

一种面向对象的软件设计模式库的设计

钟金琴¹, 辜丽川²

(1. 安徽大学 电子与信息系, 安徽 合肥 230011;

2. 安徽农业大学 计算机学院, 安徽 合肥 230036)

摘 要:成功地利用模式可以有效提高软件的开发效率 and 设计质量, 通用模式具有较强的普适性, 结合具体领域特点形成的实例模式具有更强、更具体的指导作用。为了实现一个辅助设计模式复用的 CASE 工具, 对设计模式的几个面向对象的原则进行分析, 提出了一种面向对象的分布式软件实例模式库来存储设计模式, 并给出相应的类层次库组织结构、分布式的交互管理和基于聚类的多级索引方法及面向对象的嵌套索引方法。依此存储模式以更有效地管理和利用模式。

关键词:设计模式; 实例模式库; 面向对象; 分布式实例模式库

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2008)09-0022-04

Design of an Object - Oriented Software Design Pattern Base

ZHONG Jin-qin¹, GU Li-chuan²

(1. Department of Electronics and Information, Anhui University, Hefei 230011, China;

2. School of Computer, Anhui Agricultural University, Hefei 230036, China)

Abstract: Software quality and the development efficiency can be improved by using pattern, the more general pattern can be used in a wider range. Combining with domain specific character, the case patterns have stronger and more specific guidance. To design a CASE tool for design patterns reuse, in this paper object-oriented principles of design pattern are analyzed, a object-oriented distributed software case pattern base is brought forward to store design patterns, then the class hierarchy structure, distributed interactive management, the methods of based multi-class index and object-oriented nested index of the base are described. And in order to manage and use patterns more efficiently, suggest to store patterns by the object-oriented software case pattern base.

Key words: design pattern; case pattern base; object-oriented; distributed case pattern base

0 引言

软件复用是解决软件危机最根本的途径之一。然而,随着软件的日益复杂和庞大,人们对软件可复用性的要求也越来越高了,希望所使用的软件不仅能满足当前的功能需求,而且还要有足够的灵活性以适应将来的变化。设计面向对象软件比较困难,设计可复用的面向对象软件就更加困难,而设计模式^[1]提供了一条捷径,它将许多优秀的面向对象软件中一些最重要的设计经验进行了总结并以标准的形式记录下来,它的提出有助于学习成功的设计并在更高的抽象层次上理解和讨论面向对象系统的设计,从而大大提高软件开发的效率和质量。但设计模式相对于 1995 年 Gang of Four (GoF)提出的 23 种通用的设计模式,数量已

经大大增加了。要扩大应用,就必须把设计模式纳入数据库管理中,使设计模式的记录、检索、修改等处理更趋自动化,便于用户更好地复用设计模式。

目前,进行软件模式自动化开发方法和工具研究是一个热点,最终目标是为模式提供自动化 CASE 工具支持。其中,IDE's (Integrated Development Environments)和 UML 建模软件环境已开始为模式提供支持,如 UMLStudio, togetherJ, ModelMaker 等,但还没有一个完全自动化的、代表性的复用系统可供用户使用。在这些系统中有的使用文件来存储模式,有的使用 XML 数据模型来描述并存储模式。但这两种存储方式普遍存在数据类型表达能力弱,复杂查询功能差等缺点。设计模式是面向对象设计经验的记录,因此用面向对象的数据库对其存储是最为合适的。通过对设计模式的分析,提出一种采用面向对象技术对分布式模式库进行组织和管理的方法,以改善模式库的检索和管理能力。

收稿日期: 2007-12-17

基金项目: 安徽省科研计划项目(2007jq1022, 2006jql130)

作者简介: 钟金琴(1973-),女,安徽六安人,讲师,硕士,研究方向为软件工程。

1 设计模式分析

设计模式把许多优秀的面向对象软件中一些最重要的设计经验记录下来,以提高软件开发的效率和质量。通过研究,发现尽管各个模式形态各异,目标不同,但它们都体现了相同的几个面向对象原则^[2]:

(1)封装机制:对象属性的存取是通过向其发送消息实现的,而方法封装在对象中。对象指定的公共接口是存取对象的唯一方法。

(2)类的继承:类继承是面向对象的基本特征之一,它是一个通过复用父类功能而扩展应用功能的基本机制。利用继承,可以从已存在的类中继承所需要的绝大部分功能,从而几乎无需任何代价就可以获得新的实现。

(3)优先使用对象组合:对象组合是指通过组装或组合对象来获得新的更复杂功能,对象组合要求被组合的对象具有定义好的接口,这种复用风格也叫做黑箱复用。因为对象的细节是不可见的,对象表现得就像一个“黑箱”一样。而在面向对象的数据模型中,这几个原则也包含在其核心数据模型中。

另外,Erich Gamma 和 Frank Buschmanm 的著作都是从功能、目标、范围等几个角度对模式进行分类,给出了一些普适性强、抽象度高的模式,并对模式之间的连接关系作了一定的描述。然而模式也应该有抽象度的差异、有适用范围的大小。模式不一定必须在所有的领域都适用,特定领域的在一定条件下反复出现的问题的解决方案也完全可以成为模式^[3],从 Alexander 的原始定义中可以看出这种意味,Erich Gamma 把通用模式在特定领域的具体化实现称为框架,其实框架也可以作为一种特殊的模式,因为它也是在某一个语境中重复出现的问题的解决方案,也是可以复用的,只不过抽象程度不同而已,而且对于特定领域,具体化的模式具有更易理解的领域特征,也具有更多的领域指导信息,所以也易于复用,把这种具体化的模式称为实例模式。

2 面向对象软件设计模式库的设计与实现

为了实现一个辅助设计模式复用的 CASE 工具,项目的开发人员对设计模式进行了深入地分析,并研究其复用方法和关键技术。在此过程中,提出了一种面向对象的实例模式库来存储设计模式。实例模式库采用面向对象技术进行组织和管理,这样可以使库模块和类划分更加合理,以改善模式库的检索和管理能力^[4]。

2.1 面向对象实例模式库的组织结构

利用面向对象实例模式构造类层次结构,不但可

以描述和存储实例模式,而且适合于处理嵌套集合和继承关系,从而提高了实例模式数据处理的效率。

实例模式库可以分为模式类和相应的实例两部分,两者分别存储在库中不同的物理区域,并具有相应的类-实例的映射关系。面向对象的实例模式库是通过泛化和聚集关系组成类与实例的集合。各个实例模式类包括属性和方法,消息处理需要对各类进行检索,通过问题中的对象找到所继承的类,因此实例模式类有效的组织和存储对检索具有重要作用。实例模式对象按类进行存储,而类的每一个变量都用记录型的域表示。为了提高相关类数据的检索速度,需要将相关的数据封装在一起,把属于同一类的对象存储在同一段上,相似的类存放在近邻的段上。实例对象中,类的属性按其结构进行存储,而方法是按其标示存放,因此可以将其看作一种特殊的属性。实例模式类以类标示进行存储,类之间的继承关系通过物理地址指针实现。各个继承的相似类存储在文件相邻域,文件的一部分称为实例模式存储域,相关域为实例模式关联域。其存储结构如图 1 所示。

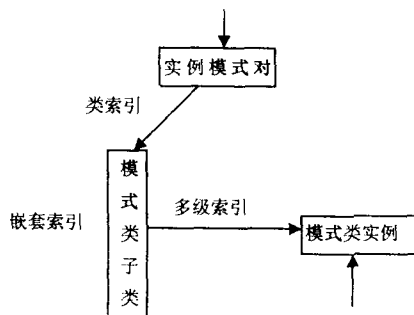


图 1 实例模式库存储结构

模式库可以是一个分布式的面向对象数据库,使用客户/服务器结构分布数据库到异质工作站的网络上^[5]。服务器端负责数据存储的管理。客户机为用户进程提供接口,并负责管理数据在应用进程的虚存空间的映射。客户和服务器端通过网络进行通信。联编程序负责把客户与数据库服务器连接起来。这种结构配置如图 2 所示。

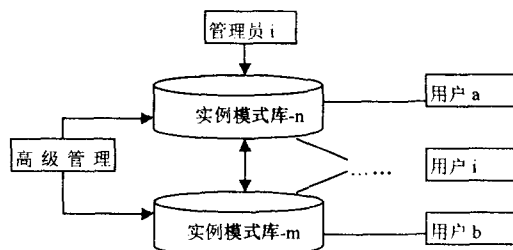


图 2 分布式实例模式库

在实例模式库中各个类方法和属性类的连接方法有两种:静态联编和动态联编。前者是将方法和属性

类直接加入到实例模式类中,而后者是将方法和属性类存放在其它位置,通过指针将方法和属性类的标识建立联系,在使用时需要进行检索,得到相应的方法和属性类。静态联编所用的存储空间较大但效率较高,而动态联编所用的存储空间较小但效率受索引的影响。因此在实例模式库中,由于方法和属性类所占空间较大,一般采用动态联编方式,通过索引等形式提高检索的效率,同时将常用属性和方法采用静态联编方式进行存储。

2.2 面向对象实例模式库的分布式交互和管理

实例设计模式的继承作为实例模式类之间的关系,也是继承类之间数据和方法的共享机制。类的定义可以在已存在类的基础上进行,把已存在类所定义的成分作为本类的属性或方法,然后加入新的内容。因此,实例模式对象的继承性自然地体现了对象间的共享,类中定义的方法和属性成为其它继承类共享的信息。实例模式对象对其方法和属性进行封装,其调用和改变均通过对象接口的消息机制进行,从而方便了实例模式类的共享。在异构环境中,对实例模式类的共享提高了系统处理能力,对象间的消息传递通过类协议的方式进行解析和处理,这样不但提高了实例模式对象的描述能力,而且方便了系统知识的扩充。

在实例模式对象的共享中,会出现多个用户同时访问或发送消息的情况,从而导致了并发冲突。面向对象实例模式类的有效封装,通过消息机制对类的属性和方法进行操作,从而避免了对对象间的并发冲突,完善了实例模式的共享机制。

2.3 面向对象实例模式库的索引方法

建立模式库的目的是为了复用,成功复用的一个关键因素是检索方法,而实例模式库的索引是提高检索效率的关键。在检索中,首先通过类的索引找到所属的类,然后根据实例索引定位实例的地址,通过实例的嵌套索引可以得到相关类的属性和方法等,因此这种实例模式库的索引方式方便了库的管理和维护。

2.3.1 基于聚类的多级索引

模式检索本质上就是从模式库中选出一个最相似的模式,基于聚类的多级索引可以在检索中获得相似的实例模式集合。在库中的索引文件很大时,需要在索引文件上再建立索引,该过程可以重复多次直到满足要求为止,则索引形式定义为多级索引。

通过对检索内容抽取的属性,定位在某个确定的存储域内,然后在所定位的域内进行检索实例模式。在多级索引中经常结合某些映射函数可以提高检索的效率。设检索的内容为 S , 映射函数为 $H(\omega, \theta)$, ω, θ 为 S 的两种属性。 F 是一个文件, A 是 F 的存储域, r 是

文件中物理地址为 a 的一条记录,通过映射函数可以建立联系:

$$H(S) \xrightarrow{F} a \xrightarrow{A} r$$

在同一索引中相似实例聚集在一起,因此根据多级索引容易得到相关实例,其关系表示为: $I_{i1} = \text{Child}(I_{i-1,j}), C(I_{ij}) \subset C(I_{i-1,j})$ 。其中, I_{ij} 表示 $I_{i-1,j}$ 的下一级索引, $\text{Child}(\cdot)$ 表示索引的下一级索引集合, $C(I)$ 表示索引 I 相关的存储实例模式。

采用一般的映射方法,可能造成数据在映射中所检索的实例过多,严重影响库的检索效率,因此需要对所存储的实例模式进行合理的划分。实例模式库提供了自组织聚类方法:在 N 个实例模式中, d_{ij} 表示第 i 个与第 j 个实例之间的距离,计算 N 个实例两两之间的距离 d_{ij} ,可以得到对称矩阵 $D(\cdot)$,表示如下:

$$\begin{bmatrix} 0 & & & & \\ d_{21} & 0 & & & \\ d_{31} & d_{32} & 0 & & \\ \vdots & \vdots & & \ddots & \\ d_{N1} & d_{N2} & & d_{NN-1} & 0 \end{bmatrix}$$

设 G_i 表示初始类,利用最短距离方法对 G_i 中距离最小的两类进行合并, D_{pq} 表示两类 G_p 与 G_q 的距离,用 D 中最小的非零元素表示:

$$D_{pq} = \min_i(d_{ij})$$

将 G_p 与 G_q 合并一类 $G_r = \{G_p, G_q\}$, 然后计算类 G_r 与其它类 $G_k (k \neq p, q)$ 的距离,取 $D_{rk} = \min\{d_{ij}\} = \min\{D_{pk}, D_{qk}\}$, 如此继续,最终得到实例模式库的多个分类。

对聚类实例模式需要进行离散程度的计算,公式为:

$$\Psi_c = \sqrt{\sum_{i=1}^n ((r_i - r_0)/(r_0 - r_g))^2}$$

其中, r_i 表示某映射下对应的所有实例模式, r_0 表示初始实例模式, r_g 表示实例模式计算的重心值。

当离散程度超过设定的阈值时,一方面需要进行实例模式的分解,另一方面可调节映射函数的动态因子,以保证实例的离散度以某一规定范围内变化,从而呈现索引的均衡状态。上下文信息,得到相关特征的相似度及范围,最终获得相似的实例模式集合。

2.3.2 面向对象的嵌套索引

实例模式库的检索主要可以分为两大类:类层次检索和实例检索。在实例模式库中,面向对象的组织结构使索引可以采用两种主要形式:一种是类的层次索引,主要用于检索实例模式类本身,或检索相关层次上直接或间接的子类。另一种索引形式是嵌套索引,

作为类复合层次上嵌套属性的索引,可以通过遍历简单的索引来实现复杂的检索和查询。

设 H 表示实例模式类的聚集 $C_1 C_2 \cdots C_n$, H 上的一条检索路径定义为 $P = C_i A_1 A_2 \cdots A_n (n > 1)$, 其中 C_i 表示 H 中的一个类, A_i 表示 C_i 中的一个属性, 则 C_i 是类 C_{i-1} 的一个属性 A_{i-1} 的值域 ($1 < i \leq n$)。

由于库中的实例模式类具有多种属性和成分, 嵌套索引同时也是对多种成分检索的有效接口。通过各个实例模式类的公共属性, 可以对多个不同的实例模式类进行联合检索。由此可见, 嵌套索引是体现面向对象特征检索的索引形式, 直接关系到面向对象的分析和检索效率。

3 结束语

随着软件设计模式的日益增多, 在各个学术和技术领域, 模式都在有意无意地应用着, 把这些人类知识的精华进行有效复用的意义是非常大的, 有效地复用模式需要首先有效地组织和管理模式, 文中对软件设计模式的特性进行分析, 提出一种面向对象的软件实例模式库来存储设计模式, 并给出相应的库组织结构、

管理和索引方法。

采用面向对象的方法对软件设计模式进行划分、组织和封装, 克服了传统数据库系统普遍存在的数据类型表达能力弱, 复杂查询功能差等缺点, 提高了模式知识的共享性和重用性, 增强了对实例模式的检索和管理能力。力图在此基础上研制出一个高效且真正实用的设计模式的复用和支持系统。

参考文献:

- [1] Gamma E, Helm R, Vlissides J J. Design Patterns: Elements of Reusable Object - Oriented Software[M]: [s. l.]: Addison - Wesley, 1995.
- [2] 王晓庆, 曾文英. 设计模式中的面向对象原则及其子模式[J]. 计算机工程, 2003, 29(9): 192 - 193.
- [3] 孟祥文, 邵维忠. 设计模式特化和模式库组织[J]. 计算机工程, 2002, 28(5): 36 - 37.
- [4] Su S, Lam H, Eddula S, et al. Osam: An object - oriented knowledge base management system for supporting advanced application[J]. ACM SIGMOD, 1993, 22(2): 12 - 40.
- [5] 雷光复. 面向对象的新一代数据库系统[M]. 北京: 国防工业出版社, 2000.

(上接第 21 页)

由表 2 中两种索引项提取方式的横向对比数据可以看出, 两种索引项提取方式在执行 TD ($N = 10$) 和 NP 任务时均可获得令人满意的效果。但专业索引项提取策略在检索效率上大大优于通用索引项提取策略, 全部结果中的无关文档数量也明显减少。

表 2 测试结果对比数据

	专业词 + 基本词索引项				基本词 + 单字索引项			
	P(10)	R	结果数量	时间(ms)	P(10)	R	结果数量	时间(ms)
1	10/10	/	32	6	8/10	/	418	28
2	4/4	/	4	≤ 5	4/10	/	372	21
3	9/9	/	9	≤ 5	9/10	/	3797	118
4	8/10	/	25	≤ 5	8/10	/	5151	162
5	5/5	/	5	≤ 5	5/10	/	52	8
6	6/6	/	6	≤ 5	6/10	/	41	11
7	3/10	/	1687	38	3/10	/	5802	181
8	/	1	2	≤ 5	/	1	30	8
9	/	1	17	≤ 5	/	1	6594	139
10	/	1	157	15	/	1	226	32

3 结束语

文中提出的索引项提取策略可在不降低基本索引项提取精度的前提下有效地提取专业索引项。并经过

实验证明, 此方法提取的索引项可以较明显地提高专业检索系统的检索效率, 降低硬件开销。

参考文献:

- [1] Nie Jian - Yun, Gao Jiangfeng, Zhang Jian, et al. On the use of words and n - grams for Chinese information retrieval[C] // In: Proceedings of the fifth international workshop on Information retrieval with Asian languages. Hong Kong, China: [s. n.], 2000: 141 - 148.
- [2] He Hongzhao, Gao Jianfeng. Finding the better indexing units for Chinese information retrieval[C] // In: Proceeding of the first SIGHAN workshop on Chinese language processing - Volume 18. [s. l.]: [s. n.], 2002: 1 - 7.
- [3] Salton G. A Vector Space Model for Automatic Indexing[J]. Communications of the ACM, 1975, 18: 613 - 620.
- [4] Fagan J L. Experiments in Automatic Phrase Indexing For Document Retrieval: A Comparison of Syntactic and Non - Syntactic Methods[D]. Upson Hall Ithaca, NY, USA: Computer Science/Arts & Eng. Dept., Cornell University, 1987.
- [5] TREC[EB/OL]. 2004. <http://nltpress.mit.edu/books/chapters/0262220733chap1.pdf>.
- [6] SEWM[EB/OL]. 2007 - 03. <http://www.cwrf.org/Evaluation/CWT.html>.