

Internet 平台下的构件描述模型^{*})

许 峰 张安勤 黄 皓 王志坚

(软件新技术国家重点实验室 南京大学计算机科学与技术系 南京 210093)

摘 要 为了使构件使用者能很容易地理解构件的功能及其属性,对构件作一个清晰的描述是非常必要的。随着构件技术和 Internet 技术的发展,构件的结构和构件软件框架发生了变化,急需一种能描述在 Internet 平台下的构件的描述模型。本文概括出一个四元组的构件描述模型,为构件异构问题的解决提供了一种方法。

关键词 Internet, 构件, 构件模型, 过程模型

Component Description Model in the Internet Environment

XU Feng ZHANG An-Qin HUANG Hao WANG Zhi-Jian

(State Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing 210093)

Abstract It is very essential to describe component clearly for the component user can understand the functions and attributes of component easily. With the evolution of component technology and Internet techniques, the great change has taken place in the structure of component and component software framework. Thus, it becomes an urgent need a component description model which can describe the component in the Internet environment. In this paper, we propose component description model of quadruple and this can present an effective solution to problem of heterogeneous component.

Keywords Internet, Component, Component model, Process model

1 引言

20 世纪 90 年代软件开发技术的一个重要进展就是构件化。现在的系统规模越来越大,一个系统要求完成的功能很多,因此软件复用和集成更加具有非同寻常的意义。为了提高软件生产率,有效地开发应用程序,设计开发人员应尽可能地利用可复用的软件构件,组装构造新的软件系统。

作为解决软件危机的一种有效手段,软件构件(简称构件,Components)技术成功地完成了系统开发的重点从程序设计到对已有构件的选取、组装和部署的转换。在构件模型的支持下,通过复用已有的构件,软件开发者可以“即插即用”地快速构造应用系统。这样不仅可以节省时间和经费,提高工作效率,而且可以产生更加规范、更加可靠的应用软件。

Internet 无疑是 20 世纪末最伟大的技术进展,其普及和发展为我们提供了一种全球范围的信息基础设施。这个不断延伸的网络基础设施,形成了一个资源丰富的计算平台。以 Internet 为主干,各类局域网或无线网为局部设施,伴随着移动通讯技术的发展,再加上各种信息处理设备和嵌入设备作为终端,构成了人类社会的信息化、数字化基础,成为我们学习、生活和工作的必备环境^[1]。

Internet 平台具有如下基本特征:无统一控制的“真”分布性;节点的高度自治性;节点链接的开放性和动态性;人、设备和软件的多重异构性;实体行为的不可预测性;运行环境的潜在不安全性;使用方式的个性化和灵活性;网络连接环境的多样性等。

软件复用是软件发展中的一个重要方向,构件化的软件生产有了很广泛的研究与应用。但是,在传统方式下人们在局域网里构建自己的软件构件库、软件构件检索机制,其工作效率和复用程度是有限的,我们要把整个复用环境放到大广域网甚至 Internet 上。而随着 Internet 的迅猛发展,基于 Internet 的应用系统的逐渐增多,构件的交互范围进一步扩大并分布在整个 Internet 上。Internet 平台和环境的出现,它的异构、开放、动态和多变的特征,对构件技术的发展、理论研究提出新的问题,也提供了新的契机。

目前已出现对 Internet 环境下构件的结构和构件软件框架的研究^[2]。现有的构件描述技术难以充分描述 Internet 环境下的构件。本文主要对构件描述方面进行研究,给出一种在 Internet 环境下的构件描述模型。

2 构件描述模型

可复用性对可复用对象的要求较高,要求可复用对象能被充分理解,方便存取,且具有通用功能。因此,为了使构件使用者能很容易地理解构件的功能及其属性,对构件作一个清晰的描述是非常必要的。一般认为描述构件的途径有三条:模型、方法和语言。构件模型是对构件进行全面描述的最佳途径。

构件模型是构件技术研究与应用的交互点,既是构件技术研究的主要内容和成果,又是构件技术运用于实际应用系统开发的主要依据。一个构件模型应该满足一些基本原则:

(1)表达能力足够强:模型是对客观对象的抽象,合适的

^{*}) 本课题得到国家自然科学基金(60473091)和国家“八六三”高技术研究发展计划项目基金(2003AA142010)资助。许 峰 博士研究生,主要研究方向为软件构件和分布式计算。张安勤 硕士,主要研究方向为软件构件。黄 皓 教授,博士生导师,主要研究领域为软件自动化。王志坚 教授,博士生导师,主要研究领域为软件自动化,分布式计算。

抽象层次十分重要,必须既能抓住本质,又不陷入细节,既要有足够的表达能力,又不能过于复杂。

(2)简单性:简单性是各种模型必须予以考虑的重要性质,简单意味着易于掌握和理解,简单性和强的表达能力是一对矛盾,必须合理权衡。

(3)一致性和完备性:模型必须具有对内的一致性和对外的完备性。一直性是指构件作为一个封装体必须具有一致的对外接口、一致的组成结构以及一致的交互方式。完备性是指不存在模型所不能描述的构件。

(4)实用性:构件模型应易于理解、方便分类和检索。

(5)扩展性:指在保持模型本身一致性和完备性的前提下,模型可以随着应用需求的增长而变化。

目前,随着研究的进展以及对构件认知的深入,已有多种构件模型出现,在学术界,比较有代表性的是 W. Tracz 提出的 3C 模型^[3],REBOOT 项目中提出的 REBOOT 模型^[4],北京大学提出的青鸟构件模型^[5]。它们是指导性模型,抽象层次比较高,用户可以根据不同的问题域来扩展模型。在产业界,以 CORBA^[6]、COM/DCOM/COM+^[7]和 EJB^[8]为代表的基于分布式对象技术的构件实现模型正在向实用化方向快速发展,它们对构件的基本构成及其体系结构的演化产生着十分重要的影响,从而成为实现级的主流构件模型。

3 Internet 平台下的构件模型

随着 Internet 技术的快速发展与普及,软件系统呈现出柔性、多目标、连续反应式的新的形态。软件实体具有独立性、主动性和自适应性,实体之间的交互以多种静态连接和动态合作的形式实现互连、互通、协作和联盟。软件不再被当做是一种产品,而被看做是一种 Internet 环境下的服务资源,面向服务的理念成为一种共识。在 Internet 环境下,构件提供的功能以服务的形式表现出来。对构件信息的准确描述是使用构件的前提,这就要求建立一种恰当的构件描述模型。

根据面向对象方法学,这里给出了构件在 Internet 环境下的形式化描述,提出用如下特征信息集来刻画一个可复用构件,用 BNF 范式表达如下:

定义 1 构件描述 CD(Component Description)抽象地表示成一个四元组 $CD ::= \{ COD, CID, CBD, CMD \}$ 。其中:

CD:构件描述信息的集合;COD:构件本体描述(Component Ontology Description);CID:构件的接口描述(Component Interface Description);CBD:构件的行为描述(Component Behaviour Description);CMD:构件的映射描述(Component Mapping Description)。

定义 2 构件本体描述 COD 包含了构件的概念术语,它抽象地表示成一个三元组 $COD ::= \{ TNI, TMI, TRI \}$ 。其中:

TNI 术语名称信息(Term Name Information)给出术语的名称;TMI 术语的含义信息(Term Meaning Information)给出术语的准确含义;TRI 术语的关系信息(Term Relation Information)是对术语概念之间关系的描述。

定义 3 构件接口描述 CID 包含了 Web 构件的基本特征信息和有关的性能属性,它抽象地表示成一个三元组 $CID ::= \{ TI, FI, AI \}$ 。其中:

TI 文字信息(Text Information)::={名称,制作日期,存放的位置,提供商的信息,开发语言,开发环境,版本号};

FI 功能信息(Function Information)::={输入,输出}。

通过输入和输出来描述构件的功能。其中:

输入指定了构件进行提供功能所需要的信息。输出指定了构件执行的结果。

AI 附加信息(Additional Information)::={算法说明,构件历史,备注文件,构件的质量,构件类型}。

其中:构件类型={需求分析构件 | 设计构件 | 代码构件 | 测试构件}。

构件的接口是它和外界的一组交互作用点。接口描述构件需求的服务和提供给外界的服务,特别是构件能够接受的消息和发出的通知。构件通过接口与连接器的角色相匹配,这样,若干构件就在彼此不知道的情况下由连接器组装成一个有机的整体,完成系统家族的共同行为。

构件可能有多个类 IDL 接口,这些公共接口是进行构件组装的基础。以上我们把构件接口分为三部分进行描述。其意义在于:一是有利于构件的合成,二是有利于构件的逻辑层次的划分。一组相关的构件合成为一个构件时,根据接口的分类描述,在提供的服务和需要的服务之间建立接口连接,而合成构件的接口依然按这三部分描述。

定义 4 构件行为描述 CBD::={组合的输入信息,组合的输出信息,组合的控制结构}。

行为是由操作次序和消息的发送和接收规则确定。引入一种过程模型^[9]来描述构件的内部流程。一个构件的流程就是一个过程,用这个过程来展现构件的内部细节。过程描述了构件执行的步骤和在执行过程中会遇到的情形,因此它是对构件的行为特性的一种描述。一个构件的过程可以调用其它构件的过程,这样一个复杂构件可以由其它构件组合。这个过程模型就提供了描述构件组合的机制。

我们把具有复杂功能的构件称为复合构件,而实现简单功能的构件称为原子构件。组合构件是由原子构件按照不同的组合模式进行组合而实现的。其中,构件的组合模式是构件的执行序列,它们是顺序、条件分支、并行、无序、选择、Repeat 循环和 Until 循环这 7 种控制结构^[10]。过程模型如图 1 所显示:构件中有两种类型的过程:原子(atomic)和复合(composite)过程。

(1)原子过程

从构件请求者的角度来看,原子过程是可以直接调用的(通过传送恰当的消息),没有子过程且在一个步骤内执行完成。也就是说,它接收一个输入消息,执行,然后返回输出消息。原子构件的执行过程对于构件请求者是不可见的。对于每个原子过程必须提供一个映射,这样才能知道构件是如何实现的。但映射的声明和过程的声明是分开的,这样做允许重用过程定义。

(2)复合过程

如果一个复合过程调用了多个子过程,它需要安排调用这些子过程的顺序和调用逻辑,这就是控制流。在编程语言中,有三种基本的控制结构:顺序、条件分支和迭代。原则上说,有了这三种控制结构,通过重复嵌套可以构成任意复杂的结构。在这三种基本控制结构的基础上,考虑了并行程序语言,一共有 7 种控制结构:顺序(Sequence)、并行(Concurrence)、无序(Unordered)、选择(Choice)、条件分支(If)、Repeat 循环(RepeatWhile)和 Until 循环(UntilWhile)。用这 7 种控制结构来描述复合过程中的控制流。

每个复合过程都有一个控制结构。每个控制结构由过程构件(ProcessComponent)组成,过程和控制结构都是过程构

件。控制结构决定了其过程构件的调用顺序和调用逻辑。下面逐一介绍这 7 种控制结构。

Sequence 的过程构件按顺序执行,Concurrence 的则是并发执行,Choice 允许选择一个构成构件执行。Unordered 允许其过程构件以任意顺序或并发执行,但所有的过程构件都必须执行。If 有一个条件(Condition)和 Then 分支,0 到 1 个 Else 分支。其中,条件(Condition)是逻辑表达式,Then 和 Else 分支分别对应一个过程构件。它的语义是,如果条件为真,则调用 Then 分支对应的过程构件,否则调用 Else 分支对应的过程构件。RepeatWhile 有一个条件(Condition)和重复(Repeat)块。其中,条件是逻辑表达式,重复块对应一个过程构件。它的语义消息是,如果条件为真,则反复执行重复块对应的过程构件,直到条件为假。

RepeatUntil 和 RepeatWhile 类似,其语义是,反复执行重复块对应的过程构件,直到条件为真。

(3) 数据流

在复合过程中,除了控制流外,还有数据流。

在定义过程时,一个过程的输入/输出参数可能与其它过程的输入/输出参数相同。

一个简单的例子就是,一个用于购买商品的原子过程。要购买的商品由某个名字或标识指定,并作为该过程的输入参数。该过程的一个输出参数也引用这个标识,作为输出消息的一部分说明这个交易是成功还是失败。类似这样的情况有很多,需要指明哪些参数是相等的,如有下面两种情形:

①将一个组合过程的输入和输出同它的子过程的输入和输出关联。

②将一个组合过程的子过程的输入和输出关联。例如,当一个组合过程的控制结构是 Sequence,它的过程构件包括多个子过程,则这些子过程执行顺序中的一个子过程的一个输出可能是其后续子过程的一个输入。

在编程语言或一个逻辑语言中,可以通过变量和赋值将这些参数关联起来。在编程语言中,变量可以是函数参数或局部变量。它们可以在函数体中被引用,通过赋值指明某个步骤的一个参数和函数的一个输入相同,或者它和前面某个步骤的输出相同。

在过程模型中,采取类似于编程语言中的方法,即使用变量和赋值。为此,定义了变量定义和赋值的机制。在一个复

合过程中定义的一个或多个变量就组成了这个过程的上下文(Context)。在一个复合过程中的多个赋值就组成了这个过程的数据流(DataFlow)。变量除了用于描述数据流外,还可以用于条件表达式,用来描述控制结构 if、RepeatWhile 和 RepeatUntil 的条件。

组合的输入信息代表了执行组合过程所需的信息,例如子构件的名称。组合的输出信息代表了组合过程执行后所提供的信息。

定义 5 构件的映射描述 CMD 说明了构件是如何实现的,它由 1 到多个操作映射描述组成。在 Internet 环境,构件以 Web 服务作为基础技术,操作定位机制指定了参加映射的 Web 服务操作。CMD 可表示成一个四元组 $CMD_i = \{ \text{服务类型}, \text{服务的位置}, \text{服务的接口}, \text{服务的操作} \}$ 。

结束语 本文在分析传统构件的描述模型和 Internet 运作机理的基础上,概括出构件的描述模型,使用四元组(构件本体的描述,构件接口的描述,构件行为的描述,构件的映射描述)来描述构件,并用 BNF 范式进行了表达。在模型中构件以 Web 服务作为基础技术,构件的操作通过映射到 Web 服务来实现。这个模型为构件异构问题的解决提供了一种方法。

参 考 文 献

- 1 杨美清. 软件工程技术发展思考[J]. 软件学报, 2005, 16(1): 1~7
 - 2 吕建, 张鸣, 廖宇, 陶先平. 基于移动 Agent 技术的构件软件框架[J]. 软件学报, 2000, 11(8): 1018~1023
 - 3 Tracz W. Implementation working group summary. In: James Baldo, ed. Reuse in Practice Workshop Summary, Alexandria, VA, April 1990. 10~19
 - 4 Weighted Term Spaces for Related Search. In: Proc. of the 1st Intl. Conf. on Information and Knowledge Management (CIKM'92), 1992/NOV. 5~8
 - 5 青鸟可复用软件开发指南: [青鸟工程项目组技术报告]. 北京大学计算机科学系, 1997
 - 6 OMG CORBA Specification. <http://www.omg.org/corba/whatiscorba.html>
 - 7 Microsoft. <http://www.microsoft.com/isapi>
 - 8 SUN. <http://developer.java.sun.com/developer>
 - 9 费立蜀, 顾庆, 陈道庆. 一种过程定义模型及其验证性分析[J]. 计算机科学, 2004, 31(1): 145~151
 - 10 王斌, 张尧学, 陈松乔. 基于语义网络的构件组织模型及实现. 小型微型计算机系统, 2004, 7: 1231~1234
-
- (上接第 181 页)
 - 结论 集值信息系统可以用来处理数据不完全的信息系统。本文通过给出集值决策信息系统中的一种新的关系定义, 给出了在这种关系下协调集值决策信息系统属性约简的判定定理和辨识矩阵, 并讨论了协调集值决策信息系统中不同类型的属性特征及每一种属性的判定定理。本文还定义了不协调集值决策信息系统的分配协调集, 给出了一种将不协调集值决策信息系统转化为广义协调近似空间的方法, 并证明了不协调集值决策信息系统的分配协调集就是广义协调近似表示空间的协调集, 说明了可以通过广义协调近似表示空间的辨识矩阵计算不协调集值决策信息系统的分配协调集, 不协调集值决策信息系统中也具有三种不同类型的属性, 并且满足相应的属性特征及判定定理。
 - 参 考 文 献
 - 1 Pawlak Z. Rough Sets. International Journal of Computer and Information Sciences, 1982, 11: 341~356
 - 2 Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishers, Boston, 1991
 - 3 张文修, 梁怡, 吴伟志. 信息系统与知识发现. 北京: 科学出版社, 2003
 - 4 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法. 北京: 科学出版社, 2001
 - 5 张文修, 仇国芳. 基于粗糙集的不确定决策. 北京: 清华大学出版社, 2005
 - 6 王国胤. 粗糙集理论与知识获取. 西安: 西安交通大学出版社, 2001
 - 7 王珏, 王任, 苗夺谦, 等. 基于 Rough Set 理论的“数据浓缩”. 计算机学报, 1998, 21(5): 393~400
 - 8 Kryszkiewicz M. Comparative study of alternative type of knowledge reduction in inconsistent systems. International Journal of Intelligent Systems, 2001, 16: 105~120
 - 9 Kryszkiewicz M. Rough set approach to incomplete information systems. Information Sciences, 1998, 112: 39~49