

演绎数据库系统的研究与发展

余金山 (华侨大学)

摘 要

Deductive database systems are now one of the most important directions in the development of database technologies because of their many attractive properties. They have received widespread intensive interest and very great attention. In this paper We first give a brief overview of the development of deductive database systems. Then We motivate as extensively and thoroughly as possible the directions and significant research topics in this field. The status, importance, trends, research methods of these topics, and the relationship among these topics and other disciplines are also discussed.

一般认为,演绎数据库(DDB—Deductive Database)“诞生”于1978年,至今仅十一年的历史。但是由于其突出的特点,演绎数据库自面世以来就引起了人们的极大重视和广泛兴趣。目前DDB已构成了数据库学科中的一门独立分支,成为数据库技术中最受重视的前沿研究领域之一。这些可见

诸于近期以来出版的大量有关文献及其国际上举行的各种有关学术会议,包括近几届的VLDB国际会议^[1,10,12,18,21]。我国对DDB的研究大约只有六年的时间。但不管在理论方面还是实践方面也都取得了不少鼓舞人心的成果^[14,15,16,19],由此可见演绎数据库的重要性。用著名学者Ullman教授在“数据库

5. 结 论

有些作者,[Kim 88]和[Dittrich 86]争辩说OODBS是一个以面向对象的数据模型为基础的DBMS。如果在更广的意义上对待数据模型的概念,尤其包括面向记录以外的其它方面,则此观点无疑相应于我们的方法。[Dittrich 1988]和[Dittrich 1988]引入了面向对象的数据模型(及随后的OODBS)的分类:若一个模型支持复杂对象,则称其是结构上面向对象的;若提供可扩充性,则称其是行为上面向对象的;一个完全面向对象的模型必须同时提供这二个特性。这个定义也需要持久性、磁盘管理、并发性及恢复;它至少蕴含地假定了大部分其它特性(是否合适,要依不同的类而定);总的说来,它比我们的方法略为宽松。但是,不管怎样大多数当前的系统和原型没有达到我们的定义规定的所有需求。所以这个分类对比较已有系

统和现行工作提供了一个有用的框架。

我们已经给出了一组关于面向对象的数据库系统特性的定义。就我们所知,本文中给出的金科玉律是目前关于面向对象的数据库系统最详细的定义。特性的选择和对它们的解释出自于对当前系统的分析和实现的经验。对面向对象的数据库设计、实现和形式化的进一步体验将无疑会修改和完善我们的观点(换言之,如果将来你听到我们中的某位作者鞭挞当今的定义请勿吃惊)。我们的目标仅提出一个具体的意见以引起科学界的争论,批评和分析。这样我们最后的规则是:你应该怀疑金科玉律。

参考书目(略)

[王宗斌 王 珊 译校自 Proceedings of the first deductive and object-oriented database, Japan, Dec. 1989, pp40-pp57]

理论的过去与未来^[4]”一文中开头的一句话来说是：“过去是Prolog，未来还是Prolog。”其中Prolog一词实际上指的就是面向逻辑的演绎数据库。关于演绎数据库的优点及其研究的重要意义，许多文献^[1,2,14,8]都作了较为详细的介绍，也是大家所熟知的，因此这里不再叙述。目前，国内对演绎数据库的研究十分活跃，文献^[14]指出：已处于高潮时期。本文将根据国内外的研究状况，并结合我们的体会与观点，全面、深入地讨论这一研究领域各种富有重要意义的研究课题，其范围包括研究内容、现状、发展趋势、研究方法、各课题之间的联系以及它们与其它学科之间的联系等等。此外，为使读者对演绎数据库的过去、现在与前景有一个比较清楚的了解，我们也将讨论一些有关基本概念，回顾一下发展历史，并展望其未来。

一、发展简史

与演绎数据库有关的研究工作最早可追溯到1969年。当时年仅二十几岁的青年人、斯坦福大学的博士研究生 Cordell Green 成功地开发了一个叫QA的问题解答系统，并发表了多篇有关的论文。QA系统的核心本质上是一个基于一阶谓词演算和Robinson归结原理的自动定理证明器。Green的工作十分重要，被认为是演绎数据库领域中的开创性工作，QA系统的出现意味着演绎数据库已开始孕育^[1]。为此ACM在1985年授与Green Grace Murray Hopper奖。

从Green的工作算起，演绎数据库的发展过程大致可分为：孕育形成、理论发展与基础研究、实现与应用等三个阶段。

1.1 孕育形成阶段 (1969—1978)

在这个时期，对演绎数据库的形成具有重大影响的工作大致包括以下几方面：

1. 逻辑程序设计(LP)思想的发展。自第一个PROLOG语言出现之后，特别是1974年在斯德哥尔摩召开的IFIP会议上，Kowalski提出了具有划时代意义的思想——谓词逻辑作为程序设计语言。尔后，逻辑程序设计的研究十分活跃，有关的成果对演绎数据库产生了巨大的影响（因为从逻辑的观点出

发，演绎数据库可被看成是一个逻辑程序）。其中包括Van Emden和Kowalski论及逻辑程序的不动点语义和操作性语义的经典性论文^[22]。实际上有许多研究是直接在这两者的交叉领域中进行的。

2. 在Green工作的影响下，1970年美国马里兰大学的J. Minker等人进行了一个他们称之为演绎数据库系统MRPPS 3.0的实验性研制工作。MRPPS 3.0实现了许多在今天看来仍然是非常有用的重要技术，如语义合一、子句索引技术、数据库查询中的完整性约束技术、编译技术等等，为演绎数据库的开发提供了宝贵的经验。

1977年，根据当时的研究现状及发展趋势，在Gallaire和Nicolas的倡导和组织下，于法国的土鲁斯城召开了第一次“逻辑与数据库”的专题讨论会。会上各位专家学者交流了他们各自的研究成果、经验和观点。其中具有重大影响的文章包括：Nicolas和Gallaire的关于数据库的模型论和证明论观点；Reiter的封闭世界假设(CWA)公理和编译技术；Clark的否定作为失败公理；Minker等人关于实际系统的开发经验与技术；Kowalski的使用逻辑描述数据的方法；Nicolas和Yazdaniyan的关于完整性约束的讨论；以及Futo等人关于应用PROLOG实现演绎数据库的方法与技术等等。这些成果后来都收入了Gallaire和Minker编辑的“逻辑与数据库^[11]”一书。该书的出版(1978年)标志着演绎数据库这一研究领域的诞生。通常也把该书的出现看作是演绎数据库研究工作的真正开始^[14,17]，因为该书首次明确地提出了演绎数据库的基本概念，全面地描述了数据库与逻辑之间的相互作用和关系。

1.2 理论发展与基础研究 (1979—1987)

这个阶段是演绎数据库和逻辑程序设计的理论研究以及为实现有关的基础研究迅速、全面地发展与完善的时期，是DDB和LP发展中最富有成果和最重要的时期。SLD推

理系统的完善,SLDNF推理系统的建立,演绎数据库形式理论(主要是证明论观点)的建立,CWA规则的推广,开世界假设(OWA)的引入,把AI中的界限(Circumscription)概念应用于DDB,公理的编译理论,递归查询处理,PROLOG与传统数据库的接口问题,用PROLOG直接实现DDBS的理论基础,对并行DDBS的探讨,层次(hierarchical)数据库、层迭(stratified)数据库以及与域无关(domain-independent)的数据库等一些特殊类型的数据库的提出等等,许多具有重大意义的研究成果都是在这个时期取得的。

1986年由马里兰大学J. Minker教授组织,在华盛顿召开了“演绎数据库与逻辑程序设计基础”的专题讨论会。会议的优秀论文均收入了1987年由Minker教授编辑出版的“演绎数据库与逻辑程序设计基础^[10]”一书。如果说文献[11]的出现标志DDB的诞生,那么文献[10]的问世则标志着演绎数据库(以及逻辑程序设计)理论的成熟^[1]。

文献[10]以及[3]、[5]、[6]、[9]总结了这一阶段所取得的几乎全部重要成果。特别是文献[10]、[3]、[6]尤为突出,它们已成为每个有志从事DDB或LP研究的人士必读的杰出文献。当然它们也是每个有关的研究工作者的不可多得优秀参考资料。

1.3 实现与应用阶段(1988年—)

到1987年,演绎数据库的基础理论虽然已相当成熟,牢固、严密的理论体系虽然已经建立,人们也从理论上认识到DDBS相对于RDBS的优越性,但存在的重大缺陷是DDBMS的实现问题。虽然已经有了许多实验性系统,但在商业上尚没有实用、成功的DDBMS产品投放市场,更谈不上应用的发展了。可见DDBS所走过的路程与当时RDBS所走过的路程是极其相似的。对RDBS来说,要是没有象System R和INGRES那样杰出的系统的研制成功,那么RDBS能否象今天一样在信息管理中占主导地位是值得怀疑的。因此,从1988年始,人们开始把注意力转移

到了DDBS的实现方面。期望在十年内,DDBS能象目前的RDBS一样成为一种被广泛应用的标准DBS。

二、什么叫演绎数据库

什么叫演绎数据库呢?目前文献上对这个问题的解释不尽相同,对演绎数据库的叫法和定义也不太一致。演绎数据库的发展历史较短当然是其原因之一,但其中也反映了人们看待演绎数据库的不同角度。因此稍作一些分析与比较是有益的。下面我们先给出几种常见的、具有代表性的定义,然后对它们进行分析和综合比较,希望能使读者对这个基本概念有更明确的认识并从中得到某些启示。

文献上常见的几种关于演绎数据库的定义如下:

(1)演绎数据库是数据库语句的有限集,数据库语句是形为

$$A \rightarrow W$$

的一阶谓词公式。其中, A 是形如 $P(t_1, t_2, \dots, t_n)$ 的原子公式, W 是任意一阶谓词公式(可为空)。 A 中的任一变量及 W 中的任一自由变量均假定是全称量化的^[3,6]。

(2)所谓演绎数据库,就是将数据库看成一个演绎系统,即一个数据库可与一个演绎理论相一致,它由一些公理组成,通过公理中的演绎规则可推导出定理^[3,14]。(这里略去了关于理论的详细定义。)

(3)将逻辑语言与数据库技术结合,使数据库系统具有推理演绎的功能,称为演绎数据库。

(4)能从已知事实推出新事实的数据库称为演绎数据库^[3,17]。

(5)演绎数据库是一种具有推理功能的数据库系统^[15]。

(6)知识数据库是以现行的DBMS为核心,增加了知识处理能力。如果它支持的是语义知识,即是“语义数据库”,若支持的是演绎知识,就是“演绎数据库”^[16]。

此外,在许多地方,演绎数据库还被称为逻辑数据库^[3,9,11]。

可以看出, (1)、(2)主要是从理论的角度出发, 给出了演绎数据库的一种严格的形式定义, 以便从理论上对它进行研究; (3)、(4)、(5)则从直观的, 工程的角度来看待演绎数据库, 明确地指出了DDB是一个数据库, 但增加了推理的功能, 因而容易让人联想到演绎数据库的实现可以通过结合现有的数据库技术和人工智能中的推理技术来达到; (6)则从知识的角度来看待演绎数据库, 因而有利于用AI中知识处理的思想来考虑问题。另外, (1)、(3)还强调了演绎数据库研究的逻辑方法, 即所谓的逻辑数据库或面向逻辑或基于逻辑的演绎数据库。

我们认为, 比较贴切且具有一般性的叫法应该是: 演绎数据库系统是一种具有演绎推理能力的数据库系统。因为, 一方面, 我们知道推理形式除了演绎推理外, 还有归纳推理等等, 而演绎数据库是不能进行归纳推理的。另一方面, 尽管逻辑因为其具有易于理解和十分完善的理论等优点, 特别是逻辑程序设计思想的发展, 在演绎数据库的研究中, 它几乎占据了绝对的统治地位 (这可能是为什么常常把演绎数据库叫做逻辑数据库的原因), 但是这并不排除用其它的方法来研究DDB。例如文献[8]就给出了一种基于代数中格论的演绎数据库模型。在推理方面, 演绎推理除了包含逻辑推理外还包含不确定性推理等等。

三、研究方向与课题

萨师煊教授指出^[20], 数据库的研究课题跨越于应用、系统软件与理论三个方面。演绎数据库是一种新的数据库技术, 其本质仍然是数据库。因此我们认为它的有关研究依然涉及这三个领域。其中, 理论方面虽然已有大量的研究, 也已相当成熟, 但是其研究历史毕竟还短, 其研究难度也较大, 因而仍有许多重要问题需要进一步探讨和完善。而系统的实现和应用方面则是当前的重点研

究方向, 已引起了人们的普遍关注和巨大兴趣^[8, 21]。下面我们分别就这三个方面对有关的重要研究课题进行介绍和讨论。

3.1 理论研究

3.1.1 建立较为理想的DDB理论体系

研究结果已经证明, 全一阶理论为DDB提供了一种优美的、严密的理论基础, 但在实现上则是困难的、低效的。因此必须从理论的严密性和实现的简单性和效率出发, 寻求一种相对合理的理论, 以指导DDB的发展。南京大学徐洁磐教授等人提出的扩充模型论EMT^[10]以及美国MCC的LDL可看作是这方面的具有开创性意义的成果。但是值得注意的是, 如何保持面向一阶逻辑的DDB所具有的能够把数据子语言与主语言完全地且简单自然地统一起来的这一突出优点。

3.1.2 并行推理与分布式DDB

已经明确, 要解决逻辑程序设计或演绎数据库效率低这一致命缺点, 并行处理技术是一种主要的, 甚至被认为是一种最有希望的方法。因而并行推理已成为实现第五代计算机的关键技术之一。然而, 关于并行逻辑程序或分布式DDB的理论模型及其语义等问题到现在为止还没有彻底弄清。显见这一研究课题具有十分重要和深远的意义。一些人已从度量空间和拓扑学的角度出发, 进行了有限的艰难探索并取得了某些初步结果^[6], 但大部分问题还未得到解决。看来难度很大, 尚需付出巨大的力量。

3.1.3 更多地利用AI的思想与技术

演绎数据库是AI与DB技术相结合而产生的一种数据库新技术。但是目前仅仅利用了AI的演绎推理中的某些概念和思想。实际上可利用的思想还很多, 特别是在不完全信息的处理方面。目前被认为可引入到DDB的最有希望的AI思想是所谓的周界和信念(belief)思想。已有人做了初步的探讨^[1]。例如, 如何具有智能性地回答诸如“谁既是张三的母亲又是张三的父亲”这样的问题是富有意义的。

3.1.4 引入非一阶逻辑 例如高阶逻辑。因为高阶逻辑具有较一阶逻辑更强的表达能力,因而使用高阶逻辑,某些问题特别是有关数据库的一些问题的表达和处理将更为方便自然。但是对这个问题存在不同意见。有些人认为,基于一阶逻辑的程序设计或演绎数据库的理论基础比较简洁易懂而且已被广泛深入地研究过,因此应该珍惜它;再者利用元程序设计技术,完全可以在一阶范围内实现高阶的功能。因此他们不主张使用高阶逻辑。到底哪一种方法为好,目前尚无定论。

应该提及的是不管是否使用高阶逻辑,元程序设计已经成为一种重要的技术。因为通过使用元程序设计技术还具有使系统的语义清晰、易于理解和修改的优点。其有关研究已引起了人们的极大兴趣。

时态逻辑是另一种十分值得探讨的逻辑。除了明确的语义和有力的演绎推理功能外,时态逻辑的一个突出特点是能够很好地表示和处理与时间有关的概念和动态知识。这些特点对演绎数据库来说无疑是十分重要和富有意义的。已有人开始研究基于时态逻辑的程序设计语言并探讨它们在数据库中的应用。但有关工作刚刚在起步。

此外还有人开始在探讨引入模糊逻辑和模态逻辑。

3.1.5 特殊类型的DDB 这个问题主要研究某种具有足够强的表达能力且又有严密的理论支持的DDB。已提出的两类最重要的DDB是所谓的层迭DDB和与域无关的DDB。

演绎数据库的理论基础大部分是基于所讨论的DDB是确定的(definite) DDB或Horn DDB这样的假设建立起来的。所谓确定的DDB是指仅由确定的数据库子句构成的DDB。确定的数据库子句是形为

$$A \leftarrow A_1 \wedge A_2 \wedge \dots \wedge A_n$$

的子句。其中 $A_1, A_2, \dots, A_n$ 均为原子。显然,这样的数据库没有足够的表达能力。本

文第二节中引用的定义(1)可看成是一种具有一般性的定义。但是要建立其所需的理论基础则比较困难。因而Apt^[10]和Chandra等人提出了层迭数据库的概念并取得了许多重要成果。所谓的层迭数据库是指这样的数据库,存在一个等级映射,使得数据库的每个子句 $p(t_1, \dots, t_n) \leftarrow w$,在 w 中出现的每一个肯定谓词其等级小于或等于 p 的等级,而以否定形式出现的谓词其等级小于 p 的等级。其中,等级映射是一个从谓词符号集到非负整数上的映射。

与域无关的DDB是Topor^[7,10]等人提出来并加以研究的。其出发点是:在什么样的情况下才能保证一个DDB的语义不随着数据库子句中的变量的定义域(domain)的变化而变化?试看下面的一个极简单的DDB D :

$$D: \begin{aligned} &p(a) \leftarrow \\ &q(x) \rightarrow \forall x p(x) \end{aligned}$$

D 是与域有关的。因为从直观上可以看出,当且仅当 a 是 x 的定义域中的唯一元素时, $q(a)$ 才是“正确的答案”(这里略去了有关概念的严格定义)。显然,与域无关的DDB具有重要的现实意义,因为我们知道在实际应用中,数据库大多都是动态的。

有关以上这些方面的研究大致可朝两个方向发展:(1)探讨其它类型的DDB,(2)对已经提出的某类DDB进行深入、全面的研究,例如在与域无关的DDB中就有许多问题值得进一步研究。欲知详情,读者可参阅文献[1]、[7]、[10]。

3.2 演绎数据库管理系统的实现

3.2.1 总体实现方法 关于演绎数据库管理系统(DDBMS)总体实现的三种流行方法:PROLOG方法、传统DBMS+演绎层方法和用专门软件来实现的方法,文献[14]中已作了较全面的介绍和分析。在这里我们针对其中的某些问题做如下讨论。

(1)用PROLOG实现。支持用PROLOG实现演绎数据库的牢固理论基础是Lloyd和Topor的研究成果^[5]。我们认为用PROLOG实现演绎数据库

不但具有一定的现实意义,而更重要的,还具有深远的意义,它应该成为最主要的方向,因为逻辑程序设计的优点之一就是它把数据库和程序设计在形式上完全而且自然地统一了起来。至于用PROLOG实现演绎数据库目前存在的效率低、增、删、改功能实现困难等问题,可以认为那是PROLOG本身存在的缺陷。PROLOG的最终目标应该是成为一种既适用于传统的程序设计任务,又适用于AI程序设计任务以及数据库任务的通用逻辑程序语言。有关PROLOG的研究与开发也正是朝这个方向努力的。例如,澳大利亚墨尔本大学研制的Mu-PROLOG和Nu-PROLOG,除了提供自动灵活的控制策略外,还提供了强有力的演绎数据库设施^[6],从而使得用它们实现的大型演绎数据库,其查询效率也可达到令人满意的实用水平,尽管仍存在有待解决如何高效且具有智能性地实现增、删、改功能等重要问题。可见用PROLOG来实现演绎数据库除了能较方便地满足实际应用的需要外,还可为PROLOG的发展和完善积累经验、提供基础。

(2)传统的DBMS+演绎层。这种方法实质上是对传统DBMS的功能扩充,即增加演绎推理能力。PROLOG语言是实现这种扩充的较好工具,而且已成为扩充的主要载体。这种方法由于可以充分利用现有的传统数据库系统,因而实现容易,见效快。国内外都已做了大量的工作。但是这种仅仅是功能性扩充的工作,专家们已普遍认为是不大可能产生新的积极成果的工作^[23]。有关研究似可作为“短、平、快”的项目,但看来不宜投入太大的人力物力。

在文献[16]中,姚卿达教授等人则指出了一条新的道路。他们把对传统DBMS的扩充提高到他们称之为知识数据库的角度来认识,在实现上采取局部增强型和整体耦合型的体系结构,同时实现了智能化和可扩充性的双重目标。可扩充的数据管理又正是目前数据库领域中另一个极其重要的研究方向。从而显出了该方法的特色和生命力。

(3)用专门软件实现。这种方法实际上是理论问题3.1.1的研究结果的应用与实现。DMTL^[15]等的研究与实现说明了这是一条有希望的新途径。但是这种方法由于一切都得从头开始,因而需要的研制周期较长,工作量也较大。

根据DDB和LP的关系、它们的发展与现状,我们认为专门软件的研究与实现,除了考虑其功能和效率外,其理论基础和语言的形式应尽量与逻辑

程序设计相一致,使得有可能把这种专门软件作为逻辑程序设计语言系统的一个数据库子系统,从而能用统一的方式完成程序设计与数据库应用任务。

3.2.2 查询算法与优化策略 查询处理与优化策略是演绎数据库在实现问题上受到重视最多的方面。已进行了大量的研究并提出了许多非常有效的算法。进一步的研究大致朝三个方向发展:(1)扩大算法的适用范围;(2)提高性能;(3)具体实现。采取的措施包括提出新算法和对现有算法的改进两方面。

(1)扩大适用范围。现有的大部分算法其适用范围虽不尽相同,但都是有限的。如仅限于线性递归规则或自底向上可求值规则等等。因而有必要扩大其适用范围。已有人把魔集算法推广到层迭数据库和容许(allowed)数据库。一个算法的适用范围当然是越大越好。

(2)提高性能。提高性能的重要性是显然的。已有的研究结果表明了,重复计算,相关事实集的大小和中间结果的存放形式是影响性能的三个最重要的因素。目前所采取的措施可分为:不结合具体的文件结构(如仅考虑处理过程和对有关规则进行变换等)、结合某种具体的文件结构和基于语义的优化三种。根据现有的状况,我们认为后两种可能是较有希望进一步发展的方法。

(3)具体实现。具体实现的问题包括降低算法实现的复杂性和代价以及对实现技术的研究两方面。因为算法若难以实现就可能失去其应用价值。另一方面,在已有的众多算法中,由于它们的适用范围及性能具有较大的差别,因此在开发DDBMS时,如何选择一种或几种妥当的算法并加以合理实现的问题无疑也是十分重要的。

在这里还值得一提的是文献[9]和[12]。其中,文献[9]对现有的大部分算法进行了全面的分析与比较。文献[12]则详细地研究了有关具体实现的问题。根据现状,已有许多很好的算法,关键是如何合理地实现之。这也与RDB发展中当时着手研制System R的情况十分相似。因此我们认为对算法具体实现的研究更加富有现实意义。

3.2.3 用户界面研究 用户界面的研究已在第14届VLDB会议上提出作为最重要的研究方向之一而加以肯定。许多研究结果已经表明逻辑程序设计语言是实现自然语言

处理的有力工具, 各种逻辑文法如DCG、XG、MSG、RLG等实际上就是语言学与LP结合的产物。演绎数据库又与逻辑程序设计紧密相关, 因此演绎数据库应该在提供良好的用户界面方面发挥其应有的潜力。

3.2.4 其它 另外一些与DDB的实现有关的重要研究课题包括: 元程序设计技术在DDB中的应用、完整性约束处理、分布式DDB的实现方法与技术、与各种特殊类型DDB有关的实现问题等等。从广泛的意义上来看, 逻辑程序设计语言与函数式语言的结合也与这一研究范畴有关, 因为在证明论的观点下, 一个DDB可看成是仅由语言和推理系统两部分构成的。

4.3 应用研究

类似于传统的DBS, 应用研究在这里专指演绎数据库的设计问题。由于迄今尚无在商业上具有竞争性的演绎DBMS投入市场, 因而有关演绎数据库设计的研究几乎完全未被触及, 也未引起足够重视。我们认为, 演绎数据库作为数据库的一种新技术, 其最终的目的是要投入实际应用的, 而且希望它的应用范围比传统的数据库更广。演绎数据库设计的重要性不亚于传统数据库设计的重要性。另一方面, 鉴于近几年来在演绎数据库实现方面投入的巨大兴趣, 有理由相信在不久的将来它将得到普遍的应用。因此, 现在开展对演绎数据库设计的研究显然具有重要的意义。下面我们提出几个我们认为值得重点考虑的问题。一些术语, 如概念结构、逻辑模型等等, 在DDB中尚无明确的定义, 为方便起见我们姑且按大意沿用传统数据库中的叫法。

(1) 演绎数据库的设计方法学与设计工具。由于其本身的特点, DDB的设计方法与工具显然不能照搬传统数据库的设计方法与工具。那么DDB的设计方法学的研究以及有关支持工具的开发显然是重要的、迫切的。

(2) 演绎数据库的体系结构。传统DBS的外模式、概念模式和内模式的三级结构对DDBS仍然适用吗? 若回答是肯定的, 那么两者之间有哪些联系

与区别? 如何充分利用DDB的特点, 又如何解决由此可能带来的新问题(如安全性问题)? 若回答是否定的, 那么DDBS的体系结构是什么?

(3) 数据模型。传统数据库设计中流行的E-R模型对DDB来说显然已不适用。在已有的其它语义数据模型中存在适合于DDB设计的模型吗? 能否找到象E-R那样受欢迎的有力模型?

(4) 逻辑模型设计。能否给出一组普遍适用的准则以帮助设计人员确定把何种信息成份定义为事实, 演绎规则或完整性约束?

(5) 与逻辑程序设计的关系。DDB与LP紧密相关。在形式上, 一个演绎数据库与一个逻辑程序没有什么差别。那么DDB的设计与逻辑程序的设计有什么联系与区别? 两者之间有哪些方面可以相互借鉴和利用?

四、结 论

演绎数据库由于其突出的优点和重要性, 从出现起十几年来一直深深地吸引着大批研究工作者为之努力工作。当今它已成为数据库技术发展的主要方向之一, 受到了极端重视。本文简要地介绍了它的发展历史, 从不同角度探讨了什么是演绎数据库这个最基本的概念, 然后力求广泛、全面地提出了有关的研究方向和研究课题, 并对各课题作了不同深度和广度的讨论。从中我们可以看到:

1. 演绎数据库的发展速度之快是令人吃惊的, 其前途是十分光明的。一些著名的专家预言, 到1993年DBMS将使用支持全部知识库系统的90%的简单推理技术进行扩展^[10], 到1997年演绎数据库系统将象RDBS在目前为标准数据库系统一样成为一种普及的标准DBS^[6]。这些预言是完全有可能实现的。当然, 它依赖于实际应用的迫切需要, 依赖于广大研究工作者的共同努力, 依赖于商业界的大力支持。

2. 演绎数据库管理系统的实现研究是当前的研究重点。DDBS当前的状况十分类似于RDBS发展中人们着手研制INGRES和System R当时的状况, 理论与技术都已相当成熟, 但缺乏在商业上具有竞争性的实用系统。根据已发表的成果看, 国内对DDB的研究正是放在实现这个重点上, 这是可喜

的。

3. 演绎数据库的设计牵涉到实际应用问题, 十分重要, 但目前对它的研究却很少, 必须引起足够重视。

4. 具有严密、牢固的理论基础是DDB的突出特点之一, 虽然它已趋于成熟, 但尚有许多问题有待进一步研究解决, 特别是在并行结构、不完备信息的处理和特殊类型的演绎数据库方面。

5. 演绎数据库、逻辑程序设计和人工智能三者紧密相关。它们在很多方面都可以互相借鉴、互相利用和互相促进, 而且三者都是计算机科学技术发展的重要方向。因此它们的交叉研究具有十分重要的意义。其实已有的许多成果正是这种交叉的产物。

6. 把DDB与数据库发展的其它两个重要方向, 即面向对象数据库以及可扩展的数据库, 结合起来研究也是可能的, 而且一些研究结果^[13,16]已表明这种结合是很有前途的。

参考文献

- [1] J. Minker, "Perspectives in deductive databases," *J. Logic Programming*, no.5, 1988.
- [2] J. W. Lloyd, "An introduction to deductive database systems," *Australian Computer J.*, no.2, 1983.
- [3] H. Gallaire et al, "Logic and databases, A deductive approach," *Computing Survey*, no.2, 1984.
- [4] J. D. Ullman, "Database theory: Past and future," in *Proc. Principles Database Theory*, ACM, 1987.
- [5] J. W. Lloyd and R. W. Topor, "A basis for deductive database systems," *J. Logic Programming*, no.2, 1985, no.1, 1986.
- [6] J. W. Lloyd, *Foundations of Logic Programming*, 2nd Edition, Springer-Verlag, 1987.
- [7] R. W. Topor, "Domain independent formulas and databases," (to appear in: *Theoretical Computer Science*)
- [8] N. Spyrtatos, "The Partition model: A deductive database model," *ACM Trans. DBS*, no.1, 1987.
- [9] F. Bancilhon et al, "An amateur's introduction to recursive query processing strategies," in *Proc. ACM SIGMOD*, 1986.
- [10] J. Minker(ed.), *Foundations of Deductive Databases and Logic Programming*, Morgan Kaufmann, Los Altos, 1978.
- [11] H. Gallaire and J. Minker(eds.), *Logic and Databases*, New York, Plenum Press, 1978.
- [12] J. D. Ullman, "Implementation of logical query languages for databases," *ACM Trans. DBS*, no.3, 1985.
- [13] F. Bancilhon, "Object-oriented database systems," in *Proc. Symposium PODB*, 1988.
- [14] 徐洁磐, "演绎数据库的进展", *软件产业*, no.11, 1988.
- [15] 徐洁磐, 魏红, "演绎数据库 EMTL 及其应用", *计算机工程*, no.1, 1989.
- [16] 姚卿达等, "数据模型智能化扩充的研究: 知识数据模型", *计算机科学*, no.3, 1989.
- [17] 蔡吟秀等, "演绎数据库初探", *计算机研究与发展*, No.3, 1988.
- [18] M. Brodie, "未来的信息系统: AI与DB技术的结合", *计算机科学*, no.3, 1989.
- [19] 马立兴等, "数据库与知识库的结合策略", *计算机科学*, no.6, 1988.
- [20] 萨师煊等, "数据库的理论与实践", *计算机科学*, no.1, 1983.
- [21] 萨师煊等, "面向新的应用领域的数据库技术", *计算机科学*, no.3, 1989.
- [22] M. Van Emden and R. Kowalski, "The semantics of predicate logic as a programming language," *J. Assoc. Comput.*, March. 23, 1976.