALE Computing in ENERGY Security, DoE SciDAC Review, 2009, www.scidacreview.org/1001/html/energy.html

[7] Россия в мировой энергетике XXI века, Велихов Е. П., Гагаринский А.Ю., Субботин С. А., Цибульский В. Ф., Москва, ИздАТ 2006

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ АГРОМЕТЕОРОЛОГИЧЕСКИХ ФАКТОРОВ ДЛЯ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ МОДЕЛЕЙ ПРОДУКТИВНОСТИ ЗЕРНОВЫХ КУЛЬТУР[1]

Н.В. Гавриловская
Алтайский государственный университет
e-mail: Gavrilovskayanv@gmail.com

Аннотация

В докладе рассмотрена актуальная научная задача, состоящая в разработке алгоритма моделирования агрометеорологических факторов для информационного обеспечения моделей продуктивности зерновых культур, основанного на принципе аналогичности. При решении задачи прогнозирования урожайности сельскохозяйственных культур с использованием математических моделей продуктивности возникает проблема оценки агрометеорологических параметров от даты прогноза до окончания вегетационного периода. Решение данной проблемы базируется на использовании технологии формирования сценариев погодных реализаций с помощью лет-аналогов и генератора погодных данных.

Интеллектуальным ядром данного алгоритма является технология определения лет-аналогов, которая является оригинальной и имеет научную и практическую значимость. Теоретическая значимость проведенных исследований, определяется возможностью применения принципа аналогичности к решению задач прогнозирования агрометеорологических факторов и оценке урожайности зерновых культур. Практическая значимость состоит в разработке математической модели и алгоритмов обработки агрометеорологической информации в условиях различного вида неопределенностей. Это способствует дальнейшему развитию и применению методов математического моделирования и современных информационных технологий для установления количественных зависимостей формирования урожая от агрометеорологических факторов, а также в области упреждающего прогнозирования урожайности зерновых культур.

1. Введение

Для поддержки сельскохозяйственных производителей и управления ценовыми рисками с сентября 2006 г. в России реализуется проект по организации рынка биржевых торгов зерном с использованием механизма форвардных и фьючерсных контрактов, которые заключаются, начиная с марта каждого года. В связи с этим субъектам аграрной сферы необходимо обладать аналитической информацией о складывающихся погодных условиях, их возможном воздействии на объекты сельскохозяйственного производства и ожидаемую продуктивность зерновых культур.

Сдерживающим моментом в решении этих задач является погодный фактор, компоненты которого (среднесуточная температура воздуха, суточная сумма осадков и т.д.) могут изменяться в широком диапазоне от года к году и в течение самого периода вегетации растений. При обосновании стратегии ведения сельского хозяйства экономический эффект должен быть получен на всем множестве возможных погодных реализаций в многолетнем разрезе. Однако на стадии планирования отсутствует любая информация о метеорологической ситуации последующего периода, и необходимо принимать решения, исходя из возможного спектра реализаций агрометеорологических факторов для данной климатической зоны.

Для решения практических задач оценки урожайности в службах Росгидромета используются достаточно простые по структуре эмпирические и физико-статические модели, с помощью которых невозможно учесть реальную изменчивость погодных условий и варьирующих в больших пределах других факторов формирования урожая. В настоящее время разработаны эффективные моделирующие комплексы продукционного процесса,

[1] Работа выполнена при поддержке ведомственно-аналитической программы «Развитие научного потенциала Высшей школы 2009-2011» №2.2.2.4/4278.

такие как AGROTOOL (Агрофизический институт, г. Санкт-Петербург, Россия), EPIC (Soil & Water Research Laboratory, USDA-ARS), AGROSIM (Centre for Agricultural Landscape Research, Müncheberg, Germany) и другие, которые еще до практической реализации того или иного агротехнологического мероприятия, способны спрогнозировать его последствия, встраиваясь непосредственно в технологию принятия решений. Однако их использование сдерживается отсутствием необходимой агрометеорологической информации будущего периода. В данной работе приведен алгоритм моделирования агрометеорологических факторов для информационного обеспечения моделей продуктивности зерновых культур.

2. Математическое моделирование погодных сценариев для прогнозирования агрометеорологических факторов

Моделирование погодных сценариев по принципу аналогичности. Решение данной задачи заключается в том, чтобы из всех подмножеств $A_k \in \Omega$ выбрать класс объектов A_{k_0}, наилучшим образом соответствующий, согласно определенным критериям, новому элементу $X^{n+1} = \{x_{i_0 j}^{n+1}\}$, $i_0 = \overline{1, l_0}$; $j = \overline{1, m}$, $l_0 < l$; l обычно принимают равным 365 дням; l_0 – номер суток, с которого производится моделирование погодного сценария. Для оценки влияния погодных условий на формирование урожая требуется на основе исследования совокупности агрометеорологических параметров классифицировать ситуацию в определенный период времени, учитывая ее влияние на состояние растений, точнее, на урожайность.

Вследствие этого сформированный класс объектов A_{k_0} образует ансамбль возможных реализаций погодных условий, который можно описать вектором $G = \left\{G^0, G^1, \ldots, G^{k_0}\right\}$, где G^0 – исследуемый год, $G^1, \ldots, G^{k_0}$ – годы-аналоги.

Тогда прогнозный сценарий $\hat{G}^0(t + l')$, $l' = \overline{l_0, l}$ для исследуемого года можно построить с помощью оптимизационной процедуры:

$$\mathrm{O} = \sum_{t=1}^{l_0 - 1} \left[\hat{G}^0(t) - G^0(t)\right]^2 \to \min, \qquad (1)$$

где $\hat{G}^0(t) = \sum_{i=1}^{k_0} \alpha_i G^i$, $\sum_{i=1}^{k_0} \alpha_i = 1$, $\alpha_i \geq 0$ – показатели аналогичности.

Использование этого метода предполагает, что при работе модели на ее вход подаются фактические погодные данные до того момента, с которого начинается прогнозирование. Для состыковки (сглаживания) фактических данных и данных лет-аналогов фиксируются отклонения фактических данных на дату прогноза и отклонения данных года-аналога. Эти отклонения сглаживаются с помощью динамического звена первого порядка, отфильтровывающего флуктуации.

Пусть отклонения фактической температуры воздуха до даты сева составляют величину $\Delta T_{real}(k)$, а присоединенных данных года-аналога – $\Delta T_{anal}(k)$. Эти отклонения сглаживаются по формуле:

$$\Delta T_{prog}(k+1) = Q \cdot \Delta T_{anal}(k) + (1 - Q) \cdot \Delta T_{real}(k),$$

где показатель сглаживания Q изменяется в пределах $0 < Q < 1$. Полученная в результате этого сглаженная поправка добавляется к климатической кривой, которая распространяется далее на весь период прогноза.

В данном случае речь идет не о прогнозе погоды в общепринятом значении этого слова, а о моделировании некоторого правдоподобного сценария ежедневных метеорологических данных на прогнозный период.

Прогноз начинается с момента заключения форвардных контрактов и уточняется по мере поступления метеорологической информации. Таким образом, ход метеоусловий «адаптируется» к особенностям конкретного года.

Второй метод имеет *стохастический (вероятностный) характер.* Источником новых погодных реализаций служит т.н. *«генератор погоды»*, в котором ежесуточные метеорологические данные, такие как максимальная и минимальная температуры воздуха, минимальная влажность воздуха, скорость ветра, осадки и коэффициент ослабления солнечной радиации моделируются как многомерный случайный нестационарный процесс. Данный подход базируется на авторегрессионной модели, введенной Ричардсоном и модифицированной в лаборатории математического моделирования агроэкосистем Агрофизического института г. Санкт–Петербурга [1].

Стохастический генератор погоды моделирует синтетические ежесуточные ряды метеорологических элементов со статистическими характеристиками, близкими к тем, что и в исторических данных фактических реализаций погоды за 20–30 прошлых лет. Событие дождя в данные сутки оказывает основное воздействие на относительную влажность, температуру и солнечную радиацию, поэтому генератор погоды, в первую очередь генерирует осадки. Это событие описывается простой цепью Маркова с условными вероятностями, а количество осадков в дождливый день рассчитывается как случайная величина, подчиняющаяся закону гамма–распределения. Хотя эти модельные временные ряды могут и не давать истинное количество выпавших на землю осадков, но интегральные характеристики модельного и фактического временного ряда схожи.

Среднесуточное значение скорости ветра также моделируется как случайная величина, имеющая другие параметры гамма–распределения. Четыре остальных метеопараметра моделируются как случайный многомерный процесс. Первоначально рассчитываются климатические средние значения величин (вычленяется постоянный годовой тренд). После чего центрированные и нормированные отклонения от этого тренда определяются с помощью специального формирующего фильтра первого порядка, характеристики которого однозначно определяются через автокорреляционные матрицы процесса, идентифицированные по фактическим наблюдениям:

$$x(k+1) = A \cdot x(k) + B \cdot e(k) .$$

Здесь k – номер дня реализации; $x(k), x(k+1)$ – 4-х мерные вектора, компонентами которых являются метеорологические параметры; A, B – матрицы коэффициентов фильтра; $e(k)$ – 4-х мерный дискретный белый шум. После вычисления они добавляются к климатическим нормам, образуя в целом составляющие погодного сценария.

Итогом моделирования погодных сценариев является «веер» возможных траекторий формирования урожая и соответствующие ему множества возможных значений потенциального вероятностного прогноза ресурсов продуктивности. Понятие выработки сценария погоды не означает, что в результате будет получена реализация метеоусловий, когда-либо встречавшихся в данной местности.

Основная цель моделирования погодных сценариев заключается в том, что эта процедура, будучи использована в качестве входных данных математических моделей продукционного процесса, даст результат, решающий задачу прогноза того или иного параметра продукционного процесса. Таким образом, полученные в результате моделирования метеорологические ситуации присоединяются к имеющимся фактическим метеоусловиям, образуя полный набор ежедневных входных данных для математических моделей продуктивности зерновых культур.

3. Апробация системы моделирования погодных сценариев для оценки урожайности яровой пшеницы в моделях продуктивности зерновых культур

Оценка урожайности яровой пшеницы осуществлялась в условиях особой аграрной зоны Алтайское Приобье на независимом материале по данным 2007–2010 гг. При этом на интервале до даты составления прогноза используются фактические погодные условия данных лет. На этапах прогноза считаем, что фактические погодные условия неизвестны.

Для поэтапной оценки урожайности были определены следующие даты прогноза: на дату заключения форвардных контрактов (a); дату сева (b); в основные периоды вегетации:

первые две декады после сева (c); на плановые даты оценки урожайности службами Росгидромета (21 июня (d), 21 июля (e)).

Предварительно была проведена проверка на корректность агрометеорологических параметров и сформирована таблица данных для проведения классификации в программном комплексе. При классификации было определено оптимальное количество кластеров и классы лет-аналогов по степени влияния фактических агрометеорологических факторов на урожайность [2]. В таблице 1 приведены классы лет-аналогов, сформированных на дату сева для, 2007–2010 гг.

Из каждого класса лет-аналогов выбраны по два представителя (таблица 1) с максимально схожими с соответствующим исследуемым годом обобщенными метеорологическими характеристиками (уточнение лет-аналогов). Значение возможной урожайности для каждого года вычислялось по формуле

$$y_p = \beta y_1 + (1-\beta) y_2 ,$$

где y_i, $i = 1,2$ – урожайность уточненных лет-аналогов, β – параметр, определяемый оптимизационной процедурой.

Таблица 1

Прогноз урожайности зерновых культур по годам-аналогам

Исследуемые годы	Годы-аналоги	Уточненные годы-аналоги	Значение параметра β	Расчетная урожайность, ц/га	Фактическая урожайность, ц/га	Ошибка прогноза, %
2007	1961, 1979, 1989, 1990, 1991, 2005	1989, 1990	0,85	18,65	17,6	6,0
2008	1980, 1988, 1993, 2003, 2006	2003,2006	0,95	19,85	18,9	5,0
2009	1985, 1987, 2006	1985, 1987	1	24,2	35,5	31,8
2010	1972, 1989, 2005	1989, 2005	0,3	21,44	19,4	10,5

На следующем этапе исследования от даты прогноза строим веер возможных погодных реализаций, смоделированных по годам-аналогам и с помощью генератора погодных данных.

Результаты численных экспериментов по прогнозированию погодных сценариев $\hat{G}^0(t+l')$, $l' = \overline{l_0, l}$ для исследуемых лет с помощью критерия (1) приведены на рис. 2 (прогноз среднесуточных температур воздуха и суммы осадков с 22.06 по 21.07. 2008 г.)

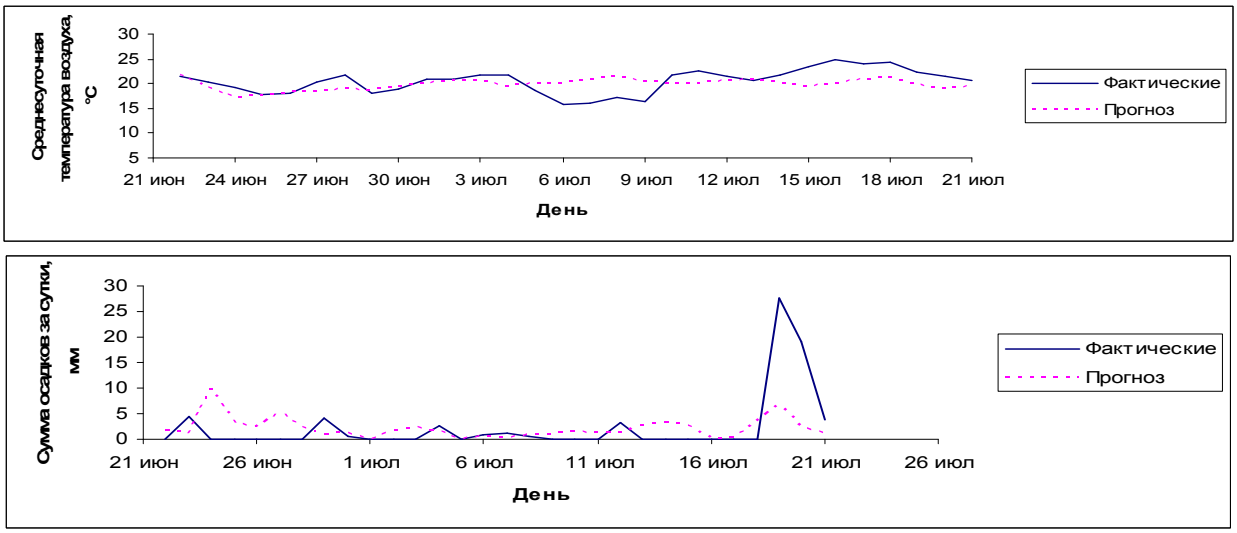


Рис. 2. Расчетные и фактические погодные сценарии

С помощью моделей продуктивности зерновых культур AGROTOOL и EPIC, адаптированных автором к условиям Западной Сибири, выполнены расчеты по развитию яровой пшеницы, и оценена урожайность в зависимости от сгенерированных различных вариантов погодных условий.

В таблице 2 представлены результаты оценки урожайности яровой пшеницы в условиях Алтайского края на дату сева.

Таблица 2

Результаты прогноза урожайности яровой пшеницы на дату сева

2007 г.		$\hat{G}^0(t+l')$	Годы-аналоги		Погодные сценарии				
			1989	1990	1	2	3	4	5
Фактическая	17,6	-	18,0	20,6	-	-	-		-
Прогноз AGROTOOL	17,1	17,2	16,9	18,4	32,1	31,1	22,7	21,1	37,4
Прогноз EPIC	19,2	16,9	21,1	19,2	29,5	30,1	16,2	22,7	38,5
ε_o, AGROTOOL/ EPIC	2,8/ 9,1	2,3/ 4,0	4,0/ 19,9	4,5/ 9,1	82,4/ 67,6	76,7/ 71,0	29,0/ 8,0	19,9/ 29,0	112,5/ 118,8
ε_a, AGROTOOL/ EPIC	0,5/ 1,6	0,4/ 0,7	0,7/ 3,5	0,8/ 1,6	14,5/ 11,9	13,5/ 12,5	5,1/ 1,4	3,5/ 5,1	19,8/ 20,9
P, AGROTOOL/ EPIC	0/0	0/0	0/0	0/0	1/1	1/1	1/0	0/1	1/1
O, AGROTOOL/ EPIC	97/91	98/96	96/80	95/91	18/32	23/29	71/92	80/71	-/-
2008 г.		$\hat{G}^0(t+l')$	Годы-аналоги		Погодные сценарии				
			2003	2006	1	2	3	4	5
Фактическая	18,9	-	15,4	25,4	-	-	-	-	-
Прогноз AGROTOOL	19,9	19,5	17,4	21,5	22,1	25,7	27,6	16,1	3,3
Прогноз EPIC	17,3	18,0	16,8	22,3	20,4	25,1	26,4	15,2	8,7
ε_o, AGROTOOL/ EPIC	5,3/ 8,5	3,2/ 4,8	7,9/ 11,1	13,8/ 18,0	16,9/ 7,9	36,0/ 32,8	46,0/ 39,7	14,8/ 19,6	82,5/ 54,0
ε_a, AGROTOOL/ EPIC	1,0/ 1,6	0,6/ 0,9	1,5/ 2,1	2,6/ 3,4	3,2/ 1,5	6,8/ 6,2	8,7/ 7,5	2,8/ 3,7	15,6/ 10,2
P, AGROTOOL/ EPIC	0/0	0/0	0/0	0/0	0/0	1/1	1/1	0/0	1/1
O, AGROTOOL/ EPIC	95/92	97/95	92/89	86/82	83/92	64/67	54/60	85/80	17/46
2009 г.		$\hat{G}^0(t+l')$	Годы-аналоги		Погодные сценарии				
			1985	1987	1	2	3	4	5
Фактическая	35,5	-	26,1	22,8	-	-	-	-	-
Прогноз AGROTOOL	25,5	31,8	27,6	25,7	21,2	25,3	34,1	30,2	34,9
Прогноз EPIC	23,1	23,7	28,3	21,2	24,1	22,7	29,8	27,6	31,2
ε_o, AGROTOOL/ EPIC	28,2/ 34,9	10,4/ 33,2	22,3/ 20,3	27,6/ 40,3	40,3/ 32,1	28,7/ 36,1	3,9/ 16,1	14,9/ 22,3	1,7/ 12,1
ε_a, AGROTOOL/ EPIC	10,0/ 12,4	3,7/ 11,8	7,9/ 7,2	9,8/ 14,3	14,3/ 11,4	10,2/ 12,8	1,4/ 5,7	5,3/ 7,9	0,6/ 4,3
P, AGROTOOL/ EPIC	1/1	0/1	1/1	1/1	1/1	1/1	0/1	1/1	0/0
O, AGROTOOL/ EPIC	72/65	90/67	78/80	72/60	60/68	71/64	96/84	85/78	98/88
2010 г.		$\hat{G}^0(t+l')$	Годы-аналоги		Погодные сценарии				
			1989	2005	1	2	3	4	5
Фактическая	19,4	-	18,0	14,1	-	-	-	-	-
Прогноз AGROTOOL	21,6	20,1	19,1	17,6	37,4	27,4	35,3	28,9	27,3
Прогноз EPIC	21,1	17,7	20,6	17,2	31,2	24,7	23,1	26,2	30,1
ε_o, AGROTOOL/ EPIC	11,3/ 8,8	3,6/ 8,8	1,5/ 6,2	9,3/ 11,3	92,8/ 60,8	41,2/ 27,3	82,0/ 19,1	49,0/ 35,1	40,7/ 55,2
ε_a, AGROTOOL/ EPIC	2,2/ 1,7	0,7/ 1,7	0,3/ 1,2	1,8/ 2,2	18,0/ 11,8	8,0/ 5,3	15,9/ 3,7	9,5/ 6,8	7,9/ 10,7
P, AGROTOOL/ EPIC	0/0	0/0	0/0	0/0	1/1	1/1	1/0	1/1	1/1
O, AGROTOOL/ EPIC	89/91	96/91	99/94	91/89	8/39	59/73	18/81	51/65	59/45

Обозначения в таблице 2: ε_o - относительная ошибка прогноза (%), ε_a - абсолютная ошибка прогноза (ц/га), P - оценка прогноза относительно допустимой погрешности (превышает - 1, не превышает - 0), O - оправдываемость прогнозов (%).

Разброс значений урожайности по погодным сценариям случайных реализаций составляет от 3,25 до 37,4 ц/га (результат не предсказуем), по годам-аналогам - от 16,9 до 31,8 ц/га. Анализ модельных расчетов (таблица 2) свидетельствует о правдоподобных результатах моделирования погодных сценариев отражающих в величине урожайности.

Результаты поэтапного прогнозирования урожайности яровой пшеницы по двум моделям (AGROTOOL и EPIC) в различные периоды приведены в таблице 3.

Таблица 3

Результаты поэтапного прогноза урожайности яровой пшеницы

Годы	По AGROTOOL					По EPIC					Факти-ческая
	(a)	(b)	(c)	(d)	(e)	(a)	(b)	(c)	(d)	(e)	
Смоделированные погодные сценарии по принципу аналогичности											
2007	19,4	17,2	18,2	17,1	17,1	16,2	16,9	20,1	17,0	17,0	17,6
2008	20,05	19,5	17,8	19,9	19,4	17,6	18,0	20,4	18,2	18,1	18,9
2009	24,2	31,8	24,7	25,5	27,3	20,5	23,7	27,5	25,7	26,1	35,5
2010	20,9	20,1	18,5	20,1	18,9	16,9	17,7	22	18,3	18,5	19,4

Оценка качества прогнозов и, соответственно, успешность разработанной технологии прогноза урожайности яровой пшеницы производилась по их оправдываемости относительно допустимой погрешности прогноза ($0{,}67\sigma, \sigma$ – среднее квадратическое отклонение) согласно инструкции по оценке оправдываемости агрометеорологических прогнозов.

4. Заключение

Верификация (калибровка) предложенного алгоритма моделирования агрометеорологических факторов была произведена с помощью технологии определения лет-аналогов по трем независимым выборкам, сформированным по различным критериям. Обоснование применимости разработанного алгоритма генерирования погодных сценариев с целью оценки урожайности зерновых культур подтверждается качественным анализом полученных результатов (таблица 4).

Таблица 4

Анализ результатов оценки урожайности зерновых культур по моделям

Факторы	AGROTOOL	EPIC
Количество составленных прогнозов	90	90
Число оправдавшихся прогнозов	84	79
Число не оправдавшихся прогнозов	6	11
Доля оправдавшихся прогнозов, %	93	88
Доля не оправдавшихся прогнозов, %	7	12
Средняя абсолютная ошибка прогноза, ц/га	1,4	2,7
Средняя относительная ошибка прогноза, %	11,3	15,4

Анализ полученных результатов показал хорошее совпадение фактической урожайности с расчетной по погодным сценариям лет-аналогов, это говорит о том, что смоделированные погодные сценарии могут успешно применяться для текущего планирования и перспективного прогнозирования урожайности яровой пшеницы на территории особой аграрной зоны Алтайское Приобье.

ЛИТЕРАТУРА

[1]. Топаж А.Г. Моделирование суточных метеоданных как входного сигнала модели продукционного процесса // Почва и растение – процессы и модели: сб. науч. тр. СПб., 1992. С. 79 - 86.

[2]. Гавриловская Н.В., Хворова, Л.А. Разработка алгоритма определения года-аналога для оценки урожайности зерновых культур в условиях Алтайского края // Известия АлтГУ. 2007. №1 (53). С. 66-67/

ОЦЕНКА И ПРОГНОЗИРОВАНИЕ ТЕХНИЧЕСКОГО СОСТОЯНИЯ ГИДРОТЕХНИЧЕСКИХ СООРУЖЕНИЙ ГЭС

Д.Б. Воробьева

Учреждение Российской академии наук Конструкторско-технологический институт вычислительной техники Сибирского отделения РАН

vorobyeva.darya@gmail.com

Е.П. Золотухин

zepzoloto@mail.ru

Аннотация

Предлагается система, предназначенная для исследования и прогнозирования поведения плотины под действием различных нагрузок, в том числе и землетрясений, оценки технического состояния плотины и поддержки принятия оперативных решений с целью обеспечения безопасности и предотвращения аварийных ситуаций.

Оценка и прогнозирование технического состояния ГТС осуществляется с помощью численных моделей и натурных наблюдений сооружения, по которым рассчитываются показатели технического состояния сооружения. В разрабатываемой системе с помощью трехмерной расчетной модели осуществляется статический и динамический анализ ГТС, вычисляются напряжения, деформации, температуры, собственные частоты, формы колебаний и др. В качестве численного метода используется метод конечных элементов [1]. На основе сравнения данных натурных наблюдений и расчетных данных при сходных внешних воздействиях происходит идентификация модели. Расчет позволяет выяснить, соответствуют ли измеренные параметры состояния плотины внешним условиям, или в сооружении возникли изменения, нарушения.

ЛИТЕРАТУРА

[1]. Пельмутер А.В., Сливкер В.И. Расчетные модели сооружений и возможность их анализа. Киев: Сталь, 2002.

НЕЙРОКИБЕРНЕТИЧЕСКИЕ ПРИНЦИПЫ ФИЗИОЛОГИЧЕСКОЙ ОРГАНИЗАЦИИ ПОВЕДЕНИЯ

К.Н. Дудкин, И.В. Чуева

Институт физиологии им. И.П. Павлова РАН, Санкт-Петербург, Россия,
cognition@pavlov.infran.ru

Формирование кибернетики в середине XX века, во многом, подготовлено предшествующими достижениями в мировой физиологии и психологии. И.М. Сеченов и И.П. Павлов и их последователи, создав русскую физиологическую школу, выявили ряд основных принципов организации поведения. Поставив задачу изучения организации поведения *«всего нераздельно животного организма»* (по Павлову), включающего психическую и физиологическую деятельность, они не только определили вектор развития физиологии, но и стали предтечами кибернетики.

Сеченов впервые (1863) ввёл в физиологию принцип зависимости восприятия (сенсорных процессов) от прошлого опыта (памяти) и рассмотрел чувствование, как *сигнал*, заложив идею *саморегуляции* в рефлекторную схему. Павлов не только открыл и изучил механизмы *условного рефлекса*, но и высказал идею о механизмах, предшествующих сенсорному анализу (*рефлекс цели*, 1916) и представляющих собой *«стремление к обладанию определённым раздражающим предметом»* (*«стремление к цели»)*. Идеи Павлова развили его последователи: «функциональная система» и прогнозирующий механизм - акцептор действия (Анохин); «физиология активности» с механизмами «программирования потребного будущего» (Бернштейн). Позднее «целеустремлённая система» стала одним из центральных понятий в кибернетике (Rosenblueth, Wiener, Bigelow, 1943; Ackoff, Emery, 1972).

Заложив основы науки о поведении, Сеченов и Павлов ввели в физиологию новые научные понятия, адекватные новым задачам. С*игнал, анализатор, анализ сигналов, центральное торможение, саморегуляция, мышление, образы, признаки* (по Сеченову), *условный и ориентировочный рефлексы, анализ и синтез сигналов, подкрепление, мотивация, рефлекс цели, рефлекс свободы, мышление, первая и вторая сигнальные системы* (по Павлову) составили понятийную основу для изучения нейрофизиологической организации поведения. Многие из этих понятий были использованы в кибернетике.

Данная работа основана на многолетнем изучении организации поведения обезьян (макак резус) и является развитием концепций Сеченова и Павлова. Полагаем, что поведение реализуется, исходя из биологических и социальных потребностей, формирующих цели, мотивации, установки, управляющие процессами восприятия и механизмами памяти. Объекты внешней среды, воздействуя на организм, определяют формирование разнообразных когнитивных структур, образующих “модель мира” в его памяти (Дудкин, 1985; 2007; Дудкин, Чуева, 2008). Нейрофизиологические принципы организации поведения, сформулированные в концепциях Сеченова и Павлова, во многом обусловливают современную методологическую базу когнитивной нейрофизиологии и нейрокибернетики. Исходя из этого, считаем необходимым - рассмотреть их в этом современном контексте, тем более, что во второй половине 20-го века во многие физиологические направления хлынули потоки новых идей, концепций, методов, понятий и новой терминологии. Большинство из них основывалось на информационном и системном (кибернетическом) подходах.

Сеченов сформулировал (1863) физиологические принципы поведения, основу которого составляет психическая деятельность. Он рассмотрел физиологический феномен - «рефлекс» в качестве глобального принципа взаимодействия организма с внешней средой, трактуя его как процесс, включающий три основные составляющие: чувствование, психическую деятельность и движение. Психические процессы по Сеченову - звено рефлекса между сенсорными входами и моторными выходами, локализованное в головном мозгу. Были определены две основные функции чувствования: инициирование характеристик действия (реакций организма) и их регуляцию. Впервые он рассмотрел чувствование, как сигнал, заложив идею саморегуляции в рефлекторную схему (обратная связь - кольцевое управление движением).

Понимание воспринимаемой информации базируется на сведениях, хранящихся в памяти, через которую осуществляется связь между программами восприятия и моторными реакциями, вызываемыми сокращениями мышц, их результаты - мышечные сигналы - посредством обратной связи направляются в центры головного мозга. В дальнейшем управлении поведением организма основную роль играют уже психические процессы. Это управление производится в соответствии с мышечными сигналами, которые отображают пространственно-временную структуру изменений во внешней среде. Сеченов впервые установил (1862), что наряду с центрами, возбуждающими рефлексы, существуют центры, которые рефлексы тормозят. Совместная деятельность этих центров определяет организацию поведения. Он считал, что «центральное торможение», локализованное в головного мозгу, является нервным механизмом психических функций – воли и мышления.

Восприятие и понимание объектов внешнего мира по Сеченову возникают благодаря памяти. При взаимодействии организма с внешней средой формируются представления, понятия, образы, признаки (в современной трактовке все эти понятия - когнитивные структуры), и эта информация хранится в долговременной памяти. Узнавание по Сеченову обусловлено способностью организмов - *выделять и запоминать наиболее характерные признаки, а процесс узнавания осуществляется посредством сопоставления по тождеству результатов текущего сенсорного анализа с репродуцируемым содержанием долговременной памяти.* Процессы узнавания включают обработку сенсорной информации и выделение признаков, в результате формируются впечатления; репродукцию из памяти содержащихся в ней признаков и образов; сопоставление по тождеству полученных впечатлений и воспоминаний о них. Признаки и образы по Сеченову – продукты физиологической структурно-функциональной организации. Эта идея - одна из основных в современной когнитивной нейрофизиологии.

Работы Павлова и его школы - новая эпоха в развитии науки о поведении – «высшей нервной деятельности». Павлов, также как и Сеченов, считал, что организм и внешняя среда – единая система, и во всех своих исследованиях руководствовался системным подходом. Он изучал не только внутренние механизмы саморегуляции физиологических процессов, но и механизмы, обеспечивающие непрерывное взаимодействие и неразрывное единство организма и окружающей среды. На протяжении всей своей научной жизни Павлов интересовался взаимодействием нервной системы с деятельностью внутренних органов. Эксперименты, связанные с изучением пищеварительной системы, в начале XX века привели его к открытию и дальнейшему всестороннему изучению условных рефлексов. Павлов и его коллеги обнаружили, что один только вид или запах пищи действуют как сигнал для образования слюны. Следовательно, *«Можно любое явление внешнего мира сделать временным сигналом раздражающего слюнные железы объекта, если раздражение слизистой оболочки этим*

объектом связать во времени один или несколько раз с действием этого явления на соответствующую воспринимающую поверхность тела» (Павлов, 1949. с. 350).

И.П. Павлов впервые экспериментально доказал, что работа пищеварительной системы зависит от нервной системы и обусловлена нервными и психическими управляющими механизмами. Эти исследования привели к открытию *условных рефлексов*. Именно за эти исследования И.П. Павлову в 1904 году была присуждена Нобелевская премия. По Павлову, *условный рефлекс* - элементарное психическое явление, которое является и чисто физиологическим явлением. После 1902 г. Павлов полностью сконцентрировался на изучении высшей нервной деятельности, лежащей в основе поведения. В этих исследованиях он опирался на работы Сеченова, посвященные введению физиологических основ в психические процессы, и неоднократно подчеркивал своё единомыслие с ним. В исследованиях Павлова с сотрудниками был выработан и окончательно сформирован новый язык описания механизмов организации поведения – *высшей нервной деятельности*. У*словный рефлекс* Павлов использовал, как методологический инструмент для выявления закономерностей поведения. Положение о взаимодействии организма с внешней средой было им сформулировано в виде одного из основных принципов рефлекторной деятельности.

В поведении животных и человека по Павлову много сходного, однако, человек имеет важное отличие - речь. Он ввёл новое понятие - *«вторая сигнальная система»,* характеризующее *«чрезвычайно сложный дополнительный инструмент»*, позволяющий использовать смысл речевых или письменных сообщений. *«Вторая сигнальная система»* - бесконечно более сложная, чем *«первая сигнальная система»* (первичные сигналы и символы, воспринимаемые животными). Положение Павлова о речи как второй сигнальной системе, следует рассматривать в качестве важного первого шага в создании современного направления, связанного с изучением возможных принципов и моделей порождения и восприятия речи человеком.

По Павлову организм взаимодействует с внешней средой на основе анализа воспринимаемой информации и синтезирующей деятельности нервной системы. Он разработал программу изучения *«анализаторов», «центральные концы»* которых локализованы в коре больших полушарий. Объекты внешней среды, воздействуя на организм по двум основным каналам (*условный и безусловный* стимулы), вступают с ним во взаимодействие. Сенсорная информация (условный стимул) проходит несколько этапов переработки: *элементарный анализ* и *элементарный синтез* в *периферических концах анализаторов*; *высший анализ и высший синтез в центральных концах анализаторов*. При анализе отображения объектов раскладываются на *единицы* (формируются описания объектов); при синтезе части вновь соединяются в целое. При выработке условного рефлекса механизмы мотивации, активированные безусловными стимулами, и механизмы реализующие *рефлекс цели*, активно участвуют в анализе сенсорной информации и синтезе результатов анализа на всех этапах, формируя *биологическое значение* индифферентного ранее объекта (условного стимула) и стремление к реализации цели.

Очевидно, Павлов был хорошо знаком с взглядами Сеченова, который более 100 лет назад предложил концепцию о признаках и образах и утверждал, что они являются *продуктами структурно-функциональной физиологической организации* (Сеченов, 1952, с.355). То, что Павлов не использовал понятия: «признаки» и «образы», хотя, в неявном виде они присутствуют в его концепциях, можно понимать, как максималистскую позицию в принципиальном

отказе от применения любых психологических терминов. Сеченов сформулировал концепцию о *конкретном предметном мышлении* (Сеченов, 1952, с. 358) при *сопоставлении по тождеству реального впечатления и воспоминания о нём при узнавании предмета.* «*…Это есть наипростейшая форма мысли, свойственная даже животным, форма, с которой начинается умственная жизнь»* (Сеченов, 1952, с. 496). По Сеченову при узнавании предметов необходимо различать их свойства и характеристики: *"...различение в предметах их свойств есть уже род мышления предметами и их свойствами..."* (Сеченов, 1952, с. 358). Фактически в своих работах Сеченов впервые дал физиологическое объяснение взаимодействия восприятия и мышления.

Информация об объектах внешнего мира по Павлову при «анализе» расчленяются на единицы, части, которые представляют собой нервные процессы. По Сеченову анализ предметов заключается в выделении их признаков и свойств. При «синтезе» по Павлову из составляющих, полученных при анализе, под управлением процессов мотивации и цели вновь формируется нечто целое (биологически значимый образ?). По Сеченову образ – представление, репродуцируемое из памяти для сравнения по тождеству реального впечатления с воспоминанием о нем. Все нервные процессы по Павлову - продукты функционального разнообразия структурной организации коры и его интересовали правила, по которым производится и анализ, и синтез в коре больших полушарий. В процессе познания Павлов, интерпретируя результаты исследований, включал в свои концепции новые понятия. *«В окончательном результате большими полушариями собаки постоянно производится в разнообразнейших степенях как анализирование, так и синтезирование падающих на них раздражителей, что можно и должно назвать элементарным, конкретным мышлением. Это мышление таким образом обусловливает совершенное приспособление, более тонкое уравновешивание организмом окружающей среды»* (Павлов, 1949, с. 473). Это положение соответствует современному понятию «обогащение внутренней модели мира» в процессах формирования поведения, ибо именно конкретное мышление лежит в основе создания и дальнейшего развития модели мира.

Положения Сеченова и Павлова являются фундаментом для многих современных физиологических и нейрокибернетических парадигм. Например, в процессе организации поведения при обучении зрительному узнаванию обычно постулируется три уровня: низкий, связанный с процессами обработки и сегментации объектов; средний, заключающийся в символическом описании свойств, выделенных в процессе сегментации; высокий, на котором происходят понимание воспринятой информации, классификация и узнавание объектов. Обучение основывается на способности организмов выделять и запоминать наиболее характерные разделительные признаки. Установлено, что условно-рефлекторное поведение основано на взаимодействии сенсорных, когнитивных и управляющих (мотивации, цели и внимания) процессов. Условные временные связи в этом поведении устанавливаются между исполнительными реакциями и когнитивными структурами (разделительными признаками), формирующимися в коре головного мозга в процессе обучения инвариантному зрительному узнаванию. Те же структуры выступают в роли установок (цели) для выделения разделительных признаков в текущем поведении. Фактически в процессе условно-рефлекторного поведения формируется целеустремлённая функциональная система – модель результата действий, она и реализует требуемый результат.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ ДЛЯ ОЦЕНОК ЭКОЛОГИЧЕСКИХ РИСКОВ В УСЛОВИЯХ ИЗМЕНЯЮЩЕГОСЯ КЛИМАТА*

В.В. Пененко
Институт вычислительной математики и математической геофизики СО РАН
e-mail:penenko@sscc.ru

Взаимосвязанные процессы экологии и климата в настоящее время являются предметом интенсивных исследований. Это обусловлено тем, что с повышением технической сложности и удельной мощности производственных объектов увеличивается опасность крупномасштабных техногенных катастроф, которые, в свою очередь, могут приводить к серьезным экологическим последствиям. Современные модели климатической системы, модели переноса и трансформации загрязняющих примесей в газовом и аэрозольном состоянии довольно сложны, поскольку в них учитываются как естественные, так и антропогенные факторы, влияющие на изучаемые процессы.

В докладе представлена концепция природоохранного прогнозирования и проектирования. Её основная идея состоит в том, чтобы с единых позиций рассматривать множество процессов различных пространственно-временных масштабов и разнородную фактическую информацию о наблюдаемом поведении климато-экологической системы. С позиций математики это обеспечивается применением вариационных принципов для формирования общей структуры комплекса математических моделей и алгоритмов их практической реализации на ЭВМ. Для решения природоохранных задач и оценок экологических рисков разработаны методы прямого и обратного моделирования и методы теории чувствительности моделей и целевых функционалов к вариациям входных данных и параметров источников воздействий. Приводятся примеры решения конкретных задач.

* Работа поддержана Программами фундаментальных исследований №4 Президиума РАН и №3 Отделения математических наук РАН, а также проектом РФФИ 11-01-00673.

Вычисление гарантированных оценок в моделях загрязнения потоков с неточно заданными данными

Рогалев А.Н.
ИВМ СО РАН, г.Красноярск
rogalyov@icm.krasn.ru

Со времен классических исследований Стритера и Фелпса (Streeter - Phelps) [1], [2] моделированию и анализу содержания растворенного кислорода (РК) было посвящено значительное внимание в моделях качества воды. Это естественно по двум причинам: с одной стороны, РК является естественным показателем жизнедеятельности водной среды, с другой стороны, поскольку многие параметры загрязнения и качества воды взаимодействуют с РК, его изменения во времени трудно охарактеризовать. Один из более известных критериев загрязнения отходами (БПК), является основной причиной уменьшения содержания РК в воде и, следовательно, ухудшения качества воды. В классической модели Стритера – Фелпса рассматривается система, состоящая из воды и растворенного в ней кислорода и органического вещества. Математически она описывается системой двух обыкновенных дифференциальных уравнений (ОДУ) с соответствующими начальными условиями. Фундаментальной особенностью этой системы является отсутствие обратной связи между концентрацией растворенного кислорода и скоростью окисления органического вещества. Это может приводить к физически некорректным решениям, в которых концентрация растворенного кислорода становится отрицательной величиной. Эта физическая некорректность остается потенциально опасной и для более сложных систем дифференциальных уравнений, полученных в результате усложнения классической модели. Одним из способов улучшения адекватности модели кислородного баланса водотока является включение [4] в классическую систему уравнений Стритера-Фелпса дополнительного уравнения, описывающего зависимость параметра от концентрации С растворенного кислорода.

Для установившихся течений влияние диффузии на загрязнение становится пренебрежимо малой, которой можно пренебречь. Коэффициенты этой системы могут изменяться вдоль потока, но остаются заключенными в заданные области. Описание их с помощью фиксированных значений, выбранных в пределах этих границ, становится недостаточно в силу нескольких причин – особенностей экспериментального определения коэффициентов в потоках. Одновременно выбор случайных коэффициентов для этих систем характеризуется несколькими недостатками, которые исправляют модели с коэффициентами – многозначными величинами. Так, например, уровень значимости оценок, полученных в стохастической модели не превосходит 66 процентов [3], что может приводить к тому, что все траектории этой системы покинут указанные области за исследуемое время. В докладе описывается применение гарантированных методов для оценки решений дифференциальных уравнений моделей, модифицирующих модель Стритера-Фелпса. Приводятся результаты вычислений и графики. Оценки, посчитанные в гарантированных методах, являются достоверными для данных областей в рамках выбранной модели.

[1] Streeter H., Phelps E. A study of the pollution and a natural purification of the Ohio River // U.S. Publ. Health Service Bull. 1925. № 146. p.1-75.
[2] Padgett W., Schultz G., Tsokos C. A random differential equation approach to the probability distribution of BOD and DO in streams. // SIAM J. Appl. Math., 1977.- v.32. - p. 467 - 483.
[3] Padgett W. A stream pollution model with random deoxygenation and reaeration coefficients. // Mathematical Biosciences, 1978.- v.42. - p. 137 - 148.
[4] Пряжинская В.Г., Ярошевский Д.М., Левит-Гуревич Л.К. Компьютерное моделирование в управлении водными ресурсами. М.: Физматлит, 2002.

МОДЕЛИРОВАНИЕ С НЕОПРЕДЕЛЕННОСТЬЮ ЭКОСИСТЕМЫ ОЗЕРА

Е.В. Пахт

ФГОУ ВПО Дальневосточный государственный технический рыбохозяйственный университет, Владивосток, 690091, Россия

e-mail: st-francis@mail.ru

В настоящей работе использовались, предложенные Л. А. Заде [1] идеи применения нечетких множеств для описания неопределенностей в исходных данных о состоянии биологического сообщества.

Моделирование с неопределенностью

Модель динамической экосистемы построена на основе системы дифференциальных уравнений:

$$\dot{x} = f(t, x) \tag{1}$$

Состояние экосистемы описывается n – мерным вектором x, компоненты которого представляют собой биомассы соответствующих блоков.

Условия (*): функция $f(t, x)$ непрерывна по (t, x) и непрерывно дифференцируема по x в окрестности точки $t = 0$. Предположим, что решение системы (1) при любом неотрицательном начальном условии существует и единственно для всех неотрицательных значений t.

Так, как начальное состояние экосистемы известно лишь с некоторой точностью, то можно предположить, что оно принадлежит замкнутому ограниченному множеству $\Omega_0 \subset D \subset R^n_+$. Таким образом, Ω_0 будем рассматривать, как нечеткое подмножество, представленное в виде набора пар «значение вектора x_t состояния экосистемы в заданный момент времени t – степень соответствия (близости к) реальному состоянию экосистемы в заданный момент времени t», где степень соответствия определяется некоторым интервалом значений функции принадлежности (характеристической функции) [1].

Обозначим через $\varphi(t, x)$ функцию принадлежности состояния экосистемы заданному значению x в момент времени t. Характеристическая функция показывает степень принадлежности значения вектора x состояния экосистемы в заданный момент времени t к нечеткому подмножеству Ω_t . Чем больше $\varphi(t, x)$, тем в большей степени вектор $x \in D$ описывает состояние экосистемы в момент времени $t \in T$.

Через $\chi(t, \tau, x_\tau)$ обозначим решение $x(t)$ системы уравнений (1) в момент времени t при условии $x(\tau) = x_\tau$. Пусть Ω_0 замкнутая, ограниченная область в D. Тогда $\Omega_t = \chi(t, t_0, \Omega_0)$ – также замкнутая, ограниченная область в D, $t_0 \in T$.

Утверждение[2]. Пусть $\varphi(t,x) \in C^1\left(R^{n+1}\right)$, задано начальное условие $\varphi(t_0,x) = \varphi_0(x)$, $\varphi_0 \in C^1\left(R^n\right)$ и $\int\limits_{\Omega_t} \varphi(t,x)dx$ не зависит от t. Тогда $\varphi(t, x)$ вычисляется по формуле:

$$\varphi(t,x) = \varphi_0\left(\chi(0,t,x)\right) \cdot \exp\left(-\int\limits_0^t div\, f(\tau, \chi(\tau,t,x))\, d\tau\right). \tag{2}$$

Утверждение доказывается на основе известной формулы Лиувилля [2, 4].

Теорема Лиувилля [6] . Пусть для системы (1) выполняются условия (*).

Пусть $V(t)$ - объем области Ω_t . Тогда

$$\dot{V}(t) = \int_{\Omega_t} div\, f(t,x)dx\,.$$

Математические модели для экосистемы озера Большой Окуненок

Озеро Большой Окуненок расположено в Лужском районе Ленинградской области (заказник «Сяберский»). Географические координаты центра озера: 58°48′22″ N (с. ш.), 29°5′50″ E (в. д.). Физико - географическая характеристика: площадь водной поверхности – 5.8 га (0.058 кв.км.), средняя глубина – 2 м, максимальная глубина – 3 м. Озеро Большой Окуненок относится к разряду высокоторофных водоемов (с высоким содержанием органики)[3].

В экосистеме озера учтено 36 компонентов: две группы бактерий (одиночные и агрегированные, в том числе входящие в состав детрита), четыре размерные группы потребляемого фитопланктона, синезеленые водоросли (цианобактерии), инфузории и коловратки разных трофических предпочтений, хищный и нехищный зоопланктон, включающий различные виды ракообразных разных размеров и стадий развития, личинки комаров разных видов и стадий развития, бентос среднего размера (хищный и нехищный), макробентос, три вида мальков рыб: карп и сиговые рыбы – пелядь и пелчир, взвешенный и донный детрит, растворенное в воде органическое вещество (РОВ).

Так как в системе (1) множество неопределенных параметров, использовался многомодельный подход [5], состоящий в постепенном дизагрегировании первоначальной модели, уточнении переменных. Простейшие математические модели позволяют подкрепить большую дизагрегированную модель некоторым качественным анализом, рассмотреть основные ее динамические свойства, что в значительной мере облегчает дальнейшую работу с ней.

Поскольку описывается один объект, то все модели должны быть согласованы между собой по структуре и по параметрам. Рассмотрены четыре модели. Время измеряется в сутках, плотности масс (биомасс) блоков – в *ккал/м*^{2}.

Первая модель отражает иерархию организмов экосистемы: выделены самые крупные ее представители – рыбы, остальные компоненты поделены на живые и «неживые» – детрит, РОВ и биогены. В схеме потоков вещества фитопланктон разного размера, синезеленые водоросли, две группы бактерий, две группы мелких коловраток, аспланхна, науплии ракообразных, две размерные группы циклопов, диаптомусы, мелкие кладоцеры, две размерные группы дафний, пять групп макробентоса представлены одним блоком (x_1) (на основе данных о биомассах за 1986 год [3]). Объединены в один блок (x_2) три вида рыб. РОВ и разнокачественный детрит (донный и взвешенный) объединены (x_3) с биогенными элементами – минеральными веществами на основе углерода, кремния, фосфора и азота, участвующими в образовании органического вещества, в процессе фотосинтеза, построении растительных организмов (фитопланктона). Таким образом, пресноводная экосистема представляется основными блоками сообществ близких видов, (x_1) и (x_2), и совокупности близких по свойствам веществ (x_3).

Модель М1 описывает идеализированную замкнутую по веществу систему, то есть сумма правых частей дифференциальных уравнений М1 равна нулю.

$$\begin{cases} \dot{x}_1 = v_{31}(x_3)\cdot x_1 - v_{12}(x_1)\cdot x_2 - m_1(x_1)\cdot x_1 \\ \dot{x}_2 = k_{12}\cdot v_{12}(x_1)\cdot x_2 + k_{23}\cdot v_{32}(x_3)\cdot x_2 - m_2(x_2)\cdot x_2 \\ \dot{x}_3 = (1-k_{12})\cdot v_{12}(x_1)\cdot x_2 + (1-k_{23})\cdot v_{32}(x_3)\cdot x_2 + \\ \quad + m_1(x_1)\cdot x_1 + m_2(x_2)\cdot x_2 - v_{31}(x_3)\cdot x_1 - v_{32}(x_3)\cdot x_2. \end{cases} \quad \text{(M1)}$$

Трофические функции скорости выедания блока i на единицу биомассы блока j обозначены $v_{ij}(x_i)$, доля k_{ij} этой выеденной биомассы, идет на рост хищника, то есть на образование продукции блоком j. Такие же функции применяются для описания потребления биогенов фитопланктоном при фотосинтезе, а также для описания потребления планктоном и рыбами РОВ и детрита. Последний формируется за счет неусвоенной пищи и отмершей биомассы, пополняющей детрит со скоростью, определяемой функцией $m_i(x_i)$ – удельные скорости отмирания.

Вторая модель сосредоточена на живой составляющей экосистемы: активно и неактивно движущиеся организмы. В результате сильных упрощений реальной картины построенная модель (М2) не дает точного описания какой-либо конкретной ситуации, но позволяет провести исследование общих свойств объекта. Вторая модель (незамкнутая) представляет взаимодействие планктона (x_1) и рыб (x_2):

$$\begin{cases} \dot{x}_1 = q(x_1) - v_{12}(x_1)\cdot x_2 - m_1(x_1)\cdot x_1 \\ \dot{x}_2 = k_{12}\cdot v_{12}(x_1)\cdot x_2 - m_2(x_2)\cdot x_2. \end{cases} \quad \text{(M2)}$$

В третьей и четвертой моделях реализованы разные гипотезы о трофических взаимодействиях.

Общей закономерностью для водных экосистем является существование двух главных путей переноса энергии – «планктонный» (фитопланктон – нехищный зоопланктон – хищный зоопланктон – рыбы) и «детритно - бактериальный» (отмирающие остатки организмов – детрит – растворенное органическое вещество – микрофлора) [3].

Третья модель. Третья модель (незамкнутая), отражающая «планктонный» путь, получена следующим преобразованием: из первого блока первой модели (М1) выделены фитопланктон и синезеленые водоросли, образующие блок (y_1), остальные компоненты объединены в блок (y_2), три вида рыб образуют блок (y_3). РОВ, разнокачественный детрит, и биогенные элементы рассматриваемые в первой модели (М1), не включены в третью модель (М3). Переменные моделей (М1) – (М2) и модели (М3) связаны следующим образом: $x_1 = y_1 + y_2$, $x_2 = y_3$.

$$\begin{cases} \dot{y}_1 = q(y_1) - v_{12}(y_1)\cdot y_2 - m_1(y_1)\cdot y_1 \\ \dot{y}_2 = k_{12}\cdot v_{12}(y_1)\cdot y_2 - v_{23}(y_2)\cdot y_3 - m_2(y_2)\cdot y_2 \\ \dot{y}_3 = k_{23}\cdot v_{23}(y_2)\cdot y_3 - m_3(y_3)\cdot y_3. \end{cases} \quad \text{(M3)}$$

Значения коэффициентов смертности $m_1^{(0)}$ и $m_2^{(0)}$ вычислены из условия равновесности наблюдаемого состояния системы и условия равенства удельной смертности модели М2 и средней взвешенной удельных смертностей модели М3.

Четвертая модель. Так как среди рыб, образующих блок (y_3) только в рацион пеляди входят синезеленые водоросли, то эта трофическая связь не была включена в модель (М3). Добавление этой связи с учетом того, что пелядь составляет 14.5% биомассы блока (y_3) и синезеленые водоросли – 15% биомассы блока (y_1), в первое и третье уравнения модели (М3) приводит к появлению новых слагаемых, соответствующих перетоку энергии из блока (y_1) в блок (y_3), получим следующую модификацию третьей модели (М3):

$$\begin{cases} \dot{y}_1 = q(y_1) - \nu_{12}(y_1)\cdot y_2 - \nu_{13}(y_1)\cdot y_3 - m_1(y_1)\cdot y_1 \\ \dot{y}_2 = k_{12}\cdot\nu_{12}(y_1)\cdot y_2 - \nu_{23}(y_2)\cdot y_3 - m_2(y_2)\cdot y_2 \\ \dot{y}_3 = k_{13}\cdot\nu_{13}(y_1)\cdot y_3 + k_{23}\cdot\nu_{23}(y_2)\cdot y_3 - m_3(y_3)\cdot y_3. \end{cases} \quad \text{(M4)}$$

Свойства функций во всех моделях предполагаются одинаковыми. Трофические функции $\nu(x)$ и функции роста биомассы $q(x)$ в соответствии с их общими свойствами [7] предполагаются неотрицательными и равными нулю в нуле, строго возрастающими, вогнутыми. Функции $\nu(x)$ и $q(x)$ характеризуют взаимодействия между различными трофическими уровнями. Описание этих взаимодействий представлено функцией Моно [8]: $\nu(x) = \nu^{(0)}\cdot g(x, x_0)$, $q(x) = q^{(0)}\cdot g(x, x_0)$, где $g(x, x_0) = x/(x_0 + x)$. Функция $m_i(x_i) = m_i^{(0)}$ [8] предполагается постоянной и определяет удельную скорость отмирания вида i по причинам не связанным с трофическими отношениями.

Предположим, что средние оценки биомасс блоков ([3], таблица № 6) определяют равновесное состояние экосистемы: суммарная плотность биомассы планктона и бентоса составляет 25.3 ккал/м^2, для рыб – 15.8 ккал/м^2.

Неизвестные коэффициенты $\nu^{(0)}$, $q^{(0)}$ и параметры смертности $m^{(0)}$ выбраны так, что это состояние является равновесием во всех моделях.

Сделаны расчеты функции принадлежности по формуле (2) для всех моделей. Для некоторого начального состояния задается эллипсоид (или эллипс). На ограниченном этой поверхностью (линией) множестве определяются ненулевые значения начальной функции $\varphi_0(x)$ принадлежности в виде параболоида, в двухмерном случае, с максимальным значением в центре и нулевым – на границе эллипса.

Расчеты

Для иллюстрации динамических свойств решений в моделях выбраны различные начальные состояния и проведено сравнение решений во всех моделях.

В модели M1 выбраны несколько начальных состояний с различным биомассами детритного блока. Во всех случаях траектории движения функции принадлежности стремятся к соответствующим асимптотически устойчивым положительным равновесиям на кривой равновесий, представляющей собой однопараметрическую кривую (с параметром x_1, меняющимся на некотором интервале положительных значений). Но нечеткость начального состояния делает эту устойчивость мало значащей, так как небольшие отклонения начальных значений дают значительные расхождения устойчивых состояний на линии равновесий. В модели M2 из всех выбранных начальных состояний решение стремится к асимптотически устойчивому равновесному решению $x_3^* = (25.3;\ 15.8)$, которое лежит на линии положительных равновесий модели M1. Начальные значения функции принадлежности характеризуют на траекторию и время перехода в устойчивое равновесное состояние. Устойчивое положительное равновесие $y_4^* = (14.7;\ 10.5;\ 15.8)$ в модели M3 совпадает (с учетом соответствий между фазовыми переменными) с равновесием $x_3^* = (25.3;\ 15.8)$ модели M2. В модели M4 два асимптотически устойчивых вырожденных равновесных решения $y_3^* = (9.2;\ 11.2;\ 0)$ и $y_5^* = (14.7;\ 0;\ 623.3)$. Модель M4 имеет также неустойчивое положительное равновесие $y_4^* = (14.7;\ 10.5;\ 15.8)$. Это приводит к медленному разделению начального односвязного множества положительных значений функции принадлежности на два непересекающихся множества (рис. 1). Каждая из частей движется к одному из устойчивых равновесий, то есть система вырождается. При этом так, как неустойчивость положительного равновесия являет-

ся слабой, то поведение решений вблизи этого равновесия оказывается похожим во всех моделях, а именно решения с начальными условиями в малой окрестности этого неустойчивого положительного равновесия остаются вблизи него и через сто лет.

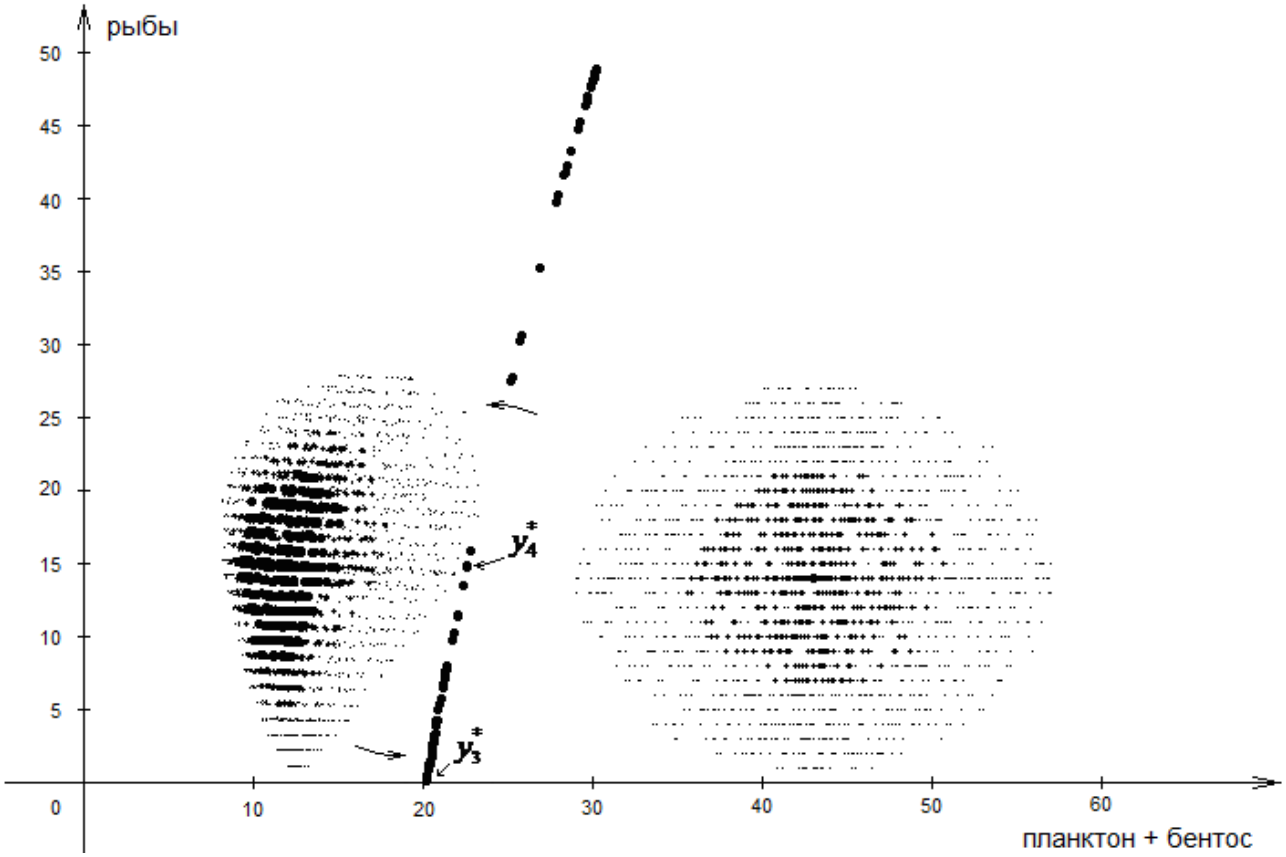


Рис. 1. Точки ненулевых значений функции принадлежности $\varphi(t, x)$ для четвертой модели М4. Координаты центра начального множества (24.9, 18.1, 14). Стрелками показаны направления смещения образа эллипсоида во времени. Время указано в тыс. суток.

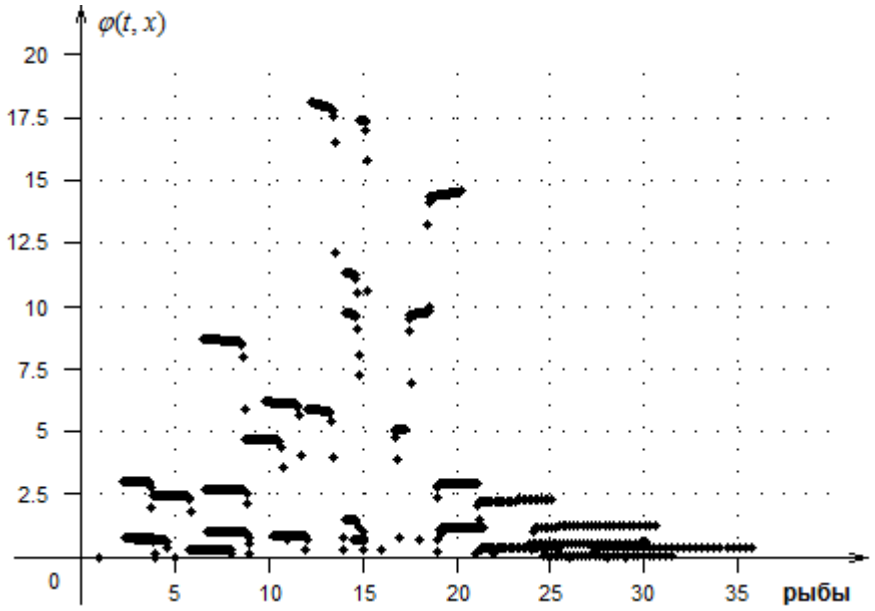


Рис. 2. Ненулевые значения функции принадлежности $\varphi(t, x)$ для четвертой модели М4. Координаты центра начального множества (24.9, 18.1, 14).

Учет неопределенности состояния системы позволяет проследить начальный момент и последующий процесс разделения множества (рис. 2).

Начальные множества положительных значений функции принадлежности «размываются» и порождают разные возможности эволюции системы – от вырождения до более или менее устойчивого сосуществования всех частей экосистемы. Заметим, время этого вырождения велико (более сотни лет).

Все четыре модели частично согласованы между собой в динамике – существует положительное равновесное решение, общее для второй, третьей и четвертой моделей и принадлежащее множеству равновесных решений первой модели. Это равновесие устойчиво в первых трех моделях и неустойчиво в последней модели. В первой модели множество возможных состояний системы распределяется вдоль многообразия положительных равновесий, во второй и третьей модели решение довольно медленно приближается к устойчивым равновесиям, а в четвертой модели решение длительное время (более сотни лет) находится вблизи положительного равновесия, хотя оно неустойчиво.

Значительный временной интервал, необходимый для достижения равновесного состояния, во всех четырех моделях показывает, что опытные данные среднесезонных значений биомасс компонентов экосистемы, рассмотренные в качестве равновесных значений, не могут достоверно характеризовать состояние изменяющейся экосистемы в фиксированный момент времени. В этих условиях некоторым специальным образом заданная функция принадлежности может указать на степень близости определенного состояния биологического сообщества к действительному в заданный момент времени.

Исследование решений в модели М1 показало значительное влияние бактериально-детритного пути переноса энергии для экосистемы озера Большой Окуненок, обусловленное тем, что разнокачественный детрит занимает важное место в рационе большей части обитателей озера, в том числе рыб.

Все расчеты показывают, что нынешнее состояние озера Большой Окуненок может незначительно меняться, но заметные изменения могут произойти только за достаточно большое время (сто лет и более).

ЛИТЕРАТУРА

1. Zadeh L.A. Biological applications of the theory of fuzzy sets and systems // Biocybernetics of the central nervous system. Ed. L.C. Proctor, L. C. Littke. Boston: Brown & Co. 1969. P. 199 – 206.
2. Абакумов А.И. Неопределенность данных в математической экологии. // Дальневосточный математический журнал. 2000. Т. 1, № 1. С. 38 – 42.
3. Казанцева Т.И. Балансовая модель экосистемы мелкого высокоэвтрофного озера. // Журнал общей биологии. 2003. Т. 64, № 2. С. 128 – 145.
4. Абакумов А.И., Пахт Е.В. Моделирование при неопределенности данных (на примере водной экосистемы). // Информатика и системы управления. 2009. № 2(20). С. 3 – 10.
5. Абакумов А.И., Гиричева Е.Е. Многомодельный подход к исследованию водных экосистем. // Известия Самарского научного центра РАН. 2009. Т.11, № 1(7). С. 1399 – 1402.
6. Chorin A.J., Marsden J.E. A mathematical introduction to fluid mechanics. Third Edition. N.Y.: Springer-Verlag New York, Inc., 1992. P. 182.
7. Свирежев Ю.М., Логофет Д.О. Устойчивость биологических сообществ. М.: Наука, 1978. С. 352.
8. Monod J. The growth of bacterial cultures. Ann. Rev. Microbiol. 1949. V. 111. N. 2. P. 371 – 394.

В работе изложен подход для формирования объектов-событий, их сравнения и выполнения операций над ними с помощью объектов той же природы в рамках внутреннего «физического» пространства информационной системы. Рассматриваются типы событий и операции над ними. В результате исследования свойств основных символических операций над событиями удалось выявить шесть сложных операций – процедур, которым присвоены следующие названия *индукция* и *дедукция*, *соединение* и *разъединение*, *анализ* и *синтез*. Эти процедуры позволяют с помощью основных операций получать из наблюдаемых *фрагментов целостное событие*, а также выделять его составные части. На наш взгляд эти сложные операции позволяют в некоторой степени формализовать ментальные процедуры, которые, вероятно, заложены в основе классического теоретико-множественного подхода, основанного на редукционизме, и системного подхода, рассматривающего объекты и явления реального мира, как *организованные целые*. Определены также процедуры, которые объединяют возможности этих методологических подходов. Исследуемый подход направлен на расширение возможностей искусственного интеллекта, кодирования, защиты информации, а также дидактики и практики обучения.

ИММУННОСЕТЕВОЕ МОДЕЛИРОВАНИЕ ЗАВИСИМОСТИ «СТРУКТУРА – СВОЙСТВО» ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ

Г.А.Самигулина, С.В.Чебейко

Разработан иммунносетевой подход к моделированию зависимостей «структура-свойство» лекарственных препаратов. Предложенная интеллектуальная технология позволяет уменьшить погрешности энергетических оценок и повысить достоверность прогноза зависимости «структура-свойство» химических соединений.

Актуальнейшим вопросом биоинформатики в настоящее время является целенаправленный поиск новых веществ и материалов с заранее заданными свойствами, в том числе разработка новых лекарственных средств на основе компьютерного молекулярного дизайна. Применение последних достижений вычислительной техники, искусственного интеллекта, инновационных суперкомпьютерных технологий открывает широкие возможности для успешного решения данной проблемы. Вполне реальным становится создание систем по типу САПР (напомним, что в технике это означает «системы автоматического проектирования») для проектирования и конструирования молекулярных устройств с заданными функциями [1].

Разработка новейших интеллектуальных технологий позволит существенно сократить временные и финансовые расходы при производстве новых лекарств. Даже простая предварительная информация о возможных способах посадки предполагаемой молекулы – лекарства на соответствующий рецепторный участок позволяет резко сократить расходы на создание новых фармакологических препаратов.

Время на создание нового лекарственного препарата составляет 12-15 лет, а общие затраты денег более 800 млн. долларов [2]. Из 100 тысяч потенциальных лекарственных соединений только 1000 могут стать основой для нового препарата, из которых только 3 смогут выйти на фармакологический рынок. С помощью компьютерного молекулярного дизайна, интеллектуальных технологий возможно создание принципиально новых компьютерных алгоритмов, программ поиска и отбора активных веществ целевого назначения на основе концепции взаимосвязи молекулярной структуры и биологической активности химических соединений.

Процесс создания нового лекарства включает в себя поиск мишени действия нового препарата; поиск биологически активного вещества, обладающего нужным фармакологическим действием; изучение этого соединения и т.д.

Количественное описание молекулярной структуры химических соединений в компьютерном молекулярном дизайне осуществляется с помощью дескрипторов [3]. Дескриптор

- это математический параметр, который характеризует структуру органического соединения, отмечая наиболее важные черты этой структуры. Существует проблема создания дескрипторов наиболее полно характеризующих рассматриваемое соединение и позволяющих в удобной форме использовать их в вычислительном процессе.

В современном компьютерном дизайне выделяется три основных этапа исследований:

-формирование обучающей выборки соединений с заданным свойством (активностью);

- описание молекулярной структуры исследуемых соединений (дескрипторов);

- установление взаимосвязи «структура – свойство (биологическая активность)» с последующим созданием устойчивых прогностических математических моделей.

Среди методов прогнозирования зависимости «структура – свойство» следует отметить рост исследований по искусственным нейронным сетям [4]. Наиболее популярна многослойная нейронная сеть прямого распространения, обучающаяся по методу обратного распространения ошибки.

Моделирование биологической активности органических соединений также возможно с помощью нового биологического направления искусственного интеллекта – искусственных иммунных систем [5] (ИИС). Под ИИС понимаются информационные методологии, использующие понятия теоретической иммуннологии для решения различных прикладных задач.

Интерес к таким задачам объясняется их сложной структурой, неполнотой и зашумленностью данных, невозможностью успешного использования классических математических подходов.

Постановка задачи формулируется следующим образом: необходимо разработать эффективную интеллектуальную информационную технологию для компьютерного молекулярного дизайна (моделирования и предсказания свойств новых лекарственных препаратов с заданными свойствами) на основе биологического подхода искусственных иммунных систем [6, 7]. Специфика состоит в том, что при построении математических моделей необходимо одновременно использовать как распознающие, так и оптимизационные способности иммунной сети [8]. Задачу прогнозирования можно рассматривать как задачу оптимизации: необходимо построить оптимальную модель (структуру), которая бы наиболее точно описывала требуемую биологическую активность (свойство) органического соединения.

Реальные данные почти всегда содержат нежелательные составляющие, которые называются шумом. Существуют различные виды неопределенностей данных, например случайные погрешности (неточности при сборе и формировании дескрипторов); систематические (причинные) погрешности; ошибки при моделировании из-за использования упрощенных моделей и т.д. Шум и избыточность данных проявляются в корреляции между пе-

ременными. За исключением простых случаев, искажения в данных не могут быть устранены полностью. Основной целью при построении систем на основе иммунносетевого моделирования является уменьшение ошибки обобщения, поскольку малая ошибка обучения гарантирует адекватность модели лишь в заранее выбранных точках.

Возможна ситуация когда почти идентичные по структуре соединения имеют диаметрально противоположные свойства и следовательно биологическую активность, что совершенно недопустимо при конструировании лекарственных соединений. Актуально применение свойств реальных белков для решения этой проблемы.

Таким образом, решение задачи минимизации ошибки обобщения позволяет повысить прогностическую способность модели и является наиболее трудной при построении данных систем.

Литература

1. Шайтан К.В. Молекулярная динамика пептидов. http://www.bioeng.ru/doc/stat/moldyn.htm.

2. Кубиньи Г. В поисках новых соединений-лидеров для создания лекарств //Российский химический журнал. 2006, № 2. С. 5-17.

3. Раевский О.А. Дескрипторы водородной связи в компьютерном молекулярном дизайне //Российский химический журнал. 2006, № 2. С. 97-108.

4. Гальберштам Н.М., Баскин И.И., Палюлин В.А., Зефиров Н.С. Нейронные сети как метод поиска зависимостей структура – свойство органических соединений //Успехи химии. 2003, № 72(7). С. 706-727.

5. Тараканов А.О. Математические модели биомолерулярной обработки информации: формальный пептид вместо формального нейрона // Проблемы информатизации. 1998. С.65-70.

6. Самигулина Г.А., Чебейко С.В. Прогнозирование зависимости структура-свойство органических соединений на основе иммунносетевого моделирования // Химический журнал Казахстана. –Алматы. 2010, № 3. С. 164-172.

7. Самигулина Г.А., Чебейко С.В. Технология иммунносетевого моделирования для компьютерного молекулярного дизайна лекарственных препаратов // Вестник Харьковского Университета. Тематический выпуск: информатика и моделирование. –Харьков. 2011, № 17. С.142-148. http://www.nbuv.gov.ua/Portal/natural/vcpi/IiM/2011_17/20pdf

8. Бидюк П.И., Литвиненко В.И., Фефелов А.А., Баклан И.В. Гибридная иммунная сеть для решения задач структурной идентификации // Искусственный интеллект. -2004, - № 3. –С.89-99.

ФИЗИЧЕСКИЕ ОСНОВЫ, МОДЕЛИ ПРЕДСТАВЛЕНИЯ И РАСПОЗНАВАНИЯ ОБРАЗОВ В МИКРОТРУБОЧКЕ ЦИТОСКЕЛЕТА НЕЙРОНА, ИХ ПРИЛОЖЕНИЯ ДЛЯ КОДЕКА ВИДЕОДАННЫХ

Е.Е. Слядников
Томский филиал ИВТ СО РАН
opi@hq.tsc.ru

Широкое применение в современных информационных технологиях получил цифровой кодер-декодер (кодек) видеоданных - устройство, специально созданное для обработки видеоданных в режиме реального времени [1]. Основной проблемой в обработке видеоданных является избыточность данных по отношению к информации, в них содержащейся. Поэтому при сжатии видеоданных необходимо применить значительно более мощные и сложные концепции, методы и средства искусственного интеллекта, способные выделять и обрабатывать информацию. Наиболее сложная проблема интеллектуальной предобработки видеоданных это проблема выбора системы и способов сегментации, распознавания образов, которые, подобно сознанию человека должны обладать ассоциативной памятью. Наиболее вероятным кандидатом динамической системы с ассоциативной памятью для интеллектуального кодека видеоданных, а также биологическим прототипом оптического нейронного процессора является дипольная система микротрубочки цитоскелета нейрона.

В последние годы наблюдается значительный интерес к гипотезе «квантово-статистической» природы человеческого сознания [2]. Есть экспериментальные свидетельства, что в нейронах мозга подходящими субстратами для «квантово-статистических вычислений», являются микротрубочки цитоскелета [3]. Недавно была предложена микроскопическая модель дипольной системы микротрубочки [4], в которой в качестве причин возникновения сегнетоэлектрического упорядочения, наряду с диполь-дипольным взаимодействием, принята во внимание роль двухямного потенциала в молекуле тубулина. Затем в [5] была исследована взаимосвязь физических и информационных характеристик в окрестности точки сегнетоэлектрического перехода в дипольной системе микротрубочки, сформулированы принципы функционирования нейрона.

Целью настоящей работы является: развитие модели микротрубочки цитоскелета нейрона [4,5] и вывод уравнения эволюции для случайного параметра порядка, которое описывает процесс представления и распознавания образов; формулировка модели подсистемы обработки данных в кодеке, обладающей искусственным интеллектом.

В ходе развития и исследования предложенной модели микротрубочки получены следующие результаты и выводы [6]:

1. Модель дипольной системы микротрубочки цитоскелета представляет собой систему дипольных моментов, в которой основную роль играют антиферромагнитное взаимодействие ближайших соседей и ферромагнитное взаимодействие следующих за ближайшими соседями. Следовательно, между парами дипольных моментов происходят случайные конкурирующие взаимодействия разных знаков – ферро- или антиферромагнитные. При таком взаимодействии в дипольной системе микротрубочки цитоскелета протекает фазовый переход, связанный с локальным замораживанием дипольных моментов. Этот фазовый переход можно описать с помощью параметров порядка: случайной функция поляризации и ее двух первых моментов: средней поляризации и среднего квадрата поляризации микротрубочки цитоскелета, которые связаны между собой уравнениями молекулярного поля.
2. Проблему распознавания образов микротрубочка цитоскелета решает в два этапа с помощью механизма фазового перехода в дипольной системе. При поступлении образа микротрубочка ставит ему в соответствие вектор, компоненты которого дипольная система микротрубочки цитоскелета определяет, производя соответствующую перестройку своей структуры. Естественным образом происходит разложение образа на характерные признаки,

в результате чего каждому признаку ставится в соответствие термодинамическое среднее значение дипольного момента микротрубочки цитоскелета на узле. На этом этапе обработки данных пространство признаков описывается узельной системой координат, а всему образу соответствует вектор образа, который в узельной системе координат образует упорядоченный набор компонент.

3. Затем, используя механизм фазового перехода, дипольная система микротрубочки релаксирует к ближайшему собственному состоянию матрицы связей, естественным путем разлагая образ на другие характерные признаки (по базису собственному векторов матрицы связей). Каждому признаку ставится в соответствие случайный параметр порядка – проекция вводимого вектора на собственный вектор матрицы связей. То есть микротрубочка устанавливает уже в другом пространстве признаков – параметров порядка, какой хранящейся в памяти микротрубочки структуре отвечает определенный микротрубочкой вектор и, следовательно, входящий образ. Таким образом, с помощью механизма фазового перехода осуществляется нелинейное преобразование (проектирование) одного пространства характерных признаков (узельного) с большой размерностью в другое пространство характерных признаков (параметров порядка) с существенно меньшей размерностью. Невысокая размерность пространства параметров порядка означает, что обработке подлежит не слишком большое количество информации.

Из предлагаемой модели вытекают важные следствия:

– Если начальный образ неполон, он может быть дополнен до полного вектора прототипа, то есть дипольная система микротрубочки цитоскелета действует как ассоциативная память.

– Если в памяти дипольной системы хранятся прототипы образов, а системе предъявляется образ, описываемый начальным вектором, то система превращает начальный образ в ближайший к нему прототип. Следовательно, дипольная система микротрубочки цитоскелета осуществляет распознавание образов.

Из описанных выше вычислительных характеристик синергетической сети микротрубочки цитоскелета нейрона можно сделать выводы о преимуществах данной сети по сравнению с другими типами искусственного интеллекта, использующимися в настоящее время:

1) локальное взаимодействие между данными (только между ближайшими соседями), так как взаимодействие между дипольными моментами в микротрубочке цитоскелета нейрона происходит локально, между ближайшими соседями.
2) параллельная обработка данных. Поскольку данные хранятся и обрабатываются при помощи локальных связей, а операции обработки данных выполняются при помощи механизма фазового перехода, который имеет коллективный и параллельный характер, то операции обработки распределены между отдельными элементами сети, поэтому они выполняются коллективно и параллельно.
3) коллективный механизм фазового перехода приводит к нелинейному взаимодействию между различными элементами сети, в результате чего обработка данных протекает оригинально и нелинейно, вследствие этого не возникает ложных образов при распознавании изображения.
4) информация в данных заключается в упорядочении локальных связей синергетической сети, поэтому микротрубочка цитоскелета нейрона обладает ассоциативной памятью.
5) колоссальное сокращение «данных» при обработке синергетической сетью, путем перехода от вектора всего изображения, имеющего размерность N, к вектору параметров порядка с размерностью M (равной числу прототипов системы), которая много меньше N. Синергетическая сеть микротрубочки цитоскелета обрабатывает уже не данные, а метаданные – информацию в них содержащуюся.

Из этих свойств следует, что время работы синергетической сети микротрубочки будет меньше, чем у других систем искусственного интеллекта.

Используя полученные вычислительные архитектуру, модель, алгоритм обработки данных, ассоциативную память микротрубочки цитоскелета, предлагается усовершенствовать существующую модель кодека видео данных [1] путем добавления к существующей системе подсистемы предобработки информации, обладающей искусственным интеллектом, и состоящей из пяти новых компонент:

- блока предобработки изображения;
- блока сегментации изображения;
- блока распознавания образов изображения;
- блока произвольных аффинных и проективных преобразований изображения;
- управляющего обработкой данных блока - синергетической сети микротрубочки цитоскелета нейрона.

В результате проведенных исследований можно сформулировать научное положение:

– Увеличение контрастности на 20-50%, сегментация и распознавание образов в кодере-декодере медиа данных, обладающем искусственным интеллектом, осуществимы с помощью модели и алгоритма, основой которых являются вычислительная архитектура, модель, механизм сегнетоэлектрического фазового перехода и ассоциативная память в микротрубочке цитоскелета нейрона, для изображений, разрешением от 16x25 пикселей и выше и для углов наклона распознаваемых символов не более 15 градусов.

1. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения / Москва: Техносфера, 2005. – 368 с.
2. Пенроуз Р. Тени разума: в поисках науки о сознании. – Ижевск: ИКИ, 2005. – 690 с.
3. Hameroff S., Penrose R. Orchestrated Reduction of Quantum Coherence in Brain Microtubules // Journal of Consciousness Studies. – 1994. – J.1. – P.91–118.
4. Слядников Е.Е. Физическая модель и ассоциативная память информационной биомакромолекулы // Журнал технической физики. – 2007. – Т.32. – В. 8. – С. 52–59.
5. Слядников Е.Е. О взаимосвязи физических и информационных характеристик в окрестности точки сегнетоэлектрического перехода в дипольной системе микротрубочки цитоскелета // Журнал технической физики. – 2009. – Т.79. – В.7. – С.1–12.
6. Слядников Е.Е. Физические основы, модели представления и распознавания образов в микротрубочке цитоскелета нейрона // Журнал технической физики. – 2011. – Т.81. – В.12. – С.1–33.

Сизиков Виктор Петрович, *к.т.н., доцент*
Омский государственный университет путей сообщения

Языки программирования нового уровня

Одна из главных проблем у языков программирования – универсализация, которая затрагивает представление знаний и функционирование. Попытки решить эту проблему, исходя из традиционных представлений о формальных грамматиках и их языках, к успеху не привели. Следует признать, что основной преградой на этом пути является изначально прописанный линейный характер формирования текстов в грамматиках, препятствующий синтезу различных языков. Необходимы грамматики, работающие с гипертекстами, включающими в себя элементы, характерные для образного мышления. Здесь неизбежно требование, что смысл несут лишь серии связей, объединённых в циклы. Именно потенциал смысловых серий должен быть положен в основу грамматик и языков нового уровня. А общность и проработка до автоматизированного воплощения таких грамматик позволяют заговорить о языках программирования нового уровня. И этот момент уже обрёл определённую практическую реализацию в рамках ДИС-технологии, развитой на базе теории динамических информационных систем (ДИС, ТДИС).

Описаны ДИС как рабочий объект ДИС-технологии, а также основные этапы и особенности работы ДИС-технологии как аппарата имитационного моделирования. Даны представления о ДИС-технологии как о языке программирования нового уровня, а также как об оболочке экспертных систем. Приведены серия закономерностей, характерных для функционирования ДИС, и их интерпретации биологической направленности.

Имеются программные базы для смысловой обработки понятий, построения качественных моделей объектов (процессов), для проведения в рамках Excel численных экспериментов и выдачи графиков функционирования ДИС. Эти базы обеспечивают междисциплинарный синтез, общность методов и интерпретаций при различных, на первый взгляд, явлениях. Они включены в рабочие программы нескольких дисциплин, предназначенных студентам всех курсов и магистрантам.

Sizikov V.P., Omsk state transport university

Programming languages of new level

One of the main problems for programming languages is a universalization which touches upon both the presentation of knowledge and functioning. If we proceed from the traditional representations about formal grammars and their languages, some attempts to solve this problem haven't resulted in success. It must be admitted that the basic obstacle in this way is initially prescribed a linear character of texts forming in grammars which interferes synthesis of various languages. We need the grammars working with the hypertexts which include some elements that are characteristic for pictorial thinking. Here it is inevitable the requirement that only the series of the communications united in cycles carry the sense. The potential of semantic series should be taken as a principle of the grammars and languages for the new level. And the generality and study up to the automated embodiment for such grammars permit to start talking about programming languages of the new level. And this moment has already found certain practical realization within the scope of the DIS-technology which has been developed on the basis of the dynamic information systems theory (DIS, TDIS).

DIS and also the basic stages and features of work of DIS-technology as a device for imitation modeling are described as a working object of DIS-technology. Representations about DIS-

technology as a programming language of the new level and also as an expert systems shell are given. We have demonstrated some series of patterns characteristic for DIS functioning and their interpretations through a biological orientation.

There are the program bases for semantic processing of concepts constructions of qualitative models of objects (processes) to carry out numerical experiments within the scope of Excel and to produce DIS-functioning graph. These bases provide interdisciplinary synthesis, a generality of methods and interpretations at various (at first sight) phenomena. They are included in working programs of the several disciplines intended for students of all courses and undergraduates.

А.А. ЛЯПУНОВ О ПРОЦЕССАХ УПРАВЛЕНИЯ В ЖИВОЙ ПРИРОДЕ И РАЗВИТИЕ КИБЕРНЕТИЧЕСКОГО ПОДХОДА В БИОЛОГИИ

А.Б. Савинов
Нижегородский государственный университет
sabcor@mail.ru

Академик А.А. Ляпунов разрабатывал вопросы строения и эволюции управляющих систем, с позиций кибернетики рассматривал процессы управления в живой природе, привлекая к таким исследованиям биологов (Ляпунов, 1963, 1973; Ляпунов, Стебаев, 1964). Под его редакцией регулярно публиковались сборники «Проблемы кибернетики», в которых имелись разделы, посвященные кибернетическим вопросам биологии. Под влиянием работ А.А. Ляпунова специалист в области молекулярной биологии профессор В.А. Ратнер (1966, 1975, 1983, 1993, 2002) развивал направление молекулярной кибернетики. В научной группе А.А. Ляпунова работал физиолог В.И. Федоров, использовавший кибернетический подход в моделировании физиологических процессов в организме человека. В настоящее время профессор В.И. Федоров активно пропагандирует методологию и методы кибернетического подхода в физиологии, развивает направление кибернетической физиологии (Федоров, 1999, 2000а, б, 2003; Федоров, Шутова, 2004).

Автор (Савинов, 2000, 2006, 2011) использует кибернетические положения для разработки биосистемологии – направления биологии, в рамках которого на основе кибернетического подходов должны изучаться вопросы организации, функционирования и эволюции биологических систем разных уровней (субклеточного, клеточного, организменного, популяционного, биоценотического) в группах различного ранга всех царств живой природы, а также принципы строения, функционирования и эволюции экосистем, в том числе и в условиях техногенеза.

В последнее время внимание автора обращено на симбиотические системы. В симбиотических системах происходят сложные процессы самоуправления, основанные на прямых и обратных связях между членами симбиоза (см. Кеннеди, 1978; Проворов, 2001). В связи с этим рационально использовать кибернетические положения об организации, функционировании и эволюции самоуправляемых систем, к которым относятся биосистемы организменного и надорганизменного уровней (Савинов, 2006).

На основании вышесказанного автором предложен и разрабатывается симбиотический подход к вопросам организации, функционирования и эволюции систем организменного и популяционного уровней с учетом кибернетических принципов самоуправления (Савинов, 2005–2011).

Литература

Савинов А.Б. Методология системно-кибернетического подхода в экологическом мониторинге // Экологический мониторинг. Ч.4. Н. Новгород: Изд-во ННГУ, 2000. С.342-366. (http://www.sabcor.dmls.ru/article/page/51/)

Савинов А.Б. Новая популяционная парадигма: популяция как симбиотическая самоуправляемая система // Вестн. Нижегород. ун-та. Сер. биол. 2005. Вып. 1 (9). С. 181–196.

Савинов А.Б. Биосистемология (системные основы теории эволюции и экологии). Н.Новгород: Изд-во ННГУ, 2006. 205 с.

Савинов А.Б. Аутоценоз и демоценоз – экологические категории организменного и популяционного уровней в свете симбиогенеза и системного подхода // Экология. 2011. № 3. С. 163-169.

РЕКОНСТРУКЦИЯ КОМПЛЕКСНОЙ МОДЕЛИ РЕГУЛЯЦИИ КРОВООБРАЩЕНИЯ У ЧЕЛОВЕКА (ПО GUYTON ET AL., 1972) С ПРИМЕНЕНИЕМ КОМПОЗИЦИОННОГО ПОДХОДА В СИСТЕМЕ BIOUML

Киселев И.Н.[1,2,*]

[1]Конструкторско-Технологический Институт Вычислительной Техники СО РАН, Новосибирск, Россия;

[2]ООО «Институт Системной Биологии», Новосибирск, Россия;

* axec@systemsbiology.ru

Создание, а тем более дальнейшее развитие математических моделей сложных биологических процессов, таких как модель регуляции кровообращения у человека [1], обычно бывает нетривиальной задачей из-за больших объемов и сложной структуры. Однако очень часто оказывается возможным разбить модель на функциональные блоки, тем самым, применив композиционный подход, которые могут рассматриваться по отдельности, что существенно облегчает понимание модели и работу с ней. С целью решения данной задачи, был разработан плагин для системы BioUML включающий:

1. Пользовательский интерфейс для графического представления, создания и редактирования комплексных моделей, состоящих из отдельных, но связанных между собой блоков.

2. Программу для автоматической генерации из комплексной блочной модели плоской, пригодной для симуляции обычными методами численного решения алгебро-дифференциальных уравнений.

Созданный плагин позволяет легко реализовывать сложные модели в виде набора связанных блоков с определенными входами и выходами, варьировать их содержание и проводить различные эксперименты. С применением описанного выше подхода реконструирована модель Гайтона, состоящая из 18 блоков. Модель валидирована путем воспроизведения экспериментов, описанных в [1,2], и доступна в системе BioUML по адресу www.biouml.org.

Работа проведена при поддержке междисциплинарного гранта №91 СО РАН.

Литература:

1. Guyton A.C., Coleman T.G., Grander H.J. Circulation: Overall Regulation. Ann Rev Physiol. 1972. 41:13-41

2. Kofránek J, Rusz J. Restoration of Guyton´s diagram for Regulation of the Circulation as a Basis for Quantitative Physiological Model Development. Physiol Res. 2010. 59(6):897-908.

Разумов Владимир Ильич, *д.ф.н., профессор*
Омский государственный университет им. Ф.М. Достоевского
Рыженко Леонид Игоревич, *к.т.н., доцент*
Омская государственная автомобильно-дорожная академия
Сизиков Виктор Петрович, *к.т.н., доцент*
Омский государственный университет путей сообщения

Смысловая экспертная система на базе ТДИС

Эффективность экспертных систем (ЭС) существенно зависит от организации базы знаний и работы с ней. Решение этой задачи затрудняется рядом обстоятельств. В организации базы знаний и устройстве решателя используются разные логики. Остаётся нерешённой фундаментальная проблема о согласовании содержательно-смыслового компонента, характеризующего предметную область ЭС, и формально-математического содержания алгоритмической базы и программного обеспечения ЭС. Развитие ЭС уместно рассматривать в русле совершенствования процедур автоматизации, где обостряется противоречие между ускоренно совершенствующими вычислительными технологиями и практически стагнацией в области автоматизации рассуждения. В целях устранения отмеченных недостатков предлагаются системно организованные и онтологически обоснованные процедуры осмысления знаний в ЭС. Это достигается переходом к гипертекстовым грамматикам и языкам, подключающим потенциал образного мышления. В таком случае устраняются принципиальные различия между организацией данных и знаний в ЭС.

Предлагается подход к построению ЭС, в которой, во-первых, используется база знаний, имеющая специфическую форму (так называемая смысловая база данных – СмИБД), во-вторых, решатель фактически совпадает с алгоритмом конструирования СмИБД. Основой для формирования такой ЭС выступает теория динамических информационных систем (ДИС, ТДИС). Осмысление знаний реализуется с использованием генетически обусловленных структур, организующих знания в смысловые базы данных. Система знаний представляется сетью ДИС-компьютеров, проинтерпретированных адекватной гиптертекстовой грамматикой. С учётом этого разработан программный продукт «Когнитивный ассистент» как экспертный редактор смысловой ЭС. Аннотированный здесь подход к формированию смысловых ЭС универсален. Он позволяет систематизировать знания различных предметных областей, поддерживая междисциплинарные исследования и проекты инновационного характера.

Razumov, Vladimir I., PhD, Professor Omsk State University. FM Dostoevsky

Ryzhenko Leonid I., Ph.D., Associate Professor Omsk State Automobile and Highway Academy

Sizikov Victor P., Ph.D., Associate Professor Omsk State Transport University

Semantic expert system based on TDIS

The effectiveness of expert systems (ES) essentially depends on the organization of the knowledge base and work with it. A number of circumstances impede the solution of this problem. For organization of the knowledge base and device solver we use various logics. Fundamental problem of harmonization of content-semantic component remains unsolved. This component defines subject domain of ES, and mathematical-formal content of algorithmic base,

and software of ES. Development of ES we consider in the movement to improving of procedures of automation where a contradiction between computing technologies and automation of reasoning is intensified. For elimination identified drawbacks we suggest system organized and ontologically founded procedures for understanding of knowledge in ES. It may be reached by transition to hypertext grammar and languages that involved potential of image thinking. In this case fundamental differences between organization of data and knowledge in ES are going to eliminate.

We suggest an approach for designing ES, where, firstly, we use knowledge base having specific form (so called, semantic data base – SmIBD), secondly, solver practically coincides with algorithm for designing of SmIBD. For forming the ES we use the theory of dynamic information systems (DIS, TDIS). For understanding of knowledge we use genetically determined structures that organize knowledge into semantic database. Knowledge system is implemented as network of DIS-computers with adequate hypertext grammar. On that basis, we developed a software product «Cognitive assistant» as an expert editor of a semantic ES. Annotated approach for forming semantic ES is universal. It allows us to systematize knowledge in various subject domains, including support of interdisciplinary research and innovation projects.

Проблемы описания документооборота в виде онтологий и потоков работ

Ярёменко Ф.В.[1]
[1] Вычислительный центр им. А. А. Дородницына РАН
e-mail: yaremenkofv@gmail.com
Бездушный А. Н.[1]
e-mail: anb@ccas.ru

Аннотация

В нашей работе мы попытались смоделировать деятельность, именуемую документооборот, описывая структурную и целостную составляющие с помощью owl-онтологий (OWL - Web Ontology Language) [1], а функциональную - потоков работ. Мы рассчитываем, что использование этих двух аппаратов позволит формализовать делопроизводство, даст возможность гибко изменять поведение системы электронного документооборота (СЭД) в соответствие с регламентами, бизнес-процессами организации.

Использование потоков работ как функциональной основы для систем электронного документооборота позволит описывать и исполнять разнообразные, сложные схемы движения документов, которые либо вообще не поддерживается многими существующими СЭД, либо поддерживается с существенными ограничениями. Формальная и стандартная спецификация owl-онтологий и бизнес-процессов обеспечивает возможность взаимодействия и, как следствие, поддерживает интеграцию систем и масштабируемость.

Обзор

В работе представлена текущая версия формальной модели документооборота и делопроизводства в виде owl-онтологий и потоков работ.

Для того чтобы спроектировать, реализовывать и внедрять систему электронного документооборота в деятельность предприятия, необходимо выявить, проанализировать и смоделировать его бизнес-процессы и потоки движения документов, затем их формализовать и описать (специфицировать). Перед описанием бизнес-процессов необходимо задать объекты, над которыми будут производиться действия в этих процессах. Объекты, например организационная структура любой сложности или иерархически связанные документы, описаны в модели с помощью owl-онтологий. На основе модели объектов предприятия создаётся конфигурация системы нормативно-справочной информации, которая является важным компонентом СЭД. Описание бизнес-процессов строится с помощью семантических элементов предметной области, которые раскрываются в подпроцессы нотации BPMN 2.0 (Business Process Model and Notation) [2]. Нотация BPMN 2.0 формализует описание бизнес-процессов на уровне детализации, позволяющем исполнять их на программных платформах, поддерживающих этот стандарт моделирования.

Описание процессов документооборота с помощью семантических элементов даёт возможность разрабатывать и изменять бизнес-процессы не только программистам, но и специалистам по документообороту и управляющим.

Разделяют несколько подходов к управлению предприятием: структурный, функциональный, процессный, проектный и т.д. [3] Как правило, системы электронного документооборота, включающие средства работы с бизнес-процессами, ориентированы на предприятия с процессным подходом к управлению. Одно из требований к разрабатываемой модели - поддержка функционального подхода к управлению, который традиционно используется в организациях на территории РФ.

Преимущества использования BPMN 2.0

Поддержка подпроцессов в BPMN позволяет использовать принцип декомпозиции. Декомпозиция даёт возможность постепенно и структурировано представлять модель процесса в виде иерархической структуры отдельных диаграмм, что делает ее менее перегруженной и более понятной для разработчиков.

Полученные модели поддерживают широкие возможности экспорта и импорта процессов в(из) распространённые(ых) форматы(ов) такие как XPDL, BPEL, UML, IDEF3 и EPC. Механизм импорта позволяет загрузить в систему и использовать ранее разработанные модели, созданные, например, в процессе стандартизации предприятия.

Спецификация BPMN 2.0 содержит описание процедуры трансляции диаграмм в язык BPEL (Business Process Execution Language) [2, 4]. Язык BPEL был стандартизован в 2004 году и на сегодняшний день широко используется. BPEL подразумевает реализацию описанных потоков работ в виде веб-сервисов. Использование веб-сервисов даёт возможность внедрять системы в территориально-распределённые организации и использование в межведомственном взаимодействии.

Большинство программных платформ исполнения бизнес-процессов включают функции мониторинга и аудита процессов, что позволяет контролировать работу как организации в целом, так и отдельных должностных лиц. Также, информация, полученная с помощью средств мониторинга, содержит необходимые данные для реинжиниринга бизнес-процессов документооборота, который в последствие позволит сократить временные затраты на обработку документов как службы ДОУ (документационного обеспечения управления) так и остальных сотрудников организации.

Заключение

Используемый в работе подход позволяет быстро изменять описание предметной области, выраженное в виде owl-онтологий. Также, предоставляется возможность легко менять описание бизнес-процессов делопроизводства, проектировать сложные бизнес-процессы организации и интегрировать в делопроизводство смежные/граничные бизнес-процессы. Одним из итогов работы является создание СЭД, в основе которой лежит платформа управления бизнес-процессами. Полученная модель позволяет реализовать шлюзы, с помощью которых консолидируется работа различных СЭД, имеющие различные системы понятий, операции и интерфейсы. Эти шлюзы можно применять, например, в территориально-распределённых организациях и межведомственном взаимодействии.

Литература

[1]. Peter F. Patel-Schneider, Patrick Hayes, and Ian Horrocks, Editors OWL Web Ontology Language Semantics and Abstract Syntax http://www.w3.org/TR/2004/REC-owl-features-20040210/
[2]. Business Process Model and Notation. Версия спецификации 2.0 http://www.omg.org/spec/BPMN/2.0/
[3]. Структурный, функциональный, процессный и проектный подходы к построению организационной структуры http://corpsys.ru/Articles/Structure/Approaches.aspx
[4]. OASIS Web Services Business Process Execution Language (WSBPEL) Technical Committee http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel

Подсистема автоматизированного проектирования дорожной разметки и знаков в составе информационно-аналитического программного комплекса «AUTOROAD»

А.В. Приступа
Томский государственный университет
E-mail: pristupa@sibmail.com

В статье рассматривается модуль автоматизированного проектирования дорожной разметки и дорожных знаков, входящий в состав ИАПК «AUTOROAD». Выделены типичные ситуации, требующие нанесения определенного вида разметки или установки определенных знаков. Предложен алгоритм универсального нанесения разметки в комплексных ситуациях, требующих принятия решения по нескольким критериям.

Ключевые слова: автоматизированное проектирование, организация дорожного движения, дорожная разметка, дорожные знаки

Логико-вероятностная модель прогнозирования характеристик паразитарной системы клещевого энцефалита, заболеваемости населения в зависимости от природных факторов

Полякова Г.Л., Лбов Г.С., В. Н. Бахвалова, О. В. Морозова, В. В. Панов, В. В. Глупов, Щучинова Л. Д., Маслов П.П.

В докладе рассматривается метод прогнозирования, основанный на обнаружении логических закономерностей, отражающих вероятностные причинно-следственные связи изучаемых факторов. Метод был использован при исследовании влияния природных факторов на характеристики паразитарной системы клещевого энцефалита (КЭ) заданного региона (численность мелких млекопитающих, клещей, индексы обилия личинок и нимф на мелких млекопитающих, уровень зараженности клещей вирусом КЭ) на уровень заболеваемости КЭ.

Методы обнаружения вероятностных логических закономерностей обладают рядом положительных свойств, например, дают более надежные результаты в условиях малой выборки, осуществляют отбор более информативных переменных и результаты статистического анализа представляются на языке, близком к естественному языку логических суждений [1]. Эффективность предложенного метода иллюстрируется результатами решения задачи исследования влияния природных факторов на заболеваемость КЭ [2].

В докладе будет приведена логико-вероятностная модель прогнозирования характеристик паразитарной системы клещевого энцефалита, заболеваемости населения в зависимости от природных факторов для трех регионов Сибирского федерального округа.

Список литературы

1. Лбов Г. С., Старцева Н. Г. Логические решающие функции и вопросы статистической устойчивости решений. // — Новосибирск: Изд-во Ин-та математики, 1999. 211 С.

2. Лбов Г. С., Полякова Г. Л., Бахвалова В. Н., Морозова О. В. Исследование влияния природных факторов на заболеваемость клещевым энцефалитом. // — Вестник Новосиб. гос. ун-та. Серия: Биология, клин. Медицина, 2010. № 3.С. 31-37.

Применение мультиэвристического подхода совместно со специальным алгоритмом архивации для сравнения последовательностей ДНК.

А.Г. Панин, Ю.Хайруллова

Тольяттинский Государственный Университет, Тольятти, Россия

ag.panin@gmail.com, ruiner@mail.ru

В работе рассматривается применение мультиэвристического подхода к проблеме определения схожести строк, кодирующих молекулы ДНК. Перед применением мультиэвристического подхода строки обрабатываются специальным алгоритмом архивации, описанным в данной работе. При сравнении получившихся строк учитываются также некоторые характеристики, полученные после архивации. Мультиэвристический подход позволяет получить целый класс метрик на множестве строк, некоторые из которых могут дать интересные результаты при применении их к сравнению ДНК. Использование его совместно с алгоритмом архивации позволяет не только анализировать сжатые данные без их распаковки, но и даёт возможность учесть некоторые характеристики исходных строк, полученные в процессе архивации.

Несколько подходов к оптимизации алгоритма вейвлет-преобразования, реализованного на графическом процессоре.

А.Г. Панин, М.Н. Точёная

Тольяттинский Государственный Университет, Тольятти, Россия

ag.panin@gmail.com, m.tochenaya@gmail.com

В данной работе описаны особенности реализации алгоритмов вейвлет-преобразований для платформы NVIDIA CUDA. Архитектура CUDA обладает сложной структурой памяти, и для того, чтобы максимально использовать вычислительную мощность графического процессора, необходимо уделить пристальное внимание оптимизации работы с памятью. В статье описаны несколько различных способов работы с памятью, приводится сравнение скорости работы. Так же из-за того, что для полноценной загрузки GPU нужны тысячи нитей, важным вопросом является разделение данных на подзадачи. В статье приводится зависимость скорости работы алгоритма от различных параметров при разделении данных. Ещё одной особенностью реализации, рассмотренной в статье, является оптимизация арифметических выражений.

ОЗОНОВЫЙ ЩИТ ЗЕМЛИ: ЛЕГЕНДЫ И РЕАЛИИ. АЭРОДИНАМИЧЕСКАЯ МОДЕЛЬ АНТАРКТИЧЕСКОЙ ОЗОНОВОЙ ДЫРЫ

В.Б. Кашкин, Т.В. Рублева, Р.Г. Хлебопрос

В.Б. Кашкин
Профессор кафедры радиотехники Сибирского Федерального Университета
rtcvbk@rambler.ru
Т.В. Рублева
Ст. преподаватель кафедры теплофизики Сибирского Федерального Университета
tvrubleva@ksc.krasn.ru
Р.Г. Хлебопрос
Профессор кафедры экономики природопользования Сибирского Федерального Университета, Директор МНЦИЭСО при Президиуме КНЦ СО РАН

Формирование ультрафиолетового климата Земли зависит от стратосферного озона. Озоносфера регулирует биологическую активность солнечного излучения, в связи с этим проблема истощения озонового слоя тесно связана с воздействием на биосферу в целом. В настоящее время почти все исследования по деградации озонового слоя и механизмов образования озоновых дыр направлены на доказательство антропогенного происхождения этих явлений.

В настоящей работе исследовалась озоновая аномалия в нижней стратосфере южного полушария – озоновая дыра (устойчивое понижение над Антарктидой общего содержания озона (ОСО) ниже климатической нормы (<220 е.Д.), существующая весной в течение 3-3,5 месяцев). Предложена аэродинамическая модель сложной циклонической системы, объединяющей циркумполярный вихрь и озоновую дыру. Разработан новый метод слежения за движениями воздушных масс в нижней стратосфере, где озон играет роль индикатора. На основе спутниковых озонных данных рассчитаны зональные и меридиональные скорости масс озона в весенний период 1996-2009г.г. Наибольшие значения зональной скорости приходятся на широтную область 45°S-55°S, в которой регистрируются максимальные значения ОСО. Выявлена связь между зональной скоростью в области накопления ОСО и дефицитом массы озона в дыре, коэффициент корреляции 0,78. Таким образом, экспериментальные результаты показывают, что аэродинамика нижней стратосферы является доминирующим фактором при образовании озоновой дыры в южном полушарии в исследуемый период.

Компания HP

Hewlett-Packard
Ленинградское шоссе, 16а, стр.3
125171 Москва, Россия
www.hp.ru

HP – мировой лидер рынка информационных технологий – создает новые возможности для использования технологий на благо людям, компаниям, правительствам и обществу в целом. HP - это крупнейшая в мире технологическая компания, которая предлагает семейство продуктов, охватывающее системы печати, персональные компьютеры, программное обеспечение, услуги и ИТ-инфраструктуру, объединяющую облачные среды, для простой, безопасной, контекстной работы во взаимосвязанном мире. Компания работает в 178 странах мира, а количество ее сотрудников составляет свыше 304 000 специалистов. Доход HP за 2010 финансовый год составил 126 млрд. долл. США. Штаб-квартира HP расположена в г. Пало-Альто (штат Калифорния, США). Компанию возглавляет Мэг Уитмен (Mag Whitman).

HP в России

Компания HP присутствует на российском рынке более 40 лет.

Сегодня:

— В штат сотрудников компании входит более 1000 человек, а партнерская сеть насчитывает почти 300 компаний, которые предлагают решения, услуги и продукты HP.

— Компания HP представлена в 11 городах, помимо Москвы,и покрывает все федеральные округа России

— HP постоянно принимает активное участие в развитие инноваций: в 2007 году был открыт научно-исследовательский центр HP Labs в Санкт-Петербурге, началась разработка программного обеспечения и производство персональных компьютеров совместно с компанией Foxconn

— Помимо ведения бизнеса в России, HP вносит вклад в поддержку образования, научных исследований и оснащение культурных объектов новыми технологиями.

© 2011 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

ИНТЕРВАЛЬНАЯ АРИФМЕТИКА: ЭМПИРИКО-СТАТИСТИЧЕСКИЙ ПОДХОД К ОЦЕНКЕ РЕЗУЛЬТАТОВ ДЕЙСТВИЙ[1]

Петрушин В.Н., Никульчев Е.В.
Всероссийская государственная налоговая академия
Министерства финансов Российской Федерации
E_mail nikulchev@mail.ru

В процессе измерений исследователи крайне редко имеют с точными значениями. В силу погрешности приборов, невозможности количественно определить воздействие всех факторов на изучаемое явление или процесс наблюдаемая величина является случайной, и, как правило, ограниченной на каком-либо сегменте или интервале.

Если ставится условие получить решение какой-либо расчетной задачи на базе экспериментальных данных с вероятностью единица, целесообразно применять интервальный анализ, достоинством которого является безразличие к виду функции распределения случайной величины. Но у этого вида оценок два существенных недостатка: при применении интервальной арифметики происходит значительное расширение интервалов, и такие расчеты не пригодны для практических целей; вторым существенным недостатком является отсутствие знания точных границ, что приводит, к появлению числа невозможных решений.

В теории эксперимента отказываются от достоверности расчетов, исследователь самостоятельно определяет необходимую надежность результата. Такой подход требует знания функции плотности вероятности случайной величины или ее закона распределения. Чаще всего, традиционно, предполагается нормальное распределение и распределения, получаемые на базе нормального при нелинейных преобразованиях. Интервальная оценка рассчитывается по принятому распределению с выбранной вероятностью. Достаточно часто совершается вероятностная ошибка: надежностью результата провозглашается выбранная вероятность попадания в интервал, но при этом забывают, что теоретическое распределения принято тоже с какой-то вероятностью. Гипотеза о виде распределения связана с расчетами параметров распределений по выборочным данным методом моментов или методом наибольшего правдоподобия, и несет за собой все погрешности этих расчетов.

Авторы поставили себе цель совместить преимущества каждого из подходов при обработке экспериментальных данных.

В докладе приведена методика поиска минимально возможного интервала оценки значений случайной величины с заданной надежностью. Рассмотрены вопросы построения гистограммы на основании выборки. Проведена интервальная оценка числовых характеристик без гипотезы о виде распределения. Приведены результаты интервальной арифметики с применением выборочных распределений.

Приведенные примеры оценок величин, их средних, результатов арифметических действий свидетельствуют о том, что эмпирический подход позволяет делать более точные вероятностные интервальные оценки для любого наблюдаемого распределения без его аналитической аппроксимации и связанных с ней погрешностями. Удалось совместить независимость классического интервального анализа от вида распределения случайной величины внутри сегмента значения и возможность выбора наиболее вероятных значений без оценки границ генеральной совокупности.

[1] Работа выполнена при финансовой поддержке РФФИ (грант № 11-07-00772-а).

DENSITY FUNCTIONAL THEORY APPLICATION TO THE STEROID HORMONES SPECTROSCOPIC BIOCLASSIFICATION

B.F. Minaev, V.A. Minaeva, G.V. Baryshnikov
Bohdan Khmelnitskij National University, Cherkassy, Ukraine
e-mail: bfmin@rambler.ru
O.P. Cherkasova
Institute of Laser Physics SB RAS, Novosibirsk, Russia
e-mail: chrom@laser.nsc.ru

Abstract

The steroid hormones carry out the most important regulatory functions and influence many physiologic processes. The steroid hormone molecules are built of a standard four-ring system (three six-member rings and one five-member ring). How these similar-type compounds can provide various complicated biological functions based on recognition of numerous receptors and their modification actions on DNA? Vibrational spectroscopy together with quantum density functional theory (DFT) calculations provide a good bioinformatics background for classification and biospecification of their structure-activity relationship.

The presence of the four rings provides a specific structure with conformational flexibility and a large number of low-frequency vibrations. Minor alterations in position of side functional groups result in essential change of binding processes with steroid receptors. In this work we present our studies of teraherz time-domain (THz) and infrared (IR) spectra of a number of hormones and compare their absorption THz-IR spectra with Raman scattering. The Raman and THz spectra were measured in the temperature range from 83 K to 297 K. Based on quantum chemical calculations of frequencies by the DFT B3LYP/6-31 G (d,p) approximation we can assign the vibrational bands in both, the THz-IR absorption and Raman spectra.

The chosen steroid hormones (progesterone, 17α-hydroxyprogesterone, cortisol, cortisone and acetate cortisone) differ in the sequential addition of OH groups to the progesterone core. Low frequency vibrational spectra of steroid hormones contain vibrational modes which include deformation of the whole molecule (30-35 cm^{-1}, twisting of steroid rings), and modes of side chain (75-77 cm^{-1}). Position and intensity of lines depends on the structure of molecules, the presence and quantity of hydrogen atoms and intermolecular bonds. The knowledge of vibrational modes presented here is important for understanding of hormone-receptor binding.

Метод фрактальной обработки кардиоинтервалограмм в диагностических целях

Технологический институт ЮФУ в г. Таганроге

Резниченко А.А.

Naos_88@mail.ru

Кардиоинтервалограмму (КИГ) в норме можно рассматривать как процесс в виде фрактальной динамической структуры с самоподобием. Критерием устойчивости такой динамической системы является отсутствие конфликтов эндогенных циклов. Естественно предположить, что для здорового организма должно существовать определенное гармоническое равновесие между ритмами разной периодичности при гомеостазе, т.е. наличие определенного инварианта и его вырождение при патологиях. Предварительные экспериментальные данные подтвердили вид спектра КИГ типа $1/f^{\beta}$ (где ß~1), а возможность выявления интервалов частот, в которых укладываются равные величины энергии, показывает достаточное согласие с параметрическими рядами, типа Фибоначчи-последовательности, что подтверждает наличие фрактальной структуры спектра, т.е. определенную самоорганизацию в кардиоритме [1].

Это позволит нам использовать измеренные ряды, полученные путем обработки КИГ, и в дальнейшем перейти к их обработкой через фрактальные показатели. Дальнейшие вычисления проводятся через индекс фрактальности. Индекс фрактальности можно использовать для традиционного анализа фрактальной структуры временных рядов, которые наблюдаются в КИГ. В частности, результаты анализа временных рядов при помощи индекса фрактальности μ (где, $\mu=D_{\mu}-1$) совпадает с традиционными результатами, согласно показателей Херста [2]. Но подобное использование индекса фрактальности позволит при одинаковом количестве данных определится в несколько раз точнее, чем показатель Херста.

При этом минимальный масштаб необходимый для определения μ с приемлимой точностью составит несколько десятков точек, тогда как для определения H необходимо несколько тысяч данных. Этот факт позволяет использовать μ в качестве локального фрактального показателя при анализе временных рядов.

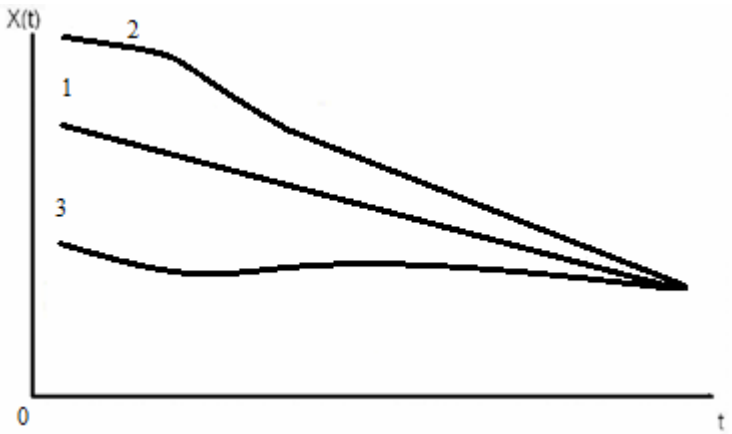


Рис. 1 Теоретическое представление спектра RR интервалов

Далее обрабатывая полученные результаты, можно получить картину распределения параметров состояния организма человека как показано на рисунке 1. Где схематически представлено как должен выглядеть график для здорового человека (1), и если работа органа нарушена то это также отразится на графике (2) и (3). Таким образом, исходя из полученной картины, можно не только наблюдать за оказываемым воздействие на организм больного, но и минимизировать процедуру, если кривая перейдет в нормальное состояние.

ЛИТЕРАТУРА

1. Алдонин Г.М. Самоорганизация в гомеостазе и донозологическая диагностика // Моделирование неравновесных систем – 98: тез. докл. l Всероссийского семинара. – Красноярск, 1998.
2. Старченко Н.В. ИНДЕКС ФРАКТАЛЬНОСТИ И ЛОКАЛЬНЫЙ АНАЛИЗ ХАОТИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ // Математическое моделирование, численные методы и комплексы программ – М., 2005.

Применение мультиэвристического подхода для анализа сигналов акустической эмиссии

Д.Л. Мерсон, А.Г. Панин

Тольяттинский Государственный Университет, Тольятти, Россия

d.merson@tltsu.ru, ag.panin@gmail.com

Метод акустической эмиссии является относительно новым и достаточно перспективным методом исследования, позволяющим получать информацию о составе, дефектности, различных свойствах исследуемых объектов, изучать динамику происходящих процессов в реальном времени. Для эффективного применения этого метода необходимо решить ряд проблем, связанных с обработкой полученных данных. Одними из важнейших являются проблема обработки большого количества акустических данных, а так же проблема поиска связей между регистрируемыми акустическими сигналами и протекающими внутри объекта исследования процессами. В данной работе предлагается применение мультиэвристического подхода для задания метрики на множестве акустических сигналов, на основе которой проводится кластерный анализ. Рассматриваются результаты применения указанного подхода совместно с вейвлет-преобразованием, которое используется для первоначальной обработки данных.

Геоинформационная система для анализа этноэкологического потенциала Республики Бурятии[1]

Т.Г. Дармаев, Г.М. Барлуков

Научно-образовательный и инновационный центр системных исследований и автоматизации Бурятского государственного университета
dtg@bsu.ru

Аннотация

В данной работе описывается геоинформационная система разработанная в целях распространения этноэкологических знаний и природосберегающих технологий районов с сохранившимися формами традиционного природопользования, ориентированная на регламентированный сбор, хранение и обработку геоинформации, на основе современных средств интеллектуального анализа данных, объектно-ориентированных баз данных и географических информационных систем.

Введение

На конференции в Рио-де-Жанейро был одобрен план действий по устойчивому развитию, получивший название «Повестка дня на XXI век». Она была принята в связи со стремительно ухудшающейся экологической ситуацией и прогнозом о возможной глобальной катастрофе в XXI веке, которая может привести к гибели всего живого на планете. Концепция устойчивого развития предполагает позитивный синтез и системное решение экологических, экономических, социальных и культурных проблем современности. Основной принцип устойчивого развития - улучшение качества жизни населения без превышения допустимых пределов нагрузки на природные ресурсы.

Таким образом, возникает потребность в создании комплексной геоинформационной системы ориентированной на регламентированный сбор, хранение и обработку разнородной информации на основе современных средств интеллектуального анализа данных, объектно-ориентированных баз данных и географических информационных систем в целях распространения этноэкологических знаний и природосберегающих технологий районов с сохранившимися формами традиционного природопользования.

Архитектура системы

Архитектура рассматриваемой системы включает в себя геоинформационный сервер (ГИС-сервер), сервер баз атрибутивных данных, подсистему широкоформатной печати, различные проблемно ориентированные надстройки. ГИС-сервер отвечает за базовые операции по манипуляции с абстрактными документами, слоями, визуализаторами векторных данных и т.п. (см. Рис 1). В надстройках описываются конкретные реализации определённых типов документов, слоёв, специальные алгоритмы анализа и обработки графической информации.

[1] Работа выполнена в рамках гранта РГНФ 2011г., проект №11-06-12043в «Этно-экологические традиции и пространственные особенности социально-экономического развития коренных этносов Республики Бурятия»

В геоинформационной системе векторные данные хранятся в памяти в виде динамических массивов для обеспечения прямого доступа ко всем объектам, при этом все объекты индексируются с помощью деревьев, что позволяет достичь при оперативной работе в системе скорости реакции. Атрибутные описания графических объектов хранятся частично вместе с графикой в специальном внутреннем формате, зависящем от типа слоя графики, при этом для всех видов векторных данных предоставляется возможность присоединения по некоторому ключевому полю дополнительных таблиц, имеющих произвольный формат, поддерживаемый через стандартные программные интерфейсы BDE, IDAPI или ODBC. Просмотр и анализ атрибутов объектов при этом может выполняться с помощью штатных средств для работы с базами данных, в том числе с использованием языка запросов SQL.

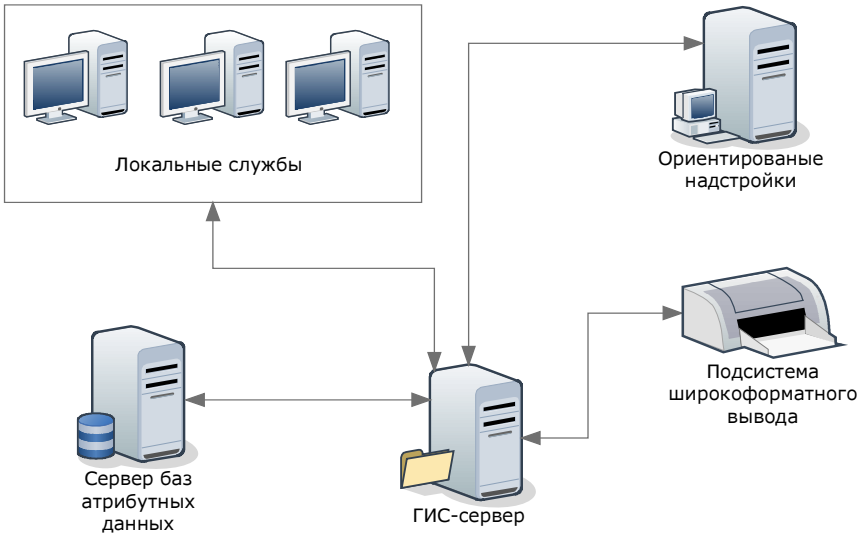


Рис. 1. Архитектурная схема геоинформационной системы

Разработка предполагает использование различных СУБД (MySQL, PostgreSQL, MSSQL), языков программирования (PHP, C++), технологий, решений и библиотек (веб-технологии и библиотеки – java, ASP.NET, AJAX, vlibtemplate, JQuery, ГИС- и инженерные технологии и решения – Mapserver, AUTOCAD и др.).

В системе под термином карта подразумевается любой графический рисунок, созданный с помощью редактора карт. Вся графическая информация на картах состоит из упорядоченной совокупности графических слоёв. В пределах одного слоя обычно объединяются данные, имеющие некоторое семантическое сходство, например, карта района может быть разбита на слои домов, дорог, гидрологии, рельефа и т.д. Большое разнообразие различных поддерживаемых видов данных позволяет системе создавать и эффективно работать с комбинированными графическими изображениями. Информация, находящаяся в слоях шейп-файлов и покрытий, отображается с помощью визуализаторов, т.е. алгоритмов, отрисовывающих точки, линии или полигоны в соответствии со своими настройками. Каждому слою может быть сопоставлено любое количество визуализаторов, что позволяет комбинированно отображать сразу несколько параметров. Например, для слоя район на карте можно отобразить разным цветом климатическое и природно-рекреационное районирование, затем точками плотности показать плотность населения, подписями указать

названия населенных пунктов, а также расставить диаграммы экологического прогноза загрязнения окружающей среды, прогноза демографической ситуации.

В геоинформационнной системе имеется широкий спектр стандартных визуализаторов, позволяющих создавать весьма выразительные тематические карты:

1. Климат - основные климатические характеристики (среднегодовое количество осадков, среднегодовые показатели температуры, господствующее направление ветров, показатели давления и т.д.);
2. Водные ресурсы - гидрография и гидрологическая изученность поверхностных вод;
3. Растительные ресурсы – современное состояние растительного покрова, лесной фонд, распространение редких и нуждающихся в охране видов растений, растительных сообществ, лекарственных, пищевых и кормовых растений;
4. Охрана природы – природоохранное зонирование, особо охраняемые природные территории, предприятия по охране природы;
5. Население – информация о структуре населения (численность населения, национальный состав, занятость, уровень образованности и т.д.);
6. Социально-экономическая инфраструктура (населенные пункты, дороги, места общественного питания, медицинские учреждения, учебно-образовательные учреждения, санаторно-курортные пансионаты и т.д.);
7. Традиционное природопользование – данные по основным пастбищам и сенокосам, запасам дикоросов, охотничьим угодьям. В этом же блоке будет сосредоточена информация о состоянии аборигенных (номадных) животных, разводимых на пилотных территориях (Окинский, Курумканский районы Бурятии);
8. Природосберегающие традиции – сбор данных о святых местах, с которыми связаны природосберегающие традиции, легенды, сказания;
9. Туризм - данные по современному состоянию туристской инфраструктуры (количество гостиниц, кемпингов, гостевых домов). Миграции (численность въезжающих и выезжающих туристов), основные маршруты и объекты посещения.

Подготовка данных системы проводится на основе комбинированного подхода от локальных рабочих мест по вводу данных на базе стандартного программного обеспечения ГИС и СУБД до Internet/Intranet-решений. В качестве топоосновы используется цифровая карта масштаба 1:1000000.

Литература

1. *Дармаев Т.Г., Бильгаева Л.П., Табхаева Б.Б. Некоторые аспекты создания региональной эколого-информационной системы (РЭИС) // «Проблемы региональной экологии». Вып. 8., Новосибирск, изд-во СО РАН, - 2000 с. 219.*

2. *Ошорова Б.В. ГИС – этноэкологический потенциал рекреационных территорий Бурятии (на примере национального парка «Тункинский») Материалы Сибирского межведомственного научно-практического семинара «Теория и практика экологического страхования: итоги и перспективы».- Улан-Удэ, 2010 с.146-148.*

А.А. Ляпунов – первый теоретик программирования в СССР

Р.К. Казакова, А.К.Платонов
Институт прикладной математики им. М.В. Келдыша РАН

Член-корреспондент АН СССР Алексей Андреевич Ляпунов в 1953 году в только что созданном М.В. Келдышем Отделении прикладной математики (ОПМ) был заведующим отделом программирования. Это было время освоения создаваемой в недрах Института точной механики и вычислительной техники АН СССР (ИТМ и ВТ) новой техники «Большой электронной счётной машины» (БЭСМ). Её создатель - будущий академик Сергей Алексеевич Лебедев был первым и сначала единственным её программистом. Ранее ещё в 50-ые годы в Киеве им была создана первая в нашей стране Малая Электронная Счетная Машина (МЭСМ).

Для большинства специалистов создаваемая БЭСМ была чем-то новым. Это было время, когда молодые сотрудники «Стекловки» постигали азы программирования и в первую очередь - двоичную и шестнадцатеричную системы счисления, на которых работала БЭСМ. Стихийно собирались группы сотрудников и учились переводить десятичные числа в двоичные, складывать, делить или умножать их «столбиком». Иногда на этих увлекательных упражнениях присутствовал М.В.Келдыш. Требовалось создание нового математического формализма для описания устройства программ и развития средств программирования.
А.А.Ляпуновым были поставлены задачи придумать формульный способ описания программ и алгебры их преобразований. Вместе с сотрудниками были найдены способы описания и преобразования программных кодов.

Ляпунов стал одним из первых ученых в СССР, который начал развивать и внедрять в сознание соратников новую в то время науку «кибернетику». Хорошо известно, что эта зарубежная наука поначалу была запрещена в нашей стране. Поэтому программирование Алексей Андреевич развивал в ОПМ, а осмысление и развитие кибернетики он организовал у себя дома в виде полноценных домашних семинаров. М.В. Келдыш поддержал А.А. Ляпунова, и своим авторитетом и рядом осторожных действий ему удалось спасти кибернетику. После переезда Алексея Андреевича в Сибирское отделение АН СССР в ОПМ был создан отдел кибернетики под руководством В.С.Яблонского. А.А. Ляпунов оставил яркую память о себе среди сотрудников ИПМ, и связь с ним не была разорвана после его ухода из института.

Вычислительная техника быстро развивалась, и знания в области информатики активно расширялись. В Институте прикладной математики (ИПМ) была создана первая в стране сеть вычислительных машин и первая связь удалённых таких машин по телеграфному (а позднее и по телефонному) каналу. Бывшие ученики А.А. Ляпунова стали создателями первых систем автоматизации программирования, позднее появились формальные языки программирования, стало развиваться системное программирование, как основное средство управления работой ЭВМ. Позднее для развития этого направления в МГУ был создан специальный факультет Вычислительная Математика и Кибернетика (ВМиК), первым деканом которого был академик А.Н.Тихонов.

В ИПМ было развито новое направление программирования - «машинная графика», была создана мощная Дисплейная Интерактивная Система Проектирования Орбит (ДИСПО) для решения задач построения траекторий полета к планетам, кометам и астероидам Солнечной системы. Стала развиваться отдельная линия программирования робототехнических систем и гибких производственных систем, где важную роль играет созданная последователями Алексея Андреевича теория дискретных систем и теория автоматов. Все эти направления программирования и кибернетики навсегда связаны с именем А.А. Ляпунова.

О ДАРВИНИЗМЕ И НОМОГЕНЕЗЕ: КИБЕРНЕТИЧЕСКИЙ СИНТЕЗ

С.Н.Гринченко

Институт проблем информатики РАН (Москва)
sgrinchenko@ipiran.ru

Аннотация

Дискуссия между сторонниками основных теорий биологической эволюции – селектогенеза (дарвинизма) и номогенеза – происходит, в основном, в русле противопоставления их одна другой, хотя высказываются и мнения, что эти концепции не взаимоисключающи, а дополнительны. В поддержку последней позиции предлагается использовать выдвинутый ранее И.И.Шмальгаузеном междисциплинарный кибернетический подход. Согласно его современному виду, живая природа представляет собой *кибернетическую систему*, имманентно включающую механизм иерархической адаптивной поисковой оптимизации её энергетики. Поведение одной части элементов и связей такой модели живого интерпретируется в терминах дарвинистских версий теории эволюции, а другой части – номогенетических.

Насколько можно судить по литературе, уже много десятилетий подряд продолжается дискуссия по поводу справедливости основных теорий биологической эволюции – селектогенеза (дарвинизма и неодарвинизма, «синтетической теории эволюции» и т.п.) с одной стороны, и номогенеза (ламаркизма, теорий направленности эволюции Л.С.Берга и др.) – с другой. В основном апологеты каждой из этих теорий настаивают на адекватности именно своего понимания эволюции живого. На этом фоне как-то теряются мнения некоторых эволюционистов, которые утверждают идею необходимости их синтеза в единую теорию, в частности: «…дискуссия между сторонниками номогенеза и дарвинистического подхода основана на непонимании того, что эти концепции дополнительны, а не взаимоисключающи» [1], и «…естественный отбор бесспорно остаётся движущим фактором эволюции, но не единственным. Дарвиновская и “недарвиновская” эволюции взаимно не противоречивы» [2], стр.606. Получается, что дискутанты не слышат как друг друга, так и предложения о синтезе их альтернативных позиций. Возможно, для выхода из этой, практически тупиковой, ситуации, стоит воспользоваться междисциплинарным подходом, а именно: привлечь соображения, базирующиеся на *кибернетическом рассмотрении живой природы* (т.е. на интерпретации её в терминах теории управления).

Является ли подобный кибернетический подход к биологической эволюции чем-то новым? Нет, не является. Ещё в 1958 г. И.И.Шмальгаузен опубликовал в Зоологическом журнале статью «Регулирующие механизмы эволюции», в которой представил регулирующий механизм эволюции популяции в виде кибернетического цикла и отметил, что сама эволюция оказывается автоматически регулируемым процессом прогрессивной адаптации, а «этот кибернетический цикл является лишь перефразировкой дарвиновского понимания эволюции» [3], стр.177. По мнению А.А.Ляпунова, высказанному в 1962 г., «управление можно объявить характеристическим свойством жизни в широком смысле» [4], стр.208.

Но дальнейшего развития эти идеи тогда не получили, и причинами этого, по-видимому, являлись не столько неготовность тогдашних биологов-эволюционистов к их восприятию, сколько и главным образом – недостаточная разработанность аппарата собственно кибернетики на тот, начальный период её развития. Опора лишь на элементарные кибернетические представления тогда – в 1950-1960-х гг. – крайне ограничивала возможности синтезируемых на их базе моделей эволюции.

Ситуация резко изменилась в этом смысле к началу 1970-х гг. с созданием в технической кибернетике аппарата *экстремального управления* (поисковой оптимизации) – универсального управляющего инструментария, применимого в том числе и при моделировании сложных природных систем. Именно этот аппарат положен в основу *информатико-кибернетической концепции* [5-7], в рамках которой живая природа рассматривается как иерархическая кибернетическая система (управления, или точнее, – *самоуправления*), представители уровней/ярусов которой перманентно реализуют процесс своего приспособительного

поведения, осуществляя это по алгоритмам *иерархической адаптивной поисковой оптимизации* (процесса движения к цели, в отличие от оптимальности – состояния достигнутой цели).

В рамках предлагаемой концепции к фундаментальным свойствам системы живой природы отнесена *обобщенная адаптивность*, или *обобщенная приспособляемость*, реализующая имманентную системе живого тенденцию к активности, экспансивности и структуризуемости посредством стремления к *гармоничности*, т.е. к согласованию внутренних интересов подсистем с потребностями их внешнего окружения. Термин *обобщенная* указывает здесь именно на то важнейшее явление, что адаптивность проявляется – причем *перманентно* – по отношению к изменениям не только *внешней* (для целостной системы живой природы) неспецифической среды, но и её *внутренней* специфической среды. Это справедливо для всех основных уровней интеграции системы живого.

Таким образом, для того, чтобы выявить особенности процессов биологической эволюции, следует:

1) построить кибернетическую модель иерархической биологической системы как целого,

2) изучить закономерности приспособительного поведения представителей всех уровней/ярусов в модельной иерархии (каждый из которых изменяется со своим характерным временем), и

3) интерпретировать соответствующие модельные переменные как эволюционные.

Основными средствами реализации поисково-оптимизационного процесса являются:

1. селекция направленная (задающая тенденцию изменения приспособительного поведения субъектов группы в направлении экстремума энергетического целевого критерия № 1),

2. селекция канализирующая (задающая предпочтительные варианты приспособительного поведения субъектов группы согласно целевому критерию № 2),

3. селекция ограничивающая (запрещающая нежелательные варианты приспособительного поведения субъектов группы согласно целевому критерию № 3),

4. селекция посредством внешнего дополнения (никак не связанная с проявлениями активности субъектов в составе группы),

5. канализируемость структурная (посредством параметрических влияний системной памяти [5] на все перечисленные четыре типа селекции).

Таким образом, эффективный и универсальный механизм/контур поисковой оптимизации (что продемонстрировано в технической кибернетике: см. напр., [8]) должен включать, по крайней мере, три основных элемента [5]:

- генератор *имманентной активности* поисковых переменных в контуре оптимизации;
- *селектор* получаемого результата, наличие которого однозначно определяет и присутствие в контуре оптимизации блока вычисления её *целевой функции*;
- *системную память*, реализующую свойство адаптивности поискового алгоритма.

Представим интерпретацию указанных элементов механизма/контура поисковой оптимизации в биологических терминах.

1) Поисковая активность. Представители всех ярусов в иерархии живого проявляют ***активность***, которую можно интерпретировать как *поисковую*. По мнению В.С.Ротенберга, «Поисковая активность – активность, направленная на изменение ситуации (или отношения к ней) при отсутствии определенного прогноза результатов этой активности, но при постоянном учете достигнутых результатов» [9]. К поисковым «рысканиям» относятся, например, тремор, флуктуации, дрожания и т.п. действия многоклеточных организмов, характерные небольшими вариациями (колебаниями, модификациями, изменениями) уровня их специфической активности относительно «среднего» значения. Об активности живого пишут Л.С.Берг («Замечательно, что организм обладает способностью активно приспособляться к среде, обнаруживая при этом как бы присутствие некоего *внутреннего регулирующего принципа*» [10]), Н.А.Бернштейн («Активность выступает как наиболее общая всеохватывающая характеристика живых организмов и систем» [11]) и др. М.Д.Голубовский отмечает: «открытия в области подвижной генетики показали, что клетка как целостная система в ходе отбора может адаптивно перестраивать свой геном. Она способна ответить на вызов среды *активным*

генетическим поиском, а не пассивно ждать случайного возникновения мутации, позволяющей выжить» [12]. Применительно к интерпретации эволюционного процесса: поисковая активность является непосредственной причиной «изменчивости» – важнейшей составляющей дарвиновской «триады».

2) Энергетические критерии поисковой оптимизации. С позиций предлагаемой концепции, биообъекты, относящиеся к «основным уровням интеграции живого» (т.е. «элементонам» – прокариотам и ультраструктурным внутриклеточным элементам, сложным клеткам-эвкариотам, многоклеточным организмам, биогеоценозам и Биогеосфере в целом), проявляют *тенденцию* вести себя таким образом, чтобы перманентно выполнялось условие: $Q_t = \frac{E_t^{(p)}}{E_t^{(\phi)}} \to \min$, где: t – интервал усреднения времени; $E_t^{(\phi)}$ – поток энергии, затрачиваемой на обеспечение специфической функциональной активности биообъекта (напр., для клетки – рецепции входных сигналов и обеспечения выходной реакции); $E_t^{(p)}$ – поток дополнительной энергии, вырабатываемой во время переходных процессов в биообъекте и затрачиваемой на перестройку его внутренней структуры. Речь идёт именно о долговременной тенденции, ибо реакция биообъектов на текущие раздражители связана с увеличением $E_t^{(p)}$. Легко показать, что минимизация Q_t (т.е. целевого критерия/целевой функции поисковой оптимизации) эквивалентна максимизации «к.п.д.» такого биообъекта (если рассматривать, с точки зрения внешней специфической среды, в качестве «полезных» $E_t^{(\phi)}$, а в качестве «бесполезных» – $E_t^{(p)}$). Близкие трактовки эволюционной роли энергетических потоков в биосистеме высказываются в ряде работ. Они опираются на представления о всеобщности применения вариационных принципов в естествознании, о которых Н.Н.Моисеев пишет: «любые законы неживого мира – вариационные принципы, из которых следуют законы сохранения в механике, электродинамике и других областях физики, второй закон термодинамики, закон минимума диссипации энергии и многие другие, описываемые в терминах «эволюционной теории» в широком смысле, поскольку все эти законы являются, по сути дела, тем или иным отбором реальных движений, реальных траекторий из числа виртуальных, т.е. мысленно возможных» [13]. И.И.Гительзон прямо указывает, что «наибольший интерес представляет стремление найти количественное выражение *энергетической зависимости эволюции* и вывести её принципы через энергетические принципы экстенсивного и интенсивного развития. Эти показатели приобретают не оценочное, а фундаментальное значение *факторов, направляющих эволюцию*, если строго доказать, что каждый элементарный шаг отбора на уровне микроэволюции дает энергетический выигрыш» [14]. С позиций предлагаемой концепции в качестве «элементарного шага отбора» следует рассматривать иерархическую совокупность серий поисковых шагов оптимизации энергетики живого, производимых на соответствующих ярусах его иерархии, а не какой-то отдельный поисковый шаг отдельного представителя того или иного яруса (могущий приводить и к локальному «проигрышу»). Применительно к интерпретации эволюционного процесса: совокупность целевых критериев поисковой оптимизации является необходимым элементом механизма «естественного отбора» – другой важнейшей составляющей дарвиновской «триады».

3) Системная память – понятие, которое было введено в работе [5] как характеристика запоминания/забывания влияний со стороны высших ярусов иерархической поисковой оптимизационной системы живой природы на процесс формирования приспособительного поведения элементами каждого из вложенных в последние (по принципу «матрешки») низших ярусов. Именно наличием межъярусной *системной памяти* живая природа отличается от неживой. Можно утверждать, что первичные механизмы поисковой оптимизации в системе природы, т.е. те механизмы, которые работали, начиная с момента её возникновения (как *системы*), могли быть только *абсолютно случайными*, с равномерным распределением вероятностей случайного шага во всех направлениях поискового пространства. Но затем, по мере *развития* системы природы и выделения в ней подсистемы *живой природы*, среди этих ме-

ханизмов с необходимостью (за счет появления средств запоминания предыстории поиска, т.е. *системной памяти*) стали появляться их адаптивные варианты. Последние, обладая существенно более высокой эффективностью, и, в частности, возможностью перманентно обеспечивать компромисс между противоречивыми требованиями к поиску (повышения его быстродействия при одновременном повышении точности отслеживания экстремума целевой функции) – находят себе все большее применение. Но при этом они не вырождаются в «полностью» детерминированные алгоритмы, сохраняя в своих основных характеристиках ту или иную степень случайности. Применительно к интерпретации эволюционного процесса: *системная память* иерархической системы живого обеспечивает ***канализируемость*** его эволюции, т.е. феномен, на который указывает ряд номогенетических теорий.

Процессы во *всех* пяти контурах поисковой оптимизации энергетики живой природы (рис. 1) реализуют процессы обобщённой адаптации. Но традиционно сравнительно быстро (по отношению к темпу нашего их восприятия) протекающие приспособительные процессы, т.е. происходящие в иерархических контурах системы живого «Сложные клетки–Элементоны» и «Многоклеточные организмы–Сложные клетки», называют *адаптацией* (соответственно *клеточной* и *организменной*). По аналогии сравнительно медленно (по отношению к темпу нашего их восприятия) протекающие приспособительные процессы, т.е. происходящие в контурах «Биогеоценозы–Многоклеточные организмы» и «Биогеосфера–Биогеоценозы», следует называть *эволюцией* (соответственно *биогеоценотической* и *Биогеосферной*). К определяемым таким образом процессам эволюции нельзя относить процессы, связанные с ростом иерархии самой системы живого. Т.е. в рамках предлагаемой концепции намечается новый ракурс исследования эволюционных процессов в живой природе, при котором микроэволюционные исследования отдельных видов выступают в качестве первичных этапов изучения фрагментов более «объёмной» биогеоценотической эволюции, а макроэволюционные – фрагментов более «объёмной» Биогеосферной эволюции.

Определив понятие «***метаэволюция***» *как процедуру наращивания числа ярусов в иерархической системе в ходе её формирования как таковой*, укажем, что именно метаэволюция описывает процессы, связанные с возникновением в отдалённом прошлом Земли представителей новых ярусов развивающейся иерархической системы живого: вначале, ориентировочно около 3,6 млрд. лет назад, прокариот, затем, около 2,59 млрд. лет назад, – сложных клеток-эвкариот, далее, около 1,58 млрд. лет назад, многоклеточных организмов, и, наконец, около 0,57 млрд. лет назад, биогеоценозов. В настоящее время идёт формирование Биогеосферы как *единой* самоуправляющейся системы [5], но пока оно находится в своей начальной фазе (т.е. эффективность оптимизационного процесса минимальна). Предлагаемая концепция базируется на том, что процесс метаэволюции живого не является приспособительным, а, напротив, происходит вполне регулярно (причём внутри указанных миллиардолетий длительности её меньших фаз подчиняются простой математической закономерности). В противоположность этому процессы как клеточной и организменной адаптаций, так и биогеоценотической и Биогеосферной эволюций являются приспособительными.

На рис. 2 приведена реконструкция схемы «дарвиновского» механизма биологической эволюции, т.е. то видение этого механизма, которое могло бы быть у современных последователей Ч.Дарвина – «неодарвинистов», если бы они рассмотрели его иерархически, причём «спроецированная» на схему механизма иерархической адаптивной поисковой оптимизации живой природы [5]. Очевидно, что сложность их несопоставима…

Предлагаемая кибернетическая иерархическая схема живой природы демонстрирует, что в эволюционном процессе участвуют представители не только четырёх ярусов иерархии живого – т.е. отдельные гены, геномы как целое, организмы и популяции, – но и представители девяти других его ярусов. При этом иерархически входящие в эволюционирующие биоценозы организмы, эвкариотические клетки и ультраструктурные внутриклеточные элементы перманентно (каждый в своем типичном времени, различающемся на несколько порядков!) *оптимизируют* – в энергетическом смысле и в иерархически достаточно автономных режимах – *свои структуры и поведения* в своих контурах оптимизации. Именно этот факт

естественным образом объясняет высокую эффективность и относительно сжатые сроки (~4,6 млрд. лет) всего интегрального биоэволюционного процесса на Земле – явление, которому креационисты не находят другого объяснения, кроме проявления Божественной воли.

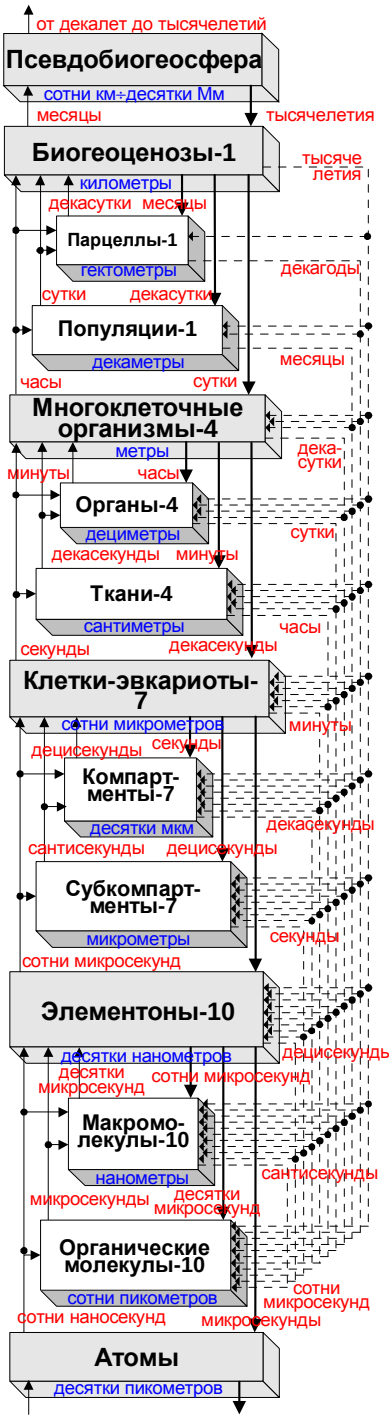
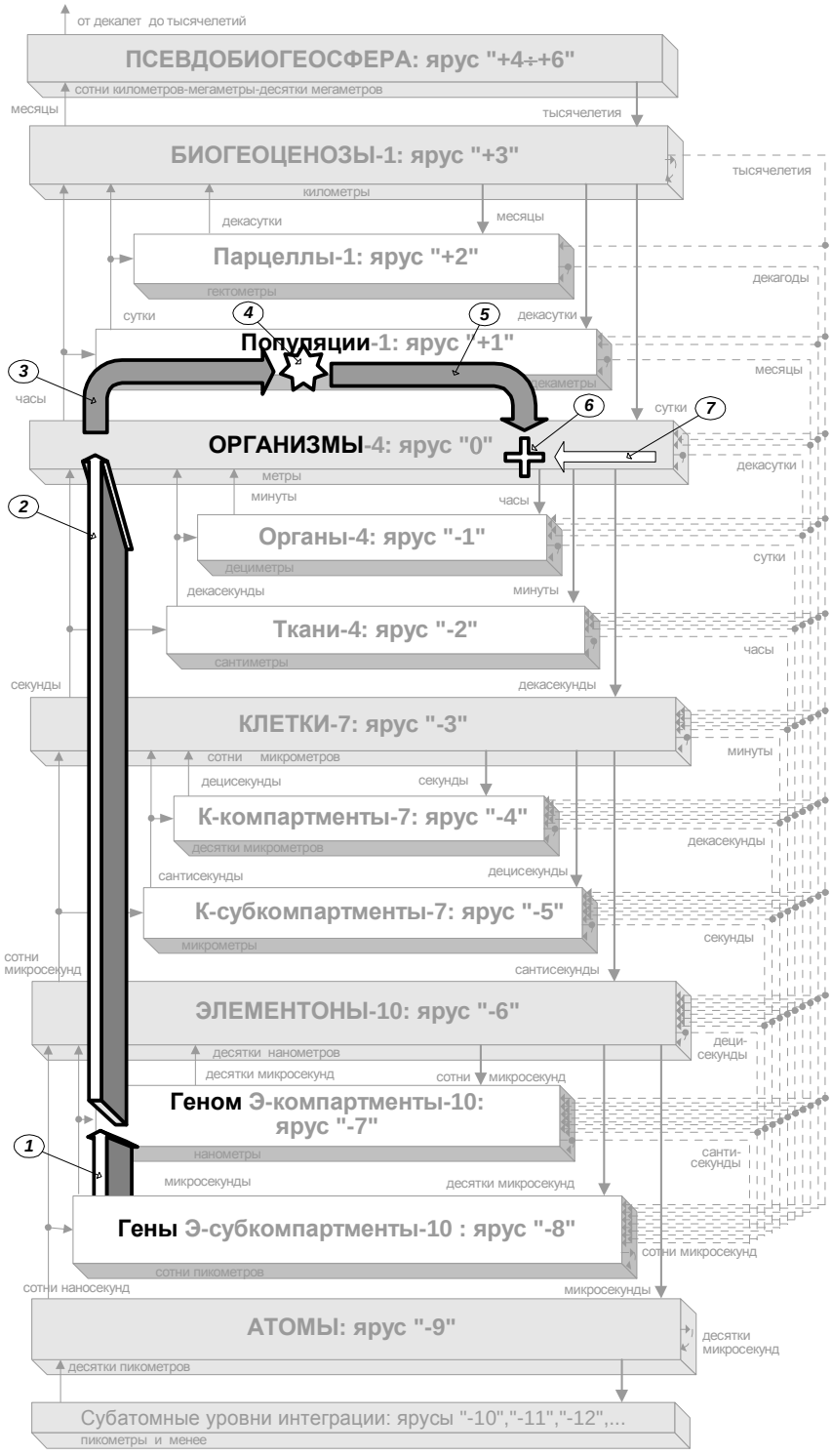


Рис. 1. Схема механизма иерархической адаптивной поисковой оптимизации живого.



Примечание: стрелки, направленные вверх, имеют структуру (отражают отношение) «многие - к одному», а вниз - «один - ко многим».

Рис. 2. Проекция реконструкции «неодарвинистской» модели биологической эволюции на схему механизма иерархической адаптивной поисковой оптимизации живой природы.

На рис.2: 1 – активность, т.е. изменчивость, проявляемая генами; 2 – процесс онтогенеза, отражающий наследственность; 3 – проявление активности организмов в рамках популяций; 4 – селектор получаемого на базе конкурентной борьбы между особями результата, т.е. естественный отбор на базе критерия «приспособленности»; 5 – процесс сообщения отрицательного результата селекции на организменный уровень; 6 – процедура изъятия организма из популяции; 7 – «внешнее дополнение», также приводящее к его устранению из популяции.

Кроме того, становится понятным, почему *конкретные события* эволюционного процесса – в отличие от его *тенденции* – принципиально не могут быть предсказуемы: случайная составляющая поиска присутствует в нём всегда (даже если поведение каждого из активных компонентов оптимизационного контура вполне регулярно с их «точки зрения», то поведение всей совокупности таких компонентов – т.е. на уровне высшего иерархического яруса – выглядит как случайное: у последнего отсутствуют возможности для фиксации поведения каждого из его компонентов).

Что же касается дарвинистской схемы, то необходимо признать, что из трёх основных составляющих поискового оптимизационного механизма биоэволюции Ч.Дарвином гениально угадано (причём около 150 лет назад, за сотню лет до создания кибернетики!) примерно «полтора»:

1) введено (хотя и неявно) понятие «генератора активности», что следует из трактовки им понятия «*изменчивости*» (которое вполне можно интерпретировать в качестве «поисковой активности») как фундаментального в его теории;
2) введён прообраз селектора – «*естественный отбор*» и прообраз целевой функции (правда, не в полной форме, включающей целевой критерий экстремального типа, а лишь в форме целевых ограничений типа неравенств и селекции посредством внешнего дополнения) – т.е. прообраз собственно контура оптимизации как такового (но всего лишь одного, а не пяти, иерархически вложенных друг в друга, как это оказывается).

В свою очередь, многие факты, которыми обосновывают теории номогенеза, хорошо ложатся в схему третьей составляющей этого механизма – системной памяти живого. Тем самым дарвинистские и номогенетические теории эволюции – точнее, *их* ***модифицированные версии*** *– объединяются в рамках единой схемы перманентной* ***иерархической адаптивной поисковой оптимизации*** *живой природы.*

ЛИТЕРАТУРА

1. Жерихин В.В., Раутиан А.С. Филоценогенез и эволюционные кризисы. 1997. – http://www.macroevolution.narod.ru/rautian2.htm
2. Воронцов Н.Н. Развитие эволюционных идей в биологии. М.: Издат. отдел УНЦ ДО МГУ, Прогресс-Традиция, АБФ, 1999.
3. Шмальгаузен И.И. Кибернетические вопросы биологии. Новосибирск: Наука, 1968.
4. Ляпунов А.А. Проблемы теоретической и прикладной кибернетики. М.: Наука, 1980.
5. Гринченко С.Н. Системная память живого (как основа его метаэволюции и периодической структуры). – М.: ИПИРАН, Мир, 2004. – 512 с. – см. также http://www.ipiran.ru/publications/publications/grinchenko/
6. Гринченко С.Н. Метаэволюция (систем неживой, живой и социально-технологической природы). – М.: ИПИРАН, 2007. – 456 с. – см. также http://www.ipiran.ru/publications/publications/grinchenko/book_2/
7. Гринченко С.Н. Теории биологической эволюции и кибернетика: новый синтез // Эволюция: Дискуссионные аспекты глобальных эволюционных процессов. Междисциплинарный альманах, вып. 3. М.: Либроком, 2011, С. 81-102.
8. Растригин Л.А. Адаптация сложных систем. Методы и приложения. Рига: Зинатне, 1981.
9. Ротенберг В.С. Поисковая активность, сон и устойчивость организма // Кибернетика живого: человек в разных аспектах. М.: Наука. 1985. С. 80-91.
10. Берг Л.С. Номогенез,или эволюция на основе закономерностей. Петербург,Госиздат,1922.
11. Бернштейн Н.А. Очерки по физиологии движений и физиологии активности. М.: Медицина, 1966.
12. Голубовский М.Д. Неканонические наследственные изменения // Природа, 2001, №8: 3-9, №9: 3-8.
13. Моисеев Н.Н. Коэволюция человека и Биосферы: кибернетические аспекты // Кибернетика и ноосфера. М.: Наука, 1986. С. 68-81.
14. Гительзон И.И. От редактора // В кн.: Печуркин Н.С. Энергетические аспекты развития надорганизменных систем. Новосибирск: Наука, 1982. С. 3–4.

ФОРМА НОСОВОЙ ПОЛОСТИ И СТРУКТУРА ТЕЧЕНИЯ ВОЗДУХА В НОСУ ЧЕЛОВЕКА. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

В.М. Фомин, В.Л. Ганимедов, М.И. Мучная, А.С. Садовский, В.Н. Шепеленко

Институт Теоретической и Прикладной Механики им. С.А. Христиановича СО РАН
630090 Новосибирск, Россия

Развитие математического моделирования и компьютерной техники, создание новых численных методов, появление доступных коммерческих пакетов прикладных программ способствовало в последние годы внедрению численного эксперимента в различные отрасли знаний, казалось бы далеких от математики. Интересы данной работы лежат в области медицины, а точнее - физиологии дыхания. Сложное анатомическое строение дыхательных путей человека является препятствием для детального изучения картины течения воздуха с помощью традиционных методов визуализации и диагностики потоков. В настоящей работе приведены результаты численного моделирования течения в носовой полости человека с помощью пакета прикладных программ "FLUENT". На его основе построены геометрические модели носовых полостей 15 конкретных людей, восстановленных по сериям томографических снимков носа. В рамках уравнений Навье - Стокса для них проведены расчеты течения воздуха при различных перепадах давления на режимах вдоха и выдоха. Получены поля скорости, давления, температуры, величина расхода воздуха. Проведен анализ полученных результатов в плане влияния строения носовой полости на структуру течения. Отмечено чрезвычайное разнообразие форм и структур. Предпринята попытка некоторого обобщения.

ИССЛЕДОВАНИЕ ЗАКОНОМЕРНОСТЕЙ, ОБНАРУЖЕННЫХ В ТАБЛИЦАХ ИСПОЛЬЗОВАНИЯ КОДОНОВ.

Кузьмина О.В.

Тюменский государственный университет

e-mail: Angel_RT@qip.ru

Семёнов Д.А.

e-mail: dasem@mail.ru

Одной из характеристик геномов является набор частот встречаемости определенных кодонов в кодирующей части генов. Представление данных в такой форме имеет сходство с традицией, существующей при статистическом анализе литературных текстов. Частота использования определенных букв может характеризовать очень общие свойства языка. Использовать эту информацию для анализа авторства текста или жанра текста уже не перспективно. О содержании текста такая информация просто ничего не говорит. Сама по себе она даже ничего не говорит о ближнем порядке в расположении букв: хотя и позволяет выдвигать некоторые гипотезы на сей счет, но проверка гипотез требует уже привлечения дополнительных данных.

С учетом всего сказанного, неочевидно, что таблицы частот использования кодонов (codon usage) будут содержать важные закономерности, носящие общий характер для всех организмов. Тем не менее, ряд закономерностей выявляется, и можно даже сказать: «бросается в глаза». Так, почти общий характер носят закономерности: 1) Доминирование гуанина над цитозаном в первой позиции кодона (G>C для 100% исследованных геномов) 2) Доминирование аденина над урацилом в среднем по кодирующей части мРНК (для 80% геномов) и для первой позиции кодона (86% геномов). 3) Доминирование урацила над аденином в третьей позиции кодона (более чем для 80% геномов). 4) «Вымывание» аргининовых кодонов CGA и CGG вследствие AT-обогащения генома (замена на AGA и AGG). 5) Пороговый характер встречаемости цитозина и урацила в третьей позиции кодона, указывающий на неустойчивость процесса.

Из всех вышеперечисленных закономерностей в данной работе важны пункты 1-3. То есть предметом нашего рассмотрения будет асимметрия распределения нуклеотидов в среднем по кодирующей части, а также в первой и третьей позиции кодона. Именно эти явления будут нами рассмотрены совместно.

Материалом для анализа первоначально служила выборка из базы данных (http://www.kazusa.or.jp/codon/) , содержащей данные о частоте встречаемости кодонов для различных организмов. Выборка была сформирована путем упорядочения всех данных по объему и отбора наиболее представительных данных. Таким образом, в выборку вошли 1100 «геномов» представленные более чем 50000 триплетов каждый.

На первом этапе были построены распределения для интересующих нас величин, характеризующих в каждом случае асимметрию встречаемости нуклеотидов: 1) доля аденина, по отношению к сумме частот аденина и урацила A/(A+U), 2) то же самое для первой позиции кодона A1/(A1+U1), 3) то же самое для третьей позиции кодона

A3/(A3+U3), 4) доля гуанина по отношению к сумме частот гуанина и цитозина G/(G+C). Исследование полученных распределений привело к выбору инструмента для последующей работы: так как было показано, что распределения значимо отличаются от нормального, то для исследования использовались непараметрические критерии, в частности критерий Колмогорова-Смирнова.

Как известно, в ДНК содержится попарно одинаковое количество аденина и тимина; цитозина и гуанина. Для синтеза мРНК выбирается толька одна из цепей двойной спирали и, как оказалось не случайно. На этот факт указывает исследованная нами асимметрия. На втором этапе была продемонстрирована статистическая значимость обнаруженных закономерностей. Что каждый раз сводилось к демонстрации значимости смещения математического ожидания от значения 0,5.

Накопленное в современной биологии количество фактических данных позволяет выстраивать направление исследования от общего к частному. Первоначально закономерности были найдены в данных по всему геному. Существование возможной специфики на уровне отдельных генов заставило сформировать выборку на основе кодирующей последовательности случайно выбранных генов. На уровне отдельных генов все три закономерности оказались статистически значимы.

На третьем этапе была сформулирована гипотеза, объясняющая все три закономерности на основе представления о совместном влиянии AT-обогащения генома и структуры генетического кода. Характеристикой обнаруженной асимметрии будем считать отношение количества аденина в кодирующей части к суммарному количеству аденина и урацила A/(A+U). Считая критерием AT-обогащения, процент AT-пар приходящийся на кодирующую часть генома, построим зависимость асимметрии от AT-обогащения. В выбранных координатах построили линейную регрессию.

Усредненная асимметрия аденина и урацила действительно хорошо согласуется с этой гипотезой (коэффициент корреляции 0,4). Для асимметрии в первой и второй букве кодона гипотеза не подтвердилась, это можно сформулировать так: даже если аденина мало, то в первой позиции кодона его больше, чем урацила; даже если урацила мало, то в третьей позиции его больше, чем аденина.

На четвертом этапе работы нас интересовало последовательное сочетание кодонов, поэтому одних данных по частоте встречаемости было не достаточно. Выборка, составленная на основе отдельных генов, позволяет проверить простую гипотезу: накопление урацила в третьей позиции кодона и аденина в первой позиции можно объяснить из предположения о наличии взаимодействия между третьей и первой позициями соседних последовательных кодонов. Иными словами, ожидается не случайное сочетание кодонов вида (NNU)(ANN). Повышенная частота таких сочетаний могла бы обеспечивать большую частоту встречаемости стоп-кодонов при сдвиге рамки считывания на две позиции вправо.

Вопреки ожиданиям, проверка данной гипотезы продемонстрировала крайне низкое число сочетаний данного типа. Даже при сдвиге рамки считывания стоп-кодоны избегаются, что противоречит работе [1]. Более того, хотя одновременно существует доминирование аденина в первой и доминирование урацила в третьей позиции кодона, но величины

характеризующие эти процессы коррелируют значимо отрицательно (-0,7). То есть сильная асимметрия в первой позиции не сопровождается сильной асимметрией по третьей позиции, что может интерпретироваться как следствие избегания комбинаций (NNU)(ANN).

Одновременно, обращает на себя внимание значительная частота сочетаний кодонов вида: (NNU)(GNN). Можно предположить, что данное сочетание влияет на фолдинг мРНК, что позволяет рассматривать данный результат в русле работ [2, 3].

Гуанин, также как и аденин значимо накапливается в первой позиции кодона (причем в 100% исследуемых геномов). На материале отдельных генов продемонстрировано накопление пар UG, где урацил находится в третьей, а гуанин в первой позиции соседних кодонов. Возникает вопрос: насколько взаимное влияние урацила и аденина позволяет объяснить общую тенденцию накопления этих нуклеотидов в данных позициях? Для ответа на этот вопрос исследуем связь двух величин: 1) разность частот гуанина в первой позиции и в среднем по геному 2) разность частот урацила в третьей позиции и в среднем по геному. Две эти величины оказываются линейно связаны (коэффициент корреляции 0,678). То есть гипотеза о взаимном влиянии урацила и гуанина в этих позициях хорошо объясняет закономерности накопления урацила в третьей, а гуанина в первой позициях кодона.

1. Г.Г. Малинецкий, С.А. Науменко, А.В. Подлазов Об экстремальных свойствах разметки генетического кода. ДАН 2007 т.414 №6 с.831-835
2. T. Warnecke L.D. Hurst GroEL dependency affects codon usage—support for a critical role of misfolding in gene evolution. Molecular Systems Biology 6:340
3. D.A. Semenov Evolution of the genetic code. Emergence of stop codons arXiv:0710.5825 2007

СРАВНЕНИЕ СИСТЕМЫ «DISCOVERY» с MICROSOFT ASSOCIATION RULES[1]

Николай Иванович Фирсов
Институт Систем Информатики им. А.П. Ершова
e-mail: <nickfirsov@gmail.com>

Abstract

We have developed a relational approach (Relational Data Mining) to Data Mining, and software system «Discovery», that realize this approach. The goal of this paper is to provide both theoretical and experimental comparisons of «Discovery» system versus Microsoft Association Rules. We show that «Discovery» system performs better for rules discovery and prediction tasks than the Association Rules, and, can analyze high-noisy data, for example, financial time series.

Аннотация

Нами был разработан реляционный подход (Relational Data Mining) к методам извлечения знаний и программная система «Discovery» [1, 2, 4], снимающие практически все ограничения с методов KDD&DM (Knowledge Discovery in Data Bases and Data Mining).

Система Discovery, обладает следующими важными теоретическими свойствами: может обнаруживать теорию предметной области, может обнаруживать все правила, имеющие максимальные условные вероятности, может обнаруживать непротиворечивую вероятностную аппроксимацию теории предметной области [2], обнаруживает все максимально специфические правила, позволяющие предсказывать без противоречий [3].

В данной работе проводится теоретическое и экспериментальное сравнение системы «Discovery» с алгоритмом Microsoft Association Rules. Мы показали, что система «Discovery» больше подходит для обнаружения закономерностей и прогнозирования, чем Association Rules, а также, в отличие от алгоритма Association Rules, метод Discovery позволяет обнаруживать знания в сильно зашумленных данных, какими, например, являются финансовые временные ряды [4].

ЛИТЕРАТУРА

[1] Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. – НГУ, Новосибирск, 2006. – 293 с.
[2] Evgenii Vityaev, Boris Kovalerchuk, Empirical Theories Discovery based on the Measurement Theory. Mind and Machine, v.14, #4, 551-573, 2004)
[3] Evgenii Vityaev. The logic of prediction. In: Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference (August 16-19, 2005, Novosibirsk, Russia), World Scientific, Singapore, 2006, pp.263-276).
[4] Kovalerchuk, B.Y., Vityaev E.E. Data mining in finance: Relational and hybrid methods, Kluwer, 2000, 308 p.

[1] Работа поддержана грантом РФФИ РФФИ **№11-07-00560а**; интеграционными проектами СО РАН № 47, 115, 119, а также работа выполнена при финансовой поддержке Совета по грантам Президента РФ и государственной поддержке ведущих научных школ (проект НШ-3606.2010.1.)

PRINCIPLES, THEORY, MODELS OF STABILITY AND CONTOLLABILITY OF EXPERIMENTAL CLOSED ECOSYSTEMS

Degermendzhy A.G., Barcev S.I., Barchatov Yu.V.
Institute of Biophysics SB RAS, e-mail: ibp@ibp.ru

Biosphere in first approximation (taking into consideration the most general characteristics) can be characterized by the inflow of light energy, the presence of differentiated alive and inert matter and high degree of thermodynamic closedness. In the planetary scale the possibility of biocycles destruction due to the formation of biologically non-degradable matters (‘‘deadlocks’’) can be absolutely real, besides other mechanisms of biosphere degradation (toxicity, warming, ozone holes, etc.). The same problem exists in constructing artificial closed ecosystems, e.g. with the aim of life support for crew of spacecrafts, underwater and arctic settlements and the possible model prototype of noosphere. Manmade ecosystems with an essentially closed material cycle can be an efficient instrument for experimental modeling of biospheric processes, including the investigation of their tolerance towards human impact. It is well known that the life support systems intended for long-duration functioning should be based on biological material cycling. A high degree of closure of the material cycle in manmade ecosystems can be only attained by selecting the system components on a sound scientific basis and by achieving the necessary cycling rates of chemical elements and components and their conjugation. However, if the designers of the system ignore the knowledge of the closure mechanisms responsible for maintaining the long-duration and steady material cycling, there may be serious harmful consequences affecting the whole system. The example of such a failure is the attempt to keep “Biosphere-2” (USA) self-contained for a long period of time. Theory and two principles of natural closure of ecosystems: “ecological” (models and experiments) and “evolutionary” are presented in detail. Theory of optimum layout of closed ecosystems, composition control mechanisms and applications of these approaches to modeling essential properties of biosphere are given.

К вопросу выражения требований нормативно-правовых актов в виде формальных спецификаций бизнес-процессов

Ярёменко Ф.В.[1]
[1] Вычислительный центр им. А. А. Дородницына РАН
e-mail: yaremenkofv@gmail.com
Бездушный А. Н.[1]
e-mail: anb@ccas.ru

Аннотация

В работе на примере Федерального Закона №94 «О размещении заказов на поставки товаров, выполнение работ, оказание услуг для государственных и муниципальных нужд» (94-ФЗ) [1] рассматривается попытка выражения требований законодательства в виде формального описания соответствующих бизнес-процессов, специфицированных в виде BPMN потоков работ (workflow). Формализация требований законов ориентирована на упрощение автоматизации их реализации в государственных и коммерческих структурах, выявление неточностей, неясностей, неполноты и противоречивости законодательства.

Актуальность задачи

В последние годы в Российской Федерации особое внимание уделяется внедрению информационных технологий в органы законодательной и исполнительной власти. С каждым годом растёт рынок программных решений для государственных и коммерческих учреждений. Федеральные законы регулируют различные сферы: государственные закупки, бухгалтерский учёт, землеустройство и прочее. Большинство систем, используемых для автоматизации деятельности и управления предприятием, построено на основе действующего законодательства. Разработчики автоматизированных информационных систем, опирающихся на нормативно-правовые акты, имею различные интерпретации законодательства, ограниченные возможностями этих систем. Разработка спецификации законодательства в нотации BPMN 2.0 (Business Process Model and Notation) [2] даёт возможность получить единый набор формальных моделей, описывающий нормативно-правовые акты, предназначенный для многократного использования и исключающий двойственность интерпретации. Описание бизнес-процессов российского законодательства позволит строить автоматизированные информационные системы, применимые и в государственных, и в коммерческих организациях.

Преимущества использования нотации BPMN 2.0

BPMN 2.0 является графической нотацией, диаграммы в ней дают наглядное представление предметной области. В результате разработанная модель легко читается аналитиками, инженерами и программистами. BPMN 2.0 поддерживает подпроцессы, что позволяет использовать принцип декомпозиции. Декомпозиция даёт возможность постепенно и структурировано представлять модель процесса в виде иерархической структуры отдельных диаграмм, что делает ее менее перегруженной и более понятной для разработчиков.

При изменении законодательства необходимо внести поправки в разработанный набор моделей для актуализации сценария работы всех систем, построенных на нём, с незначительными доработками на местах. В итоге обеспечивается оперативность реагирования на обновление нормативно-правовых актов.

Спецификация BPMN 2.0 содержит описание процедуры трансляции диаграмм в язык BPEL (Business Process Execution Language) [3]. Язык BPEL был стандартизован в 2004 году и на сегодняшний день широко используется. BPEL подразумевает реализацию описанных потоков работ в виде веб-сервисов. Использование веб-сервисов даёт возможность внедрять системы в организациях с территориально-распределённой структурой и использовать их в межведомственном взаимодействии.

Заключение

Полученные описания бизнес-процессов были применены в разработке электронной торговой площадки для задания порядка торгов, их контроля и мониторинга. Были внесены поправки в модели в соответствии с изменениями в законодательстве и протестирована площадка, оценена корректность работы. В будущем планируется использовать описанный подход при реализации информационных систем для автоматизации различных сфер деятельности.

Литература

[1]. Федеральный Закон №94 «О размещении заказов на поставки товаров, выполнение работ, оказание услуг для государственных и муниципальных нужд» от 21 июля 2005 г. http://www.consultant.ru/online/base/?req=doc;base=LAW;n=100596

[2]. Business Process Model and Notation. Версия спецификации 2.0 http://www.omg.org/spec/BPMN/2.0/

[3]. OASIS Web Services Business Process Execution Language (WSBPEL) Technical Committee http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel

ТЕХНОЛОГИЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ В ДИАГНОСТИКЕ ПРОФЕССИОНАЛЬНЫХ НЕЙРОТОКСИКОЗОВ

Работа выполнена при поддержке гранта РГНФ № 10-06-12121в.

Иванов А.Г
Ангарский филиал УРАМН ВСНЦ экологии человека СО РАМН – НИИ медицины труда и экологии человека
e-mail: iag2009@yandex.ru
Дьякович М.П.
e-mail: marik914@rambler.ru

Рассмотрены основы информационной технологии поддержки принятия решений в диагностике профессиональных нейротоксикозов. Показана концепция, методический состав, модельная структура и порядок применения экспертной системы, реализующей такую технологию.

The basics of the information technology for decision support at occupational neurointoxications diagnosis are reviewed. For the expert system implementing such a technology, the concept, methodical solution, model structure and usage procedure are revealed.

Хронические профессиональные нейротоксикозы (ПН) развиваются у работников вследствие накопления токсикантов в организме и вызываемых ими нарушений. В современных химических производствах клинические формы ПН проявляются при стаже работы с токсикантом свыше 15 лет, и завершаются дегенеративными явлениями в обменных процессах, сердечно-сосудистой и пищеварительной системах через 6-7 лет болезни. К этому времени патологические изменения нервной системы приобретают стойкий характер [1,2] с формированием психоорганического синдрома [3] со значительными личностными нарушениями, свидетельствующих о социально-психической дезадаптации [4], что обусловливают сложность реабилитации таких больных. Хронические нейротоксикозы, составляющие небольшую долю в структуре профессиональной заболеваемости РФ, имеют большую социальную значимость, так как приводят к тяжелой инвалидизации лиц трудоспособного возраста. Среди ПН, установленных на территории Иркутской области, хроническая ртутная интоксикация (ХРИ) занимает ведущее место, что связано с использованием ртути в технологическом процессе химических предприятий.

Ранняя диагностика ХРИ и прогноз развития заболевания являются актуальными задачами в профпатологической практике, качественные улучшения которых будут способствовать сохранению трудового потенциала промышленных предприятий региона, чем обусловливается социально-экономическая значимость указанных задач.

Своевременная идентификация у работников нарушений, вызванных воздействием производственных токсикантов, позволит повысить качество лечебно-реабилитационного процесса, чему будет способствовать внедрение комплексной информационной технологии (ИТ) поддержки клинической диагностики ПН. Ее основой является человеко-машинный процесс использования формализованного профессионального врачебного знания с целью идентификации ПН и прогноза его развития.

Для реализации ИТ нами разрабатывается информационно-аналитическая система (ИАС), в которой предусматривается функционирование подсистем регистрации и систематизации результатов мониторинга дозовой токсической нагрузки и поликлинического скрининга в условиях производства [5].

Для идентификации ПН и прогноза его развития на основании фактуального материала, извлекаемого из указанных подсистем, служит входящая в состав ИАС гибридная экспертная система «ONTIS» («Occupational NeuroinToxication Identification System») (ЭС). В основе модельной иерархии ЭС лежит набор дескрипторов состояния здоровья, каждый из которых содержит формализованное клиническое описание ПН в рамках дифференциальной диагностики, в объеме, необходимом и достаточном для формирования байесовских сетей доверия (БСД), каждая из которых синтезируется на основании единственного порождающего дескриптора. Такой подход позволяет увеличить гибкость байесовской модели, а также ее прозрачность и интерпретируемость с точки зрения врача-профпатолога.

Инициирующей процедурой применения ЭС для диагностики ПН является формирование пространства наблюдений, описывающего актуальное клиническое состояние обследуемого. Такое пространство включает идентификатор группы профессионального риска, определяемый на основе превышения токсической нагрузки за время индивидуальной экспозиции предельную безопасную нагрузку (накопление токсиканта при условии его воздействия на уровне предельно допустимой концентрации за это же время), а также комплекс значений клинических индексов, охватывающий анамнестический, биохимический, психоневрологический и электрофизиологический блоки базы данных АИС. Поступлению описанных исходных данных на вход байесовского решателя предшествует их дискретизация на основе диапазонов нормы, содержащихся в дескрипторе состояния здоровья.

Диагностическое заключение в ЭС формируется как аргументированная условная вероятностная индивидуальная мера риска ПН и предназначаются для обоснования принятия решения врачом-профпатологом о направления работника на углубленное обследование в стационаре клиники профессиональных заболеваний. Указания и рекомендации для развития формальных знаний системы могут быть сформулированы как в процедуре верификации результатов вывода ЭС экспертами (профпатологическим консилиумом), так и в процессе совершенствования дифференциальной диагностики ПН на основе поисковых научных исследований.

ЛИТЕРАТУРА

1. Трошин В.В. Профессиональные нейротоксикозы//Медицинский альманах. 2010.№2. С.52-61.
2. Пивень Б.Н. Экзогенно-органические заболевания головного мозга. М.: Медицина; 1998. 144с.
3. Колесов В.Г., Мещерягин В.А., Лахман О.Л., Шевченко О.И. Психопатологические проявления в отдаленном периоде профессиональных нейротоксикозов. Журнал неврологии и психиатрии. 2005. № 1. С. 25-29.
4. Дьякович М.П., Казакова П.В. Ценностно-мотивационная сфера и качество жизни больных с хронической ртутной интоксикацией профессионального

генеза. Известия Самарского научного центра Российской академии наук. 2008. № 2. С. 32-35.

5. Иванов А.Г. Бахвалов С.В. Организация яруса данных информационно-аналитической системы идентификации профессиональных нейроинтоксикаций / Вестник ИрГТУ. 2011. №8. С.81-86.

Критерии выбора технологических процессов для реализации виртуальных лабораторных стендов

Н. В. Журавлева
Институт Автоматики и Электрометрии СО РАН
E-mail: zhuravleva@iae.nsk.su

Одна из важнейших задач современного учебного заведения не только дать необходимую информацию, но и помочь овладеть и усвоить ее. При этом знания являются не столько целью обучения, сколько средством для решения задач и выполнения некоторой профессиональной деятельности. Исходя из того, что у человека в памяти остается 10% того, что он услышал, 50% того, что увидел и 90% того, что он выполнил сам, практическая работа является важным фактором усвоения учебного материала [1].

При изучении языков программирования по созданию управляющих алгоритмов практические занятия крайне необходимы, т.к. невозможно получить практический опыт без применения средств, на которых можно было бы отрабатывать навыки программирования и иметь возможность наблюдать за результатом выполнения учебных программ. Поэтому лабораторные работы должны отображать типовые задачи из области промышленной автоматизации. Для реализации этого компонента обучения потребуются значительные затраты, таким образом, качество обучения во многом зависит от возможности учебного заведения.

Внедрение современных информационных технологий дает возможность создавать виртуальные лабораторные стенды (ВЛС)[2], лишенные таких недостатков физических средств как:

- значительные временные и денежные затраты по созданию и поддержке работоспособности стенда;
- аварийная опасность;
- ограниченное количество экземпляров для работы;
- привязка к определенному месту.

ВЛС позволяет реализовать дорогие и опасные технологические процессы, включить в процесс обучения игровой элемент и реализовывать «нештатные ситуации» (авария, взрыв, пожар). Для разработки лабораторного практикума необходимо выбрать набор типовых технологических процессов.

Фокусирование на проектировании технологического процесса (ТП) позволяет из одного процесса создать несколько сценариев работы путем вариации требований. Такой подход позволит индивидуализировать лабораторную работу и постепенно наращивать сложность управляющих алгоритмов. Для создания хорошо продуманного, адекватного и полезного ВЛС необходимо рассмотреть реальные, интересные ТП.

При выборе ТП для дальнейшей реализации были выработаны следующие критерии отбора:

- логическое управление;
- работа с временными интервалами;
- возможность задания различных сценариев для работы;
- работа с параллельными потоками управления.

Желательно, чтобы ТП соответствовал хотя бы одному критерию.

Для придания реалистичности работе ВЛС при проектировании ТП необходимо учитывать специфику его развития, определяемую количеством нештатных ситуаций (поломка оборудования, реакция на аварию соседних звеньев) и возможностью отработки действия специалиста для ее устранения.

Литература

1. С. В. Ланько. Индивидуально-методологический подход в технологии обучения информатике. Известия Российского государственного педагогического университета им. А.И. Герцена.: Изд-во РГПУ им. А. И. Герцена, №98, 2009.
2. Зюбин В.Е., Калугин А.А. Виртуальные лабораторные стенды: обучение программированию задач промышленной автоматизации // Промышленные АСУ и контроллеры. 2009. № 2. С. 39-43.

ИСПОЛЬЗОВАНИЕ ЧЕТЫРЕХЗНАЧНЫХ ЛОГИЧЕСКИХ СИСТЕМ ДЛЯ АНАЛИЗА ДНК- ВЫЧИСЛЕНИЙ

Л.А.Лютикова
НИИ ПМА КБНЦ РАН, г. Нальчик
lylarisa@yandex.ru

Аннотация.
В данной работе рассматривается вопрос о формализации свойств ДНК методами четырехзначной логики. Исследуется вопрос о возможных построениях ДНК-логических алгоритмов для решения задач распознавания, предлагается алгоритм решения поставленной задачи и логическая функция реализующая предложенный алгоритм.

ДНК-вычисления — это раздел области молекулярных вычислений на границе молекулярной биологии и компьютерных наук. Основная идея ДНК-вычислений — построение новой парадигмы, создание новых алгоритмов вычислений на основе знаний о строении и функциях молекулы ДНК и операций, которые выполняются в живых клетках над молекулами ДНК при помощи различных ферментов.

Одно из преимуществ ДНК-процессоров в сравнении с обычными кремниевыми процессорами заключается в том, что они могут производить все вычисления не последовательно, а параллельно, что обеспечивает выполнение сложнейших математических расчетов буквально за считанные минуты. Традиционным компьютерам для выполнения таких расчетов потребовались бы месяцы и годы

В молекулах ДНК имеется четыре базовых основания: аденин (A), гуанин (G), цитозин (C) и тимин (T), связанных друг с другом в цепочку.

Здесь используется не двоичная, а четверичная логика. И подобно тому, как в двоичной логике любую информацию можно закодировать в виде последовательности нулей и единиц, в молекулах ДНК можно кодировать любую информацию путем сочетания базовых оснований.

- A – 0 C – 1 G - 2 T - 3

Комплементарность оснований заключается в том, что образование водородных связей.

При соединении одинарных цепочек ДНК в двойную цепочку возможно только между парами A – T и G - C. Введем функцию К(x),характеризующую комплементарные основания, заданную следующей таблицей:

К(x)=3-x …Таблица 1

x	К(x)
0	3
1	2
2	1
3	0

Процесс соединения двух одинарных цепочек ДНК путем связывания комплементарных оснований в регулярную двойную спираль называется ренатурацией,

описывается следующей логической функцией Re*(x1,x2),* которая имеет приведенную ниже таблицу истинности

$$j_3(x_1 \oplus x_2) = x_1 \Leftrightarrow \overline{x_2} = \mathrm{Re}(x_1, x_2)$$

Re*(x1,x2)* *Таблица 2*

x_1/x_2	0	1	2	3
0	0	0	0	1
1	0	0	1	0
2	0	1	0	0
3	1	0	0	0

Обратный процесс, то есть разъединение двойной цепочки и получение двух одинарных цепочек, — денатурацией.

Копирование, или размножение, ДНК-молекул осуществляется в ходе полимеразной цепной реакции (Polymerase Chain Reaction, PCR). Процесс копирования можно разделить на несколько стадий. Он происходит лавинообразно. На первом шаге из одной молекулы образуются две, на втором — из двух молекул — четыре, а после n-шагов получается уже 2n молекул. $Cop(\mathrm{x}) = \mathrm{x} \vee \mathrm{x}$

Задача распознавания образов имеет следующую формулировку:

Соответствие множества объектов характеризующим их признаков может быть представлено следующей таблицей:

Таблица 3

x_1	x_2	x_3 …… ..	x_n	W
$x_1(w_1)$	$x_2(w_1)$	$x_3(w_1)$ ..	$x_n(w_1)$	w_1
$x_1(w_2)$	$x_2(w_2)$	$x_3(w_2)$ ..	$x_n(w_2)$	w_2
$x_1(w_m)$	$x_2(w_m)$	$x_3(w_m)$	$x_n(w_m)$	w_m

$X_j = \{x_1(w_j), x_2(w_j) \ldots\ldots x_n(w_j)\}$ -вектор качественных признаков, каждый элемент которого - фиксированный признак характеризуемого объекта.

$W = \bigcup_{j=1}^{m} w_j$ - множество характеризуемых объектов.

Вид функции *W = f(X)* не задан. Требуется восстановить неизвестную зависимость по наблюдениям.

Каждый соответствующий признак $x_j(w_j)$ в общем случае кодируется предикатом k_i-значности [0,1,2,3]

Для нахождения значения функции *W= f(X)* системе важно обращаться к базе знаний, которая по запросу $X_j = \{x_1, x_2 \ldots\ldots x_n\}$ выдаст w_j, в случае если w_j принадлежит исследуемой предметной области, или объект (группу объектов), наиболее соответствующих данном запросу/

Для нахождения наиболее соответствующего заданному запросу $X_j = \{x_1, x_2 .... x_n\}$ объекта w_j, предлагается мследующий алгоритм:

Запрос формируется в виде комплементарной цепочки к исходному запросу.

Копируется необходимое количество цепочек запроса.

Добавляются необходимые ферменты и через несколько секунд в результате реакции ренатурации синтезируются всевозможные ответы на данный запрос.

Выбирается максимальная из двойных цепочек ДНК.

$$F(X) = \max[\mathrm{Re}(Cop_{j=1}^{m/2}(K(\overset{n}{\underset{i=1}{x_i}})) \vee_{j=1}^{m} X_j]$$

Литература:

1. Малинецкий Г.Г., Митин Н.А., Науменко С.А. Нанобиология и синергетика. Проблемы и идеи. ИПМ им. М.В.Келдыша РАН. Москва, 2005

2. Паун Г., ДНК-компьютер: Новая парадигма вычислений

Выделение информации и анализ данных для слабоформализованных предметных областей

М.А. Гузев
Институт прикладной математики ДВО РАН
guzev@iam.dvo.ru
Е.Ю. Никитина
Дальневосточный федеральный университет
nikitina@imcs.dvgu.ru

Во многих слабо-формализованных предметных областях выделение информации из накопленных данных часто проводится с целью определения влияющих факторов, а также изучения поведения моделируемых объектов предметной области во времени. Трудность решения этой проблемы требует разработки новых методов и подходов для анализа накопленных данных, результаты которых могут использоваться при построении моделей предметных областей с целью получения информации и поддержки процесса принятия решений. Для этого в анализируемой предметной области выделим классы объектов, каждый из которых обладает набором характеристик, зависящих от набора параметров и факторы, которые могут влиять на значения характеристик объектов предметной области. Основная идея предлагаемого нами подхода состоит в построении ранговых распределений для функций, задающих значения характеристик и факторов и сравнение полученных распределений. При анализе информации рассматривается проекция каждой построенной функции на некоторую ось, задающую значения параметра характеристики или фактора. Предполагается, что фактор влияет на характеристику, если распределение его ранговых значений совпадает с распределением значений характеристики. Приложение сформулированного выше подхода было выполнено для различных предметных областей: биология (выявление факторов, влияющих на структуру популяций лососевых), криминология (выявление влияющих факторов на состояние преступности, анализ изменения Уголовного Кодекса РФ), библиометрия (изучение эффективности научных публикаций). Проведенный анализ структуры распределений наборов данных позволяет выделить лакунарные области в наборах характеристик, выдать рекомендации по корректировке и получению новых данных.

Полиномиально вычислимые Σ–спецификации иерархизированных моделей реагирующих систем.

Глушкова В.Н.

ДГТУ, lar@sfedu.ru

Аннотация. Рассматривается полиномиально-реализуемый класс $\Delta_0 T$– формул, интерпретируемый на многосортных моделях с иерархическими надстройками. Структура надстройки описывается произвольной КС-грамматикой. Для теорий из $\Delta_0 T$– квазитождеств, обладающих свойством нётеровости и конфлюентности, можно построить константную модель. Предикаты и функции сигнатуры модели интерпретируются на исходном КС-списке, достраиваемом в процессе интерпретации. Этот класс формул можно использовать для моделирования систем реального времени. Приводится пример логической спецификации управляющего устройства поведением робота.

The polynomial realized $\Delta_0 T$ –formulas are considered with quantifiers acting on hierarchy lists described by CF-grammars. These formulas are interpreted on a many-sorted model with the hierarchy list superstructure. The constant model is constructed for the Noether confluent theory based on quasi-identities. The signature predicates and functions are interpreted on input CF-list extended on the interpretation process. The $\Delta_0 T$ –formulas theories might be used to model real-time systems. As example a logic specification of robot behaviour is given.

1.Введение

Метод формальной верификации model cheking, успешно используемый в технологии разработки программных систем, основан на темпоральных логиках [1]. В этом подходе реальная система моделируется разновидностью конечного автомата, состояния которого помечаются набором атомных формул исчисления высказывания. Для спецификации функционирования моделируемой системы во времени ее поведение дискретизируется и конкретные значения времени игнорируются. Поэтому средствами темпоральной логики сложно выразить свойства о количественных характеристиках системы. Использование для моделей непрерывного времени общепринятого формализма временных автоматов [2] затрудняется экспоненциальными оценками алгоритмов верификации.

В предлагаемом подходе для спецификации моделируемых систем используется класс Δ_0 – формул многосортного языка ИП, выделенных в концепции семантического программирования [3]. В этой концепции, основанной на теоретико-модельном подходе, особую роль играют многосортные модели с надстройкой из конечных списков, формируемых из элементов исходной модели. В сигнатуру модели с надстройкой явно вводится сорт “ *list*” (список) и стандартные операции и отношения для элементов этого сорта. Списки используются для ограничения области изменения квантифицируемых переменных. А именно, вводятся ограниченные кванторы вида $\forall\, x\in t$, $\exists\, x\in t$, $\forall\, x\subseteq t$, $\exists\, x\subseteq t$, где x – переменная произвольного сорта, t – терм *list*-сорта (t не содержит x); $\in$– отношение принадлежности элемента списку, $\subseteq$ – отношение включения для списков.

2. **Используемые понятия**

Определим многосортную модель M сигнатуры $\sigma = < \hat{I}, C_M, F, R>$, где $\hat{I} = I \cup \{list\}$ – множество сортов, C_M, F, R – множества констант, функций и предикатов соответственно. Каждому сигнатурному символу приписан тип: n- местная функция $f \in F$, $n \geq 0$ имеет тип $<i_1, \ldots i_n, i>$, где $i_1, \ldots i_n, i \in \hat{I}$, i – тип значения функции, остальные символы – типы аргументов. Предикат $r \in R$ имеет тип $<i_1, \ldots i_n>$. Универс M состоит из индексированного семейства носителей $U_j = C_j$, $j \in I$, где C_j – множество констант сорта j; f: $U_{i_1} \times \ldots \times U_{i_n} \to U_i$,

$r \subseteq U_{i_1} \times ... \times U_{i_n}$.

Списочная надстройка $U_{list} = D_G(C)$ модели M состоит из иерархизированных списков, структура которых задается некоторой КС-грамматикой $G = (V, P)$, где V, P – множества символов и правил соответственно; $V = N \cup T$, где N, T – множества нетерминальных и терминальных символов. Каждому символу $A \in V$ приписан атрибут сорт $\rho(A)$ из множества сортов I. Множество $D_G(C)$ определяется как наименьшее множество всех списков $<t_1, ..., t_n>$, формируемых для каждого правила $A \to X_1 ... X_n \in P$, $n \geq 1$, $X_i \in I$ следующим образом: если $X_i \in T$, то t_i – произвольная константа из $C_{\rho(X_i)}$, в противном случае t_i – произвольный список сорта $\rho(X_i)$. Список $<t_1, ..., t_n>$ имеет сорт $\rho(A)$.

Δ_0–формулы определяются с использованием всех логических связок и ограниченных кванторов. Будем использовать лишь ограниченные кванторы вида $\forall x \in t$, $\forall x \overset{*}{\in} t$, где t – переменная сорта $list$, отношение $\overset{*}{\in}$ является рефлексивным транзитивным замыканием $\in$. Пусть $\bar{x}$ обозначает индексированную последовательность переменных x (многосортных), $\overset{\circ}{\in}$ – отношение $\in$ или $\overset{*}{\in}$, $\prec$ – отношение «левее» для списочных объектов.

Определение 2. Δ_0–формула вида

$$(\forall x_1 \overset{\circ}{\in} t_1) \ldots (\forall x_m \overset{\circ}{\in} t_m)\ (y_1 \prec z_1) \ldots (y_p \prec z_p)\ \Phi(\bar{x}, \bar{t}),\ m \geq 1,\ p \geq 0 \quad (1)$$

называется $\Delta_0 T$ – формулой, если $y_j, z_j \in (\bar{x}, \bar{t})$, $1 \leq j \leq p$, и переменные префикса удовлетворяют условию: для всех $1 \leq i \leq m$, $k \leq i$ выполняется: $t_{i+1} = t_i$ или $t_{i+1} = x_k$. Если $t_{i+1} = x_k$, то для всех $l \leq i$: $x_{i+l} \neq x_l$, $x_{i+l} \neq t_l$.

Если представить все переменные префикса узлами с дугами, ведущими от t_i к x_i, то получится дерево с корнем t_1.

Для спецификации реагирующих систем, функционирующих неограниченно долго, выразительных возможностей Δ_0–формул не достаточно.

Определение 3. Формула, полученная из $\Delta_0 T$– формулы навешиванием неограниченного квантора всеобщности, называется ΠT – формулой.

Модель M строится по теории Th из квазитождеств с отрицаниями вида:

$$(\forall x_1 \overset{\circ}{\in} t_1) \ldots (\forall x_m \overset{\circ}{\in} t_m)\ (y_1 \prec z_1) \ldots (y_p \prec z_p)(\varphi(\bar{x}, \bar{t}) \to \psi(\bar{x}, \bar{t})), \quad (2)$$

Формула $\psi(\varphi)$ – это конъюнкция атомных формул (или их отрицаний) вида: r, $\tau_1 = \tau_2$, $(f = \tau)$, τ_1, τ_2 – термы сигнатуры σ.

Для теорий, обладающих свойством нётеровости и конфлюентности, можно построить индуктивно вычислимую модель из констант C_M на основе правила вывода "modus ponens". При интерпретации аксиом теории арифметические операции считаются встроенными. Отрицание для предикатов интерпретируется по принципу "замкнутого мира". Интерпретация начинается, исходя из Th_0 –"подтеории" фактов, а именно –формул вида $r(\bar{c})$ или $\neg r(\bar{c})$, $f(\bar{c}) = c_n$, $\bar{c} \in C_M^*$, $c_n \in C_M$ и списка s_0 - начального значения переменной t_1 из (2). Списку s_0 в грамматике G соответствует вывод $A \overset{*}{\Rightarrow} \alpha_0$. Некоторые предикаты сигнатуры σ согласованы с нетерминальными символами из N, например, совпадают по имени. Выделенные аксиомы Ax теории Th также согласованы с правилами грамматики G и им приписана последовательность правил $p \in P^*$. Последовательность правил p применяется к выводу в грамматике G, построенному до момента интерпретации аксиомы Ax.

Пусть правая часть (2) содержит предикаты $Q_1, ... Q_m$ (в порядке их следования) и предикат $Q \in \{Q_1, ... Q_m\}$ определен на области $D_1 \times .... \times D_n$, тогда $Q \overset{*}{\Rightarrow} имя(D_1) ... имя(D_n)$.

Если это не вызывает противоречия, можно считать, что *имя*(D_j)=D_j, $1 \le j \le n$, тогда $Q \overset{*}{\Rightarrow} D_1 \dots D_n$. Если предикат Q имеет тип $<i_1, \dots i_n>$, то $\rho(D_j) = i_j$. Пусть $\overline{d} = <d_1, \dots d_n>$ составляет набор констант, на котором интерпретируется предикат Q в аксиоме Ax. Последовательность правил p составлена так, что $Q \overset{*}{\underset{p}{\Rightarrow}} X_1 \dots X_n$, $X_j \in T$, $1 \le j \le n$, $\rho(X_j) = \rho(d_j)$, поэтому в качестве терминальных символов X_j в последнем выводе берутся константы d_j. Таким образом, список, на котором интерпретируется теория *Th,* расширяется и дерево, соответствующее исходному списку, достраивается. В результате получается дерево, отражающее последовательность шагов интерпретации предикатов и функций, в результате которой область их определения расширяется до получения "неподвижной точки" вычисления.

Ограничивая зависимость переменных, входящих в префикс $\Delta_0 T$- формул, можно выделить класс полиномиально реализуемых формул относительно «размера» списка, определяемого количеством узлов соответствующего дерева вывода в G. Верификация модели M состоит в проверке свойств, которые формализуются теорией из произвольных $\Delta_0 T$ – формул, проверяемых за полиномиальное время. Причем степень полинома зависит от вида грамматики и префикса формул. Справедлива теорема.

Теорема. Произвольная $\Delta_0 T$–формула (*1*) с префиксом длины m интерпретируется на модели M с временной сложностью $O(n^{m+1})$ относительно мощности списков.

3. Пример

Приведем логическую спецификацию из $\Delta_0 T$ и ΠT – формул поведения управляющего устройства на примере робота, обслуживающего технологический автомат [4]. Роботизированный комплекс состоит из робота, позиции загрузки заготовок (*lp*), позиции разгрузки деталей (*ulp*) и технологического автомата (*ap*). Напротив позиций находятся соответствующие датчики. Цикл функционирования робота начинается с позиции загрузки.

ЦИКЛ:

1. На позиции *lp* для загрузки детали робот выдвигает исполнительный механизм (за время 3 сек.); берет заготовку, сжимая захват (1 сек.), вдвигает исполнительный механизм (3 сек) и начинает перемещение вправо к автомату до срабатывания датчика положения *ap*.
2. На позиции *ap* с целью установки заготовки на автомате робот выдвигает исполнительный механизм, разжимает захват (1 сек.), вдвигает исполнительный механизм. После ожидания 3 мин. робот разгружает автомат, повторяя те же процедуры, что и на позиции *lp*. Далее робот двигается влево к позиции разгрузки до срабатывания датчика положения *ulp*.
3. Деталь разгружается на транспортер за время 7 сек. Робот перемещается влево до срабатывания датчика *lp* на позицию загрузки и цикл работы комплекса повторяется.

Спецификация системы состоит из нескольких уровней. Поведение робота зависит от сигналов с датчиков его положения:*lp*, *ulp*, *ap* ($\neg lp, \neg ulp, \neg ap$, отрицание означает, что сигнал не вырабатывается). Эта последовательность сигналов представляется списком $mc=<x, y, z...>$, где $x = lp \ (\neg lp)$, $y=ulp \ (\neg ulp)$, $z= ap \ (\neg ap)$ и легко описывается конечным автоматом с начальным состоянием, помеченным *x*. Этот автомат можно построить по праволинейной грамматике с правилами:

$C \to lp\ C_1 \mid \neg lp\ C_1 \mid \varepsilon$

$C_1 \to ap\ C_2 \mid \neg ap\ C_2$

$C_2 \to ulp\ C \mid \neg ulp\ C$

Обозначим через *Dc* – множество списков, составленных из цепочек символов, порожденных этой грамматикой.

Количество переходов в автомате задаёт "внешнее" дискретное время (переменная *n* в логической спецификации). На втором уровне используется *КС* – грамматика, описываю-

щая последовательность действий *(Act)* робота: *L* , *La* – загрузка робота заготовкой в позиции *lp* и *ap* соответственно; *Ul* ,*Ula* – разгрузка детали в позиции *ulp* и *ap* ; *Rmove*, *Lmove* – движение робота направо и налево; *Stopl*, *Stopa*, *Stopul* – остановка робота в соответствующей позиции; *Wait* – ожидание; *Br* – поломка устройства управления робота. Действия робота влияют на его состояния (символ *St*). В данном примере состояние характеризуется "непрерывным" временем *Ctime* и дискретным *Dtime*, задаваемым натуральным числом. Значением сорта *Ctime* являются сегменты вида $<t_1,\ t_2>$, $t_1,\ t_2$ – константы, причем < заменяется на (или [в зависимости от того, включена левая граница в сегмент времени или нет, аналогично для >.

При спецификации поведения робота абстрагируются от значения времени движения робота от одной позиции до другой. Положение робота определяется сигналами с датчиков положения, являющимися входными к устройству управления робота. Выходными являются сигналы, подаваемые на исполнительные механизмы робота, для осуществления перемещений и работы захвата.

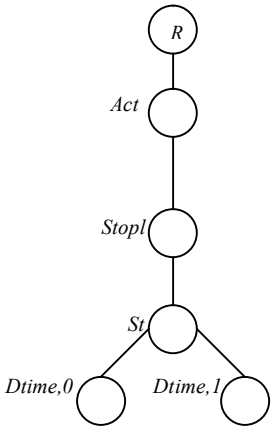
Правила грамматики *G* имеют вид:

1. $R \rightarrow \{Act\}^*$
2. $Act \rightarrow L \mid La \mid Ul \mid Ula \mid Rmove \mid Lmove \mid Stopl \mid Stopa \mid Stopul \mid Wait \mid Br$
3. $L \rightarrow St$
4. $La \rightarrow St$
5. $Ul \rightarrow St$
6. $Ula \rightarrow St$
7. $Rmove \rightarrow St$
8. $Lmove \rightarrow St$
9. $Stopl \rightarrow St$
10. $Stopa \rightarrow St$
11. $Stopul \rightarrow St$
12. $Wait \rightarrow St$
13. $Br \rightarrow St$
14. $St \rightarrow Ctime\ Dtime \mid Dtime\ Dtime$
15. $Ctime \rightarrow Dtime \mid (Dtime, Dtime) \mid [Dtime, Dtime) \mid (Dtime, Dtime] \mid [Dtime, Dtime]$

Dtime – класс лексем, значениями которых являются натуральные числа, вычисляемые во время интерпретации теории *Th.*

В теории *Th* переменные в формулах обозначаются мнемонично в соответствии с их сортом: $\rho(st) = \rho(St)$, $\rho(act) = \rho(Act)$, $\rho(n) = \rho(t) = \rho(Dtime)$, $\rho(ct) = \rho(Ctime)$. Предикаты *Lp, Ap, Ulp* определены на множестве *Dtime*,*Lp* (*n*) истинен, если робот находится на позиции загрузки, аналогично для *Ap* (*n*) – на позиции обрабатывающего автомата, *Ulp* (*n*) – на позиции разгрузки. Перечислим области определения остальных предикатов: *Stopl, Stopa, Stopul, Rmove, Lmove, Br* $\subseteq Dtime \times Dtime$; *L, La ,Ul, Ula, Wait* $\subseteq Ctime \times Dtime$. В формулах используются стандартные функции *head* $(<x_1,...x_n>) = x_1$, *tail* $(<x_1,...x_n>) = <x_2,...x_n>$ и функция *Mc*: $Dtime \rightarrow Dc$, Dc – множество списков из сигналов датчиков.

В начальный момент времени $t = 0$, $n = 1$ и на *1*-ом шаге вычисления выполняется предикат *Stopl*(*0,1*); *Mc* (*1*) =*mc*, где $mc \in Dc$. В формулах 1-11 переменные *t,n* и переменная *ct* связаны ограниченным квантором $\forall\ st \in^* act,\ \forall\ t,n,\ ct \in st$. В формулах 12-17 переменная *n* связана неограниченным квантором $\forall$; $s_0 = <<<<\ 0,1>>>>$ – начальное значение списка, на котором интерпретируется теория; составляющие его списки в порядке глубины вложенности имеют сорта $\rho(R)$, $\rho(Act)$, $\rho(Stopl)$, $\rho(0) = \rho(1) = \rho(Dtime)$. Списку s_0 соответствует дерево T_0 :



В формулах теории справа в квадратных скобках приводится последовательность правил грамматики G, достраивающих дерево T_0.

Аксиомы теории:

1. $Stopl(t,n)$, $Lp\ (n) \rightarrow L([t\ ,t+7),\ n)$, $Mc\ (n+1) = tail(Mc\ (n))$ [1; 2.1; 3; 14.1; 15.3]
2. $L(ct,n)$, $Ap(n+1) \rightarrow Rmove\ (ct\ [2],\ n+1)$, $Stopa\ (ct\ [2],\ n+1)$ [1; 2.5; 7; 14.2;1; 2.8; 10; 14.2]
3. $Stopa\ (t,n) \rightarrow Ula\ ([t\ ,t+7],\ n)$ [1; 2.4; 6; 14.1; 15.2]
4. $Ula\ (ct,n) \rightarrow Wait\ ((ct\ [2],\ ct\ [2]+180),\ n)$ [1; 2.10; 11; 14.1; 15.2]
5. $Wait\ (ct,n) \rightarrow La\ ([ct\ [2],\ ct\ [2]+3),\ n)$, $Mc\ (n+1) = tail(Mc\ (n))$ [1; 2.2; 4; 14.1; 15.2]
6. $La\ (ct,n)$, $Ulp\ (n+1) \rightarrow Lmove\ (ct\ [2],\ n+1)$, $Stopul\ (ct\ [2],\ n+1)$ [1;2.6; 8; 14.2; 1; 2.9; 11; 14.2]
7. $Stopul\ (t,n\ ) \rightarrow Ul\ ([t,t+7),\ n)$, $Mc\ (n+1) = tail(Mc\ (n))$ [1; 2.3; 5; 14.1; 15.2]
8. $Ul\ (\ ct,n)$, $Lp(n+1) \rightarrow Lmove\ (ct\ [2],\ n+1)$, $Stopl\ (ct\ [2],\ n+1)$ [1; 2.6; 8; 14.2; 1; 2.7; 9; 14.2]
9. $Ul\ \ (ct,n)$, $\neg\ Lp\ (n+1) \rightarrow Br\ (ct\ [2],\ n+1)$ [1; 2.11; 13; 14.2]
10. $La\ (\ ct,n)$, $\neg\ Ulp\ (n+1) \rightarrow Br\ (ct\ [2],\ n+1)$ [1; 2.11; 13; 14.2]
11. $L\ (\ ct,n)$, $\neg\ Ap(n+1) \rightarrow Br\ (ct\ [2],\ n+1)$ [1; 2.11; 13; 14.2]
12. $head\ (Mc\ (n)) =$ “lp” $\rightarrow Lp\ (n)$
13. $head\ (Mc\ (n)) =$ “$\neg\ lp$” $\rightarrow \neg\ Lp\ (n)$.
14. $head\ (Mc\ (n)) =$ “ap” $\rightarrow Ap(n)$
15. $head\ (Mc\ (n)) =$ “$\neg ap$” $\rightarrow \neg\ Ap(n)$
16. $head\ (Mc\ (n)) =$ “ulp” $\rightarrow Ulp\ (n)$
17. $head\ (Mc\ (n)) =$ “$\neg ulp$” $\rightarrow \neg\ Ulp\ (n)$

Теория Th обладает свойством нётеровости, т.к. изменение переменной под квантором $\forall$ ограничено k – количеством элементов в исходном списке mc и $head\ (Mc\ (k+1))$ не определено, т.к. $Mc\ (k+1) = <\ >$. Отметим, что цепочки “$\neg\ lp$” и др. в правой части аксиом 12-17 имеют сорт $string$ и ”$\neg$ ”рассматривается не как логическая операция, а как символ.

Для начального значения функции $Mc\ (1) = <lp,\ ap,\ ulp,\ lp,\ \neg\ ap,\ ulp>$ получаем множество следствий: $Stopl(0,1)$, $Lp\ (1)$, $L\ ([0,7),\ 1)$, $Mc\ (2) = <\ ap,\ ulp,\ lp,\ \neg\ ap,\ ulp>$, $Rmove\ (7,2)$, $Stopa(7,2)$, $Ula([7,14],\ 2)$, $Wait([14,\ 194],\ 2)$, $La([194,\ 197),\ 2)$, $Mc\ (3) = <ulp,\ lp,\ \neg\ ap,\ ulp>$, $Lmove(197,\ 3)$, $Stopul\ (197,\ 3)$, $Ul([197,\ 204],\ 3)$, $Mc\ (4) = <\ lp,\ \neg\ ap,\ ulp>$, $Lmove(204,\ 4)$, $Stopl(204,4)$, $L\ ([204,\ 211),\ 4)$, $Mc\ (5) = <\ \neg\ ap,\ ulp>$, $Br(211,\ 5)$. Полученное множество следствий иерархизируется в соответствии с выводом в грамматике G, полученным в результате правил, приписанных к интерпретируемым аксиомам. В соответствии с ними к дереву T_0 добавляются справа от узла, помеченного символом Act, еще 12 вершин, помеченных этим же символом, связанных ребрами с корнем. К новым вершинам подвешиваются поддеревья с корнями, помеченными символами Lp, $Rmove$, $Stopa$ и т.д. с их состояниями и константами сорта $\rho\ (Dtime)$, полученными в результате интерпретации.

На построенной модели можно проверять истинность произвольных $\Delta_0 T$– формул. Например, формализуем утверждение: если робот стоял на позиции загрузки в момент вре-

мени t на шаге n цикла его функционирования, то на шаге $n+2$ через 197 сек. он начинает разгрузку в течение 7 сек. Ниже приведенная формула верифицируется на заданном списке act сорта $\rho^{*}(Act)$:

$(\forall st \in act)\ (\forall t \in st)(\ \forall n \in st\ )\ (Stopl(t,n) \rightarrow\ Ul([t+197,\ t+204], n+2\ ).$

ЛИТЕРАТУРА

[1]Э.М. Кларк, мл. О. Грамберг, Д. Пелед. Верификация моделей программ. Изд-во Московского центра непрерывного математического образования. Москва, 2002.

[2] R. Alur, C. Courcoubetis, and D.L. Dill. Model-cheking for real-time system. In Proceedings of the 5th Annual Symposium on Logic in Computer Science. IEEE Computer Society Press, 1990, p.414-425.

[3]Goncharov S.S. and Sviridenko D.I. Theoretical aspects of Σ- programming. Mathematical Methods of Specification and Synthesis of Software Systems' 85. Proceed. of the Internat. Spring School. Springer-Verlag, April, 1985. pp. 169-179.

[4]. В.А. Горбатов, М.И. Смирнов, И.С. Хлытчиев. Логическое управление распределенными системами. М., Энергоатомиздат. 1991. 288 с.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ СУПЕРЭВМ ЭКЗАФЛОПСНОЙ ПРОИЗВОДИТЕЛЬНОСТИ

Б.М. Глинский, А.С. Родионов, М.А. Марченко

По прогнозам экспертов реальные экзафлопсные компьютеры с производительностью 1000 PFlops появятся ориентировочно в 2018 году. Предполагается гибридная архитектура их построения, содержащая от 10 до 100 млн. вычислительных ядер. Создание таких суперкомпьютеров потребует решения целого ряда проблем: технических (энергопотребление, охлаждение, надежность, память большого объема и др.); создания алгоритмического и программного обеспечения (переработка ОС, среды исполнения, компиляторов, библиотек, сред программирования и пр.).

В докладе рассматривается вариант применения системы моделирования для экзафлопсных вычислений, которая позволит прогнозировать организацию и функционирование будущих компьютеров. Предлагается использовать хорошо зарекомендовавший себя для масштабного распределённого моделирования агентно-ориентированный подход. Показана организация модели, основанная на делегировании управляющих функций в локальные области вычислений с использованием иерархической системы распределённого управления.

В качестве примера масштабной задачи, хорошо распараллеливаемой на произвольное количество вычислителей, используется задача численного статистического моделирования. Предлагаемый агентно-ориентированный подход к построению имитационной модели экзафлопсного компьютера, не требующий глобального централизованного управления, хорошо соответствует методике распределенного статистического моделирования.

Рассматривается вариант реализации мультиагентного моделирования для предотвращения сбоев и отказов вычислительных узлов. Динамическая система предотвращения сбоев состоит из агентов различного назначения и экспертной системы, содержащей свод правил поведения системы для предотвращения сбоев вычислительных узлов. Для каждой поступающей из очереди задачи инициируется система имитационного моделирования, которая прогнозирует и предотвращает сбойные ситуации.

Кодирование сообщений, порожденных неизвестным Марковским источником, при различных длительностях передаваемых сигналов.

В.К.Трофимов, Т.В. Храмова
ГОУ ВПО «Сибирский государственный университет телекоммуникаций и информатики»
e-mail: trofimov@sibsutis.ru, e-mail: tvkhramova@gmail.com

Аннотация

Предложен метод оптимального универсального кодирования сообщений, порожденных марковским источником с бинарным алфавитом для каналов с различной стоимостью букв кодового алфавита. Получены асимптотически совпадающие оценки избыточности универсального кодирования сообщений порожденных неизвестным марковским источником. Для получения верхней оценки избыточности кодирования предложен конструктивный метод такого кодирования.

Ключевые слова: энтропия, стоимость кодирования, пропускная способность, избыточность.

Coding of messages generated by unknown Markov's source with varying durations of the transmitted signals.

V.K.Trofimov, T.V.Khramova

"Siberian State University of Telecommunications and Information Sciences"

e-mail: trofimov@sibsutis.ru, e-mail: tvkhramova@gmail.com

Abstract

A method of optimal universal coding of binary Markov's sources is proposed for channels with variable durations of coding symbols. Asymptotically matching estimates of redundancy of encoding messages generated by unknown Markov's source are obtained. Constructive method of the coding is proposed for obtaining upper estimation of encoding redundancy.

Key words: entropy, the cost of encoding, capacity, redundancy

ЛОКАЛЬНАЯ ИНФРАСТРУКТУРА ПРОСТРАНСТВЕННЫХ ДАННЫХ ИРКУТСКОГО НАУЧНОГО ЦЕНТРА СО РАН

И.В. Бычков, Г.М. Ружников, А.Е. Хмельнов
bychkov@icc.ru, rugnikov@icc.ru, alex@icc.ru

Для включения больших объёмов научных пространственных данных, накопленных учреждениями СО РАН, в процесс проведения междисциплинарных научных исследований Байкальской природной территории актуально создание инфраструктуры пространственных данных (ИПД) Иркутского научного центра СО РАН с целью доступа к традиционным и электронным ресурсам, метаданным, имеющим координатную и/или географическую привязку к внешним информационным объектам, доступным по стандартным протоколам.

Структура локальной ИПД ИНЦ СО РАН включает компоненты: базовые пространственные данные, базы пространственных метаданных, национальные стандарты на пространственные данные, технологическое обеспечение. Формирование инфраструктуры осуществляется путём создания сети взаимосвязанных геопорталов, содержащих каталоги метаданных для поиска геоинформационных ресурсов.

Локальная ИПД ИНЦ СО РАН включает внутренний сегмент (ЛВС учреждения ИНЦ СО РАН) и внешний сегмент (узел ИРНОК), который обеспечивает открытость и взаимодействие с ИПД СО РАН.

Сформирована методика формирования компонент локальной ИПД ИНЦ СО РАН и проводятся работы по их созданию.

Функционально данная ИПД ориентирована на поддержку создания, структурирования, архивирования и поиска пространственных данных при проведении мониторинга и научных исследований Байкальской природной территории.

ОПЕРАЦИИ НАД МНОЖЕСТВАМИ, БУЛЕАНАМИ И ГИПЕРГРАФАМИ, СВЯЗАННЫЕ С ИДЕМПОТЕНТНОЙ МАТЕМАТИКОЙ

С.Л. Блюмин
Липецкий государственный технический университет
slb@stu.lipetsk.ru

Аннотация

Предложено формальное перенесение в алгебру множеств некоторых подходящим образом модифицированных шагов деквантования Маслова полуполя положительных чисел в идемпотентное полуполе. На этой основе определены некоторые нестандартные операции над множествами, позволяющие дополнить известные соотношения между операциями над множествами и булеанами. Определены соответствующие операции над гиперграфами.

OPERATIONS OVER SETS, BOOLEANS AND HYPERGRAPHS CONNECTED WITH IDEMPOTENT MATHEMATICS

S.L. Blyumin
Lipetsk State Technical University
slb@stu.lipetsk.ru

Abstract

The formal transference in algebra of sets is proposed for some suitably modified steps of Maslov dequantization of positive numbers semifield in idempotent semifield. Some nonstandard operations over sets are defined on this base permitting to supplement known relations between operations over sets and Booleans. Corresponding operations over hypergraphs are defined.

МОДЕЛИРОВАНИЕ СТРОЕНИЯ КАТИОНОВ ИНДОПОЛИКАРБОЦИАНИНОВЫХ КРАСИТЕЛЕЙ МЕТОДОМ ТЕОРИИ ФУНКЦИОНАЛА ПЛОТНОСТИ

Г. В. Барышников, Б. Ф. Минаев, В. А. Минаева, А. Т. Подгорная

Черкасский национальный университет им. Б. Хмельницкого, Украина
E-mail: glebchem@rambler.ru

Аннотация

Полиметиновые красители (ПК), обладающие уникальным набором спектрально-люминесцентных свойств, являются перспективными фотосенсибилизаторами для фотоэлектрических преобразователей, а также широко используются в качестве активных сред в квантовой электронике. Особенности спектрально-люминесцентных свойств многих ПК в средах различной полярности подробно изучены и описаны в литературе. Рентгеноструктурному анализу строения некоторых ПК и их кристаллической упаковки также уделяется немало внимания, однако теоретические исследования строения ПК обычно ограничены расчетом их равновесных геометрий с дальнейшим их применением для расчета оптических свойств. Нами проведено комплексное квантово-химические моделирование ряда катионов симметричных индополикарбоцианиновых красителей – классических представителей класса ПК, и сделаны выводы о их внутримолекулярной стабилизации за счет водород-водородных связывающих взаимодействий.

Поиск равновесной геометрии основного, первого синглетного и первого триплетного состояний исследуемых катионов симметричных индополикарбоцианинов (ИПКЦ) проведен на уровне теории функционала плотности с использованием гибридного обменно-корреляционного функционала B3LYP в базисе атомных орбиталей 6-31G(d) в вакуумном приближении. О нахождении истинного минимума энергии катионов ИПКЦ свидетельствует отсутствие мнимых частот в рассчитанных ИК спектрах поглощения. На основе полученных равновесных геометрий проведен топологический анализ функции распределения электронной плотности $\rho(r)$ на уровне теории Бейдера «Quantum Theory Atoms in Molecules» (QTAIM).

MODELING OF THE INDOPOLYCARBOCYANINES DYES CATIONS STRUCTURE BY THE DENSITY FUNCTIONAL THEORY METHOD

G. V. Baryshnikov, B. F. Minaev, V. A. Minaeva, A. T. Podgornaya

B. Khmelnitsky National University, Cherkasy, Ukraine
E-mail: glebchem@rambler.ru

Abstract

Polymethine dyes (PM) with a unique set of spectral-luminescent properties are promising photosensitizers for photovoltaic cells, as well as they widely used as active media in quantum electronics. Features of spectral-luminescent properties of many PCs in the different polarities solutions are studied in detail and described in the literature. The X-ray structure analysis of some PMs and of their crystal packing also obtained much attention, but theoretical studies of the structure of the PM are usually limited to the calculation of their equilibrium geometries for their further use for calculations of the optical properties. We have performed a comprehensive quantum chemical modeling of a number of symmetrical indopolycarbocyanines dyes cations – the classic PM representatives, and draw conclusions about their stabilization by intramolecular hydrogen-hydrogen bonding interactions.

The search of the equilibrium geometry of the ground state, first singlet and triplet excited states of the studied symmetrical indopolycarbocyanines dyes cations (IPCC) is performed at the density functional theory level using the hybrid B3LYP exchange-correlation functional in the 6-31G (d) basis set in the vacuum approximation. The absence of imaginary frequencies in the calculated IR absorption spectra indicates a finding of the IPCC cations true energy minimum. On the basis of equilibrium geometries the topological analysis of the electron density distribution function $\rho(r)$ by the Bader's theory «Quantum Theory Atoms in Molecules» (QTAIM) is performed.

МОДЕЛИ ПРОИЗВОДИТЕЛЬНОСТИ ПРОЦЕССОРОВ

В.Г. Митихин
Московский государственный текстильный университет им. А.Н.Косыгина
e-mail: mvgmia@mail.ru

Оценивание эффективности функционирования компьютерной системы (КС) является многофакторной и многокритериальной задачей, отражающей известное обстоятельство, что КС представляет собой пример сложной системы. Сложные системы характеризуются следующим принципиальным положением: не существует единственной модели такой системы, напротив, необходим набор моделей, каждая из которых важна для изучения определенного круга вопросов, связанных со структурой и поведением системы.

Выбор концептуальной модели функционирования процессора, который, в свою очередь, является сложной системой и, в то же время, фундаментальным элементом КС, является необходимым этапом и предопределяет дальнейшие этапы проектирования, исследования и производства. Для сложных систем процесс формирования набора базисных характеристик представляет собой слабо формализуемый этап и, как правило, опирается на опыт и интуицию исследователя. При этом часто используют экспериментально установленные соотношения между исходными характеристиками функционирования КС (например, «закон Мура»), то есть, соотношения регрессионного (статистического) типа. Но известны простые примеры того, что статистические оценки искажают информацию о связи между характеристиками системы, а также трудности содержательной интерпретации и экстраполяции результатов, полученных на основе статистики.

Одним из естественных и фундаментальных способов, снимающих отмеченные выше принципиальные трудности, является подход, основанный на анализе размерностей характеристик концептуальных моделей сложных систем.

В настоящей работе обсуждаются концептуальные модели производительности современных процессоров, позволяющие формировать модели зависимости производительности от безразмерных параметров, полученных на основе наборов исходных размерных характеристик функционирования процессора. Содержательная интерпретация полученных моделей и соответствующих безразмерных параметров содержит как уже реализованные архитектурные решения и технологии: RISC-архитектура, технология Hyper-Threading, многоядерные процессоры, решения для системной шины и кэш-памяти, так и возможные пути повышения производительности процессоров.

Частотно - временной анализ многоканальных ЭЭГ мозга, ориентированный на поиск признаков заболевания Паркинсона на ранней стадии

Обухов Ю.В., Королев М.С.
Институт радиотехники и электроники им. В.А. Котельникова РАН

Важной проблемой является диагностика болезни Паркинсона (БП) на ранней стадии. Одним из методов диагностики является электроэнцефалография (ЭЭГ). ЭЭГ пациентов с диагнозом БП было принято анализировать с помощью Фурье преобразования. Понижение частоты доминирующего ритма на поздних стадиях заболевания было исследовано. Но Фурье анализ ЭЭГ не дает динамики электрической активности. Существует предположение о том, что такая динамика может дать значимую информацию для диагностики БП на ранней стадии.

Описан новый метод частотно - временного анализа многоканальных ЭЭГ мозга, ориентированный на поиск признаков заболевания Паркинсона на ранней стадии. Суть метода заключается в исследовании частотно-временного распределения экстремумов «вспышек» вейвлет – спектрограмм.

Метод апробирован на данных ЭЭГ обследования группы пациентов на ранней стадии паркинсонизма и контрольной группы. Впервые обнаружено повышение частоты и доминирующего ритма альфа диапазона на ранней стадии паркинсонизма. Метод позволяет обнаружить межполушарную асимметрию и разброс по частоте доминирующего ритма, что соответствует односторонней форме на ранней стадии и представлениям о дезорганизации электрической активности мозга при паркинсонизме.

Работа поддержана Программой Президиума РАН «Фундаментальные науки - медицине»

АГЕНТНОЕ МОДЕЛИРОВАНИЕ СЕРЕДЕЧНО-СОСУДИСТОЙ СИСТЕМЫ ЧЕЛОВЕКА

Киселев И.Н.[1,2,*], *Семисалов Б.В.*[3], *Бибердорф Э.А.*[3], *Леонова Т.И.*[2], *Шарипов Р.Н.*[2,4], *Блохин А.М.*[3], *Колпаков Ф.А.*[1,2]

[1]Конструкторско-Технологический Институт Вычислительной Техники СО РАН, Новосибирск, Россия;

[2]ООО «Институт Системной Биологии», Новосибирск, Россия;

[3]Институт Математики им. С.Л. Соболева СО РАН, Новосибирск, Россия;

[4]Институт Цитологии и Генетики СО РАН, Новосибирск, Россия;

* axec@systemsbiology.ru

Компьютерное моделирование сердечно-сосудистой системы (ССС) человека является одной из наиболее актуальных задач современной науки. Создание ее полной модели позволит перенести эксперименты в область *in silico* и увеличить производительность научных исследований в области создания лекарств нового поколения.

В рамках интеграционного проекта СО РАН мы создали комплексную модель ССС человека на основе разработанных ранее моделей: одномерной модели артериальной системы человека, созданной нами на основе [2], модели сердечных сокращений, рассматривающей кратковременные эффекты в ССС человека (ОДУ) [3] и модели долговременных эффектов регуляции сердца и почки (АДУ) [4]. Объединение моделей разного типа стало возможным благодаря специально разработанному на основе программы Ascape [4] плагину для системы BioUML.

Созданная комплексная модель ССС человека протестирована и получены предварительные результаты для некоторых классических патологий. Модель доступна в системе BioUML по адресу www.biouml.org.

Работа проведена при поддержке междисциплинарного гранта №91 СО РАН.

Литература:

1. Lamponi D. One dimensional and multiscale models for blood flow circulation. Pour l'obtention du grade de docteur es sciences. Ecole Polytechnique Federale De Lausanne. 2004.

2. Proshin A. P. and Solodyannikov Yu.V. Mathematical modeling of blood circulation system and its practical application. Automat Remote Contr. 2006. 2: 174–188.

3. Karaaslan F. et al. Long-term mathematical model involving renal sympathetic nerve activity, arterial pressure, and sodium excretion. Ann Biomed Eng. 2005. 33(11): 1607-1630.

4. Inchiosa, M.E. and M.T. Parker Overcoming Design and Development Challenges in Agent-based Modeling Using Ascape. PNAS, 2002. 99, Suppl. 3: 7304-7308.

МОДЕЛЬ ЭКОЛОГИЧЕСКОЙ СИСТЕМЫ “ХИЩНИК-ЖЕРТВА

Т.А. Аверина
Институт Вычислительной Математики и Математической Геофизики СО РАН,
Новосибирск, Россия
ata@osmf.sscc.ru

В последние годы повышенный интерес у математиков и физиков вызывают хаотические колебания в динамических системах различной природы. Так как детерминистический хаос наблюдается в нелинейных системах обыкновенных дифференциальных уравнений (ОДУ) размерности N>2, то возрастает роль численного моделирования при их анализе. Численное моделирование колебательных систем ОДУ, находящихся под воздействием случайных флуктуаций, сводятся к статистическому моделированию траекторий решений нелинейных систем стохастических дифференциальных уравнений (СДУ). Для решения таких систем предлагается малотрудоемкий алгоритм переменного шага, в основу которого положен трехстадийный обобщенный метод Рунге-Кутты. С помощью построенного метода анализируется модель экологической системы “хищник-жертва” с двумя шумящими параметрами.

Model of ecological system a "predator-victim"

Recently mathematicians and physicists have been getting highly interested in chaotic fluctuations in dynamic systems of the various nature. As the deterministic chaos is observed in nonlinear systems of the ordinary differential equations (ODEs) of dimension N>2, the role of numerical modeling increases at their analysis. Numerical modeling of oscillatory systems of ODE which are under the influence of casual fluctuations is reduced to statistical modeling of trajectories of solutions of nonlinear systems of the stochastic differential equations (SDEs). For the solution of such systems, the algorithm of a variable step is offered which is based on the three-phase generalized method of Runge-Kutty. By means of the constructed method the model of ecological system "predator-victim" with two nuisance parameters is analyzed.

А.А.ЛЯПУНОВ И СТАНОВЛЕНИЕ ИНФОРМАТИКИ В РОССИИ

А.М.Федотов
Новосибирский государственный университет,
Институт вычислительных технологий СО РАН,
E-mail: fedotov@nsc.ru

В настоящее время, когда информационные технологи, а вместе с ним кибернетика и информатика, заняли чуть ли не основные позиции как в научных исследованиях, так и в обществе, необходимо четко понимать, какой путь развития они перед этим прошли. Отсутствие достоверной информации, позволяющей адекватно представить историю науки, весьма опасно. Понимание истории позволяет учиться на событиях прошлого, увеличивать эффективность исследований и прикладных работ. Сказанное относится к любой науке, но к информатике — особенно. Это в первую очередь связано с тем, что возникновение информатики как науки было инициировано «атомным проектом», вследствие чего многое в ее истории замалчивается, а иногда как это часто бывает в истории - фальсифицируется.

Нет сомнения в том, что историю делают личности. В этом смысле нам очень повезло: в истории нашей страны и нашей науки есть замечательная личность - Алексей Андреевич Ляпунов. Свыше сорока лет своей жизни отдал Алексей Андреевич служению науке, обогащая её первоклассными результатами в области математики, кибернетики, математической биологии и другими знаниями.

Исключительно велика роль А.А. Ляпунова в становлении отечественной кибернетики. А.А. Ляпунов занялся кибернетикой в начале 50-х годов. С именем Ляпунова неразрывно связана борьба за признание кибернетики в нашей стране, первые шаги этой новой науки и дальнейший её расцвет в 60-е годы.

Начав с задач теоретического программирования (в частности, создав операторный метод программирования, позволяющий проводить анализ функций программы, абстрагируясь от ее конкретного содержания), А.А. Ляпунов очень быстро переходит к решению кибернетических задач. К этому времени он был сложившимся учёным, известным своими работами в области дескриптивной теории множеств, математической статистики, теории стрельбы, геофизики. Эрудированность в сочетании с многосторонними научными интересами позволила ему возглавить новое направление в науке.

С годами из кибернетики выросла такая обширная область знания как информатика. Значительность вклада А.А. Ляпунова в информатику давно признана в нашем отечестве.

Современные вычислительная техника и информационные технологии предоставляют исследователю мощный аппарат для «манипулирования данными», а не информацией. Данные, переведенные в электронную форму, приобретают новое качество, обеспечивая им более широкое распространение и эффективное использование. Однако применение информационных технологий должно основываться на использовании различных моделей (феноменологических, информационных, математических и др.). Как неоднократно отмечал А.А.Ляпунов: «нет модели - нет информации». Перефразируя слова А.А.Ляпунова следует отметить, что «конечная цель всей работы, связанной с применением информационных технологий - является понимание того или иного явления, а не получение каких-либо чисел или картинок».

УДК 581.4 582.5/9 +519.712 519.717.2

В.Е. Харченко
Луганский национальный аграрный университет,
кафедра биологии растений, городок ЛНАУ,
Луганск, Украина, 91008
E-mail: viktoriakharchenko@rambler.ru

МОДУЛЬНЫЙ ПОДХОД К МОРФОГИЧЕСКОМУ АНАЛИЗУ РАСПОЛОЖЕНИЯ ЦВЕТКОВ НА РАСТЕНИЯХ

В статье речь идёт о разработке формального описания расположения цветков на растениях, путём поэтапного выделения соподчинённых гомологичных морфологических блоков (модулей). Определение состава и структуры модулей соподчинения разного уровня проводили на основании морфогенетических и филогенетических сопоставлений.

Расположение цветков на растении принадлежит к числу важнейших направлений идиоадаптации *Angiospermae* (*Magnoliophyta*) [1, 2]. Для описания расположения цветков на растениях используется целый ряд разноплановых подходов, которые не всегда охватывали одинаковый круг вопросов [3-5]. Изначально в морфологи растений сформировался физиономический подход, основанный на использовании ассоциативных характеристик [6-11]. На анализ расположения цветоносных осей растения направлен типологический подход W. Troll (1935), который основывается на использовании принципа изменяющихся пропорций и структурных отношений побегов [12-13]. Ритмологический поход основывается на последовательности формирования составляющих элементов цветоносного побега и направлен на установление границы соцветия [3, 14, 15]. В рамках ритмологического подхода был введён термин: «флоральная единица», который позволяет рассматривать соцветие как одно целое (модуль), не зависимо от того, сколько цветков в его составе один цветок или множество [3-5, 14, 15]. В связи с тем, что смена фаз развития обуславливает морфологические изменения, поэтому представляется, целесообразным проведение поэтапного анализа расположения цветков на растениях [3-5]. В связи с чем, Т.В. Кузнецовой (1991) был предложен комплементарный подход к анализу расположения цветков на растениях, предполагающий синтез всех существующих концепций, к описанию расположения цветков на осевой системе побегов на основании терминов принятых в рамках физиономического подхода [3-5].

Но приходится констатировать, что в терминологии морфологии соцветий существует множество разночтений. Ассоциативные характеристики, принятые в рамках физиономического подхода, не всегда в достаточной мере могут характеризовать структуру побега и соцветия. В частности, разнообразные формы монохазиев и дихазиев бывает трудно отличить от кисти [16-17]. Терминология типологического подхода не достаточно универсальна [3, 13]. Проблемы с терминологией в морфологии соцветий отчасти связаны с тем, что структура соцветия является выражением равновесия между эндогенными процессами роста и лимитирующим действием окружающей среды, это обуславливает её вариабельность [18]. В результате морфологически сходные структуры могут иметь разный генезис [19]. Поэтому, для морфологического анализа расположения цветков на растения требуется разработка формальной системы описания.

В связи с тем, что побеги растений обладают свойствами самовоспроизводящихся структур, то есть в процессе развития растения воспроизводятся на нём с уменьшением масштабов подобно фракталам [20]. В связи с тем, что "фундаментальным понятием кибернетики является понятие различия, то есть различимы два объекта по существу, либо один объект изменяется во времени" [21], так же как и для морфологического анализа побегов растения представляется перспективным использование кибернетического подхода для морфологического анализа растений путём выделения на нём системы соподчинённых модулей [3, 22-31]. На поэтапном воспроизведении морфологических блоков основаны работы по компьютерному моделированию растений [32]. Но эффективность модульного анализа растений зависит от того, будет ли выделенный на растении морфологический блок (модуль) состоять из гомологичных элементов [23-25]. Поэтому при выделении на растении модулей требуется проведение морфогенетических и филогенетических сопоставлений.

В ходе исследований были использованы растения *Angiospermae* из 725 видов, 81 семейства, а так же 12 мутантных линий *Arabidopsis thaliana*. При выделении на растениях

гомологичных морфологических блоков использовали морфогенетические и филогенетические сопоставления.

У растений постоянно происходит смена фаз развития, которая обуславливает поэтапные морфологические изменения [33-35]. Поэтому растение можно сравнить с динамической структурой, которая постоянно изменяется подобно фракталу [20]. В связи с этим морфологический анализ расположения цветков на растении целесообразно проводить поэтапно, путём выделения на нём соподчинённых гомологичных морфологических блоков (модулей), которые в ходе морфогенеза закономерно возобновляются с уменьшением масштаба. В связи с тем, что в ряде случаев трактовка терминов для описания расположения цветков на растениях может быть спорной, для обозначения выделенных на растении морфологических блоков (модуей) целесообразно использование минонимов, которые позволят избежать разночтений. В частности, в современной ботанической литературе, для обозначения соцветия (или цветка) используется миноним – ФЕ (флоральная единица) [3-5]. Флоральная единица может рассматриваться как элементарный модуль (*Modulus elementarius* - *ME*) [3]. При поэтапном морфологическом анализе структура и состав морфологического блока другого уровня не имеет значения. Соцветие входит в состав побега, то есть является его соподчинённым элементом. Поэтому побег может расцениваться как исходный модуль (*Modulus initialis* - *MI*). Соцветия состоят из метамеров морфологические блоки, которые повторяются при его развитии - модуль метамеров (*Modulus metamus* - *MM*). Таким образом, побег (*MI*) рассматривается как исходное множество, в составе которого имеется подмножество – флоральная единица (*ME*), состоящая из подмножеств – метамеров соцветия (*MM*). Под влиянием факторов внутреннего и внешнего происхождения состав множеств и подмножеств может варьировать ([1; +∞)). Таким образом, побег состоит из нескольких нечётких множеств, обобщёных в смысле Гогена [36-37].

На начальном этапе представляется целесообразным определение структуры и состава побегов (*MI*) главного и боковых, а затем границы флоральной единицы (*ME*). При поэтапном морфологическом анализе структура и состав морфологического блока (модуля) другого уровня не имеет значения. Поэтому на этапе анализа состава и расположения побегов (*MI*) структура и состав элементарного модуля (*ME*) безразличны.

Побеги (*MI)* и флоральные единицы (*ME*) образованы метамерами (*MM*). Способ (*modus*) воспроизведения метамеров на растении обусловлен его генезисом. При переходе к цветению побег изменяет свою специализацию, в результате происходит субституция: листья замещаются прицветниками, прилистники - прицветничками, боковые побеги – цветками (или боковыми соцветиями). Любой элемент метамера может изменяться (редуцироваться, трансформироваться и др.) автономно, при этом у одних элементов может наблюдаться пролонгация, а у других аббревиация и др. Кроме того, в составе метамеров соцветия могут возникать элементы, которые были свойственны предковым формам.

С целью формализации данных морфологического анализа генеративных побегов предлагается использование системы соподчинённых модулей. Побег растения рассматривается как исходное множество (*MI*), в его составе имеется подмножество (*ME*) – флоральная единица, которое в свою очередь образовано подмножеством метамеров (*MM*).

M – Modulus – модуль;
MI – Modulus initialis - основной модуль - побег:
cMI – communis MI - общий;
pMI – primaria MI - главный;
lMI – lateralis MI - боковой;
rMI – rosulata MI - розеточный;
1 - lMI – 1 - lateralis MI - первого порядка;
2 - lMI – 2 - lateralis MI - второго порядка;
3 - lMI – 3 - lateralis MI - третьего порядка;
ME - Modulus elementarius - модуль элементарный – флоральная единица (цветок, соцветие):
В зависимости от интенсивности и характера ветвления:
ME – simplex ME - простое;
cME – compositus ME - сложное;
aME – aggregatae ME - составное;
по положению в соцветии**:**

tME - terminalis ME – верхушечный *ME*;
lME – lateralis ME - боковой *ME*;
1 - lME - боковой 1-го порядка;
2-lME - боковой 2-го порядка;
3-lME - боковой 3-го порядка и т.д.
MM - Modulus metamus – модуль метамеров.
на основании гомологизации прицветных листьев:
BMM - Bracteae MM - с прицветниками;
bMM– bracteolae MM - с прицветничками;
BbMM - Bracteae bracteolae MM - с прицветниками и прицветничками.

Качественная оценка причин изменения электрической активности в моделях возбудимых клеток

А.Д. Васильева[1], О.Э. Соловьева[1,2]

[1]Институт иммунологии и физиологии УрО РАН (Екатеринбург)

[2]Уральский федеральный университет имени первого Президента России Б.Н.Ельцина (Екатеринбург)

e-mail: azelgadiss@yandex.ru, o.solovyova@iip.uran.ru

Аннотация:

При изменении внешних условий происходят изменения в функционировании механизмов, формирующих потенциал действия (ПД), отражающиеся в изменении формы и длительности ПД. Качественное описание влияния клеточных механизмов на генерацию трансмембранного потенциала стало возможным благодаря различным математическим методам. На основе использования этих методов нами был разработан подход, позволяющий проанализировать причины изменения генерации ПД в моделях возбудимых клеток.

Ключевые слова: изменение генерации потенциала действия, метод ведущего потенциала, метод интегралов разности токов.

Qualitative evaluation of the electrical activity changes in excitable cells models

A.D. Vasilyeva[1], O.E. Solovyova[1,2]

[1]Institute of Immunology and Physiology of the RUS (Ekaterinburg)

[2]Ural Federal University named after the first President of Russia B.N.Yeltsin (Ekaterinburg)

Abstract:

Under exchanging external conditions alters appear in the operating of mechanisms to form the action potential (AP), reflecting in changes of AP form and duration. Qualitative description of cellular mechanisms which took part in the generation of action potential became possible due to various mathematical methods. Using these methods we have developed an approach to identify the origins of action potential generation changes in models of excitable cells.

Keywords: action potential generation changes, leading potential analysis, difference – current integral method

Математические модели возбудимых клеток содержат большое количество переменных, которые сложным образом зависят друг от друга. Эта сложная нелинейная обратная связь создает трудности при оценке вклада различных клеточных механизмов в генерацию потенциала действия (ПД). При изменении нормальной деятельности клетки изменяется и ПД. Качественный анализ вклада клеточных компонентов в изменение генерации ПД необходим для понимания особенностей функционирования клетки.

Существуют математические методы, которые позволяют оценить вклад различных клеточных механизмов в генерацию ПД [1, 2]. На основе использования этих методов нами был разработан подход, позволяющий на качественном уровне проанализировать причины изменения генерации трансмембранного потенциала.

В данной работе в качестве примера было рассмотрено изменение формы и длительности ПД в нервных клетках за счет изменения максимальной проводимости калиевых каналов в рамках классической модели Ходжкина – Хаксли [3].

Уравнение, описывающее скорость развития ПД, в данной модели записывается следующим образом [3]:

$$\frac{dV_m}{dt} = -\frac{1}{C_m}\left(I_{stim} + g_K \cdot n^4 \cdot (V_m - E_K) + g_{Na} \cdot m^3 \cdot h \cdot (V_m - E_{Na}) + g_L \cdot (V_m - E_L)\right), \quad (1)$$

где V_m – мембранный потенциал; C_m – емкость мембраны; I_{stim} – стимулирующий ток; g_K – максимальная проводимость калиевых каналов; g_{Na} – максимальная проводимость натриевых каналов; g_L – проводимость мембраны для тока утечки; n^4 – вероятность активации калиевых каналов; m^3 – вероятность активации натриевых каналов; h – вероятность инактивации натриевых каналов; E_K – реверсный калиевый потенциал; E_{Na} – реверсный натриевый потенциал; E_L – реверсный потенциал тока утечки.

В классической модели максимальная проводимость калиевых каналов (g_K) равна 36 мСм/см^2. В работе моделировалась генерация ПД при следующих значениях g_K: 36 мСм/см^2, 52 мСм/см^2, 62 мСм/см^2 и 72 мСм/см^2 (рис.1). Аналогичное изменение развития ПД во времени было получено при уменьшении E_K и при увеличении g_K совместно с уменьшением E_K (результаты не показаны).

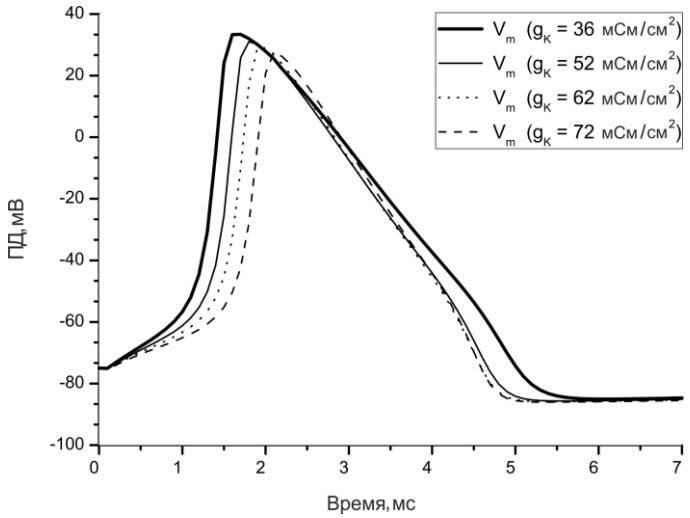


Рис.1. Временной ход ПД при разных значениях максимальной проводимости калиевых каналов

Взяв за основу метод «ведущего потенциала» японских авторов [1], при помощи которого можно оценить количественный вклад каналов и обменников в скорость развития потенциала действия, мы усовершенствовали его и оценили, какой вклад в изменение скорости развития потенциала действия вносит изменение проводимости каналов, которое

обусловлено, в нашем случае, увеличением максимальной проводимости калиевых каналов. Второй метод интегралов разности токов [2], используемый в работе, позволяет рассчитать количественный вклад для изменения тока каждого канала в собственно изменение потенциала действия. Таким образом, усовершенствованный первый метод дает оценку механизмов изменения ПД на уровне «проводимостей», второй метод – на «токовом» уровне.

При увеличении максимальной проводимости калиевых каналов от 36 мСм/см^2 до 72 мСм/см^2 были получены следующие результаты:

- фаза деполяризации ПД замедляется (t~0-2 мс). Согласно первому методу, в эту фазу основную роль в изменении скорости развития ПД играет изменение проводимости натриевых каналов (изменение вероятности активации натриевых каналов), в основном за счет более медленного нарастания вероятности их активации. Согласно второму методу, за изменение величины ПД отвечает изменение калиевого тока (фаза ранней деполяризации) и натриевого тока (фаза поздней деполяризации). Нам показался интересным тот факт, что при изменении параметра калиевых каналов в медленную фазу деполяризации за изменение скорости развития ПД отвечает изменение проводимости натриевых, а не калиевых каналов, что, вероятно, связано с большей чувствительностью натриевых каналов к мембранному потенциалу в этом диапазоне значений потенциала.

- фаза реполяризации ускоряется (t~2.0-5.5 мс). Согласно обоим методам в начальную фазу реполяризации этот эффект связан с изменениями функционирования калиевых каналов: первый метод указывает на основную роль изменения временного хода активации калиевых каналов в увеличении скорости реполяризации ПД, а второй метод – на основной вклад изменения самого калиевого тока в изменение временного хода ПД. В финальную фазу быстрой реполяризации согласно первому методу начинает превалировать вклад изменения проводимости натриевых каналов (а именно, ускорение инактивации натриевых) по сравнению с вкладом изменения проводимости калиевых каналов в увеличение скорости реполяризации ПД. Согласно второму методу, за изменение величины ПД отвечает изменение калиевого тока.

Таким образом, объединив два метода, мы смогли представить более полную картину внутренних механизмов, которые вносят вклад в изменение ПД, чем при использовании каждого метода в отдельности. Однако ни первый, ни второй методы, взятые по отдельности, не дают ответ на вопрос, поставленный в ходе работы – изменение какого конкретно параметра канала является пусковым механизмом, который приводит к изменению ПД. При сочетании двух методов можно оценить последствия изменения параметров модели и сравнить вклад различных механизмов в наблюдаемые изменения временного хода развития ПД. В то же время, сочетание двух методов может позволить высказать разумные предположения о возможных причинах возникновения отклонений во временном ходе развитии ПД. По результатам расчетов, проведенных с использованием первого метода, мы получили, что в фазу ранней деполяризации (до t~1.4 мс) вероятность активации калиевых каналов практически не меняется, а из расчетов, проведенных с использованием метода разности токов – что изменение калиевого тока вносит основной вклад в изменение ПД. Зная уравнение для калиевого тока из модели Ходжкина – Хаксли (см. второе слагаемое из

уравнения (1)), можно сделать вывод, что замедленная деполяризация ПД обусловлена либо изменением максимальной проводимости калиевого канала, либо изменением реверсного потенциала калиевого канала.

Работа поддержана грантом 09-М-14-2001 Президиума УрО РАН, грантом РФФИ 11-04-00785-а.

Список используемой литературы:

[1] *Cha C.Y., Himeno Y., Shimayoshi T, Amano A., Noma A.* A novel method to quantify contribution of channels and transporters to membrane potential dynamics. Biophysical Journal. 2009. 97(12), 3086-94

[2] *Solovyova O., Vikulova N., Markhasin V.S., Kohl P.* A novel method for quantifying the contribution of different intracellular mechanisms to mechanically induced changes in action potential characteristics. Lecture Notes in Computer Science. 2003. Vol. 2674 - P. 7-16.

[3] *Hodgkin A.L. and Huxley A.F.* A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve. *The Journal of Physiology*. 1952. 117, 500-544

Нижняя оценка сложности реализации в классе π-схем q-ичного счётчика кратности q [1]

К. Л. Рычков
Институт математики им. С.Л. Соболева СО РАН
rychkov@math.nsc.ru

Под q-ичной параллельно-последовательной контактной схемой (π-схемой) понимается обычная (двоичная) π-схема, контактам которой приписаны символы x_i^δ, $i = 1, ..., n$; $\delta = 0, 1, ..., q-1$. Но при этом символ x_i^δ мы понимаем не как булеву переменную или её отрицание, а как функцию одной переменной x_i, определённую на множестве $B_q = \{0, 1, ..., q-1\}$ и принимающую значения из множества $\{0, 1\}$. Значение функции x_i^δ равно 1, если $x_i = \delta$, и равно 0, если $x_i \neq \delta$.

Функция $f : B_q^n \to \{0, 1\}$ проводимости q-ичной π-схемы определяется по аналогии с двоичным случаем: по определению q-ичная π-схема реализует функцию $f(x_1, ..., x_n) = \bigvee\limits_C K_C$, где дизъюнкция берётся по всем простым (без самопересечений) цепям, соединяющим полюсы схемы, а K_C это конъюнкция всех функций $x_{i_1}^{\delta_1}, ..., x_{i_k}^{\delta_k}$, приписанных контактам цепи C. Как и в двоичном случае мы говорим, что контакт, помеченный x_i^δ, замкнут на наборе $(\alpha_1, ..., \alpha_n) \in B_q^n$, если $\alpha_i = \delta$, и разомкнут в противном случае.

Сложностью $L(S)$ q-ичной π-схемы S называется число контактов в S. Сложностью $L_\pi(f)$ функции $f : B_q^n \to \{0, 1\}$ в классе π-схем называется $\min\limits_S L(S)$, где минимум берётся по всем q-ичным π-схемам, реализующим f.

Заметим, что при $q \geq 3$ в отличие от двоичного случая сложность функции $f : B_q^n \to \{0, 1\}$ может отличаться от сложности её отрицания $\overline{f}$. Так сложность функции x_i^δ равна 1, а сложность её отрицания $\overline{x_i^\delta} = x_i^0 \vee, ..., \vee x_i^{\delta-1} \vee x_i^{\delta+1} \vee, ..., \vee x_i^{q-1}$ равна $q-1$. При этом, очевидно, имеет место неравенство

$$L_\pi(f) \geq \frac{1}{q-1} L_\pi(\overline{f}).$$

Действительно, если S это минимальная q-ичная π-схема, реализующая функцию f, и F — соответствующая S формула, то "навесив"на F отрицание, "опустив"его по законам де Моргана до переменных и заменив все $\overline{x_i^\delta}$ на $x_i^0 \vee, ..., \vee x_i^{\delta-1} \vee x_i^{\delta+1} \vee, ..., \vee x_i^{q-1}$, получим формулу а значит и некоторую π-схему S' для $\overline{f}$. При этом, очевидно, выполнены соотношения

$$L_\pi(f) = L(S) = \frac{1}{q-1} L(S') \geq \frac{1}{q-1} L_\pi(\overline{f}).$$

На множестве B_q^n определим следующую функцию:

$$\varphi_q(x_1, ..., x_n) = \begin{cases} 1, & \text{если } x_1 + ... + x_n = 0 \pmod q; \\ 0 & \text{в противном случае.} \end{cases}$$

[1] Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 09-01-00528а).

В [1] было установлено, что при $q \geq 2,\ n \geq 2$ справедливо неравенство

$$L_\pi(\overline{\varphi_q}(x_1, ..., x_n)) \geq \max\ \{\frac{qn \log q}{2},\ (q-1)n^2\}.$$

Это неравенство в силу сделанного замечания влечёт следующую нижнюю оценку сложности функции $\varphi_q(x_1, ..., x_n)$:

$$L_\pi(\varphi_q(x_1, ..., x_n)) \geq \max\ \{\frac{qn \log q}{2(q-1)},\ \ n^2\}.$$

Результатом настоящей работы является теорема, улучшающая эту нижнюю оценку сложности.

ТЕОРЕМА. Для любых натуральных $q, n,\ q \geq 2, n \geq 3$, справедливы следующие неравенства:

$$L_\pi(\varphi_q(x_1, ..., x_n)) \geq \max\ \{2qn - q,\ \ \frac{q^n n^2}{q^n - (q-1)^n - (-1)^n(q-1)}\},\ \text{при } n = 1 \pmod q;$$

$$L_\pi(\varphi_q(x_1, ..., x_n)) \geq \max\ \ \{2qn - q,\ \frac{q^n n^2}{q^n - (q-1)^n + (-1)^n}\},\ \text{при } n \neq 1 \pmod q.$$

Литература

1. Рычков К. Л. О сложности обобщённых контактных схем// Дискрет. анализ и исслед. операции. Сер. 1. 2009. Т. 16, № 5. С. 78-87.

Графическое отображение некоторых понятий суперкомпиляции

(трассировщик суперкомпилятора SCP4)

(Тезисы доклада)

Андрей П. Немытых

Институт программных систем РАН
nemytykh@math.botik.ru

Суперкомпиляция есть метод специализации программ. Суперкомпилятор SCP4 был разработан и реализован в конце 1990-х – начале 2000-х гг. для функционального языка программирования рефал-5. Для данной программы `p` и её частично определенной входной точки `f(`x_0`,y)` суперкомпилятор разворачивает потенциально бесконечное дерево всех возможных вычислений `p`, начинающихся в `f(`x_0`,y)`, анализирует и оптимизирует это дерево и сворачивает его в граф, который представляет результат суперкомпиляции.

В рамках поддержки реализации и развития суперкомпилятора SCP4, совершенствования технологии суперкомпиляции автором разработаны принципы построения графического трассировщика суперкомпилятора SCP4 и реализована его первая версия. В данной статье описываются принципы реализации этого трассировщика и демонстрируется часть его инструментария.

Суперкомпилятор SCP4 реализован на функциональном аппликативном языке программирования рефал-5 и является оптимизатором программ, написанных на этом же языке, основанным на технологии суперкомпиляции, основополагающие идеи которой были заложены В. Ф. Турчиным. Суперкомпиляция выполняет метаинтерпретацию исходных программ на их частично заданных входных данных. Стартовая версия экспериментального суперкомпилятора SCP4 была разработана в конце 1990 годов автором данной статьи – под научным руководством В. Ф. Турчина. Руководство пользователя суперкомпилятора SCP4 было написано в 1999 г. А. В. Корлюковым. Первые публикации по SCP4 появились в начале 2000-х. В 2007 году А. П. Немытых опубликовал монографию «Суперкомпилятор SCP4: Общая структура». Название – SCP4, отражающее историю суперкомпиляции, было предложено В. Ф. Турчиным. Одна из старых версий суперкомпилятора SCP4 может быть вызвана в режиме on-line с электронной страницы http://www.botik.ru/pub/local/scp/refal5/ . Позднее посредством SCP4 был произведен ряд удачных экспериментов по верификации некоторого класса недетерминированных парметризованных коммуникационных протоколов с условиями глобальной корректности. На данный момент суперкомпилятор SCP4 является единственным экспериментальным суперкомпиляторм для реального языка программирования рефал-5 (без каких-либо ограничений на язык); в то же время существуют несколько простых суперкомпилято-

ров для модельных языков программирования или простейших фрагментов реальных языков программирования.

Процесс суперкомпиляции автоматически развивает и анализирует потенциально бесконечное дерево всех возможных вычислений исходной программы `P` в контексте задачи `Z` на специализацию программы `P`, поставленной суперкомпилятору SCP4. При этом, как исходная задача `Z`, так и промежуточные задачи, возникающие перед суперкомпилятором, могут автоматически разлагаться на композицию подзадач. То есть, суперкомпилятор работает со сложной структурой – лесом деревьев, помеченных информацией о параметризованных состояниях в точках ветвления этих деревьев и условиями выбора ветвей при конкретном вычислении исследуемой программы `P`. Задача суперкомпилятора эффективно свернуть этот лес потенциально бесконечных деревьев в конечный граф, который и будет представлять результат суперкомпиляции. Линейно-текстовое представление такой сложной структуры данных крайне трудоемко для анализа процесса суперкомпиляции и, как правило, сводится к ручному рисованию этой структуры. Разработка графического отображения трассы вычисления суперкомпилятора является принципиальным шагом как для развития и совершенствования самой технологии суперкомпиляции, как таковой, так и для расширения областей применения суперкомпиляции. Анализ графического представления трассы суперкомпилятора SCP4 позволяет повысить эффективность: процесса отладки суперкомпилятора SCP4 при его дальнейшем развитии; улучшения существующих алгоритмов суперкомпиляции; генерации новых идей, которые должны позволить существенно расширить границы применения суперкомпилятора SCP4; процесса обучения студентов основам технологии суперкомпиляции.

Качество остаточных программ[1] существенно зависит от стратегий, выбранных для конкретного процесса суперкомпиляции. Выбор наиболее сильных с точки зрения преобразований стратегий может привести: к большому времени исполнения этого процесса для некоторых исходных (входных) программ; более того, к незавершаемости процесса суперкомпиляции. В суперкомпиляторе SCP4 реализованы несколько уровней подробности трассировки, предоставляющие возможности изучения алгоритмов суперкомпиляции с целью их улучшения или другими целями (например, для извлечения формального доказательства в случае удачной попытки верификации или для изучения причин: почему попытка верификации оказалось неуспешной).

Ранее графическая поддержка отображения трассы суперкомпиляции была реализована И.Г. Ключниковым в простейшем модельном суперкомпиляторе SCPS для простейшего подмножества языка Scala.

1 Принципы реализации графического трассировщика

Говоря точнее, описываемый нами трассировщик является графически-текстовым. Введено понятие шага графического трассировщика, которое со-

[1] Программ поученных в результате суперкомпиляции.

ответствует логическому шагу изменения ориентированного графа, анализируемого и преобразуемого суперкомпилятором. На каждом шагу трассировщика генерируется графическое описание состояния суперкомпилятора в языке описания графов Dot, разработанном компанией AT&T. Язык Dot является абстрактным языком, независимым от конкретного графического формата. Преобразователь состояний суперкомпилятора SCP4 в язык Dot реализован, как и сам SCP4, на языке рефал-5. Далее используется свободно распространяемая система Graphviz, которая позволяет получить из данного Dot-описания конкретную картинку, поддерживает все широко-распространенные графические форматы. Система Graphviz вызывается непосредственно из суперкомпилятора. Масштабирование и навигация по графам реализована с использованием свободно распространяемой системой ZGRViewer. Графические состояния текущего и предыдущего шага трассировщика отображаются на экране посредством браузера[2] в двух соседних окнах, что позволяет пользователю проанализировать результат исполнения текущего шага суперкомпилятора. Для отображения свойств узлов и ребер графа состояний суперкомпилятора используются как текстовая информация, так и цветовая. Управление трассировщиком, в частности, осуществляется посредством текстовой консоли, которая также, при необходимости, дает возможность более детального изучения состояния суперкомпилятора SCP4 в терминах внутренних структур (кодировок) суперкомпилятора. Графическое Dot-описание состояния текущего шага трассировщика строится на основании уже имеющегося Dot-описания состояния предыдущего шага трассировщика, – заменой описания изменившихся частей рисунка: генерация заново всего описания графического состояния не производится.

В полной версии данной статьи показано отображение ряда понятий суперкомпиляции средствами графического трассировщика суперкомпилятора SCP4. В данных тезисах ограничимся графическим отображением результата прогонки и дадим краткие пояснения к нему.

2 Пример (результат прогонки)

Пусть дан некоторый язык программирования L. Любой интерпретатор `Int` языка L обычно имеет следующую структуру.

```
Int(p,d) {
  инициализация;
  while (вычисления не завершены)
            { STEP; }
  return (результирующее состояние);
}
```

Здесь `STEP` суть минимальный логически замкнутый шаг, включающий в себя все механизмы необходимые для интерпретации любой L-программы `p`

[2] Браузер может быть выбран пользователем.

на любых входных данных `d`. Скорее невозможно, чем трудно, реализовать какой-нибудь интерпретатор другим способом[3]. `STEP` изменяет глобальное состояние интерпретатора; спецификация этого состояния зависит от L. Если физическое время вычисления `STEP` равномерно ограничено по размеру его входных данных, тогда число шагов, необходимое для вычисления `p` на `d`, можно рассматривать как логическое время вычислений, которое *адекватно* отражает соответствующее физическое время вычислений.

Вернёмся к языку рефал. Одним из базовых механизмов суперкомпиляции является прогонка данной параметризованной конфигурации (параметризованного состояния стека вызовов функций специализируемой рефал программы). Алгоритм прогонки является метарасширением интерпретации одного шага рефал машины, преобразующего данное конкретное состояние стека вызовов (в терминах рефала – конкретного поля зрения) в следующее конкретное состояние стека вызовов. Результатом прогонки является конечное дерево возможных вычислений данного на входе прогонки параметризованного (т.е. частично неизвестного) стека вызовов функций. На рисунке 1 показан пример снимка экрана с результатом первого вызова алгоритма прогонки. Серые кластеры показывают структуры параметризованных стеков в листьях грозди прогонки. Зеленым цветом отмечены полностью вычисленные параметризованные части стеков (т.е. в них нет вызовов функций), соответствующих пути вычисления из корня грозди в данный лист грозди. Вершины грозди прогонки, обозначенные точками, не учитываются суперкомпилятором SCP4 при анализе «истории вычислений» – развертки метадерева возможных вычислений поставленной задачи на суперкомпиляцию. По этой причине параметризованное состояние вычислительной системы не показано в вершинах-точках. Остальные вершины называются опорными и могут в процессе анализа стать входными точками функций (подпрограмм) результата суперкомпиляции. Стартовая параметризованная конфигурация, задача на суперкомпиляцию, являющаяся корнем метадерева возможных вычислений, выделена указанием на неё стрелки-ребра, в исходящей вершине которого нет никакой конфигурации. Ребра грозди прогонки помечены предикатами выбора/расщепления конкретных значений параметров (выбора конкретного рефал шага в момент исполнения преобразуемой программы). Отметим, что, в отличие от интерпретации конкретного (не параметризованного) стека вызова функций, разложение синтаксической композиции вызовов функций в стек, линейную последовательность их исполнения, может быть нетривиальным и зависит от стратегии суперкомпиляции. Примерами таких стратегий могут быть «ленивое» (call by need) и аппликативное (call be value) вычисление в процессе суперкомпиляции. Интерпретатор рефала-5 является консольным приложением. В нижнем правом углу экрана находится окно консоли интерпретатора рефала-5, исполняющего суперкомпилятор SCP4, трасса которого (результат первого вызова прогонки) отображена в правом верхнем окне – это результат первого шага

[3] Это следует из того, что операционная семантика языка L описывается текстом конечной длины.

графического трассировщика суперкомпилятора SCP4. Для получения следующей пары картинок необходимо отдать команду в командной строке на консоли рефал-5 интерпретатора.

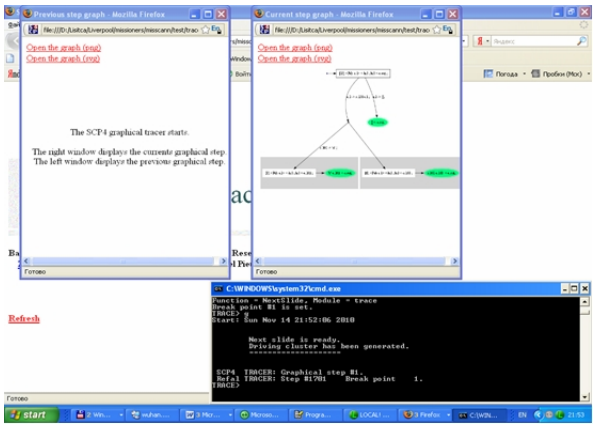


Рис. 1. Результат первого вызова алгоритма прогонки

Рисунок 2 показывает результаты первого и второго шагов трассировщика – левый и правый экраны соответственно. Результатом второго графического шага суперкомпилятора SCP4 является развитие дерева возможных вычислений задачи на суперкомпиляцию (пары – преобразуемой программы и ее параметризованной входной точки): результат второго вызова алгоритма прогонки был подклеен вместо листа левой ветви грозди прогонки, построенной предыдущим вызовом прогонки.

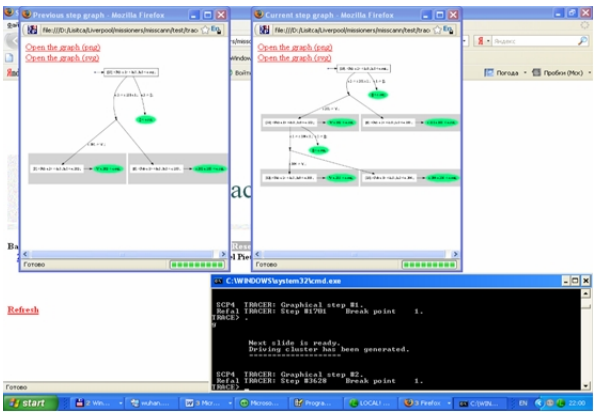


Рис. 2. Результаты первого и второго шагов трассировщика

Эти картинки дают хороший повод для пояснения – почему алгоритмы обобщения не были вызваны суперкомпилятором SCP4 между этими двумя вызовами прогонки. Конфигурация, которой помечен самый левый лист грозди прогонки, построенной первым вызовом прогонки, может быть транзитной и это свойство данной конфигурации есть предмет анализа второго вызова алгоритма прогонки. Другими словами, второй узел прогонки, возможно, построит единственную ветвь, исходящую из рассматриваемого листа. В этом случае представляется шанс для оптимизации исходной программы: два шага рефал машины могут быть соединены в один шаг. С другой стороны, если из узла грозди прогонки исходят несколько ветвей, тогда некоторые из них могут быть обрезаны последующими вызовами алгоритма прогонки. Таким образом, в данной точке алгоритма суперкомпиляции может работать стратегия, которая должна решать: сколько последовательных вызовов прогонки должны быть выполнены между двумя последовательными вызовами алгоритма свёртки.

В левых верхних углах окон шагов графического трассировщика красным шрифтом обозначены ссылки, позволяющие увеличить представленный граф (в png или svg формате). На рисунке 3 показан результат такого увеличения.

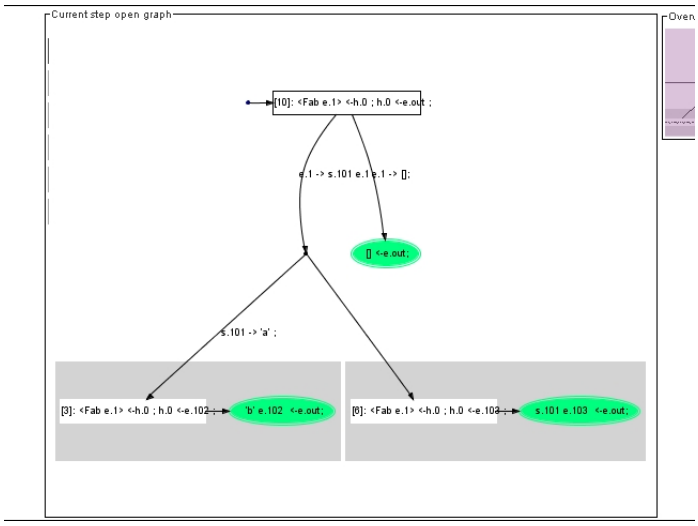


Рис. 3. Результат первого вызова прогонки (крупный масштаб)

Специализированные программные системы для проведения аналитических выкладок как инструмент решения вероятностных задач со случайным разбиением интервала[1]

А.Л.Резник[2], В.М.Ефимов[2], А.А.Соловьев[2], А.В.Торгов[2]

[2]Институт автоматики и электрометрии, Сибирское Отделение Российской Академии Наук, 63090 г. Новосибирск, проспект академика Коптюга, 1, Россия,
E-mail: reznik@iae.nsk.su, torgov@iae.nsk.su, solowey@rambler.ru

Многочисленные научные и технические приложения, относящиеся к обработке случайных дискретных изображений, теории массового обслуживания, технической диагностике и другим дисциплинам [1-4], приводят к следующей вероятностной задаче, связанной со случайным разбиением интервала:

"Пусть n точек случайно брошены на интервал (0,1), т.е. имеется n независимых испытаний случайной величины, равномерно распределенной на интервале (0,1). Необходимо найти вероятность $P_{n,k}(\varepsilon)$ того, что внутри интервала (0,1) не содержится ни одного подынтервала Ω_ε длиной ε, содержащего более k точек."

Несмотря на кажущуюся простоту поставленной задачи, ее аналитическое решение известно (Parzen, E. (1960) Modern Probability Theory and Its Applications, John Wiley and Sons, Inc., New York-London) [5] лишь для случая, когда k=1:

$$P_{n,1}(\varepsilon) = (1-(n-1)\varepsilon)^n, \quad (0 \le \varepsilon \le 1/(n-1)). \tag{1}$$

Наиболее доступный способ получения этой формулы состоит в представлении искомой вероятности в виде повторного интеграла

$$P_{n,1}(\varepsilon) = n! \int_{(n-1)\varepsilon}^{1} dx_n \left\{ \int_{(n-2)\varepsilon}^{x_n-\varepsilon} dx_{n-1} \ldots \left\{ \int_{2\varepsilon}^{x_4-\varepsilon} dx_3 \left\{ \int_{\varepsilon}^{x_3-\varepsilon} dx_2 \left\{ \int_{0}^{x_2-\varepsilon} dx_1 \right\} \right\} \right\} \right\} \tag{2}$$

с дальнейшим последовательным интегрированием по всем переменным x_1, x_2,..., x_n (множитель $n!$ перед интегралом связан с количеством перестановок из n элементов, поскольку запись (2) предполагает, что точки x_1, x_2,..., x_n упорядочены по возрастанию).

К сожалению, для значений k>1 вероятность $P_{n,k}(\varepsilon)$ не удается представить в такой же компактной форме, как при k=1. Более того, в общем случае вероятность $P_{n,k}(\varepsilon)$, в отличие от случая k=1, описывающегося формулой (1), представляет собой не единое полиномиальное выражение, а набор полиномов от параметра ε, имеющих степень n и непрерывно «сшитых» в нескольких местах.

Из-за сложностей, возникающих при попытке найти общий вид соотношений для вероятности $P_{n,k}(\varepsilon)$, нами были предприняты усилия по поэтапному решению

[1] Настоящая работа частично поддержана Российским фондом фундаментальных исследований (проект № 10-01-00458), Президиумом РАН (проект № 228/2009) и Президиумом СО РАН (интеграционный проект № 71/2009).

сформулированной задачи. В первую очередь эти усилия были направлены на то, чтобы попытаться найти частные формулы для вероятности $P_{n,k}(\varepsilon)$ при фиксированных значениях n и k, которые в дальнейшем могли бы помочь в отыскании общих закономерностей их образования и тем самым «навести» исследователя на правильный ответ.

Вообще говоря, общее решение задачи может быть представлено в виде многомерного интеграла

$$P_{n,k}(\varepsilon) = n! \int\limits_{D_{n,k}(\varepsilon)} \dots \int dx_1 \dots dx_n, \tag{3}$$

где область интегрирования $D_{n,k}(\varepsilon) \subset R^n$ описывается системой линейных неравенств

$$\begin{cases} 0 < x_1 < x_2 < \dots < x_{n-1} < x_n < 1, \\ x_{k+1} - x_1 > \varepsilon, \\ x_{k+2} - x_2 > \varepsilon, \\ \vdots \\ x_n - x_{n-k} > \varepsilon. \end{cases} \tag{4}$$

Суть первого из предложенных нами подходов, который мы назвали методом прямого аналитического интегрирования, такова. Интеграл (3) по области (4) записывается в эквивалентной форме:

$$P_{n,k}(\varepsilon) = n! \int\limits_{-\infty}^{+\infty} \dots \int 1[x_1]1[x_2] \dots 1[x_n - x_{n-1}]1[1 - x_n]1[x_{k+1} - x_1 - \varepsilon] \times$$

$$\times 1[x_{k+2} - x_2 - \varepsilon] \times \dots \times 1[x_n - x_{n-k} - \varepsilon] dx_1 \dots dx_n, \tag{5}$$

т.е. интегрирование уже ведется по всему пространству, а индикаторная функция области интегрирования $D_{n,k}(\varepsilon)$ трансформируется в произведение единичных функций $1[z] = \begin{cases} 0, & z \le 0, \\ 1, & z > 0, \end{cases}$ входящих в подынтегральное выражение многократного интеграла (5).

Затем n-мерный интеграл (5) с помощью последовательного применения соотношения

$$\left(\prod_{j=1}^{l} 1[x_r - \alpha_j] \right) \left(\prod_{i=1}^{m} 1[\beta_i - x_r] \right) =$$

$$= \sum_{j=1}^{l} \sum_{i=1}^{m} 1[x_r - \alpha_j]1[\beta_i - x_r] \left\{ 1[\beta_i - \alpha_j] \left(\prod_{\substack{q=1 \\ q \ne j}}^{l} 1[\alpha_j - \alpha_q] \right) \left(\prod_{\substack{s=1 \\ s \ne i}}^{m} 1[\beta_s - \beta_i] \right) \right\} \tag{6}$$

сводится к набору повторных интегралов с уже расставленными пределами интегрирования и с указанием на границы изменения параметра ε, к которым относится каждый из них. На заключительном этапе нужно лишь провести собственно интегрирование и объединить результаты.

Предложенный подход позволяет конструктивно вычислять частные формулы $P_{n,k}(\varepsilon)$ при конкретных значениях n и k, но вычислительная сложность алгоритма такова, что для значений $n>4$ практически невозможно провести расчеты «вручную». Поэтому изложенный метод был нами абсолютно формализован и на его основе в полном соответствии с описанной выше схемой создан пакет программ для *компьютерного* проведения всех необходимых аналитических преобразований, включая расчет пределов интегрирования, проверку промежуточных систем неравенств на совместность, упорядочение по мультииндексам и непосредственное интегрирование. Используя их, были вычислены формулы $P_{n,k}(\varepsilon)$ вплоть до значений n =14 (некоторые примеры таких расчетов приведены в Таблице 1).

Табл.1

n	k	Область изменения ε	$P_{n,k}(\varepsilon)$
3	2	(0,1)	$1-3\varepsilon^2+2\varepsilon^3$
4	2	(0,1/2)	$1-12\varepsilon^2+24\varepsilon^3-14\ \varepsilon^4$
		(1/2,1)	$2-8\varepsilon+12\varepsilon^2-8\varepsilon^3+2\varepsilon^4$
	3	(0,1)	$1-4\varepsilon^3+3\varepsilon^4$
….			
14	2	(0,1/8)	$1 - 1092\varepsilon^2 + 16744\varepsilon^3 + 156156\varepsilon^4 - 6498492\varepsilon^5 + 63159096\varepsilon^6$ $- 56566224\varepsilon^7 + 295079404\varepsilon^8 + 338060148\varepsilon^9 + 1741401056\varepsilon^{10}$ $- 79779000\varepsilon^{11} - 522127532\varepsilon^{12} - 701324068\varepsilon^{13} + 1460968152\varepsilon^{14}$
		(1/8,1/7)	$65 - 6552\varepsilon + 309036\varepsilon^2 - 8975512\varepsilon^3 + 178510332\varepsilon^4 - 2564958396\varepsilon^5$ $+ 27418999608\varepsilon^6 - 221377405392\varepsilon^7 + 1356313562604\varepsilon^8$ $- 6275548683404\varepsilon^9 + 21590662949856\varepsilon^{10} - 53575501829304\varepsilon^{11}$ $+ 90670532458324\varepsilon^{12} - 93708297788196\varepsilon^{13} + 44634761108184\varepsilon^{14}$
		(1/7,1/6)	$429 - 36036\varepsilon + 1405404\varepsilon^2 - 33729696\varepsilon^3 + 556539984\varepsilon^4 - 6678479808\varepsilon^5$ $+ 60106318272\varepsilon^6 - 412157611008\varepsilon^7 + 2163827457792\varepsilon^8$ $- 8655309831168\varepsilon^9 + 25965929493504\varepsilon^{10} - 56652937076736\varepsilon^{11}$ $- 84979405615104\varepsilon^{12} - 78442528260096\varepsilon^{13} + 36182263971840\varepsilon^{14}$

Пример работы программы прямого аналитического интегрирования.

Для определения асимптотического поведения вероятности $P_{n,k}(\varepsilon)$ на первом участке (при $n\varepsilon$ <<1) нами был разработан еще один метод, в основу которого положено простое соображение, что для восстановления полинома степени n достаточно знать его значение и значения всех его производных до степени n в какой-либо одной точке (например, в нуле), т.е. алгоритм вычислений основан на полиномиальном тождестве

$$P(\varepsilon) = \sum_{i=0}^{n} \frac{1}{i!} \frac{d^i P}{d\varepsilon^i}(0)\varepsilon^i . \qquad (7)$$

Значение самого полинома $P_{n,k}(\varepsilon)$ в точке ε=0, естественно, равно 1, так как первоначальная система неравенств (4) при ε=0 совместна и налагает ограничения лишь

на порядок следования переменных x_1, x_2,..., x_n. Далее находим первую производную

$$\frac{dP(\varepsilon)}{d\varepsilon} = n!\int...\int \frac{d}{d\varepsilon}\{1[...]\times...\times 1[...]\}dx_1...dx_n. \qquad (8)$$

Для этого нам потребуются следующие простые соотношения:

$$\frac{d}{dz}1[z] = \delta(z), \qquad (9)$$

$$\int_{-\infty}^{+\infty} \delta(z)f(z)dz = f(0), \qquad (10)$$

где δ (z) – обычная дельта-функция.

Проводя дифференцирование подынтегрального выражения в формуле (8) по обычным правилам дифференцирования произведения функций и используя при этом (9), исходное соотношение для первой производной переводится в сумму интегралов, каждый из которых можно упростить, осуществив (с использованием (10)) интегрирование по одной из переменных, входящих в аргумент соответствующей дельта-функции. Такая схема замечательна тем, что не требует расчета пределов интегрирования, а сам вычислительный процесс сводится к многократно повторяющимся операциям подстановки. Поскольку нас интересует лишь значение производной полинома в точке ε=0, фактическое интегрирование заменяется проверкой каждого из подынтегральных выражений на совместность: если нет нарушения в порядке следования переменных, то интеграл равен единице, в противном случае он равен нулю. В результате последовательного выполнения таких преобразований выражение (8) принимает вид суммы интегралов, кратность которых на единицу меньше по сравнению с первоначальным интегралом. Поскольку при этом структура подынтегральных выражений остается неизменной, сохраняется возможность повторного использования описанной процедуры для вычисления производных более высокого порядка без изменения принципиальной схемы.

Для нахождения общего решения $P_{n,2}(\varepsilon)$ на всем интервале изменения параметра ε нами был разработан специальный рекурсивно-комбинаторный метод. Краткая суть алгоритма такова. Интервал $(0,L)$ интерпретируется как совокупность r равных дискретов. Случайное бросание n точек на интервал $(0,L)$ интерпретируется как случайное бросание n неразличимых шаров по r ящикам. Множество из l смежных ячеек служит аналогом подынтервала длиной ε. Исход бросания, когда ни один из таких l-подынтервалов, содержащихся внутри исходного r-интервала $(0,L)$, не имеет более 2 точек, считается "успешным", а отношение общего числа "успешных" бросаний $Q(r,n,l)$ к общему числу исходов опыта $Q(r,n)$ принимается в качестве целочисленного аналога вероятности $P_{n,2}(\varepsilon)$. В принципе, нахождение замкнутого аналитического выражения для общего числа "успешных бросаний" эквивалентно полному решению задачи для $k=2$, поскольку в этом случае вероятность $P_{n,2}(\varepsilon)$ вычисляется простым предельным переходом

$$P_{n,2}(\varepsilon) = \lim_{\substack{r\to\infty \\ (l/r)\to\varepsilon}} \frac{Q(r,n,l)}{Q(r,n)}, \qquad (11)$$

а выражение для общего числа бросаний *Q(r,n)* известно (Feller, W. (1967) An Introduction to Probability Theory and Its Applications (in Russian), Mir, Moscow) [6]:

$$Q(r,n) = C_{n+r-1}^{r-1} = \frac{(n+r-1)!}{n!(r-1)!}. \qquad (12)$$

Для реализации этой идеи нами были получены рекуррентные соотношения, позволившие осуществить расчет формул *Q(r,n,l)* при последовательно возрастающих значениях *n*. Вся процедура, как и в предыдущих алгоритмах, была полностью формализована и запрограммирована.

С использованием всех описанных алгоритмов нами были проведены расчеты, с помощью которых установлено, что для произвольных *n* на интервале *0 < ε <1/(entier((n+1)/2))* вероятность $P_{n,2}(\varepsilon)$ (т.е. вероятность $P_{n,k}(\varepsilon)$ при *k=2*) представляется в виде:

$$\begin{aligned} P_{n,2}(\varepsilon) &= C_n^0 + C_n^2(-n+2)\varepsilon^2 + C_n^3(4n-10)\varepsilon^3 + C_n^4(3n^2-37n+86)\varepsilon^4 + C_n^5(-40n^2+394n- \\ &922)\varepsilon^5 + C_n^6(-15n^3+625n^2-5171n-12086)\varepsilon^6 + C_n^7(420n^3-10724n^2+79996n-187002)\varepsilon^7 \\ &+ C_n^8(105n^4-10570n^3+205499n^2-1426841n+3336406)\varepsilon^8 + C_n^9(5040n^4-155708n^3 \\ &+ 2267664n^2 - 17317506n^1 + 52315558) + C_n^{10}(-945n^5+189000n^4-15794625n^3 \\ &+ 389687181n^2 - 3798029823n + 12998966646)\varepsilon^{10} + o(\varepsilon^{10}). \end{aligned} \qquad (13)$$

Кроме того, проведенные расчеты позволили усмотреть общую закономерность и высказать гипотезу [4], заключающуюся в том, что. при *k=2* для четных *n* на интервале *1/(n/2) < ε <1/((n/2)-1)* справедлива формула

$$P_{n,2}(\varepsilon) = (2/n)C_n^{(n/2)-1}(1-((n/2)-1)\varepsilon)^n. \qquad (14)$$

Впоследствии нам удалось провести строгое математическое доказательство этой формулы [7].

Чтобы продвинуться далее в этой задаче, в качестве дополнительного метода нами предложен еще один алгоритмический подход к вычислению вероятностей $P_{n,k}(\varepsilon)$. Суть его в том, что с помощью прямого интегрирования непосредственно в общей аналитической форме последовательно отыскиваются формулы для вероятностей $P_{n,n-1}(\varepsilon)$, $P_{n,n-2}(\varepsilon)$, $P_{n,n-3}(\varepsilon)$ и т.д. Проведенные в соответствии с разработанным алгоритмом расчеты позволили установить следующие формулы:

$$P_{n,n-1}(\varepsilon) = 1 - \varepsilon^n - n\varepsilon^{n-1}(1-\varepsilon). \qquad (15)$$

$$P_{n,n-2}(\varepsilon)=\begin{cases}1-2C_n^2\varepsilon^{n-2}(1-\varepsilon)^2-2\varepsilon^n, & 0\le\varepsilon\le(1/2);\\ 1-2\varepsilon^n+(2\varepsilon-1)^n-2C_n^2\varepsilon^{n-2}(1-\varepsilon)^2, & (1/2)\le\varepsilon\le 1\end{cases} \tag{16}$$

$$P_{\substack{n,n-3\\(n>6)}}(\varepsilon)=\begin{cases}1-2\varepsilon^n+C_n^1(6\varepsilon^n-4\varepsilon^{n-1})+C_n^2(-3\varepsilon^n+\varepsilon^{n-2})+\\ \quad+C_n^3(9\varepsilon^n-18\varepsilon^{n-1}+12\varepsilon^{n-2}-3\varepsilon^{n-3}),\\ \qquad 0\le\varepsilon\le(1/2);\\ 1-2\varepsilon^n+(2\varepsilon-1)^n+C_n^1(1-\varepsilon)(-2\varepsilon^{n-1}+2(2\varepsilon-1)^{n-1})+\\ \quad+C_n^2(1-\varepsilon)^2(\varepsilon^{n-2}+(2\varepsilon-1)^{n-2})-3C_n^3\varepsilon^{n-3}(1-\varepsilon)^3,\\ \qquad (1/2)\le\varepsilon\le 1.\end{cases} \tag{17}$$

Отметим, что, в отличие всех ранее приводившихся соотношений, формулы (15)-(17) рассчитаны вручную. Формализация алгоритма для получения аналитических соотношений $P_{n,n-4}(\varepsilon)$, $P_{n,n-5}(\varepsilon)$,… программным путем представляется весьма сложной (так как выкладки необходимо программировать при двух свободных параметрах – n и ε), но вполне осуществимой задачей.

Заключение

Проведенные нами исследования показали, что, несмотря на внешнюю простоту поставленной задачи, ее решение представляет собой сложно формализуемый и весьма трудоемкий в вычислительном плане процесс. В ходе проделанной работы нам удалось провести идею, высказанную в свое время Дж. фон Нейманом: исследователь сталкивается с задачей, которую не в состоянии решить, привлекает ЭВМ для проведения трудоемких расчетов, которые способны натолкнуть его на «правильный» ответ, и в случае удачи (т.е. зная подсказанное компьютером решение) проводит строгое и конструктивное доказательство. В нашем случае формула (14) была «подсказана» компьютером, а лишь впоследствии строго доказана. Интересно также отметить, что коэффициенты, входящие в полученную нами формулу (14), являются числами Каталана, которые впервые встречаются в работах Л.Эйлера и возникают при решении огромного числа вероятностно-комбинаторных и прикладных задач.

Список литературы:

[1] *Ефимов В.М., Искольдский А.М., Лившиц З.А., Крендель Ю.М.* О характеристиках различных методов считывания изображений дискретных структур // Автометрия, 1973, №1, С.3-7.

[2] *Ефимов В.М., Резник А.Л.* Аналитическое вычисление на ЭВМ объемов, ограниченных системой гиперплоскостей в *n*-мерном пространстве //Автометрия, 1976, №1, С.116-119.
[3] *Ефимов В.М., Резник А.Л.* Аналитическое определение с помощью ЭВМ статистических характеристик процесса щелевого сканирования потока Бернулли. //Автометрия, 1977, №4, С.49-51.
[4] *Резник А.Л.* Моделирование на ЭВМ непрерывного считывания изображений дискретной структуры // Автометрия, 1981, №6, С.3-6.
[5] *Parzen, E.* Modern Probabillity Theory and Its Applications, John Wiley and Sons, Inc., New York-London, 1960
[6] *В.Феллер*. Введение в теорию вероятностей и ее приложения. М.: Мир, 1966.
[7] *A. L. Reznik, V. M. Efimov, A. V. Torgov, and A. A. Solov'ev*. Computer Analytics in Problems with a Random Partition of the Interval // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications, 2011, Vol.21, No.2, pp.202-205.

К ВОПРОСУ СОЗДАНИЯ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ МЕСТОРОЖДЕНИЙ РЕДКИХ МЕТАЛЛОВ

В.В.Костин, А.Н.Бездушный
В.В.Костин
Вычислительный Центр им. А.А. Дородницына РАН
kosvic11@mail.ru
А.Н.Бездушный
Вычислительный Центр им. А.А. Дородницына РАН
anb@ccas.ru

Аннотация

В работе рассматривается инициатива по реализации информационно-аналитической системы месторождений редких металлов. С этой целью совместно с прикладными специалистами были проанализированы сущности, их взаимосвязи, сопутствующие понятию месторождения. Отдельно рассмотрены специализации месторождений для редких металлов. Предложена онтологическая (owl) спецификация понятий, их свойств и связей предметной области. Реализованы хранилище данных (база данных добываемых природных ресурсов), среда визуального анализа сведений о месторождениях, позволяющая проводить OLAP анализ. В онотологии на работу с этими данными была настроена машина вывода Pellet, заданы таблицы фактов и сформулированы логические правила, что позволило пополнить массив хранимых данных вычисленными истинными фактами (по системе условий целостности). По результатам визуального OLAP анализа получен ряд достоверных фактов о соотношениях сущностей системы, сформулирован и проверен ряд гипотез о свойствах месторождений редких металлов.

Введение

На сегодняшний день существует спрос на анализ экономической эффективности добычи полезных ископаемых, в частности редких металлов, в отвалах месторождений. В работе рассматривается инициатива по реализации информационно-аналитической системы месторождений редких металлов.

OWL онтология

С этой целью совместно с прикладными специалистами на основе работ [1] и [2] были проанализированы сущности, определена четкая система понятий. В ходе работы выделены ключевые с точки зрения анализа сущности, заданы их взаимосвязи, сопутствующие понятию месторождения, формализованы и составлены полные словари экземпляров сущностей. Отдельно рассмотрены специализации месторождений для редких металлов, выделены географическая, геологическая и геоэкономическая составляющие. Предложена онтологическая (owl) спецификация понятий, их свойств и связей предметной области. Заданы ограничения целостности сущностей. Определены мощности связей.

OLAP анализ

Реализованы хранилище данных (база данных добываемых природных ресурсов), позволяющее сократить время обработки имеющейся информации, среда визуального анализа сведений о месторождениях, позволяющая смотреть многомерные срезы OLAP куба. Имеется возможность анализировать изменяющуюся во времени информацию. Создан дружественный интерфейс работы со средой. Проведен OLAP анализ месторождений.

Выделены закономерности в зависимостях расположения крупных месторождений лития и геологических характеристик. Отдельно проведен анализ в геоэкономическом аспекте, проанализирована экономическая целесообразность добычи редких элементов в отвалах других месторождений. В систему интегрирован инструмент, позволяющий преобразовывать информацию из таблиц хранилища данных в тройки «субъект»-«предикат»-«объект» сетевой онтологии и обратно. Успешно настроены элементы визуальной оценки системы сущностей онтологии и их свойств.

Машина Вывода

Опираясь на работу [3], проведен анализ имеющихся машин вывода. В онтологии для работы с этими данными была интегрирована и настроена машина вывода Pellet (См. [4]). Заданы таблицы фактов и сформулированы логические правила, что позволило получить новые истинные данные (по системе условий целостности). На основе этих данных был вновь проведен визуальный OLAP анализ и получен ряд достоверных фактов о соотношениях сущностей системы, сформулирован и проверен ряд гипотез о свойствах месторождений редких металлов. На основе описываемых элементов было построено html приложение, позволяющее ознакомиться с возможностями системы удаленно.

Заключение

Система, созданная в ходе данной работы, продемонстрировала высокую эффективность. Проведенные испытания показали эффективность при поиске скрытых взаимосвязей характеристик месторождения и при оценке экономической эффективности добычи. В случае использования на более крупных массивах данных система позволит выделить скрытые взаимосвязи между характеристиками месторождений и спрогнозировать экономическую эффективность ведения добычи как первичных, так и вторичных полезных ископаемых.

ЛИТЕРАТУРА

[1] Т.Ю. Усова, Н.А. Архипова, Е.А. Калиш, М.Ф. Комин, Д.С. Ключарев. Редкие металлы – сырьевое обеспечение инновационных технологий// ФГУП «ИМГРЭ», Москва.

[2] Вершинин А.В., Дьяконов И.А., Ряховский В.М., Шкотин А.В. Архитектура распределенной геоинформационной среды на основе формальных онтологии пространственных данных и сервисов. «Геоинформатика», №2, 2008.

[3] J Bock, P Haase,Q Ji,Volz. Benchmarking OWL Reasoners. // In: Proc. of the ARea2008 Workshop, 2008.

[4] E. Sirin, B. Parsia, B. Cuenca-Grau, A. Kalyanpur, Y. Katz. Pellet: A practical OWL-DL reasoner. // Journal of Web Semantics, 2007.

Parallel Algorithm of the Direct Simulation Monte Carlo Method for graphic card

Laricheva M.E., Kashkovsky A.V.
Khristianovich Institute of Theoretical and Applied Mechanics SB RAS
e-mail: laricheva@itam.nsc.ru

Key words*: CUDA, DSMC, GPU,* graphic card

Motivation and Aim: Use of graphic processors (GPU) allows increasing productivity of mathematical calculations considerably. However, the program for calculations on GPU should be adapted for this architecture. The Direct Simulation Monte-Carlo technique (DSMC) [1] is used for numerical analysis of rarefied gas flows. DSMC method is very resource-consuming, therefore there was a requirement to realize algorithm of DSMC on GPU. CUDA [2] or Compute Unified Device Architecture is used for this purpose.

Methods and Algorithms: Algorithm similar to MPI technology has been developed. Fixed number of threads is used; each of them is considered as the independent processor.

Results: Results of calculations on GPU have shown that some stages of DSMC can be accelerated from 10 to 30 times. At the same time, process of the information exchange between threads appeared extremely slow (80 % of general time of calculations).

Conclusion: The algorithm of DSMC method on GPU has been developed. We expect an increase in speed of calculations up to 20 times using extensive modification of the process of data exchange between threads.

Availability: For calculations video card GeForce GTX 460 SE has been used. The interface with GPU was provided by CUDA technology .

References:
1. G.A. Bird. (1994) *Molecular Gas Dynamics and the Direct Simulation of Gas Flows* (Oxford)
2. CUDA SDK http://www.nvidia.ru/object/cuda_home_new_ru.html

Исследование библиотек технического зрения
для построения системы учёта движения на перекрёстках

А. Е. Соколов
ИАиЭ СО РАН
e-mail: ktonibud@gmail.com

Аннотация

The number of cars on the roads has increased in the last several years. There're many traffic jams in the cities, which make citizens' life very uncomfortable. One of the reasons of traffic jams is low channel capacity of crossroads. It can be raised by the reconstruction of the crossroads. The reconstruction is based on the statistical information about the crossroad usage. It is needed to know what types of cars use the crossroad, how many cars move in different directions, their speed, etc.

One of the ways to gather the statistical information is to ask people about the directions they drive their cars, but its accuracy is very low and the price is high. The other way is to use computer vision. The road CV systems which are used nowadays aren't eligible for this task, because they are set only on few lanes, while complex analysis of the whole traffic is needed here.

In this work we propose to use video stream recorded from above. We defined requirements, proposed architecture for the system. To choose which tools to use to implement the system, we compared NI LabVIEW IMAQ functions to those of OpenCV.

За последнее десятилетие в России произошёл резкий рост количества автомобилей. Особенно это заметно в больших городах, таких как Москва, Санкт-Петербург, Новосибирск, где на многих перекрёстках образуются пробки, сильно осложняющие жизнь как водителям, так и пешеходам. Одна из причин пробок — низкая пропускная способность перекрёстков. Радикальный способ увеличения пропускной способности перекрёстка — его перестройка. Перепланировка проблемных дорожных развязок выполняется на основании результатов обследования дорожной развязки, исследования реального автомобильного потока, основных направлений движения, скоростного режима и типа автотранспорта.

Один из способов сбора этой статистики — анкетирование местных жителей, однако полученная информация неточна, а обработка анкет трудоёмка. Другой способ — использовать автоматические системы видеонаблюдения. Однако уже имеющиеся средства стационарного видеонаблюдения неприменимы для таких целей, поскольку они охватывают только полосы движения, критичные с точки зрения безопасности, а также организации «зелёной волны» [1]. Их дооснащение неприемлемо из экономических соображений.

Цель работы — разработать систему учёта движения на перекрёстках, которая

позволит собирать данную статистику, что даст возможность оптимизировать перекрёсток и тем самым уменьшить пробки.

В работе предлагается собирать статистическую информацию о перекрёстке на основе видеозаписи перекрёстка, получаемой средствами аэросъёмки [2].

Были определены требования к системе: необходима работа при изменяющихся погодных условиях, высокая производительность системы, низкая погрешность измерений. Система должна компенсировать дрожания и корректировать вносимые им искажения съёмки. Для каждого транспортного средства должны быть определены: размеры, дорога, по которой автомобиль въехал на перекрёсток, дорога, по которой выехал с перекрёстка, время въезда и время выезда. Система должна предоставлять оператору возможность задать расположение линий примыкания дорог к перекрёстку, а также указать названия улиц.

Предложено разделить систему на два функциональных компонента. Первый компонент производит анализ видеопотока, в результате чего составляется база автомобилей. Второй компонент предоставляет статистику о транспортных потоках в требуемом виде по формируемому запросу.

С целью определения инструментальных средств реализации были проведены замеры времени исполнения базовых функций систем обработки видео LabView IMAQ [3] и OpenCV [4]. Для этого были выбраны классы функций: чтение файлов .avi, фильтрация, рисование примитивов, извлечение каналов, вывод на экран и пр. Из каждого класса было выбрано по функции, и были проведены замеры времени исполнения этих функций. Каждая функция исполнялась 1000 раз в цикле. Для соблюдения равных условий функции библиотеки IMAQ и библиотеки OpenCV запускались из среды LabView, при этом для удобства была написана легковесная библиотека, упрощающая использование OpenCV из LabView. Каждый тест

запускался 3 раза. Сравнение показало более высокую (от 3 до 10 раз) эффективность библиотеки OpenCV для функций чтения файлов, фильтрации, рисования примитивов, извлечения каналов и пр. IMAQ оказался производительнее при копировании одного изображения в другое (на 10%) и при выводе на экран (в 5 раз). На основании того, что в задаче не требуется отображение обрабатываемого видео на экране, было принято решение использовать OpenCV в качестве базового инструмента.

Функция	OpenCV, мсек	IMAQ, мсек	Отношение времени IMAQ к времени OpenCV, раз
Чтение кадра из .avi	9866	75197	7,62
Вывод изображения на экран	14207	2635	0,19
Трёхцветное размытие	40455	166364	4,11
Перевод в серую шкалу	6038	18755	3,10
Перевод в серую шкалу с последующим размытием	19495	62145	3,18
Копирование изображения	4413	3933	0,89
Рисование отрезка (1000 раз*)	4053	40318	9,95
Извлечение зелёного канала	2591	8098	3,13
Преобразование из RGB в HLS	124105	396111	3,19
Копирование цветов пикселей из одного отрезка в другой (500 раз**)	2872	15454	5,38

* Учитывая, что функция запускалась в цикле 1000 раз, всего отрезок рисовался 1000000 раз.

** Соответственно, всего 500000 раз.

Итак, были определены требования к системе, была предложена архитектура, было проведено сравнение библиотек IMAQ и OpenCV.

В настоящее время работа ведётся над исследованием алгоритмов, необходимых для решения поставленной задачи при выполнении выявленных требований.

Научный руководитель — канд. техн. наук Зюбин Владимир Евгеньевич.

ЛИТЕРАТУРА

1. Д. А. Колтышев Мониторинг автомобильного трафика // XLVIII Международная научная студенческая конференция «Студент и научно-технический прогресс», 2010.

2. А. Е. Соколов Разработка системы технического зрения для учёта движения на перекрёстках // XLIX Международная научная студенческая конференция «Студент и научно-технический прогресс», 2011.

3. IMAQ Vision for LabView User Manual

http://www.csun.edu/~rd436460/Labview/IMAQ-Manual.pdf

4. Интернет-портал библиотеки OpenCV http://opencv.willowgarage.com/wiki/

ОБ ИНДИВИДУАЛЬНОМ БЫСТРОДЕЙСТВИИ АБОНЕНТСКОГО БЕСПРОВОДНОГО СОЕДИНЕНИЯ

П.А. Михеев,
Томский государственный университет
doka.patrick@gmail.com
С.П.Сущенко
ТНЦ СО РАН
ssp@inf.tsu.ru

Аннотация

В работе предложена математическая модель беспроводной сети для оценки индивидуального быстродействия абонента. На примере двух абонентов рассчитано среднее время передачи протокольного блока данных и средняя скорость передачи данных. При формализации процесса доступа к разделяемой среде показана возможность ее захвата одним из абонентов на неопределенно долгое время. Найдена вероятность захвата среды при первой попытке передачи данных. Предложены меры исключения эффекта захвата.

Рассмотрим беспроводную локальную вычислительную сеть (ЛВС), основывающуюся на стандарте 802.11 [1]. Протокол, обеспечивающий работоспособность беспроводных ЛВС в распределенном режиме доступа (DCF) [1, 2], использует механизм доступа к среде, который называется «множественный доступ с контролем несущей и предотвращением коллизий» *(carrier sense multiple access with collision avoidance, CSMA/CA).* Данный механизм основан на том, что передающая станция проверяет, присутствует ли в среде сигнал несущей и, прежде чем начать отправку кадра, ожидает освобождения среды передачи данных. Беспроводные станции стандарта 802.11, в отличие от проводных станций Ethernet, не способны обнаруживать коллизии в среде передачи данных [3, 4]. В силу этого обнаружение коллизий и бесконфликтных передач протокольных блоков данных основано на механизме тайм-аутов и алгоритме положительной решающей обратной связи. Цикл передачи кадра данных от станции отправителя к станции получателя согласно протокольной процедуре содержит следующую последовательность действий. Прежде всего, станция-отправитель прослушивает среду для определения ее незанятости. Далее по истечении межкадрового интервала запускается алгоритм случайной задержки для выбора номера слота, в котором можно начать передачу данных. Номер слота равновероятно выбирается в пределах размера конкурентного окна, измеренного в слотовых интервалах. Ширина конкурентного окна не может превышать максимального значения установленного стандартом. Для всех физических уровней и способов модуляции стандарт 802.11 устанавливает максимальную ширину конкурентного окна равную 1024 [1]. Номер выбранного слота присваивается значению таймера отсрочки, после чего начинают отсчитываться слотовые интервалы. В конце каждого слотового интервала таймер отсрочки уменьшается на единицу, при этом прослушивается среда передачи данных. Как только фиксируется занятость среды, таймер отсрочки замораживается до тех пор, пока не освободится среда передачи данных. После освобождения среды таймер запускается со значения, зафиксированного непосредственно перед замораживанием. По истечении таймера отсрочки станция-отправитель начинает передачу кадра данных. По окончании передачи отправитель ждет квитанции в течение времени тайм-аута, по истечении которого считается,

что была коллизия и станции, попавшие в коллизию, увеличивают значение номера повторной передачи на единицу, а действия, направленные на отправку данных, повторяются. Но ширина конкурентного окна удваивается с каждой попыткой передать кадр данных, пока не достигнет максимального значения, а с каждой последующей попыткой отправить данные ширина конкурентного окна остается равной максимальному размеру до тех пор, пока данные не будут успешно переданы, после чего ширина окна принимает начальное значение. Таким образом, технология беспроводного доступа в силу невозможности обнаружения коллизии в среде передачи данных имеет три существенных отличия от случайного метода доступа, реализуемого в проводной среде. Во-первых, беспроводный метод передачи использует механизм положительной обратной связи (положительных квитанций). Во-вторых, в отличие от случайного метода доступа в проводных сетях беспроводная технология WiFi уже при первой передаче использует механизм случайной отсрочки. И наконец, беспроводный протокол доступа использует механизм «замораживания» счетчика отсрочки при обнаружении факта занятости среды до истечения таймера случайной задержки.

Рассмотрим функционирование беспроводной локальной сети до первой безошибочной передачи кадра и получения квитанции об успешной доставке данных рассматриваемым абонентом [3]. Исходя из этого, найдем основные операционные характеристики системы. Предположим, что в беспроводной ЛВС имеется K станций – источников данных. Считаем, что все источники независимы, равноправны, всегда имеют кадры данных для отправки, а все интервальные промежутки выражены в слотовых интервалах. Пусть все станции обмениваются кадрами одинакового размера. Тогда согласно последовательности протокольных действий элементарный цикл отправки кадра получателю определится размером межкадрового промежутка, периодом случайной отсрочки, длительностью «заморозки» таймера случайной отсрочки, временем передачи информационного кадра, а также величиной тайм-аута ожидания положительной квитанции, которая складывается из короткого межкадрового промежутка и времени передачи положительной квитанции. Следует отметить, что среди указанных компонент цикла времена отсрочки и заморозки являются функциями номера повторной передачи. Среднее время передачи кадра складывается из взвешенной суммы средних времен ожидания неудачных отправлений и времени, затраченном на успешную передачу. Отметим, что при неудачных передачах кадра получатель не отправляет подтверждений.

Начнем анализ с изучения процесса соперничества двух станций беспроводной ЛВС ($K = 2$). Обозначим соперничающие (конфликтующие) станции через A и B. Найдем ВВХ процесса передачи данных станцией A. Средние условные времена до неудачной и успешной n–ой попытки передачи данных складываются из средней длительности случайной отсрочки (среднее количество слотов до начала передачи) и среднего количества заморозок из-за передачи данных станцией B. Численные исследования среднего времени передачи кадра абонентом показывают, что этот показатель имеет ярко выраженный минимум по параметру, определяющему начальный размер конкурентного окна и, как следствие, степень рассеяния станций по длительностям отсрочки перед началом процедуры соперничества. Очевидно, что значение размера конкурентного окна, минимизирующее среднее время передачи кадра, максимизирует индивидуальную пропускную способность. Кроме того, уже на этапе формализации задачи стал очевиден эффект захвата среды передачи данных одним из абонентов, о котором упоминается в [4]. Особенно сильно этот эффект проявляется при

малых значениях конкурентного окна. Уже при первой попытке соперничества двух станций возможен захват среды передачи данных (например станцией *B*), вероятность которого определится вероятностями того, что у одной из станций (*B*) длительность отсрочки окажется меньше длительности отсрочки другой станции (*A*), а затем у «успешной» станции (*B*) будет выпадать отсрочка нулевой длительности, чередуясь с отсрочками меньшими, чем оставшееся значение таймера отсрочки станции. Отсюда нетрудно видеть, что вероятность захвата в значительной мере определяется начальной шириной конкурентного окна. После нескольких конфликтов возможность захвата среды «успешной» станцией становится еще более вероятной. Основной причиной эффекта захвата среды передачи данных является протокольное действие «замораживание отсрочки», поскольку именно оно приводит к тому, что после бесконфликтной передачи станция может неопределенно долго удерживать среду передачи данных, попадая в интервал отсрочки от 0 до остаточного значения отсрочки других станций. Другой причиной эффекта захвата среды передачи данных после нескольких конфликтов одним из абонентов являются различные размеры конкурентного окна для станций, вышедших из конфликта и станций, продолжающих разрешение конфликта в состоянии ожидания истечения времени отсрочки и периодов заморозки. После положительного разрешения конфликта одной из станций размер ее конкурентного окна кратно сокращается до начальной ширины, что дает этой станции преимущественное право при дальнейшем соперничестве за «эфир» с «конфликтующими» станциями, поскольку меньшая длительность случайной отсрочки для такой станции имеет существенно большую вероятность по сравнению с аналогичным операционным показателем «конфликтующей» станции. Очевидно, что для исключения эффекта захвата «эфира» на неопределенно долгое время необходимо модифицировать протокольные параметры. С одной стороны, разумно зафиксировать размер конкурентного окна для первой и всех последующих передач, оптимизировав его к числу активных абонентов, а с другой – длительность случайной отсрочки следует выбирать на интервале от единицы до правой границы размера конкурентного окна, исключив тем самым отсрочку нулевого размера. При этом захват «эфира» одной станцией никогда не превысит количества успешных передач до очередного конфликта или его разрешения, равного ширине конкурентного окна, уменьшенной на два.

ЛИТЕРАТУРА

1. IEEE Std 802.11 – 2007, Revision of IEEE Std 802.11 – 1999. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Computer Society, 2007. 1184 p.
2. Олифер В.Г., Олифер Н.А. Компьютерные сети. СПб.: Питер, 2006. 958 с.
3. Новиков Ю.В., Кондратенко С.В. Локальные сети: Архитектура, алгоритмы, проектирование. М.: ЭКОМ, 2000. 312 с.
4. Вишневский В.М., Ляхов А.И., Портной С.Л., Шахнович И.В. Широкополосные беспроводные сети передачи информации. М.: Техносфера, 2005. 592 с.

Математическая физиология. Моделирование функции сердечной мышцы – от молекул до органа.

О.Э. Соловьева[1,2], Мархасин В.С.[1,2]

[1]Институт иммунологии и физиологии УрО РАН (Екатеринбург)

[2]Уральский федеральный университет имени первого Президента России Б.Н.Ельцина (Екатеринбург)

e-mail: o.solovyova@iip.uran.ru, v.markhasin@iip.uran.ru

Аннотация:

Математическое моделирование является уникальным и мощным инструментом исследования физиологических процессов, позволяющим существенно углублять наши знания об изучаемых явлениях, формировать принципиально новые количественные представления об этих явлениях, выявлять широкий спектр откликов системы, изменяя параметры модели, формулировать конкретные количественные гипотезы, которые могут быть проверены в эксперименте и, наконец, предсказывать и выявлять принципиально новые классы явлений. Уже имеются учебники по математической физиологии (см., например, *Keener J., Sneyd J.* Mathematical physiology. Springer-Verlag New York, 1998), а статьи по математическому моделированию охотно принимают известные физиологические журналы. Можно утверждать, что наряду с экспериментальной физиологией имеет место и бурно развивается самостоятельная ветвь физиологии – математическая физиология, являющаяся специфическим источником новых знаний о природе физиологических процессов. Правда, последняя находится в стадии становления, особенно в России. Опираясь на собственный многолетний опыт математического моделирования механических и электрических явлений в миокарде, мы на конкретных примерах дадим подтверждение перечисленным выше уникальным возможностям математического моделирования в физиологии. Будут приведены примеры, иллюстрирующие применение математических моделей для описания и понимания сложных процессов, а также предсказания новых явлений, протекающих на различных уровнях организации сердечной мышцы – от молекулярного до тканевого и органного. Будет показано, что реалистичные математические компьютерные модели являются средством для интегративного описания миокардиальной системы в целом с учетом взаимосвязей между процессами на разных уровнях и временных шкалах. Хотя возможности математического моделирования будут проанализированы в докладе применительно к миокарду, они, по мнению авторов, имеют универсальный характер и относятся к моделированию любых физиологических процессов.

Работа поддержана грантами Президиума УрО РАН 09-М-14-2001, РФФИ 11-04-00785-а.

Ключевые слова: математическое моделирование сердечной мышцы, миокард, динамические системы

НЕЛОКАЛЬНАЯ КРАЕВАЯ ЗАДАЧА С ЛОКАЛЬНЫМ СМЕЩЕНИЕМ ДЛЯ ОБОБЩЕННОГО УРАВНЕНИЯ МАК КЕНДРИКА-ФОН ФЕРСТЕРА

Д. А. Шоранова
Учреждение Российской Академии наук Научно - исследовательский институт прикладной математики и автоматизации КБНЦ РАН
(г. Нальчик)

Пусть $u(x,t)dx$ - количество клеток в единице объема в возрастном интервале от $x \geq 0$ до $x + dx$ в момент времени t. Если убыль клеток определяется естественной гибелью, вымыванием в случае проточного культивирования и убылью из данного возраста x при делении родительской клетки, то при интуитивно ясных предположениях относительно других характеристик другой популяции можно записать:

$$u_t(x,t) + u_x(x,t) = c(x,t)u(x,t).$$

Рассмотрим в области $G = \{0 < x < l, 0 < t < T\}$ уравнение[1]

$$\partial_{0t}^{\alpha} u(x,t) + \partial_{0x}^{\beta} u(x,t) = c(x,t)u(x,t), \tag{1}$$

где $\partial_{0z}^{\mu} u(x,t)$- регуляризованная производная по Капуто порядка μ по переменной z, $0 < \alpha < 1$, $0 < \beta < 1$,

Задача. Найти регулярное решение $u(x,t)$ уравнения (1) в области G удовлетворяющее условиям

$$u(x,0) = \varphi(x), 0 < x < l \tag{2}$$

$$u(0,t) = \sum_{i=1}^{n} \alpha_i u(x_i,t), 0 < t < T \tag{3}$$

где φ -заданная функция, α_i - неотрицательные коэффициенты.

Краевые задачи для уравнения (1) исследовались в [см. 2, с.61 и библиографию там]. Отметим также [3], где для уравнения (1) исследовалась нелокальная краевая задача. Задача (1)-(3)относится к классу внутрикраевых задач с локальным смещением [4,с.29]

Вопрос о резрешимости задачи (1)-(3) сводится к решению системы интегральных уравнений Вольтерра II рода.

$$\begin{cases} u(x_1,t) = F(x_1,t) + \sum_{i=1}^{n} \alpha_i \int_0^t u(x_i,\eta) w(x_1, t-\eta) d\eta, \\ u(x_2,t) = F(x_2,t) + \sum_{i=1}^{n} \alpha_i \int_0^t u(x_i,\eta) w(x_2, t-\eta) d\eta, \\ ... \\ u(x_j,t) = F(x_j,t) + \sum_{i=1}^{n} \alpha_i \int_0^t u(x_i,\eta) w(x_j, t-\eta) d\eta \end{cases}$$

где

$$F(t) = \int_0^x \varphi(s) \Big(- \frac{1}{x-s} \Big) e_{\alpha,\beta}^{0,1} \Big(- \frac{(x-s)^\alpha}{t^\beta} \Big) ds dx +$$

$$+ \int_0^x \int_0^t c(s,\eta) e_{\alpha,\beta}^{\alpha,0} \Big(\frac{(x-s)^\alpha}{(t-\eta)^\beta} \Big) \frac{(x-s)^{\alpha-1}}{t-\eta} d\eta ds dx,$$

$$w(x_j,t) = \frac{1}{t-\eta} e_{\alpha,\beta}^{1,0} \Big(- \frac{x_j^\alpha}{(t-\eta)^\beta} \Big), \qquad (4)$$

где $e_{\alpha,\beta}^{\mu,\delta}(z)$- функция типа Райта [2], $j = \overline{1,n}$.

С биологической точки зрения $u(x,t)$ - численность популяции возраста $x \in (0,l)$ в момент времени $t \in (0,T)$, а уравнение (1) моделирует динамику численности эволюционирующих популяций.

После определения функций $u(x_i,t)$, функция $u(x,t)$ находится как решение краевой задачи для уравнения (1) [2],[3].

Список литературы

[1] Нахушев А. М. Уравнения математической биологии. — М.: Высш. шк., 1995. —301 с.

[2] Псху А.В. Краевые задачи для дифференциальных уравнений с частными производными дробного и континуального порядка. —Нальчик, 2005. —185 с.

[3] Шоранова Д.А.Нелокальная краевая задача для уравнения рождаемости Мак Кендрика-фон Ферстера с дробной производной Капуто по времени. Материалы Международной Российско - Болгарской VIII школы молодых ученных "Нелокальные краевые задачи и проблемы современного анализа и информатики"Нальчик - Хабез, 2010 г. — 120-122 с.

[4] Нахушев А. М. Задачи со смещением для уравнений в частных производных. —М.: Высш. шк., 2006. —287 с.

ОСОБЕННОСТИ ДИНАМИЧЕСКИХ РЕЖИМОВ ТОЧЕЧНЫХ ОТОБРАЖЕНИЙ[1]

Шлюфман К.В.
Институт комплексного анализа региональных проблем ДВО РАН, Биробиджан, Россия
e-mail: Shlufman@mail.ru
Фишман Б.Е.
e-mail: bef942@mail.ru

В исследовании точечных отображений наибольшее внимание уделяется изучению точек бифуркаций и определению областей значений параметров устойчивых циклов. При этом области значений параметров, при которых исследуемое отображение, не имеет устойчивых циклов, относятся к областям апериодической динамики. Такие области, на наш взгляд, представляют наибольший интерес.

Было замечено, что точечные отображения могут иметь разные по характеру апериодические динамические режимы, причем эти различия можно однозначно определить по особенностям статистических характеристик. Это было детально исследовано на ряде популяционных моделей, для которых были рассмотрены такие характеристики, как гистограмма, фазовые портреты и автокорреляционные функции.

Использованные характеристики указали на возможность качественно различать апериодические режимы на два типа. Первый тип динамического режима представляет собой циклы конечной длины (больше единицы), элементы которых, флуктуируя в ограниченных диапазонах, образуют взаимно непересекающиеся интервалы значений решения уравнения, и имеющий незатухающие автокорреляционные функции. Предложено такую динамику называть интервально периодической. Второй тип динамического режима характеризуется единственным выделяемым интервалом значений решения и затухающей автокорреляционной функцией.

Введено формальное определение интервально периодической динамики системы.

На основании предложенного определения разработан алгоритм, позволяющий по фрагменту решения уравнения обнаруживать интервальную периодическую динамику, а также определять ее период и выделять интервалы значений.

С помощью разработанного алгоритма в пространстве параметров моделей были выделены области, в которых наблюдаются интервально периодические динамики.

[1] Исследования проведены при частичной финансовой поддержке РФФИ (№ 11-01-98512-р_восток_а)

STOCHASTIC SIMULATION OF NEURONAL SPIKE TRAINS AND DISLOCATIONS IN CRYSTALS

K.K. Sabelfeld
Институт вычислительной математики и математической геофизики СО РАН
sabelfeld.karl@yahoo.de

Abstract

In this talk I present a simulation method for the analysis of neuronal spike trains considered as stationary stochastic random processes with a given power spectrum. In rapidly growing classes of neurophysiological experiments, the train of impulses ("spikes") produced by a nerve cell is a subject of statistical treatment involving the time intervals between spikes.

A neuronal spike train is the sequence of nerve impulses, or action potentials, produced by a neuron, typically observed over a long period of time. The analysis of spike trains has been of increasing interest to neurophysiologists in recent years, stimulated by wide availability of automatic data-processing equipment. Mathematically, the neuronal spike train can be considered as a random point process defined as a generalized stochastic process. As a simple example, widely used in practice, we mention a renewal process defined as a sum of delta-impulses happened at random times having a Poisson distribution. In a generalized renewal process the random time intervals between two spikes may be arbitrarily distributed but arranges in a Markovian structure.

We evaluate explicitly the correlation function of this type of spike trains. Much more difficult is the inverse problem: we are given a correlation function B(t) of the spike train, construct a renewal process having the desired correlations B(t). We construct a stochastic algorithm which solves this problem for the renewal processes as well as for other point processes, both for stationary and nonstationary random processes. Interesting applications of the spike trains are in the analysis of dislocations in crystals, e.g., see [1]. To calculate the x-ray diffraction peaks on dislocations, we have to simulate spatial spike trains distributed along an edge of a crystal, so there is a full analogy with the neuronal spike trains. We discuss in this talk other applications of the spike trains.

References

[1] Vladimir M. Kaganer and Karl K. Sabelfeld. X-ray diffraction peaks from correlated dislocations: Monte Carlo study of the dislocation screening. Acta Crystallographica, A66, 2010, 703-716.

МОДЕЛИРОВАНИЕ ДИНАМИКИ ЧИСЛЕННОСТИ И ОПТИМИЗАЦИЯ ПРОМЫСЛА ДЛЯ ПОПУЛЯЦИИ С ВОЗРАСТНОЙ И ПОЛОВОЙ СТРУКТУРОЙ[1]

О.Л. Ревуцкая, Е.Я. Фрисман
Учреждение Российской Академии наук
Институт комплексного анализа региональных проблем ДВО РАН
О.Л. Ревуцкая
e-mail: oksana-rev@mail.ru
Е.Я. Фрисман
e-mail: frisman@mail.ru

В большом многообразии внутрипопуляционных процессов важную роль играет регуляция возрастной и половой структуры популяций. В данном сообщении рассмотрена модель динамики численности популяции, которая может быть представлена совокупностью трех групп: младшей, включающей неполовозрелых особей, и двух старших, состоящих из самок и самцов, участвующих в размножении. Исследованы сценарии переходов динамических режимов при изменении параметров системы. Показано, что потеря устойчивости равновесного решения может произойти только при прохождении пары комплексно-сопряженных корней характеристического уравнения через единичную окружность. В результате в системе наблюдаются нерегулярные колебания численности, а в фазовом пространстве модельных переменных возникает предельная замкнутая инвариантная кривая. При изменении параметров системы квазипериодические колебания приобретают хаотический характер. В результате изучения системы показано, что увеличение средней индивидуальной приспособленности особей (т.е. увеличение коэффициентов плодовитости и выживаемости) в моделях экологически лимитированных популяциях может приводить к потере устойчивости и возникновению хаотических аттракторов, структура и размерность которых меняются при изменении параметров модели. Половая асимметричность коэффициентов выживаемости способствует увеличению размерности аттракторов.

Для исследуемой системы рассмотрена задача оптимизации промысла в стационарном режиме. В общем случае эта задача может быть решена численными методами при известных значениях популяционных параметров. Проанализированы частные случаи, когда изъятие ведется дифференцированно и только из одной половозрастной группы. Показано, что увеличение коэффициента, характеризующего интенсивность рождаемости, и коэффициента выживаемости взрослых самок, а также уменьшение величины, характеризующей зависимость рождаемости от соотношения полов, приводят к увеличению значения оптимальных долей промыслового изъятия из трехкомпонентной популяции. Однако изъятие на максимальном уровне весьма опасно для популяции, поскольку небольшое снижение интенсивности рождаемости или выживаемости самок приводит в этом случае к быстрому снижению численности популяции вплоть до ее полного вырождения.

[1] Исследования проведены при частичной финансовой поддержке РФФИ (проекты № 09-04-00146-а, № 11-01-98512-р_восток_а) и ДВО РАН (проект № 11-III-B-01M-002)

Об одной экстремальной функции, связанной с неравенством А.А. Маркова

В.П. Скляров (Саратов)

Аннотация

Будут представлены к обсуждению графики экстремальных функций, связанных с равномерно взвешенным неравенством Маркова для нормы производной алгебраического многочлена.

Функция

$$M_n(x) = \sup_{p \in P_n, p \neq 0} \frac{|\rho(x)p'(x)|}{\|\rho p\|}, \tag{1}$$

впервые была рассмотрена А.А. Марковым [1] в конце девятнадцатого столетия для веса $\rho(x) \equiv 1$, $< a, b >= [-1, 1]$ и нормы $\|f\| = \max\limits_{x \in <a,b>} |f(x)|$. Здесь P_n – множество алгебраических многочленов, степени не выше n. Позднее в двадцатом столетии в работах С.Н. Бернштейна, Е.В. Вороновской, В.С. Виденского и многих других исследовались различные свойства таких функций аналитическими методами. Насколько известно автору, в наше время нет общепринятой вычислительной схемы, позволяющей численно строить график функции (1). Использованная в рассматриваемом случае схема вычислений ($\rho(x) = e^{-|x|}$ и промежуток $< a, b >= (-\infty, \infty)$) в основном опирается на одно из утверждений работы [1], суть которого заключается в том, что экстремальным в неравенстве Маркова, по необходимости, будет многочлен Золотарева. Подробное изложение этого вопроса в случае равномерно взвешенной нормы можно найти в [4]. Отметим, что полученное А.А. Марковым аналитическое выражение для $M_2(x)$ позволяет судить о точности проводимых численных экспериментов. Оценки для максимальных значений функции $M_n(x)$ для веса Лагерра в случа оси и полуоси можно найти в работах [2, 3].

Список литературы

[1] Марков А.А., Об одном вопросе Д. И. Менделеева. Избранные труды.ГИТТЛ М.-Л., 1948, 51-75.

[2] Levin A. L. and Lubinsky D. S., L_∞ Markov and Bernstein inequalities for Freud weights, SIAM J. Math. Anal. **21** (1990), no. 4, 1065–1082.

[3] Скляров В. П., О точной константе в неравенстве Маркова для веса Лагерра. Матем. сб., 200:6 (2009), 109-118.

[4] Schönhage A.. Approximationstheorie. Walter de Gruyter & Co., Berlin-New York 1971, 212 pp.

Функция разметки состояний и базисные слова для конечных автоматов

© 2011

Крайнюков Н.И., к.т.н., доцент, Тольяттинский государственный университет, кафедра «Прикладной математики и информатики»

nik9kr@gmail.com

Мельников Б.Ф., д. ф-м. н, профессор Тольяттинский государственный университет, кафедра «Прикладной математики и информатики»

B.Melnikov@tltsu.ru

Ключевые слова: конечные автоматы, сложность, представление конечных автоматов, произведение регулярных языков.

Аннотация: Представление конечных автоматов в виде базисных слов с использованием функции разметки и базисного автомата [2] позволяет в ряде случаев оценивать сложность языка, представить его произведения как количество состояний базисного автомата при процедуре детерминизации.

1.Введение

Конечные автоматы, как детерминированные так и недетерминированные, нашли широкое применение в информатике, компьютерных науках. Представление конечных автоматов в виде перезаписывающей системы на основе базисных слов позволяет исследовать алгебраические свойства множества, как последовательности вызовов операторов в программных системах.

2.Основная часть

Введем необходимые определения и обозначения.

Пусть X – конечный алфавит, X^* - свободный моноид над алфавитом $X = \{x_1, x_2, ...x_m\}$. Язык $L \subset X^*$ - является подмножеством классов эквивалентности гомоморфизма $\psi(X^*)$ из свободного моноида X^* в синтаксический моноид $M = M(L)$, порожденный синтаксической конгруэнцией $\approx_L$ множества L.

Обратно, по заданной конгруэнции $\approx$ на X^* мы можем построить [3] автомат $K_{\approx} = (Q_{\approx}, X, \delta_{\approx}, q_{\varepsilon})$ следующим образом, слову $w \in X^*$ сопоставим класс эквивалентности $[w]_{\approx}$, множеством состояний $Q_{\approx}$ будет множество всех классов эквивалентности $\approx$, пустому слову ε, будет соответствовать начальное состояние q_{ε} и для любого $x \in X$ определим функцию переходов

$\delta_{\approx}([w]_{\approx}, x) = [wx]_{\approx}$. Если число классов эквивалентности конгруэнции $\approx$ на $X*$ конечно, тогда автомат, порожденный этой конгруэнцией, будет иметь конечное число состояний. Построенный конечный автомат $K_{\approx}$, будет всюду определенным детерминированным автоматом, но без множества заключительных состояния $F_{\approx}$. Регулярный язык $L_{\approx}$, допускаемый автоматом $K_{\approx} = (Q_{\approx}, X, \delta_{\approx}, q_{\varepsilon}, F_{\approx})$, определяется выбором множества допустимых состояний $F_{\approx}$.

Недетерминированный конечный автомат (НДА) $K = (Q, X, \delta, I, F)$, где Q - конечное множество состояний, X – входной алфавит, δ – функция переходов состояний, под действием символов входного алфавита X, $\delta : Q \times X \to 2^Q$, I – множество начальных состояний, F – множество заключительных состояний, 2^Q - множество всех подмножеств состояний автомата K.

Конечный автомат K – является детерминированным (ДКА), если он имеет одно начальное состояние и $|\delta(q, x)| \leq 1$, для любых $q \in Q$ и $x \in X$, в противном случае, он является недетерминированным (НКА).

Для НДА K, используя алгоритмы детерминизации и минимизации, можно построить минимальный детерминированный автомат $\tilde{K}$, который будем называть каноническим [2].

Пусть канонический автомат $\tilde{K} = (Q, X, \delta, q_s, F)$ для регулярного языка L имеет N состояний $q_i \in Q$, $i = 1,..N$ и определяет правую конгруэнцию $\approx_L$ на свободном моноиде X^*. Расширим определение функции переходов стандартным образом, т.е. $\delta(q, \varepsilon) = q$ и $\delta(q, wx) = \delta(\delta(w, q), x)$.

Выберем в каждом классе эквивалентности по представителю $\tilde{w}_i$, $i = 1,..N$, введем представляющую функцию $P : Q \to X^*$, которая сопоставляет каждому состоянию $q \in Q$, такое слово w, что $\delta(q_s, w) = q$., т.е. слово w - приводит из начального состояния q_s в состояние q. Понятно, что если ввести упорядочение слов в алфавите X, то для значений представительной функции можно выбрать слова, наименьшие в этом порядке. Будем называть такие слова **базисными**, множество базисных слов обозначим.

Пример 1.

Рассмотрим язык над алфавитом $X = \{a, b\}$, заданный регулярным выражением X^*ab. Ниже приведена таблица и граф переходов конечного автомата A_1, допускающего этот язык.

	1	2	3
a	2	2	2
b	1	3	1

Начальное состояние: [1]

Заключительное состояние: [3]

Таб. 1. Таблица переходов автомата, допускающего язык X^*ab.

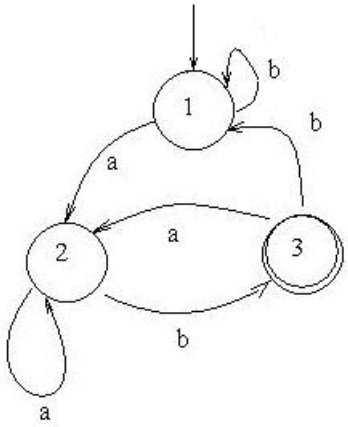


Рис 1. Граф переходов автомата, построенного по таблице 1.

Представительные функции $P_1(q)$ $P_2(q)$ могут быть заданы следующим образом:

Состояние q	$P_1(q)$	$P_2(q)$
1	$P_1(1) = b$	$P_2(1) = bb$
2	$P_1(2) = ba$	$P_2(2) = a$
3	$P_1(3) = bab$	$P_2(3) = ab$

Очевидно, что представительная функция канонического автомата всегда будет инъективной и ее значения – это представители классов эквивалентности $[w]_{\approx}$ по конгруэнтности $w \approx v$ тогда и только тогда $\delta(w, q_s) = \delta(v, q_s)$, где $w, v \in X^*$

Используя представительную функцию для канонического автомата $\tilde{K} = (Q, X, \delta, q_s, F)$ легко вычислить функцию разметки [2] для любого автомата $N = (Q', X, \delta', q_s', F')$, допускающего язык L:

$$\varphi^{in}_{\tilde{K}N}(q) = \delta'(q_s', P(q)) \qquad (1)$$

Для недетерминированного автомата представительная функция может быть не инъективной, Детерминизация обеспечивает инъективность представительной функции.

С каноническим автоматом, можно связать переписывающую систему [4]. Напомним определение, переписывающая система – это пара $\gamma = (X, T)$, где X - алфавит, а $T \subset X^* \times X^*$ множество пар (w, v) (они могут записываться в виде $w \to v$), который называются перезаписывающими правилами/продукциями. Для слов $s, t \in X^*$, будем писать $s \Rightarrow_\gamma t$, если $s = s_1 u s_2$, $t = s_1 v s_2$ для некоторой продукции $w \to v \in T$.

Построим для конечного канонического автомата $\tilde{K} = (Q, X, \delta, q_s, F)$ с фиксированной представительной функцией $P : Q \to X^*$ переписывающую систему $\gamma = (X, T)$. Определить автомат с помощью перезаписывающей системы можно, указав правила перехода из состояния в состояние по буквам входного алфавита. Правила $E = \{(wx, P(\delta(q_s, wx))\}$, где $w \in B$ - базисное слово. Автоматная конгруэнция $\approx$ определяет в правилах перезаписывающей системы правила вида $\{(wx, vx)\}$, где (w, v) правила для базисных слов, $x \in X^*$.

Пример 2.

Для автомата A_1, из примера 1, продукции перезаписывающей системы для представляющих функций $P_1(q)$, $P_2(q)$.

Состояние q	$P_1(q)$	$w \to v$	$P_2(q)$	$w \to v$
1	$P_1(1) = b$	$(ba, ba), (bb, b)$	$P_2(1) = bb$	$(bab, ba), (bbb, bb)$
2	$P_1(2) = ba$	$(baa, ba), (bab, bab)$	$P_2(2) = a$	$(aa, a), (ab, ab)$
3	$P_1(3) = bab$	$(baba, ba), (babb, b)$	$P_2(3) = ab$	$(aba, a), (abb, bb)$

Продукции для представляющей функции $P_1(q)$ дополняются правилами (bbv, bv), $(baav, bav)$, ,$(babv, babv)$, $(babav, bav)$, $(babbv, bv)$ для любого слова $v \in X^*$.

3.Заключение

Представление автомата в виде базисных слов, представляющей функции и функции разметки позволяет для исследования сложности регулярных языков использовать методов ассоциативных исчислений, перезаписывающих систем.

4.Список литературы

1.Хорпкрофт Дж., Р.Мотвани, Дж. Ульман «Введение в теорию автоматов. Языков и исчислений», Из-во «Вильямс», 2008 – 528с.
2.Мельников Б.Ф. Недетерминированные конечные автоматы.Тольятти. ТГУ, 2009 – 161 с.
3.Лаллеман Ж. Полугруппы и комбинаторные приложения. – М. Мир, 1985-440 с.
4.Паун Г., Розенберг Г., Саломаа А. ДНК-компьютер. Новая парадигма вычислений. – М.: Мир, 2004 – 528 с.
5.Брауэр В. Введение в теорию конечных автоматов. М.: Радио и связь, 1987.-392 с.

НА ПУТИ К ВЕРИФИКАЦИИ СИ-ПРОГРАММ: СТАНДАРТНАЯ БИБЛИОТЕКА

Промский А.В.[1]

Институт систем информатики им. А.П. Ершова СО РАН
630090 Новосибирск, просп. Лаврентьева, 6
E-mail: promsky@iis.nsk.su

Аннотация

Для решения важной задачи верификации Си-программ в лаборатории теоретического программирования ИСИ СО РАН развит двухуровневый подход. Входной язык C-light является представительным подмножеством стандарта Си. Его формальное определение задано в виде операционной семантики. Суть двухуровневого подхода в том, что в языке C-light выделено ядро — язык C-kernel, в которое транслируются исходные программы. Были разработаны правила перевода из C-light в C-kernel, а также аксиоматическая семантика языка C-kernel, используемая для верификации. Были формально доказаны корректность перевода и непротиворечивость аксиоматической семантики относительно операционной.

Успешная разработка теоретических методов позволяет перейти к решению ряда практических задач, одной из которых является разработка логических спецификаций для стандартной библиотеки языка Си. Заметим, что эта проблема практически не отражена в литературе. Причиной является как низкий уровень многих функций библиотеки, так и отсутствие (до недавнего времени) удобного языка спецификаций Си-программ.

В данной работе описан подход к специфицированию библиотеки посредством языка ACSL. Изученное подмножество, названное Lib-light, включает такие важные средства, как файловый ввод-вывод, обработка строк, работа с памятью и математические функции. Спецификации включают в себя логические определения типов (в том числе рекурсивные), используемых в этих библиотеках, пред- и постусловия для библиотечных функций, а также инварианты циклов. Достоинством языка ACSL является нотация на основе самого языка Си, поэтому спецификации будут понятны не только теоретикам, но и программистам-практикам. Ряд библиотечных функций были не просто специфицированы, но и верифицированы в рамках двухуровневого метода.

1 Работа поддержана Лаврентьевским молодежным грантом СО РАН «*Верификация программ на языке С с эффективной локализацией ошибок*» и частично поддержана грантом РФФИ № 11-01-00028-а.

Оптимизация рестриктивной динамической системы с запаздывающим управлением на примере инерционного рынка одного товара

В.В. Поддубный
Томский государственный университет, vvpoddubny@gmail.com
О.В. Романович
hjkm@ngs.ru

Рассматривается рестриктивная (подчиняющаяся ограничениям типа неравенств) управляемая динамическая система, описываемая в каждый момент дискретного времени t неуправляемой переменной x_t, $x_{\min} \le x_t \le x_{\max}$, и управляемой переменной y_t, подчиняющейся рестриктивному рекуррентному соотношению $y_{t+1} = y_t + u_{t-\tau} - \min(f(x_t), y_t + u_{t-\tau})$, где f – монотонно убывающая функция x_t, $u_{t-\tau} \ge 0$ – запаздывающее на время τ управление. Найдена структура оптимального управления такой системой по локальному критерию максимума вогнутой по x_t кусочно дифференцируемой целевой функции $J(x_t, y_t, \min(f(x_t), y_t + u_{t-\tau}))$. В качестве примера приведено решение задачи математического описания и имитационного моделирования инерционного рынка одного товара при оптимальном управлении поставкой товара на рынок в условиях запаздывания поставок. В этой модели x_t – цена товара в момент времени t, y_t – остаток товара на рынке, $u_{t-\tau}$ – объём товара, заказываемого в момент времени $t-\tau$ для поставки на рынок в момент t (стратегия поставок), $f(x_t)$ – спрос на товар при цене x_t, $\min(f(x_t), y_t + u_{t-\tau})$ – объём продаж, $J = \min(f(x_t), y_t + u_{t-\tau})x_t - c_1 u_{t-\tau} - c_2 y_t - (R/2)(x_t - x_{t-1})^2$ – прибыль продавца, равная разности между выручкой от продажи товара и затратами на его приобретение и хранение. Последнее слагаемое (с коэффициентом $R > 0$) выражает «штрафные санкции» за изменение цены товара и определяет инерционность рынка.

Ключевые слова: динамическая система, ограничения типа неравенств, управление, запаздывание, оптимизация, критерий, рынок одного товара, имитационное моделирование.

Population Modeler: программа для индивидуум-ориентированного моделирования сообществ взаимодействующих особей*

Б. Ю. Пичугин

ОФИМ им. С. Л. Соболева СО РАН

e-mail: bpichugin@mail.ru

11-14 октября 2011

Аннотация

В данном докладе рассказывается о программе Population Modeler, разработанной для проведения вычислительных экспериментов с вероятностными моделями сообществ особей. Формулируются базовые предположения, в рамках которых строятся модели. Описываются особенности реализации программы. Работа программы продемонстрирована на примере модели распространения туберкулёза органов дыхания и модели травления особей сообщества токсичным веществом.

Программа Population Modeler предназначена для проведения численных экспериментов с моделями сообществ особей, удовлетворяющих следующим предположениям:

- все сообщество поделено на несколько популяций;
- каждая особь сообщества охарактеризована своим набором параметров; количество и семантика параметров определяются популяцией, к которой принадлежит особь;
- изменение состояния сообщества (рождение и гибель особей, изменение параметров особей) происходит скачкообразно в случайные моменты времени;
- скачки состояния сообщества могут быть вызваны *взаимодействиями* или *переходами*;
- взаимодействия — это "незапланированные"изменения состояния сообщества:

*Работа выполнена при финансовой поддержке РФФИ (проект №09-01-00098-А) и СО РАН (интеграционный проект №26)

  - моменты возникновения взаимодействий описываются пуассоновским потоком, интенсивность которого зависит от состояния сообщества;
  - вероятность выбора особи в качестве участника взаимодействия и результат взаимодействия могут зависеть от параметров особей;
- переход — это «запланированное» изменение состояния сообщества:
  - моменты возникновения переходов описываются ветвящимся процессом типа Беллмана–Харриса с несколькими типами частиц, в котором тип частицы интерпретируется как тип перехода;
  - цепь переходов может быть инициирована в момент рождения особи или в момент взаимодействия особей;
  - распределение времени между переходами и результат перехода могут зависеть от параметров особей.

В частности, при помощи данной программы можно рассчитывать реализации широкого класса ветвящихся случайных процессов с взаимодействием частиц и марковских случайных процессов как с дискретным, так и с непрерывным временем.

Особенностями реализации программы Population Modeler:

- вычисления распараллеливаются по реализациям как в рамках одной ЭВМ, так и между несколькими ЭВМ, соединенными в локальную сеть;
- алгоритм поиска ближайшего момента перехода, построен на базе алгоритма сортировки кучей;
- алгоритм выбора особи для взаимодействия, построен на базе почти полного бинарного дерева, для представления которого использован линейный массив со специальной индексацией;
- использован мультипликативный датчик случайных чисел с модулем 2^{128};
- эффективное использование памяти позволяет рассчитывать сообщества в несколько миллионов особей;
- использован алгоритм накопления ошибки для представления моментов времени (фактически этот алгоритм дает 128-битную точность представления времени);
- программа реализована в виде библиотеки java-классов (интерфейса пользователя пока нет).
- кросплатформенность.

Тестирование производилось на известных моделях, допускающих в некоторых случаях аналитическое выражение таких характеристик, как математическое ожидание, его асимптотика, вероятность вырождения и т.п. Например, в качестве тестовых моделей

были использованы ветвящийся процесс Беллмана–Харриса, модель процесса регулируемого размножения нейтронов. Результаты всех проведенных расчетов согласуются с аналитическими выражениями.

При помощи программы Populations Modeler было произведено исследование индивидуум-ориентированной модели распространения туберкулеза органов дыхания [5] и модели динамики популяций, развивающихся в условиях воздействия токсичных веществ [6].

Литература

[1] Б. Ю. Пичугин, Н. В. Перцев. Статистическое моделирование популяций взаимодействующих частиц с произвольным распределением времени жизни // Математические структуры и моделирование. — Омск: ОмГУ, 2001. Вып. 7. С. 67–78.

[2] N. V. Pertsev, B. J. Pichugin. Stochastic modeling of the individual's community with their transformation and interaction // Proceedings of the International Conference on Computational Mathematics. Part I. — Novosibirsk: ICM&MG Publisher, 2002. P. 249–253.

[3] Б. Ю. Пичугин. Стохастическая модель сообщества взаимодействующих особей, охарактеризованных набором параметров // Труды международной конференции по вычислительной математике МКВМ-2004. Ч. I / Под ред. Г. А. Михайлова, В. П. Ильина, Ю. М. Лаевского. Новосибирск: Изд. ИВМиМГ СО РАН, 2004. С. 303–309.

[4] Н. В. Перцев, Б. Ю. Пичугин. Применение метода Монте—Карло для моделирования динамики сообществ взаимодействующих индивидуумов // Вестник Воронежского государственного технического университета. Серия «Вычислительные и информационно-телекоммуникационные системы», Т. 2, № 5, 2006, С. 70–77.

[5] Н. В. Перцев, Б. Ю. Пичугин. Индивидуум-ориентированная стохастическая модель распространения туберкулеза // Сиб. журн. индустр. математики. 2008. Т. 12. № 2(38). С. 97–110.

[6] Н. В. Перцев, Б. Ю. Пичугин, К. К. Логинов. Статистическое моделирование динамики популяций, развивающихся в условиях воздействия токсичных веществ // Сиб. журн. индустр. математики. 2011. Т. 14, № 2. С. 84—94.

Влияние параметров на устойчивость биосферной модели

А.М. Федотов, С.Б. Медведев, А.И. Пестунов

Институт Вычислительных Технологий

В работе продолжено исследование минимальной модели углеродного цикла [1]. Ранее были найдены условия [2], при которых происходит потеря устойчивости стационарного решения модели. Для исследования механизма переключения из устойчивого состояния в неустойчивое проводится анализ влияния параметров на динамику системы. В соответствии с теоремой Ляпунова об устойчивости по линейному приближению для решения поставленной задачи достаточно рассмотреть вещественную часть ранее найденных собственных значений и определить ситуации, в которых происходит смена знака.

Цель исследования заключается в определении параметров, критически значимых для прогнозирования сценариев биосферной динамики, ошибка в оценке или естественное изменение которых может существенно повлиять на результаты моделирования.

Литература

1. Барцев С.И., Дегерменджи А.Г., Ерохин Д.В. Глобальная минимальная модель многолетней динамики углерода в биосфере // ДАН, Геофизика, 2005, том 401, №2, с. 233-237.
2. Федотов А.М., Медведев С.Б., Пестунов А.И., Пестунов И.А. О нестандартном поведении минимальной модели углеродного цикла // Вестник НГУ. Серия: Информационные технологии. 2011. Том 9, выпуск 1. С. 82-88.

НЕЛИНЕЙНЫЕ РЕЖИМЫ ДИНАМИКИ ЛИМИТИРОВАННОЙ ПОПУЛЯЦИИ С РАЗЛИЧНОЙ РОЖДАЕМОСТЬЮ У ОСОБЕЙ РЕПРОДУКТИВНОГО ВОЗРАСТА[1]

Г.П. Неверова, Е.Я. Фрисман
Учреждение Российской Академии наук
Институт комплексного анализа региональных проблем ДВО РАН
Г.П. Неверова
e-mail: galina.nev@gmail.com
Е.Я. Фрисман
e-mail: frisman@mail.ru

Открытие циклических и хаотических режимов в простейших моделях динамики биологических популяций дало новый важный импульс исследованиям природы популяционных колебаний. Наиболее впечатляющими здесь оказываются различные эффекты нелинейности, проявляющиеся уже в моделях с возрастной структурой, представленной несколькими возрастными группами особей.

В данной работе исследуется трехкомпонентная модель динамики численности популяции с сезонным характером размножения, которая может быть применена к описанию и анализу динамики трехлетних растений, некоторых видов рыб, насекомых и мелких млекопитающих. Акцент делается на изучение влияния плотностно-зависимых факторов на популяции с различной рождаемостью у особей разных возрастов, составляющих репродуктивную группу. Проведено подробное исследование всех типов плотностной регуляции, возникающей в процессе внутривидовой конкуренции между тремя возрастными классами. Для каждого случая показано, что возможно единственное нетривиальное стационарное состояние системы. Определена область его устойчивости, проанализирован характер потери устойчивости и сценарии переходов динамических режимов. Сделаны выводы о динамике численности популяции в зависимости от значений параметра выживаемости молоди.

Показано, что в лимитированных популяциях с различной рождаемостью у особей разных возрастов, составляющих репродуктивную группу, параметрическая область нерегулярной динамики тем больше, чем больше возрастные различия в коэффициентах воспроизводства. При этом рост репродуктивного потенциала в некоторых ситуациях способен обеспечить переход «от хаоса к порядку» и даже привести к устойчивым динамическим режимам.

[1] Исследования проведены при частичной финансовой поддержке РФФИ (проекты № 09-04-00146-а, № 11-01-98512-р_восток_а) и ДВО РАН (проект № 11-III-В-01М-002)

О суперкомпиляции

(к 80-тилетию со дня рождения В. Ф. Турчина)

(Тезисы доклада)

Андрей П. Немытых

Институт программных систем РАН
nemytykh@math.botik.ru

В статье описывается история и обзор современного состояния развития идей В.Ф. Турчина в области автоматического преобразования программ, известных как суперкомпиляция, и дается анализ этих идей в общем контексте автоматической специализации программ. Исследования в области построения систематических методов специализации программ по отношению к фиксированным свойствам их аргументов были начаты в 1970-х годах А.П. Ершовым («смешанные вычисления»), В.Ф. Турчиным («суперкомиляция») и Ё. Футамурой («generalized partial computation»).

Идеи суперкомиляции в основном изучались в предметной области функционального языка рефал [40], хотя ряд результатов отшлифовывался на экспериментальной базе функционального языка LISP и его потомков. В настоящее время, наряду с несколькими примитивными моделями суперкомпиляторов [32], [19], [14], [16], построенными для простейших чисто теоретических языков, существует единственный экспериментальный суперкомпилятор SCP4 для реального языка программирования рефал-5 [40], [44]. Ниже нам понадобится следующее определение:

Определение 1. *Реализацией функционального языка программирования* $\Re$ *назовём четвёрку* $\langle P, D, U, T\rangle$*, где множество* P *называется множеством* $\Re$*-программ, множество* D *называется множеством* $\Re$*-данных; частично рекурсивные функции* U*:* $P\times D \mapsto D$ *и* $T : P \times D \mapsto \mathbb{N}$ *называются соответственно универсальной функцией (или семантикой) и сигнализирующей функцией времени языка* $\Re$*. Здесь* $\mathbb{N}$ *– множество натуральных чисел.*

Будем использовать обозначение p(x) для U(p,x). В литературе рассматриваются две постановки задачи специализации как таковой. Сформулируем их. Пусть дана реализация функционального языка программирования $\langle$P,D,U,T$\rangle$, где $\mathtt{D} = \bigcup_{n\in\mathbb{N}} M^n$ для некоторого множества M.

Задача №1. Пусть $\mathtt{p(x,y)} \in \mathtt{P}$ реализует *частично рекурсивную функцию* $F(x, y) : \mathtt{D}\times\mathtt{D} \mapsto \mathtt{D}$. Зафиксируем значение первого аргумента этой функции $x_0 \in \mathtt{D}$. В задаче специализации требуется построить $\mathtt{q(y)} \in \mathtt{P}$ такую, что

$$\forall \mathtt{y} \in \mathtt{D}.(\mathtt{q(y)} = \mathtt{p}(\mathtt{x}_0, \mathtt{y})) \wedge (\mathtt{T(q,y)} \leq \mathtt{T(p,}\mathtt{x}_0, \mathtt{y})),$$

где значение q(y) определено тогда и только тогда, когда определено $\mathtt{p}(\mathtt{x}_0,\mathtt{y})$. В случае неопределённости $\mathtt{p}(\mathtt{x}_0,\mathtt{y}_0)$ и $\mathtt{q}(\mathtt{y}_0)$ их типы неопределённости должны совпадать: программа q(y) представляет ту же самую *частично рекурсивную функцию*, что и программа $\mathtt{p}(\mathtt{x}_0,\mathtt{y})$, а именно – $F(\mathtt{x}_0,\mathtt{y})$.

Задача №2. Пусть $\mathtt{p(x,y)} \in \mathtt{P}$ реализует *общерекурсивную функцию* $F(x,y) : X \times Y \mapsto \mathtt{D}$, где $X \subset \mathtt{D}, Y \subset \mathtt{D}$. Зафиксируем значение первого аргумента этой функции $x_0 \in X$. Требуется построить $\mathtt{q(y)} \in \mathtt{P}$ такую, что

$$\forall \mathtt{y} \in Y.(\mathtt{q(y)} = \mathtt{p}(\mathtt{x}_0, \mathtt{y})) \wedge (\mathtt{T(q,y)} \leq \mathtt{T}(\mathtt{p},\mathtt{x}_0,\mathtt{y})).$$

Т.е., во второй задаче программа q представляет некоторое *продолжение общерекурсивной функции* $F(x_0, y) : Y \mapsto \mathtt{D}$ по второй компоненте.

Программу q называют *остаточной* программой. *Содержательная сторона задачи состоит в построении оптимальной q (по времени исполнения).* Разные уточнения понятия оптимальности (функциии T) определяют конкретные аппроксимации задачи специализации как таковой. Первая задача, грубо говоря, требует от q сохранения операционной семантики исходной программы p. Вторая задача более естественна: пользователя обычно не интересует каким будет операционное поведение q на входных данных, не принадлежащих к области определения его предметной области. Условия второй задачи оставляют бóльшую свободу методам специализации, что часто позволяет строить более оптимальные остаточные программы по сравнению с методами действующими в рамках первой задачи.

1 Обзор результатов в специализации программ

На пути реализации идей, сформулированных А.П. Ершовым, В.Ф. Турчиным и Ё. Футамурой, встретились серьёзные трудности. Н.Д. Джонс предложил ослабить первоначально поставленные цели за счёт упрощения методов специализации [11], [10]. Эта техника «частичных вычислений» наиболее разработана к данному моменту. Она решает первую задачу специализации. Пытаясь решить задачи самоприменения специализатора, также сформулированные в 70-х годах независимо вышеуказанными тремя исследователями, Н.Д. Джонс с сотрудниками сделали ещё один принципиальный шаг в сторону упрощения, но теперь уже не постановки задачи, а ограничения подходов к её решению. В 1983 г. им удалось решить аппроксимирующие задачи самоприменения частичного вычислителя mix. Заметим, что всегда существует функция времени T, делающая любую задачу специализации программы $\mathtt{p}(\mathtt{x}_0,\mathtt{y})$ бессодержательной: функция T, которая позволяет построить следующую остаточную программу $\mathtt{q(y)}$ { = $\mathtt{p}(\mathtt{x}_0,\mathtt{y})$; }. Первые результаты самоприменения mix, содержательно, незначительно отличались от приведённой тривиальной остаточной программы. Как пишет Н.Д. Джонс [10], длина остаточной программы, полученной в результате решения простейшей задачи самоприменения $\mathtt{mix}(\underline{\mathtt{mix}(\mathtt{p}_0,\mathtt{x},\mathtt{y})})$[1] mix по данной трёхстрочной программе $\mathtt{p}_0\mathtt{(x,y)}$, была пятьсот страниц текста. Здесь значения y неизвестны обеим копиям mix; значение аргумента x известно специализируемому mix, но неизвестно специализирующему mix. Анализируя остаточную программу, группа Джонса предложила понятия «online» и «offline» методов специализации [11]. Выбор более простых offline методов

[1] Здесь подчёркивание означает кодировку.

позволил в 1984 г. решить более приемлемые аппроксимирующие задачи самоприменения частичного вычислителя `mix` [11], [30]. Введя инструменты повышения местности специализируемых программ в рамках offline подхода, С.А. Романенко [26], [28], [29] удалось в 1987 г. существенно улучшить временны́е и структурные характеристики остаточных программ задач самоприменения частичного вычислителя `unmix`.

Offline специализация разделяет в отдельные стадии анализ исходной программы `p` и метаинтерпретацию локальных шагов этой программы, которые могут быть выполнены без знания *конкретных* значений неизвестной части аргументов этих шагов. На вход первой стадии, которая называется связыванием по времени (BTA[2]), подаётся `p` и указание – какие из её аргументов будут известны на второй стадии преобразований, а не сами значения этих аргументов, и какие неизвестны. Первые называются статическими, вторые динамическими. На выходе у BTA размеченная программа $\mathtt{p}^{\mathtt{ann}}$, в которой каждое элементарное действие помечено как статическое, если его можно однозначно проинтерпретировать без знания *конкретных* значений динамической части входов этих действий-шагов. Аргументы каждого шага также размечаются на статические и динамические; анализ «движения» статической информации по `p` производится BTA. Задача, поставленная перед BTA, как таковая алгоритмически неразрешима. Повсюду здесь имеется в виду некоторая аппроксимация BTA-задачи. На вход второй стадии, которая собственно и называется при этом подходе «*специализацией*», подаётся результат BTA – $\mathtt{p}^{\mathtt{ann}}$ и значения её статических аргументов. «Специализация» логически проста и алгоритмизируема: все содержательные проблемы перенесены в BTA. Группа Н.Д. Джонса, как и С.А. Романенко, решила не оригинальную классическую задачу самоприменения, а задачу $\mathtt{mix}(\underline{\mathtt{mix}^{\mathtt{ann}}(\mathtt{p}_0^{\mathtt{ann}},\mathtt{x},\mathtt{y})})$[3], которая существенно проще классической: обе копии `mix` производят лишь вторую стадию преобразований, не занимаясь связыванием по времени. Здесь важно, что входные данные каждого шага программы просматриваются только один раз на этапе «специализации».

Online специализация производит метавычисления шагов преобразуемой программы `p` по ходу дела анализа свойств этой программы; никак не ограничивая, вообще говоря, *a priori* себя ни в средствах, ни в количестве проходов по программе `p`. Каждый проход даёт цикл, который в общем случае алгоритмически нераспознаваем, даже если он работает вхолостую, не производя преобразований, а только занимаясь поиском какого-то свойства и не находя его. Следовательно, этот цикл в общем случае будет присутствовать в остаточной программе, ухудшая её временну́ю сложность. При попытке решения задач самоприменения, проходы по данным, анализирующие являются ли эти данные статическими или динамическими, будут наблюдаться преобразующей копией специализатора, что существенно осложняет его логику. Алгоритмически неразрешимая и разрешимая части этой логики никак не разделены. Резюме: разработка методов online специализации

[2] Binding time analysis.

[3] Здесь в задаче решённой С.А. Романенко нужно заменить оба `mix` на `unmix`.

сложнее разработки методов offline специализации. По определению, online специализация менее ограничена в методах, чем offline специализация, и, потому, потенциально значительно более сильная. Самым важным является то, что при offline специализации происходят преобразования только исходной программы p; а при online специализации могут происходить также преобразования подпрограмм, построенных самим специализатором, а не только подпрограмм программы p, написанных человеком.

Суперкомпиляция [39], [23] и generalized partial computation (GPC) [4], [6], как наборы методов online специализации, дают намного более сильные механизмы автоматического анализа и преобразования программ, чем частичные вычисления. И, как следствие, ставят много более трудные задачи, – чисто познавательные, алгоритмические и технологические. Методы суперкомпиляции [37], [39], [24], [24], [23], [20] и GPC [4], [5], в отличии от методов частичных вычислений: позволяют иногда понижать порядок временно́й сложности специализируемых программ; строят остаточные программы полностью на основании метаинтерпретации специализируемой программы, а не на её пошаговой чистке. По-видимому, наименее разработанными на данный момент нужно считать идеи GPC. В отличие от частичных вычислений и суперкомпиляции, подход GPC к специализации незамкнут.

2 Обзор развития методов суперкомпиляции

Мы именует основные этапы истории развития идей суперкомпиляции следуя терминологии В.Ф. Турчина [38], [39]. Первая публикация В.Ф. Турчина «Эквивалентные преобразования рекурсивных функций, описанных на языке рефал» датируется 1972 г. [33]. Рефал изначально проектировался Турчиным как метаязык, ориентированный на преобразование программ. В этой статье описывается подмножество рефала, в котором время отождествления входных данных функции с образцом равномерно ограничено по размеру входных данных. На синтаксис образцов накладывается ограничение. Эти ограниченные образцы названы L-выражениями. Вводится понятие прогонки L-выражений. В.Ф. Турчин формулирует исчисление эквивалентных преобразований ограниченных рефал программ. Идеи этого исчисления заложили основу методов суперкомпиляции. Второй важной работой В.Ф. Турчина стал Courant Computer Science отчёт №20 (1980 г., [35]): работа изобилует примерами неалгоритмизированных преобразований программ, постановками интересных задач, многие из которых до сих пор не решены.

SCP1. Первая простейшая модель суперкомпилятора была реализована в Нью-Йорке в 1981 г. В.Ф. Турчиным, Б. Ниренбергом и Д.В. Турчиным [43] на одном из диалектов рефала. SCP1 работал в диалоговом режиме, запрашивая у человека результат обобщения конфигураций: основная задача аппроксимации алгоритмически неразрешимой части логики суперкомпиляции осталась полностью за скобками рассмотрения. SCP1 представлял важный шаг в шлифовке алгоритма прогонки, которая производила метавычисления вызовов по необходимости на этапе суперкомпиляции. Авторы проспециализировали несколько простых примеров: SCP1 решал *Задачу №2.*

SCP2 был разработан В.Ф. Турчиным в 1984 г. Статья В.Ф. Турчина «The concept of a supercompiler», опубликованная в 1986 г. [37] и описывающая некоторые идеи реализации суперкомпилятора SCP2, стала основным классическим трудом по методам суперкомпиляции. В языке описания параметризованных конфигураций программы появился конструктор отрицания, что позволило решить методами суперкомпиляции классическую задачу преобразования наивного алгоритма `p(s,x)` поиска подстроки `s` в строке `x` в алгоритм КМР [17]; посредством специализации исходной программы по первому аргументу $\texttt{p(s}_0\texttt{,x)}$. Показана возможность использования SCP2 для автоматического доказательства простых теорем существования. Алгоритм обобщения работает *ad hoc*, и для получения интересных преобразований требуется вмешательство человека в работу этого алгоритма.

Во второй половине 80-х гг. на одном из рефал семинаров А. Веденов объявил о скором завершении им реализации суперкомпилятора рефала. Публикаций и сообщений о фактической реализации не последовало. В 1987 г. выходят два препринта ИПМ АН, рассматривающие проблемы суперкомпиляции: «Рефал-4 – расширение рефала-2, обеспечивающее выразимость прогонки» (С.А. Романенко, [27]) и «Метавычислитель для языка рефал, основные понятия и примеры» (Анд. В. Климов, С.А. Романенко, [15]).

Второй по значимости работой В.Ф. Турчина стала статья «The algorithm of generalization in the supercompiler» (1988 г., [36]), в которой описывается алгоритм обобщения стека вызовов функций и доказывается теорема об остановке этого алгоритма. Алгоритм получил название «Обнинский алгоритм разрезания стека». Этот алгоритм является одним из самых важных алгоритмов суперкомпиляции. SCP2 был расширен этим алгоритмом [41].

В 1989 г. была сделана первая реальная попытка самоприменения суперкомпилятора. Базисными конфигурациями называются параметризованные обобщённые конфигурации специализируемой программы `p`, в терминах которых можно описать соответствующую `p` остаточную программу. В результате изучения трассировки циклящегося процесса самоприменения SCP2 в ручную были построены конечные множества базисных конфигураций для нескольких простых задач самоприменения. Эти множества базисных конфигураций обеспечивали конечность процесса суперкомпиляции на данных задачах без использования алгоритма обобщения. Удалив из SCP2 алгоритм обобщения и сообщая на входе суперкомпилятору SCP2 для каждой задачи множество базисных конфигураций, соответствующее этой задаче, удалось провести удачные эксперименты с несколькими простыми задачами самоприменения [8]. Таким образом, было достигнуто самоприменение алгоритмически разрешимой части задачи специализации как таковой, – без участия аппроксимирующего алгоритма обобщения.

В 1990 г. Н.В. Кондратьев предпринял попытку реализации суперкомпилятора рефала, которая осталась незавершённой. Внутренним языком преобразований являлся вариант языка рефал графов. Аналогичную незавершённую попытку в 1992 г. сделали С.М. Абрамов и Р.Ф. Гурин. В 1993 г. Анд. В. Климов и Р. Глюк опубликовали статью [7], в которой алгоритм

прогонки излагался в терминах данных языка LISP. Целью этой публикации было ознакомление зарубежных исследователей с некоторыми простыми идеями суперкомпиляции, демонстрируя их на бинарных деревьях.

SCP3. В 1993 г. В.Ф. Турчин решил ограничить объектный язык суперкомпилятора алгоритмически полным подмножеством базисного рефала-5, в котором не допускаются явные синтаксические конструкции композиции вызовов функций и синтаксис образцов гарантирует, что время отождествления равномерно ограничено по размеру входных данных. Критическим шагом было расширение языка параметров: введение дополнительной типизации параметров, позволяющей более точно формулировать задачи на самоприменение. В 1994 г. удалось достигнуть *полностью автоматического самоприменения* SCP3 на нескольких простых задачах. В 1996 г. появилась статья [24] (В.Ф. Турчин, А.П. Немытых, В.А. Пинчук) с изложением основных идей, позволивших провести эти удачные эксперименты. Алгоритм обобщения всё еще не имел под собой прочной теоретической основы, хотя и работал *полностью автоматически.*

В 1995 г. М. Соренсен [31] предложил использовать отношение Хигмна-Краскала [9], [18] для принятия аппроксимирующего решения: *Обобщать или не обобщать две данные конфигурации?* Это поставило на теоретическую основу алгоритм обобщения «положительной» части конфигураций, описанной *без использования логического отрицания.* В 1996 г. вышла статья М. Соренсена, Р. Глюка и Н.Д. Джонса, описывающая модельный суперкомпилятор для простейшего подмножества LISPа: язык описания параметризованных конфигураций не использовал связки отрицания. В 1995 г. выходит монография С.М. Абрамова «Метавычисления и их применение» [1], в которой дан алгоритм обобщения части параметризованной конфигурации, описанной с использованием связки отрицания, но заданной только в терминах данных *единичного размера.*

SCP4. Результатом продолжительных (под руководством В.Ф. Турчина) исследований автора данной статьи была разработка и реализация экспериментального суперкомпилятора SCP4 (1999-2003) для реального языка программирования рефал-5. В процессе этой работы были разработаны и реализованы принципиально новые алгоритмы преобразований. Самым ключевым из которых стал алгоритм автоматического построения выходного формата функции F по ходу дела суперкомпиляции; что позволяет сразу использовать построенный формат для специализации по нему как других функций, вызывающих F, так и функции F. До настоящего времени SCP4 является единственным экспериментальным суперкомпилятором, который доступен также в режиме online. В 2007 г. вышла монография А.П. Немытых «Суперкомпилятор SCP4: Общая структура» [23]. В полной версии данной статьи даётся общая структура SCP4.

В 2008-2010 гг. [19], [14], [16] были разработаны и реализованы простые суперкомпиляторы для модельных языков высшего порядка. В них не реализован ключевой Обнинский метод обобщения параметризованных стеков.

Список литературы

1. Абрамов С.М. Метавычисления и их применения. 1995, Наука-Физматлит.
2. Библиотека по языку рефал. http://refal.botik.ru/library/library.htm
3. Ershov A. P. Mixed computation: potential applications and problems for study. In: Theoretical Computer Science. Vol. **18**, no. **1**, (1982), pp:41–67.
4. Futamura Y., Nogi K. Generalized partial computation In the Proc. of the IFIP TC2 Workshop, (1988) 133–151. Amsterdam: North-Holland Publishing Co.
5. Futamura Y., Nogi K., Takano A. Essence of generalized partial computation. Theoretical Computer Science. **90** (1991) 61–79. North-Holland Publishing Co.
6. Futamura Y., Konishi Z., Glück R. Program Transformation System Based on Generalized Partial Computation. New Generation Computing. Vol. **90** (2002) 75–99. Ohmsha Ltd. and Springer-Verlag.
7. Glück R., Klimov And. V. Occam's razor in metacomputation: the notion of a perfect process tree. In Proc. of the Static Analysis Symposium, LNCS, Vol. **724** (1993) 112–123, Springer-Verlag.
8. Glück R., Turchin V. F. Application of metasystem transition to function inversion and transformation. In the Proc. of the ISSAC'90 (1990), 286–287. ACM Press.
9. Higman G. Ordering by divisibility in abstract algebras. Proc. London Math. Soc. **2(7)** (1952) 326–336.
10. Jones N.D. MIX ten years later. In the Proc. of the ACM SIGPLAN PEPM'95, (1995) 24–38, ACM Press.
11. Jones N.D., Gomard C.K., Sestoft P. Partial Evaluation and Automatic Program Generation. (1993) Prentice Hall International.
12. Jones N.D., Sestoft P., Søndergaard H. An experiment in partial evaluation: the generation of a compiler generator. In Proc. of Conf. on Rewriting Techniques and Applications, LNCS, **202** (1985), 125–140. Springer-Verlag.
13. Jones N.D., Sestoft P., Søndergaard H. Mix: A Self-Applicable Partial Evaluator for Experiments in Compiler Generation. Lisp and Symbolic Computation, **2(1)** (1989) 9–50.
14. Jonsson P.A., Nordlander J. Positive Supercompilation for a higher order call-by-value language. ACM SIGPLAN Notices., (2009), Vol. **44**, no. **1**, pp:277–288.
15. Климов Анд. В., Романенко С. А. Метавычислитель для языка РЕФАЛ, базисные понятия и примеры. Препринт № **71** (1987), ИПМ АН СССР, Москва.
16. Ключников И. Г. Суперкомпилятор HOSC 1.5: гомеоморфное вложение и обобщение для выражений высшего порядка. Препринт № **62** (2010), ИПМ РАН.
17. Knuth D. E., Morris J. H., Pratt V. R. Fast Pattern Matching in strings. SIAM J. Comput., Vol. **6(2)** (1977) 323–350.
18. Kruskal J.B. Well-quasi-ordering, the tree theorem, and vazsonyi's conjecture. Trans. Amer. Math. Society, **95** (1960) 210–225.
19. Mitchell N., Runciman C. A supercompiler for core Haskell. In Implementation and Application of Functional Languages, LNCS, Vol. **5083**, 147–164. Springer.
20. Nemytykh A.P. A Note on Elimination of Simplest Recursions.. In Proc. of the ACM SIGPLAN Asian Symposium on Partial Evaluation and Semantics-Based Program Manipulation, (2002) 138–146, ACM Press.
21. Nemytykh A.P. The Supercompiler SCP4: General Structure (extended abstract). LNCS, **2890** (2003) 162–170, Springer-Verlag.
22. Nemytykh A.P. Playing on REFAL. In Proc. of the International Workshop on Program Understanding, (2003) 29–39, A.P. Ershov Institute of Informatics Systems, Syberian Branch of Russian Academy of Sciences. Accessible via: ftp://www.botik.ru/pub/local/scp/refal5/nemytykh_PU03.ps.gz

23. Немытых А. П. Суперкомпилятор SCP4: общая структура. Монография, М: Издательство УРСС, 2007, ISBN 978-5-382-00365-8. - 152 с.
24. Nemytykh A. P., Pinchuk V. A., Turchin V. F. A Self-Applicable Supercompiler. In Proc. the PEPM'96 LNCS, Vol. **1110** (1996) 322–337, Springer-Verlag.
25. Nemytykh A.P., Turchin V.F. The Supercompiler SCP4: sources, on-line demonstration, http://www.botik.ru/pub/local/scp/refal5/, (2000).
26. Романенко С. А. Генератор компиляторов порожденный само-применимым специализатором может иметь ясную и естественную структуру. Препринт № **20** (1987), ИПМ АН СССР, Москва.
27. Романенко С. А. РЕФАЛ-4 – расширение РЕФАЛа-2, обеспечивающее выразимость прогонки. Препринт № **147** (1987), ИПМ АН СССР, Москва.
28. Romanenko S. A. A compiler generator produced by a self-applicable specializer can have a surprisingly natural and understandable structure. In The Proc. of the IFIP TC2 Workshop, Partial Evaluation and Mixed Computation, (1988) 445–463, North-Holland Publishing Co.
29. Romanenko S. A. Arity raiser and its use in program specialization In the Proc. of the ESOP'90, LNCS, Vol. **432** (1990) 341–360, Springer-Verlag.
30. Sestoft P. The structure of a self-applicable partial evaluator. In the Proc. of the Programs as Data Objects, LNCS, Vol. **217** (1986) 236–256, Springer-Verlag.
31. Sørensen M. H., Glück R. An algorithm of generalization in positive supercompilation. Logic Programming: Proc. of the 1995 Int. Symposium (1995) 486–479, MIT Press.
32. Sørensen M. H., Glück R., Jones N. D. A positive supercompiler. Journal of Functional Programming, Vol. **6(6)** (1996) 811–838, MIT Press.
33. Турчин В.Ф. Эквивалентные преобразования рекурсивных функций, описанных на языке РЕФАЛ. В сб.: Труды симпозиума «Теория языков и методы построения систем программирования», Киев-Алушта: 1972. Стр. 31 – 42.
34. Turchin V.F. The use of metasystem transition in theorem proving and program optimization.LNCS, Vol. **85** (1980) 645–657, Springer-Verlag.
35. Turchin V.F. The language Refal – The Theory of Compilation and Metasystem Analysis. Courant Computer Science Report, Num. **20** (1980), New York University.
36. Turchin V.F. The algorithm of generalization in the supercompiler. In the Proc. of the IFIP TC2 Workshop, (1988) 531–549.
37. Turchin V.F. The concept of a supercompiler. ACM Transactions on Programming Languages and Systems. **8** (1986) 292–325, ACM Press.
38. Turchin V.F. Metacomputation: Metasystem transition plus supercompilation. In Proc. the PEPM'96, LNCS, Vol. **1110** (1996) 481–509, Springer-Verlag.
39. Turchin V.F. Supercompilation: Techniques and results. In the Proc. of PSI'96, LNCS, Vol. **1181** (1996) 227–248, Springer-Verlag.
40. Turchin V.F. Refal-5, Programming Guide and Reference Manual. Holyoke, Massachusetts. (1989) New England Publishing Co.
(electronic version: http://www.botik.ru/pub/local/scp/refal5/ ,2000)
41. Турчин В. Ф. Metacomputation in the language Refal. (1990) (Электронная версия: Библиотека по языку рефал, http://refal.botik.ru/library/library.htm)
42. Turchin V. F., Nemytykh A. P. Metavariables: Their implementation and use in Program Transformation, Technical Report CSc. **TR 95-012** (1995), City College of the City University of New York.
43. Turchin V. F., Nireberg R., Turchin D. V. Experiments with a supercompiler Conference Record of the ACM Symposium on LISP and Functional Programming, (1982) 47–55, ACM Press.

44. Turchin V.F., Turchin D.V., Konyshev A.P., Nemytykh A.P. Refal-5: sources, executable modules. URL: http://www.botik.ru/pub/local/scp/refal5/, (2000)

Győző KOVÁCS Dr. hc.
John von Neumann Computer Society, Hungary

The Hungarian „Ljapunov": Dr. Tihamér Nemes (1895-1960).

He lived earlier as Alexey Andreevich Lyapunov, but I found several similarities in their life, e.g. he was never well accepted by the officials of the Hungarian scientific life, he followed several research activities, but he was never appointed as a scientist, but he was. He never was refused with his activities, but his ideas were not understood by the scientific public opinion (communities).

During his life, he developed and patented a lot of *cybernetical l*ike instruments mainly in the thirties, before Norbert Wiener (1894-1964) declared the bacic concepts of the cybernetics (1948). His aim was to realize (to copy) the human act and thinking with engineering methodes, mechanical elements, and electeronical circuits. Only a little bit of his ideas were realized, because the level of the technical and electronical tools were not enough developed in this time.

He patented and constructed e.g. several logical machines, walking machines, chess machiness etc. He was one of the constructor of the Hungarian television system, because he studied the methode of the human eyes.

I am a fortunate young person, because in the last two-three years of his life, I was in an ordinally connetion with him. We constructed the first Hungarian computer the M-3 between 1957-60, and he visited many times me, and questioned us about our computer. I was 24 years old, he was about 60, and I was happy to answer his questions, he was a very useful adviser. He was a very excellent speaker, too, he delivered several lectures in different institutions, and the audience enjojed his presentations.

Unfortunatelly he never saw running our computer. He died.

Items:

- Alexey Andreevich Lyapunov, educational problems, and popular science or
- General problems of cybernetics.

БЫСТРОЕ ВЫЧИСЛЕНИЕ СЛОЖНОСТИ ПО АРНОЛЬДУ ДВОИЧНЫХ СЛОВ ДЛИНЫ 2^n *

Ю. В. Мерекин

Институт математики им. С. Л. Соболева СО РАН

merekin@math.nsc.ru

Для произвольного двоичного слова $w = x_1x_2 \dots x_{2^n}$, длины $|w| = 2^n$, $n \geq 1$, возможно быстрое вычисление сложности по Арнольду $A(w)$ [1],[2]. Для этого вычисления предлагается алгоритм и получена верхняя оценка функции Шеннона $Sh(n)$, характеризующая число применения булева оператора $\oplus$ ко всем словам w длины 2^n.

Теорема. *Для слов* (w), $|w| = 2^n$, $n \geq 5$, *справедливы оценки*

$$Sh(n) \leq \begin{cases} \lfloor n - 2\sqrt{n} + 1 \rfloor, & \text{когда двоичное представление числа } A(w) \text{ нечётное;} \\ \lfloor n - 2\sqrt{n} + 2 \rfloor, & \text{в противном случае.} \end{cases}$$

Литература

1. Yu. V. Merekin, On the Computational Complexity of the Arnold Complexity of Binary Words, *Asian-European Journal of Mathematics,* **2**(4), 641–648, (2009).
2. Yu. V. Merekin, On the Computation of Arnold Complexity of of Length 2^n Binary Words, *Asian-European Journal of Mathematics,* **4**(2), 295–300, (2011).

*Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 11-01-00997)

ГЕНЕТИЧЕСКАЯ ОСНОВА МАКРОЭВОЛЮЦИОННЫХ ПРЕОБРАЗОВАНИЙ: ИССЛЕДОВАНИЕ РЕЖИМОВ МОЛЕКУЛЯРНОЙ ЭВОЛЮЦИИ ОРТОЛОГИЧНЫХ БЕЛКОВ ПОЗВОНОЧНЫХ И БЕСПОЗВОНОЧНЫХ

Работа выполнена при поддержке грантом РФФИ №09-04-01641-а.

К.В. Гунбин, В.В. Суслов

Институт Цитологии и Генетики СО РАН, Новосибирск, Россия

E-mail: genkvg@bionet.nsc.ru, valya@bionet.nsc.ru

Д.А. Афонников

Институт Цитологии и Генетики СО РАН, Новосибирский государственный Университет, Новосибирск, Россия

E-mail: ada@bionet.nsc.ru

В 2008 г. было четко показано, что естественный отбор белок-кодирующих генов неоднороден по своей природе и связан с двумя разными явлениями [1]: оптимизацией уровня экспрессии гена и оптимизацией структурно-функциональной организации белка. Очевидно, что второе явление связано прежде всего с аминокислотными заменами, по-разному влияющими на структуру и функцию белка. На этой основе нами был создан простой метод анализа режимов эволюции белков, основанный на Марковском моделировании эволюции белков. Используя созданный метод, в настоящей работе мы провели анализ режимов эволюции ортологичных белков позвоночных и беспозвоночных.

В анализ брались ортологичные группы белков из базы данных MetaPhOrs [2] содержащие по одному наиболее подтвержденному ортологу из 20 таксономических групп позвоночных (Amphibia, Tetraodontiformes, Smegmamorpha, Cypriniformes, Aves, Squamata, Prototheria, Metatheria, Afrotheria, Xenarthra, Insectivora, Carnivora, Equidae, Chiroptera, Artiodactyla, Catarrhini, Hominidae, низшие Primates, Lagomorpha и Rodentia) и 9 таксономических групп беспозвоночных (Lophotrochozoa, Caenorhabditis, низшие Chromadorea, Arachnida, Paraneoptera, Lepidoptera, Hymenoptera, Drosophilina и Culicomorpha). В качестве аутгруппы были использованы таксоны Cnidaria и Tunicata. На основе известной топологии деревьев позвоночных и беспозвоночных (Рис. 1) и вычисленных нами специфических для каждой ортологической группы матриц скоростей аминокислотных замен (с помощью программы MODELESTIMATOR [3]) в

каждом внутреннем узле дерева с помощью PAML [4] реконструировались предковые последовательности белков. Эти последовательности использовались для подсчета числа наблюдаемых аминокислотных замен. Для поиска нетипичных типов аминокислотных замен на каждой ветви дерева мы сравнивали число наблюдаемых замен с их ожидаемыми значениями для каждого типа замен в предположении о стационарном Марковском процессе эволюции. Ожидаемые значения замен подсчитывались на основе 1000 компьютерных симуляций молекулярной эволюции белков при помощи пакета INDELible [5], с учетом особенностей исследуемых белков. Сравнение числа ожидаемых и наблюдаемых замен каждого типа производилось с помощью перестановочного теста (10^5 перестановок). В ходе теста подсчитывалось число случайных выборок M, в которых частота ожидаемых замен определенного типа больше частоты наблюдаемых замен. Величина $M/10^5$ оценивала вероятность p, с которой встречаемость замен определенного типа, наблюдаемая в исходной выборке генов, могла возникнуть по случайным причинам.

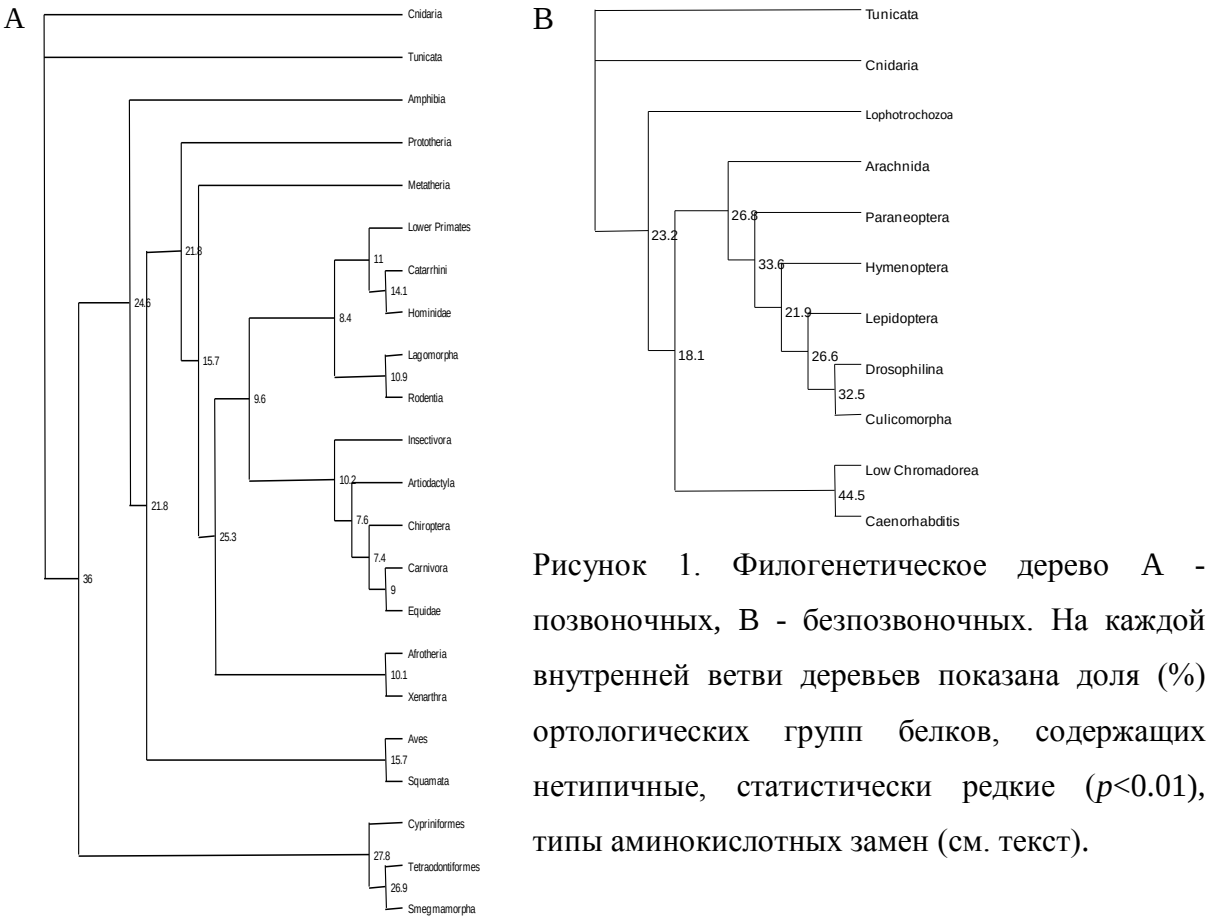


Рисунок 1. Филогенетическое дерево А - позвоночных, В - безпозвоночных. На каждой внутренней ветви деревьев показана доля (%) ортологических групп белков, содержащих нетипичные, статистически редкие (p<0.01), типы аминокислотных замен (см. текст).

В результате проведенного анализа было показано, что внутренние ветви, содержащие >20% ортологичных групп белков содержащих нетипичные (p<0.01), аминокислотные замены соответствуют основным ароморфозам позвоночных (Рис 1.): 1) периодам полногеномных дупликаций при формировании позвоночных и рыб, 2) выходу на сушу, 3) формированию амниот, 4) дивергенции млекопитающих, а также группы плацентарных. Также было показано значительное по сравнению с позвоночными (>12%) и с кембрийскими дивергенциями (>7%) увеличение числа ортологичных групп белков, содержащих нетипичные, аминокислотные типы замен (p<0.01) в группах насекомых и нематод (Рис 1.). Абсолютный максимум этого числа характерен для формирования группы нематод, интерсивно теряющей доменное разнообразие белков [6]. Показано, что формирование групп Insecta и Diptera сопровождаемое увеличением числа ортологичных групп белков, содержащих нетипичные, аминокислотные типы замен может быть связано с формированием тесного экологического сообщества насекомых и покрытосеменных растений [7]. В настоящей работе также проведен сравнительный анализ функций ортологичных групп белков содержащих нетипичные замены на каждой внутренней ветви дерева позвоночных и беспозвоночных, что позволило связать режимы молекулярной эволюции определенных ортологичных групп с морфологическими преобразованиями в палеонтологической истории таксонов Bilateria (Рис. 1).

Использованная литература

1. Drummond D.A., Wilke C.O. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell. 2008. Vol. 134, № 2. P. 341-352.

2. Pryszcz L.P. et al. MetaPhOrs: orthology and paralogy predictions from multiple phylogenetic evidence using a consistency-based confidence score. Nucl. Acids Res. 2010. doi: 10.1093/nar/gkq953.

3. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 2007. Vol. 24, № 8. P. 1586-1591.

4. Arvestad L. Efficient methods for estimating amino acid replacement rates, J. Mol. Evol. 2006. Vol. 62, № 6. P. 663-673.

5. Fletcher W., Yang Z. INDELible: a flexible simulator of biological sequence evolution, Mol. Biol. Evol. 2009. Vol. 26, № 8. P. 1879-1888.

6. Zmasek C.M., Godzik A. Strong functional patterns in the evolution of eukaryotic genomes revealed by the reconstruction of ancestral protein domain repertoires, Genome Biol. 2011. Vol. 12, № 1. P. R4.

7. Grimaldi D., Engel M.S. Evolution of the insects. NY: Cambridge University Press, 2005. 755p.

DEEP INSIDE INTO VERTEBRATES AND INVERTEBRATES MACROEVOLUTION: THE MOLECULAR EVOLUTION MODES OF STRICT ORTHOLOGOUS PROTEIN SEQUENCES

This work was supported by RFBR grant No. 09-04-01641-a.

K.V. Gunbin, V.V. Suslov

Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia

E-mail: genkvg@bionet.nsc.ru, valya@bionet.nsc.ru

D.A. Afonnikov

Institute of Cytology and Genetics SB RAS, Novosibirsk State University, Novosibirsk, Russia

E-mail: ada@bionet.nsc.ru

In 2008, it was clearly demonstrated that the natural selection of protein-coding genes is heterogeneous in nature and associated with two different phenomena [1]: optimization of the gene expression level and optimization of protein structural and functional organization. It is of known, that the second phenomenon is primarily associated with amino acid substitutions having different effects on the protein structure and function. Thus, simple method for the analysis of protein evolution modes was developed. This method is based on Markov simulation of protein sequence evolution. This study, we analyzed the evolution modes of vertebrate and invertebrate orthologous protein sequences using our novel method.

The molecular evolution modes of 752 orthologous protein groups (OPGs) of vertebrates and of 1736 OPGs of invertebrates were analyzed. In our analysis the OPGs were taken from MetaPhOrs database [2]. The OPGs containing at least one strictly confirmed orthologous sequence without ambiguous characters from each of 20 vertebrate species groups and 9 invertebrate species groups were taken into analysis (Figure 1). Cnidaria and Tunicata were served as outgroup taxa. The ancestral protein reconstruction in each internal tree node of OPG was made using PAML 4 CODEML program [3] (marginal reconstruction) on the basis of the known tree topology (Figure 1) and amino acid replacement matrices calculated by MODELESTIMATOR [4] program. These ancestral sequences were used to calculate the number of observed amino acid substitutions. For each amino acid replacement type we compared the observed number of changes with expected ones under the assumption of a stationary Markov process of protein evolution. Expected replacement numbers were calculated using 1000 computer simulations of OPG molecular evolution by the INDELible

program [5], taking into account the peculiarities of the investigated OPGs. Comparison of expected and observed numbers of each replacement type was performed using permutation tests (10^5 permutations). We count the number of random samples, *M*, in which the frequency of expected changes of a certain type higher than the frequency of observed ones. Thus, the value $M/10^5$ is the occurrence probability, *p*, of a certain amino acid replacement type observed by chance.

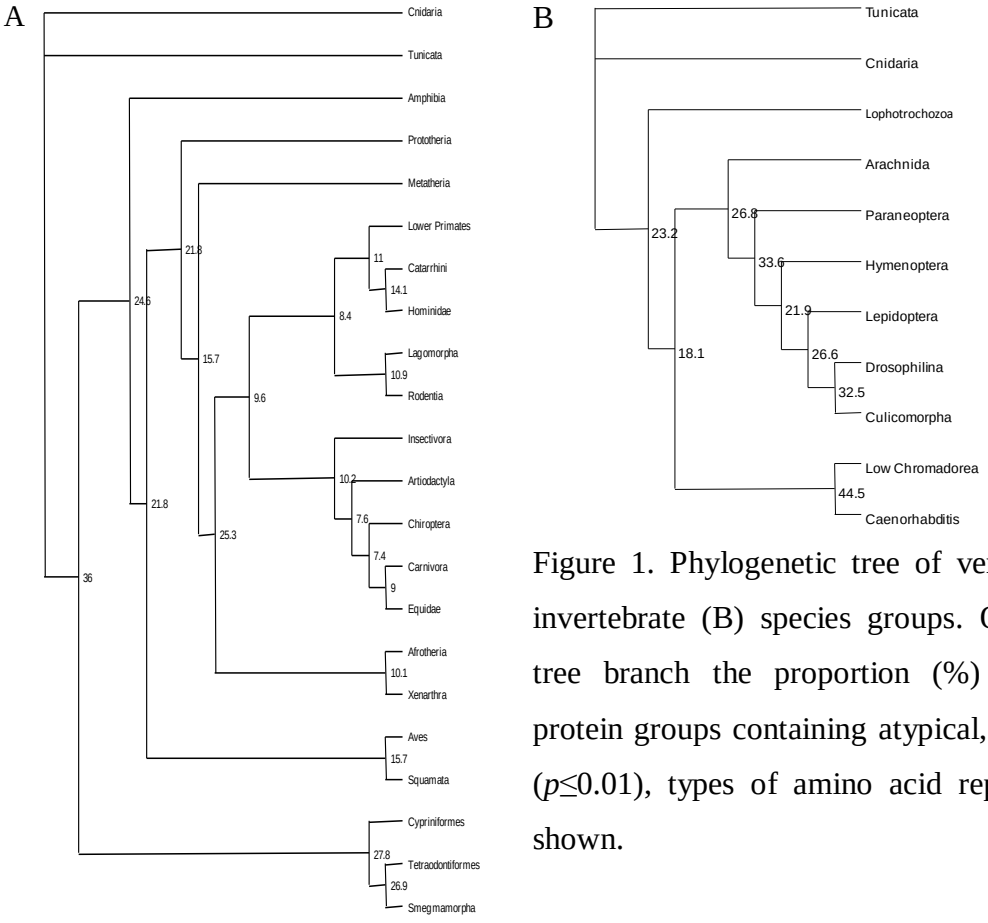


Figure 1. Phylogenetic tree of vertebrate (A) and invertebrate (B) species groups. On each internal tree branch the proportion (%) of orthologous protein groups containing atypical, statistically rare ($p \leq 0.01$), types of amino acid replacements were shown.

The analysis of *R* value, the proportion (%) of OPGs with atypical amino acid replacements ($p \leq 0.01$), showed that internal tree branches of vertebrate tree with $R \geq 20\%$ strictly correspond to aromorphoses in the vertebrate evolution (Figure 1): 1) the full genome duplications in early stages of vertebrate evolution and actynopterigian fish evolution, 2) the adaptation to terrestrial environments, 3) the origin of Amniota, 4) the divergence of primitive mammals and placental mammals. It is of big interest that the molecular evolution of insects and nematodes were characterized by significant increase in *R* value in comparison with its mean value for vertebrates (Figure 1). The absolute maximum of *R* value is typical for nematodes - a group with intensive lost of protein domain diversity [6]. It was shown that the

divergence of Insecta and Diptera accompanied by increasing of *R* (Figure 1) may be due to the emergence of insects-angiosperms ecosystems and to the formation of the characteristic Diptera morphology [7]. This study we also conducted the functional comparison of OPGs containing atypical amino acid replacements that allowed us to uncover various features of gene networks molecular evolution on each internal branches of vertebrate and invertebrate trees (Figure 1).

References

1. Drummond D.A., Wilke C.O. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell. 2008. Vol. 134, № 2. P. 341-352.

2. Pryszcz L.P. et al. () MetaPhOrs: orthology and paralogy predictions from multiple phylogenetic evidence using a consistency-based confidence score. Nucl. Acids Res. 2010. doi: 10.1093/nar/gkq953.

3. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 2007. Vol. 24, № 8. P. 1586-1591.

4. Arvestad L. Efficient methods for estimating amino acid replacement rates, J. Mol. Evol. 2006. Vol. 62, № 6. P. 663-673.

5. Fletcher W., Yang Z. INDELible: a flexible simulator of biological sequence evolution, Mol. Biol. Evol. 2009. Vol. 26, № 8. P. 1879-1888.

6. Zmasek C.M., Godzik A. Strong functional patterns in the evolution of eukaryotic genomes revealed by the reconstruction of ancestral protein domain repertoires, Genome Biol. 2011. Vol. 12, № 1. P. R4.

7. Grimaldi D., Engel M.S. Evolution of the insects. NY: Cambridge University Press, 2005. 755p.

О равновесиях Вальраса и Эджворта в многорегиональных экономических системах*

В.А. Васильев
Институт математики им. С.Л. Соболева СО РАН, Новосибирск
vasilev@math.nsc.ru

Расширенные тезисы

Введение

Работа продолжает исследования автора по вопросам эквивалентности равновесий Вальраса и Эджворта в неклассических рынках, начатые в [1]. Излагаемые в докладе результаты относятся к многорегиональным экономическим системам, изучавшимся в работах [2 - 5] и представляющим не только теоретический, но и значительный прикладной интерес. Отметим сразу же, что рассматриваемая проблематика тесно связана с задачей обоснования известной гипотезы Эджворта об асимптотической эквивалентности неблокируемых и равновесных распределений [6,7]. Показано, что найденные автором условия строгой автаркичности и ненасыщенности гарантируют совпадение нечетких ядер и вальрасовских планов анализируемых неклассических рыночных систем. На основании этого факта и полученного ранее представления равновесий Эджворта в виде элементов так называемого нечеткого $\mathbb{Q}$-ядра [5], установлен центральный результат доклада: строгая автаркичность и ненасыщенность достаточны и для совпадения множеств равновесных планов Вальраса и Эджворта. Обсуждаются также некоторые применения полученных теорем эквивалентности к отысканию условий существования вальрасовских равновесий рассматриваемых многорегиональных систем.

1 Модель $\mathcal{M}$

В докладе рассматривается модель экономического взаимодействия регионов из [2], имеющая следующий вид:

$$\mathcal{M} =< R, \{A^s, G^s, H^s, b^s, d^s\}_{s\in R} >,$$

где $R = \{1, \ldots, r\}$ - множество регионов; A^s - прямоугольная матрица размера $n_s \times l_s$, характеризующая производственный сектор региона $s \in R$; G^s и H^s - прямоугольные

*Работа выполнена при финансовой поддержке Междисциплинарного интеграционного проекта СО РАН "Развитие теории, методологии и прикладных экономико-математических исследований"и Российского фонда фундаментальных исследований, № 10-06-00168а.

матрицы размера $n_s \times n$, оисывающие способы вывоза и ввоза в регионе $s \in R$; b^s - вектор-столбец размерности n_s, характеризующий имеющийся ресурсно-технологический потенциал региона $s \in R$; d^s - вектор-столбец размерности n_s, описывающий затраты ресурсов и продукции, связанные с достижением целей развития региона $s \in R$.

Ресурсно-технологические возможности Z_s региона $s \in R$ определяются формулой

$$Z_s := \{z^s = (x^s, u^s, v^s, \lambda_s) \in \mathbb{R}^{l_s}_+ \times \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \mathbb{R}_+ \mid A^s x^s + G^s u^s + H^s v^s \geq b^s + \lambda_s d^s\},$$

где неотрицательные вектор-стобцы $x^s = (x^s_i)_{i=1}^{l_s}$, $u^s = (u^s_j)_{j=1}^n$, $v^s = (v^s_j)_{j=1}^n$ указывают объёмы производства, вывоза и ввоза, соответственно, а число $\lambda_s \in \mathbb{R}_+$ - степень достижения целей регионального развития для $s \in R$ (как обычно, символом $\mathbb{R}$ обозначается множество вещественных чисел, а неравенство для векторов понимается в обычном покомпонентном смысле: $x \geq y \Leftrightarrow x_i \geq y_i,\ i = 1, \ldots, m$ для любых $x = (x_1, \ldots, x_m)$ и $y = (y_1, \ldots, y_m)$ из $\mathbb{R}^m$). Элементы множества Z_s будем называть *планами* региона s.

Для оценки качества планов $z^s \in Z_s$ в дальнейшем используются функции t_s, сопоставляющие каждому вектору $z^s = (x^s, u^s, v^s, \lambda_s)$ его последнюю компоненту λ_s:

$$t_s(z^s) = t_s(x^s, u^s, v^s, \lambda_s) := \lambda_s, \quad (x^s, u^s, v^s, \lambda_s) \in Z_s, \quad s \in R$$

(другими словами, отображения $t_s : Z_s \to \mathbb{R}$ - целевые функции участников $s \in R$, характеризующие степень достижения целей их регионального развития).

Положим $Z_{\mathcal{M}} := \prod_{s \in R} Z_s$ и через $Z_{\mathcal{M}}(R)$ обозначим совокупность *сбалансированных планов* модели $\mathcal{M}$:

$$Z_{\mathcal{M}}(R) = \{(x^s, u^s, v^s, \lambda_s)_{s \in R} \in Z_{\mathcal{M}} \mid \sum_{s \in R} u^s \geq \sum_{s \in R} v^s\}.$$

Важную роль в дальнейшем играют так называемые *строго автаркические планы*, под которыми понимаются элементы множеств

$$\hat{Z}_{\mathcal{M}}(s) := \{(x^s, u^s, v^s, \lambda_s) \in Z_s \mid u^s \gg v^s\}, \quad s \in R$$

(как обычно, сокращение $x \gg y$ для векторов $x, y \in \mathbb{R}^m$ означает выполнение строгих неравенств $x_i > y_i,\ i = 1, \ldots, m$). Регион $s \in R$ будем называть *строго автаркическим*, если $\hat{Z}_{\mathcal{M}}(s) \neq \emptyset$.

Помимо строгой автаркичности при стоимостной характеризации коалиционно стабильных планов полезен следующий аналог ненасыщаемости индивидуальных предпочтений. Для каждого $s \in R$ положим $\widetilde{Z}_s := \mathrm{Pr}_{Z_s} Z_{\mathcal{M}}(R)$. Регион $s \in R$ будем называть *ненасыщенным,* если для него выполняется неравенство

$$\sup_{z^s \in Z_s} t_s(z^s) > \sup_{z^s \in \widetilde{Z}_s} t_s(z^s).$$

2 Равновесие Вальраса и нечеткое ядро

Следуя [2], введем одно из основных понятий работы - определение вальрасовского равновесия в модели межрегионального взаимодействия $\mathcal{M}$.

Определение 1. *Будем говорить, что план $\bar{z} = (\bar{x}^s, \bar{u}^s, \bar{v}^s, \bar{\lambda}_s)_{s \in R} \in Z_{\mathcal{M}}(R)$ является вальрасовским равновесием модели $\mathcal{M}$, если существует ненулевой вектор цен $\bar{p} \in \mathbb{R}^n_+$ такой, что $\bar{p} \cdot \bar{u}^s \geq \bar{p} \cdot \bar{v}^s$ для всех $s \in R$, и при этом для любых $s \in R$ и $z^s = (x^s, u^s, v^s, \lambda_s) \in Z_s$ справедлива импликация $\lambda_s > \bar{\lambda}_s \Rightarrow \bar{p} \cdot u^s < \bar{p} \cdot v^s$ (как обычно, выражение $x \cdot y$ означает скалярное произведение векторов x и y).*

Совокупность вальрасовских равновесий модели $\mathcal{M}$ будем обозначать через $W(\mathcal{M})$ (элементы множества $W(\mathcal{M})$ будем называть также *вальрасовскими планами* модели $\mathcal{M}$).

Для изложения основной теоремы эквивалентности приведем формулировку еще одного фундаментального понятия оптимальности в рассматриваемом классе моделей - понятия нечеткого ядра модели $\mathcal{M}$. Для формализации понятия нечеткого блокирования напомним [1,2], что элементы множества

$$\sigma_F := \{\ f = (f_1, \ldots, f_r) \mid f \neq 0,\ f_s \in [0,1],\ s \in R\}$$

называются *нечеткими коалициями.* Величина компоненты f_s нечеткой коалиции f указывает степень участия региона $s \in R$ в координации усилий "большой коалиции" R. Через $R(f)$ будем обозначать носитель нечёткой коалиции $f = (f_1, \ldots, f_r)$, определяемый равенством $R(f) := \{s \in R \mid f_s > 0\}$.

Определение 2. *Будем говорить, что план $\bar{z} = (\bar{z}^s)_{s\in R} \in Z_{\mathcal{M}}(R)$ блокируется нечёткой коалицией $f = (f_1, \ldots, f_r)$, если существуют региональные планы $z^s = (x^s, u^s, v^s, \lambda_s) \in Z_s,\ s \in R(f)$, такие, что* (1) $t_s(z^s) > t_s(\bar{z}^s),\ s \in R(f)$, *и* (2) $\sum_{s\in R(f)} f_s u^s \geq \sum_{s\in R(f)} f_s v^s$.

Совокупность сбалансированных планов модели $\mathcal{M}$, не блокируемых никакой нечёткой коалицией f, будем обозначать через $C_F(\mathcal{M})$ и называть *нечётким ядром модели* $\mathcal{M}$.

Один из главных результатов работы заключается в следующей теореме эквивалентности вальрасовских планов и планов, не блокируемых никакой нечеткой коалицией.

Теорема 1. *Если регионы модели $\mathcal{M}$ строго автаркические и ненасыщенные, то ее нечеткое ядро $C_F(\mathcal{M})$ совпадает с множеством вальрасовских планов $W(\mathcal{M})$.*

3 Нечеткое ядро и равновесия Вальраса и Эджворта

Переходя ко второй теореме эквивалентности, устанавливающей условия совпадения множеств вальрасовских и эджвортовских планов, напомним необходимые определения из [5], где было введено одно из основных понятий оптимальности для рассматриваемого класса моделей - понятие равновесия Эджворта. Несколько модифицируя классическое определение реплики в моделях типа Эрроу-Дебре [6,7], введём сначала определение *дробления (измельчения) модели* $\mathcal{M}$. Зафиксируем $k \geq 1$ и положим

$$R_{[k]} := \{(s,m) \mid s \in R,\ m = 1, \ldots, k\}. \tag{1}$$

Далее, для каждого элемента $(s,m) \in R_{[k]}$ (понимаемого в дальнейшем как "номер" m-ой доли региона $s \in R$ исходной экономики $\mathcal{M}$) введём матрицы A^{sm}, G^{sm}, H^{sm} и векторы b^{sm}, d^{sm}, полагая

$$A^{sm} := A^s,\ G^{sm} := G^s,\ H^{sm} := H^s,\ b^{sm} := \frac{1}{k} b^s,\ d^{sm} := d^s. \tag{2}$$

Определение 3. *Модель многорегиональной экономической системы*

$$\mathcal{M}_{[k]} :=< R_{[k]}, \{A^{sm}, G^{sm}, H^{sm}, b^{sm}, d^{sm}\}_{(s,m)\in R_{[k]}} >, \tag{3}$$

параметры которой определяются соотношениями (1) *и* (2), *будем называть дроблением ранга k (k-дроблением) модели $\mathcal{M}$.*

Отметим, что в силу соотношний между параметрами модели $\mathcal{M}$ и её k-дробления $\mathcal{M}_{[k]}$, для множеств

$$Z_{sm} = \{z^{sm} = (x^{sm}, u^{sm}, v^{sm}, \lambda_{sm}) \geq 0 \mid A^{sm}x^{sm} + G^{sm}u^{sm} + H^{sm}v^{sm} \geq b^{sm} + \lambda_{sm}d^{sm}\}$$

ресурсно-технологических возможностей регионов $(s, m) \in R_{[k]}$ выполняются равенства

$$Z_{sm} = \frac{1}{k}Z_s \quad \text{для каждого } \ s \in R \ \text{ и } m = 1, \ldots, k.$$

Обратим внимание также на то, что как и в исходной модели $\mathcal{M}$, целевые функции $t_{sm} : Z_{sm} \to \mathbb{R}$ регионов (s, m) дробления $\mathcal{M}_{[k]}$ определяются последней компонентой плана z^{sm}, согласно формуле $t_{sm}(x^{sm}, u^{sm}, v^{sm}, \lambda_{sm}) := \lambda_{sm}, \quad (s, m) \in R_{[k]}$.

Как видно из самого определения дробления ранга $k+1$, в $\mathcal{M}_{[k+1]}$ наряду с участниками исходной системы с уменьшенными в $k + 1$ раз ресурсами, в равном количестве из k экземпляров представлены идентичные им доли соответствующих регионов, что усиливает уровень конкуренции в новой системе и позволяет (согласно известной гипотезе Эджворта [1, 6]) более точно оценить равновесные распределения с помощью элементов ядер дроблений. В этом смысле вводимое далее понятие равновесия Эджворта является "предельным"вариантом развития указанной идеи аппроксимации вальрасовских состояний неблокируемыми в моделях типа (3), изображающих рост конкуренции при возрастании ранга дробления регионов.

Избегая громоздкой символики, положим $Z_{[k]} := Z_{\mathcal{M}_{[k]}} = \prod_{(s,m)\in R_{[k]}} Z_{sm}$ и, как и ранее, через $Z(R_{[k]})$ будем обозначать совокупность сбалансированных планов модели $\mathcal{M}_{[k]}$:

$$Z(R_{[k]}) := \{(z^{sm})_{(s,m)\in R_{[k]}} \in Z_{[k]} \mid \sum_{(s,m)\in R_{[k]}} u^{sm} \geq \sum_{(s,m)\in R_{[k]}} v^{sm}\}.$$

Наконец, для $k \geq 1$ черз $z_{[k]}$ обозначим *k-дробление плана* $z = (z^s)_{s\in R} \in Z_{\mathcal{M}}(R)$, представляющее собой сбалансированный (что будет видно из дальнейших пояснений) план модели $\mathcal{M}_{[k]}$, построенный из региональных составляющих z^s плана z по формуле:

$$(z_{[k]})^{sm} := \frac{1}{k}z^s, \quad s \in R, \ m = 1, \ldots, k.$$

Как вытекает непосредственно из определения дробления $z_{[k]}$, задаваемый им план устанавливает одинаковые региональные планы $\frac{1}{k}z^s$ однотипным[1] участникам модели $\mathcal{M}_{[k]}$. Отсюда, в силу сбалансированности z, и получается неравенство

$$\sum_{(s,m)\in R_{[k]}} u^{sm} = \sum_{s\in R} u^s \geq \sum_{s\in R} v^s = \sum_{(s,m)\in R_{[k]}} v^{sm},$$

доказывающее сбалансированность плана $z_{[k]}$ в модели $\mathcal{M}_{[k]}$.

Завершая описание конструкции, лежащей в основе понятия равновесия Эджворта, напомним [2,4] еще определение обычного (то есть порождаемого обычными коалициями $T \subseteq R$) блокирования и обычного ядра в модели $\mathcal{M}$.

Определение 4. *Коалиция $T \subseteq R$ блокирует $\bar{z} = (\bar{z}^s)_{s\in R} \in Z_{\mathcal{M}}(R)$, если у неё имеется такой план $(z^s)_{s\in T} = (x^s, u^s, v^s, \lambda_s)_{s\in T} \in \prod_{s\in T} Z_s$, для которого выполняются соотношения* (1) $t_s(z^s) > t_s(\bar{z}^s), \ s \in T$, *и* (2) $\sum_{s\in T} u^s \geq \sum_{s\in T} v^s$.

[1]Участники (s, m) и (s', m') модели $M_{[k]}$ называются *однотипными,* если $s = s'$

Множество всех планов $z \in Z_{\mathcal{M}}(R)$, которые не блокируются никакой коалицией $T \subseteq R$, будем обозначать через $C(\mathcal{M})$ и называть *ядром модели* $\mathcal{M}$.

Определим, наконец, одно из главных понятий оптимальности, изучаемых в работе - равновесие Эджворта модели межрегионального взаимодействия[2].

Определение 5. *План $\bar{z} \in Z_{\mathcal{M}}(R)$ называется равновесием Эджворта модели $\mathcal{M}$, если все его k-дробления содержатся в ядрах соответствующих дроблений модели $\mathcal{M}$:*

$$\bar{z}_{[k]} \in C(\mathcal{M}_{[k]}) \quad \textit{для каждого} \quad k \geq 1.$$

Совокупность равновесий Эджворта модели $\mathcal{M}$ будем обозначать через $E(\mathcal{M})$.

Учитывая, что для нечеткого ядра и равновесий Эджворта имеет место вложение $C_F(\mathcal{M}) \subseteq E(\mathcal{M})$ (см.[5]), нижеследующая теорема, устанавливающая эквивалентность вальрасовских и эджвортовских равновесий для обширного класса межрегиональных экономических систем, в формальном плане является усилением теоремы 1.

Теорема 2. *Если все регионы системы $\mathcal{M}$ строго автаркические и ненасыщенные, то множество ее вальрасовских равновесий совпадает с множеством равновесий Эджворта*

$$W(\mathcal{M}) = E(\mathcal{M}).$$

4 Некоторые приложения теорем эквивалентности

Помимо отмечавшейся уже значимости теорем эквивалентности в вопросах кооперативной характеризации вальрасовских равновесий (в терминах неблокируемости по отношению к некоторым классам коалиций), большой интерес представляют и возможности получения новых теорем существования равновесий Вальраса, основанные на применении нетрадиционного для этой проблематики аппарата теории кооперативных игр. В частности, как показывает работа [5], исследование условий существования равновесий Эджворта допускает прямое применение традиционных средств равновесного анализа и теории игр, включая известную теорему Скарфа об условиях непустоты обычных ядер [7]. Следовательно, новые условия существования равновесий Вальраса могут быть получены комбинацией теоремы 2 настоящей работы и теоремы 2 из [5].

Переходя к формулировке теоремы существования равновесия Вальраса, опирающейся на кооперативную характеризацию равновесных планов, указанную в теоремах 1 и 2, приведем некоторые условия непустоты множества $E(\mathcal{M})$, найденные в [5]. Напомним [4,5], что однородная составляющая $\mathcal{M}_0$ модели $\mathcal{M}$ отличается от нее только тем, что ресурсно-технологический потенциал каждого из регионов модели $\mathcal{M}_0$ равен нулю: $b^s = 0$ для каждого $s \in R$.

Определение 6. *Будем говорить, что модель $\mathcal{M}$ не имеет "рога изобилия", если множество сбалансированных планов однородной составляющей $\mathcal{M}_0$ этой модели исчерпывается нулевым планом: $Z_{\mathcal{M}_0}(R) = \{0\}$.*

Отметим, что отсутствие "рога изобилия" в смысле определения 6 означает, как и в классических моделях равновесного анализа, что при нулевом экономическом потенциале возможна лишь нулевая хозяйственная активность.

[2]Аналогичное понятие вводилось и изучалось ранее лишь для классической модели Эрроу-Дебре и некоторых её модификаций, указанных, например, в [1,7]

Еще одно используемое в дальнейшем условие представляет собой некоторое ослабление требования ограниченной трансферабельности модели $\mathcal{M}$ из [5], эквивалентное так называемой $\mathcal{P}$-*регулярности* этой модели. Напомним, что план $\bar{z} \in Z_{\mathcal{M}}(R)$ является Парето-оптимальным (слабо Парето-оптимальным), если не существует плана $z \in Z_{\mathcal{M}}(R)$ такого, что $t(z) > t(\bar{z})$ ($t(z) \gg t(\bar{z})$). Здесь используется сокращение $t(z) := (t_1(z^1), \ldots, t_r(z^r))$, а соотношение $t(z) > t(\bar{z})$ означает, как обычно, что $t(z) \geq t(\bar{z})$ и $t(z) \neq t(\bar{z})$.

Определение 7. *Будем говорить, что модель $\mathcal{M}$ является $\mathcal{P}$-регулярной, если $P'(\mathcal{M}) = P(\mathcal{M})$, где $P(\mathcal{M})$ - множество Парето-оптимальных, а $P'(\mathcal{M})$ - слабо Парето-оптимальных планов из $Z_{\mathcal{M}}(R)$.*

Опираясь на теорему 2 и упоминавшийся уже результат из [5], касающийся условий существования равновесия Эджворта модели $\mathcal{M}$, получаем новые условия существования вальрасовских равновесий в многорегиональных экономических системах, не включающие, в отличие от [8], требования ограниченности множеств $Z_s,\ s \in R$.

Теорема 3. *Если $\mathcal{M}$ является $\mathcal{P}$-регулярной и не имеет "рога изобилия", а ее регионы - строго автаркические и ненасыщенные, то в $\mathcal{M}$ существует равновесие Вальраса.*

ЛИТЕРАТУРА

[1]. Vasil'ev V.A. On Edgeworth equilibria for some types of nonclassic markets. Siberian Advances in Mathematics. 1996. V. 6, № 3. P. 96–150.

[2]. Рубинштейн А.Г. Моделирование экономических взаимодействий в территориальных системах. Новосибирск: Наука. Сиб. отд-ние. 1983.

[3]. Гранберг А.Г., Суслов В.И., Суспицын С.А. Многорегиональные системы: экономико-математическое исследование. Новосибирск: Наука. Сиб. Науч. Изд-во. 2007.

[4]. Васильев В.А., Суслов В.И. О неблокируемых состояниях многорегиональных экономических систем. Сибирский журнал индустриальной математики. 2009. Т. XII, № 4(40). С. 23–34.

[5]. Васильев В.А., Суслов В.И. Равновесие Эджворта в одной модели межрегиональных экономических отношений. Сибирский журнал индустриальной математики. 2010. Т. XIII, № 1(41). С. 18–33.

[6]. Гильденбранд В. Ядро и равновесие в большой экономике: Пер. с англ. М.: Наука. 1986.

[7]. Алипрантис К., Браун Д., Бёркеншо О. Существование и оптимальность конкурентного равновесия. М.: Мир. 1995.

[8]. Васильев В.А. О существовании вальрасовского равновесия в модели межрегиональных экономических отношений. Дискретный анализ и исследование операций (в печати).

СИНТЕЗ АРХИТЕКТУР ФУНКЦИОНАЛЬНО-ОРИЕНТИРОВАННЫХ ПРОЦЕССОРОВ НА ОСНОВЕ ОЦЕНОК АППАРАТНОЙ И ВРЕМЕННОЙ СЛОЖНОСТИ ВЫЧИСЛЕНИЙ В БАЗИСЕ СХЕМ ИЗ ФУНКЦИОНАЛЬНЫХ ЭЛЕМЕНТОВ

Лукин Н.А.
Институт машиноведения Уральского отделения РАН
e-mail: nicklookin@mail.ru

Функционально-ориентированные процессоры (ФОП) представляют собой специализированные вычислители, которые эффективно реализуют конкретные классы алгоритмов или даже один алгоритм. Решение проблемы синтеза оптимальных архитектур актуально как с точки зрения развития науки (создание теоретических основ проектирования ФОП), так и со стороны разработки конкретных аппаратно-программных решений (создание САПР архитектур ФОП). В работе рассматривается задача синтеза рациональных либо оптимальных архитектур ФОП, состоящих из двумерных массивов процессорных элементов и реализация которых предполагается целиком в виде одного кристалла СБИС или его фрагмента.

Предлагается новая методология синтеза процессорных архитектур, которая основана на верхних оценках сложности вычислений в базисе схем из функциональных элементов. Основные этапы синтеза архитектуры ФОП, оптимизированной по критерию минимума аппаратной сложности ($L_h \Rightarrow min,\ L_t^* = const$):

- построение максимальной канонической ярусно-параллельной формы графа алгоритма;
- трансформация графа, достижение D_{min}, H_{max}, где D, H – ширина и высота графа соответственно, с помощью эквивалентных dh-преобразований графа, введенных в настоящей работе;
- взвешивание вершин графа (D_{min}, H_{max}) параметрами $\{l_h, l_t\}_{1,\ldots,N}$, где l_{hi}, l_{ti} – верхние оценки аппаратной и временной сложности i-ой вершины графа соответственно. Для алгоритма в целом имеем $L_h = \sum_{i=1}^{N} l_{hi}; L_t = Max(\sum_{j=1}^{M} l_{tj})$, где M – число всех простых ориентированных путей в графе;
- построение определяющего соотношения вида $L_h = f(L_t, \hat{A})$, где f – функция, $\hat{A}$ - непустое множество информативных параметров алгоритма (например, разрядность переменных, емкость таблицы и т.п.);
- итерационные dh-преобразования графа $\Rightarrow [(D+d), (H-h)] \Rightarrow [(L_h + \Delta L_h), (L_t - \Delta L_t)$. Итерации прекращаются при достижении $L_t \le L_t^*$;
- реализация каждого из преобразований вершин графа с помощью процессорного элемента. Архитектура, оптимизированная по критерию аппаратурных затрат, синтезирована.

Данная методика описывается на примерах синтеза трех конкретных архитектур ФОП, предназначенных для вычисления в реальном времени математических функций, для интегрирования одного класса дифференциальных уравнений и для распознавания образов.

Параллельный алгоритм для решения нестационарной задачи мантийных течений

Лазарева Г.Г., Корнеев В.Д., Бабичев А.В.
Лазарева Г.Г.
ИВМиМГ СО РАН, Новосибирск, Россия
lazareva@ssd.sscc.ru
Корнеев В.Д.
ИВМиМГ СО РАН, Новосибирск, Россия
korneev@ssd.sscc.ru
Бабичев А.В.
ОИГГМ СО РАН, Новосибирск, Россия
babichev@igm.nsc.ru

Разработана параллельная версия программы для моделирования течений в мантии Земли. Нестационарная модель мантийных течений описывает сжимаемую среду с сильно изменяющимися реологическими и транспортными свойствами и основана на решении системы уравнений Навье-Стокса. Численная модель содержит как явные, так и неявные конечно-разностные схемы, реализованные векторными прогонками. При разработке параллельного алгоритма важно знать потенциальные возможности ускорения вычислений и накладные расходы, связанные организацией параллельной работы потоков, их взаимодействием и синхронизацией. Кроме того, важно знать показатели эффективности работы параллельного алгоритма на вычислительной системе, позволяющий сравнивать его с другими параллельными алгоритмами, а так же оценить возможности параллельной реализации на вычислительных системах с общей памятью с большим количеством процессорных ядер. Для численной реализации рассматриваемого алгоритма в качестве вычислительной системы с общим полем памяти используется вычислительный узел кластера ИВЦ НГУ. Вычислительные узлы кластера состоят из двух 4-х ядерных процессоров Intel Xeon 5355, работающих на частоте 2.66 ГГц, и с 16-ю ГБайт общей оперативной памяти.

В работе проведен детальный анализ параллельного алгоритма, позволяющего получать близкое к линейному ускорение, не смотря на использование векторной прогонки. Приведены основные характеристики параллельного алгоритма. Получены общее время счета и счета внутреннего цикла, связанного с вычислениями скоростей мантийных течений вдоль координат вычислительного пространства, при разной заданной точности вычислений и разных размерах сеточного пространства. Показана зависимость ускорения и эффективности от размеров сеточного вычислительного пространства при слабой зависимости этих параметров от заданной точности вычислений. В результате расчетов на многопроцессорной вычислительной системе получены параметры процесса плавления и диапиризма в нижней коре, определена структура течения всплывающей гранитной магмы.

НИР выполнена при финансовой поддержке интеграционного проекта СО РАН № 2, в рамках реализации ФЦП "Научные и научно-педагогические кадры инновационной России" на 2009 - 2013 годы.

РЕДУКЦИЯ МОДЕЛИ CD95-ИНДУЦИРУЕМОГО АПОПТОЗА

Е. О. Кутумова

Конструкторско-технологический институт вычислительной техники СО РАН,
e.o.kutumova@gmail.com

В работе рассмотрена математическая модель CD95-индуцируемой регуляции программируемой смерти клетки (апоптоза), разработанная M. Bentele и соавт. [1]. На основании методов редукции [2] при помощи программного комплекса BioUML (http://www.biouml.org) выявлены быстрые и медленные реакции модели, исследованы короткоживущие промежуточные комплексы, произведено упрощение реакций на основе кинетики Михаелиса-Ментен. Помимо этого, исследованы процессы, биохимические механизмы которых до конца не изучены. В модели таким процессам соответствуют “черные ящики” с заданными математическими функциями. В результате упрощения модели количество ее элементов уменьшилось в два раза, а число реакций сократилось с 78 до 40. При этом результаты симуляции модели для оставшихся элементов практически не изменились. Примененный подход позволяет оптимизировать объемные модели и упростить вычисления, что является важным моментом при создании на их основе более сложных комплексных моделей биологических систем.

ЛИТЕРАТУРА

[1]. Bentele M., Lavrik I., Ulrich M., Stößer S., Heermann D. W., Kalthoff H., Krammer P. H., Eils R. Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. J Cell Biol. 2004. V. 166, №6. P. 839-851.

[2]. Gorban A. N., Radulescu O., Zinovyev A. Y. Asymptotology of Chemical Reaction Networks. Chem Eng Sci. 2009. V. 65. P. 2310-2324.

МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОЙ ДИНАМИКИ СООБЩЕСТВ СВЯЗАННЫХ МИГРАЦИЕЙ[1]

М.П. Кулаков

Институт комплексного анализа региональных проблем ДВО РАН, Биробиджан

e-mail: k_matvey@mail.ru

В данной работе предлагается и исследуется модель решетчатого типа, предназначенная для описания пространственной динамики локальных сообществ, между которыми возможно небольшое миграционное взаимодействие. Локальная динамика описывается отображением на основе обобщенной зависимости запас-пополнение Рикера, учитывающим возникающие между видами сообщества взаимодействия разных типов: конкуренция, отношения типа хищник-жертва, симбиоз и т.п. Изучаются механизмы формирования нерегулярного и периодического поведения в отдельных локальных сообществах, а так же влияние миграционного взаимодействия между ними на формирование пространственной динамики. Предлагается методика определения условий, приводящих к полной синхронизации динамики всех связанных очагов либо отдельных групп сообществ.

[1] Исследования проведены при финансовой поддержке РФФИ (проекты №09-04-00146-а, №11-01-98512-р_восток_а) и ДВО РАН (конкурсные проекты №09-I-П15-01, №09-III-А-09-498)

Построение таблиц стилей текстовых произведений с использованием алгоритмов классификация на основе деревьев решений

А.И. Кубарев
Томский государственный университет, kubarev_ai@mail.ru
В.В. Поддубный
vvpoddubny@gmail.com
О.Г. Шевелев
oshevelyov@gmail.com
О.В. Кукушкина
Московский государственный университет им. М.В. Ломоносова, kukush@orc.ru

(Работа поддержана грантом РФФИ № 11-07-00776-а)

С использованием алгоритмов классификации текстов на основе деревьев решений предложен алгоритм оптимального по информационному критерию последовательного бинарного разбиения n-мерного признакового пространства стилей текстовых произведений на 2^n непересекающихся n-мерных интервалов, образующих таблицу стилей текстов, определяющую «стилевой портрет» (профили стилей) корпуса текстов. Предполагается, что признаковое пространство является частотным, т.е. образовано частотами появления в текстах наборов служебных слов, словосочетаний, биграмм и т.п. Алгоритм реализован программно в системе «СтилеАнализатор», предназначенной для комплексного исследования корпусов текстов различных типов. На материале различных корпусов текстов проведено сравнительное исследование качества классификации текстов по авторам, жанрам, стилям и другим характеристикам текстов по алгоритмам деревьев решений и таблицам стилей текстов. Получаемые при обучении алгоритма профили стилей текстов могут быть использованы для идентификации стиля предъявляемого текста неизвестного автора, что позволяет, в частности, определять наиболее вероятное авторство текста.

Ключевые слова: классификация текстов, деревья решений, таблицы стилей, профили стилей, идентификация текстов

О системах четверок Штейнера малого ранга, вложимых в расширенные совершенные коды

Д. И. Ковалевская [1]
Институт математики им. С. Л. Соболева
daryik@rambler.ru

Ф. И. Соловьева [2]
Институт математики им. С. Л. Соболева
Новосибирский Государственный Университет
sol@math.nsc.ru

Аннотация

Известно, что кодовые слова веса 4 расширенного совершенного кода, содержащего нулевой вектор, образуют систему четверок Штейнера. В работе показано, что система четверок Штейнера порядка 2^t, полученная методом свитчингов из Хэмминговой системы четверок Штейнера, вложима в расширенный совершенный код, построенный методом свитчингов $ijkl$-компонент из двоичного расширенного кода Хэмминга.

1 Введение

Пусть F^n – n-мерное метрическое пространство над полем Галуа $GF(2)$ с метрикой Хэмминга. *Двоичным кодом* длины n называется произвольное подмножество метрического пространства F^n. *Параметры* произвольного двоичного кода C из F^n обозначаются через $(n, |C|, d)$, где n – длина кодовых слов (элементов кода), $|C|$ – мощность кода, d – кодовое расстояние (т.е. минимальное хэммингово расстояние между кодовыми словами). Двоичный код C длины n с расстоянием $d = 2d' + 1$ называется *совершенным*, если для любого $x \in F^n$ существует единственный x' из C, такой, что расстояние Хэмминга $d(x, x') = (d-1)/2$. Известно (см [1]), что нетривиальный двоичный совершенный код, исправляющий одну ошибку (упоминаемый далее как совершенный), существует тогда и только тогда, когда $n = 2^t - 1$ для некоторого t.

Если V – множество, состоящее из v элементов, то t-(v, k, λ)-*схемой* называется такое размещение v различных элементов по блокам, что каждый блок содержит точно k различных элементов, любое t-элементное подмножество из V появляется точно в λ блоках. *Системой троек Штейнера порядка* v (обозначим ее $STS(v)$) и *системой четверок Штейнера порядка* v (обозначаемой как $SQS(v)$) называются 2-$(v, 3, 1)$ и 3-$(v, 4, 1)$ схемы соответственно. Известно (см. [2]), что система четверок $SQS(m)$ существует тогда и только тогда, когда $m \equiv 2, 4 (mod\ 6)$.

Пусть $\bar{C}$ – *расширенный совершенный код* длины $n = 2^t$, полученный из совершенного кода C длины $2^t - 1$ добавлением общей проверки на четность. Далее

[1] Работа выполнена при поддержке гранта Президента РФ для молодых российских ученых (грант МК-1700.2011.1)

[2] Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (проект 10-01-00424-а).

будем рассматривать только расширенные совершенные коды, содержащие нулевой вектор. Известно (см [1]), что кодовые слова веса 3 в коде C образуют систему троек Штейнера $STS(2^t-1)$, а кодовые слова веса 4 в коде $\bar{C}$ образуют систему четверок Штейнера $SQS(2^t)$.

Говорят (см. [3]), что код $C' = (C \backslash M) \cup M'$ получен *свитчингом* множества M на множество M' в двоичном коде C, если код C' имеет те же параметры, что и C. Такое множество M называется *i-компонентой* кода C, если $M' = M \oplus e_i$ для некоторого $i \in \{1, 2, \ldots, n\}$, где e_i – вектор веса 1 с единицей в i-ой координатной позиции. *Рангом* кода C называется размерность линейного подпространства пространства F^n, образованного векторами из C.

Существует множество открытых вопросов, касающихся систем троек и систем четверок Штейнера. В том числе – проблема классификации систем троек и четверок Штейнера, проблема вложимости произвольной системы троек (четверок) Штейнера в совершенный (расширенный совершенный) код. В работе [4] П. Р. Остергардом и О. Поттоненом доказано, что только 33 из 80 неизоморфных систем троек Штейнера порядка 15 являются вложимыми в совершенный код, и только 15590 из 1054163 систем четверок Штейнера порядка 16 вложимы в расширенный совершенный код.

Известно, что ранг системы троек Штейнера $STS(2^t-1)$ больше либо равен $2^t - t - 1$, ранга кода Хэмминга, см. [5]. Ранг системы четверок Штейнера $SQS(2^t)$ также не меньше, чем $2^t - t - 1$ (см. [6]). В статье [7] В. Д. Тончевым найдена нижняя оценка числа различных систем троек Штейнера порядка 2^t-1 ранга $2^t - t$, что на 1 превышает минимально возможный ранг. Тем же автором в работе см. [8] приведена аналогичная формула для числа различных систем четверок Штейнера порядка 2^t ранга $2^t - t$. В работе [9] В. А. Зиновьевым и Д. В. Зиновьевым приведен метод построения систем четверок Штейнера порядка $N = 2^t$ произвольного ранга, а также найдена нижняя оценка числа всех различных систем четверок Штейнера $SQS(2^t)$ ранга не более $2^t - t + 1$, построенных с помощью этого метода из системы четверок Штейнера порядка $N/4$ ранга $2^t - t - 1$:

$$6^{N(N-4)/2^5} \cdot (3^3 \cdot 2^{11})^{N(N-4)(N-8)/(3 \cdot 2^9)}. \qquad (1)$$

В [10] показано, что известный класс систем троек Штейнера порядка $2^t - 1$, полученный свитчингами Пэш-конфигураций, вложим в класс совершенных кодов, построенных методом ijk-компонент, и приведена нижняя оценка числа систем троек Штейнера порядка $2^t - 1$ ранга не более $2^t - t + 1$.

В данной работе для полноты изложения приводится конструкция системы четверок Штейнера $SQS(N)$ (модификация конструкции Ханани), построенная из произвольной системы четверок Штейнера $SQS(m)$, $m = 2^t$, $N = 4m$. Разбиение такой $SQS(N)$ на определенного вида компоненты соответствует некоторому разбиению на $ijkl$-компоненты расширенного совершенного кода, и такая система четверок Штейнера вложима в расширенный совершенный код, построенный известным методом $ijkl$-компонент. Приведена нижняя оценка числа различных систем четверок Штейнера $SQS(N)$ ранга не более $N - log(N) + 1$, вложимых в расширенный совершенный код.

2 Системы четверок Штейнера SQS(4m), вложимые в расширенный совершенный код

В данном разделе рассмотрим конструкцию системы четверок Штейнера $SQS(4m)$ порядка $4m$, которая строится из системы четверок Штейнера $SQS(m)$ порядка m с помощью конструкции Ханани (см. [11]). Из конструкции будет следовать, что некоторые такие $SQS(4m)$ вложимы в расширенный совершенный код.

Пусть $M = \{1, 2, 3, \ldots, m\}$ – множество мощности m, на котором задана произвольная система четверок Штейнера $SQS(m)$, $m \equiv 2, 4(mod\ 6)$. Для построения системы четверок Штейнера порядка $4m$ на множестве элементов $M \bigcup \{i_1, \ldots, i_m, j_1, \ldots, j_m, k_1, \ldots, k_m\}$, в дальнейшем упоминаемой как Q_N, где $N = 4m$, рассмотрим следующую таблицу:

$T_M =$

1	2	3	...	m
i_1	i_2	i_3	...	i_m
j_1	j_2	j_3	...	j_m
k_1	k_2	k_3	...	k_m

Сначала рассмотрим, как строится $SQS(4m)$ в самом частном случае, когда $m = 4$. Пусть, например, $SQS(4) = \{(a, b, c, d)\}$. В этом случае система четверок Штейнера $SQS(4m)$ имеет порядок 16, и таблица T_M принимает вид

a	b	c	d
i_a	i_b	i_c	i_d
j_a	j_b	j_c	j_d
k_a	k_b	k_c	k_d

Обозначим полученную таблицу через T_{abcd}. Для построения $SQS(16)$ воспользуемся конструкцией Ханани. Включим в нее все строки и столбцы таблицы T_{abcd}, четверки конструкции Ханани вида

(2)

примененные к каждой паре строк и столбцов, и все миноры второго порядка, т.е. элементы вида

$$(h_1, h_2, t_{h_1}, t_{h_2}), (t_{1_{h_1}}, t_{1_{h_2}}, t_{2_{h_1}}, t_{2_{h_2}}), t, t_1, t_2 \in \{i, j, k\}, h, h_1, h_2 \in \{a, b, c\}. \quad (3)$$

Кроме того, в $SQS(16)$ добавим всевозможные сочетания элементов, находящихся в разных строках и столбцах таблицы T_{abcd}, т.е. множество вида

$$\{(a, i_b, j_c, k_d), (a, i_b, j_d, k_c), (a, i_c, j_b, k_d), (a, i_d, j_b, k_c), (a, i_c, j_d, k_b), (a, i_d, j_c, k_b),$$
$$(b, i_a, j_c, k_d), (b, i_a, j_d, k_c), (b, i_c, j_a, k_d), (b, i_d, j_a, k_c), (b, i_c, j_d, k_a), (b, i_d, j_c, k_a),$$
$$(c, i_a, j_b, k_d), (c, i_a, j_d, k_b), (c, i_b, j_a, k_d), (c, i_d, j_a, k_b), (c, i_b, j_d, k_a), (c, i_d, j_b, k_a),$$
$$(d, i_a, j_b, k_c), (d, i_a, j_c, k_b), (d, i_b, j_a, k_c), (d, i_c, j_a, k_b), (d, i_b, j_c, k_a), (d, i_c, j_b, k_a)\} \quad (4)$$

Общее количество получившихся четверок равно $4+4+2\cdot 6\cdot C_4^2+6\cdot C_4^2+4\cdot 6=$ $=140$, что совпадает с количеством четверок в $SQS(16)$. Из построения $SQS(16)$ видно, что каждая неупорядоченная тройка элементов содержится в единственном блоке. Таким образом, построена $SQS(16)$ из $SQS(4)$.

Пусть m – произвольное, такое что существует $SQS(m), m \equiv 2,4(mod\ 6)$. Тогда в Q_N включим все столбцы, а также для любой пары столбцов – все миноры вида (3) и блоки конструкции Ханани (2). Таким образом, получим $m+6\cdot C_m^2+6\cdot C_m^2 = m+6m(m-1)$ четверок. Далее, для любой четверки (a,b,c,d) из $SQS(m)$ рассмотрим подматрицу T_{abcd}. Для этой матрицы в Q_N включим (a,b,c,d), оставшиеся строки, блоки конструкции Ханани вида (2), примененные к каждой паре строк, а также четверки вида (4).

Нетрудно видеть, что каждой матрице вида T_{abcd} соответствуют $1+3+6\cdot C_4^2+$ $4\cdot 6=64$ четверок в Q_N. Количество таблиц совпадает с количеством четверок в $SQS(m)$ и равно $m(m-1)(m-2)/24$. Следовательно, общее количество четверок в конструкции равно $m+6m(m-1)+64\cdot m(m-1)\cdot(m-2)/24=4m(4m-1)\cdot$ $\cdot(4m-2)/24=|Q_N|$.

Из построения множества четверок легко видеть, что каждая неупорядоченная тройка элементов встречается ровно в одной четверке. Таким образом, построена система четверок Штейнера Q_N порядка $N=4m$ из системы четверок Штейнера $SQS(m)$ порядка m.

Теорема 1. *Из произвольной системы четверок Штейнера порядка m можно построить систему четверок Штейнера порядка $4m$.*

Если M – i-компонента совершенного кода C длины n, содержащего нулевой вектор, то $\bar{M}$ будем называть *il-компонентой расширенного совершенного кода* $\bar{C}$ длины $N=n+1$. Если же $\bar{M}$ является il, jl и kl-компонентой расширенного совершенного кода $\bar{C}$, то $\bar{M}$ называется *$ijkl$-компонентой расширенного совершенного кода* $\bar{C}$.

Известно, что справедлива следующая теорема.

Теорема 2. *(См. [12]) Всякий расширенный совершенный двоичный код Хэмминга длины N можно представить в виде объединения непересекающихся $ijkl$-компонент R_{ijkl}^t, каждая из которых, в свою очередь, может быть представлена как объединение непересекающихся il-компонент R_{il}^{pt}:*
$\mathcal{H}^N=\bigcup_{t=1}^{N_2} R_{ijkl}^t=\bigcup_{t=1}^{N_2}\bigcup_{p=1}^{N_1} R_{il}^{pt}$, *где* $N_1=2^{(N-4)/4}$, $N_2=2^{(N+4)/4-logN}$.

Эти разбиения позволяют делать свитчинги расширенного кода Хэмминга и в результате получить широкий класс расширенных совершенных кодов.

Систему четверок Штейнера порядка N, соответствующую двоичному расширенному коду Хэмминга $\mathcal{H}^N$, будем называть *Хэмминговой системой четверок Штейнера* $SQS(\mathcal{H}^N)$.

Множество Q называется *il-компонентой Хэмминговой системы четверок Штейнера* $SQS(\mathcal{H}^N)$, если Q – подмножество векторов веса 4 из il-компоненты расширенного кода Хэмминга $\mathcal{H}^N$ длины N. Если же il-компонента системы четверок Штейнера Q является также jl-компонентой и kl-компонентой, то Q называется *$ijkl$-компонентой Хэмминговой системы четверок Штейнера* $SQS(\mathcal{H}^N)$.

Утверждение 1. *Хэммингова система четверок Штейнера $SQS(\mathcal{H}^N)$ представима в виде объединения подмножеств $1 + N(N-4)(N-8)/(3 \cdot 2^9)$ непересекающихся $ijkl$-компонент, каждая из которых, в свою очередь, является объединением подмножеств $N/4 + (N-4)(N-8)/2^5$ либо 8 непересекающихся il-компонент.*

Рассмотрим систему четверок Штейнера $SQS(m)$ и построенную из нее по Теореме 1 систему четверок Штейнера Q_N. Заметим, что если $SQS(m)$ является Хэммингговой системой четверок Штейнера $SQS(\mathcal{H}^m)$, то система Q_N является Хэммингговой системой четверок Штейнера $SQS(\mathcal{H}^N)$. Справедлива

Теорема 3. *Система четверок Штейнера, полученная методом свитчингов $ijkl$-компонент из системы $SQS(\mathcal{H}^N)$, является вложимой в совершенный код, полученный из кода Хэмминга $\mathcal{H}^N$ методом свитчингов $ijkl$-компонент.*

Приведем нижнюю оценку числа различных систем четверок Штейнера порядка N ранга не более $N - logN + 1$, вложимых в расширенный совершенный код, построенный методом свитчингов $ijkl$-компонент. Число таких систем четверок Штейнера обозначим через $R(N)$. Справедлива

Теорема 4. *Число $R(N)$ различных систем четверок Штейнера $SQS(N)$ порядка N ранга не более $N - logN + 1$, вложимых в расширенный совершенный код, не меньше, чем*

$$(3^2 \cdot 2^8 - 8)^{N(N-4)(N-8)/(3\cdot 2^9)} \cdot 2^{N(N-4)/2^5} \cdot \tfrac{N(N-1)(N-2)}{2^3} \cdot D(N/4),$$

где $D(N/4)$ – число различных Хэммингговых систем четверок Штейнера порядка $N/4$.

Полученная оценка меньше (1), и вопрос о том, все ли такие системы четверок Штейнера из [9] являются вложимыми в расширенные совершенные коды, остается открытым.

Список литературы

[1] *Ф. Дж. Мак-Вильямс, Н. Дж. Слоэн.* "Теория кодов, исправляющих ошибки": Пер. с англ. М.: Связь. 1979. 744 с.

[2] *М. Холл.* "Комбинаторика": Пер. с англ. М.: Мир. 1970. 424 с.

[3] *Ф. И. Соловьева.* "Введение в теорию кодирования": Учеб. пособие // Новосиб. гос. ун-т. Новосибирск, 2006. 124 с.

[4] *P. R. Östergård, O. Pottonen* The Perfect Binary One-Error-Correcting Codes of Length 15: Part 1 – Classification. // IEEE Trans. Inform. Theory. 2009. № 55. P. 4657–4660.

[5] *J. Doyen, X. Hubaut, M. Vandensavel* Ranks of incidence matrices of Steiner triple systems. // Math. 1978. S. Z. № 163. P. 251–259.

[6] *L. Teirlinck* On projective and affine hyperplanes. // J Combin. Theory. 1980. S. A. № 28. P. 290–306.

[7] *V. D. Tonchev.* A mass formula for Steiner triple systems $STS(2^n-1)$ of 2-rank 2^n-n. // Journal of Combin. Theory. 2001. Series A **95**, P. 197–208.

[8] *V. D. Tonchev.* A formula for the number of Steiner quadruple systems on 2^n points of 2-rank 2^n-n. // Journal of Combin. Designs. 2003. № 11. P. 260–274.

[9] *В. А. Зиновьев, Д. В. Зиновьев.* О разрешимости систем Штейнера $S(v=2^m,4,3)$ ранга $r \leq v-m+1$ над F^2. // Пробл. передачи информ. 2007. Т. 43. Вып. 1. С. 39–55.

[10] *Е. С. Глухих.* Вложимость систем троек Штейнера в совершенные коды. // Магистерская диссертация. Новосибирск. 2005.

[11] *H. Hanani.* The Existence and Construction of Balanced Incomplete Block Designs. // Ann. Math. Statist. 1961. V. 32, № 2 P. 361–386.

[12] *С. В. Августинович, Ф. И. Соловьева.* Построение совершенных двоичных кодов последовательными сдвигами $\tilde{\alpha}$-компонент. // Пробл. передачи информ. 1997. Т. 33. Вып. 3. С. 15–21.

НОВЫЙ МЕТОД МОДЕЛЬНОГО ФУНКЦИОНАЛА ЭЛЕКТРОННОЙ ПЛОТНОСТИ ДЛЯ ОПИСАНИЯ ЭЛЕКТРОННОГО СТРОЕНИЯ СВЕРХБОЛЬШИХ НАНО- И БИОНАНОСТРУКТУР

Поддержано РФФИ, проект 11-07-00470

Н.А. Аникин, М.Б Кузьминский

Н.А. Аникин
Учреждение Российской Академии Наук
«Институт органической химии им. Н.Д. Зелинского РАН»
Москва
e-mail: nikan@swf.chem.ac.ru

М.Б Кузьминский
e-mail: kus@free.net

Аннотация

Предложен новый метод квантовохимического моделирования больших и сверхбольших (тысячи и десятки тысяч атомов) наноструктур и биомолекул. Метод использует модельный одноэлектронный гамильтониан, аппроксимирующий гамильтониан современных методов функционалов плотности (DFT), и при приемлемой точности позволяет достигнуть на порядки более высокого быстродействия, характерного для более грубых полуэмпирических методах типа PM3. Для прямого нахождения разреженной матрицы плотности быстрыми линейно масштабируемыми методами планируется применение графических ускорителей (GPU).

Одной из актуальных задач современного этапа развития нано- и бионанотехнологий является сокращение времени разработок, для чего нужно быстро получать адекватную информацию об электронных и квантовых свойствах и пространственном строении вновь создаваемых и уже известных нано- и бионаноструктурах.

Широко применяемый метод функционала плотности (DFT) обычно дает приемлемую точность, но не обеспечивает необходимое быстродействие для больших молекулярных систем. Авторами предложены два концептуально сходных варианта быстрого и достаточно точного квантовохимического метода, способных дать информацию об электронном строении, и проводить массовое компьютерное моделирование крупных нано- и бионаноструктур.

Принципиальное отличие предлагаемого подхода от других более быстрых, чем DFT, приближенных методов (SCC-DFTB и полуэмпирические методы типа PM3) состоит в воспроизведении (с необходимой точностью) эталонного одноэлектронного гамильтониана ($\hat{H}$) метода DFT на структурах, типичных для нано- и биомолекулярных систем. Полученный

приближенный метод построения $\hat{H}$ затем применяется в расчетах нано- и биообъектов. Заметим, что $\hat{H}$ однозначно определяет волновую функцию, матрицу плотности и все свойства системы, а вычисление $\hat{H}$ является одной из самых трудоемких стадий расчета электронной структуры.

Указанные выше приближенные методы содержат параметры, которые подбираются сопоставлением лишь отдельных рассчитываемых из $\hat{H}$ свойств с экспериментальными и эталонными расчетными данными. По сравнению с этими методами наш подход позволяет кардинально улучшить точность и надежность (получение корректных результатов для молекулярных систем любого типа) описания электронного строения.

Методика

Гамильтониан $\hat{H}$ состоит из нескольких физически различных членов-операторов: оператор кинетической энергии $\hat{T}$, кулоновского члена $\hat{Q}$, и обменно-корреляционного члена (потенциала) $\hat{X}$: $\hat{H}=\hat{T}+\hat{Q}+\hat{X}$

Эти операторы имеют различную физическую природу и различную зависимость от геометрической и электронной структуры данной молекулы и составляющих ее атомов и связей. Матрицы оператора $\hat{T}$ и интегралов перекрывания $\hat{S}$ вычисляются точно. Аппроксимации применяются или к самому оператору $\hat{X}$ (это оператор умножения на функцию), или в формулах для вычисления матричных элементов (оператора $\hat{Q}$) в базисе, в котором осуществляется расчет.

При определении используемых в аппроксимации $\hat{H}$ параметров для его максимально аккуратного воспроизведения мы используем эталонную информацию обо всех матричных элементах $\hat{H}$ (типично - тысячи элементов на молекулу) для набора в несколько тысяч эталонных молекулярных структур.

Матричные элементы $\hat{H}$ (мы используем стандартные дираковские обозначения)

$H_{ij}=\langle \chi_i|\hat{H}|\chi_j\rangle$

рассчитываются в базисе χ, где χ - или атомные орбитали (АО) φ, или ортонормированные АО (ОАО) ϕ:

$$\phi_i=\sum_j^{AO}\Theta_{ij}\times\varphi_j,$$

где $\Theta=S^{-0.5}$ матрица, используемая для симметричной ортогонализации по Левдину, $S_{ij}=\langle\varphi_i|\varphi_j\rangle$ - симметричная положительно определенная матрица интегралов перекрывания ($S_{ii}=1$); очевидно, $\langle\phi_i|\phi_j\rangle$ является единичной матрицей.

Более простые и легче вычисляемые матричные элементы T_{ij} оператора кинетической энергии $\hat{T}$ и матрицы перекрывания S_{ij} от АО рассчитываются напрямую, без введения подгоняемых параметров. Для ускорения расчетов T_{ij}, S_{ij} приближаются сплайнами как функции межатомных расстояний для всех пар атомов и АО. Мы используем тот факт, что T_{ij} и

S_{ij} зависят только от (фиксированных нами заранее) АО каждого типа атома, и от расстояния между центрами АО соответствующих атомов.

Матричные элементы Q_{ij} кулоновского оператора $\hat{Q}$ содержат одноэлектронные вклады притяжения ко всем ядрам, и двухэлектронные вклады электрон-электронного отталкивания (зависящие от электронной плотности). Они рассчитываются с применением аппроксимаций для всех пар АО $\varphi_i(\boldsymbol{r})\times\varphi_j(\boldsymbol{r})$ (и вычисляемой из них функции общей электронной плотности молекулы) в виде линейной комбинации заранее подобранных вспомогательных функций плотности , определяемых для каждого типа атома, и центрированных на тех же ядрах, где и АО i, j. Коэффициенты этой линейной комбинации подбираются из условия минимума среднеквадратичной погрешности. При расчетах молекул по нашему методу (аналогично T_{ij} и S_{ij}) эти коэффициенты также задаются предварительно подобранными сплайнами от межатомного расстояния. Для ускорения расчетов Q_{ij} гигантских молекул используется известный метод быстрых мультиполей (Fast Multipole Method [1]).

Остающийся обменно-корреляционный потенциал $\hat{X}$ приближаются суммой атомных вкладов $\hat{X}_m$, содержащих заранее подбираемые эмпирические параметры EP_{mn} атома “m” (“n” – номер параметра внутри этого атома):

$$\hat{X}=\sum_m \hat{X}_m ,$$

$$\hat{X}_m(\boldsymbol{r})\approx\sum_n EP_{mn}\times\exp\left(-\zeta_{mn}\times(\boldsymbol{r}-\boldsymbol{R_m})^2\right) ,$$

где ζ_{mn} – фиксированные показатели экспонент, $\boldsymbol{R_m}$ – вектор координат атома “m”.

Эти EP_{mn} характеристичны для каждого типа атома и подбираются по МНК для минимизации отклонения наших приближенных матричных элементов X_{ij} от эталонных (вычисляемых по DFT в заданном базисе для тысяч типичных эталонных молекул) H_{ij}^{ET} за вычетом не содержащих параметров T_{ij} и Q_{ij}:

$$X_{ij}=H_{ij}^{ET}-T_{ij}-Q_{ij}=\langle\varphi_i|\hat{X}|\varphi_j\rangle\approx\sum_m\sum_n EP_{mn}\times\langle\varphi_i|\exp\left(-\zeta_{mn}\times(\boldsymbol{r}-\boldsymbol{R_m})^2\right)|\varphi_j\rangle$$

В настоящее время написаны программы для определения параметров метода и осуществляется их отладка. Собственно расчеты электронной структуры нано- и биомолекул по предложенному методу будут проводиться с применением этих параметров.

Схема расчета предполагает определение из $\hat{H}$ идемпотентной матрицы плотности $\boldsymbol{P}$ в базисе АО φ, и соответственно - электронного распределения в молекуле (определение $\boldsymbol{P}$ также является лимитирующей время расчета). Это делается либо через диагонализацию H_{ij} в ортонормированном базисе ОАО ϕ (стандартный путь), либо напрямую получая блочно-разреженную матрицу $\boldsymbol{P}$ из (также блочно-разреженной) матрицы $\boldsymbol{H}$ с использованием авторского метода LocalSCF [2] и усовершенствованной нами версии метода PDM (Purification of Density Marix, "очистка" матрицы $\boldsymbol{P}$ от неидемпотентности) [3]), отработанному авторами для полуэмпирических расчетов [4]. Для гигантских молекул это – самый быстрый путь, он дает линейное масштабирование времени расчета с размером молекулы. Далее из полученной МП $\boldsymbol{P}$

строится новый гамильтониан, и т.д. до самосогласования $\boldsymbol{H}$ и $\boldsymbol{P}$ (кроме описанной ниже неитерационной версии нашего метода).

Для малополярных систем (фуллерены, нанотрубки и т.д.) возможно проведение еще более быстрых неитерационных расчетов в рамках предлагаемой упрощенной (более быстрой) версии нашего метода построения $\hat{H}$. Она также эффективна для построения начального приближения для итерационного метода. Основные особенности неитерационного метода описаны ниже.

а) Применяется базис только валентных ОАО, ортогонализованных ко всем остовным АО; аналогичный базис применяется в полуэмпирических методах. Это позволяет уменьшить размер базиса и соответственно ускорить расчеты.

б) Матричные элементы гамильтониана H_{ij} в этом базисе рассчитываются в функциональном виде, сходном с описанным выше, но без зависящего от электронного распределения кулоновского оператора. $H_{ij}-T_{ij}$ аппроксимируются в базисе ОАО ϕ (причем для $\phi_i(\boldsymbol{r})\times\phi_j(\boldsymbol{r})$ производится только двухатомная ортогонализация АО на центрах φ_i и φ_j).

в) Для аппроксимации двухцентровых (двухатомных) блоков элементов H_{ij} (с ОАО $\phi_i(\boldsymbol{r})\times\phi_j(\boldsymbol{r})$) определяются иные параметры EP_{mn}, чем для одноатомных блоков.

Получаемые H_{ij} не зависят от электронного распределения, соответственно не нужно их итерационное уточнение. Для повышения точности этого метода предлагается учет влияния ближайших атомов с использованием индексов химического окружения атомов (предложенных и апробированных ранее авторами для других целей, проект РФФИ 07-07-12090-офи).

Предлагаемые нами методы (итерационный и неитерационный) позволят физически адекватно воспроизводить гамильтониан конкретных функционалов DFT (B3LYP, PBE и др.), а не подгонять лишь отдельные результаты расчетов на его базе - теплоты образования, равновесные геометрии, дипольные моменты и др., что характерно для различных параметризаций типа AM1, PM3, PM6 и т.д. Во всех этих методах одно и то же значение какого-либо свойства может быть получено из множества различных гамильтонианов (включая физически некорректные), что принципиально затрудняет однозначный и физически разумный выбор калибруемых параметров. Рассчитанные свойства "в среднем" оказываются вполне сопоставимы с DFT, но сам приближенный гамильтониан фактически "не сопоставим" с гамильтонианом "эталонного" метода DFT.

Но именно точность воспроизведения гамильтониана важна для надежности (получения корректных результатов для молекулярных систем любого типа) и физической обоснованности новых методов.

В целом предлагаемый итерационный метод имеет некоторое сходство с получившим распространение в последние годы методом SCC-DFTB (Self-Consistent-Charge Density-Functional-Tight-Binding) [5], а наш неитерационный метод - с неитерационным NCC-DFTB (Non-Consistent-Charge Density-Functional-Tight-Binding) [5]. NCC-DFTB более прост и менее точен, чем SCC-DFTB для систем с неоднородным зарядовым распределением и сильными электрическими полями, но требует меньше расчетного времени.

Отчасти приведенная выше критика полуэмпирических методов справедлива также для

SCC-DFTB и NCC-DFTB, где вообще нет прямой связи матричных элементов гамильтониана SCC-DFTB с гамильтонианом DFT (тем более - с гамильтонианом конкретного функционала DFT), и отсутствует зависимость некулоновской части H_{ij} от валентного окружения данного атома или пары атомов - центров AO i, j.

Нами предлагается физически более обоснованный математический вид предлагаемых аппроксимаций. В итерационном методе хотя и приближенно, но напрямую рассчитываются кулоновские взаимодействия (в SCC-DFTB их описание является очень грубым приближением). В неитерационном методе матричные элементы гамильтониана H_{ij} зависят от близлежащих окружающих атомов, а не фиксированы для данной пары атомов, как в NCC-DFTB.

Оба предложенных нами метода хорошо совместимы друг с другом и с эталонным гамильтонаном DFT в случае их одновременного применения для расчета разных частей одной молекулы. Это позволяет, как и в методах типа QM/QM, малую "активную" часть молекулы считать более точным методом, а остающуюся большую часть молекулы - более быстрым приближенным методом. «Стыковка» на границе раздела этих частей у нас точнее и надежнее, т.к. оба наших метода - это приближения к одному и тому же гамильтониану DFT.

Предлагаемые нами более простые, чем DFT, методы требуют гораздо меньших затрат вычислительных ресурсов и будут хорошим начальным приближением для итерационных методов более высокого уровня (от нашего неитерационного метода вплоть до точного DFT) как в рамках всей молекулы, так и (для QM/QM) ее части.

Реализация предлагаемых нами методов осуществляется в среде Linux/x86-64 и предполагает возможность применения нового поколения графических ускорителей (GPU) с архитектурой NVIDIA Fermi и средств PGI CUDA Fortran. В частности, для указанного выше метода прямого расчета разреженной матрицы плотности из гамильтониана используются подпрограммы умножения матриц из оптимизированных библиотек BLAS – CUBLAS и MAGMA (http://icl.cs.utk.edu/magma).

Для исследования накладных расходов из-за передачи данных в GPU и обратно с целью последующей оптимизации программ для GPU NVIDIA Tesla C2050 (установленной в слоте PCI-E x16 2.0, пиковая производительность платы на двойной точности 515 GFLOPS) проведен ряд тестов производительности, в т.ч. с использованием средств SHOC [6]. Использован двухпроцессорный сервер с Intel Xeon E5520/2.3 ГГц, работающий с OpenSuSE Linux версии 11.1. Полученные данные показали величину задержки на уровне 10 мкс (что соответствует более ранним данным [7]), и предельную пропускную способность на уровне 6 Гбайт/с.

Результаты для умножения матриц показали уменьшение производительности из-за расходов на обмен данными с GPU в 1.2 раза (для больших размерностей). Измеренная на модифицированных для GPU тестах STREAM пропускная способность графической памяти GPU составила до 130 Гбайт/с. Эти данные могут быть полезны для оптимизации и других расчетных программ, использующих GPU данного типа.

Заключение

Предложены два концептуально сходных, принципиально новых метода (итерационный и неитерационный) для исследований электронной структуры больших нано- и биомолекул, основанные на аппроксимации гамильтониана распространенных методов DFT. Методы позволят повысить уровень квантовохимического молекулярного моделирования благодаря сочетанию высокого быстродействия (близкого к наиболее быстрым, но недостаточно точным полуэмпрических методам), и точности, близкой к методам DFT (имеющим, однако, неприемлемо большое время расчета). Для проведения расчетов молекул по предложенным методам предполагается применение линейно масштабируемых с размером молекулы прямых процедур построения матрицы плотности, используя графические ускорители с архитектурой CUDA.

ЛИТЕРАТУРА

[1]. Greengard L., Rokhlin V.
J. Comput. Phys. 1987. v. 73, № 1. p. 325-348.
[2]. Anikin N.A., Anisimov V.M., Bugaenko V.L., Bobrikov V.V., Andreev A.M.
J. Chem. Phys. 2004. v. 121, № 3. p. 1266-1270.
[3]. Daniels A., Scusseria G.
J. Chem. Phys. 1999. v. 110, № 3. p. 1321-1328.
[4]. Чернецов А.М., Шамаева О.Ю., Кузьминский М.Б.
в сб. “Высокопроизводительные параллельные вычисления на кластерных системах. Материалы восьмого международного научно-практического семинара и всероссийской молодежной школы”, Казань. 2008. с. 347-349.
[5]. Zheng G., Lundberg M., Jakowski1 J., Vreven T., Frisch M.J., Morokuma K.
Int. J. Quant. Chem. 2009. v. 109, № 9. p. 1841–1854.
[6]. Danalis A., Marin G., McCurdy C., Meredith J., Roth P., Spafford K., Tipparaju V., Vetter J.
Proc. 3rd Workshop on General-Purpose Computation on Graphics Processors. 2010. p. 63-74.
[7]. Lawlor O. S. Message Passing for GPGPU Clusters: cudaMPI
IEEE Cluster PPAC Workshop. 2009.

КАЧЕСТВЕННЫЙ АНАЛИЗ И РЕШЕНИЯ-УТКИ НА МЕДЛЕННОМ МНОГООБРАЗИИ В ЗАДАЧЕ ХИМИЧЕСКОЙ КИНЕТИКИ

Л. И. Кононенко
Институт математики им. С. Л. Соболева СО РАН
e-mail: larakon2@gmail.ru
Е. П. Волокитин
Институт математики им. С. Л. Соболева СО РАН
e-mail: volok@math.nsc.ru *

Рассматривается пятимерная система дифференциальных уравнений, описывающая реакцию окисления CO на иридии. Система изучается с использованием техники интегральных многообразий, позволяющей вместо качественного анализа полной системы ограничиться исследованием строения многообразия медленных движений и качественным анализом систем меньшей размерности на листах этого многообразия.

Сначала пятимерная система сводится к трехмерным, рассматриваемым на четырех листах медленной поверхности, а затем в одной из трехмерных систем выделяются две быстрые и одна медленная переменные и исследуется медленное уравнение на медленном параметризованном многообразии (кривой), заданном двумя алгебраическими уравнениями.

Далее изучается геометрия медленной кривой с использованием найденной параметризации. Проведен качественный анализ трехмерной системы на одном из устойчивых листов; при этом была показана возможность существования решений-уток в окрестности самопересечения медленной кривой.

*Работа поддержана РФФИ (проект 09-01-0070) и Сибирским отделением РАН (междисциплинарный интеграционный проект №107, 119)

МОДЕЛИРОВАНИЕ ВНУТРИВИДОВОЙ И МЕЖВИДОВОЙ КОНКУРЕНЦИИ В РАЗНОВОЗРАСТНЫХ ДРЕВЕСНЫХ СООБЩЕСТВАХ[1]

А.Н. Колобов

Институт комплексного анализа региональных проблем ДВО РАН, Биробиджан

e-mail: alex_0201@mail.ru

В работе приводится описание индивидуально-ориентированной модели динамики древесных сообществ. Рассматриваются результаты моделирования внутривидовой и межвидовой конкуренции в разновозрастных древостоях.

Результаты моделирования внутривидовой конкуренции показали, что в ходе формировании структуры древостоя даже при однородных внешних условиях происходят процессы хаотической самоорганизации, приводящие к образованию сложно структурированных неоднородных (пятнистых) пространственных распределений растительных сообществ. Сложный мозаичный характер распределения деревьев приводит к установлению в системе квазистационарного режима динамики суммарных показателей древостоя, обеспечивая устойчивость сообщества. Возникновение такой неоднородности можно объяснить только причинами внутренней пространственной конкуренции за ресурсы жизнедеятельности, в частности конкуренцией за свет.

Проведенные вычислительные эксперименты по моделированию межвидовых взаимодействий теневыносливых и светолюбивых пород, показали, что рассматриваемые темнохвойные виды со временем вытесняют светолюбивую березу, которая к концу второго поколения практически полностью исчезает. Моделирование динамики взаимодействия темнохвойных видов ели, пихты и кедра демонстрирует их устойчивое совместное сосуществование. Взаимодействие темнохвойных и светолюбивых пород определяется также способом распространения семян. В случае, когда семена всех видов равномерно распределены по участку, возможен сценарий присутствия светолюбивых пород, возобновляемых в окнах.

[1] Исследования выполнены при финансовой поддержке РФФИ (проект № 09-04-00146-а) и ДВО РАН (проект №09-III-А-09-498, 11-III-В-06-130)

АНАЛИЗ ПРОПУСКНОЙ СПОСОБНОСТИ ТРАНСПОРТНОГО ПРОТОКОЛА В НАГРУЖЕННОМ ТРАКТЕ ПЕРЕДАЧИ ДАННЫХ

В.В. Кокшенев,
Томский государственный университет
vkokshenev@intant.ru
С.П.Сущенко
ТНЦ СО РАН
ssp@inf.tsu.ru

Аннотация

В работе предложена модель виртуального соединения транспортного протокола, проложенного в многозвенном тракте передачи данных с очередями пакетов в транзитных узлах. Протокольная процедура управления переносом трафика формализована в виде марковизированного процесса с дискретным временем, введением дополнительной переменной, задающей размер «мешающей» очереди перед потоком исследуемого виртуального соединения. Предложенная математическая модель позволяет оценить влияние распределения размера очереди в транзитных узлах, ширины окна, длительности тайм-аута ожидания подтверждения и характеристик каналов связи на быстродействие виртуального соединения.

В современных сетях, осуществляющих передачу данных, голоса и видео через единую инфраструктуру, существенно повышаются требования к наличию доступной полосы пропускания и, как следствие, к повышению эффективности ее использования. В работе предложена математическая модель протокола транспортного уровня в многозвенном нагруженном тракте передачи данных, формализующая режимы селективного и группового отказа [1] двумерной цепью Маркова с дискретным временем. В модели учтена также природа волоконно-оптических линий связи, имеющих, возможно, существенно разные показатели надежности в прямом и обратном направлениях передачи. Данная ситуация нередко провоцируется неправильной инсталляцией оптических коннекторов. Несмотря на то, что большинство современных стеков протоколов TCP/IP имеют, согласно RFC 793, групповую процедуру отказа, в последнее время начинают появляться реализации с функцией селективного подтверждения для TCP, описанной в RFC 2883. Примером может служить стек протоколов TCP/IP NG операционных систем Windows Server Longhorn и Windows Vista. В связи с данным фактом возрастает актуальность исследования показателей пропускной способности селективного режима отказа.

Рассмотрим обмен данных между узлами, соединенными однозвенным трактом передачи данных. Предположим, что выполняются следующие допущения:

1. Узлы тракта соединены дуплексными каналами связи, имеющими одинаковыми пропускные способности в обоих направлениях.
2. Длина тракта передачи данных, выраженная в количестве участков переприема, равна *D*.
3. Заданы вероятности потери кадра в канале связи для прямого и обратного направления передачи каждого участка переприема.
4. Передача данных на каждом участке переприема выполняется в соответствии с алгоритмом решающей обратной связи.
5. Время обработки пакетов в узлах тракта одинаково.
6. В каждом узле-отправителе имеется неограниченный поток пакетов для передачи, и обмен выполняется информационными пакетами одинаковой длины.

7. Подтверждения получателя о корректности приема принимаемых данных переносятся в информационных пакетах встречного потока.
8. Потерь кадров из-за отсутствия буферной памяти в узлах тракта не происходит.
9. Задана функция вероятностей $b(n)$ того, что поток анализируемого соединения в транзитном узле встретит очередь размера $n \le N$, где N – максимальный размер очереди, определяемый емкостью буферных пулов транзитных узлов.

Будем считать, что размер окна управляющего протокола определяется величиной W, а $S > W$ – задает длительность тайм-аута ожидания подтверждения. После передачи очередного сегмента, протокол копирует его в очередь переданных, но не подтвержденных данных и запускает тайм-аут. Как только размер очереди становится равным W, управляющий протокол приостанавливает передачу в ожидании получения квитанции или истечения тайм-аута ожидания подтверждения S. При получении подтверждения, из очереди удаляются сегменты, дошедшие до адресата без искажений. При истечении тайм-аута S соответствующий сегмент передается повторно, и тайм-аут запускается вновь. Будем называть тактом время t, необходимое для вывода кадра в линию. Такт определяется суммой времени вывода пакета в линию, времени распространения сигнала в канале связи и времени обработки пакета принимающим узлом. Селективный и групповой методы управления надежной передачей отличаются способом формировании квитанций. Селективная процедура отказа явно указывает все полученные сегменты данных, а групповая – сообщает номер первого неполученного сегмента. Таким образом, селективная управляющая процедура передает повторно лишь потерянные и искаженные кадры, в то время как групповая – все сегменты, начиная с первого недошедшего до адресата, что потенциально влечет ненужные повторные передачи и загрузку канала связи «холостым» трафиком.

Динамика очереди переданных, но не подтвержденных сегментов на узле-отправителе для различных режимов функционирования управляющего протокола может быть описана двумерной цепью Маркова с дискретным временем и числом состояний по одному измерению, равным длительности сквозного тайм-аута S, а по другому – увеличенной на единицу максимальной длине очереди: $N+1$. Очевидно, что длительность тайм-аута должна быть достаточной для того, чтобы пакет с сегментом данных по прямому каналу достиг адресата и подтверждение получателя по обратному каналу было принято отправителем потока. Отсюда следует, что размер тайм-аута, выраженный в длительностях тактов t, должен быть не меньше суммы двойной длины пути и размера встреченной очереди в транзитном узле $S \ge 2D+n$. С учетом возможных повторных передач информационных пакетов основного (прямого) потока и пакетов с подтверждениями во встречном потоке из-за искажений в отдельных звеньях тракта размер тайм-аута S целесообразно выбирать с «запасом» на повторные передачи. Пропускная способность виртуального соединения $Z(W,S)$, управляемого транспортным протоколом, определяется как отношение среднего объема данных, передаваемых между двумя последовательными получениями квитанций, к среднему времени получения квитанции [2]. Вклад в быстродействие виртуального соединения дают те состояния цепи Маркова, для которых возможно получение квитанции.

Анализ математической модели показывает, что индекс быстродействия виртуального соединения, нормированного на единицу, при абсолютно надежных каналах связи в отдельных звеньях тракта является функцией длины очереди, размера окна и длины тракта передачи данных. Данный показатель определяет потенциально достижимую скорость передачи данных в виртуальном канале. Для размера окна W, изменяющегося от 1 до $2D-1$, длительности тайм-аута $S \ge 2D+n$ и детерминированной очереди длины $n>0$ ($b(n)=1$) нормированная скорость составит $Z(W,S)=1/(2D-W+2+n)$. Если на пути основного потока мешающего трафика нет ($b(0)=1$), то быстродействие определится соотношением $Z(W,S)=1/(2D-W+1)$. При ширине окна $W \ge 2D$ и детерминированной очереди размера n

скорость обмена данными в виртуальном соединении определится соотношением $Z(W,S)=2/(2+n)$, $n\geq 0$. Численные исследования быстродействия виртуального канала в различных режимах повторной передачи показывают, что скорость передачи в канале монотонно растет с увеличением размера окна и выходит в режим насыщения при $w>2D$. Из анализа следует, что быстродействие при селективном режиме отказа значительно превышает скорость обмена при групповом отказе, однако с ростом надежности каналов преимущество селективной процедуры отказа в сравнении с групповой снижается. Полученные результаты позволяют утверждать, что при заданном размере окна показатель пропускной способности виртуального соединения возрастает с увеличением длительности тайм-аута ожидания подтверждения и практически достигает теоретического предела при насыщении по протокольному параметру W для значений S, превосходящих ширину окна на 2-5 тактов длительности t.

ЛИТЕРАТУРА

1.*ОлиферВ.Г., Олифер Н.А.* Компьютерные сети. Принципы, технологии, протоколы. – Спб.: Питер, 2006. – 958 с.

2.*Сущенко С. П.* Аналитические модели асинхронных процедур управления звеном передачи данных // Автоматика и вычислительная техника. – 1988. – № 2. – С. 32-40.

Определение местоположения высокой точности для одночастотных приёмников спутниковой навигации с использованием инерциальных датчиков

Д.Ю.Першин, А.С.Щербаков
Институт систем информатики СО РАН

Аннотация

Местоположение приемника в системе спутниковой навигации вычисляется пересечением сфер от разных спутников. Радиусы сфер определяются неточно из-за ионосферной и тропосферной задержек, релятивистского эффекта, эффекта приема отраженного сигнала и др. факторов. Без каких-либо коррекций точность определения местоположения составляет 30-50метров.

В данной работе рассматривается модель, в которой учитывается максимальное количество дополнительных данных для коррекции ошибок. Для исправления ионосферной задержки используются данные IONEX о состоянии ионосферы. Все используемые данные хранятся на международных серверах.

Также используются данные с инерциальных MEMS-датчиков (акселерометр, гироскоп, магнитометр). Помимо улучшения точности, эти данные позволяют определять местоположение, не имея данных со спутников (потеря сигнала и др.) в течении некоторого времени.

В работе описан кинематический метод высокой точности, для минимизации ошибок приборов спутниковой навигации и инерциальных датчиков использован фильтр Калмана.

Ключевые слова: *спутниковая навигация, ГЛОНАСС, GPS, IMU, инерциальные MEMS-датчики, фильтр Калмана, одночастотный приёмник, точные орбиты и часы спутников*

Теория мультимножеств и ее применение для групповой классификации и упорядочения многопризнаковых объектов

А.Б.Петровский
Институт системного анализа РАН
e-mail: pab@isa.ru

В принятии решений, распознавании изображений, искусственном интеллекте, теории формальных языков, математическом программировании, теории сетей Петри и других областях имеется достаточно много задач, где анализируемые объекты характеризуются многими разнородными признаками, которые могут быть и количественными, и качественными, и смешанными. При этом сами объекты могут существовать в нескольких экземплярах, имеющих, в частности, и противоречивые описания, которые должны рассматриваться и анализироваться как единое целое, а свертка значений признаков или невозможна, или математически некорректна. Примерами подобных задач служат классификация и упорядочение объектов, оцененных несколькими экспертами по многим качественным критериям, распознавание графических символов, обработка текстовых документов.

Такие многопризнаковые объекты можно представить как мультимножества или множества с повторяющимися элементами. Возможность многократного вхождения элементов в мультимножество создает новое качество, которое отличает мультимножество от обычного «ординарного» множества и порождает существенно большее, чем у множеств, разнообразие видов и особенностей мультимножеств. В работе представлены основные понятия теории мультимножеств и теории метрических пространств мультимножеств [1], приведены примеры их практического использования для решения задач групповой классификации и упорядочения многопризнаковых объектов [2,3].

Работа поддержана программами фундаментальных исследований президиума РАН «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация» и ОНИТ РАН «Информационные технологии и методы анализа сложных систем», РФФИ (проект 11-07-00398), РГНФ (проект 11-02-00131).

ЛИТЕРАТУРА

[1]. Петровский А.Б. Пространства множеств и мультимножеств. – М.: Едиториал УРСС, 2003.

[2]. Петровский А.Б. Многокритериальное принятие решений по противоречивым данным: подход теории мультимножеств. // Информационные технологии и вычислительные системы, 2004, № 2, С.56-66.

[3]. Петровский А.Б. Методы групповой классификации многопризнаковых объектов (части 1 и 2). // Искусственный интеллект и принятие решений, 2009, № 3, С.3-14; № 4 С.3-14.

О. И. Ильин

ОБ ОДНОЙ ИГРОВОЙ ЗАДАЧЕ УПРАВЛЕНИЯ В МОДЕЛИ ПОПУЛЯЦИИ

Камчатский научно-исследовательский институт рыбного хозяйства и океанографии
683000, Россия, Камчатский край, г. Петропавловск-Камчатский, ул. Набережная, 18
e-mail: ilin@kamniro-avacha.kamchatka.ru

Аннотация

В статье рассматривается игровая задача управления эксплуатируемой биологической популяцией. Динамика распределения численности (биомассы) популяции описывается моделью типа МакКендрика - Фон Фёрстера. Популяция подвергается управлению со стороны некоторого набора управляющих, каждый из которых имеет свой критерий качества. Управляющие независимы между собой и имеют информацию о состоянии популяции. Используемым принципом оптимальности является оптимальное по Нэшу управление.

Популяции морских рыб, как правило, облавливаются несколькими промысловыми флотилиями (управляющими). Каждый из управляющих стремится максимизировать доход от промысла, в этой связи между ними могут возникать конфликтные ситуации. Вопрос о принятии решений в подобных ситуациях приводит к игровым задачам управления.

Рассматривается задача управления, в которой популяция подвергается управлению со стороны n управляющих, каждый из которых имеет орудие лова, отличное от других. Динамика распределения численности (биомассы) популяции описывается моделью типа МакКендрика - Фон Фёрстера [1]:

$$\begin{gathered}\frac{\partial x}{\partial t}+\frac{\partial x}{\partial \tau}=rx-\sum_{i=1}^{n}q_iu_ix,\ (t,\tau)\in G=[0;\,T]\times[0;\,\hat{\tau}],\\ x(0,\tau)=x_0(\tau),\ \tau\in[0;\,\hat{\tau}];\\ x(t,0)=\Phi(s),\ s=s(t)=\int\limits_0^{\hat{\tau}}\alpha(t,\tau)x(t,\tau)\,d\tau,\ t\in[0;\,T],\end{gathered}\tag{1}$$

где t — время, τ — возраст, $x(t,\tau)$ — плотность численности (биомассы) популяции, $u_i(t)$ — интенсивность промыслового усилия i-го управляющего (число промысловых операций в единицу времени), $q_i(\tau)$ — функция, характеризующая избирательную способность орудий лова i-го управляющего. В задаче (1) функция $r(t,\tau)$ описывает процессы роста и смертности, $\alpha(t,\tau)\geq 0$ характеризует долю особей, участвующих в размножении, интеграл s представляет собой биомассу (численность) родителей. Функция $\Phi(s)\geq 0$ во внутреннекраевом условии описывает процессы рождаемости и выживаемости особей на ранних стадиях жизненного цикла. Функция $x_0(\tau)$ задает начальное распределение численности (биомассы).

Каждый из управляющих имеет свой критерий полезности $F_k(x,u_k)$ и выбирает свою управляющую функцию из условия максимизации этого критерия:

$$F_k(x,u_k)\to\sup_{u_k}.\tag{2}$$

Все управляющие независимы между собой и имеют информацию о состоянии эксплуатируемой популяции. Критерии оптимизации имеют интегральный вид:

$$F_k(x,u)=\int\limits_0^T\int\limits_0^\tau c_kq_ku_kx\,dtd\tau$$

и выражают доход от промыслового изъятия особей всех возрастов за время T (выигрыш) для каждого из управляющих. Функции $c_k = c_k(t,\tau)$ выражают стоимость одной единицы численности (биомассы) особей возраста τ, выловленных k-м игроком в момент времени t.

Максимизация функционалов F_k достигается выбором вектора управляющих воздействий $u(t)$ при заданных ограничениях:

$$0 \le u_k(t) \le U_k, k = 1, \dots, n. \tag{3}$$

Задача (1)-(2)-(3) относится к теории игр. Подобные задачи применительно к моделям однородных популяций рассматривались в ряде работ [2,3,4].

Оптимальное решение $(\hat{x}, \hat{u})$ задачи

$$\begin{cases} F_1(\hat{x}, \hat{u}_1) \ge F_1(\hat{x}, u_1), \\ \dots, \\ F_n(\hat{x}, \hat{u}_n) \ge F_n(\hat{x}, u_n) \end{cases}$$

при ограничениях (1), (3) будем называть решением оптимальным по Нэшу.

С использованием принципа максимума Понтрягина [5] для задачи оптимальной эксплуатаци популяции [6] устанавливаются необходимые условия оптимальности по Нэшу.

Утверждение 1. *Если $(\hat{x}, \hat{u})$ — оптимальное по Нэшу управление в задаче (1)-(2)-(3), то $\hat{u}_k(t)$, $k = 1, \dots, n$ определяются из условий*

$$\hat{u}_k(t) = \begin{cases} 0, \textit{если} \int\limits_0^{\hat{\tau}} q_k(\tau)(\psi_k(t,\tau) - c_k(t,\tau))x(t,\tau)\,d\tau > 0, \\ \textit{произвольное } w_k(t), 0 \le w_k \le U_k, \textit{если} \int\limits_0^{\hat{\tau}} q_k(\psi_k - c_k)x d\tau = 0, \\ U_k, \textit{если} \int\limits_0^{\hat{\tau}} q_k(\tau)(\psi_k(t,\tau) - c_k(t,\tau))x(t,\tau)\,d\tau < 0, \end{cases}$$

где x — решение системы (1), а ψ_k — решение краевой задачи

$$\frac{\partial \psi_k}{\partial t} + \frac{\partial \psi_k}{\partial \tau} = -r\psi_k + \psi_k \sum_{i=1}^{n} q_i u_i - c_k q_k u_k - \psi_k(t,0)\frac{d\Phi}{ds}\alpha;$$
$$\psi_k(T,\tau) = 0,\ \psi_k(t,\hat{\tau}) = 0,$$

соответствующие управлению $\hat{u} = (\hat{u}_1, \dots, \hat{u}_n)$.

ЛИТЕРАТУРА

1. von Foerster H. In: The Kinetics of cellular prolifiration. N.Y., 1959.

2. Захаров В.В., Петросян А.А. Теоретико-игровой подход к проблеме окружающей среды //Вестник Ленингр. ун-та. вып. 1. №1. 1981. с.26-32.

3. Мазалов В.В., Реттиева А.Н. Об одной задаче управления биоресурсами // Обозрение прикладной и промышленной математики. 2002. Т. 9. Вып. 2. С.293-306.

4. Реттиева А.Н. Методы динамических игр в задаче управления биоресурсами: подход с введением заповедной зоны: Автореф. дисс. ... канд. физ.-мат. наук. Петрозаводск: Ин-т прикладных математических исследований Карельского научного центра РАН, 2004. 23 с.

5. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М: Гос. изд-во физ.-мат. лит., 1961. 391 с.

6. Ильин О.И. Об оптимальной эксплуатации популяции рыб с возрастной структурой // Сибирский журнал индустриальной математики, т. 10, № 3 (31), 2007. С. 43–57.

Решения HP в области построения гибридных кластеров

С.В. Стрижак, А. Б. Юдин

В докладе рассматриваются вопросы построения HPC гибридных кластеров на базе типовых решений HP (S6500 шасси, сервера серии HP SL390) и графических процессоров Nvidia Tesla. Приведены примеры построения HP кластеров для заказчиков в области образования, науки и промышленности, как для простейшей конфигурации, так и для конфигурации уровня петафлопной производительности. Обсуждаются особенности нагрузочного тестирования подобных решений. На примере программы ‘Университетский кластер’ (www.unicluster.ru) и действующей web-лаборатории (www.unihub.ru) обсуждаются возможности использования открытых пакетов (Salome, OpenFOAM, Paraview) для решения задач механики сплошной среды при проведении облачных вычислений.

HP solutions for HPC hybrid clusters

S.V.Strijhak, A.B.Yudin

This presentation deals with the questions of HPC hybrid clusters construction on the basis of HP typical solutions (S6500 chassis, servers of series HP SL390) and GPU Nvidia Tesla. In the report there are listed examples of HPC solutions for customers in the area of education, science and industry both for the elementary configuration, and for a level configuration of PetaFlops performance are resulted. Peculiarities of benchmarking and testing of similar decisions are discussed. The possibilities of using Open Source software (Salome, OpenFOAM, Paraview) to solve the problems for continuous mechanics at carrying out in cloud computing are discussed on the example of a program ‘ University Cluster ’ (www.unicluster.ru) and operating web-laboratory (www.unihub.ru).

О разбиениях n-куба на совершенные двоичные коды [1]

Г. К. Гуськов

Институт математики им. Соболева СО РАН,
e-mail: m1lesnsk@gmail.com

Ф. И Соловьёва

Институт математики им. Соболева СО РАН,
Новосибирский Государственный Университет,
e-mail: sol@math.nsc.ru

В настоящей работе приводится свитчинговый метод построения разбиений n-куба на совершенные двоичные коды и отвечающая ему нижняя оценка числа различных разбиений для любых допустимых $n \geq 7$.

[1]Настоящая работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (проект 10-01-00424-а).

УПРАВЛЯЮЩИЕ КИБЕРНЕТИЧЕСКИЕ СИСТЕМЫ ЛЯПУНОВА – ЯБЛОНСКОГО И РЕШЕНИЕ ПАРАДОКСА НЕЗАВИСИМОСТИ

М.А. Федоткин

Нижегородский государственный университет им. Н.И. Лобачевского
Национальный исследовательский университет
E-mail: fma5@rambler.ru

Аннотация

На сегодня классические методы создания моделей эволюционных систем с управлением исчерпываются тремя подходами [1, 2]. Первый подход основан на построении модели с позиции «черного ящика». Второй подход предполагает представление управляющей системы в виде составляющих её элементов. При этом каждый такой элемент является более простой управляющей системой, для изучения которой можно применять первый подход. В работах [1, 2] отмечаются недостатки указанных подходов. В связи с этим в этой работе доказывается целесообразность использования кибернетического подхода [3] для решения хорошо известного парадокса независимости в задаче Мостеллера [4, 5].

Задача Мостеллера. Построение вероятностных моделей реальных эволюционных систем с управлением является первоначальной задачей их изучения. С целью иллюстрации построения вероятностной модели статистически устойчивых экспериментов с управлением с позиции «черного ящика» рассмотрим формулировку и решение задачи Мостеллера [4].

Чтобы наградить сына, играющего в теннис, отец обещает ему приз, если он выиграет подряд, по крайней мере, две теннисные партии против своего отца и мастера. Мастер играет лучше отца. Сын выигрывает у мастера с вероятностью p и у отца — с вероятностью $q > p$. Выигрыши сына независимы в совокупности. Сын имеет право выбрать один из двух вариантов очерёдности игры: 1) мастер затем отец и снова мастер; 2) отец затем мастер и снова отец. Какой вариант поведения следует выбрать сыну с точки зрения получения приза с наибольшей вероятностью?

Для проверки адекватности основных выводов различных подходов решения задачи Мостеллера необходимо большое число раз реализовать теннисную игру сына с отцом и мастером. Этот эксперимент не представляется легко осуществимым. Поэтому вместо теннисной игры рассмотрим аналогичный эксперимент, который автор работы имел возможность многократно проводить при приёме экзаменов по теории вероятностей на факультетах Нижегородского госуниверситета им. Н.И. Лобачевского.

Задача об экзаменах. Студенты сдают экзамен по теории вероятностей, каждый из которых должен ответить по билету на три вопроса раздельно профессору и ассистенту. Студенты знают, что вопросы первый и третий являются теоретическими, а второй заключается в решении практической задачи. Студент сдаст экзамен, если он два раза подряд положительно отвечает на вопросы. Профессору студент отвечает положительно на любой вопрос с вероятностью p и ассистенту — с вероятностью $q > p$. Студенту предлагается выбрать один из двух вариантов поведения: 1) сначала студент отвечает профессору, затем — ассистенту и снова — профессору; 2) на первый вопрос студент отвечает ассистенту, на второй вопрос — профессору и на последний вопрос — ассистенту.

Рассмотрим классическое решение задачи Мостеллера на основе принципа «черного ящика», которое было предложено в работах [4, 5]. Одним из основных предметов теории вероятностей является построение адекватной вероятностной модели $(\Omega, \Im, \mathbf{P}_r(\cdot))$ статистически устойчивого случайного эксперимента W_r с управлением r из некоторого множества $\mathcal{R}$. Произвольный элемент ω из множества Ω определяет с помощью некоторого языка описание так называемого элементарного исхода эксперимента W_r. Подмножество $\Im$ априори заданного множества всех допустимых исходов является σ-алгеброй и содержит все наблюдаемые исходы $D \subset \Omega$ эксперимента W_r. Наконец, вероятностная функция $\mathbf{P}_r(\cdot)$: $\Im \to [0,1]$ задается на σ-алгебре $\Im$ и зависит от выбранного управления $r \in \mathcal{R}$. Если используется под-

ход «черного ящика», то W_r представляется в виде объекта управления E_r и системы управления C_r. Поэтому эксперимент W_r есть упорядоченная пара (E_r, C_r), и будем это записывать в виде $W_r = (E_r, C_r)$. Пусть отображения $\varphi_r(\omega)$ и $\psi_r(\omega)$ измеряют выход объекта управления E_r и выход системы управления C_r, а функционал $\Phi(\varphi_r(\omega), \psi_r(\omega))$ определяет доход эксперимент W_r от управления r. Если при управлении r математическое ожидание дохода равно $\mathrm{M}_r\Phi(\boldsymbol{\varphi}_r, \boldsymbol{\psi}_r)$, то цель эксперимента W_r состоит в использовании оптимального управления $r' \in \mathcal{R}$, для которого имеет место соотношение $\mathrm{M}_{r'}\Phi(\boldsymbol{\varphi}_{r'}, \boldsymbol{\psi}_{r'}) = \sup\{\mathrm{M}_r\Phi(\boldsymbol{\varphi}_r, \boldsymbol{\psi}_r)\colon r \in \mathcal{R}\}$.

Пусть теперь для задачи об экзаменах объект управления E_r есть последовательные ответы студента преподавателям на три вопроса, а система управления C_r означает выбор студентом одного из двух вариантов поведения (управления). Здесь соотношение $r = 1$ означает выбор первого варианта управления, и равенство $r = 0$ означает выбор второго варианта поведения. Отсюда следует, что множество $\mathcal{R} = \{0, 1\}$. Заметим, что в этой задаче результаты объекта управления E_r не влияют на результаты системы управления C_r. Тогда приведенное ранее решение Мостеллером задачи об экзаменах на содержательном уровне фактически основано на представлении эксперимента W_r в виде управляющей системы (E_r, C_r) без канала обратной связи. Эта ситуация отображена на рис. 1.

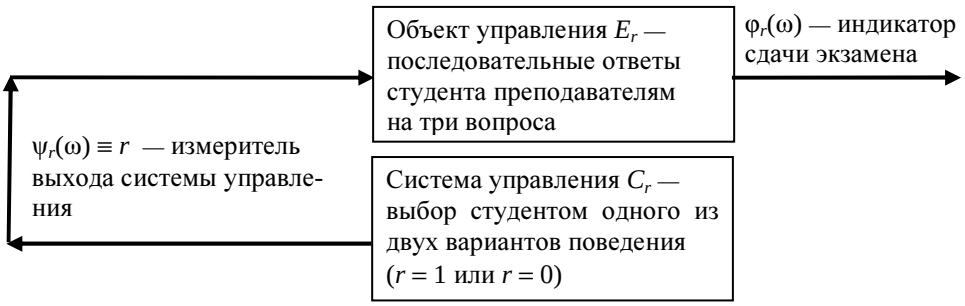


Рис. 1. Модель без канала обратной связи для эксперимента об экзаменах

Здесь величины $\varphi_r(\omega)$ и $\psi_r(\omega)$ измеряют элементарный исход $\{\omega\}$ эксперимента W_r с точки зрения объекта управления E_r и соответственно с точки зрения системы управления C_r. Значение случайной величины $\varphi_r(\omega)$ равно единице, если студент сдал экзамен, и равно нулю в противном случае, а значение величины $\psi_r(\omega)$равно выбранному управлению $r = 0, 1$.

В качестве модели эксперимента W_r можно предложить вероятностное пространство $(\Omega, \Im, \mathbf{P}_r(\cdot))$. Здесь $\Omega = \{\omega = (u_1, u_2, u_3)\colon u_1, u_2, u_3 = 0, 1\}$ и $\Im = \{D\colon D \subset \Omega\}$. Для $i = 1, 2, 3$ полагаем $u_i = 1$, если студент положительно ответил на i-й вопрос. В противном случае принимаем $u_i = 0$. Например, событие $\{(1, 1, 1)\}$ означает, что студент положительно ответил на все три вопроса. В силу независимости ответов вероятности $\mathbf{P}_r(\{(1, 1, 1)\}) = pqp$ при $r = 1$ и $\mathbf{P}_r(\{(1, 1, 1)\}) = qpq$ при $r = 0$. Если событие A означает, что студент сдаст экзамен, то $A = \{(1, 1, 0), (0, 1, 1), (1, 1, 1)\}$ и $\mathbf{P}_r(A) = \mathbf{P}_r(\{(1, 1, 0)\}) + \mathbf{P}_r(\{(0, 1, 1)\}) + \mathbf{P}_r(\{(1, 1, 1)\}) = rpq(q - p) + pq(2 - q)$. Так как $\varphi_r(\omega)$ — индикатор случайного события A, то математическое ожидание $\mathrm{M}_r\boldsymbol{\varphi}_r(\omega) = \mathbf{P}_r(A)$. Из неравенства $q > p$ и условия оптимальности $\mathrm{M}_{r'}\boldsymbol{\varphi}_{r'}(\omega) = \sup\{\mathbf{P}_r(A)\colon r \in \mathcal{R}\}$ получим, что оптимальное управление $r' = 1$. Значит, студентам целесообразно пользоваться первым вариантом поведения, хотя он два раза должен отвечать профессору. Мостеллер этот вывод объясняет важностью для студента умение решать практические задачи. Напротив, Секей говорит о том, что полученное решение этой задачи и выводы существенно опирались на факт независимости ответов студента. Однако статистические наблюдения автора этой работы показывают, что студенты, которые в течение 1967—1992 гг. сдавали экзамены, в основном выбирали первый вариант поведения. После 1992 г. студенты всё чаще выбирают более разумный на их взгляд второй вариант поведения. Чтобы объяснить такое поведение студентов, приведём нетрадиционный способ [1, 2] в проблеме изучения и оптимизации случайных экспериментов с управлением.

Задача Мостеллера как управляющая кибернетическая система. Приведенный выше подход изучает статически устойчивый эксперимент с управлением с позиции «черного ящика», как это обычно делается в известных задачах автоматического регулирования. Основной недостаток этого подхода заключается в том, что величины $\psi_r(\omega)$ и $\varphi_r(\omega)$ являются всего лишь математической моделью измерителей входа и соответственно выхода объекта управления E_r, а не вероятностной моделью статистически устойчивого эксперимента W_r в целом. Поэтому простой и доступный по структуре принцип «черного ящика» не требует детального изучения эксперимента W_r, и, значит, не позволяет в полной мере одновременно учитывать как детерминированную, так и вероятностную его природу.

При изучении реальной управляющей системы всегда возникают трудные вопросы эффективного способа описания и методов изучения с подробным учётом её конкретной физической природы и цели функционирования. В силу этого на реальную систему целесообразно смотреть не с позиции «черного ящика», а с точки зрения её общих фундаментальных свойств и методологического понятия управляющей системы, впервые данного в математической кибернетике [5]. Эта точка зрения по существу была поддержана и последовательно развивалась в работах [1, 2]. В основе кибернетического подхода при построении, анализе и оптимизации модели управляющей системы лежат следующие фундаментальные положения:

— любая управляющая система обладает такими свойствами как схемой, информацией, координатами и функцией;

— принцип дискретности актов функционирования управляющей системы во времени τ_i, $i \geq 0$, где точечный процесс $\{\tau_i;\ i \geq 0\}$ задает на $[0, \infty)$ шкалу тактов времени;

— принцип строения схемы из структурных блоков: внешней среды, входных и выходных полюсов, внешней и внутренней памяти, устройств по переработке информации внешней и внутренней памяти.

Рассмотрим теперь применение кибернетического подхода для решения парадокса независимости [4, 5]. Задача Мостеллера или задача об экзаменах может быть адекватно отображена посредством управляющей системы, блочная схема которой приведена на рис. 2.

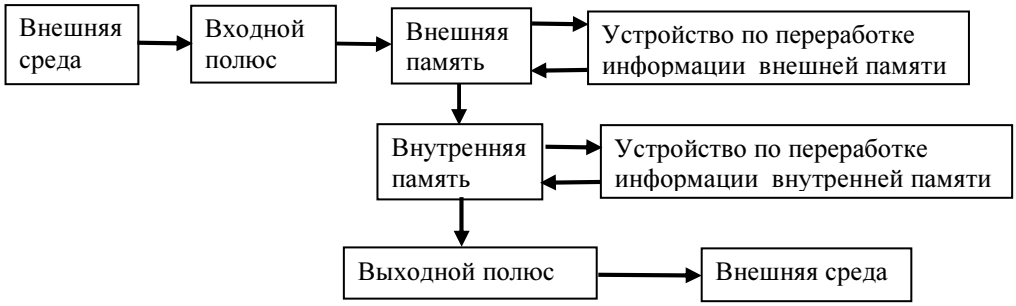


Рис. 2. Схема управляющей системы для задачи об экзаменах

Выясним неформальный смысл всех структурных блоков схемы управляющей системы, соответствующей задаче об экзаменах. Для этого обозначим на оси времени через τ_0 момент получения студентом выбранного им билета с тремя вопросами и объявления первого варианта поведения с вероятностью $r \in [0, 1]$ или второго варианта с вероятностью $1 - r$. Пусть τ_i есть момент завершения ответа студентом на i-й вопрос, где $i = 1, 2, 3$. При этом рассмотрим общий случай задачи, когда не предполагается независимость в совокупности последовательных ответов студента на вопросы для любой стратегии его поведения. Сначала пунктом 1) определим математическое описание и свойства внешней среды, а затем пунктами 2) — 7) описание и свойства каждого из остальных структурных блоков схемы.

1) Процессы формирования внешней среды занимают значительное время и являются консервативными. Внешняя среда представляет собой условия перед экзаменом, которые навязывают студенту на интуитивном уровне случайно выбрать конкретный вариант поведения. Отсюда вытекает, что внешняя среда имеет два состояния: а) профессор — ассистент — профессор (b_1); б) ассистент — профессор — ассистент (b_2). Будем считать, что внешняя сре-

да в некоторый момент τ_0 принимает значение b_1 с вероятностью $r \in [0, 1]$ и значение b_2 с вероятностью $1 - r$. Другими словами, студент выбирает план ответов вида: профессор — ассистент — профессор с вероятностью r и стратегию ответов вида: ассистент — профессор — ассистент с вероятностью $1 - r$. Информация внешней среды есть множество {профессор — ассистент — профессор, ассистент — профессор— ассистент} = $\{b_1, b_2\}$ из двух состояний внешней среды. Координаты внешней среды — номера 1, 2 её состояний b_1 и b_2. Состояние внешней среды в момент τ_i есть случайный элемент $\chi_i \equiv \chi_0$ при $i = 0, 1, 2, 3$. Код информации внешней среды есть случайный вектор $\chi = (\chi_0, \chi_1, \chi_2, \chi_3)$. При этом распределение вероятностей $\mathbf{P}(\chi_0 = b_1) = r$, $\mathbf{P}(\chi_0 = b_2) = 1 - r$ случайного элемента χ_0 определяет распределение случайного вектора χ.

2) Входной полюс означает получение студентом в момент τ_0 три последовательных вопроса w_1, w_2 и w_3 билета. Информация входного полюса есть множество $\{(w_1, w_2, w_3)\}$ из одного состояния (w_1, w_2, w_3) входного полюса. Вектор $\alpha_i \equiv (w_1, w_2, w_3)$ определяет состояние входного полюса в момент τ_i, где $i = 0, 1, 2, 3$. Код информации входного полюса есть вектор $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ с равными компонентами.

3) Внешняя память фиксирует очередные вопросы, на которые студент должен отвечать. Информация внешней памяти есть множество$\{(w_1, w_2, w_3), (w_2, w_3), w_3, \varnothing\}$ из четырёх состояний $d_1 = (w_1, w_2, w_3)$, $d_2 = (w_2, w_3)$, $d_3 = w_3$ и $d_4 = \varnothing$ внешней памяти. Координаты внешней памяти — номера 1, 2, 3, 4 её состояний d_1, d_2, d_3, d_4. Состояние внешней памяти в момент τ_i есть $\beta_i = d_{i+1}$ при $i = 0, 1, 2, 3$. Код информации внешней памяти — вектор вида $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$.

4) Устройство по переработке информации внешней памяти детерминировано удаляет вопросы, на которые студент закончил отвечать и, тем самым, изменяет номер состояния внешней памяти в последовательности 1, 2, 3 и 4. Информация блока по переработке внешней памяти есть множество $\{\varnothing, w_1, (w_1, w_2), (w_1, w_2, w_3)\}$ из четырёх состояний $e_1 = \varnothing$, $e_2 = w_1$, $e_3 = (w_1, w_2)$ и $e_4 = (w_1, w_2, w_3)$ этого блока. Координаты блока по переработке внешней памяти суть номера 1, 2, 3 и 4 состояний e_1, e_2, e_3 и e_4 блока. Состояние блока по переработке внешней памяти в момент τ_i есть $\zeta_i = e_{i+1}$, $i = 0, 1, 2, 3$. Код информации блока по переработке внешней памяти есть вектор $\zeta = (\zeta_0, \zeta_1, \zeta_2, \zeta_3)$.

5) Внутренняя память фиксирует как выбранный заранее студентом вариант поведения, так и запоминает последовательность из трёх результатов его ответов преподавателям на три вопроса. Информация блока внутренней памяти есть множество {студент выбрал первую схему ответа (f_1), студент выбрал вторую схему ответа (f_2), студент ответил на вопрос (f_3), студент не ответил на вопрос (f_4)} из четырёх состояний f_1, f_2, f_3 и f_4 внутренней памяти. Координаты внутренней памяти суть номера 1, 2, 3 и 4 состояний f_1, f_2, f_3 и f_4. Состояние внутренней памяти в момент τ_i есть случайный элемент η_i при $i = 0, 1, 2, 3$. Здесь $\eta_0 \in \{f_1, f_2\}$, $\eta_1 \in \{f_3, f_4\}$, $\eta_2 \in \{f_3, f_4\}$, $\eta_3 \in \{f_3, f_4\}$. Код информации внутренней памяти есть случайный вектор $\eta = (\eta_0, \eta_1, \eta_2, \eta_3)$. Напомним, что согласно условиям задачи об экзаменах последовательные ответы студента происходят случайным и, в общем случае, зависимым образом. Эта зависимость при заданном значении параметра $\varepsilon \geq 0$ определяется следующими соотношениями для условных вероятностей случайных элементов η_0, η_1, η_2, η_3:

$$\mathbf{P}(\eta_0 = f_1) = r,\ \mathbf{P}(\eta_1 = f_3|\eta_0 = f_1) = p,\ \mathbf{P}(\eta_1 = f_3|\eta_0 = f_2) = q,\ \mathbf{P}(\eta_2 = f_3|\eta_0 = f_1, \eta_1 = f_3) =$$
$$= q + \varepsilon - \varepsilon/p \geq p,\ \mathbf{P}(\eta_2 = f_3|\eta_0 = f_1, \eta_1 = f_4) = \mathbf{P}(\eta_3 = f_3|\eta_0 = f_2, \eta_1 = f_3, \eta_2 = f_4) =$$
$$= \mathbf{P}(\eta_3 = f_3|\eta_0 = f_2, \eta_1 = f_4, \eta_2 = f_4) = q + \varepsilon < 1,\ \mathbf{P}(\eta_3 = f_3|\eta_0 = f_2, \eta_1 = f_3, \eta_2 = f_3) = q,$$
$$\mathbf{P}(\eta_2 = f_3|\eta_0 = f_2, \eta_1 = f_3) = \mathbf{P}(\eta_2 = f_3|\eta_0 = f_2, \eta_1 = f_4) = \mathbf{P}(\eta_3 = f_3 \mid \eta_0 = f_1, \eta_1 = f_3, \eta_2 = f_3) =$$
$$= \mathbf{P}(\eta_3 = f_3 \mid \eta_0 = f_1, \eta_1 = f_3, \eta_2 = f_4) = \mathbf{P}(\eta_3 = f_3 \mid \eta_0 = f_1, \eta_1 = f_4, \eta_2 = f_3) =$$
$$= \mathbf{P}(\eta_3 = f_3 \mid \eta_0 = f_1, \eta_1 = f_4, \eta_2 = f_4) = p.$$

Эти ограничения на вероятности формализуют случайный механизм выбора последовательности ответов студента, условия задачи Мостеллера, естественную зависимость его

последовательных ответов и, наконец, модель обучения плохо подготовленных студентов на экзаменах.

6) Устройство по переработке информации внутренней памяти определяет по некоторому алгоритму условные вероятности очередного положительного ответа студента в зависимости от предыдущих его результатов, и реализовывает случайный механизм оценки ответов на вопросы. Информация блока по переработке внутренней памяти есть множество вида $\{r, (p, q), (q + \varepsilon - \varepsilon/p, q + \varepsilon, p, p), (p, p, p, p, q, q + \varepsilon, q + \varepsilon, q + \varepsilon)\}$, которое состоит из четырёх состояний: $g_1 = r$, $g_2 = (p, q)$, $g_3 = (q + \varepsilon - \varepsilon/p, q + \varepsilon, p, p)$, $g_4 = (p, p, p, p, q, q + \varepsilon, q + \varepsilon, q + \varepsilon)$. Координаты блока по переработке внутренней памяти суть номера 1, 2, 3, 4 состояний g_1, g_2, g_3, g_4. Состояние блока в момент τ_i есть $\theta_i = g_{i+1}$ при $i = 0, 1, 2, 3$. Код информации блока — вектор $\theta = (\theta_0, \theta_1, \theta_2, \theta_3)$.

7) Выходной полюс определяет наблюдаемые исходы экзамена студента. Информация этого блока есть пространство $\Omega = \{\omega = (x_0, x_1, x_2, x_3)$: $x_0 \in \{f_1, f_2\}$, $x_1, x_2, x_3 \in \{f_3, f_4\}\}$ из 16 состояний: $\omega_1 = (f_1, f_3, f_3, f_3)$, $\omega_2 = (f_1, f_3, f_3, f_4)$, $\omega_3 = (f_1, f_3, f_4, f_3)$, $\omega_4 = (f_1, f_3, f_4, f_4)$, $\omega_5 = (f_1, f_4, f_3, f_3)$, $\omega_6 = (f_1, f_4, f_3, f_4)$, $\omega_7 = (f_1, f_4, f_4, f_3)$, $\omega_8 = (f_1, f_4, f_4, f_4)$, $\omega_9 = (f_2, f_3, f_3, f_3)$, $\omega_{10} = (f_2, f_3, f_3, f_4)$, $\omega_{11} = (f_2, f_3, f_4, f_3)$, $\omega_{12} = (f_2, f_3, f_4, f_4)$, $\omega_{13} = (f_2, f_4, f_3, f_3)$, $\omega_{14} = (f_2, f_4, f_3, f_4)$, $\omega_{15} = (f_2, f_4, f_4, f_3)$, $\omega_{16} = (f_2, f_4, f_4, f_4)$. Координаты блока выходного полюса суть номера 1, 2, …, 16 состояний ω_1, ω_2, …, ω_{16}. Состояние блока в момент времени τ_3 есть случайный элемент ξ, где $\xi(\omega) \equiv \omega$ при $\omega = (x_0, x_1, x_2, x_3) \in \Omega$. Код информации блока выходного полюса есть случайный вектор $\xi = (\xi_0, \xi_1, \xi_2, \xi_3)$, где $\xi_i(\omega) \equiv x_i$, $i = 0, 1, 2, 3$.

Теперь нетрудно построить вероятностную модель $(\Omega, \Im, \mathbf{P}_r(\bullet))$ эксперимента W_r. Здесь $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{16}\}$, $\Im = \{B: B \subset \Omega\} = \{B_1, B_2, \ldots, B_{65536}\}$ и вероятностная функция $\mathbf{P}_r(\bullet))$ задается равенствами: $\mathbf{P}_r(\{\omega_1\}) = rp(q + \varepsilon - \varepsilon/p)p$, $\mathbf{P}_r(\{\omega_2\}) = rp(q + \varepsilon - \varepsilon/p)(1 - p)$, $\mathbf{P}_r(\{\omega_3\}) = rp(1 - q - \varepsilon + \varepsilon/p)p$, $\mathbf{P}_r(\{\omega_4\}) = rp(1 - q - \varepsilon + \varepsilon/p)(1 - p)$, $\mathbf{P}_r(\{\omega_5\}) = r(1 - p)p \times (q + \varepsilon)$, $\mathbf{P}_r(\{\omega_6\}) = r(1 - p)(q + \varepsilon)(1 - p)$, $\mathbf{P}_r(\{\omega_7\}) = r(1 - p)(1 - q - \varepsilon)p$, $\mathbf{P}_r(\{\omega_8\}) = r(1 - p) \times (1 - q - \varepsilon)(1 - p)$, $\mathbf{P}_r(\{\omega_9\}) = (1 - r)qpq$, $\mathbf{P}_r(\{\omega_{10}\}) = (1 - r)qp(1 - q)$, $\mathbf{P}_r(\{\omega_{11}\}) = (1 - r)q \times (1 - p)(q + \varepsilon)$, $\mathbf{P}_r(\{\omega_{12}\}) = (1 - r)q(1 - p)(1 - q - \varepsilon)$, $\mathbf{P}_r(\{\omega_{13}\}) = (1 - r)(1 - q)p(q + \varepsilon)$, $\mathbf{P}_r(\{\omega_{14}\}) = (1 - r)(1 - q)p(1 - q - \varepsilon)$, $\mathbf{P}_r(\{\omega_{15}\}) = (1 - r)(1 - q)(1 - p)(q + \varepsilon)$, $\mathbf{P}_r(\{\omega_{16}\}) = (1 - r) \times (1 - q)(1 - p)(1 - q - \varepsilon)$. Например, элементарное событие $\{\omega_8\} = \{(f_1, f_4, f_4, f_4)\} \in \Im$ означает, что студент выбрал первый вариант последовательности ответов и отрицательно ответил на все три вопроса. При этом вероятность $\mathbf{P}_r(\{\omega_8\}) = r(1 - p)(1 - q - \varepsilon)(1 - p)$. Построенная вероятностная модель $(\Omega, \Im, \mathbf{P}_r(\bullet))$ эксперимента W_r позволяет выписать функциональную связь между блоком внешней среды и выходным полюсом в виде равенств: $\chi_0(\omega_1) = \chi_0(\omega_2) = \ldots = \chi_0(\omega_8) = b_1$, $\chi_0(\omega_9) = \chi_0(\omega_{10}) = \ldots = \chi_0(\omega_{16}) = b_2$. Аналогично найдём функциональную связь между внутренней памятью и выходным полюсом. Например, функциональная связь между внутренней памятью в моменты τ_3 и выходным полюсом задается равенством $\eta_3(\omega_s) = f_3$ при нечетном значении s. Приведенные здесь и остальные функциональные связи являются поточечным заданием случайных элементов $\chi(\omega)$, $\eta(\omega)$ и $\xi(\omega)$. Поэтому их законы распределения и связанные с этими элементами условные вероятности могут быть найдены по правилам теории вероятностей. Так, $\mathbf{P}_r(\eta_2 = f_3 \mid \eta_0 = f_1) = \mathbf{P}_r(\eta_0 = f_1, \eta_2 = f_3)/\mathbf{P}_r(\eta_0 = f_1) = (\mathbf{P}_r(\{\omega_1\}) + \mathbf{P}_r(\{\omega_2\}) + \mathbf{P}_r(\{\omega_5\}) + \mathbf{P}_r(\{\omega_6\}))/r = pq + p\varepsilon - \varepsilon + (1 - p)(q + \varepsilon) = q$. Однако, условная вероятность $\mathbf{P}_r(\eta_3 = f_3 \mid \eta_0 = f_2) = \mathbf{P}_r(\eta_0 = f_2, \eta_3 = f_3)/\mathbf{P}_r(\eta_0 = f_2) = (1 - r)^{-1} \times (\mathbf{P}_r(\{\omega_9\}) + \mathbf{P}_r(\{\omega_{11}\}) + \mathbf{P}_r(\{\omega_{13}\}) + \mathbf{P}_r(\{\omega_{15}\})) = q + \varepsilon(1 - pq) > q$. Значит, произошло обучение студента на экзаменах.

Математическое описание внешней среды, входного полюса, внешней памяти, устройства по переработке информации внешней памяти, внутренней памяти, устройства по переработке информации внутренней памяти и выходного полюса выполняется векторами $\chi(\omega)$, $\alpha(\omega)$, $\beta(\omega)$, $\zeta(\omega)$, $\eta(\omega)$, $\theta(\omega)$ и $\xi(\omega)$ на вероятностном пространстве $(\Omega, \Im, \mathbf{P}_r(\bullet))$. Функция эксперимента W_r об экзамене заключается в случайном выборе студентом одного из двух ва-

риантов поведения и в его ответах на три последовательных вопроса с целью определения объективной оценки преподавателями знаний студента.

Оптимальное поведение студента и решение парадокса Мостеллера—Секея. Пусть случайное событие $A = \{\omega_1, \omega_2, \omega_5, \omega_9, \omega_{10}\ \omega_{13}\}$ означает, что студент сдал экзамен. Тогда вероятность $\mathbf{P}_r(A) = \mathbf{P}_r(\{\omega_1\}) + \mathbf{P}_r(\{\omega_2\}) + \mathbf{P}_r(\{\omega_5\}) + \mathbf{P}_r(\{\omega_9\}) + \mathbf{P}_r(\{\omega_{10}\}) + \mathbf{P}_r(\{\omega_{13}\}) = r[pq(q-p) - \varepsilon(1 + p^2 - p - pq)] + qp(2-q) + p\varepsilon(1-q)$. Формулировка задачи оптимизации для эксперимента W_r заключается в определении такого $r' \in [0, 1]$, для которого выполняется условие оптимальности: $\mathbf{P}_{r'}(A) = \sup\{\mathbf{P}_r(A): 0 \le r \le 1\}$. При определении оптимальной стратегии выбора последовательности ответов в задаче Мостеллера об экзаменах с обучением необходимо учитывать ограничения $0 < p, q < 1$, $q > p$, $\varepsilon \ge 0$, $q + \varepsilon < 1$, $q + \varepsilon - \varepsilon/p \ge p$ на параметры p и q, которые определяют уровень подготовки студента, и на параметр обучения ε. Теперь нетрудно показать следующие утверждения.

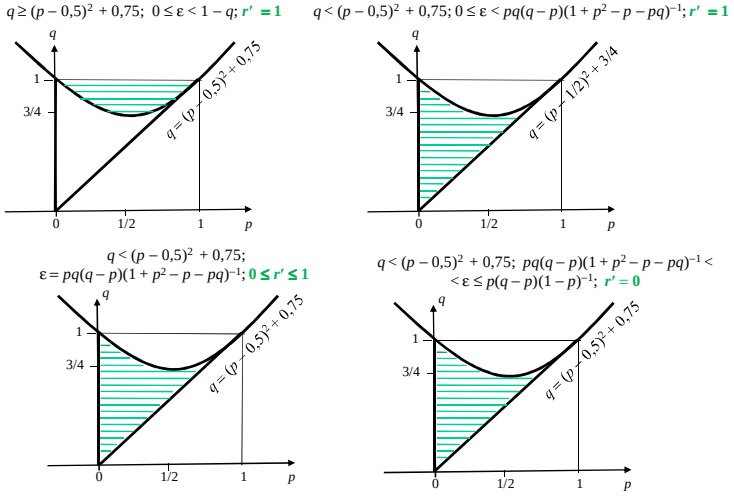
Лемма 1. Если выполняется неравенство $q \ge (p - 0{,}5)^2 + 0{,}75$ и $q + \varepsilon < 1$, то $r' = 1$.

Лемма 2. Пусть $q < (p - 0{,}5)^2 + 0{,}75$ и $0 \le \varepsilon < pq(q-p)(1 + p^2 - p - pq)^{-1}$, то $r' = 1$.

Лемма 3. Если $q < (p - 0{,}5)^2 + 0{,}75$ и $\varepsilon = pq(q-p)(1 + p^2 - p - pq)^{-1}$, то $0 \le r' \le 1$.

Лемма 4. Пусть имеет место $pq(q-p)(1 + p^2 - p - pq)^{-1} < \varepsilon \le p(q-p)(1-p)^{-1}$ и $q < (p - 0{,}5)^2 + 0{,}75$. Тогда оптимальный план $r' = 0$.

На следующих рисунках приведены заштрихованные области p, q и промежутки изменения ε, в которых следует применять оптимальный план ответов студента на экзамене.



Приведём решение парадокса о независимости. Так для отлично подготовленных студентов, ответы которых, скорее всего, будут независимыми случайными событиями ($\varepsilon = 0$) или удовлетворяют условию $q \ge (p - 0{,}5)^2 + 0{,}75$, целесообразно пользоваться первым вариантом поведения. Напротив, для плохо подготовленных студентов и при хорошей их обучаемости в случае, когда $0 < p, q < 1$, $q > p$, $q < (p - 0{,}5)^2 + 0{,}75$ и $pq(q-p)(1 + p^2 - p - pq)^{-1} < \varepsilon \le p(q-p)(1-p)^{-1}$, целесообразно применять второй вариант поведения. Такая ситуация, по-видимому, сложилась после 1992 г. не только в Нижегородском госуниверситете.

ЛИТЕРАТУРА

[1]. Федоткин М.А. Процессы обслуживания и управляющие системы. Математические вопросы кибернетики. М.: Наука. 1996. Вып. 6. С. 51–70

[2]. Федоткин М.А. Нелокальный способ задания управляемых случайных процессов. Математические вопросы кибернетики. М.: Наука. 1998. Вып. 7. С. 332–344.

[3]. Ляпунов А.А., Яблонский С.В. Теоретические проблемы кибернетики. Проблемы кибернетики. М.: Физматгиз. 1963. Вып. 9. С. 5–22.

[4]. Mosteller F. Fifty challenging problems in probability with solutions. Reading, MA: Addison-Wesley. 1965.

[5]. Székely G. Paradoxes in probability theory and mathematical statistics. Budapest: Akadémiai Kiadó. 1986.

THE PATTERN OF SNPs DISTRIBUTION IN THE HUMAN GENOME [1]

E.V.Ignatieva

Institute of Cytology and Genetics, Novosibirsk

eignat@bionet.nsc.ru

V.G.Levitsky

levitsky@bionet.nsc.ru

N.S.Yudin

yudin@bionet.nsc.ru

SNPs density is known to vary along a nucleotide sequence. The genome-scale patterns of SNP distribution have been poorly investigated yet. We have analyzed the pattern of SNPs distribution among the different regions of human genome and DNA context features of nucleotide sequences with adjacent SNPs using the data from the 1000 Genomes Project. It was found that SNPs are localized non-uniformly among human genome. About 1.3% of SNPs occur in neighboring positions (adjacent SNPs). It is 3 times more than expected accident frequency for two adjacent SNPs . In 0.8% of cases SNPs are separated by one nucleotide. SNPs density was dependent on their localization across the different parts of the gene. Low SNPs density was found in the vicinity of transcription start sites, 5'- and 3'- parts of introns. This observation is in good agreement with important functional roles of these regions. General DNA context features of the stretches with adjacent SNPs in different regions of genes (exons, 5'-UTRs, 3'-UTRs) were detected: the frequency of CpG was found to be higher among this two adjacent SNPs and AT-content was found to be higher in 5' and 3' flanks (3-5 bp). Analysis of the SNPs distribution is of great importance for experimental researches because sensitivity and specificity of most experimental methods of SNPs typing are strongly dependent from the presence of additional SNPs variants in the neighborhood with the target SNP.

This research was supported by the Ministry of Education and Science of RF (contr. № P721), SB RAS (Project № 119) and RAS (Programs №№ B.25, B.27, A.II.6)

Удалено: E.V.Ignatieva¶ *Institute of Cytology and Genetics, Novosibirsk* . eignat@bionet.nsc.ru . V.G.LEVITSKY . levitsky@bionet.nsc.ru . N.S.Yudin .

Удалено: Mutation frequencies are

Удалено: not

Удалено: systematically

Удалено: T

Удалено: was analyzed

Удалено: 1000 genome project

Отформатировано: английский (США)

Удалено: and 1.5 times more then the observed number of SNPs separated by one nucleotide.

Отформатировано: английский (США)

Отформатировано: английский (США)

Удалено: SNPs

Отформатировано: английский (США)

Отформатировано: английский (США)

Отформатировано: английский (США)

Отформатировано: английский (США)

Отформатировано: английский (США)

Отформатировано: английский (США)

Удалено: ends

Отформатировано: английский (США)

Отформатировано: английский (США)

Отформатировано: английский (США)

Отформатировано: английский (США)

Удалено: the

Удалено: central positions

Отформатировано: английский (США)

Удалено: positions adjacent to the boundary of the stretch.

Отформатировано: английский (США)

Удалено: decreased

БАЗА ЗНАНИЙ ПО РЕГУЛЯЦИИ ТРАНСКРИПЦИИ ГЕНОВ ЭУКАРИОТ[1]

Е.В.Игнатьева

Учреждение Российской академии наук Институт Цитологии и Генетики Сибирского Отделения РАН

eignat@bionet.nsc.ru

Н.Л.Подколодный

Учреждение Российской академии наук Институт вычислительной математики и математической геофизики Сибирского отделения РАН

pnl@bionet.nsc.ru

Транскрипционная активность генов эукариот определяется типом клетки, ткани и органа, стадией развития организма, стадией клеточного цикла и этапом дифференцировки, многочисленными индукторами и репрессорами ит.д. Тонкая и сложная регуляция транскрипции генов эукариот осуществляется при участии большого количества регуляторных белков, комбинация которых определяет разнообразие регуляторных механизмов. Нами разработана пилотная версия базы знаний по регуляции транскрипции генов эукариот, включающая: i) иерархическую классификацию регуляторов транскрипции; ii) классификацию стадий транскрипции; iii) сведения о транскрипционных факторах трех видов млекопитающих (человека, мыши, крысы); iv) словари по элементарным молекулярным процессам, обеспечивающим регуляцию транскрипции. База знаний может быть использована для i) реконструкци гипотетических механизмов регуляции транскрипции с учетом информации о функциях и структурах регуляторных белков, присутствующих в заданных клетках или тканях на определенной стадии развития, а также закономерностей строения регуляторных районов коэкспрессирующихся генов; ii) для функциональной интерпретации микрочиповых данных, отражающих уровни экспрессии генов.

[1] Работа выполнена при частичной поддержке Минобрнауки (госконтракт с ФАО №П721), СО РАН (проект №119), президиума РАН (программы №№ Б.25, Б.27, А.II.6)

Identification of the systems of linear difference equations

V. Demidenko

Our work is devoted to studying identification problems for linear difference equations and control theory systems. Such identification problems arise in vast variety of applications including geophysics, engineering, economics, bioinformatics etc. There are many popular modern approaches for solving such problems which are based on having relatively large number of observations. However, in real applications the number of observations can be quite small and the number of unknowns can be relatively large. In the present work we consider the problems with such features. Our main results are: algorithms for constructing approximations of the solutions of identification problem having the small number of observations based on modifications of Prony's method; the proof of convergence of the iteration process and the estimates of rate of convergence; the stability estimates of the solutions given the perturbations in the observation data. For this work we have developed the program implementation for solving the considered identification problems optimized for modern high-performance multiprocessor systems.

ВЕБ-ОРИЕНТИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА ФУНКЦИОНАЛЬНОЙ ДИАГНОСТИКИ*

Б. Болдсайхан
Монгольский университет науки и технологии, Монголия, Улан-Батор
bbld@must.edu.mn
Т.Г. Дармаев, Ф.В. Хандаров, Б.В. Хабитуев
Бурятский государственный университет, Россия, Улан-Удэ
dtg@bsu.ru

Аннотация

В данной работе излагаются подходы к реализации информационной системы по сбору и обработке данных функциональной диагностики на основе методов тибетской и монгольской медицины.

Введение

В тибетской медицине нарушения в человеческом организме трактуются нарушением баланса трех составляющих (доша) – ветра, желчи и слизи; методы диагностики организма ставят своей целью анализ этого баланса, а методы лечения – восстановление приобретенного нарушения. [1-2]. Определенный интерес представляет исследование зависимостей между диагнозом врача-тибетолога и «европейским» диагнозом.

Данная статья рассматривает подходы к реализации методов тибетской медицины и ее монгольского ответвления в рамках информационной системы функциональной диагностики. Одной из основных особенностей подобной системы помимо решения задач сбора, хранения и обработки диагностических данных, полученных различными способами, также должна стать возможность постановки гипотез о взаимосвязи между данными.

1. Общие требования к системе функциональной диагностики

Данные по функциональному состоянию организма могут быть получены в результате обследования человека в четырех «технологически» различных направлениях. Архитектура рассматриваемой информационной системы должна включать соответствующие подсистемы (см. Табл.1)

Таблица 1.

Описание подсистем, отвечающих направлениям обследования функционального состояния организма человека.

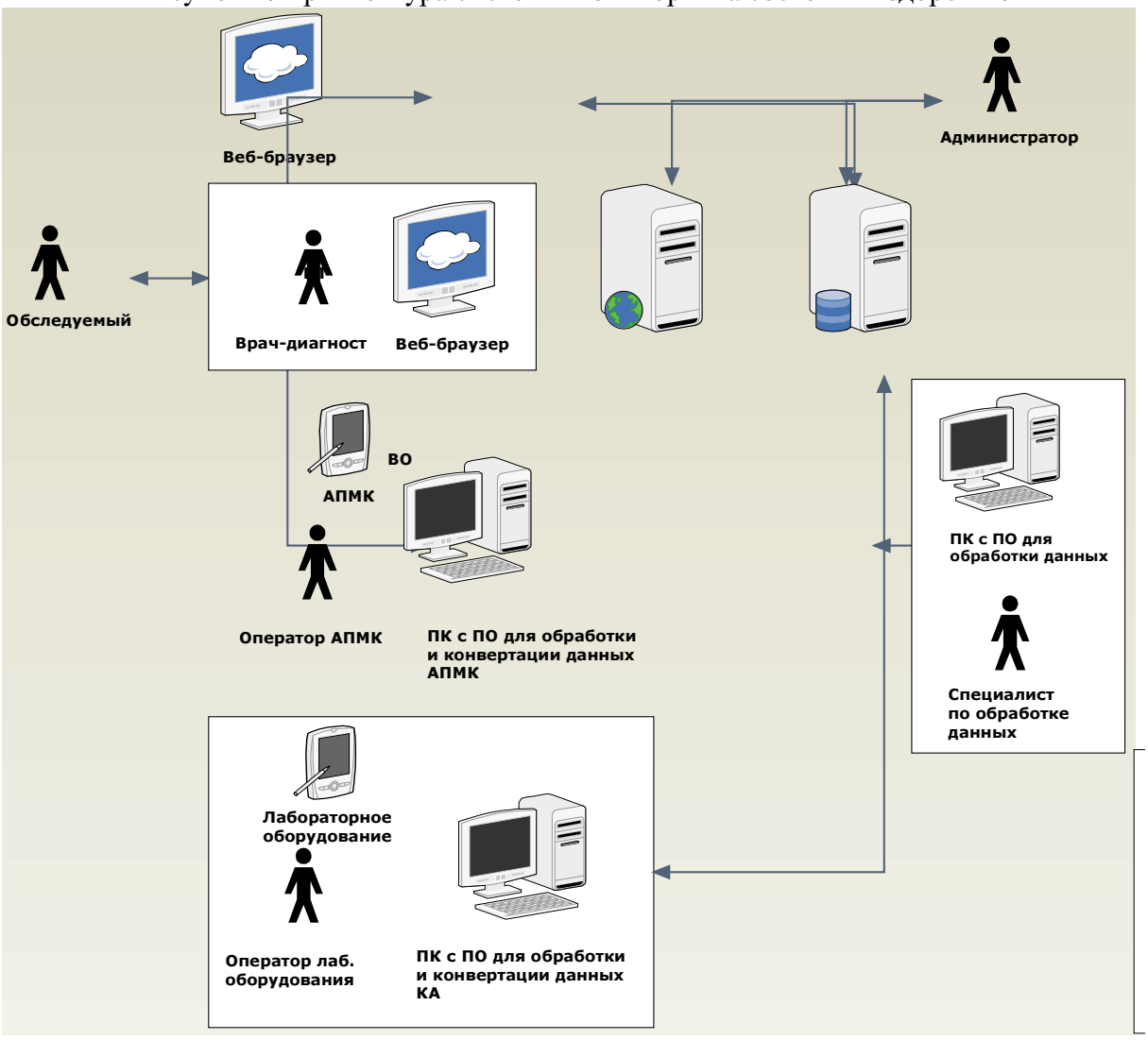
№	Наименования направления	Назначение, требования к реализации
1	Интерактивное тестирование (ИТ)	Подсистема, позволяющая создавать электронные вопросники различного типа.
2	Врачебный осмотр (ВО)	Подсистема, позволяющая создавать автоматизированные рабочие места врачей-специалистов в области функциональной диагностики, обеспечивающая гибкость при расширении (предполагается, что у каждого врача свой собственный формат карточки медицинского осмотра), доступ к большому количеству специализированных справочников (МКБ-10, справочники по препаратам, справочники по тибетской медицине, прочие специализированные справочники).
3	Аппаратно-программная диагностика (АПД)	Подсистема, позволяющая создавать автоматизированные рабочие места операторов медицинских диагностических аппаратов, обеспечивающая широкие возможности по вводу в систему новых аппаратов.

* Работа выполнена при поддержке РФФИ. Проект № 11-07-92202-Монг_а

3	Лабораторно-клинические анализы (ЛКА)	Подсистема, позволяющая создавать автоматизированные рабочие места для работников клинической лаборатории, обеспечивающая широкие возможности по вводу в систему результатов новых видов клинических анализов.

Общая архитектурная схема всей системы представлена на Рис.1.

Рисунок 1. Архитектура системы мониторинга состояния здоровья.



2. Система диагностики с учетом подходов тибетской и восточной медицины

Если методы лечения в тибетской медицине и ее монгольском ответвлении отличаются достаточно сильно (в силу различий природной фармакологической базы), то методы диагностики в основном схожи. Одним из основных диагностических методов тибетской медицины является диагностика по пульсу, также интерес представляет диагностика по состоянию языка и мочи. В настоящее время недостаточно формализованы.

Основной целью данной работы является формализация рассматриваемых диагностических методов с последующей разработкой АРМ врачей-экспертов тибетской медицины.

Для достижения поставленной цели ставятся следующие задачи:

1) организация сбора и хранения в течение продолжительного периода биомедицинских данных и результатов врачебных осмотров пациентов врачами-экспертами;

2) возможность «самомодификации», «обучения» системы при поступлении новых данных – выявление взаимосвязей между результатами современной функциональной диагностики и диагностических методов тибетской медицины;

3) реализация возможности обработки данных с использованием выявленных взаимосвязей.

Разрабатываемая информационная система должна удовлетворять следующим требованиям:

1) двуязычность – русский и монгольский языки;
2) веб-ориентированность – доступ к системе должен обеспечиваться из географически удаленных точек;
3) широкие возможности по расширению – вводу новых диагностических методов.

3. Система мониторинга состояния здоровья студентов и сотрудников Бурятского государственного университета

В настоящее время изложенные подходы частично реализованы в действующей системе мониторинга состояния здоровья студентов и сотрудников Бурятского государственного университета. [3-4]

Конкретизация методов представлена в Табл.2.

Таблица 2.

Конкретизация «технологических направлений» информационной системы для системы мониторинга состояния здоровья студентов БГУ.

№	Вид обследования	Напр-е обслед-я	Назначение
1	Общая анкета	ИТ	Сбор общих сведений о пациенте, данные о перенесенных, хронических заболеваниях, хирургических операциях
2	Типы тела [5]	ИТ	Выявление типа тела по канонам тибетской медицины
3	Психотипы [6]	ИТ	Выявление информации о психотипе пациента
4	Симптомы	ИТ	Выявление текущего состояния пациента по канонам тибетской медицины
5	Опрос «Вчера-сегодня»	ИТ	Выявление текущего состояния пациента (прием пищи, сон, заболевания и т.п.)
6	Система диагностики и коррекции организма (СДК)[7]	АПМК	Пульсовая диагностика
7	Аппаратно-программный пульсодиагностический комплекс (АПДК)[8]	АПМК	Пульсовая диагностика
8	«Варикард»	АПМК	Диагностика сердечно-сосудистой системы
9	ДиаДЭНС	АПМК	Электроакупунктурная диагностика
10	Врач-пульсодиагност	ВО	Пульсовая диагностика
11	Врач-терапевт	ВО	Давление, температура, «европейский» диагноз
12	Эли-Тесты А.Б.Полетаева	ЛКО	Диагностика иммунной системы

На сегодняшний день с помощью данной системы накоплена определенная база данных обследований. Проведенные на собранных материалах исследования позволили

прийти к ряду результатов, в том числе, произвести сравнение диагностических комплексов тибетской медицины [9], исследовать вариабельность некоторых параметров периферического пульса в зависимости от степени прижима артерии [10], определить взаимовлияние органов и систем организма человека [11-12], предложить методики определения типа тела по канонам тибетской медицины в отсутствие эксперта-пульсодиагноста и некоторые другие.

Запланированы дальнейшие этапы работы:

1) продолжение накопления базы обследований;
2) отладка и модификация информационной системы;
3) введение в систему диагностических методов монгольского ответвления тибетской медицины, в частности, диагностики по пульсу и по состоянию языка и мочи;
4) реализация непосредственно в системе методов непараметрической статистики и искусственных нейронных сетей для проведения экспресс-анализа на наличие взаимосвязей между различными методиками.

ЛИТЕРАТУРА

[1]. Чжуд-ши. Канон тибетской медицины / Пер. с тибетского яз. Дашиева Д.Б. М.: Изд. фирма «Восточная литература» РАН, 2001 – 768 с.

[2]. Жамбалдагбаев Н.Ц., Занданова Г.И. Место представлений о психофизиологических типах человека в клинической практике тибетской медицины / Сб. науч. трудов НПЦ ТМГ Минздрава России – Москва, 2001. – С. 34 – 37.

[3]. Занданова Г.И., и др. Разработка информационной системы для сбора и обработки данных функциональной диагностики / Материалы международной конференции «Вычислительная математика, дифференциальные уравнения, информационные технологии» - Улан-Удэ, 2009 – С. 181-186.

[4]. Аюшеева Л.В. и др. Информационная система мониторинга состояния здоровья студентов и сотрудников Бурятского государственного университета //Вестник Бурятского государственного университета. Сер. Математики и информатика. - 2010. – Вып.9 – С. 244-251.

[5]. Айзенк Г., Вильсон Г. Как измерить личность / Пер. с англ. М.: "Когито-центр", 2000.

[6]. Чойжинимаева С.Г. Диагностика в тибетской медицине. М.:Аст, Астрель, 2010.

[7]. Дудин С.А. Система диагностики и коррекции организма человека. Сборник трудов Международной научной конференции «Методы и алгоритмы принятия эффективных решений» (ТТИ ЮФУ (ТРТУ)) Таганрог, 2009.

[8]. Азаргаев Л.Н., Бороноев В.В., Тарнуев В.А. Методика работы на автоматизированном пульсодиагностическом комплексе тибетской медицины. Изд-во БГУ, Улан-Удэ. 2000.

[9]. Дудин С.А. Сравнение диагностических комплексов тибетской медицины // Инфокоммуникационные и вычислительные технологии и системы (ИКВТС-2010) : материалы III Междунар. конф., 6-11 сент. 2010 г., г. Улан-Удэ - оз. Байкал. - Улан-Удэ : Изд-во Бурят. госун-та, 2010. - С. 212-217.

[10]. Дудин С.А. Вариабельность некоторых параметров периферического пульса от степени прижима артерии // // Инфокоммуникационные и вычислительные технологии и системы (ИКВТС-2010) : материалы III Междунар. конф., 6-11 сент. 2010 г., г. Улан-Удэ - оз. Байкал. - Улан-Удэ : Изд-во Бурят. госун-та, 2010. - С. 126-133.

[11]. Цыбиков А.С., Дармаев Т.Г., Занданова Г.И. Обработка данных электропунктурной диагностики с ипользованием многомерных методов математической статистики // Инфокоммуникационные и вычислительные технологии и системы (ИКВТС-2010) : материалы III Междунар. конф., 6-11 сент. 2010 г., г. Улан-Удэ - оз. Байкал. - Улан-Удэ : Изд-во Бурят. госун-та, 2010. - С. 279-282.

[12]. Цыбиков А.С. и др. Оценка состояния здоровья студентов методами восточной медицины с учетом психофизиологического типа // Вопросы гуманитарных наук. – 2010. - № 1 – С. 181-187.

СИБИРСКИЙ СУПЕРКОМПЬЮТЕРНЫЙ ЦЕНТР ИВМиМГ СО РАН: ВЫСОКОПРОИЗВОДИТЕЛЬНЫЕ ВЫЧИСЛЕНИЯ, ПРИКЛАДНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ.

Б.М. Глинский
И.М. Куликов
И.Г. Черных
ИВМиМГ СО РАН
chernykh@ssd.sscc.ru

Многие современные задачи из различных областей науки и промышленности требуют значительные вычислительные мощности: будь то прогноз погоды, сложные инженерные расчеты и техническое моделирование, обработка геофизических данных и анализ больших массивов информации. Решения таких задач невозможно без использования высокопроизводительных вычислительных систем.

Сибирский Суперкомпьютерный Центр на базе ИВМиМГ СО РАН имеет аппаратное обеспечение производительностью выше 16 TFLOPS. На аппаратном комплексе установлены последние версии компиляторов и параллельных отладчиков от Intel. Также на нем установлен широкий круг коммерческого и Open Source программного обеспечения. Среди них можно выделить такие пакеты как Ansys Fluent (вычислительная газо- и гидродинамика), Gaussian (квантовая химия). В последнее время библиотека доступного программного обеспечения стала пополняться разработками пользователей суперкомпьютерного центра, например библиотека PARMONC (PARallel MONte Carlo) для распараллеливания широкого круга трудоемких приложений метода Монте-Карло (автор Марченко М.А., ИВМиМГ СО РАН), готовятся пакеты по вычислительной астрофизике (группа авторов под руководством Вшивкова В.А., ИВМиМГ СО РАН).

В рамках работы Сибирского суперкомпьютерного центра решается широкий круг фундаментальных и прикладных задач. Один из примеров таких задач – численное моделирование 3-х мерного газодинамического течения в химическом реакторе с учетом химических реакций и одновременного ввода энергии через стенки реактора и излучения. В результате численного моделирования с помощью пакета Fluent получены следующие величины: распределение плотности мощности излучения в химическом реакторе с учетом поглощения части мощности, реагирующей газовой смесью, распределения температуры в реакторе для различных вариантов подогрева реагирующей газовой смеси. Также можно выделить задачу численного моделирования движения галактик в плотных скоплениях. Основной интерес представляет изучение пространственной динамики менее доступного для наблюдения газового компонента галактик, который играет одну из определяющих ролей в изучаемом процессе. Поэтому возникает необходимость суперкомпьютерного моделирования задачи столкновения галактик в газодинамическом приближении.

На примере этих задач авторы демонстрируют возможности программно-аппаратного комплекса Сибирского Суперкомпьютерного Центра ИВМиМГ СО РАН.

О ПРИБЛИЖЕННОМ РЕШЕНИИ УРАВНЕНИЯ РОСТА БИОМАССЫ

М.М. Бухурова
Учреждение Российской Академии наук Научно - исследовательский институт прикладной математики и автоматизации КБНЦ РАН
(г. Нальчик)

При разработке различных систем автоматизированного прогнозирования урожайности важное место занимают математические модели роста и развития основных сельхозкультур, участвующих в севообороте. В качестве математической модели рассматривается обобщенное уравнение Ферхюльста-Вольтерра [1]:

$$\frac{d}{dt}m(t) = m(t)[\varepsilon(t) + q(t)m(t) + \lambda D_{0t}^{-\nu}m(t)], \qquad (1)$$

где $D_{0t}^{-\nu}$ - оператор дробного интегрирования Римана-Лиувилля порядка - ν, $0 < \nu < 1$, $\varepsilon(t), q(t)$- непрерывные на временном сигменте $0 \leq t \leq T$ функции, учитывающие различные факторы среды обитания сельскохозяйственной культуры, $D_{0t}^{-\nu}m(t)$ - выражает влияние наследственных факторов отдельной культуры.

Задача Коши для уравнения (1) ставится следующим образом.

Задача Коши. *Найти положительное решение $m = m(t)$ уравнения (1), непрерывное на $0 \leq t \leq T$ и удовлетворяющее начальному условию*

$$m(0) = m_0. \qquad (2)$$

Рассмотрим уравнение

$$\frac{d}{dt}m(t) - \overline{m}q(t)m(t) = \overline{m}[\varepsilon(t) + \lambda D_{0t}^{-\nu}m(t)], \qquad (3)$$

где

$$\overline{m} = \frac{1}{T}\int_0^T m(t)dt. \tag{4}$$

Следуя А.М. Нахушеву [2] точное решение задачи (2) -(3) можно интерпритировать как приближенное решение задачи (1)-(2). Функция $m(t)$, являющаяся решением задачи (2)-(3) удовлетворяет следующему интегральному уравнению

$$m(t) = f(t) - \mu\int_0^t K(t,\eta)m(\eta)d\eta = f(t), \tag{5}$$

где $K(t,\eta) = \int_\eta^t \frac{exp[a(t)-a(\xi)]}{(\xi-\eta)^{1-\nu}}d\xi$,
$f(t) = m_0 exp[a(t)] + \overline{m}\int_0^t exp[a(t)-a(\xi)]\varepsilon(\xi)d\xi$,
$a(t) = \overline{m}\int_0^t q(\xi)d\xi, \mu = \frac{\overline{m}\lambda}{\Gamma(\nu)}$.

Уравнение (5) является уравнением Вольтерра 2 рода, единственное решение которого имеет вид:

$$m(t) = f(t) + \mu\int_0^t R(t,\eta;\mu)f(\eta)d\eta, \tag{6}$$

где $R(t,\eta;\mu)$-резольвента ядра $K(t,\eta)$.

Функция $m(t)$, определяемая равенством (6), можно считать приближенным решением задачи (1)-(4), при этом значение $\overline{m}$ однозначно определяется из (3) и (6).

Список литературы

[1] Нахушев А.М., Казиев В.М., Энеева Л.А., Ананчина Г.А., Нахушева З.А. К вопросу автоматизированного прогнозирования урожайности основных сельхозкультур в условиях орошения и степной зоны. Сб.: САПР и АСПР в мелиорации, Нальчик, 1983, с.156.

[2] Нахушев А.М. Об одном приближенном методе решения краевых задач для дифференциальных уравнений и его

приложения к динамике почвенной влаги и грунтовых вод
// Дифференц. уравнения. 1982. Т.18,№1.С.77-81.

THE DEVELOPMENT OF WINALT SIMULATION SYSTEM

M. Ostapkevich*, S. Piskunov
The Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia, email: ostap@ssd.sscc.ru
*Corresponding author

Key words: fine-grain parallelism, simulation system, parallel programming

Motivation and Aim: In recent years, simulation models of structures and algorithms with fine-grain parallelism have gained wide acceptance in describing physical, chemical, biological and social processes. Also as the technology of VLSI construction evolves, they are becoming more relevant to the design of computing devices with massive parallelism. Models with large amounts of data that take into consideration the essential features of simulated objects are interesting from both scientific and practical points of view. The participation of various specialists and the use of supercomputers are required in order to construct models of this kind. WinALT simulation system is constructed with these factors kept in mind. It is developed as a system with free access to modification of functions by a user and with an active use of the Internet both by an individual user and by a group of developers. The description of WinALT architecture, its simulation language, technology of model construction, a set of libraries that can be extended by user in order to tune WinALT to her or his problem domain are described in the report. The samples of models are presented as well.
Methods and Algorithms: An algorithmic system [1], which worked well in describing fine-grain algorithms and structures, was put into the basis of WinALT language. An open modular architecture makes it possible for a user to actively participate in the extension of its functionality by adding new modules. It is demonstrated how external modules can be used with the help the technology of component assembly, for example for the construction of such an application component as the WinALT site. A component for simulation on the basis of clusters of Windows hosts was implemented in order to improve the efficiency of simulation. Also, the requirements to its new version are formulated and its architecture and current state of its implementation is outlined. The new version provides significantly wider functionality. It supports parallelization of a bigger set of models with fine-grain parallelism (the older version supported only models of cellular automata including those with Margolus neighborhood). It can use all main kinds of contemporary parallel architectures (MPP, SMP, NOW, GPU) and some of their combinations for simulations.
Results: The stable release of WinALT simulation system is formed. The distributive package of the system is published at the WinALT site (http://winalt.sscc.ru/).The system has been tested on numerous samples, some of which are documented and published at the WinALT site.
Conclusion: The report discusses ways to improve system's functionality in order to turn it into such an active and versatile simulation environment (from the point of view of capabilities of simulation) with Web access that includes WinALT itself and resources of the Siberian Supercomputer Center, SB RAS
Availability: Free distributive, documentation and samples as well as overviews of various issues concerning fine-grain parallelism is available at the WinALT site.
References: 1. S.M. Achasova, O.L. Bandman, V.P. Markova, S.V. Piskunov: Parallel substitution algorithm. Theory and Application. World Scientific, Singapore (1994)

Существование нерассеивающих оболочек для модели анизотропной акустики

Г.В. Алексеев

Недавно был опубликован ряд работ, посвященных описанию так называемых нерассеивающих оболочек, т.е. пространственных структур, которые делают любой расположенный внутри указанной структуры объект свободным от электромагнитного либо акустического рассеяния. Эти оболочки играют роль зоны невидимости (либо неслышимости), позволяющей скрыть любой объект, расположенный внутри указанной оболочки. Соответствующее явление получило за рубежом название “electromagnetic or acoustic cloaking”, а сами оболочки стали называться “electromagnetic or acoustic cloaking shells”.

В настоящей работе возможность существования нерассеивающих оболочек для модели анизотропной акустики исследуется в рамках теории обратных задач. В соответствии с этим подходом задача построения нерассеивающей конструкции формулируется как обратная задача нахождения неизвестных коэффициентов обобщенного уравнения Гельмгольца, отвечающего рассматриваемой модели анизотропной акустики, по дополнительной информации о рассеяном искомой оболочкой поле. Для исследования указанной коэффициентной обратной задачи применяется подход, основанный на ее сведении к соответствующей задаче управления. Он заключается во введении функционала качества, адекватно отвечающего рассматриваемой обратной задаче, и сведении исходной обратной задачи к задаче минимизации указанного функционала качества. Последнее позволяет применять для ее решения хорошо развитые методы условной оптимизации. На основе указанного подхода развивается эффективный численный алгоритм решения рассматриваемой задачи. Обсуждаются результаты вычислительных экспериментов.

О бесконечномерных квазигруппах конечных порядков

В. Н. Потапов
Институт математики им. С. Л. Соболева СО РАН,
Новосибирский государственный университет, Новосибирск
vpotapov@math.nsc.ru

Пусть Σ — конечное множество мощности k, $\mathbb{A}$ — некоторое бесконечное множество индексов переменных, $\mathcal{F} \subseteq \Sigma^{\mathbb{A}}$ — подмножество состоящее из наборов с конечным носителем. Функция $f : \Sigma^{\mathbb{A}} \to \Sigma$ называется $\mathbb{A}$-квазигруппой порядка k, когда $f(\overline{y}) \neq f(\overline{z})$ для упорядоченных наборов $\overline{y}$ и $\overline{z}$, различающихся ровно в одной позиции. Доказано, что $\mathbb{A}$-квазигруппа f порядка 4 является разделимой (представимой в виде суперпозиции) или полулинейной на каждом смежном классе по $\mathcal{F}$.

Предложена конструкция неизмеримых по Лебегу множеств основывающаяся на бесконечномерных квазигруппах.

Ключевые слова: n-арная квазигруппа, разделимость, полулинейные квазигруппы.

EFFICIENT AND FAST PROTEIN FUNCTION PREDICTION

N. Pavlovic*[*], *I. Ivanoska, S. Kalajdziski
Ss. Cyril and Methodious University, Faculty of Electrical Engineering and Information Technologies, Karpos 2 bb, 1000 Skopje, Macedonia
e-mail: natasa.pavlovik@feit.ukim.edu.mk
[*]Corresponding author

Keywords: *Gene Ontology, CLUS, Ray protein descriptor, Voxel protein descriptor, Predicting protein function.*

Motivation and Aim: Proteins are the most important parts of all living organism. They can be found in all living systems, starting from bacteria and viruses, to the human. They are responsible for the biological processes that happen in the cell. Therefore, the knowledge on protein function is of great significance. Mainly protein function is determined in laboratory, it is expensive and time consuming task. Our motivation was to apply data mining methods to speed up the protein function determination for newly discovered proteins. Once the protein structure is known, the potential function of each protein can be predicted by comparing their structures. Two similar structures are very likely to have similar functions. The computer can not recognize the diversity of different protein structures; therefore we will use protein descriptors for unifying the representation of the structure.
Methods and Algorithms: Hierarchical multi-label classification (HMC) is a classification where instances may belong to multiple, hierarchically organized classes at the same time. It often happens that one protein has more than one parent. Therefore, as a system for structural and hierarchical representation of proteins and gene products which supports the DAG hierarchy, Gene Ontology (GO) is used. Here, we use two protein descriptors: Voxel and Ray-based descriptor. Voxel protein descriptor transforms the protein's tertiary structure into N–dimensional feature vector, and additionally gives some other protein structural features. Ray-based protein descriptor transforms the protein backbone into M–dimensional feature vector. CLUS is a system for predicting protein function that created the model (decision tree) trained by using the protein descriptor datasets. As an algorithm for hierarchical multi-label classification, the CLUS-HMC algorithm is used.
Results: The datasets for both the Voxel and Ray-based protein descriptors were consisted of 24.502 proteins. 10 fold cross validation was used for training and then testing the model. The CLUS system calculates the error measures which appear in the hierarchy. A score often used for comparison is the area between PR curve and the recall axis (AUPRC). The close the AUPRC is to 1.0, the better the model is. When Voxel protein descriptor is used, average AUPRC (weighted)=0.90, while when Ray-based descriptor is used, average AUPRC (weighted)=0.83. By comparison of the scores gained for Voxel and Ray-based protein descriptors, a new direction for further predicting of protein function is given.
Conclusion: By comparison of these values (0.90 > 0.83), can be concluded that the PR curve gained from the Voxel protein descriptor dataset covers bigger area with the PR axis. That means that by using the Voxel descriptor in generating a dataset used for protein's function prediction, better and more accurate results are gained than using the Ray-based descriptor.
Availability: The software and data are available on request from the authors.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ВОДНЫХ ЭКОСИСТЕМ

А.И. Абакумов
Институт автоматики и процессов управления ДВО РАН, Владивосток
e-mail: abakumov@iacp.dvo.ru

Аннотация

В 60-ых годах XX века А.А.Ляпунов сформулировал и развил ряд идей о модельных исследованиях водных экосистем, прежде всего, океанов и морей. В докладе предлагается система согласованных математических моделей для водных экосистем и сообществ водных организмов. С точки зрения биопродуктивности экосистемы наибольшее значение имеет функционирование фитопланктона. Рассмотрены модели динамики биомасс фитопланктонного сообщества, исследованы качественные свойства решений. Распределенные модели для фитопланктона позволяют использовать данные дистанционного зондирования океана, в частности, спутниковую информацию. Представлена модель вертикального распределения фитопланктона с учетом динамики содержания минерального питания и факторов внешней среды. Вычислительные эксперименты показывают, что с применением спутниковых данных о концентрации хлорофилла и минеральных веществ в поверхностном слое океана модель вертикального распределения позволяет оценить биомассу фитопланктона и выявить закономерности динамики основных характеристик фитопланктона в определенных районах океана.

Ключевые слова: математическая модель, водная экосистема, фитопланктон, дифференциальное уравнение, устойчивость.

MATHEMATICAL MODELLING OF THE AQUATIC ECOLOGICAL SYSTEMS

A.I. Abakumov
Institute of Automation and Control Processes, Far Eastern Branch of the RAS, Vladivostok

A.A. Lyapunov has formulated and has developed the ideas about modelling researches of aquatic ecosystems in oceans and in seas at sixtieth years of the twentieth century. The system of the coordinated mathematical models for aquatic ecosystems and communities of water organisms in the report is offered. The phytoplankton functioning has the greatest value for the biological productivity of the ecosystem. Models of the biomass dynamics of the phytoplankton are considered. Qualitative properties for solutions of these models are investigated. The distributed models for the phytoplankton may use the data of distant sounding of the ocean and the satellite information. The model of vertical distribution of the phytoplankton with taking into account of the dynamics of the concentration of mineral nutrients and of factors of an environment is presented. Calculative experiments show that the model of vertical distribution can estimate of the phytoplankton's biomass. This model can find dynamics regularities for basic properties of the phytoplankton on the base of the satellite data about concentration of chlorophyll and of mineral substances in a surface layer of the ocean.

Keywords: mathematical model, aquatic ecosystem, phytoplankton, differential equation, stability.

НЕЛОКАЛЬНАЯ КРАЕВАЯ ЗАДАЧА С ЛОКАЛЬНЫМ СМЕЩЕНИЕМ ДЛЯ ОБОБЩЕННОГО УРАВНЕНИЯ МАК КЕНДРИКА-ФОН ФЕРСТЕРА

Д. А. Шоранова
Учреждение Российской Академии наук Научно - исследовательский институт прикладной математики и автоматизации КБНЦ РАН
(г. Нальчик)

Пусть $u(x,t)dx$ - количество клеток в единице объема в возрастном интервале от $x \geq 0$ до $x + dx$ в момент времени t. Если убыль клеток определяется естественной гибелью, вымыванием в случае проточного культивирования и убылью из данного возраста x при делении родительской клетки, то при интуитивно ясных предположениях относительно других характеристик другой популяции можно записать:

$$u_t(x,t) + u_x(x,t) = c(x,t)u(x,t).$$

Рассмотрим в области $G = \{0 < x < l, 0 < t < T\}$ уравнение[1]

$$\partial_{0t}^{\alpha} u(x,t) + \partial_{0x}^{\beta} u(x,t) = c(x,t)u(x,t), \tag{1}$$

где $\partial_{0z}^{\mu} u(x,t)$- регуляризованная производная по Капуто порядка μ по переменной z, $0 < \alpha < 1$, $0 < \beta < 1$,

Задача. Найти регулярное решение $u(x,t)$ уравнения (1) в области G удовлетворяющее условиям

$$u(x,0) = \varphi(x), 0 < x < l \tag{2}$$

$$u(0,t) = \sum_{i=1}^{n} \alpha_i u(x_i,t), 0 < t < T \tag{3}$$

где φ -заданная функция, α_i - неотрицательные коэффициенты.

Краевые задачи для уравнения (1) исследовались в [см. 2, с.61 и библиографию там]. Отметим также [3], где для уравнения (1) исследовалась нелокальная краевая задача. Задача (1)-(3)относится к классу внутрикраевых задач с локальным смещением [4,с.29]

Вопрос о резрешимости задачи (1)-(3) сводится к решению системы интегральных уравнений Вольтерра II рода.

$$\begin{cases} u(x_1,t) = F(x_1,t) + \sum_{i=1}^{n} \alpha_i \int_0^t u(x_i,\eta) w(x_1, t-\eta) d\eta, \\ u(x_2,t) = F(x_2,t) + \sum_{i=1}^{n} \alpha_i \int_0^t u(x_i,\eta) w(x_2, t-\eta) d\eta, \\ ... \\ u(x_j,t) = F(x_j,t) + \sum_{i=1}^{n} \alpha_i \int_0^t u(x_i,\eta) w(x_j, t-\eta) d\eta \end{cases}$$

где

$$F(t) = \int_0^x \varphi(s) \Big(- \frac{1}{x-s} \Big) e_{\alpha,\beta}^{0,1} \Big(- \frac{(x-s)^\alpha}{t^\beta} \Big) ds dx +$$

$$+ \int_0^x \int_0^t c(s,\eta) e_{\alpha,\beta}^{\alpha,0} \Big(\frac{(x-s)^\alpha}{(t-\eta)^\beta} \Big) \frac{(x-s)^{\alpha-1}}{t-\eta} d\eta ds dx,$$

$$w(x_j,t) = \frac{1}{t-\eta} e_{\alpha,\beta}^{1,0} \Big(- \frac{x_j^\alpha}{(t-\eta)^\beta} \Big), \qquad (4)$$

где $e_{\alpha,\beta}^{\mu,\delta}(z)$- функция типа Райта [2], $j = \overline{1,n}$.

С биологической точки зрения $u(x,t)$ - численность популяции возраста $x \in (0,l)$ в момент времени $t \in (0,T)$, а уравнение (1) моделирует динамику численности эволюционирующих популяций.

После определения функций $u(x_i,t)$, функция $u(x,t)$ находится как решение краевой задачи для уравнения (1) [2],[3].

Список литературы

[1] Нахушев А. М. Уравнения математической биологии. — М.: Высш. шк., 1995. —301 с.

[2] Псху А.В. Краевые задачи для дифференциальных уравнений с частными производными дробного и континуального порядка. —Нальчик, 2005. —185 с.

[3] Шоранова Д.А.Нелокальная краевая задача для уравнения рождаемости Мак Кендрика-фон Ферстера с дробной производной Капуто по времени. Материалы Международной Российско - Болгарской VIII школы молодых ученных "Нелокальные краевые задачи и проблемы современного анализа и информатики"Нальчик - Хабез, 2010 г. — 120-122 с.

[4] Нахушев А. М. Задачи со смещением для уравнений в частных производных. —М.: Высш. шк., 2006. —287 с.

Сравнение сложности вычисления функций q-значной логики двумя классами ветвящихся программ

М. С. Громов
Новосибирский Государственный университет
mgtriffid@gmail.com

Детерминированной ветвящейся программой от переменных $x_1, \ldots, x_n$ называется ориентированный граф без контуров с одной входной вершиной и двумя выходными вершинами, одна из которых помечена нулем, а другая – единицей. Все невыходные вершины помечены переменными из множества $\{x_1, \ldots, x_n\}$. Под сложностью детерминированной ветвящейся программы понимается число вершин программы, помеченных переменными. Булева функция $f(x_1, ..., x_n)$ принимает значение 1 на наборе $(a_1, \ldots, a_n)$, если существует путь от входной вершины к выходной вершине, помеченной единицей, который из вершин, помеченных переменной x_i, проходит по дугам, помеченным a_i.

Это определение можно естественным образом расширить на случай q-значной логики. Теперь число выходных вершин равно q, из каждой вершины, кроме выходных, выходит ровно q дуг. Такая ветвящаяся программа будет реализовывать функцию $f : \{0, \ldots, q-1\}^n \to \{0, \ldots, q-1\}$.

Будем рассматривать ветвящиеся программы, у которых порядок переменных, которые встречаются на пути от входной вершины до выходной, не зависит от пути, то есть для любого аргумента в процессе работы программы переменные встречаются в фиксированном порядке. Программы такого вида называются упорядоченными ветвящимися программами (англ. Ordered Decision Diagram)

Другое ограничение на структуру программы – ограничение на число проверок переменных в цепи, когда в любой цепи, идущей от входной вершины к выходной, вершины, помеченные любой переменной, встречаются не более k раз. Такие программы называются ветвящимися k-программами (англ. Read-k-times-only decision diagram).

Пусть K – некоторый класс ветвящихся программ. Сложностью вычисления функции f в классе K называется сложность минимальной программы из класса K, вычисляющей функцию f. Эта величина обозначается $C_K(f)$

Рассмотрим два класса ветвящихся программ: $1ODD$ — класс упорядоченных ветвящихся 1-программ и $2ODD$ — класс упорядоченных ветвящихся 2-программ.

Ясно, что класс $1ODD$ является подмножеством класса $2ODD$, поскольку если переменная встречается на пути не более одного раза, то тем более и не более двух раз.

Рассмотрим последовательность функций $F_1, F_2, \ldots, F_N, \ldots$ q-значной логики, $q \geqslant 3$, где функция F_N определяется следующим образом: F_N зависит от N^2 переменных и принимает значение 1, если аргумент как слово в алфавите $\{0, \ldots, q-1\}$ содержит под-

слово вида XX длины $2N$, и принимает значение 0, если такого подслова в аргументе нет.

Например, функция $F_4(x_1, x_2, \ldots, x_{15}, x_{16})$ принимает значение 1 на наборе $(1,0,4,1,3,2,4,0,$ поскольку в слове 1041324032401131 имеется подслово 32403240, но она принимает значение 0 на наборе $(1,4,5,1,2,4,3,1,2,1,3,5,1,0,0,2)$, несмотря на наличие подслова 00, так как его длина 2, а не 8, как требуется.

В настоящей работе даны полиномиальная верхняя оценка сложности функции F_N в классе упорядоченных ветвящихся 2-программ и сверхполиномиальная нижняя оценка сложности функции F_N в классе упорядоченных ветвящихся 1-программ.

Полиномиальная верхняя оценка. Сложность $C_{2ODD}(F_N)$ вычисления функции F_N в классе упорядоченных ветвящихся 2-программ не превосходит $(N(q+1)+1)(N^2-2N)$.

Сверхполиномиальная нижняя оценка. Сложность $C_{1ODD}(F_N)$ вычисления функции F_N в классе упорядоченных ветвящихся 1-программ превосходит 2^{N-1}.

Очевидным следствием этих двух оценок является следующая

ТЕОРЕМА Существует последовательность функций q-значной логики ($q \geqslant 3$), для которых сложность реализации в классе упорядоченных 1-программ превосходит в сверхполиномиальное (от количества переменных) число раз сложность реализации в классе упорядоченных 2-программ.

РАЗРАБОТКА МНОГОФАКТОРНОЙ СИСТЕМЫ ПРОГНОЗИРОВАНИЯ ДЛЯ УПРАВЛЕНИЯ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

Иванюк В.А.
Институт проблем управления им. В.А. Трапезникова РАН, г. Москва
Ivaver6@gmail.com
Цвиркун А.Д.
Институт проблем управления им. В.А. Трапезникова РАН, г. Москва
tsvirkun@ipu.ru

Целью данной работы является построение инновационной, самообучающейся вычислительной системы для комбинированного многофакторного анализа и прогнозирования с использованием косвенных данных и учетом явлений, влияющих на исследуемый показатель.

В разработанной системе применен метод анализа и прогнозирования данных на основе генетического алгоритма. Система в отличие от аналогов может выдавать экономическую оценку и прогнозы с более высокой точностью.Основными блоками программы являются блок статистического анализа, блок нейронного анализа, блок генетического алгоритма (Рис 1).

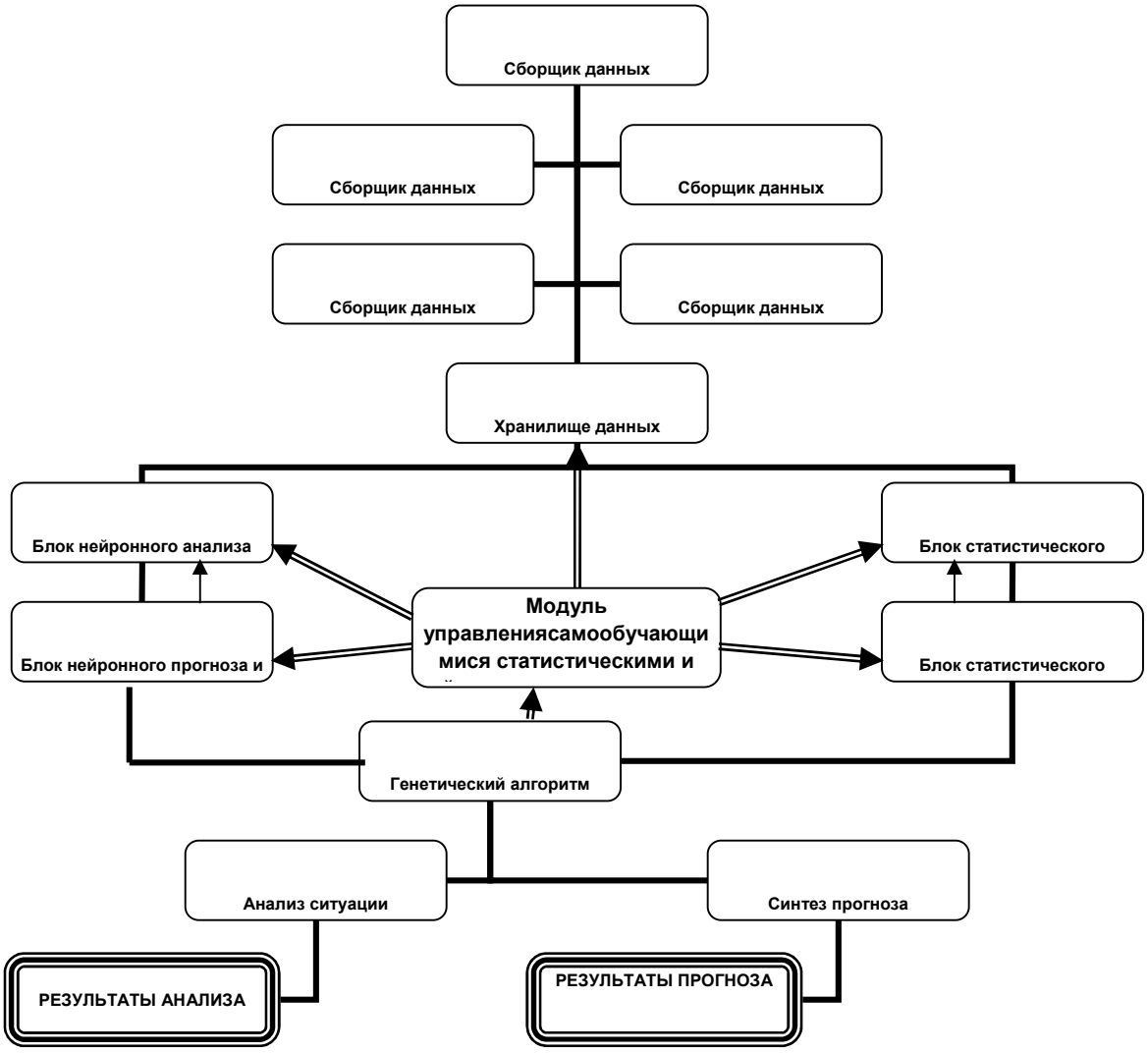


Рис.1