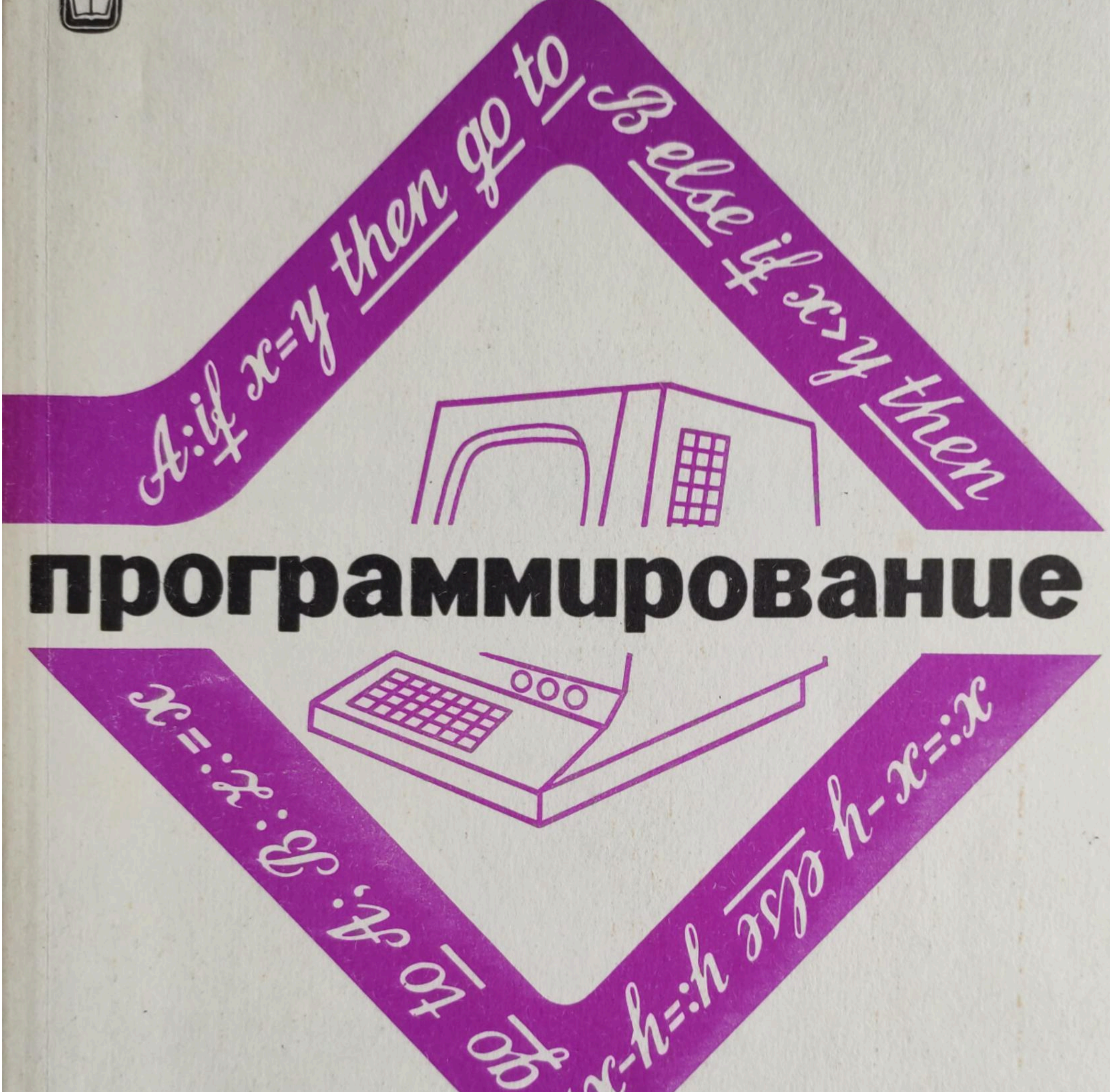
Российская академия наук

ISSN 0132-3474



программирование

4

1994

РОССИЙСКАЯ АКАДЕМИЯ НАУК

ПРОГРАММИРОВАНИЕ

4

ИЮЛЬ – АВГУСТ

1994

Журнал основан	Выходит	Москва
в январе 1975 г.	6 раз в год	"Наука"

СОДЕРЖАНИЕ

Машинная графика

Клименко С.В. От редколлегии 3

Турлапов В.Е. (Н.Новгород), **Якунин В.И.** Эффективный базис моделирования очертания огибающих, порожденных движением поверхностей вращения 5

Крицкий С.П., **Заставной Б.А.** (Ростов-на-Дону). Параллельная печать текста интерпретатором языка PostScript 14

Гёбель М. (Германия), **Клименко С.В.** (Протвино). Научная визуализация в виртуальном окружении 29
Клименко С.В., Никитин И.Н., Таланов В.В. (Протвино). Визуализация особенностей на мировых листах релятивистской струны 47
Флямер М.Г., Клименко С.В. (Протвино). Визуализация процесса светосбора в сцинтилляционных детекторах 58
Кочин В.Н., Куликов В.А. Программные средства визуализации данных на персональном компьютере 72

Компьютерная алгебра

Спиридонова М., Байчева Б. (Болгария). Символьное выполнение преобразования Лежандра 77
Абрамов С.А. Поправка 86

Сообщение

Ершов Н.М. Оценка времени выполнения вложенных операторов цикла .. 87

Обзор периодических изданий АСМ

Максимов А.В., Петренко А.К., Юфа В.М. Обзор публикаций АСМ . 92

Зав. редакцией *О.С.Разумовская*

Технический редактор *Е.А.Красина*

Сдано в набор 25.05.94. Подписано к печати 1.08.94. Формат бумаги 70×100 1/16
Печать офсетная. Усл.печ.л. 7,8. Усл.кр.-отт. 17,1 тыс. Уч.-изд.л. 92
Бум.л. 3,0. Тираж 2122 экз. Зак. 1521.

Адрес редакции: 119899, Москва, В-234, МГУ, факультет ВМиК. Тел. 939-55-74
Московская типография N^o2 РАН, 121099, Москва, Г-99, Шубинский пер., 6

УДК 519.685.3

© 1994 г. Н.М.Ершов

ОЦЕНКА ВРЕМЕНИ ВЫПОЛНЕНИЯ ВЛОЖЕННЫХ ОПЕРАТОРОВ ЦИКЛА

Рассматривается проблема статического определения времени выполнения вложенных операторов цикла. Предлагается подход, основанный на замене многомерных сумм, через которые выражается данная оценка, на многомерные интегралы. Рассматривается и сравнивается ряд приближенных методов вычисления получаемых интегралов.

Важной проблемой, возникающей при решении задачи статического крупноблочного распараллеливания больших вычислительных программ, является определение времен выполнения (весов) их различных частей — операторов, функций, подпрограмм [1,2]. Наиболее вычислительноемкими в подобных программах являются конструкции из вложенных операторов цикла [2]. Оценка их веса может быть выражена через многомерную сумму — каждый оператор цикла соответствует одной из сумм, при этом индекс цикла соответствует переменной суммирования, пределы цикла — пределам суммирования. Когда каждый оператор цикла имеет постоянные пределы (т.е. не зависящие от индексов охватывающих его операторов циклов), решение задачи является тривиальным [2]. Однако, если ряд циклов имеет пределы, которые зависят от индексов внешних циклов, задача определения веса становится существенно сложнее. Эта проблема в большинстве работ, связанных со статическим анализом вычислительных программ, либо просто игнорируется, либо для ее решения предлагаются методы, дающие достаточно грубые оценки. Отметим, что не требуется абсолютно точного вычисления получаемых многомерных сумм, так как сама задача статического анализа является приближенной, в силу невозможности учесть при анализе текста программы все факторы, влияющие на время ее выполнения. Поэтому в настоящей работе предлагается приближенный подход к определению времени выполнения конструкций из вложенных операторов циклов, в основе которого лежит идея замены указанных сумм многомерными интегралами. Смысл такой замены состоит в значительном упрощении задачи, так как существуют достаточно развитые теория и методы приближенного вычисления интегралов [3]. Рассматривается и сравнивается ряд таких приближенных методов применительно к задаче определения веса вложенных операторов циклов.

1. Будем рассматривать циклы вида

(1) $$DOi = A, B, HT(i),$$

где i является индексом цикла; A и B — нижним и верхним пределами; H — шагом, предполагается, что число итераций, равное $N = (B - A)/H$, есть целое число; $T(i)$ — телом цикла, которое может содержать любую последовательность операторов, в том числе и операторов цикла того же вида. В дальнейшем, под циклами будем подразумевать циклы вида (1).

Назовем цикл внешним, если он не содержится в другом цикле, и внутренним в противном случае. Мощностью [2] внешнего цикла назовем максимальное число вложенных друг в друга циклов в теле $T(i)$, индексной мощностью — максимальное число вложенных друг в друга циклов в $T(i)$, пределы которых явно или неявно зависят от индекса i. Смысл введения параметра индексной мощности состоит в том, что она в некоторой степени является характеристикой сложности определения весов внешних циклов (см. введение). В таблице 1 приведена относительная доля внешних циклов с различными значениями этих мощностей среди проанализированных программ вычислительной физики плазмы. Отметим, что практически все рассмотренные циклы имеют индексную мощность, меньшую или равную 2. Полученные данные хорошо согласуются с результатами специального статистического анализа вычислительных программ, проведенного в [2].

Таблица 1

	0	1	2	3	4
Мощность	62.9 %	27.2 %	8.8 %	0.8 %	0.3 %
Индексная мощность	93. %	5.6 %	0.7 %	0 %	0 %

Анализ тех же программ по типу зависимости пределов внутренних циклов от индексов внешних циклов показал, что эта зависимость является линейной в 93% случаев, и нелинейной в 7% случаев.

2. Пусть вес тела $T(i)$ равен $C(i)$, тогда вес самого цикла (1) равен

$$C = C(A) + C(A+H) + C(A+2H) + \ldots + C(B) = \sum_{i=0}^{N} C(A+iH) \tag{2}$$

В случае $H = 1$, формула (2) имеет более естественный вид

$$C = \sum_{i=A}^{B} C(i) \tag{3}$$

В качестве примера в настоящей работе будут рассматриваться гнезда циклов вида

$$\begin{aligned} &DO i = 0, N, 1 && (4)\\ &DO j_1 = 0, i, 1 \\ &\ldots \\ &DO j_K = 0, i, 1 && (5)\\ &T(i, j_1, \ldots, j_K), \end{aligned}$$

у которых вес тела $T(i, j_1, \ldots, j_K)$ равен 1. Вес таких циклов равен

$$C = \sum_{i=0}^{N} \sum_{j_1=0}^{i} \ldots \sum_{j_K=0}^{i} 1 = \sum_{i=0}^{N} i^K \tag{6}$$

Отметим, что для циклов (3) мощность и индексная мощность равны K. Вообще, для циклов с линейными по i пределами в теле $T(i)$, вес $C(i)$ представляет собой полином, степень которого равна индексной мощности цикла (1).

3. Основной идеей подсчета сумм вида (2,2') является замена операции суммирования операцией интегрирования

$$\tilde{C} = \frac{1}{H}\int_{A-H/2}^{B+H/2} C(i)di. \tag{7}$$

Смысл данной замены поясняется на рис.1. Площадь прямоугольников равна искомой сумме (2), умноженной на шаг H, и эта площадь приближенно оценивается площадью под кривой $C(i)$, которая и равна $\tilde{C}$.

Можно оценить погрешность такой замены, если учесть, что формула (2) есть не что иное, как формула прямоугольников приближенного вычисления интеграла (5) с шагом H. Поэтому справедлива следующая оценка

$$\Psi_0 = |C - \tilde{C}| \le \frac{(B-A+1)}{24} H^2 M_2, \text{где} M_2 = \sup_{i=A,B} |C''(i)|. \tag{8}$$

Следствием формулы (6) является то, что для циклов с линейными пределами в $C(i)$ и индексной мощностью меньшей 2 интегральный переход (5) является точным, так как в этом случае $C(i)$ линейна, и константа M_2, следовательно, равна 0.

4. Как следствие данных, приведенных в пункте 1, для большинства внешних операторов циклов, в силу линейности пределов у содержащихся в них внутренних циклов, возможно точное вычисление интеграла (5), так как все подинтегральные функции в этом случае есть полиномы. Однако автоматизация этого процесса требует проведения достаточно сложных символьных вычислений. Поэтому были рассмотрены приближенные методы вычисления интегралов (5). Далее, не ограничивая общности рассуждений, будем считать, что $H = 1$.

В методе средней точки интеграл (5) приближенно оценивается как

$$\tilde{C} \approx (B-A+1)C(\frac{B+A}{2}). \tag{9}$$

Так как это есть частный случай формулы прямоугольников с шагом $(B-A+1)$, то погрешность этого метода оценивается как

$$\Psi_1 \le \frac{(B-A+1)^3}{24} M_2. \tag{10}$$

Метод требует вычисления функции $C(i)$ только в одной точке.

Интеграл (5) также может быть приближенно вычислен через формулу прямоугольников

$$\tilde{C} \approx h\sum_{j=0}^{L-1} C(A + \frac{h-1}{2} + jh), \text{где} h = \frac{B-A+1}{L}. \tag{11}$$

При $L = 1$ формула переходит в (7). Данный метод имеет смысл применять только для $L \ll B - A + 1$, в противном случае проще и точнее подсчитать саму исходную сумму (2). Погрешность метода определяется как

$$\Psi_2 \le \frac{(B-A+1)^3}{24L^2} M_2. \tag{12}$$

Метод требует вычисления функции $C(i)$ в L точках.

По методу Симпсона для одного отрезка интеграл (5) оценивается следующим образом:

$$\tilde{C} \approx \frac{(B-A+1)}{6}\left(C(A-\frac{1}{2}) + 4C(\frac{B+A}{2}) + C(B+\frac{1}{2})\right). \tag{13}$$

Погрешность этого метода

$$\Psi_3 \leq \frac{(B-A+1)^5}{2880} M_4, \text{где} M_4 = \sup_{i=A,B} |C^{(IV)}(i)|. \tag{14}$$

Данный метод дает точное значение интеграла (5) для циклов (1) с линейными по i пределами у вложенных циклов и индексной мощностью, меньшей 4, и требует вычисления функции $C(i)$ в трех точках.

5. Для циклов (3) был проведен сравнительный анализ предложенных методов. Его результаты для $N = 100$ сведены в таблице 2, в которой $\delta_0 = \Psi_0/C$ (%) есть относительная погрешность замены (5); $\delta_i = \Psi_i/\tilde{C}$ (%), $i = 1, 3$ — относительные погрешности предложенных методов (7,9,11). В методе прямоугольников полагается $L = 10$. Также в таблице приведены примеры методов линейной алгебры, которые содержат циклы индексной мощности K.

Таблица 2

K	δ_0, %	δ_1, %	δ_2, %	δ_3, %	Примеры методов
0	0	0	0	0	Скалярное умножение векторов
1	0	0	0	0	Решение треугольных систем
2	0.0024	25.4	0.25	0	LU-разложение
3	0.0049	50.5	0.50	0	—
4	0.0120	69.0	0.84	4	—

На основании проведенного исследования можно сделать следующие выводы. Погрешность замены (5) для циклов (1) является удовлетворительной для циклов с большим числом итераций. Для большинства (90%) внешних циклов в проанализированных вычислительных программах (циклы с линейными пределами у внутренних циклов и индексной мощностью, меньшей двух) формула средней точки дает точное значение интеграла, и при этом требует наименьших вычислений. Однако, на практике имеет смысл применять формулу Симпсона, так как она является точной для большего числа циклов, и требует не намного больше вычислений по сравнению с формулой средней точки. Формула прямоугольников с $L > 1$, по-видимому, не имеет практического интереса, так как ее преимущество перед формулой парабол начинается с циклов с индексной мощностью, большей 3, а такие циклы встречаются чрезвычайно редко (см. таблицу 1).

В настоящее время описанный подход реализован в системе построения взвешенных графов вычислительных алгоритмов [1] по текстам их реализующих OccAM-программ. Получаемая при этом информация о весах в дальнейшем используется при решении задачи оптимальной загрузки мультитранспьютерных вычислительных систем.

Список литературы

1. *Ершов Н.М., Попов А.М.* Оптимизация параллельных вычислений с учетом архитектуры и быстродействия связей в вычислительной системе // Вестник Московского университета. Сер.15. 1993. N^o 1. с.24–30.

2. *Воеводин Вл.В.* Статистические оценки возможности выявления параллельной структуры последовательных программ // Программирование. 1990. N° 4. с.44–54.

3. *Самарский А.А., Гулин А.В.* Численные методы. М. Наука. 1989.

Поступила в редакцию 18.04.94

N.M.Ershov

AN ESTIMATION OF THE NESTED LOOPS PERFORMANCE TIME

Problem of the static evaluation of the nested loops performance time is considered. An approach, based on the replacement of many-dimensional sums, from which this estimation depends on, to many-dimensional integrals is suggested. A number of approximate methods of these integrals calculation is considered and compared.

Ершов Николай Михайлович, окончил Московский университет в 1991 году. В настоящее время является аспирантом кафедры автоматизации научных исследований факультета ВМиК МГУ. Научные интересы — параллельные вычисления в задачах вычислительной физики.