

文章编号: 1005-8451 (2003) 10-0009-03

E-R 模型特殊超类 / 子类结构转换方法研究

韩 驹, 李之堂

(华中科技大学 计算机科学与技术学院, 武汉 430074)

摘要: 分析了该系统的概念结构中出现的特殊超类 / 子类结构, 即: 超类实体对不同的子类实体具有不同的主码, 其主属性仅仅是所有子类实体共同的主属性, 各子类实体拥有各自特定的主属性。

关键词: E-R 模型; 超类; 子类; 关系模式

中图分类号: TP301

文献标识码: A

Research on translation method of Special Superclass/Subclass Structure in E-R Model

HAN Ju, LI Zhitang

(Computer Science & Technology College, HuaZhong University of Science and Technology, Wuhan 430074)

Abstract: It was analyzed the special superclass/subclass structure, i.e. the superclass entity has different primary code according to different subclass entity, the prime attributes of the superclass entity were only the common attributes of all subclasses, each subclass had its unique prime attributes.

Keyword: E-R Model; superclass; subclass; relation schema

1 引言

随着高速铁路建设的研究、城市轨道交通的建设和城际轨道交通的尝试, 铁路建设事业正跨入一个崭新的时代。开发铁路设计中负责前期工作的“经济运量”与“行车组织”2个专业的信息管理系统, 实现铁路勘测设计前期工作所涉及和必需的有关全国铁路网中信息的统一管理, 显得非常迫切和重要。

2 超类 / 子类结构到关系模式的转换方法

关于超类 / 子类结构的表示方法有很多。

在关系数据库的设计过程中, 从概念结构到逻辑结构的转换对超类 / 子类结构的处理, 目前国内外多种文献, 看法均不尽相同, 还未形成统一的意见, 现暂引用较著名的几种方法进行说明。国内一本较著名的关于数据库系统原理的文献中^[4]在假设实体型C是实体型E1、E2、……Em的超类, k、A1、A2、……、An是C的属性, k是C的主码的前提下将超类 / 子类联系的变换分为4种方法; 斯坦福大学1997年的教材一书中介绍了2种方法^[2], (1) 每个实体建立一个关系, 并且只将该实体的属性赋予这个关系, 为了识别与每个元组有关的实体集, 就需要包括每个实体共

同的键码属性; (2) 使用NULL值合并关系, 该方法将超类 / 子类结构仅用一个关系表示, 关系的属性为结构中所有实体的属性集合, 与某元组所表示的对象无关的属性的值设置为NULL; 而美国另外一本关于数据库的基础、设计与实现的文献中论述的方法与上述2种文献的方法大体上相同^[5], 其中还单独论述了当子类实体有自己的主码时的处理方法, 即将超类实体的主码作为属性置于子类实体的关系中。

综观这些方法, 均有一个前提, 就是要求在超类 / 子类结构中, 超类实体均有主码, 当子类实体没有主码时, 则继承超类实体的主码作为其相应关系的主码, 当子类实体有自己的主码时, 则其相应的关系的主码为自己的主码。

3 E-R 模型中特殊超类 / 子类结构的分析

在铁路设计项目信息管理系统的概念结构中有如下部分(非主属性略)。

这是有关铁路线路区段的分 E-R 模型, 是铁路项目信息管理系统 E-R 模型的一部分, 其中线路区段是超类实体, 既有线路区段与设计线路区段是子类实体, 与实体线路区段联系的还有2个实体能力及主要技术标准, 均为一对一的联系。实体能力和主要技术标准表示线路区段的能力和主要技术标准, 都是弱实体。

收稿日期: 2003-05-21

作者简介: 韩 驹, 在读硕士研究生; 李之堂, 教授。

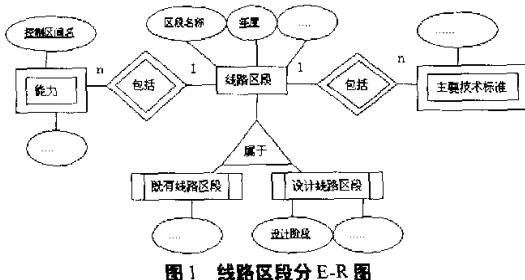


图1 线路区段分 E-R 图

图上各个实体除已经表示出的主属性外，超类实体线路区段还有属性区段类型，所属线路，区段起点，区段终点，区段长度，线路允许速度；子类实体既有线路区段还有属性列车换长；子类实体设计线路区段还有属性设计者，车站数目，最大站间距，最小站间距。

图中实体线路区段的属性区段类型表示既有区段和设计区段2种情况。表面上看起来，这个结构中超类实体线路区段有主码（区段名称、年度），子类实体设计线路区段有主码设计阶段，弱实体能力有主属性控制区间名。但是对于表示线路项目的信息，实际情况却并不完全如此。当线路区段为既有的线路区段时，其各种信息均可由区段名称和年度的值唯一标识，由于没有其它的候选码，所以超类实体线路区段和子类实体既有线路区段的主码均为区段名称和年度，并且子类实体可以继承超类实体的主码；设计阶段分规划、预可行性研究、可行性研究、初步设计、施工设计等几种，而且每个设计阶段均对应多个年度，分为设计初期、近期、远期及远景，然而既有线路区段仅表示线路的既有状况，不存在设计阶段和年度之分，所以子类实体设计线路区段的主码应该是由区段名称、年度、设计阶段共同构成，其中前2项继承超类实体的主属性，子类实体设计线路区段的主属性同时分布超类实体和子类实体中。由此看出，超类实体对不同的子类实体具有不同的主码，其主属性仅仅是所有子类实体共同的主属性而已。

4 超类 / 子类结构到关系模式转换方法局限性的分析

目前国内外各种文献中论述的将超类 / 子类机构向关系模型转换的方法，虽然其间也有一些差别，但大体上可以分为2种：(1) 将结构中的所有超类实体和子类实体建立一个关系；(2) 结构中每个超类实体

和子类实体各建立一个关系。不管采用哪一种方法将线路区段 E-R 图转换为逻辑模式，超类实体线路区段对应的关系的主码都有2种，即由主属性区段名称、年度构成和由主属性区段名称、年度、设计阶段构成，其中前者对应子类实体既有线路区段的主码，而后者则对应子类实体设计线路区段的主码。如果采用前者作为超类实体对应关系的主码，那么在表示设计线路区段的信息时，主码将不能唯一确定一个元组，因为相对于相同的区段名称和年度有多个设计阶段，对应的数据互不相同；如果采用后者作为其主码，则在表示既有线路区段的信息时，主属性设计阶段为空，违背了关系模型的实体完整性规则。并且不管采用前者还是后者，当2个弱实体能力和主要技术标准要将所联系的强实体线路区段的主码作为自己的关系的主码或主码的一部分时，将带来的问题是一样的，即要么主属性将为空而违背关系模型的实体完整性规则，要么主码不能够唯一标识一个元组从而与主码的定义相悖。

5 特殊值方法

从本文的实例来看，出现了构成子类实体主码的主属性相交但不完全相同的情况，这种情况超出了目前各种文献论述的方法的适用范围。这种情况的实质是各个子类实体相同的主属性由于是各个子类实体的共同的抽象而分布于超类实体中，而互不相同的主属性分布于对应的子类实体中，从而造成了整个实体集合对于不同的子类实体具有不同的主码并且互相不能代替的情况。

解决这个问题可以采用设置特殊值的办法。结合本文实例，由于子类实体设计线路区段具有自身特定的主属性“设计阶段”，它的主码将由从超类实体线路区段中继承的主属性区段名称、年度和自身主属性设计阶段共同构成，那么为整个超类 / 子类结构建立一个关系，将子类实体设计线路区段的主码区段名称、年度、设计阶段作为关系的主码，为主属性设计阶段设计一个特殊值“既有”作为当线路区段为既有线路区段时的值，所以关系中相应的元组也就表示子类实体既有线路区段的一条信息。关系如下：

线路区段 (区段名称, 年度, 设计阶段, 区段类型, 所属线路, 区段起点, 区段终点, 区段长度, 线路允许速度, 列车换长, 设计者, 车站数目, 最大站间距, 最小站间距)

这种方法就避免了当某元组表示既有线路区段的一个对象时，主属性设计阶段为空，违背实体完整性的情况，并且弱实体“能力”和“主要技术标准”完全可以将此关系的主码作为自己的关系的主码或主码的一部分。但是这种建立一个关系的方法具有一个明显的缺陷，就是由于非主属性数量多，关系中将出现大量的空值，造成存储空间的浪费。如：当某元组表示既有线路区段信息，则属性设计者、车站数目、最大站间距、最小站间距的值均为空值。

由于有了特殊值，还可以为每个实体建立一个关系，但是超类实体线路区段对应的关系的主码是由区段名称、年度和子类实体既有线路没有的主属性设计阶段共同构成。各个关系如下：

线路区段（区段名称，年度，设计阶段，区段类型，所属线路，区段起点，区段终点，区段长度，线路允许速度）。

既有线路区段（区段名称，年度，设计阶段，列车换长）。此关系中的主码为继承的超类实体线路区段的主码，并且关系中每个元组的主属性设计阶段的值都为“既有”。

设计线路区段（区段名称，年度，设计阶段，设计者，车站数目，最大站间距，最小站间距）。此关系的主码的前两个主属性为继承超类实体的主属性。

这种方法弥补了上一种方法的不足之处，但是它也有自身的缺陷，即主码的数据重复出现在各个关系中，造成数据冗余，浪费空间。

6 结束语

对超类 / 子类结构向关系模型转换的方法做了一些探索，提出了特殊值方法。铁路设计中负责前期工作的“经济运量”和“行车组织”两个专业的信息管理系统所涉及的数据及其间的关系十分复杂，采用特殊值方法可以完成该系统从 E-R 模型到关系模型的转换。所述的两种转换方法又各有优缺点，在设计中采用哪种方法较好，要根据实际情况而定。如果各个实体的非主属性较多而主属性相对很少，则可以采用方法二；反之则可以采用方法一比较合适。

[参考文献]

- [1] 萨师煊, 王珊. 数据库系统概论 (第三版)[M]. 北京: 高等教育出版社出版, 2000.
- [2] Abraham Silberschatz, Henry F. Korth, S. Sudarshan. Database System Concepts[M]. 北京: 机械工业出版社, 1999.
- [3] 李建中, 王珊. 数据库系统原理[M]. 北京: 电子工业出版社出版, 1998.
- [4] David M.Kroenke, 数据库处理—基础、设计与实现 (第七版)[M]. 施伯乐、顾宁、刘国华. 北京: 电子工业出版社, 2001.
- [5] Jeffrey D.Ullman, Jennifer Widom. 数据库系统基础教程[M]. 吴嘉权. 北京: 清华大学出版社, 1999.
- [6] 刘云生, 卢正鼎, 卢炎生. 数据库系统概论 (第二版)[M]. 武汉: 华中理工大学出版社, 1996.



▲ 2003年10月16日，由铁道科学研究院等单位主办的“第六届中国国际现代化铁路装备展览会”在京召开。

在开幕式上，铁道部工程师王麟书、铁道科学研究院院长陈国芳分别发表重要讲话和致辞。



▲ 展会上，铁道部总工程师王麟书、科技司司长耿志修等部领导在铁道科学研究院党委书记陈春阳、院长陈国芳、副院长董守清的陪同下来到铁道科学研究院展位，参观并询问了铁道科学研究院电子计算技术研究所研发的客票专用设备应用推广情况。

摄影 / 文 / 本刊记者 国敏