时态数据库的发展及其未来

史鑫宇

（北京师范大学研究生院珠海分院，北京 100875）

**摘要**：本文主要从时态数据库的发展历史，现状成果以及未来趋势三方面进行论述，简单介绍了三种经典时态数据库模型和模型的标准化及应用，并详细地阐述了时态数据库发展过程中所面临的问题以及未来研究方向。

**关键词**：时态数据库；时态数据模型；标准化；未来发展

**1 引言**

传统的数据库系统如层次、网状和关系型数据库，并未对时态数据加以重视，仅仅将其作为一种普通的属性值进行存储和管理。传统数据库中的数据，只能反映事物当前状态，不能联系它的过去和未来。而现实中，如自然灾害，如人事、财务、金融方面，我们需要这些历史数据去反应事物发展的过程，揭示事物发展的规律。数据库系统中的事务处理的时刻、时间区间、多用户系统中对封锁解锁的申请，排队、协调资源竞争的时标等，这些数据有助于提高数据库系统的功能和效率。因此用于管理时态信息的数据库逐渐发展起来。

**2 时态数据库的发展**

时态信息的研究始于20世纪70年代。1970年，G.Wieder-hold和J.F.Friies研制的医疗信息处理系统在处理时态信息方面做了最早的尝试，但存储技术的落后制约了时态信息处理的发展。八十年代，存储技术的发展推动了时态信息管理系统的再次发展。到九十年代，时态数据库在理论研究和原型研究上已取得丰硕的成果。

**3 时态数据库模型**

**3.1 时间模型**

在对时态信息的研究和推理过程中，主要集中在时间模型和时态属性上。

主要的时间模型有：时刻、时段、时间跨度和时间集合。时态属性主要涉及到时间模型的连续和离散两种方式。连续时间与实数集合同构；离散时间与整数集合同构。

1983年，Allen提出了以区间代数理论表示的时段代数系统。其模型认为，每个时间段包含两个端点，通过比较两个时间段的端点之间的关系，区分互不相交且联合完备的基本关系。如图3.1所示。

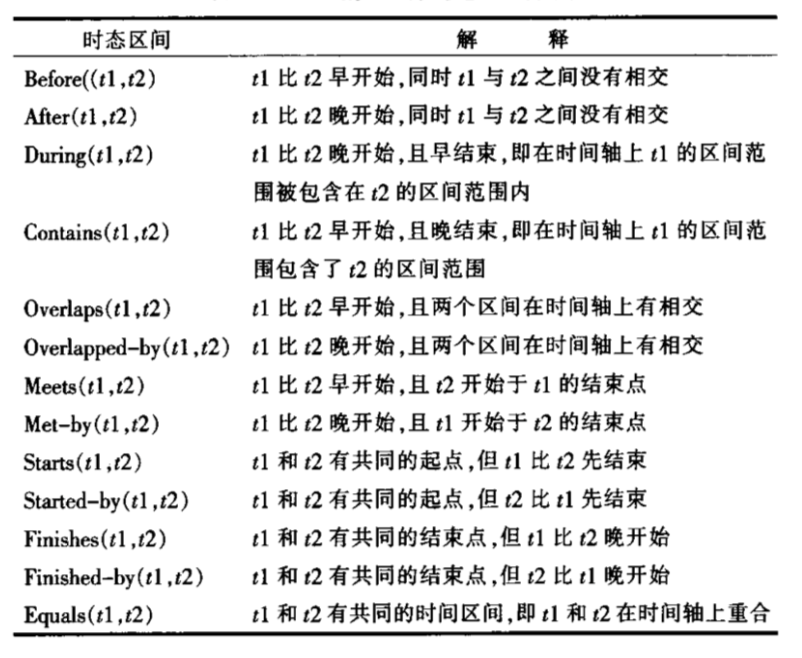


图3.1 13种时态区间关系

**3.2 典型时态数据库模型**

从七十年代到九十年代时态数据库发展的20多年里，各地学者们设计出了上百种时态数据模型和查询语言。经过演化，逐渐归并为13种模型。如图3.2所示。这13种时态数据模型从不同需求和观点出发，各自独立的建立了其概念、术语和数学关系，并形成了各自的理论体系。

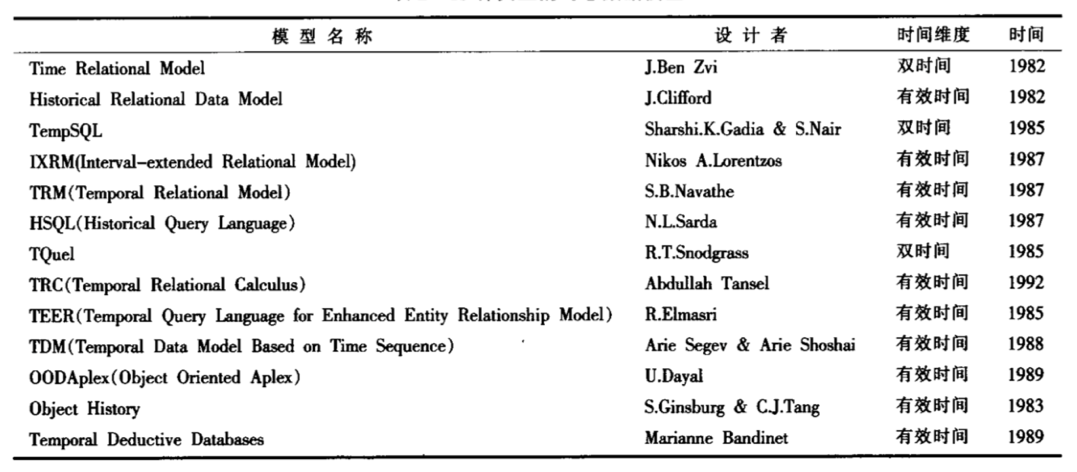


图3.2 13种典型的时态数据模型

**3.2.1 历史关系数据库模型**

历史关系数据库模型HRDM是由James Clifford等学者于1982年提出的，是时态数据库的先驱模型之一。如图3.3所示，为了能有效的解决保存快照的时间间隔的选取问题，引入了“任职年限”属性。属性“任职期限”的值是区间或区间集合，称为元组的有效时间或生命周期。在时态数据库管理系统中，该值由系统管理而不是作为普通字段来进行管理。当一个元组录入关系，则该元组的生命周期开始;当对元组作一个时态删除，则指定一个时刻或取当前时刻作为该元组生命周期的终止点，并不作物理删除。一般而言，时态数据库中的历史不能删除、不能修改，只能追加和查询。

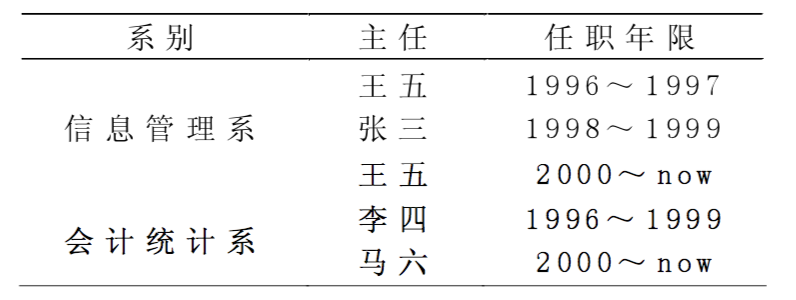


图3.3 科系管理系统

时态数据库中属性被扩展为时态属性，一个时态属性可以是一个普通属性和一个时间区间组成的二元组。生命周期是系统时间域的子集，生命周期经过集合运算后，其值仍是生命周期。

**3.2.2 TempSQL模型**

Shashi K.Gadia于1985年建立了同时性关系模型，该模型因其实现了查询语言 TempSQL 而成名。TempSQL是一个类似SQL的时态查询语言，它兼容了SQL的主要功能，能查询被管理对象的历史、数据库本身增删改的历史以及用户和数据库本身出错的历史。TempSQL保持了时态数据库和静态数据库的无缝连接。快照数据库可以看作是TempSQL中时态关系的生命周期缩小为一个时间量子[now，now]时的特例。由于TempSQL扩展时态功能的成功，人们常以TempSQL指代该TDB模型。TempSQL中非关键字属性值可以随时间变化，但关键字不随时间变化。为了模型简单实用，TempSQL 规定一个元组中各个属性的生命周期一致，称为满足同时性条件。

**3.2.3 双时态数据模型**

TQuel是Richard Snodgrass在1985年左右开发的时态查询语言。TQuel建立在双时态数据库模型上，该模型中，对象的生命周期称为有效时间，数据库事务执行时间称为事务时间，并认为这两个时间轴是正交、互为独立的。该模型建立了完备、复杂的理论，为TQuel语言制订了详细的句法规范，在理论上较深入研究了时态闭包、完全性、有效时间的不确定性、时态查询优化、时态约束，研究了TQuel和元组演算的语义关系，并得出一些较深刻的定理。TQuel是Quel语言的时态扩展，Quel语言作为元组语言的具体实现，在关系数据库理论和系统研究上占有重要地位。

**3.3 时态数据模型标准化及应用**

1993年，国际时态数据库理论架构研究小组在基于扩展SQL的基础上，开始了第2代时态查询语言的研究工作——TSQL2。1994年，TSQL2语言标准发表。TSQL2模型综合了其他模型的优点，是迄今为止最容易理解的、文档资料完备的、并被广泛接受的时态查询语言。然而由于ISO委员会内部没有就SQL中该如何处理时态关系取得一致意见，2001年底有关计划被迫取消，TSQL2工作草案处在废弃的边缘。

根据时态系统处理时态属性的方式和能力，可以将时态系统的应用方式分为3类：混合式、嵌入式和完全的时态应用系统。如图3.4所示。

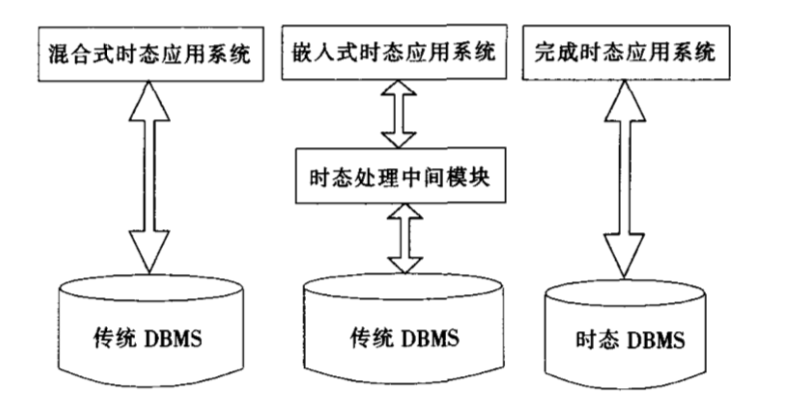


图3.4 时态应用系统分类

混合式时态应用系统。是指将时态处理技术与应用信息处理技术结合，实现应用系统中对时态属性的支持，其底层数据库采用传统的RDBMS。是当前使用最广泛的方法。这种方法容易实现，但时态查询功能受到限制，重用性差。

嵌入式时态应用系统，是指通过采用时态处理中间模块，在传统RDBMS与时态应用系统之间进行通讯。来解释时态查询、时态推理等功能。这是当前普遍接受的时态应用系统设计方法，既满足了时态信息处理的需要，又支持传统RDBMS。其缺点是不能直接应用已有的时态存储结构、索引方法以及连接、接合等时态代数运算。

完全的时态应用系统，是基于时态数据库系统开发的应用，对时态信息的处理完全由底层数据库来支持。由于目前还没有商业化的时态数据库系统，因此这种模式还无法实现。

从短期效果看，采用混合式和嵌入式方法更容易实现。然而从长远考虑，随着商业数据库对时态处理功能的提高，完全式的应用更具有推广价值。

**4 时态数据库存在的问题及未来方向**

**4.1 存在的问题**

虽然在理论上已经证明时态数据库拥有传统数据库所无法企及的优点，但迄今为止，在实践中还没有真正实现一个既具有传统数据库的功能，又能实现时态处理的完整的时态数据库。也没有出现过能够进行时态处理的商业数据库管理系统。这就使得时态数据库的理论得不到实际应用的检验，阻碍了时态数据库的推广和进一步的发展。其存在的突出问题集中表现为：理论研究多、实际应用少；数据模型多、软件工具少；以模型划分的门派多，支持时态处理的厂商少。

(1)时态数据库研究初期，几乎每篇博士论文都提出自己的模型、理论体系。这在研究初期是可以的，但最近一些年仍有部分新出现的、并无明显特色和优点的模型。这对时态数据库的研究是不利的。而且这些模型大多处于理论研究阶段，没有给出具体的应用方法。

(2)现有模型特点不突出，优势不明显，标准有待于进一步完善。大部分时态数据查询语言是通过扩展SQL实现，查询功能有限，效率比较低，而且提供的查询工具少，基本处于试验阶段。

(3)模型分类多，许多模型处于演化和改进阶段。数据库厂商难以决断究竞选择哪种模型更适合实际应用。目前主流数据库产品基本处于快照数据库阶段，对时态的支持有了改进，然而还不能真正支持时态处理，专门时态数据库产品更为少见。

(4)大多数研究局限在数据库的时态属性，而忽视了其它信息的时态属性，例如知识库的时态特征和应用。时态逻辑和推理方面的研究与时态数据库和时态信息处理研究脱离。

(5)由于没有成熟的时态模型和产品，大部分与时态相关的应用，只是借鉴当前一些时态数据模型，其时态部分的解释是由应用程序而非数据库本身来进行。

**4.2 未来方向**

时态数据库领域研究已经基本完成了理论奠基工作。在将来的工作中，应侧重以下几个方向:

(1)集中精力研究现有模型的整合与归并，为时态数据库和时态查询语言的标准化提供支持。

(2)传统数据库的应用已经非常普及，因此时态数据库的研究要以其为基础，而不应与其脱离，在其基础上逐步加入时态信息处理机制。事实上，在短期内，是不能取代现有数据库系统而开发出全新的时态数据库的。利用现有数据库系统，并与第三方开发的时态信息处理中间件结合是一个比较好的解决方案。

(3)充分理解时态数据库的时问语义，利用先进的软件思想和技术，设计合理的、能够完全支持时态数据库设计的模型方法以及设计工具。

(4)时态数据库旨在管理数据的历史，随着时间的流逝，新的数据源源不断的进入数据库，为了保证其检索效率，必须研究高效的数据存储和索引机制。

(5)面向对象思想更符合人类的逻辑思维语义，因此在数据库领域得到广泛应用。采用面向对象方法实现支持时态信息处理的时态面向对象数据库将具有很好的发展前景。

(6)大力研究时态数据库技术在数据挖掘系统中的应用。时态数据库中的有效时间反映了事物发生发展的过程，有助于揭示事物发展的本质规律：事务时间反映了系统中元事件的时态信息，有助于提高数据库系统的可靠性和效率。

(7)研究时态数据模型在数据仓库中的应用。数据仓库的奠基者Inmon认为，时态变化是数据仓库的4个重要特征之一。因此时态信息的研究对数据仓库的研究至关重要。

(8)将时态研究范围拓展到空间维，设计同时支持时间和空间的模型和数据库系统。时空模型和时空数据库系统的研究一直是相关领域的热点。时间与空间的结合带来许多新的挑战。如，时态数据库与全球定位系统(GPS)、地理信息系统(GIS)相结合产生的位置服务系统LBS(Location Based Service)就是这方面的一个新的研究领域。

**5 总结**

时态数据库发展较早，且有强大的理论体系和丰富的模型。时态数据库有着关系数据库无法比拟的优点，比如在时间时态上的处理更为优秀。

与此同时，时态数据库在20世纪末期发展鼎盛，但步入21世纪发展却停滞不前。时态数据库的发展理论模型较多，但无法取得一致的意见标准，所以无法形成成熟模型和产品。时态数据的查询功能有限，效率较低，且对时态数据库的研究仅局限于单一方面。

目前时态数据库的研究方兴未艾。无论在理论，原型研究和系统实现方面都有待深化。重点在于整合管理现有模型，形成统一标准，并以传统数据库为基础。同时，将时态数据的研究范围扩大，可以与数据挖掘、数据仓库等大数据时代的技术相结合；也可以结合其他数据库，如空间数据库，时间和空间的结合也将带来更大的机遇和挑战。

参考文献

[1]唐常杰, 吴子华. 时态数据库研究的过去,现状和发展方向[J]. 计算机应用, 1993(1):1-3.

[2]高艳云. 时态数据库及­其发展趋势刍议[J]. 民营科技, 2009(11):20-20.­

[3]黄楠. 时态数据库管理系统的研究与实现技术[J]. 电­­脑知识与技术:学术交流, 2012(11):7402-7403.­

[4]姜晓轶, 蒋雪中, 周云轩. 时态数据库研究进­­展[J]. 计算机工程与应用, 2005, 41(24):27-30.

[5]黄楠, 刘爱琴. 时态数据库技术[J]. 计算机技术与发展, 2002, 12(1):24-26.