УДК 004.4

ОБОБЩЕНИЕ ОЦЕНКИ СЛОЖНОСТИ ПРОГРАММНЫХ СРЕДСТВ

В.В. БАХТИЗИН, А.А. КУЗИКОВ

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 28 октября 2009

Разработана модель метрик сложности, объединяющая различные методики оценки сложности с целью сопоставления результатов измерения. Предложен математический аппарат для подготовки метрик к совместной оценке сложности. Предоставлен инструмент оценки качества преобразований метрик после проведения такой подготовки.

Ключевые слова: качество программного обеспечения, метрики сложности.

Введение

Совершенствование информационных технологий и, в частности, автоматизация бизнес-процессов с целью повышения эффективности работы и снижения производственных затрат привела к широкому применению программного обеспечения во многих сферах деятельности. Высокий спрос на системы автоматизации и управления производством вызвал развитие рынка программных средств (ПС). Результатом этого стало появление ряда программных продуктов, предназначенных для решения сходных задач. Поэтому на сегодняшний день успех и конкурентоспособность ПС всецело зависят от их качества.

Характерной особенностью современных ПС является высокая сложность, которая увеличивает время разработки программных продуктов, приводит к срыву графиков выполнения работ, нередко является причиной переноса выпуска очередной версии на более поздний срок, увеличивает количество ошибок и значительно затрудняет сопровождение. Таким образом, сложность, оказывая непосредственное влияние на надежность, сопровождаемость, эффективность, определяет качество ПС, регламентируемое стандартом ISO/IEC 9126-1 [1, 2]. Сложность оценивается рядом метрик качества, тем не менее, не являясь отдельной характеристикой ПС. Ввиду отсутствия единого, четко формализованного определения понятия сложности [3], для объективной численной оценки различных ее аспектов было разработано множество метрик, которые в совокупности могли бы сформировать представление о сложности в целом. Однако они оперируют разными свойствами ПС. Поэтому вычисленные значения метрик часто несравнимы, что препятствует выработке результирующего заключения о сложности.

Основные определения

Методика измерения сложности ПС в общем случае основывается на расчете значений метрик сложности. В вычислении значения такой метрики могут участвовать величины как однотипные (операторы модулей ПС), так и достаточно разнородные (глобальные переменные модулей и вершины и дуги управляющего графа ПС). Различные метрики сложности по-разному учитывают особенности анализируемого ПС. Поэтому наиболее полную картину о сложности ПС можно получить путем расчета для него различных метрик сложности и сопоставления результатов измерений. Для того, чтобы сопоставить значения разных метрик сложности, необходимо определить их сходные компоненты и разработать процедуру обобщения.

Определение 1. Множество D_k однотипных величин вида k, единообразно характеризующих какой-либо аспект ПС на протяжении его жизненного цикла, называется *измерением ПС*.

При выделении измерений ПС для каждой метрики сложности из набора метрик, сформированного для проведения оценки, следует руководствоваться процедурой выделения измерений ПС, которая включает следующую последовательность шагов:

1) выделить свойства ПС, задействованные в расчете конкретной метрики;

2) поставить в соответствие каждому уникальному свойству отдельное измерение;

3) руководствуясь принципом наименьшей зависимости измерений, ограничить состав выделенных в шаге 2 измерений ПС наиболее общими и мало зависимыми измерениями. Например, измерения "операторы" и "предикаты" во многих языках программирования сводимы к общему измерению "операторы";

4) при объединении списков измерений, полученных для заданного набора метрик сложности, необходимо исключить измерения-дубликаты.

Определение 2. Множество всех измерений D_k ПС называется полным семейством измерений R_B. При этом

$$R_B = \bigcup_{k=1}^{N} \{D_k\}, \tag{1}$$

где N — количество измерений ПС.

В соответствии с (1) любое семейство измерений R является подмножеством полного семейства измерений R_B. Таким образом, если I_N — некоторое подмножество индексов $I_N \subseteq \{1, 2, \ldots, N\}$, то соответствующее семейство R задается выражением

$$R = \bigcup_{k \in I_N} D_k \quad . \tag{2}$$

Определение 3. Декартово произведение всех измерений ПС, входящих в состав некоторого семейства измерений R, называется *пространством измерений X* для данного семейства R. При этом "R порождает X" и, учитывая (2),

$$\begin{pmatrix} R \to X \\ R = \bigcup\limits_{k \in I_N} D_k \end{pmatrix} \overset{def}{\Leftrightarrow} \left(X = \prod_{k \in I_N} D_k \right). \tag{3}$$

В формировании любой метрики сложности μ участвуют величины, принадлежащие хотя бы одному измерению из полного семейства измерений R_B. Таким образом, метрика μ имеет областью определения некоторое пространство измерений X, то есть $E(\mu)=X$. Пространство X порождается семейством измерений R (подмножеством полного семейства измерений R_B) и определяется в соответствии с (3).

Для определенности будем считать, что порядок измерений, формирующих семейство (и, соответственно, порядок компонентов порожденного пространства), неизменен для однажды выбранного набора метрик. Все последующие измерения, включаемые в полное семейство измерений при расширении первоначального набора метрик, помещаются в конец семейства.

Индекс пространства (семейства) есть десятичное представление двоичного кода. При этом, номер измерения соответствует номеру бита в таком индексе при нумерации битов справа налево, а значение бита фиксирует наличие (единица) либо отсутствие (ноль) соответствующего измерения D_k в данном пространстве (семействе).

Определение 4. Функция $n(i)$ индекса i, который единственным образом определяет наличие либо отсутствие каждого измерения в структуре пространства (семейства), ставит в соответствие своему аргументу размерность i-го пространства X_i (семейства R_i) и называется *функцией индекса*.

Необходимо отметить, что равенство значений функций индекса $n(i)$ и $n(j)$ не обусловливает совпадение соответствующих пространств X_i и X_j (семейств R_i и R_j). Так, например, n(2)=n(4)=1, однако $(010_2=2_{10})\neq(100_2=4_{10})$. Этот факт необходимо учитывать при сопоставлении различных метрик, имеющих одинаковую размерность пространств.

Таким образом, каждая метрика сложности μ, определенная на некотором пространстве измерений X_i в соответствии с формулой (3), устанавливает размерность данного пространства по правилу соответствия n(i), согласно определению 4.

Определение 5. Величины, участвующие в вычислении значения метрики μ, которая определена на пространстве измерений X_i, и сгруппированные по соответствующим измерениям D_k пространства X_i формируют n(i)-мерную точку пространства X_i. Множество значений координат такой точки определяется программой (модулем) P. Характеристикой данной точки служит n(i)-мерный вектор $\bar{\mathbf{x}}^{\mathrm{n}(i)}(P)$ пространства X_i, компоненты которого содержат координаты данной точки, то есть

$$\begin{pmatrix} \mathrm{E}(\mu) = X_i \\ X_i = \prod\limits_{k\in \mathrm{I_N}} D_k \end{pmatrix} \overset{def}{\Leftrightarrow} \begin{pmatrix} \exists \bar{\mathbf{x}}^{\mathrm{n}(i)}(P) = \overrightarrow{(x_1,\ldots,x_j,\ldots,x_{\mathrm{n}(i)})}, \\ \exists\alpha, \exists\beta, \forall k, \quad (\alpha \le k \le \beta), \quad (\alpha, k, \beta) \subseteq I_N, \\ x_1 \in D_\alpha,\ldots,x_j \in D_k,\ldots,x_{\mathrm{n}(i)} \in D_\beta \end{pmatrix}. \tag{4}$$

Порядок следования компонентов вектора $\bar{\mathbf{x}}^{\mathrm{n}(i)}(P)$ во множестве координат $\{x_1,\ldots,x_{\mathrm{n}(i)}\}$ соответствует порядку измерений D_k, формирующих семейство R_i и соответствующее порожденное пространство X_i.

Согласно определению 5, а также основываясь на том, что метрика задает правило вычисления своего значения, метрику μ можно выразить следующим образом:

$$\mu = \mathrm{C}\left(\bar{\mathbf{x}}^{\mathrm{n}(i)}(P)\right), \tag{5}$$

где С — правило связи координат вектора $\bar{\mathbf{x}}^{\mathrm{n}(i)}(P)$ для вычисления значения метрики.

Вследствие особенностей выделения измерений ПС, функция С в (5) может не совпадать по структуре с формулой вычисления метрики сложности. Поэтому передаваемые в формулу значения параметров могут отличаться от переданных в функцию С компонентов вектора.

Преобразование координат характеристических векторов

Как следует из (5), задача сопоставления значений различных метрик сложности сводима к задаче преобразования координат характеристических векторов. Для того, чтобы сопоставить значение метрики μ_1, заданной вектором $\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)$, со значением метрики μ_2, заданной вектором $\bar{\mathbf{x}}_2^{\mathrm{n}(j)}(P)$, в общем случае необходимо совершить операцию преобразования координат вектора $\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)$ из пространства X_i в координаты результирующего вектора $\bar{\mathbf{y}}_1^{\mathrm{n}(j)}$ из пространства X_j:

$$\mathrm{T}: \quad \bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P) \to \bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P)$$

или

$$\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P) = \mathrm{T}\left(\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)\right), \tag{6}$$

где Т — правило преобразование координат.

Определение 6. Пространство измерений X_C является общим для метрик μ_1 и μ_2, определенных на пространствах X_i и X_j, соответственно, если справедлива следующая система выражений:

$$\begin{cases} R_i \to X_i, \; R_j \to X_j \\ R_C = R_i \bigcap R_j, \; X_C = \prod\limits_{\forall D_k \in R_C} D_k \end{cases}. \tag{7}$$

Определение 7. Преобразование координат (6) вектора $\bar{\mathbf{x}}_1^{n(i)}(P)$ из пространства X_i в целевое пространство X_j называется *однозначным* и обозначается $\mathrm{T_U}$, если каждая координата вектора $\bar{\mathbf{y}}_1^{n(j)}(P)$, принадлежащая конкретному измерению D_k пространства X_j, определяется по некоторому правилу координатой вектора $\bar{\mathbf{x}}_1^{n(i)}(P)$ из пространства X_i, принадлежащей тому же измерению:

$$\begin{pmatrix} \bar{\mathbf{x}}_1^{n(i)}(P) \in X_i \\ \bar{\mathbf{y}}_1^{n(j)}(P) \in X_j \\ \bar{\mathbf{y}}_1^{n(j)}(P) = \mathrm{T_U}\left(\bar{\mathbf{x}}_1^{n(i)}(P)\right) \end{pmatrix} \overset{def}{\Leftrightarrow} \begin{pmatrix} \exists k \in I_N, \; \exists m \in \left[1, \mathrm{n}(i)\right], \; \exists \mathrm{g}_{lm} \\ (x_{1,m}, y_{1,l}) \subseteq D_k, \; \forall l \in \left[1, \mathrm{n}(j)\right] \\ y_{1,l} = \mathrm{g}_{lm}(x_{1,m}) \end{pmatrix}, \tag{8}$$

где g_{lm} — правило зависимости (функция уточнения) l-й координаты вектора $\bar{\mathbf{y}}_1^{n(j)}(P)$ от m-й координаты вектора $\bar{\mathbf{x}}_1^{n(i)}(P)$.

В матричном виде формулу (8) можно записать следующим образом:

$$\bar{\mathbf{y}}_1^{n(j)}(P) = \mathbf{G}\left(\bar{\mathbf{x}}_1^{n(i)}(P)\right), \tag{9}$$

где $\mathbf{G}$ — матрица размерности $\mathrm{n}(j)\times\mathrm{n}(i)$ уточняющих функций g_{lm}. С учетом (8) компоненты вектора $\bar{\mathbf{y}}_1^{n(j)}(P)$, рассчитанного по формуле (9), можно записать так:

$$y_{1,l} = \begin{cases} \mathrm{g}_{lm}(x_{1,m}), & \exists k \in I_N, \; (x_{1,m}, y_{1,l}) \subseteq D_k \\ \gamma_l, & \forall k \in I_N, \; (x_{1,m}, y_{1,l}) \not\subset D_k \end{cases}, \tag{10}$$

где γ_l — предопределенные константы, которые указывают на отсутствие зависимости между координатами, принадлежащими разным измерениям. Например, можно полагать $\gamma_l=0$, если элементы соответствующего измерения D_k — скаляры, и $\gamma_l=\varnothing$ в случае элементов-множеств.

Следствие. Преобразование вектора $\bar{\mathbf{x}}_1^{n(i)}(P)$ из исходного пространства измерений X_i в конечное X_j является однозначным, если общее пространство совпадает с конечным и все измерения, формирующие оба пространства, независимы:

$$\begin{pmatrix} R_i \to X_i, \; \bar{\mathbf{x}}_1^{n(i)}(P) \in X_i \\ R_j \to X_j, \; \bar{\mathbf{y}}_1^{n(j)}(P) \in X_j \\ (R_i \bigcap R_j) \to X_j \end{pmatrix} \Rightarrow \left(\exists \mathrm{T_U}, \; \bar{\mathbf{y}}_1^{n(j)}(P) = \mathrm{T_U}\left(\bar{\mathbf{x}}_1^{n(i)}(P)\right)\right). \tag{11}$$

Действительно, совпадение общего пространства измерений с конечным $(R_i \bigcap R_j) \to X_j$ ()$\to X_j$ предполагает, что $R_i \bigcap R_j = R_j$, то есть $R_j \subseteq R_i$. Следовательно, в случае невырожденного семейства R_j имеется непустое множество общих измерений D_k таких, что: $R_i \supseteq \prod\limits_{\forall D_k \in (R_i \cap R_j)} D_k = R_j$.

Независимость измерений внутри каждого семейства гарантирует, что каждый компонент, принадлежащий некоторому измерению D_k из семейства R_j, зависит только от компонента, принадлежащего тому же измерению, но из семейства R_i, то есть

$$y_{1,l} = \mathrm{g}_{lm}(x_{1,m}), \quad \forall l \in \left[1, \mathrm{n}(j)\right].$$

Учитывая тот факт, что множество общих измерений D_k повторяет структуру семейства R_j, то по определению 7 преобразование (6) является однозначным, то есть $\mathrm{T}=\mathrm{T}_\mathrm{U}$. Следовательно, формулировка (11) справедлива.

Рассмотрим преобразование (6) в случае невырожденных семейств, когда $R_i \bigcap R_j=\varnothing$. В соответствии с процедурой выделения измерений ПС рекомендуется сформировать набор независимых измерений для всего набора привлеченных к оценке метрик. Тем не менее, не исключена вероятность формирования семейств, в рамках которых измерения независимы, однако, это условие не выполняется для полного семейства в целом либо для вновь включаемых в набор метрик, ввиду выявления новых взаимосвязей между измерениями. В таком случае, каждая координата вектора $\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P)$, принадлежащая конкретному измерению из пространства X_j, будет определяться набором координат вектора $\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)$ из пространства X_i в соответствии с некоторым правилом. Такое преобразование называется функциональным и обозначается T_F. Ввиду независимости измерений в рамках отдельного семейства (например, X_i) можно ограничиться линейной комбинацией функций координат вектора $\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)$ при вычислении каждой координаты вектора $\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P)$.

$$\left(\begin{array}{l} \bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P) \in X_i \\ \bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P) \in X_j \\ \bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P) = \mathrm{T}_F\left(\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)\right) \end{array}\right) \overset{def}{\Leftrightarrow} \left(\begin{array}{l} \forall l = \left[1, n(j)\right],\ \exists \mathrm{f}_{lm} \\ y_{1,l} = \mathop{\mathrm{S}}\limits_{m=1}^{\mathrm{n}(i)} \mathrm{f}_{lm}(x_{1,m}) \\ (x_{1,m}, y_{1,l}) \not\subset D_k,\ \forall k \in I_N \end{array}\right), \tag{12}$$

где f_{lm} — правило зависимости l-й координаты вектора $\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P)$ от m-й координаты вектора $\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)$; $\mathbf{S}$ — оператор связи координат. Например, в качестве $\mathbf{S}$ может применяться Σ в случае скалярных $x_{1,m}$ или $\bigcup$, если $x_{1,m}$ — множества.

В матричном виде формула (12) примет вид

$$\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P) = \mathbf{F}\left(\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)\right)^{\mathrm{T}}, \tag{13}$$

где $\mathbf{F}$ — матрица размерности $\mathrm{n}(j)\times\mathrm{n}(i)$ функций f_{lm}.

В общем случае, пространства X_i и X_j могут иметь общие измерения, но соответствующие семейства R_i и R_j не являются подмножествами друг друга. Тогда, учитывая выражения (9) и (13) и возможное нарушение условия независимости измерений в рамках полного семейства измерений, преобразование (6) можно записать в виде комбинации однозначных и функциональных преобразований

$$\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P) = \mathrm{S}_1\left(\mathbf{G}\left(\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)\right)^{\mathrm{T}}, \mathbf{F}\left(\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)\right)^{\mathrm{T}}\right). \tag{14}$$

Функция S_1 играет роль оператора связи координат $\mathbf{S}$ из (12). В общем случае, операции над координатами вектора $\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)$ из (14) соответствуют преобразованиям (10) и (12):

$$y_{1,l} = \begin{cases} \mathrm{g}_{lm}(x_{1,m}), & \exists k \in I_N,\ (x_{1,m}, y_{1,l}) \subseteq D_k \\ \mathop{\mathrm{S}}\limits_{m=1}^{\mathrm{n}(i)} \mathrm{f}_{lm}(x_{1,m}), & \forall k \in I_N,\ (x_{1,m}, y_{1,l}) \not\subset D_k \end{cases}.$$

В результате преобразования (14) координат вектора метрики $\mu_1 = \mathrm{C}_1\left(\bar{\mathbf{x}}_1^{\mathrm{n}(i)}(P)\right)$ из пространства X_i в пространство X_j будет получена метрика $\mu_1' = \mathrm{C}_2\left(\bar{\mathbf{y}}_1^{\mathrm{n}(j)}(P)\right)$ (аналог метрик μ_1 для пространства X_j), значение которой можно сопоставлять со значением метрики $\mu_2 = C_2\left(x_2^{n(j)}(P)\right)$.

Оценка однозначности преобразования координат

Если задача определения зависимостей **F** трудноразрешима, то в качестве значений от действия матрицы **F** на компоненты характеристического вектора в (12) допустимо использовать некоторое предопределенное значение (константу), то есть $\sum_{m=1}^{n(i)} f_{lm}\left(x_{1,m}\right) = \gamma_l$. Однако подобное решение вносит неоднозначность в преобразования (13) и (14). Для оценки однозначности преобразования координат характеристического вектора $\bar{\mathbf{x}}_1^{n(i)}(P)$ при переходе из пространства X_i в пространство X_j можно воспользоваться *метрикой качества преобразования координат* $\xi_{X_i \to X_j}$:

$$\xi_{X_i \to X_j} = \frac{n(i \otimes j)}{n(i \otimes j) + n(i \otimes j \oplus j)}, \qquad (15)$$

где $n(i \otimes j)$ — размерность пространства, общего для пространств X_i и X_j; $n(i \otimes j \oplus j)$ — размерность пространства, образованного измерениями из пространства X_j, отсутствующими в пространстве X_i; $\otimes$ — операция поразрядного двоичного умножения; $\oplus$ — операция поразрядного сложения по модулю два.

Формула (15) показывает, что чем большее число общих измерений имеют пространства X_i и X_j, то тем выше однозначность преобразования координат и с тем большей достоверностью можно сопоставлять значения метрик, определенных на данных пространствах.

Заключение

Приведенную в работе модель можно применять для обобщения различных типов метрик сложности с целью сопоставления их значений и получения более достоверной информации о сложности анализируемого ПС. Модель предоставляет аппарат для преобразования метрик и позволяет произвести априорную оценку допустимости и точности такого преобразования посредством метрики качества преобразования координат. Результаты оценки сложности можно использовать для управления качеством разрабатываемого ПС.

GENERALISING A SOFTWARE COMPLEXITY EVALUATION

V.V. BAKHTIZIN, A.A. KUZIKAU

Abstract

A complexity metrics model generalising different complexity metrics for the purpose of collating measurement results has been developed. A mathematical apparatus to prepare the metrics for a joint complexity evaluation in practice is presented. Means of a metric conversion quality evaluation is provided, once the metrics' preparation phase has been completed.

Литература

1. Разработка программного обеспечения. Качество продукции. Часть 1. Модель качества: ISO/IEC 9126-1:2001. — ISO/IEC JTC 1/SC 7 Разработка систем и программного обеспечения.
2. *Бахтизин В.В., Глухова Л.А.* Стандартизация и сертификация программного обеспечения. Минск, 2006.
3. *Lloyd S.* Programming the Universe. NY, 2006.