

УДК 621.325.5

А.В. ДРОЗД

Одесский национальный политехнический университет, Украина

РАБОЧЕЕ ДИАГНОСТИРОВАНИЕ В ОБРАБОТКЕ ПРИБЛИЖЕННЫХ ДАННЫХ

Рассмотрены особенности рабочего диагностирования вычислительных устройств в условиях противостояния сложившейся модели точных данных и возрастающего значения обработки приближенных данных. Показано, что рабочее диагностирование разработано для случая точных данных, что не согласуется с реалиями окружающего мира, несуществующего вне допусков, описываемых приближенными числами. Под действием модели точных данных произошло искажение цели рабочего диагностирования, неправильно оценивается достоверность его методов. По мере развития компьютерной техники отмеченное противоречие усиливается, ограничивая возможности решения вычислительных задач.

рабочее диагностирование, вычислительные устройства, модель точных данных, приближенные вычисления, достоверность результатов, существенные и несущественные ошибки, достоверность методов рабочего диагностирования

Постановка проблемы

Использование вычислительной техники, ее методов и средств преследует единую цель – решение вычислительной задачи с получением достоверных результатов при выполнении всего объема требуемых вычислений в установленный срок.

Этой цели должны быть подчинены как сам вычислительный процесс, так и все этапы, предшествующие ему, включая проектирование средств решения задачи и различные виды диагностирования.

Реально все составляющие, обеспечивающие решение задачи, обособились в отдельные направления, ограниченные локальными целями, которые могут существенно расходиться с единой целью. Примером может служить совершенствование компьютера, которое стало самоцелью, т.е. цель «решение задачи» подменяется улучшением инструмента для ее решения.

Одной из важных причин, приводящих к такому расхождению, является противоречие, существующее между ограниченными возможностями компьютера представлять числовую информацию и реалиями нашего мира.

Числовые данные, записанные в разрядную сетку компьютера, могут быть отождествлены с их порядковыми номерами. Например, в 3-разрядную сетку можно записать двоичные коды от 000_2 до 111_2 и пронумеровать их от нулевого значения до седьмого. Номера элементов множеств являются единственным представителем точных данных, т.е. целых по своей природе. Таким образом, компьютер поддерживает модель точных данных, которая состоит в том, что любые числа рассматриваются как точные независимо от их настоящей природы. Это можно отнести и к представлению приближенных данных в форматах с плавающей точкой, где мантисса также записывается в разрядной сетке, т.е. целым по своей природе, точным данным [1].

В плену модели точных данных оказалось все, связанное с компьютерным представлением данных, включая проектирование и диагностику вычислительных устройств и систем. Их развитие состоялось и продолжается в рамках обработки точных данных. Однако нас окружает мир приближенных данных, которые содержат ошибку от рождения.

С первых классов школьного образования складывается стереотип, определяющий негативное от-

ношение к ошибке, как нарушению, которое следует исключать из практики. Создается устойчивое убеждение, что это возможно. Между тем, ошибка является одной из основ нашего мира, вне которой этот мир просто не существует.

Ошибка – это разница между абсолютной и относительной истинами, т.е. посредством ошибки осуществляется познание мира.

Методом проб и ошибок происходит развитие мира. В живой природе такой ошибкой является мутация, которая при определенных условиях закрепляется в ходе естественного отбора как основной путь дальнейшей эволюции.

Ошибка определяет также способ существования живого и неживого в этом мире. Любое изделие изготавливается с учетом допусков, в пределах которых оно существует, а за пределами – отбраковывается, т.е. ошибка составляет жизненное пространство существования этого изделия. Человек имеет право на ошибку в том смысле, что в пределах некоторого множества ошибок остается человеком, а не утрачивает себя в каком-либо из смыслов. Ошибки, позволяющие существовать, являются несущественными, т.е. не противоречат существованию. Вне ошибок (а именно такого их типа как несущественные ошибки) нет существования.

Все перечисленное указывает на наш мир как источник приближенных данных – результатов измерений, выполняемых в пределах допусков (несущественных ошибок), и их обработки.

По мере развития компьютер, как и все остальное, структурируется под реалии приближенного мира, что проявляется в неуклонном увеличении доли обработки приближенных данных. Этот процесс поддерживается расширением круга решаемых задач, разработкой и совершенствованием средств и методов их решения. Это ведет к усилению противоречия между господствующей сегодня моделью точных данных и объективными проявлениями приближенного мира.

Человечество вступило в эру принятия ответственных решений. От способности просчитывать их последствия зависит, насколько долго продлится эта эра. Защищенность общества все чаще определяется умением решать задачи на пределе возможностей, которые основываются на правильной оценке допусков, способности различать существенные и несущественные ошибки.

Неприятие ошибки как способа существования мира значительно ограничивает возможности решения вычислительных задач. Поскольку точные данные являются номерами элементов множеств, то ошибка в точном данном – это ошибка адресации к элементу множества. Для точного данного любая ошибка является существенной, поскольку все его разряды относятся ко множеству верных разрядов, и искажение любого из них делает число недостоверным [2]. Модель точных данных исключает использование несущественных ошибок, лишая число жизненного пространства.

Степень заблуждения, вызванного моделью точных данных, можно оценить на примере развития рабочего диагностирования (РД) вычислительных устройств.

Рабочее диагностирование вычислительных устройств

К концу 60-х годов прошлого столетия в теории самопроверяемых цифровых схем были сформулированы основные положения РД, сложившиеся к этому времени [3, 4]. К ним следует отнести:

- цель РД – обнаружение неисправностей цифровых схем в процессе выполнения вычислений, т.е. на фактических данных;
- требование к методам РД – обнаруживать неисправность цифровой схемы по первой ошибке результата вычислений.

Эти положения определили развитие РД и являются действующими по сей день, не вызывая никаких сомнений по поводу их корректности [5, 6].

Однако объявленная цель противоречит здравому смыслу и нереализуема в рамках сформулированного требования к методам РД. К тому же не соответствует практике использования РД. Действительно, поиск неисправностей в процессе вычислений не согласуется с целью решить вычислительную задачу. При каждой «удаче» в реализации объявленной цели откладывается достижение единой цели.

Настоящая роль РД в решении вычислительной задачи состоит в оценке достоверности вычисляемых результатов, что позволяет исключить из вычислений недостоверные результаты.

Ошибки возникают под действием сбоев или отказов, причем сбои происходят значительно чаще отказов [7]. Поэтому метод РД, отвечающий поставленному требованию, по первой ошибке, как правило, обнаружит сбой – кратковременную самоустранившуюся неисправность, которая не позволяет судить об исправности цифровой схемы в следующий момент времени. Отказ и вовсе невозможно распознать по одной ошибке – то ли одиночной, вызванной сбоем, то ли являющейся началом порождаемой отказом серии ошибок.

Почему же на практике стремятся обнаруживать первую ошибку? – Только для получения ответа на вопрос: «Является ли вычисленный результат достоверным или его следует пересчитать?»

Таким образом, настоящая цель РД – оценка достоверности вычисляемых результатов.

Почему же объявлена другая цель и насколько эти цели различаются?

По первой части вопроса следует отметить, что ошибка, обнаруженная методом РД, одновременно указывает и на наличие неисправности в цифровой схеме и на вычисление недостоверного результата. Однако последнее имеет место только для случая точных данных. Следовательно, метод РД под видом обнаружения неисправностей цифровой схемы реально оценивает достоверность результатов обработки точных данных.

Поэтому для точных данных объявленная и настоящая цели РД вычислительных устройств по существу совпадают.

Ответ на вторую часть вопроса может быть получен при рассмотрении возможностей обнаружения ошибки, вызванной неисправностью, на квадрате с единичной длиной сторон и, соответственно, единичной площадью (рис. 1).

	P_C	$P_H = 1 - P_C$
P_O	$P_{OC} = P_O P_C$ 1	$P_{OH} = P_O P_H = P_O (1 - P_C)$ 2
$P_{\Pi} = 1 - P_O$	3 $P_{\Pi C} = P_{\Pi} P_C = (1 - P_O) P_C$	4 $P_{\Pi H} = P_{\Pi} P_H = (1 - P_O) (1 - P_C)$

Рис. 1. Вероятности обнаружения и пропуска ошибок

На горизонтальной стороне квадрата откладываются вероятности P_C и $P_H = 1 - P_C$ того, что ошибка является существенной и несущественной, а на вертикальной стороне – вероятности P_O и $P_{\Pi} = 1 - P_O$ ее обнаружения и пропуска. При этом квадрат разбивается на четыре части, определяющие вероятности P_{OC} и P_{OH} обнаружения существенных и несущественных ошибок, а также вероятности $P_{\Pi C}$ и $P_{\Pi H}$ их пропуска, для которых выполняется соотношение

$$P_{OC} + P_{OH} + P_{\Pi C} + P_{\Pi H} = 1,$$

где $P_{OC} = P_O P_C$;

$$P_{OH} = P_O P_H = P_O (1 - P_C);$$

$$P_{\Pi C} = P_{\Pi} P_C = (1 - P_O) P_C;$$

$$P_{\Pi H} = P_{\Pi} P_H = (1 - P_O) (1 - P_C).$$

Согласно объявленной цели, метод РД является достоверным, если обнаруживается неисправность, независимо от того, является ли вызванная ею ошибка существенной или нет. Это определяет показатель достоверности D_O метода РД как сумму частей 1 и 2 квадрата по следующей формуле:

$$D_O = P_{OC} + P_{OH}.$$

Согласно настоящей цели, метод РД является достоверным, если правильно оценивает вычисляемый результат: достоверный результат определяет достоверным, а недостоверный результат – недостоверным. Оценка достоверности результата производится путем обнаружения или пропуска ошибки, а действительным признаком достоверного или недостоверного результата является отсутствие или наличие существенной ошибки.

В части 1 квадрата метод РД обнаруживает существенную ошибку, справедливо определяя результат недостоверным.

В части 2 квадрата обнаруживается несущественная ошибка, что приводит к отбраковке достоверного результата – метод РД дает ошибочную оценку результата.

Часть 3 квадрата также указывает на ошибку метода РД, поскольку пропущена существенная ошибка, и недостоверный результат не выявлен.

В части 4 квадрата метод РД игнорирует несущественную ошибку, справедливо подтверждая достоверность вычисляемого результата.

Поэтому показатель достоверности метода РД складывается из частей 1 и 4 квадрата и определяется по следующей формуле:

$$D = P_{OC} + P_{OH}.$$

Объявленная и настоящая цели различаются в оценке достоверности метода РД: показатели достоверности D_O и D отличаются друг от друга на части 2 и 4 квадрата.

Каковы размеры различающих частей квадрата?

Для ответа на этот вопрос следует принять во внимание высокую вероятность обнаружения ошибок, достигнутую в современных методах РД (контроль по паритету, по модулю [8, 9]), т.е. $P_O \gg P_{\Pi}$.

В случае точных данных вероятность $P_C = 1$, поскольку все ошибки являются существенными.

Из этого следует, что

$$P_{OH} = P_O (1 - P_C) = 0;$$

$$P_{\Pi H} = (1 - P_O) (1 - P_C) = 0.$$

Иными словами, для случая точных данных части 2 и 4 исключаются из квадрата, как это показано на рис. 2, а показатели достоверности D_O и D , описываемые частью 1 квадрата, совпадают и являются высокими.

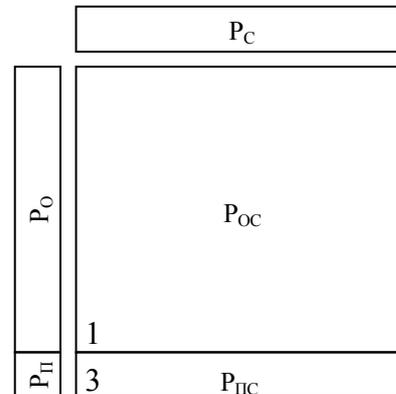


Рис. 2. Случай точных данных

Это еще раз свидетельствует о том, что РД разработано для частного случая обработки точных данных и в этих рамках обеспечивает высокую эффективность.

Для общего случая необходимо учесть соотношение $P_C \ll P_H$, присущее обработке приближенных данных [10, 11].

На рис. 3 показана реальная картина разбиения квадрата на части.

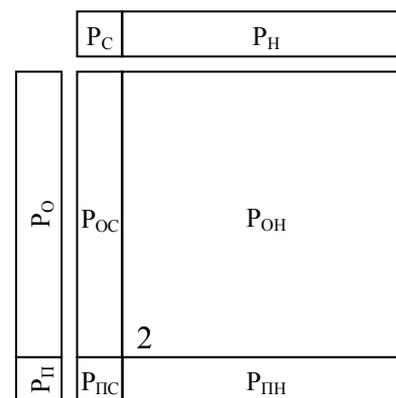


Рис. 3. Реальная оценка обнаружения и пропуска ошибок

Размеры части 1 квадрата значительно сокращаются вследствие малого значения вероятности P_C , а

часть 2, напротив, увеличивает свои размеры и становится основной частью квадрата, описывая совершенно новое качество методов РД отбраковывать достоверные результаты при обнаружении несущественных ошибок.

Ранее, т.е. в рамках действия модели точных данных, достоверные результаты вычислений могли быть отбракованы только при условии, что ошибка находится в контрольных кодах и вызвана неисправностью средств контроля (ошибка второго рода). Сам метод РД не мог отбраковать достоверный результат по причине отсутствия несущественных ошибок в результате обработки точных данных.

Реальная оценка вероятностей обнаружения и пропуска существенных и несущественных ошибок позволяет сделать следующие выводы:

- объявленная и настоящая цели РД различаются на самую большую часть 2 квадрата, которая практически совпадает с самим квадратом, т.е. для приближенных данных цели становятся диаметрально противоположными;

- показатель достоверности методов РД является высоким в рамках объявленной цели и низким для настоящей цели.

Поскольку квадрат описывает возможности обнаружения уже произошедшей ошибки результата, то сделанные выводы имеют влияние на решение вычислительных задач в зависимости от вероятности появления ошибок.

При решении вычислительной задачи на пределе возможностей, ограничением степени распараллеливания вычислений является сложность средств решения задачи, а ограничением сложности – их надежность.

Таким образом, решение сложных задач происходит в условиях плотного потока неисправностей, что приводит к значительной доле результатов, посчитанных с ошибками.

Пересчет ошибочных результатов может существенно удлинить время решения задачи, исключая

самую возможность ее решения в реальном масштабе времени.

Известная организация вычислений с использованием контрольных точек предполагает повторение операций от предыдущей контрольной точки, если в данной контрольной точке вычислены ошибочные результаты. Для обеспечения высокой вероятности выполнения безошибочных вычислений между соседними контрольными точками необходимо уменьшать количество таких вычислений и соответственно увеличивать количество контрольных точек, на которые разбивается процесс решения вычислительной задачи. Это также ведет к увеличению времени решения вычислительной задачи и ограничивает возможности решения сложных задач.

Несущественные ошибки и их высокая вероятность P_H составляют важную предпосылку для решения сложной задачи, поскольку большинство результатов, посчитанных под действием проявившихся в виде ошибки неисправностей, являются достоверными.

Однако для реализации этой предпосылки необходимо высвободить методы РД вычислительных устройств из «плена» модели точных данных. Следует разрабатывать методы РД, которые признают за числом его право на существование в пределах допусков, т.е. право оставаться достоверным при возникновении несущественных ошибок. Методы РД должны уметь различать существенные и несущественные ошибки результатов обработки приближенных данных.

Заключение

Компьютерное представление данных ограничивается возможностью разрядной сетки отображать числа двоичными кодами или кодами, записанными в конечном алфавите, отличном от двоичного. Такие коды являются номерами элементов множеств и относятся к целым по своей природе, т.е. точным данным. В этих условиях действует модель точных

данных, в рамках которой все числа рассматриваются как точные, игнорируя их настоящую природу. Развитие компьютерной техники происходит в рамках данной модели. Это находится в противоречии с реалиями нашего мира, где все существует в пределах допусков – допустимых отклонений в пределах нормы, т.е. несущественных ошибок. По мере развития компьютерная техника структурируется под реалии нашего мира, что, в частности, выражается в увеличении доли обработки приближенных данных. Это ведет к усилению указанного противоречия. Под действием модели точных данных происходит подмена понятий, искажение цели и задач, решаемых отдельными направлениями, снижается эффективность методов и средств компьютерной техники.

Примером может служить РД вычислительных устройств, для которого в качестве цели рассматривается обнаружение неисправностей цифровых схем, в то время как настоящей целью, реализуемой на практике, является оценка достоверности вычисляемых результатов. Такая подмена цели вполне допустима в рамках модели точных данных, а за ее пределами объявленная и настоящая цели расходятся принципиально. Показатель достоверности современных методов РД, оцененный относительно объявленной цели, близок к единице, а согласно настоящей цели – близок к нулю. Большинство сегодняшних вычислительных задач могут решаться в рамках модели точных данных, однако, эти рамки не позволят решать сложные задачи завтра.

Литература

1. ANSI/IEEE Std 754-1985. IEEE Standard for Binary Floating-Point Arithmetic. – IEEE, New York, USA, 1985. – 18 с.
2. Демидович Б.П., Марон И.А. Основы вычислительной математики. – М.: Физматгиз, 1966. – 664 с.
3. Carter W., Schneider P. Design of Dynamically Checked Computers // Proc. IFIP Congress 68. – Edinburgh, Scotland. – 1968. – P. 878-883.
4. Anderson D.A., Metz G. Design of Totally Self-Checking Circuits for n-out-of-m Codes // IEEE Trans. on Computers. – 1973. – Vol. C-22. – P. 263-269.
5. Согомоян Е.С., Слабаков Е.В. Самопроверяемые устройства и отказоустойчивые системы. – М.: Радио и связь, 1989. – 208 с.
6. Favalli M., Metra S. Optimization of Error Detecting Codes for the Detection of Crosstalk Originated Errors // Proc. of IEEE Design, Automation and Test in Europe, Munich, Germany. – 2001. – P. 290-296.
7. Noufal I., Nicolaidis M. A CAD Framework for Generating Self-Checking Multipliers Based on Residue Codes // Proc. of IEEE Design, Automation and Test in Europe, Munich, Germany. – 1999. – P. 122-129.
8. Saposhnikov V., Dmitriev M., Goessel M, Saposhnikov V. Self-dual parity checking – a new method for on-line testing // Proc. IEEE VLSI Test Symp. – 1996. – P. 162-168.
9. Sparmann U., Reddy S.M. On the Effectiveness of Residue Code Checking for Parallel Two's Complement Multipliers // Proc. 24th Fault Tolerant Computing Symposium. – Austin (USA). – 1994. – P. 261-265.
10. Drozd A. On-line Testing of Computing Circuits at Approximate Data Processing // Radioelectronics & Informatics. – 2003. – № 3. – P. 113-116.
11. Drozd A., Lobachev M., Drozd J. The problem of on-line testing methods in approximate data processing // Proc. 12th IEEE International On-Line Testing Symposium, Como, Italy. – 2006. – P. 251-256.

Поступила в редакцию 20.02.2007

Рецензент: д-р техн. наук, проф. С.А. Нестеренко, Одесский национальный политехнический университет, Одесса.