

基于本体的领域特征建模过程研究

顾剑华 赵文耘 彭 鑫

(复旦大学计算机科学与工程系 上海 200433)

摘 要 软件复用的研究和实践表明特定领域的软件复用活动相对容易取得成功,这主要是由于领域的内聚性和稳定性。而系统、全面的领域分析是成功的领域级复用活动的前提和关键。在领域分析和建模的相关研究工作中,已在特征建模方法的基础上引入本体作为特征建模的基础。在此基础上对基于本体的领域分析和建模过程进行了研究,提出了一种领域分析和建模方法 OBDA (Ontology Based Domain Analysis)。该方法使用专家分析和自动分析方法相结合,分析业务动作之间的分解、泛化、刻面和依赖等关系以及各种关系的可变性,从而得到完整的领域特征模型。

关键词 领域建模 领域分析 本体

RESEARCH ON ONTOLOGY-BASED DOMAIN FEATURE MODELING PROCESS

Gu Jianhua Zhao Wenyun Peng Xin

(Department of Computer Science and Engineering, Fudan University, Shanghai 200433, China)

Abstract Research and practice on software reuse indicate that specified domain software reuse activity can be successful easily, and this comes out of the cohesion and stableness of the domain. Systematic and comprehensive domain analysis is the prerequisite and key step of domain-level reuse activity. Ontology is introduced as the base of feature modeling in the research of domain analysis and modeling. OBDA (Ontology-Based Domain Analysis), a domain analysis and modeling method is presented, which combines domain experts analysis with automation analysis. The decomposition relation, generation relation, facet relation, and dependence relation between business actions and the variables in these relations are analyzed to obtain the intact domain feature model.

Keywords Domain modeling Domain analysis Ontology

0 引 言

软件复用的研究和实践表明,特定领域的软件复用活动相对容易取得成功^[1]。而领域工程是其中的关键,即可复用软件资产(包括体系结构和构件等)的生产阶段,主要包括领域分析^[1]、领域设计和领域实现这三个活动。领域分析是对特定的领域进行需求工程的活动。它覆盖了对领域需求的获取、分析、规约和检验/验证的整个过程^[2]。

建立基于本体的模型依赖于专家分析和模型自动识别相结合的方法,分成两个阶段:建立领域特征层次关系模型、识别领域特征依赖关系。其中第一阶段又分成三个步骤:(1)识别捕获业务动作;(2)识别业务动作刻画;(3)识别模型泛化关系。

1 相关工作

1.1 领域分析相关方法

领域分析中最早提出面向特征的领域分析的方法是 FODA (Feature-Oriented Domain Analysis)^[3],并且由 FODA 方法逐渐演化成基于产品线(Product Line)的软件复用方法^[2]。韩国浦项科学与技术大学在 FODA 基础上提出了面向特征的复用方法 FORM (Feature-Oriented Reuse Method)^[4],惠普实验室将 FODA 方法与 RSEB (Reuse-Driven Software Engineering Business) 结合

提出的 FeaturSEB 方法^[5]。

国内在这个领域内的研究最多的是北京大学软件工程研究所,他们提出了面向对象的领域工程方法^[1]。在领域需求分析方面,北京大学软件工程提出了基于需求聚类建立领域模型的方法^[6]。能够在客户需求的基础上,通过聚类算法自动识别特征模型树。

上海复旦大学软件工程实验室将领域工程与本体相结合,提出了基于本体的软件构件开发方法。包括基于本体的领域分析,领域建模,领域构件识别和组装。本文提出的方法是基于本体的领域需求分析建立特征模型的方法。

1.2 本体相关概念

使用本体作为领域模型建模语言有两个原因:

(1) 本体本身良好的表述能力。通过三元组的描述,能够化繁为简,将模型间复杂交错的关系网,转化成三个三元组的集合,每一个三元组用最直观和最简明的方式表述领域内的一条关系。

(2) 本体具有推理能力,通过定义本体推理规则,本体能够根据现有的三元组,推导出新的三元组。利用本体推理能力,领域建模师能够完成复杂的模型检查。

收稿日期:2006-04-28。国家 863 计划项目(2005AA113120),上海市科委项目(055115005),上海市科委科技攻关项目(04DZ15022)。
顾剑华,硕士,主研领域:软件工程,领域分析。

2 基于本体领域建模相关定义

领域本体是用于描述指定领域知识的一种专门本体,可以用于构建特定领域业务知识体系,为领域内的构件语义提供描述基础。

基于本体的特征元模型如图1所示。在基于本体的领域模型中这些特征被进一步细分为业务动作、动作刻面和术语等建模元素。相关概念定义如下,基于本体的特征建模方法参见我们的相关工作。

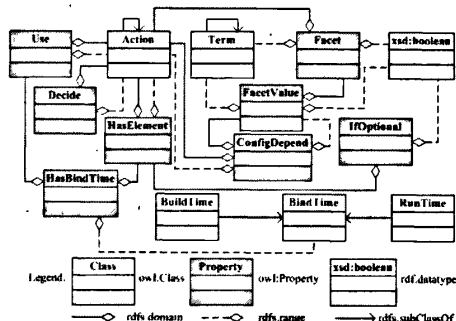


图1 基于本体的特征元模型

定义1 业务动作 (Action owl:Class) 是业务动作的语义主体 (即动作化描述的特征)。

定义2 动作刻面 (Facet owl:Property), 是业务动作的语义刻面值 (即特征属性)。

定义3 术语 (Term owl:Class), 是刻面的取值 (即特征值)。

定义4 子关系 (SubClassOf owl:Property), 描述业务动作之间或术语之间的父子关系。

定义5 分解关系 (HasElement owl:Property), 描述业务动作之间的分解组合关系。

定义6 泛化关系, 业务动作之间的父子关系。

定义7 调用关系 (Use owl:Property), 描述业务动作之间的调用关系。

定义8 决定关系 (Decide owl:Property), 描述业务动作之间对于领域可变性的确定关系。

定义9 绑定时间关系 (HasBindTime owl:Property), 描述领域内可变性的确定时间。

定义10 刻面值关系 (FacetValue owl:Property), 不同的业务动作在同一动作刻面上可能取不同的值, 刻面值关系由一个业务动作和动作刻面唯一决定。

定义11 配置依赖关系 (ConfigDepend owl:Property), 描述业务动作、刻面值关系之间相互的依赖决定关系。

3 ODBA 建模方法

ODBA 建模方法分成两个阶段: 领域特征层次关系建模和领域依赖关系模型。第一阶段建立领域需求模型的产品是由业务动作、动作刻面、术语以及三者之间分解、泛化、动作刻面关系组成的领域需求层次模型。第二阶段调整领域模型的产品是在层次模型的基础上, 分析调用关系、决定关系、绑定时间关系等关系, 形成完整的领域模型。

3.1 第一阶段建立领域特征层次关系模型

3.1.1 步骤1: 识别捕获业务动作

识别捕获业务动作的方法有两种: (1) 通过分析应用项目获得, 通过分析已经存在的领域内应用项目的业务动作, 来获得领域业务动作; (2) 基于领域专家的分析。基于领域专家的分析是由领域专家进行归纳总结, 分析领域内应该存在的业务动作, 然后由建模人员进行建模。

在 ODBA 中识别捕获业务动作, 首先识别领域业务动作树的根节点, 每一个领域都由一个领域业务动作根节点开始, 领域业务动作根节点从最粗的粒度描述了领域内最基本的业务功能。其次从根节点开始进行分解, 对已经存在的领域内应用项目进行分析, 结合领域业务专家的意见, 将根节点按照需求功能划分为领域业务动作树的第二层业务动作节点。第二层业务动作节点之间是并列关系, 与第一层节点之间是分解关系, 第二层业务动作通过合作能够完成顶层的业务动作功能。然后采用递归的方法, 分别对业务动作树中存在的业务动作节点进行分解, 直到业务动作树中的业务动作叶节点已经达到建模人员要求的粒度为止。

在建立领域业务动作树的过程中, 在将业务动作加入到领域业务动作树之前, 必须对将该业务动作与领域辞典内已存业务动作进行比较, 如果不同则在业务动作加入到业务动作树之后, 将该业务动作输入到领域辞典中。避免由于对领域内业务动作不同的理解而产生的术语不统一问题。

建模人员确定模型底层业务动作粒度的依据主要根据业务动作的可变性原则, 如果一个叶子业务功能只可能成为父节点的子功能, 与其他业务动作之间不存在其他依赖关系, 并且在调用的时候该业务动作本身不存在变化性, 而它的父亲节点也同样是没有可变性的业务节点, 则该业务动作粒度过细。如果一个叶子业务动作还能够分解成若干个子业务动作, 并且存在具有可变性的子业务动作, 根据可变性从细粒度到粗粒度提升的原则, 该业务动作也具有可变性, 则该业务动作粒度较粗, 可以被进一步分解。

如图2中业务动作1的粒度比较大, 在领域建模的时候, 需要被进一步地分解为业务动作2和业务动作3。业务动作2能够被进一步分解为业务动作4和业务动作5, 如果业务动作4和业务动作5只能作为业务动作2的一部分被调用, 则业务动作2就不需要被进一步分解为业务动作4和业务动作5。因为业务动作4和业务动作5不可能被分别生成构件, 而是和业务动作2一起作为一个构件。同理对于业务动作3。在图2中, 以业务动作2和业务动作3作为叶子节点是比较合适的粒度。

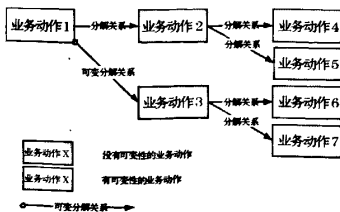


图2 业务动作分解粒度示意图

3.1.2 步骤2: 识别领域业务动作刻面

识别领域业务动作刻面是建立在领域业务动作层次模型基础上的, 由领域专家结合领域应用项目, 对业务动作的变化性进行分析后, 将变化性用动作刻面和术语的形式进行描述。识别

领域业务动作刻画可以从上而下进行,也可以从下而上进行。

从上而下进行的业务动作刻画识别是从业务动作树根节点开始,直到业务动作树叶节点为止。对于抽象的业务动作变化性可以通过头脑风暴法组织领域专家进行场景描述,分析不同场景下,业务动作之间的变化性和属性,或者通过对现有领域内应用的归纳分析和对领域未来发展的分析提取业务动作的变化性和属性。

从下而上进行的业务动作刻画是应用动作刻画识别方法进行自动识别,从业务动作树叶节点开始,分析领域业务动作的变化性。与从上而下进行的业务动作刻画识别不同,业务动作树父节点不仅要分析本身的动作刻画,还需要分析子节点的动作刻画,如果子节点所拥有的动作刻画属于父节点典型动作刻画,则动作刻画需要上浮。

从下而上进行的业务动作刻画上浮选择典型动作刻画方法有两种:数量比例法和关键节点法。

(1) 数量比例法。对于 Action,令 $Num = Action$ 拥有的分解关系子节点数量,令 $NumFacet[i] = Action$ 的分解关系子节点中拥有动作刻画 $Facet[i]$ 的数量。如果 $NumFacet[i]$ 与 Num 的比值大于动作刻画上浮阈值,则动作刻画 $Facet[i]$ 也是业务动作 Action 拥有的动作刻画。如图3所示:业务动作1被分解为业务动作1.1、业务动作1.2、业务动作1.3,业务动作1.1拥有动作刻画1、动作刻画2、动作刻画3,业务动作1.2拥有动作刻画2,业务动作1.3拥有动作刻画3。对于业务动作1来说,被分解成3个业务动作,3个子业务动作和3个动作刻画相关:Num动作刻画[1]为1,Num动作刻画[2]为2,Num动作刻画[3]为2。如果设定动作刻画上浮阈值为50%,则Num动作刻画[2]、Num动作刻画[3]与业务动作1分解数3之间的比值大于50%,所以动作刻画2和动作刻画3也是业务动作1的动作刻画,而动作刻画1就不需要上浮成为业务动作1的动作刻画。

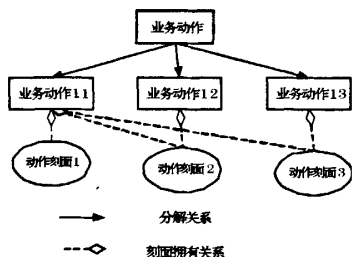


图3 业务动作刻画上浮数量比例法

(2) 关键节点法。子节点中有一些节点是本身节点的关键节点。如果关键子节点拥有某一动作刻画,则该动作刻画需要上浮。关键节点的选择依据是判断子节点是否描述了父节点的核心功能。例如,评分的业务动作可以被分解为:试卷准备、试卷评分、评分分析3个子业务动作,其中试卷评分和试卷准备要根据不同类型的试卷(图片类、语音类)分别存在不同类型的试卷准备和试卷评分,而评分分析和试卷的类型无关,所以试卷准备、试卷评分可以被看作是评分节点的关键节点,它们的动作刻画需要上浮到评分业务动作上,而评分分析则不属于关键节点,所以它的动作刻画不需要上浮。

3.1.3 步骤3:识别领域模型泛化关系

在识别领域动作刻画后,需要识别领域内的泛化关系业务动作。识别领域内的泛化关系方法也有两种,自动识别法和专家分析法。两种方法都是从领域中找出相对重要的动作刻画,

然后根据动作刻画下术语的变化性,产生泛化关系业务动作。

自动识别法:相对重要的动作刻画在领域内的影响力也比较大,表现为有较多的业务动作拥有该动作刻画。通过遍历领域内所有业务动作,能够计算出动作刻画在领域内的影响力。例如,对于业务动作A1,如果A1相关动作刻画有F1、F2和F3,则F1、F2、F3在领域模型中的影响力EffectF1、EffectF2和EffectF3分别累加1。在遍历完所有的业务动作后,如果动作刻画F1影响力大于设定的泛化关系识别阈值,则把F1认为是领域内关键动作刻画,并在对于所有该动作刻画所包含的业务动作根据动作刻画下术语变化性进行泛化关系分解。

专家分析法:专家分析法是通过领域专家对领域内现有应用项目、领域发展方向等的分析,标识领域内相对比较重要的动作刻画。

自动识别法和专家分析法的不同点在于,自动识别法识别速度较快,可以通过软件算法完成,设定不同的泛化关系识别阈值可能获得不同的泛化关系,然而自动识别法识别出的泛化关系算法可能不具有现实应用意义。专家分析法分析的速度较慢,且不能通过软件辅助方法解决,但是分析的结果准确性比较高,具有现实应用意义。

3.2 第二阶段调整领域需求模型

建模第一阶段只能够完成分解关系、泛化关系和动作刻画,不能够完成其他关系的生成。建模人员可以在已经生成的领域模型基础上,根据领域内业务动作之间实际的含义添加决定关系、调用关系和依赖关系。由于这三种关系比较复杂,难以利用系统自动分析方法,而且三者之间没有必然联系,在第一阶段完成领域模型框架后,建模人员必须根据领域专家的意见,分析领域内不同业务动作之间的交互关系,从而完成完整的领域模型。

3.3 ODBA 建模原则

- 领域现有应用和领域发展分析相结合

领域内的现有应用情况代表领域现有状况,而对领域发展的分析则能够对领域未来的发展做出预测。仅仅分析前者使领域模型缺乏弹性,很难演化出除领域内现有应用以外的应用。而脱离领域应用的领域模型将失去对领域应用的指导作用。进行领域建模的时候,建模人员应该包括领域专家和领域建模师。其中领域专家提供领域内应用情况、提供领域发展趋势,领域建模师根据领域专家的意见,进行领域建模。

- 专家分析法和系统半自动法相结合

专家法能够对领域现有的状况得出一个概括性的分析,分析的准确度较高,但是分析速度较慢;系统自动分析法根据模型内部的数据进行分析,能够帮助模型建立人员快速地获得模型,但是分析精度不高,常常生成无实际意义的业务动作或动作刻画。有效地结合专家分析法和系统自动法,能够综合两种方法的优势,以比较少的成本,获得比较精确的模型。

4 应用算法例子研究——阅卷领域建模

阅卷领域包括传统试卷的阅卷和英语口语上机考试的阅卷两大类,一般阅卷领域程序会包括以下一些功能。

GradeDomain: 阅卷领域的根节点。

Prepare: 阅卷准备工作,包括试卷收集入库、试卷分配。

Grade: 阅卷给分工作,包括试卷下载、给分、评分上传。

Review: 阅卷分析汇总工作。

(下转第34页)

参考文献

- [1] 何蓓, 李辉. Delphi 应用程序与 Matlab 接口技术的实现及应用[J]. 计算机系统应用, 2000(8): 40-42.
- [2] 程云艳, 陈桦. 用 Delphi 面向对象开发模式实现 Matlab 的调用[J]. 计算机时代, 2004(4): 30-31.
- [3] 韩守红, 唐力伟, 等. Delphi 和 Matlab 接口的实现方法研究[J]. 微计算机信息, 2001, 17(10): 45-46.
- [4] 韩守红, 唐力伟, 等. Delphi 和 Matlab 动态数据交换技术的实现[J]. 微计算机信息, 2002, 18(1): 66-67.
- [5] 游佳, 何健鹰. Delphi 与 Matlab 接口以及脱离 Matlab 运行[J]. 计算机与数字工程, 2004, 32(6): 21-23.
- [6] 岳玉芳, 尤忠生, 等. 基于 COM 的 VC 与 Matlab 混合编程[J]. 微机发展, 2005, 15(5): 47-48.
- [7] Matlab GOM Builder user's Guide [Z]. The Mathworks Inc, 2002.

(上接第 9 页)

PaperPrepare: 试卷收集入库工作。

PaperAssign: 根据试卷数量和老师数量分配阅卷工作。

PaperDownload: 阅卷老师下载属于自己批阅的试卷。

PaperGrade: 阅卷老师批阅试卷。

ScoreUpload: 阅卷老师上传已经批阅完成试卷的分数。

FinalDeal: 阅卷后期处理工作。

ScoreStatics: 阅卷分析工作, 包括按照试题的分数汇总、按照学生的分数排序、批阅老师的批阅速度及准确率排序。

在进行领域建模的时候, 首先识别业务动作, 从上到下递归地建立领域业务动作分解树。如图 4 所示。

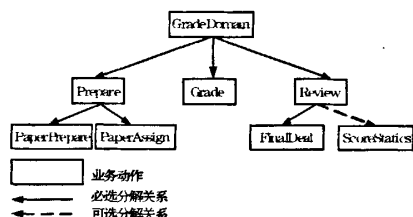


图4 阅卷领域业务动作树

然后识别业务动作刻画面, 这里采用从下到上递归半自动地搜寻刻画面。从分解关系树的叶子节点开始搜寻动作刻画面。结果如图 5 所示。

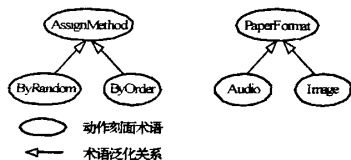


图5 阅卷领域动作刻画面示意图

最后采用自动识别法识别模型泛化关系, 遍历领域词典中所有的动作刻画面: PaperFormat 和 AssignMethod。在增加完所有的泛化关系业务动作后, 在按照分解关系中的层次关系, 连接泛化关系业务动作之间的分解关系。

5 工具支持

为了能够有效地支持给予本体领域建模, 我们实现了领域

建模工具 OntoFeature。该工具能够通过图形化的界面, 帮助领域建模师完成基于本体领域建模、分析工作。如图 6 所示。

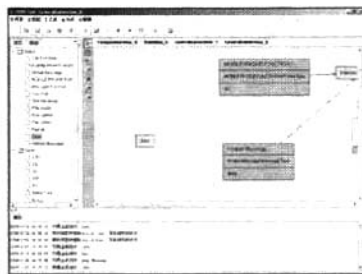


图6 OntoFeature 界面

工具支持四种视图, 泛化视图、分解视图、规则视图和综合视图。四种视图共享一个领域模型, 具有内部联系。工具能够支持建模元素增加、修改和删除过程中自动检查, 也能够支持在模型建立完成后的模型合理性检查。

6 结束语

本文提出了一种结合本体和领域分析的方法 ODBA, ODBA 解决了一般特征建模中存在关系非一阶实体的问题, 在 ODBA 中可变性不仅能够出现在特征上, 而且能够分析特征或业务动作之间关系的可变性, 使关系成为特征模型中的一阶实体。

在本文中主要提出了如何利用 ODBA 进行建模的步骤和方法, 分成模型建立和模型调整两个阶段, 其中模型建立阶段又分成: 识别业务动作、识别业务动作刻画面和识别模型泛化关系 3 个步骤。建模人员在领域专家的帮助下, 运用本文提出的半自动建模方法, 可以比较简单地完成模型建立阶段。

在 CASE 工具 OntoFeature 的帮助下, 用户能够进行图形化 ODBA 建模, 并且在建模后依靠本体推导完成从领域建模生成领域本体文件(owl 文件), 领域本体模型文件可以直接应用到领域体系结构的设计中, 完成领域体系结构的分析和设计。

参考文献

- [1] 张伟, 梅宏. 一种面向特征的领域模型及其建模过程. 软件学报, 2003, 14(8): 1345-1356.
- [2] Gray Chastek, et al. Product Line Analysis: A Practical Introduction. SET Technical Report CMU/SEI-2001-TR-001, June 2001.
- [3] Kyo C Kang, et al. Feature-Oriented Domain Analysis Feasibility Study. SEI Technical Report CMU/SEI-90-TR-21, November 1990.
- [4] Kyo C Kang, et al. FORM: A feature-oriented reuse method with domain-specific architecture. Annals of Software Engineering, 1998, V5: 143-168.
- [5] Martin L Griss, John Favaro, Massimo d'Alessandro. Integrating Feature Modeling with the RSEB. in Proceedings of 5th International Conference on Software Reuse, Victoria, Canada, June, 1998, IEEE: 76-85.
- [6] Kun Chen. An Approach to Constructing Feature Models Based on Requirements Clustering. Proceedings of the 2005 13th IEEE International Conference on Requirements Engineering (RE'05).
- [7] Peng Xin, Zhao Wenyan, Xue Yunjiao, Wu Yijian. Ontology-Based Feature Modeling and Application-Oriented Tailoring. Proceeding of the 9th International Conference on Software Reuse, ICSR2006, Torino, Italy.