

ПРАВИЛА ДЕКОМПОЗИЦІЇ ПРИ ФОРМУВАННІ МОДЕЛІ ПРЕДМЕТНОЇ ОБЛАСТІ ДЛЯ НАВЧАЛЬНИХ КУРСІВ

В. В. Любченко

Кандидат технічних наук, доцент
Кафедри системного програмного забезпечення
Одеський національний політехнічний університет
пр. Шевченка, 1, м. Одеса, Україна, 65044
Контактний тел.: (048) 734-8566
e-mail: vira.lyubchenko@gmail.com

Розглянута задача визначення правила зупинки процесу декомпозиції для формування предметних областей навчальних курсів. Сформульовані і описані два підходи: базований на використанні числа Міллера і базований на властивостях розподілу значень міри асоціативного зв'язку

1. Вступ

При проектуванні різноманітних інформаційних систем, зокрема гіпермедіа систем та систем підтримки навчального процесу, виникає потреба в створенні моделі предметної області. Одним з підходів до системного будівництва такої моделі є рекурсивна декомпозиція поняття предметної області. Очевидно, що зі збільшенням глибини декомпозиції зростає інформативність одержуваної структури, але при цьому зростає її складність. Можна припустити, що на певному етапі продовження виконання декомпозиції стане нераціональним, тому що інформація про предметну область стає більш докладною, ніж це необхідно. Це призведе до одержання значної складності дерева декомпозиції та збільшення трудомісткості її виконання. До того ж надлишкова докладність знижує наочність моделі, ускладнює аналіз структури курсу та істотно ускладнює подальші його модифікації. Тобто має існувати певний оптимальний з погляду співвідношення показників інформативність-складність рівень деталізації, при досягненні якого стає недоцільним продовження рекурсивного виконання декомпозиції.

Якщо як модель декомпозиції взяти дерево, то складність декомпозиції можна оцінити, як складність дерева

$$Cplx = \sum_{v_i \in V} cplx(v_i), \quad (1)$$

де V — множина вершин дерева декомпозиції. При цьому складність вершини $cplx(v_i)$ визначається як добуток чис-

ла ребер, які входять до цієї вершини, та числа вершин, які належать її кущу [1]. Можемо також ввести в розгляд відносний показник — питому складність дерева:

$$sCplx = \frac{Cplx}{|V|}, \quad (2)$$

Проте, навіть при переході до відносної оцінки, вибір граничного значення, що його можна використати як критерій зупинки рекурсивної декомпозиції, уніфікований бути не може, оскільки функція оцінки монотонно зростає зі збільшенням глибини декомпозиції. Отже актуальною є задача визначення правила зупинки процесу декомпозиції.

Постановка задачі. Метою використання рекурсивної декомпозиції є формування моделі предметної області навчальних курсів, результатом чого є структура, математичною моделлю якої є дерево. Необхідно визначити правило зупинки процесу декомпозиції, результатом використання якого є визначення найкращої з погляду співвідношення показників інформативності і складності рівня деталізації структури.

2. Підхід базований на числі Міллера

Вивчаючи питання психології сприйняття, Дж. Міллер виявив, що «людина одночасно може утримувати в пам'яті, в фокусі уваги не більше ніж 7 ± 2 об'єкти» [2]. Пізніше виявився зв'язок числа Міллера з відомою психологам і фізіологам «кривою забування». Сьогодні можна говорити про наявність «визначеної за-

кономірності сприйняття» як теорії, що відбулася [3]. Стосовно до нашого випадку, це обмеження говорить, що не більше 7 ± 2 понять (концептів та зв'язків між ними) мають знаходитися у фокусі уваги. Більш того, чим менше (без втрати суті) понять, тим краще вони і сприймаються, і запам'ятовуються.

Покажемо на основі числа Міллера, що трирівнева декомпозиція для базових концептів є достатньою.

Нехай дано базовий концепт c_i . Перший рівень декомпозиції утвориться концептами, що пояснюють c_i . В найпростішому випадку цих концептів два, наприклад, c_{i1} і c_{i2} , і вони певним чином пов'язані між собою та базовим концептом. Назвемо цей зв'язок відношенням деталізації $R(c_i, c_{i1}, c_{i2})$, яке поглинає всі типи зв'язків, що їх використовують для декомпозиції, зокрема відношення «бути типом», «бути класом», «бути етапом процесу».

Другий рівень декомпозиції утворюють концепти, які пояснюють концепти першого рівня c_{ij} (в найпростішому випадку їх буде по два для кожного c_{ij}), а також два відношення $R(c_{i1}, c_{i11}, c_{i12})$ та $R(c_{i2}, c_{i21}, c_{i22})$. Таким чином, пояснення базового концепту c_i утворюють 6 концептів і 3 відношення, що в сумі дає 9 понять – обмеження Міллера.

Зауважимо, що розглядався найпростіший випадок – бінарна декомпозиція, проте ускладнення навіть цієї конструкції просуванням у глибину є недоречним. Отже можна зробити висновок про достатність трирівневої (базовий рівень та два рівні деталізації) декомпозиції для базових концептів предметної області.

Відзначимо, що питання балансу між широтою і глибиною структури досліджувалося в предметній області веб-дизайну [4]. Дослідження показують, що в загальному випадку «широкі» структури працюють краще, ніж «глибокі».

Сформулюємо правило. Якщо при виконанні рекурсивної декомпозиції для формування моделі предметної області відомі основні концепти, то зупинити процедуру слід не пізніше заглиблення на два рівні деталізації.

Розглянемо приклад. В предметній області SWB-OK – «Інженерія програмного забезпечення» [5] одним з основних концептів є концепт C_7 – «Управління програмним проектом». До пояснюючих його концептів входить C_{72} – «Планування програмного проекту» (см.рис.1).

Подальша декомпозиція концепту C_{72} приводить до одержання списку концептів, де є C_{725} – «Метод критичного шляху», який в свою чергу може бути представлений двома концептами C_{7251} – «Метод activity-on-node» і C_{7252} – «Метод activity-on-arrow». Декомпозиція цих концептів приводить до розгляду елементарних концептів «Критична подія», «Критичний шлях», «Повний часовий резерв» тощо. Подальша декомпозиція концептів недоречна, оскільки приводить до руйнування цілісності понять.

Якщо аналізувати предметну область оглядового курсу «Основи програмної інженерії», то базовими концептами будуть основні концепти SWBOK, зокрема C_7 . Відповідно, при декомпозиції понять предметної області для цього курсу доречно спуститися на два рівні щодо цього концепту (для наведених в прикладі концептів до C_{725}), але спускатися до елементарних концептів потреби немає.

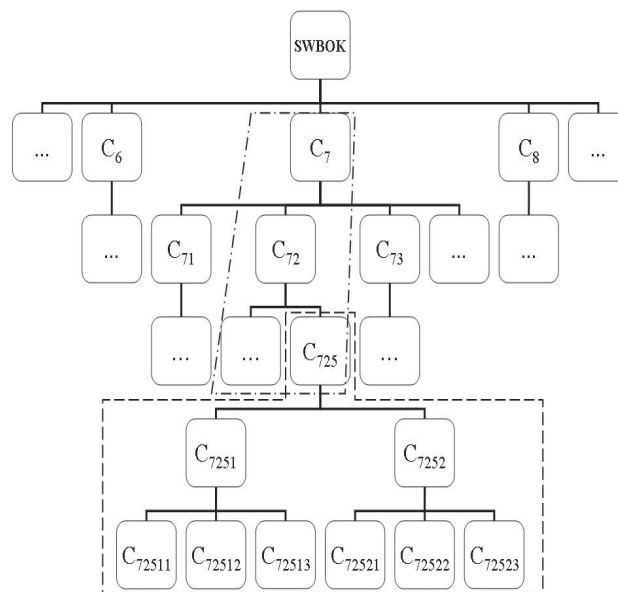


Рисунок 1. Фрагмент схеми областей знань SWBOK

Інша ситуація виникає в спеціалізованому курсі «Менеджмент проектів програмного забезпечення». Для нього одним з базових концептів є концепт C_{725} . Відповідно до підходу, базованому на числі Міллера, при моделюванні предметної області цього курсу необхідно в декомпозиції дійти до елементарних концептів.

3. Підхід базований на властивостях розподілу

В рамках вирішення поставленої задачі застосуємо підхід, в основі якого лежить використання властивостей грануляції моделі предметної області, тобто об'єднання атомарних понять для утворення концептів.

В [6] були сформульовані три критерії формування предметної області інтелектуальної системи. На їхній основі з використанням кількісної міри асоціативного зв'язку між концептами [7] сформулюємо три умови конструктивного визначення концептів:

Умова зайвої відособленості концепту c_i

$$\forall j, k : \text{ass}(c_k, c_i) < \text{ass}_{\min} , \tag{3}$$

де ass – міра асоціативного зв'язку. Виконання цієї умови означає, що концепт c_i практично не пов'язаний з іншими концептами предметної області. Найчастіше це викликано надмірною загальністю концепту (в цьому випадку доцільно розбити концепт на складові частини), але можуть бути і інші причини.

Умова надмірної докладності концептів

$$\exists i, j : \text{ass}(c_i, c_j) > \text{ass}_{\max} , \tag{4}$$

Виконання цієї умови означає, що є сенс об'єднати концепти c_i та c_j .

Умова збалансованого визначення концептів

$$\forall i, j : \text{ass}_{\min} \leq \text{ass}(c_i, c_j) \leq \text{ass}_{\max} , \tag{5}$$

Виконання цієї умови означає, що як результат грануляції отримано модель предметної області, яка не має ані надмірно відокремлених, ані надмірно пов'язаних концептів.

Для того, щоб скористатися умовою (5) слід визначити чисельні значення для ass_{min} і ass_{max} , що, очевидно, можна зробити методом експертного оцінювання. Але експертне оцінювання призведе до значної суб'єктивності в оцінці, спричиненої як особистими характеристиками експерта, так і особливостями предметної області.

Розглянемо множину значень міри асоціативного зв'язку для певної моделі предметної області як вибірку випадкових значень. На значення кількісної міри асоціативного зв'язку між концептами впливає багато факторів, отже вибірка, що розглядається, описується нормальним розподілом. Тоді для визначення граничних значень ass_{min} та ass_{max} застосуємо правило трьох сигм, яке стверджує про те, що імовірність відхилення випадкової величини від середнього значення більш ніж на 3σ , складає менше 0,003. При цьому якщо справжня величина середнього значення невідома, як σ слід використовувати стандартний відхил випадкової величини. Стосовно до умови збалансованого визначення концептів це правило формулюється як:

$$ass_{min} = ass_m - 3\sigma, ass_{max} = ass_m + 3\sigma, \quad (6)$$

де ass_m та σ — відповідно, середнє значення та стандартний відхил кількісної міри асоціативного зв'язку в аналізованій моделі предметної області.

Сформулюємо правило. Якщо умови (3)–(5) використовуються для перевірки конструктивності визначення концептів, то граничні значення ass_{min} та ass_{max} для цих умов слід визначати відповідно з (6).

4. Висновки

Метою використання рекурсивної декомпозиції є формування такої моделі предметної області навчальних курсів, яка має найкращий з погляду співвідношення показників інформативність-складність рівень деталізації.

В роботі сформульовано правило достатності тривірневої декомпозиції базових концептів моделі, доведення якого базується на використанні числа Міллера, яке дозволяє при відомих основних концептах визначити момент зупинки виконання рекурсивної декомпозиції.

Крім цього в роботі сформульовані умови конструктивного визначення концептів, які базуються на

дослідженні імовірнісних характеристик розподілу міри асоціативного зв'язку між концептами, і які дозволяють перевірити якість виокремлення концептів в моделі предметної області.

Слід зауважити, що розглянуті в роботі підходи доповнюють один одного. Підхід, базований на числі Міллера, за своєю суттю є структурним підходом, тоді як підхід, базований на властивостях розподілу, — семантичним. Отже, якщо при будівні моделі предметної області навчального курсу використовувалося правило тривірневої декомпозиції, доречно виконати для концептів, які складають модель, перевірку за умовами (3)–(5).

Література

1. Рыженко Н.Г., Жигарева Н.А. Структуризация и систематизация сюжетных задач по сложности их решения [Электронный ресурс] // Вестник Омского университета. — 1998. — Вып. 4. — С. 111 — 114. — Режим доступа: <http://www.omskreg.ru/vestnik/articles/y1998-i4/a111/article.html>
2. Miller G.A. The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information [Электронный ресурс] // Psychological Review. — 1956. — Vol.63. — P.81 — 97. — Режим доступа: <http://psychclassics.yorku.ca/Miller/>.
3. Букалов А.В. Феномен структурирования психоинформационного пространства: иерархия объемов человеческого внимания, памяти и мышления [Электронный ресурс] // Соционика, ментология и психология личности. — № 2. — 1999. — Режим доступа: <http://www.socionics.ibc.com.ua/t/magic.html>.
4. Kalbach J. The Myth of «Seven, Plus or Minus 2» [Электронный ресурс] // Dr. Dobb's Journal. — 2002. — No 1. — Режим доступа: <http://www.ddj.com/184412300>.
5. Guide to Software Engineering Base of Knowledge (SWE-BOK) [Электронный ресурс]. — Режим доступа: <http://www.swebok.org/>
6. Krissilov V., Shabadash D. Towards the Problems of an Evaluation of Data Uncertainty in Decision Support Systems // Information Theories & Applications. — Vol.3. — P. 376 — 380.
7. Любченко В., Крисилов В. Использование ассоциативных сетей для грануляции материалов курса // Proceedings of the Second International Conference «Modern (e-) Learning» — Varna, 2007. — Sofia: FOI ITHEA, 2007. — P. 61 — 64.