

数据库逻辑设计中消除规范化处理问题^①

邢 泉

(浙江工业大学之江学院 经贸管理分院, 杭州 310024)

摘 要: 关系数据库设计要求规范化处理的过程以消除数据冗余和更新异常问题. 范式化理论的算法与数据依赖有着直接的联系, 数据依赖中重要的是函数依赖, 识别该依赖既耗时又复杂. 讨论如何在采用 ER 模型的数据库逻辑设计步骤中消除规范化处理过程, 在概念结构设计过程中遵循一组规则将产生满足 3NF 范式的初始关系模式集, 以节省数据库设计者在导出初始关系模式集后的规范化处理步骤.

关键词: ER 模型; 关系模式; 规范化; 函数依赖; 逻辑设计

Eliminating Process of Formalization in Logical Database Design

XING Quan

(Economic and Trade Management Branch, Zhijiang College, Zhejiang University of Technology, Hangzhou 3100024, China)

Abstract: the relational database standardization design requirements for handling process to eliminate data redundancy and update anomalies. Fan Shihua theory of algorithms and data dependence has a direct connection, relying on data dependence is the most important thing in function, and identify the dependence is both time-consuming and complex. This article discusses how to in the ER model database logic design step eliminate standardization process, in the process of conceptual structure design to follow a set of rules will lead to meet the initial relationship model set of 3 nf paradigm, to save the database designer after initial relational schema set derived the standardization of the processing steps.

Key words: ER model; relationship model; standardization; function dependence; logic design

关系数据库设计有二种方法, 一种是采用 ER 模型的概念模型设计和逻辑模型设计, 包括 ER 建模、导出初始关系模式集和规范化处理三个步骤. 一种是采用范式化理论设计数据库, 从最原始的、包含所有数据项的一个关系模式开始, 按照范式化理论逐步分解这个关系模式, 直到满足范式基本要求为止. 在实际数据库设计中第一种方法应用得更为广泛. 二种方法都需要规范化处理过程, 具体来说, 就是设计一个商业上可用的关系数据库, 它是消除了数据冗余和由数据冗余引起的更新异常问题的起码满足第三范式(3NF)的关系模式集. 一个关系需要识别出所有的函数依赖, 并且将不满足 3NF 的关系无损分解且保持依赖地分解成 3NF 的算法要求出函数依赖的最小依赖集, 这是一个复杂且耗时的工作, 因此有必要考虑能否将规范化处理步骤消除, 本文讨论如何在 ER 建模过程中, 提

高 ER 模型的质量, 使之能够转换为最终的关系模式集, 消除第三步的规范化处理步骤.

1 ER图转换成关系模式集的算法

ER 图中的主要成分是实体类型和联系类型, 转换算法就是如何把实体类型、联系类型转换成关系模式. 算法:

步骤 1(实体类型的转换): 将每个强实体类型转换为一个关系模式, 实体的简单属性即为关系模式的属性, 实体标识符即为关系模式的键. 将每个弱实体类型转换为一个关系模式, 弱实体的简单属性和它的强实体转换成的关系模式的键即为关系模式的属性, 弱实体的实体标识符和强实体的实体标识符即为关系模式的键.

步骤 2(联系类型的转换): 根据不同的情况做不同的处理.

^① 收稿时间:2012-11-19;收到修改稿时间:2012-12-28

2.1 若实体间的联系是 1:1, 可以在两个实体类型转换成的两个关系模式中任意一个关系模式的属性中加入另一个关系模式的键和联系类型的属性.

2.2 若实体间的联系是 1:N, 则在 N 端实体类型转换成的关系模式中加入 1 端实体类型的键和联系类型的属性.

2.3 若实体间的联系是 M:N, 则将联系类型也转换成关系模式, 其属性为两端实体类型的键加上联系类型的属性, 而键为两端实体键的组合.

2 违反范式的情况

在从 ER 模型转换为关系模式集时有违反各级范式的情况出现, 我们考虑每种情况下如何在概念建模阶段就消除产生问题的根源, 以便产生的关系模式集自动符合规范化的要求.

2.1 1NF 的违反情况

如果一个关系模式 R 的所有属性都是不可分的基本数据项, 则称该关系模式符合第一范式, 记作 $R \in 1NF$. 在任何一个关系数据库系统中, 第一范式是一个对关系模式最起码的要求, 如果一个实体的属性是多值属性, 则该实体转换成的关系模式中的一些分量包含多值数据项, 成为不满足第一范式的非规范化的关系模式.

ER 模型中含有复合属性时, 由于根据转换算法, 总是将复合属性转换成简单属性, 也就是转换成了具有原子值的属性, 因此不会出现具有组合数据项的关系模式.

考虑一个关系模式: $R(A_1, A_2, \dots, A_k, A_{k+1}, \dots, A_n)$

任一属性 A_k 是多值属性, 则改进后的 ER 模型可以转换成满足 1NF 的关系模式集的规则 1 如下:

将原来的多值属性用几个新的单值属性来表示, 或者将原来的多值属性用一个新的弱实体类型表示, 这个新实体类型和原来的实体类型之间是 1:N 联系. 下面举例说明.

考虑一个关系模式 R(零件编号, 零件名, 销售价格), 概念结构设计阶段的 ER 模型如图 1 所示.

根据规则 1, 可以改进该 ER 图的质量, 得到图 2 和图 3, 它可以根据算法转换成以下关系 R1(零件编号, 零件名, 经销价格, 代销价格, 零售价格)或者 R21(零件编号, 零件名)和 R22(零件编号, 销售性质, 售货价

格), 这些关系都满足了 1NF.

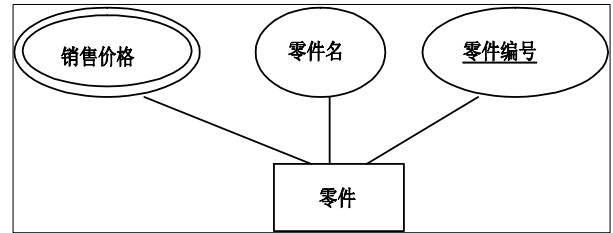


图 1 违反 1NF 的 ER 图

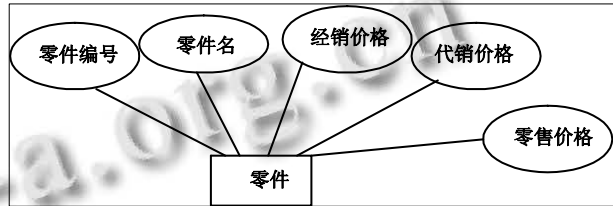


图 2 多值属性用多个单值属性表示

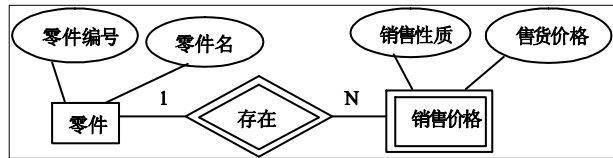


图 3 多值属性用弱实体表示

2.2 2NF 的违反情况

考虑一个关系模式: $R(A_1, A_2, \dots, A_k, A_{k+1}, \dots, A_n)$
它的主键为 $K=\{A_1, A_2, \dots, A_k\}$, 非主属性为 $NK=\{A_{k+1}, \dots, A_n\}$

对于 $FDK \rightarrow NK$, 如果 $X \subset K$ 和 $A \subset NK$, 存在 $X \rightarrow A$, 那么存在局部函数依赖, 因此 R 违反了第二范式. 如果说关系模式 R 是通过实体 E 转换而来, 该函数依赖并没有在 ER 模型中体现, 对应于该函数依赖在 ER 模型中应该有一个实体类型 E' . X 成为 E' 的实体标识符, 属性 X 和 A 属于实体类型 E' . 因为 X 是实体类型 E 标识符中的属性, 现在移到了 E' 中, 因此 E 不再是强实体, E 变成了 E' 的弱实体, E 的部分主键是 Y, 其中 $Y=K-X$. 改进过的 ER 模型可以转换成满足 2NF 的关系模式集的规则 2 如下:

对于函数依赖 $X \rightarrow A$, X 属于实体 E 的标识符中的属性, A 属于实体 E 的非标识符中的属性, 建立一个实体类型 E' , 它的属性由 X 和 A 组成, 它的实体标识符是 X; 实体类型 E 是 E' 的弱实体, 它的属性由除去 X 和 A 的实体类型 E 的属性组成, 新实体类型 E' 和原来实体变成的弱实体 E 之间是 1:N 联系. 下面举例说明.

考虑一个关系模式 $R(\text{学生编号}, \text{活动项目}, \text{活动成绩}, \text{费用})$, 概念结构设计阶段的 ER 模型如图 4 所示:

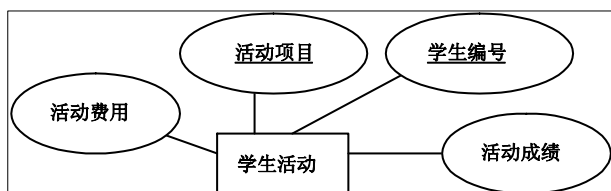


图 4 违反 2NF 的 ER 图

用户的语义如下, 一个学生可以参加许多活动项目, 一项活动只能收取一种费用. 在关系 R 中, 识别出二个函数依赖, 即 $FD_1: (\text{学生编号}, \text{活动项目}) \rightarrow \text{活动成绩}$, $FD_2: \text{活动项目} \rightarrow \text{费用}$, 由于 FD_2 的存在, 因此关系 R 不满足 2NF. 根据规则 2, 可以改进该 ER 图的质量, 得到图 5, 它可以根据算法转换成以下两个关系: $R_1(\text{学生编号}, \text{活动项目}, \text{活动成绩})$ 和 $R_2(\text{活动项目}, \text{费用})$, 这两个关系都满足了 2NF.

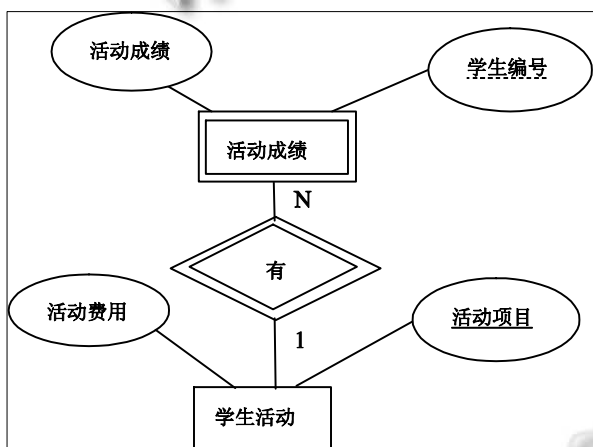


图 5 图 4 改进后的 ER 图

2.3 3NF 的违反情况

考虑一个关系模式: $R(A_1, A_2, \dots, A_k, A_{k+1}, \dots, A_n)$

它的主键为 $K=\{A_1, A_2, \dots, A_k\}$, 非主属性为

$NK=\{A_{k+1}, \dots, A_n\}$

对于 $FD K \rightarrow NK$, 如果至少存在一种情况, $X \subset NK$ 和 $A \subset NK$, 有 $X \rightarrow A$, 那么存在传递函数依赖, 因此 R 违反了第三范式. 如果说关系模式 R 是通过实体 E 转换而来, 该函数依赖并没有在 ER 模型中体现, 对应于该函数依赖在 ER 模型中应该有一个实体类型 E' . X 成为 E' 的实体标识符, 属性 X 和 A 属于实体类型 E' . 因为 X 不是实体类型 E 标识符中的属性, 因

此 E 还是强实体. 改进过的 ER 模型可以转换成满足 3NF 的关系模式集的规则 3 如下:

对于函数依赖 $X \rightarrow A$, X 和 A 属于实体 E 的非标识符中的属性, 建立一个实体类型 E' , 它的属性由 X 和 A 组成, 它的实体标识符是 X ; 在 E 和 E' 之间建立一个 $N:1$ 的联系类型; 实体类型 E 的属性由除去 X 和 A 的实体类型 E 的属性组成. 下面举例说明.

考虑一个关系模式 $R(\text{学生编号}, \text{学生姓名}, \text{宿舍}, \text{费用})$, 概念结构设计阶段的 ER 模型如图 6 所示:

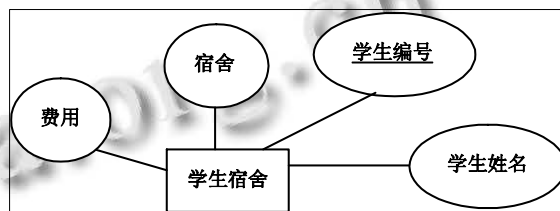


图 6 违反 3NF 的 ER 图

用户的语义如下, 一个学生只能住在一栋宿舍楼里, 一栋宿舍楼只能收取一种费用. 在关系 R 中, 识别出二个函数依赖, 即 $FD_1: \text{学生编号} \rightarrow (\text{学生姓名}, \text{宿舍}, \text{费用})$, $FD_2: \text{宿舍} \rightarrow \text{费用}$, 由于 FD_2 的存在, 因此关系 R 不满足 3NF. 根据规则 3, 可以改进该 ER 图的质量, 得到图 7, 它可以根据算法转换成以下两个关系: $R_1(\text{学生编号}, \text{学生姓名}, \text{宿舍})$ 和 $R_2(\text{宿舍}, \text{费用})$, 这两个关系都满足了 3NF.

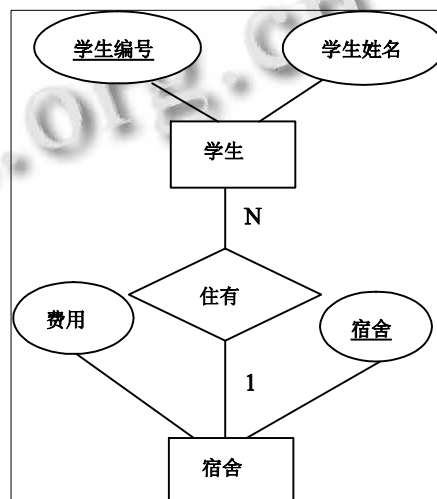


图 7 图 6 改进后的 ER 图

3 结论

通过以上分析可以看到, 在概念结构设计阶段的 ER 建模中, 由于忽视了引起 1NF, 2NF 和 3NF 的函数

(下转第 203 页)

果给出建议。

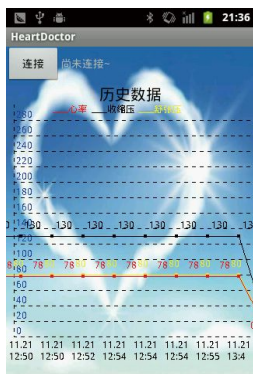


图 7 开始界面



图 8 收到测试数据



图 9 测量结果发送短信

5 结语

本文设计的基于 mega328p 和 android 的数字血压计不仅在硬件上采用合适滤波电路使得脉搏波形噪声小易于识别,在软件上还采用均值滤波、异常值过滤等处理,使得结果准确,工作稳定。

与传统的电子血压计相比,本系统能够通过蓝牙连接并发送测量结果到手机。当今的智能手机拥有强大的功能,通过手机这个平台,本系统可以将测量结果发送短信至其他手机。在下一步的研究工作中,可以通过手机将测量数据上传至“云”端,提供个人信息保存的功能或为相关单位或部门做数据挖掘提供数据来源。

参考文献

- 1 郭岳,熊合金.基于 AVR 单片机的数字血压计设计.计量与测试技术,2010,37(2):25-26.
- 2 康华光.电子技术基础:模拟部分.第 5 版.北京:高等教育出版社,2006.36-37,413-429.

(上接第 181 页)

依赖,因此逻辑设计根据转换算法把 ER 模型转换成关系模式集时需要进一步做规范化处理,使之成为商业上可用的关系数据库,由于这是一个耗时且费力的工作,因此有必要考虑节省这一步骤,在概念结构设计中即改进 ER 模型,改进后的 ER 模型可以根据本文得出的规则,通过增加弱实体类型或一般实体类型得到.数据依赖中除了函数依赖,还有多值依赖、连接依赖和包含依赖,如何将这类型的数据依赖体现在概念结构设计中,真正消除规范化处理步骤是作者今后要进一步研究的问题。

参考文献

- 1 施伯乐,丁宝康,汪卫.数据库系统教程.第 3 版.北京:高等教育出版社,2008.
- 2 潘文林.ER 模型转换为关系模式的实用规则.计算机工程与应用,2006,(1):169-173.
- 3 Hussain T, Shamail S, Awais MM. Eliminating Process of Normalization in Relational Database Design. Proceedings IEEE INMIC 2003.
- 4 Kroenke DM. Database Processing Fundamentals, Design & Implementation. Prentice Hall, 1999.