

一种扩展的 ORM2 数据库建模方法

刘 超, 刘琳岚, 王红奎, 付 峥

(南昌航空大学 计算机学院/软件学院, 江西 南昌 330063)

摘 要: 在分析、比对目前常用数据库建模方法的基础上, 提出一种扩展的面向角色建模方法, 按照分割规则将大型信息管理系统分割成规模可控的小型应用, 对每个小型应用采用概念化模式设计过程(CSDP)设计子模型, 再利用 NORMA 工具将这些子模型集成逻辑模型或 DDL 语句, 最终生成物理数据库。实例说明, 采用提出的方法建立数据库模型, 可以提高信息管理系统开发效率。

关键词: 数据库; 数据库建模; 面向角色建模; 概念化模式设计过程

中图分类号: TP31

文献标识码: A

文章编号: 1000-7180(2010)01-0042-04

An Extended ORM2 Database Modeling Method

LIU Chao, LIU Lin-lan, WANG Hong-kui, FU Zheng

(School of Computing, School of Software, Nanchang Hangkong University, Nanchang 330063, China)

Abstract: This paper proposes an extended object role modeling method after analyzing the current data based modeling method. It divides large scale information management system into small scale-controlled system according to partition policy. It employs Conceptual Schema Design Procedure (CSDP) to design sub models, which are transformed to logical model or DDL sentences, so as to create database. The practice shows that the database model built by this method can improve development efficiency of information management system.

Key words: database; database modeling; ORM; CSDP

1 引言

数据库建模是信息管理系统设计中的重要部分, 它可以帮助人们创建有效的形式模型, 为数据库的合理设计提供保障。常用数据库建模方法有实体联系(ER-Entity Relationship)建模, 面向对象建模(UML-Unified Modeling Language)^[1]和面向角色建模(ORM-Object Role Modeling)^[2]。

ER 建模最早由 Peter Chen 于 1976 年提出, 它是一种面向现实世界的概念建模技术^[3]。ER 模型由实体、联系、属性三个概念组成数据结构概念集合, 定义了实体之间的基本关系, 只可以表达业务领域中的大多数简单数据结构, 而且 ER 表达方式不统一, 难以标准化, 建模流程也不细致。

UML 建模是 OMG 组织于 1997 年发布的面向

对象软件过程整个生命周期的建模^[4], 用类图表示静态数据结构和操作。UML 克服了 ER 标准不统一的问题, 信息表达能力比 ER 强, 但 UML 类图将应用领域看作是对象及对象间的联系, 本质上与 ER 相同, 仍然存在对约束和业务规则的表达力比较差、建模流程粗糙的问题。

Terry Halpin 于 1998 年提出 ORM 建模, 于 2005 年提出 ORM 2 (the second generation of ORM)^[5]建模, 着重解决上述两种建模方法存在的问题, 提供了概念性的、易于理解的模型化数据的方法, 可以帮助建模人员更好把握业务规则。例如, 有以下业务规则: 员工属于部门, 员工领导部门, 领导某部门的员工(部门经理)必须属于该部门, 这是一个子集约束。用 ER 和 UML 无法表达这个子集约束, 而用 ORM2 表达这个子集约束如图 1 所示。

收稿日期: 2008-11-25; 修回日期: 2009-01-25

基金项目: 国家自然科学基金项目(60773055)

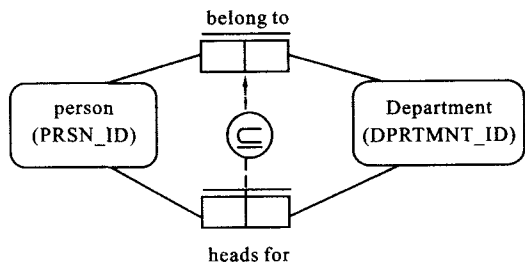


图 1 ORM 图表示的子集约束

文中分析 ORM2 相对于 ORM1 的改进之处,并基于 ORM2 提出一种扩展的面向角色建模(X-ORM)方法。

2 ORM2 与 ORM1 的比较

尽管 ORM 模型比 UML 和 ER 表达能力更强,但是 ORM1 占用空间太大、功能不够全面,ORM2 通过对图形符号的改进,使得功能也得到了一定的扩展,主要表现在以下几个方面:

(1) ORM1 用椭圆来表示对象类型,但是椭圆比矩形更占空间,尤其是对对象名比较长的时候;ORM2 默认用圆角矩形来表示对象类型,并提供矩形和较小的椭圆作为可选形状,使得所建模型更紧凑。

(2) ORM1 用带箭头的直线表示内部唯一性约束,而 ORM2 用简单的直线,更节省图形空间;ORM1 外部唯一性约束用圆内加 U 来表示,ORM2 用带圆圈的横线表示,如果是主要约束则是带圈的双横线,而 ORM1 用带圈的字母 P;ORM1 内部主要约束是唯一性约束基础上再加字母 P,而 ORM2 是在角色方框上加双横线。

(3) ORM1 用英文缩写来表示环约束,例如,ir (irreflexive)表示非自反环约束,非英语国家的人不能直观看出这些英文缩写代表什么约束,且用英文缩写占图形空间;ORM2 用比较直观的图标来表示不同类型的环约束。

(4) ORM2 的相等约束支持 N 元事实类型,而 ORM1 只支持二元事实类型。

(5) ORM1 限制了只有值列表相对稳定并且值的所选范围较少才可以添加值约束,对于无限范围的值约束只支持[a1,a2]这种包含 a1 和 a2 的类型,而 ORM2 的无限范围值约束支持各种类型。

(6) ORM1 只能对二元事实类型对象化,而 ORM2 可以对各种事实类型进行对象化。

3 X-ORM 数据库建模方法

文中对 ORM2 建模进行扩展,提出 X-ORM 数

据库建模方法,按照分割规则将大型信息管理系统分割成规模可控的小型应用,对每个小型应用采用概念化模式设计过程(Conceptual Schema Design Procedure, CSDP)设计子模型,以建立更清晰准确的数据库模型,再利用 NORMA 工具将这些子模型集成逻辑模型或 DDL 语句,最终生成物理数据库。

3.1 CSDP 建模过程

CSDP 是创建 ORM 图表的过程的名字。在 CSDP 过程中,以面向用户的概念开始,然后遵循以下步骤:

Step1:把熟悉的信息实例转变成基础事实,并进行品质检验;

Step2:绘制事实类,并用成员集检查;

Step3:检验应该合并的实体类型,对所有的算术推导加以注释;

Step4:添加唯一性约束,检查事实类型的元数;

Step5:添加强制角色约束,检查逻辑推导;

Step6:添加值、集合、比较和子类型约束;

Step7:添加其他约束,并执行最终检查。

该过程的前三步关系到确定的事实类,后四步添加事实类型的约束,整个过程通过检验来检测导出的事实。

3.2 扩展的建模过程

(1) 简化的事实编辑方法

国内关于 ORM 建模的研究,其模型图中的文本一般是中文^[2,6],因为 ORM1 一般是用作概念层上的推理模型。在实际应用中,如果用中文 ORM 模型生成物理数据库,所得的数据库字段名同样是中文的,给应用程序设计带来一定的不便。以往建模人员编辑一个事实类型,要先分开编辑事实中的对象类型,然后再用角色框和连接线将其连接起来,并编辑角色类型,这样的事实编辑方法比较繁琐。

2008 年,微软发布了最新的支持 ORM2 的 NORMA(Neumont ORM Architect for Visual Studio)建模工具,具有事实编辑器、样本集编辑器等工具窗口,利用这些工具窗口可以简化建模过程,建立更准确清晰的 ORM 模型。根据一定语法规则在事实编辑器中编辑英文事实,可以直接将事实类型转化为包括对象类型和角色类型的图形,这种事实编辑方法可以简化 CSDP 的 Step2,而且所建模型生成物理数据库后的数据库字段名是英文的,有利于应用程序设计。

(2) ORM 模型图的分割与集成

CSDP 只适合小型应用,而对于大型应用,如果

所有的事实都在一个 ORM 模型中表达,则图形占用空间太大,建模的准确度也会降低,所以必须将大型应用域分割成小型应用,为每一个小型应用设计子模型,然后将子模型集成或聚集形成一个覆盖整个应用的全局模型.这个自顶向下的设计过程可用伪代码描述如下:

```

if (largeUoD)
{ Divide(LargeUoD, SmallUoD); //把大型应用分割成小应用
while (! IsEmpty(UndesignedList))
{ CSDP(SmallUoD); //应用 CSDP 于小应用
Design(SmallUoD, SubModel); //为小应用设计子模型
}
Integrate(SubModel, GlobalModel); //将子模型集成全局模型
}

```

根据一定的分割规则,可以将一个大的 ORM 模型图分成若干个 ORM 子模型图.分割规则应该根据应用域的类型而定,对于大型信息管理系统,由于包含较多功能模块和报表,信息量比较大,文中提出按功能模块来分割,如果是比较复杂的模块还要按一定规则再分割,直到每个子模型都比较清晰为止.具体分割规则如下:

(1) 每个模块必须分割,命名格式是模块代码 + “-” + 可识别的名称.其中,将基础数据维护模块中的事实类型分布于其他各个模块中,不另做一个子模型;

(2) 如果某个复杂的模块需要分成比较小的部分,则在模块分割的基础上继续分割,命名格式为模块代码 + 二级模块代码 + “-” + 可识别的名称;

(3) 如果(2)中所分的某个二级模块还需要再分成更小的部分,则继续在此基础上分割.依次类推,直到每个 ORM 子模型都比较清晰.

对于最后的集成,可以利用 NORMA 的分页功能,将每个子模型设计成全局模型上的一个 ORM 页面,一个模型上的不同页面在生成逻辑模型或者 DDL(Data Definition Language)语句的时候是可以自动集成的.

4 X-ORM 建模实例

X-ORM 数据库建模方法适合大型信息管理系统的数据库设计,可以提高所建数据库模型的准确度,提高应用程序的开发效率.下面以江西某上市公司的人力资源管理系统为例,详细介绍用此方法设计数据库模型的过程.

首先根据分割规则将系统分割成机构管理、基本工资等 15 个功能模块如图 2 所示.

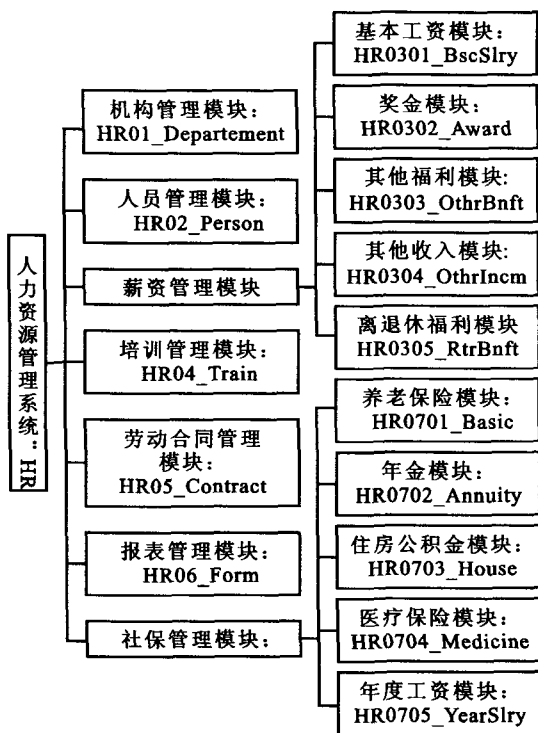


图 2 系统功能模块图

以人员管理模块为例,介绍采用 CSDP 设计人员管理子模型的过程.

(1) Step1: 把熟悉的信息实例转变成基础事实,并进行品质检验.

从客户调研得到员工花名册,从中可以抽取:“序号为 22 的员工工号为 0102”,“序号为 22 的员工姓名为周晓华”等 19 个基础事实.对这些基础事实进行检验,发现还缺少“员工有身份证号”、“员工有进入本公司时间”、“员工有工龄”、“员工领导部门”这些事实.

(2) Step2: 画出事实类,并用成员集检查

经分析,序号用人员 ID 这个实体类型表示,还有部门、职称、学历、学位为实体类型,其他用值类型表示.在 Fact Editor 中输入 Person(PRSN-ID) has DPRTMNT-ID() 等事实类型后初步的 ORM 图已经形成.

(3) Step3: 检验是否有实体类型需要合并,对所有的算术推导加以注释.

两个实体类型是否需要合并要看它们是否都含有相同的基于单位的引用模式.本例没有要合并的实体类型.本例的工龄是推导事实类型,工龄 = 当前日期 - 参加工作日期.

(4) Step4: 添加唯一性约束, 检查事实类型的元数.

每一个基本事实类型必须至少有一个唯一性约束.“员工有身份证号”、“员工有工号”和“员工领导部门”这三个事实类型为一对一的关系, 员工属于部门和员工有职务是多对多关系, 其余是多对一关系.

(5) Step5: 添加强制角色约束, 检查逻辑推导.

从样本集中可以分析得出, 一个员工有且仅有一个姓名、工号、身份证号, 一个员工至少属于一个部门, 至少有一个职务, 这些都是强制角色.

(6) Step6&Step7: 添加值、集合比较和子类型约束, 添加其他约束并最后检验.

性别这个对象类型的值可以限定为{男、女}. 员工属于部门和领导部门这两个角色的关系是领导某部门的员工(部门经理)必须属于该部门, 所以部门经理是该部门员工的子集, 添加这两个角色的子集约束. 本例没有其他约束.

人员管理模块的 ORM 模型如图 3 所示.

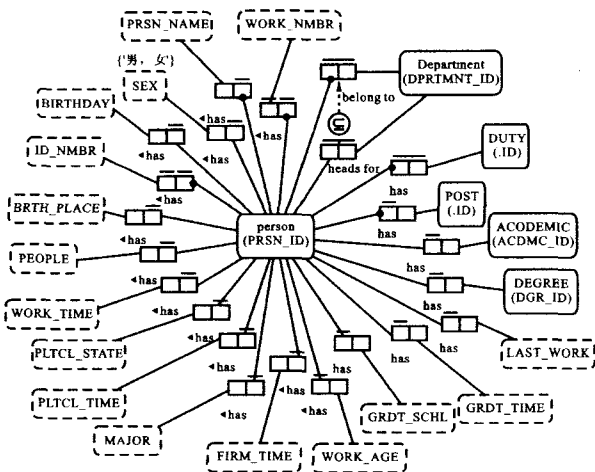


图 3 人员管理模块的 ORM 模型

该模型只是人力资源管理系统 ORM 模型的一个人员管理子模型, 其中职称、职务等对象类型属于基础数据维护模块, 根据分割规则, 需要在此模型上对部门、职位等对象类型进行扩展, 最终的人员管理子模型将包含职称、职务等基础数据信息. 其他功能模块按上述步骤生成相应模块的子模型, 最后所有

的子模型可以集成整个人力资源管理系统的全局模型, 进一步生成逻辑模型或 DDL 语句, 最终生成物理数据库.

5 结束语

从实例可以看出, 使用 X-ORM 进行数据库建模, 可以表达出 ER 和 UML 难以表达的业务规则, 例如子集约束, 并且为信息管理系统设计更准确、清晰的数据库模型, 也有利于生成物理数据库. 但是这种方法限制了图形中所显示的所有对象类型名和角色类型名都必须是英文的, 对于非英语国家的建模人员还有一定局限性, 这是需要改进的地方, 而且该方法对动态建模支持不够, 这将是下一步研究的重点.

参考文献:

- [1] 游晓明, 刘升, 陈传波. 基于 UML 的地理数据模型研究[J]. 微电子学与计算机, 2004, 21(5): 90-92.
- [2] 胡鸿明, 解帆. 一种全新的 ORM 数据建模方法[J]. 计算机工程, 2003, 29(2): 6-8.
- [3] 唐忠, 汪连栋, 刘东玉. 雷达有源干扰设备数据库 ER 概念模型建模[J]. 航天电子对抗, 2006, 22(6): 51-54.
- [4] 熊曾刚, 张学敏, 陈建新, 等. 基于 UML 语音发票查询系统的研究及应用[J]. 微电子学与计算机, 2004, 21(12): 42-45.
- [5] Terry Halpin. ORM 2. On the move to meaningful internet systems 2005 [C]//OTM 2005 Workshops. Berlin: Springer LNCS 3762, 2005: 676-87.
- [6] 王志英, 葛世伦. 面向角色的信息系统建模方法研究[J]. 计算机与数字工程, 2006, 34(4): 77-80.

作者简介:

刘超 女, (1984-), 硕士研究生. 研究方向为数据库建模技术.

刘琳岚 女, (1968-), 硕士生导师, 教授. 研究方向为分布式系统、软件工程.

王红奎 男, (1980-), 硕士, 讲师. 研究方向为数据库建模技术.

付 峥 女, (1973-), 硕士, 讲师. 研究方向为数据库建模技术.