

地质矿产点源数据库系统的模型库及其分类体系^{*}

汪新庆 刘 刚 韩志军 袁艳斌

(中国地质大学地矿信息系统研究所, 武汉 430074)

摘 要 机助软件工程模型库及其分类体系在地质矿产点源数据库系统(GMPDBS)的设计、开发和运行中起着重要的作用。模型库是各个软件设计阶段的程序集成器和信息传递转换器,模型库分类体系的建立使全部模型库中错综复杂的属性、关系和结构信息条理化、程序化,使整个系统脱离混乱状态而变得有序,实现了开发—运行—维护一体化。机助软件工程模型库及其分类体系的应用,是系统结构、系统功能、模型结构 and 应用模型等实现参数化的方法基础,也是建立地矿点源数据库管理系统机助软件工程(GDCASE)的技术基础。论述了机助软件工程模型库及其分类体系的内涵、功能、结构组成、设计思路与方法,以及它在地矿点源数据库系统(GMPDBS)中的地位。这一技术成果在其他MIS的开发中,同样具有广阔的应用前景。

关键词 数据库,机助软件工程模型库,模型库分类体系,地质矿产点源信息系统,地质矿产点源数据库系统。

中图法分类号 P628.4

第一作者简介 汪新庆,男,副教授,1964年生,1985年毕业于武汉地质学院地球物理勘探系,现从事地矿信息系统工程方面的研究和教学工作。

地质矿产点源数据库系统(geology and mineral point - source database system,简称 GMPDBS)是地质矿产点源信息系统的核心^[1,2]。每一个 GMPDBS 通常都有数百个数据子模式,开发 GMPDBS 管理系统的关键,是解决数据库管理系统与数据库相对独立的问题,将功能模块与数据分离,并且建立功能模块与数据库之间的有机联系,让任一功能模块均可操作任一数据库。这样做既可大大减少 GMPDBS 中的程序冗余,又能使系统适应环境的变化,实现本身不断修改和完善,还能让用户方便地使用和维护系统。为此,必须建立机助软件工程模型库及其分类体系,并且开发出功能强劲的数据库机助软件工程(GDCASE)技术。

1 机助软件工程模型库及其分类体系

1.1 机助软件工程模型库在 GDCASE 系统中的地位

GDCASE 系统可以分成 3 个基本部分:前端、模型库和后端^[3]。前端部分对应软件生存期的前端

或前期阶段,即分析和设计阶段,还对应了 GDCASE 软硬件平台的相应部分。前端的 GDCASE 工具提供了支持分析和设计工作的功能,例如建立原型和检查规格说明。后端部分对应软件生存期的后端或后期阶段,即程序的实现阶段和维护阶段。后端也对应 GDCASE 软硬件平台的相应部分。后端 GDCASE 工具可用于自动实现编码、测试、数据库生成、数据规范化和现存系统的效果分析等任务。

机助软件工程模型库(GDCASE model base,以下简称模型库)是把 GDCASE 系统的前端和后端部分连接起来的一个通讯机构,其基本功能是实现:(1) GDCASE 工具的集成;(2) 系统规格说明的一致性和完整性控制;(3) 文档标准化;(4) 系统文档的生成;(5) 软件的可复用性;(6) 项目的管理与控制;(7) 系统信息的共享。概言之,模型库既是 GDCASE 系统的核心,又是各个软件设计阶段的程序集成器和信息传递转换器。

1.2 模型库的窗口

GDCASE 模型库能自动地跟踪并维护所有存贮在库中的信息,并通过一组标准的报告,让开发人员了解有关情况。例如,每个数据元在什么地方被用到?是谁存贮了这些信息?什么时候修改了它们的

1997 年 11 月 29 日收稿。

* 地质矿产部九五重点攻关项目(No. 9506002)及地质矿产部矿产资源定量预测与勘查评价开放研究实验室基金联合资助。

定义?一共修改了多少次?等等。模型库报告的基本类型有:(1)内容/列表报告;(2)交叉引用/使用地点报告;(3)分析报告;(4)扩展报告;(5)错误报告。

1.3 模型库的信息管理功能

模型库存贮信息的主要目的,是为了对全部系统信息实现自动化管理。在许多人员协同进行软件开发时,使用非模式化的手工方法来维护系统信息的完整性是不可能的。为了保证开发人员能安全、有序地共享系统的信息,需要有一种模式化和自动化的信息管理设施。

GDCASE 模型库的信息管理工具为开发人员提供了以下功能。(1)安全性控制。用密码口令来控制对模型库的访问,并用权限来约束访问内容的范围和深度。(2)访问优先级控制。赋予不同开发人员只读或全面更新的权利。(3)型式控制。跟踪某个系统展开的全过程,并允许多个当前的型式存在,如生产、测试,以及开发人员的个人工作集。(4)多用户控制。允许多个开发人员同时访问模型库中的同一个信息,但不能同时更新一个信息。(5)变更控制。自动跟踪模型库中信息修改情况,并能自动地综合多次的变更。

如前所述,GDCASE 模型库并非简单的数据词典,因为需要存贮许多类型的信息目标及目标之间的关系,并且需要强有力的信息管理设施。辅助设计字典实际上也是一种有 GMPDBS 支持的关系数据库,能够由外部的词典和数据库自动地生成。模型库与外部词典和数据库之间的信息流是双向的,GDCASE 系统必须为进出模型库的系统信息提供恰当的途径。

1.5 模型库分类体系的作用

为了实现向设计者和用户提供规格说明信息的存贮、访问、更新、分析和报告功能,模型库必须保存所有用于建立、修改、展开和管理整个软件系统的全部信息。这些信息的范围涉及如下方面:(1)所要解决的问题;(2)问题的定义域;(3)形成的解系统;(4)所采用的软件过程;(5)项目的资源和历史;(6)文件组织的环境。换言之,结构化图形、屏幕及菜单的定义、报告格式、记录说明、处理逻辑、数据模型、处理模型、源代码、事务规则、项目管理形式、数据元素,以及系统信息模块之间的关系等,都是可以存入模型库中的信息目标。

模型库中的每个信息目标,都是以其属性来描述的,即:(1)标识符(ID);(2)代名(号);(3)类型;

(4)文字描述;(5)功能模块;(6)目标所在的文件;(7)范围值;(8)编辑与推导规则。此外,还有每个目标同其他目标的关系及审计信息,例如,包含某个目标所有扩展层的关系信息,所有使用了该目标的目标和交叉引用信息等。审计信息及目标所有权认定与变更控制规则有关。模型库中的内容会随时间推移而不断增加,但对同一组织或同一应用领域来说,许多项目管理信息、组织模式、数据模型、数据实体,甚至某些程序的处理模型都可能相同;因此,随着更多的信息被存贮,建立一个新的软件系统所需的新信息会越来越少。

GDCASE 有多个模型库,而各模型库中的属性、目标、关系和模式如此错综复杂,需要用一个完整的体系来加以总体描述,这就是 GDCASE 模型库的分类体系。模型库分类体系是系统的框架,它包含 GMPDBS 的全部数据结构、系统结构、功能结构 and 应用结构的描述,使整个系统脱离混乱状态而变得有序。凭借该分类体系,设计者和用户能够科学、合理、灵活、快速、方便地使用 GMPDBS 中的全部数据库。这一技术方法也适合于其他 MIS 系统。

2 模型库分类体系的组成结构

模型库分类体系本身也可以转化成为一个内容丰富、功能强劲的软件系统。为便于开发与应用,在模型库设计时,对其结构和功能进行了合理分割,设置了 5 个模型库:系统功能模型库、数据库项目模型库、系统环境模型库、系统维护模型库、辅助处理模型库。模型库中有多个模型子库,子库中包含多个模型集。各模型库的结构如图 1~5 所示。

3 主要模型库的设计方法

模型库的建立,使整个系统的分析、设计和实现 3 阶段工作既相对独立又紧密连接,真正实现了结构化程序设计的方法和思想。它不仅保存了整个系统的功能菜单和用户界面信息,而且把系统实现阶段的模块层次划分、层次间关系处理和接口参数设置,从系统设计中分离出来,使型式的划分、实施和升级变得非常容易。

3.1 数据结构代码模型库

数据结构代码模型库也称为属性字典。

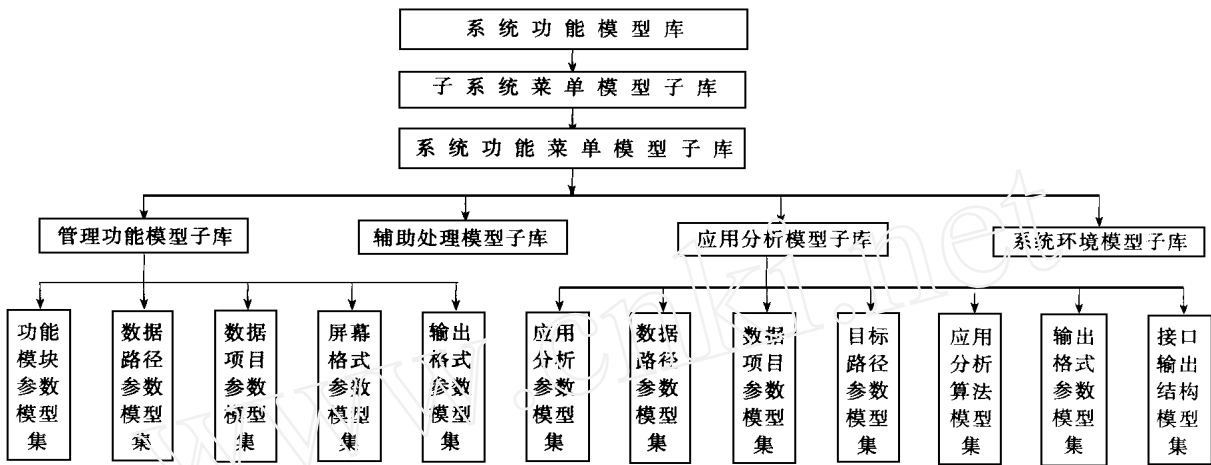


图 1 系统功能模型库

Fig. 1 The function model base of system



图 2 数据库项目模型库

Fig. 2 The item model base of database

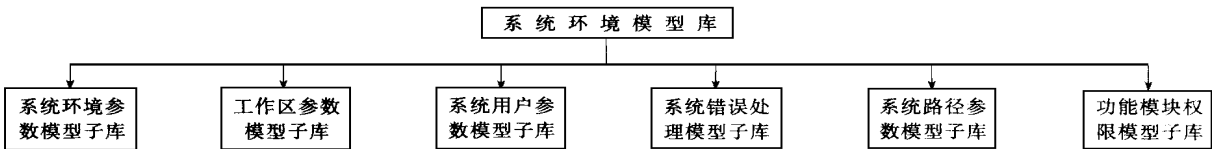


图 3 系统环境模型库

Fig. 3 The environmental model base of system

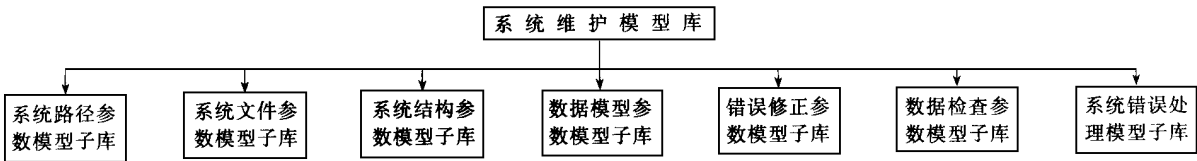


图 4 系统维护模型库

Fig. 4 The safeguard model base of system

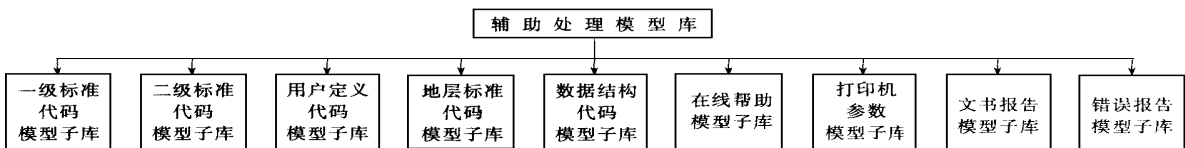


图 5 辅助处理模型库

Fig. 5 The assist process model base of system

GMPDBS 全部子模式所含的属性通常有数千项。属性名、属性说明、属性类型、单位,以及有关属性域值完整性字典存放着这些属性的全部信息,其中包括:属性完整性、一致性、安全性的约束条件信息。对于定性属

性采用标准化描述方式,并存放其关键字索引约束条件和专业字典约束条件.对于定量属性,通常以属性域的范围,作为其约束条件,并存放在属性字典中.在关系子模式维护、修改、扩充、重组过程中,属性字典起着恢复和传输各关系子模式属性参数的作用.

在属性字典中,不同类型的数据采用不同的存贮方式.描述型数据可用多数据或单数据的代码和术语存贮,数字型数据中的准确数据可直接按数值方式存贮.数字型数据中的一些不确定值或只给出值域的数据,可借用描述型数据方式,将它们当成字符来存贮.在 GMPDBS 中,规定了不同类型数据的处理方式:“ ”代表描述型、单项数据代码存贮;“,”代表描述型、多项离散数据代码存贮;“~”代表描述型、多项连续数据代码存贮;“*”代表描述型、单项数据术语存贮;“#”代表描述型、多项离散数据术语存贮;“=”代表描述型、多项连续数据术语存贮;“%”代表数字型、不确定值或值域数据当作字符存贮.

3.2 数据库项目菜单模型库

地质矿产点源数据库的全部关系子模式可分为十大类(可称为数据库文件集或数据子库):(1)勘查区概况;(2)勘查开发史;(3)地理条件;(4)水文条件;(5)岩心编录;(6)采样化验;(7)测井分析;(8)生产管理;(9)综合数据;(10)备用数据^[3].每一大类又分出若干个数据子模式(或称为数据库文件).数据库项目菜单是 GMPDBS 与数据子模式连接的桥梁,为了完整地表达本数据库系统数据模式的逻辑结构,特将项目菜单分为两级:一级模式菜单——主模式菜单,可引导用户进入上述十大类勘查区数据;二级模式菜单——子模式菜单,可在一级模式菜单引导下调用各数据库文件集的全部数据文件.

设计项目菜单需考虑关系子模式的可修改性、可扩充性和可移植性,并且实现模式菜单随着关系子模式及其参数改变而自动修改.其实现方法同样是建立概念模型字典,将关系子模式的有关参数传递到用户视图及应用专题中,且提供系统维护服务.在模式菜单中传输的关系子模式参数是:项目号、名称、关系名、关键字、关键词等.

3.3 数据库文件结构参数模型库

数据库子模式设计的首要任务,是把这些经过系统设计和系统分析建立起来的数据模型,通过系统提供的操作模式,定义为本系统的数据关系模式,然后利用数据库基础管理系统的原有功能,转换成物理模式.一般地说,每一个被定义的关系模式,包

括关系名(数据库文件名)、属性名(字段名)、属性域的类型和长度,以及数据关系模式的主要关键字.在 GMPDBS 中,我们把根据这些数据模型所建立的关系模式称为数据库文件,而把其他关系模式称为模型库.数据库文件的结构要素是:属性名(字段名);属性类型(字段类型);属性长度(字段长度);属性小数位(字段小数位).

在建立全部数据库文件的同时,系统将自动建立其相应的数据结构参数模型库(也称为数据模式结构描述字典).数据结构参数模型库对数据关系子模式进行了详细描述,其内容包括:属性名(字段名)、属性说明(字段说明)、属性类型(字段类型)、字段长度、小数位、单位、数据项约束、属性域、计算方法,等等.属性域参数由系统自动从数据结构代码模型库中传输过来,其他有关参数则在进行数据库文件结构描述时输入.

3.4 数据存取路径算法模型库

数据路径算法模型库中储存的主要参数有:工作区号、项目名称、关键字、关键字汉字名、项目文件名、索引文件名、数据结构参数模型库文件名、索引控制、关联关系、别名、关键字个数、处理约束字段栏号、约束字段名和控制处理方式.数据文件之间关联关系的属性,必须是进入当前工作区之前的数据模式中所有的.关键字的个数可由系统自动识别.由于数据的查询是多文件、多关系、多途径的操作,所以,算法模型不但在存和取的实现上有差别,而且在不同的功能模块中的表现也有差别.

数据库文件之间大多数用关键字关联,少数用一般数据项(字段)关联.关键字可以是单字段的,也可以是多字段的.在涉及多个数据库文件时,我们将其中一个含有主要查询内容的文件称为主数据文件,而将另一个或几个含有次要查询内容的文件称为副数据文件,并且按其关联方式,把数据库文件的连接关系分为星型、串型、混合型和隐蔽型 4 种.“星型”连接指所有副数据文件都与主数据文件存在关联关系,副数据文件之间没有前后顺序,可以随意排列.“串型”连接指副数据文件与主数据文件、副数据文件与副数据文件之间,均存在传递关联关系,必须严格按照关联顺序排列.只有当所有关联关键字都相同时,“串型”连接才能与“星型”连接等同.“混合型”连接指数据文件之间,同时存在“星型”和“串型”连接关系,副数据文件必须根据关键字的前后顺序排列.在实际应用中只要了解数据文件与前一数据

文件的连接关系,而不必知道前一数据文件是哪个。具有“星型”、“串型”和“混合型”连接关系的数据文件之间,都存在着对应的记录关系,只要设计好连接模型库,系统就可以根据存取算法模型库自动建立连接关系,并把多个数据库文件连接起来。“星型”和“串型”连接,是“混合型”的两种特例,因此,在 GMPDBS 的设计中,只需考虑“混合型”连接关系,便可具备通用性。“隐蔽型”连接指两个数据文件之间的关联关系是隐蔽的,要通过特定的计算处理,才能连接起来;因此,也称为“处理型”连接关系。这种连接关系需要编写特定的连接处理程序,有关的数据文件顺序需与处理程序的顺序一致。在“隐蔽型”关系中,连接关系的处理方式分为名义-有序型和间隔-比例型两种类型;与此相应,其连接处理方法也分为精确连接处理和模糊连接处理两种类型。

4 结语

地矿点源数据库的机助软件工程模型库及其分类体系,是将 GMPDBS 或 MIS 中除数据以外的各种信息,按照其面向对象和数据的功能进行归类,并使其有序化的技术系统。其中,模型库是 GDCASE 技术的核心,它用于存放开发人员设计 GMPDBS 必备的各类信息,以及系统运行过程中产生的各类信息,实现了系统功能与数据分离,并且将开发人员和用户距离拉近。它的存在也是 GMPDBS 对各种环境

有自适应能力的重要原因。采用模型库及其分类体系来开发 GMPDBS,不仅可以缩短开发周期,还改变了开发方式,使得 GMPDBS 的开发不再采用“生命周期法”和“快速原形法”^[4~6],而采用“实验运行法”^[3],实现了系统开发与系统运行、系统维护同步,使用户能够参与开发、修改和扩充,大大地增强了 GMPDBS 的生命力。

显然,模型库是 GMPDBS 开发与运行的生命线。鉴于数据库管理系统开发的相似性,本模型库及其分类体系在其他 MIS 中,同样具有广阔的推广应用前景。

参 考 文 献

- 1 吴冲龙,金友渔,王仁铎等.聚煤盆地地质计算机处理的途径与方法.北京:地质出版社,1992. 5~10
- 2 吴冲龙,汪新庆,刘刚等.论地质矿产点源信息系统.见:中国地质学会数学地质专业委员会主编.中国数学地质(6).北京:地质出版社,1995. 120~126
- 3 吴冲龙,汪新庆,刘刚等.地质矿产点源信息系统设计原理及应用.武汉:中国地质大学出版社,1996. 52, 59~63
- 4 郭凤鸣.地质管理信息系统.北京:地质出版社,1992. 50, 58
- 5 刘云生,卢正鼎,卢炎生.数据库系统概论.武汉:华中理工大学出版社,1992. 62~64
- 6 冯玉才.数据库系统基础.第二版.武汉:华中理工大学出版社,1993. 162

THE CASE MODEL BASE AND ITS CLASSIFICATION SYSTEM OF GEOLOGICAL AND MINERAL POINT - SOURCE DATABASE SYSTEM

Wang Xinqing Liu Gang Han Zhijun Yuan Yanbin

(Institute of Geology and Mineral Resources Information System,
China University of Geosciences, Wuhan 430074)

Abstract Computer aided software engineering model base and its classification system are of great importance in the design, development and application of geological and mineral point - source database system (GMPDBS). This model base is the program integrater and information transformer at every stage of software design. The foundation of the model base classification system enables the whole properties, relations and structure of the model base to become formulary and programmable, and the whole program system to be free from the disordered state and become manageable, thus fulfilling the integration of software development, running and maintenance. The application of this model base and its classification system are the base of para-

metric method of the design of system structure , system function , model structure and application model , and also the technique base of the foundation of geological and mineral point - source database management system computer aided software engineering (GDCASE) . In this paper the connotation , function , structure composing , mind and method of design , and the position in geological and mineral point - source database system (GMPDBS) of computer aided software engineering model base and its classification system are discussed. Also this technology achievement has a wide application value in other MIS development.

Key words database , computer aided software engineering model base , model base classification system , geological and mineral point - source information system , geological and mineral point - source database system.

* * * * *

简 讯

“973”国家重大基础研究项目研讨会在我校召开

1998 年 2 月 22 ~ 23 日在我校召开了由国家教委委托的国家重大基础研究项目研讨会. 会议由校长殷鸿福院士主持, 张锦高书记致了开幕词. 赵鹏大院士和殷鸿福院士分别就所建议的项目的主题、目标、立项意义及要求作了主题发言, 莫宣学副校长就国家科委重大基础研究项目的立项精神及申请事宜进行了通报.

来自西北大学、四川联合大学、北京科技大学、南京大学、北京大学、上海交通大学、中国科技大学、中南工业大学、广州海洋地质调查局、冶金部中南冶勘局、桂林矿产地质研究所、地矿部经济研究院、地矿部勘查技术院、中国地质大学(北京、武汉)的 50 位专家学者, 就如下两个主题进行了分组讨论: (1)

赵鹏大院士提出的“非传统矿产资源的发现、开发基础研究”; (2) 殷鸿福院士提出的“中国中央造山带组成、结构、演化、资源环境与可持续发展”.

通过两天的研讨, 大家一致认为这两个项目均面向 21 世纪, 注重多学科的联合与交叉, 为保证下世纪我国不可再生资源的可持续供给, 提供新型矿产资源支撑, 以及为发展新的矿产资源具有理论意义、经济意义及社会意义. 大家表示将通力合作, 互相支持, 争取项目申报成功.

科研处 张俐

1998 年 2 月 22 日