Сборник докладов конференции «Технические и программные средства систем управления, контроля и измерения». М.: Институт проблем управления, 2008.

СЕРТИФИКАЦИЯ СИСТЕМ ЛОГИЧЕСКОГО УПРАВЛЕНИЯ НА ОСНОВЕ КОЛИЧЕСТВА ВЫПОЛНЯЕМЫХ ОПЕРАЦИЙ И ОБЪЕМА ОБРАБАТЫВАЕМЫХ ДАННЫХ

Выхованец В.С.
Институт проблем управления им. В.А. Трапезникова РАН, г. Москва
valery@vykhovanets.ru

Ключевые слова: логическое управление, дискретная обработка данных, метрики качества программ, сертификация систем управления.

Введение

Измерение качества программ является важной составляющей частью в процессе обеспечения качества систем логического управления [1].

Основным способом обеспечения и подтверждения качества программных средств предлагается считать использование процедур сертификации на соответствие стандартам серии ISO 9000, но этого явно не хватает для реального обеспечения качества. Практический опыт свидетельствует об относительной действенности сертификации [2].

Наибольшую трудность в процессе управления качеством программных средств вызывает установление, выявление причин (факторов), которые ухудшают те или иные характеристики (показатели) качества. Еще более сложной и в то же время необходимой процедурой является их оценка.

В настоящее время в мировой практике используется несколько сотен метрик программ, образующих систему измерений их качества. Эти измерения не могут быть проведены без субъективных оценок свойств программ, а если такое и случается, то оценка качества зависит от субъективной интерпретации получаемых данных [3].

Измерить параметры логической обработки данных, реализованной в виде программного средства – значит получить объективную основу для выработки оперативных и экономически эффективных управляющих воздействий. Поэтому актуальным является выбор и внедрение методов, обеспечивающих выявление конкретных характеристик качества и методов, позволяющих принимать обоснованные решения и осуществлять конкретные мероприятия.

Для устранения субъективности в оценке качества программ и устройств предлагается использовать верхние оценки трудоемкости вычислений для задач дискретной обработки данных конечной размерности, для чего служит разработанная методика абсолютной оценки эффективности дискретных моделей, которая применима не только при программной, но и для аппаратурной реализации логической обработки данных.

1. Качество систем управления

Качество системы управления – это совокупность свойств, определяющих ее полезность в соответствии с ее функциональным назначением и предъявленными к ней требованиями [4]. Характеристика качества – это показатель, отражающий отдельные факторы, влияющие на качество системы и поддающиеся измерению. Критерий качества – численный показатель, характеризующий степень, в которой системе присуще оцениваемое свойство. Критерий, как правило, должен [5]:

– численно характеризовать оцениваемую функцию системы;

– обеспечивать возможность определения затрат, необходимых для достижения требуемого уровня качества, а также степени влияния на показатель качества различных внешних факторов;

– быть по возможности простым, хорошо измеримым и иметь малую дисперсию.

Для измерения характеристик и критериев качества используют метрики. Метрика качества получается путем измерения характеристик качества. Эти измерения могут проводиться как на уровне критериев качества, так и на уровне отдельных характеристик качества. В первом случае система измерений позволяет непосредственно сравнивать различные системы по качеству. При этом сами измерения не могут быть проведены без субъективных оценок свойств системы. Во втором случае измерения характеристик можно выполнить объективно и достоверно, но оценка качества системы в целом будет связана с субъективной интерпретацией получаемых оценок. Таким образом, метрика – это комбинация конкретного метода измерения характериситки системы и шкалы измерения. Метрика определяет меру характеристики – переменную, которой присваивается значение в результате измерения.

В исследовании метрик различают два основных направления [6]:

– поиск метрик, характеризующих специфические свойства систем управления, т.е. метрик оценки свойств;

– использование метрик для оценки технических характеристик и факторов разработки систем управления, т.е. метрик оценки условий разработки.

По виду данных, получаемых при оценке качества систем управления метрики можно разбить на три группы [7]:

– метрики, оценивающие отклонение от нормы характеристик исходной системы (они устанавливают полноту заданных технических характеристик разработанной системы);

– метрики, позволяющие прогнозировать качество разрабатываемой системы (они заданы на множестве возможных вариантов решений поставленной задачи и определяют качество системы, которое будет достигнуто в итоге);

– метрики, по которым принимается решение о соответствии законченной системы заданным требованиям (они позволяют оценить соответствие разработанной системы необходимым требованиям).

Качественные оценки можно сгруппировать по шести направлениям [8]:

– оценки топологической, структурной и логической сложности системы;

– оценки надежности системы, позволяющие прогнозировать отказы;

– оценки производительности системы и повышения эффективности системы путем выявления ошибок проектирования;

– оценки уровня формальных средств, используемых для описания системы, и особенностей их применения;

– оценки трудности восприятия и понимания технической документации, ориентированные на психологические факторы, существенные для сопровождения и модернизации системы;

– оценки производительности труда разработчиков для прогнозирования сроков разработки и планирования соответствующих работ.

Существует множество различных метрик, которые могут представлять интерес для оценки качества систем управления. Критерии качества, в этом случае, могут включать следующие характеристики: экономичность, документированность, гибкость, модульность, надёжность, обоснованность, тестируемость, ясность, точность, модифицируемость, эффективность, легкость сопровождения и т.д. [9].

В современных системах управления большая доля технических средств приходится на программные средства. При оценке сложности программ, как правило, выделяют три основные группы метрик [10]: метрики размера программ, метрики сложности потока управления программ, метрики сложности потока данных программ. Традиционной характеристикой размера программ является количество строк исходного текста, где под строкой понимается любой оператор программы. Метрики сложности потока управления программ оценивают плотностью управляющих переходов внутри программ, либо взаимосвязями этих переходов. В том и в другом случае стало традиционным представление программ в виде управляющего ориентированного графа. Последняя группа метрик – метрики сложности потока данных, т.е. использования, конфигурации и размещения данных в программах. Здесь измеряют такие характеристики как число обращений к локальным и глобальным переменным, характер использования переменных из списков ввода-вывода, и др.

Современная программная индустрия накопила множество моделей и метрик, оценивающих отдельные производственные и эксплуатационные свойства программного обеспечения [11, 12]. Однако наличие субъективного фактора в оценке качества программ приводит к необходимости рассматривать их в комплексе и только так они могут служить отправной точкой для принятия объективных решений при сертификации программ [13], и построенных на их основе систем управления.

2. Верхняя оценка сложности дискретной обработки данных

Для синтеза формального описания системы логического управления дискретные функции декомпозируются и представляются в виде композиции функций меньшей размерности. Особенностью описываемого подхода является использование алгебраической декомпозиции дискретных функций в наиболее общей постановке задачи, когда с целью синтеза эффективных математических моделей дискретной обработки данных переменные и функции принимают значения на произвольных конечных множествах, а выбор алгебраических операций не ограничен каким-либо их подмножеством.

Пусть имеется некоторая спецификация системы логического управления. Представим эту спецификацию в виде таблицы, описывающей дискретную функцию (систему дискретных функций), возможно недоопределенную,

(1) $y = f(x_0, x_1, \ldots, x_{n-1})$,

где функция f имеет значность k_f, а переменные x_i – значности k_i ($i = \overline{0,\ n-1}$). Переход от исходной формы к табличному представлению осуществим путем определения значения функции при заданных значениях входных переменных.

Разложим функцию f в фундаментальной (конечном поле, целостном кольце) или аддитивной алгебре [14] по некоторой системе функций θ_i ($i = \overline{0, k' - 1}$),

(2) $$f(x', x'') = \sum_{i=0}^{M-1} \theta_i(x') \times a_i(x''),$$

где множество переменных функции $X = \{x_0, x_1, \ldots, x_{n-1}\}$ разеделено на два непересекающихся множества $X' = \{x'_0, x'_1, \ldots, x'_{n'-1}\}$ и $X'' = \{x''_0, x''_1, \ldots, x''_{n''-1}\}$ с числом переменных n' и n'', которые образовали обобщенные переменные x' и x'' со значностями k' и k'' соответственно,

$$k' = \prod_{(x_j \in X')} k_j = \prod_{j=0}^{n'-1} k'_j,\ k'' = \prod_{(x_j \in X'')} k_j = \prod_{j=0}^{n''-1} k''_j,\ k'k'' = m.$$

Сложность разложения определим количеством слагаемых M, $M \le k'$. Под сложностью представления L будем понимать количество операций, необходимых для вычисления функции по формуле (2), $L \le nM$. Не нарушая общности рассмотрим бесповторную бесскобочную аналитическую конструкцию функций θ_i, заданную с точностью до порядка вхождения переменных,

(3) $\theta(x') = x'_0 \circ_0 x'_1 \circ_1 x'_2 \circ_2 \ldots \circ_{n'-2} x'_{n'-1}$,

где $\circ_j$ – произвольные бинарные операции. Подсчитаем количество различных функций N_θ, порождаемых конструкцией (3),

$$N_\theta(k, k'_0, k'_1, \ldots, k'_{n'-1}) = k^{k'_0 k'_1 + (k-\alpha)\sum_{i=2}^{n'-1} k'_i} = k^{(k-\alpha)(\sum_{i=2}^{n'-1} k'_i - \beta)}, \ \beta = k'_0 + k'_1 - \frac{k'_0 k'_1}{k-\alpha},$$

где α – порождающая способность аналитической конструкции, $0 \le \alpha < 1$. Зависимость порождающей способности α от количества переменных n показана на рис. 1 [15].

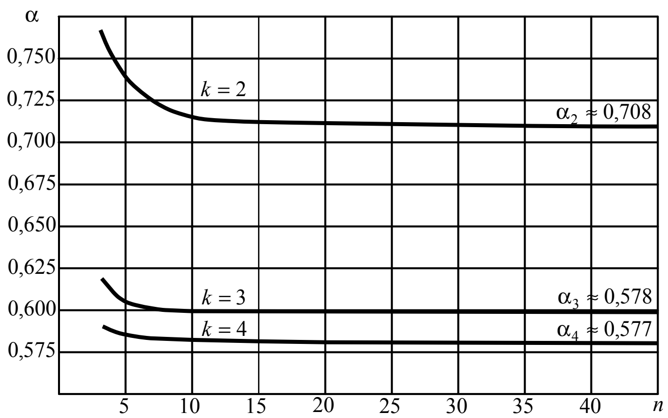

Рис. 1. Порождающая способность бесповторных формул

Для оценки M предварительно иследуем спектральное представление

$$f(x) = \sum_{i=0}^{M-1} \theta_i(x) \times a_i ,$$

где a_i – константы, или спектральные коэффициенты. Для порождения всех функций необходимо, чтобы

$$k^m = \sum_{i=1}^{M} C^i_{N_\theta} (k-1)^i , \ m = \prod_{i=0}^{n-1} k_i ,$$

где C^q_p – число сочетаний из p элементов по q. Отсюда находим оценку сложности спектрального представления,

(4) $$M = \frac{1}{k-\alpha} \frac{\prod_{i=0}^{n-1} k_i}{\sum_{i=0}^{n-1} k_i - \beta} .$$

Теперь найдем сложность представления функции (4) при произвольном разделении переменных. Для этого в качестве первого приближения системы функций $\{\theta_i\}$ выберем столбцы прямоугольной матрицы функции размерности $k' \times k''$, порожденной некоторым разделением переменных на два непересекающихся множества (рис. 2).

Для формирования $\{\theta_i\}$ возможно использование различных линейных комбинаций столбцов матрицы функции. Количество таких комбинаций равно $k^{k''} - 1$, что при $k'' \approx k'$ позволяет сформировать практически произвольную систему функций, в том числе и представимую бесповторными бесскобочными формулами с наименьшим числом операций. При уменьшении k'' количество слагаемых в разложении уменьшается, а в пределе имеем спектральное разложение. Однако, в общем случае, разлагаемая функция не имеет бесповторного представления. Следовательно, существует некоторое минимальное k'', при котором система $\{\theta_i\}$ все еще представима бесповторными формулами.

$$\left[\begin{array}{cccc:cc} \theta_0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 \\ \theta_0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 \\ \theta_0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 \\ \theta_0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 \\ \theta_0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 \\ \theta_0 & \theta_1 & \theta_2 & \theta_3 & 0 & 0 \end{array}\right] \times \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hdashline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{cccc} \theta_0 & \theta_1 & \theta_2 & \theta_3 \\ \theta_0 & \theta_1 & \theta_2' & \theta_3 \\ \theta_0 & \theta_1 & \theta_2' & \theta_3 \\ \theta_0 & \theta_1 & \theta_2' & \theta_3 \\ \theta_0 & \theta_1 & \theta_2' & \theta_3 \\ \theta_0 & \theta_1 & \theta_2' & \theta_3 \end{array}\right]$$

Рис. 2. Разделительная декомпозиция

Так как при достаточно большом $k'' \le k'$ функции $\{\theta_i\}$ могут быть, фактически, произвольными, представим их как спектральное разложение некоторой неизвестной функции. Тогда из (4) следует условие, при котором гарантируется представимость системы из k'' функций длины k' бесповторными формулами, а сложность разложения M станет равна k'',

$$(5)\quad \prod_{j=0}^{n''-1} k_j'' = \frac{1}{k-\alpha'} \frac{\prod_{i=0}^{n'-1} k_i'}{\sum_{i=0}^{n'-1} k_i' - \beta'}.$$

Подвергнем функции a_i, в свою очередь, спектральному разложению. Тогда из (4) и (5) получаем

$$(6)\quad M = \frac{1}{(k-\alpha')(k-\alpha'')} \frac{\prod_{i=0}^{n-1} k_i}{(\sum_{i=0}^{n'-1} k_i' - \beta')(\sum_{j=0}^{n''-1} k_j'' - \beta'')}.$$

На следующем шаге найдем разделение переменных, которое обеспечивает наименьшую сложность разложения, полагая $\alpha' \approx \alpha''$ и $\beta' \approx \beta''$. Для этого сформулируем задачу поиска условного экстремума (6):

$$(7)\quad \begin{cases} s = \sum_{i=0}^{n'-1} k_i' + \sum_{j=0}^{n''-1} k_j'', \\ m = (k-\alpha)\Big(\sum_{i=0}^{n'-1} k_i' - \beta\Big) \prod_{j=0}^{n''-1} k_j''^2, \\ (\sum_{i=0}^{n'-1} k_i' - \beta)(\sum_{j=0}^{n''-1} k_j'' - \beta) \to \max; \end{cases}$$

где s и m – константы, первое и второе тождество задает зависимость переменных, а последнее выражение определяет целевую функцию. Решение (7) методом неопределенных множителей Лагранжа выявляет следующую связь между значностями переменных,

$$\sum_{i=0}^{n'-1} k_i' - \sum_{j=0}^{n''-1} k_j'' = \mu(2n''-1)$$

где μ – некоторая константа, не зависящая от разделения переменных. Полагая $\mu(2n''-1)=\rho n$, имеем

$$(8)\quad M=\frac{4}{(k-\alpha)^2}\frac{\prod_{i=0}^{n-1}k_i}{(\sum_{i=0}^{n-1}k_i-2\beta)^2-\rho^2n^2},\ \rho=\frac{1}{n}\Big(\sum_{i=0}^{n'-1}k_i'-\sum_{j=0}^{n''-1}k_j''\Big)$$

где ρ – параметр, характеризующий разделение переменных, для вычисления оптимального значения которого воспользуемся условием (5), заданным для переменных со средней арифметической и средней геометрической значностью $\bar{k}$ и $\tilde{k}$,

$$(9)\quad \rho n=\bar{k}\log_{\tilde{k}}\left(\frac{1}{2}(k-\alpha)(\bar{k}n-2\beta+\rho n)\right).$$

Результаты решения трансцендентного уравнения (9) при одинаковых значностях переменных, равных k, приведены на рис. 3.

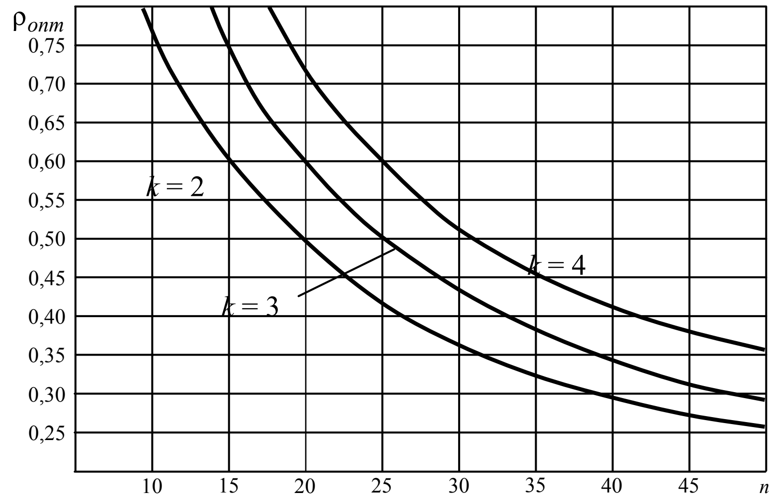

Рис. 3. Оптимальное разделение переменных

В итоге из (8) получаем верхнюю оценку сложности представления дискретной функции (1) в виде формулы в базисе произвольных дискретных операций [16]:

$$(10)\quad L=\frac{4}{n(k-\alpha)^2}\frac{\prod_{i=0}^{n-1}k_i}{\frac{1}{n^2}(\sum_{i=0}^{n-1}k_i-2\beta)^2-\rho^2},$$

где L – количество дискретных операций ($L=nM$), k – значность используемых бинарных операций ($k\ge k_f$, $k\ge k_i$), α – порождающая способность бесповторной аналитической конструкции ($0\le\alpha<1$), β – начальная порождающая способность ($\beta\approx k$), ρ – оптимальность разделения переменных ($0<\rho<0{,}8$). Оценка (10) определяет количество операций, которых достаточно для реализации (вычисления) произвольной функции вида (1).

3. Методика оценки абсолютной эффективности

Формула (10) позволяет по объему обрабатываемых данных и выполненному при этом количеству операций L' оценить эффективность произвольной дискретной обработки данных, а в том случае, если количество операций L' будет меньше, чем L, то определить и степень минимизации выполненной реализации дискретной обработки данных, равную отношению L к L'.

Пусть имеется вычислительное средство с разрядностью обрабатываемых элементов данных, равной r двоичных разрядов. Последнее означает, что максимальная значность переменных k, выполняемой этим устройством дискретной обработки данных, равна 2^r, $k = 2^r$.

Пусть на вход устройства подаются данные, объемом D_x двоичных разрядов, а с выхода устройства получаю данные, объемом D_y двоичных разрядов. В этом случае может быть вычислено количество переменных n, от которых зависит получаемый результат, и количество дискретных функций s, вычисляемых устройством,

(11) $n = \dfrac{D_x}{r}$, $s = \dfrac{D_y}{r}$; или $D_x = nr$, $D_y = sr$.

Максимально необходимое количество операций L для дискретной обработки данных, выраженной функциями значности k от n переменных той же значности, очевидно,

(12) $$L = \frac{4s}{n(k-\alpha)^2}\frac{2^{D_x}}{\frac{1}{n^2}(n2^r - 2\beta)^2 - \rho^2} \approx \frac{4s}{n(k-\alpha)^2}\frac{2^{D_x}}{2^{2r}} = \frac{4s}{n(k-\alpha)^2}2^{D_x-2r}.$$

Выражая s и n в соответствии с (11) и пренебрегая слагаемыми α, β и ρ, которые значительно меньше k, из (12) получим итоговое выражение для максимального количества операций L, которое требуется выполнить для реализации произвольной дискретной обработки данных,

(13) $$L \approx 4\frac{D_y}{D_x}2^{D_x-4r}.$$

Из (13) следует, что для вычислительного средства с разрядностью r эффективность обработки данных E при достаточно большом их объеме $D_x >> r$ может быть подсчитана по следующей формуле,

(14) $$E = \frac{L}{L'} \approx \frac{1}{L'}\frac{D_y}{D_x}2^{D_x-4r+2},$$

где L (L') – максимально необходимое (реальное) количество операций в программе или количество элементов в устройстве обработки данных, D_x (D_y) – объем входных (выходных) данных в битах, r – разрядность процессора или разрядность логических элементов в битах, E – достигнутая эффективность обработки данных.

Если в результате измерения количества операций L', которые выполняются системой логического управления для вычисления результата объемом D_y при обработке входных данных объемом D_x окажется, что эффективность реализации E, вычисленная по формуле (14), меньше единицы ($E < 1$), то это означает, что имеет место некачественная реализации логического управления. При $E = 1$ имеем удовлетворительную реализацию в том смысле, что почти все законы логического управления требуют такого объема операций. Если $E > 1$, то получена минимизированная реализация системы логического управления, что, вероятнее всего, свидетельствует об избыточности входных данных, так как число дискретных функций, имеющих такую реализацию, пренебрежимо мало [15].

Заключение

На основе верхних оценок сложности представления дискретных функций разработана методика оценки абсолютной эффективности и степени минимизации дискретной обработки данных путем определения объемов обрабатываемых данных и количества выполняемых при этом операций. Последнее позволяет ввести в использование новую метрику качества систем логического управления, повышающую объективность сертификации такого рода систем.

Литература

1. *Амбарцумян А.А., Искра С.А.* Технология проектирования ПТК систем безопасности, ориентированная на корректность // Матер. Межд. конф. «Автоматизация проектирования дискретных систем». Минск, 1997. С. 34-40.
2. *Voas J.* The Software Quality Certification Triangle // Crosstalk. 1998, No 11. PP. 12-14.
3. *Goodman P.* Software metrics: Best practices for successful IT management. Brookfield: Rothstein Associates Inc., 2004. 264 p.
4. ISO 9001:1994 Quality systems – Model for quality assurance in design, development, production, installation and servicing.
5. *Jones C.* Applied Software Measurement. Global Analisys of Productivity and Quality. New York: McGraw-Hill, 2008. 696 p.
6. *Pandian C.R.* Software Metrics – A guide to Planning, Analysis and Application. London: Auerbach Publications, 2004. 312 p.
7. *Kan S.Y.* Metrics and Models in Software Quality Engineering. Addison Wesley, 2002. 560 p.
8. Best practices in software measurement / Ebert C., Dumke R., Bundschuh M., Schmietendorf A. Berlin: Springer, 2005. 299 p.
9. *Липаев В.В.* Программная инженерия. Методологические основы. М.: ТЕИС, 2007. 609 с.
10. *Андон Ф.И., Коваль Г.И., Коротун Т.М., Суслов В.Ю.* Основы инженерии качества программных систем. Киев: Академпериодика, 2007. 670 с.
11. *Laird L.M., Brennan V.C.* Software measurement and estimation. A practical approach. Hoboken: Wiley & Sons, 2006. 276 p.
12. *Mubson J.C.* Software engineering measurement. New York: CRC Press LLC, 2003. 443 p.
13. *Ebert C., Dumke R.* Software measurement. Establish, Extract, Evaluate, Execute. Berlin: Springer-Verlag, 2007. 568 p.
14. *Выхованец В.С.* Спектральные методы в логической обработке данных // Автоматика и телемеханика. 2001. № 10. С. 28-53.
15. *Выхованец В.С.* Алгебраическая декомпозиция дискретных функций // Автоматика и телемеханика. 2006. № 3. С. 20-53.
16. *Выхованец В.С.* Оптимальный синтез логического управления // Тезисы докладов 3-ой Международной конференции по проблемам управления. М., 2006. Т. 2. С. 105.