维普资讯 http://www.cqvip.com

广西大厂矿田地质矿产数据库的研究与系统开发

毛先成，戴塔根，邹艳红，吴湘滨
（中南大学地学与环境工程学院，长沙 410083）

[摘 要]大厂矿田地质矿产数据库及其管理系统实现了大厂矿田历年来积累的大量地质矿产资料的数字化存储和管理，是矿山企业信息化和专题研究的数据基础平台。通过资料现状与数据需求分析，将地质矿产资料划分为表格、文档、图件和目录簿四类。论文提出资料表（数据表、文档表、图件表）和分类树型目录结构属于两种不同的数据模型，可分别对原始资料（表格、文档、图件）和目录簿进行描述和存储；资料表可直接用关系数据库的二维关系表来表示。但目录簿具有与关系表的线型结构完全不同的非线性树型结构，对此作者建立了孩子关系表示法，从而实现了关系表对目录簿的树型结构的间接表示。通过数据需求分析、概念模型定义、逻辑结构设计、数据库的存储与运行维护设计，建立了大厂矿田地质矿产中央资源数据库GEODBS，开发了相应的数据库管理系统DCDBMAN。

[关键词]大厂矿田 矿产资源 数据库
[中图分类号]P628 [文献标识码]A [文章编号]0495-5331(2003)05-0072-05

大厂矿田是全国重要的锡多金属产地，已探明锡储量的90%赋存在大厂矿区。大厂矿区以长坡—铜坑、巴力、龙头山3个矿体为主体，在中深部存在数量众多的锡石—硫化物型矿体。

大厂矿田自20世纪50年代开采以来形成了海量的地质资料，数据种类繁多，包括各类地质矿产数据和统计数据，如钻孔工程数据、物化探数据等。所有的资料数据按大厂矿田数据和区外资料进行分类，而大厂矿田数据又可分为东带资料、西带资料、中带资料、钻孔资料数据、物化探专项数据、水文资料数据等十三类，其中，东带资料、西带资料、中带资料又按各自的地名再进行分类，资料室则存放各类数据的目录簿，查找资料时，由目录簿统一进行索引。由此可知，资料的存在形式表现为两种：(1)原始的图件、报告、表格资料数据；(2)分类目录簿。

大厂矿田地质矿产数据库及其管理系统的建立可有效地解决大厂矿田地质矿产资料的利用率低、数据共享程度差、数据深层次的挖掘尚未展开等问题，是矿山企业信息化和专题研究（如隐伏矿体定量预测[1]）的数据基础平台。系统设计充分结合了现有资料分类检索习惯和流通管理模式，考虑了数据库系统的存储与处理要求、安全性与完整性要求和数据可扩展性要求。

1 地质矿产资料的数据模型

1.1 数据模型与地质矿产资料存储

数据模型是现实世界数据特征的抽象。数据模型根据其应用目的不同分为两类：其一，概念模型，是按用户的观点来对数据和信息建模，用于数据库设计；其二，数据模型，是按计算机系统的观点对数据建模，用于数据库管理系统（DBMS）的实现，主要包括网状模型、层次模型、关系模型和面向对象模型[2]。目前，最为流行的数据库管理系统Oracle、SQL Server、Sybase、DB2等都是基于关系模型实现的，所以，地学中的各种数据存储与管理系统主要是基于关系型数据库系统的[3]。我们选择关系型数据库系统SQL Server 2000[4]作为大厂矿田地质矿产资料管理的数据库平台。

关系型数据库是将数据按表结构形式进行组织的，通常被用于处理线性结构的数据。但对地质矿产资料进行存储管理时，不仅要对表格数据等大量线性结构数据进行处理，还要对目录簿等非线性结构数据进行处理。若要用关系模型处理非线性结构（树型结构、网状结构）的数据，就不仅需要对数据本身进行存储，还需要利用合适的方法对数据间的关系进行存储。因此，为利用数据库实现地质矿产

[收稿日期]2003-06-26；[修订日期]2003-07-10；[责任编辑]曲丽莉。
[第一作者简介]毛先成(1963年-)，男，1986年毕业于中南工业大学，获硕士学位，在读博士生，副教授，现主要从事地学信息系统的研究工作。

维普资讯 http://www.cqvip.com

资料的存储与管理,我们必须首先研究地质矿产资料的数据模型表达方法。

地质矿产资料的存储与管理在数据上有4种表现形式:表格、报告、图件、目录簿。所以,对地质矿产资料数据的关系模型表达,实际上是对表格、报告、图件、目录簿4种形式数据的关系模型的表达。

1.2 地质矿产资料表格

地质矿产资料主要是以数据表格形式存在的,有采样试料总帐、样品化验分析单、钻孔弯曲测斜、储量计算表格、地质编录数据、测量数据、原生晕分析数据、磁测数据等。

这些表格型数据的数据结构是线性的,它们可直接用关系模型的逻辑结构即二维表来表示,称为数据表。

1.3 地质矿产资料报告与图件

地质矿产资料报告与图件主要包括各种普查报告、勘探报告、储量总结报告、设计报告、科研报告、中段地质图、勘探线剖面图、综合地质图、地形图、物化探成果图、设计图等。报告的电子文档一般以Microsoft Word、WPS、Adobe PDF等格式文件保存,图件的电子文档一般以TIFF、MapInfo、AutoCAD等光栅和矢量格式文件保存。

这些图件和报告的电子文档的处理由专门的应用软件(Word、MapInfo、Photoshop、AutoCAD、ArcInfo)完成,本系统主要应完成在地质矿产资源数据库中的存储和管理。但这些电子文档的存储量都较大,不能用关系模型中普通数据类型来表达。但SQL Server、Oracle等大型关系数据库系统都提供了二进制大对象(BLOB)类型数据的操作与处理支持,所以我们可以将报告与图件的文档文件、栅格图像文件、矢量图形文件的以文件的形式上载进数据库,保存在相关二维表格的BLOB类型的字段中。对于某些报告或图件,若以多个文件存储,我们可以将这些文件作为多个二进制大对象保存在数据库表格的多条记录的BLOB字段中,或用压缩工具(如WinZip、WinRar等)打包为单个文件进行上载存储。当需要使用存储在数据库中的图件和报告时,我们可以从数据库下载到本地保存为相应的文件,然后使用本地的专用软件进行处理。在系统实现上,对于含有二进制数据的文档或图件关系表,界面中设有图形上载、显示和下载的命令按钮,利用ADO字段对象提供的GetChunk和AppendChunk方法[4]编程实现这些功能。

多个同类型的报告或图件保存在一个二维关系表格中,称为报告表或图件表。报告表、图件表、数据表统称为资料表。

1.4 地质矿产资料分类与目录簿

地质矿产资料目录簿保存了各种地质矿产资料的分类目录与索引信息,这些信息的数据结构是非线性的树型结构(又称层次结构)(图1)。

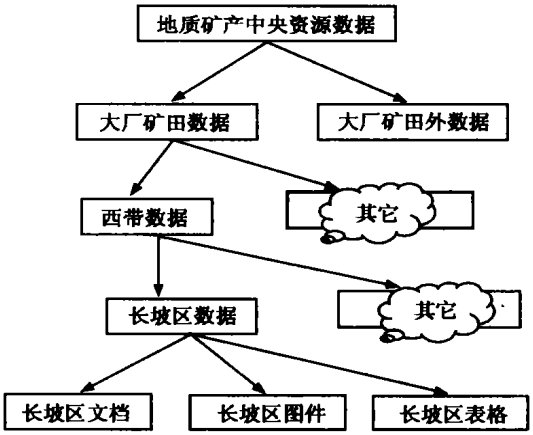

图1 大厂矿田地质矿产资料分类树型目录结构

非线性的树型结构与层次模型的树型数据结构相一致,但与关系模型的关系型数据结构不一致,因而基于关系模型的关系型数据库不直接支持树型结构数据的存储和访问。

树型结构数据的数据库表达,在当前技术条件下,有两种合适的途径:(1)采用层次数据库系统的层次数据模型直接表达;(2)采用关系数据系统的关系数据模型间接表达。但为了各种结构类型数据存储的统一,采用关系数据模型间接地描述树型结构数据则是最佳选择。树型结构数据表达的主要方法有:双亲表示法、孩子表示法、孩子兄弟表示法[5]。经过反复试验,我们构建了一种与孩子表示法相似的方法,可用关系数据库提供的二维关系表来表示树型结构数据,我们称之为孩子关系表示法。

在该方法中,所有二维关系表都采用统一的关系模式(Relation Schema)[2],一个关系表表示树型结构中的一个结点。表示树型结构结点的关系表的关系模式定义为:

目录表(编号,名称,表名,备注)

式中的编号、名称、备注均是树型目录结构的本级目录结点的属性描述;表名代表描述某个下级子目录的关系表的表名,显然,它是一个指向下级子目录结点的指针。描述一个目录结点的关系表(即目录表)可以包含多条记录,一条记录对应于一个指向孩子结点的指针域,所以,一个目录表可以描述指

维普资讯 http://www.cqvip.com

向多个子目录结点的多重指针，类似于孩子表示法中的多重链表。

以下给出对应于图 1 所示树型目录结构的孩子关系表示（图 2）。在该树型结构中，分支结点[5]表示目录表，其指针域指向下级目录表或资料表（数据表、报告表、图件表）；叶子结点表示资料表。须注意的是，在该表示方法中，分支结点的关系模式是相同的，即分支结点是同构[2]的，但叶子结点与叶子结点和分支结点的关系模式有可能不同，因而不是同构的。

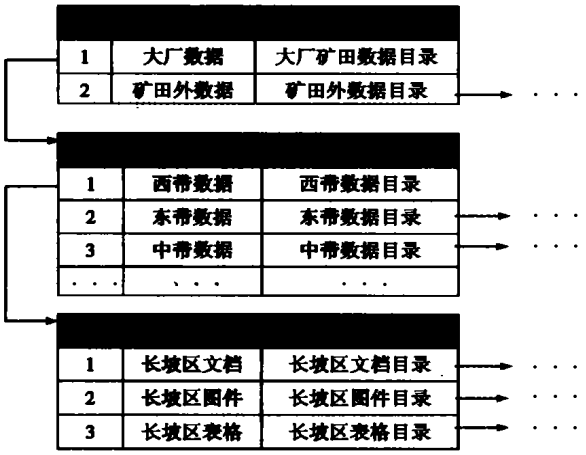

图 2 地质矿产资料树型目录结构的孩子关系表示

2 地质矿产资源数据库设计

地质矿产资源数据库是系统功能实现的基础，是整个系统的核心，数据的有效组织是系统实现的关键。建立大厂矿田地质矿产中央资源数据库之前，必须对已有的原始资料进行分类整理，根据用户的需求和后期科研应用数据需求进行数据库的设计，定义严谨的数据库结构，在计算机中有效地组织数据。在设计时结合已有的资料分类检索习惯和其流通管理的模式，考虑了用户原有的数据分类、数据库系统的存储与处理、安全性与完整性要求以及数据的可扩展性要求，进行了用户数据需求分析、概念模型定义、逻辑结构设计、数据库的存储与运行维护设计。

2.1 数据流图

根据资料管理和利用的业务流程，地质矿产资料管理数据流可从资料入库、资料登记查询、资料出库等主要加工环节来描述。地质矿产资料管理数据流图见图 3。

图 3 所示数据流图中包括了：(1) 地质矿产资料管理的 5 个数据加工环节（核查成果、审批、登记查询资料、入库、出库）；(2) 8 个数据流（成果表、有效成果表、入库资料、库存资料、出库资料、有效出库资料、借阅单、有效借阅单）。本研究以上述资料、加工环节、数据流为研究对象和研究基础。我们对上述全部资料进行了分类整理，建立与资料室目录簿有较好对应关系的树状分类体系，建立了全部资料的数据字典和数据库设计，建立了和资料室资料仓库对应的中央资源数据库，并进一步导出了隐伏矿体预测专题数据库、建立矿床数学模型和进行隐伏矿体立体定量预测。

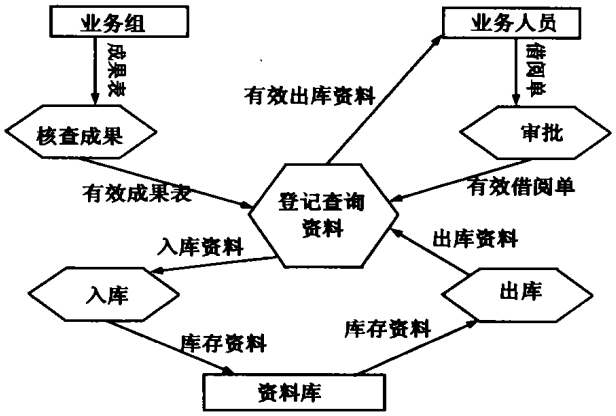

图 3 地质矿产资料管理数据流图

采用与原始纸介质资料分类相一致的分类体系结构，在计算机上实现该分类体系结构，参照规范、统一规格地设计大厂地质矿产数据字典，是成功地建立地质矿中央资源数据库的关键。

2.2 数据字典设计

数据字典作为收集、整理数据的分析归档方式[6]，是系统中各类数据描述的集合。

基于上述地质矿产资料数据模型，按用户的资料分类及数据需求，考虑地质矿产数据在数据库中建表的实现方式，将数据字典定义为三类：目录表、资料表、转换表。

2.2.1 目录表的数据字典设计

在尽量保持与原始人工分类方法相似的前提下，我们对资料进行分级归类，形成资料的规范化的树型结构分类组织，称为数据的树型分类目录组织方法。依据该方法，对大厂矿田地质矿产资料设计五级目录表和一级基本表（资料表），目前共建立的目录表有 134 个，分类树型目录结构见图 1，其孩子关系表示法实现见图 2。

树型分类目录组织方法中的目录结构可任意扩充，支持数据库管理员在以后数据输入过程中进行

维普资讯 http://www.cqvip.com

必要扩充。

表名:目录表的名称

描述:本目录表在树型目录结构中的级别以及用途等信息

定义:目录表 = 下级名称(下级目录表或数据表称呼) + 下级表名 + 备注

2.2.2 资料表的数据字典设计

根据数据的分类整理与建库思路,进行数据归整,对已有资料数据进行综合预处理,按原有资料的分类方式,最后将每一类资料按(地名)类别分别整理成表格、文档和图件三种类型。对于每一类资料按(地名)类别的资料表格,建立了相应的数据表结构。

表名:资料表的名称

描述:存储的数据所在的资料名称和类别等信息

表结构:资料表 = 属性1 + 属性2 + …… + 属性n

字段定义:属性字段的编号、名称、字段类型、长度以及有关的描述信息。

2.2.3 转换表的数据字典设计

大厂矿田已少部分原始资料录入计算机中,主要以.dbf文件和.txt文件格式存储。

为利用按这些格式存储的原始数据,我们采用以转换表为中间表的方法,将原始数据导入到数据库中。

为完成数据转换,设计了转换表的数据字典。转换表的数据字典与上述资料数据表数据字典的设计方法类似,但还附有原数据的存储描述与有关说明。

对于原有的以*.dbf文件格式保存在计算机中的钻孔、探槽、地表原生晕与次生晕分析成果,按文件名与表名对应的关系转换为转换表;对于磁测中精测剖面.txt与测线.txt文件,按文件名与表名对应的关系转换为转换表。

3 数据库管理系统实现

地质矿产原始数据库及其管理系统的建立,是地质找矿工作向计算机化、定量化、标准化、高效益迈进的关键和基础。大厂地质矿产数据库是开展各项专题研究的基础,为各项研究、找矿、生产和信息化工作的开展提供一个统一的、标准化的数据资源平台,可大幅度地加快各项工作的进程和效益,各项研究和生产工作可根据其特点目标,通过大厂地质矿产数据库管理系统选取满足其要求的数据视图,从地质矿产原始数据库析取所需的数据资源。

系统开发的总目标是提供一个以地质矿产原始数据库为核心,按树形目录的数据组织方式,具有各类数据的输入、分类查询与显示、数据修改与维护、报表输出以及用户权限管理等功能的数据库管理系统。同时,大厂地质矿产数据库是开展各项专题研究的基础,为各项研究、找矿、生产和信息化工作的开展提供一个统一的、标准化的数据资源平台。

"大厂矿田地质矿产中央资源数据库管理信息系统DCDBMAN"是在对大厂矿田地质矿产原始资料进行分类整理,按规范建立地质矿产原始数据库的基础上而开发的相应的管理信息系统。

通过用户调查、系统需求分析、概要设计、详细设计、编码和测试,实现了大厂矿田地质矿产数据库管理系统提供对地质、物化探、重砂、同位素、勘探工程等各种原始数据提供组织、录入、存储、检索、报表、维护等管理功能。

系统基于Microsoft SQL Server 2000为数据库系统平台,用Rational Rose、ErWin、Visio、Project 2000、Access 2000等工具进行系统设计、项目管理和数据录入,用Visual Studio 6等进行代码编写和软件调试。

系统以现行资料的查询管理机制为蓝本。按目前资料的管理形式可将系统功能分为两大部分:资料数据的管理与资料目录簿的查询管理。资料数据的管理分为数据输入、存储、检索、报表输出、维护等管理功能;资料目录簿的查询管理主要是数据树型目录的组织与显示、数据类的扩充与修改等功能。另外,根据管理需要,系统还需提供权限管理、用户跟踪管理等实用功能。

系统设计如下功能模块:

1)数据类的管理:数据类的保存、备份、还原与修改;

2)数据输入:单记录的输入与整表输入;

3)查询统计:单记录查询、整表查询、模糊查询与SQL查询;

4)数据维护:数据类的维护与数据表结构的维护;

5)图档载入:图件、文档的上载与下载;

6)实用工具:报表输出、用户跟踪与有关的计算工具;

7)权限管理:权限设置管理与用户跟踪管理。

系统模块结构图如图4所示。

维普资讯 http://www.cqvip.com

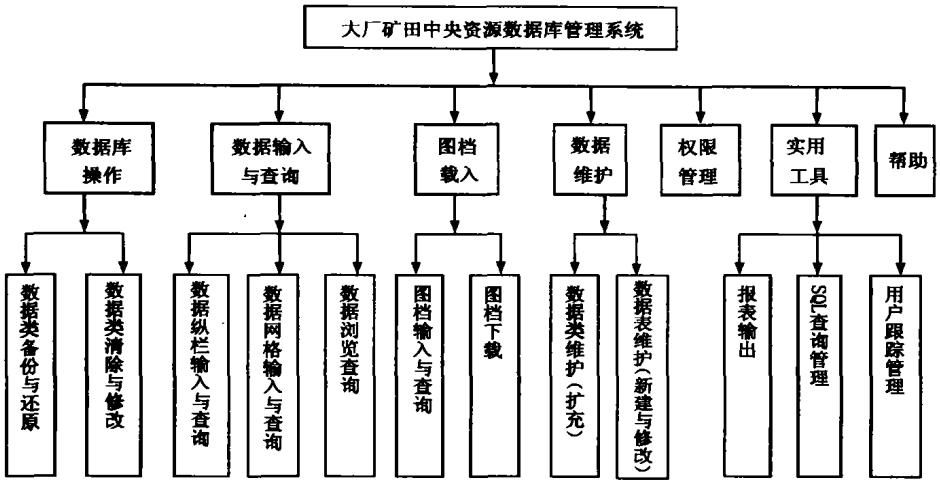

图 4 大厂矿田中央资源数据库管理系统模块结构图

4 结 论

1)地质矿产数据库和管理系统是专题研究、找矿勘探、生产和信息化工作的统一数据资源信息平台,是隐伏矿体立体定量预测研究、企业信息化等研究和工作的基础。

2)提出用关系数据库描述地质矿产资料的数据模型:数据表、文档表、图件表和树型目录结构;数据表和资料表直接地用二维关系表实现,树型目录结构按本文首次提出的孩子关系表示法间接地用二维关系表实现。

3)进行地质矿产资源数据库和管理系统的设计与开发,建立了大厂矿田地质矿产中央资源数据库 GEODBS,实现了地质矿产数据库管理系统 DCDBMAN。

[参考文献]

[1] 毛先成,陈国光. 香花岭锡矿床数学模型及立体定量预测初探[J]. 地质与勘探,1988,24(10):25~31.

[2] 萨师煊,王珊编著. 数据库系统概论,第 3 版[M]. 北京:高等教育出版社,2000,21~33,392~408.

[3] D S Schmatz, M Engi, J E Lieberman. ParaDIS: a relational database for the consistent documentation and analysis of metamorphic mineral assemblages[J]. Computer & Geoscience. 1995, 21, (9):1031~1041.

[4] 李真文. SQL Server 2000 开发人员指南[M]. 北京:北京希望电子出版社,2001,206~210.

[5] 严蔚敏,吴伟民. 数据结构(C 语言版)[M]. 北京:清华大学出版社,1997,135~137.

[6] J R A Giles, D J Lowe, K A Bain. Geological dictionaries——critical elements of every geological database[M]. Computer & Geoscience. 1997,23(6),621~626.

RESEARCH AND SYSTEM DEVELOPMENT OF GEOLOGICAL MINERAL DATABASE OF THE DACHANG OREFIELD

MAO Xian-cheng, DAI Ta-gen, ZOU Yan-hong, WU Xiang-bing

(*Central South University, Changsha* 410083)

Abstract: The geological mineral database of the Dachang orefield and its management system have been brought to success the digitized storage and management of the enormous geological and mineral data that have been accumulated for many years. The database and system are the data infrastructure for mine enterprise information system and subject research. Based on the analysis of the data status and the data requirements, the geological and mineral information has been classified into four types, table, document, drawing and catalog directory. This paper puts forward that the information tables (data table, document table and drawing table) and the classification tree-type directory structure have different data models which can be respectively used to describe and store raw data (tables, documents and drawings) and catalog directories. An information table can be represented by a 2 dimensional relation table of the relational database, but a catalog directory owns a non-linear tree structure wholly different from the linear structure of a relation table. For the non-linear structure, a method called Child Relation Representation Method to represent indirectly the tree structure of the catalog directory has been invented. Through the analysis of data requirement, the definition of concept model, the design of logical structure, and data storage and maintenance, the Dachang geological mineral database GEODBS is established and the database management system DCDBMAN is developed.

Key words: Dachang orefield, mineral resource, database