

Оптимальный метод перераспределения общей памяти для двухприоритетной очереди, представленной в виде двух последовательных циклических FIFO-очереди

Аксенова Е. А., Соколов А. В.

(Петрозаводск, Институт прикладных математических исследований КарНЦ РАН, Петрозаводский государственный университет)

aksenova@krc.karelia.ru, avs@krc.karelia.ru

Во многих приложениях используется структура данных, в которой основными операциями являются вставка элемента и удаление элемента с наибольшим приоритетом. Такую структуру данных называют приоритетной очередью. Основными методами реализации приоритетной очереди являются упорядоченные и неупорядоченные списки, массивы, бинарные деревья, пирамиды [1]–[3]. Эти методы сравниваются с точки зрения стоимости (времени) выполнения основных операций. В работах [4]–[7] предлагались математические модели для последовательного циклического и связанного способов представления очередей и решались задачи оптимального начального распределения памяти для разных критериев оптимальности.

Целью данной работы является изучение динамики изменения длины очереди в случае представления n -приоритетной очереди в виде n последовательных FIFO-очереди, когда все элементы с одинаковыми приоритетами помещаются в одну FIFO-очередь. Каждой FIFO-очереди выделен свой участок памяти, поэтому нет затрат памяти на хранение приоритетов. Одна из FIFO-очереди может исчерпать свою память, а у остальных очередей в этот момент еще будет свободная память. При этом допускается возможность переполнения FIFO-очереди, что, например, в маршрутизаторах является обычной ситуацией. В данной работе предлагается математическая модель и алгоритм оптимального перераспределения памяти после переполнения одной из FIFO-очереди. В качестве критерия оптимальности рассмотрено максимальное среднее время работы до переполнения.

Рассмотрим очередь с двумя приоритетами, расположенную памяти размера m единиц. Будем считать, что время дискретно, и в

каждый момент времени происходит одна из следующих операций:

- вставить элемент с приоритетом i с вероятностью p_i ($i = 1, 2$),
- удалить элемент с наибольшим приоритетом с вероятностью q ,
- найти элемент с наибольшим приоритетом с вероятностью r , где $p_1 + p_2 + q + r = 1$.

Вероятности не зависят от времени и количества элементов в очереди. Наивысший приоритет равен 2, наименьший — 1. Работа начинается с пустой очереди и не завершается при исключении элемента из пустой очереди. Исключение элемента из очереди происходит по наивысшему приоритету. Это означает, что пока вторая FIFO-очередь не пуста, с вероятностью q исключение элементов происходит из этой очереди. Как только вторая FIFO-очередь станет пустой, с вероятностью q исключение элементов будет происходить из первой FIFO-очереди. Предполагаем, что при переполнении одной из FIFO-очереди происходит перераспределение свободной памяти между очередями и работа с программной системой продолжается.

Выделим первой FIFO-очереди s единиц памяти, тогда для второй останется $m - s$ единиц. Обозначим текущие длины очередей x_1 и x_2 . В качестве математической модели рассмотрим случайное блуждание по целочисленной решетке в прямоугольной области на плоскости $0 \leq x_1 < s + 1$, $0 \leq x_2 < m - s + 1$. Блуждание начинается в точке $x_1 = 0, x_2 = 0$. Переполнению первой FIFO-очереди соответствуют точки на прямой $x_1 = s + 1$, переполнению второй — точки на прямой $x_2 = m - s + 1$. Прямые $x_1 = -1$ и $x_2 = -1$ — отражающие экраны, то есть при исключении элемента из пустой очереди работа не завершается.

Требуется оптимально перераспределить свободную память между FIFO-очередями после переполнения одной из них. То есть, при попытке включения элемента в заполненную FIFO-очередь, когда $x_1 = s$ (или $x_2 = m - s$), требуется определить новую область блуждания $0 \leq x_1 < s^* + 1$, $0 \leq x_2 < m - s^* + 1$, то есть такое значение s^* , где $s^* > s$ (или $m - s^* > m - s$), чтобы время до следующего переполнения какой-либо FIFO-очереди было максимальным. Этот процесс перераспределения можно продолжать до полного исчерпания свободной памяти. Такой подход оправдан, если при переполнении одной из FIFO-очереди, в области памяти, выделенной для других очередей, есть достаточно свободной памяти. В алгоритме Гарвика [1] для

работы с n последовательными стеками и очередями при переполнении какой-либо структуры данных приблизительно 10% свободной памяти делится поровну между структурами данных, а оставшиеся 90% делятся пропорционально росту размеров структур данных с момента предыдущего распределения памяти.

Математическая постановка задачи сводится к решению задачи нелинейного целочисленного программирования, с критерием оптимальности, заданным алгоритмически. Случайное блуждание по целочисленной решетке будем рассматривать как конечную однородную поглощающую цепь Маркова с матрицей вероятностей переходов P [8]. Для решения задачи требуется матрица Q вероятностей переходов из невозвратных состояний в невозвратные (Q — подматрица матрицы P). Вводится нумерация состояний Марковской цепи, это позволяет определить структуру матрицы Q для любого размера памяти m и любого значения s . Среднее время работы до переполнения будем искать с помощью фундаментальной матрицы $N = (E - Q)^{-1}$. Предполагаем, что в самом начале работы, когда обе очереди пустые $x_1 = 0, x_2 = 0$, память между ними тоже нужно разделить оптимально, то есть выбрать такое s , чтобы время работы с очередями до переполнения было максимально.

Алгоритм:

- 1) Ввод вероятностных характеристик приоритетной очереди, размера памяти m ;
- 2) Для каждого значения s , $0 \leq s \leq m$, генерируем матрицу Q , вычисляем матрицу N ;
- 3) Для каждого значения s суммируем элементы матрицы N в строке, соответствующей состоянию $x_1 = 0, x_2 = 0$;
- 4) Выбираем такое значение s , которому соответствует максимальная сумма элементов в строке соответствующей матрицы N ;
- 5) Для каждого состояния, в котором одна из FIFO-очереди заполнила выделенную память, то есть $x_1 = s$ или $x_2 = m - s$, перебираем значения $0 \leq s^* \leq m$, где $s^* > s$, если $x_1 = s$, и $m - s^* > m - s$, если $x_2 = m - s$. Для каждого s^* генерируем новую матрицу Q , вычисляем матрицу N ;

- 6) Для каждого значения s^* суммируем элементы матрицы N в строке, соответствующей рассматриваемому состоянию ($x_1 = s$ или $x_2 = m - s$);
- 7) Выбираем такое значение s^* , которому соответствует максимальная сумма элементов в строке соответствующей матрицы N . Для каждого состояния, соответствующего тому, что одна из очередей заполнила выделенную память, получаем оптимальное значение s^* .

Шаги 5–7 можно повторять до полного исчерпания свободной памяти (если перед каждым новым прераспределением текущее значение s^* обозначить как s).

Вычисление критерия оптимальности в данной задаче является очень ресурсоемким за счет обращения матрицы $(E - Q)$ большого размера. Матрицу N можно представить в виде $N = (E - Q)^{-1} = E + Q + Q^2 + \dots = \sum_{k=0}^{\infty} Q^k$. Для данной задачи такое представление матрицы дает выигрыш в размере памяти для вычислений. Поскольку для алгоритма важна сумма элементов в определенной строке матрицы N , то запись в виде суммы ряда позволяет вычислять элементы конкретной строки, не вычисляя остальных элементов матрицы.

Работа выполнена при финансовой поддержке РФФИ, грант 09-01-00330.

Список литературы

- [1] Кнут Д. Искусство программирования для ЭВМ. — М.: Вильямс, 2001.
- [2] Седжвик Р. Фундаментальные алгоритмы на C++. — К.: Диасофт, 2001.
- [3] Боллапрагада В., Мэрфи К., Уайт Р. Структура операционной системы Cisco IOS. — М.: Вильямс, 2002.
- [4] Аксенова Е. А., Соколов А. В. Некоторые задачи оптимального управления FIFO-очередями // Труды Второй Всероссийской научной конференции «Методы и средства обработки информации». — М.: Изд. отдел ВМК МГУ им. М. В. Ломоносова, 2005. — С. 318–322.

- [5] Аксенова Е. А. Оптимальное управление FIFO-очередями на бесконечном времени // Межвузовский сборник «Стохастическая оптимизация в информатике». — СПб: Изд-во С.-Петербургского ун-та, 2006. — С. 71–76.
- [6] Аксенова Е. А., Драц А. В., Соколов А. В. Об оптимальном управлении FIFO-очередями на бесконечном времени // Обзорные прикладной и промышленной математики. — 2009. Т. 16. Вып. 3. — С. 401–415.
- [7] Аксенова Е. А., Драц А. В., Соколов А. В. Оптимальное управление n FIFO-очередями на бесконечном времени // Информационно-управляющие системы. — 2009. № 6. — С. 46–54.
- [8] Кемени Дж., Снелл Дж. Конечные цепи Маркова. — М.: Наука, 1970.

Построение и анализ детерминированных методов прогнозирования

Андреев А. В., Пытьев Ю. П. (Москва, МГУ
им. М. В. Ломоносова)

alvlandr@gmail.com, yuri.pytyev@gmail.com

В работе рассмотрены математические методы краткосрочного прогнозирования.

Целью данной работы являлось построение и исследование методов, обеспечивающих высокую точность прогноза стоимостей акций на l шагов вперёд («на l следующих дней»).

Прогнозирование стоимости акций на l «следующих дней»

Сформирован портфель из акций пяти различных компаний. Известна динамика изменения курса этих акций — стоимость акций за определённый период времени: n дней. Требуется оценить стоимость акции каждой компании в период с $(n+1)$ -го по $(n+l)$ -ый день. В работе представлены три подхода, позволяющие решить поставленную задачу.

Метод с предиктором.

Данный метод является методом линейного прогнозирования, аналогичным методу наименьших квадратов (МНК). В его основе лежит решение следующей задачи на минимум:

$$\|YX - Z\| \sim \min_{X \in R^M} \quad (1)$$

где Y — матрица, каждая строка которой составлена из пяти временных рядов (по количеству типов акций) одинаковой длины, элементами (уровнями) которых являются стоимости акций, следующая строка получается сдвигом предыдущей на один шаг вперёд (таким образом, что в каждой следующей строке кроме первой используется пять новых значений относительно предыдущей), X — предсказывающая матрица или предиктор, Z — матрица, каждая строка которой состоит из $5l$ значений стоимостей акций, опережающих значения в

соответствующей строке матрицы Y на число шагов от 1 до l . Проще говоря, при умножении одной строки матрицы Y на предиктор появятся $5l$ будущих значений стоимостей акций, составляющих матрицу Z .

Вероятностный аналог данного метода рассмотрен в [2].

Метод, минимизирующий среднеквадратичное отклонение.

Пусть A_k — матрица размера $5 \times m$, каждая из пяти строк которой (по числу типов акций) представляет собой временной ряд длины m , $\alpha_k: 5 \times l$ — вектор, составленный из цен акций, опережающих соответствующие значения в A_k на один шаг (на один день). Введём векторы $x: m \times l$ и $y: 5 \times l$ такие, что преобразование $A_k x + y$ было как можно ближе в среднеквадратичном к α_k . Таким образом, необходимо решить следующую задачу:

$$\frac{1}{s-1} \sum_{k=m+1}^{s+m} \|\alpha_k - A_k x - y\|^2 \sim \min_{x,y} \quad (2)$$

Требуется решить задачу на минимум (2) и найти x, y , если известны вектор α_k и матрица A_k .

Анализ методов.

В результате проведённого численного исследования построенных методов был получен ответ на вопрос: сколько предыдущих значений следует использовать при прогнозе будущего. Данная характеристика соответствует значению величины m . Усреднённая относительная ошибка прогноза оказалась наименьшей при $m = 2$. Однако такое значение параметра являлось не равномерно наилучшим, поэтому были разработаны два адаптивных алгоритма выбора m . Первый алгоритм вычислял прогноз при нескольких значениях m , и то значение, при котором относительная ошибка оказывалась наименьшей, выбиралось для прогноза на следующий период. Во втором алгоритме подсчитывалось, как часто то или иное значение m обеспечивало наименьшую ошибку, и в дальнейшем значения параметра разыгрывались, то есть применялась рандомизация. Реализация этих алгоритмов показала, что такой адаптивный выбор m не позволил повысить качество прогноза.

Применение сглаживающих сплайнов в задаче прогнозирования.

Проведённый анализ, показавший, что качество прогноза получается в среднем наилучшим при выборе параметра $m = 2$, позволяет предположить, что для решения поставленной задачи можно использовать сплайны. В частности, при решении задачи интерполяции через точки $(t_1, r_1), \dots, (t_n, r_n)$ требуется провести в известном смысле самую гладкую кривую. Можно показать, что решением этой задачи является сплайн. Однако гладкость сплайна, проходящего через заданные n точек, может оказаться недостаточной. В таком случае не следует требовать, чтобы сплайн точно проходил через указанные точки. Это тем более разумно, если точки получены из приблизительных измерений. Компромисс между гладкостью кривой и точностью аппроксимации при $t = t_1, t_2, \dots, t_n$ можно реализовать, решая задачу на минимизацию:

$$\min \left\{ \int_a^b (f^{(q)})^2(t) dt + \rho \sum_{j=1}^n (f(t_j) - r_j)^2 : f \in H^q \right\} \quad (3)$$

при должным образом выбранном $\rho > 0$, где H^q — гильбертово пространство функций на $[a, b]$ с абсолютно непрерывной $q - 1$ производной и суммируемых в квадрате q производной:

$$\int_a^b (f^{(q)})^2(t) dt < \infty \quad (4)$$

Решением этой задачи также является сплайн, называемый сглаживающим, его гладкость тем выше, чем меньше ρ .

Результаты и выводы

Предложены методы прогнозирования изменения динамики курсов акций, обеспечивающие точность предсказания сравнимую с уже известными методами. Для сравнения методов между собой использовалась величина $E^{(k)}$, значение которой для каждого метода характеризует качество прогноза, получаемого при использовании соответствующего метода:

$$E^{(k)} = \frac{1}{N} \sum_{i=1}^N \frac{|k_i - k_i^{pr}|}{k_i}, \quad (5)$$

где k_i — стоимость акции в i -ый день, k_i^{pr} — полученное в результате прогноза значение стоимости акции в i -ый день, N — количество сделанных прогнозов.

Результаты использования описанных выше методов приведены в таблице 1. Помимо разработанных подходов в сравнении участвовали методы Брауна и Хольта. Полученные значения ошибок указывают на то, что все эти методы сопоставимы, и из них нельзя выделить один равномерно наиболее точный. В тоже время можно заметить, что средняя ошибка прогноза метода с предиктором равномерно выше ошибки метода, минимизирующего среднеквадратичное отклонение. Разница между первым и вторым методами такая же, как между МНК и математической редукцией, и проигрыш метода с предиктором в качестве прогноза связан с тем, что в первом случае мы минимизируем невязку, а во втором случае — ошибку.

	Сбер-банк	7 Кон-тинент	Север-сталь	МСРК	Лукойл
МП	0.039	0.032	0.036	0.025	0.027
ММСК	0.042	0.025	0.033	0.022	0.026
МС	0.038	0.030	0.034	0.025	0.028
МБ	0.049	0.034	0.043	0.033	0.034
МХ	0.036	0.027	0.032	0.024	0.027

Таблица 1. Значения $E^{(k)}$ для каждого метода. Параметр $m = 2$. МП — метод с предиктором, ММСК — метод, минимизирующий с.к. отклонение, МС — метод, использующий сглаживающие сплайны, МБ — метод Брауна, МХ — метод Хольта.

Судя по полученным результатам, не следует ожидать улучшения качества прогноза без учёта модели данных. Если модель известна приближённо, то можно поставить задачу оптимального предсказания. Приведём пример модели данных, в которых предсказание на основе обучения невозможно. Рассмотрим двоичное представление

$\sqrt{2}$: $e = 1.0110\dots 1_n\dots$, $e_0 = 1, e_1 = 0, \dots, e_n = 1$ и последовательность $\{e_i\}$. Известно, что элементы этой последовательности независимы в совокупности, поэтому невозможно предсказать значение в любом разряде (e_n), на основе значений элементов последовательности в старших разрядах (e_0, \dots, e_{n-1}). Иначе говоря, методы построения предиктора, рассмотренные выше, не позволят решить задачу предсказания. Тем не менее, если модель известна, то значение e_n определяется точно, причём без использования e_0, \dots, e_{n-1} .

В общем случае, зная точную модель, задача отыскания оптимального предиктора, сводится к следующей:

$$X = \arg \min_X Err(X), \quad (6)$$

где $Err(X)$ — ошибка прогноза. Однако если модель известна, то приближённо. Пусть имеется параметрический класс моделей и класс предикторов. Выберем произвольно модель и предиктор из класса и вычислим ошибку прогноза. Из всех моделей выберем наихудшую, то есть ту, для которой ошибка максимальна:

$$Err(X, \theta) \sim \max_{\theta}. \quad (7)$$

Решением задачи прогнозирования, будет минимаксный предиктор X , являющийся решением следующей задачи на минимум:

$$\max\{Err(X, \theta)|\theta\} \sim \min_X. \quad (8)$$

Список литературы

- [1] Бокс Дж., Дженкинс Г. Анализ временных рядов: прогноз и управление. 1974.
- [2] Цыплаков А. Введение в прогнозирование в классических моделях временных рядов // Квантиль. № 1. С. 3–19.
- [3] Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов. — М.: Финансы и статистика, 2003.
- [4] Дегтярёв К. Прогнозирование валютных котировок с использованием модифицированного стационарного метода, основанного на нечетких временных рядах.

Конфликты между методами математики и механики и техногенные катастрофы

Величенко В. В. (Москва, Институт машиноведения РАН, МГУ им. М. В. Ломоносова)

vlad.velichenko@mail.ru

Рассматриваются технические и теоретические проблемы техногенных катастроф. Детально анализируются конфликты между методами математики и механики, приводящие при их формально правильном использовании к ошибочным выводам. Приводятся примеры катастрофических ошибок в технических расчетах в результате использования формально правильных методов математики и механики.

Введение

Аварии и катастрофы свирепствуют во всех средах нашего обитания и во всех областях техники — на земле и под землей, на поверхности моря и в его глубинах, в воздухе и в космосе. Техника и наука осознают необходимость борьбы с этим злом, но усилия, предпринимаемые в направлении борьбы с катастрофами, совершенно недостаточны, и полностью перекрываются усилиями тех же техники и науки по воспроизводству все новых катастроф. Создаются конструкции все большей высоты, протяженности и объема, все большей мощности — и этим уже закладываются будущие глобальные катастрофы с массовыми жертвами.

Назрела необходимость в развитии специальных разделов науки и техники, занимающихся проблемами борьбы с катастрофами. Эти проблемы более сложны, чем все предыдущие, но надежды на их решение силами современных «развитых» науки и техники — нет. Их работники уже заняты — производством все новых техногенных катастроф, и у них уже нет личного физического времени для решения этих долговременных проблем.

Единственная надежда на создание этих новых областей науки и техники — сегодняшние студенты. Им автор и адресует эту статью, которую можно считать введением в эту новую специальность. Специальность эта заключается в борьбе со страшным противником, и

автор говорит о нем неприкрашенную правду. Но рассказать студентам, почему именно утонули такие знаменитые океанские корабли, как «Титаник» и «Конкордия», и почему упали такие знаменитые космические корабли, как «Челленджер» и «Фобос», в этой статье невозможно — не хватит места, и требуются предварительные специальные знания. Мы расскажем здесь о вещах более доступных, уже отчасти известных студентам, — и еще более важных.

Теория и техника борьбы с катастрофами еще не созданы. В этой будущей науке будет много специальных разделов. Здесь мы подробно рассмотрим только один из самых первых. Как раз тот, который уже находится в руках студентов — будущих исследователей этой интереснейшей и благородной области знаний. Этот важнейший раздел помещается в тех самых учебниках математики и механики, которые вы открывали и читали много раз, не подозревая, что напечатанные в них безупречно верные истины могут превратиться — в бомбы разрушительного теоретического терроризма.

Материал этой статьи — почти элементарный, доступен студентам 2–3 курсов, но затрагивает фундаментальные вопросы математики, механики и техники.

1. Технические и теоретические проблемы техногенных катастроф

Экспертные заключения по разбору многочисленных, резко участвовавших в последнее время аварий сложных технических объектов и техногенных катастроф отмечают две их главные причины: «техническую» — поломки техники, и «человеческий фактор» — неправильное управление объектом. Каждая из них, в свою очередь, является сложной, состоящей из многих компонентов. Чтобы перечислить причины поломок техники, нужно указать все этапы ее проектирования и все технологические процессы ее изготовления. «Человеческий фактор» тоже сложен, и его описание совпадает с общим описанием человека и его жизнедеятельности — его работой и отдыхом, образованием и психологией, физической подготовкой и дисциплиной, и т. д. На границе между этими причинами находится нарушение регламентов эксплуатации, когда работники вынуждены «правильно»

использовать технику, уже пришедшую в негодность, и находясь при этом в состоянии запредельного утомления.

Эти причины лежат на поверхности. Глубже лежат причины менее очевидные. Например, когда уважаемые ученые под давлением «государственной необходимости», а на самом деле примитивного экономического поощрения, «обоснованно» продлевают сроки ресурсов уже отработавших свой регламент опасных установок и производств, в том числе и таких, как атомные электростанции. Богатую питательную среду для роста катастроф культивируют экономика и политика. Авиационные корпорации в погоне за прибылью соревнуются в изготовлении огромных самолетов, — и в непредсказуемой атмосфере, запечатанные в этих ненадежных гигантах сотни (в последних моделях — почти тысяча) пассажиров загадывают итог *каждого* рейса: пронесет — не пронесет. Конкурирующие автомобильные магнаты рекламируют и производят автомобили все большей, ничем и никем не обоснованной и не ограниченной мощности, и обыватель, получив в свои руки ракету на колесах, становится убийцей на большой дороге, — и самоубийцей. Физика создает энергостанции все большей мощности — обещания будущих глобальных катастроф. Ученые обсуждают на конгрессах негативные последствия потепления климата нашей планеты — а техника судорожно тратит остатки природных ресурсов на ее нагрев. Военные техника и наука занимаются профессиональным производством катастроф — вообще не задумываясь об их последствиях. Правители получили в свои руки оружие, способное на любом расстоянии от их безопасных кабинетов поразить любую цель, и превращают слабые страны в руины, — разрушая тем самым и свою будущую безопасность. Эти причины — предмет для исследований прокуроров и геополитиков.

К предмету настоящей статьи ближе еще одна важнейшая причина аварий и катастроф — ошибки проектирования и расчета потерпевшего катастрофу объекта. Эта «теоретическая» причина остается практически неотмеченной в протоколах аварийных комиссий, и никогда публично не обнародуется, хотя неявно всегда учитывается во временном запрете эксплуатации аналогичных объектов до выяснения причин аварии. Вместе с тем, она является одной из главных: поломки техники являются результатом неправильных расчетов, и

причинами ошибок человека-оператора являются неправильно сконструированные системы управления. Эта теоретическая причина, в свою очередь, тоже является сложной и многоплановой, и ее основные составляющие хорошо известны специалистам: это ошибки методик и конкретных расчетов.

В связи с конструктивным усложнением техники, в качестве важной причины выступает и недостаточное обеспечение проектов вычислительными средствами. Эта причина, по-видимому, — вечная. Инженеру всегда не хватало возможностей грифельной доски, логарифмической линейки, сегодня — компьютера. Эта причина также лежит на поверхности, и очевидна. Однако в последнее время появилась еще одна, неявная и противоположная причина — достаточное, или избыточное обеспечение проекта вычислительными средствами. Возможно, именно она становится основной опасностью для современной техники, развивающейся столь быстро, что ее не успевает даже грамотно описать современная математика. — Ошибочная методика, пропущенная через многодневные вычисления, и снабженная многочисленными таблицами и убедительными графиками, приобретает респектабельный облик достоверной теории! А запущенные в практику, такие расчеты в лучшем случае приводят к неработоспособным изделиям, а в худшем — ждут подходящего случая, чтобы взорвать и уничтожить внешне успешно работающий объект.

Ошибки проектирования и расчетов ведут к авариям и катастрофам объектов, но одновременно разрушают эти проекты и расчеты. Имея в виду эти две деструктивных характеристики, их можно назвать *катастрофами проектов* и *катастрофами расчетных методов*.

Однако недостатки плохих проектов с течением времени выявляются, неверные методики и ошибочные вычисления в конце концов все же могут быть исправлены. Мы же рассмотрим здесь причину самую глубокую, *полностью скрытую от наблюдения* — систематические методы получения *неверных* выводов из *правильных* вычислений при *правильном* использовании *верных* теорий математики и механики. Это двойная катастрофа: неверные результаты расчетов ведут к техническим катастрофам, а неверный результат правильных теорий является катастрофой этих теорий. Мы назовем такие скрытые ошибки *теоретическими катастрофами*.

Эти вопросы должны быть предметом специального изучения профессиональных математиков и механиков. Но для них они представляют только теоретический интерес, тогда как в прикладных областях, в руках инженеров, они непосредственно ведут к реальным катастрофам. Эти прикладные области в условиях современных требований совершенствования и интенсификации техники чрезвычайно сложны, требуют обширных и глубоких специальных знаний, и умения владеть многими экспериментальными и теоретическими инструментами. Среди прочих, конечно, математикой и классической механикой — самыми надежными средствами науки, веками проверенными практикой. Перечислениям имен ученых, разработавших мощные инструменты математики и механики, и их достижениям посвящены энциклопедии, и мы тоже будем ссылаться на гениев этих наук — Галилея, Ньютона, Лагранжа, Эйлера [1]–[3], [4]. Но наша цель здесь другая: не продемонстрировать еще раз замечательные возможности этих инструментов математики и механики, а, наоборот, указать на *катастрофические конфликты* между этими безупречными науками.

Эта статья адресована студентам разных специальностей — теоретических, и технических. Мы говорим здесь о теоретических методах математики и механики, но их непосредственными потребителями является техника, она в первую очередь заинтересована в доброкачественности этих теоретических продуктов, и должна поэтому относиться к ним предельно критически. Сами математика и механика, развивавшиеся в прошлом в атмосфере соревнования, острой критики и бурных дискуссий, с завершением их основных теорий практически полностью утратили способность самокритики. В них сегодня царят тишь и благодать, редко нарушаемые «крупным математическим достижением» в решении частной задачи, которой для весомости присваивается термин «проблема». А механика вообще считается завершенной дисциплиной, в которой остается лишь решать конкретные задачи методами, завещанными ее классиками столетия назад. Здесь критика стала вообще неуместной, как в храме, или на кладбище.

Поэтому эта критическая статья автора-математика представлена как взгляд на методы математики и механики «со стороны», с точки зрения их потребителей — инженеров. Все разработки современной

техники начинаются и завершаются математическими методами, а основным их объектом являются механические системы. В статье это обстоятельство подчеркивается несколькими примерами грубейших ошибок в формировании моделей механических объектов, к которым приводят безупречные методы математики и механики.

Помимо этих объективных фактов, целью статьи является также указание на необходимость изменения психологии инженера-разработчика. Если математика и механика являются его инженерными *инструментами*, к ним и относиться нужно как к любым другим *рабочим* инструментам: следует проверять их работоспособность перед употреблением, не использовать их не по назначению, и в любом случае не доверять слепо навешанным на них формальным биркам-теоремам.

Наша публикация отличается от традиционных еще в одном отношении. Все научные исследования и отчеты об инженерных разработках с энтузиазмом повествуют об их значительных, важных, или, в крайнем случае, полезных результатах. Наша мрачная и заупокойная публикация, наоборот, рассказывает об ошибках, неудачах, катастрофах и трагедиях. И при этом утверждает, что это вовсе не ложка дегтя в бочке меда, а совсем наоборот, — бочка дегтя в ложке меда. Наука, если говорить о ее механико-математических отраслях, всегда отставала, а сегодня особенно, трагически отстала от уровня развития техники, и систематически обманывает технику своими мнимо достоверными рекомендациями и обоснованиями. В реальном сложном и нелинейном мире она решает «проблемы линеаризации», вырывает из неразрывной системы характеристик объекта одно свойство, например, «устойчивость», или «оптимальность», и вместо их исследования в полном объеме, предлагает решения частных задач. То есть — ищет то, что «лежит под фонарем», а не то, что нужно.

Техника за 250 лет после трудов Лагранжа опутала Землю сетью железных дорог и авиалиний, разъединила материки и соединила океаны каналами, достигла планет, а теоретическая механика все та же — Ньютонова и Лагранжева. Выдающиеся успехи самой передовой отрасли техники — космонавтики обеспечены теорией движения *одной материальной точки* Ньютона и теорией движения *одного*

твердого тела Эйлера — трехвековой давности! И это — на фоне гигантских успехов, достигнутых за это время другими науками.

Это мнение автора этой статьи.¹ Но если оппоненты скажут ему, что физика, химия и биология, достигшие за время летаргического сна механики грандиозных успехов, проще сложной механики, — он не поверит. Его опыт в решении математических проблем биологии свидетельствует об обратном.

Приведенные пессимистичные оценки автор в равной степени относит и к себе. Тема настоящей статьи является фрагментом общей проблемы «Моделирование и управление катастрофами с целью минимизации наносимого ими ущерба» [5]–[10]. Катастрофы свирепствуют с нарастающей силой, а мы все еще разрабатываем проект будущей «общей теории катастроф», конструируя детали инструментов для исследования сложных и рваных динамических процессов [11]–[18]. Некоторые попытки приложений этих инструментов в аварийной и катастрофической динамике летательных и космических аппаратов [19]–[23], и в задачах управления катастрофами человеческого организма [24]–[28] оказались относительно успешными. Однако в анализе катастрофического развала и возможностей стабилизации экономики [29]–[31] научные рекомендации остаются благими пожеланиями. А в проблеме автоматизации управления катастрофами, где требуются мгновенные и жесткие решения, которые неспособен принять эмоциональный человек в момент внезапной опасности, сделаны лишь первые шаги [32]–[38].

¹Содержание этой работы в большой степени является отражением пессимизма автора. Коллеги, просмотревшие рукопись статьи, указали ему на нежелательность публичной демонстрации его личного пессимизма на общем фоне научных публикаций с бодрыми отчетами об успешных исследованиях. Автор с ними согласился, и переработал текст, оставив в нем только сухие математические факты. Но тут опять случилась серия крупных катастроф, в том числе с космическими ракетами, и потерпели крушение еще несколько кораблей в реках, морях и океанах. И автор вернулся к своему первоначальному тексту. После участия в экспертизах чудовищных катастроф избавиться от пессимизма невозможно, и он посчитал необходимым поделиться им со своими оптимистичными коллегами — и студентами. Что же касается априорных оценок любой работы, то существуют убедительные примеры спорности *априорных оценок реальности* [51]. Они тем более спорны при обсуждении *катастрофической реальности*.

Представленный в настоящей статье фрагмент этой объемной темы — особенный. Все указанные выше работы посвящены решению задач моделирования и управления *уже возникшими* катастрофами; здесь рассматриваются проблемы *предотвращения* катастроф. Все методы общей темы используют сложные математические инструменты, адекватные сложности рассматриваемых проблем; здесь, в основном, обсуждаются простые инструменты «студенческого» уровня. Но эти простые вопросы не являются тривиальными и, тем более, очевидными — ранее в теоретических и прикладных математических и механических исследованиях они не отмечались, и не обсуждались. Вместе с тем они, возможно, более важны, чем предшествующие сложные и трудные исследования по этой теме. Скрытые от внешнего наблюдения, они оказываются «миной замедленного действия», ждущей в респектабельной упаковке достоверных, обоснованных и проверенных теорий и надежно рассчитанных технических конструкций лишь удобного случая, чтобы уничтожить эти теории, и разметать эти конструкции.

В свете сказанного, мы далее будем параллельно обсуждать технические катастрофы, и их причины — проектно-расчетные и теоретические катастрофы.

2. Технические и проектно-расчетные катастрофы

Проблема технических катастроф неразрывно связана с историей техногенной цивилизации. В конце прошлого века в регистре страховой лондонской компании Ллойда было около 20 000 кораблей, за год гибло около 2000. Сегодняшняя статистика потерь кораблей, самолетов и вертолетов нам неизвестна, но ежедневные сообщения о них утрашающие. Все разумные пределы превосходят автомобильные аварии.

Приведем несколько примеров. Часть их общеизвестна, часть взята из личных наблюдений автора.

1. В заключениях о трагедии технического шедевра англичан прошлого века — океанского суперлайнера «Титаник» в 1912 г. аварийные комиссии были единодушны — причинами были природные обстоятельства и ошибки судоводителей. В последующем были пред-

приняты большие усилия по научному анализу этих ошибок, все более глубоких по мере развития средств математического анализа и моделирования. Некоторые из этих вопросов рассматривались нами в [5], [10].

Перевернулся, и буквально раздавил экипаж — не в суровом северном океане с плавающими айсбергами, а на расстоянии нескольких метров от берега в мелкой бухте Севастополя в 1955 г. линкор «Новороссийск», построенный итальянцами. Причиной этой катастрофы была вражеская мина.

Вблизи берега Новороссийской бухты, в 1961 г. встретились и не разминулись два корабля — пассажирский лайнер «Адмирал Нахимов» (немецкой постройки) и некий сухогруз. Лайнер ушел на дно с сотнями пассажиров. Капитаны судов, не согласовавших свои курсы, остались невредимыми и — получили по 15 лет тюрьмы. Настоящая причина катастрофы — задержка отправления пассажирского корабля в ожидании высокопоставленного чиновника не афишировалась.

Недавно, в конце 2011 г. в реке утонул корабль «Булгария», и тоже унес с собой человеческие жизни. Причина — неисправность двигателя, дисбаланс при загрузке топлива, волнение речки от подувшего ветра. Прокуроры ищут виноватых.

Автор не успел закончить эту статью, а услужливая статистика предоставила ему еще один пример — ровно через сто лет после «Титаника», в январе 2012 г. еще более огромный и представительный лайнер с тысячей кают — «Коста Конкордия» сел на риф в Средиземном море — и лег на борт, раздавив часть пассажиров.

Это — *очевидные причины* произошедших катастроф. Но! Если бы корпуса всех этих судов были спроектированы правильно, все эти очевидные причины не привели бы к гибели кораблей! Очевидно, что для огромных кораблей «Титаника» и «Новороссийска» полученные ими пробоины были бы всего лишь неприятным случаем, если бы их корпуса были спроектированы с минимальной грамотностью. Кто же должен сидеть в тюрьме в ответе за тысячи погибших людей? Капитаны, или — проектировщики? Которые уже в проектах заложили те самые «мины замедленного действия», которые должны были «взорваться» и заполнить водой трюмы этих заранее обреченных кораблей. Те мины, которые ждали только *внешних* обстоятельств, за-

пустивших в действие их взрыватели. Между тем, проектировщиков никто не думал судить и сажать в тюрьму.

Произошли потрясшие страну трагедии с субмаринами «Комсомолец» и «Курск». Доступа к техническим причинам этих катастроф у научных работников нет, но возможности расчетных ошибок видны уже из фактов, оказавшихся известными в связи с работами иностранных спасателей. В многочисленных отечественных проектах подъема затонувшей ПЛ «Курск» основными были предложения подвесить ее на десятках и сотнях тросов. При этом все они игнорировали тот *элементарный* факт, что напряженными в любом положении объекта могли находиться всего 6 тросов, в то время как остальные десятки и сотни должны были бы болтаться свободными. Очевидно, что реализация таких проектов привела бы к еще одной катастрофе. Квалифицированные иностранные спасатели решили проблему, подсоединив десятки тросов своей системы к воздушным домкратам-демпферам, то есть, по существу, подвесив поднимаемую лодку на регулируемых пружинах, и успешно справились с задачей.

Но *почему утонули* подлодки? Кто виноват — их экипажи, военные противники, или — проектировщики, заложившие «мины замедленного действия» в неверных расчетах их проектов? Или те скрытые враги в профессорских мантиях, которые снабдили этих добросовестных проектировщиков своими неправильными теоретическими методами? Известий о судебных процессах над профессорами-теоретиками не поступало.

Если подвести итог этим морским и речным катастрофам, то он двойной: это катастрофы техники и — катастрофа *проекта железного корабля*, который должен плавать по воде. Катастрофы техники трагичны, но обыденны, катастрофа проекта исторична — она перечеркивает многотысячелетнюю концепцию «плавающего корыта». Каждый корабль, построенный в соответствии с этой концепцией, в независимости от того, как называется это корыто — сухогрузом, линкором, Титаником, или Конкордией, что именно навешано на это плавающее корыто — грузы, пушки, или рестораны, — обязан утонуть! Ждет только подходящего случая.

2. Замечательный проект дирижабля — «воздушной подводной лодки» завершился катастрофами — и умер. Но катастрофический

проект летательного аппарата тяжелее воздуха продолжает жить в самолетах и вертолетах. А каждый такой аппарат — обязан упасть! Ждет только подходящего случая, и этот случай выпадает. Экономика и политика настойчиво заставляют инженеров повторять этот опровергнутый физикой и статистикой проект, и пренебрегают уже проверенными и реализованными, принципиально более безопасными проектами кораблей на подводных крыльях и экранолетов.

3. В последние годы обрушились, и продолжают падать многочисленные кровли бассейнов, стадионов и рынков. Комиссии во всех случаях находят *неисправные* поломанные детали, и привлекают к ответственности их изготовителей. Но кровля стоит на опорах, число которых модные архитекторы безосновательно уменьшают, и при неверном расчете, при отсутствии нужных методик для таких сложных сооружений, эти опоры концентрируют напряжения, которые не в состоянии выдержать *исправные* детали.

4. Уравнения движения современных сложных технических объектов — сложные. Современная вычислительная техника позволяет их *решить*, но никакая вычислительная техника не позволяет их *исследовать*. Для этого требуется знать *всё о всех* траекториях объекта, а компьютер может вычислить только *некоторые* из них. Практика подсказала выход: линеаризовать нелинейные уравнения! Но их все равно оказывается слишком много, и умельцы в КБ научились их «грамотно» сокращать, выбрасывая те, которые *кажутся* им «несущественными». И в авиационном КБ — забыли об исходной задаче динамики *реальных* летательных аппаратов: вместо этого создана «научная школа» дальнейшего упрощения тех уже линеаризованных уравнений, которые никакого отношения к реальным самолетам и вертолетам не имеют. И летают *нелинейные* самолеты и вертолеты под управлением беспринципно упрощенных *линейных* систем ручного управления и автопилотов. А аварийные комиссии ищут ошибки управления пилота, которому в руки дали штурвал не реального самолета, а его *выдуманной* «фанерной» имитации.

Виноваты ли в этом расчетчики, у которых нет необходимых математических инструментов для исследования нелинейных систем? Конечно, нет, — виновата наука, отставшая от техники. А техника делает вид, что доверяет науке и пользуется ее результатами, но, по

существо, — обходится без нее. После изготовления самолета все начинается сначала, и новенький самолет разламывают, чтобы узнать его настоящую, а не расчетную прочность, а летчики-испытатели, не доверяя никаким расчетам, осторожно и понемногу приручают самолет вручную, и месяцами и годами исправляют его недостатки — часто ценой собственной жизни.

5. Существует международный научно-технический проект тросовых космических систем, состоящий в создании систем космических аппаратов, связанных тросами, и двигающихся на значительном расстоянии друг от друга. Проект сулит значительные выгоды: победу над невесомостью за счет вращения связки, возможность выработки электрической энергии за счет движения в космических магнитных полях, и другие. Но — даже *собрать* тросовую систему в неинерциальном поле Земли существующими методами невозможно: наши расчеты показали, что выпускаемые тросы вовсе не хотят направляться к предназначенным им целям, а — наматываются «бабушкиными клубками» на выпускающий их аппарат [21]! А гипотетическая тросовая система, отбирающая энергию из магнитного поля Земли, тормозится пропорционально отбираемой энергии, и — *должна упасть*, как детская игрушка, у которой кончилась батарейка.

6. На научной конференции специалист рассказывает о решении им сложной и ответственной задачи, подчеркивая, что это сделано с помощью современного (!) импортного (!) математического обеспечения, позволившего проникнуть в самую суть задачи путем приведения многомерных матриц кинетической и потенциальной энергии, и матрицы демпфирования к диагональному виду. Доклад получает одобрение. Но . . .

Перечисление этих «но» требует отдельной статьи. Упоминание о матрицах означает *линеаризацию*, а компьютерная программа — использование *численных методов*. Итого в докладе специалиста присутствует *четыре* объекта: 1 — реальный объект, 2 — его нелинейная математическая модель, 3 — его линеаризованная математическая модель, 4 — результаты расчетов с использованием компьютерной программы. Докладчик убеждает аудиторию, что он изучил объект 1, но — показывает ей объект 4. А дистанция от первого до четвертого объекта огромна!

Некоторые факты здесь очевидны, некоторые вскрываются только в жарких дискуссиях на защитах математических диссертаций. Например, все *численные* методы разрушают *физические* характеристики механических объектов. Численный метод либо отбирает механическую энергию из модели системы, либо, наоборот, добавляет ее. Значит, можно не долететь до цели, можно — перелететь. Но *попасть* в цель с помощью таких расчетов — нельзя!

А стандартная компьютерная программа — кем она изготовлена? Тонким знатоком новейших достижений алгебры, еще не вошедших в классические руководства (которые позволяют привести к диагональной форме одновременно только *две* специальных, но не *три* матрицы), или ее разработал ловкий программист? Если первое — то должны быть указаны исключительные свойства обрабатываемых программой матриц, а этого в описании пакета программ нет. А наш докладчик — толковый инженер, но не профессиональный математик, — как он проверит свои матрицы на соответствие этим уникальным критериям? И какие цели преследует иностранная компьютерная корпорация, поставляющая на рынок свои великолепные программы — помочь разработчикам наших ракет, или приумножить свои сверхприбыли? Эти программы — достижения математики, или математический блеф? Или — математический терроризм?

И объект, в который заложена эта «математическая мина», красиво упакованная в обертку компьютерных расчетов и снабженная этикеткой «сделано по импортным технологиям», отправится на работу, ожидая удобного случая, чтобы взорваться, сгореть и — унести с собой «достижения» современных теорий и вычислительной техники. А аварийные комиссии будут искать дефектные детали, подозревать пилота, диспетчера, и — спишут все на «погодные условия».

7. Не переложил ли здесь автор черной краски? Статистика катастроф говорит, что виноват в них не автор. Однако следует отметить и объективные обстоятельства. Обсуждение примеров в научной публикации принято предварять, или завершать их теоретическим обоснованием. Не будем отступать от этого правила, тем более что теория техногенных катастроф должна, наконец, ставиться на формализованные основания. Приведем основное положение этой будущей теории.

Теорема неизбежности катастроф. *При увеличении срока эксплуатации любой исправный технический объект должен потерпеть катастрофу.*

Приведем эскиз ее обоснования. В первой части используем математику. Множество всех *возможных* начальных условий и режимов работы любого объекта составляет *континуум*. Вместе с тем множество его *расчетных* начальных условий и режимов работы, и множество регламентированных операций по управлению этим объектом — *конечно*. Следовательно, с течением времени объект, *всегда* функционирующий в среде непредсказуемых возмущений, *всегда* выходит в режим, не предусмотренный его расчетами, и не подвластный управлениям оператора.

Вторая часть обоснования техническая. С увеличением сроков эксплуатации объекта среди континуума указанных выше его нерегламентированных и неуправляемых режимов всегда найдется такой, при котором он потерпит аварию, или катастрофу.

В формализованных частных задачах теоремы во всех публикациях автора доказываются строго, и детальнее. Однако в *общей проблеме катастроф* с неформализованным определением самого понятия *катастрофа*, детальнее обосновать приведенную теорему катастроф, по-видимому, невозможно. Важно то, что она — справедлива, что, к сожалению, подтверждает трагическая практика.

Для объективности отметим, что теорема катастроф, оставаясь справедливой, — может быть не реализована. Наступление предсказанного этой мрачной теоремой трагического события может предотвратить еще более трагический, и уже не требующий обоснований факт — смерть. Объект заведомо не потерпит катастрофу, которая ожидает его в будущем, если к этому времени он уже завершит свое существование, закончив свою штатную работу и отправившись в утиль.

Эта теорема, сформулированная автором несколько лет тому назад, на днях получила независимое подтверждение. При освещении для прессы многократных катастроф российских информационных спутников и космических ракет, причины которых нельзя было списать ни на «неисправность техники», ни на «человеческий фактор», ни на «погодные условия», была, наконец, хотя и нехотя, но все же

публично названа их настоящая причина: — «объекты вошли в нерасчетный режим».

Теперь мы должны рассмотреть причину техногенных катастроф, лежащую в теоретических основаниях математики и механики. Эту причину мы назовем *теоретическим терроризмом*, и в нескольких случаях прямо укажем на исполнителей и заказчиков этого научного вредительства. Все они, конечно, оказываются террористами поневоле, в силу объективных, не зависящих от них обстоятельств — в силу конфликтов, противоречий *правильных* методов математики и механики между собой. В теории этих дисциплин эти конфликты — *теоретические катастрофы*, и мы придем к заключению, что теоретические катастрофы приводят к техническим катастрофам.

3. Краткая история вечных конфликтов между математикой и механикой

Но прежде чем перейти к конкретному обсуждению этих вопросов, отметим, что хотя мы будем заниматься делом не совсем обычным, оно в науке вовсе не оригинально. Конфликты в математике и механике, и между математикой и механикой возникли не сегодня, имеют тысячелетнюю историю, привлекали и вовлекали в свой оборот выдающихся ученых, и были — мощным двигателем науки.

Первый конфликт механики с математикой возник в первый день ее рождения. Знаменитый греческий историк Плутарх (45–127 н. э.) свидетельствует:

«Механика, предмет искания и прославления, есть изобретение Евдокса (408–355 до н. э.) и Архита (428–365 до н. э.). Они хотели некоторым образом иллюстрировать геометрию (дать геометрии внешнюю прикраску) и основать на чувственных и материальных предметах теоремы, которые трудно решить помощью рассуждений и научных доказательств. Но скоро Платон (428–348 до н. э.) в негодовании стал упрекать их, что они портят геометрию, лишают ее достоинства, обращают ее в беглого раба, заставляя ее от изучения бестелесных и умственных вещей переходить к чувственным предметам и прибегать, кроме рассуждений, к помощи тел, рабски изготовленных работою руки».

С тех пор конфликты между математикой и механикой, и конфликты в математике и механике, — традиционны, и происходят на всех уровнях, от философских оснований и общетеоретических построений, до интерпретации конкретных расчетов. Примерами могут служить конфликт между теорией дифференциальных уравнений и механикой в трактовке уравнений связей в механических системах [3], знаменитая многовековая скандальная дискуссия о «принципе наименьшего действия» [39]–[44], или обвинения известного в свое время геометра Пуансо [45] против гениального математика и механика Лагранжа в незнании последним основ геометрии! Но все они относятся к обсуждению ошибок, которые рано или поздно, иногда через века после смерти насмерть спорящих оппонентов, исправляются. Принцип наименьшего действия, покрывший в свое время позором его создателя Мопертюи [39], через 250 лет был реабилитирован в нашей работе [42]–[44] с помощью современной теории управления, о возникновении которой через будущие века спорщики того времени не могли и подозревать. Конфликт между теорией дифференциальных уравнений и механикой Лагранж своевременно разрешил, введя в механическую систему реакции связей [3], которых не существует в теории дифференциальных уравнений. Но обвинения Пуансо Лагранж не мог опровергнуть, потому что они появились после его смерти. И скандальная статья Пуансо [45] путешествует по времени и пространству вместе со всеми переизданиями гениального трактата Лагранжа [3] на всех языках — в качестве его приложения, попала и в Россию в русском переводе. Этот заочный спор имеет отношение к теме настоящей статьи — к криволинейным пространствам с косоугольными ненормированными координатными базисами, в которых *существует* аналитическая механика, но о которых не знал ее создатель Лагранж, и о которых не мог подозревать его критик Пуансо.

Таковыми противоречиями между *правильными* и *неверными* методами и результатами заполнена вся история математики и механики. Однако эту благородную тему исправления *ошибок* мы приводим только в качестве предостережения, и обсуждать ее в этой публикации не будем. Мы здесь рассматриваем тему гораздо более глубокую и скрытую от внешнего наблюдения — конфликты между *правильными* методами математики и механики.

4. Скрытые конфликты между корректными методами математики и механики

Механика — математическая наука, но в отличие от других наук, пользующихся математическими методами, она неразрывно срослась с математикой, фактически превратилась в раздел математики. Галилей первым сделал попытку ее аксиоматического построения [1]. Ньютон, по существу, превратил механику в теорию дифференциальных уравнений [2]. Лагранж, создавая аналитическую механику, поставил и практически реализовал цель сделать ее «новой отраслью анализа» [3]. В последующем все теории механики развивались как теории преобразований уравнений движения и поиска их первых и полных интегралов. В том числе, к поиску интегралов уравнений движения сводятся достижения теории Солнечной системы Ньютона [2], все общие и частные достижения аналитической механики Лагранжа [3], и теории движения твердого тела Эйлера [4]. Используя точные методы математики, все различные теории механики, безотносительно к различиям в их формализме, демонстрируют полное совпадение своих результатов во всех теоретических и прикладных исследованиях и задачах. Это считается непреложным фактом, не подлежащим обсуждению в той же мере, в какой не подлежит обсуждению неизбежность методов математики. В результате все теории механики и все методы математики во всех исследованиях используются совместно, как инструменты единой мастерской.

Цель настоящего исследования — проанализировать достоверность этого непреложного согласия. Мы, наоборот, ищем факты расхождения, конфликтов между методами математики и механики, и между различными методами механики. При этом в нашем обсуждении не затрагиваются ни прошлые дискуссии среди математиков и механиков, ни высшие разделы этих дисциплин, такие как неевклидова геометрия, или теория относительности. Мы рассматриваем только основные простые факты математики и механики, их традиционные и проверенные методы.

О проблемах, обсуждаемых в этой статье, нигде не говорится, автору, во всяком случае, с их обсуждением встречаться не приходилось. В стане точных наук, в отличие от приближенно-описательных

гуманитарных, царят формальная логика и железная дисциплина, и поэтому они являются для других непревзойденным эталоном праведности и справедливости, сформулированном в знаменитом афоризме — «В каждой науке столько истины, сколько в ней математики». Однако в действительности, на кухне первых глав математики и классической механики с научными рецептами не все ладно. Слово *кухня* здесь к месту: мы обсуждаем *простые* частные факты, всем известные в академической обертке теорем и доказательств, но одновременно приглашаем читателя заглянуть в подсобное помещение, где из них изготавливаются совместные *сложные* продукты сомнительного качества. Это — самая опасная кухня. Ингредиенты абсолютно безвредны, потому что проверены веками. Но смешанные вместе, они могут неожиданно превратиться в гремучую смесь. Для техники взрывного дела это обычная практика, и первая основа профессии. Но, оказывается, и теоретическая механика может служить взрывчатым веществом, способным разрушить созданную ею же сколь угодно прочную конструкцию, если при строительстве смешивать ее с математикой не в безопасных пропорциях. Чтобы непригодность такого блюда распробовать, нужны гуманитарные средства — *подозрение* и *недоверие*, из точных наук давно изгнанные. Автор старается долю этих необходимых на каждой кухне горьких специй в науку вернуть.

Руководства по механике учат нас, *как нужно делать*, мы же говорим о том, *как не нужно делать*. С точки зрения последнего, и первое переходит из одномерного измерения *это правильно* в многомерное пространство *это — правильно?*²

²Один из коллег указал автору, что его статья будет только запутывать студентов, и окажется вредной для их «несформировавшегося еще профессионального сознания». Автор, со своей стороны, убежден, что формирование сознания студентов самоуверенными профессионалами, которые наводнили мир своими опасными изделиями — бесперспективно. Если это сознание не изменить, стремительный рост катастроф победить не удастся. Образование студентам, разумеется, необходимо, но обязательно — критическое, к чему и призывает эта статья. Науку всегда двигали вперед не отличники-зубрилы, а «плохие ученики» с «несформировавшимся еще профессиональным сознанием», плохо заучившие наизусть *старые* истины, и поэтому вынужденные создавать *новое* знание.

Родителей студентов о вреде нашей статьи мы предупредили. А что касается самих студентов, то 50-летний опыт преподавания автора в МФТИ и МГУ, во Владивостокском и зарубежных университетах показал, что им интереснее все-

Метод изложения этих вопросов мы выбираем самый простой — обсуждаем их с помощью примеров, большей частью элементарных. Изложение фундаментальных фактов механики простыми средствами — это метод ее «неграмотных» классиков, учившихся ее грамоте, и создававших эту грамоту на простых примерах. Автор настоятельно советует своим молодым читателям прочесть в качестве образца педагогический и художественный шедевр Галилея [1]. Простые примеры были ценны при строительстве основ механики Галилеем и Ньютоном, и они снова оказываются ценными при анализе ее уже построенных конструкций. Примеры общедоступны и являются лучшим введением в сложные теории. Но главное, — элементарные примеры способны разрушить сколь угодно сложную теорию, не предъявляя ей никаких обвинений по существу.

Нашей целью не является предъявление обвинений. Мы обращаем внимание на то, что незамеченная несовместимость *корректных* «деталей» может привести к созданию *некорректных* теоретических и прикладных методов и конструкций. Здесь, по крайней мере, нужны первые основы правописания, хрестоматия контрпримеров, начало которой, возможно, положит наше простое исследование.

го то, что *вредно*, и *не разрешается*. Так стоит ли продолжать вдалбливать им нудные математику и механику, или — лучше рассказывать им захватывающие «страшилки по Эдгару По», к сожалению, выслеснувшиеся из его книжек в нашу действительность? Не вернуться ли от скучного, многократно переписанного под разными фамилиями одного и того же учебника механики — к механическим поэмам, гимнам, молитвам Галилея, Ньютона, Эйлера, Мопертюи? Чтобы воспитать у студентов (простите за гуманитарный термин) художественный вкус, позволяющий различить музыку Бетховена в научных и богословских трактатах Ньютона и Мопертюи, и Моцарта — в математических симфониях Лагранжа! Чем еще может привлечь наука в наше меркантильное время, как не красотой.

5. Примеры конфликтов между методами Математики и Механики

И так как перед вами, читатель, не трактат, а лишь несколько случайных соображений, которые должны послужить предисловием к моему не совсем обыкновенному рассказу, то я пользуюсь случаем заявить, что непритязательная игра в шашки требует куда более высокого умения размышлять и задает уму больше полезных задач, чем мнимая изощренность шахмат.

Эдгар По. Убийство на улице Морг

Любой механический объект является механической системой. Мы рассмотрим в качестве такой системы простейший фундаментальный объект механики, *основной предмет исследований Галилея, Ньютона и Лагранжа* [1]–[3] — *материальную точку*. Заставим ее играть серьезную роль «механической системы», будем наблюдать за ее поведением в этом качестве, задавать ей серьезные системные вопросы, и критически оценивать ее ответы, относя их к ответственности той теории, которую эта точка в контексте этой теории будет представлять. В качестве декораций возьмем вторую в истории механики (после рычага) теоретическую задачу Галилея о движении по наклонной плоскости [1]. Галилей с ее помощью открыл *все законы классической механики*, известные нашей цивилизации, но сама эта задача все еще не утратила своего познавательного значения. Задача эта *элементарная*, но посмотрим, является ли она *простой*.

Задача 1. *Составить уравнения движения материальной точки, движущейся под действием силы тяжести по неподвижной гладкой прямой спице l , отклоненной от вертикали на 45° .*

Решение. Исходные математические уравнения этой задачи, очевидно, можно записать так:

$$\text{А) } m\ddot{x}=0, \quad \text{Б) } m\ddot{y}=-mg, \quad \text{В) } l=l(x, y)=x+y-c=0, \quad c=\text{const.} \quad (1)$$

Поручим решить эту задачу математику, и механику. Это — демонстрационный пример, и мы намеренно попросим участвовать в его анализе математика-формалиста, не знакомого с теоретической механикой.

5.1. Беспомощность методов теории дифференциальных уравнений в механике

Приступив к решению Задачи 1, математик-теоретик уже при виде ее системы трех уравнений (1) вернет их Галилею, как некорректные: уравнение В) *должно быть* интегралом уравнений А) и Б), чего, очевидно, быть не может.

Это наблюдение *математика* является важным общим фактом *механики*:

Предупреждение 1. *Уравнение механической связи первым интегралом уравнений движения механической системы не является, и для понижения их порядка использоваться не может.*

5.2. Конфликт между методами теории дифференциальных уравнений и теоретической механики

Для механика противоречий в системе уравнений (1) нет: уравнение В) описывает *механическую связь*, которая воздействует на уравнения А) и Б) дополнительной силой своей реакции. Сила реакции в этой задаче одним уравнением связи В) определена быть не может. Следует использовать дополнительное условие ее *гладкости*. Это условие требует, чтобы сила реакции связи была к ней ортогональной.

Чтобы избежать вычисления силы реакции косо́й связи В), подкованный в математике механик предпочтет упростить уравнение этой связи, выровняв ее с помощью преобразования системы уравнений (1) к новым переменным

$$\text{а) } u = x + y, y = y; \text{ либо б) } u = x + y, x = x. \quad (2)$$

В этих новых переменных исходные данные Задачи (1) запишутся так:

$$\text{А) } m\ddot{u} = -mg, \text{ Б) } m\ddot{y} = -mg, \text{ В) } l = l(u) = u - c = 0; \quad (3-1)$$

$$\text{А) } m\ddot{u} = -mg, \text{ Б) } m\ddot{x} = 0, \text{ В) } l = l(u) = u - c = 0. \quad (3-2)$$

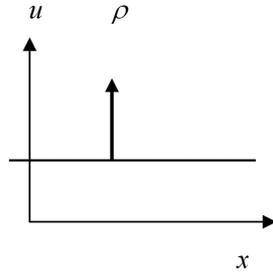


Рис. 1.

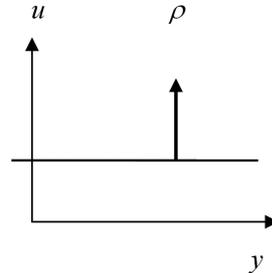


Рис. 2.

Никаких рекомендаций и инструкций о правилах преобразований уравнений движения, отличных от правил корректных преобразований дифференциальных уравнений, в теоретической механике нет, поэтому будем обращаться с уравнениями (3-1) и (3-2) как с уравнениями обычных задач механики.

Связь (1В) превратилась после преобразований (2) в связь (3-1В), (3-2В), и стала здесь горизонтальной, а сила реакции ρ этой *гладкой* связи, в соответствии с рис. 1 и 2, — вертикальной.

Эта сила реакции связи (3-1В) и (3-2В) войдет, следовательно, только в уравнения (3-1А) и (3-2А) для координаты u , и уравнения (3-1) и (3-2) с учетом этой силы запишутся, соответственно, в виде

$$\text{А) } m\ddot{u} = -mg + \rho, \text{ Б) } m\ddot{y} = -mg, \text{ В) } \ddot{u} = 0; \quad (4-1)$$

$$\text{А) } m\ddot{u} = -mg + \rho, \text{ Б) } m\ddot{x} = 0, \text{ В) } \ddot{u} = 0. \quad (4-2)$$

Единственными решениями алгебраических уравнений (4-1) и (4-2) относительно их переменных \ddot{u}, \ddot{y}, ρ и \ddot{u}, \ddot{x}, ρ являются

$$\text{а) } \ddot{u} = 0, m\ddot{y} = -mg, \rho = mg; \text{ б) } \ddot{u} = 0, m\ddot{x} = 0, \rho = mg. \quad (5)$$

Преобразование решений (5), обратное к (2), дает два решения Задачи 1:

$$m\ddot{x} = mg, \quad m\ddot{y} = -mg, \quad \rho = mg; \quad (6-1)$$

$$m\ddot{x} = 0, \quad m\ddot{y} = 0, \quad \rho = mg. \quad (6-2)$$

Два *разных* решения элементарной Задачи 1, полученные *правильными* вычислениями, вызывают вопросы. Читатель может (и должен)

искать арифметические и методические ошибки в приведенных вычислениях. В этом ему поможет и то очевидное обстоятельство, что оба эти решения — *вздор*. Мы же ограничимся следующим выводом:

Предупреждение 2. Формальное преобразование уравнений движения механических систем методами теории дифференциальных уравнений — недопустимо.

Негативное утверждение ускользает от прямой ответственности. Его нельзя схватить за руку, проверив его доказательство — оно же отсутствует! Это всего лишь разрушающий пример. Но, раз здесь *что-то не так*, подозрение падает не только на почтенные методы математики и механики, на которые жаждет бросить тень автор этой статьи, но и на самого автора. Это и является нашей целью — вызвать у читателя *подозрения*. Поможем ему: возможно, в приведенных примерах виноваты не математика и механика, а автор этой статьи, который неверно понял и применил метод Лагранжа замены механической связи ее силой реакции? Чтобы разрешить этот вопрос, следует обратиться к самому компетентному арбитру — к самому Лагранжу. Кстати, наткнемся еще на один конфликт в этом безупречном семействе *точных* наук.

5.3. Конфликт между методами теории дифференциальных уравнений и аналитической механики

Нужно вернуться к тому пункту предыдущих выкладок, который предшествует введению силы реакции связи, то есть, к уравнениям (3-1), (3-2).

Выпишем для этих двух задач Общую формулу динамики Лагранжа [3]. В ней вместо явной силы реакции ρ в уравнения движения (3-1) и (3-2) следует ввести виртуальный дифференциал δl связи l (3-1В), (3-2В) с помощью множителей Лагранжа, которые обозначим здесь λ_1 и λ_2 :

$$(m\ddot{u} + mg)\delta u + (m\ddot{y} + mg)\delta y + \lambda_1\delta u = 0; \quad (7-1)$$

$$(m\ddot{u} + mg)\delta u + m\ddot{x}\delta x + \lambda_2\delta u = 0. \quad (7-2)$$

Приравнивая в этих уравнениях множители при независимых дифференциалах переменных к нулю, и добавляя для определения «лиш-

них» переменных λ_1 и λ_2 последнее уравнение (3-1), (3-2), получаем две системы уравнений

$$m\ddot{u} + mg + \lambda_1 = 0, \quad m\ddot{y} + mg = 0, \quad \ddot{u} = 0; \quad (8-1)$$

$$m\ddot{u} + mg + \lambda_2 = 0, \quad m\ddot{x} = 0, \quad \ddot{u} = 0. \quad (8-2)$$

Уравнения Лагранжа (8-1) и (8-2) при $\lambda_1 = -\rho$, $\lambda_2 = -\rho$ совпадают с уравнениями механика (4-1) и (4-2), и он, следовательно, ни в чем не ошибся. Но эти уравнения Лагранжа приводят к тем же *разным* решениям (6-1) и (6-2). Следовательно, вину за их абсурдность теперь можно снять с автора этой статьи, и обоснованно отнести на счет методов математики и механики, теперь — аналитической механики Лагранжа.

Читатель по-прежнему может (и мы ему это советуем) не доверять автору. Мы же подытожим наше элементарное исследование следующим предупреждением:

Предупреждение 3. Формальное преобразование уравнений движения методами теории дифференциальных уравнений в аналитической механике — недопустимо.

5.4. Высший арбитраж Ньютона

Обратимся к высшему авторитету в Механике — Ньютону. Уравнения движения материальной точки Ньютона — единственное, что надежно знают студенты. Поэтому приведем решение Задачи 1, полученное студентом.

Студент нарисовал линию движения точки l и, в соответствии с рис. 3, разложил вектор вертикальной силы тяжести на ортогональную и касательную к этой линии составляющие, обе величины $mg/\sqrt{2}$.

Затем он уничтожил ортогональную компоненту по третьему Закону Ньютона реакцией опоры. Повторно разложив оставшуюся касательную компоненту силы на горизонтальную и вертикальную составляющие в соответствии с Рис. 4, в результате он получил

$$m\ddot{x} = mg/2, \quad m\ddot{y} = -mg/2. \quad (9)$$

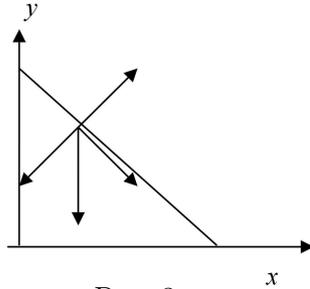


Рис. 3.

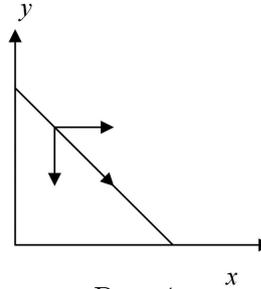


Рис. 4.

Студент получил *третий* результат, отличный от обоих результатов механика и Лагранжа (6). Вопросы стало больше, и ясно, что в точной науке Механике приходится использовать еще один, кроме «предупреждения», гуманитарный термин — *подозрение*. Можно подозревать, что ошибаются *аналитические* методы, а прав студент, использовавший *наглядные* методы Ньютона. Следовательно, нужно еще раз проверить эти аналитические методы.

5.5. Решение Задачи 1 методами аналитической механики Лагранжа

Попросим Лагранжа решить Задачу 1 без математических усовершенствований механика, который ввел преобразования (2), но, чтобы избежать разночтений, сделать это путем того же предварительного преобразования переменных (2).

Рассмотрим для краткости случай (2б). Лагранж *не пользуется* методами теории дифференциальных уравнений (как будто он уже прочел наше Предупреждение 3!), а производит преобразования переменных в уравнениях движения *собственными* методами, с помощью своей универсальной Общей формулы динамики [3]. Ее нужно записать для исходных уравнений (1), и затем произвести в ней преобразования (2б):

$$\begin{aligned}
 m\ddot{x}\delta x + (m\ddot{y} + mg)\delta y + \lambda\delta\xi(x, y) &= \\
 = m\ddot{x}\delta x + [m(\ddot{u} - \ddot{x}) + mg](\delta u - \delta x) + \lambda\delta l(u) &= \\
 = (2m\ddot{x} - m\ddot{u} - mg)\delta x + [m(\ddot{u} - \ddot{x}) + mg + \lambda]\delta u &= 0. \quad (10)
 \end{aligned}$$

Извлекая из (10), при произвольных независимых виртуальных перемещениях δx и δy , два уравнения, и добавляя к ним уравнение связи $\ddot{u} = 0$, получаем систему уравнений

$$2m\ddot{x} - m\ddot{u} - mg = 0, \quad m(\ddot{u} - \ddot{x}) + mg + \lambda = 0, \quad \ddot{u} = 0, \quad (11)$$

решение которой дает результат — совпадающий с решением студента (9):

$$m\ddot{x} = mg/2, \quad \ddot{u} = 0, \quad \lambda = -mg/2, \quad \text{откуда } m\ddot{x} = mg/2, \quad m\ddot{y} = -mg/2. \quad (12)$$

Легко проверить, что к этому же результату (9), (12) приводят уравнения Лагранжа I рода, и уравнения Лагранжа II рода, при использовании в качестве обобщенной координаты в этой одностепенной задаче как любой из координат x, y , так и u . К этим же уравнениям (9), (12) приводят наши матричные уравнения [46]–[48] в декартовых, и в обобщенных координатах.

Комментарии. Результаты студента (9) и Лагранжа (12) — верные, а результаты механика (6) — ошибочные, хотя никаких *математических* нарушений при их получении — нет! И правильность его вычислений даже была выше подтверждена Лагранжем! Это — парадокс, и причина этого парадокса заключается, очевидно, не в вычислительных ошибках. Она в *конфликте* методов математики и механики: преобразование (2) уравнений (1) — недопустимо! Этот факт, уже отмеченный выше в Предупреждениях 2 и 3, оказывается общим для *всех* методов механики, и отличать ее отдельные разделы уже не обязательно:

Предупреждение 4. *Формальное преобразование дифференциальных уравнений движения механики методами теории дифференциальных уравнений — недопустимо.*

Заявление ответственное, но автор ответственность за него с себя полностью снимает — оно следствие примера, который автор всего лишь публично продемонстрировал, и который невозможно опровергнуть в силу его тривиальности. Но наш маленький пример — это медицинский анализ «по капле крови», выявляющий скрытую болезнь, внутренний конфликт гигантов — математики и механики.

5.6. Конфликт между методами преобразования уравнений в теории дифференциальных уравнений и в аналитической механике Лагранжа

Преобразования (2) системы уравнений (1) в системы уравнений (3) было произведено по всем правилам теории дифференциальных уравнений: они невырожденные, взаимно однозначные, и сохраняют порядок дифференциальных уравнений.

Лагранж *простыми* приемами этой *математической* теории пренебрег, и преобразование (2) системы уравнений (1) в систему уравнений (11) производит *сложным, странным и загадочным образом*: он собирает *все* старые уравнения (1) в *единственное* уравнение (10), и — из этого *единственного* уравнения получает новую *систему* уравнений (11)! Глубоко развитая и подробно изученная теория дифференциальных уравнений такого фокуса не знает! В ней уравнения можно складывать, но, сложив два уравнения, одно из слагаемых обязательно нужно сохранить, и нельзя потерять ни одной производной! А Лагранж, волшебник математических превращений, показывает нам *открытые* карты одной масти, собирает их в *одну* колоду, на наших глазах ее тасует, и — раскладывает перед нами карты *другой* масти!

Это не *математический*, а *механический* способ преобразования дифференциальных уравнений, и это замечательно, потому что он должен расширить теорию дифференциальных уравнений. Но такого расширения не произошло, потому что эти методы конфликтуют между собой, и метод Лагранжа не дополняет, а *вытесняет* методы теории дифференциальных уравнений из аналитической механики! Потому что последние, как показывают Предупреждения 2–4, для преобразований дифференциальных уравнений механики — непригодны. Здесь применимы только методы преобразований Лагранжа!

Предупреждение 5. *Для преобразования уравнений движения механических систем в механике могут быть использованы только методы Лагранжа.*

К числу этих методов, помимо использующих Общую формулу динамики Лагранжа, относятся методы преобразования переменных

в функциях Лагранжа и Гамильтона. Последние присутствуют в активном арсенале аналитической механики [3], [4].

Взаимоотношения Предупреждения 5 с методами преобразования уравнений аналитической механики в теории Гамильтона для обыкновенных дифференциальных уравнений и в теории Гамильтона-Якоби для уравнений в частных производных мы предоставляем обдумать читателю.

6. Примеры конфликтов между методами аналитической и теоретической механики

Продолжим обсуждать простые примеры, теперь — примеры конфликтов разных методов механики между собой.

Рассмотрим замечательную методическую задачу Галилея с еще одной стороны.

Задача 2. *Составить уравнения движения в задаче 1 в полярных координатах.*

Решение. В этой задаче два этапа: первый состоит в преобразовании ее исходных уравнений, второй — в исключении уравнения связи. Выполним первый этап методом теоретической механики, а второй — методом аналитической механики.

6.1. Преобразования уравнений движения методами теоретической механики

Перейдем в исходных уравнениях задачи (1) к полярным координатам r и φ , проектируя ускорение точки на радиальное и трансверсальное направления. В результате первый этап приводит к уравнениям

$$\begin{aligned} m(\ddot{r} - r\dot{\varphi}^2) &= -mg \sin \varphi, \\ m(r\ddot{\varphi} + 2\dot{r}\dot{\varphi}) &= -mg \cos \varphi, \\ l = r(\cos \varphi + \sin \varphi) + c &= 0. \end{aligned} \tag{13}$$

6.2. Решение преобразованных уравнений методами аналитической механики

На втором этапе записываем для уравнений (13) Общую формулу динамики Лагранжа

$$m(\ddot{r} - r\dot{\varphi}^2 + g \sin \varphi)\delta r + m(r\ddot{\varphi} + 2\dot{r}\dot{\varphi} + g \cos \varphi)\delta\varphi + \lambda\delta[r(\cos \varphi + \sin \varphi) + c] = 0. \quad (14)$$

Расщепление уравнения (14) при независимых δr и $\delta\varphi$ дает два уравнения

$$\begin{aligned} m(\ddot{r} - r\dot{\varphi}^2 + g \sin \varphi) + \lambda(\cos \varphi + \sin \varphi) &= 0, \\ m(r\ddot{\varphi} + 2\dot{r}\dot{\varphi} + g \cos \varphi) + \lambda r(\cos \varphi - \sin \varphi) &= 0. \end{aligned} \quad (15)$$

Решать уравнения (15) не потребуется. Они — *некорректны*: множитель Лагранжа λ должен принимать в них *одинаковое* значение *разной* размерности!

Комментарий. Решение указанных двух этапов одним методом, как теоретической, так и аналитической механики, приведет, как легко проверить, к правильному результату. В абсурдности формул (15) виновато *смешение* разных методов.

Предупреждение 6. *Методы теоретической механики для использования в алгоритмах аналитической механики — непригодны.*

Здесь действует также Предупреждение 5.

Против нашего метода вывода *общих* заключений из элементарных *частных* примеров можно возражать. Но, например, возражение против Предупреждения 6 потребует анализа на совместимость *всех* методов теоретической механики со *всеми* алгоритмами аналитической механики. А они, в силу нашего примера, *в целом* заведомо несовместимы. И выбор *частных* случаев совместимости должен быть основан на общем *теоретическом* анализе: примеры могут опровергать, но не могут сказать ничего позитивного. В нашу задачу такой обширный анализ не входит. Предупреждение 6, всего лишь, — предупреждает. И если указанный анализ будет когда-либо проведен, он это Предупреждение только уточнит, но не опровергнет: пример Задачи 2 этого не позволит.

7. Конфликт между аналитической механикой Лагранжа и теорией твердого тела Эйлера

Рассмотрим еще один пример, тоже методический, но уже не учебный. Он описывает важную техническую задачу, используемые в нем теории присутствуют во многих теоретических исследованиях и вычислительных алгоритмах, а приложения этих теорий уже привели, и еще приведут к многочисленным катастрофам.

Задача 3. Составить уравнения движения груза, переносимого на тросовой подвеске вертолетом.

Решение. Принципы решения этой задачи очевидны. Нужно выписать уравнения движения свободного вертолета и свободного груза, затем записать уравнение троса, и это уравнение троса исключить из уравнений движения вертолета и груза, с тем, чтобы получить уравнения движения *связанной* системы вертолет-груз *в целом*.

Эту методическую задачу мы предельно упростим, считая вертолет цельным твердым телом, а груз — материальной точкой. Наиболее широко в теоретических и технических расчетах для описания динамики твердых тел используются уравнения Эйлера.

Выпишем их для вертолета (вращение — в системе главных осей инерции тела ξ, η, ζ , координаты поворота — углы Эйлера ψ, θ, φ ; координаты центра масс x_B, y_B, z_B — в системе неподвижных осей x, y, z):

$$\begin{aligned} \Psi &= A\dot{p} + (C - B)qr - M_\xi = 0, \\ p &= p(\psi, \theta, \varphi, \dot{\psi}, \dot{\theta}, \dot{\varphi}), \quad \dot{p} = \dot{p}(\psi, \theta, \varphi, \dot{\psi}, \dot{\theta}, \dot{\varphi}, \ddot{\psi}, \ddot{\theta}, \ddot{\varphi}), \\ \Theta &= B\dot{q} + (A - C)rp - M_\eta = 0, \\ q &= q(\psi, \theta, \varphi, \dot{\psi}, \dot{\theta}, \dot{\varphi}), \quad \dot{q} = \dot{q}(\psi, \theta, \varphi, \dot{\psi}, \dot{\theta}, \dot{\varphi}, \ddot{\psi}, \ddot{\theta}, \ddot{\varphi}), \\ \Phi &= C\dot{r} + (B - A)pq - M_\zeta = 0, \\ r &= r(\psi, \theta, \varphi, \dot{\psi}, \dot{\theta}, \dot{\varphi}), \quad \dot{r} = \dot{r}(\psi, \theta, \varphi, \dot{\psi}, \dot{\theta}, \dot{\varphi}, \ddot{\psi}, \ddot{\theta}, \ddot{\varphi}). \end{aligned} \quad (16-1)$$

$$X_B = m\ddot{x}_B - F_x = 0, \quad Y_B = m\ddot{y}_B - F_y = 0, \quad Z_B = m\ddot{z}_B - F_z = 0, \quad (16-2)$$

Затем выписываем уравнения движения груза

$$X_\Gamma = m\ddot{x}_\Gamma - f_x = 0, \quad Y_\Gamma = m\ddot{y}_\Gamma - f_y = 0, \quad Z_\Gamma = m\ddot{z}_\Gamma - f_z = 0. \quad (17)$$

Трос l одним концом прикреплен к точке a движущегося вертолета, другим — к грузу b . После элементарных, но громоздких вычислений уравнение троса-связи представляется в функции координат вертолета, груза и времени формулой

$$l = |r_a - r_b| = l(x_B, y_B, z_B, \psi, \theta, \varphi, x_G, y_G, z_G, t) = 0. \quad (18)$$

Попросим решить Задачу 3 нашего студента, который решал Задачу 1, но теперь уже окончил ВУЗ и стал инженером. Наш пример методический, и нашего инженера мы попросим сыграть две методические роли — расчетчика из КБ и младшего научного сотрудника из НИИ.

7.1. Инженерный расчет, основанный на методах теоретической механики

В КБ традиционно используются теоретическая механика и общие методы математики. Аналитическая механика считается развлечением теоретиков из университетов.

Использование уравнения троса (18) позволяет исключить из уравнений вертолета (16-1), (16-2) и груза (17) одну из их координат. Например, координату вертолета φ . Ее можно исключить из (16-1) с помощью самого уравнения (18), а ее производные $\dot{\varphi}$ и $\ddot{\varphi}$, соответственно, — с помощью производных этого уравнения

$$\begin{aligned} \frac{dl}{dt} = \frac{\partial l}{\partial x_B} \dot{x}_B + \frac{\partial l}{\partial y_B} \dot{y}_B + \frac{\partial l}{\partial z_B} \dot{z}_B + \frac{\partial l}{\partial \psi} \dot{\psi} + \frac{\partial l}{\partial \theta} \dot{\theta} + \frac{\partial l}{\partial \varphi} \dot{\varphi} + \\ + \frac{\partial l}{\partial x_G} \dot{x}_G + \frac{\partial l}{\partial y_G} \dot{y}_G + \frac{\partial l}{\partial z_G} \dot{z}_G + \frac{\partial l}{\partial t} = 0, \quad (19) \end{aligned}$$

$$\begin{aligned} \frac{d^2l}{dt^2} = \frac{\partial l}{\partial x_B} \ddot{x}_B + \frac{\partial l}{\partial y_B} \ddot{y}_B + \frac{\partial l}{\partial z_B} \ddot{z}_B + \frac{\partial l}{\partial \psi} \ddot{\psi} + \frac{\partial l}{\partial \theta} \ddot{\theta} + \frac{\partial l}{\partial \varphi} \ddot{\varphi} + \\ + \frac{\partial l}{\partial x_G} \ddot{x}_G + \frac{\partial l}{\partial y_G} \ddot{y}_G + \frac{\partial l}{\partial z_G} \ddot{z}_G + \dots = 0. \quad (20) \end{aligned}$$

Отбрасывание уже ненужного последнего уравнения (16-1) приводит к системе оставшихся первых двух уравнений (16-1), не содержащих

переменной φ и ее производных, трех уравнений центра масс вертолета (16-2) и трех уравнений груза (17).

Работа завершена, и подробные вычисления остается свести в технический отчет, который наш инженер-расчетчик может использовать в материалах своей кандидатской диссертации по техническим наукам.

Комментарии. Инженеру в его расчетной работе можно помочь. Редукцию системы уравнений (16-1) можно провести отбрасыванием не только ее последнего уравнения, но и любого другого. Можно также исключить одну из координат системы уравнений (16-2), или (17). В итоге можно получить не одну, а 9 редуцированных систем и, соответственно, 9 решений Задачи 3!

Мы предоставляем эти расширенные возможности инженеру-расчетчику, а он, в свою очередь — конструктору вертолета. Что посоветовать конструктору — какую ему из этих 9 систем уравнений движения механической системы вертолет-груз выбрать, чтобы заказать не ее основе систему управления вертолетом?

Попросим ответить читателя, а сами пока посетим НИИ.

7.2. Научный расчет, основанный на методах аналитической механики

В современном НИИ для решения Задачи 3 естественно воспользоваться методами аналитической механики. Общая формула динамики Лагранжа для уравнений (16-1), (16-2), (17) с механической связью (18) записывается в виде

$$\begin{aligned} & \Psi\delta\psi + \Theta\delta\theta + \Phi\delta\varphi + X_B\delta x_B + Y_B\delta y_B + Z_B\delta z_B + X_G\delta x_G + Y_G\delta y_G + Z_G\delta z_G + \\ & + \lambda \left(\frac{\partial l}{\partial \psi}\delta\psi + \frac{\partial l}{\partial \theta}\delta\theta + \frac{\partial l}{\partial \varphi}\delta\varphi + \frac{\partial l}{\partial x_B}\delta x_B + \frac{\partial l}{\partial y_B}\delta y_B + \frac{\partial l}{\partial z_B}\delta z_B + \right. \\ & \left. + \frac{\partial l}{\partial x_G}\delta x_G + \frac{\partial l}{\partial y_G}\delta y_G + \frac{\partial l}{\partial z_G}\delta z_G \right) = 0. \quad (21) \end{aligned}$$

Расчленение уравнения (21) на множители при независимых виртуальных дифференциалах приводит к 9 уравнениям

$$\begin{aligned}
 X_{\text{в}} + \lambda \frac{\partial l}{\partial x_{\text{в}}} = 0, \quad Y_{\text{в}} + \lambda \frac{\partial l}{\partial y_{\text{в}}} = 0, \quad Z_{\text{в}} + \lambda \frac{\partial l}{\partial z_{\text{в}}} = 0, \\
 \Psi + \lambda \frac{\partial l}{\partial \psi} = 0, \quad \Theta + \lambda \frac{\partial l}{\partial \theta} = 0, \quad \dots,
 \end{aligned}
 \tag{22}$$

к которым следует добавить еще уравнение связи (20).

Выражения (16-1), (16-2), (17), которые должны быть подставлены в (22), и уравнение (20) линейны по вторым производным всех 9 координат системы вертолет-груз, и решение этой линейной системы 10 уравнений (22), (20) относительно 10 переменных $\ddot{x}_{\text{в}}, \ddot{y}_{\text{в}}, \ddot{z}_{\text{в}}, \dot{\psi}, \dot{\theta}, \dot{\varphi}, \ddot{x}_{\text{г}}, \ddot{y}_{\text{г}}, \ddot{z}_{\text{г}}$ и λ дает искомое решение Задачи 3.

Работа завершена, и подробные вычисления наш инженер — м.н.с. из НИИ может изложить в научном отчете, и использовать их в кандидатской диссертации, теперь по физико-математическим наукам.

Комментарии. Можно обсуждать достоинства и недостатки представленных «инженерного» и «научного» методов: первый приводит к дифференциальным уравнениям движения 16 порядка, второй — к уравнениям 18 порядка. И *разных* решений теперь еще больше — к 9 решениям КБ можно добавить решение НИИ. Теперь у конструктора вертолета выбор необходимых уравнений движения для заказа системы управления еще больше. Только это — выбор вида катастрофы, который неизбежно потерпит вертолет в первом же полете. Потому что *все* эти уравнения — ошибочные. Точнее, они не имеют никакого отношения к истинным уравнениям движения системы вертолет-груз.

Пока читатель объяснит конструктору вертолета, какие именно ошибки допустили его помощники — расчетчик из КБ и м.н.с. из НИИ, автор пояснит, что приведенные сценарии решения Задачи 3 взяты им из личного опыта общения с работниками КБ и сотрудниками НИИ. И приведет дословно свой диалог с еще одним действующим лицом этой научно-технической драмы, слишком часто завершающейся трагедией, — с вертолетчиком, переносившим трансформаторы и опоры электропередач над бездорожьем тайги.

7.3. Диалог автора с пилотом вертолета об эффективных методах механики

Автор: — Как ведет себя груз в полете?

Пилот: — Болтается.

Автор: — Какие маневры вы предпринимали, чтобы уменьшить колебания груза при сильной болтанке?

Пилот: — Рубил трос!

7.4. Разбор несостоявшихся полетов

Ошибки «инженерного» решения.

На вздорность «инженерного» решения 6.1 указывает Предупреждение 1: уравнение связи (18) *не является* первым интегралом системы уравнений (16-1), (16-2), (17) и *не может* быть использовано в этом качестве для понижения ее порядка. Это независимо подтверждает алгебра. Общее число уравнений (16-1), (16-2), (17) и (20) равно 10, а неизвестных $\dot{\psi}, \dot{\theta}, \dot{\varphi}, \ddot{x}_B, \ddot{y}_B, \ddot{z}_B, \ddot{x}_G, \ddot{y}_G, \ddot{z}_G$ в этих уравнениях — 9. Очевидно, что эти уравнения совместными быть не могут, и поэтому любые их преобразования бессмысленны. Таким образом, не заглядывая в конкретные расчеты нашего инженера из КБ, мы убеждаемся, что его отчет — блеф.

К каким последствиям приведет использование его вздорного решения на практике — очевидно. Кем является в этом случае наш расчетчик, очевидно тоже — террористом! Но он в этом не виноват. В ВУЗе он был отличником, усердно изучал курс механики и, одновременно, курс дифференциальных уравнений, на котором ему рассказали, что *любое* конечное соотношение между переменными системы дифференциальных уравнений можно (и нужно!) немедленно использовать для понижения ее порядка. В результате в нашей задаче, обладая хорошей *математической* подготовкой, он сделал все *правильно* с точки зрения теории дифференциальных уравнений. Но пятерка по этому *математическому* предмету оказывается — двойкой *по механике*! И виноватым оказывается не этот невольный террорист, а — конфликт между математикой и механикой!

Ошибки «научного» решения.

Обнаружить ошибки в «научном» решении 6.2 труднее. Здесь наш м.н.с. тоже оказался в банде техногенных террористов, но настоя-

щим заказчиком этого преступления оказывается вредитель, надежно скрывшийся — под профессорской мантией! Вот что говорят о нем в научных изданиях: «Выдающийся ученый, создавший капитальный труд, по обилию материала, полноте и строгости изложения представляющий собой по существу энциклопедию знаний в области классической механики . . . ». Этот, действительно выдающийся труд замечательного французского ученого П. Аппеля [49] содержит *безобидное* утверждение, что *уравнения движения твердого тела Эйлера являются уравнениями Лагранжа* ([49], п. 443, стр. 283)! Это утверждение, полученное на основании красивой, но неполной и потому *ошибочной* аналогии, широко растиражировано в многочисленных современных отечественных учебниках. Но безобидное в энциклопедии по механике начала прошлого века, с появлением современных сложных технических систем связанных тел, требующих использования методов множителей Лагранжа, оно становится орудием математического терроризма. Мы специально подчеркнем это в следующем предупреждении:

Предупреждение 7. *Уравнения движения твердого тела Эйлера уравнениями Лагранжа не являются.*

Но почему следует называть терроризмом всего лишь разницу в *названиях* разных видов уравнений движения? Сам Лагранж утверждает [3], что его метод множителей является *универсальным*, и пришел в механику из анализа, где он используется повсеместно. Представляется *очевидным*, что его, следовательно, можно использовать в любых уравнениях движения теоретической и аналитической механики, и что наш математически подкованный м.н.с. вполне грамотно применил его для уравнений движения вертолета в форме Эйлера (16-1).

Попробуем разобраться по существу. Однако здесь следует обращаться уже не к анализу, а к геометрии. Уравнения (21) и (22), вполне уважаемые с точки зрения формальных методов аналитической механики и математики, с геометрической точки зрения — бессмысленны. Они уравнивают между собой два 9-мерных вектора. Одним является вектор, составленный из левых частей уравнений движения

(16-1), (16-2) и (17), а вторым — градиент уравнения троса (18) по координатам этих уравнений. Но это векторы *разных* пространств!

В частности, три значения Ψ , Θ , Φ первого вектора составляют компоненты трехмерного вектора, разложенного по осям связанной с телом подвижной системы координат Эйлера ξ, η, ζ . А соответствующие им три значения $\lambda \partial l / \partial \psi$, $\lambda \partial l / \partial \theta$, $\lambda \partial l / \partial \varphi$ составляют вектор, разложенный по ортам $\psi^0, \theta^0, \varphi^0$ системы координат, которой Эйлер никогда не пользовался. Это трехгранник, орты которого являются локальными касательными к криволинейной координатной сетке углов Эйлера ψ, θ, φ подпространства, которому в 9-мерном пространстве этой задачи принадлежит точка прикрепления r_a троса l (18) к вертолету.

Эти *разные* пространства с разными системами координат конфликтуют, не позволяют складывать свои векторы и, следовательно, *запрещают* писать как уравнение (21), так и уравнения (22).

Этот запрет оказывается общим, и ведет к неустранимому конфликту между методами Лагранжа и Эйлера. Векторы разных систем координат, разумеется, можно перепроектировать. Но перепроектирование вектора $\lambda \partial l / \partial \psi$, $\lambda \partial l / \partial \theta$, $\lambda \partial l / \partial \varphi$ из системы координат $\psi^0, \theta^0, \varphi^0$ в систему координат ξ, η, ζ уничтожает метод множителей Лагранжа. В свою очередь, перепроектирование уравнений вращения Эйлера Ψ , Θ , Φ из системы координат ξ, η, ζ в систему координат $\psi^0, \theta^0, \varphi^0$ уничтожает их как уравнения Эйлера: общая геометрическая теория преобразований уравнений движения [46]–[47] показывает, что такое перепроектирование превращает их в уравнения Лагранжа.

7.5. Почему же иногда благополучно летают вертолеты с грузом?

Ответ очевиден: потому что пилоты вертолетов — первоклассные артисты, *чувствующие* всеми своими органами чувств свой *цирковой снаряд*, и ведомые своей великолепной интуицией. Цирки должны приглашать их для выступлений на манеже!

7.6. Конфликт между методами Лагранжа и Эйлера

Но где же здесь наука? Спустимся на землю, и подытожим сказанное. В динамике связанных тел можно пользоваться теорией Эй-

лера — но без метода множителей Лагранжа. При желании в механических системах со связями пользоваться методом множителей Лагранжа, уравнения движения тел тоже следует записывать в форме Лагранжа. Пример приведен в [46]–[47].

Предупреждение 8. Теория уравнений движения твердого тела Эйлера и метод множителей Лагранжа — несовместимы.

Ряд фундаментальных руководств по механике, не приводя при этом обоснований, в общем, следуют Предупреждениям 7 и 8. Например, в руководстве Г. К. Сулова [4] при использовании уравнений движения Эйлера множители Лагранжа заимствуются из уравнений движения Лагранжа. Э. Дж. Раус [50] составляет уравнения движения несвободного тела с множителями Лагранжа только в форме Лагранжа.

В современных теоретических исследованиях и компьютерных алгоритмах метод виртуальных перемещений и множители Лагранжа формально используются без каких-либо обоснований, при этом в большинстве случаев именно в системах уравнений твердого тела в форме Эйлера. Опасность этой теоретико-террористической деятельности высока — потому что скрыта от традиционных средств ее выявления.

Заключение

Подытоживая, можно сказать, что классическая механика представлена набором теорий, пользующихся *инструментами* математики, но с *методами* математики и между собой — не согласованных.

Отмеченные конфликты между методами математики и механики требуют внимания. Все результаты теоретической и аналитической механики получены путем *преобразования* и *последующего* интегрирования уравнений движения исходных элементов — материальных точек, тел, и механических систем. Какие из этих результатов верные? Они традиционно обосновываются корректностью *отдельных* используемых при их получении инструментов. Но на возможность несовместимости и конфликтов этих корректных по отдельности инструментов *между собой* внимания никогда не обращалось.

Наши тривиальные примеры бросают тень даже на общие законы движения механики. Эти законы устанавливаются путем формального преобразования уравнений движения элементов системы, а Предупреждение 2 на основе примера Задачи 1 утверждает, что это может привести к абсурду. Теория твердого тела Эйлера повсеместно используется без ограничений, но такие ограничения, как показывают Предупреждения 6 и 7, существуют, и должны учитываться. О современном состоянии этих теорий можно сказать, что они получены в результате «правдоподобных рассуждений», отчасти проверенных впоследствии экспериментом и практикой, отчасти так и оставшихся «правдоподобными».

Подчеркнем, что расчетные ошибки, возникающие вследствие обсуждаемых конфликтов между корректными математическими методами, принципиально не могут быть выявлены ни при традиционных текущих проверках правильности вычислений, ни аварийными комиссиями, поскольку получены *правильными* вычислениями на основе *правильных* методов. Эта область может быть исследована только средствами общей теории, не доверяющей порознь ни методам математики, ни методам механики.

На указанные здесь частные замечания и вопросы должны существовать общие ответы. В случае теории твердого тела Эйлера можно дать более глубокие и исчерпывающие ответы [47], другие мы оставим без ответа: цель этого исследования — обратить внимание на важные проблемы с помощью простых примеров. Ответы на выявленные вопросы требуют более глубокого теоретического анализа.

Оправдана ли позиция автора, написавшего эту «антинаучную» статью, бросающую тень на заслуженные методы математики и механики и, попутно, на своих коллег — инженеров и научных работников. И оправдан ли его прием обвинений без предъявления конкретных доказательств — судить читателю. Автор же еще раз напомнит, что корабли тонут, поезда терпят крушения, и самолеты падают с увеличивающейся частотой, несмотря на все более тщательную работу по их инженерному *проектированию*, совершенствованию технологий *изготовления* и улучшению режимов их *эксплуатации*. Значит, должны быть еще другие, общие и еще не выявленные причины!

Поэтому автор поступил так, как обычно поступают при экспертизах катастроф: пока не найдена ее истинная причина, выдвигаются гипотезы, и проверяются на соответствие реальным фактам. Наша цель состояла в том, чтобы число таких гипотез, по возможности, расширить, и на нескольких подозреваемых мы здесь указали. Одного знания математических инструментов для работы в простейшей среди всех современных точных наук, но вовсе не простой механике — недостаточно. Нужно еще уметь ими правильно пользоваться — как и инструментами в любой другой профессии.

Это только фрагмент большой проблемы ревизии теорий прикладных наук. Классическая механика лежит в их основании, ее проблемы оказываются общими, а в этих науках возникают дополнительно и свои аналогичные «проблемы истины». Если эти гипотезы, допускающие *возможность* катастрофы, разыскивать, и выявленные причины, как военных противников Техники и Цивилизации, беспощадно уничтожать, безопасность будет повышаться, а когда будет уничтожена последняя из них, корабли перестанут тонуть, поезда — сходить с рельсов, а самолеты — падать. Конечно, это процесс бесконечный, но нужно постараться сделать его, по крайней мере, монотонно убывающим.

Для молодых читателей есть образцы, с которых нужно брать пример — это Галилей, Ньютон, Лагранж, и наши соотечественники М. В. Келдыш, М. А. Лаврентьев, Л. С. Понтрягин. Выдающиеся математики, они показали, как именно нужно использовать *математику* для решения *содержательных* проблем и задач *механики* и *техники*.

Список литературы

- [1] Галилей Г. Избранные труды. — М.: Наука, 1964.
- [2] Ньютон И. Математические начала натуральной философии // Собрание трудов академика А. Н. Крылова. Т. 7. — М.–Л., 1936.
- [3] Лагранж Ж. Аналитическая механика. — М.–Л.: Гостехиздат, 1950.
- [4] Суслов Г. К. Теоретическая механика. — М.–Л.: Гостехиздат, 1944.

- [5] Величенко В.В. К проблеме управления катастрофами // ДАН. — 1996. Т. 349. № 6. — С. 732–735.
- [6] Величенко В.В. Управление дискретно-непрерывными моделями катастроф // ДАН. — 1996. Т. 350. № 1. — С. 9–11.
- [7] Величенко В.В. К проблеме управления катастрофами // Безопасность России. Раздел 1. — М., 1998. — С. 379–382.
- [8] Величенко В.В. Математическая оболочка для моделирования и управления в типовых сценариях техногенных аварий и катастроф // Science. Вып. 3. — М.: ИМАШ РАН, 1999. — С. 112–113.
- [9] Velichenko V.V. Catastrophe Control Problem // Proceedings of First FLINS'94 Intern. Workshop on Fuzzy Logic and Intelligent Technology in Nuclear Science. — Belgium, Mol: «World Scientific», 1994. — P. 117–121.
- [10] Velichenko V.V. Variational Analysis and Control of Catastrophic Dynamical Systems // Nonlinear Analysis, Theory, Methods & Applications. — Oxford, England: Elsevier Science Ltd, 1997. Vol. 30. No. 4. — P. 2065–2067.
- [11] Величенко В.В. О задачах оптимального управления для уравнений с разрывными правыми частями // Автоматика и телемеханика. — 1966. № 7. — С. 20–30.
- [12] Величенко В.В. Численный метод решения задач оптимального управления // ЖВМ и МФ. — 1966. Т. 6. № 4. — С. 636–647.
- [13] Величенко В.В. Оптимальное управление составными системами // ДАН СССР. — 1967. Т. 176. № 4. — С. 754–756.
- [14] Величенко В.В. Об инвариантности разрывных систем // Автоматика и телемеханика. — 1973. № 8. — С. 155–160.
- [15] Величенко В.В. Задачи оптимального управления с промежуточными условиями // Исследование операций. — ВЦ АН СССР, 1974. Вып. 4. — С. 126–145.
- [16] Величенко В.В. Сильно нелинейные динамические системы: ударная динамика механических систем с односторонними связями // Межд. Конгресс «Нелинейный анализ и его приложения». Тез. докл. — М., 1998. — С. 107.

- [17] Величенко В. В. Матричная теория удара в механических системах с переменным составом связей // 8-й Всероссийский съезд по теоретической и прикладной механике. Аннотации докладов. — Пермь, 2001. — С. 149–150.
- [18] Величенко В. В. Механика трансформирующихся систем // ДАН. — 2003. Т. 388. № 6. — С. 757–760.
- [19] Величенко В. В., Сыромятников В. С. и др. Синтез управления стыковкой космических аппаратов на основе принципа максимальной скорости сближения в метрике расстояний между фазовыми траекториями // Межд. Конгресс «Нелинейный анализ и его приложения». Тез. докл. — М., 1998. — С. 83.
- [20] Величенко В. В. Ударная динамика штатных и аварийных стыковок космических кораблей и станций // Труды XII Байкальской междун. конф. — Иркутск, 2001. Т. 6. — С. 62–67.
- [21] Величенко В. В., Сыромятников В. С. и др. Управление динамикой развертывания космической тросовой системы // 2-я Межд. конф. «Устойчивость и управление для нелинейных трансформирующихся систем». Тез. докл. — М., 2000. — С. 64.
- [22] Величенко В. В., Валувев А. М. О задаче синтеза траекторий полета самолета в трансформирующейся воздушной обстановке // 2-я Межд. конф. «Устойчивость и управление для нелинейных трансформирующихся систем». Тез. докл. — М., 2000. — С. 62.
- [23] Величенко В. В., Козьминых В. А. Управление возвратом воздушно-космического аппарата в атмосферу // 2-я Межд. конф. «Устойчивость и управление для нелинейных трансформирующихся систем». Тез. докл. — М., 2000. — С. 63.
- [24] Величенко В. В., Притыкин Д. А. Искусственный интеллект и социальная динамика ВИЧ-инфицированной иммунной системы человека // Интеллектуальные системы. — 2001. Т. 6. Вып. 1–4. — С. 5–24.
- [25] Величенко В. В., Притыкин Д. А. Численные методы оптимального управления динамикой ВИЧ-инфекции // Известия РАН. Теория и системы управления. — 2006. № 6. — С. 53–64.

- [26] Величенко В. В., Пристыкин Д. А. Управление лечением СПИДа // Автоматика и телемеханика. — 2006. № 6. — С. 166–185.
- [27] Величенко В. В. К проблеме компьютерного интеллекта: решение задачи на выживание, предложенное и реализованное компьютером // XV Межд. конф. Проблемы теоретической кибернетики. — Казань, 2008. — С. 14.
- [28] Величенко В. В., Пристыкин Д. А. Возможности искусственного интеллекта и компьютерных технологий в построении программ лечения сложных иммунных заболеваний // Журнал фундаментальной и прикладной математики. — 2010. Т. 15. Вып. 5.
- [29] Величенко В. В. Принципы технического интеллекта в проблеме управления сложными экономическими системами // Интеллектуальные системы. — 1997. Т. 2. Вып. 1. — С. 5–34.
- [30] Величенко В. В. Принципы и методы современной теории управления в задачах стабилизации и развития экономики // Новая парадигма развития России. Комплексные исследования проблем устойчивого развития / Ред. В. А. Коптюг. — М.: Academia, 1999. — С. 322–328, 436–442.
- [31] Величенко В. В. Интеллектуальные ресурсы в проблеме предотвращения стратегической экономической катастрофы России // Моделирование и управление процессами регионального развития. — М.: Физматлит, 2001. — С. 351–392.
- [32] Величенко В. В. Технический интеллект // Интеллектуальные системы. — 1996. Т. 1. Вып. 1. — С. 5–18.
- [33] Величенко В. В. К методологии интеллекта. Творчество Компьютера // Сб. трудов 2-й Межд. научн. конф. «Философские проблемы математики». — М.: МГУ, 2009.
- [34] Величенко В. В. Интеллектуальный компьютер — теоретик механики. Компьютерные методы в машиноведении // Сборник лекций Межд. научной школы для молодых ученых «Компьютерные методы анализа инженерных задач механики — 2009». — М.: ИМАШ РАН, 2009. — С. 12–19.

- [35] Величенко В. В. Математическое конструирование нейросетевых вычислителей // Сборник докладов научно-технического семинара «Кибернетические системы». Ч. 2. — М., 1999. — С. 29–32.
- [36] Velichenko V. V. Neural Computer's Educational Levels: from Intuition to Formal Mathematics // YU-INFO 99 Symposium. — Yugoslavia, Karaonic, 1999. — P. 1–7.
- [37] Величенко В. В. Нейровычислительные алгоритмы интегрирования обыкновенных дифференциальных уравнений // V Всерос. конф. «Нейрокомпьютеры и их применение». Сборник докладов. — М., 1999. — С. 211–216.
- [38] Величенко В. В. Интуитивные и формализованные методы в нейроматематике. Нейрокомпьютерное интегрирование дифференциальных уравнений // Нейрокомпьютеры: разработка и применение. — 2001. № 2. — С. 9–29.
- [39] Мопертюи П. Законы движения и покоя, выведенные из метафизического принципа // Вариационные принципы механики / под ред. Л. С. Полака. — М.: Физматгиз, 1959. — С. 41–55.
- [40] Якоби К. Лекции по динамике. Общетеchnическая литература. — Л.–М., 1936.
- [41] Полак Л. С. Вариационные принципы механики, их развитие и применения в физике. — М.: Физматгиз, 1960.
- [42] Величенко В. В. О достаточных условиях оптимальности для многолистных полей экстремалей // ДАН СССР. — 1976. Т. 226. № 4. — С. 757–760.
- [43] Величенко В. В. Минимальная интерпретация принципа наименьшего действия // ДАН СССР. — 1979. Т. 248. № 3. — С. 553–556.
- [44] Величенко В. В. Об условиях абсолютной минимальности в принципе наименьшего действия // Устойчивость движения, аналитическая механика, управление движением. — М., 1981. — С. 100–109.
- [45] Пуансо. Об основном положении «Аналитической механики» Лагранжа / Приложение к [3]. — С. 307–318.

- [46] Величенко В. В. Матрицы, геометрия, механика и ЭВМ. — М.: МФТИ, 1984.
- [47] Величенко В. В. Матрично-геометрические методы в механике с приложениями к задачам робототехники. — М.: Наука, 1988.
- [48] Величенко В. В. Матричная механика: возвращение к основаниям Ньютона // Проблемы машиноведения. Сборник трудов. — М.: ИМАШ РАН, 2008. — С. 134–138.
- [49] Апфель П. Теоретическая механика. Т. 2. Динамика системы. Аналитическая механика. — М.: Физматгиз, 1960.
- [50] Раус Э. Дж. Динамика системы твердых тел. Т. I. — М.: Наука, 1983.
- [51] Тарасенко Ф. П. Четыре типа связей в моделировании реальности // Труды XIII Межд. конф. «Проблемы управления и моделирования в сложных системах». — Самара, 2011. — С. 110–114.

ЦЛП-формулировка для совмещенной задачи выбора и планирования инструкций в генераторе кода

Вьюкова Н. И., Галатенко В. А., Самборский С. В.

(Москва, Научно-исследовательский институт системных исследований РАН)

niva@niisi.msk.ru, galat@niisi.msk.ru, sambor@niisi.msk.ru

Введение

Фаза генерации кода в современных оптимизирующих компиляторах включает последовательное выполнение выбора инструкций, планирования инструкций и распределения регистров. Поскольку эти задачи взаимосвязаны, результат их последовательного решения может быть неоптимальным.

В настоящем докладе рассматривается подход, обеспечивающий одновременное выполнение выбора и планирования инструкций с учетом ограничений на число доступных регистров. Он позволяет также в случае дефицита регистров автоматически осуществлять их спиллинг (временное сохранение значения из регистра в память с последующим восстановлением). Предлагаемый метод генерации кода предполагает совместное описание задач выбора и планирования инструкций в виде задачи целочисленного линейного программирования. Он может быть использован как для линейных участков, так и для программной конвейеризации циклов.

Метод отложенного выбора инструкций

Рассмотрим основные моменты предлагаемого метода, который получил название *генерация кода с отложенным выбором инструкций* [1]. Входными данными для генератора кода служат грамматика, описывающая систему команд целевого процессора, и дерево (или лес деревьев), задающее входной линейный участок. Далее мы будем использовать следующий простой пример грамматики:

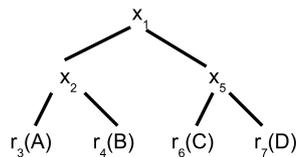
```
N = {R, M}           /* нетерминальные символы */
A = {x, r}           /* терминальные символы */
```

```

P = {
    /* Правила */
    1 R = r :2 /* Загрузка значения на регистр */
    2 R = M :2 /* Чтение из памяти в регистр */
    3 M = R :2 /* Запись регистра в память */
    4 R = x(R,R) :3 /* Выполнение операции x */
}

```

Здесь нетерминальный символ **M** соответствует памяти, **R** — регистр. Терминальный символ **x** соответствует некоторой арифметической инструкции процессора, а **r** описывает значение в памяти или константу. Целые константы, заданные через двоеточия после правил, описывают латентность соответствующих машинных команд. Пример входного дерева для выражения $(A * B) * (C * D)$ показан на рисунке.



Индексы задают нумерацию узлов дерева. Каждому узлу сопоставляется множество виртуальных регистров: V_1, M_1 для узла 1, V_2, M_2 для узла 2, и т. д., исходя из того, что результат подвыражения в узле i может быть вычислен либо в регистре (R_i) , либо в ячейке памяти (M_i) .

Генерация кода состоит из двух шагов: 1) первичный выбор инструкций и 2) планирование с окончательным выбором инструкций.

Мы используем модификацию известного метода BURG [2], одного из декларативных методов выбора команд, обзор которых можно найти в [3]. Алгоритм BURG основан на методах синтаксического разбора, где командам процессора соответствуют правила вывода в некоторой контекстно-свободной грамматике (без однозначности разбора). Применение правила имеет определенную цену; цена дерева разбора — сумма цен примененных правил, что позволяет находить оптимальный разбор методом динамического программирования.

Модификация алгоритма BURG заключается в том, что запоминаются все возможные команды, применимые в узлах дерева (а не

только наиболее дешевые). Для нашего примера будут сгенерировано множество команд (Com):

Узел	Команды			
1	$R1 = x(R2, R5)$	$M1 = R1$	$R1 = M1$	
2	$R2 = x(R3, R4)$	$M2 = R2$	$R2 = M2$	
3	$R3 = r3$	$M3 = R3$	$R3 = M3$	
4	$R4 = r4$	$M4 = R4$	$R4 = M4$	
5	$R5 = x(R6, R7)$	$M5 = R5$	$R5 = M5$	
6	$R6 = r6$	$M6 = R6$	$R6 = M6$	
7	$R7 = r7$	$M7 = R7$	$R7 = M7$	

Команды вида $M_i=R_i$ и $R_i=M_i$ обычно являются избыточными. Но в условиях дефицита регистров планировщик может включить их в расписание для реализации спиллинга.

Окончательный выбор инструкций выполняется во время планирования. Формулировка ЦЛП-задачи планирования с выбором команд в целом аналогична [4]; отличия касаются методов подсчета числа требуемых регистров и требования об однократности выполнения каждой входной команды. Здесь оно заменяется требованием о необходимости выполнения подмножества команд из Com , достаточного для вычисления входного выражения.

В формулировке ЦЛП-задачи используются константы T и $Nreg$. $T > 0$ — задает число тактов, за которое должна выполняться программа. На стадии планирования последовательно делаются попытки составить расписание, укладывающееся в $T = T_{min}, T_{min} + 1, \dots$ тактов, пока решение ЦЛП-задачи не закончится успешно. Константа $Nreg$ указывает число доступных физических регистров.

Из-за ограничений на размер тезисов мы не приводим полную формулировку ЦЛП-задачи. Рассмотрим результат ее решения в условиях острого дефицита физических регистров в предположении, что процессорная архитектура допускает запуск двух команд за такт без дополнительных условий их совместимости.

Не вытаскивая регистры в память, нельзя вычислить выражение $(A * B) * (C * D)$ менее, чем на 3 регистрах (не меняя заданный скобками порядок операций). Попробуем решить построенную ЦЛП задачу

для $Nreg = 2$. При $T < 13$ решения нет, но выделив 13 тактов удастся сгенерировать код:

S	Вирт. регистры хранимые "недо- писанные"	Ассемблерный код
1: R6 = r6 R7 = r7		load f0,C; load f1,D
2:	R6 R7	nop
3: R3 = r3 R5 = x(R6,R7)	R6 R7	load f0,A; mul f1,f0,f1
4:	R3 R5	nop
5:	R3 R5	nop
6: R4 = r4 M5 = R5	R3 R5	load f1,B; store f1,tmp_R5
7:	R3 R4 M5	nop
8: R2 = x(R3,R4) R5 = M5	R3 R4 M5	mul f0,f0,f1; load f1,tmp_R5
9:	R2 R5	nop
10:	R5 R2	nop
11: R1 = x(R2,R5)	R2 R5	mul f0,f0,f1
12:	R1	nop
13:	R1	nop

Легко видеть, что сгенерированный код действительно использует всего два физических регистра. Это достигается за счет того, что на шестом такте содержимое регистра R5 сохраняется в память (и замещается значением переменной B), а на такте 8 загружается обратно. Таким образом, полностью автоматически выбран регистр для спиллинга и сгенерирован спилл-код.

Заключение

При одновременном выборе инструкций и их планировании несложно учесть ограничения по числу регистров и реализовать спиллинг регистров, сохранение значений на регистрах другого типа (менее дефицитных), или повторное вычисление некоторого значения, если это дешевле, чем хранение. Также автоматически реализуется дублирование значений, если это необходимо для параллельной обработки.

Обычно эти преобразования реализуются в компиляторах по отдельности, и редко полностью. Часто реализуется единственное преобразование без которого нельзя обойтись (спиллинг). При этом нет гарантии, что будет оптимально выбран регистр для выталкивания в память. В предлагаемом подходе все управление регистрами происходит комплексно и гарантировано оптимальным образом.

Для этого достаточно в формулировке совмещенной задачи выбора и планирования инструкций разрешить повторное вычисление любого виртуального регистра, а также исполнение любой инструкции более одного раза.

Список литературы

- [1] Вьюкова Н. И. Генерация кода с отложенным выбором инструкций. — М.: НИИСИ РАН, 2010. — С. 80–96.
- [2] Fraser C. W., Henry R. R., Proebsting T. A. BURG — Fast optimal instruction selection and tree parsing // SIGPLAN Notices. — 1992. 27. 4. — P. 68–76.
- [3] Вьюкова Н. И., Галатенко В. А., Самборский С. В. К проблеме выбора машинных инструкций в генераторах кода. — М.: НИИСИ РАН, 2009. — С. 3–31.
- [4] Самборский С. В. Формулировка задачи планирования линейных и циклических участков кода // Программные продукты и системы. — 2007. 3 (79). — С. 12–16.

Сверхразрешение и робастность динамических матриц сенсоров

Григорьева А. М., Пытьев Ю. П.

(Москва, МГУ им. М. В. Ломоносова)

elf1ka@mail.ru, yuri.pytyev@gmail.com

Пусть на рисунке 1а) V — одномерное поле зрения, S — строка сенсоров. f — распределение интенсивности потока, падающего на поле зрения V . g — выходной сигнал матрицы сенсоров, причем выходной сигнал каждого сенсора прямопропорционален потоку интенсивности, падающего на сенсор, и не содержит информации о вариациях потока в его пределах.

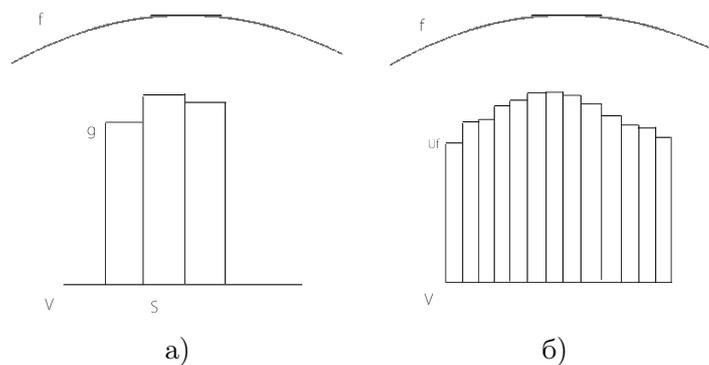


Рис. 1. V — поле зрения, f — распределение интенсивности потока, падающего на V , s — строка сенсоров. а) g — выходной сигнал строки сенсоров, б) s — выходной сигнал строки сенсоров, покрывающей всё поле зрения, и состоящей из сенсоров, размеры которых равны шагу перемещения строки сенсоров.

Определим геометрическую разрешающую способность строки сенсоров как размер светочувствительного элемента. Чем меньше размер сенсора, тем выше геометрическая разрешающая способность строки сенсоров. При регистрации сигнала g будем перемещать строку сенсоров по полю зрения с шагом, меньшим чем размер сенсора, и фиксировать сигнал. Зарегистрированный сигнал методами матема-

тической редукции можно свести к виду, свойственному выходному сигналу Uf (рис. 1б)) строки сенсоров, покрывающей всё поле зрения V , и состоящей из сенсоров, размеры которых равны шагу перемещения строки сенсоров. Таким образом решается задача сверхразрешения — восстанавливается вариация интенсивности потока f до размера, фиксированного шагом перемещения строки сенсоров по полю зрения. Такая система регистрации обладает так же робастностью. Представим, что один из сенсоров строки вышел из строя. Благодаря тому, что строка сенсоров перемещается, интенсивность потока света, падающего на область неработающего сенсора будет зарегистрирована другими, работающими сенсорами. При этом, по виду сигнала Uf нельзя определить какой сенсор вышел из строя, и система регистрации будет продолжать работать.

Для компьютерного моделирования данной системы удобнее регистрировать сигнал не строкой сенсоров, а одним сенсором. Тогда строка сенсоров будет смоделирована таким образом: зафиксируем положения и время нахождения сенсоров строки, затем будем последовательно регистрировать сигнал, перемещая сенсор по полю зрения и задерживаясь в каждой позиции на соответствующее время. Система с одним сенсором универсальна с той точки зрения, что мы можем регулировать время нахождения сенсора в заданной позиции поля в зависимости от того, на сколько хотим понизить уровень шума в данной области поля зрения V . Но для регистрации динамических потоков света лучше объединять сенсоры в матрицу, так как регистрация сигнала одним сенсором требует больше времени.

Данная система регистрации описывается следующей схемой измерений:

$$\begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_p \end{pmatrix} f + \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_p \end{pmatrix}, \quad (1)$$

где g и ν — независимые случайные векторы с математическим ожиданием равным нулю и ковариационными операторами P и K соответственно. Вектор f моделирует распределение интенсивности светового потока, падающего на поле зрения. A_i — оператор, моделирующий отклик матрицы сенсоров на световой поток f в каждом акте

регистрации, $i = 1, \dots, p$. Вектор ν моделирует аддитивный шум. g_i — выходной сигнал матрицы сенсоров. Нужно оценить выходной сигнал Uf матрицы, состоящей из сенсоров с размерами равными шагу сканирования. Для этого минимизируем функционал:

$$\|Rg - Uf\|^2 \sim \min_R, \quad (2)$$

где R — оператор редукции.

Для двумерного поля зрения задача большой, поэтому для вычисления оператора R воспользуемся рекуррентной редукцией:

$$Uf_i = Uf_{i-1} + \frac{UP_iA_i(g_i - A_i f)}{\sigma_i^2 + A_i^* P_i A_i}, \quad (3)$$

$$P_i = P_{i-1} + \frac{P_{i-1}A_i(P_{i-1}A_i)^*}{\sigma_i^2 + A_i^* P_i A_i}, \quad (4)$$

где $i = 1, \dots, p$, $\hat{g}_0^T = 0$, $P_0 = P$, $K = \text{diag}(\sigma_1^2, \dots, \sigma_p^2)$ — ковариационный оператор шума.

На рисунке 2 показано изображение, которое получено с помощью статической матрицы сенсоров с низким разрешением, покрывающей всё поле зрения.

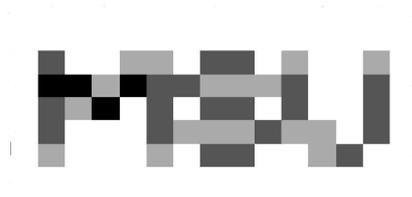


Рис. 2. Изображение, полученное статической матрицей сенсоров, покрывающей всё поле зрения.

На рисунках 3а),б) представлено изображение, полученное методом математической редукции сигнала, зарегистрированного динамической матрицей сенсоров, причем рисунку 3а) соответствует большее время регистрации сигнала, чем изображению, представленному на рисунке 3б).

Рисунки 4а),б) демонстрируют робастность описанной выше системы регистрации изображений. Изображение на рисунке 4а) было получено методом математической редукции сигнала, зарегистрированного стопроцентно работающей матрицей, а на рисунке 4б) матрицей, часть сенсоров которой вышли из строя.



Рис. 3. Изображение, полученное методом математической редукции сигнала.



Рис. 4. Изображение, полученное методом математической редукции.

Рассмотренная система регистрации подобна зрительной системе человека. Глазное яблоко при регистрации изображений совершает непрерывные неконтролируемые человеком движения такие как дрейф, скачки и тремор. И разрешающая способность зрительной системы действительно выше, чем геометрическая разрешающая способность сетчатки, которая определяется размерами колбочек. Причем гибель отдельных колбочек не приводит к заметным дефектам зрения.

Список литературы

- [1] Ярбус А. Л. Роль движений глаз в процессе зрения. — М.: Наука, 1965.
- [2] Демидов В. Е. Построение оценок. — М.: Наука, 2008.
- [3] Пытьев Ю. П. Методы математического моделирования измерительно-вычислительных систем. — М.: ФИЗМАТЛИТ, 2011.

**Задача минимизации числа параллельных процессоров
для системы заданий с ограничениями
предшествования**

Григорьева Н. С. (Санкт-Петербург, СПбГУ)

gns@interzet.ru

Рассмотрим систему заданий $U = \{u_1, u_2, \dots, u_n\}$, на которой отношение частичного порядка \prec задано графом $G = \langle U, E \rangle$, в котором есть дуга $e = (u_i, u_j) \in E$ тогда и только тогда, когда $u_i \prec u_j$. Если $u_i \prec u_j$ то выполнение задания u_j может быть начато только после завершения задания u_i . Задано время выполнения каждого задания $t(u_i)$, которое будем считать целым.

Множество заданий выполняется на параллельных идентичных процессорах, любое задание может выполняться на любом процессоре, и каждый процессор может выполнять не более одного задания в каждый момент времени. Прерывания выполнения заданий не допускаются. Для данной модели можно рассмотреть две задачи составления расписаний [1]. В задаче 1 задано число процессоров, а время выполнения заданий требуется минимизировать. В задаче 2 общее время выполнения заданий задано, требуется минимизировать количество процессоров. Эти задачи, за исключением некоторых частных случаев, являются *NP*-трудными [2]. Для решения этих двух задач предлагается метод ветвей и границ в сочетании с бинарным поиском. Алгоритм решения задачи 1 рассмотрен в [3], ниже рассмотрим задачу 2. Общей подзадачей этих двух задач будет задача построения допустимого расписания при заданном числе процессоров и общем времени выполнения.

Построить расписание, значит найти для каждого задания u_i время начала выполнения задания $\tau(u_i)$ и номер процессора $num(u_i)$, на котором оно выполняется. Длиной расписания S называется величина

$$T_S = \max\{\tau(u_i) + t(u_i) | u_i \in U\}.$$

Требуется определить минимальное число процессоров M_{opt} при котором существует расписание S заданной длины T_S .

Алгоритм нахождения минимального числа процессоров

Пусть M_{opt} минимальное количество процессоров. Определим интервал $(a, b]$ такой, что $a < M_{opt} \leq b$. Нижняя оценка значения целевой функции в данной задаче $LB(M) = \lceil \sum_{i=1}^n t(u_i)/T_S \rceil$. Поэтому положим $a = LB(M) - 1$. В качестве верхней оценки возьмем $b = n$. Тогда $M_{opt} \in (a, b]$.

Алгоритм построения расписания $SCH(U, T_S, a, b; S, M_{opt})$

Пока $b - a > 1$ положим $m := \lceil (a+b)/2 \rceil$. Решим задачу построения допустимого расписания для m процессоров процедурой $BB(U, m; S)$. Если расписание S для m процессоров найдено, то $b := m$, в противном случае $a := m$. В результате определяется минимальное количество процессоров M_{opt} , при котором существует расписаний S заданной длины.

Метод ветвей и границ построения допустимого расписания $BB(U, m; S)$

Требуется построить допустимое расписание S длиной T_S при заданном количестве процессоров m или определить невозможность построения такого расписания.

Для каждого задания u определяется раннее время начала выполнения задания $v_{min}(u)$, и позднее время начала задания $v_{max}(u)$. Возможное время простоя процессоров $I = m \cdot T_S - \sum_{i=1}^n t(u_i)$.

Для частичного решения $\sigma_k = (u_{i_1}, u_{i_2}, \dots, u_{i_k})$ известно время освобождения процессоров после выполнения заданий, включенных в частичное решение, $time_k[1 : m]$ и $t_{min}(k) = \min\{time_k[i] | i \in 1 : m\}$. Для каждого задания $u \in \sigma_k$ определено время начала его выполнения $\tau(u)$, номер процессора $num(u)$, на котором оно выполняется, время простоя процессора перед его началом $r(u)$, резерв простоев $I_k = I - \sum_{i=1}^k r(u_i)$ и R_k — множество заданий запрещенных на уровне перебора k , тех заданий, которые уже добавлялись к частичному решению и привели к недопустимым частичным решениям.

Основным способом сокращения перебора будет как можно более раннее установление недопустимости частичного решения.

Условия недопустимости частичного решения σ_k

- 1) Если существует задание $u_{cr} \notin \sigma_k$, такое что $v_{max}(u_{cr}) < t_{min}(k)$, то частичное решение σ_k недопустимо.
- 2) Рассмотрим интервалы времени $[t_1, t_2] \subseteq [t_{min}(k), T_S]$. Для каждого интервала определим общую мощность процессоров, которые свободны в данном интервале $MP(t_1, t_2)$. Для каждого задания $u_i \notin \sigma_k$, вычислим $v_i = \max\{v_{min}(u_i), t_{min}(k)\}$ возможное время начала задания и найдем интервалы $x_i = [v_i, v_i + t(u_i)] \cap [t_1, t_2]$, и $y_i = [v_{max}(u_i), v_{max}(u_i) + t(u_i)] \cap [t_1, t_2]$. Обозначим $L([t_1, t_2])$ длину интервала $[t_1, t_2]$. Найдем $M_k(t_1, t_2)$ минимальную потребность заданий в ресурсах на интервале $[t_1, t_2]$, и $M_k(t_1, t_2) = \sum_{u_i \notin \sigma_k} \min\{L(x_i), L(y_i)\}$. Если нашелся интервал $[t_1, t_2]$, такой что $M_k(t_1, t_2) > MP(t_1, t_2)$, то частичное решение σ_k недопустимо.
- 3) Частичное решение σ_k недопустимо, если нет заданий $u \notin R_k$, для которых все предшественники u включены в частичное решение σ_k , и выполнено $v_{min}(u) - t_{min}(k) \leq I_k$.
- 4) Если существует задание $u_{cr} \notin \sigma_k$, такое что $v_{max}(u_{cr}) < t_{min}(k)$, тогда для любого задания u , такого что

$$\max\{t_{min}(k-1), v_{min}(u)\} + t(u) > v_{max}(u_{cr}),$$

частичное решение $\sigma_{k-1} \cup u$ недопустимо.

- 5) Если существует задание $u_{cr} \notin \sigma_k = \sigma_{k-1} \cup u_k$, такое что $v_{max}(u_{cr}) < t_{min}(k)$, и выполнены неравенства $v_{max}(u_k) < t_{min}(k)$ и $t_{min}(k-1) + t(u_{cr}) > v_{max}(u_k)$, тогда недопустимо частичное решение σ_{k-1} .

Доказательства утверждений приведены в [3]. Второе условие является модификацией оценок, предложенных в [4].

Алгоритм построения допустимого расписания $BB(U, m; S)$

$k = 0$; $time[i] = 0$; $i \in 1 : m$ $\sigma_0 = \emptyset$;

Для частичного решения σ_k выполнить:

- 1) Проверить для всех заданий $u \notin \sigma_k$, выполнение условия 1. Если оно не выполнено, то $u_{cr} := u$ и перейти к п.6.
- 2) Если не выполнено условие 2, то перейти к п.6.
- 3) Если не выполнено условие 3, то перейти к п.5.
- 4) Выбрать задание u_0 , и добавить к частичному решению σ_k . $k := k + 1$. Если $k = n$, то конец алгоритма, иначе перейти к п.1.
- 5) Аннулировать список заданий, запрещенных на k -ом уровне перебора $R_k := \emptyset$.
- 6) Если $k = 0$, то конец алгоритма. Шаг назад: от частичного решения $\sigma_k = \sigma_{k-1} \cup u_t$ вернуться к решению σ_{k-1} . Задание u_t перевести во множество R_k и $k := k - 1$.
- 7) Если выполнено условие 4, то запретить группу заданий, если выполнено условие 5, то сделать еще один шаг назад. Перейти к п.1.

Если $k = 0$, то допустимого расписания не существует, если $k = n$, то получили допустимое расписание $S = \sigma_n$, с длиной T_S .

Приближенный алгоритм

Для уточнения верхней оценки числа процессоров можно применить приближенный алгоритм. Определим минимально возможное количество процессоров m и будем строить допустимое расписание S заданной длины T_S , проверяя на каждом шаге выполнение условий 1 и 2 допустимости расписания. Если будет установлено, что частичное расписание σ_k недопустимо, то увеличим число процессоров на единицу и начнем построение расписания S сначала уже для большего числа процессоров.

Вычислительный эксперимент подтвердил работоспособность алгоритмов.

Список литературы

- [1] Теория расписаний и вычислительные машины / Под ред. Коффмана Э.Г. — М.: Наука, 1984. (Computer and job-shop Scheduling theory / ed. by E. G. Coffman, J.)

- [2] Ullman J. D. NP-complete scheduling problems // *Journal Comput. System Sci.* — 1975. 10. — P. 384–393.
- [3] Григорьева Н. С. Алгоритм ветвей и границ построения расписания для многопроцессорной системы // *Вестник СПбГУ. Сер. 10.* — 2009. Вып. 1. — С. 44–55.
- [4] Fernandez E., Bussell B. Bounds the number of processors and time for multiprocessor optimal schedules // *IEEE Trans. on Computers.* — 1973. — P. 745–751.

Экспериментальное исследование метода активных контуров

Жизневский А. Н. (Казань, Казанский (Приволжский)
Федеральный Университет)

write2aracon@gmail.com

Введение

При дешифрировании изображений часто используются геометрические признаки. При этом обязательным этапом дешифрирования является получение проекций объектов (сегментация).

В 1988 г. в работе [1] предложен новый подход к задаче поиска границ, названный методом активных контуров (или методом «змей»). В соответствии с ним на плоскости задается семейство кривых, содержащее границу заданного объекта. На этом семействе задается целевая функция, принимающая минимальное значение на границе. Граница выявляется методом последовательных приближений.

Вслед за этой публикацией разными авторами был предложен ряд модификаций и новых методов на основе идеи активных контуров [2–7].

К сожалению, практическое применение этого метода наталкивается на ряд серьезных трудностей, обсуждению которых посвящена настоящая работа.

1. Описание метода

Пусть $v(s) = (x(s), y(s))$, $s \in [0, 1]$ — кривая, параметризованная по длине. Тогда целевая функция (энергия) определяется следующим образом:

$$E(v) = \int_0^1 \left(\frac{1}{2} \alpha(s) \|v'_s(s)\|^2 + \frac{1}{2} \beta(s) \|v''_{ss}(s)\|^2 + E_{ext}(v) \right) ds.$$

Первое слагаемое в скобках растёт при увеличении длины кривой, второе — при увеличении кривизны. Коэффициенты $\alpha(s)$ и $\beta(s)$ определяют вклад длины и кривизны в общую энергию кривой.

В [1] предлагаются различные виды $E_{ext}(v)$. Для поиска границ предлагается использовать градиент сглаженного изображения:

$$E_{ext}(v) = -\|\nabla[G_\sigma(x(s), y(s)) * I(x(s), y(s))]\|^2.$$

Необходимым условием экстремума E является выполнение уравнения Эйлера [8]:

$$-(\alpha(s)v'_s(s))'_s + (\beta(s)v''_{ss}(s))''_{ss} + \nabla E_{ext}(v) = 0.$$

Это уравнение решается численно методом последовательных приближений. Никаких рекомендаций по выбору начального приближения и критерию останова вычислений не предлагается.

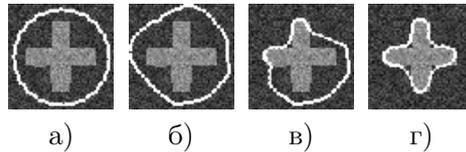


Рис. 1. Работа метода активных контуров. $\alpha = 0,03$, $\beta = 0,0001$, $\sigma = 2,0$.

На рис. 1 представлен пример движения активного контура по изображению размера 60×60 пикселей. Объект (крест) имеет размер 40×40 пикселей и расположен в центре изображения. Здесь и далее значения энергетической яркости точек объекта и фона имеют нормальное распределение со стандартным отклонением 20 и средним значением 60 (фон) и 140 (объект). В качестве начального приближения для метода выбрана окружность (рис. 1, а). На рис. 1 (б,в,г) показаны положения контура после 1000, 3000 и 5000 итераций соответственно. На последующих итерациях значительного изменения в положении контура не наблюдалось.

2. Результаты экспериментов

Важным достоинством подхода является учет замкнутости границы и ограничение кривизны. Кроме того, найденная проекция автоматически получается односвязной. В настоящее время часто используется задание границы с помощью множества уровня (level-set),

позволяющее искать объекты с многосвязными проекциями (см., например, [9]).

При моделировании выявлены два серьезных недостатка метода. Первый связан с зависимостью результата сегментации от выбора начального приближения. На рис. 2 показаны результаты сегментации изображения при двух вариантах выбора начального приближения. В качестве начальных приближений были выбраны окружности, рас-

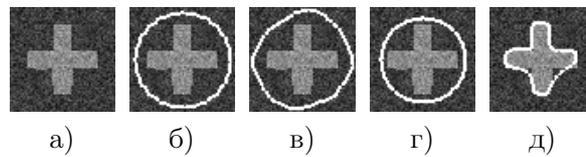


Рис. 2. Зависимость результата сегментации от выбора начального приближения. $\alpha = 0.005$, $\beta = 0.00005$, $\sigma = 1.0$.

положенные в центре изображения, с диаметром 54 пиксела (б) и 48 пикселей (г). Результаты сегментации представлены соответственно на рис. 2, (в, д). Как можно видеть, они существенно отличаются. Методы из более поздних работ, например, GVF [3], позволяют уменьшить зависимость результата от начального приближения.

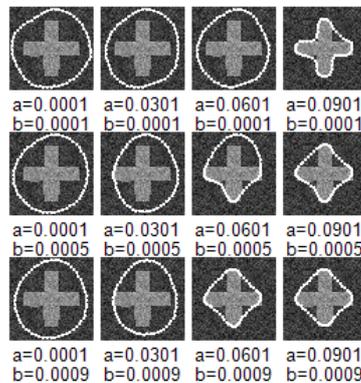


Рис. 3. Зависимость результата сегментации от выбора значений коэффициентов α и β . $\sigma = 1, 0$.

Второй недостаток метода заключается в отсутствии объективного способа выбора коэффициентов α и β , которые могут сильно повлиять на результаты сегментации. На рис. 3 показаны положения активного контура после 30000 итераций при различных значениях α и β . Во всех экспериментах использовалось одно и то же изображение и одно и то же начальное приближение в виде окружности с диаметром 54 пиксела, расположенной в центре изображения. Результаты сегментации для разных α и β существенно отличаются.

Описанные недостатки проявляются и в случае объектов более простых форм (например, квадрат).

Заключение

Проведено экспериментальное исследование метода активных контуров применительно к сегментации изображений. Выявлены зависимость результата сегментации от выбора начального приближения и коэффициентов α и β .

Список литературы

- [1] Kass M., Witkin A., Terzopoulos D. Snakes: Active contour models // *Int. J. Comput. Vis.* — 1988. № 1. — С. 321–331.
- [2] Cohen L.D. On active contours models and balloons // *Computer Vision, Graphics, and Image Processing: Image Understanding.* — 1991. 53 (2). — С. 211–218.
- [3] Xu C., Prince J.L. Gradient vector flow: A new external force for snakes // *IEEE Proc. Conf. on Comp. Vis. Patt. Recog. (CVPR'97).* — 1997. — С. 66–71.
- [4] Williams D.J., Shah M. A fast algorithm for active contours and curvature estimation // *CVGIP: Image Underst.* — 1992. 55 (1). — С. 14–26.
- [5] Amini A.A., Weymouth T.E., Jain R.C. Using dynamic programming for solving variational problems in vision // *IEEE Transactions on pattern analysis and machine intelligence.* — 1990. Vol. 12. № 9. — С. 855–867.

- [6] Melonakos J., Pichon E., Angenent S., Tannenbaum A. Finsler Active Contours // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2008. Vol. 30. № 3. — С. 412–423.
- [7] Caselles V., Kimmel R., Sapiro G. Geodesic active contours // International journal of computer vision. — 1997. Vol. 22. № 1. — С. 61–79.
- [8] Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление. — М.: Наука, 1965.
- [9] Chan T. F., Vese L. A. Active Contours Without Edges // IEEE Trans. on Image Processing. — 2001. № 2. — С. 266–277.

Локальный элиминационный алгоритм и параллельные вычисления

Лемтюжникова Д. В., Щербина О. А.

(Симферополь, Таврический национальный университет
им. В. И. Вернадского)

darabbt@gmail.com, oshcherbina@gmail.com

Из-за высокой вычислительной трудоемкости, а также возможности распараллеливания вычислительных процессов, присущей алгоритмам дискретной оптимизации (ДО), разработка параллельных алгоритмов для решения задач ДО представляет значительный интерес. Параллельные вычислительные системы обладают возможностями параллельного использования большого числа процессоров для обработки информации. Разумное использование вычислительного потенциала посредством распараллеливания вычислительного процесса существенно ускоряет решение практически важных задач ДО большой размерности. Современные достижения параллельных методов комбинаторной оптимизации описаны в [1]. Декомпозиционные алгоритмы ДО являются первоочередными кандидатами для параллелизации [2, с. 225, 257].

Для решения задач ДО с помощью декомпозиционных методов представляют интерес *системы с распределенной памятью*. Параллельное программное обеспечение (ПО) для этих систем использует явную декомпозицию решаемой задачи на подзадачи и назначение их процессорам, а также быструю коммуникационную сеть, обеспечивающую синхронный обмен данными между процессорами. Наиболее типичной архитектурой является Бевульф кластер, важные свойства которого: однородность, коммуникационная сеть типа «все-со-всеми», отсутствие общей памяти и отсутствие глобальных часов для синхронизации вычислений. В такой архитектуре все процессы передачи сообщений и синхронизации должны осуществляться с помощью явного обмена сообщениями между процессорами, используя «message-passing» протокол. Программирование для системы с распределенной памятью осуществляется с помощью стандартного языка высокого уровня типа C++ или Fortran с явной передачей со-

общений. Имеются два стандартных API для передачи сообщений: Message-Passing Interface (MPI) и Parallel Virtual Machine (PVM).

Остановимся на вопросах разработки ПО для параллельных алгоритмов решения задач ДО. SYMPHONY [3] — это параллельная система ПО для решения задач смешанного целочисленного линейного программирования (СЦЛП), подобная COIN/BCP [4]. COIN/BCP и SYMPHONY объединены в новом решателе ALPS [5]. Кроме того, существуют такие параллельные решатели, как PUBB [6], PPBLib [7] и PICO [8]. PARINO [9] и FATCOP [10] — параллельные решатели общего назначения для решения задач СЦЛП, причем FATCOP предназначен для грид-систем. ParaLEX [11] — это параллельное расширение решателя CPLEX для задач СЦЛП, предназначенное для работы в распределенной вычислительной среде. Более современной системой является программная система CHiPPS (COIN-OR High Performance Parallel Search), созданная для параллелизации поисковых алгоритмов на дереве [12].

При параллельной реализации локального элиминационного алгоритма (ЛЭА) [13] предпочтительны методы top-down, работающие с макрографами потоков данных, так как структура графа потоков данных известна (элиминационное дерево или древовидная декомпозиция). Высокая вычислительная сложность, один из основных недостатков ЛЭА, может быть снижена с помощью параллельных вычислений. ЛЭА основан на использовании элиминационного дерева, которое обладает необходимой информацией о зависимости между данными в задаче ДО. Высота элиминационного дерева представляет собой (грубую) меру вычислительной работы по параллельной элиминации. Элиминационное дерево управляет процессом элиминации переменных и блоков переменных и указывает потоки данных, а также характеризует возможности распараллеливания при решении задачи ДО с помощью ЛЭА. В случае вычислительных машин с распределенной памятью после построения элиминационного дерева далее необходимо решить задачу отображения переменных на процессоры. Основными целями этого отображения являются хорошая балансировка нагрузки и низкая межпроцессорная коммуникация, причём они могут конфликтовать друг с другом. При параллельной реализации блочного ЛЭА используемый граф потоков данных является

деревом, которое должно обрабатываться от листьев к корню. Когда для вершины дерева найдено частичное решение, множество промежуточных данных передается в родительскую вершину. Как только промежуточная информация от всех вершин-потомков собрана в родительской вершине, она используется при решении соответствующей родительской вершине задачи ДО. Таким образом, независимые ветви дерева могут обрабатываться параллельно; будем называть этот тип параллельности (древовидной) параллельностью первого типа. В общем случае, он может быть использован более эффективно в нижней части дерева, нежели вблизи ее корневой вершины. Второй тип параллельности соответствует параллельным задачам для одного блока. При этом используется парадигма «Master-Worker», в которой имеются два различных типа процессоров: мастер-процессор и рабочие процессоры. При этом управление алгоритмом осуществляется мастер-процессором: он разбивает задачу на отдельные подзадачи и назначает их для решения рабочим процессорам, учитывая взаимосвязи между ними. Такая парадигма особенно полезна, когда задача разбита на легко решаемые подзадачи, а также при слабой взаимосвязи между этими подзадачами. Мастер-процессор выбирается в начале решения пакета задач, соответствующих блоку древовидной декомпозиции. Далее, рабочие процессоры выбираются мастером динамически из списка процессоров-кандидатов, используя критерий хорошей балансировки нагрузки.

Грид-компьютинг (или метакомпьютинг) в общем описывает параллельные вычисления на географически распределенной и гетерогенной платформе [14]. При этом могут использоваться рабочие станции с общей памятью, узлы кластеров ПК, кластеры и суперкомпьютеры. Потенциал грид-компьютинга [14] используется в настоящее время в ДО лишь частично. Декомпозиционные алгоритмы решения задач дискретной оптимизации (включая ЛЭА) хорошо подходят для грид-платформ.

При разработке параллельных версий ЛЭА представляется перспективным использование грид-компьютинга с использованием интерфейса Condor, а также технология PVM.

Список литературы

- [1] Ferreira A., Pardalos P. (eds.) Solving Combinatorial Optimization Problems in Parallel: Methods and Techniques // LNCS 1054. State-of-the-Art Surveys. — Springer-Verlag, 1996.
- [2] Bertsekas D. P., Tsitsiklis J. N. Parallel and Distributed Computation: Numerical Methods. — Englewood Cliffs: Prentice-Hall, 1989.
- [3] Ralphs T. K., Ladanyi L., Salzman M. J. Parallel branch, cut and price for large-scale discrete optimization // Mathematical Programming. — 2003. B. 98. — P. 253–280.
- [4] Ralphs T. K., Ladanyi L. COIN/BCP User's Guide. — 2001. URL: <http://www.coin-or.org> (дата обращения 30.07.2011).
- [5] Ralphs T. K., Ladanyi L., Saltzman M. J. A library hierarchy for implementing scalable parallel search algorithms // Journal of Supercomputing. — 2004. V. 28 (2). — P. 215–234.
- [6] Shinano Y., Higaki M., Hirabayashi R. Control schemas in a generalized utility for parallel branch and bound // Proc. of the 1997 Eleventh International Parallel Processing Symposium. — Los Alamitos: IEEE, Computer Society Press, 1997.
- [7] Tschoke S., Polzer T. Portable parallel branch-and-bound library PPBBLib user manual. — Department of computer science, Univ. of Paderborn, 1996.
- [8] Eckstein J., Phillips C. A., Hart W. E. PICO: An Object-Oriented Framework for Parallel Branch and Bound / RUTCOR Research Report 40–2000. — Rutgers University, Piscataway, 2000.
- [9] Linderoth J. T. Topics in Parallel Integer Optimization / Ph.D. Dissertation. — Georgia Institute of Technology School of Industrial and Systems Engineering, 1998.
- [10] Chen Q., Ferris M. C., Linderoth J. T. Fatcop 2.0: Advanced features in an opportunistic mixed integer programming solver // Annals of Operations Research. — 2001. V. 103. — P. 17–32.
- [11] Shinano Y., Fujie T. ParaLEX: A parallel extension for the CPLEX mixed integer optimizer // Recent Advances in Parallel Virtual

- Machine and Message Passing Interface / F. Cappelto, T. Herault, J. Dongarra (eds.) — Springer, 2007. — P. 97–106.
- [12] Xu Y., Ralphs T.K., Ladanyi L., Saltzman M.J. Computational experience with a software framework for parallel integer programming // The INFORMS Journal on Computing. — 2009. V. 21. — P. 383–397.
- [13] Щербина О. А. Локальные элиминационные алгоритмы решения разреженных дискретных задач // Журнал вычислительной математики и математической физики. — 2008. Т. 48. № 1. — С. 161–177.
- [14] Foster I., Kesselman C. (eds). The Grid: Blueprint for a New Computing Infrastructure. — San Francisco: Morgan Kaufmann Publishers Inc., 1998.

**О частотно-временных признаках многоканальных
электроэнцефалограмм мозга при заболевании
Паркинсона на ранней стадии**

Обухов Ю. В., Королев М. С.
(ИРЭ им. В. А. Котельникова РАН)

obukhov@cplire.ru, korolevclub@mail.ru

Введение

Важной проблемой является диагностика болезни Паркинсона (БП) на ранней стадии. Одним из методов диагностики является электроэнцефалография (ЭЭГ). Спектры сигналов ЭЭГ и магнитной энцефалографии пациентов с диагнозом БП было принято анализировать с помощью Фурье преобразования [1, 2]. Но Фурье анализ ЭЭГ не позволяет исследовать динамику электрической активности мозга. Следует отметить, что характерной чертой паркинсонизма обычно признается синдром дезинтеграции, проявляющийся на разных системных уровнях (двигательные нарушения, вегетативная, нейрогуморальная дезинтеграция, эмоциональные и психические нарушения) [3, 4]. Полагают, что при ранних формах БП эти нарушения носят преимущественно функциональный нейродинамический характер. Предполагается, что такая дезинтеграция может отражаться также в динамике электрической активности мозга. В данной работе описан метод анализа вейвлет спектрограмм ЭЭГ, направленный на поиск частотно-временных признаков раннего паркинсонизма. Суть метода заключается в том, что исследуется распределение локальных максимумов всплесков вейвлет преобразования по частоте и времени.

Метод

В работе исследовалась фоновая электрическая активность мозга, то есть испытуемый сидел в кресле с закрытыми глазами в расслабленном состоянии. Было обследовано шестнадцать людей с БП 1-ой стадии по шкале Хен-Яра (ХЯ) [5], четырнадцать людей со 2-ой стадией, восемь людей 2,5–3 стадией, и одиннадцать здоровых испы-

туемых. ЭЭГ записывалась с электродов со стандартным расположением 10x20 с референтными электродами на ушах. Оцифрованные записи ЭЭГ были обработаны фильтром Баттерворта 4-ого порядка для удаления частот 50, 75 и 100 Гц, а также частот ниже 1 Гц. Исследование частотно-временной динамики электрической активности проводилось с помощью вейвлет преобразования ЭЭГ:

$$W(\tau, T) = \frac{1}{\sqrt{T}} \int x(t) \psi\left(\frac{t-\tau}{T}\right) dt$$

В качестве материнской функции используется комплексный вейвлет Морле:

$$\psi(\eta) = \frac{1}{\sqrt{\pi F_b}} e^{2i\pi F_c \eta} e^{-\frac{\eta^2}{F_b}}$$

где, $F_b = F_c = 1$.

Вейвлет спектрограмма (рис. 1) состоит из временных серий всплесков, расположенных в различных частотных диапазонах — тета, альфа, бета и др. Суть метода заключается в том, что исследуется распределение локальных максимумов всплесков вейвлет преобразования по частоте и времени. В качестве примера на Рис. 1 представлено вейвлет преобразование сигнала в центральном отведении СЗ здорового испытуемого 27 лет. Для нахождения местоположения всплесков на плоскости частота — время делается две проекции вейвлет преобразования — одна на плоскость частота — амплитуда, другая — на плоскость время — амплитуда. С помощью первой проекции выделяются частотные диапазоны хребтов вейвлет спектрограмм. Далее, в каждом из диапазонов частот вейвлет спектрограмма проецируется на плоскость время — амплитуда. Это необходимо для того, чтобы всплески высоких амплитуд одного хребта не перекрывали всплески других диапазонов. Далее вычисляются координаты положения всех всплесков. Таким образом, задача поиска локальных максимумов вейвлет-преобразования на плоскости частота — время сводится к поиску локальных максимумов в двух плоскостях — время — амплитуда и частота — амплитуда.

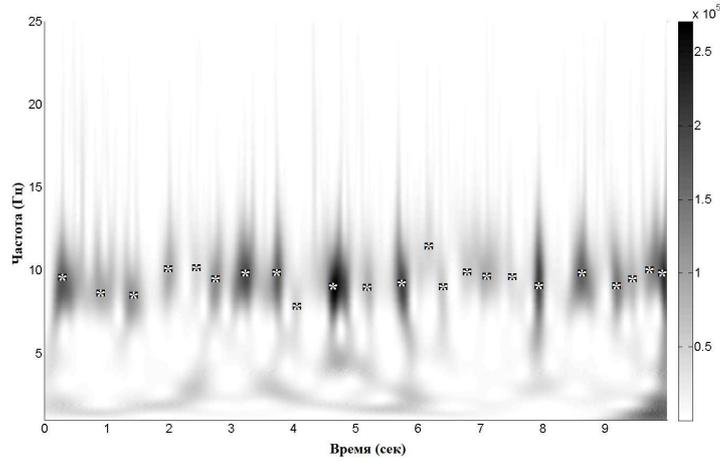


Рис. 1. Вейвлет спектрограмма ЭЭГ здорового испытуемого с отведения СЗ в полосе частот 1–25 Гц.

В результате, получены 2 координаты локальных максимумов и значения их амплитуд. На рис. 2 представлены положения локальных максимумов на плоскости время — частота с симметричных пар отведений СЗ (в виде крестиков) и С4 (в виде ноликов), отвечающих за моторику, здорового испытуемого и пациента с заболеванием Паркинсона на 1-ой стадии.

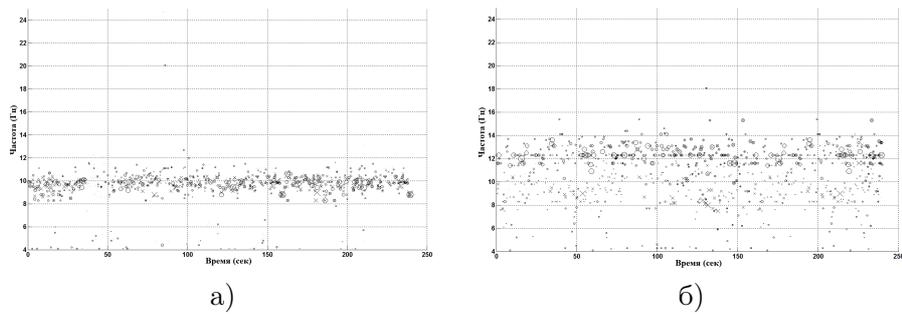


Рис. 2. а). Испытуемый К. (здоровый человек, 27 лет), б) Испытуемый Е. (пациент на 1-ой стадии БП, 29 лет) X — отведение СЗ, O — отведение С4.

Для анализа динамики изменения частоты доминирующего ритма во времени в каждой полосе частот 0,5 Гц суммируются амплитуды локальных максимумов вспышек в окне 16 секунд и сдвигом окна на 16 секунд на всем временном интервале продолжительностью 240 секунд. Соответствующие гистограммы распределений частот вспышек с симметричных пар отведений ЭЭГ отображаются на одном графике (рис. 3).

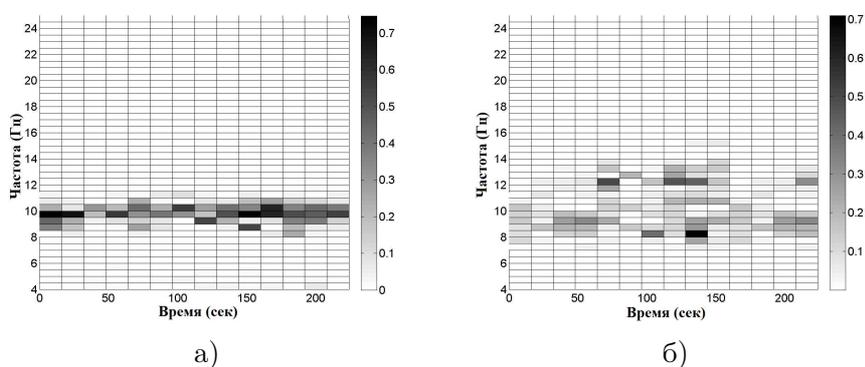


Рис. 3.

Из рис. 3 видно, что у здорового человека вспышки распределены в узком диапазоне частот от 9 до 11 Гц. На ранней стадии паркинсонизма происходит увеличение частоты основного ритма в левом полушарии (X), уширение диапазона частот и проявляется межполушарная асимметрия электрической активности.

Заключение

Разработан новый метод частотно-временного анализа электрической активности мозга, ориентированный на поиск признаков заболевания Паркинсона на ранней стадии. 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 заболевания и

представлениям о дезорганизации электрической активности мозга при паркинсонизме.

Работа поддержана Программой Президиума РАН «Фундаментальные науки — медицине»

Список литературы

- [1] Stoffers D., Bosboom J.L.W., Dejen J.B., Wolters E.C., Berendse H.W., Stam C.J. Slowing of oscillatory brain activity is a stable characteristic of Parkinson's disease without dementia // Brain. — 2007. 130. — P. 1847–1860.
- [2] Soikkeli R., Partenen J., Soininen H., Paakkonen A., Riekkonen P. Sr. Slowing of EEG in Parkinson's disease // EEG Clin. Neurophysiol. — 1991. 79. — P. 159–165.
- [3] Вейн А. М., Голубев В. Л., Яхно Н. Н. Паркинсонизм с позиций функционально-неврологического анализа // Паркинсонизм: Вопросы клиники, патогенеза и лечения. — М., 1974. — С. 57–65.
- [4] Голубев В. Л., Левин Я. И., Вейн А. М. Болезнь Паркинсона и синдром паркинсонизма. — М.: МЕДпресс, 1999.
- [5] Movement Disorder Society Task Force on Rating scales for PD. Unified Parkinson's disease rating scale: status and recommendations // Mov. Disord. — 2003. 18. — P. 738–750.

Модель информационно-поискового сайта филиала МГУ

Осокин В. В. (Москва, МГУ им. М. В. Ломоносова)

osvic@mail.ru

Ставится задача построения универсальной информационно-поисковой системы, на базе которой можно реализовать веб-сайт любого филиала МГУ, факультета или кафедры. Система должна включать в себя объекты разных типов, таких как, например, сотрудник, студент, курс, практикум и т. д. В дальнейшем будем называть такие объекты *сущностями*. Система должна обеспечивать авторизованным пользователям интерфейс добавления произвольных *типов сущностей* с произвольными полями. Также система должна обеспечивать авторизованным пользователям интерфейс добавления произвольных сущностей введенных в систему типов.

Базовым типом сущности системы является *веб-страница*. Сущности такого типа имеют всего два поля: название и html-описание сущности. Все остальные типы сущности «наследуют» данный тип: в дополнение к указанным двум полям для каждого типа может быть указано неограниченное число дополнительных полей. Так, для сущностей типа «человек» логично добавить поле «дата рождения».

Сущности могут связываться друг с другом *отношениями*. Так, несколько сущностей типа «человек» могут быть связаны отношением «студенты группы 507». В системе подразумеваются различные *типы отношений*. В отношениях типа «группа» все члены-сущности равноправны, в отношениях типа «процесс» все члены-сущности упорядочены по дате и так далее. Система должна обеспечивать авторизованным пользователям интерфейс добавления произвольных отношений и типов отношений.

Существует специальный тип отношений — *структурные отношения*. В структурные отношения могут входить только сущности типа «веб-страница», причем любая «веб-страница» входит лишь в одно структурное отношение. Структурные отношения позволяют строить сайты на базе сущностей типа «веб-страница» в стандартном понимании термина «сайт». Так, структурному отношению ставится

в соответствие сущность, которая задает шаблон всем сущностям-членам структурного отношения. Эту сущность можно включить в новое структурное отношение и тем самым обеспечить иерархию сущностей-страниц сайта.

Еще одним объектом системы является так называемый *виджет*. Виджеты вставляются в html-поля сущностей и могут отображать другие сущности или отношения. Так, если мы хотим на одной странице сайта (то есть сущности типа «веб-страница») сначала отобразить некоторую сущность типа «человек», скажем, профессора, а ниже отобразить все курсы, читаемые этим профессором (отношение типа «группа», связывающее его сущности-курсы), то мы в описание этой сущности-веб-страницы вставляем два виджета: виджет, отображающий сущность (человека) и виджет, отображающий отношение (список курсов).

Особенностью системы является возможность давать различные доменные имена различным сущностям-веб-страницам системы. Так, у филиала МГУ в рамках системы может быть одно доменное имя, у одного его факультета — второе, у второго — третье, и так далее. При этом, например, одна и та же сущность типа «человек» может быть представлена на всех трех сайтах тремя различными сущностями-веб-страницами при помощи виджетов. При изменении настоящей сущности изменится и ее отображение на всех этих сайтах.

В реализации системы возможность неограниченного добавления сущностей, отношений, типов сущностей и типов отношений достигается путем применения NOSQL СУБД MongoDB. В отличие от стандартных реляционных SQL СУБД, в MongoDB вместо таблиц используются коллекции, а вместо строк таблиц — документы, причем структура документа не определяется коллекцией. Другими словами, каждый документ может иметь свои уникальные поля. При этом даже поля, присутствующие лишь в некоторых документах, можно индексировать. В результате, мы получаем единое хранилище сущностей и отношений, оптимизированное для поиска в нем.

Среди других используемых в системе технологий стоит выделить серверный язык php (фреймворк Yii) и клиентский язык JavaScript (библиотека jQuery).

Один из первых сайтов, реализованных в рамках системы — сайт ташкентского филиала МГУ <http://uz.msu.ru>.

Модель гипертекстовой организации информации в информационных системах для исторических исследований

Перевертень В. А. (Москва, РГГУ)

dpva@mail.ru

Как отмечают специалисты по применению информационных технологий в исторической науке, создание информационных систем для исторических исследований является более сложной и специальной задачей, чем построение традиционных информационных систем. Разработка таких систем требует решения целого ряда научных и технологических задач, центральное место среди которых занимает задача организации связанной с историческими исследованиями информации.

Историко-исследовательская информация должна быть организована так, чтобы обеспечить определенную комфортность, качество и эффективность познавательно-информационной деятельности историков. Организация информации будет тем лучше, чем больше она будет соответствовать специфике исторических исследований и особенностям информационной деятельности историков как исследователей.

Ряд работ, собственные наблюдения и беседы с историками дают основания гипотетически утверждать, что с информационной точки зрения в исторических исследованиях процесс познания субъектом объекта состоит в накоплении информации фрагментами и объединении и связывании образов (моделей) этих информационных фрагментов.

Фрагмент информации либо вычленяется исследователем из имеющегося в его распоряжении массива информации, либо порождается им самим. Рассматривая фрагмент по частям и воспринимая его содержание с различных точек зрения, с одним и тем же информационным фрагментом исследователь может связывать множество образов, по-разному отражающих этот фрагмент.

Для борьбы с «проклятием размерности», которое выражается в огромном количестве информационных фрагментов и их образов, на-

капливающихся в процессе решения исследовательской задачи, применяются следующие два приема. Первый прием заключается в выделении части образов и объединении их в некоторые совокупности, исходя из самых произвольных соображений познающего субъекта. Второй прием состоит в связывании образов друг с другом также по воле субъекта.

Исходя из рассмотренного представления об информационной деятельности историка, по нашему мнению, наиболее подходящей является гипертекстовая организация историко-исследовательской информации. Именно гипертекст позволяет упорядочить информацию, которую невозможно формализовать и выделить в ней необходимые для построения баз данных и баз знаний категории, у которой нет априорной структуры, с которой исследователь может работать, лишь выделяя в ней фрагменты и связывая их.

Предлагаемая нами модель гипертекстовой (гипермедиа) организации историко-исследовательской информации несколько отличается от «классического» представления о гипертексте. В ее основе лежат следующие понятия: информационный объект (ИО), образ ИО, объект гипертекста (ОГТ), ассоциация и связь.

Информационный объект — это фрагмент информации, который может объективно существовать или порождаться как единица информации, произвольно выделяться из имеющегося информационного массива или формироваться познающим субъектом. На содержание, форму представления, объем, структурную организацию и носитель информационного фрагмента, понимаемого как информационный объект, не накладывается никаких ограничений.

Образ ИО — это обозначение одного или нескольких аспектов содержания ИО, это некоторые признаки, отражающие содержание ИО, это вторичная информация, являющаяся моделью содержания ИО.

Образ ИО, дополненный уникальным идентификатором его и указателем на ИО (уникальным идентификатором ИО), к которому этот образ относится, называется *объектом гипертекста*. ОГТ выступают в качестве узлов гипертекстовой сети и играют роль «представителей» ИО в гипертексте. С одним ИО может быть связано несколько ОГТ (один — первичный, а остальные — вторичные), представляю-

щих его с разных сторон. При такой организации сам ИО, будучи включенным в гипертекст, является пассивным компонентом гипертекста и существует независимо от него.

Ассоциациями мы называем группы ОГТ, которые исследователь формирует исходя из объективных свойств образов ИО или своих субъективных соображений. Каждая ассоциация имеет уникальное имя. Один и тот же ОГТ может входить в несколько ассоциаций.

Под *связью* в предлагаемой модели гипертекстовой организации информации подразумевается симметричное бинарное отношение, устанавливаемое между двумя ОГТ относительно определенных ассоциаций, в которых связываемые ОГТ находятся. Именно в таком определении связей заключается существенное отличие нашей модели от модели «классического» гипертекста.

В формализованном виде нашу модель гипертекстовой организации информации (*HT*-модель) можно определить как пару:

$$HT = (HTS, HTP),$$

где *HTS* — множество допустимых гипертекстовых структур;

HTP — множество допустимых операций над этими структурами.

Пусть *NH* — множество идентификаторов ОГТ, *IM* — множество образов ИО, *ID* — множество идентификаторов ИО, *IO* — множество ИО, *NA* — множество имен ассоциаций. Допустимые гипертекстовые структуры (*HTS*-структуры) определяются соответствующими отношениями, заданными на перечисленных множествах и ограничениями целостности. В предлагаемой нами модели *HTS*-структура — это кортеж:

$$HTS = (HTO, IDIO, ASS, LNK, HTSL),$$

где *HTO* — множество ОГТ;

IDIO — отношение идентификатор ИО — ИО;

ASS — отношение ассоциации;

LNK — множество связей;

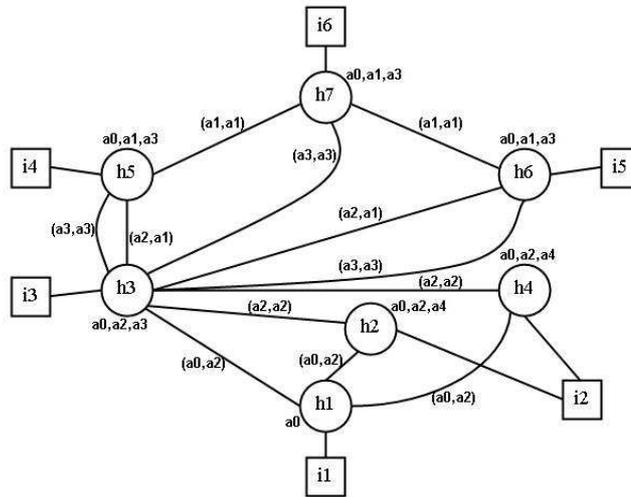
HTSL — ограничения целостности гипертекстовой структуры.

$HTO \subset NH \times IM \times ID$, $IDIO \subset ID \times IO$, $ASS \subseteq NH \times NA$, $LNK \subseteq NH \times NA \times NH \times NA$, при этом $\pi_1 HTO \leftrightarrow \pi_{2,3} HTO$, $\pi_1 IDIO \leftrightarrow \pi_2 IDIO$, элементы отношения LNK представляют собой кортежи вида $(h1, a1, h2, a2)$, где $h1, h2 \in \pi_1 HTO$, $a1 \in ASS(h1)$, $a2 \in$

$ASS(h2)$, причем $h1 \neq h2$ и для каждого кортежа в LNK имеется симметричный ему кортеж $(h2, a2, h1, a1)$.

Начальное состояние HTS -структуры соответствует утверждению: $(HTO = \emptyset) \& (IDIO = \emptyset) \& (ASS = \emptyset) \& (LNK = \emptyset)$. В состояниях, отличных от начального, она должна удовлетворять ограничениям целостности $HTSL$, которые заключаются в утверждении: $(\pi_1 IDIO = \pi_3 HTO) \& (\pi_1 ASS = \pi_1 HTO) \& (\pi_{1,2} LNK \subseteq ASS)$.

Наглядно HTS -структуру можно представить в виде графа, пример которого показан ниже ($h1$ – $h7$ — идентификаторы ОГТ, $i1$ – $i6$ — идентификаторы ИО, $a0$ – $a4$ — имена ассоциаций):



В множество операций HTP входит 15 операций трех типов:

- операции формирования гипертекста, которыми определяются изменения состояния HTS -структуры;
- операции выборки элементов гипертекстовой структуры, предназначенные для отбора тех ее элементов, с участием которых будут выполняться некоторые операции;
- операции выделения части гипертекстовой структуры в соответствии с заданными условиями.

Примером информационной системы для исторических исследований, в которой организация информации основывается на представленной выше *HT*-модели, может служить созданная автором система «Просис» [1, 2].

Список литературы

- [1] Гутнов Д. А., Перевертень В. А. Российские историки XVIII–начала XX вв.: проект и информационная система // Круг идей: новое в исторической информатике. Труды I конференции Ассоциации «История и компьютер». — М., 1994. — С. 39–50.
- [2] Гутнов Д. А., Перевертень В. А. Просопографическая информационная система «Просис»: версия 2.0 // Информационный бюллетень Ассоциации «История и компьютер». — М., 1994. № 10. — С. 17–18.

Математическое моделирование субъективных суждений модельера-исследователя о модели объекта исследования

Пытьев Ю. П. (Москва, МГУ им. М. В. Ломоносова)

yuri.pytyev@gmail.com

Рассмотрены методы математического моделирования неполного и недостоверного знания модели $M(x)$ объекта исследования, выраженного в форме субъективных суждений модельера-исследователя (м.-и.) о возможных значениях неизвестного параметра $x \in X$. Математическая модель «неизвестного параметра» определена как неопределенный элемент (н. э.) \tilde{x} , характеризующий (как неопределенная высказывательная переменная) субъективные суждения м.-и. об истинности каждого значения $x \in X$ значениями мер правдоподобия $Pl^{\tilde{x}}$ равенства $\tilde{x} = x$ и доверия $Vel^{\tilde{x}}$ неравенства $\tilde{x} \neq x$.

Предисловие

Вероятностные и возможность методы применяются при моделировании многих аспектов как неясности и неопределенности, отражающих неполноту и недостоверность информации, так и случайности, нечеткости и неточности, относящихся к её содержанию. Модель случайности и нечеткости обычно считается вероятностной или возможностьной, а неясность и неопределенность ассоциируются с неполным знанием последней, но «моделируются», как правило, вербально [1, 2, 3]. Обычно «неполное знание» вероятностной $(\Omega, \mathcal{A}, Pr(\cdot; x))$ или возможностьной $(\Omega, \mathcal{P}(X), P(\cdot; x))$ моделей обусловлено их зависимостью от *неизвестного* параметра $x \in X$. В научной, инженерной и прочей исследовательской и творческой деятельности невозможно исключить использование неполной, противоречивой и недостоверной информации, ассоциированной с опытными фактами, с практикой их использования и с полученными знаниями. В работе рассмотрен метод математического моделирования подобной информации о возможных значениях неизвестного параметра $x \in X$, выраженной в форме субъективных суждений [4, 5].

Существо метода иллюстрирует конструкция *неопределенного случайного элемента*, заданного на произведении пространств: вероятностного $(\Omega, \mathcal{A}, \text{Pr}(\cdot; x))$, известного с точностью до значения $x \in X$ и моделирующего *неопределенный объект исследования*, и пространства $(X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ с мерами правдоподобия $\text{Pl}^{\tilde{x}}$ и доверия $\text{Bel}^{\tilde{x}}$, моделирующего неопределенный элемент (н. э.) \tilde{x} , характеризующий субъективные суждения исследователя об истинности каждого значения параметра $x \in X$ значениями мер правдоподобия $\text{Pl}^{\tilde{x}}(\tilde{x} = x)$ и доверия $\text{Bel}^{\tilde{x}}(\tilde{x} \neq x)$.

В работе показано, как такая модель «н. э. \tilde{x} » позволяет *вычислять* правдоподобие и доверие *любых* суждений исследователя о *любых* свойствах объекта, обусловленных его моделью $(\Omega, \mathcal{A}, \text{Pr}(\cdot; x))$, $x \in X$.

Неопределенный элемент. Меры правдоподобия Pl и доверия Bel

Пусть семейство моделей $M(x)$, $x \in X$, — *неопределенная модель* объекта исследования. Модель «неизвестного параметра» $x \in X$ — *неопределенный элемент* (н. э.) \tilde{x} , характеризующий субъективные суждения исследователя об истинности каждого значения $x \in X$ значениями мер правдоподобия $\text{Pl}^{\tilde{x}}(\tilde{x} = x)$ и доверия $\text{Bel}^{\tilde{x}}(\tilde{x} \neq x)$.

В модели $(X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ н. э. \tilde{x} меры правдоподобия $\text{Pl}^{\tilde{x}}(\cdot) : \mathcal{P}(X) \rightarrow \mathcal{L}$ и доверия $\text{Bel}^{\tilde{x}}(\cdot) : \mathcal{P}(X) \rightarrow \hat{\mathcal{L}}$, где \mathcal{L} и $\hat{\mathcal{L}}$ суть шкалы их значений, определены равенствами

$$\begin{aligned} \text{Pl}^{\tilde{x}}(E) &\stackrel{\text{def}}{=} \text{Pl}^{\tilde{x}}(\tilde{x} \in E) = \sup_{x \in E} t^{\tilde{x}}(x), \quad \text{Bel}^{\tilde{x}}(E) \stackrel{\text{def}}{=} \text{Bel}^{\tilde{x}}(\tilde{x} \in E) = \\ &= \inf_{x \in X \setminus E} s^{\tilde{x}}(x), \quad (1) \end{aligned}$$

в которых $E \in \mathcal{P}(X)$ и

$$t^{\tilde{x}}(x) \stackrel{\text{def}}{=} \text{Pl}^{\tilde{x}}(\tilde{x} = x), \quad s^{\tilde{x}}(x) \stackrel{\text{def}}{=} \text{Bel}^{\tilde{x}}(\tilde{x} \neq x). \quad (2)$$

— значения правдоподобия равенства $\tilde{x} = x$ и доверия неравенства $\tilde{x} \neq x$, $x \in X$. Функции $t^{\tilde{x}}(\cdot) : X \rightarrow [0, 1]$ и $s^{\tilde{x}}(\cdot) : X \rightarrow [0, 1]$ —

распределения правдоподобий и доверий значений н. э. \tilde{x} . Определения (1), (2) и их содержательная интерпретация суть следствия условий:

- Модельер-исследователь *может*, основываясь на своих неполных и недостоверных априорных знаниях свойств объекта, считать $x \in X$ значениями н. э. \tilde{x} и *предложить его модель* $(X, \mathcal{P}(X), Pl^{\tilde{x}}, Bel^{\tilde{x}})$, указав в (2), насколько, по его мнению, *относительно правдоподобны* равенства $\tilde{x} = x, x \in X$, и насколько *следует относительно доверять* неравенствам $\tilde{x} \neq x, x \in X$, где «относительно» означает, что в $(X, \mathcal{P}(X), Pl^{\tilde{x}}, Bel^{\tilde{x}})$

1) численные значения $Pl^{\tilde{x}}(E)$ и $Bel^{\tilde{x}}(E), E \in \mathcal{P}(X)$, в (1), *отличные от нуля и единицы, не могут быть содержательно истолкованы, а существенна лишь их упорядоченность*,

2) меры $Pl^{\tilde{x}}(\cdot)$ и $Pl^{\tilde{x}}(\cdot)$ ($Bel^{\tilde{x}}(\cdot)$ и $Bel^{\tilde{x}}(\cdot)$) считаются эквивалентными, если $\exists \gamma(\cdot) \in \Gamma \forall E \in \mathcal{P}(X) \gamma(Pl^{\tilde{x}}(E)) = Pl^{\tilde{x}}(E) (\exists \gamma(\cdot) \in \Gamma \forall E \in \mathcal{P}(X) \gamma(Bel^{\tilde{x}}(E)) = Bel^{\tilde{x}}(E))$, где Γ — группа непрерывных, строго монотонных функций $\gamma(\cdot) : [0, 1] \rightarrow [0, 1], \gamma(0) = 0, \gamma(1) = 1$, с групповой операцией « \circ », $\gamma \circ \gamma'(a) \stackrel{\text{def}}{=} \gamma(\gamma'(a)), a \in [0, 1]$.

Условия 1), 2) в терминах свойств шкал \mathcal{L} и $\hat{\mathcal{L}}$, означают, что

- класс Γ определяет *группу автоморфизмов шкал* $\mathcal{L} = ([0, 1], \leq, +, \times) \equiv ([0, 1], \leq, \max, \min)$ и $\hat{\mathcal{L}} = ([0, 1], \hat{\leq}, \hat{+}, \hat{\times}) \equiv ([0, 1], \hat{\leq}, \min, \max)$ значений мер $Pl^{\tilde{x}}(\cdot) : \mathcal{P}(X) \rightarrow \mathcal{L}$ и $Bel^{\tilde{x}}(\cdot) : \mathcal{P}(X) \rightarrow \hat{\mathcal{L}}$, то есть

$\forall \gamma(\cdot) \in \Gamma$ в шкалах $\gamma \mathcal{L}$ и $\gamma \hat{\mathcal{L}}: \gamma([0, 1]) = [0, 1], \forall a, b \in [0, 1] a \leq b \Leftrightarrow \gamma(a) \leq \gamma(b), a \hat{\leq} b \Leftrightarrow \gamma(a) \hat{\leq} \gamma(b)$, и $\gamma(a * b) = \gamma(a) * \gamma(b)$, где $*$ — символ любой из *операций*: сложения $+$, $\hat{+}$ и умножения $\times, \hat{\times}$.

Равенства $a + b = a \hat{\times} b = \max\{a, b\}, a \times b = a \hat{+} b = \min\{a, b\}, a, b \in [0, 1]$, следуют из требований [3] *непрерывности* и коммутативности операции $*$ как отображения $[0, 1]^2 \rightarrow [0, 1]$ и свойств *нейтральных элементов* 0 и 1 шкал \mathcal{L} и $\hat{\mathcal{L}}$ и группы $\Gamma: \forall a \in [0, 1] a + 0 = a \hat{\times} 0 = a \times 1 = a \hat{+} 1 = a, a + 1 = a \hat{\times} 1 = 1$ и $a \times 0 = a \hat{+} 0 = 0$.

Поскольку шкалы $\gamma \mathcal{L}, \gamma \in \Gamma$, и $\gamma \hat{\mathcal{L}}, \gamma \in \Gamma$, *изоморфны*, исследователи могут формулировать модели н. э. \tilde{x} в *произвольных* шкалах $\mathcal{L}, \hat{\mathcal{L}}$. Будучи сформулированными в шкалах $\mathcal{L}', \hat{\mathcal{L}}'$ и $\mathcal{L}'', \hat{\mathcal{L}}''$, модели считаются *эквивалентными*, если *существуют* шкалы $\mathcal{L} = \gamma' \mathcal{L}' = \gamma'' \mathcal{L}''$ и $\hat{\mathcal{L}} = \hat{\gamma}' \hat{\mathcal{L}}' = \hat{\gamma}'' \hat{\mathcal{L}}''$, $\gamma', \gamma'', \hat{\gamma}', \hat{\gamma}'' \in \Gamma$, в которых их формулировки совпадают, а *содержательно истолкованы* могут быть только те, форму-

лировки которых *не зависят* от выбора шкал \mathcal{L} , $\widehat{\mathcal{L}}$, то есть *одинаковы для всех исследователей*.

В терминах операций сложения и умножения в шкалах \mathcal{L} и $\widehat{\mathcal{L}}$
 $\forall E \in \mathcal{P}(X) \text{Pl}^{\tilde{x}}(E) \stackrel{\text{def}}{=} \text{Pl}^{\tilde{x}}(\tilde{x} \in E) = \bigoplus_{x \in E} t^{\tilde{x}}(x)$, $\text{Bel}^{\tilde{x}}(E) \stackrel{\text{def}}{=} \text{Bel}^{\tilde{x}}(\tilde{x} \in E) = \bigoplus_{x \in X \setminus E} s^{\tilde{x}}(x)$, причем для $E = X$ и $E = \emptyset$ принимаются независимые от выбора шкал \mathcal{L} и $\widehat{\mathcal{L}}$ условия $\text{Pl}^{\tilde{x}}(X) = \bigoplus_{x \in X} t^{\tilde{x}}(x) = 1$, $\text{Bel}^{\tilde{x}}(\emptyset) = \bigoplus_{x \in X} s^{\tilde{x}}(x) = 0$, означающие, что среди значений $x \in X$ есть истинное, определяющее модель объекта исследования.

Согласно (1), (2) если $\varphi(\cdot) : X \rightarrow Y$, $\varphi^{-1}(Y) = X$, — функция, задающая н. э. $\tilde{y} = \varphi(\tilde{x})$, то пространство $(Y, \mathcal{P}(Y), \text{Pl}^{\tilde{y}}, \text{Bel}^{\tilde{y}})$ — модель \tilde{y} , в которой

$$\begin{aligned} \text{Pl}^{\tilde{y}}(A) &= \text{Pl}^{\tilde{x}}(\varphi(\tilde{x}) \in A) = \sup_{y \in A} t^{\tilde{y}}(y), \\ \text{Bel}^{\tilde{y}}(A) &= \text{Bel}^{\tilde{x}}(\varphi(\tilde{x}) \in A) = \inf_{y \in Y \setminus A} s^{\tilde{y}}(y), \quad A \in \mathcal{P}(Y), \end{aligned} \quad (3)$$

где

$$\begin{aligned} t^{\tilde{y}}(y) &\stackrel{\text{def}}{=} \text{Pl}^{\tilde{y}}(\tilde{y} = y) = \text{Pl}^{\tilde{x}}(\varphi(\tilde{x}) = y) = \sup_{\substack{x \in X \\ \varphi(x) = y}} t^{\tilde{x}}(x), \\ s^{\tilde{y}}(y) &\stackrel{\text{def}}{=} \text{Bel}^{\tilde{y}}(\tilde{y} \neq y) = \text{Bel}^{\tilde{x}}(\varphi(\tilde{x}) \neq y) = \inf_{\substack{x \in X, \\ \varphi(x) = y}} s^{\tilde{x}}(x). \end{aligned} \quad (4)$$

Заметим, что исследователь *в любом случае может* считать $x \in X$ значениями н. э. \tilde{x} и *предложить* его модель $(X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$, ибо может воспользоваться моделями «полного знания» и «полного незнания» модели $M(x)$, $x \in X$, и свойств объекта.

Модель «полного незнания» модели $M(x)$, $x \in X$, и свойств объекта определяется распределениями: $t^{\tilde{x}}(x) = 1$, $x \in X$, (все значения н. э. \tilde{x} равноправдоподобны), $\sup_{x \in X} t^{\tilde{x}}(x) = 1$, и $s^{\tilde{x}}(x) = 0$, $x \in X$, (любому неравенству $\tilde{x} \neq x$, $x \in X$, доверять нельзя), $\inf_{x \in X} s^{\tilde{x}}(x) = 0$. В этих случаях такие же распределения в (4) будет иметь любая функция $\varphi(\tilde{x}) = \tilde{y}$, $\varphi^{-1}(Y) = X$: $t^{\tilde{y}}(y) = 1$, $s^{\tilde{y}}(y) = 0$, $y \in Y$.

Модель «полного знания» модели объекта определяется распределениями:

$$t^{\tilde{x}}(x) \stackrel{\text{def}}{=} \text{Pl}^{\tilde{x}}(\tilde{x} = x) = \begin{cases} 1, & x = x_0, \\ 0, & x \neq x_0, \end{cases} \quad x \in X, \quad (x_0 \text{ — единственное прав-}$$

доподобное значение параметра)

$$\text{и } s^{\tilde{x}}(x) \stackrel{\text{def}}{=} \text{Bel}^{\tilde{x}}(\tilde{x} \neq x) = \begin{cases} 1, & x \neq x_0, \\ 0, & x = x_0, \end{cases} \quad x \in X, \quad (x_0 \text{ — единственное зна-}$$

чение параметра, при котором неравенству $\tilde{x} \neq x_0$ доверять нельзя), а в (4) в этом случае

$$t^{\tilde{y}}(y) = \begin{cases} 1, & y = y_0, \\ 0, & y \neq y_0, \end{cases} \quad y \in Y, \quad s^{\tilde{y}}(y) = \begin{cases} 1, & y \neq y_0, \\ 0, & y = y_0, \end{cases} \quad y \in Y, \quad \text{где } y_0 = \varphi(x_0).$$

Разумеется, модели «полного незнания» и «полного знания» не зависят от выбора шкал $\gamma\mathcal{L}$ и $\hat{\gamma}\hat{\mathcal{L}}$, $\gamma, \hat{\gamma} \in \Gamma$, то есть являются таковыми для любого исследователя.

Неопределенный элемент как неопределенная высказывательная переменная

Н.э. \tilde{x} моделирует *высказывания* исследователя о его значениях, истинность и ложность которых не абсолютны. Такая интерпретация н.э. основана на *теоретико-множественном представлении логики высказываний*, согласно которому в $(X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ X — множество элементарных высказываний (э.в.), любое высказывание a взаимно однозначно представлено множеством $A \in \mathcal{P}(X)$ э.в. $x \in X$, каждое из которых влечет a : $a \leftrightarrow A = \bigcup_{\substack{x \in X \\ x \rightarrow a}} \{x\} \equiv \{x \in X, x \rightarrow a\}$, где \leftrightarrow и \rightarrow суть символы взаимно однозначного соответствия и логической импликации. Каждое э.в. $x \in X$ представлено в X одноточечным множеством $\{x\}$, $x \leftrightarrow \{x\}$, и выделено среди всех высказываний условием: любое э.в. $x \in X$ не следует ни из какого высказывания, кроме x и всегда ложного высказывания $\mathbf{0}$. При этом если $a \leftrightarrow A$, $b \leftrightarrow B$, то $a \& b \leftrightarrow A \cap B$, $a \vee b \leftrightarrow A \cup B$, $\neg a \leftrightarrow X \setminus A$, $a \rightarrow b \equiv (\neg a) \vee b \leftrightarrow (X \setminus A) \cup B$, $\mathbf{1} \leftrightarrow X$, $\mathbf{0} \leftrightarrow \emptyset$.

В таком представлении $t^{\tilde{x}}(x) = \text{Pl}^{\tilde{x}}(\tilde{x} = x)$ ($\text{Pl}^{\tilde{x}}(\tilde{x} \in E)$) — правдоподобие истинности неопределенного высказывания (н.в.) $x \leftrightarrow \{x\}$

($e \leftrightarrow E$), согласно которому $\tilde{x} = x$ ($\tilde{x} \in E$). Соответственно $s^{\tilde{x}}(x) = \text{Bel}^{\tilde{x}}(\tilde{x} \neq x)$ ($\text{Bel}^{\tilde{x}}(\tilde{x} \in E)$) — доверие истинности н. в. $\neg x \leftrightarrow X \setminus \{x\}$ ($e \leftrightarrow E$), согласно которому $\tilde{x} \neq x$ ($\tilde{x} \in E$), $x \in X$.

Предложив модель $(X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ н. э. \tilde{x} , исследователь согласно (3), (4) может вычислять правдоподобия и доверия истинности *любых* н. в. о значениях *любых* характеристик объекта как функций н. э. \tilde{x} . Пусть, например, $M(x)$, $x \in X$, — семейство произвольных моделей, среди которых есть модель $M(x_0)$ объекта, но x_0 неизвестно. Если $(\Omega, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ — его модель н. э. \tilde{x} , то $M(\tilde{x})$ — неопределенная модель объекта, и для любой неопределенной характеристики объекта $\varphi(\tilde{x}) \stackrel{\text{def}}{=} \mathbf{C}(M(\tilde{x}))$ правдоподобие и доверие истинности высказываний, согласно которым $\varphi(\tilde{x}) = y$ и $\varphi(\tilde{x}) \neq y$, $y \in Y$, определены в (4).

Далее $M(x) = (\Omega, \mathcal{A}, \text{Pr}(\cdot; x))$, $x \in X$, характеристики $C(M(\tilde{x}))$ модели $M(\tilde{x})$ суть неопределенные вероятности $\widetilde{\text{Pr}}(A) \stackrel{\text{def}}{=} \text{Pr}(A; \tilde{x})$, $A \in \mathcal{A}$, неопределенные числовые характеристики случайных величин, заданных на $(\Omega, \mathcal{A}, \widetilde{\text{Pr}})$ и т. п.

Неопределенные случайный элемент, вероятность и математическое ожидание

Для пары пространств: вероятностного $(Y, \mathcal{B}, \text{Pr}^\eta)$ и с правдоподобием и доверием $(X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ рассмотрим отображение $q(\cdot, \cdot) : (Y, \mathcal{B}) \times (X, \mathcal{P}(X)) \rightarrow (\Omega, \mathcal{A})$, \mathcal{B}, \mathcal{A} измеримое при каждом $x \in X$: $\forall x \in X \forall A \in \mathcal{A} \{y \in Y, q(y, x) \in A\} \in \mathcal{B}$ и $\forall x \in X q^{-1}(\Omega, x) = Y$.

Определение. Функцию $\tilde{\xi} = q(\eta, \tilde{x})$ случайного η и неопределенного \tilde{x} элементов назовем *неопределенным случайным элементом, заданным на произведении пространств $(Y, \mathcal{B}, \text{Pr}^\eta) \times (X, \mathcal{P}(X), \text{Pl}^{\tilde{x}}, \text{Bel}^{\tilde{x}})$ и принимающим значения в (Ω, \mathcal{A}) .*

При $\tilde{x} = x$ $\tilde{\xi} \Big|_{\tilde{x}=x} \stackrel{\text{def}}{=} \xi(x) = q(\eta, x)$ — случайный элемент, определенный на $(Y, \mathcal{B}, \text{Pr}^\eta)$ со значениями в $(\Omega, \mathcal{A}, \text{Pr}(\cdot; x))$, $x \in X$, где

$$\text{Pr}(A; x) \stackrel{\text{def}}{=} \text{Pr}^{\xi(x)}(\xi(x) \in A) = \text{Pr}^\eta(q(\eta, x) \in A) = \int_{\substack{y \in Y, \\ q(y, x) \in A}} \text{Pr}^\eta(dy), \quad A \in \mathcal{A},$$

вероятность, определяющая семейство вероятностных пространств $(\Omega, \mathcal{A}, \text{Pr}(\cdot; x))$, $x \in X$, *неопределенную вероятность* $\widetilde{\text{Pr}}(A) \stackrel{\text{def}}{=} \text{Pr}(A; \tilde{x})$, $A \in \mathcal{A}$, и неопределенное вероятностное пространство $(\Omega, \mathcal{A}, \widetilde{\text{Pr}})$ как неопределенную модель объекта исследования.

Согласно (4) $t^{\widetilde{\text{Pr}}(A)}(\text{pr}) \stackrel{\text{def}}{=} \text{Pl}^{\tilde{x}}(\text{Pr}(A; \tilde{x}) = \text{pr}) = \sup\{t^{\tilde{x}}(x) \mid x \in X, \text{Pr}(A; x) = \text{pr}\}$ — правдоподобие истинности н. в., согласно которому неопределенная вероятность $\widetilde{\text{Pr}}(A)$ события A равна pr . Соответственно $s^{\widetilde{\text{Pr}}(A)}(\text{pr}) \stackrel{\text{def}}{=} \text{Bel}^{\tilde{x}}(\text{Pr}(A; \tilde{x}) \neq \text{pr}) = \inf\{s^{\tilde{x}}(x) \mid x \in X, \text{Pr}(A; x) = \text{pr}\}$ — доверие истинности н. в., согласно которому $\widetilde{\text{Pr}}(A) \neq \text{pr}$, $\text{pr} \in [0, 1]$.

Рассмотрим, например, семейство вероятностей $\text{Pr}(\cdot; x) : \mathcal{P}(\Omega) \rightarrow [0, 1]$, $x \in X = (0, \infty)$, где $\Omega = \{\omega_1, \omega_2, \dots\}$,

$$\text{Pr}(\{\omega_i\}; x) \stackrel{\text{def}}{=} \text{pr}_i(x) = x/(1+x)^i, \quad i = 1, 2, \dots, \quad (5)$$

x — значение н. э. \tilde{x} , и неопределенное вероятностное пространство $(\Omega, \mathcal{P}(\Omega), \widetilde{\text{Pr}})$, в котором $\widetilde{\text{Pr}}(A) \stackrel{\text{def}}{=} \text{Pr}(A; \tilde{x})$, $A \in \mathcal{P}(\Omega)$, как неопределенную модель объекта.

Правдоподобие $\text{Pl}^{\tilde{x}}(\forall i \in \{1, 2, \dots\} \tilde{\text{pr}}_i = \text{pr}_i)$ н. в., согласно которому $\forall i = 1, 2, \dots \tilde{\text{pr}}_i \stackrel{\text{def}}{=} \text{pr}_i(\tilde{x}) = \text{pr}_i$, не равно нулю лишь на семействе $\text{pr}_i = \text{pr}_i(y) = y/(1+y)^i$, $i = 1, 2, \dots$, $y \in (0, \infty)$, поэтому $\text{Pl}^{\tilde{x}}(\forall i \in \{1, 2, \dots\} \tilde{\text{pr}}_i = \text{pr}_i(y)) = \text{Pl}^{\tilde{x}}(\tilde{x} = y) = t^{\tilde{x}}(y)$ и, соответственно, $\text{Bel}^{\tilde{x}}(\exists i \in \{1, 2, \dots\} \tilde{\text{pr}}_i \neq \text{pr}_i(y)) = \text{Bel}^{\tilde{x}}(\tilde{x} \neq y) = s^{\tilde{x}}(y)$ суть правдоподобие н. в., согласно которому $\widetilde{\text{Pr}}(\cdot) = \text{Pr}(\cdot; y)$, и доверие н. в., согласно которому $\widetilde{\text{Pr}}(\cdot) \neq \text{Pr}(\cdot; y)$, $y \in (0, \infty)$.

Поэтому $t^{\tilde{x}}(y)$ — *правдоподобие истинности н. в.*, согласно которому $(\Omega, \mathcal{P}(\Omega), \text{Pr}(\cdot; y))$ *есть модель объекта*, $s^{\tilde{x}}(y)$ — *доверие истинности н. в.*, согласно которому $(\Omega, \mathcal{P}(\Omega), \text{Pr}(\cdot; y))$ *не есть его модель*, $y \in (0, \infty)$.

Рассмотрим неопределенное математическое ожидание $\widetilde{\text{E}}\xi_\lambda \stackrel{\text{def}}{=} \text{E}_{\tilde{x}}\xi_\lambda(\cdot)$ случайной величины $\xi_\lambda(\cdot) : \Omega \rightarrow [0, 1]$, зависящей от $\lambda > 0$, заданной равенствами $\xi_\lambda(\omega_i) = \lambda^{i-1}e^{-\lambda}/(i-1)!$, $i = 1, 2, \dots$ на семействе $(\Omega, \mathcal{P}(\Omega), \text{Pr}(\cdot; x))$, $x \in X$. Так как согласно (5) $\text{E}_x\xi_\lambda(\cdot) = \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} e^{-\lambda} \frac{x}{(x+1)^i} = \frac{x}{x+1} \exp\left(-\frac{\lambda x}{x+1}\right)$, $x \in X$, то $t^{\widetilde{\text{E}}\xi_\lambda}(m) \stackrel{\text{def}}{=} \text{Pl}^{\tilde{x}}(\text{E}_{\tilde{x}}\xi_\lambda(\cdot) =$

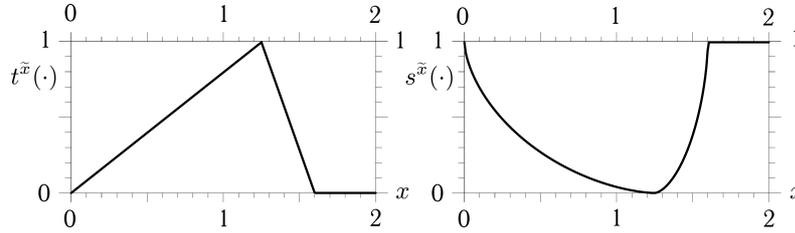


Рис. 1. Графики распределений $t^{\tilde{x}}(x)$ и $s^{\tilde{x}}(x)$, $x \in X = [0, 2]$, (2), предложенных исследователем.

$m) = \sup\{t^{\tilde{x}}(x) \mid x \in (0, \infty), x(x+1)^{-1} \exp(-\lambda x(x+1)^{-1}) = m\}$, $m \in (0, \infty)$, — распределение правдоподобий н. в., согласно которым $\tilde{E}\xi_\lambda = m$, $s^{\tilde{E}\xi_\lambda}(m) \stackrel{\text{def}}{=} \text{Bel}^{\tilde{x}}(\mathbf{E}_{\tilde{x}}\xi_\lambda(\cdot) \neq m) = \inf\{s^{\tilde{x}}(x) \mid x \in (0, \infty), x(x+1)^{-1} \exp(-\lambda x(x+1)^{-1}) = m\}$, $m \in (0, \infty)$, — распределение доверий н. в., согласно которым $\tilde{E}\xi_\lambda \neq m$. см. рис. 1, 2.

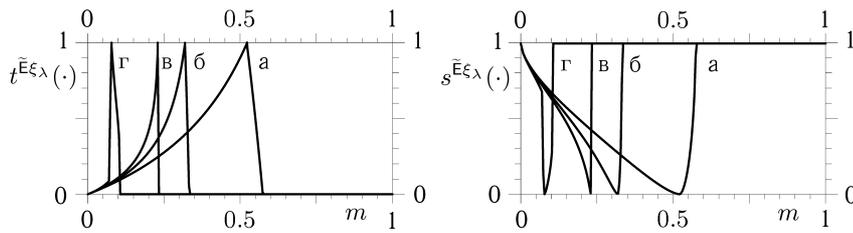


Рис. 2. Графики распределений $t^{\tilde{E}\xi_\lambda}(m)$ и $s^{\tilde{E}\xi_\lambda}(m)$, $m \in (0, \infty)$, для $\lambda = 0.11$ (а), $\lambda = 1$ (б), $\lambda = 1.58$ (в), $\lambda = 3.55$ (г).

Оптимальные значения н. э.

Иногда н. э., в частности, $\tilde{E}\xi_\lambda$, удобнее охарактеризовать не распределениями $t^{\tilde{E}\xi_\lambda}(\cdot)$ и $s^{\tilde{E}\xi_\lambda}(\cdot)$, а указав его, в известном смысле, *оптимальное значение*. В таком случае исследователь должен построить семейство пространств $(L, \mathcal{P}(L), \text{Pl}^{\tilde{l}}(\cdot|m, m'), \text{Bel}^{\tilde{l}}(\cdot|m, m'))$, $(m, m') \in (0, \infty)^2$, в котором m — значение $\tilde{E}\xi_\lambda$, m' — его оценка, L — пространство *элементарных потерь*, $\mathcal{P}(L)$ — класс всех подмножеств L , $\text{Pl}^{\tilde{l}}(\cdot|\cdot, \cdot)$, $\text{Bel}^{\tilde{l}}(\cdot|\cdot, \cdot)$ суть переходные правдоподобие и доверие для пространств $((0, \infty)^2, \mathcal{P}((0, \infty)^2))$, $(L, \mathcal{P}(L))$, и для каждого $\lambda \in (0, \infty)$

указать множество $V(\lambda) \in \mathcal{P}(L)$ существенных для него элементарных потерь.

Пусть $\tilde{t}^l(l|m, m')$, $l \in L$, и $\tilde{s}^l(l|m, m')$, $l \in L$, распределения $Pl^l(\cdot|m, m')$ и $Bel^l(\cdot|m, m')$, $m, m' \in (0, \infty)$. Тогда $Pl^l(V(\lambda)|m, m') = \sup_{l \in V(\lambda)} \tilde{t}^l(l|m, m') \stackrel{\text{def}}{=} pll_{m, m'}$ и $Bel^l(V(\lambda)|m, m') = \inf_{l \in L \setminus V(\lambda)} \tilde{s}^l(l|m, m') \stackrel{\text{def}}{=} bell_{m, m'}$ суть правдоподобие $pll_{m, m'}$ и доверие $bell_{m, m'}$ потерь, сопутствующих решению $\tilde{E}\xi_\lambda = m'$ исследователя о значении неопределенного математического ожидания $\tilde{E}\xi_\lambda$, в то время как на самом деле $\tilde{E}\xi_\lambda = m$, $m, m' \in (0, \infty)$.

Соответственно $\int_{m \in (0, \infty)} (pll_{m, m'} \times t^{\tilde{E}\xi_\lambda}(m)) \stackrel{\text{def}}{=}} pl_{t^{\tilde{E}\xi_\lambda}}(pll_{., m'}) = Pl$ (потерь, сопутствующих решению $m' \in (0, \infty)$). Аналогично $\int_{m \in (0, \infty)} (bell_{m, m'} \times s^{\tilde{E}\xi_\lambda}(m)) \stackrel{\text{def}}{=}} bel_{s^{\tilde{E}\xi_\lambda}}(bell_{., m'}) = Bel$ (потерь, сопутствующих решению $m' \in (0, \infty)$).

Исследователь принимает оптимальные решения m'^* и m'_* как минимизирующие правдоподобие $pl_{t^{\tilde{E}\xi_\lambda}}(pll_{., m'})$ и доверие $bel_{s^{\tilde{E}\xi_\lambda}}(bell_{., m'})$ потерь на множестве решений $m' \in (0, \infty)$:

$$\begin{aligned} pl_{t^{\tilde{E}\xi_\lambda}}(pll_{., m'^*}) &\stackrel{\text{def}}{=} \sup_{m \in (0, \infty)} \min\{pll_{m, m'^*}, t^{\tilde{E}\xi_\lambda}(m)\} \equiv \\ &\equiv \int_{m \in (0, \infty)} (pll_{m, m'^*} \times t^{\tilde{E}\xi_\lambda}(m)) = \min_{m' \in (0, \infty)} \sup_{m \in (0, \infty)} \min\{pll_{m, m'}, t^{\tilde{E}\xi_\lambda}(m)\}, \end{aligned}$$

$$\begin{aligned} bel_{s^{\tilde{E}\xi_\lambda}}(bell_{., m'_*}) &\stackrel{\text{def}}{=} \inf_{m \in (0, \infty)} \max\{bell_{m, m'_*}, s^{\tilde{E}\xi_\lambda}(m)\} \equiv \\ &\equiv \int_{m \in (0, \infty)} (bell_{m, m'_*} \times s^{\tilde{E}\xi_\lambda}(m)) = \min_{m' \in (0, \infty)} \inf_{m \in (0, \infty)} \max\{bell_{m, m'}, s^{\tilde{E}\xi_\lambda}(m)\}, \end{aligned}$$

где pl - и bel -интегралы $pl_{t^{\tilde{E}\xi_\lambda}}(pll_{., m'})$, $m' \in (0, \infty)$, и $bel_{s^{\tilde{E}\xi_\lambda}}(bell_{., m'})$, $m' \in (0, \infty)$, предполагаются полунепрерывными снизу.

Автор признателен Ю. М. Нагорному, рассчитавшему зависимости на рис. 2 и подготовившему электронный вариант рукописи.

Список литературы

- [1] Пытьев Ю. П. Методы математического моделирования измерительно-вычислительных систем. — М.: Физматлит, 2011 (третье издание).
- [2] Дюбуа Д., Прад А. Теория возможностей. — М.: Радио и связь, 1990.
- [3] Пытьев Ю. П. Возможность как альтернатива вероятности. — М.: Физматлит, 2011 (второе издание).
- [4] Пытьев Ю. П. Неопределенные нечеткие модели и их применения // Интеллектуальные системы. — 2004. Т. 8. Вып. 1–4. — С. 147–310.
- [5] Пытьев Ю. П. Эмпирическое восстановление мер возможности и правдоподобия возможности в моделях экспертных решений // Автоматика и телемеханика. — 2010. № 3. — С. 131–146.

Математический метод определения каталитической активности ферментов в сложных биологических растворах

**Садовничий В. А., Ветров Д. П., Вишневский В. В.,
Галатенко А. В., Галатенко В. В., Зыкова Т. В.,
Коршунов А. А., Лебедев А. Е., Лукашенко Т. П.,
Подольский В. Е., Политов А. В.**
(Москва, ИМИСС МГУ им. М. В. Ломоносова)

vgalat@imscs.msu.ru

Информация о каталитической активности ферментов в сложных биологических растворах (в частности, в крови) важна для корректной постановки медицинского диагноза и выбора правильной стратегии лечения. При наличии априорной информации о присутствующих в растворе ферментах эта информация может быть получена стандартными методами, а именно, применением ингибиторов и сведением к задаче определения каталитической активности фермента в чистом растворе (то есть в растворе, содержащем только один активный фермент). Однако интерес представляет и постановка, в которой априорных данных о присутствующих в сложном растворе ферментах нет.

Математическая часть этой задачи может быть сформулирована следующим образом. Имеется реализация функции $v(S)$, сопоставляющая каждому неотрицательному значению S соответствующее ему значение функции v . При этом известно, что аналитически функция $v(S)$ представима в виде

$$v(S) = \sum_{n=1}^N \frac{v_{max}^{(n)} S}{K_M^{(n)} + S}$$

(K_M^n — попарно различные положительные константы, представляющие собой константы Михаэлиса входящих в раствор ферментов, v_{max}^n — положительные константы, характеризующие максимальную скорость реакции для чистого раствора отдельного фермента), но параметры N , $K_M^{(n)}$ и $v_{max}^{(n)}$ неизвестны. Химически это соответствует мо-

дели Михаэлиса — Ментен (см., например, [1], [2]) и отсутствию взаимного влияния ферментов. Задача заключается в восстановлении по реализации функции $v(S)$ параметров N , $K_M^{(n)}$ и $v_{max}^{(n)}$.

Предлагаемый метод решения этой задачи состоит из трех этапов: интегральное преобразование, комплексно-аналитическое продолжение и разложение по ортогональной системе в пространстве почти периодических функций.

Перед перечисленными содержательными шагами осуществляется переход от функции $v(S)$ к реально исследуемой функции $f(t)$. Этот переход состоит в вычитании из $v(S)$ величины $v_\infty = \lim_{S \rightarrow \infty} v(S)$, умножении результата на минус единицу и переходе от переменной S к переменной $t = \sqrt{S}$. В результате функция $f(t)$ имеет вид

$$\sum_{n=1}^N \frac{v_{max}^{(n)} K_M^{(n)}}{K_M^{(n)} + t^2}.$$

Применение к функции $f(t)$ интегрального преобразования Фурье

$$F(z) = \int_{-\infty}^{+\infty} f(t) e^{-itz} dt$$

(см., например, [3, Гл. VIII, § 4,5]), с учетом равенства

$$\int_{-\infty}^{+\infty} \frac{e^{-i\lambda x}}{x^2 + a^2} dx = \frac{\pi}{a} e^{-a|\lambda|}$$

(см., например, [3, Гл. VIII, § 4, п. 1, пример 3]), переводит функцию $f(t)$ в функцию

$$F(z) = \sum_{n=1}^N v_{max}^{(n)} \gamma_n \pi e^{-\gamma_n z},$$

где $\gamma_n = \sqrt{K_M^{(n)}}$.

Следует отметить, что вместо стандартного (экспоненциального) преобразования Фурье возможно (без изменения дальнейшей сути метода) и применение косинус-преобразования.

Далее осуществляется комплексно-аналитическое продолжение функции $F(z)$ с луча $[0; +\infty)$ действительной прямой на квадрант $\operatorname{Re} z \geq 0, \operatorname{Im} z \geq 0$ комплексной плоскости. Такое продолжение осуществимо в силу аналитичности функций $e^{-\gamma_n z}$ во всей комплексной плоскости. Комплексно-аналитическое продолжение может быть реализовано на основе стандартных конечно-разностных схем и определения комплексной производной или же эквивалентных комплексной дифференцируемости условий Коши — Римана: оценки производных могут осуществляться рассмотрением горизонтальных приращений, а затем, на основе рассмотрения вертикальных приращений, могут быть восстановлены значения функции на следующем уровне сетки.

После комплексно-аналитического продолжения осуществляется переход к функции действительного переменного $L(x)$: $L(x) = \operatorname{Re} F(ix)$ ($x \geq 0$). Аналитически функция $L(x)$ представляет собой сумму

$$\sum_{n=1}^N v_{max}^{(n)} \gamma_n \pi \cos(\gamma_n x).$$

Наконец, для оценки характеристик присутствующих в растворе ферментов используются базовые факты теории почти периодических функций (см., например, [4], [5, Доп.: Почти периодические функции]).

На пространстве почти почти периодических функций определено скалярное произведение

$$(f, g) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x) \overline{g(x)} dx,$$

и система $\{\cos(\lambda x)\}_{\lambda > 0}$ является ортогональной относительно этого скалярного произведения.

Соответственно, для оценки количества ферментов и их характеристик функция $L(x)$ разлагается по системе $\{\cos(\lambda x)\}_{\lambda > 0}$ или, иными словами, строится функция $C(\gamma) = (L(x), \cos(\gamma x))$ ($\gamma > 0$). Функция $C(\gamma)$ отлична от нуля лишь в точках γ_n ($n = 1, 2, \dots, N$), причем

$$C(\gamma_n) = \frac{v_{max}^{(n)} \gamma_n \pi}{2}.$$

Таким образом, число слагаемых (число ферментов) N совпадает с числом точек, в которых функция $C(\gamma)$ совершает скачек (принимает ненулевое значение), и каждая такая точка позволяет восстановить значение $K_M^{(n)}$ ($K_M^{(n)} = \gamma_n^2$) и значение $v_{max}^{(n)}$ ($v_{max}^{(n)} = \frac{2C(\gamma_n)}{\pi\gamma_n}$).

Так как в приложении перебор континуального множества (проведение вычислений для всех действительных γ) неосуществим, реальные вычисления осуществляются лишь для сетки аргументов. В этом случае по данным о порядках максимально и минимально возможного значения K_M для присутствующих в растворе ферментов и данным о минимально возможном различии между значениями K_M для различных ферментов определяется мелкость сетки, а также параметр T_{max} , используемый при оценке скалярных произведений в качестве значения для T (а переход к пределу при $T \rightarrow \infty$ не осуществляется). Тогда при значениях γ , близких к γ_n , у функции $C(\gamma)$ будет наблюдаться «горб» с высотой порядка

$$\frac{v_{max}^{(n)} \gamma_n \pi}{2},$$

а при значениях γ , далеких от всех γ_n , функция $C(\gamma)$ будет принимать малые значения (то есть будет наблюдаться шум сравнительно малой амплитуды). По этим данным (количество «горбов», их центры и высоты) и осуществляется оценка значений N , $v_{max}^{(n)}$ и $K_M^{(n)}$.

Следует отметить, что для определения каталитической активности ферментов вместо интегрального преобразования Фурье в качестве интегрального преобразования может использоваться обратное преобразование Лапласа (см., например, [3, Гл. VIII, § 6]). В этом случае интегральное преобразование осуществляется после комплексно-аналитического продолжения.

Работа выполнена в рамках Проекта по развитию пост-геномных исследований и технологий «Разработка и реализация математических методов для определения каталитической активности ферментов в сложных биологических растворах» (в соответствии с приказом ректора МГУ имени М. В. Ломоносова № 484 от 25 мая 2011 года).

Список литературы

- [1] Яковлев В. А. Кинетика ферментативного катализа. — М., 1965.
- [2] Уэбб Л. Ингибиторы ферментов и метаболизма. — М., 1966.
- [3] Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. — М., 1976.
- [4] Левитан Б. М. Почти-периодические функции. — М., 1953.
- [5] Демидович Б. П. Лекции по математической теории устойчивости. — М., 1967.

Получение, обработка и воспроизведение медицинской тактильной информации

Садовничий В. А., Соколов М. Э., Бармин В. В.,
Буданов В. М., Галатенко А. В., Галатенко В. В.,
Коршунов А. А., Козорезов Ю. Ю., Подольский В. Е.
(Москва, МГУ им. М. В. Ломоносова)

vgalat@imscs.msu.ru

Наряду со зрением и слухом тактильные ощущения являются для человека одним из основных источников информации о внешнем мире. Тактильная информация существенно отличается от информации, получаемой посредством зрения и слуха. Связано это, в частности, с разным характером ощущений (тактильные ощущения являются контактными, а не дистантными).

К настоящему времени для визуальной информации (изображения, видео-поток без звука) и аудио-информации, а также для объединения этих видов информации (видео-поток со звуком) разработана и активно используется формальная база, включающая четкие стандарты и позволяющая строго говорить об обработке, хранении, воспроизведении и анализе такого рода данных и в аналоговом, и в цифровом форматах. В этой связи ограничимся лишь упоминанием цифровых стандартов JPEG/JPEG 2000 (см., например, [1]), MP3 (MP3-1/2 Audio Layer 3, см., например, [2], [3]) и MPEG-2 (см., например, [4]). Для тактильной же информации до последнего времени не было разработано ни формальной модели, ни устройств, обеспечивающих получение и воспроизведение такого рода информации.

Однако в последние годы был разработан медицинский тактильный эндохирургический комплекс, обеспечивший возможность получения, хранения, воспроизведения и анализа медицинской тактильной информации. Основными составляющими комплекса являются тактильный механорецептор (см., например, [5]), тактильный дисплей и управляющая станция.

Тактильный механорецептор представляет собой круговой цилиндр с диаметром основания 10–20 мм, на одном из оснований которого расположена рабочая головка — электронная «таблетка», со-

держающая датчики давления и обеспечивающая регулярное снятие показателей датчиков (в текущей реализации — 100 раз в секунду) и передачу этих показателей управляющей станции. Дополнительно механорецептор может быть оборудован системой, позволяющей точно определять текущее положение и ориентацию этого прибора. Мобильная версия тактильного механорецептора предназначена для получения информации в ходе эндоскопических операций. Стационарная версия предназначена для исследования удаленных тканей. В настоящее время механорецептор успешно проходит клинические испытания.

Тактильный дисплей представляет собой устройство, позволяющее воспроизводить тактильную информацию, в частности, информацию, полученную от тактильного механорецептора. Основной частью тактильного дисплея является экран, состоящий из рабочих элементов — пикселей. Каждый рабочий элемент (независимо от других) может изменять свое текущее положение и текущую жесткость, а также получать и передавать информацию о силе воздействия (текущем давлении) на этот элемент.

Управляющая станция обеспечивает получение информации от тактильного механорецептора и тактильного дисплея, передачу команд на тактильный дисплей, хранение, обработку и анализ тактильной информации.

Медицинский тактильный эндохирургический комплекс позволяет решить ряд важных медицинских и смежных задач. Перечислим некоторые из них.

Объективизация тактильных данных. Переход от субъективного описания тактильных характеристик ткани («мягкая», «умеренно жесткая» и т. п.) к объективным данным обеспечивает возможность сопоставления результатов различных тактильных исследований, в частности, дает возможность объективно оценивать динамику изменений.

Получение тактильной информации при проведении эндоскопических операций. При проведении эндоскопических операций визуальной информации может оказаться недостаточно, а получение тактильной информации стандартным образом в данном случае невозможно. Особенно это актуально в грудной хирургии, так как

при выполнении операции легкое спадается. В этих условиях различить патологический очаг, расположенный в глубине легочной ткани, например, опухолевый узел, бывает практически невозможно из-за его смещения.

Кроме того, исследование тканей тактильным механорецептором в ходе операции значительно превосходит по скорости срочное гистологическое исследование. Последнее, к тому же, требует круглосуточной работы специальной лаборатории, что по организационным причинам оказывается неосуществимым во многих стационарах в ходе экстренных операций, проводимых в ночное время. Такая ситуация часто приводит к ошибкам в постановке диагноза и тактическим промахам, в то время как использование тактильного механорецептора делает возможным проведение исследования членом операционной бригады без привлечения других специалистов.

Таким образом, применение тактильного механорецептора интраоперационно является эффективным способом получения информации, необходимой для правильной диагностики и выработки плана дальнейшего ведения операции.

Телемедицина. Медицинский тактильный комплекс позволяет осуществлять передачу и воспроизведение тактильной информации по локальным и глобальным сетям в реальном масштабе времени, обеспечивая тем самым возможность удаленного консультирования во время проведения операций.

Обучение. Использование тактильного дисплея позволяет показывать студентам-медикам свойства тканей с различными патологиями в условиях отсутствия реальных образцов таких тканей.

Автоматизированный анализ. Интраоперационный автоматизированный анализ тактильных данных позволяет в реальном масштабе времени диагностировать наличие неоднородностей, выявлять ткани с заданными тактильными характеристиками. Также воспроизведение на тактильном дисплее специальным образом обработанной тактильной информации (например, информации с усиленной тактильной контрастностью) позволяет увеличить качество диагностики в сложных случаях.

Таким образом, медицинский тактильный комплекс предоставляет новые для медицины возможности, эффективно дополняя существующее оборудование.

Следует отметить, что в дальнейшем представляется естественным применять тактильный комплекс не только в медицине, но и в других приложениях. Также важной представляется естественная задача построения сравнительно универсальной модели тактильной информации, позволяющей говорить об обработке и анализе тактильных данных и при этом не ограниченной спецификой текущей реализации медицинского тактильного эндохирургического комплекса.

Работа выполнена в рамках комплексного проекта 2010–218–01–345 Министерства образования и науки РФ «Организация производства медицинских и биологических устройств с тактильными возможностями».

Список литературы

- [1] Taubman D. S., Marcellin M. W. JPEG2000: image compression fundamentals, standards, and practice. — Norwell, Massachusetts: Kluwer Academic Publishers, 2004.
- [2] Brandenburg K.-H., Stoll G. et al. ISO-MPEG-1 Audio: A Generic Standard for Coding of High Quality Digital Audio // JAES. — 1994. 42 (10). — P. 780–792.
- [3] Information Technology: Generic coding of Moving pictures and associated audio — Audio Part — International standard ISO/IEC 13818-3. — 1995.
- [4] Watkinson J. MPEG 2. — Oxford: Focal Press, 1999.
- [5] Садовничий В. А., Буданов В. М., Соколов М. Э. и др. Математические задачи и методы в тактильной диагностике. — М.: МАКС Пресс, 2008.

Математические методы автоматизации тактильной диагностики

Садовничий В. А., Соколов М. Э., Ветров Д. П.,
Галатенко А. В., Галатенко В. В., Зыкова Т. В.,
Лебедев А. Е., Лукашенко Т. П., Подольский В. Е.,
Политов А. В.

(Москва, МГУ им. М. В. Ломоносова)

vgalat@imscs.msu.ru

Тактильные ощущения являются важным источником информации для человека. В ряде случаев, в том числе в медицинской диагностике, именно тактильные данные являются основой, на которой строится оценка ситуации, выработка гипотез и принятие решений.

В то же время, до последнего времени, в отличие от визуальных образов, аудио- и видео-данных, тактильные данные носили исключительно субъективный характер, и модель, позволяющая строго говорить об обработке и анализе тактильных данных, отсутствовала. Однако после появления специального устройства (тактильного механорецептора), используемого в хирургии, в том числе, при проведении эндоскопических операций, появилась возможность говорить об объективных, численно измеряемых тактильных данных: результатом исследования, основанного на использовании тактильного механорецептора, является поле давлений $P(x_n, t_m)$, снимаемое с датчиков x_n рабочей головки механорецептора в процессе исследования. В результате возникла задача об анализе этих данных, в первую очередь, задача автоматизации медицинской диагностики.

Неформально задача автоматизации диагностики может быть сформулирована следующим образом. Известно, какая именно ткань обследуется тактильным механорецептором и какие диагнозы возможны. Требуется на основе результатов тактильного исследования из множества возможных диагнозов выбрать наиболее вероятный (или наиболее вероятные). При этом можно считать, что для каждого возможного диагноза имеется, во-первых, экспертное описание тактильных характеристик ткани, свойственных этому заболеванию, и, во-вторых, набор образцов тканей с данным диагнозом.

Описанная задача может быть формализована следующим образом. Имеется банк данных с информацией о тактильных свойствах тканей. Более конкретно, имеется совокупность диагнозов, и для каждого варианта диагноза имеется экспертное описание свойств тканей с таким диагнозом, а также совокупность результатов обследований образцов тканей с таким диагнозом. Имеется результат обследования исследуемой ткани. Для каждого диагноза, данные о котором есть в банке данных, требуется указать одно или несколько чисел, характеризующих достоверность (или недостоверность) данного диагноза применительно к исследуемой ткани. В результате диагностики упорядочиваются по убыванию достоверности, и можно увидеть, является ли какой-либо диагноз бесспорным, есть ли среди диагнозов возможные, или же ни один диагноз из банка данных не применим к исследуемой ткани.

Для эффективного применения в клинической практике метод автоматизации диагностики должен удовлетворять следующим требованиям: результаты диагностики должны быть формально обоснованы; метод должен иметь низкую вычислительную сложность, позволяющую применять его в реальном масштабе времени; метод должен быть устойчив к изменениям входных данных, вызванным технологическими погрешностями; желательна возможность сочетания экспертных знаний и экспериментальных данных (то есть результатов обследования образцов с известным диагнозом); желательно сопровождение результатов показателями достоверности.

Эти требования делают невозможным использование для автоматизации диагностики ряда стандартных методов решения задач классификации. Так, применение экспертных систем (см., например, [1], [2]) не дает возможности эффективно использовать экспериментальные данные, а также требует регулярного обновления банка правил в «ручном» режиме. Применение нейронных сетей (см., например, [3], [4]), наоборот, затрудняет использование экспертных данных, а также не удовлетворяет требованию о формальной обоснованности вывода. Классификация на основе тестов (см., например, [5]) также приводит к невозможности эффективного сочетания экспертных знаний и экспериментальных данных: при автоматическом создании набора тестов не могут быть эффективно использованы экспертные зна-

ния (а также нарушаются требования о формальной обоснованности вывода), а при неавтоматическом создании набора тестов возникает сложность учета экспериментальных данных (и возникает необходимость регулярного обновления набора тестов в «ручном» режиме).

В рассматриваемой ситуации оптимальным подходом к решению задачи автоматизации диагностики представляется методика, сочетающая в себе идеи сразу нескольких классических подходов, в первую очередь, классификации, основанной на кластерном анализе (см., например, [6], [7]), а также теории вероятностей и математической статистики (вероятностные модели).

Основные шаги разработанного метода автоматизации диагностики заключаются в следующем.

Понижение размерности. Осуществляется переход от непосредственных результатов тактильного исследования к совокупности из сравнительно небольшого числа характеристик, каждая из которых представляет собой функцию от результатов исследования. На характеристики налагаются требования корректной зависимости от результатов исследования и наличия естественной трактовки, понятной медикам и другим специалистам. Именно второе требование делает предлагаемый способ понижения размерности в рассматриваемой ситуации более привлекательным по сравнению с широко используемым методом главных компонент (см., например, [7]).

Предобработка банка данных. Результаты обследований банковских образцов ткани с известным диагнозом рассматриваются как точки пространства характеристик. Набор полученных точек пополняется точками, построенными на основе экспертных знаний (это осуществимо благодаря наличию и характеристик естественных трактовок). Для каждого диагноза на основе полученного множества точек автоматически строится функция (функции), сопоставляющая точкам пространства характеристик оценку вероятности того, что ткань, исследование которой порождает эту точку, соответствует данному диагнозу. Построение этой функции основано на определении для каждого диагноза метрики в пространстве характеристик, учитывающей особенности этого диагноза. Допускается возможность «ручной» коррекции построенной функции.

Предсказание диагноза. Осуществляется переход от непосредственных результатов обследования ткани с неизвестным диагнозом к точке пространства характеристик, и для этой точки (и для каждого возможного диагноза) вычисляются функции, характеризующие достоверность диагноза, после чего диагнозы упорядочиваются по убыванию достоверности. Помимо основных показателей достоверности того, что рассматриваемый диагноз применим или не применим к ткани, для каждой из используемых характеристик вычисляется показатель согласованности с типичной для этого диагноза картиной. Совокупность таких показателей согласованности также представляет собой полезную информацию, показывающую, какие характеристики подтверждают диагноз, а какие опровергают.

Разработанный метод позволяет при настройке системы эффективно сочетать экспертные знания и экспериментальные данные; дает возможность естественной интерпретации результатов диагностики, а также обеспечивает обоснованность принятия решения; позволяет корректно обрабатывать ситуации, когда к изучаемому образцу ткани не применим ни один из диагнозов банка данных, а также ситуации, когда к изучаемому образцу применимы сразу несколько диагнозов; обладает низкой вычислительной сложностью и высокой устойчивостью. На эвристическом уровне можно также отметить, что разработанный метод автоматизации диагностики в некотором смысле близок к тому, что на самом деле делает врач.

Апробация на модельном банке данных показала эффективность разработанного метода автоматизации медицинской диагностики, осуществляемой на основе данных тактильного исследования.

Работа выполнена в рамках комплексного проекта 2010–218–01–345 Министерства образования и науки РФ «Организация производства медицинских и биологических устройств с тактильными возможностями».

Список литературы

- [1] Джексон П. Введение в экспертные системы. 3-е издание. — М.: Изд. дом «Вильямс», 2001.

- [2] Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование. — М.: Изд. дом «Вильямс», 2007.
- [3] Уоссерман Ф. Нейрокомпьютерная техника. Теория и практика. — М.: Мир, 1992.
- [4] Псиола В. В. Обзор основных нейросетевых моделей // Интеллектуальные системы. — 1999. Т. 4. Вып. 3–4.
- [5] Кудрявцев В. Б., Андреев А. Е., Гасанов Э. Э. Теория тестового распознавания. — М.: Физматлит, 2007.
- [6] Мандель И. Д. Кластерный анализ. — М.: Финансы и статистика, 1988.
- [7] Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика. Классификация и снижение размерности. — М.: Финансы и статистика, 1989.

Алгоритмы упорядочивания переменных в локальном элиминационном алгоритме

Свириденко А. В., Щербина О. А.

(Симферополь, Таврический национальный университет
им. В. И. Вернадского)

oleks.sviridenko@gmail.com, oshcherbina@gmail.com

Для решения разреженных задач дискретной оптимизации (ДО) путем вычисления *глобальной информации* с помощью *локальных вычислений* на основе анализа окрестностей элементов задачи [1] может быть использован класс локальных элиминационных алгоритмов (ЛЭА) [2].

Анализ публикаций, посвященных этой проблеме, позволяет сделать вывод, что в настоящее время экспериментальное исследование поведения алгоритма несериального динамического программирования (НСДП) [3, 4], являющегося локальным алгоритмом элиминации переменных, не проводилось.

НСДП последовательно элиминирует переменные в порядке, задаваемом упорядочением переменных, которое существенно влияет на время счета. На практике для нахождения элиминационной последовательности переменных используются всевозможные эвристики (такие как метод минимальной степени, рекурсивного разбиения и др.).

Нерешенным является вопрос экспериментального исследования поведения алгоритма НСДП в зависимости от использования алгоритма упорядочивания переменных.

Целью настоящей работы является анализ влияния каждого из пяти алгоритмов упорядочивания переменных, оказываемое ими на время решения разреженной задачи ДО с помощью алгоритма НСДП.

Рассмотрим разреженную задачу ДО, структура которой задается графом взаимосвязей переменных [3, 4]. В графе взаимосвязей задачи ДО вершины соответствуют переменным, причем две вершины соединяются ребром, если соответствующие переменные имеются

в одном и том же ограничении или/и в одном компоненте целевой функции.

Алгоритм НСДП может быть кратко записан так:

- 1) Перенумеровать переменные согласно порядку α (в данном случае $\alpha : \{x_1, \dots, x_n\}$).
- 2) Для $i = 1, \dots, n$ исключить x_i , изменяя граф взаимосвязей путем добавления ребер для превращения окрестности исключаемой вершины в клику.

Алгоритм минимальной степени MD

В алгоритме минимальной степени (minimum degree (MD)) выбирается вершина v графа G с минимальной степенью. Далее строится граф G' , получаемый путем создания клики из вершины v и ее соседей с последующим удалением v и инцидентных ей ребер. Рекурсивно из G' с помощью эвристики создается *хордальный* суперграф H' . И, наконец, получают связный суперграф H из G , путем добавления v и инцидентных ей ребер из G к H' . Будучи алгоритмом локальной минимизации, алгоритм MD не всегда дает упорядочение с минимальным пополнением для графа в целом. Для получения упорядочения при помощи алгоритма минимальной степени MD в данной работе была выбрана функция `minimum_degree_ordering()` из библиотеки BOOST [5].

Алгоритм рекурсивного разбиения ND

Алгоритм рекурсивного разбиения ND [6] находит сепаратор, то есть множество вершин S , разделяющее граф на две части A и B , помещая S в упорядочении последним. Алгоритм применяется рекуррентно к частям графа A и B , пока их размеры не станут меньше, чем некоторое пороговое значение. Для вычисления упорядочения с помощью алгоритма рекурсивного разбиения ND в настоящей работе использовалась функция `METIS_EdgeND()` библиотеки Metis [7].

Алгоритм MCS

В [8] был предложен алгоритм MCS (Maximum Cardinality Search — поиск по максимальной степени). MCS на некотором графе G производит за время $O(n + m)$ полное упорядочение множества

вершин следующим образом. Из произвольной вершины, выбирается любая, еще не пронумерованная вершина, смежная максимальному числу уже пронумерованных вершин. MCS упорядочение было получено при помощи пакета CHOMPACK [9].

Алгоритм Minimum Fill-in

Эвристика минимального пополнения (Minimum Fill-in) [10] работает практически так же, как и описанная выше эвристика минимальной степени MD, с той лишь разницей, что здесь на каждом шаге выбирается такая вершина, чтобы число добавляемых к ней ребер, необходимых для получения клики, было минимальным.

Алгоритм Lex-BFS

Rose, Tarjan, и Lueker [11] предложили алгоритм, вычисляющий хорошую элиминационную последовательность за линейное время, который был назван *лексикографическим поиском в ширину* (Lex-BFS). Суть данного метода заключается в следующем. Вершины нумеруются от n до 1 (нумерация фиксирует позиции переменных) в упорядочении. Далее, для каждой вершины создается метка, содержащая множество чисел, записанных по убыванию. Таким образом вершины могут быть лексикографически упорядочены согласно их меткам.

Описание вычислительного эксперимента

Алгоритм НСДП реализован на языке программирования Python, который был выбран в связи с наличием в нем средств для быстрого построения и описания задач ДО, а также ввиду обширного количества имеющихся библиотек для работы с графами и алгоритмами упорядочивания.

Описание множества тестовых задач

Тестовые задачи ДО генерировались на основе уже существующих гиперграфов из библиотеки задач удовлетворения ограничений [12]. Для построения ограничения i бралось очередное гиперребро гиперграфа из библиотеки [12], содержащее множество переменных X_{S_i} , входящих в строящееся ограничение. Далее с помощью

процедуры, использующей датчик случайных чисел, строились коэффициенты A_{S_i} при соответствующих переменных, тогда левая часть i -го ограничения имела вид $A_{S_i} X_{S_i}$. Правая часть i -го ограничения имеет вид $\sigma \sum A_{S_i}$, где σ — случайное число из интервала $(0, 1)$. Коэффициенты c_j целевой функции $\sum_{j=1}^n c_j x_j \rightarrow \max$ также строились с помощью процедуры, использующей датчик случайных чисел. Далее для каждой тестовой задачи ДО строился соответствующий граф взаимосвязей. Для каждого графа взаимосвязей применялись алгоритмы упорядочивания MD, ND, MCS, MIN-FILL и LEX-BFS, после чего задачи ДО решались с помощью алгоритма НСДП с соответствующими найденными упорядочениями.

Анализ результатов вычислительного эксперимента

Для проведения вычислительного эксперимента были взяты пять групп тестовых задач: 'dubois', 'bridge', 'adder', 'pret' и 'NewSystem' из библиотеки [7], содержащие 33 тестовые задачи. Все вычисления были проведены на базе процессора Intel Core 2 Duo @ 2.66 GHz, 2 GB ОЗУ и операционной системы Linux, версия ядра 2.6.35-24-generic. Согласно результатам вычислительного эксперимента, на множестве тестовых задач для алгоритма ND минимальное время работы алгоритма НСДП было достигнуто 0 раз (0%), для MD — 2 раза (6,0%), LEX-BFS — 3 раза (9,1%), MCS — 9 раз (27,3%) и MIN-FILL — 19 раз (57,6%).

Список литературы

- [1] Журавлев Ю. И. Избранные научные труды. — М.: Магистр, 1998.
- [2] Щербина О. А. Локальные элиминационные алгоритмы решения разреженных дискретных задач // Журнал вычислительной математики и математической физики. — 2008. Т. 48. № 1. — С. 161–177.
- [3] Bertele U., Brioschi F. Nonserial Dynamic Programming. — New York: Academic Press, 1972.

- [4] Щербина О. А. О несериальной модификации локального алгоритма декомпозиции задач дискретной оптимизации // Динамические системы. — 2005. Вып. 19. — С. 179–190.
- [5] BOOST, minimum degree ordering algorithm. — URL: <http://www.boost.org>
- [6] George J. A. Nested dissection of a regular finite element mesh // SIAM J. Numer. Anal. — 1973. V. 10. — P. 345–367.
- [7] Karypis G., Kumar V. MeTiS — a software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices. Version 4. — University of Minnesota, 1998.
- [8] Tarjan R. E., Yannakakis M. Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs // SIAM J. Comput. — 1984. V. 13. — P. 566–579.
- [9] CHOMPACT, maximum cardinality ordering algorithm. — URL: <http://abel.ee.ucla.edu/chompact/>
- [10] Jégou P., Ndiaye S. N., Terrioux C. Computing and exploiting tree-decompositions for (Max-)CSP // Proceedings of the 11th International Conference on Principles and Practice of Constraint Programming (CP-2005). — 2005. — P. 777–781.
- [11] Rose D., Tarjan R., Lueker G. Algorithmic aspects of vertex elimination on graphs // SIAM J. on Computing. — 1976. V. 5. — P. 266–283.
- [12] Musliu N., Samer M., Ganzow T., Gottlob G. A csp hypergraph library. Technical Report. DBAI-TR-2005-50. — Technische Universität Wien, 2005.

Достаточное условие эффективной разрешимости задачи Open shop в терминах суммарной нагрузки

Севастьянов С. В., Черных И. Д. (Новосибирск, ИМ СО РАН)

seva@math.nsc.ru, idchern@math.nsc.ru

Классическая задача теории расписаний open shop формулируется следующим образом. Множество работ $\mathcal{J} = \{J_1, \dots, J_n\}$ должно быть выполнено на машинах из $\mathcal{M} = \{M_1, \dots, M_m\}$. Выполнение операции O_{ji} работы J_j машиной M_i занимает время p_{ji} . Никакие две операции одной работы или на одной машине не могут выполняться одновременно. Требуется построить допустимое расписание минимальной длины, которая определяется как время завершения наиболее поздней операции (C_{\max}). Задача полиномиально разрешима в случае $m = 2$ и является NP-трудной при $m \geq 3$ [1]. При нефиксированном числе машин (когда m является частью входа) задача NP-трудна в сильном смысле [2].

Обозначим множество входов m -машинной задачи open shop через \mathcal{I}_m . Для каждого $I \in \mathcal{I}_m$ будем использовать следующие обозначения:

- $\ell_i(I) \doteq \sum_j p_{ji}$ — нагрузка машины M_i ;
- $d_j(I) \doteq \sum_i p_{ji}$ — длина работы J_j ;
- $\ell_{\max}(I) \doteq \max_i \ell_i(I)$, $d_{\max}(I) \doteq \max_j d_j(I)$ — наибольшая машинная нагрузка и наибольшая длина работы, соответственно;
- $\Delta(I) \doteq \sum_{i,j} p_{ji} = \sum_i \ell_i(I) = \sum_j d_j(I)$ — суммарная нагрузка;
- $\bar{C}(I) \doteq \max\{\ell_{\max}(I), d_{\max}(I)\}$ — нижняя оценка оптимума для входа I ;
- через $\mathcal{J}(I)$ и $\mathcal{M}(I)$ будем обозначать множества работ и машин во входе I .

Обозначение входа I будем опускать в случаях, когда это не вносит разночтений.

Для входа I допустимое расписание длины $\bar{C}(I)$ будем называть *нормальным*; также *нормальным* будем называть вход I , для которого существует нормальное расписание. Подмножество входов \mathcal{K} , состоящее только из нормальных входов, будем называть *нормальным классом*. Если для некоторого нормального класса \mathcal{K} существует

полиномиальный алгоритм построения нормального расписания для любого из его входов, то такой класс будем называть *эффективно-нормальным*.

В статье [3] рассматриваются достаточные условия нормальности входов, описываемые в терминах неравномерности нагрузок машин, и описываются соответствующие (достаточно широкие) эффективно-нормальные классы. Целью данной работы является выявление эффективно-нормальных классов, описываемых в терминах ограничений на суммарную нагрузку Δ .

Рассмотрим класс входов $\mathcal{K}_m(\alpha) \doteq \{I \in \mathcal{I}_m \mid \Delta(I) \leq \alpha \bar{C}(I)\}$. Для каждого $m \geq 3$ исследуем, при каком наибольшем значении α класс $\mathcal{K}_m(\alpha)$ является нормальным, то есть найдем наиболее широкие нормальные классы вида $\mathcal{K}_m(\alpha)$. Заметим, что для случая $m = 2$ такая постановка вопроса неинтересна, поскольку множество всех входов \mathcal{I}_2 является эффективно-нормальным классом [1].

В дальнейшем будем пользоваться следующими преобразованиями входов.

Определение. Будем говорить, что вход I' получен из входа I с помощью *склеивания множества машин \bar{M}* , если $\mathcal{J}(I') = \mathcal{J}(I)$ и $\mathcal{M}(I') = \mathcal{M}(I) \setminus \bar{M} \cup \{M_{i'}\}$, где $p_{ji'} = \sum_{M_i \in \bar{M}} p_{ji}$.

Заметим, что допустимое расписание для «склеенного» входа I' можно интерпретировать как допустимое расписание такой же длины для исходного входа I . Справедлива следующая

Лемма. Для любого $m \geq 3$ и $\alpha > 2$ класс $\mathcal{K}_m(\alpha)$ не является нормальным.

Доказательство. Рассмотрим следующий вход с четырьмя работами. Работа J_1 имеет три операции на машинах M_1 , M_2 и M_3 длительности $1/3$, операции на остальных машинах — нулевые. Каждая из работ J_k , $k = 2, 3, 4$, имеет единственную ненулевую операцию длительности $1/3 + \varepsilon$ на машине M_{k-1} . Для этого входа $\bar{C} = 1$, $\Delta = 2 + 3\varepsilon$. Докажем, что он не является нормальным. Предположим обратное, что для этого входа существует расписание длины \bar{C} . Поскольку $d_1 = \bar{C}$, операции работы J_1 выполняются в этом расписании одна за другой, без промежуточных ожиданий. При этом одна из них выполняется (на некоторой машине M_k) в интервале $[1/3, 2/3]$. Ясно

однако, что ненулевую операцию работы J_{k+1} невозможно выполнить на этой же машине ни в одном из оставшихся свободных интервалов $[0, 1/3], [2/3, 1]$, что означает невозможность выполнения совокупности работ в интервале $[0, \bar{C}]$. Полученное противоречие доказывает лемму.

Рассмотрим объединение классов $\mathcal{K}(\alpha) \doteq \bigcup_m \mathcal{K}_m(\alpha)$. Справедлива следующая

Теорема. $\mathcal{K}(2)$ является эффективно-нормальным классом.

Доказательство. Рассмотрим вход $I \in \mathcal{K}(2)$. В силу симметричности понятий машина/работа в системе open shop, без ограничения общности можем считать, что $\bar{C}(I) = \ell_1(I)$. Поскольку $\sum \ell_i(I) \leq 2\bar{C}(I)$, справедливо $\sum_{i=2}^m \ell_i(I) \leq \bar{C}(I)$. Проведем склеивание множества машин $\bar{M} = \{M_2, \dots, M_m\}$. Поскольку $\ell_1(I') = \ell_1(I) = \bar{C}(I)$, $\ell_2(I') \leq \bar{C}(I)$, а длины работ нового входа I' совпадают с длинами соответствующих работ из I , имеем $\bar{C}(I') = \bar{C}(I)$.

Поскольку $I' \in \mathcal{I}_2$, оптимальное расписание длины $C(I')$ может быть построено алгоритмом Гонзалеза-Сани [1] за время, линейное от числа работ. Это расписание можно интерпретировать как допустимое расписание длины $C(I)$ для исходного входа I .

Заметим, что в силу леммы и теоремы любой нормальный класс вида $\mathcal{K}_m(\alpha)$ является подмножеством $\mathcal{K}(2)$, то есть для любого $m \geq 3$ наиболее широким нормальным классом вида $\mathcal{K}_m(\alpha)$ является $\mathcal{K}_m(2)$.

Работа выполнена при финансовой поддержке РФФИ, грант № 08-01-00370.

Список литературы

- [1] Gonzalez T., Sahni S. Open shop scheduling to minimize finish time // J. Assoc. Comp. Math. — 1976. 23. — P. 665–679.
- [2] Williamson D. P. et al. Short shop schedules // Oper. Res. — 1997. V. 45. N. 2. — P. 288–294.
- [3] Kononov A., Sevastianov S., Tchernykh I. When difference in machine loads leads to efficient scheduling in open shops // Annals of Oper. Res. — 1999. 92. — P. 211–239.

Параллельное разложение Холецкого разреженной матрицы

Старостин Н. В., Сафонова Я. Ю.

(Нижний Новгород, Нижегородский государственный университет
им. Н. И. Лобачевского)

nvstar@mail.ru, safonova.yana@gmail.com

1. Введение

Системы линейных уравнений $Ax = b$ с симметричной положительно определенной матрицей A возникают при численном решении задач математической физики. Для решения систем такого рода применяют разложение Холецкого вида $A = U^T U$, где U — верхнетреугольная матрица. Тогда решение исходной системы эквивалентно решению двух систем: $U^T y = b, Ux = y$. Классический алгоритм нахождения разложения, предложенный в [1], последовательно формирует элементы новой строки, начиная с первой, и использует при этом значения в найденных строках. Такой подход не обладает параллелизмом, так как независимый расчет строк приведет к некорректному результату. Однако, если матрица A является разреженной, то некоторые строки могут не зависеть от значений предыдущих строк, и выделение независимых друг от друга групп позволяет организовать параллельный расчет. В [2] дано определение дерева исключения матрицы, как отражения ее структуры разреженности, определяющей зависимость между строками, и сформулированы теоремы, позволяющие сконструировать параллельный алгоритм нахождения разложения по известному дереву исключения. Задача построения дерева исключения совпадает по сложности с задачей определения портрета матрицы U [3], при этом вид полученного дерева может не допускать эффективного распараллеливания. В этом случае предлагается предварительно переупорядочивать матрицу A с помощью алгоритмов [1], в ходе выполнения которых дерево исключения матрицы становится известным и хорошо сбалансированным. К таким алгоритмам относят методы семейства вложенных сечений. В данной статье рассматривается параллельная схема общего вида для реше-

ния поставленной задачи, позволяющая достичь ускорения $\log_2 p$, где p — число независимых вычислителей.

2. Связь дерева исключения с переупорядочиванием матрицы

Пусть $G(A) = (V, E)$ — граф матрицы A . Деревом исключения называется дерево, чье множество вершин совпадает с V , а ребро (v_i, v_j) существует тогда и только тогда, когда $i = \min\{k : u_{jk} \neq 0 \ \& \ k > j\}$. Таким образом, v_i является родителем v_j , если первый недиагональный элемент строки j располагается в столбце i . Такое определение дает понятие зависимости между строками. Если вершина v_i является предком вершины v_j , то значения в строке i должны быть вычислены после значений в строке j , так как зависят от них. Ресурсом для получения ускорения является параллельное вычисление строк, соответствующие вершины которых не являются по отношению к друг другу предком или потомком. При этом, чем меньше высота дерева исключения при заданном количестве вершин, тем более высокое ускорение можно получить. Переупорядочивание матрицы методом вложенных сечений дает бинарное дерево исключения [4], причем настройка параметров алгоритма позволяет получать сбалансированность между поддеревьями одного уровня. Другим критерием при выборе алгоритма переупорядочивания является заполнение, возникающее в матрице U , так как при решении большеразмерных систем сохранение разреженности более важная задача, чем распараллеливание. В статье приведены результаты, демонстрирующие, что для типовых задач, возникающих на практике, метод вложенных сечений дает не худшее заполнение, чем другие известные алгоритмы [1].

3. Программная реализация

На основе приведенных рассуждений предлагается программная реализация, основанная на предварительном переупорядочивании исходной матрицы методом вложенных сечений. Итогом переупорядочивания является матрица перестановок и дерево исключения. В силу эквивалентности задач определения дерева исключения и символической факторизации будем считать, что портрет матрицы U является известным. Для дальнейшего параллельного выполнения необходи-

мо сделать разбиение дерева исключения. Один из способов разбиения — выделение поддеревьев одинаковой высоты. Высота поддеревьев — эвристический параметр, влияющий на балансировку между потоками, и его оптимальное значение зависит от конфигурации вычислительной системы. Для матриц с порядком от 50000 до 1000000 рекомендуется выбирать его в диапазоне от 50 до 100. Между поддеревьями необходимо установить отношение предок — потомок, и каждое поддерево связать с задачей. Под задачей понимается атомарный набор операций, который будет последовательно выполняться на одном вычислителе для всех вершин поддерева, ей соответствующего.

Пример задачи:

- 1) Для всех потомков текущего поддерева запустить параллельное выполнение соответствующих задач.
- 2) Синхронизировать окончание выполнения порожденных задач.
- 3) Для каждой вершины поддерева выполнить шаги вычислительной фазы.

Запуск параллельного алгоритма инициализирует задача, соответствующая корневному поддереву. Данная схема подходит для численной факторизации и решения системы $U^T y = b$ при обходе вершин от листьев поддерева к корню. Решение системы $Ux = y$ соответствует схеме с последовательностью шагов 3, 1, 2 и обратным порядком обхода вершин в поддереве для шага 3.

4. Результаты вычислительных экспериментов

В первой части экспериментов сравнивались показатели заполнения для разных алгоритмов переупорядочивания. Результаты показывают, что метод вложенных сечений дает сравнимое с другими алгоритмами заполнение.

Название	CM	RCM	King	Sloan	ND
shallow_water2	0,00656	0,00656	0,00739	0,00654	0,00063
parabolic_fem	0,00163	0,00163	0,00163	0,00163	0,00019
tmt_sym	0,00242	0,00236	0,00266	0,00225	0,00018
thermomech_dM	0,00181	0,0017	0,00178	0,00512	0,00045

где CM — прямой алгоритм Катхилл — Макки, RCM — обратный алгоритм Катхилл — Макки, King — алгоритм Кинга, Sloan — алгоритм Слоуна, ND — метод вложенных сечений.

Во второй части экспериментов тестировалась описанная выше схема распараллеливания. Следует отметить, что верхняя оценка для ускорения такой схемы на p вычислителях $\log_2 p$. Результаты решения СЛАУ на 8 потоках показывают ускорение, близкое к теоретически возможному:

Название	N	S_1	S_2	H	H^*
shallow_water2	81920	1,65	2,62	930	100
parabolic_fem	525825	1,84	2,83	2826	75
tmt_sym	726713	1,62	2,59	3932	100
thermomech_dM	204316	2,79	3,16	1293	50

где N — порядок матрицы, S_1 — ускорение численной факторизации, S_2 — ускорение при решении систем, H — высота дерева исключения, H^* — высота поддеревьев в разбиении.

Список литературы

- [1] Писсанецки С. Технология разреженных матриц. — М.: Мир, 1988. — С. 76–79, 115, 163.
- [2] Heggerness P. Minimizing fill-in size and elimination tree height in parallel Cholesky factorization. — Bergen: Department of Informatics University of Bergen, 1992. — P. 15–17.
- [3] Van Grondelle J. Symbolic Sparse Cholesky Factorization Using Elimination Trees. — Utrecht: Department of Mathematics Utrecht University, 1999. — P. 3–6.
- [4] Сафонова Я. Ю. Использование графовых моделей при распараллеливании метода Холецкого при решении разреженных симметричных СЛАУ // Материалы XVI международной конференции Проблемы теоретической кибернетики. — Нижний Новгород: Изд-во ННГУ, 2011. — С. 426–430.

Формализация задачи поиска объектов на векторной сцене

Фофанов В. Б., Жизневский А. Н.

(Казань, Казанский (Приволжский) Федеральный Университет)

viatcheslav.fofanov@ksu.ru, write2aracon@mail.ru

Введение

Поиску на сцене объектов определенного вида по ее изображению посвящено довольно большое количество публикаций. Предлагаемые в них методы являются, как правило, эвристическими. В настоящей работе излагаются формализация и решение этой задачи в рамках теоретико-вероятностного подхода. При этом предполагается, что объекты являются пятнами, а исходной информацией о сцене служит набор ее изображений. Поиск объектов предлагается проводить в три этапа. На первом определяются квадратные фрагменты сцены, названные зонами интереса. Каждая зона содержит один объект и его окружение (фон). На втором этапе (сегментации) проводится классификация пикселей зоны интереса на два класса. Пиксели, оказавшиеся в классе с именем объект, используются на третьем этапе для вычисления геометрических признаков объекта и принятия окончательного решения о его присутствии на сцене.

1. Модель сцены

Пусть (Ω, A, P) — вероятностное пространство, $Y = \{0, 1, \dots, n - 1\}$ — множество из $|Y| = n$ объектов и A — конечное подмножество на целочисленной решетке Z^2 . Назовем объектом сцены с проекцией A семейство $\xi_A = (\xi_a)_{a \in A}$ векторных случайных величин $\xi_a = (\xi_a^j)_{1 \leq j \leq \nu}$, определенных на (Ω, A, P) и принимающих значения в $\mathbf{Y} = Y^\nu$. Изображением объекта ξ_A будет называться семейство $\mathbf{x}_a = (\mathbf{x}_a)_{a \in A}$ из \mathbf{Y}^A . Объект считается заданным, если на множестве \mathbf{Y}^A его изображений задано распределение вероятностей $P_{\mathbf{Y}^A} = (p_{\mathbf{Y}^A}(\mathbf{x}_A))_{\mathbf{x}_A \in \mathbf{Y}^A}$. Совокупность объектов назовем векторной сценой, если их проекции попарно не пересекаются, а их сумма равна Z^2 . Таким образом, сцена является векторным случайным полем

$(\xi_z)_{z \in Z^2}$. В частном случае, когда случайные величины являются скалярными, сцена также будет называться скалярной. В качестве изображения $\mathbf{x}_z = (\mathbf{x}_z)_{z \in Z^2}$ сцены естественно рассматривать реализации (выборочные поверхности) случайного поля. Существование указанных сцен вытекает из следующей теоремы.

Теорема 1. Пусть на Z^2 задано разбиение, состоящее из конечных попарно непересекающихся подмножеств, называемых проекциями, и пусть каждой проекции $A \subset Z^2$ поставлено в соответствие распределение вероятностей $P_{\mathbf{Y}^A}$ на множестве \mathbf{Y}^A . Тогда существует вероятностное пространство (Ω, A, P) и векторная сцена $(\xi_z)_{z \in Z^2}$ такая, что

$$P\{\omega \in \Omega : \xi_A(\omega) = \mathbf{x}_A\} = p_{\mathbf{Y}^A}(\mathbf{x}_A)$$

для любой проекции A и любого $\mathbf{x}_A \in \mathbf{Y}^A$. Кроме того, если A и B — проекции разных объектов, то для любых $a \in A$ и $b \in B$ случайные величины ξ_a и ξ_b — независимы.

Далее предполагается, что каждый объект ξ_A является фрагментом однородного случайного поля с вектором средних значений $\mathbf{m}_A \in \mathbf{Y}$, для которого выполняются условия эргодической теоремы Слущкого. Это позволяет использовать среднее арифметическое значение $\bar{\mathbf{x}}_A = \frac{1}{|A|} \sum_{a \in A} \mathbf{x}_a$, вычисленное по изображению объекта \mathbf{x}_A в качестве статистической оценки для \mathbf{m}_A .

2. Поиск зон интереса

Пусть d — евклидово расстояние на Z^2 . Точку a из A назовем граничной, если $d(a, A^C) = 1$. Совокупность $Fr(A)$ граничных точек A будет называться его границей. Объект ξ_A , проекция которого является односвязным множеством, назовем пятном, если существует квадрат Q на Z^2 такой, что $A \subset (Q \setminus Fr(Q))$, если $E\xi_z = \mathbf{m}_{Q \setminus A}$, $z \in Q \setminus A$, и если $d(\mathbf{m}_A, \mathbf{m}_{Q \setminus A}) > 0$. Семейство $\xi_{Q \setminus A}$ будет называться окрестность пятна, а ξ_Q — его зоной интереса.

Пусть A — проекция пятна ξ_A , $d(A)$ — его диаметр, а $B(a, r)$ — окрестность (круг с центром a и радиусом r), принадлежащая A . Легко видеть, что граница квадрата $Q \subset Z^2$ с центром a и стороной l , $l \geq 2(d(A) - r + 1)$, не имеет общих точек с A . Если ξ_Q —

зона интереса для ξ_A , то изображение \mathbf{x}_{Fr} границы можно использовать для получения оценок среднего $\mathbf{m}_{Q \setminus A}$ фона. В самом деле, пусть $n = |B(a, r)|$, разделим границу $Fr(Q)$ на s непересекающихся связных частей Fr_j , $1 \leq j \leq s$, по n точек и вычислим s средних арифметических значений $\bar{\mathbf{x}}_j = \frac{1}{n} \sum_{z \in Fr_j} \mathbf{x}_z$, $1 \leq j \leq s$. С другой стороны, среднее арифметическое значение $\bar{\mathbf{x}}_a = \frac{1}{n} \sum_{z \in B(a, r)} \mathbf{x}_z$ будет оценкой среднего \mathbf{m}_A пятна. Сопоставим каждому квадрату Q со стороной l и центром $a \in Z^2$ признак $f_{\{a\}}$, определенный равенством

$$f_{\{a\}}(\mathbf{x}) = \sum_{j=1}^s I_{]0, +\infty[}(d(\mathbf{x}_a, \mathbf{x}_{Fr}) - d(\mathbf{x}_j, \mathbf{x}_{Fr})).$$

Если Θ_1 — множество зон интереса, а Θ_2 — квадраты из пикселей фона, то для сцен, полученных скользким суммированием по окрестности с радиусом \hat{r} , имеет место следующий результат.

Теорема 2. Пусть $P(\Theta_1)$ и $P(\Theta_2)$ — априорные вероятности, Fr_j , $1 \leq j \leq s$, — связные фрагменты границы $Fr(Q)$, содержащие по $n = |B(a, r)|$ пикселей, такие, что $d(Fr_i, Fr_j) > 2\hat{r}$, $f_{\{a\}}$ — признак квадрата ξ_Q , соответствующий r и s . Тогда семейство $h_n^s : Y_f \rightarrow \{1, 2\}$ решающих правил вида

$$h_n^s(f_{\{a\}}(\mathbf{x})) = \begin{cases} 1, & f_{\{a\}}(\mathbf{x}) = s \\ 2, & f_{\{a\}}(\mathbf{x}) < s \end{cases}$$

классификации квадратов из $\Theta_1 + \Theta_2$ стремится при $r \rightarrow +\infty$ и $s \rightarrow +\infty$ к байесовскому решающему правилу с вероятностью ошибки, равной нулю.

3. Сегментация

Пусть ξ_Q — зона интереса векторной сцены. Пусть Fr_j , $1 \leq j \leq s$, — семейство связных попарно не пересекающихся фрагментов из точек на $Fr(Q)$. По теореме Слуцкого, средние арифметические значения

$$\bar{\mathbf{x}}_{Fr} = \frac{1}{|Fr(Q)|} \sum_{t \in Fr(Q)} \mathbf{x}_t \text{ и } \bar{\mathbf{x}}_j = \frac{1}{|Fr_j|} \sum_{t \in Fr_j} \mathbf{x}_t, \quad 1 \leq j \leq s,$$

являются оценками координат вектора $\mathbf{m}_{Q \setminus A}$. Пусть $B(z, r)$ — окрестность из $Q \setminus Fr(Q)$, а $\bar{\mathbf{x}}_z = \frac{1}{|B(z, r)|} \sum_{t \in B(z, r)} \mathbf{x}_t$ — среднее арифметическое значение. Если $f_{\{z\}}(\mathbf{x}) = s$, то пиксель ξ_z будем относить к объекту ξ_A , а при $f_{\{z\}}(\mathbf{x}) < s$ — к фону. Присвоим номер 1 подмножеству пикселей, образующих объект, и номер 2 — подмножеству всех остальных пикселей зоны интереса. Тогда изложенный способ сегментации можно рассматривать как классификацию ее пикселей на два класса с использованием решающего правила

$$h_r^s(f_{\{z\}}(\mathbf{x})) = \begin{cases} 1, & f_{\{z\}}(\mathbf{x}) = s \\ 2, & f_{\{z\}}(\mathbf{x}) < s \end{cases}$$

Пусть A° и $(Q \setminus A \setminus Fr(Q))^\circ$ — внутренние пиксели объекта и фона соответственно. Для сцен, полученных скользящим суммированием по окрестности с радиусом \hat{r} , имеет место следующий результат.

Теорема 3. Пусть ξ_Q — зона интереса, $P(A^\circ)$ и $P((Q \setminus A \setminus Fr(Q))^\circ)$ — априорные вероятности классов, Fr_j , $1 \leq j \leq s$, — связанные фрагменты границы $Fr(Q)$, состоящие из $n = |B(a, r)|$ пикселей каждый, и такие, что $d(Fr_i, Fr_j) > 2\hat{r}$, $f_{\{z\}}$ — признак, соответствующий r и s . Тогда семейство h_r^s решающих правил стремится при $r \rightarrow +\infty$ и $s \rightarrow +\infty$ к байесовскому решающему правилу h с вероятностью ошибки $e(h) = 0$.

Математические модели экономических процессов

Черемных Ю. Н. (Москва, МГУ им. М. В. Ломоносова)

Содержательные области и проблематика экономико-математического моделирования исторически формировались постепенно: проблемы ценообразования в различных рыночных структурах (от чистой конкуренции до чистой монополии), рационального поведения потребителя на рынке, максимизации прибыли, спроса и предложения, относящиеся традиционно к микроэкономике. К макроэкономике относятся вопросы, связанные со статикой и динамикой инвестиций, процентной ставки, безработицы и валового национального продукта. Отдельно следует отметить проблему рыночного равновесия (на микро- и макроэкономическом уровнях). Сюда необходимо добавить вопросы структуризации данных и их преобразования в модельную информацию, анализа влияния отдельных факторов на поведения ключевых экономических показателей (предмет математической статистики и эконометрики), совокупность задач, связанных с операциями на финансовых рынках (то есть на рынках ценных бумаг и финансовых услуг).

Применяемые математические методы и средства для решения разнообразных экономических задач отличаются большим разнообразием, что связано с широтой круга содержательных экономических областей. В дополнительных комментариях не нуждается то обстоятельство, что экономическая материя много сложнее различных областей естественно-научной материи.

Теория систем линейных алгебраических уравнений — основной инструмент анализа статической межотраслевой модели, предложенной в 1930-е гг. американским экономистом В. В. Леонтьевым (Премия по экономике имени Нобеля, 1973). Статическая межотраслевая модель (модель затраты — выпуск) представляет собой развитие идеи межотраслевого описания национальной экономики Советской России, реализованной группой советских экономистов в середине 1920-х гг.

Метод неопределенных множителей, опубликованный французским математиком Ж. Л. Лагранжем на рубеже 18–19 вв., активно использовался и используется экономистами для решения разнооб-

разных задач — в основном микроэкономики: задачи рационального поведения потребителя на рынке, задача рационального распределения ограниченных ресурсов и задача издержек при фиксированном объеме выпускаемой продукции. В этих задачах множитель Ж. Л. Лагранжа имеет глубокий экономический смысл. В теории потребления он связывает величину дохода потребителя с уровнем полезности, в теории производства — лимит на ресурсы с объемом выпускаемой продукции.

Линейное программирование возникло под давлением экономической проблематики (проблема рационального использования ограниченных ресурсов) и было отмечено Премией по экономике имени Нобеля, которую получили советский математик Л. В. Канторович и американский математик Т. Купманс. На основе симплексного метода, предложенного американским математиком Дж. Данцигом, и на базе использования ЭВМ было решено большое число оптимизационных задач с экономическим и военным содержанием.

Линейное программирование долгое время представляло собой интенсивно развивающуюся область прикладной математики (нелинейное программирование).

Теория обыкновенных дифференцированных и разностных уравнений широко востребована экономистами для анализа разнообразных эволюционных процессов, в частности, динамики рыночных цен и ее качественного анализа с использованием теории устойчивости А. М. Ляпунова.

Положения теории оптимального управления находят активное применение в решении и анализе многих задач макроэкономики.

Для решения широких классов разнообразных экономических задач интенсивно применяются разделы теории вероятностей и математической статистики.

Следует отметить, что примерно до середины 20 в. уровень применения математики в экономических исследованиях заметно отставал от аналогичного уровня применения математики в естественнонаучных областях. После публикации в 1954 г. журналом «Эконометрика» статьи «Существование равновесия в конкурентной экономике» американского экономиста К. Эрроу (лауреат Премии по экономике имени Нобеля (1972) и французского математика Ж. Дебре

(лауреат Премии по экономике имени Нобеля (1983)) ситуация заметно изменилась благодаря высокому математическому уровню статьи К. Эрроу и Ж. Дебре. Уровень был вполне сопоставим с математическим уровнем самых продвинутых статей естественно-научного профиля. Добавим, что Ж. Дебре был членом группы Н. Бурбаки, работавшей во Франции после окончания Второй мировой войны.

Исторически первой монографией экономико-математического содержания была работа «Исследования математических принципов теории богатства» французского математика А. Курно (1838). А. Курно предложил модель олигополии и равновесие этой модели, которое позже получило имя равновесия А. Курно. В модели А. Курно фирмы принимают решения об объеме выпускаемой продукции в условиях полной неопределенности, ибо не знают, какие решения примут конкуренты. В связи с этим вопрос о статусе оптимального решения этой модели является до сих пор открытым. Равновесие А. Курно — это субоптимальное решение модели, которое обладает свойством, что в одиночку ни одной фирме не выгодно выходить из этого равновесия.

Позднее, в 1950 г., американский математик Дж. Нэш (лауреат Премии по экономике имени Нобеля, 1994) предложил понятие равновесия некооперативной игры конечного числа игроков, которое позже получило его имя и которое обобщает равновесие А. Курно.

Во второй половине 19 в. французский экономист Л. Вальрас предложил понятие конкурентного равновесия в многопродуктовой экономике, состоящей из двух сфер: сферы производства и сферы потребления. Сфера производства состоит из конечного числа фирм, каждая из которых максимизирует свою прибыль. Сфера потребления состоит из конечного числа потребителей, каждый из которых максимизирует свою функцию полезности. Суть конкуренции по Л. Вальрасу в том, что фирмы и потребители каждый сам по себе решают свою задачу максимизации. Вопрос формулировался так: существуют ли такие цены (называемые ценами равновесия), ориентируясь на которые каждая фирма и каждый потребитель решают задачу максимизации и при этом не будет дефицита ни по одному продукту. Полное положительное решение этой задачи К. Эрроу и Ж. Дебре (как уже отмечалось выше) опубликовали в 1954 г. Они

использовали теорему о неподвижной точке точечно-множественного отображения японского математика Ш. Какутани, опубликованную в 1941 г.

В конце 19 в. итальянский экономист В. Парето предложил фундаментальное понятие, называемое сейчас Парето-эффективностью (Парето-оптимальностью), суть которого в следующем: совокупность потребительских наборов Парето-эффективна, если не существует другой совокупности потребительских наборов, на которых значения функции полезности не меньше, чем на первоначальных наборах и, по крайней мере, одно из этих неравенств должно быть строгим.

В 1945 г. на английском языке был опубликован перевод с немецкого статьи «Модель общего экономического равновесия» американского математика Дж. фон Неймана, в которой было предложено понятие динамического равновесия общей модели производства в матричной форме, охватывающей производственную и монетарную сферы. Статья Дж. фон Неймана оказала большое влияние на развитие математической экономики после окончания Второй мировой войны.

В 1958 г. в книге «Линейное программирование и экономический анализ» трех американских авторов — Дорфман Р., Сэмуэльсон П. (лауреат Премии по экономике имени Нобеля, 1970), Солоу Р. (лауреат Премии по экономике имени Нобеля, 1987) — было описано магистральное свойство оптимальных траекторий валовых выпусков и цен, суть которого состояла в том, что в случае продолжительного временного промежутка модели ее оптимальная траектория выпусков (для оптимальной траектории цен ситуация аналогична) состоит из трех участков. Первый участок — это отрезок оптимальной траектории, который приближается к траектории максимального постоянного пропорционального роста, расположенной на луче, называемом магистралью; второй участок — это отрезок оптимальной траектории, который близок (в смысле углового расстояния) к траектории максимального постоянного пропорционального роста; третий участок — это отрезок оптимальной траектории, который может отойти от траектории максимального постоянного пропорционального роста. Основные характеристики траектории максимального постоянного пропорционального роста (темп роста и структура) определяются технологическим множеством модели, то есть эндогенно. Тео-

ретическое значение магистрального свойства состоит в том, что в случае продолжительного временного горизонта модели оптимум достигается через максимальный постоянный пропорциональный рост, практическое значение магистрального свойства в том, что оно позволяет дать рациональное решение «проблемы хвоста» и проблемы сглаживания оптимальной траектории.

В заключение подчеркнем наличие взаимного влияния экономической проблематики и отдельных разделов математики: математика использовалась и используется для решения задач экономической теории и хозяйственной практики, а экономические задачи выступали в качестве генераторов новых областей как прикладной (линейное программирование), так и чистой математики (выпуклый анализ).

Бурное развитие компьютеризации привело к повышению эффективности применения математики в решении прикладных экономических задач.

Список литературы

- [1] Экономико математический энциклопедический словарь. — М.: Большая Российская энциклопедия, Изд. дом «ИНФРА-М», 2003.

Достаточное условие разрешимости двухмашинной задачи Open shop с маршрутизацией и разрешением прерываний

Черных И. Д. (Новосибирск, ИМ СОРАН)

Кузеванов М. А. (Новосибирск, НГУ)

idchern@math.nsc.ru, theredstar@mail.ru

Рассматриваемая задача обобщает две классические задачи дискретной оптимизации: метрическую задачу комивояжера и задачу Open Shop с прерываниями. Имеется n работ J_1, \dots, J_n , каждая из которых должна быть обработана каждой из m машин M_1, \dots, M_m в произвольном порядке. Операция O_{ji} обработки работы J_j машиной M_i занимает p_{ji} единиц времени. Операцию можно прерывать в любой момент времени и возобновить ее выполнение позднее. Работы расположены в вершинах транспортной сети, описываемой реберно-взвешенным графом $G = \langle V, E \rangle$, расстояние между вершинами v_i и v_j обозначается через τ_{ij} . Множество работ, расположенных в вершине v_k , обозначим через \mathcal{J}_k . Одна из вершин является *базой*; изначально все машины находятся в базе и должны вернуться туда после выполнения всех своих операций. Машины передвигаются с единичной скоростью. Требуется составить допустимое расписание, минимизирующее время возвращения всех машин на базу после выполнения всех операций. Задача рассматривается в двух постановках: с заданной базой или с выбираемой базой, обозначаемых соответственно $ROm|pmtn|F_{\max}$ и $ROm|pmtn, free - depot|F_{\max}$.

Задача является NP-трудной в сильном смысле даже в случае $m = 1$ (в этом случае она эквивалентна метрической задаче комивояжера). Двухмашинная задача является полиномиально разрешимой в случае, когда транспортная сеть состоит из двух вершин [1]. Алгоритмическая сложность двухмашинной задачи на треугольнике является открытым вопросом. Двухвершинная задача с нефиксированным числом машин является NP-трудной в сильном смысле [1]. Целью данной работы является описание полиномиально-разрешимых подклассов задач $RO2|pmtn|F_{\max}$ и $RO2|pmtn, free - depot|F_{\max}$.

Поскольку задача включает в себя задачу комивояжера, во всех ее полиномиально разрешимых подслучаях граф G и/или функция расстояний τ должны позволять решить задачу комивояжера за полиномиальное время. Например, число вершин графа может быть фиксированным, или структура графа позволяет упростить поиск кратчайшего обхода.

Введем следующие обозначения. Через $\ell_i \doteq \sum_j p_{ji}$ обозначим *нагрузку* машины M_i через $d_j \doteq \sum_i p_{ji}$ обозначим *длину* работы J_j . Длину кратчайшего обхода графа G обозначим через T^* . Величина $\tilde{F} \doteq \max\{T^* + \max \ell_i, \max d_j\}$ является нижней оценкой оптимума задачи в обоих постановках.

Задача с выбираемой базой

Справедлива следующая

Теорема 1. Пусть кратчайший обход графа G может быть найден за время t_{TSP} . Пусть суммарная длина работ в одной из вершин не меньше \tilde{F} . Тогда для задачи $RO2|pmtn, free - depot|F_{\max}$ оптимальное расписание имеет длину \tilde{F} и такое расписание с не более чем одним прерыванием может быть найдено за время $O(n + t_{TSP})$.

Опишем идею алгоритма. В качестве базы выберем вершину с наибольшей суммарной длиной работ. Найдем кратчайший обход R графа G и переобозначим вершины в порядке обхода R начиная с базы. Заметим, что поскольку $\sum_j d_j = \ell_1 + \ell_2 \leq 2(\tilde{F} - T^*)$, справедливо

$$\sum_{J_j \notin \mathcal{J}_1} d_j \leq \tilde{F} - 2T^*. \quad (1)$$

Проведем процедуру «огрубления» исходного примера следующим образом. Для каждой вершины v_k , $k \neq 1$, заменим все работы из \mathcal{J}_k одной новой работой J'_k , операции которой имеют длительность $p'_{ki} = \sum_{J_j \in \mathcal{J}_k} p_{ji}$. В вершине v_1 проведем «склеивание» работ следующим образом. Пусть $\mathcal{J}_1 = \{J_1, \dots, J_r\}$. Если $\sum_{J_j \in \mathcal{J}_1} d_j = \tilde{F}$, заменим все

работы из базы на одну так же, как в остальных вершинах. В противном случае выберем наибольшее l такое, что $\sum_{j=1}^l d_j \leq \tilde{F}$. Заменяем работы J_1, \dots, J_l на одну работу J'_{1-1} , переобозначим работу J_{l+1} через J'_{1-2} , и заменим работы J_{l+2}, \dots, J_r на новую работу J'_{1-3} (если $l+1 < r$). Заметим, что в новом примере длины всех работ не превосходят \tilde{F} . Без ограничения общности считаем, что в новом примере в базе ровно три работы (при необходимости добавим фиктивные работы с операциями нулевой длительности).

Обозначим операции работы J'_j нового примера через a_j, b_j . Теми же буквами будем обозначать длительности этих операций. При необходимости, перенумеруем работы из базы и/или машины так, чтобы выполнялось

$$a_{1-2} \geq b_{1-1}, \quad a_{1-3} \geq b_{1-2}. \tag{2}$$

Построим расписание по следующей схеме. Машина M_1 сначала выполняет операции работ J'_{1-1}, J'_{1-2} и J'_{1-3} , потом следует по маршруту R , выполняя операции в каждой вершине, и возвращается на базу. Машина M_2 сначала переезжает в вершину v_2 , выполняет соответствующую операцию и продолжает движение по маршруту R , выполняя операции в каждой вершине. После возвращения на базу выполняет операции работ J'_{1-1}, J'_{1-2} и J'_{1-3} . При этом для работ из базы выполняется сначала операция первой машины, потом второй, для остальных работ — в обратном порядке.

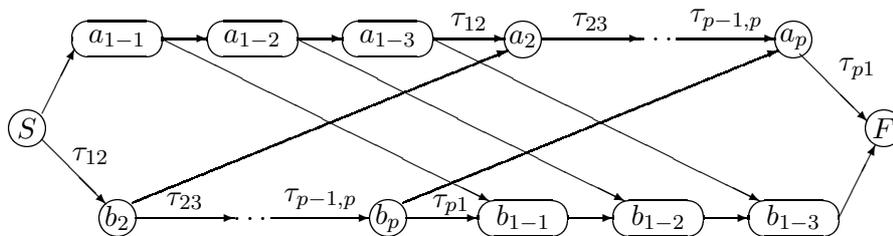


Рис. 1. Схема N построения расписания.

Длина раннего расписания S , построенного по схеме N , совпадает с длиной некоторого критического пути в N из S в F . Заметим, что если критический путь не содержит работ из базы, то в силу (1) его

длина не превосходит \tilde{F} . Если критический путь содержит только операции одной машины, то его длина по определению \tilde{F} также не превосходит \tilde{F} .

Среди оставшихся путей из S в F в силу (2) критическим может быть только $S \rightarrow a_{1-1} \rightarrow a_{1-2} \rightarrow a_{1-3} \rightarrow b_{1-3} \rightarrow F$. Пусть его длина превышает \tilde{F} . Тогда расписание S имеет следующий вид:

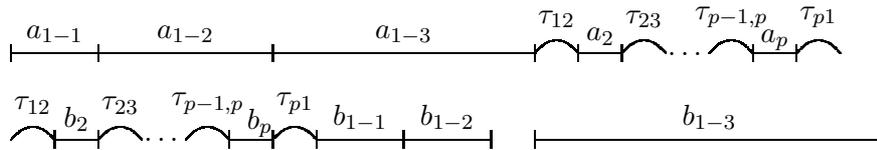


Рис. 2. Вид неоптимального расписания S .

Это расписание несложно перестроить в расписание длины \tilde{F} следующим образом. Разделим операцию b_{1-3} на два фрагмента, b'_{1-3} и b''_{1-3} так, что $b''_{1-3} = T^* + \sum_{j=2}^p a_j$. Фрагмент b'_{1-3} выполним в момент времени 0, сдвинув вправо интервалы выполнения операций машины M_2 и, при необходимости, операцию a_{1-3} и последующие операции машины M_1 . Длина полученного расписания будет равна $\max\{\ell_1 + T^*, \ell_2 + T^*, d_{1-3}\} \leq \tilde{F}$.

Задача с заданной базой

Без ограничения общности считаем, что вершина v_1 является базой. Для задачи $ROm|pmtn|F_{\max}$ величина $\bar{F} \doteq \max\{\max_i \ell_i + T^*, \max_k \{\max_{j \in \mathcal{J}_k} d_j + 2\tau_{1k}\}\}$ является нижней оценкой оптимума.

Вершину v_k назовем *перегруженной*, если $\sum_{j \in \mathcal{J}_k} d_j \geq \bar{F} - 2\tau_{1k}$.

Справедлива следующая

Теорема 2. Пусть кратчайший обход R графа G может быть найден за время t_{TSP} , v_2 и v_p — соседние с базой вершины в R . Если одна из вершин v_1 , v_2 и v_p является перегруженной, то для задачи $RO2|pmtn, free - depot|F_{\max}$ оптимальное расписание имеет длину

\bar{F} и такое расписание с не более чем одним прерыванием может быть найдено за время $O(n + t_{TSP})$.

Доказательство этой теоремы аналогично доказательству теоремы 1.

Работа выполнена при финансовой поддержке РФФИ, грант № 08-01-00370.

Список литературы

- [1] Пяткин А. В., Черных И. Д. Задача открытого типа с маршрутизацией и разрешением прерываний на двухвершинной сети // Материалы IV Всероссийской конференции «Проблемы оптимизации и экономические приложения». Омск, 29.06–4.07.09. — С. 158.