计算机科学1996 Vol.23№.1

关于BMI模型中规则可分解性的研究

邱玉辉 (西南师范大学 重庆630715)

罗旭东 (国家智能计算机研究开发中心)(北京轻工业学院)

蔡经球 (厦门大学)

摘 要 对于我们在文[1]中所建立的基于区间估计的BMI模型及MYCIN确定性因子模型，本文说明了：1)规则是不可分解的；2)顺序组合运算对平行组合运算不满足分配律，因而使用时要谨慎小心。

关键词 区间估计，确定性因子，规则，可分解性，分布式专家系统。

一、引 言

我们在文[1]中提出了一种新的基于区间估计的不精确推理模型BMI。BMI模型是从MYCIN的不精确推理模型[2]出发，通过线性变换和Fuzzy数学的扩展原理[3]推导出来的。这个模型与文[4]模型一样能够表示"不知道"的信息，但又继承了MYCIN的模型能够很好区分证据的出现导致的是对假设的信任还是怀疑的优点，并且还能够退化为点估计的MYCIN模型。文[2]分析了基于区间估计的不精确推理模型应满足的直观约束，并说了BMI模型基本上满足这些约束。

本文将进一步探讨BMI模型中规则的可分解性，这个问题在MYCIN的确定性因子模型中也存在。我们的研究结果提示AI工作者在运用这些模型时要谨慎小心。

二、BMI模型简介

为了讨论方便，这里我们首先简单介绍一下BMI模型。

1.不确定性表示

设E为一断言，其不确定性估计为

$$CF(E)=[l_E,u_E]\subseteq[0,1] \qquad (1)$$

这里l_E和u_E分别表示在观察背景之下，对断言E的可信度的悲观估计和乐观估计。

规则E→H的强度为一对值(f(H,E),f(H,～E))，它们分别为：

$$f(H,E)=[f_{l_E}(H,E),f_{u_E}(H,E)]\subseteq[0,1] \qquad (2)$$

$$f(H,\sim E)=[f_{l_E}(H,\sim E),f_{u_E}(H,\sim E)]\subseteq[0,1] \qquad (3)$$

这里，f(H,E)的两个分量含义如下：$f_{l_E}(H,E)>0.5$时，它表示证据E的出现导致H的可信程度的最悲观估计，而$f_{u_E}(H,E)$则表示最乐观的估计；$f_{u_E}(H,E)<0.5$时，它表示证据E的出现导致H的被怀疑程度的最低限度的估计，而$f_{l_E}(H,E)$则是最大限度的估计。类似地，可以解释f(H,～E)。

上面用于估计不确定性的区间，不妨令为[a,b]，也应有一定的要求，即要么被[0,0.5]所包含，要么被[0.5,1]所包含。换言之，对于不确定性的估计，表现出某种趋向性：要么趋于信任，要么趋于怀疑。不表示出这样的趋向性的不确定估计是没有意义的。这类似于投票模型，投赞成票的人数或反对票的人数，要达到或超过半数才有效一样。此外，b-a显然便是不知道的成分，即不知道的程度，相当于投弃权票人数的百分比。

2.不确定性的传播算法

(1)顺序传播 设规则E→H的强度的两个分量各为：

$$\begin{aligned} f(H,E)&=[f_{l_E}(H,E),f_{u_E}(H,E)],\\ f(H,\sim E)&=[f_{l_E}(H,\sim E),f_{u_E}(H,\sim E)] \end{aligned} \qquad (4)$$

又已知 $CF(E)=[l_E,u_E]$，则

$$CF(H)=\begin{cases} [(f_{l_E}(H,E))-0.5(2l_E-1)+0.5,(f_{u_E}(H,E)-0.5)(2u_E-1)+0.5] \\ \quad 当\ CF(E)\subseteq[0.5,1],CF(H,E)\subseteq[0.5,1] \\ [(f_{l_E}(H,E)-0.5)(2u_E-1)+0.5,(f_{u_E}(H,E)-0.5)(2l_E-1)+0.5] \\ \quad 当\ CF(E)\subseteq[0.5,1],CF(H,E))\subseteq[0,0.5] \\ [(0.5-f_{l_E}(H,\sim E)(2l_E-1)+0.5,(0.5-f_{u_E}(H,\sim E)(2u_E-1)+0.5] \\ \quad 当\ CF(E)\subseteq[0,0.5],f(H,\sim E)\subseteq[0,0.5] \\ [(0.5-f_{l_E}(H,\sim E))(2u_E-1)+0.5,(0.5-f_{u_E}(H,\sim E)(2l_E-1)+0.5] \\ \quad 当\ CF(E)\subseteq[0,0.5],f(H,\sim E)\subseteq[0.5,1] \end{cases} \qquad (5)$$

(2)平行传播 设

$CF_1(H)=[l_H^1,u_H^1]$，$CF_2(H)=[l_H^2,u_H^2]$，则

CF(H)=

$$
\begin{cases}
[2(l_H^1+l_H^2-l_H^1 l_H^2)-1, \\
\quad 2(u_H^1+u_H^2-u_H^1 u_H^2)-1] \\
\quad 当\ CF_1(H)\subseteq[0.5,1], CF_2(H)\subseteq[0.5,1] \\
[2l_H^1 l_H^2, 2u_H^1 u_H^2] \\
\quad 当\ CF_1(H), CF_2(H)\subseteq[0,0.5] \\
[l_H^1+l_H^2-0.5, u_H^1+u_H^2-0.5] \\
\quad 当\ CF_1(H)\subseteq[0.5,1], CF_2(H)\subseteq[0,0.5] \\
\quad\quad 或\ CF_1(H)\subseteq[0,0.5], CF_2(H)\subseteq[0.5,1]
\end{cases}
\tag{6}
$$

(3)前提的逻辑组合　设 $CF(E_1)=[l_{E_1}, u_{E_1}]$, $CF(E_2)=[l_{E_2}, u_{E_2}]$, $CF(E)=[l_E, u_E]$，则

$$CF(E_1 \wedge E_2)=[\min\{l_{E_1}, l_{E_2}\}, \min\{u_{E_1}, u_{E_2}\}] \tag{7}$$

$$CF(E_2 \vee E_2)=[\max\{l_{E_1}, l_{E_2}\}, \max\{u_{E_1}, u_{E_2}\}] \tag{8}$$

$$CF(\sim E)=[1-u_E, 1-l_E] \tag{9}$$

三、规则的可分解性

在用 PROLOG 语言实现专家系统时，人们常将如下的规则 $E_1 \vee E_2 \xrightarrow{RS} H$ 分解成两条规则：

$$E_1 \xrightarrow{RS} H, \quad E_2 \xrightarrow{RS} H$$

这里 RS 是规则的强度。此时要保证专家系统在分解后运行结果等价，必须使得：

$$CF(E_1 \vee E_2) ** RS = (CF(E_1) ** RS) ++ (CF(E_2) ** RS) \tag{10}$$

成立。这里，$**$ 和 $++$ 分别表示公式(5)和(6)表示的运算。我们称(10)为规则的可分解律。

命题1　规则的可分解律在 BMI 模型中不成立。

事实上，令 $CF(E_1)=[0.3, 0.4]$, $CF(E_2)=[0.7, 0.8]$，规则 $E_1 \vee E_2 \to H$ 的强度 $RS=([0.8, 0.9], [0.5, 0.8])$，于是

$$CF(E_1 \vee E_2) ** RS = [0.7, 0.8] ** RS = [0.62, 0.74]$$

但

$$(CF(E_1) ** RS) ++ (CF(E_2) ** RS) = [0.62, 0.8]$$

可见不满足(10)，所以规则的可分解不成立。

由上可见，在用 PROLOG 语言实现专家系统时，对于我们的模型 BMI，不允许进行上述形式的规则分解。

BMI 模型是由 MYCIN 确定因子模型推导而来，那么，我们要问，在 MYCIN 确定因子模型中是否允许上述的规则分解呢？也不能。我们来看一个简单的例子。设 $CF(E_1)=0.3$, $CF(E_2)=0.7$ $RS=0.8$，显然 $\max\{CF(E_1), CF(E_2)\}\cdot RS=\max\{0.3, 0.7\}\times 0.8=0.56$, $CF(E_1)\cdot RS+CF(E_2)\cdot RS-CF(E_1)\cdot RS\cdot CF(E_2)\cdot RS=0.6656$，即可分解律不成立。因此，采用 PROLOG 语言实现专家系统时，不能随意对 MYCIN 确定因子模型使用分解规则。

四、基于 BMI 模型的分布式专家系统的分解性

在两个结点 ES_1 和 ES_2 构成的分布式专家系统中，假定 ES_1 的知识库中有三条规则

$$E_1 \xrightarrow{RS_1} H_1, E_2 \xrightarrow{RS_2} H_1, H_1 \xrightarrow{RS_3} H$$

如果 ES_1 能利用 $E_2 \to H_1$ 的信息，依赖于另一个结点专家系统 ES_2，那么在已知信息足以利用 $E_1 \to H_1$ 推得 H_1 时，显然，ES_1 无需等待 ES_2 给出可利用 $E_2 \to H_1$ 的信息而继续进行推理，利用 $H_1 \to H$ 推得 H。当然也有可能 ES_1 所获得的信息足以同时利用 $E_1 \to H_1$ 和 $E_2 \to H_1$，然后再利用 $H_1 \to H$ 推得 H。此时当然应该要求 ES_1 所用的不精确推理模型能保证上述两种不同的方式得出的 H 之确定性估计应相等。在 BMI 模型中，即要求：

$$(CF_1(H_1) ** RS_3) ++ (CF_2(H_1) ** RS_3) = (CF_1(H_1) ++ CF_2(H_2)) ** RS_3 \tag{11}$$

成立。我们称它为 $**$ 对 $++$ 的分配律。

命题2　$**$ 对 $++$ 的分配律在 BMI 模型中不成立。

事实上，令 $CF_1(H_1)=[0.7, 0.8]$, $CF_2(H_1)=[0.9, 1]$, $RS_3=([0.8, 0.9], [0.5, 0.5])$，则 $(CF_1(H_1) ** RS_3) ++ (CF_2(H_1) ** RS_3)=[0.8, 0.95]$。可见 $**$ 对 $++$ 的分配律不成立。

上述命题说明，在分布式专家系统中，若某一结点专家系统的知识库中有如下的三条规则：

$$E_1 \xrightarrow{RS_1} H_1, E_2 \xrightarrow{RS_2} H_1, H_1 \xrightarrow{RS} H$$

且该结点利用 $E_2 \to H_1$ 的信息依赖于其它结点，而获得可利用 $E \to H_1$ 的信息往往在获得可利用 $E_2 \to H_1$ 的信息之前时，这个结点的专家系统一般是不能使用我们的不确定推理模型 BMI。

同样地，我们会向 MYCIN 确定因子模型适合于分布式专家系统吗？对此，我们有

命题3　顺序组合运算 $**$ 对平行组合运算 $++$ 的分配律在 MYCIN 确定性因子模型中不成立。

事实上，令 $CF_1(H_1)=0.7$, $CF_2(H_1)=0.8$, $RS_3=0.9$，则

$$(CF_1(H_1) ** RS_3) ++ (CF_2(H_1) ** RS_3) = 0.8964$$

另一方面，

$$(CF_1(H_1) ++ CF_2(H_1)) ** RS_3 = 0.846$$

可见，$**$ 对 $++$ 的分配律不成立。因此，在这一原理下，MYCIN 确定性因子模型不适合于分布式专家系统。

（下转第79页）

法甚至会在某些情况下失效。跳出局部极小区的一般做法是：增加隐单元数，或改变学习速率，或同时从多个初始点开始学习。这些做法需试凑多次才能成功，而每次学习时间又很长，致使训练很困难。王小同等(1994)提出了一种避免局部极小问题的方法。这一方法从寻找全局极小点的思路出发，将全局优化方法运用于前向网络学习算法，只需在原来学习算法中加入一个由全局优化方法形成的初值点选择模块，以选择好初始权值，从而自动地避免了局部极小问题的发生。他们分别采用了隧道效应法、填充函数法、测度论法三种全局优化方法，经实验验证，认为测度论法在求解 BP 算法全局极小值问题时更为有效。为了求解全局极小问题，可以采用一些优化理论的方法，如卡尔曼滤波、同伦优化等。S. Singhal 等(1989)利用最优估计理论中的卡尔曼方法，把 BP 算法的网络权值作为滤波的状态变量，从而利用推广卡尔曼滤波来实现非线性网络的学习，不仅避免了局部极值，而且大大提高了学习速度。J. Chow 等(1991)将一 BP 网络误差函数最小化问题转化为一非线性代数方程的求解问题，然后将连续同伦思想用于非线性代数方程的求解，建立了相应的同伦 BP 网络理论和学习算法。同伦 BP 算法不但是大范围收敛的，同时具有良好的收敛速度和可克服病态能力，在梯度法不收敛时，它仍能给出满意的结果。

5 BP 网络研究的有关问题

BP 网络理论已经引起了许多领域科学家的兴趣和关注，但它还处在迅速发展中，因此，还有许多工作要我们去做。概括起来，至少对下述问题的研究是有意义的：

——开发更为简便的具有全局最小和快速运算的 BP 算法。

——BP 网络中隐单元层数、隐单元个数的选取需要进一步严密的理论上的指导。

——自适应拓扑结构和 BP 算法更具有适应性。

——开发可用于 VISI 的 BP 网络结构与算法。

——BP 网络的研究不仅其本身正在向综合性发展，而且愈来愈与其它相关学科密切结合起来，发展出性能更强的结构。

人工神经网络的研究经历了七十年代低潮期以后，进入八十年代开始复苏，并掀起了第二次研究热潮。神经网络有两个与传统方法进行信息处理完全不同的性能：第一，神经网络是自适应和可以被训练的，它有自修改能力。第二，神经网络结构本身就决定了它是大规模并行机制，就是说神经网络从原理上就比传统方法快得多。因此，它的研究与应用已成为科学技术研究中的又一新的热点。BP 网络做为人工神经网络中最基本的和使用最广泛的网络，不仅在理论研究上日臻成熟，其应用也取得了令人鼓舞的进展。如家用电器、故障诊断、模式识别、图像处理、工业控制、专家系统、管理系统、运输系统、财政金融、VLSI 芯片等，已广泛采用了 BP 网络技术。(参考文献共15篇略)

(上接第81页)

五、结论

本文对 BMI 模型有关的两个问题进行了讨论，我们的结论是

(1)运用 PROLOG 实现基于 BMI 模型的专家系统时，形如 $E_1 \vee E_2 \xrightarrow{RS} H$ 的规则不能随意地分解为 $E_1 \xrightarrow{RS} H$ 和 $E_2 \xrightarrow{RS} H$。

(2)BMI 模型应用于分布式专家系统时，在某些情况下可能会导致系统运行结果不等价。

上面两个结论，对 MYCIN 的确定性因子模型仍成立。因此，我们在使用它们时必须谨慎小心。

参考文献

[1]罗旭东、蔡经球、邱玉辉，一种新的基于区间估计的不确定推理模型，西南师范大学学报(自然科学版)，1994

[2]张为群，基于区间估计的不确定推理模型 BMI 的性质分析，计算机科学，Vol. 22，№5. 1995

[3] Shortliffe, E. H., Computer-Based Medical Consultations: MYCIN, American Elsevier Publishing Inc., New York, 1976

[4]Zadeh, L. A., The Concept of a Linguistic Variable and its Application to Approximate Reasoning, Part I, Inform. Sci., Vol. 8, 199-249; Part Ⅱ, Inform. Sci., Vol. 8, 301-357; Part Ⅲ, Inform. Sci. Vol. 9, 43-80, 1975

[5]Baldwin, J. F., Support Logic Programming, Int. J. Intell. Syst., 1, 73-104, 1986