

# 知识库系统及关系知识库机 系统结构的研究\*

王志英 慈云桂

(计算机系)

**摘要** 文中首先以数据库为基础研究了知识库系统KBS-1,然后基于关系知识模型,研究了一台关系知识库模型机,并以它为硬核研究了知识库系统KBS-2,最后进行了性能分析。

**关键词** 数据库,知识库系统,模型机

**分类号** TP 392

知识库系统是一种用电子化手段对知识进行全面组织、管理和使用的系统。知识库系统具有通用性,可处理不同领域的知识,可提供多种知识表示方式。对知识有较为完善的管理手段,包括一致性、完整性处理等。

本文的研究工作包括:

- (1) 对知识库系统的若干基本问题进行研究。
- (2) 以 VAX 11/780 为宿主机研究知识库系统 KBS-1, 对以数据库为基础建立知识库系统的方法进行研究。
- (3) 研究分析关系知识模型, 主要是其并行性的挖掘, 为硬件支撑系统的设计打下理论基础。
- (4) 研究关系知识库机系统结构, 设计一个新的系统, 达到高效的目的。
- (5) 实现一台关系知识库模型机, 用它来分析各种设计中的问题, 并为新一代计算机的系统结构研究提供模拟环境。在此基础上研究一个以它为硬核的知识库系统 KBS-2。
- (6) 对模型机系统进行性能评价, 从而对系统结构的设计作出结论。

## 1 KBS-1 知识库系统

PROLOG与关系式数据库管理系统的结合从系统结构的角考虑大体上有松耦合、

紧耦合两种。我们以紧耦合方式来进行研究。

知识库系统是知识获取，知识管理和推理机构的有机结合。知识获取部分对新知识按完整性约束要求进行检测后装入库中。知识管理部分负责对存在库中的知识进行操作管理，知识的存储以系统所规定的方式进行。推理机构是知识利用部分，它针对用户提出的各种询问，对知识库中的内容进行推理导出解答。

KBS-1的系统结构框图如图1所示。

系统中数据结构由二类构成，一类是演绎推理所涉及到的数据结构，另一类是库管理所涉及到的数据结构。

知识消化由知识的可证明性检查，矛盾性检查，冗余性检查和独立性检查组成。知识消化系统在元推理和演绎机制支撑下工作，进行上述检查之前，首先要对新输入的知识进行语法上的检查排错。

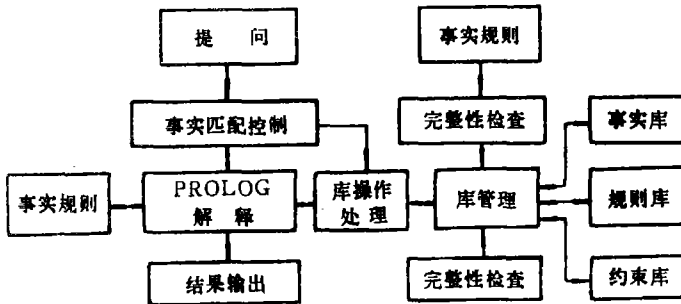


图1 KBS-1系统结构

KBS-1中所有知识获取部分都是以 PROLOG 写出的。消化新知识过程是通过谓词来进行的。它需要很强的元级知识处理能力，但问题是效率低，关键是 demo 谓词在知识库知识推理操作之中，仅有元级库的索引。这样导致了 demo 做了大量的无用操作。我们提出一种新的元知识表示，以过程名为索引、原有的知识表示形式为：

kbnlname (knowledge)

改进后的形式为：

procedure name (kbnlname knowledge)

原来形式下同一库中的所有知识都用指针连成一个长链，每进行一次 demo 操作，将查遍链上的每一个知识项。改进后的形式下，以知识的过程名不同而构成了若干链，这时的 demo 谓词操作只查询有关过程链下的内容。这样使所要查询的内容大为减少，而只涉及到有关部分。如果只存在一个库名时，demo 的执行速度将相当于二级索引的速度。

推理机构是基于GKD-PROLOG来实现的。它运行效率高，我们对它进行了扩充，这主要包括二大部分：元功能扩充和知识操作谓词。知识库管理部分是基于 RDB/VMS实现的。它对系统中的知识进行检索，更新和其它管理。对应于规则，事实和完整性约束三种知识内容，我们在原先数据库基础上建立了事实库，规则库和约束库。

## 2 关系知识模型及其分析

关系知识模型基于一个全新的概念，这就是一致化检索 (Retrieval By Inification)。它在同一层上以宽度优先的方式同时求出全部解，较好地实现了并行操作模型的重要之处在于项结构的引入。常规数据库处理的常量不足以表达目标世界，这导致了在模型所处理的项中引入变量，产生出项关系及有关的一致化关系操作。

关系知识库从项关系中按用户的提问条件进行检索求解，通过反复地将一致化连接和一致化选择操作作用于项关系，求出检索的结果。本文提出一种新的一致化检索环境，这就是基于集成化存放知识的特点，引入过程索引。这使检索操作能迅速定位。

模型具有很高的并行性，包括项关系之间的并行，多个操作之间的并行，项关系内单个操作的并行。我们还从 I/O、调度和基本操作这几个重要方面分析了模型的瓶颈问题。关系知识模型易于实现系统结构上的支持，这是由于模型所需操作种类少，简单且并行性好，并行操作之间必需交换的信息很少，便于调度。关系知识模型中知识是以 horn 子句形式存储的，经过一定转换之后，它可支持用户使用一阶逻辑之外的其它知识表示。

## 3 关系知识库机系统结构及模型机

关系知识库机应当能够充分开发检索过程中的并行性，实现真正的并行处理。这就意味着它是一种多处理机或多计算机系统结构。我们研究了一种关系知识库机系统结构如图 2 所示。

系统中有较大的带宽，磁盘可同时提供数据，系统中数据流是由磁盘的原始数据开

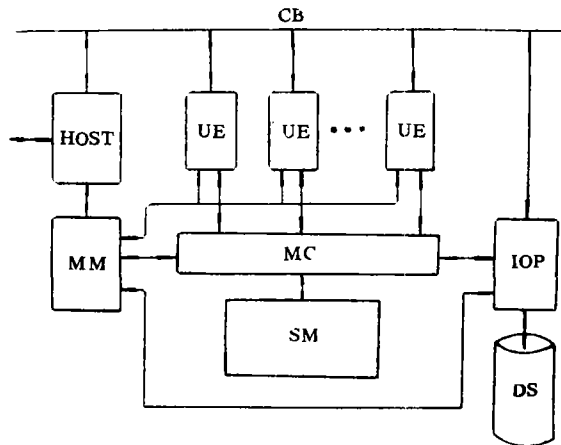


图 2 关系知识库机系统结构

始，送往共享存储器，由共享存储器提供了一致化处理单元的操作数据，操作的结果又送回共享存储器。系统中控制流都是来自于主控制处理机，其它各个部件均是被动的在主控制处理机的命令下进行各种工作。这种命令可以通过控制总线或共享存储器中的信箱来进行。共享存储器为多端口。

系统中有三级存储结构，即共享存储器，局部存储器和通用寄存器。这样一致化的

中间操作内容可以在局部存储器中,有效地减少访问共享存储器的频度。

对应于上述系统结构的模型机物理结构主要部分包括一个控制处理机,四个处理单元,一个共享存储器和一个磁盘系统。

系统中软件配置有两部分,一个是在处理单元中的,主要由监控程序,同步通讯程序,基本的一致化操作程序构成,另一部分是在主控处理机上的核心软件,它分为知识处理,一致化检索和资源管理三层,它们在操作系统的控制下协调地工作。模型机的操作系统是分布式的,它是在86/310微机原配置操作系统rmx86的基础上改造而成的,模型机中通信同步采用了信箱方式,在每个处理单元与主机之间设立一对信箱,一个用于主机发送、处理单元接收,另一个用于处理单元发送、主机接收。主机向处理单元发送命令采用播送方式。结果回收采取的方针是先来先服务。

知识处理是面向用户的顶层,它负责知识的装入,更新和用户的询问处理。输入的新知识将在这里转化成为内部的项关系形式。一致化检索层接受对项关系处理的指令,把一致化检索操作或循环操作转换成逻辑设备指令。资源管理层负责物理设备的分配,各设备的并行执行和通信同步等控制。面向设备的将是逻辑设备指令经确定转换后的物理设备指令。各层软件均以PL/M语言实现。

为具有高效,必须选用恰当的数据结构,关系知识库本身的各知识元组是以树形结构来表示的,它将知识的各部分分解成树上的结点,这样在进行操作时容易迅速查找定位。而各结点上的内容则采用了加标志的数据结构。知识之间是顺序存放的,相同过程名的知识通过排序后连在一起,以一个总的指针进行索引,系统对库中各类不同的知识定位通过索引机构来进行。

性能评价结果表明了关系知识库模型机具有良好的性能,当有4个处理单元时,对于能发挥后端关系库机的长处的查询来说,可望获得比PROLOG运行提高20倍以上的性能。通过全面的分析和评价得出结论:

(1) 关系知识模型具有操作简单的特点,易于实现硬件支持,易于用VLSI来支持。

(2) 关系知识模型具有良好的并行性,可用多个处理单元并发来提高其性能。这种方式实际可行,符合计算机并行处理系统发展的方向。

(3) 关系知识库机运行速度快,效率高。尤其对大量知识存取并且查询所需要推理较少时效果特别明显,是构造未来大型知识处理系统的一种有效方式。

关系知识模型的不足之处在于对某些操作宽度过大的查询,需要对宽度进行一些有效的限制。对这提出二种方式予以解决:一种是建立强有力的存储管理,另一种是以宽度-深度相结合方式推进搜索求解。

# Study of Knowledge Base System and Relational Knowledge Base Machine Architecture

Wang Zhiying Ci Yungui

## Abstract

A knowledge base system KBS-1 is first studied based on a database. Then a relational knowledge base prototype machine is built and another knowledge base system KBS-2 is studied. At last a performance evaluation is given.

**key words:** database, knowledge base system, prototype machine