多段规则推理模型

吴伟民　李吉桂　李冠英

（华南师范大学计算机科学系）

提要

本文将[1]提出的统一框架推广到多段规则的情况，并把这种多段规则推理归结为关系运算，从而提高了效率。文中还探讨了多段规则推理的性质及其与现行处理方法的关系。

一、问题的提起

在现实系统中，常存在以下我们称之为“多段规则”的情况：

首先，前件X存在l种状态，即有

$$X=(x_1\ \ x_2\ \ \cdots x_l)$$

例如，断言X（温度）可分为高、中、低等三种状态，即有

$$X=(x_{高}\ \ x_{中}\ \ x_{低})$$

其次，对应地，后件A也存在k种状态，即

$$A=(a_1\ \ a_2\cdots\ \ a_k)$$

例如，断言A（颜色）可以分为白、浅红、红、暗红等四种状态，即有

$$A=(a_1\ \ a_2\ \ a_3\ \ a_4)$$

所谓“多段规则”$X\to A$就是给出了前件断言X与后件断言A的关系。现行的处理方法是将其分解为$l\times k$条普通规则，根据专家的经验，给出每条规则$x_i\to a_j$的确定性因子d_{ij}（$i=1,\ 2,\ \cdots,\ l$；$j=1,\ 2\cdots k$），然后按照通常的方法（见[1]）进行综合。易知，特别是在多级推理的场合，这种处理方法的效率是很低的。本文给出一种综合方法，即不必将其分解为$l\times k$条规则，而是将“多段规则”的确定性因子表示为一种“关系”，从而将推理过程归结为关系的运算。其结果，对于最低层断言，则便于知识的成块表示；对于其它层次的断言，则不但提高了推理的效率，而且更客观地反映了人类对于客观过程的认识。

二、多段规则的确定性矩阵及其性质

定义1　由规则$x_i\to a_j$（$i=1,\ 2,\ \cdots l$；$j=1,2\cdots k$）的确定性因子$CF[a_j,\ x_i]=a_{ij}$组成的矩阵称为多段规则X→A的确定性矩阵，记为

$$CF[A,\ X]=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & & & \cdots \\ a_{l1} & a_{l2} & \cdots & a_{lk} \end{vmatrix} \qquad (1)$$

其中$X=(x_1\ \ x_2\cdots\ \ x_l)$，$A=(a_1\ \ a_2\cdots\ \ a_k)$，为了方便，并且规定$\sum_{j=1}^{k}a_{ij}=1$。

定义2　最低层次的断言的确定性矩阵（向量）由专家给出。其它层次断言的确定性矩阵由下式定义

$$CF[A]=CF[X]\circ CF[A,\ X] \qquad (2)$$

易知，当$l=k=1$且（2）式的运算“$\circ$”为普通乘法运算时，（2）式变为[1]的（2）式。由（2）式可知，所谓“多段规则”推理，实质是一种关系运算。这种处理方法的另一方便之处是可以根据需要及可能而灵活地定义运算“$\circ$”的。

例　设已知断言$X=(x_1\ x_2\ x_3)$的确定性矩阵为$CF[X]=(0.5\ 0.4\ 0.1)$，多段规则X→A和A→B的确定性矩阵分别为

$$CF[A,\ X]=\begin{vmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0 & 0.3 & 0.7 \end{vmatrix}$$

$$CF[B,\ A]=\begin{vmatrix} 0.1 & 0.2 & 0.6 & 0.1 \\ 0.2 & 0.7 & 0.1 & 0 \\ 0 & 0.1 & 0.3 & 0.6 \end{vmatrix}$$

且设运算"○"为普通矩阵乘法。

由(2)式可推知断言A的确定性矩阵为

$$CF[A]=CF[X]\bigcirc CF[A, X]$$
$$=(0.43\quad 0.37\quad 0.20)$$

同理，CF[B]的确定性矩阵为

$$CF[B]=CF[A]\bigcirc CF[B, A]$$
$$=(0.117\quad 0.365\quad 0.355\quad 0.163)$$

还应指出，上例将(2)式的运算"○"规定为常规矩阵运算，实质是将确定性因子a_{ij}看作一种"权"，而运算结果则反映了"求加权和"的特性。

此外，由于常规矩阵乘法满足结合律，所以(2)式便利了多段推理。例如$X\to A\to B$可以表示为

$$CF[B]=CF[A]\bigcirc CF[B, A]$$
$$=(CF[X]\bigcirc CF[A, X])\bigcirc CF[B, A]$$
$$=CF[X]\bigcirc(CF[A, X]\bigcirc CF[B, A])$$

即等效于确定性矩阵为$CF[A, X]\bigcirc CF[B, A]$的多段规则$X\to B$的一级推理。

现行系统中，为了提高推理速度而希望进行并行处理。如果将规则按定义1进行划分，则按定义2的(2)式进行运算本身就是一种并行算法，而且只须进行"划分"的匹配(选择多段规则)。因此，(2)式的引入对于提高一类基于规则系统的推理速度和"知识求精"都是有效的。

三、多段规则推理的综合

如上所述，多段规则推理结果的确定性因子是一个矩阵(向量)。根据预先规定的阈值，可以确定入选分量(全都不能入选)。以下把[1]的几种典型综合方法推广到多段规则的推理。

1、AND型

两个断言AND结合，其结果的确定性矩阵（向量）定义为

$$CF_P=CF[X\ AND\ Y]=\min(CF[X],CF[Y])\qquad(3)$$

如果$CF[X]=(a_1, a_2\cdots, a_m)$且$CF[Y]=(b_1, b_2\cdots, b_n)$，则$CF_P$是一个$m\cdot n$元向量，它的第$(i-1)n+j$个分量定义为

$$c_{(i-1)n+j}=CF[x_i\ AND\ y_j]=\min(a_i, b_j)\qquad(4)$$

其中，$i=1, 2, \cdots, m$且$j=1, 2, \cdots, n$。

2、OR型

两个断言OR结合，其结果的确定性矩阵(向量)定义为

$$CF_P=CF[X\ OR\ Y]=\max(CF[X], CF[Y])\qquad(5)$$

即CF_P也是一个$m\cdot n$元向量，它的第$(i-1)n+j$个分量定义为

$$c_{(i-1)n+j}=CF[x_i\ OR\ y_j]=\max(a_i, b_j)\qquad(6)$$

易知，上述定义在确定性矩阵退化为标量时，与[1]的定义完全一致。

3、COMB型

设有n支推理路径推出同一结果断言，其确定性向量分别为$(C_{i1}\ C_{i2}\cdots C_{im})$，$i=1, 2, \cdots, n$，需要进行综合。如[1]所述，对于非多段规则，两支以上推理路径的结果综合，一般都归结为两两综合，逐步导致最终结果。这不但存在效率问题，而且需要作无序假设。这里对多段规则给出一种效率高，且形式上和(2)式相似的综合方法。我们认为，这种方法将更客观地反映人们的推理过程。

定义3 n支推理路径推理结果的综合定义为

$$CF[C]=(l_1\ l_2\ \cdots\ l_n)\bigcirc C\qquad(7)$$

其中矩阵C由各支路径的结果确定性向量组成，即有

$$C=\begin{vmatrix}C_{11} & C_{12}\cdots C_{1m}\\ C_{21} & C_{22}\cdots C_{2m}\\ \cdots & \cdots\\ C_{n1} & C_{n2}\cdots C_{nm}\end{vmatrix}$$

向量$L=(l_1 l_2\cdots l_n)$称为"权重向量"，它反映了各支路径的重要程度。规定$0\leqslant l_i\leqslant 1$，$i=1,2,\cdots,n$。

由以上假设易知$C_{ij}\geqslant 0$。当(7)式的运算"○"为通常的矩阵乘法，且$l_i=1(i=1,2,\cdots,n)$时，可得

$$C_j=\sum_{i=1}^{n}C_{ij}\qquad(8)$$

其中C_j为CF[C]的第j个分量。

如果设由(8)式和[1]的(8)式进行综合的结果分别为$CF_多$和CF_{MYCIN}，则有下述不等式

$$CF_{MYCIN}\leqslant CF_多\leqslant 1\qquad(9)$$

在具有这种多段规则的多级推理系统中，如果每一支路径由一台处理机来完成，则(7)式便可看作是对这种并行处理的综合，且可以灵活地给定各支路的权。因此，定义3也便于并行处理。

应该指出，在(1)式中$\sum_{j=1}^{k}a_{ij}=1$的规定，在知识收集时，可能会给专家带来某些不便。如果在收集过程中不管这一规定，最后再按下式进行规格化，则可解决这一问题。

$$a_{ij}=a'_{ij}/\sum_{j=1}^{k}a'_{ij}\qquad(10)$$

其中a'_{ij}为知识收集时得到的值。

如[1]所述，在确定性理论中，规定确定性因子的值不大于1。本文在前面规定$\sum_{j=1}^{k}a_{ij}=1$的一个

(下转第17页)

$$(\sum_{i=1}^{n}T_i+I)\to\sum_{i=1}^{n}T_i'+T_0+I'+E \qquad (4)$$

(4) 式是调节操作

$$(\sum_{i=1}^{n}T_i+I)\to\sum_{i=1}^{n}T_i'+T_0+E_1$$

与同化操作

$$(T_0+I)\to I'+E_2$$

的合写形式。其中，I是短时记忆中的试用概念；T是长时记忆中需要调整的已掌握概念，它满足条件δ (body(T)) $\not\subseteq$I；T_i 是长时记忆中的已掌握概念，它满足条件δ(body(T_i)) $\subseteq$I；T′是对已掌握概念T调整的结果，它代替了长时记忆中T的内容，并满足条件δ (body (T′)) $\subseteq$I；T_i′是根据I对T_i施行调整的结果，满足条件δ(body (T_i′)) $\not\subseteq$I；I′是对长时记忆调节后，系统对I实施同化操作的结果；T_0是长时记忆中根据试用概念I构造的新概念，它满足条件δ(body(T_0)) $\subseteq$I；E是有关拟合程度的启发式信息。

参考资料

1. Michalski, R. S.: A Theory and Methodology of Inductive Learning (见Michalski, R.S. Carbonell, J.G. Mitchell, T.M. 主编：Machine Learning: An Artificial Intelligence Approach Tioga Publishing Co.) p.83—134. 1983
2. Simon, H.A著 荆其诚 张厚粲译：人类的认知——思维的信息加工理论 科学出版社 1986
3. Piaget, J. Inhelder, B.著 吴福元译：儿童心理学 商务印书馆 1980
4. Piaget, J.著：皮亚杰的理论
5. 宗智贤：皮亚杰儿童思维心理学简介 北京师范大学学报（社会科学通讯）№.1 p.71—82, 1980
6. 李彦，机器的认知学习 全国高教系统第十次AI学术讨论会，全国第二次自动推理研讨会论文 陕西 1989.9

（上接第19页）

主要原因是为了满足这一要求。然而，这可能会引起另一问题，即在多级推理时，可能会使确定性因子的值递减过快，难以定出合理的取舍阈值。为了解决这一问题，可以改变(2)中的运算“○”.例如：

设$0\leqslant a_{ij}\leqslant 1$，规定运算“○”的定义为：首先按普通矩阵乘法进行计算，然后将其结果压缩到[0, 1]区间，即有

$$CF[A]_i=\min(1,\ CF[A]^{(2)}_i) \qquad (11)$$

其中$CF[A]^{(2)}_i$为利用普通矩阵乘法进行计算的结果。易知，这样的运算不但可使确定性因子不会递减过快，而且仍然满足乘法结合律。

同一推理系统，采用不同方法进行综合，所得结果不同，这和人们从不同的角度观察同一事物，可能得出不同结论一样。我们建议使用不同方法进行试验，然后设法让系统进行“自学”，从而不断提高其精度。

四、结束语

产生式规则过于精细的结构直接影响了大型知识库的知识检索、推理和维护的效率。多段规则实质上是根据知识的自然结构集成的阵列规则。引入这种块状结构，可以提高知识库的性能指标。这种集成化的规则表示也可以作为知识整理的工具，指导成块知识的获取、表示和存贮，有利于知识获取的系统化和完整化，以及全面综合考虑确定性因子（矩阵）的设定。多段规则的确定性矩阵集中表示了规则推理中的处理对象，可以直接利用现有的数值处理系统进行有效的推理综合，容易实现并行处理。现有的各种推理综合方法各有适用范围。我们可以根据知识的内在联系，对每个多段规则指定具体的综合方法，实现多种综合方法的混合使用。

参考文献

[1] 李冠英，李吉桂，确定性理论及主观Bayes方法的一个统一框架，华南师范大学学报（自然科学报），1988年第1期
[2] Frederick Hayes-Roth, Rule-based Systems, CACM, vol.28, №.9, 1985
[3] Richard Fikes, Tom Kehler, The Role of Frame-based Representation in Reasoning, CACM, vol.28, №.9, 1985
[4] 游子墨，赵一鸣，产生式系统的划分和并行处理，计算机科学，1988年第1期