

对城市活断层探测项目中地震地质数据建模的探讨

葛伟鹏^{1,2}, 袁道阳^{1,2}, 郭华^{1,2,3}

(1. 中国地震局地震预测研究所兰州科技创新基地, 甘肃兰州 730000; 2. 中国地震局兰州地震研究所, 甘肃兰州 730000; 3. 中国科学院南京地理与湖泊研究所, 江苏南京 210008)

摘要:在城市活断层探测项目中, 基于地震地质和活动构造研究的数据模型研究是相关应用领域中的重要问题。本文依据“中国地震活动断层探测技术系统”规程应用统一建模语言讨论和建立了地震地质数据模型, 着重论述了地震地质数据模型与活动断层探测相关数据的组织方法, 给出了地震地质基于 ArcGIS 的 UML 模型的空间和属性类结构图、类关系图表。在此基础上分析了地震地质 UML 模型的特点。

关键词:地震地质; 数据模型; 统一建模语言; 计算机辅助软件工程

中图分类号: TP392; P315.2 文献标识码: A 文章编号: 1000-0844(2006)02-0134-06

Discussion on Seismogeologic Data Modeling for Urban Active Fault Surveying Project

GE Wei-peng^{1,2}, YUAN Dao-yang^{1,2}, GUO Hua¹

(1. Lanzhou Base of Institute of Earthquake Prediction, CEA, Lanzhou 730000, China;

2. Lanzhou Institute of Seismology, CEA, Lanzhou 730000, China;

3. Nanjing Institute of Geography and Limnology, CAS, Nanjing 210008, China)

Abstract: In urban active fault surveying project, research on the spatial data modeling based on Seismogeology and active tectonics is very important for interrelated application fields. According to the stipulation of Technical System for China Earthquake Active Fault Surveying, the Seismogeologic Data Model is studied and created using Unified Modeling Language (UML) in this paper. It emphasizes the method of data integration about active fault surveying and the structure of Seismogeologic Data Model, and it also presents the class structure graphs and the relations chart of classes based on ArcGIS. Meanwhile the attributes of seismogeologic UML data model are analyzed.

Key words: Seismogeology; Data Model; UML (Unified Modeling Language); Computer-Aided software/System engineering tool

0 引言

目前, 国际上的地质学家和地球物理学家已着手对地质图和地球物理探测等数据进行数据建模和编码实现。其中有针对地球科学逻辑和语法的抽象层面的项目, 如地学标记语言 (GML)、北美地质图数据模型、数字地球科学空间模型、探测和采矿标记语言、澳大利亚政府地学信息政策顾问委员会地质

数据模型等, 采用了基于 XML Schema 和 UML 的建模工具^[1]; 而针对概念和物理实现层面的项目有维多利亚地学数据、西澳大利亚地学数据、新南威尔士地质数据模型、美国地质调查局地质图数据库建设——北美地质图数据模型 (C1)^[2] 等, 这些项目主要是基于 UML 建模工具和基于 ArcGIS 的空间数据

收稿日期: 2006-04-05

基金项目: 中国地震局“十五”国家重大建设项目 (1-4-28-1); 中国地震局兰州地震研究所青年地震科学基金 (Qj0603); 中国地震局兰州地震研究所论文著号: LC20060020

作者简介: 葛伟鹏 (1981—), 男 (汉族), 山西稷山人, 硕士生, 主要研究方向为地理信息系统在活动构造研究中的应用。

库。

“中国地震活动断层探测技术系统”是国家发展与改革委员会批准的中国地震局“十五”国家重大建设项目《中国数字地震网络工程》的重要组成部分^[3],包括基础地理信息、地震地质、地球物理和地球化学等不同的探测数据,涉及到众多的空间数据表和属性数据表。因此建立高效稳定的数据库是活断层探测项目中最终实现数据管理和数据演示的重要环节,也是城市活断层信息系统开发的关键。而使用适当的、高效的数据库建模技术并且正确地建立和使用面向对象的空间数据模型是重点研究目标。

根据《中国地震活动断层探测技术系统技术规程》中的相关规定,建立探测项目的信息系统所涉及到的空间数据以 ArcGIS 的相关数据格式存储,并且采用国家基础信息数据的格式,所以 ESRI 的 ArcGIS 软件已经成为实际应用中的标准软件。ESRI 公司也提出了与各行业有关的基于 ArcGIS 软件产品的 Geodatabase 数据格式的数据模型,比如 Geology Data Model。此模型是由英国地质调查局、荷兰地球科学应用研究所、美国地质调查局、加拿大地质调查局与 ESRI 公司共同开发^[4],但模型中没有涉及到与活动断层及活动断层探测手段相关的地震地质数据,如地质填图数据、探槽数据、钻孔数据等。又如 USGS2002 数据库扩展部分的地质断层参数数据库设计模型—SCEC2005 Fault Data Model 1.0^[5],提出关于断层参数估计的类型与断层研究有关的点状、线状类型的数据,如古地震探槽数据等,但没有提出面状数据,如用全站仪测量的冲沟断错位移,断层危险性评价中的等震线图等面状数据。因此有关断层面状数据的研究,将丰富和完善活动断层研究的内容。

城市活断层探测项目的空间数据模型按照探测方法不同可分为地震地质数据模型、地球物理数据模型和地理信息数据模型。其中地理信息数据模型的相关研究标准已较完善;而探测项目中所涉及到的多种地球物理探测手段较复杂,其数据模型将在后续的研究中尝试出概念设计模型;本文仅重点探讨了地震地质数据模型的建模问题,期望为城市活断层探测项目数据建模提供一些参考依据。

1 模型实现的关键技术与方法分析

1.1 UML 面向对象建模

UML 是用于面向对象软件设计的概念层建模

的新兴标准之一^[6]。它是一种可视化、文档化、基于面向对象方法的标准系统建模语言,用于概念层对结构化模式和动态行为进行建模。它给出了一套完整的符号体系和描述规则^[7]。

UML 已经被 ANSI(美国国家标准协会, American National Standards Institute)和 OMG(对象管理组织, Object Management Group)等国际标准管理机构采用为标准,也是被大多数公司与研究机构所采用的标准^[8]。

UML 模型分为静态结构和动态结构。静态建模机制是 UML 的基础,包括例图、类图、对象图、包、构件图和配置图等;状态图用来描述一特定对象所有可能状态(初态、终态、中间状态、复合状态)及其引起状态转移的事件,是对类图的补充。就地震地质数据模型设计而言,我们只关心数据模型的静态结构。在静态结构中,类、对象和它们之间的关联是面向对象技术中最基本的元素。UML 中类和对象模型分别用类图和对对象图表示,类图描述了类和类的静态关系,显示了信息的结构和系统的行为。

应用 UML 所设计的数据库概念模型即可以转化为 XML(eXtensible Markup Language,可扩展标记语言)文档,可实现在 CASE 工具中设计的数据模型从逻辑概念模型到物理结构的转化。XML 元数据交换(XML Metadata Interchange, XMI)是一个非常复杂的规范,XMI 并没有规定一种 XML 词汇表,没有定义 Class、Attribute、Association,而是定义了从元模型生成词汇表的一种算法^[9]。

1.2 地学数据模型设计环境分析

数据模型设计环境是一个软件设计工具与结构化开发方法的结合体,软件设计工具可以自动执行软件过程,而结构化的开发方法则是定义一个被自动执行的软件过程。软件设计工具可以为信息系统的设计与开发节省时间与投入,同时也能够提高地理信息系统的稳定性和易用性。

数据模型设计环境要满足数据建模应有的两个需求:一是模型概念的设计应能够更加精确地反映现实世界;二是能够快速实现概念模型到物理模型的转换。

目前,针对地球科学语义建模的方法和软件还不是很成熟,如前面所提到的 GML 等建模方法,仅作为一种研究性质的内容还在不断的改进中,提出的方法和开发的软件还无法应用到大型的工程项目中^[10-13]。而现今的工程项目中经常用到的基于统一建模语言的建模工具主要是目前比较流行的专注

于非地质领域内建模的 CASE 软件。

1.3 地震地质数据模型实现技术与方法

在地质数据模型实现过程中,地质数据建模设计环境是如何基于软件工程方法的?在现有的商用软件中,仅有 ArcGIS 的 Geodatabase 空间数据库可以用 UML 方法对空间数据库进行面向对象的设计与管理^[14-15],利用它可以建立面向对象的空间数据库,或者说实体关系型数据库。当前的商用关系数据库管理系统中,通过特定的数据引擎如 ArcSDE,就可以实现空间数据和非空间数据的统一存储和集中管理。

UML 概念模型的建立为物理模型及活断层探测数据库的建立提供了一个基础架构,通过 ESRI 公司提供的基于 Rational Rose 或者 Microsoft Visio2003 等 CASE 建模工具的 ArcInfo UML Model 数据模型的模板即可建立地震地质数据模型的 UML 概念模型,结合 XMI 则可以建立基于 Geodatabase 数据格式的 Microsoft Access (个人数据库) 或者 ArcSDE 的空间数据库。

2 地震地质数据模型设计

通过对 ArcGIS Geology Data Model、North American Geologic Map Data Model1.0^[16]、SCEC2005 Fault DataModel1.0 等模型的研究和借鉴,并结合活断层探测中所涉及到的地震地质数据提出 Seismogeologic Data Model(地震地质数据模型)的概念模型。地震地质数据模型结合了现今中国地震活断层探测项目与活动构造研究的特点而建立的,综合各种活动断层研究技术与鉴定方法,以下对活断层探测项目的信息系统建设的空间数据组织进行探讨。实现了对活断层探测项目数据组织和管理。

2.1 地震地质概念数据模型设计的特点

North American Geologic Map Data Model(北美地质图数据模型)从地质事件、地质过程、地质属性、地质关系、地质结构和地质单元等角度展示了对自然界的地质认识到地质概念模型的全过程,是现今国外对地质科学研究中数据建模的综合与总结,它是由亚利桑那州地质数据模型^[17]和爱达荷州地质数据模型^[18-19]等演化而来的。其设计的出发点是基于普通的地质研究领域,这些模型的共同特点均是基于普通地质学基础。本文在研究城市活断层地质数据模型的基础上,结合与地震有关的地质数据模型中的部分属性,探讨建立城市地震活断层探测地震地质领域的数据库模型,从而满足信息管理与

数据库建设的要求。地震地质数据模型的研究重点是与活动构造有关系的的研究方法与探测手段。

地震地质数据模型包含了地层学、岩石学、年代学、地貌学、构造地质学、活动构造学、古生物学、地图制图学、古地震及历史地震研究等领域的研究内容。其中,地层学与岩石学的内容包含与活动构造研究有关的地层时代、岩性、上下地层之间的关系等;与地震研究相关的地貌学研究内容主要为断错地貌类型及特点、位移量等;古生物学和年代学在确定地层年代方面起着重要作用。其实,地震地质数据模型关注的就是与地震有关的地质现象与地质过程,注重研究与活动构造有关的数据结构表达。

2.2 地震地质数据模型 UML 图

UML 的静态模型描述了系统类的框架结构,给出类之间的静态关系,描述了空间数据库的静态类结构。空间数据库的静态结构反映的是实现该功能需要哪些类来参与,这些类之间的关系如何等,这奠定了空间数据库基本表的框架结构。当然这里筛选出来的类并不一定是最终的类,随着分析的深入有些类可能消失,有些新类还会引入。内容将设计期的类转换成空间数据库中的基本表以及一些关系描述表,就可完成空间数据库的静态结构设计。

建立 ArcGIS 数据模型的目标是为 GIS 工程项目提供行业应用模型,地震地质数据模型实现了在应用模型的基础建立适用于地震行业 GIS 项目的特殊应用模型。ESRI 设计了基于 Geodatabase 数据库结构的 UML 规范的数据模板:ArcInfo UML Model,应用 Microsoft visio 软件设计与空间数据库有关的类。其类的集合是由 ClassExtension、ComplexEdgeFeature、ComplexJunctionFeature、EdgeFeature、Feature、FeatureClassExtension、FeatureClassExtension、NetworkFeature、Object、ObjectClassExtension、Row 等类组成。地震地质数据模型的空间数据顶层类为 Feature,定义了 ESRI Types: : esriFieldTypeGeometry 类型的 shape 特性,模型中的断层、盆地、地震带等空间数据类均派生于 Feature 类。在各个派生类中可以定义标记值的几何类型来确定类的空间特性。如 GeometryType 可以分为: esriGeometryPoint、esriGeometryMultiPoint、esriGeometryPolyline、esriGeometryPolygon;同样,在各派生类的标记值中定义了空间数据要素类的投影方式。

(1) 图 1 表示与地震地质空间数据研究有关的内容,共定义了 44 个空间数据要素类 (Feature Class)、9 个自定义域 (Domain)、24 个泛化 (Genera-

tion) 和 19 个二元关联 (Binary Association)。其中二元关联如图 2 所示。图表中的类所描述的地震地质数据模型所涉及到的主题范围主要由描述地形、地层、钻孔、野外地质调查、断层、年代、历史地震、震源深度估计和地震区划等要素类构成,要素类继承于 ESRI Class:Feature。由于建立的 UML 数据模型的每一个类的特性较多且复杂,所以文章中的类仅列出每个类的 ID 字段和较重要的自定义域字段。模型涉及的研究内容由表 1 所示。

表 1 地震地质空间数据模型所涉及的数据

类	涉及的数据类型
地形	盆地(面)、隆起(面)、褶皱(线)、地表破裂带(面)等
地层	地层(面)、岩性、岩层(面)、地层等值线(线)
钻孔	钻孔剖面数据(点)、钻孔数据(点)、跨断层钻孔剖面(线)
断层	地质调查点(点)、探槽(面)、断层测量点(点)、断层段(线)、全站仪测量数据(线)、全站仪测站点(点)、震源深度估计(点)
年代	样品(点)、化石(点)等
历史地震	历史地震烈度(点)、等震线图(线)
地震区划	地震区(面)、地震带(线)

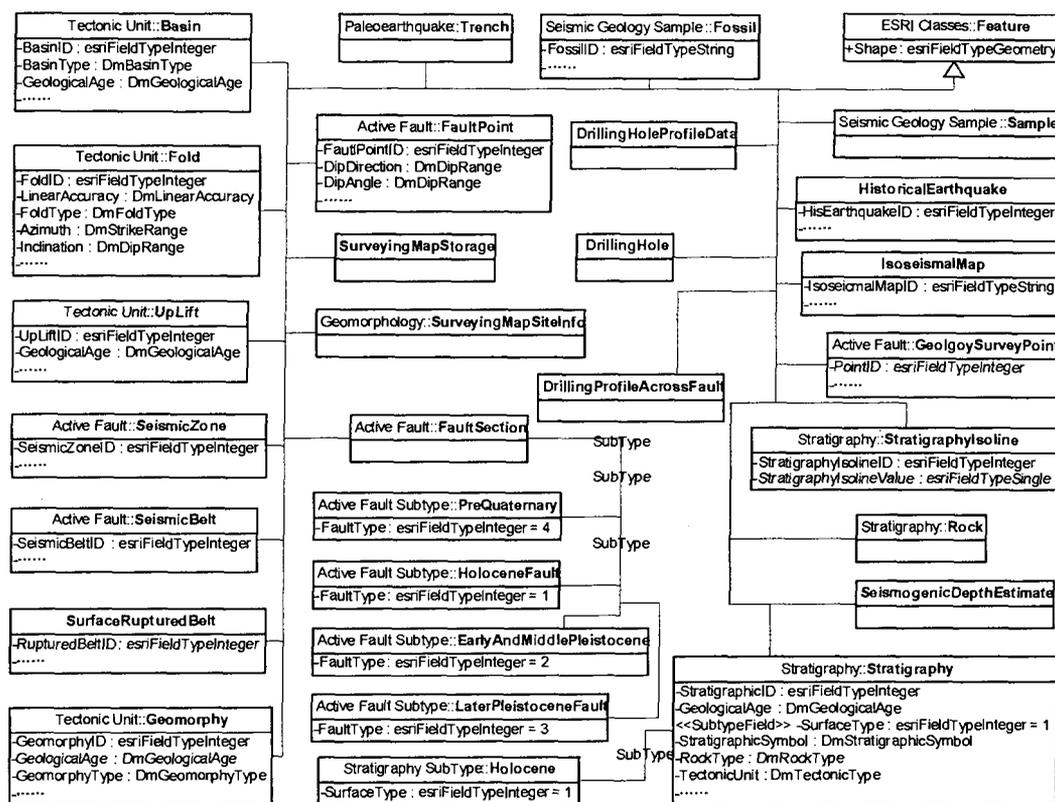


图 1 地震地质空间数据模型简单要素类图

Fig. 1 Simple Chart of feature classes of the seismo-geology spatial data model.

在地层和断层段要素类中分别定义子类型字段 SurfaceType 和 FaultType,且仅可定义为 esriFieldTypeInteger 类型。建立子字段类别的过程是通过二元关联的关系建立父类与子类字段相关的建模过程。在解决大系统问题时,常将许多类集成一个更高层次的单位,形成一个高内聚、低耦合的类集合,在 UML 中称这种分组机制为包。建模过程中,大致定义 ActiveFaultSubType 和 StratigraphySubType 两个包。在 ActiveFaultSubType 包中建立用于区分活断层时代的 4 个子类: HoloceneFault(全新世断层)、LaterPleistoceneFault(晚更新世断层)、EarlyAndMiddlePleistoceneFault(早中更新世断层) ProQuaternary-

Fault(前第四纪断层)。这样,就可以避免过多的建立不同时代的活断层要素类,实现 UML 模型的可读性。在 StratigraphySubType 包中则定义了区分不同地层类型的 15 个地层子类型,模型图中仅列出 Holocene(全新世地层)。

在 FaultSection 要素类中, FaultAttribute 字段类型为 DmFaultAttribute 域,域在 UML 建模过程中被定义为版类,可以为定义字段值和类型所用。Dm-FaultAttribute 定义了活断层性质,如正断裂、逆断裂、走滑断裂等。ESRI 的 UML 数据模板的域类型分为三类: TemplateCodedValueDomain、TemplateGeometricNetwork、TemplateRangeDomain。

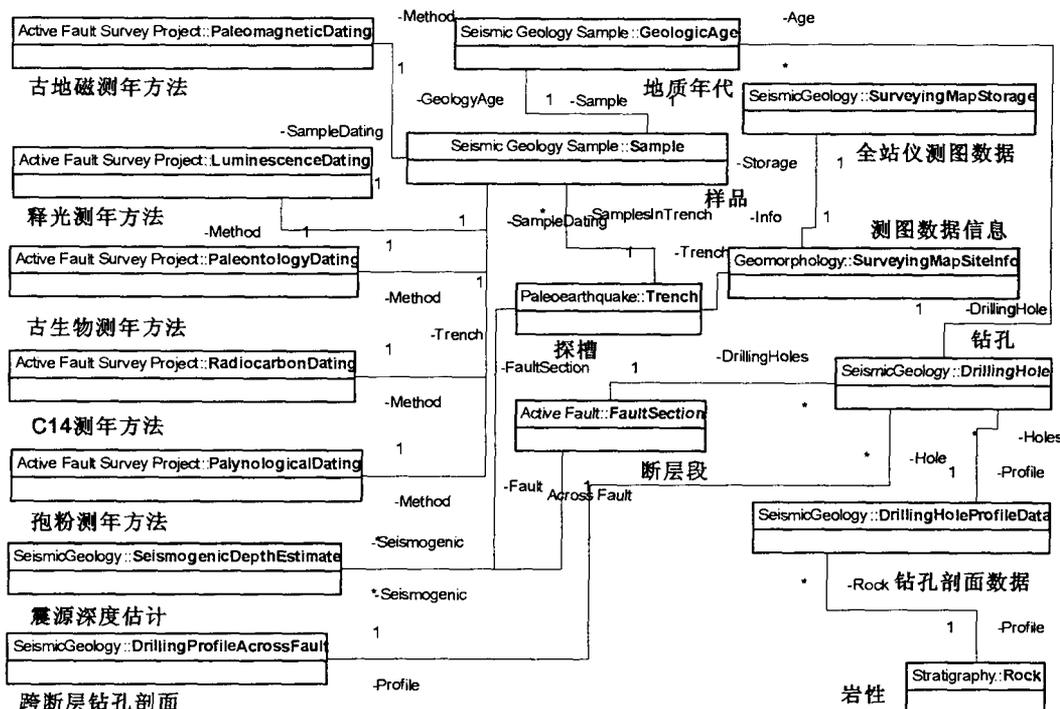


图2 地震地质空间数据模型要素类简单二元关联图

Fig.2 Chart of binary association of the seismo-geology spatial data model.

在 19 个二元关联中,分别为 15 个地层子类与父类地层的关联和 4 个断层段子类与父类断层段的关联。

(2) 图 2 定义了地震地质空间数据模型中 16 个要素类之间的二元关联。UML 关联描述了要素类和对象类之间的关系类,关联是模型元素的一种语义连接,关系被专业化为通用化、相关、关联、转移和链接,关联表示两个类之间存在某种意义上的联系。关系图表中的二元关联类型的依赖关系主要分为两种:1-1 和 1-M,在 1-M 的依赖关系中定义了 OriginClass、OriginPrimaryKey、OriginForeignKey、DestinationPrimaryKey 等标记值,二元关联的建立对于空间数据和属性数据关联查询起着重要的作用。

(3) 图 3 表示与地震地质的属性数据有关的数据模型内容,定义了 8 个属性类,分别继承于 ESRI Class: Object。其中 DesriptionAbs 为抽象类,DataQualit 类继承了 DesriptionAbs 类的所有特性。其中,年代学数据根据测年方法不同分为四类,分别为: LuminescenceDating (释光测年)、Palynological-Dating(孢粉测年)、PaleomagneticDating (古地磁测年)、RadiocarbonDating (¹⁴C 测年)。根据图 3 中所定义的属性类在 Geodatabase 数据库中转换所得到数据类型与继承于 ESRI Class: Feature 所得到的要素集不同,而是得到数据表(table)用于存储非空间

数据,诸如,在属性类字段定义的 esriFieldTypeBlob (Binary Large Object) 可以存储各种图片、文献、探测设计书等大型二进制数据。

3 结论与讨论

本文详细调研了北美地质图模型^[20]以及在北美地质图模型基础上发展的其它国家和地区的地质图数据模型,如澳大利亚的维多利亚地质图数据模型^[21]等,对正在实施的城市活断层探测项目中的地震地质数据模型的建立过程中类的选取与设计、类的属性设计及类之间的关联关系等几个关键问题进行探讨,城市活断层地震地质模型更侧重于活动断层探测的数据组织与管理。目前,本文所建立的地震地质数据模型仅是一个初步的、理想化的、概括性的活断层探测地震地质数据模型,期望能为全国城市活断层探测数据建模提供一些参考。

与活断层探测项目的探测规范所定义的数据表有小部分的出入,且只有与活动断层探测项目中将要建立的地球物理数据模型有机的组合,才能够真正地发挥数据模型在数据组织、管理及信息系统的建设过程的高效性地建模能力;模型中没有考虑探测规范中所提到的各种图片、文献、探测设计书等非空间数据的建模。在后续工作中,还有待于进一步地改善和研究地震地质 UML 数据模型。

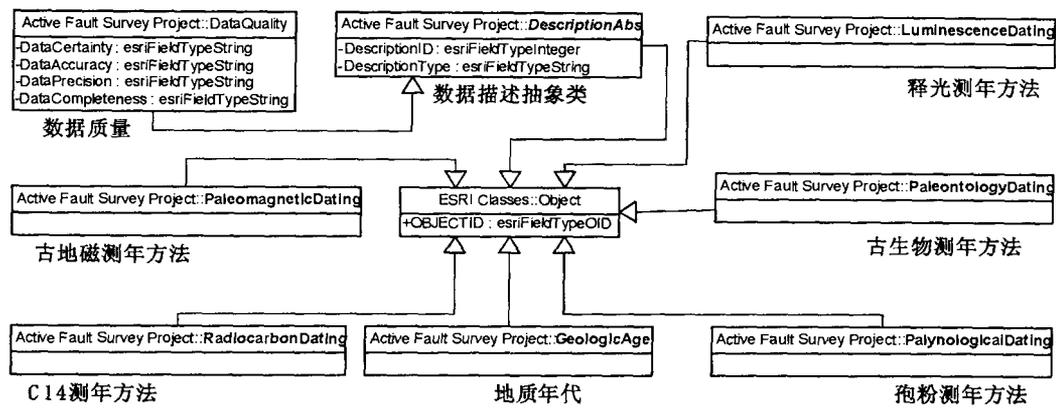


图 3 地震地质属性数据模型图

Fig. 3 UML data model of the seismo-geology attribute data.

探讨如何将地震地质数据的概念模型转换为基于商用数据库的物理模型,即如何应用 CASE 工具建立通过 OMG 定义的 UML 和 XMI 转换地震地质数据模型建立基于 ArcGIS 的 Geodatabase 数据库结构的地震地质数据库,将是以后研究的重点。

致谢:感谢中国地震局兰州地震研究所刘百篪研究员在论文写作过程中提出的关于古地震、历史地震及年代学等数据模型设计的意见和建议。

[参考文献]

[1] Simon Cox. CGIModel. ModelInventory, Inventory of geoscience data modelling projects [EB/OL]. <https://www.seegrid.csiro.au/twiki/bin/view/CGIModel/ModelInventory>, 2005.

[2] 中国地震局. 中国地震活动断层探测技术系统技术规程 [M]. 北京:地震出版社, 2004.

[3] Bruce R, Johnson, Boyan Brodaric. DIGITAL GEOLOGIC MAP DATA MODEL Version 4.3 [EB/OL]. <http://geology.usgs.gov/dm/>, 1999.

[4] Steve Grisé. ESRI Redlands, ArcGIS Geology Data Model [EB/OL]. <http://support.esri.com/>, 2002.

[5] Ned Field. SCEC 2005 CONSENSUS GEOLOGIC FAULT PARAMETER DATABASE [EB/OL]. <http://www.relm.org/mail-active-archive/msg00025.html>, 2005.

[6] Booth G, Rumbaugh J, Jacobson I. The unified modeling language user guide [M]. U. S. A. Addison Wesley, 1999.

[7] Shashi Shekhar, Sanjay Chawla, 著. 谢昆青, 等译. 空间数据库 [M]. 北京:机械工业出版社, 2004.

[8] Hans-Erik Eriksson, Magun Penker, Brian Lyons, David Fado, et al. 著. 余安萍, 等译. UML 2 Toolkit [M]. 北京:电子工业出版社, 2004.

[9] XML. Metadata Interchange (XMIO®) 2.1 [EB/OL]. http://www.omg.org/technology/documents/modeling_spec_catalog.htm#XML.

[10] Jochen Schmidta, Holger Gaertnera, Andreas Bergmannb. Object-oriented modeling of data sources as a tool for the integration of heterogeneous geoscientific information [J]. Computers &

Geosciences. 2001, 27:975-985.

[11] Ca'lin Arens, Jantien Stoter, Peter van Oosterom. Modelling 3D spatial objects in a geo-DBMS using a 3D primitive [J]. Computers & Geosciences. 2005, 31:165-177.

[12] Vijay Khatria, Sudha Ramb, Richard T. Snodgrass. On augmenting database design - support environments to capture the geo-spatio-temporal data semantics [J]. Information System, 2006, 31:98-133.

[13] C Parent, S Spaccapietra. The MurMur project: Modeling and querying multi-representation spatio-temporal databases [R]. Information System, Available online, 7 April 2005.

[14] 程昌秀, 周成虎, 陆锋. ArcInfo 8 中面向对象空间数据模型的应用 [J]. 地球信息科学, 2002, (1): 86-90.

[15] 李芳, 赵俊三. UML 和象形图建立 GIS 数据库概念模型初探 [J]. 测绘与空间地理信息, 2005, 28(3): 12-16.

[16] NADM Steering Committee Data Model Design Team. North American Geologic Map Data Model 1.0 [EB/OL]. <http://www.nadm-geo.org/>, 2004.

[17] Stephen M Richard. Geological Concept Modeling, with Examples for lithology and some other basic geoscience features [EB/OL]. <http://pubs.usgs.gov/of/1999/of99-386/richard.html>, 1999.

[18] Loudon R Stanford. Idaho Geological Survey's Digital Geological Map Data Model, Version 2.0 [EB/OL]. <http://www.nadm-geo.org/dmdt/pdf/IDAHO-Variant-Description.pdf>, 2001.

[19] Loudon R Stanford. A Brief Documentation of the Idaho Geological Survey's Digital Geologic Map Data Model, Version 2.1: A Variant of the North American Digital Geologic Map Data Model, Version 4.3 [EB/OL]. <http://www.idahogeology.org/Lab/data-model.htm>, 2002.

[20] Gerald A Weisenfluh. Map Unit Descriptions and the North American Data Model [EB/OL]. <http://pubs.usgs.gov/of/2001/of01-223/weisenfluh.html>, 2001.

[21] Bruce Simons, GeoDATA. Vic - Managing Victoria's Geoscientific MapData [EB/OL]. <http://www.dpi.vic.gov.au/dpi/nrenmp.nsf/>, 2005.