

数据库系统中的软件集成技术

田沧海 林钧海

(南京航空学院计算机系)

摘要: 本文通过三级模式的一致性研究, 提出了数据库系统中的软件集成方法, 目的是使得数据库系统的开发更为简便, 应用面更为广泛。

1 引言

在数据库领域中, 软件集成占有相当重要的地位。这是因为随着数据库技术应用的拓广, 传统的DBMS已经不能满足许多特殊的需求, 如图形、图像、嵌套表、CAD/CAM/CIM等数据库系统。显然, 为每一个特殊的应用开发一个专门的DBMS是不合适的^[1]。这不仅浪费许多开发时间, 而且由于各应用之间数据不能共享, 冗余度很大。目前, 虽然有一些集成化的数据库系统, 如Lotus1-2-3、Informix等, 但这些系统只是功能的集成, 而在软件组织结构上, 集成化程度仍然很低。例如要实现功能 f , 其中 f 中包含若干个子功能: f_1, f_2, \dots, f_n , 一般的实现方法如图1所示。

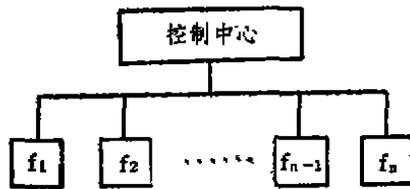


图1 功能集成方法

其中, 每一个子功能相应开发一个子模块, 通过控制中心分别调用, 由于许多相同的子程序不能共享, 因而系统占用内存量大。更重要的, 子程序冗余会产生潜在的运行结果不一致。目前, 软件组织结构的集成(图2)得到了越来越多的关注, 它使得系统在程序、模型和数据等方面能得到充分的共享。

从三级模式着手, 可以有效地实现软件组织结构集成, 使用户接口友好, 数据结构简单, 且数据

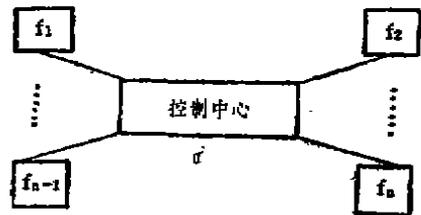


图2 软件组织结构集成方法

冗余度很小。

2 三级模式与软件集成

目前, 大部分数据库系统在软件集成方面, 忽略了三级模式的一致性研究。典型的例子是扩充外模式(图3), 虽然功能增加, 但概念模式不能很好地描述外模式, 使得系统结构不一致, 程序冗余。

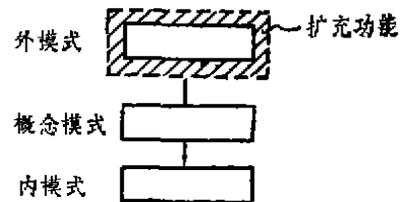


图3 扩充外模式

为了有效地支持多种功能的数据库系统, 就要认真研究三级模式的一致性, 也就是说, 概念模式

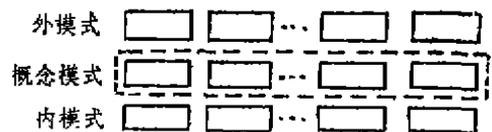


图4 三级模式的简单扩充

要能够完整地描述具有多种功能的外模式并处理多种数据的内模式。简单的功能集成(图4)会在内模式上产生许多问题。

(1) 概念模式: 各模式之间相互没有联系, 因而数据结构繁多, 系统组织结构庞大, 且由于程序模块的冗余会导致运行结果的不一致。

(2) 外模式: 用户界面不友好, 用户必须十分了解与系统相关的许多概念才能进行编程, 加重了用户的负担。

(3) 内模式: 数据大量冗余, 无形中削弱了系统的功能和可靠性。

因此, 为使软件集成系统具有生命力, 必须实现软件组织结构的集成, 概念模式是软件组织结构集成的核心(图5)。以一种数据模型为基础, 通过构造可以定义其它多种模型, 由于系统的功能都可以在概念模式上加以构造和描述, 并且数据存取一致, 因而能用最少的编程量最大限度地实现集成化的数据库系统。

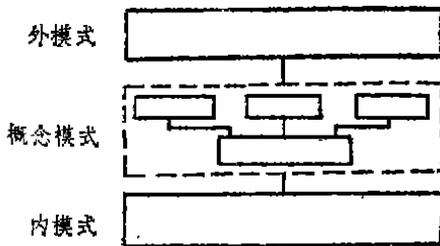


图5 概念模式的构造和集成

3 实例

根据上述方法, 我们用C语言在IBM PC机上实现了一个软件组织结构高度集成化的第四代语言系统NH₄GL⁽²⁾。目前, 该系统可以支撑关系表格、嵌套表格、图形生成及屏幕设计。用数据字典(DD)支持概念模式, 在关系模型的基础上进行构造, 使系统的三级模式结构保持一致(图6)。可以看到, 随着功能的扩充, 三级模式仍能维持一致性。

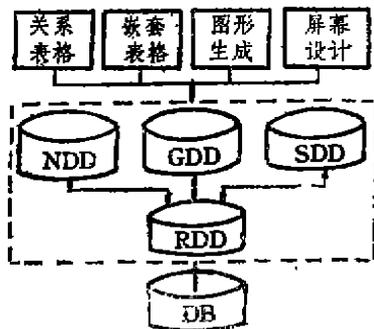


图6 构造方法

其中, RDD、NDD、GDD、SDD分别表示关系(Relation)、嵌套表(Nested table)、图形(Graph)和屏幕(Screen)的数据字典。关系数据字典(RDD)的描述结构包含⁽³⁾：

(1) 关系的名字、存放此关系中每个元组所占字节数以及每个元组含有的属性个数。

(2) 每个属性的名字、类型以及此属性存放位置所占字节数。

(3) 联系: 建立关系数据字典与其它数据字典之间的联系。

从图6可以看出, 所有的功能模块都建立在关系模型基础之上, 通过数据字典(如NDD、GDD、SDD)的构造, 可以定义其功能, 以嵌套表的数据字典(NDD)描述结构nt-dict为例,

```
struct nt-dict
{
    char nt-name[20];
    int nt-delete;
    int nt-number;
    char nt-type;
    char nt-relationship[20];
}
```

其中:

nt-name[20]记录嵌套表及其子表的名字。

nt-delete为删除标志, 当该嵌套表或其子表被删除时, 标志位置1。

nt-number指出嵌套表或其子表的成员数。

nt-type为嵌套表所允许有的数据类型, 包括基本数据类型(如整型I、实型F、字符型C、日期型D等)和复杂数据类型(如关系R、属性A、嵌套表N等)。

nt-relationship[20]记录嵌套表或其子表与关系数据字典的联系。

如果用户在关系 $T_1(A_1, A_2, A_3)$ 和 $T_2(A_4, A_5)$ 上定义嵌套表 $NT(A_1, T_2(A_4, A_5))$, 其中 A_1 来源于 T_1 , 数据类型为I, 从而在嵌套表的数据字典(NDD)中, 只需完整地存储信息 $NT(T_1,$

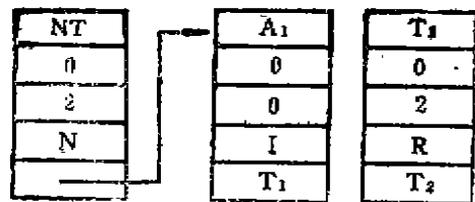


图7 信息 NT (T₁, A₁, T₂) 的存储方式
(下转第28页)

tive formation of orthogonalizing filters and associative memory in recurrent memory in recurrent networks of neuron-like elements, *Biol. Cybernetics* 21, 85—95, 1976.

- [3] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. of the National Academy of Science USA* 79, 2254—2258, 1982.
- [4] J. J. Hopfield, Neurons with graded response have collective computational properties like those of two-state neurons, *Proc. of the National Academy of Science USA* 81, 3088—3092, 1984.
- [5] J. J. Hopfield, D. W. Tank, Neural computation of decisions in optimization problems, *Biol. Cybernetics* 62, 141—162, 1985.
- [6] G. E. Hinton, T. J. Sejnowski,

ki, Optimal perceptual inference, *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 448—453, 1983.

- [7] D. H. Ackley, G. E. Hinton et al, A learning algorithm for Boltzmann Machine, *Cognitive Sci.* 9, 147—169, 1985.
- [8] N. Metropolis, et al, Equation of state calculations for fast computing machines, *Journal of Chemical Physics* 6 1953.
- [9] S. Kirkpatrick, et al, Optimization by simulated annealing, *Science* 220, 671—680, 1983.
- [10] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images, *IEEE Trans. PAMI-6*, 721—741, 1984.
- [11] H. Szn, R. Hartley, Fast simulated annealing, *Physics Letters A*, Vol. 112, 157—162, 1987.

(上接第58页)

A_1, T_2), 如图7所示。

显然, 当用户需要时, 根据RDD可以自动地形成嵌套表的结构形式, 生成嵌套表NT (图8)。

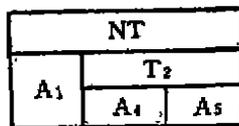


图8 嵌套表NT

嵌套表NT中的数据可以从关系 T_1 和 T_2 中得到。

用类似的方法, 可构造图形和屏幕的数据字典, 以实现相应的功能。

系统功能的扩充也是相当方便的, 如果要支撑CAD功能, 只需先以关系的形式存放点、线、面, 然后在其基础之上用相应的数据字典构造CAD的功能。

4. 结束语

从三级模式的一致性出发, 研制数据库系统的

软件集成, 不仅使得系统占内存少、功能强、用户界面友好, 数据冗余度小, 而且对于研究复杂对象的构造以及多介质数据库都是大有裨益的。

参考文献

- [1] D. S. Batory, T.Y. Leung and T.E. Wise, "Implementation Concepts for an Extensible Data Model and Data Language", *ACM Trans. on Database System*, Vol.13, No.3, 1988.
- [2] 田沧海、林钧海, "Fourth Generation Language NH₄GL—Towards an integration for software organization", *ICCC Symposium '89 on Computer Communication*, 1989.
- [3] 田沧海、郭红英、林钧海, "基于数据字典的用户友好接口", *小型微型计算机系统*, Vol.9, No.9, 1988.