

数据库

多媒体数据库

索引

查询

⑨

计算机科学 1998 Vol. 25 No. 5

17-23

97 数据库研究新进展

The Frontier of Database Research 1997

田增平 王宇君 施伯乐

(复旦大学计算机系 上海 200433)

TP311.13

摘要 With the development of computer networks and multimedia technologies, new applications are emerging in large numbers. Database technologies are facing more challenges and opportunities. In this paper, we investigate the proceedings of famous international database conferences such as PODS and SIGMOD. Combined with our research results, we summarize new developments of database research from several aspects such as multimedia databases, WWW and databases etc.

经过三十余年的开发与研究,数据库已成为一种相对成熟的计算机软件技术,其产品已大量出现且被广泛应用,研究工作也取得了丰硕的成果。然而,数据库技术的研究并没有停滞,特别是近几年来,多媒体技术、网络技术和新型应用的迅猛发展,对数据库技术提出了新挑战,为了迎接挑战,更大地发挥数据库技术对计算机应用的支持作用,数据库领域的科研人员针对新技术和新应用所提出的问题,进行了大量的研究工作。我们从如下几方面介绍近年来的主要研究工作:多媒体数据库、WWW与数据库、约束数据库、半结构化数据处理、演绎数据库(知识库)、主动数据库,以及视图与更新等。

1. 多媒体数据库技术

“多媒体”一词首次出现于1983年在意大利佛罗伦萨举行的第九届VLDB会议上,它为多媒体数据库的研究与开发作了思想概念上的准备。此后,在1984年的VLDB会议和1985年的SIGMOD会议上,都举办了有关多媒体数据库的研讨班,对多媒体数据库日益频繁的讨论和开发活动,预示着它已作为数据库领域的一项新技术逐渐受到人们的注意。

八十年代后期,出现了大量多媒体数据库原型系统。其中有法国GRENoble大学开发的历史多媒体数据库[1986]、加拿大多伦多大学开发的MINOS系统[1986]、意大利多所大学联合开发的MULTOS系统[1988]、美国MCC开发的多媒体信息管理器MIM[1986]、德国GMD-IPSI开发的VODAK系统[1988]和新加坡信息系统研究所开发的

多媒体数据库管理系统[1989]等。这些系统为多媒体数据库系统的研究与开发提供了宝贵的经验。

自九十年代以来,媒体的数字化技术得到了极大发展,为多媒体信息的计算机处理和应用提供了可能。硬件技术的发展奠定了设备基础,各种独立多媒体数据管理技术的发展提供了软件技术保障。大量多媒体应用的出现将研究推向新高潮。

多媒体数据的新特点给多媒体数据库研究提出许多问题,可从以下四个方面考察。

1) 内容查询技术

传统数据库查询基于对属性值谓词布尔组合的准确计值。多媒体数据库中,基于属性值的确定查询已远不能满足要求,用户希望根据图形、动态视频影像和文本等非结构化对象所富含的语义信息进行查询,即基于内容的查询。这种查询与传统数据库查询不同,其本质是近似和模糊的。内容查询涉及的问题有:

(1) 查询形式。即用户采用什么形式将查询要求提交给系统。最基本的两种方式是图示法(板样法)和查询语言。

(2) 特征提取技术。基于内容的查询也可认为是基于多媒体信息某些特征的查询。特征是由基本数据提取的、可测度的、代表一定数据语义信息的概念。多媒体数据库的特征可以静态地提取,也可以在查询处理中动态提取。

(3) 索引及查询优化。多媒体数据的特征信息只有按一定的索引结构组织,才能较好地支持查询优化,提高系统效率。

2) 视频信息建模技术

在视频信息系统中,视频数据类型与常用的数据类型具有同样的地位,系统提供一组工具,用户用以能够对视频数据进行存储、管理和检索等传统数据库操作。在视频信息系统的研究与开发中,视频信息建模技术是一个关键,对它的研究需注意下列问题:

(1)概念建模。通常,人们对视频数据的认识是:一段视频信息(电影、电视节目等)是一个帧的序列。以电视节目为基本建模单位,用户只能按照名称来查找相应的节目,这正是早期连续媒体服务器的功能;而仅以帧为建模单位,则概念太小,用户很难想出要在帧上查找什么,因而必须建立合适的概念帮助视频数据建模。

(2)内容建模。视频信息系统的目的在于支持用户对视频数据的存储、组织和检索,不但应当支持传统的数据库查询(基于视频信息的标准属性),而且要支持基于视频信息内容的查询。因此,多媒体信息的内容建模同样非常重要。Y. Francis 等人提出一种应用无关的视频语义有向图(VSDG)模型组织视频数据内容,客观描述了视频数据的空间和时间特征,与面向对象的概念模型结合,支持用户基于内容的查询操作。

(3)查询语言。查询语言是系统与用户之间的接口,是用户操纵视频数据的工具。视频数据模型的概念及能力反映在查询语言中,同时,查询语言的功能也是对视频数据模型能力的体现。目前已有许多视频查询语言的研究工作。

3) 多媒体信息的表现技术

表现是多媒体信息的重要形式,具有时间和空间特征,多媒体数据库不仅要存储和管理多媒体数据,而且要管理数据表现的时、空特性。涉及的表现问题主要有以下几方面:

(1)表现形式描述。多媒体信息的表现形式与表现内容具有同样重要的作用,著作者通过两者来传递自己要表达的思想。因此,多媒体数据库系统不但要支持用户对多媒体对象内容(信息数据)的定义,还要支持对表现形式的定义。表现形式描述工具的研究中涉及到表达能力、分布特性、动态交互和表现约束的描述等问题。

(2)表现调度。表现调度定义表现活动中各个多媒体元素开始及结束表现的时刻表。

(3)表现控制。表现进行过程中,需要某些动态管理机制保证表现质量。表现形式描述和表现调度

生成中有一个基本假设,即,表现数据在需要的时候便可获得。但是,实际情况往往并非如此,表现系统无法保证网络传输和缓冲区等资源总能够满足要求。这就需要系统设置相应的动态管理机制,对数据短缺情况进行适当处理,保证系统服务质量。

4) 连续媒体的存储与管理技术

由于连续媒体数据量大,表现过程中具有实时性,因而其存储、传输与管理比较复杂,在研究与开发连续媒体服务系统中会涉及到许多问题:

(1)数据放置技术。连续媒体存储机制不但要保证数据的可靠存储,而且要保证它的实时存取。数据在存储设备上的放置策略是影响存储速度的重要因素。常用的方法有连续放置、交错放置和条块放置(striping)等。

(2)磁盘调度策略。磁盘调度策略决定了系统进行磁盘操作的原则,包括可靠性、有效性及对各用户公平等。

(3)请求接受控制(Admission Control)。连续媒体服务系统给客户提供服务时要消耗一定的资源,如网络带宽、存储设备存取带宽和缓冲区等。系统中这些资源是有限的,这就意味着服务器能同时服务的客户也是有限的。因此,系统中应当有专门的机制来决定哪些服务请求可以接受,哪些服务请求延迟接受或不接受。

(4)传输控制。连续媒体数据量大、实时性强的特点决定了对其传输过程必须施加严格控制。传输控制涉及同步控制、流量控制和缓冲区管理等问题。

(5)容错技术。连续媒体数据量巨大,常用多个存储设备存储,增加了存储设备发生故障的机率。例如,若每个磁盘的平均故障时间(MTTF)为300,000小时,使用200个磁盘时,MTTF为1500小时。因此,研究连续媒体存储系统的容错技术具有重要的意义。

2. WWW 与数据库

在过去的二十年内,计算机和通信工具每年以迅猛速度增长。与此同时,计算机通讯网络也以惊人的速度发展,将世界的各个角落连接成为一个巨大的信息网。WWW网是最为流行的全球性信息资源网,它连通各个国家拥有数以万计的网络服务器、数以千万计的网络用户和无数的网络信息资源,形成了一个覆盖全球的、巨大的资源网络,为信息的生产、传输和利用开辟一个新的空间。WWW网络的日益流行和发展,对计算机技术的各个领域都产生

了极大影响。不同领域的技术人员都从各自的角度出发,展开对 WWW 网络上相关技术的研究与开发。数据库领域的工作主要有:

1) 信息的组织与集成技术

Web 网络拥有众多的用户,每个用户都可以通过它发布信息或获取信息。因此,可以形象地称这种网络为一个巨大的“信息自由市场”。要使它能够有条不紊地运作,每个用户都能顺畅地发表自己的消息,每个信息消费者都能获其所需,会遇到许多问题。

(1)大量多媒体信息的集成。信息量巨大是 Web 网络信息的重要特点。大量的数据对网络传输的带宽、服务质量及存储和缓冲技术提出新要求,此外,媒体种类繁多,数据结构复杂又是 Web 网络信息的另一特点。不同媒体、不同质量、具有不同的访问和表现技术,同一种媒体,在不同系统环境中其表现会具有不同的效果,因而网络信息的统一组织和处理技术的研究具有重要意义。

(2)语义异构。Web 网络延伸至全球各个角落。信息提供者按照自己的习惯和知识来组织信息。不同的用户对相同的信息的理解会产生偏差。从用户工具方面来说,应当提供方便的用户工具及系统接口,使不同用户能够方便地互访信息,从语义理解方面来说,应当建立适当的机制,尽量减少语义异构带来的影响。

(3)系统扩充。Web 网络是一个巨大的动态系统,它的网点数目、用户数量及信息量等,每时每刻都在不断地增加。系统的平滑扩充所需要注意的问题是:其一系统的开放性,是否能保证新加入的网点能通过良好定义的接口与已存在的网点进行系统交换;其二信息的质量问题,Web 上的大量信息可能会存在不正确、不一致和不完全等问题。

2) 信息访问技术

Web 网络蕴含着丰富的信息资源,然而要使广大用户能高效、准确地获取各自所需的信息,就必须提供适当的信息查询工具和查询处理机制。

(1)信息查询工具。目前普遍使用的是 Mosaic、Gopher 和 Netscape 等浏览器。它们的界面简单、使用方便。然而作为信息获取工具,一是查询能力不够强,用户根据自己的知识和判断,通过 URL 追踪,寻找所需信息,很容易“迷失”在信息海洋中;二是仅能提供交互式访问,构造复杂应用比较困难。因此 Web 查询语言成为数据库界研究的一个热点。

(2)信息查询处理。信息传输速度和查询处理效

率是影响用户查询响应速度的重要因素。随着网络技术的发展及通信技术的提高,前者的影响逐渐降低,后者则成为提高查询响应速度的关键。查询处理的效率与查询处理机制的性能及信息的索引技术关系密切。

3) 非技术问题

Web 网络开辟了交流信息的新空间。人类在通常信息交流中遇到的问题在此同样存在,如信息的安全保密、版权、收费、个人隐私、色情等。解决这些问题可以从两方面入手:一是开发有关的技术,缓解或处理这些问题,如信息加密、防拷贝、信息检查等;二是制订有关法律、法规,约束人们在 Web 上从事信息交流的行为,这需要国际社会的共同努力。

3. 约束数据库技术

约束查询语言和约束数据库的概念是 Kanel-lakis 等在 1990 年的 PODS 会议上首先提出来的,近年来国际上一些著名的数据库会议关于这方面的报告在不断增多。围绕约束数据库的工作极大地推进了数据库理论的研究。约束数据库可以看成是对关系数据模型的扩充,其核心思想是,将关系数据库中的每个元组看成是一个等词约束的合取式。这种等词约束可推广到其它约束,这样约束元组就是定义在一个变量集上的约束合取式。根据约束元组,可以类似关系数据库定义约束关系和约束数据库。根据定义约束元组中约束种类及论域的不同,可定义不同的约束数据库。目前正在研究的有:有理数或实数域上的密序约束、线性序约束和实多项式序约束等。

由于每个约束元组可以描述一个可能是无限的点集,因此约束数据库模型可以用约束的形式来统一表示时空信息或不完整的信息。与现在流行的各种数据库相比,表示相应信息时更灵活、更一致。约束数据库的研究成果可从如下几个方面来讨论。

1) 查询语言的表达能力

查询语言表达能力的研究集中在研究某个或某类查询语言的语义,它所能表达的查询及查询的合理性。约束查询语言从本质上来说,都是某种查询语言扩充了某类约束而形成的,自然会问,约束查询语言比原来的基本查询语言的表达能力强多少?是否能表达所有“合理的”查询?等。这方面的研究成果有:

(1)基本理论成果。它们适合于各种约束数据库,主要有:约束查询语言的形式定义、约束数据库

的逻辑基础研究和用于研究约束数据库的基本工具和理论等。

(2) 几种约束语言的成果。S. Grumbach 和 J. Su 研究了密序约束数据库查询语言的能力, 给出了带密序约束的一阶语言 FO, 带线性约束的一阶语言 FO⁺ 和带密序约束的、带否定的、基于膨胀语义的 Datalog 语言的表达能力的研究结果。J. Pardaens 等人证明了线性序约束并没有增加纯关系演算的表达能力。M. Benedikt 和 G. Dong 等人证明了实多项式约束下的关系演算中, 奇偶校验和传递闭包等查询是不可表达的, 并进一步说明当考虑的输入和输出为一般关系时, 在关系演算中增加实多项式约束并没有增加其表达能力等。

2) 查询语言的扩充

基本的约束查询语言可从以下几方面进行扩充。

(1) 复杂对象。最初的约束数据库是一种基于“平”数据结构的数据模型。将其与复杂对象理论相结合, 就产生了约束复杂对象模型。S. Grumbach 和 J. Su 研究了含有元组和集合构造子的约束复杂对象模型, 提出一种基于逻辑的约束查询语言 C-CALC [1996], 并给出了一些有关其表达能力的结果。A. Brodsky 等人将“平”的约束查询语言扩充到了面向对象数据库查询语言, 提出一种约束对象查询语言 Lyric [1996]。

(2) 聚集函数。关系数据库的实践表明, 查询语言中增加聚集函数是非常有用的, 如 Sum, Average 等。在约束数据库中增添聚集函数会带来问题: 一是某些传统的聚集函数在约束数据库中可能没有意义, 如 Count 函数; 其二是封闭性问题, 即某种约束查询语言中使用聚集函数后的结果不能由此种约束查询语言表达。G. M. Kuper 等人给出解决这些问题的三种途径: 限制查询语言的使用、限制聚集函数类的使用、限制聚集函数的使用方式。

(3) 代数查询语言。提高查询描述的“说明性”是约束查询语言的一个目标, 然而, 任何数据库查询语言都必须具有高效的查询处理和优化机制。P. C. Kanellakis 等人首先提出一种支持密序约束的代数查询语言, 定义了六种基本代数操作。他们后来又将这种语言推广到单调双变量约束上。

3) 查询处理

查询描述的“说明性”和查询处理的“有效性”是约束数据库研究的两个重要目标。在查询优化方面, 包括对约束代数语言及求值技术的研究。通常, 基于

约束的空间数据库查询计值分为两个阶段: 第一是过滤, 先选出可能相关的空间对象; 第二是求精, 确定哪些对象为真正的答案。在索引技术方面, 任何数据模型的成功实现都有赖于对辅存的有效管理, 而索引技术是组织数据的有效方法。对约束数据模型索引技术的研究表明: 一维索引的重要问题是辅存上的区间管理。在这方面虽然有 P. C. Kanellakis 和 S. Ramaswamy 等人提出的 M 树和路径缓冲等数据结构及技术, 但是还有待于进一步的工作。

4. 半结构化数据处理技术

近年来, Web 的流行以及基因数据库的流行, 使用半结构化数据的组织和访问技术成为数据库领域研究的一个热点。半结构化数据又称为“自描述”数据, 有些半结构化数据没有独立的结构描述信息(模式信息), 如文本数据及基因信息等序列数据, 而有些数据虽然有结构描述信息, 但是数据与其结构描述之间的关系是很松散的, 如 HTML 文本之类的“平”文档数据。关于半结构化数据处理的研究工作有:

1) 数据组织模型

包括两个方面: 一是半结构化数据库模型, 它将整个 Web 视为一个大的数据库, 为其中的数据建模, 不但考虑一个 Web 主页中各成份的关系, 还要考虑各个 Web 页之间的关系。在这类数据模型中, 大都将半结构化数据表示成某种图或树的结构。一般的定义是, 由叶子结点表示“值”, 由中间结点表示“类名称”, 边常标注上“属性”; 二是“平”文档模型, 它只考虑某种结构内含文档内各元素的关系。将“平”文档的结构以合适形式描述为数据库能够接受的模式是管理“平”文档首先要解决的问题。

2) 查询语言

半结构化数据查询语言主要包括以下三个方面:

(1) 半结构化数据库查询语言。这类查询语言通常采用两种途径来研究: 一种是以 SQL 或 OQL 语言为基础, 增添必要的机制, 使其能够表达一组有用的查询; 另一种途径是以某种基于半结构化数据形式计算概念的语言为基础, 将其语法进行适当的变形, 成为一种便于使用的查询语言。应当说明的是, 根据两种途径所设计出的语言非常相似。

(2) “平”结构文档查询机制。采用数据库技术管理“平”文档, 使得“平”文档的应用具有并发、恢复等特点。对“平”文档中的结构信息, 可方便地利用数据

库提供的查询功能进行检索。但对于非结构信息,数据库缺乏有效的检索机制。这就要求管理“平”文档的多媒体数据库系统扩充基于内容查询等信息检索机制,以方便“平”文档的查询。

(3) 序列数据查询语言。传统的文本数据库查询以模板匹配为基础,仅能使用一组预先定义的算子对文本操作,而且,这些操作也缺乏一个理论框架。它难于满足基因数据库的需求。这种新型的序列数据库的查询语言应当满足:既能表达模板匹配也能表达序列重构,具有自然的语法和语义,在有序列重构时操作是安全的。

3) 查询处理

关于半结构化数据库的查询处理取决于系统的界面是基于半结构化数据模型还是某种(如面向对象)结构化模型。前者可使用传统的非结构数据的索引来优化查询处理,而后者也可使用传统数据库的优化技术。关于“平”文档的优化和更新操作有下列研究:

(1) 查询优化操作。查询优化是数据库系统长期以来研究的一个问题,是提高系统性能的关键。对于“平”文档管理数据库系统也存在同样的问题。文档结构描述提供了文档结构与数据库模式之间的映射,用户按照数据库模式定义的查询操作,必须作用于实际文档数据,得出查询结果,返回给用户,即用户如同对数据库进行查询一样。

(2) “平”文档更新操作。更新与查询优化一样,一直是数据库系统研究的问题。它比查询优化的难度更大。文档更新的问题在于:用户依照文档结构描述定义的数据库模式给出更新要求后,系统如何在实际文档中实现更新要求。

5. 演绎数据库(知识库)技术

演绎数据库是逻辑程序语言与关系数据库相结合的产物。七十年代初,Green等人以逻辑推理作为问题回答系统的基础,开创了用逻辑来表达和处理知识的先河。Kowalskit和Minker等人结合前人在逻辑程序语言和关系数据库方面的研究成果,以谓词逻辑作为统一的理论框架,研究数据描述、约束表示和用户查询的表达,开创了演绎数据库技术。

演绎数据库较早的工作集中在逻辑程序的语义研究方面,主要有Reiter的封闭世界假设[1978]、Clark的程序完全化[1978]和Emden与Kowalski的最小模型语义[1978]等。

八十年代以来,演绎数据库的研究主要集中在

逻辑程序的语义和查询优化技术两方面。前者根据规则体中负文字对语义的影响,研究的最小模型、理想模型、稳定模型和良基模型就形成逻辑程序语言语义的一个谱系。后者随着斯坦福大学的NAIL系统和MCC的LDL系统的研制,取得的成果包括基于程序语义等价性的优化、规则改写技术、运行优化等。演绎数据库研究的新进展主要集中在如下四个方面。

(1) 语义研究。前期的研究工作表明,良基语义似乎是逻辑程序最为“自然”的语义,但良基语义是3值的,而且缺乏必要的语法判定手段。K. A. Ross在1994年将局部分层进行推广,定义了一种2值良基语义,而且,根据规则以及程序事实在单调域上的某些约束,能够从语法上判定一个逻辑程序是否具有此种语义。1995年,他又将单调域约束的条件推广到任何域,从而进一步扩大到具有2值良基语义的程序类。

(2) 优化技术。语义查询优化是一种利用完整性约束(无头的Datalog规则)来优化逻辑程序求值的技术。它在递归查询及不能转换为合取查询之并的查询上的使用一直是一个研究的问题。A. Y. Levy等人在92~95年的工作给出了递归规则中使用有序约束(如选择)的方法。P. T. Wood在1995年将J. F. Naughton等人提出的魔因子法[1989]可以使用的程序类扩充到一类可以含有多个IDB的闭包程序上,有效地扩大了这种方法的使用范围。

(3) 逻辑语言中的更新。关系数据库的实践表明,更新语言应当与查询语言具有某些相同的性质,即面向集合和确定性。当允许逻辑程序中使用更新原子时,就会出现两个问题:一是查询部分与更新部分可能会互相影响,使得结果取决于求值规则及原子的顺序;二是对同一个规则不一致的更新操作(插入和删除相同的事实)会导致结果的不确定性。E. Bertino等人提出利用滞后语义和更新的约束表达来解决这两个问题,并针对带更新的Datalog(U-Datalog)进行了研究。

(4) 演绎数据库与面向对象数据库的结合。演绎数据库具有较完备的理论基础和说明性数据语言,然而它通常是基于平关系的,不支持具有复杂对象;而面向对象数据库则是应用驱动的,能处理复杂对象,但是缺乏完备的理论基础。两者互相补充,已产生演绎的面向对象数据库(DOOD)这一新的研究方向。研究成果有IQL,ROL和F-logic等。

6. 主动数据库技术

主动数据库技术的研究与开发工作始于七十年代中期。近来,主动数据库技术引起数据库厂商及研究者的极大兴趣,已有大量的项目着眼于关系或面向对象的主动数据库系统的设计与开发。数据库产品(如 Oracle, DB2, Sybase, Informix 等)都引入触发子机制以实现商业规则的主动处理功能。SQL3 标准也在考虑数据库的规则处理特性。

主动数据库并非是一种全新的数据库技术,而是给数据库系统中引入规则描述处理机制(如触发子),以便描述和自动维护应用中的商业规则。数据库产品大多已具有这些机制,只不过是各功能强弱不同。数据库系统不提供主动规则处理时,用户所写的数据库应用是一种被动应用,在这种应用中,只能利用数据库的创建、检索、修改和删除数据的功能,对于应用涉及的商业规则的维护只能由应用来处理。

1) 主动数据库系统具有的潜在优点

触发子机制为信息系统的商业规则提供一种统一、集中式的描述方法。利用它能够自动实现商业规则的维护,它具有较高的可靠性。这是由于规则的维护及处理是由数据库系统而不是应用程序来完成的。利用触发子机制可能提高应用的性能,可以集中描述某些应用语义,便于优化,同时可以用于应用的性能调节。

2) 主动数据库技术的现状及问题

虽然在主动数据库技术的研究中已出现了很多成果,而且不少得以实际应用。当今的大部分数据库产品也都或多或少地具有主动处理用户定义规则的能力,但是,当主动数据库系统应用到实际应用,特别是大型应用的开发过程中时,仍存在许多问题。

(1)当前的各类 DBMS 提供的触发子语言缺乏足够的表达能力。触发子语言的基本语句是定义事件-条件-动作规则。例如,要定义一个简单的商业规则“若某雇员的收入超出其经理的收入,则发出提示”,可能需要定义多个规则来描述它。这是由于引起违反此约束的操作很多,而每个操作都要定义一条事件-条件-动作规则。此外,某些系统中限制与一个关系相关联的规则数目,而且一个条件所能触发的最大规则数也有所限制。

(2)触发子语言缺乏简单和标准的定义,主要表现为:一是现有触发子语言的语法和语义差别较大,使得用户涉及触发子的应用难于在不同的系统

间移植,这有待未来的 SQL3 标准给予缓解;二是触发子的作用粒度(是元组级还是语句级)不同;三是在同时支持触发子和完整性断言的系统中,两种机制的差别不清楚。

(3)安全性、可靠性和不可判断性等问题。主要表现在:一是规则设计的方法和指导原则,用户很难确定定义和规则在实际运行时是否遵循自己的意图;二是涉及触发子的事务的行为难于理解。

(4)性能问题。性能不高是大型应用开发者不愿意使用触发子的主要原因。其关键在于:触发子规则设计缺乏指导方法,应用程序中将事务程序与规则定义分离时,使得全局优化难于进行,导致系统效率降低。

7. 视图与更新等技术

这方面的工作有视图的应用,更新操作的研究和异构数据库的集成问题。

1) 视图

(1)视图的查询重写。研究的问题有:给定一个查询和一组预定义的视图,是否能由这些视图来表达查询?它是早期有关合取查询的等价、包含等研究工作的继续,对查询优化和多数数据源的集成等应用具有重要意义,一般说来,是否能由任意一组视图来改写某个查询的问题是无可判定的。对于合取查询虽然是可判定的,但是 NP-完全的。A. Y. Levy 等人[1995]年提出一种对不带内部谓词的合取查询,在某些情况下能给出极小重写(含有极小子目标)的多项式算法;O. M. Duschka 等人[1997]提出一种仅用视图改写合取查询的方法,并证明此改写判定问题具有双指数复杂度;C. Beeri 等人[1997]将视图改写技术用于复杂对象视图(用描述逻辑来表示),并证明改写判定问题为指数级(O^{ns})(m 为描述级, k 为描述嵌套深度, s 为构造子数)。

(2)视图的维护。视图的更新往往与数据模型、视图的定义及更新操作的语义具有密切关系。近来, H. V. Jagadish 等人针对交易信息数据库视图的维护进行了有益的研究,提出一种数据模型,将概括信息定义为交易信息的视图,利用提出的 SCA 代数,对这些视图进行增量更新。

2) 更新操作

(1)对象更新。更新问题是数据库系统研究的一个热点与难点,在面向对象数据库中,该问题更为复杂,面向集合的对象更新技术是一个研究方向。在更新对象集合时会遇到两个问题:一是对单个对象的

更新会导致多个对象的实质更新;二是对集合中的多个对象更新时,更新的次序,即并行更新还是顺序更新。

(2)复本数据更新。近年来,随着网络及分布式计算等的发展,复本数据的使用越来越多。它为早期复本数据处理的有关研究注入新的活力。

3) 异构数据库的集成

R. Hull 在 PODS'97 上作了关于数据库集成中语义异构问题的特邀报告,包括:

(1)数据库互操作的体系结构。Hull 总结为四类:协调者模式、联邦模式、带更新的协调者模式和工作流模式。

(2)数据源透明。解决这种问题的方法形成了一个谱系。一个极端是开发统一的数据表达形式和数据模型;另一个极端是允许不同的系统具有异构性,仅给用户及程序员提供处理异构的工具。其它各种处理策略都介于两者之间。

(3)数据表示。语义异构问题是由于采用多种方式来表达相同或重迭数据而造成的,因此比较不同数据模型和数据库模式的信息表达能力是一个重要研究问题。

(4)只读访问的数据集成。在这种集成策略中,全局访问只查询而不更新各种数据。这方面的工作主要集中在研究由局部模式来构造全局模式的方

法,输入和集成各数据源的选定部分及局部数据依赖的集成方法等。

(5)读写访问的数据库集成。在这种集成策略中,要着重考虑的是更新问题。它与视图更新很相似。视图更新问题较复杂,涉及到全局事务处理等技术。

(6)操作元数据。在数据库集成过程中,开发访问各数据源的元数据,即数据库模式及其语义信息的方法和工具非常重要,在此方面已有许多研究成果。

8. 其它研究问题

数据库领域的其它研究工作还有查询的复杂度理论[C. H. Papadimitriou 1996, 1997; M. Y. Vardi 1995; C. Beeri 1996], 一阶增量求值系统(FOIES) [G. Dong 1995]和 O. I. Keridar 关于增强的三阶段提交协议(E3PC)[1995]等。

参考文献

- [1] PODS: Proc. of ACM Symposium on Principles of Database System, 1993-1997
- [2] SIGMOD: ACM SIGMOD Conf. on Management of Data, 1993-1997
- [3] Jon Doyle et al., Strategy Direction in Artificial Intelligence, ACM Computing Survey, 28(4)1996
- [4] B. Grosz and R. Davis, A Report ARPA on 21st Century Intelligent System, AI Magazine, 15 (3) 1994
- [5] D. E. Heckerman, Probabilistic Similarity networks, MIT Press, Cambridge, MA, 1991
- [6] E. Horvitz et al., A decision-theoretic approach to the display of information for time-critical decision: The Vista Project, In Proc. of SOAR-92 Conf. On Space Operations Automation and Research, 1992
- [7] D. E. Heckerman et al., Decision-theoretic troubleshooting, CACM 38(3)1995
- [8] R. F. Lyon and L. S. Yaeger, On-line hand-printing recognition with neural networks, In Fifth Intl. Conf. on Microelectronics for Neural Networks and Fuzzy System, IEEE Computer Society Press, Los Alamitos CA, 1996
- [9] T. Jochem et al., Life in the fast lane: the evolution of an adaptive vehicle control system, AI Magazine, 17(2)1996
- [10] T. E. Senator et al., The financial crimes enforcement network AI system (FAIS): Identifying potential money laundering from reports of large cash transactions, AI Magazine, 16(4)1995
- [11] D. R. Smith et al., Synthesis of planning and scheduling software, In Advanced Planning Technology, AAAI Press, Menlo Park, CA 1996
- [12] J. Goebel et al., A Bayesian classification of the L-RAS LRS Atlas, Astron. Astrophys, 222, 1989

(上接第 28 页)

经开始了下一个五十年的远航,我们为它的前景所振奋,渴望同其它学科一起合作,并且对它对计算研究的贡献、相关性和中心地位充满自信。

参考文献

- [1] Jon Doyle et al., Strategy Direction in Artificial Intelligence, ACM Computing Survey, 28(4)1996
- [2] B. Grosz and R. Davis, A Report ARPA on 21st Century Intelligent System, AI Magazine, 15 (3) 1994
- [3] D. E. Heckerman, Probabilistic Similarity networks, MIT Press, Cambridge, MA, 1991
- [4] E. Horvitz et al., A decision-theoretic approach to the display of information for time-critical decision: The Vista Project, In Proc. of SOAR-92 Conf. On Space Operations Automation and Research, 1992
- [5] D. E. Heckerman et al., Decision-theoretic troubleshooting, CACM 38(3)1995
- [6] R. F. Lyon and L. S. Yaeger, On-line hand-printing recognition with neural networks, In Fifth Intl. Conf. on Microelectronics for Neural Networks and Fuzzy System, IEEE Computer Society Press, Los Alamitos CA, 1996
- [7] T. Jochem et al., Life in the fast lane: the evolution of an adaptive vehicle control system, AI Magazine, 17(2)1996
- [8] T. E. Senator et al., The financial crimes enforcement network AI system (FAIS): Identifying potential money laundering from reports of large cash transactions, AI Magazine, 16(4)1995
- [9] D. R. Smith et al., Synthesis of planning and scheduling software, In Advanced Planning Technology, AAAI Press, Menlo Park, CA 1996
- [10] J. Goebel et al., A Bayesian classification of the L-RAS LRS Atlas, Astron. Astrophys, 222, 1989