

空间数据库查询语言的特征

李霖

(武汉测绘科技大学土地科学学院,武汉市珞喻路 39号,430070)

摘要 简介了目前空间数据库查询语言研究情况和所具有的形式,按空间数据的特点和查询语言的一般要求,认为空间数据库查询语言应具有以下特征:能理解“空间概念”,能描述查询结果的表达方式和查询描述的非过程化。指出了目前空间数据库查询语言集中的3个研究方面各自的特点及待研究的主要问题。

关键词 空间数据;数据库;查询语言

分类号 TP 311.13

空间数据库是一种特殊的数据库,它以空间(地理)目标作为存储集,与一般数据库的最大不同是它包含“空间”(或几何)概念,因而,理解空间概念是空间数据库查询语言必备的前提,尽管关于空间的概念很复杂,查询语言难以全部包含,但至少应该理解一些普通空间概念。从这个意义讲,空间数据库比一般数据库所体现的语义应更具体一些,显然标准的SQL不支持空间概念(如空间关系)。

要实现空间数据库的查询,必须增加空间或几何概念,许多文献^[2, 6, 8-12]对空间数据库查询的不同方面作了很有意义的研究和讨论。大部分研究的基础是关系型数据库的查询语言,通过适当扩展实现空间数据库的查询功能。

1 查询语言的表现形式

关系数据库以其概念简洁、使用简单在商业上取得了巨大的成功,其管理软件目前在市场上占统治地位,关系数据模型的概念和方法对许多实际运行系统的设计和开发有着不可估量的影响。但是,当把关系数据模型用于空间数据(或地理数据)时,就暴露出其天生的弱点^[7]。

关系数据库的结构化查询语言(SQL)以其非过程化的描述、简洁的语法对空间数据库查询语言的研究有很大的影响。为了使查询语言能完成空间数据的查询,通常可以在SQL上扩充谓词集,使之包含空间关系谓词,并增加一些空间操作。如Egenhofer^[8]根据空间数据库的特点以及空间数据表示的要求,在关系型的SQL上发展了一套空间结构化查询语言或空间SQL(Spatial

SQL),在查询结果表示方面,提供了包含6个显示参数的显示环境。这种Spatial SQL实质上是由基于SQL风格的空间数据库查询语言和图形表示语言(GPL)两部分组成。例如查询所有高速公路并用红虚线表示,用Spatial SQL表示为:

```
SET COLOR red
PATTERN dashed
FOR SELECT geometry
FROM roads
WHERE type= "Highway"
```

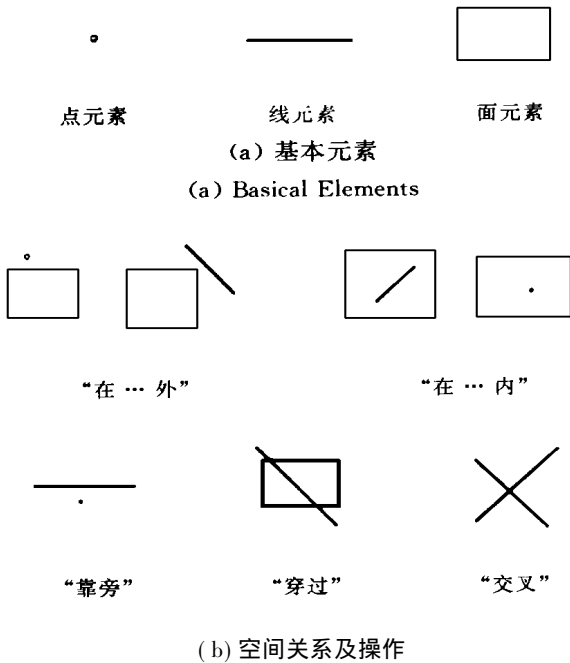
GeoSAL^[1]也是一种基于关系模型查询的扩展语言,这种查询语言可以用于空间查询和某些空间分析,它是在称之为Cantor的关系数据库系统(RDBMS)查询语言SAL上扩展而来,Cantor是由瑞士位于Stockholm的国防研究所(FOA)开发的,用于统计数据的处理和分析。为了能处理空间概念,GeoSAL增加了空间函数和操作来表示空间数据的处理,并且在数据模型中增加了集合类型。这种查询语言不仅可以用来提取空间数据库的数据,还可以用来描述空间操作(如叠置、插值等),但它的表示具有函数形式,没有SQL那样简洁。例如查询所有通过B省的第二级公路,用GeoSAL表示为:

```
OBJECT Roads2 Lineset [nα Integer]:
= * (α Roads, b Provinces)
WHERE [a. class= 2) AND (b. name=
' B )
AND (b. region INTERSETCS
a. geometry) ]
[geometry:: a. geometry, nα a. no];
```

GeoSAL查询的描述是将数据查询操作与数

据定义结合起来,查询结果生成一个新的目标集。

可视查询^[3,12,14]主要是将查询语言的元素,特别是空间关系,用直观的图形或符号表示,因为某些空间概念用二维图形表示比一维文字语言描述更清晰易理解。例如,某些空间操作可用图 1 表示



(b) Spatial Relations and Operations
图 1 可视化查询的语言元素

Fig. 1 Visual Query Language Elements

但是以图形、图像或符号为语言元素的可视查询,仅是对查询直观形象化的描述,在这个意义下,它仅是空间查询语言的一个子集。由于只有部分空间概念可用与人类空间观念一致或接近的图形或图像表示,某些查询很难用人类易理解的图形表示,而且这种查询最终要落实在某一查询模型上,因此,目前意义上的可视查询不可能表达所有的空间查询。

为了使查询语言的描述更接近人的自然语言,在查询语言中引入自然语言的概念可以使数据库查询更轻松自如^[2,4]。对 GIS 的应用来说,很多地理概念是模糊的,比如地理区域的划分实际上并不像境界那样有明确的界线,而查询语言往往设计用来表示精确的概念。Wang^[4]用模糊数学的方法先将模糊概念量化为确定的范围,实现具有能理解某些模糊概念的查询。例如,查询高气温的城市名称,首先表示为:

```
SELECT name
FROM Cities
```

WHERE temperature is high

通过对许多数据的统计处理和分析,按模糊

公式计算,城市气温 $\geq 33.75^{\circ}\text{C}$ 时认为是高气温,因此上述查询的表示在实施查询操作时转化为确定的查询描述:

```
SELECT name
FROM Cities
WHERE temperature  $\geq 33.75$ 
```

这类查询,其模糊概念有很强的专业性和区域性,语义很强。如概念“高(high)”对城市气温这个属性,其阈值为 33.75,而对人的身材高矮,则其阈值是另外完全不同的值;即使对城市气温,统计范围(区域)也很有关系,北半球“高气温”阈值不一定适合于南半球。因此,具有自然语言的查询显然不能作为通用数据库查询语言设计的目标,它只能适合于某个专业领域的数据库应用。

2 空间数据库查询语言的特征

对于空间数据库,其查询语言中最关键的部分是空间概念的描述,理想的情况是其语言完全能表示人所能理解的空间含义,目前这显然是不可能的事情。对空间数据库的查询而言,应具有什么样的特征呢? Frank 和 Mark^[2]认为空间查询应具备两个基本功能: (1) 选取用户所需要的子集; (2) 以有意义的形式将查询结果呈现给用户。

Egenhofer^[9]根据一般数据库查询语言的特点以及空间数据的特征,归纳总结了空间查询语言应满足 11 项具体的要求。根据这些要求, Egenhofer 比较了 5 个比较有名的空间数据库查询语言,其结果如表 1

表 1 5 个基于 SQL 扩充的空间数据库查询语言特征的比较

Tab. 1 Evaluation of Five SQL Extensions for Spatial Data

	GEOQL	Extended SQL	PSQL	KGIS	TIGRIS
1. 抽象数据类型“空间”	√	√ ^{1,2}	√	√ ²	√ ²
2. 查询结果的图形表示	√	√	√	√	√
3. 查询的组合	—	—	—	√ ³	—
4. 带上下文的图形表示	—	—	√ ⁴	√	—
5. 多种查询组合的内容	—	—	—	—	—
6. 扩展的对话方式	√	—	—	√	—
7. 修改显示方式	—	—	√ ⁴	—	—
8. 描述性的图例	—	—	—	—	—
9. 适当的注记	—	—	√	—	—
10. 显示比例尺	—	—	√	—	—
11. 区域的限定	—	—	—	√	—

注: 1 只有空间关系, 2 无数据定义, 3 只对上下文, 4 作为列在 ON-语句图片的一部分。

从表中可以看出,这5种语言没有一个全部满足这11项要求,多数只满足少数几项要求,有些非常有意义的特征被多数查询语言所忽略,即使像“支持空间概念”这一基本特征,尽管这些语言都具有,但它们理解空间的概念很有限,这说明空间数据库查询语言有许多方面亟待研究。

从空间数据库与一般数据库的关系以及空间数据描述实际地理现象的特点看,空间数据库查询语言的特征可以概括为3条:

- 查询语言能理解“空间概念”;
- 查询语言能描述查询结果的表达方式;
- 查询描述的非过程化

3 空间数据库查询语言待研究的问题

查询语言的目的就是使语言环境能尽可能地满足各种用户的要求。对空间数据库而言,查询语言不仅能描述有关空间位置的查询,还应能表达数据间空间关系的查询,以及数据的空间操作结果。然而数据的空间特性很复杂,空间数据库的建立总是基于某一比例尺,数据的尺度性^[13]也属空间特征,目前的空间数据模型还没有有效工具来表示空间目标的尺度性。

空间数据库查询的结果亦为空间数据,它代表了现实世界中某一空间现象或空间现象的分布规律,也在某种程度上反映人们的认识水平。空间数据的不同表示会给人们不同的感受,从而形成不同的概念(或观念),因此结果的表示是查询语言不可缺少的部分。若没有适当的表示,即使内部查询到所需结果,如果用户不能理解或产生误解,那么毫无益处。

目前空间数据库查询语言的研究主要集中在以下方面:

—— 扩展 SQL 关系数据库的用户占主导地位,SQL 已作为工业标准,对其 SFW 风格用户已习惯,扩展 SQL 可以马上得到用户的认可和接受,但受其关系模型的限制,表达空间查询的能力很有限;

—— 可视化查询语言: 语言本身具有空间特征,表达某些空间查询简单、明了,但查询能力有限,很难为语言建立形式化的基础;

—— 基于自然语言的查询: 用户使用简单、自如,没有学习查询表达负担,但数据库查询中的概念与语义背景有关,仅限于专业数据库查询。

由此可知,空间数据库查询语言至今还没有

达到完善的程度,有许多工作要做,其中最主要的方面有:(1)如何完善空间数据模型,达到支持“完备”的空间概念;(2)拓展查询模型,使查询空间尽可能地与空间数据表述的现实空间接近。

参 考 文 献

- 1 Huang Z. Design of GeoSAL. A Database Language for Spatial Data Analysis [Dissertation Thesis]. School of Surveying, Royal Institute of Technology, 1993.
- 2 Frank A U, Mark D M. Language Issues of GIS. In: Geographical Information Systems Principles and Applications. London: Longman, 1991. 147~ 163
- 3 Du H, Azmoodeh M. SQL: A Graphical Database Language Using Pattern Images. In: New Trends in Computer Graphics Proceedings of CG International 88. Springer-Verlag, 1988. 611~ 620
- 4 Wang F. Towards a Natural Language User Interface: An Approach of Fuzzy Query. INT. J. GIS, 1994 (2): 143~ 162
- 5 Raper J, Bundock M. Development of a Generic Spatial Language Interface for GIS. In: Geographical Information Handling-Research and Applications. John Wiley & Sons Ltd, 1993. 113~ 144
- 6 Becker L, Guting R H. Rule-Based Optimization and Query Processing in an Extensible Geometric Database System. ACM Trans. Database System, 1992, 17 (2): 247~ 303
- 7 Worboys M F, Hearnshaw H M, Maguire D J. Object-Oriented Modelling for Spatial Databases. INT. J. GIS 1990, 14 (4): 369~ 383
- 8 Egenhofer M J. Extending SQL for Graphical Display. Cartography and Geographical Information Systems, 1991, 18 (4): 230~ 245
- 9 Egenhofer M J. Spatial SQL: A Query and Presentation Language. IEEE Trans. Knowl. Data Eng., 1994, 6 (1): 86~ 95
- 10 Boursier P. Analysis of Urban Geographic Queries. In: Magnenat Thalmann N, Thalmann D, eds. New Trends in Computer Graphics Proceedings of CG International 88. Springer-Verlag, 1988. 601~ 610
- 11 Kerr R L, Phillip C-Y. Sheu, SPARQ: Spatial Relationship Query Based on Spatial Decomposition. Computer and Graphics, 1991, 15 (2): 267~ 284
- 12 Lee Y C, Chin F L. An Iconic Query Language for Topological Relationships in GIS. INT. J. GIS, 1995, 9 (1): 25~ 46
- 13 李霖,李德仁. GIS 二维空间目标的非原子性和尺度性.测绘学报,1994, 23 (4): 315~ 321
- 14 刘东.空间查询语言的原理与应用. [学位论文].

北京: 北京大学遥感所, 1996.

The Features of Spatial Database Query Languages

Li Lin

(School of Land Science, W TU SM, 39 Luoyu Road, Wuhan, China, 430070)

Abstract Query languages are the most sensitive functions to users. Spatial database query languages should completely supply users with needed data which describe some spatial phenomena. This papaer will give a brief description of the current situation on the spatial database query languages and their main forms. Through analysis of the requirements for spatial database query languages, we boil down, based on characteristics of spatial data and general requirements of query languages, that the query languages must be equipped with the three features "understanding "spatial concepts", descriptive representation of queried results and declarative query process. In the final part of this paper, we outline the main features of three main aspects in spatial database query languages and the key research topics on the query languages in the future.

Key words spatial data; databases; query languages

(上接第 95页)

Some Thoughts about Geographic Information Theory

Li Deren

(The Secretariat of President, W TUSM, 39 Luoyu Road, Wuhan, China, 430070)

Abstract Geographic information systems are very important vehicles running on Information Superhighway and serve many different users with geo-spatial data. The two characters of GIS are technology-driven and application-oriented. With the rapid development of computer technique, communication technique, space positioning technique, remote sensing and visualization technique GIS hardware and software environment and the methods of spatial data acquisition, storage, access, analysis, representation and visualization will be improved continually and rapidly. Facing this situation the author try to think about geographic information theory. Main discussion in this paper includes optimal representation of spatial data with attribute data in computer, automatic acquisition and updating of spatial data, spatial data quality and uncertainty, spatial query and analysis of GIS data, visualization symbolization and multi-scaling representation of spatial data etc. On the discussion we pay more attentions to compare and to distinguish the differences between computer based GIS and paper map based "analogue GIS" and to consider the developing tendency of tomorrow's computer technique and information science.

Key words geographic information theory; spatial database; cartographic database