第24卷 第2期
1998年3月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 24, No. 2
March., 1998

研究简报

EMM：一种用于产生式系统的高速匹配算法

张大力 毛祖楫 阎平凡
（清华大学自动化系 北京 100084）

关键词 产生式系统，编码，规则库，推理机

1 引言

产生式系统是专家系统普遍采用的一种结构，其规则集采用“IF……THEN……”的形式. 产生式系统的工作过程是一个推理和搜索的过程，匹配算法极为费时. 改善匹配算法是提高系统性能的关键之一[1-4].

针对产生式系统现有匹配算法的不足，本文从信息编码的角度提出一种全新思路的匹配算法——编码映射匹配法 EMM(Encoding and Mapping for Match)，该算法通过对规则的前提因素编码直接映射规划的动作因素进行规则匹配，极大地提高了匹配速度.

2 编码映射匹配法(EMM)

产生式规则的一般形式为

$$\text{IF}(\text{situation})\text{THEN}(\text{action}), \tag{1}$$

或 $IF(S_1=Z_1 \text{ AND } S_2=Z_2 \text{ AND} \cdots \text{AND } S_m=Z_m)$

$$THEN(a_1=A_1, \cdots\cdots, a_n=A_n), \tag{2}$$

其中 $S_i(i=1,2,\cdots\cdots,m)$称为规则的前提因素，$a_j(j=1,2,\cdots\cdots,n)$称为规则的动作因素.

2.1 EMM 法的规则表示

1) 用有序整数集合对前提因素和动作因素的状态进行映射

假设某规则的前提因素集合和动作因素集合分别用 S 和 A 表示，即

$$S = \{s_1, s_2, \cdots\cdots, s_m\}, \tag{3}$$

$$A = \{a_1, a_2, \cdots\cdots, a_n\}. \tag{4}$$

某前提因素 $S_i(i=1,2,\cdots\cdots,m)$的所有可能状态集合用 U_i 表示，即

$$U_i=\{u_{i1}, u_{i2}, \cdots\cdots, u_{iP_i}\}, \quad (i=1,2,\cdots\cdots,m) \tag{5}$$

其中 P_i 为前提因素 S_i 的所有可能状态数.

收稿日期 1995-12-06

定义一个有序整数集合 M_i

$$M_i=\{0,1,2,\cdots\cdots,P_i-1\},\quad (i=1,2,\cdots\cdots,m),\tag{6}$$

则存在一个映射 f_i

$$f_i:U_i \rightarrow M_i,\quad (i=1,2,\cdots\cdots,m).\tag{7}$$

使得 U_i 中的元素 $u_{i1},u_{i2},\cdots\cdots,u_{iP_i}$ 分别与 M_i 中的元素$0,1,\cdots\cdots,P_i-1$一一对应. M_i 称为前提因素 S_i 的状态编号集合.

同理,对动作因素 a_j 可以定义相应的集合和映射

$$V_j=\{v_{j1},v_{j2},\cdots\cdots,v_{jQ_j}\},\tag{8}$$

$$N_j=\{0,1,\cdots\cdots,Q_j-1\},\tag{9}$$

$$g_j:V_j \rightarrow N_j,\quad (j=1,2,\cdots\cdots,n).\tag{10}$$

2) 用某规则的 m 个前提因素的状态编号进行编码

设(2)式表示的规则的 m 个前提因素的状态分别为 $u_{1K_1},u_{2K_2},\cdots\cdots,u_{mK_m}$,对应的状态编号为 $k_1,k_2,\cdots\cdots,k_m$,则可用编码公式(11)将整数序列$\{k_1,k_2,\cdots\cdots,k_m\}$表示成一个整数 I_k,这里称 I_k 为规则(2)的 m 个前提因素的编码(EC).

$$I_k=\sum_{i=1}^{m}k_i\cdot W_i,\tag{11}$$

其中 $W_i(i=1,2,\cdots\cdots,m)$称为前提因素 S_i 的权,用递推公式(12)来计算 W_i 的值.

$$\begin{cases}W_m=1,\\ W_i=P_{i+1}\cdot W_{i+1}\quad (i=m-1,\cdots\cdots,2,1),\end{cases}\tag{12}$$

其中 P_i 为前提因素 S_i 所有可能的状态数.

3)用一维数组存放所有规则的动作因素的编号.

对应规则(2)的 n 个动作因素定义一维数组 $C_j(j=1,2,\cdots\cdots,n)$,称为动作编号数组,其大小等于 m 个前提因素的组合数:$P_1\times P_2\times\cdots\cdots\times P_m$. 对规则库的每条规则用(11)式计算其前提因素的编码 I_k,将该规则的 n 个动作因素的状态编号分别存放在 n 个动作编号数组 $C_j(j=1,2,\cdots\cdots,n)$的下标为 I_k 的元素 $C_j[I_k]$中.

2.2 EMM 法的规则匹配方法

1) 系统初始化时,用2.1所述方法对规则库中的每条规则计算其前提因素的编码;将其动作因素状态编号分别存放在相应的动作编号数组元素中. 初始化结果存在内存中以备调用.

2) 系统工作过程中,根据不同的前提条件,不断计算前提因素编码 I_k,由 I_k 直接映射 $C_j[I_k]$;根据映射关系(10)式,读取动作状态 V_{jl_j},并执行其所规定的动作.

2.3 对 EMM 法的讨论

由于 EMM 法不存在搜索,极大地缩短了匹配时间和占用的内存空间. 该算法与前提因素、动作因素的具体表达模式无关,即可以是数值表达式、逻辑表达式、模糊表达式或其他语言表达式,因此具有较大的灵活性.

但由于 EMM 法采用直接映射方法,丢失求解过程,不利于解释功能和人机交互;另外,由于该方法不存在推理过程,因而没有自学习能力.

3 应用实例

将EMM法用于合成氨过程中的氢氮比智能控制系统中，得到了快速匹配的效果. 为了与穷尽搜索法、分级搜索法进行比较，这里给出各种方法搜索一条规则所用的平均时间如表1所示.

表1

匹配方法	平均匹配时间
穷尽搜索法	约1 210μs
分级搜索法	约110μs
EMM 法	约8. 27μs

参 考 文 献

1 McDermott J, Newell A, Moore J. The efficincy of certain production system implementations. In: Waterman. D A and Hayes-Roth F(Eds), Pattern-directed inference systems, New York: Academic Press, 1978, 155—176

2 张大力，马革新等. 实时模糊神经网络及其在过程智能控制中的应用. 清华大学学报，1995，**35**(S3)：106—111

3 梁军，杨永耀，吕勇哉. DMES：一个实时控制专家系统. 信息与控制，1993，**22**(2)：65—70

4 刘亚彬，孙优贤，周春辉. 专家系统在大型工业过程中的应用. 信息与控制，1990，**19**(1)：47—49

EMM: A HIGH SPEED MATCH ALGORITHM FOR PRODUCTION SYSTEM

ZHANG DALI MAO ZUJI YAN PINGFAN

(*Department of Automation, Tsinghua Unviersity, Beijing* 100084)

Key words production system, coding, rule base, inference machine