第9卷第11期
1998年11月
软件学报
JOURNAL OF SOFTWARE
Vol.9, No.11
Nov. 1998

复合模糊命题的真值及其信息量计算的新方法*

刘纯武[1] 孙即祥[2]

[1](长沙工学院学机电工程与仪器系 长沙 410073)

[2](长沙工学院电子工程学院 长沙 410073)

摘要 从模糊信息量入手,对复合模糊命题的真值及其模糊信息量的计算进行了探讨,提出了一种新方法.此方法既考虑了子命题的逻辑关系,又考虑了子命题的相对重要程度而不丢失过多信息.实验证明,这种方法克服了封闭性、信息丢失、二义失效和全同失效的缺陷,能获得较合理的结果.

关键词 模糊信息量,加权模糊逻辑,模糊命题,真值.

中图法分类号 TP18

在模糊诊断和分析问题中,各个子命题对结论支持的程度(或相对重要程度)和子命题之间关系是不一样的.此时,运用"max-min"算子进行运算,各子命题的重要程度都等同地看待,忽略了各子命题的重要程度,这是大家公认的[1].另外,还存在着信息丢失、二义失效和等同失效的缺陷.针对上述问题,人们又提出了加权模糊逻辑运算方法.[2]这种方法克服了"max-min"算子忽略子命题的相对重要程度的问题,并且在数据库的模糊检索与匹配问题[2,3]和模糊诊断问题[4]中得到了应用.但是它扬弃了子命题之间的逻辑关系;不论"合取"或"析取"逻辑关系[1],其运算公式是一样的[2],显然这不符合常理.可认为这是从"max-min"算子的一个极端走向了另一个极端.

本文提出的新方法旨在既能体现上述运算的优点,同时又能克服它们的缺陷.由于各个子命题的相对重要程度不同,它们对复合命题的支持程度或所起的作用也不同.依据常识,每个子命题对复合命题都起一定的作用,且使复合命题获得一定量的信息.早在1968年,扎德就对这个问题进行了研究[5],提出了模糊熵的概念,并给出了模糊熵的计算式.但是还存在一些缺点,令人费解.[6,7]

1 模糊信息量

一般诊断分析型问题较复杂,一个复合命题往往由多个子命题组成.每个子命题的真值是推理过程中当前事实与已有知识的相似度或贴近度[6],它具有客观的和主观的两方面因素.实际中需考虑的不仅仅是某事件发生的概率,而更多的是考虑这事件或命题包含某一概念或意义的程度.

设事件(信源)集合为子命题构成的模糊集合:$A=\{a_1,a_2,...,a_m\}$,每个子命题的权因子$W=\{w_1,w_2,...,w_m\}$,且所有权因子之和等于1.结论(信宿)集合为复合命题构成的模糊集合:$B=\{e\}$.A'可看成是为真的子命题构成的模糊子集:$A'=\{A_1,A_2,...,A_m\}$;其中A_i对应着集合A中元素a_i的项构成的模糊子集,而$Q(A_i|a_i)$是集合A_i在集合A中元素a_i存在的条件下的主观条件概率.显然,它与子命题的真值$T(A_i)$的意义是一致的.

根据随机集落影理论[8]、广义信息论[6]中的后验广义熵或广义条件熵及自信息量的定义[9~11]可得,针对复合命题而言,某个子命题所提供的模糊信息量为

$$L(a_i)=H(e|a_i)=-w_i*\log T(A_i). \tag{1}$$

由式(1)可知:当真值$T(A_i)$一定时,子命题所提供的信息量$L(a_i)$随w_i增加而增加.显然这是合乎常理的,并且此模糊信息量包含了客观和主观两方面的信息量.

众所公认,运用"max-min"算子得到复合命题真值的结果基本能表达各子命题的逻辑组合情况.从模糊信息量的角度看,此时结果所含的模糊信息量包含了各子命题所能提供的大部分,同时也丢失了一小部分.由于复合命题获得的信息只能来自组成它的各子命题,因而真值最小(或最大)的子命题给复合命题提供的模糊信息量最多(或最少),其

---

* 本文研究得到国防预研基金资助.作者刘纯武,1963年生,讲师,主要研究领域为人工智能,多媒体技术.孙即祥,1946年生,教授,主要研究领域为计算机智能视觉,人工智能,模式识别.

本文通讯联系人:刘纯武,长沙410073,长沙工学院机电工程与仪器系801室

本文1997-05-15收到原稿,1997-10-14收到修改稿

余的就只提供少量模糊信息量.这说明子命题给复合命题提供的模糊信息量与子命题之间的逻辑关系存在着一定的联系.

传统的加权模糊逻辑考虑所有子命题对复合命题所起的作用,即其运算结果包含了所有子命题提供的模糊信息.但它不论子命题所提供的模糊信息的大小和多少,一概接受,这对于人类思维来说,尤其是解决复杂问题时,几乎是不可能的.如果把这种加权模糊逻辑的运算说成是过分注重局部的话,则运用“max-min”算子的逻辑运算就是过分地注重整体.因而,这种逻辑运算缺乏整体概念,扬弃了模糊子命题之间的逻辑关系.显然,将注重整体和局部的两种运算有机地结合起来才是最佳的.

当子命题之间具有“合取”关系时,子命题的真值越小,则它给复合命题提供的模糊信息量就越多,反之就少.同样,当子命题之间具有“析取”关系时,子命题的假值越小,则它给复合命题提供的模糊信息量就越多,反之就少.

设模糊命题的真值在[0,1]上,则其假值为:$F(A_i)=1-T(A_i)$.故当子命题之间具有“析取”关系时,可得

$$L(a_i)=-\omega_i*\log F(A_i)=-\omega_i*\log(1-T(A_i)).\tag{2}$$

当 $T(A_i)=0$(或 $F(A_i)=0$)时,式(1)(或式(2))的值为无穷大.此时,正好对应“max-min”算子的封闭性.为了克服封闭性的缺陷,实际运用中以很小的数来代替 $T(A_i)$ 或 $F(A_i)$;使(1)和(2)两式的值不为无穷大,此数值将根据具体情况确定.当式(1)对所有 $T(A_i)$ 都有 $T(A_i)=1$ 或者式(2)对所有 $F(A_i)$ 都有 $F(A_i)=1$ 时,这两式的值都为0.即每个子命题给复合命题提供的模糊信息量为0.实际运用中,一般用很小的数 β 代替这两式的值,消除全同的缺陷.根据上述分析,定义模糊信息量计算式为:

定义. 假设复合命题 $E=\{A_i|i=1,2,\ldots,m\}$ 由 m 个子命题组成,子命题的真值为 $T(A_i)$,权因子为 $W=\{w_1,w_2,\ldots,w_m\}$,且权因子之和为1,则子命题给复合命题提供的模糊信息量为(其中 L_∞ 是一个有限的数):

(1) 若子命题 i 与其他子命题之间存在“合取”关系,

$$L_{and}(A_i)=\begin{cases}-\omega_i*\log T(A_i) & 0<T(A_i)\leqslant 1\\ L_\infty & T(A_i)=0\end{cases},\tag{3}$$

(2) 若子命题 i 与其他子命题之间存在“析取”关系,

$$L_{or}(A_i)=\begin{cases}-\omega_i*\log(1-T(A_i)) & 0\leqslant T(A_i)<1\\ L_\infty & T(A_i)=1\end{cases}.\tag{4}$$

2 复合命题真值的演算

设 L_{max} 为所有子命题中给复合命题提供的最大模糊信息量,若 L_{max} 不存在,则用一有限数 α 来表示,有

$$L_\infty=\begin{cases}(1+\sigma)*L_{max} & L_{max}\text{存在}\\ \alpha & L_{max}\text{不存在}\end{cases}.\tag{5}$$

式(5)中的 σ 称作容忍率,其取值范围为$(0,\infty)$.式(6)给出了具有 m 个子命题的复合命题所能获得的模糊信息量(不存在 L_∞ 时),

$$L_m=\begin{cases}\sum_{i=1}^{m}L_{and}(A_i)\\ \sum_{i=1}^{m}L_{or}(A_i)\end{cases}=\begin{cases}-\sum_{i=1}^{m}\omega_i\log T(A_i) & \text{“and” logic}\\ -\sum_{i=1}^{m}\omega_i\log(1-T(A_i)) & \text{“or” logic}\end{cases}.\tag{5}$$

若将“max-min”运算过程也看成是“抽取”,则这种运算“抽取”一个子命题的真值就作为复合命题的真值,因而它存在一个抽取阈值,只是此阈值等于子命题中最大或最小的真值.为了将“整体”和“局部”的运算结合起来,这里定义一个抽取率 φ 来表达进行抽取前后子命题给复合命题提供的模糊信息量之比.假如某复合命题具有 m 个子命题,则所有子命题给复合命题提供的模糊信息总量为 $L=L_m$,且 L 不等于0;抽取后,若有 n 个子命题被抽取,则被抽取的子命题给复合命题提供的模糊信息量为 $L'=L_n$.这样,抽取率可写为 $\varphi=L'/L*100\%$.显然有 $0<\varphi\leqslant 1$.当抽取率 φ 大时,就说明复合命题的结果中丢失的信息少,反之则多.其中 $\varphi\neq 0$ 说明被抽取子命题至少有一个.当 $\varphi=1$ 时,抽取了所有子命题,恰好与传统的加权模糊逻辑运算过程是一致的.

当 $L=0$ 时,所有子命题给复合命题提供的模糊信息量之和为0,子命题的真值全部相同,且其真值皆为1(或0).此时,运用“max-min”算子将遇到全同失效.为了克服全同失效,设每个子命题给复合命题提供的模糊信息量皆为 β(小常数),且 $L=m*\beta$,则此时抽取率为 $\varphi=n/m$.

算法过程:(1) 计算每个子命题给模糊复合命题提供的模糊信息量及其总信息量.(2) 对子命题按其所提供的模

糊信息量从大到小进行排序.(3) 从已排序的队列中抽取一个子命题.(4) 判断抽取率是否满足，若满足则转(5)；否则，再抽取下一个子命题.(5) 运用相对权重模糊合/析取式[1,4]进行复合命题的真值计算.

由算法过程可知，在计算各子命题的模糊信息量的同时，又考虑了子命题之间的逻辑关系.其抽取过程类似于“max-min”算子的运算中的“max”运算和“min”运算，不同的是“max-min”算子的运算只“抽取”一个子命题，而新方法根据子命题提供的模糊信息量的多少抽取一个或多个子命题.第(5)步与传统的加权模糊运算方法类似，不同的是新方法不是针对所有子命题，而是只针对被抽取的部分子命题.因此，这种运算方法将“max-min”算子的运算与传统的加权模糊逻辑运算有机地融合在一起，既消除了“全同失效”和“二义失效”的缺陷，又考虑了各个模糊子命题之间的相对重要程度.由实验验证，复合命题的真值不仅与子命题的真值、权因子及子命题之间的逻辑关系有关，而且还与抽取比有关，显然这是符合实际的.

图 1 表示在只有抽取率不同的情况下，运用“max-min”算子、传统的加权模糊运算以及本文所提的新算法的运算结果比较，由此可知，运用新算法所得结果当抽取比较小时接近运用“max-min”的运算结果；而抽取比较大时接近运用传统的加权模糊运算的结果.这与理论上的分析是一致的.由实验可知，运用“max-min”算子时一般只抽取了总信息量的 30～40%.同时，当抽取率取 76%左右时，运用新算法所得结果最佳，此时充分兼顾了子命题的真值、相对重要程度及其逻辑关系.

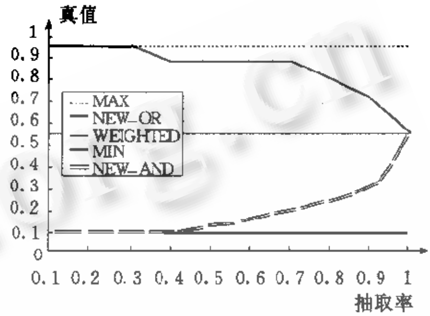

图1 不同抽取比后计算结果的比较

3 结 论

综上所述，这种新算法不仅考虑了各个子命题的真值，而且还考虑了子命题的相对重要程度及其逻辑关系.同时，由于它从各个子命题给复合命题提供的信息量出发，使得其运算结果丢失的信息不多，克服了“max-min”算子的二义失效和全同失效.实验表明：对于既要考虑子命题的真值及其逻辑关系，又要考虑命题的相对权重的诊断和分析问题，新算法的运用虽然复杂些，然而其效果优于运用“max-min”算子及传统的加权模糊逻辑算法.本文所述的算法已应用于心电图智能诊断及图象处理，获得了令人满意的结果.[12]

参考文献

1 何新贵.模糊知识处理的理论和技术.北京：国防工业出版社，1994.23～32，178～182
(He Xin-gui. The Theory and Technology of Fuzzy Knowledge Processing. Beijing: National Defense Industry Press, 1994. 23～32, 178～182)

2 何新贵.加权模糊逻辑及其广泛应用.计算机学报，1989，12(6)：458～464
(He Xin-gui. Weighted fuzzy logic and its applications in different fields. Chinses Journal of Computers, 1989, 12(6): 458～464)

3 欧阳为民.加权模糊逻辑及其应用.微型计算机，1994，14(6)：13～15
(Ou-Yang Wei-ming. Weighted fuzzy logic and its applications. Microcomputer, 1994, 14(6): 13～15)

4 陈颖等.加权模糊逻辑的研究.微型计算机，1994，14(6)：15～16
(Chen Ying *et al.* The research of weighted fuzzy logic. Microcomputer, 1994, 14(6): 15～16)

5 Zadeh L A. Probability measure of fuzzy events. Journal of Mathematical Analysis and Application, 1968, 23: 421～427

6 鲁晨光.广义信息论.合肥：中国科学技术大学出版社，1983.30～74
(Lu Chen-guang. The Generalized Information Theory. Hefei: Chinese Science and Technology University Press, 1983. 30～74)

7 Nikhil R P, Bezdek J C. Measuring fuzzy uncertainty. IEEE Transactions on Fuzzy Systems, 1994, 2(2): 107～118

8 汪培庄.模糊集和随机集落影.北京：北京师范大学出版社，1984
(Wang Pei-zhuang. Fuzzy Set and Random Set Projection. Beijing: Beijing Normal University Press, 1984)

9 Shannon C E. A mathematical theory of communication. Bell System Technical Journal, 1948, 27(3): 379～429

10 屈荫生.FUZZY 信息量及其在计算机辅助诊断中的应用.模糊数学，1982，6(2)：77～88
(Qu Yin-sheng. Fuzzy information amount and their application in computer-aided diagnosis. Fuzzy Mathematics, 1982, 6(2):

77～88)

11 孟庆生.信息论.西安:西安交通大学出版社,1986
(Meng Qing-sheng. Information Theory. Xi'an: Xi'an Traffic University Press, 1986)

12 刘纯武.一种新型的加权模糊推理模型及其在心电图智能诊断中的应用[硕士论文].长沙工学院,1996
(Liu Chun-wu. A new weighted fuzzy inference model and its application of ECG diagnosis [M. S. Thesis]. Changsha Institute of Technology, 1996)

A New Algorithm to the Truth Value and Its Information Amount of Compound Fuzzy Proposition

LIU Chun-wu[1] SUN Ji-xiang[2]

[1](*Department of Mechatronics Engineering and Instrumentation Changsha Institute of Technology Changsha* 410073)

[2](*Institute of Electronic Engineering Changsha Institute of Technology Changsha* 410073)

Abstract In this paper, the algorithm to the truth value and information amount of compound fuzzy proposition from fuzzy information amount is presented. The new algorithm is effective because it is relative to logic and relative importance between sub-propositions, and loses a little information. Through the experiments, the new algorithm eliminates the drawback of losing information, closing and invalidates of whole equality and ambiguity, and the reasonable results can be obtained.

Key words Fuzzy information amount, weighted fuzzy logic, fuzzy proposition, truth value.