《智能信息处理》课程考试

基于本体的项目知识集成

王雪莹

考核	到课[10]	作业[20]	考试[70]	课程成绩[100]
得分				

基于本体的项目知识集成

王雪莹

(大连海事大学信息科学技术学院 辽宁大连 110319)

摘 要:针对产品生命周期中知识表达和集成的需求,提出一种基于中间层的产品知识集成框架,采用本体技术,具有显式定义语义、统一表达和动态集成产品知识的能力 .元本体层定义了本体关系原语,提供通用的本体描述方 法;本体层扩展元本体为共享本体、领域本体、过程本体和应用本体,支持本体的协作定义和集成映射;知识层提供 了相应的逻辑结构,支持对知识实例的结构化、标准化管理,并给出使用Web 本体语言(OWL)形式化表达的方法.以汽车减振器为实例说明该方法能有效提高产品数据管理(PDM)系统对知识集成的能力.

关键词:本体:元本体:映射:知识集成:生命周期

Ontology based project knowledge integration

Wang xueying

(Institute of Computer Science and Technology, Dalian Maritime University, Dalian 110319, China)

Abstract: Aimed at the requirements of know ledge representation and integration in product lifecycle, a framework of product know ledge integration based on middle-layer was proposed. Using the techniques of onto logy, the framework has the abilities of explicit definition of semantics, uniform representation and dynamic integration of product know ledge. In the meta-onto logy layer, relationship primitives in onto logy are defined to provide a generic representation method of ontology. In the ontology layer, meta-ontology is extended to share ontology, domain ontology, process ontology and application onto logy to support the collaborative definition and the integrated mapping of ontology. In the know ledge layer, corresponding logical structures are provided for the structured and standardized management of know ledge instances, and the methods of using Web ontology language(OWL)for formal representation are given. The application of a shock absorber demonstrated that the approach can improve the product data management(PDM) system of integrating know ledge.

Key words: ontology; meta-ontology; mapping; knowledge integration; lifecycle

0 引言

激烈的市场竞争、经济的全球化发展给当今制造企业带来了新的挑战,企业信息化和知识化是提高产品开发效率和质量的必然途径.协同的产品信息集成机制、高效的产品数据管理环境、有效的产品开发知识支持是其中需要解决的关键问题[1].已有的产品数据管理(product data management, PDM)、计算机辅助技术(computer aided X, CAX)、供应链管理(supply chain management, SCM)、企业资源计划(enterprise re

source planning, ERP)、客户关系管理(customer relationship management, CRM)等已部分实现了信息化的产品管理与开发[2],并获得巨大的经济效益,但作为单元技术其信息表达和共享层次低,只侧重于几何数据表达和产品某阶段基本信息管理,缺乏对产品生命周期如概念设计、工艺设计等阶段知识集成的支持,而这些知识对产品重用与创新却是至关重要的;另外,各个应用系统之间信息共享困难,不能满足现代企业集成产品开发的需求[3].因此对产品知识集成与共享的研究具有重要意义[4],本文引入知识工程中本体的概念[5],并基于已有的汽车减振器 PDM 系统,扩展其

Vol.25 No.2

Dec. 2020

知识表达和集成功能,构建可动态扩展、可重用、可共享、高度一致的产品知识集成框架[6].

1本体概述

信息是组织或结构化的数据,知识是语义化 的信息,为了支持计算机对信息的语义理解和自 动处理,本体这一哲学术语被引入知识工程,它能 够支持知识的表达、处理和交换.本体是对共享概 念模型明确的形式化规范说明,其内涵包括以下 4 个方面:本体是客观世界本质上的抽象概念模 型:概念及其关系具有明确含义和精确定义:可形 式化被计算机处理;本体知识是使用者共同认可 的.本体的这些特性使它在产品知识集成中具有 显著优势,不仅可以支持标准化的知识表达和动 态的数据结构,还易于构建业务逻辑以实现知识 重用和推理.现代产品开发面临着网络化和分布 式的环境,本体在 Web 上的应用可以实现 Web 环 境下信息语义上的互操作和智能应用,为 Web 环 境下产品知识集成提供了基础.本体是知识工程 和人工智能方面研究的热点,但它在工程领域的 研究与应用还处于初步阶段,本文结合本体及其 描述语言OWL来完善现有产品数据管理系统在 知识集成方面的不足,主要有下面几点:

1)产品开发各阶段的信息互相交叉、重叠、 关联和依赖,本体对产品元数据及其关系语义进 行规范定义,可以从根本上保证产品信息语义级 的一致性,避免不一致引起的工程错误和数据冗 余.

2)具有柔性的数据结构,支持产品开发过程数据结构的动态更改和扩展.

3)使用机器可读、基于逻辑的语言表达,便于知识重用和推理,为产品开发应用提供知识支持,构建基于逻辑的智能化应用.

4)支持 Web 应用,便于分布式系统和异构系统的知识共享和交换.

2 产品知识集成框架

2.1 框架概述

基于中间层的产品知识集成是一种开放式的集成方式,它在数据源和多个应用程序之间建立了统一的中间层,具有统一的知识表达和构建业务逻辑能力,屏蔽了底层数据处理的复杂性,面向新应用提供了通用的集成平台.对于中间层具有很高的要求.既要易于扩展新的应用集成,又要

能够尽可能覆盖以往遗留系统的结构.文献[7]中建立了一个构建于语义 Web 之上的中间层,并进一步划分为本体层和信息访问层,重点研究了EXPRESS 模式到 DAML+OIL 的语义映射方法,但仍需进一步详细研究中间层的表达和实现方法.本文在以上研究的基础上,重点研究了以本体集成为基础的中间层的构建和表达,并以 PDM 作为集成平台,在广度上延伸其生命周期,在深度上扩展其知识表达能力,为产品生命周期提供知识集成的能力,如图 1 所示.

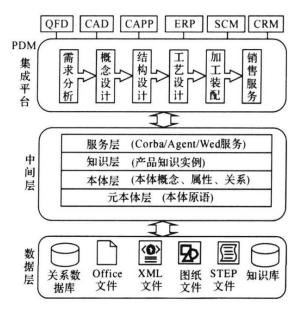


图 1 产品知识集成框架

Fig.1 Framework of product knowledge integration

整个框架分为 3 层,数据层集成了多种数据源,它们是产品开发中各种信息和知识存储的地方,面向不同的应用,应用不同表达和存储结构也各不相同.中间层需要尽可能的覆盖这些异构的数据结构,提供一致的表达方法来描述这些数据源的语义,它由元本体层、本体层、知识层和服务层组成:元本体层提供通用的本体描述方法;本体层定义产品生命周期相关本体,提供本体集成管理机制、知识层表达和管理知识实例;服务层基于现有一些成熟的信息技术,如 CORBA、Agent 技术、Web 服务等提供多种分布式的访问模式[7],它们共同服务于面向产品生命周期和由

PDM集成的应用平台层.PDM系统管理和集成了产品从需求到销售、回收等整个生命周期的过程、信息和知识,各阶段有各种相应的支持工具辅助开发,并由中间层提供统一知识共享和集成

服务.可见,中间层在知识集成框架中起着承上启下的关键作用,它通过语义显式表达和本体集成多种数据源和应用,提供统一的知识表达和服务,以下对其关键技术作详细描述.

2.2 元本体层

元本体是具有普遍适应性和通用性的抽象概念模型,从最底层显式且明确地描述了客观世界概念间基本关系和表达方法,它是用来描述本体的基本原语,可以扩充描述产品各领域本体的概念及其关系,并可通过OWL实现形式化表达和共享.

定义 1 元本体可表达为一个五元组:O=(C,P,I,R,F).其中:C 是概念集;P 是属性集;I 是概念实例集;R 是关系集;F 是公理和规则集.

元本体体现为与领域无关的一种通用描述 机制.C、P、I 是 3 个核心元素,C 是对 I 子集在概念上的抽象和分类,I 是 C 中概念的具体实现,P 可对 C 中概念进行详细的特性描述,并通过属性赋值来实例化概念.R 和 F 表达了这些元素之间的关系语义和约束规则.元本体层体现为概念层通用的关系原语,如:分类关系、组合关系、关联关系、映射关系、约束关系等,这种显式的语义表达为建立语义级的知识集成和可互操作的业务模型提供了基础,是构成本体通用的表达方法和约束规则.

2.3 本体层集成

本体层在基于元本体的表达方法和语义约 束下,形成概念集、属性集和关系集,这需要解决 两个关键问题:本体定义和管理.本体建模是一项 复杂的系统工程,目前有自顶向下和自底向上两 种建模方式.自顶向下的方式系统性强,一致性好, 但对建模人员要求高,扩展性差;自底向上的方式 扩展性强但可维护性差,它面向大型应用时容易 产生一致性问题.而知识系统是一个不断积累、更 新的动态系统,它面向不同的领域、过程和应用, 需要领域专家、技术人员和用户共同完成,这里把 本体层从功能和应用的角度分解为共享本体、领 域本体、过程本体和应用本体,在统一的元本体描 述框架下由各合作者共同完成产品生命周期相 关知识本体定义,可结合自顶向下和自底向上两 种建模方式的优点,通过本体映射和进化保持知 识的一致性和不断更新.本体层集成如图 2 所示.

2.3.1 共享本体

由系统设计师定义的面向产品全生命周期

的全局本体,具有全局唯一性,主要定义了产品最 基本的工程知识和管理知识相关本体,可嵌套递 归形成树状的本体分类体系.工程知识是产品生

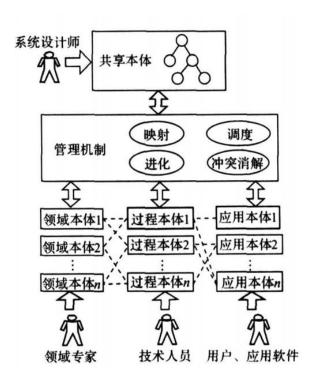


图 2 本体层集成 Fig.2 Integration in ontology layer

命周期中相关的技术知识,如产品、部件、零件、特征、功能、形状、材料等;管理知识是组织开发和生产相关的知识,如项目、任务、人员、角色、权限等.

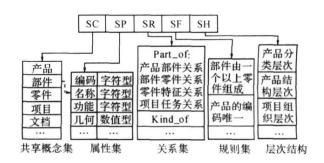


图 3 共享本体 Fig.3 Sharing ontology

共享本体主要用于定义通用本体和维护本体的一致性,在元本体的基础上,扩展其概念、属性和关系,形成产品知识的共享本体,将其定义

为:SO=(SC,SP,SR,SF,SH).如图 3 所示,其中 SC 表示共享本体概念集;SP 表示其属性集,图中示意了部件的部分属性;SR 表示关系集,细化了元本体中的基本关系;SF 表示公理和规则集,定义了产品集成开发和管理的一些基本规则;SH 表示它们的层次结构,它实际上是元本体中关系和规则的组合,这里把它作为产品知识的一种基本表达和组织方式.共享本体是抽象的概念描述,不包含实例描述,元本体中提及的实例集在下节中的知识实例层中得以体现.

2.3.2 领域本体

由领域专家定义的面向某一领域的通用知识本体,领域本体是与具体产品无关的领域通用概念,它基于本体协作定义的思想,考虑到本体建模涉及多领域、多学科,既方便了各个领域专家分别定义,也可作为企业的一种可重用知识储备,服务于多个产品和其他企业.领域本体在元本体的基础上增加了领域的范畴,产品生命周期过程中涉及多个领域的知识,如工程领域、信息科技领域、生产管理领域等,大领域还可以再划分为多个子领域,如结构设计领域、工艺设计领域、加工制造领域等.不同于共享本体,领域本体可包含各领域知识的常用实例,其目的是为了便于对知识的理解和典型实例的重用.图 4 示意了典型机械产品生命周期中涉及到的领域及其概念本体.

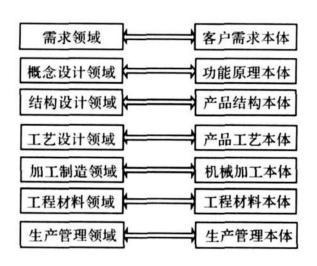


图 4 领域本体 Fig.4 Domainontology

2.3.3 过程本体

由技术人员定义的面向具体产品生命周期 过程的知识本体,如某类产品设计阶段的常用术 语和概念规范化的定义.通常一个过程会用到几个领域或其他过程的知识,通过管理机制可以调度领域本体和其他过程本体以供使用.基于元本体的本体构建模式,并参考产品相关业务流程的自身特点,可统一将业务流程概念化描述为一系列过程单元(processunit,PU)及它们之间拓扑关系的组合,体现了过程间的本质关系.如图5,PU包含4个关键元素:输入、输出、事件和参与人员,PU之间的拓扑结构可分为顺序结构、反馈结构、选择结构(or、xor)、并行结构(and)等基本结构,通过这些基本结构可组合复杂的业务流程.

2.3.4 应用本体

用户的知识背景和应用软件对知识的表达 千差万别,应用本体通过语义标注并集成这些不 同的应用知识源.应用本体可采用本体引用在已 有本体基础上构建,并通过本体映射可将表达同 一语义的不同应用本体映射到同一个共享本体、 领域本体或过程本体上,而对于应用本体语义上 的冲突,可通过冲突消解和本体合并以达到语义 上的一致理解.本体集成要解决以上本体间的异

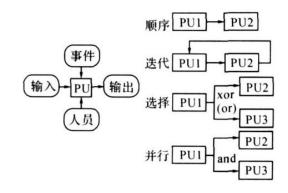


图5 过程本体

Fig. 5 Process ontology

图 5 过程本体 Fig.5 Processontology

构,实现互操作,主要可采用以下几种策略:

1)本体引用.产品在其生命周期中处于复杂的知识网络中,这些知识可能是跨学科、地域和组织的,通过本体引用可避免重复定义,集成并重用这些已有的知识.目标本体可简单引用源本体,源本体中所有概念都可在目标本体中使用,如共享本体、领域本体就经常在过程本体构建中被引用.本体引用要解决的关键问题是分布式环境下的

本体定位和声明,OWL 采用通用资源标志符 (universal resource identifier, URI)和命名空间提供了 Web 环境下的引用机制.

2)本体映射.在元本体层中提到映射是产品 知识对象间的一种基本关系,本体层基于这种映 射关系来解决以上各种本体间的结构和语义异 构.扩展映射定义为

Mapping (E1, E2, F, T, V).

式中:E1 和 E2 表示源本体和目标本体中的元素 集合,可以是一个或多个概念、属性和关系;F表示 映射函数,是映射关系的表达式:T 表示映射的类 型,主要包括概念之间、属性之间的映射,实际应 用环境中由于概念、属性、关系之间不同的定义 和划分,也可能存在概念、属性和关系之间的相互 映射,根据映射元素数量可分为1:1、1:n、m: n 等 3 种:V 表示映射的信任度,可归一化取值为 (0,1),主要用来度量映射元素间的相似性.产品知 识集成中知识表达形式多样,各应用系统本体间 存在多种结构、语法和语义上的映射关系,需要采 用多种策略支持多种类型的映射建立.如图 6 所 示.映射表存储了本体间元素的映射关系.它通过 交互映射、半自动映射和自动映射三种方法针对 不同的情况来获得.交互映射主要是在各本体由 相应人员建立时,根据应用需求和自身知识手工 建立的映射关系,如:专业术语在不同环境中的映 射,当前应用为集成其他应用而建立的概念、属

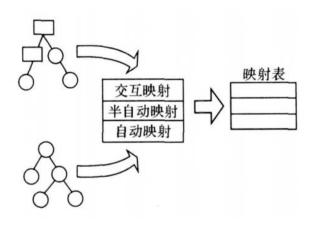


图 6 本体映射 Fig.6 Ontology mapping

性映射关系.半自动映射由设计人员预先定义建立映射条件的启发式规则,在规则的约束下可辅助建立一些映射关系,如:整机厂和配件厂对部件编码时规则的对应关系;在建立功能与结构的映

射关系时,根据具体产品可预先限制功能、结构的 类型和关键属性,能有效降低本体映射算法的复 杂性,提高映射准确性.自动映射根据元素名称、 结构相似性来自动建立映射关系,可采用比较成 熟的算法,如字符串匹配、图结构匹配来实现.

2.4 知识实例层

元本体层和本体层为产品知识提供了抽象的表达和集成机制,知识层则面向产品生命周期为具体的产品知识实例提供有效的组织管理和形式化表达.采用Web本体语言OWL可以增量式和分布式地定义知识,便于Web环境下的知识集成和共享.

1)基于树结构的分类管理.基于分类关系的 嵌套递归形成树结构的分类体系及产品知识实 例的组织方式,实例位于树的叶子节点,它们的上层节点是和本体直接对应的概念节点.其规则描述如下:

规则1对于某个概念C1,它是概念C2的父类,同时C2又是概念C3的父类,则C1也是C3的父类,则C1也是C3的父类.可逻辑描述为

$(\lor C1,C2,C3)$ Kind_of(C1,C2) \lor Kind_of(C2,C3) \rightarrow Kind_of(C1,C3)

规则 1 是构建概念分类系统的基础,为知识集成系统在管理级上提供产品数据的分类管理,在技术级上为设计人员和智能系统提供知识导航.

规则2对于某个概念C1,它是概念C2的父类,同时实例 I1 又是 C2 的一个实例,则 I1 也是 C1 的一个实例.即

(C1,C2,I1)Kind_of(C1,C2) \vee Ins_of(C2,I1) \rightarrow Insof(C1,I1)

规则2反映了知识实例与不同抽象层次概念之间的联系.管理层面上可用于产品实例数据不同层次的归类,技术层面上可支持知识实例不同粒度的重用.

具体应用包括面向开发的产品分类树、零部件分类树、文档分类树、工装分类树、工艺分类树等;面向管理的项目分类树、部门人员分类树、任务分类树、角色分类树等.OWL通过关键字Class和Sub Class Of表达分类关系,并以XML格

式记录实例数据,部分片断如下:

<owl:Class rdf:ID= "GG 型 A 系列减振器">
<rdf:subClassOf>
<owl:Class rdf:ID= "A 系列减振器"/>
</rdf:subClassOf>

2)基于树结构的产品组成结构.基于组合关系的嵌套递归形成树状的产品组成结构本体,它体现了产品、部件、零件、特征之间的内在组成关系.其规则可描述如下:

规则3对于某个概念C1,概念C2是其组成部分,同时概念C3又是C2的组成部分,则C3也是C1的组成部分,即

(C1,C2,C3)Part_of(C1,C2) \vee Part_of(C2,C3) \rightarrow Part_of(C1,C3).

规则 3 反映了复杂概念分解和组合的层次关系,是产品知识集成系统中产品组成结构表达和应用的基础,此规则对产品实例同样适用.产品结构贯穿于产品的整个生命周期,基于产品结构本体描述的语义关系,可以派生和映射出各阶段的产品 Bom 实例数据,如设计 Bom、工艺 Bom、销售 Bom 等.OWL 中没有直接的关键词和 Part of相对应,但可以通过 Object Property 自定义,并可通过 Sub Property Of 来区分是产品、部件间的组成关系还是部件、零件间的组成关系,部分片断如下:

<owl: Thing rdf: about="GG 型A 系列减振器">

< GG 型 A 系列减振器: hasSubComponent rdf: resource="活塞杆总成"/>

< GG 型 A 系列减振器: hasSubPart rdf: resource="工行缸"/>

</owl: Thing>

7 own: Timing

3)基于表结构的属性管理.属性是描述对象特性的元数据,最简单的方式是通过对象、属性、属性值三元组来定义,一个知识对象可通过一组属性组成的表结构来组织和表达,属性依附于对象类,属性值依附于具体的对象实例.基于规则 4可建立分类树上节点间属性的继承和扩展关系,

其规则可描第 180510 期江伟光,等:基于本体的产品知识集成述如下:规则 4 对于某个概念 C1,它是概念 C2 的父类,同时 P1 又是 C1 的一个属性,则 P1 也是 C2 的一个属性.即

(C1,C2,P1)Kind_of(C1,C2) \vee Proof(C1,P1) \rightarrow Proof(C2,P1).

规则 4 反映了属性在不同抽象层次的概念之间的继承和扩展关系,可用于描述产品特性的通用性和差异性,便于知识表达和重用.OWL 中可通过 DataTypeProperty 来定义属性和属性值,部分片断如下:

4)基于有向图的关联网络.知识实例间的关 联关系最为多样化,一个对象可能关联多个同类 对象,也可能关联多个不同类的对象,即具有不同 的关联语义.用有向图的节点来表示知识实例,有 向边表示关联语义可以有效表达和组织这种关 系,形成多重有向的关联网络.有向图的优点在于 不仅可以表达对象间有无关联,还可以明确表达 存在什么样的关联,以产品的关联为例,有向边可 以定义为:产品和文件之间的一般关联、进一步细 化为概念设计阶段的产品与设计说明书之间的 文档关联、结构设计阶段与 CAD 图纸之间的文 档关联,以及工艺阶段与工艺文件之间的关联.产 品在其生命周期中还存在着和模型、项目、部门、 人员等多种对象之间的关联,OWL 中可通过 Object Property 和 Sub Property Of 自定义这些关 联关系.以下是采用OWL对多种文档的关联语义 描述片断:

<SpecDocof rdf: resource= " #AHN30433s3648-2005-12-05 "/>

< CADDocof rdf; resource = " #AHN30433-000A"/>

.....

产品生命周期实例知识内容虽然各不相同, 但都具有以上类似的语义结构,在实际应用中可 采用相似的逻辑结构和存储结构,以便于实现产 品知识的标准化和集成化管理.

3应用实例

基于以上研究,以汽车减振器为对象,建立了汽车减振器的产品知识集成模型,有效地增强了现有 PDM 系统中知识表达和集成的能力.如图 7 是型号为 AHN60153 的产品实例的本体知识模型的一部分,在该产品生命周期会有更丰富的信息和知识需要记录和使用,可采用类似的结构和语义表达方法.从图中可以看出,该产品不再是以往 PDM 系统所采用的孤立的特性描述和刚性的关系模式,而是以丰富的知识情境为背景,集成其

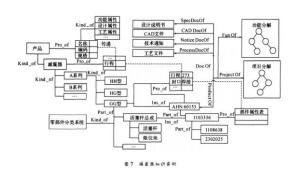


图 7 减震器知识实例

Fig.7 Knowledg einstance of shock absorber

类型层次、结构组成层次、属性传递关系、属性表和关联语义有向图等相关知识,这种结构化的知识本体和实例便于集成和重用于多种应用系统.基于此产品知识集成模型,在 PDM 系统中实现了对这些知识和实例的管理,采用.Net 开发工具提供的 Web 服务提供知识与信息的调用,可服务于面向局域网的 C/S 应用和面向 Internet 的 B/S 应用.图 8 是其中知识实例层相对应的部分应用界面.

结语将

本体引入信息集成系统,有效弥补了现有系统在知识表达与集成方面的不足.本体显式定义语义的能力,使语法级的信息集成进化为语义级的知识集成,延伸了产品生命周期信息的管理与应用范畴;OWL增量式和分布式的特点提供了面向Web的形式化表达能力.文中提出的中间层不仅提供了一种开放式的集成框架,也详细分析了知识本体表达的原语、本体集成的方法和知识实例的组织结构,这些方法具有一定的通用性,可满足产品全生命周期对产品知识的需求.本文侧重于信息管理中对知识集成的描述,其通用的知识表达方法和结构化的知识组织方法可进一步为知识重用提供服务,从而使有效的管理能够促进高效的开发

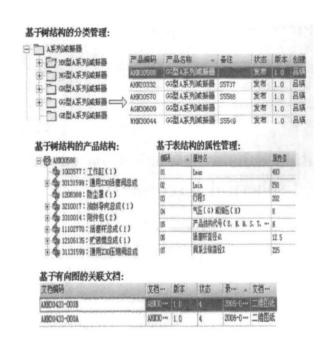


图 8 知识实例界面

Fig.8Interface of knowledgeinstance4

参考文献(References):

[1] 黄双喜,范玉顺.产品生命周期管理研究综述[J]. 计算机集成制造系统,2015,10(1):1-9.HUANG Shuang-xi, FAN Yu-shun.Overviewofproduct lifecycle management[J].Computer Integrated Manufacturing Systems, 2015, 10(1):1-9.

- [2] CHANE, YU KM.A concurrency control model for PDM systems [J]. Computers in Industry, 2007, 58(8-9):823-831.
- [3] DANESIF, GARDANN ,GARDANY, etal.P4LM: a methodology for product lifecycle management [J] .Computers in Industry, 2016 , 59(2-3):304-317 .
- [4] 林慧苹, 范玉顺, 黄琛.产品设计中基于元模型的知识管理 [J].计算机集成制造系统, 2015, 13(4):663-667. LIN Hui-ping, FAN Yu-shun, HUANG Chen .Metamodel-based knowledge management in product design[J].Computer Integrated Manufacturing Systems, 2015, 13(4):663-667.
- [5] 王生发,顾新建,郭剑锋,等.面向实例推理的产品设计 本体建模研究及应用[J].机械工程学报,2007,43(3): 112-117. WANG She ng-fa, GUXin-jian, GUO Jian-feng, etal. Research and a pplication of ontology modeling for product design based on case reasoning [J]. Chinese Journal of Mechanical Engineering, 2007, 43(3):112-117.
- [6] 顾巧祥, 祁国宁, 纪杨建, 等.基于元数据的产品数据本 体建模技术[J].浙江大学学报:工学版, 2007, 41(5): 736-741. GU Qiao-xiang, QI Guo-ning, JI Yang-jian, et al.Modeling techniques for product data ontology based on me ta-data [J].Journal of Zhejiang University:Engineering Science, 2007, 41(5):736-741.
- [7] 郭鸣, 李善平, 董金祥, 等.基于本体论及语义 Web 的产 品信息模型研究[J].浙江大学学报:工学版, 2004, 38 (1):22-28. GUO Ming, LI Shan-ping, DONG Jin-xiang, et al. Research on product information model based on ontology and semantic Web [J].Journal of Zhejiang University: Engineering Science, 2004, 38(1):22-28.