

基于状态和变化的统一时空数据模型*

郑扣根, 谭石禹, 潘云鹤

(浙江大学 人工智能研究所, 浙江 杭州 310027)

E-mail: zkg@cs.zju.edu.cn

http://www.zju.edu.cn

摘要: 提出了一个基于状态和变化的统一时空数据模型 SCUDM (state and change based unified spatio-temporal data model). 时空数据库的状态是关于对象、空间 and 时间的三元组, 时空数据库的变化就是状态的变化. 按照对象域、空间域和时间域将时空数据库中的变化分别投影分解, 显式地表达在对象、空间和时间上的变化. 而事件是在某一时刻前后, 时空实体的变化关系, 在事件表达中蕴涵了空间实体之间的时间拓扑关系. SCUDM 不仅支持状态和变化的表达, 而且支持时间拓扑和空间拓扑的表达.

关键词: 时空数据模型; 时空数据库; 时空语义; 状态变化; 时空拓扑

中图法分类号: TP311 **文献标识码:** A

时间数据库和空间数据库已得到了广泛研究, 然而同时支持时间-空间的数据建模和时空数据库只是在近 10 年才引起人们的注意^[1]. 在传统的面向记录的数据模型(关系模型)中, 数据库由一系列的定长记录表达实体和关系. 它虽然有一套成熟的商用的数学基础和形式化、规范化的建模方法, 但却缺乏丰富的数据语义描述能力和灵活的数据组织能力, 不适合于复杂的时空数据模型.

在已有的文献中, 时空数据建模大多采用面向对象的思想, 语义表达和建模能力都得到了很大的加强. 但针对时空的特殊性, 一般的对象数据模型则缺乏两方面的特性.

(1) 时空结构化的组织数据. 在面向对象的数据模型中, 侧重于对对象的描述, 一般采用抽象数据类型(abstract data type, 简称 ADT)机制构造时间、空间和时空数据类型. 时空数据库对时空的所有支持其实是针对特定的时空抽象数据类型的支持. 对于对象而言, 时空属性和其他属性是没有什么区别的. 换言之, 在构造数据时, 没有按时空语义来组织数据.

(2) 时空变化的语义. 在时空数据建模中, 时空的特定语义是问题的关键, 总的思想是提升抽象的层次, 从基于时间戳的方法^[2,3]到基于事件的方法^[4,5]. 时间的本质是变化, 在时空数据建模中, 变化可以是关于时空实体的属性、位置、形状的变化, 也可以是关于拓扑关系的变化, 这使得完整地表达空间的变化变得困难了. 虽然已有的时空数据建模方法基本上都强调对变化的描述能力, 但至今在理论上仍然缺乏一种支持变化的统一的数据模型, 特别是不能表达空间拓扑的变化.

注意到在对真实世界的时空现象建模时, 空间、时间和对象的独立性包含了丰富的语义, 本文对时间和空间的语义进行抽象和形式化, 定义了时空数据库中的状态、变化和事件的概念, 在此基础上提出了一个基于状态和变化的统一时空数据模型 SCUDM (state and change based unified spatio-temporal data model).

* 收稿日期: 1999 09 13; 修改日期: 2000 05 26

基金项目: 国家 863 高科技发展计划资助项目(863-306-04-03-3)

作者简介: 郑扣根(1964-), 男, 江苏镇江人, 博士, 副教授, 主要研究领域为计算机图形/图像, GIS; 谭石禹(1970-), 男, 江西景德镇人, 博士, 主要研究领域为时空数据库, 时空推理研究; 潘云鹤(1946-), 男, 浙江杭州人, 教授, 博士生导师, 中国工程院院士, 主要研究领域为计算机图形学, 人工智能, 认知科学, GIS.

1 状态和变化表达

SCUDM 数据模型是在时空复合概念^[2]的基础上发展起来的,为了下文叙述方便,我们先定义几个概念:

定义 1. 时空元素是时空立方体范围内最大的共同时空单元.

定义 2. 空间实体是在时空立方体范围内一定时间段上不变的空间状态.

定义 3. 时空对象是时空数据库中的语义对象,它由语义上的特征标示自己,与特定的时间和空间状态无关.

将时空对象与具体的空间和时间状态相分离,体现了空间、时间相对于对象的独立性.在时空数据库中定义的时空对象的集合称为数据库的对象域.同样,我们把数据库的空间定义范围称为数据库的空间域,数据库的时间定义范围称为数据库的时间域.

定义 4. 时空数据库的状态是一个三元组 $state = (o, s, t)$, $o \in OBJ$, $s \in SPACE$, $t \in TIME$; OBJ , $SPACE$ 和 $TIME$ 分别是时空数据库的对象域、空间域和时间域.

定义 5. 时空数据库的变化是关于数据库状态的变化,是从一个数据库状态到另一个状态的函数 $f^{OST}: state \rightarrow state$, f 的上标 OST 表示状态可在对象域、空间域和时间域之间变化.变化过程是一系列的变化 $F^{OST}: f_1^{OST}, f_2^{OST} \dots f_m^{OST}$.

状态的变化可能会很复杂,状态三元组中对象变量、空间变量和时间变量都可能在它们各自的域内独立地变化,描述起来可能很复杂,甚至很困难.为了方便描述时空现象中所发生的变化,我们把 f^{OST} 和 F^{OST} 从 3 个域上的变化降到两个域上的变化,将状态的变化分别沿着对象域、空间域和时间域进行投影,这样,数据库的变化就分解到以下的 3 个变化:

- 对象域上的变化, $f^{ST}: State|_o \rightarrow State|_o, F^{ST}: f_1^{ST}, f_2^{ST} \dots f_m^{ST}$.
- 空间域上的变化, $f^{OT}: State|_s \rightarrow State|_s, F^{OT}: f_1^{OT}, f_2^{OT} \dots f_m^{OT}$.
- 时间域上的变化, $f^{OS}: State|_t \rightarrow State|_t, F^{OS}: f_1^{OS}, f_2^{OS} \dots f_m^{OS}$.

2 数据模型

SCUDM 数据模型是一个 3 层的数据模型,如图 1 所示.时空对象处在模型的最上层,它的空间属性指向多个空间实体.换句话说,空间实体构成它的多个历史版本.空间实体处在中间层,在同一时间点的空间实体构成此时间的静态图层.但空间实体的几何描述并不是在静态图层本身表达,而是指向多个时空元素.在这个意义上说,它是一个代理.时空元素处在模型的最底层,由时空元素构成的时空复合图层是一个单一的图层,把多个时间点上的静态图层复合在一起,它承担着空间几何描述和空间拓扑描述的具体任务.

SCUDM 数据模型用以下 3 种模式表达对象域、空间域和时间域上的状态变化:

- 对象域模式, $Object-Domain-Schema = \langle o, \langle s_1, t \rangle, \dots, \langle s_n, t_n \rangle \rangle; n \geq 1, n \in N$.
- 空间域模式, $Space-Domain-Schema = \langle s, \langle o_1, t_1 \rangle, \dots, \langle o_n, t_n \rangle \rangle; n \geq 1, n \in N$.
- 时间域模式, $Time-Domain-Schema = \langle t, \langle o_1, s_1 \rangle, \dots, \langle o_n, s_n \rangle \rangle; n \geq 1, n \in N$.

在对象域模式层,采用时空对象表描述时空对象上的空间变化.表中每一个记录收集了一个时空对象在它的生命期(lifespan)内所有的历史版本,描述了这个时空对象空间范围的演变过程.时空对象表是建立在对象域模式上的,我们用 st_object_table (Object-Domain-Schema) 表示表格 st_object_table 在时空数据库中可以提供模式 Object-Domain-Schema.

<pre> st object table; Obj1:t1 A/t2 C/t3 E// Obj2:t1 B/t2 D/t3 E// </pre>			Object-Domain-Schema-Layer ^①

<pre> st entity table; A:t1 t2;Obj1:(c d)// B:t1 t2;Obj2:(e f)// C:t2 t3;Obj1:(c)// D:t2 t3;Obj2:(d e f)// E:t3 now;Obj1:(c f)// F:t3 now;Obj2:(d e)// </pre>	<pre> st event table; t1:// t2:A C B D// t3:C E D F// </pre>	Time-Domain-Schema-Layer ^②	

<pre> st polygon table; c:t1 A/t2 C/t3 E// d:t1 A/t2 D/t3 F// e:t1 B/t2 D/t3 F// f:t1 B/t2 D/t3 E// </pre>	<pre> st arc table; 3:t1 AP/t2 AI/t3 AI// 4:t1 AI/t2 AP/t3 AP// 5:t1 AP/t2 AP/t3 AP// 6:t1 AP/t2 AP/t3 AI// 7:t1 AI/t3 AI/t3 AP// </pre>	<pre> st node-table; n1:t1 NS/t2 NS/t3 NS// n2:t1 NS/t2 NS/t3 NS// </pre>	Space-Domain-Schema-Layer ^③

①对象域模式层,②时间域模式层,③空间域模式层.

Fig. 1 SCUDM data schema

图 1 SCUDM 数据模式

在时间域模式层,采用时空实体表描述时间上的空间-对象变化,或者说在不同的静态时间上的空间对象分布,即静态传统概念的静态图层.时空实体表在表达上没有完全遵照时间域模式,因为一个静态图层一般有数万个以上的时空实体,把它们编列到一个记录上完全没有必要.但时空实体表和时间域模式完成的任务是一样的,可用于描述时间上的空间-对象变化,故时空实体表对数据库可以提供时间域模式, *st_entity table* (Time-Domain-Schema).

在空间域模式层,采用时空面表用来描述空间几何形状上的变化.表中的每一个记录是关于一个时空元素的时间序列信息,描述随时间发展,这个时空元素所表达的空间区域为不同的时空实体所占用的变化.同时,通过时空实体这个代理,与时空对象发生联系.故时空面表与数据模型中的其他表合作,对数据库可以提供空间域模式, *st_polygon table* (Space-Domain-Schema).

注意到时空数据库中的变化是状态的变化,当 $n > 1$ 时,模式描述的是状态的变化;当 $n = 1$ 时,对象域(空间域,时间域)模式描述的是变化的特例,即单状态本身.通过对象域、空间域和时间域这 3 种模式来表达状态数据.虽然从数据库状态描述的角度来看,数据有冗余,但这种数据冗余不是由描述状态本身引起的,而是由对象域、空间域和时间域这 3 种模式表达状态的变化引起的.另一方面,这些状态数据之间的一致性是由数据库系统来维护的,不会造成用户的复杂性.同时,时空数据库中提供了不同的角度,以方便用户对时空的理解和建模.

3 拓扑表达

SCUDM 数据模型不仅是支持状态和变化显式地表达,而且支持拓扑关系的显式表达,包括空间拓扑和时间拓扑的表达.在 SCUDM 数据模型中,静态图层的几何描述由空间域模式层中的时空复合图层完成,同时,它的空间拓扑描述也是由时空元素构成的单一时空复合图层来完成,这里利用了我们在文献[6]中定义的复合图层拓扑和时变拓扑的概念.

在 SCUDM 模型的空间域模式层,点序列列表和线序列列表用来描述空间拓扑关系的变化.在线

(点)序列表中,每一个记录是关于一条线(点)在时间序列上,每一个时间点的拓扑变化分类,从而显式地描述了空间拓扑的时变信息.另外,空间域模式层中描述时空元素几何信息和复合图层拓扑信息的表格,限于篇幅,这里不再给出.

定义 7. 时空数据库的状态变化事件是从一个数据库状态子集到另一个状态的函数 $g^{OS}: State|_{t-} \rightarrow State|_{t+}$. 其中 g 的上标 OS 表示状态子集可在对象域和空间域变化; $t-$ 和 $t+$ 表示 t 时间点前和 t 时间点后. 它表达在时空数据库中, t 时间点发生事件前后,不同的时空对象在空间上的变化关系.

对应于状态变化事件 g^{OS} , 时空数据库的表达模式为

$$\text{事件模式, Event-Schema} ::= \langle t, \langle o_1, s_1 \rangle, \langle o_2, s_2 \rangle \rangle.$$

SCUDM 数据模型支持事件模式,注意到 SCUDM 中已在状态的其他模式中提供了对数据库状态的表达,因此,事件模式可以简化为以下两个事件模式中的一种:

- 空间事件模式, $Space-Event-Schema ::= \langle t, s_1, s_2 \rangle;$
- 对象事件模式, $Object-Event-Schema ::= \langle t, o_1, o_2 \rangle.$

在 SCUDM 数据模型的时间域,采用空间事件表显式地表达时间域发生的空间变化事件.表中的每一条记录是关于在一个时间点上的空间变化事件,这些事件按事件发生的时间排序构成数据库的事件序列.事件描述了空间变化前后,各时空实体的变化关系,如在时间点 t_2 时的事件中,时空实体 A 变化为 B ,时空实体 C 变化为 D .换句话说,事件也显式地给出了时空实体之间的时间拓扑关系,用 Allen 给出的 13 种时间拓扑关系^[7]描述,事件 $(t_2; A|C/B|D//)$ 表达了 B meet A , A end C , D meet C , B start D 这 4 个时间拓扑关系.所以,SCUDM 数据模型通过对 Event Schema 的支持, $st_event_table(Event-Schema)$ 提供了对时间拓扑的显式表达.

也可以考虑对象事件模式,把事件对时空实体的变化关系描述转换到各时空对象上的变化关系上的描述,在时空实体表中, $(A; t_1 t_2; Obj1; (c d)//), (B; t_1 t_2; Obj2; (e f)//), (C; t_2 t_3; Obj1; (c)//), (D; t_2 t_3; Obj2; (d e f)//)$. 这样,事件 $(t_2; A|C/B|D//)$ 即表示 $(t_2; Obj1; (c d) | (c) / Obj2; (e f) | (d e f) //)$, 描述了在此空间变化事件中,空间范围 d 从时空对象 $obj1$ 中分离出来,并合并到时空实体 $obj2$ 中,如图 2 所示.由于利用了对象数据库中的聚集(aggregation)和联合(association)概念,事件的语义可以表达得很丰富.

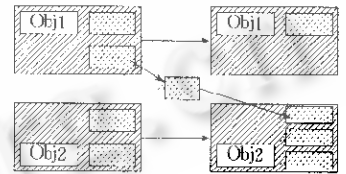


Fig. 2 Object change at t_2
图2 t_2 时的对象变化

4 数据模型的形式化

$$\langle \text{Spatio-Temporal-Database Schema} \rangle ::= \langle \text{Object-Domain-Schema} \rangle | \langle \text{Space-Domain-Schema} \rangle | \langle \text{Time-Domain-Schema} \rangle | \langle \text{Event-Schema} \rangle | \{ \langle \text{class} \rangle \};$$

$$\langle \text{Object-Domain-Schema} \rangle ::= \{ \langle o, \langle s_1, t_1 \rangle, \dots, \langle s_n, t_n \rangle \rangle \};$$

$$\langle \text{Space-Domain-Schema} \rangle ::= \{ \langle s, \langle o_1, t_1 \rangle, \dots, \langle o_n, t_n \rangle \rangle \};$$

$$\langle \text{Time-Domain-Schema} \rangle ::= \{ \langle t, \langle o_1, s_1 \rangle, \dots, \langle o_n, s_n \rangle \rangle \};$$

$$\langle \text{Event-Schema} \rangle ::= \{ \langle t, \langle o_1, s_1 \rangle, \langle o_2, s_2 \rangle \rangle \};$$

$\langle o \rangle ::=$ a object value in object domain of spatio-temporal database;

$\langle s \rangle ::=$ a space value in space domain of spatio-temporal database;

$\langle t \rangle ::=$ a time value in time domain of spatio-temporal database;

$\langle n \rangle ::= a$ integer value in integer domain, $n > 0$.

SCUDM 数据模型支持的查询表达形式为 $\{\langle x_1, x_2, \dots, x_n \rangle | P(x_1, x_2, \dots, x_n)\}$, 其中 x_1, x_2, \dots, x_n 可以是对象域、空间域或时间域中的变量, P 代表了由原子构造的公式.

一个原子可以是以下形式中的任意一种(表达中的 $r(R)$ 表示 r 是模式 R 在数据库中的表达):

- $\langle o, \langle s_1, t_1 \rangle, \dots, \langle s_n, t_n \rangle \rangle \in r(\text{Object-Domain-Schema});$
- $\langle s, \langle o_1, t_1 \rangle, \dots, \langle o_n, t_n \rangle \rangle \in r(\text{Space-Domain-Schema});$
- $\langle t, \langle o_1, s_1 \rangle, \dots, \langle o_n, s_n \rangle \rangle \in r(\text{Time-Domain-Schema});$
- $\langle t, \langle o_1, s_1 \rangle, \langle o_2, s_2 \rangle \rangle \in r(\text{Event-Schema});$

• $x \Theta y$, 其中 x 和 y 为同一个域(对象域、空间域、时间域)上的变量, Θ 是相关域的比较操作, 对象域时为(=, \neq)等, 时间域时为(before, equal, meet, after)等时间拓扑操作, 空间域时为(=, \neq , contain, overlap, meet)等空间拓扑操作;

• $x \Theta c$, 其中 x 为域(对象域、空间域、时间域)的变量, Θ 是相关域的比较操作, c 为与 x 变量是同一个域的常量.

我们运用以下规则构造公式:

- 原子是公式.
- 如果 p_1 是公式, 那么 $\neg p_1$ 和 (p_1) 也是公式.
- 如果 p_1 和 p_2 是公式, 那么 $p_1 \wedge p_2, p_1 \vee p_2, p_1 \Rightarrow p_2$ 也是公式.
- 如果 $p_1(x)$ 是关于 x 的公式, x 是域变量, 那么 $\exists x(p_1(x))$ 和 $\forall x(p_1(x))$ 也是公式.

注意到如果 $(\text{Spatial-Temporal-Database-Schema}) ::= \{\langle class \rangle\}$ 就是一般的对象数据库了. 一般的对象数据库、一般对象数据库中的扩展操作(如 average, sum 等)以及时空数据库中的扩展操作(如 area, distance 等)在此不作讨论.

5 实例

在地籍管理应用中, 需要对土地所有人的土地、宗地和宗地交易的概念进行建模. 在现实世界中, 所有者对宗地进行交易. 在交易过程中, 所有人购买、出售或交换宗地, 所有人拥有的宗地发生变化. 交易完成后, 土管局编造新的宗地记录, 但旧的宗地记录不删除, 归为历史的宗地记录.

采用 SCUDM 模型对地籍管理进行建模, 土地所有者拥有的土地定义为时空数据库中的时空对象, 时空对象的同一为所有者所标定(如 Obj1: 张山; Obj2: 李明). 宗地定义为时空实体, 宗地的交易定义为空间变化事件. 时空元素是模型的底层支持, 承担空间几何和空间拓扑关系表达的任务.

例 1: 找出李明在 t_2 时间拥有的宗地:

$$\{\langle s \rangle | \langle o, \langle s, t \rangle \rangle \in r(\text{Object-Domain-Schema}) \wedge o = \text{“李明”} \wedge t = t_2\}.$$

对于同样的要求, 也可以给出另外两种不同的查询方法:

$$\{\langle s \rangle | \langle t, \langle o, s \rangle \rangle \in r(\text{Time-Domain-Schema}) \wedge o = \text{“李明”} \wedge t = t_2\};$$

$$\{\langle s \rangle | \langle s, \langle o, t \rangle \rangle \in r(\text{Space-Domain-Schema}) \wedge o = \text{“李明”} \wedge t = t_2\}.$$

这是因为 SCUDM 提供对象域、时间域和空间域这 3 种模式理解状态. 采用这 3 种状态模式对同一个状态查询时有不同的访问路径, 系统实现的访问效率也是不同的, 因此, 在查询处理时可以利用表达式等价性原理进行查询优化.

例 2: 找出与宗地 C 在历史上空间相邻的宗地:

$$\{\langle s_1 \rangle | \exists o_1, s_1, o_2, t_2 (\langle s_1, o_1, t_1 \rangle \in r(\text{Space-Domain-Schema}) \wedge$$

$$\langle s_2, o_2, t_2 \rangle \in r(\text{Space-Domain-Schema}) \wedge s_2 = C \wedge s_1 \text{ meet } s_2 \}.$$

例 3:找出在空间区域 d 范围内相关的所有的宗地以及宗地所有者的变化过程:

$$\{ \langle s, o_1, t_1, \dots, o_n, t_n \rangle \mid \langle s, o_1, t_1, \dots, o_n, t_n \rangle \in r(\text{Space-Domain-Schema}) \wedge s \text{ overlap } d \}.$$

6 总 结

由于时空数据建模既要考虑一般的数据建模的问题,又涉及到空间、时间这种普遍适用而又极难说明的建模难题,由于研究者对空间和时间的理解不同,侧重点也会有所不同,使得已有的时空数据建模的思路和方法表现不一,层出不穷.本文希望引入和定义基本的时空语义概念,通过形式化的分析,可以形成一个大家讨论的基础.其目的在于:① 更强的建模能力.在形式化的基础上,模型应尽可能统一更多的表达能力;② 更方便、快捷的建模过程.模型应更符合人类对复杂时空现象的抽象认知方式,使用户对数据库中的时空数据易于理解,并方便采用此模型对时空现象的建模.

本文是我们建立时空数据库工作的一部分,为突出讨论的重点,在此没有涉及对象属性在时间上的变化.一个基于 SCUDM 模型的时空数据库原型工作正在对象-关系数据库 PostgreSQL 上实现.

References:

- [1] Al-Taba, K. K., Snodgrass, R. T., Soo, M. D. Bibliography on spatiotemporal databases. ACM SIGMOD Record, 1993, 22(1):59~67.
- [2] Langran, G. Time in Geographic Information Systems. London: Taylor & Francis, 1992. 45~54.
- [3] Worboys, M. F. A unified model for spatial and temporal information. Computer Journal, 1994, 37(1):26~34.
- [4] Peuquet, D. J., Duan, N. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. International Journal of Geographical Information Systems, 1995, 9(1):7~24.
- [5] Rapper, J., Livingstone, D. Development of a geomorphological spatial model using object-oriented design. International Journal of Geographical Information Systems, 1995, 9(4):359~384.
- [6] Tan, Shi-yu, Zheng, Kou-gen, Pan, Yun-he. Time-Varying topology. Computer Research and Development, 2000, 37(7):769~775, (in Chinese).
- [7] Allen, J. F. Maintaining knowledge about temporal intervals. Communications of the ACM, 1983, 26(11):832~843.

附中文参考文献:

- [6] 谭石禹,郑扣根,潘云鹤.时变拓扑.计算机研究与发展.2000,37(7):769~775.

A Unified Spatio-Temporal Data Model Based on State and Change*

ZHENG Kou gen, TAN Shi yu, PAN Yun he

(Institute of Artificial Intelligence, Zhejiang University, Hangzhou 310027, China)

E-mail: zkg@cs.zju.edu.cn

http://www.zju.edu.cn

Abstract: In this paper, a unified spatio-temporal data model (SCUDM) based on state and change is proposed. A state in spatio-temporal database is represented by a ternary tuple about object, space and time. Thus, a change in spatio-temporal database is represented by a change over a state, which can be projected to object domain, space domain and time domain separately in order to explicitly represent change over object, space and time. An event in spatio-temporal database is a relation of changes about spatio-temporal entities, implying the temporal topology of spatio-temporal entities. SCUDM model can support both representation of state and change and representation of spatial-temporal topology.

Key words: spatio-temporal data model; spatio-temporal database; spatio-temporal semantic; state change; spatio-temporal topology

* Received September 13, 1999; accepted May 26, 2000

Supported by the National High Technology Development Program of China under Grant No. 863 306 04 03 3