第 27 卷第 11 期
2010 年 11 月

控 制 理 论 与 应 用
Control Theory & Applications

Vol. 27 No. 11
Nov. 2010

文章编号: 1000−8152(2010)11−1518−07

带参数聚合算子的模糊联想记忆网络

李 鹰[1], 徐蔚鸿[1,2], 唐良荣[1]

(1. 长沙理工大学 计算机与通信工程学院, 湖南 长沙 410076; 2. 吉首大学 数学与计算科学学院, 湖南 吉首 416000)

摘要: 基于最大运算Max和t–范数T的神经网络模型Max-T FAM是B.Kosko提出的经典模糊联想记忆(FAM)网络的一种重要的广义形式, 其性能有多处不足. 本文利用一种参数化聚合算子$\vee^{\lambda}$, 提出了一种计算简单、易于硬件实现的广义模糊联想记忆(GFAM)网络, 其连接算子从$\{\vee^{\lambda}|\lambda\in[0,1]\}$中选取; 从理论上严格证明了GFAM具有一致连续性, 比所有Max-T FAM的映射能力和存储能力强很多; 接着运用模糊关系方程理论提出和分析了GFAM的一种所谓的Max-Min-λ学习算法; 最后用实验对GFAM和Max-T FAM的完整可靠存储能力进行了比较, 并示例了GFAM在图像联想方面的应用.

关键词: 模糊神经网络; 模糊联想记忆; 学习算法; t–范数; 模糊关系方程
中图分类号: TP183 　　**文献标识码**: A

Fuzzy associative memory network based on parameterized gathering operator

LI Ying[1], XU Wei-hong[1,2], TANG Liang-rong[1]

(1. College of Computer and Communications Engineering, Changsha University of Science and Technology, Changsha Hunan 410076, China;
2. College of Mathematics and Computer Science, Jishou University, Jishou Hunan 416000, China)

Abstract: The neural network model Max-T FAM with the maximum operation and a t-norm T is an important generalized form of the classical fuzzy associative memory(FAM) network proposed by B.Kosko. This model has several disadvantages in its properties. Using a parameterized aggregating operator $\vee^{\lambda}$, we present a new generalized fuzzy associative memory(GFAM) network which is simple in computation and easy in implementation by hardware. All conjunctive operators of the interconnections of GFAM are chosen from a cluster $\{\vee^{\lambda}|\lambda\in[0,1]\}$. The strict theoretical study reveals that the GFAM is uniformly continuous and has much higher mapping ability and stronger storage capability than all Max-T FAMs. From the theory of fuzzy relational equations, we derive and analyze a so-called Max-Min-λ learning algorithm for GFAM. Experimental comparisons of the storage capability have been made between GFAM and all Max-T FAMs. An application of GFAM to associative images is illustrated.

Key words: fuzzy neural network; fuzzy associative memory; learning algorithm; t-norm; fuzzy relational equation

1 引言(Introduction)

不确定性人工智能已经成为21世纪人工智能的研究热点和重大前沿课题[1]. 模糊性是主观和客观世界中最普遍的不确定性形式之一. 自1987 年B.Kosko提出了模糊联想记忆网络(FAM)以来, 模糊神经网络(FNNs)得到了大量研究, 已成为带有模糊不确定性的智能系统的一种重要建模工具, 并在模式识别、自动控制和预测系统等领域取得大量的成功应用[2~5]. 神经网络和模糊系统均属无模型估计器和非线性动力学系统. 神经网络适合于处理非结构化信息, 模糊系统对处理结构化知识更为有效, 而FNNs是模糊系统和神经网络的互补式结合. FAM及其演变模型的理论和应用研究方兴未艾[2,3,6~10]. 研究表明: 采用模糊联想记忆网络(FAM)的模糊控制系统具有良好的鲁棒性, 而将微分竞争学习(DCL)机制引入FAM所形成的自适应FAM, 可以在输入、输出样本不甚完备的条件下, 以相对少的计算量获得较好的系统性能. FAM的研究非常活跃的广义模型是以下的Max-T FAM模型: $y_j=\bigvee\limits_{i\in I}(x_i\mathrm{T}w_{ij}), \forall j\in J$. 其中T是某个t–范数, $I=\{1,2,\cdots,n\}$, $J=\{1,2,\cdots,m\}$, w_{ij}是连接权值, $(x_1,x_2,\cdots,x_n)$是输入向量, $(y_1,y_2,\cdots,y_m)$是输出向量. 当T为取小运算Min时, Max-T FAM模型即传统的FAM模型. 然而Max-T FAM模型仍存在有

收稿日期: 2008−10−15; 收修改稿日期: 2010−05−12.
基金项目: 国家自然科学基金资助项目(6003302); 教育部重点科研基金资助项目(208098).

多种不足:

1) 网络簇{Max-T FAM|T是t–范数}中各具体网络的共性和个性还不很清楚. 实际应用中选择合适的t–范数T有困难.

2) t–范数是一类模糊算子簇, 其中不少具体的t--范数表达式复杂、运算量大, 且不便于硬件电路实现, 或不适宜应用于工业过程实时控制.

3) Max-T FAM模型的映射能力和存储能力等不尽人意. 例如, 当存在i_0使得$y_{i_0} > \bigvee\limits_{i\in I} x_i$时, 对任意t-范数T, Max-T FAM不能可靠地存储模式对$((x_1, x_2, \cdots, x_n)\ (y_1, y_2, \cdots, y_m))$.

1975年L.W.Fung和K.S.Fu提出了的一种有用的带参数λ的聚合算子$\vee^\lambda$. 本文将基于该算子提出一种不属于Max-T FAMs簇的结构新颖且性能好的模糊联想记忆网络(GFAM). 鉴于网络模型的连续性、映射能力、存储能力与网络的泛化能力、鲁棒性、容错性等实用性指标有着本质的直接联系, 所以本文将严谨地分析GFAM的连续性、映射能力和存储能力, 基于模糊关系方程理论提出它的一种学习算法, 并与Max-T FAMs簇进行比较.

2 相关定义和引理(Related definitions and lemmas)

在模糊集理论中, 为了弥补算子∧和∨在一些实际应用中对信息的处理太粗糙, 而将它们分别推广为t–范数和t-余范数. t–范数已在模糊理论中获得了广泛研究和应用. 例如, t–范数常被应用于推广传统的模糊联想记忆网络FAM、模糊双向联想记忆网络FAMB、模糊Hopfield网络等模型.

定义 1　二元运算T: [0,1]×[0,1]→[0,1]为t–范数, 如果对于任意$a, b, c \in [0, 1]$, T满足以下4个条件:

1) $a\mathrm{T}1 = a$(边界性);

2) 如果$b \leqslant c$, 则$a\mathrm{T}b \leqslant a\mathrm{T}c$(非减性);

3) $a\mathrm{T}b = b\mathrm{T}a$(交换性);

4) $a\mathrm{T}(b\mathrm{T}c) = (a\mathrm{T}b)\mathrm{T}c$(结合性).

据定义1, 对于任意a ∈[0,1]有$a\mathrm{T}0 = 0$.

若存在t–范数T使得$\forall a, b \in [0, 1], a\mathrm{S}b = 1-(1-a)\mathrm{T}(1-b)$, 则算子S称为t-余范数.

定义 2　对于任意给定的$\lambda \in [0, 1]$, 二元运算

$$\vee^\lambda : [0, 1] \times [0, 1] \to [0, 1]$$

定义为

$$a \vee^\lambda b = (a \wedge b) \vee (a \wedge \lambda) \vee (b \wedge \lambda),$$

或定义为

$$a \vee^\lambda b = \begin{cases} a \wedge b, & a, b > \lambda\ , \\ \lambda, & \lambda \in [a, b]或\lambda \in [b, a], \\ a \vee b, & a, b < \lambda. \end{cases}$$

易发现这种带有参数λ的算子仅由∨和∧两种算子构成, 易用硬件实现. 显见, 当$\lambda \neq 0$时, 算子$\vee^\lambda$不是t–范数, 而当$\lambda \neq 1$时$\vee^\lambda$也不是t-余范数. 算子$\vee^\lambda$拥有如下优良的运算性质:

1) 如果$\lambda \neq 0$且$a < \lambda$, $a \vee^\lambda 1 = a \vee \lambda \neq a$; 如果$\lambda \neq 1$且$a > \lambda$, $a \vee^\lambda 0 = a \wedge \lambda \neq a$;

2) 如果$b \leqslant c$, 则$a \vee^\lambda b \leqslant a \vee^\lambda c$(非减性);

3) $a \vee^\lambda b = b \vee^\lambda a$(交换性);

4) $(a \vee^\lambda b) \vee^\lambda c = a \vee^\lambda (b \vee^\lambda c)$(结合性);

5) $a \vee^\lambda b = (a \wedge b) \vee [\lambda \wedge (a \vee b)]$;

6) $a \vee^1 b = a \vee b, a \vee^0 b = a \wedge b$;

7) 当$\lambda_1 \leqslant \lambda_2$且$a, b \in [0, 1]$时, 有$a \vee^{\lambda_1} b \leqslant a \vee^{\lambda_2} b$.

当参数λ从0变化到1时, $a \vee^\lambda b$的值在区间$[a \wedge b, a \vee b]$上连续变化, $\vee^\lambda$算子便具有了自适应能力.

引理 1[6]

1) 不等式$|\bigvee\limits_{i\in I} a'_i - \bigvee\limits_{i\in I} a_i| \leqslant \bigvee\limits_{i\in I} |a'_i - a_i|$成立;

2) 如果$\delta > 0$且$|a'_i - a_i| \leqslant \delta$, $i \in I$, 则$|\bigwedge\limits_{i\in I} a'_i - \bigwedge\limits_{i\in I} a_i| \leqslant \delta$成立;

3) 不等式$|\bigwedge\limits_{i\in I} a'_i - \bigwedge\limits_{i\in I} a_i| \leqslant \bigvee\limits_{i\in I} |a'_i - a_i|$总成立.

引理 2　对任意$a_1, a_2, b_1, b_2, \lambda_1, \lambda_2 \in [0, 1]$,

$$|a_1 \vee^{\lambda_1} b_1 - a_2 \vee^{\lambda_2} b_2| \leqslant$$
$$|a_1 - a_2| \vee |b_1 - b_2| \vee |\lambda_1 - \lambda_2|$$

成立.

证　可利用引理1简单得证, 此处略.

3 一种基于带参数聚合算子的GFAM(A GFAM based on a parameterized gathering operator)

3.1 GFAM的定义(Definition of the GFAM)

如图1所示, 基于算子$\vee^\lambda$的GFAM, 从第i个输入节点到所有输出节点的连接算子是$\vee^{\lambda_i}$. 第i个输入节点到第j个输出节点的权值记为w_{ij}, 其中$i \in I$, $j \in J$. 对于任意输入向量$X = (x_1, x_2, \cdots, x_n) \in [0, 1]^n$, GFAM的输出将被逐点表示为

$$y_j = \bigvee_{i\in I} (x_i \vee^{\lambda_i} w_{ij}),\ \forall j \in J.$$

每个参数$\lambda_i \in [0, 1]$根据GFAM在实际应用中的需要进行调整. 根据算子$\vee^\lambda$的性质, y_j的另一种表达式为

$$
\begin{aligned}
y_j = &\bigvee_{i\in I}(x_i \vee^{\lambda_i} w_{ij}) = \\
&\bigvee_{i\in I}[(x_i \wedge w_{ij}) \vee (\lambda_i \wedge (x_i \vee w_{ij}))] = \\
&\bigvee_{i\in I}(x_i \wedge w_{ij}) \vee \bigvee_{i\in I}[\lambda_i \wedge (x_i \vee w_{ij})],\ \forall j \in J.
\end{aligned}
$$

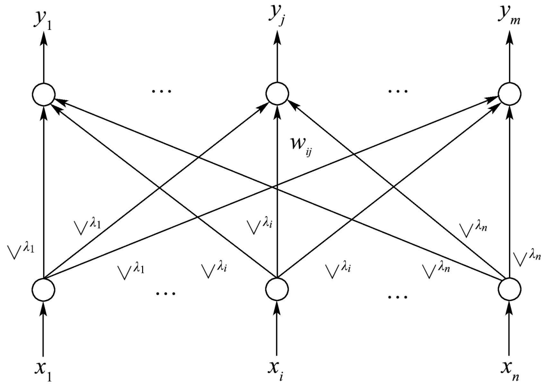

图 1 带有聚合算子$\bigvee^{\lambda_i}$的GFAM拓扑结构

Fig. 1 Architecture of the GFAM with operator $\bigvee^{\lambda_i}$

显然GFAM本质上只涉及$\vee$和$\wedge$两种运算, 故易于硬件实现. 如果所有λ_i都置为0, GFAM就变成为传统FAM. GFAM的各连接算子可能不同, 它们都是从算子簇$\{\vee^{\lambda}|\lambda \in [0,1]\}$中选取. 当参数$\lambda \neq 0$时, GFAM并不是Max-T FAM的一个特例.

3.2 GFAM的连续性(Continuity of the GFAM)

模糊神经网络常被用作模糊推理机. 当它被看作向量函数, 它的逼近能力与它是否关于输入向量具有连续性直接相关. 通常没有连续性的模糊神经网络不能很好地模拟真实系统, 有时其基本性能也难保证.

在以下部分, 当任意给定输入向量$X = (x_1, x_2, \cdots, x_n)$后, GFAM记为$\mathrm{GFAM}_X(\lambda_1, \lambda_2, \cdots, \lambda_n, w_{ij})$; 当任意给定输入向量$(x_1, x_2, \cdots, x_n)$和连接权矩阵$W$后, GFAM记为$\mathrm{GFAM}_{X,W}(\lambda_1, \lambda_2, \cdots, \lambda_n)$; 当任意给定连接权矩阵$W$, 参数向量$\varLambda = (\lambda_1, \lambda_2, \cdots, \lambda_n)$后, GFAM记为$\mathrm{GFAM}_{W,\varLambda}(x_1, x_2, \cdots, x_n)$.

定理 1

$$
\begin{aligned}
&\mathrm{GFAM}_{W,\varLambda}(x_1, x_2, \cdots, x_n),\\
&\mathrm{GFAM}_{X,W}(\lambda_1, \lambda_2, \cdots, \lambda_n),\\
&\mathrm{GFAM}_{X}(\lambda_1, \lambda_2, \cdots, \lambda_n; w_{ij}).
\end{aligned}
$$

分别是空间$[0,1]^n$, $[0,1]^n$和$[0,1]^{n+n\times m}$上的一致连续函数.

证 显然有$\mathrm{GFAM}_{W,\varLambda}(x_1, x_2, \cdots, x_n) \in [0,1]^m$, 又知所有的范数都是相互等价的, 所以只需以典型范数$||\cdot||_\infty$为例来证明定理1.

令: $G_1 = \mathrm{GFAM}_{W,\varLambda}(x1_1, x1_2, \cdots, x1_n)$,

$$G_2 = \mathrm{GFAM}_{W,\varLambda}(x2_1, x2_2, \cdots, x2_n),$$

则

$$
\begin{aligned}
&||G_1 - G_2||_\infty = \bigvee_{i\in I}|y1_j - y2_j| = \\
&\bigvee_{j\in J}|\bigvee_{i\in I}(x1_i \vee^{\lambda_i} w_{ij}) - \bigvee_{i\in I}(x2_i \vee^{\lambda_i} w_{ij})| \leqslant \\
&\bigvee_{j\in J}\bigvee_{i\in I}|x1_i \vee^{\lambda_i} w_{ij}) - (x2_i \vee^{\lambda_i} w_{ij})|(\text{据引理1}) \leqslant \\
&\bigvee_{j\in J}\bigvee_{i\in I}(|x1_i - x2_i| \vee |w_{ij} - w_{ij}| \vee \\
&|\lambda_i - \lambda_i|)(\text{据引理2}) = \bigvee_{j\in J}\bigvee_{i\in I}|x1_i - x2_i| = \\
&\bigvee_{i\in I}|x1_i - x2_i| = ||X1 - X2||_\infty.
\end{aligned} \tag{1}
$$

对于任意给定的$\varepsilon > 0$, 令正数$\delta(\varepsilon) = \varepsilon$, 任意给定$X1, X2 \in [0,1]^n$, 其满足不等式$||X1 - X2||_\infty < \delta(\varepsilon)$. 根据不等式(1), 有

$$||G_1 - G_2||_\infty \leqslant ||X1 - X2||_\infty < \delta(\varepsilon) = \varepsilon.$$

因此$\mathrm{GFAM}_{W,\varLambda}(x_1, x_2, \cdots, x_n)$关于输入向量$(x_1, x_2, \cdots, x_n)$是一致连续函数.

$\mathrm{GFAM}_{X,W}(\lambda_1, \lambda_2, \cdots, \lambda_n)$和$\mathrm{GFAM}_X(\lambda_1, \lambda_2, \cdots, \lambda_n; w_{ij})$分别是空间$[0,1]^n$和$[0,1]^{n+n\times m}$上的一致连续函数可类似证明, 版面所限, 此处略.

证毕.

定理1表明GFAM两个近似输入向量会产生两个近似的输出向量; GFAM两个近似的连接权矩阵对任意给定的同一输入向量, 会产生两个近似的输出向量. 定理1实际上已经保证了GFAM具有良好的抗噪性和鲁棒性. 其严格证明, 因版面限制此处略.

3.3 GFAM的映射能力(Mapping ability of the GFAM)

假设模式A与模式B是关联的, 如果模糊神经网络模型不能将A映射到B, 那么这种模型不能可靠存储模式对(A, B), 也不能从A准确地预测出B. 因此神经网络的映射能力在一定程度上影响了它的实用性. 如: 对任意t–范数T, 模式对$((0, 0, \cdots, 0)\ (0.5, 0, \cdots, 0))$不能被可靠地存储在Max-T FAM中. 这是因为对任意的Max-T FAM, 模式$(0, 0, \cdots, 0)$只能映射到$(0, 0, \cdots, 0)$上.

当神经网络看作从一个向量空间到另一个向量空间的映射时, 它的映射能力与它的值域紧密相关. 假设对于前向神经网络FNN, 当输入向量$X = (x_1, x_2, \cdots, x_n)$送入FNN时, FNN$(X)$代表FNN的输出向量, 则对任意给定的$X$, 令

$$
\begin{aligned}
&D_X(\mathrm{GFAM}) = \{\mathrm{GFAM}(X)|\forall W, \forall \varLambda,\ X\},\\
&D_X(\text{MAX-T FAM}) = \{\text{MAX-T FAM}\ (X)\,|\,X,\ \mathrm{T}\},\\
&D_{X,W}(\mathrm{GFAM}) = \{\mathrm{GFAM}(X)|\forall \varLambda,\ X,\ W\}.
\end{aligned}
$$

它们分别代表相应的值域, 其中T是任意t–范数.

定理 2

1) $D_{X,W}(\text{GFAM}) = \prod_{j=1}^{m}[\bigvee_{i\in I}(x_i\wedge w_{ij}), \bigvee_{i\in I}(x_i\vee w_{ij})]$;

2) $D_X(\text{GFAM}) \supseteq \bigcup_{\forall \text{T}} D_X(\text{MAX-T FAM})$.

特别地, 如果输入向量$X=(x_1,x_2,\cdots,x_n)$满足$\bigvee_{i\in I} x_i < 1$, 那么

$$D_X(\text{GFAM}) \supset \bigcup_{\forall \text{T}} D_X(\text{MAX-T FAM}).$$

证　因版面限制, 本文只给出2)的证明. 对

$$\forall(y_1',y_2',\cdots,y_m') \in \bigcup_{\forall \text{T}} D_X(\text{MAX-T FAM}).$$

必存在某些w_{ij}'和t–范数T′使得$y_j'=\bigvee_{i\in I}(x_i\text{T}'w_{ij}')(j\in J)$成立. 又根据t–范数的属性有

$$x_i\text{T}'w_{ij}' \leqslant x_i\text{T}'1=x_i,$$

所以

$$x_i\wedge(x_i\text{T}'w_{ij}')=x_i\text{T}'w_{ij}'.$$

现在令$\lambda_i=0(i\in I)$且$w_{ij}=x_i\text{T}'w_{ij}'(i\in I,j\in J)$, 此时

$$\bigvee_{i\in I}(x_i\vee^{\lambda_i}w_{ij})=\bigvee_{i\in I}[x_i\vee^0(x_i\text{T}w_{ij}')]=$$
$$\bigvee_{i\in I}[x_i\wedge(x_i\text{T}'w_{ij}')]=\bigvee_{i\in I}(x_i\text{T}'w_{ij}')=y_j',\ \ j\in J.$$

这就意味着$(y_1',y_2',\cdots,y_m')\in D_X(\text{GFAM})$成立. 另外, 当向量$(x_1,x_2,\cdots,x_n)$满足$\bigvee_{i\in I}x_i<1$时, 构造集合

$$D_{1,X}=\{(y_1^*,y_2^{*'},\cdots,y_m^*)|\ y_j^*=\bigvee_{i\in I}(x_i\vee^{\lambda_i}w_{ij}^*),$$
$$\lambda_i>\bigvee_{i\in I}x_i,\ w_{ij}^*>\bigvee_{i\in I}x,\ \forall i\in I,\ j\in J\}.$$

显然$D_{1,X}\subseteq D_X(\text{GFAM})$, 其中任意$(y_1^*,y_2^*,\cdots,y_m^*)\in D_{1,X}$, 这时有

$$y_j^*=\bigvee_{i\in I}(x_i\vee^{\lambda_i}w_{ij}^*)=\bigvee_{i\in I}(\lambda_i\vee w_{ij}^*)>\bigvee_{i\in I}x_i,\ \ j\in J.$$

然而, 对于每个

$$(y_1',y_2',\cdots,y_m')\in\bigcup_{\forall \text{T}} D_X(\text{MAX-T FAM}),$$

满足

$$y_j'=\bigvee_{i\in I}(x_i\text{T}w_{ij})\leqslant\bigvee_{i\in I}(x_i\text{T}1)=\bigvee_{i\in I}x_i,\ \ j\in J,$$

这样就有

$$(y_1^*,y_2^*,\cdots,y_m^*)\notin\bigcup_{\forall \text{T}} D_X(\text{MAX-T FAM}).$$

显然有当输入向量$X=(x_1,x_2,\cdots,x_n)$满足$\bigvee_{i\in I}x_i<1$时,

$$D_X(\text{GFAM})\supset\bigcup_{\forall \text{T}} D_X(\text{ MAX-T FAM})$$

成立.　　证毕.

定理2表明GFAM比模糊神经网络簇Max-T FAMs的映射能力强得多.

3.4 GFAM的存储能力(Storage capacity of the GFAM)

一个神经网络的可靠存储模式对的能力是网络的一个重要性能指标, 而B.Kosko提出的FAM及其广义形式Max-T FAMs簇的存储能力不尽人意. 以下假设$K=\{1,2,\cdots,p\}$是一个非空子集.

定理 3　对任意给定的t–范数T, 如果模式对集合$\{(A_k,B_k)|k\in K\}$能完整可靠地存储在Max-T FAM上, 那么它也能完整可靠地存储在GFAM上. 然而, 存在一些模式对集合能完整可靠地存储在GFAM上, 但不能完整可靠地存储在任何Max-T FAM上.

证　假设模式对集合$\{(A_k,B_k)|k\in K\}$能用带权值w_{ij}的Max-T FAM可靠地存储, 那么等式

$$b_{kj}=\bigvee_{i\in I}(a_{ki}\text{T}w_{ij}),\ \ \forall j\in J,\ \ \forall k\in K$$

成立. 下文构造能可靠存储该模式对集的GFAM的$\{w_{ij}'\}$和$\{\lambda_i'|i\in I\}$.

选择$\lambda_i'=0,\forall i\in I$. 对$\forall k\in K,\forall i\in I,\forall j\in J$, 首先令

$$U_{kj}=\{i|b_{kj}=a_{ki}\text{T}w_{ij},i\in I\}.$$

因为等式$b_{kj}=\bigvee_{i\in I}(a_{ki}\text{T}w_{ij})$成立, 故$U_{kj}\neq\varnothing$. 定义$w_{ij}'$如下:

1) 如果$i\in U_{kj}$, 则令$w_{ij}'=a_{ki}\text{T}w_{ij}(=b_{kj})$;

2) 如果$i\notin U_{kj}$, 则令$w_{ij}'=0$. 则

$$b_{kj}=\bigvee_{i\in I}(a_{ki}\vee^{\lambda_i'}w_{ij}')=$$
$$\bigvee_{i\in I}(a_{ki}\vee^0 w_{ij}')=\bigvee_{i\in I}(a_{ki}\wedge w_{ij}')=$$
$$[\bigvee_{i\in U_{kj}}(a_{ki}\wedge w_{ij}')]\vee[\bigvee_{i\in I-U_{kj}}(a_{ki}\wedge w_{ij}')]=$$
$$[\bigvee_{i\in U_{kj}}(a_{ki}\wedge(a_{ki}\text{T}w_{ij}))]\vee[\bigvee_{i\in I-U_{kj}}(a_{ki}\wedge 0)]=$$
$$\bigvee_{i\in U_{kj}}(a_{ki}\text{T}w_{ij})=b_{kj}(a_{ki}\geqslant a_{ki}\text{T}w_{ij}).$$

因此集合$\{(A_k,B_k)|k\in K\}$能可靠地存储在权值为w_{ij}'且$\lambda_i'=0,\forall i\in I$的GFAM上.

构造模式对集合$\{(A_k^*,B_k^*)|k\in K\}$, 使它满足

$$\forall k\in K,\ \ \forall i\in I,\ \ 0\leqslant a_{ki}^*<1,$$

又令

$$\forall i\in I,\ \ \lambda_i^*=\lambda^*,\ \ 1\geqslant\lambda^*>\bigvee_{k\in K}\bigvee_{i\in I}a_{ki}{}^*.$$

再令 $\forall k\in K,\ \forall j\in J,\ b_{kj}^*=\lambda^*.$

则集合$\{(A_k^*,B_k^*)\}$必能可靠地存储在带参数向量$(\lambda^*,\lambda^*,\cdots,\lambda^*)$和权值$w_{ij}^*$的GFAM上, 其中:

$$\forall i\in I,\ \ \forall j\in J,\ \ 1\geqslant w_{ij}^*\geqslant\lambda^*>\bigvee_{k\in K}\bigvee_{i\in I}a_{ki}^*.$$

这是因为有

$$\forall\ k\in k,\ \forall j\in J,$$

$$\bigvee_{i\in I}(a_{ki}^*\vee^{\lambda_i^*}w_{ij}^*)=\bigvee_{i\in I}\lambda_i^*=\lambda^*=b_{kj}^*.$$

然而集合$\{(A_k^*,B_k^*)|k\in K\}$不能完整可靠地存储在任何Max-T FAM上, 这是由于对带任何权值矩阵$\{w_{ij}\}$的Max-T FAM有

$$b_{kj}=\bigvee_{i\in I}(a_{ki}^*\mathrm{T}w_{ij})\leqslant\bigvee_{i\in I}a_{kj}<\lambda^*=b_{kj}^*.$$

证毕.

定理3表明GFAM比Max-T FAMs可靠存储多模式对的能力强得多. 有趣的是, 通过仔细分析定理3的证明过程, 可发现当$\forall\ i\in I,\lambda_i=\lambda$的特例GFAM模型就比Max-T FAM簇存储能力强很多.

4 GFAM的一种学习算法(A Learning algorithm for the GFAM)

神经网络的BP学习算法是一种基于梯度的比较通用的学习算法,但它有可能限于局部极小值点、收敛速度慢, 甚至不一定收敛等不足. 对于一个具体的神经网络模型, 只有结合它的结构特征, 才有可能开发出一个快速有效的学习算法. 这通常是困难而充满技巧性的. 对于Max-T FAMs已经提出了一些解析型学习算法. 下面提出的GFAM学习算法利用了模糊关系方程, 其求解方法研究一直活跃[10~12].

命题 1[12] 假设已知$R=(r_{ij})_{n\times m}$和$B=(b_1,b_2,\cdots,b_m)$, 则关于$X=(x_1,x_2,\cdots,x_n)$的模糊关系方程$X\circ R=B$有解的充分必要条件是等式$\bar{X}\circ R=B$成立, 其中$\bar{x}_i=\bigwedge_{j\in J}\{b_j|r_{ij}>b_j\}$, 且当方程有解时, $\bar{X}$是最大解. 对于空集$\varnothing$, 此处约定$\wedge\varnothing=1$. 算子“$\circ$”代表模糊关系合成运算.

在命题1中, 模糊关系方程$X\circ R=B$是方程$\bigvee_{i\in I}(x_i\wedge r_{ij})=b_j(j\in J)$的简写形式. 注意到$\bar{X}$的表达式在命题1中已直接给出, 可以直接用该表达式编程以检查模糊关系方程是否有解. 近几年人们利用模糊关系方程或关于t–范数T的广义模糊关系方程的解直接给出了FAM甚至Max-T FAM的解析型学习算法[10]. 例如, 对于FAM已有一个有效的所谓Max-Min学习算法. 只要这些模式对能完整可靠地存储在FAM上, 它就能为一组训练模式对直接提供合适的连接权值w_{ij}. 这种算法的成功归于对模糊蕴涵算子φ的巧妙利用, 其中φ表示为

$$(a\,\varphi\,b)=\begin{cases}1,\ a\leqslant b,\\ b,\ a>b.\end{cases}$$

在模糊集理论和技术中模糊蕴涵算子[13]是非常有用的概念, 如构造Max-T FAM的学习算法的常用方法就是在学习算法中利用某个合适的蕴涵算子.

下面同样利用模糊蕴涵算子, 为GFAM提出一个Max-Min-λ学习算法:

步骤 1 基于第k个模式对(A_k,B_k)和蕴涵算子φ, 给出了一种$n\times m$的临时权值矩阵W_k, $W_k=(w_{ij}^{(k)})=a_{ki}\,\varphi\,b_{kj}(k\in K)$.

步骤 2 将上述临时权值矩阵W_k合并成为最终的权值矩阵$\bar{W}$,

$$\bar{W}=\bigcap_{k=1}^{p}W_k=(\bigwedge_{k=1}^{p}(a_{ki}\,\varphi\,b_{kj})).$$

步骤 3 用下述方法来确定GFAM的参数$\lambda_i(i\in I)$:

验证对于任意$k\in K$, 等式$A_k\circ\bar{W}=B_k$是否成立, 即验证式(2)是否为真:

$$\bigvee_{i\in I}(a_{ki}\wedge\bar{w}_{ij})=b_{kj},\ \forall k\in K,\ \forall j\in J.\qquad(2)$$

这里GFAM的权值矩阵$\bar{W}$由步骤1与步骤2给出.

a) 如果式(2)成立, 令λ_i在区间$[0,\bigvee_{i\in I}(a_{ki}\wedge\bar{w}_{ij})]$上任意取值, 其中$i\in I$.

b) 如果式(2)不成立, λ_i的取值如下:

$$\lambda_i=\bigwedge_{k\in K,j\in J}\{b_{kj}|r_{k_{ij}}>b_{kj}\}.$$

其中: $r_{k_{ij}}=a_{ki}\vee\bar{w}_{ij},\ k\in K,\ i\in I,\ j\in J.$

假设对于空集$\varnothing$, $\bigwedge_{k\in K,j\in J}\varnothing=1$(参见上述命题1).

本文以下部分有如下约定: 对于任意模式对集合$\{(A_k,B_k)|k\in K\}$, 两个矩阵$[A_1\ A_2\cdots A_p]^{\mathrm{T}}$和$[B_1\ B_2\cdots B_p]^{\mathrm{T}}$分别记为$A$和$B$. 对于任意两个$h1\times h2$的矩阵$E=(e_{ij})$和$F=(f_{ij})$, $E\subset F$意味着对于任意i和j不等式$e_{ij}<f_{ij}$成立.

根据命题1易知, 如果关系$A\circ\bar{W}\subset B$成立, 训练模式集$\{(A_k,B_k)|k\in K\}$不能完整可靠地存储在FAM上, 这里权值矩阵$\bar{W}$由上述GFAM的学习算法的步骤1与2得到.

定理 4 假设GFAM采用了上述Max-Min-λ学习算法, 且训练集$\{(A_k,B_k)|k\in K\}$满足关系$A\circ\bar{W}$

$\subset B$, 那么当且仅当等式$\bar{\Lambda}\circ R_k = B_k, \forall k \in K$成立时, 训练集就能被完整可靠地存储在连接权值为$\bar{w}_{ij}$的GFAM上, 这里权值矩阵$\bar{W}$由步骤1和2确定, 且:

$$\bar{\Lambda} = (\bar{\lambda}_1\ \bar{\lambda}_2 \cdots \bar{\lambda}_n),\ \ \bar{\lambda}_i = \bigwedge_{k\in K, j\in J} \{b_{kj} | r_{k_{ij}} > b_{kj}\},$$

$$R_k = (r_{k_{ij}})_{n\times m} = (a_{ki} \vee \bar{w}_{ij})_{n\times m}.$$

证　集合$\{(A_k, B_k)|k \in K\}$能完整可靠地存储在连接权值为$\bar{w}_{ij}$的GFAM上, 当下列带有多个未知数$\lambda_1, \lambda_2, \cdots, \lambda_n$的方程有且仅有一个解时

$$\bigvee_{i\in I} (a_{ki} \vee^{\lambda_i} \bar{w}_{ij}) = b_{kj},\ \ \forall k \in K,\ \ \forall j \in J, \tag{3}$$

即

$$\begin{cases} \bigvee_{i\in I} [(a_{ki} \wedge \bar{w}_{ij})] \vee \bigvee_{i\in I} [\lambda_i \wedge (a_{ki} \vee \bar{w}_{ij})] = b_{kj}, \\ \forall k \in K,\ \forall j \in J. \end{cases} \tag{4}$$

因有已知条件$(A_k \circ \bar{W})_j < b_{kj}$, 所以方程组(4)能变换为如下形式:

$$\bigvee_{i\in I} [\lambda_i \wedge (a_{ki} \vee \bar{w}_{ij})] = b_{kj},\ \forall k \in K,\ \forall j \in J. \tag{5}$$

根据命题1, 式(5)有解当且仅当等式$\bar{\Lambda}\circ R_k = B_k$, $\forall k \in K$成立. 因此集合$\{(A_k, B_k)|k \in K\}$能完整可靠地存储在连接权值为$\bar{w}_{ij}$的GFAM上, 当且仅当$\bar{\Lambda}\circ R_k = B_k$, $\forall k \in K$成立.　　证毕.

本学习算法直接给出了网络权值和λ_i的解析表达式, 易于进行数学性质分析, 无需迭代, 时间效率高, 整个学习过程只涉及取大(∨)和取小(∧)两种运算, 易于硬件实现. 近来提出了基于遗传算法的求解模糊关系方程的方法. 就本GFAM而论, 该方法需要迭代, 只能求近似解, 过程复杂, 不易硬件实现.

对于输入为n维向量且输出为m维向量的GFAM和Max-T FAM模型, 训练后进行输出计算时, GFAM计算复杂度为$m\cdot n$次$\vee^{\lambda}$运算加$m\cdot(n-1)$次取大∨运算, 而Max-T FAM计算复杂度为$m\cdot n$次T模运算加$m\cdot(n-1)$次取大∨运算. 因为$\vee^{\lambda}$本质上是由3次取小∧运算和2次取大∨运算构成, 而常见的T模运算(除T = ∧外)一般都比$\vee^{\lambda}$运算耗时. 所以训练后, 通常GFAM比Max-T FAM产生输出的计算复杂度要小, 甚至从本质上小很多. 例如, 当

$$\mathrm{T}(x, y) = 1 - \min[1, ((1-x)^p + (1-y)^p)^{1/p}].$$

且n和m比较大时.

5　实验(Experiment)

5.1　比较GFAM和任意Max-T FAM的存储能力(Storage capacity comparison between the GFAM and any Max-T FAM)

本实验中使用由6个模式对组成的训练模式对集合TRAIN_SET = $\{(A_k, B_k)|k = 1, 2, \cdots, 6\}$(见表1). 首先注意到, 对任意的t–范数T和相应的Max-T FAM的任意连接权矩阵$W = (w_{ij})_{6\times 2}$, 当送入A_4给网络Max-T FAM时, 得到的输出B_4^*不等于B_4, 这是因为

$$b_{41}^* = \bigvee_{i=1}^{6} [a_{4i} \mathrm{T} w_{i1}] \leqslant \bigvee_{i=1}^{6} [a_{4i} \mathrm{T} 1] = 0.7 < b_{41} = 0.8.$$

所以, TRAIN_SET不能被任意Max-T FAM模型完整可靠存储, 但TRAIN_SET能被GFAM完整可靠存储. 当GFAM采用本文提出的Max-Min-λ学习算法时, 相应的连接权矩阵$\bar{W}$和λ_i的取值见表2.

表 1　本实验用的训练多模式对

Table 1　The training pattern pairs used in the experiment

	A_k	B_k
1	0.7, 0.2, 0.4, 0.6, 0.3, 0.8	0.8, 0.8
2	0.9, 0.4, 0.4, 1.0, 0.7, 0.0	1.0, 0.95
3	0.3, 0.6, 0.5, 0.5, 0.2, 0.9	0.8, 0.8
4	0.1, 0.7, 0.4, 0.7, 0.7, 0.7	0.8, 0.8
5	0.3, 0.6, 0.4, 0.8, 0.2, 0.4	0.8, 0.8
6	0.7, 0.9, 0.6, 0.1, 0.7, 0.4	0.9, 0.9
7	0.8, 0.1, 0.6, 0.5, 0.4, 0.3	0.8, 0.8

表 2　GFAM采用Max-Min-λ学习算法得到的λ_i和连接权矩阵

Table 2　The obtained λ_i and weight matrix of the GFAM using the Max-Min-λ learning algorithm

i	λ_i	$(\bar{w}_{ij})_{6\times 2}$	i	λ_i	$(\bar{w}_{ij})_{6\times 2}$
1	0.8	1.0, 1.0	4	0.8	1.0, 0.95
2	0.8	1.0, 1.0	5	0.8	1.0, 1.0
3	0.8	1.0, 1.0	6	0.8	0.8, 0.8

5.2　GFAM应用于图象联想(Application of the GFAM in an image association)

本实验给出GFAM在图像联想方面的应用(由一个小学女生联想到她同桌男生). 当女生图像出现一些噪声时, 通过GFAM依然能基本联想到她同桌男生.

图2中两同桌学生图像的灰度级为256、大小为64×64. 图2中左、右图像的上半部分分别作为一个训练模式对的前件和后件, 由左、右图像的下半部分构成另一训练模式对. 构建相应GFAM, 采用本文提出的学习算法训练后存储图2中两图像. 将图2女生图像加入灰度不等的7个噪声斑块后成为图3女生图像, 并将其送入GFAM, 输出图象为图3的男生图像, 其与图2男生图像基本一致.

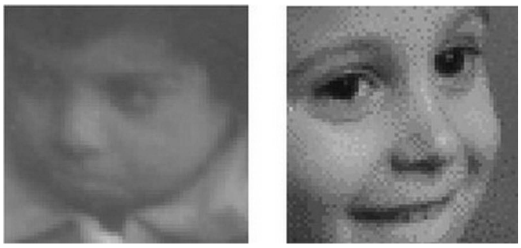

图 2　被GFAM存储的女生和男生图象

Fig. 2　Images of a girl and a boy stored by GFAM

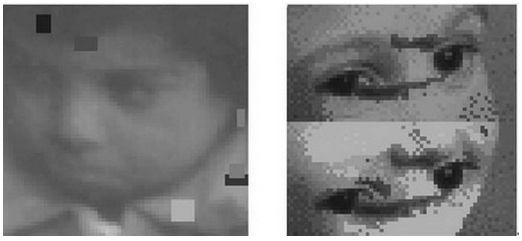

图 3　送入含噪声的女生图象给GFAM后联想到的男生图象

Fig. 3　The associated boy image when the noisy girl image is presented to the GFAM

6　结论(Conclusions)

聚合算子$\vee^{\lambda}$能统一算子$\vee$和$\wedge$, 特别地, 由于参数λ使它具有灵活性和适应性. 得益于这种算子的很好的性质, GFAM结构简单, 易于硬件电路实现, 具有一致连续性, 其映射能力和存储能力比模糊神经网络簇Max-T FAMs强得多, GFAM的Max-Min-λ学习算法是一种解析型的学习算法, 易于性质分析和软件实现. 在本文的结论和分析技巧的基础上, 可对GFAM的泛化能力、容错性、鲁棒性、学习算法的设计和实际应用作进一步研究. GFAM良好性能预示了其良好的应用潜力.

参考文献(References):

[1] 李德毅, 刘常昱. 不确定性人工智能[J]. 软件学报, 2004, 15(11): 1583 – 1594.
(LI Deyi, LIU Changyu. Artificial intelligence with uncertainty[J]. *Journal of Software*, 2004, 15(11): 1583 – 1594.)

[2] VALLE M E , SUSSNER P. A general framework for fuzzy morphological associative memories[J]. *Fuzzy Sets and Systems*, 2008, 159(7): 747 – 768.

[3] COH H , LIM J H, QUEK C. Fuzzy associative conjuncted maps network[J]. *IEEE Transactions on Neural Networks*, 2009, 20(8): 1302 – 1319.

[4] 乔俊飞, 王会东. 模糊神经网络的结构自组织算法及应用[J]. 控制理论与应用, 2008, 25(4): 703 – 707.
(QIAO Junfei, WANG Huidong. Structure self-organizing algorithm for fuzzy neural networks and its applications[J]. *Control Theory & Applications*, 2008, 25(4): 703 – 707.)

[5] 张化光, 何希勤. 模糊自适用控制理论及其应用[M]. 北京: 北京航空航天大学出版社, 2002.
(ZANG Huaguang, HE Xiqin. *Theory and Application of Fuzzy Self-adaptive Control*[M]. Beijing: Beijing Aeronautics and Astronaotics University Press, 2002.)

[6] 徐蔚鸿, 宋鸾姣. 训练模式的摄动对模糊双向联想记忆网络的影响及其控制[J]. 计算机学报, 2006, 29(2): 337 – 344.
(XU Weihong, SONG Luanjiao. Influences and control of perturbation of training pattern pairs on fuzzy bidirectional associative memories[J]. *Chinese Journal of Computers*, 2006, 29(2): 337 – 344.)

[7] CHUNG F L, LEE T. On fuzzy associative memory with multiple-rule storage capacity[J]. *IEEE Transactions on Fuzzy Systems*, 1996, 4(3): 375 – 384.

[8] SUSSNER P, VALLE M E. Implicative fuzzy associative memories[J]. *IEEE Transactions on Fuzzy Systems*, 2006, 14(6): 793 – 807.

[9] LIU P Y. The fuzzy associative memory of max-min fuzzy neural network with threshold[J]. *Fuzzy Sets and Systems*, 1999, 107(2): 147 – 157.

[10] STAMOU G B, TZAFESTAS S G. Neural fuzzy relational systems with a new learning algorithm[J]. *Mathematics and Computers in Simulation*, 2000, 51(1): 301 – 314.

[11] PEDRYCZ W. S–t fuzzy relational equations[J]. *Fuzzy Sets and Systems*, 1993, 59(2): 189 – 195.

[12] 谢季坚, 刘承平. 模糊数学方法及其应用[M]. 武汉: 华中科技大学出版社, 2000.
(XIE Jijian, LIU Chengping. *Fuzzy Mathematics and Its Applications*[M]. Wuhan: China Central University of Science and Technology Press, 2000.)

[13] YING M S. Implication operators in fuzzy logic[J]. *IEEE Transactions on Fuzzy Systems*, 2002, 10(1): 88 – 91.

作者简介:

李　鹰　(1961—), 男, 副教授, 主要研究方向为不确定性人工智能、灰色理论和数据库技术, E-mail: liying5196@163.com;

徐蔚鸿　(1963—), 博士, 教授, 博士生导师, 主要研究方向为不确定性人工智能、模糊系统、模式识别和软件工程, E-mail: xwhxd@163.com;

唐良荣　(1963—), 讲师, 主要研究方向为人工智能与数据库技术, E-mail: tlr1688@163.com.