

# 空间数据仓库的认知过程

吴金华<sup>1</sup>, 祝国瑞<sup>2</sup>

(1. 长安大学 地球科学与国土资源学院, 陕西 西安 710054; 2. 武汉大学 资源与环境科学学院, 湖北 武汉 430079)

[摘要] 从理论上探讨了空间数据仓库的认知过程, 主要是十一个层次的世界模型通过十个算子转换的过程, 对十一个层次世界及十个算子概念进行了较详细地讨论, 并用代数系统对其进行了定义。

[关键词] 空间数据仓库; 世界模型; 转换算子

[中图分类号] P28 [文献标识码] A [文章编号] 1672-6561(2004)04-0067-04

[作者简介] 吴金华(1965—), 女, 陕西西安人, 副教授, 博士, 现从事地图制图学与地理信息工程教学与研究工作。

## 1 空间数据仓库

传统的GIS应用系统一般是面向某一个具体应用, 由日常的工作流程驱动的, 数据往往处于采集时的原始状态, 系统应用也只是对业务数据进行增、删、改等事物处理操作和简单的空间查询与分析。为了更好地适应当今全球变化和可持续发展研究的需要, 需要用一个统一的信息视图将来自不同学科的相关数据按照相应的主题转换成统一的格式, 集成、存储在一起, 然后通过各种专业模型从多个角度去了解这个世界, 空间数据仓库正是为了更好地满足这一需要而产生的。空间数据仓库的复杂性就在于数据的空间特性, 其核心问题是多源数据融合、联机分析和数据挖掘。

空间数据仓库和一般的空间数据库在物理本质上均是对数据高效地存贮。空间数据仓库是建立在传统的空间数据库管理系统之上, 依靠它们管理数据的存贮, 而不管它们是集中式的, 还是分布式的。二者之间的差别在于面向的应用不同, 因此在数据的组织、集成上存在着较大地差异。空间数据库(源数据库)负责原始数据的日常操作性应用, 提供简单的空间查询和分析; 空间数据仓库则根据主题通过专业模型对不同源数据库中的原始业务

数据进行抽取和聚集, 形成一个多维视角, 为用户提供一个综合的、面向分析的决策支持环境。这样, 空间数据仓库和空间数据库各自只做自己所擅长的, 系统的性能达到最佳。

## 2 空间数据仓库认知过程

当代地学问题的提出和解决常常要求多学科多部门之间的协作。因此, 空间信息的共享与互操作、分析与综合在对地学综合问题的研究中便显得日趋重要。人类对地理对象的认识是一个复杂的过程, 由于职业、专业、志向、社会地位等的不同, 人们往往在所关心的问题、研究的对象、期望的结果等方面存在着差异, 因而对于现实世界有着不同的看法、观念和兴趣, 对现实世界的描述和抽象亦是不同的, 形成了不同的用户视图。而地理信息的共享则是要对同一现实世界有同一的认识, 那么, 需要考虑用户需求的共性, 用统一的语言描述和综合、集成各用户视图, 形成对现实空间世界的概念模型。在建立此概念模型的基础上, 才能实现把一系列具体对象抽象为在计算机上能够表示的信息。然后将其有效地管理、组织, 并用于指导工作。这就是空间数据仓库的认知与抽象过程, 它可通过十一个层次的世界模型转换(图1)。

由图1可以看出, 空间数据仓库的认知过程主要就是十一个层次的世界通过十个算子转换的过程, 这个认知过程指导了空间数据仓库系统的实现。

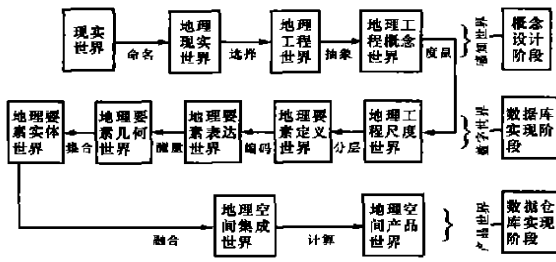


图 1 空间数据仓库的认知过程

Fig. 1 The cognition process of spatial data warehouse

## 2.1 层次世界模型

十一个层次世界模型描述的内容不同, 对其进行表述是空间数据仓库认知过程的重要部分。

### 2.1.1 现实世界模型

现实世界中, 在地球表面能看到一系列物质和现象, 对于这些物质和现象, 不管是否能叫上名字, 其都是客观存在的, 并且相互之间通过其组成形成了自然界的千差万别, 能将这些物质和现象用一定的代数形式进行描述就是现实世界模型。

定义 1: 现实世界模型。设  $A_1, A_2, \dots, A_n$  为  $n$  个地球表面上的不同物质和现象,  $R_A$  表示物质和现象之间的关系,  $\Omega_A = \{A_1, A_2, \dots, A_n\}$  表示地球表面上所有的物质和现象的组成集合, 那么现实世界模型可表示为  $Wor = \{\Omega_A, R_A\}$ , 其中  $0 \leq n \leq \infty$ 。

### 2.1.2 地理现实世界模型

在现实世界中, 包含地球表面上的所有物质和现象, 在这些物质和现象中就包含地球表面上的不同地理现象, 因此, 地球表面上的地理现象是地球表面上的所有物质和现象的一个子集。

定义 2: 地理现实世界模型。设  $B_1, B_2, \dots, B_m$  为  $m$  个地球表面上的不同地理现象,  $R_B$  表示地理现象之间的关系,  $\Omega_B = \{B_1, B_2, \dots, B_m\}$  表示地球表面上所有的地理现象的组成集合, 那么地理现实世界模型可表示为  $Gis = \{(\Omega_B, R_B) \mid (\Omega_B \in \Omega_A) \vee (R_B \in R_A)\}$ , 其中  $0 \leq m \leq n$ 。

### 2.1.3 地理工程世界模型

在地理现实世界模型中, 包含地球表面上的所有地理现象, 而地理工程世界模型是特指一个或若干个地球表面上的地理现象, 因此, 地理工程世界模型中的地理现象是地理现实世界模型中所有地理现象的一个子集。

定义 3: 地理工程世界模型。设  $e$  表示为地球表面上的某个特指地理现象, 那么地理工程世界模型

可表示为  $Ein = \{e \mid e \in Gis\}$ , 且  $\bigcup_{i=1}^m Ein_i = Gis$ 。

### 2.1.4 地理工程概念世界模型

概念世界是对现实世界的抽象, 是现实世界在人们头脑中的反映。在地理工程世界模型中特指一个或若干个地球表面上的地理现象, 要用计算机来描述这些地理现象, 就必须对这些地理现象进行抽象, 即首先给地理工程世界中可以识别出的地理现象赋予“名称”; 其次构架出描述地理现象的结构, 即构造出既方便人们认知又适合计算机解释和处理的实现模式, 概念世界模型是建立现实世界某些现象的抽象描述的过程。

定义 4: 地理工程概念世界模型。设  $o_1, o_2, \dots, o_k$  为地理工程世界模型  $Ein$  中某个特指地理现象的若干地理空间目标,  $R_o$  表示这些地理空间目标之间的关系,  $\Omega_o = \{o_1, o_2, \dots, o_k\}$  表示所有的地理空间目标组成集合, 那么地理工程概念世界模型可表示为  $Con = \{\Omega_o, R_o\}$ , 其中,  $o_i = \{(a_i, x_{i1}, y_{i1}, z_{i1}), \dots, (a_i, x_{il}, y_{il}, z_{il})\}$ ,  $a_i$  表示地理空间目标的属性集合,  $(x_{i1}, y_{i1}, z_{i1})$  表示地理空间目标的坐标集合,  $l$  表示地理空间目标坐标的个数。

### 2.1.5 地理工程尺度世界模型

通过地理工程概念世界的操作, 可以将地理工程世界中的地理现象抽象成概念模型, 用于在计算机中描述该地理现象, 但是仅仅有这些还远远不够, 因为在现实世界中所有的地理现象都是有度量的, 这些度量包括欧几里德几何坐标系、数学基础。

定义 5: 地理工程尺度世界模型。设  $Ein$  中欧几里德几何坐标系  $Coor = \{c \mid c \in 0, 1, 2, 3 \text{ 维}\}$ , 数学基础  $Math = \{Z, H, P\}$ , 其中  $Z = \{\text{旧 1954 年北京坐标系, 新 1954 年北京坐标系, 1980 年西安坐标系, 2000 年坐标系}\}$  表示坐标系基准集合,  $H = \{1956 \text{ 年黄海高程基准, 1985 年黄海高程基准}\}$  表示高程系基准集合,  $P = \{\text{无投影, 地图投影}\}$  表示地图投影集合, 则地理工程尺度世界模型  $Med = \{Coor, Math\}$ 。

### 2.1.6 地理要素定义世界模型

按照 GIS 理论, 地球表面上的地理现象最终都是通过地理要素来表达的, 但这些地理要素众多, 如何对其进行合理地设计和划分就显得特别重要。根据 ARC/INFO 的分层理论, 只有将这些地理现象划分成地理要素层, 才能高效地处理它们。因此, 地理要素定义世界中有一个基本概念必须搞清

楚,就是地理要素层,一组在地理意义上相关的地理要素的集合。

定义 6:地理要素定义世界模型。设  $E_{in}$  中地理要素分层表示为  $L = \{l_1, l_2, \dots, l_{17}\}$ ,  $l_i = \{l_{i1}, l_{i2}, \dots, l_{ij}\}$ ,  $l_i$  代表某一地理要素分层,  $l_{ij}$  代表该分层中的小类,  $i, j$  为正整数, 则地理要素定义世界模型为  $F_{ea} = \{l_i \mid l_i \in L\}$ 。

2.1.7 地理要素表达世界模型

在上面的工作完成后,所需做的一项重要事情是如何在计算机中表达出已抽象出的具有度量的地理要素。众所周知,要使计算机能识别和处理现实世界中的事物和现象,就必须给这些事物和现象进行分类编码,即用一串数字来表示它们,该分类编码就成为某一事物和现象在计算机中的唯一标识,以便计算机能识别和处理。

定义 7:地理要素表达世界模型。设  $E_{in}$  中地理要素分层表示为  $L = \{l_1, l_2, \dots, l_{17}\}$ ,  $l_i = \{l_{i1}, l_{i2}, \dots, l_{ij}\}$ ,  $l_i$  代表地理要素分层,  $l_{ij}$  代表该分层中的小类,  $i, j$  为正整数;地理要素编码表示为  $Mod = \{m_1, m_2, m_3\}$ ,  $m_1, m_2, m_3$  分别代表该地理要素的大类码、小类码和顺序码, 则地理要素表达世界模型为  $Exp = \{l_i \mid (l_i \in L) \wedge (l_i \in Mod)\}$ 。

2.1.8 地理要素几何世界模型

地理要素的整体或局部在数字世界的反映就是基本逻辑单元或叫地理要素几何目标,地理要素几何目标包括基本目标、复合目标和注记目标。基本目标:地理要素按地理实体的空间特征划分为点状目标、线状目标、面状目标、体状目标和表面状目标等 5 种基本目标;复合目标:由同一要素层中的基本目标有机集合构成的目标,复合目标可以嵌套;因此,最终地理要素几何目标分为点、线、面、表面和体五类空间目标。

定义 8:地理要素几何世界模型。设  $E_{in}$  中地理要素的点状目标集合为  $Po$ , 线状目标集合为  $Lo$ , 面状目标集合为  $Ao$ , 体状目标集合为  $To$ , 表面目标集合为  $So$ , 则地理几何世界模型表示为  $G_{eo} = \{Po, Lo, Ao, To, So\}$ , 其中,  $Po = \{x, y, z\}$ ,  $Lo = \{e \mid e \in Po\}$ ,  $Ao = \{e_1 \mid e_1 \in Po, e_2 \mid e_2 \in Lo\}$ ,  $To = \{e_1 \mid e_1 \in Po, e_2 \mid e_2 \in Lo, e_3 \mid e_3 \in Ao\}$ ,  $So = \{e_1 \mid e_1 \in Po, e_2 \mid e_2 \in Lo, z\}$ 。

2.1.9 地理要素实体世界模型

地理实体是地球上的一种真实地理现象,在概念世界中,将不能再细分的同种类型的地理现象称

为地理要素实体。在计算机数字世界中,地理要素实体就是具有同一地理特征的地理要素几何目标集合,这些地理要素几何目标分为点、线、面、表面和体五类空间目标。

定义 9:地理要素实体世界模型。设  $E_{in}$  中地理要素点状目标集合为  $Po$ , 线状目标集合为  $Lo$ , 面状目标集合为  $Ao$ , 体状目标为  $To$ , 表面目标表示为  $So$ ,  $Atr$  为地理要素的某一地理特征集合, 则地理要素实体世界模型为  $Ent = \{e \mid (Po, Lo, Ao, To, So) \in Atr\}$ 。

2.1.10 地理空间集成世界模型

在空间数据仓库中,地理空间数据种类有 DLG, DRG, DOM, DEM 和大地成果,将这些种类的地理空间数据进行集成,具有特别重要的意义,它能为用户提供许多地理空间数据产品。地理空间数据的集成就是在统一的地理空间数学基础上实现正射影像、地面高程模型、栅格地形图和矢量地形图的叠加,以达到相互比照和综合利用之目的。

定义 10:地理空间集成世界模型。设  $E_{in}$  中地理空间数据种类为 DLG, DRG, DOM, DEM 和大地成果数据,它们的数据集合分别表示为  $Dlg, Drg, Dom, Dem$  和  $Ddd$ , 地理空间数据集成集合为  $\Omega_{int} = \{(Dlg \cap Dom), (Dlg \cap Drg), (Dom \cap Dem), (Drg \cap Dem), (Dlg \cap Dom \cap Dem)\}$ , 则地理空间集成世界模型  $Int = \{\Omega_{int} \mid \Omega_{int} \in Mea\}$ , 其中  $Mea$  为地理工程尺度世界模型。

2.1.11 地理空间产品世界模型

多个种类的地理空间数据集成后,具有统一的地理空间数学基础,对单个或叠加后的地理空间数据种类进行专题计算处理后得到的产品为地理空间数据产品。这些地理空间数据产品划分为单一的数字产品、融合的数字产品、派生的数字产品、加工的数字产品和关联的数字产品。

定义 11:地理空间产品世界模型。设  $E_{in}$  中地理空间数据种类为 DLG, DRG, DOM, DEM 和大地成果数据,专题计算算法集合为  $Fru$ , 则地理空间产品世界模型  $Pdu = \{\Omega_{int} \mid (\Omega_{int} \in Mea, \Omega_{int} \in Fru)\}$ 。

2.2 转换算子

在空间数据仓库的认知过程中,这十一个层次世界的变换离不开这些转换算子,其分别是命名、选择、抽象、度量、分层、编码、测量、集合、融合和计算这十个算子,由它们实现每两个世界模型的转

换。虽然这些转换算子的具体定义不同,但都是实现每两个世界模型的转换,从数学的定义上说就是某个世界模型通过函数转换到另一个世界模型上,因此这些转换算子的数学定义是一致的。

定义 0: 转换算子。  $X$  和  $Y$  是两个世界模型的集合表示,  $F$  是  $X$  到  $Y$  的一个转换算子, 如果对于每一个  $x \in X$ , 有唯一的  $y \in Y$ , 使得  $\langle x, y \rangle \in F$ , 则称  $F$  为转换算子, 记作:  $F: X \rightarrow Y$ 。

在代数定义中,  $X$  或  $Y$  就代表着现实世界、地理现实世界、地理工程世界、地理工程概念世界、地理工程尺度世界、地理要素定义世界、地理要素表达世界、地理要素几何世界、地理要素实体世界、地理空间集成世界和地理空间产品世界模型,  $F$  代表着命名、选择、抽象、度量、分层、编码、测量、集合、融合和计算算子。

### 3 结束语

空间数据仓库的认知过程, 主要就是十一个层次的世界通过十个算子转换的过程, 这个认知过程

指导了空间数据仓库系统的实现。但是文中只从理论上对其进行了初步讨论, 并用代数系统对这十一个层次的世界模型及转换算子进行了定义。如何用代数系统描述它们之间的代数运算关系是下一步需要研究的一个深层次问题。

#### [ 参 考 文 献 ]

- [ 1 ] 承继成, 赵永平. 地理信息及其元数据标准化[ J ]. 遥感学报, 1998, 2(2): 149 ~ 154.
- [ 2 ] 王珊. 数据库技术与联机分析处理[ M ]. 北京: 科学出版社, 1998.
- [ 3 ] 赵永平. 基于国家空间信息基础设施的元数据研究及其共享示范体系的建立[ D ]. 北京: 北京大学, 1998.
- [ 4 ] 周成虎, 李军. 地球空间元数据研究[ J ]. 地球科学——中国地质大学学报, 2000, 25(6): 579 ~ 584.
- [ 5 ] David Danko. The ISO Metadata Standard for Geographical Data [ A ]. Geo-Informatics Conference of the International Eurasian Academy of Sciences and the Fourth International Workshop on Geographical Information System[ C ]. Beijing: 1979. 98 ~ 104.
- [ 6 ] Kleinrock L. Technology Issues in the Design of the NREN. In Building Information Infrastructure (B. Kahin Ed.) [ M ]. New York: McGraw-Hill. 1992.

## Research on the cognition process of spatial data warehouse

WU Jin-hua<sup>1</sup>, ZHU Guo-rui<sup>2</sup>

(1. School of Earth Sciences and Resources Management, Chang'an University, Xi'an 710054, China ;

2. School of Resource and Environment Science, Wuhan University, Wuhan 430079, China)

**Abstract** The paper discusses the cognition process of spatial data warehouse in theory, and it is considered primarily the transforming process of eleven levels world by ten operators. The concepts about the eleven levels world and the ten operators are discussed. All of them are designed with algebra system.

**Key words** spatial data warehouse; world model; operator

[ 英文审定: 马智民 ]

(上接第 48 页)

## Design of lattice frame anchor structure for landslide control

LIU Ni-na, LIU Cong, LI Xun-chang, MEN Yu-min, SHI Liang

(School of Geological Engineering and Surveying Engineering, Chang'an University, Xi'an 710054, China)

**Abstract** Taking shizishan landslide control in Lueyang Country as an example, the whole design process of lattice frame anchor structure are studied. Firstly, according to the geological conditions and technical code for design, safety criterion for landslide control is confirmed. Secondly, by making use of the features of lattice frame anchor structure which can protect slope and retain stability of landslide together with its small space between beams and light anchor, the calculation model and formulas of slope force are dealt with and the design process of lattice beam and pillar as well as anchor and plate are introduced carefully. Finally, the advice and request for construction are specific ally put forward.

**Key words** loosen body landslide; lattice frame anchor structure; design; landslide control

[ 英文审定: 苏生瑞 ]