

# 具有不精确推理能力的信息管理系统

AN INEXACT REASONING INFORMATION MANAGEMENT SYSTEM

汪 庭 飏

Wang Tingkai

(自动化系)

**摘 要** 本文将人工智能中的一些基本思想和方法引入信息管理系统,提出了一种能够根据带可信度的知识进行不精确推理的信息管理系统的设计方案。由该方案设计出的系统不仅具有一定的不精确推理能力,同时还具有适应面广、易变更和扩展等特点。本文介绍了这一设计方案及程序实现技巧。

**关键词** 人工智能;信息管理;数据库/知识库;不精确推理;可信度;产生式系统;知识表示

中国图书资料分类法分类号 TP315; TP316; TP391

**ABSTRACT** The author introduces some basic ideas and methods of artificial intelligence into an information management system, and therefore proposes a design for the information management system which has the ability of inexact reasoning based on the certainty knowledge. According to the design, the system can also be made with the flexibility of change and expansion, and good adaptability for application. This paper describes the design and its programing skill in detail.

**KEY WORDS** artificial intelligence; information management; data libraries; knowledge base; inexact reasoning; certainty factor; production system; knowledge representation

## 一、引 言

自1954年美国商业界首先用计算机进行工资管理以来,计算机在管理方面的应用得到了迅速而广泛的发展。特别是近十年来,以dBASE II/III以及FoxBASE为工具软件开发生的信息管理系统在我国各部门的管理工作中,发挥着越来越重要的作用。然而目前大量开发的各

本研究得到国家自然科学基金资助。

本文于1988年12月9日收到。

种管理软件水平普遍较低。其主要特点是：(1) 其功能主要限于简单的手工业务流程的计算机实现，基本上没有知识处理能力或决策推理能力<sup>[1]</sup>；(2) 各功能软件与数据库结合得太紧，修改数据库的结构或增加新的数据库都必须增加或修改功能软件。这样的软件不仅局限性大，适应面窄，变更、扩展均较为困难，而且极大地限制了管理系统的科学水平。本文将人工智能中的一些基本思想和方法引入信息管理系统，构成了一种基于知识库的信息管理系统的设计方案。用该方案设计出来的系统不仅具有一定的知识处理和决策推理能力，同时又能在变更或扩充数据库时无需修改或增加具体程序，而只需修改相应的知识库和增加相应的数据库，具有适应面广，易变更和扩充等特点，本文将介绍这一设计方案及其程序实现技巧。

## 二、系统的总体结构与知识库

人之所以能解决各种复杂的问题是因为他们具有各种知识，那么一个管理系统要具有一定的推理能力并能适应各种不同的情况，一种较好的办法就是使该管理软件自身具有某种人的智能，也就是必须以人的某种知识作为其工作的基础。因此，在设计信息管理软件时，除考虑一般数据库和管理机构（包括功能软件）外，还要为其设计相应的知识库及相应的推理机制，这样系统的结构如图1所示。

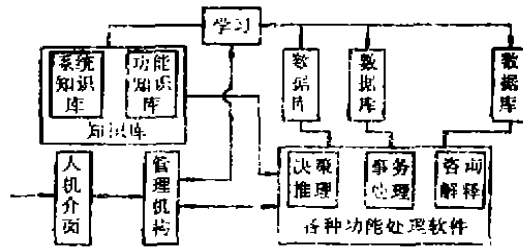


图1 系统结构示意图

该系统在工作时，根据操作者的要求，首先访问知识库，由知识库的系统结构知识可寻找到欲完成该操作所需的数据所在的数据库名、字段名以及操作序列等特别处理事项，据此访问相应的数据库，取出数据并依据操作序列依次访问功能知识库，根据功能知识库中的知识进行相应的推理（或决策），以得到所需的结论。在该系统中，各功能软件可以与数据库的结构完全无关。这样在对数据库的结构进行修改或增加新的数据库以及修改功能软件中的功能等变更时都无需修改或增加相应的功能软件，使得该系统具有适应面广、易变更和扩充等特点。

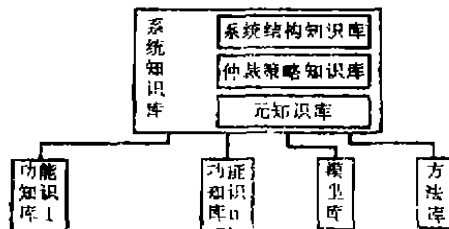


图2 知识库的分层结构

由此可知，知识库是该系统的重要组成部分，为更有效的利用管理知识，在本系统中根据知识的性质和用途分成两个层次的结构<sup>[2]</sup>，如图2所示：(1) 系统管理机构的知识库，该知识库除包含系统结构知识、资源知识及各具体功能知识库的有关信息以外，还包含各种仲裁策略、特殊问题的解决次序等各种较为高层次知识；(2) 各功能知识库，这些知识库主要存放具体政策、经验、事实模型及方法等具体知识。这两个层次的知识构成一个完整的系统知识库整体。

知识的表示在人工智能中已有较多的研究<sup>[3]</sup>，目前常用的方法有：逻辑模式、框架语义

知识的表示在人工智能中已有较多的研究<sup>[3]</sup>，目前常用的方法有：逻辑模式、框架语义

网络,产生式系统,状态空间法,剧本等方法,由于本系统所涉及的知识比较杂乱并且互相之间联系不大,同时还经常需要修改和扩充,因此,本系统采用带有可信度的产生式系统作为其知识表示的形式,产生式系统的基本结构是一个规则库,其中每条规则是一个“条件—结果”产生式,具体形式如下:

$$\text{If } A \text{ Then } B \text{ CF}$$

其中 $A$ 为触发事实, $B$ 为结论, $CF$ 为这条产生式规则的可信度因子。

在系统具体实现时,我们将数据库的前 $n$ 个字段用于存放触发条件,后 $m$ 个字段存放满足该触发条件的结论以及该条件规则的可信度,从而构成了一个规则库。功能知识库主要存放

表 1

条件 1	条件 2	条件 3	条件 4	结论	CF
重点客户	有原料			接恰	0.90
重点客户				接恰	0.40
主要客户	有原料	有生产能力	价格大于等于标准价	接恰	0.80
主要客户	有原料	有生产能力		接恰	0.30
有原料	生产忙		价格大于标准价	接恰	0.40
无原料				接恰	-0.90
生产很忙				接恰	-0.40
生产较忙	原料紧张			接恰	-0.70

功能软件中所涉及的处理知识。该知识库的字段名分别为 $y_1, y_2, \dots, y_n, y_{n+1}, \dots, y_{n+m}$ ,其中 $y_1, \dots, y_n$ 分别存放触发条件, $y_{n+1}, \dots, y_{n+m}$ 分别存放满足该触发条件所得的结论(包括 $CF$ 值)。表 1 为重庆某厂业务管理系统中某功能知识库中的部分知识。系统结构知识库主要存放该系统的结构、资源分配等方面的知识,具体的讲就是存放该系统中所有的知识库、数据库的库名、字段名以及适用范围方面的知识,其具体字段名为: $y_0, y_M, y_N, y_1, y_2, \dots, y_n, y_{n+1}$ , $y_0$ 表示数据库或知识库的编号, $y_M$ 表示数据库或知识库名, $y_N$ 表示该库索引文件名, $y_1$ 表示该库的字段数。 $y_2, \dots, y_n$ 分别表示相应数据库或知识库中字段 $X_i$ ( $i = 2, \dots, n$ )所表示的中文名或意义, $y_{n+1}$ 表示适用范围。

### 三、系统的推理机制

推理是任何智能系统都应具有的一种功能,它是根据一个或一些数据得出另一个判断的思维过程,目前使用的推理方法较多,大致可分为两类:精确推理和不精确推理。精确推理是把领域知识表示成必然的因果关系,逻辑关系,推理的结论是肯定的,这方面的理论有三段论,归纳反演系统和基于规则的演绎系统等。不精确推理是指推理过程中所使用的各种信息和知识是不确定的,不精确的和不完全的,其理论有 $Bayes$ 概率理论、 $Shafer$ 的证据理论、 $Zadch$ 的可能性理论,可信度方法,采用非数值方法的非单调推理模型等<sup>[4]</sup>。在本系统中,系统结构知识库内的知识是一种确定的因果关系,因此这部分推理我们采用了基于规则的演绎系统的方式,而功能知识库的知识往往属于人们的主观判断,是不精确和含糊的,甚至是

不完全的, 用不精确推理较为理想。在众多的不精确推理理论中可信度方法是一种较为简单的方法<sup>[5]</sup>, 其基本思想是: 每个事实和产生式规则都与一个确定性因子(CF)相关, CF是区间 $[-1, 1]$ 中的一个实数, 表示该事实或规则的可信程度,  $-1$ 表示绝对否定,  $+1$ 表示绝对肯定, 由于规则是独立的, 在知识库中是顺序存放的, 所以推理过程中在对规则库的搜索时可能会发生同一证据支持多个结论或者同一结论有多个证据支持的情况。因此, 我们引进了下列定义和公式:

$$CF[H, E] = MB[H, E] - MD[H, E]$$

其中 $CF[H, E]$ 是在证据条件 $E$ 成立时, 结论 $H$ 的可信程度,  $MB[H, E]$ 是证据条件 $E$ 成立时, 确信结论 $H$ 为真被增强的程度,  $MD[H, E]$ 是证据条件 $E$ 成立时, 不确信结论 $H$ 为真被增强的程度。

公式<sup>(1)</sup>

$$\begin{aligned} MB[H, E_1 \& E_2] &= MB[H, E_1] + MB[H, E_2](1 - MB[H, E_1]) \\ MD[H, E_1 \& E_2] &= MD[H, E_1] + MD[H, E_2](1 - MD[H, E_1]) \end{aligned}$$

公式<sup>(2)</sup>

$$\begin{aligned} MB[H_1 \& H_2, E] &= \text{MIN}(MB[H_1, E], MB[H_2, E]) \\ MD[H_1 \& H_2, E] &= \text{MAX}(MD[H, E], MD[H_2, E]) \end{aligned}$$

公式<sup>(3)</sup>

$$\begin{aligned} MB[H_1 \vee H_2, E] &= \text{MAX}(MB[H_1, E], MB[H_2, E]) \\ MD[H_1 \vee H_2, E] &= \text{MIN}(MD[H_1, E], MD[H_2, E]) \end{aligned}$$

公式<sup>(4)</sup>

$$\begin{aligned} MB[H, E_1] &= MB'[H, E_1] \cdot \text{MAX}(0, CF(E_1, E)) \\ MD[H, E_1] &= MD'[H, E_1] \cdot \text{MAX}(0, CF(E_1, E)) \end{aligned}$$

其中 $MB'[H, E_1]$ 为规则中前提证据 $E_1$ 支持结论 $H$ 为真被增强程度,  $MD'[H, E_1]$ 为规则中前提证据 $E_1$ 否定结论 $H$ 为真被增强程度,  $CF[E_1, E]$ 为前提证据 $E$ 中, 证据 $E_1$ 的可信度。

可信度从某种程度上讲是体现了Bayes理论中关于条件概率的概念, 然而又避免了复杂的概率运算。现以下例说明该方法的推理过程。

重庆某厂为某产品是否现在生产 $H$ 进行决策, 对功能知识库搜索支持该产品现在投产 $H$ 的规则有两条(为简明起见)。第一条为现在该厂的生产能力超过实际生产水平, 支持该产品现在生产的 $CF[H, E_1] = 0.4$ , 第二条为原料供应状况支持现在投产的 $CF[H, E_2] = 0.3$ , 计算该产品现在生产的最后可信度 $CF = [H, E_1 \& E_2]$

$$CF[H, E_1 \& E_2] = MB[H, E_1 \& E_2] - MD[H, E_1 \& E_2]$$

据公式<sup>(1)</sup>

$$\begin{aligned} MB[H, E_1 \& E_2] &= MB[H, E_1] + MB[H, E_2](1 - MB[H, E_1]) \\ &= 0.4 + 0.3(1 - 0.4) = 0.58 \\ MD[H, E_1 \& E_2] &= MD[H, E_1] + MD[H, E_2](1 - MD[H, E_1]) \\ &= 0 + 0 = 0 \end{aligned}$$

所以

$$CF[H, E_1 \& E_2] = 0.58 - 0 = 0.58$$

#### 四、数据库的结构及系统的初始化

在具体地实现该系统时,这两方面的问题是值得注意的,其一是数据库的结构,其二是系统的初始化。下面简要介绍这两问题实现技巧。

为了使本系统具有通用和易扩展等特性,我们将数据库的结构规范化,规定各数据库的字段名依次为 $X_1, X_2, \dots, X_n$ ,对不同的数据库 $n$ 可以不同,并指定 $X_1$ 代表该项记录的编号, $X_i (i > 1)$ 为未加定义的形式变量,其代表的意义由系统结构知识库给出。这一规定在程序与数据库分离中起着关键性的作用。

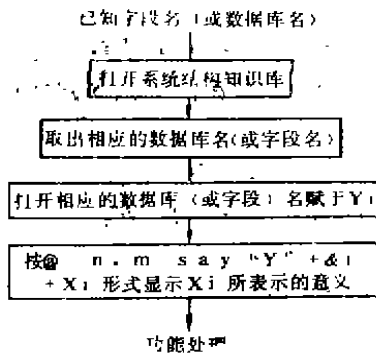


图3 系统初始化示意框图

由于数据库中的字段名 $X_i$ 为未加定义的形式变量,而具体的意义是存放在系统结构知识库之中的,因此,在实现具体的功能处理时必须将数据中的数据与该数据代表的意义(在知识库中)建立一种联系,这一工作我们就称其为系统的初始化。系统的初始化过程中要设置一定的公共变量并通过大量而巧妙地进行宏代换来完成,图3给出了一种初始化流程。系统在初始化的基础上即可进行各种功能处理。

#### 五、结 论

本文介绍的系统将数据库技术和人工智能思想紧密地结合,产生并形成了一种具有一定不精确推理能力的信息管理系统,该系统具有以下几个显著特点:

(1) 建立了以系统知识和功能处理知识为基础的具有分层特点的知识库,并以此作为推理的依据。

(2) 确立了以可信度因子为基础的不精确推理规则,并据此进行推理。

(3) 建立了通用数据库的结构,以及大量而巧妙的使用宏代换命令,使程序和数据库完全分离,给系统性能的扩展及变更带来了极大的便利,提高了系统的适应能力。

笔者将该系统的设计思想应用于重庆某厂的业务管理系统,在该系统中成功地对40~60个不断变更和扩充的数据库和知识库进行查询、统计、维护、工价估算及决策推理等,取得了令人满意的效果。

#### 参 考 文 献

- [1] 吴泉源、钱放,智能化情报自动化系统的实践与研究,全国办公自动化学术会议论文集,1988,9,沈阳
- [2] 汪庭凯、李祖枢,企业计算机管理系统的智能化研究,中国人工智能学会计算机视觉与智能控制学会首届学术年会论文集,1989,4
- [3] 付京孙、蔡自兴、徐光佑,人工智能及其应用,北京,清华大学出版社,1987
- [4] 李凡、徐剑,图象分析专家系统中的不精确推理,信息与控制,1988,(3): 30—36
- [5] 史忠植,知识工程,北京,清华大学出版社,1988