

5.9 小结

函数依赖是一个关系模式的属性间的一种数据依赖，反映了现实世界事物属性间的约束关系。关系模式属性间的部分函数依赖和传递函数依赖导致关系模式存在着数据冗余、更新异常和数据不一致问题。

数据依赖满足一定约束（性质）的关系模式的集合称为范式，根据约束的程度，范式有多个级别。在函数依赖的范畴内，主要有1NF、2NF、3NF、BCNF。范式从低到高，对关系模式的约束逐渐加强，范式之间是一种包含关系。

在函数依赖范畴内，从低范式到高范式，通过不断消除决定因素非码的非平凡函数依赖，也逐渐消除了不好的函数依赖给关系数据库带来的数据冗余、更新异常和数据不一致问题。使得实现了数据库中数据的逐渐分离。

一个关系模式R上的函数依赖集F所蕴含的函数依赖的集合称为F的闭包 F^+ 。Armstrong公理为计算 F^+ 提供了一个有效且完备的基础理论。属性集X关于函数依赖集F的闭包 是 F^+ 中所有函数依赖于属性集X的所有属性的集合，可利用属性集闭包来判断给定的函数依赖 $X \rightarrow Y$ 是否属于 F^+ 。

若两个函数依赖集的闭包相等，则这两个函数依赖集等价。每一个函数依赖集至少等价一个最小函数依赖集。最小函数依赖集必须具备三个条件，一是F中每一个函数依赖的右部为单一属性，二是F中不存在可被F中其他函数依赖所逻辑蕴含的函数依赖；三是函数依赖是左部不可约的。寻找与一个函数依赖集等价的最小函数依赖集的过程就是对函数依赖集不断进行等价变换，使其满足这三个条件的一个过程。

在最小函数依赖集的基础上，可确定候选键，判断范式级别，并利用分解算法来实现关系模式的规范化。模式分解应尽可能具备无损连接性和保持函数依赖，使分解前后的两个数据库模式等价，这种等价包括数据等价和语义等价两个方面。但为了保持数据等价，分解的结果有时不得不在3NF 和BCNF中做选择。

关系模式规范化理论为数据库设计提供了理论指南和工具，数据库设计人员应熟悉规范化理论。