

多体存贮器有效频带分析

吴涛 陈福接

(计算机系)

摘要 本文以排队论为工具建立了紧耦合多处理机系统存贮器的排队模型。通过对存贮体忙期的分析, 得出了存贮器有效频带的解析表达式, 研究了在保证存贮器有效频带最大的情况下, 处理机的访存请求率 λ 、存贮体的服务率 μ 和存贮体数 m 三者之间的关系。

关键词 存贮器, 排队论, 模型/有效频带, 排队模型

分类号 TP301

计算机已经进入多处理机时代。多处理机系统有紧耦合和松散耦合两种结构。特别是共享主存的紧耦合多处理机系统对存贮器提出了更高的要求。如著名的巨型机系统Cray X—MP, Cray-2等都对存贮器的结构和性能做了很好的设计和优化。存贮器在多处理机系统中占有重要的位置。

在多处理机系统中, 存贮器是由多个存贮体组成的, 系统通过对多个体的并行访问

1989年10月5日收稿

The System Architecture of an Intelligent Machine

Wang Pu Zhang Chenxi Hu Shouren

(Department of Computer Science)

Abstract

This paper puts forward the architecture and organization of a high performance prolog-oriented intelligent machine system, which comprises a conventional computer as its host. A powerful prolog processor supported by a special hardware is its back-end machine. Although the architecture of the system is simple, the performance is very high. Simulation results show that the performance of the system is about 500 KLIPS for logical inference and 17 MIPS for numerical calculations.

Key Words computer system, artificial intelligence, PROLOG/prolog machine, WAM mode, unification operation, dereference operation

来实现数据的高速吞吐。频带是衡量存储器性能的重要指标。存储器的频带分理想频带和有效频带,人们更关心有效频带,因为有效频带客观地反映了存储器的性能。对于不同的作业类型,存储器所反映的有效频带是有差异的,但为了得到存储器在一般情况下的有效频带,通常使用的方法有模拟的方法、标准程序测试方法和分析的方法。本文基于紧耦合多处理机系统的一般模型,用排队论来分析存储器的有效频带。

1 紧耦合多处理机系统的结构

1.1 紧耦合多处理机的结构

在图 1 所示的紧耦合多处理机结构中有 n 个处理机、 m 个存储体和 r 台外围处理机。各处理机通过 $m \times (n+r)$ 互连网共享 m 个存储体。

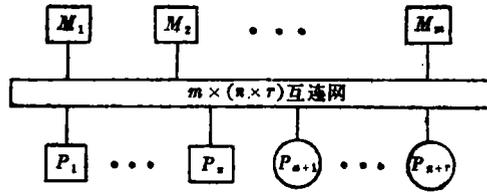


图 1 一般紧耦合多处理机的结构

$P_i (1 \leq i \leq n)$: 第 i 个处理机, $M_j (1 \leq j \leq m)$: 第 j 个存储体, $P_k (n+1 \leq k \leq n+r)$: 第 k 台外围处理机

在图 1 中 $m \times (n+r)$ 互连网是一组交叉开关 (Crossbar), 它使得任一处理机和任一存储体之间都有直接的地址和数据通路。但当处理机的个数和存储体的个数增加时,交叉开关的规模迅速增长。所以在一般实际的系统中,为了避免交叉开关的过大开销,处理机的个数和存储体的个数不能太多。当多个处理机同时访问一个存储体时,会发生体冲突。遇到体冲突,存控根据确定的优先顺序规则来完成冲突请求所要求的读写,在存储体中也可以使用多端口技术或多子体的二维存储结构技术来缓解访问冲突和提高存储体的频带。

1.2 系统频带的概念

定义 1.2.1 频带是指各部件单位时间内处理的字数。

常用 B_m 表示存储器的频带,用 B_p 表示处理机的频带,用 B_d 表示外围处理机 (或设备) 的频带。 B_m 、 B_p 、 B_d 也叫做相应部件的理想频带。

定义 1.2.2 有效频带是指在程序运行过程中各部件实际表现出的频带。用 B_m^* 、 B_p^* 和 B_d^* 分别表示存储器、处理机和外围处理机的有效频带。

根据如上的定义有

$$\begin{cases} B_m \geq B_m^* \\ B_p \geq B_p^* \\ B_d \geq B_d^* \end{cases} \quad (1)$$

在程序的运行过程中,要求系统的频带要平衡,即系统中各部件并行工作,且都经常保持忙碌的状态。在以存储器为中心的系统中,频带平衡的概念要求系统满足 (2) 式:

$$B_m^* = B_p^* + B_d^* \quad (2)$$

(2) 式中的 B_m^* 、 B_p^* 和 B_d^* 在图 1 所示的结构中分别表示 m 个存储体的总有效频带、 n 个处理机的总有效频带和 r 台外围处理机的总有效频带。

2 紧耦合多处理机存储器的排队模型

2.1 系统中的访存流

通过对处理机 $P_i (1 \leq i \leq n+r)$ 的运行分析可知,由于指令缓冲部件的作用,指令中操作数访存的随机性及转移指令的影响,使得一般程序在 P_i 上运行时其访存过程是随机的。人们常用Poisson流来模拟这一过程。为了建立图1所示结构的排队模型,提出如下假设:

假设 2.1.1 $P_i (1 \leq i \leq n+r)$ 在运行过程中发出的访存请求流服从Poisson分布,即,

$$P_k(t) = \frac{(\lambda_i t)^k}{k!} \cdot e^{-\lambda_i t} \quad k=0, 1, 2, \dots, i=1, 2, \dots, n+r$$

假设 2.1.2 $P_i (1 \leq i \leq n+r)$ 在运行过程中发出的访存请求等概率地访问 $M_j (1 \leq j \leq m)$,即 P_i 发出的访存请求以 $\frac{1}{m}$ 的概率访问 M_j 。

假设 2.1.3 $P_i, P_j (1 \leq i, j \leq n+r, i \neq j)$ 发出的访存请求流是相互独立的。

2.2 存储器的排队模型

由如上的假设已知 P_i 发出的访存请求流是一Poisson过程,且它们之间相互独立,那么 $n+r$ 个请求流合起来服从什么分布呢? P_i 发出的访存请求等概率地访问 M_j ,那么到达单个存储体 M_j 的请求流服从什么分布呢?下面的定理回答了如上的问题。

定理 2.2.1 设有 l 个独立的请求源,每个源发出的请求流是抵达率为 λ_k 的Poisson流, $k=1, 2, \dots, l$,把它们合起来可得到一个抵达率为 λ 的Poisson流,其中

$$\lambda = \sum_{k=1}^l \lambda_k$$

证明 源 $i (1 \leq i \leq l)$ 的请求流为

$$P_k(t) = \frac{(\lambda_i t)^k}{k!} \cdot e^{-\lambda_i t} \quad i=1, 2, \dots, l$$

(以下 $P_k(t)$ 简写为 P_k)

其母函数 $G_i(z)$ 为

$$\begin{aligned} G_i(z) &= \sum_{k=0}^{\infty} P_k z^k \\ &= \sum_{k=0}^{\infty} \frac{(\lambda_i t)^k}{k!} \cdot e^{-\lambda_i t} \cdot z^k \\ &= \exp[-\lambda_i t(1-z)] \end{aligned}$$

来自所有独立输入源的总抵达请求数有如下的母函数:

$$\begin{aligned} G(z) &= \prod_{i=1}^l G_i(z) \\ &= \exp[-\lambda_1 t(1-z)] \cdot \exp[-\lambda_2 t(1-z)] \cdots \exp[-\lambda_l t(1-z)] \\ &= \exp[-\lambda t(1-z)] \quad \text{其中, } \lambda = \sum_{i=1}^l \lambda_i \quad (\text{证毕}) \end{aligned}$$

定理2.2.1说明了Poisson流的可选加性。对于图1所示的结构,若 P_i 发出的访存Poisson流的到达率为 λ_i ,则系统总的访存请求流仍为Poisson流,其到达率 $\lambda' = \sum_{i=1}^{n+r} \lambda_i$ 。

定理 2.2.2 设有一个到达率为 λ 的Poisson流,把它分成 q 个支流,每个支流独立地选择,其选择概率为 $\alpha_k (k=1, 2, \dots, q)$,则第 k 个支流是具有到达率为 $\alpha_k \lambda$ 的Poisson过程,其中 $\sum_{k=1}^q \alpha_k = 1$ 。

证明 设 $n(t)$ 表示在 $(0, t)$ 期间输入抵达的顾客数目, $n_k(t)$ 表示在 $(0, t)$ 期间从第 k 条支路输出的顾客数目。已知 $n(t) = n$,则 $n_k(t)$ 的条件联合分布是多项式的分布:

$$\begin{aligned} P\{n_1(t) = n_1, n_2(t) = n_2, \dots, n_q(t) = n_q | n(t) = n\} \\ &= C_{n_1}^{n_1} \alpha_1^{n_1} \cdot C_{n_2}^{n_2} \alpha_2^{n_2} \dots C_{n_q}^{n_q} \alpha_q^{n_q} \\ &= \frac{n!}{n_1! (n - n_1)!} \alpha_1^{n_1} \cdot \frac{(n - n_1)!}{n_2! (n - n_1 - n_2)!} \alpha_2^{n_2} \dots \frac{(n - \sum_{i=1}^{q-1} n_i)!}{n_q! (n - \sum_{i=1}^q n_i)!} \alpha_q^{n_q} \\ &= \frac{n!}{n_1! n_2! \dots n_q!} \alpha_1^{n_1} \alpha_2^{n_2} \dots \alpha_q^{n_q} \end{aligned}$$

乘以随机变量 $n(t) = n$ 的概率,即

$$\begin{aligned} P\{n_1(t) = n_1, n_2(t) = n_2, \dots, n_q(t) = n_q\} \\ &= \frac{n!}{n_1! n_2! \dots n_q!} \alpha_1^{n_1} \alpha_2^{n_2} \dots \alpha_q^{n_q} \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\ &= \prod_{i=1}^q \frac{(\alpha_i \lambda t)^{n_i}}{n_i!} \cdot e^{-\alpha_i \lambda t} \end{aligned}$$

故,每个独立分支的顾客流都为Poisson流,其中到达率为 $\alpha_i \lambda$ 。(证毕)

推论 2.2.1 在图1所示的结构中,到达每个存贮体的访存请求流是Poisson流,即

$$P_k(t) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}$$

其中, $\lambda = \frac{1}{m} \sum_{i=1}^{n+r} \lambda_i$, m 为存贮体的个数, λ_i 为 P_i 发出的访存请求率。

证明 (略)。

存贮体 M_i 对于每个存贮请求的服务时间都相同,也就是说 M_i 服务时间的分布是定长分布。由此可知 M_i 的排队模型是一个 $M/D/1$ 系统,为不失一般性,假设服务时间为独立同分布,即访存请求的服务时间 x_1, x_2, \dots 之间及它们与输入之间均相互独立。设 x_i 的分布函数为 $B(x)$,则

$$P\{x_i \leq x\} = B(x) \quad i = 1, 2, \dots$$

记
$$\bar{x} = \frac{1}{\mu} = \int_0^{\infty} x dB(x)$$

其中 \bar{x} 为访存请求的平均服务时间, μ 为存贮体 M_j 的服务率。因此,可以把 M_j 的排队模型看作 $M/G/1$ 系统。 $M/D/1$ 系统是 $M/G/1$ 系统的特例。

由推论2.2.1可知紧耦合多处理机存贮器的排队模型如图2所示。

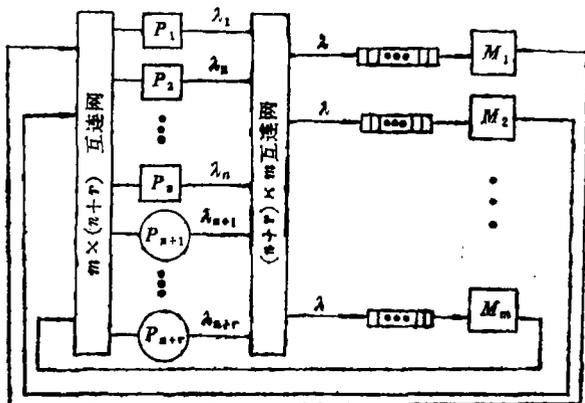


图2 紧耦合多处理机存贮器的排队模型

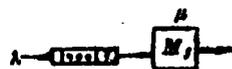


图3 M_j 的排队模型 $M/G/1$

从图2知,紧耦合多处理机存贮器是一种独立的 m 个 $M/G/1$ 并联的排队系统。由于 m 个 $M/G/1$ 系统的输入流和服务时间的分布完全相同,因此下面重点分析其中一个分支 M_j 组成的系统。 M_j 的排队模型如图3所示。

3 M/G/1系统忙期的分析

3.1 系统的转移概率

为了方便地描述 $M/G/1$ 系统的特性,引入如下的参数:

- C_n : 进入系统的第 n 个顾客;
- τ_n : C_n 到达的时刻;
- x_n : C_n 的服务时间;
- q_n : 当 C_n 离开系统时系统中的顾客数;
- ν_n : 在 C_n 服务期间到达的顾客数。

由如上的参数定义知,

$$\begin{cases} q_{n+1} = q_n - 1 + \nu_{n+1} & q_n > 0 \\ q_{n+1} = \nu_{n+1} & q_n = 0 \end{cases}$$

令

$$\Delta_k = \begin{cases} 1 & k=1, 2, \dots \\ 0 & k=0 \end{cases}$$

则

$$q_{n+1} = q_n - \Delta_{q_n} + \nu_{n+1} \tag{3}$$

容易看出, $\{\nu_n\}$ 是一串相互独立、相同分布的随机变量,其分布为

$$\begin{aligned} \alpha_k &= P\{\nu_k = k\} \\ &= \int_0^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} dB(x) \end{aligned} \tag{4}$$

同时显见, v_{n+1} 与 (q_1, q_2, \dots, p_n) 相互独立, 称 $\{q_n\}$ 是随机服务系统 $M/G/1$ 的嵌入马尔可夫链, 其转移概率 P_{ij} 为

$$\begin{aligned}
 P_{ij} &= P\{q_{n+1}=j \mid q_n=i\} \\
 &= P\{q_n - \Delta q_n + v_{n+1}=j \mid q_n=i\} \\
 &= P\{v_{n+1}=j-i + \Delta_i\} \\
 &= \begin{cases} \alpha_{j-i+1} & i \geq 1, j \geq i-1 \\ \alpha_j & i=0, j \geq 0 \\ 0 & \text{其它} \end{cases} \quad (5)
 \end{aligned}$$

3.2 忙期的分析

考察一个存贮体的状态, 无外乎两种情况: 访存请求服务 (处于忙期), 或没有访存请求 (处于闲期)。从时间角度来看, 存贮体总是处于忙期、闲期, 忙期、闲期这样周期地变化。

定义 3.2.1 $U(t) \triangleq$ 系统在 t 时刻没有完成的工作;

$G(y) = P\{Y_n \leq y\} \triangleq$ 忙期的分布

图 4 说明, 有顾客流 C_1, C_2, \dots, C_{11} 进入系统, C_1 启动了一个忙期 Y , 在 C_1 服务期间进入系统的顾客按照 LCFS (Last Come First Served) 规则给予服务并形成相应的子忙期, 在子忙期中进入系统的顾客同样按照 LCFS 给予服务。下面的 BSB (Busy Sub-Busy) 算法描述了图 4 中的顾客流。在 BSB 中用栈 A 和栈 B 来支持两处使用的 LCFS 服务规则。

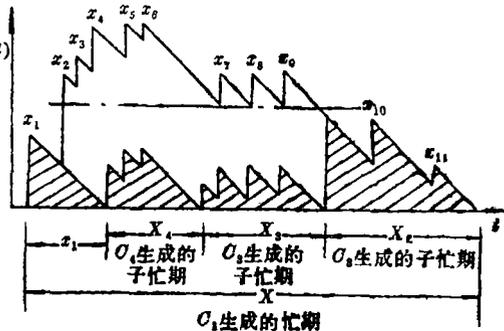
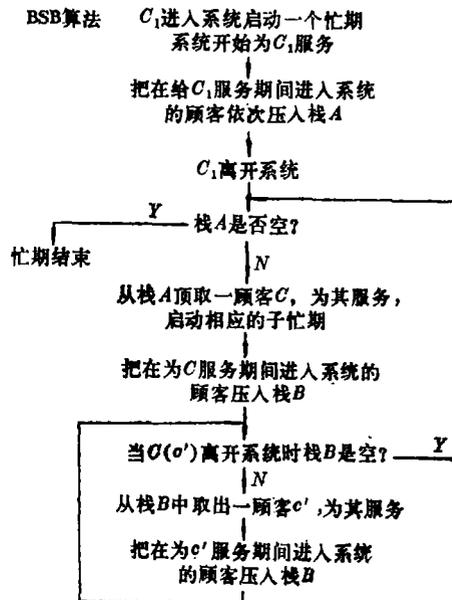


图 4 忙期及LCFS顾客流



由BSB算法知,若 C_1 服务期间有 ν 个顾客进入系统,则

$$Y = x_1 + X_{\nu+1} + X_2 + \dots + X_3 + X_2 \quad (6)$$

$G(y)$ 的拉普拉斯变换为

$$G^*(S) = \int_0^{\infty} e^{-Sy} dG(y) \triangleq E[e^{-SY}]$$

定理 3.2.1 若 $G(y) = P\{Y \leq y\}$, 则,

$$G^*(S) = B^*[S + \lambda - \lambda G^*(S)]$$

证明

$$\begin{aligned} E[e^{-SY} | x_1 = x, \nu = k] \\ &= E[e^{-S(x + X_{k+1} + X_k + \dots + X_3 + X_2)}] \\ &= E[e^{-Sx} \cdot e^{-SX_{k+1}} \dots e^{-SX_3} \cdot e^{-SX_2}] \end{aligned}$$

由于 x 是常数,且子忙期的分布相互独立,子忙期和忙期有相同的分布

$$\begin{aligned} E[e^{-SY} | x_1 = x, \nu = k] \\ &= E[e^{-Sx}] \cdot E[e^{-SX_{k+1}}] \dots E[e^{-SX_3}] \cdot E[e^{-SX_2}] \\ &= e^{-Sx} \cdot [G^*(S)]^k \end{aligned}$$

用概率表示出 $\nu = k$ 的条件,

$$\begin{aligned} E[e^{-SY} | x_1 = x] &= \sum_{k=0}^{\infty} E[e^{-SY} | x_1 = x, \nu = k] \cdot P\{\nu = k\} \\ &= \sum_{k=0}^{\infty} e^{-Sx} [G^*(S)]^k \frac{(\lambda x)^k}{k!} e^{-\lambda x} \\ &= e^{-x[S + \lambda - \lambda G^*(S)]} \end{aligned}$$

同样消去条件 $x_1 = x$,

$$\begin{aligned} G^*(S) &= \int_0^{\infty} e^{-x[S + \lambda - \lambda G^*(S)]} dB(x) \\ &= B^*[S + \lambda - \lambda G^*(S)] \quad (\text{证毕}) \end{aligned}$$

由定理3.2.1知, $G^*(S)$ 是一个隐函数,不易求出它的解析解。但我们往往关心随机变量的各阶矩。设 g_1 为忙期 Y 的一阶矩,由拉普拉斯变换的性质知

$$\begin{aligned} g_1 &= -G^{*'}(0) \\ &= -B^{*'}(0) \frac{d}{dS} [S + \lambda - \lambda G^*(S)] |_{s=0} = 0 \end{aligned}$$

整理得,
$$g_1 = \frac{\bar{x}}{1 - \rho} \quad (\text{其中 } \rho = \lambda \bar{x} = \lambda / \mu) \quad (7)$$

4 存储器的有效频带

4.1 存储器有效频带的表达式

首先来分析一个存储体 M_j 的有效频带。

M_j 的状态总是处于忙期、闲期,忙期、闲期这样周期地变化。设周期的平均值为

T , 若在一个忙期中平均服务的顾客数为 n , 由(7)式知,

$$n = \frac{g_1}{\bar{x}} = \frac{1}{1 - \rho} \quad (8)$$

由于请求流服从 Poisson 分布, 所以请求的平均到达间隔为 $\frac{1}{\lambda}$, 则有

$$T = g_1 + \frac{1}{\lambda} - \bar{x} \quad (9)$$

其中 \bar{x} 为忙期中最后一个顾客的服务时间。

由式(8), (9)知 M_j 的有效频带 B_{jm}^* 为

$$B_{jm}^* = \frac{n}{T} = \frac{1}{g_1 - \frac{1}{\lambda} - \bar{x}}$$

整理得

$$B_{jm}^* = \frac{\lambda}{\rho^2 - \rho + 1}$$

对于图 3 所示的模型, $\lambda = \frac{1}{m} \sum_{i=1}^{n+r} \lambda_i$, $\rho = \frac{1}{m\mu} \sum_{i=1}^{n+r} \lambda_i$

$$\begin{aligned} B_{jm}^* &= \frac{\frac{1}{m} \sum_{i=1}^{n+r} \lambda_i}{\left(\frac{1}{m\mu} \sum_{i=1}^{n+r} \lambda_i\right)^2 - \frac{1}{m\mu} \sum_{i=1}^{n+r} \lambda_i + 1} \\ &= \frac{m\mu^2 \sum_{i=1}^{n+r} \lambda_i}{\left(\sum_{i=1}^{n+r} \lambda_i\right)^2 - m\mu \sum_{i=1}^{n+r} \lambda_i + m^2\mu^2} \end{aligned} \quad (10)$$

具有 m 个体的存贮器的有效频带 B_m^* 为,

$$B_m^* = \frac{m^2\mu^2 \sum_{i=1}^{n+r} \lambda_i}{\left(\sum_{i=1}^{n+r} \lambda_i\right)^2 - m\mu \sum_{i=1}^{n+r} \lambda_i + m^2\mu^2} \quad (11)$$

假如各处理机发生的访存请求率相等, 且为 λ , 则(11)式变为,

$$B_m^* = \frac{m^2\mu^2\lambda(n+r)}{(n+r)^2\lambda^2 - (n+r)m\mu\lambda + m^2\mu^2} \quad (12)$$

(12)式是一个重要的结论, 它给出了多处理机系统中存贮器有效频带的解析表达式, 并说明 B_m^* 是 m 、 λ 和 μ 的函数。

4.2 m 、 λ 和 μ 之间的关系

经分析知, B_m^u 关于 m 、 λ 和 μ 的极值不存在, 但 B_m^u 的最大值存在, 即为理想频带 B_m .

本文所讨论的系统是稳定的, 系统的服务强度 $\rho \leq 1$, 即有

$$(n+r)\lambda \leq m\mu \quad (13)$$

在 $\rho=1$, 即 $(n+r)\lambda = m\mu$ 的临界条件下, $B_m^u = B_m$, 所以把

$$(n+r)\lambda = m\mu \quad (14)$$

称为最大值条件。根据该条件, 任知 m 、 λ 和 μ 的其中两个, 都可以确定出第三个, 并在此情况下存储器取得最大的有效频带。

5 结束语

著名的计算机专家、美籍华人黄凯教授曾用无冲突访问序列的概念分析了多体存储器的有效频带, 得出了有效频带 B_m^u 的一个简单表达式,

$$B_m^u = \frac{B_m}{\sqrt{m}} \quad (15)$$

但人们往往认为, 该表达式的值比存储器实际的有效频带要低。我们用(12)式验证了银河(YH)巨型机、CRAY X-MP等巨型机存储器的有效频带, 验证结果和设计目标很吻合, 其 B_m^u 的计算值比用(15)式的计算值高8%。这说明我们的分析结果更接近于存储器的实际有效频带。

参 考 文 献

- [1] 徐光辉. 随机服务系统. 科学出版社, 1980. 1~100
- [2] Leonard Kleinrock. Queuing System. A Wiley Interscience Publication, New York, 1975
- [3] Takacs L. A Single-Server Queue with Poisson Input. Operations Research, 1962, 10
- [4] Kai Hwang, Faye A. Briggs. Computer Architecture and Parallel Processing. McGraw-Hill, New York, 1984. 13~16
- [5] 吴涛. 巨型机存储器结构的分析. 第三届全国研究生论文研讨会, 长沙, 1988

The Analysis of Utility Bandwidth of Memory with Multibank

Wu Tao Chen Fujie

(Department of Computer Science)

Abstract

Using the queuing theory, the queue model of memory with multibank

一种具有知识评估和知识求精的知识获取(KER)算法*

杨 莉 胡守仁

(计算机系)

摘 要 本文给出了一种获取多类知识的决策树算法,该算法根据所给定的属性的优先级和取值类型进行分类知识的获取。为了保证获得知识的有效性,根据科恩(Cohen)的归纳概率提出了一种证据支持程度来对所获得的知识进行评价,并相应地给出了一种知识求精的方法。

关键词 人工智能,知识获取,算法/决策树,知识评估,知识求精,归纳概率

分类号 TP311.13, TP18

智能系统要求拥有大量专门技术领域的知识,即需要知识获取。如何使智能系统具有知识获取能力呢?我们从人类获取知识的方法中可获得一些启发。人类在生产和科学实践中,往往对大量的事实进行观察记录,通过分析这些分散的甚至粗看起来十分杂乱的观察现象或数据来得出一般的规律,即人类获取知识的过程是一个归纳推理过程。为了使智能系统具有归纳推理以获取知识的能力,就是要设计出种种归纳推理算法,以使智能系统能使用这些算法,从大量的分散数据中总结出规律。由于归纳推理是由已观察到的事件过渡到尚未观察到事件的推理,即:

已观察到 e_1, e_2, \dots, e_n 是 c 和 p

没有观察到任何的 c 不是 p

\therefore 所有的 c 都是 p .

* 此项研究由国家自然科学基金资助
1989年11月5日收稿

in tightly-coupled multiprocessor system is constructed. The analytic expression of memory utility bandwidth is given by analysing the busy period of a bank. The relationship among the rate λ of the accessing memory of processors, the serving rate μ of a bank and the number m of banks is studied under the condition of keeping the memory utility bandwidth maximum.

Key Words memory, queueing theory, model/utility bandwidth, queue model