

一种粗逻辑神经网络的构造和学习方法

陈锦荣¹ 丁世飞^{1,2}

(中国矿业大学计算机科学与技术学院 徐州 221116)¹

(中国科学院计算技术研究所智能信息处理重点实验室 北京 100080)²

摘要 传统的粗逻辑神经网络可以对信息系统及粗推理决策进行研究,能够深入地揭示粗糙集理论实质,但对于处理非单值输入问题不能取得良好的效果。粗糙神经元的上边界和下边界恰好能解决这一问题,且随着粗集理论不断发展,上下边界的概念得到了广泛的应用。综合两个方面的优点,提出了一种粗逻辑神经网络的构造与学习方法。它主要由传统粗逻辑神经网络和粗糙神经元的思想(模式中每一个特征变量都包含上界和下界两个边界)构成:边界粗逻辑神经网络。首先给出了粗糙神经元和粗逻辑及决策的基本知识,然后提出了边界粗逻辑神经网络的结构和学习方法及两种模型并比较了模型间的优缺点。与传统粗逻辑神经网络相比,这类神经网络能更有效地处理非单值和连续近似域函数问题。最后提出可以进一步优化的方向。

关键词 粗糙神经元,粗糙集,神经网络,粗逻辑神经网络

Construction and Learning Method of Rough Logic Neural Network

CHEN Jin-rong¹ DING Shi-fei^{1,2}

(School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)¹

(Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China)²

Abstract The traditional rough logic neural network can do research in information systems and decision-making and reveal the substance of rough set theory, but cannot get good results when dealing with the problem of non-single-value input. The rough neuron with upper and lower boundary can deal with the above problem, and with the development of rough set, the concept of upper and lower boundaries has been widely used. Comprehending the above advantages, this paper propounded the construction and studying of a kind of rough logic neural network. It is made up of rough logic neural network and rough neurons (each variable in this pattern has both upper and lower bounds), which is called boundary rough logic neural network. First the paper gave the basic knowledge about rough neuron, rough logic and decision-making, and then propounded the structure of boundary rough logic neural network and learning methods, then gave the two models about it and compared the advantages and disadvantages between them. It indicated that this type of neural network, compared with traditional rough logic neural network, can be more efficiency when dealing with the problem non-single-valued and continuous approximation function. At last it proposed the optimized direction.

Keywords Rough neuron, Rough set, Neural network, Rough logic neural network

1 引言

粗糙集理论是由波兰学家 Pawlak^[1] 提出的一种对不确定、不完整、不精确知识进行处理的数学理论。该理论凭借强大的定性分析能力,已成为除模糊集、神经网络之外的一种新的有效的不确定性计算工具,在模式识别、机器学习、知识发现等领域得到广泛应用。但是粗糙集本身抗噪声性能差,推广能力弱,在应用中往往需要与其它软计算方法如模糊集、神经网络进行集成。由于神经网络具有抗噪声性能好,泛化能力较强的特点,与粗糙集理论具有很强的互补优势,因此近年来,粗集理论与神经网络集成技术得到了积极的研究。其中利用粗糙集的信息零损失的属性约简并从中提取规则的能

力的粗糙神经网络得到了广泛的应用^[2-5];但以往的粗集神经网络都是以单存的粗糙集数据分析和约简为基础的,对其内在决策原理缺乏深入的阐述。粗逻辑^[6,7]从逻辑学的观点,利用逻辑工具,对信息系统及粗推理决策进行研究,能够深入地揭示粗糙集理论的实质。

数据的上界和下界概念在人工智能领域中使用得非常普遍,在粗糙集理论中尤为明显,随着粗集理论不断发展,上下边界的概念得到更大范围的应用。粗神经元的概念最早是 Lingras^[8]在1996年提出来的,粗糙神经元的定义和粗模式有关,模式中每一个特征变量都包含上下边界,这使基于粗糙神经元的神经网络能够较好地处理非单值输入的情况。张东坡^[9]等人提出了基于近似域划分的可变离散精度粗逻辑网

到稿日期:2011-01-22 返修日期:2011-04-23 本文受国家自然科学基金项目(60975039),江苏省自然科学基金项目(BK2009093)资助。

陈锦荣 男,硕士生,主要研究方向为粗糙神经网络, E-mail: mylikeme@126.com; 丁世飞 男,教授,博士生导师,主要研究方向为模式识别与人工智能、机器学习与数据挖掘、粗糙集与软计算、粒度计算。

络,即通过对可能性区域的进一步划分,提高处理这方面数据的精度和减少网络训练时间。本文从粗逻辑的角度,结合了粗糙神经元的思想,研究了基于粗糙神经元的粗糙逻辑神经网络。

2 粗糙神经元和粗逻辑

2.1 粗糙神经元

粗神经元的概念最早是 Lingras 在 1996 年提出来的,粗神经元的定义和粗模式有关,模式中每一个特征变量都包含上界和下界两个边界,如与气候相关的温度(日最高温度、日最低温度)、降雨量(年最高雨量、年最少雨量)这些特征参数都不能用一个精确的数值来描述,而用两个边界值或某一个变化区间来描述更合适,因此粗神经元比普通单输入输出的神经元更适合处理这些信息。图 1 是一个粗糙神经元的结构图^[10]。

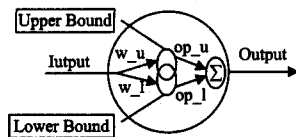


图 1 粗糙神经元结构图

每个粗神经元都由一对交叠上神经元 r (Upper Neuron) 和下神经元 r_l (Lower Neuron) 组成,与粗糙集中的上下近似的概念相呼应。粗神经元之间有全连接、兴奋连接、抑制连接 3 种常见的连接形式。上下神经元的交叠表示两者之间存在信息交换。

基于神经元激励函数 F (为 $\log \sin a$ 型函数),粗神经元的输出为:

$$(output)_{lower} = \min(F(input_{lower}), F(input_{upper}))$$

$$(output)_{upper} = \max(F(input_{lower}), F(input_{upper}))$$

2.2 粗逻辑及决策

已知一个信息系统 $S=(U, R)$, 信息系统如果可以以关系表的形式表示,则行对应要研究的对象,列为对应对象的属性。当属性集 R 由条件属性集 C 和决策属性集 D 构成时 $C \cap D = \Phi$, 此时信息系统对应的关系表称之为决策表。粗逻辑是一种基于粗糙集理论,通过逻辑的工具对决策表进行分析处理的理论。

由于知识库中往往存在冗余和相关信息,并非所有的知识都是必要的,因此引出了知识约简的概念,对于与知识信息系统对应的粗逻辑决策算法而言,知识约简意味着粗逻辑决策算法的简化。一个决策算法中所有的决策规则并非都是必要的,针对每一个决策类可以删除多余的决策规则,删除这些规则后并不影响决策的制定,去掉的规则的功能可以由留下来的规则完成。设 $S=(U, R)$ 是一个信息系统, F 是一个基本算法, F_ψ 表示算法 F 中所有具有后件 ψ 的基本规则集合, P_ψ 表示 F_ψ 中所有规则的前件组成的集合。算法 F 中规则 $\phi \rightarrow \psi$ 是多余的,如果 $\forall P\psi = \bigvee (P\psi - \{\phi\})$, 否则称该规则不是多余的,如果 F_ψ 中所有规则都不是多余的,则其是独立的。 F_ψ 的子集 $F\psi'$ 成为 F_ψ 的一个约简,如果 $F\psi'$ 是独立的,且 $\forall P\psi = \bigvee P\psi'$ 。决策规则集 $F\psi$ 称为既约的,如果它自身是一个约简。基本算法 F 称为极小的,如果其中每一个规则是既约的且对每一个决策规则 $\phi \rightarrow \psi, F\psi$ 是既约的。

3 边界粗逻辑神经网络

粗逻辑推理的优势在于可以通过上述粗逻辑理论来简化决策算法,为实际问题找到简捷高效的解决算法。具体可以分为 4 个步骤:属性域知识编码、知识获取、决策算法的简化、决策算法的实现。最后从最简决策表中提取决策规则,仿造模糊逻辑神经网络的结构,构建相应粗逻辑神经网络,在整合粗逻辑的决策能力和神经网络的推广泛化能力的基础上,解决实际问题。

3.1 边界粗逻辑神经网络结构及学习

在已获得的最小决策算法的基础上,类似于模糊逻辑神经网络和粗逻辑神经网络,构建使用于边界的粗逻辑神经网络,网络总共有 4 层构成,如图 2 所示。

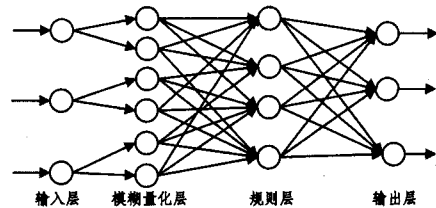


图 2 边界粗逻辑神经网络结构

第 1 层为输入层,输入向量 $X=(x_1, x_2, \dots, x_n)^T, x_i=(i=1, 2, \dots, n)$ 。

第 2 层为模糊量化层,依据某种离散化方法,分别将输入变量 $x_i=(i=1, 2, \dots, n)$ 离散化为 r_i 个不同值,该层神经元激励函数可取为高斯函数:

$$\mu_i^j = \exp\left(-\frac{x_i - c_{ij}}{\delta_{ij}}\right)^2$$

式中, c_{ij}, δ_{ij} 分别是被划分到第 j ($j=1, 2, \dots, r_i$) 档的变量 x_i 的类中心和方差,初始权值由训练样本进行估算。

第 3 层为规则层,其中每一个节点代表一条规则,规则的获取和描述的理论依据是粗逻辑理论,该层神经元和第 2 层、第 4 层神经元的连接关系由规则的前件和后件来决定,该层神经元本身完成规则的适应度计算、第 k 个神经元 R_k 的输出:

$$O_k^3 = \mu_1^{j_1}, \mu_2^{j_2}, \dots, \mu_n^{j_n}$$

式中, $\mu_1^{j_1}, \mu_2^{j_2}, \dots, \mu_n^{j_n}$ 表示与该节点连接的第 2 层节点的输出值, $j_1 \in \{1, 2, \dots, r_1\}, j_2 \in \{1, 2, \dots, r_2\}, \dots, j_n \in \{1, 2, \dots, r_n\}$, $k=1, 2, \dots, s, s$ 为规则数。

第 4 层为输出层,也就是边界决策层。

由于把粗糙神经元的思想引入到粗逻辑神经网络中,使得该网络的输出分为上边界输出和下边界输出,最后综合输出:

$$Output = (output)_{lower} + (output)_{upper}$$

上边界输出为:

$$(output)_{upper} = \max(f(input_{lower}), f(input_{upper}))$$

下边界输出为:

$$(output)_{lower} = \min(f(input_{lower}), f(input_{upper}))$$

其中函数 f 为:

$$y_c = O_c^4 = f_4\left(\sum_{l=1}^p w_{cl} O_l^3\right)$$

式中, $c=1, 2, \dots, C, l=1, 2, \dots, p, C$ 为决策类数, p 是与第 c 类相关的规则数, f_4 是输出层的激励函数,取为对数 S 函数

logsig。第3层和第4层神经元之间的连接权表示规则的合成算法,初始值可由规则的支持度来确定。

网络需调节的权值参数为 w_k 及等价类划分的类中心 c_{ij} 与方差 σ_{ij} , 可由方向传播 BP 算法迭代进行训练。但对于粗糙逻辑神经网络有如下需要解决的问题。

首先,对于粗糙逻辑神经网络来说,知识的等价类划分对网络性能影响很大,而在设计网络时,这往往是事先给定的,存在较大的主观性;但实际的情况下,上、下边界的确定不是太容易,因而会影响到粗糙逻辑难以得到最优结果。其次,粗糙逻辑神经网络设计过程中,可能会得到不相容的决策规则,这是由于条件属性过少造成的,可以通过增加条件属性来获得相容的决策表,但这不一定可行,因为需要增加的属性是未知的,这会造成获取的代价过高,另外属性过多的网络之规模和复杂度大大增加;这可以通过均衡决策,即从冲突本身给出的统计信息,通过调节第3层和第4层间的权值来实现。

粗糙逻辑神经网络的训练过程就是不断地调节其等价类划分(c_{ij}, σ_{ij})和规则合成算法(w_k)来实现模式分类。

3.2 构造边界粗糙逻辑神经网络

边界粗糙逻辑神经网络具有比传统的粗糙逻辑神经网络更好的性能,尤其是在处理不确定性区域的问题上。图3—图5是传统的粗糙逻辑神经网络到边界神经网络的结构示意图。

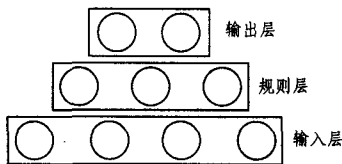


图3 传统的粗糙逻辑神经网络模型

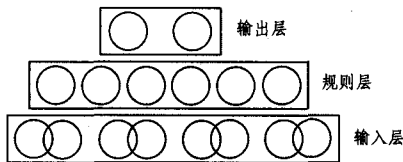


图4 边界粗糙逻辑神经网络模型一

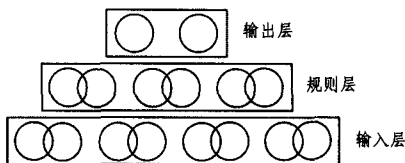


图5 边界粗糙逻辑神经网络模型二

图3为常规的粗糙逻辑神经网络模型,其要求的是单值的输入,在处理非单值输入时就要通过改变神经网络参数来处理,这样会增加神经网络的复杂程度和增加神经网络的训练时间。传统的粗糙逻辑神经网络在处理近似区域时,由于粗糙逻辑神经网络的精度很大程度上取决于离散化时的不可分辨区间的精细程度,因此划分的知识粒度越小,精度越高,但相应的规模越大。这都是传统的粗糙逻辑不能很好解决的问题。图4是边界粗糙逻辑神经网络的一种模型,其输入层的结构采用了粗糙神经元的思想,有上输入和下输入,这样就可以比较好地处理非单值输入的问题和近似区域的问题,但其仍用的是传统的规则层,这种边界粗糙逻辑神经网络并非最好的模型。图5的模型为图4的改进,此粗糙逻辑神经网络具有规则的上下层,其规则层能很好地与输入层对应起来,这样处理的结果

会比图4模型更加准确,同时输入层和决策层都用到了粗糙神经元的思想,这时就要考虑它们之间的连接问题,两个粗糙神经元部分连接的兴奋或者抑制特性是由连接权动态决定的。

主要用边界模糊神经网络处理非单值输入的情况。非单值输入可以分为两种情况:第一是输入值变量包含上界和下界,如与气候相关的温度(日最高温度、日最低温度)、降雨量(年最高雨量、年最少雨量)等特征参数都不能用一个精确的数值来描述,而用两个边界值或某一个变化区间来描述更合适,此时就可以把它们直接作为粗糙逻辑神经网络输入,不需要做太多的改动;另外一种连续含有近似域的函数,对于含有近似域的函数,如果在离散化时对其分得太细,会使网络的负担太大,反而得不到比较精确的结果,这在一定程度上会影响神经网络的适应程度,这时把上、下界的思想引入到离散化的过程中:以离散化后的结果为中心,以(离散化后的值+常数,离散化后的值-常数)为上、下边界,这就适合了粗糙逻辑神经网络的输入。这样当出现连续近似数据时,由于存在着上、下边界,离散化函数时就不要对其化得太细,以减少神经网络的负担。在处理最后结果时由于存在着上、下规则,最后可以综合结果,从而得到最后输出的精确结果。为了推广这种神经网络模型的应用范围,使常规的粗糙逻辑处理的问题也能使用该模型,这时要处理的问题为:神经网络的输入,可以参考连续函数近似域的方法来处理,通过人为的构造输入的上、下界来确定输入;如何处理常规的规则和含上、下层的规则层,可以参考图4和图5的模型,处理一般问题时,只用常规的规则层。

结束语 本文将粗糙神经元的思想和粗糙逻辑神经网络有机地结合,提出了基于粗糙神经元的粗糙逻辑神经网络,即粗边界逻辑神经网络用以有效地处理非单值和连续近似域函数的输入。这种神经网络既继承了传统粗糙逻辑神经网络的特点:用规则层代替隐层,使神经网络易于理解;又结合了粗糙神经元上下边界的思想,解决了传统粗糙逻辑神经网络不能处理非单值输入的问题和换了一种思路解决连续近似函数。但是仍有许多需要改进的地方,如粗糙逻辑神经网络性能在很大程度上取决于等价类划分,划分粒度越小,计算精度越高,但同时网络结构复杂化,计算量和训练时间会大量增加,如何取得较好的平衡仍是一个待进一步研究的问题;边界粗糙逻辑神经网络如何进一步融合传统粗糙逻辑,最后取代传统粗糙逻辑神经网络。

参考文献

- [1] Pawlak Z. Rough Set[J]. International Journal of Computer and Information Sciences, 1982, 11(15): 341-356
- [2] Banerjee M, Mitra S, Pal S K. Rrough fuzzy MLP: Knowledge encoding and classification[J]. IEEE Trans. on neural networks, 1998, 9(6): 1203-1216
- [3] Cui Bao-xia, Qu Xing-yu, Duan Yong. Study of ball mill material measure based on rough sets and RBF neural network data fusion[C]//Proceeding of 2009 International Conference on Intelligent Human-Machine Systems and Cybernetics. Zhejiang, 2009: 237-240

(下转第 268 页)

的融合图像清晰,原 MRI 图像中所含的软组织部分丢失了部分细节信息。与医学图像融合中常用的图 1(d)~图 1(g)的融合方法相比,本文算法得到的融合图像更清晰,视觉效果更好。

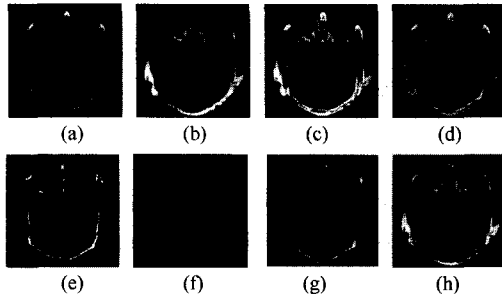


图 1 实验实例图

为了更好地评价融合图像性能,本文采用图像的信息熵、平均梯度、峰值信噪比和均方根误差作为评价指标对融合图像进行客观比较。信息熵反映了其包含的信息量的多少,熵值越大,说明融合效果越好;平均梯度反映了图像细节反差程度和纹理变化特征,一般来说,平均梯度值越大,表明图像越清晰,融合效果越好;峰值信噪比反映的是图像信噪比变化情况的统计平均,即整个图像的失真程度,一般情况下,峰值信噪比越高,说明融合效果和质量越好;均方根误差反映了融合图像与标准参考图像之间的差异程度,值越小,说明融合图像和参考图像越接近,融合效果越好。

表 1 给出了 5 种融合算法融合的图像(图 1(d)~图 1(h))的评价指标的计算结果。从表 1 可以看出,本文算法得到的融合图像的信息熵、平均梯度和峰值信噪比与其他几种算法相比,都是最大的,均方根误差与其他几种算法相比是最小的。这说明本文提出的融合规则能有效克服基于金字塔及小波变换融合忽略图像边缘信息的不足,能够从源图像中提取更多的边缘及细节信息。

表 1 图 1 中不同融合算法的性能

	熵	平均梯度	峰值信噪比	均方根误差
图(d)	6.0532	10.1187	19.8444	25.9610
图(e)	2.0603	6.1502	10.8312	73.2791
图(f)	6.3064	7.6025	17.0288	35.9005
图(g)	6.1326	9.7429	16.4699	38.2866
图(h)	6.7562	13.6662	20.0497	25.3546

结束语 医学图像的特殊性需要边缘及细节更加清晰的医学图像融合方法,而医学图像本质上的模糊性影响了医生

准确判断病灶区域。本文提出的基于提升小波变换的融合方法可以使医学图像边缘细节信息更加清晰,达到了很好的融合效果。本文方法能为医生临床诊断提供准确判断病灶的依据,具有良好的应用前景。

参考文献

- [1] 周朋,奚日辉. 基于小波变换的医学图像融合技术的实现[J]. 中国图象图形学报, 2006, 11(11): 1720-1723
- [2] Pajares G, de la Cruz J M. A wavelet-based image fusion tutorial [J]. Pattern Recognition, 2004, 37: 1855-1872
- [3] 周志光. 基于多尺度变换的图像融合方法研究[D]. 大连: 辽宁师范大学, 2006
- [4] 董卫军, 樊养余, 周明全. 基于提升格式小波变换的医学图像融合技术[J]. 计算机工程与应用, 2008, 44(13): 185-187
- [5] Zhu Yang-ming, Cochhoff S M. An object-oriented framework for medical image registration, fusion, and visualization Computer Methods and Programs [J]. Biomedicine, 2006, 82(3): 258-267
- [6] 张德峰. MATLAB 小波分析[M]. 北京: 机械工业出版社, 2009: 120-125
- [7] 李玲玲, 丁明跃. 一种基于提升小波变换的快速图像融合方法[J]. 小型微型计算机系统, 2005, 26(4): 667-670
- [8] 牛晓晖, 贾克斌. 基于 PCA 和自适应区域方差的图像融合方法[J]. 计算机应用研究, 2010, 27(8): 3179-3181
- [9] Kong Jun, Zheng Kai-yuan, Zhang Jing-bo, et al. Multi-focus image fusion using spatial frequency and genetic algorithm [J]. International Journal of Computer Science and Network Security, 2008, 8(2): 220-224
- [10] Burt P J, Adelson E H. The Laplacian Pyramid as a Compact Image Code[J]. IEEE Trans on Communication, 1983, 4(31): 532-540
- [11] Toet A, Van Ruyven L. Merging thermal and visual images by a contrast pyramid[J]. Optical Engineering, 1989, 28(7): 789-792
- [12] Burt, Kolczynski R J. Enhanced Image Capture Through Fusion [C] // Fourth International Conference on Computer Vision, 1993: 173-183
- [13] Hill P R, Bull D R, Canagarajah C N. Image fusion using a new framework for complex wavelet transforms [A] // Proceedings IEEE International Conference on Image Processing [C]. Genova, Italy, 2005, 2: 1338-1341

(上接第 238 页)

- [4] Gong Wei. Application of rough set and fuzzy neural network in information handling [C] // Proceeding of international Conference on Networking and Digital Society. Guizhou, 2009: 36-39
- [5] 夏红霞, 王惠营, 胡磊. 基于粗糙集的神经网络结构优化方法[J]. 计算机与数字工程, 2010, 38(5): 49-51
- [6] Pawlak Z. Rough Logic [J]. Bulletin of the Polish Academy of Sciences, Technical Science, 1987, 35(26): 253-258
- [7] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2001

- [8] Lingras P. Rough neural networks [C] // Proc. of the 6th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU'96). IEEE, 1996: 1445-1450
- [9] 张东坡. 基于近似域的可变离散精度粗逻辑网络及其遥感图像分类应用[J]. 电子与信息学报, 2007, 29(11): 2720-2723
- [10] Kothari A, Keskar A. Rough neuron based neural classifier [C] // Proceeding of first international conference on emerging trends in engineering and technology. Maharashtra, India, 2008: 624-628