第 25 卷 第 1 期 电子与信息学报 Vol.25 No.1
2003 年 1 月 JOURNAL OF ELECTRONICS AND INFORMATION TECHNOLOGY Jan. 2003

二进神经网络的模式匹配学习 [1]

陆 阳 魏 臻 韩江洪 樊玉琦

(合肥工业大学计算机与信息学院 合肥 230009)

摘 要 二进神经网络的知识提取需要了解每个神经元的逻辑意义．一般来说，对二进神经网络学习结果的分析是困难的．该文提出了一种基于线性可分结构系结构分析的学习算法，采用这种方法对布尔空间的样本集合进行学习，得到的二进神经网络隐层神经元都归属于一类或几类线性可分结构系，只要这几类线性可分结构系的逻辑意义是清晰的，就可以分析整个学习结果的知识内涵．

关键词 二进神经网络，线性可分，模式匹配

中图号 TN-052

1 引 言

对于 n 维布尔空间中的样本集合，如果用不同的二进神经网络学习算法对它进行学习，所得到的学习结果是不一样的，表现在其中各二进神经元的权值和阈值不同；一个二进神经元等价于一个线性可分函数，因此上述问题的实质是不同的二进神经网络学习算法导致二进神经元表达了不同的线性可分函数，这些函数在布尔空间中具有不同的空间结构，它们的知识内涵也不相同．虽然目前对于某些线性可分结构已经具有一般判别方法，并且也可以“理解”其中的逻辑意义，例如汉明球 [1] 和 PSP(Positive Successive Product) 函数 [2]，但并不是每一个二进神经元所表达的线性可分结构都是可判别的，也不是每一种线性可分结构的逻辑意义都是清晰的，例如文献 [3] 中定理 2 所表述的线性可分结构、文献 [4] 中定理 2 所表述的线性可分结构以及 Jung H. Kim 等提出的几何学习算法 [2] 的隐层神经元等都存在这样的问题．在这种情况下，如果能构造出一种二进神经网络学习算法，它在学习过程中能将二进神经网络的每个神经元导向可判别的和可“理解”的二进神经元，那么这种学习算法的整体学习结果也将是一个可“理解”的二进神经网络；通过这种算法，就可以对二进神经网络进行知识分析，从而建立一条从二进神经网络中进行知识提取的新途径．

$\boldsymbol{F}_2$ 表示二元域， f 是 $\boldsymbol{F}_2^n \to \boldsymbol{F}_2$ 的任意映射，称 f 为 n 元布尔函数， $\boldsymbol{F}_2^n$ 构成了 n 维布尔函数样本空间．对所有 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)\in \boldsymbol{F}_2^n$，以 $f(x_1,x_2,\cdots,x_n)$ 表示 f 在 $\boldsymbol{X}$ 处的取值，简记为 $f(\boldsymbol{X})$，称 $\boldsymbol{X}$ 为 $\boldsymbol{F}_2^n$ 中的样本， $f(\boldsymbol{X})$ 为样本值， $f(\boldsymbol{X})\in \boldsymbol{F}_2$．通常情况下布尔函数二元域取 $\{0,1\}$，在本文的讨论中为使讨论的结果更具一般性， $\boldsymbol{F}_2$ 定义为 $\{-1,1\}$，可以通过变换 $x_i=2c_i-1$ 将 $\{0,1\}$ 映射为 $\{-1,1\}$．

2 几类线性可分结构系

线性可分结构系是指具有相同结构属性的线性可分结构的集合。线性可分结构系中的结构属性是抽象概念，它是用于描述和定义某类线性可分结构系的一组可表达的关系 (穷举表达方式除外)．

根据布尔空间线性可分函数的概念，设 $f(x_1,x_2,\cdots,x_n)$ 是定义于 $\{-1,1\}$ 的线性可分布尔函数，$\sum_{i=1}^n w_ix_i-T=0$ 是 f 的分类超平面，则 $f(x_1,x_2,\cdots,x_n)=1$ 时，$\sum_{i=1}^n w_ix_i-T\ge 0$；$f(x_1,x_2,\cdots,x_n)=-1$ 时， $\sum_{i=1}^n w_ix_i-T<0$．如果二进神经元有如下结构：

[1] 2001-07-12 收到， 2001-12-17 改回

安徽省重点科研计划 (No.01041175) 资助项目

$$S = U(\sum_{i=1}^{n} w_i x_i - T)$$

$$U(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

则称 S 表达布尔函数 $f(x_1, x_2, \cdots, x_n)$ ，记 S 为 $S(\boldsymbol{W}, \boldsymbol{X}, T)$ ，其中 $\boldsymbol{W} = (w_1, w_2, \cdots, w_n)$ ，$\boldsymbol{X} = (x_1, x_2, \cdots, x_n)$ 。

定义 1　集合 $f^{-1}(1) = \{\boldsymbol{X} \in \boldsymbol{F}_2^n | f(\boldsymbol{X}) = 1\}$ 。

在下面的讨论中，将用符号 x_i^* 表示 x_i 或者 $\overline{x}_i$ ，x_i 称为 x_i^* 的正态，$\overline{x}_i$ 称为 x_i^* 的负态。

定义 2　$\boldsymbol{F}_2^n$ 空间中，函数 $f(\boldsymbol{X}) = x_{s1}^* \wedge x_{s2}^* \wedge \cdots \wedge x_{sm}^*$ ，$m \leq n$ ，则由 $f^{-1}(1)$ 构成的空间结构称为正超立方体结构；所有正超立方体组成的集合称为正超立方体结构系。

定义 3　$\boldsymbol{F}_2^n$ 空间中，函数 $f(\boldsymbol{X}) = x_{s1}^* \wedge x_{s2}^* \wedge \cdots \wedge x_{sl}^* \wedge (x_{t1}^* \vee x_{t2}^* \vee \cdots \vee x_{th}^*)$ ，$l+h \leq n$ ，$h \neq 0$ ，则由 $f^{-1}(1)$ 构成的空间结构称为凹超立方体；所有凹超立方体组成的集合称为凹超立方体结构系。

定义 4　$\boldsymbol{F}_2^n$ 空间中，函数 $f(\boldsymbol{X}) = x_{s1}^* \vee x_{s2}^* \vee \cdots \vee x_{sl}^* \vee (x_{t1}^* \wedge x_{t2}^* \wedge \cdots \wedge x_{th}^*)$ ，$l+h \leq n$ ，$h \neq 0$, 则由 $f^{-1}(1)$ 构成的空间结构称为凸超立方体结构；所有凸超立方体组成的集合称为凸超立方体结构系。

定义 5　在 $\boldsymbol{F}_2^n$ 中，对于样本 $\boldsymbol{X}^c = (x_1^c, x_2^c, \cdots, x_n^c)$ ，如果存在集合 $U(d) = \{\boldsymbol{X} = (x_1, x_2, \cdots, x_n) \in f^{-1}(1) | d_H(\boldsymbol{X}^c, \boldsymbol{X}) \leq d\}$ ，d 是位于区间 $[0, n]$ 中的整数，则 $U(d)$ 称为以 $\boldsymbol{X}^c$ 为中心，半径为 d 的汉明球，$d_H(\boldsymbol{X}^c, \boldsymbol{X})$ 为 $\boldsymbol{X}^c$ 与 $\boldsymbol{X}$ 之间的汉明距离；所有汉明球组成的集合称为汉明球系。

有关研究表明 [1,5] ，正超立方体、凹超立方体、凸超立方体和汉明球都是线性可分结构，并且这些线性可分结构都具有明确的逻辑意义。

3 线性可分结构系的分层表达

对于不同的线性可分结构系，表达它的结构属性是不一样的，但一般说来，每一种线性可分结构系的结构属性都是可以分层表达的。虽然划分线性可分结构系结构层次的方式根据不同线性可分结构系的结构特点而不同，但仍然存在一些基本的类似之处。下面通过对凹超立方体结构系的结构属性分层来说明这个问题。

定义 6　凹超立方体结构系中，称一个满足条件的二元组 $< l, h >$ 为凹超立方体结构系的一个结构样式 (Structure style) 。每个结构样式中由 $x_{s1}^*, x_{s2}^*, \cdots, x_{sl}^*$ 和 $x_{t1}^*, x_{t2}^*, \cdots, x_{th}^*$ 形成的每一种组合，称为该结构样式中的一种结构类型 (Structure type) 。

定义 7　一种线性可分结构称为一种模式 (Pattern) 。

对于凹超立方体而言，每个结构类型中由各输入变量的正态和负态形成的每一种组合，都构成一个模式。结构样式和结构类型的概念是在凹超立方体结构系中定义的，而模式的概念适用于所有类型的线性可分结构系。

以 8 维布尔空间中的凹超立方体 $f(\boldsymbol{X}) = x_1 \wedge x_4 \wedge \overline{x}_5 \wedge (\overline{x}_2 \vee x_3 \vee x_7 \vee \overline{x}_8)$ 为例，在结构样式这一个层次 $l = 3$ ，$h = 4$, 满足 $l + h \leq n$ ；在结构类型这一个层次，由 x_1^* ，x_4^* ，x_5^* 构成其中的"与"逻辑部分，由 x_2^* ，x_3^* ，x_7^* ，x_8^* 构成其中的"或"逻辑部分；在模式这一个层次，由各变量的正态、负态 x_1 ，$\overline{x}_2$ ，x_3 ，x_4 ，$\overline{x}_5$ ，x_7 ，$\overline{x}_8$ 构成一个确定的模式。所以，一个模式对应一个线性可分结构，结构样式的全体构成了凹超立方体结构系。

虽然结构样式和结构类型的定义只适用于凹超立方体结构系，但所采用的思路是可以借鉴到其它线性可分结构系中的．例如对于汉明球系，如果不对具体的层次概念进行定义，可以给出以下的结构分层的方法．

(1) 以不同的汉明球半径 d，划分汉明球系的第一个结构层次；

(2) 以不同的汉明球中心 $\boldsymbol{X}^c=(x_1^c,x_2^c,\cdots,x_n^c)$，划分汉明球系的第二个结构层次．

在凹超立方体结构系中，结构层次可以划分为三层，而在汉明球中只需划分二层结构，就可以明确一个具体的汉明球，这说明不同的线性可分结构系结构属性数目可能是不一样的．

4 模式匹配学习

4.1 理论基础

连通集是布尔空间中线性可分函数的必要条件 [6]，模式匹配学习算法就是在线性可分函数这一重要的理论基础上，通过对凹超立方体结构系的结构性质进行分析而提出的．

定义 8 $\boldsymbol{A}$ 是 $\boldsymbol{F}_2^n$ 空间的连通集，称 $\boldsymbol{A}$ 中的样本个数为连通集 $\boldsymbol{A}$ 的连通度．$\boldsymbol{I}=f^{-1}(1)$ 是 $\boldsymbol{F}_2^n$ 空间的输入样本集，$\boldsymbol{I}$ 是数个连通集的并集，称其中含有最多样本个数的连通集为 $\boldsymbol{I}$ 的最大连通集，称最大连通集的连通度为最大连通度．

定义 9 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 是 $\boldsymbol{F}_2^n$ 空间的一种模式，如果 T 是该模式中样本值为“1”的样本集合，则称集合 T 中的样本个数为模式 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 的覆盖度．

定义 10 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 是 $\boldsymbol{F}_2^n$ 空间的一个模式，$\boldsymbol{A}$ 是 $\boldsymbol{F}_2^n$ 空间的一个连通集，如果对 $\forall\boldsymbol{X}$，$S(\boldsymbol{X},\boldsymbol{W},\theta)=1$，都满足 $\boldsymbol{X}\in\boldsymbol{A}$，并且 $\forall\boldsymbol{X}$，$S(\boldsymbol{X},\boldsymbol{W},\theta)=-1$，都满足 $\boldsymbol{X}\notin\boldsymbol{A}$，则称模式 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 匹配 $\boldsymbol{A}$．

定理 1 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 是 $\boldsymbol{F}_2^n$ 空间的一个模式，$\mathrm{CV}(S)$ 是 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 的覆盖度，$\boldsymbol{A}$ 是输入样本集 $\boldsymbol{I}$ 中的一个连通集，$\mathrm{CN}(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的连通度，若 $\mathrm{CV}(S)>\mathrm{CN}(\boldsymbol{A})$，则 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 不能匹配 $\boldsymbol{A}$．

证明 采用反证法；假设 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 能匹配 $\boldsymbol{A}$，则根据定义 10，对 $\forall\boldsymbol{X}$，$S(\boldsymbol{X},\boldsymbol{W},\theta)=1$，都满足 $\boldsymbol{X}\in\boldsymbol{A}$．

设 T 是模式 $S(\boldsymbol{X},\boldsymbol{W},\theta)$ 中样本值为“1”的样本集合，则因为 $\mathrm{CV}(S)>\mathrm{CN}(\boldsymbol{A})$，即 $\boldsymbol{T}$ 中的样本个数多于 $\boldsymbol{A}$ 中的样本个数，所以一定存在样本 $\boldsymbol{X}$，满足 $S(\boldsymbol{X},\boldsymbol{W},\theta)=1$，并且 $\boldsymbol{X}\notin\boldsymbol{A}$，与假设矛盾． 证毕

4.2 算法描述

模式匹配学习是用已知的一类或几类线性可分结构系中的所有线性可分结构 (模式) 对输入样本集进行结构匹配，直至所有的样本都至少包含在一种线性可分结构 (模式) 中．匹配后的每一个线性可分结构 (模式) 构成一个二进神经元．在本文中采用的线性可分结构系是凹超立方体结构系．

定理 2 $\boldsymbol{F}_2^n$ 空间中，对于由二元组 $<l,h>$ 构成的凹超立方体结构系的一个结构样式，它所包含的所有模式的覆盖度都等于 $(2^h-1)\times 2^{n-l-h}$；称 $(2^h-1)\times 2^{n-l-h}$ 为对应于 $<l,h>$ 的结构样式覆盖度．

证明 求取模式的覆盖度就是求取模式中样本值为“1”的样本个数．

对于凹超立方体结构系，“或”项 $x_{t1}^*\vee x_{t2}^*\vee\cdots\vee x_{th}^*$ 中为“1”的组合数目是 2^h-1，“与”项 $x_{s1}^*\wedge x_{s2}^*\wedge\cdots\wedge x_{sl}^*$ 中为“1”的组合数目是 1，而整个逻辑表达式中还有 $n-l-h$ 个无关变量，这些无关变量的组合数目是 2^{n-l-h}，所以由二元组构成的凹超立方体中样本值为“1”的样本个数为 $(2^h-1)\times 2^{n-l-h}$．

在上述求取各组合数的过程中，与变量 $x_{s1}^*, x_{s2}^*, \cdots, x_{sl}^*, x_{t1}^*, x_{t2}^*, \cdots, x_{th}^*$ 的排列无关，并且也与它们的逻辑正态、逻辑负态无关，所以对于符合由二元组 $<l, h>$ 构成的结构样式的任何模式，它们样本值为“1”的样本个数都是一样的。　证毕

定义 11　$\boldsymbol{I} = f^{-1}(1)$ 是 $\boldsymbol{F}_2^n$ 空间的输入样本集，GCN($\boldsymbol{I}$) 是 $\boldsymbol{I}$ 的最大连通度，对于凹超立方体结构系，如果一种结构样式的覆盖度小于或等于 GCN($\boldsymbol{I}$)，则称这种结构样式是对于 $\boldsymbol{I}$ 的一个有效结构样式。$S(\boldsymbol{X}, \boldsymbol{W}, \theta)$ 是一个模式，如果对 $\forall \boldsymbol{X}$，$S(\boldsymbol{X}, \boldsymbol{W}, \theta) = 1$ 都满足 $\boldsymbol{X} \in \boldsymbol{I}$，则称 $S(\boldsymbol{X}, \boldsymbol{W}, \theta)$ 是对于 $\boldsymbol{I}$ 的一个有效模式。

至此，可以给出模式匹配学习的算法描述如下：

(1) 计算凹超立方体中每一种结构样式的覆盖度；

(2) 计算输入样本集的最大连通度；

(3) 删除覆盖度大于最大连通度的结构样式，剩下的是有效结构样式；

(4) 在有效结构样式中寻找有效模式；

(5) 用有效模式对样本集合进行匹配，匹配时覆盖度大的模式优先。

4.3 举例分析

这里用二个例子来说明这种学习算法，一个例子是文献 [7] 中的“圆环结构”问题，它的卡诺图如图 1；另一个问题是卡诺图中的“梯形结构”，如图 2 。采用模式匹配学习算法学习后，分别得到表 1 和表 2 所示的学习结果。

表 1　“圆环结构”学习结果

Neuron	w_{i1}	w_{i2}	w_{i3}	w_{i4}	w_{i5}	w_{i6}	T_i
1	−2	2	1	−2	2	1	7
2	2	−2	−1	−2	2	1	7
3	−2	2	1	2	−2	−1	7
4	2	−2	−1	2	−2	−1	7

表 2　“梯形结构”学习结果

Neuron	w_{i1}	w_{i2}	w_{i3}	w_{i4}	w_{i5}	w_{i6}	T_i
1	3	3	1	0	1	1	4
2	1	3	−1	0	3	−1	7

对于“圆环结构”问题的学习结果是产生了 4 个隐元，隐元数目和文献 [7] 的学习结果相同；这 4 个隐元所对应的凹超立方体是：$f(\boldsymbol{X}) = \overline{x}_1 \wedge x_2 \wedge \overline{x}_4 \wedge x_5 \wedge (x_3 \vee x_6)$，$f(\boldsymbol{X}) = x_1 \wedge \overline{x}_2 \wedge \overline{x}_4 \wedge x_5 \wedge (\overline{x}_3 \vee x_6)$，$f(\boldsymbol{X}) = \overline{x}_1 \wedge x_2 \wedge x_4 \wedge \overline{x}_5 \wedge (x_3 \vee \overline{x}_6)$ 和 $f(\boldsymbol{X}) = x_1 \wedge \overline{x}_2 \wedge x_4 \wedge \overline{x}_5 \wedge (\overline{x}_3 \vee \overline{x}_6)$ 。“梯形结构”的学习结果是产生了 2 个隐元，这 2 个隐元所对应的凹超立方体是：$f(\boldsymbol{X}) = x_1 \wedge x_2 \wedge (x_3 \vee x_5 \vee x_6)$，$f(\boldsymbol{X}) = x_2 \wedge x_5 \wedge (x_1 \vee \overline{x}_3 \vee \overline{x}_6)$ 。这表明在某些情况下，采用凹超立方体结构系的模式匹配学习算法对布尔空间的样本集合进行学习，学习结果的表达效率还是很高的，最重要的是学习结果有明确的逻辑意义，也就是具有“透明性”。

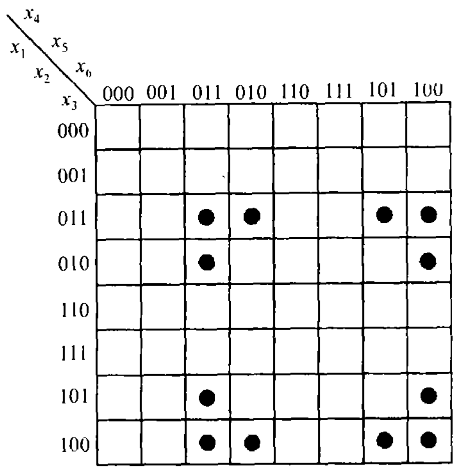

图 1 "圆环结构" 卡诺图

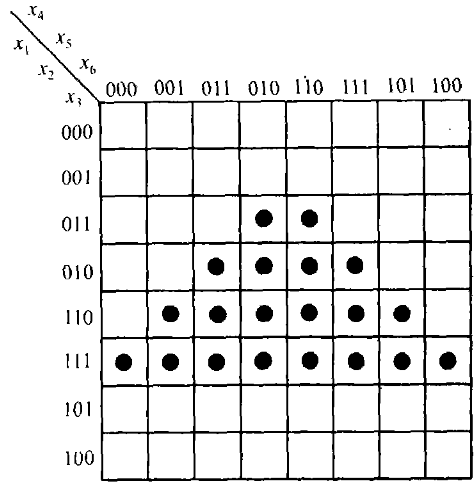

图 2 卡诺图中的 "梯形结构"

4.4 算法复杂度分析

在采用凹超立方体结构系的模式匹配学习算法中，学习的过程主要分三部分：有效结构样式的建立、有效模式的建立和模式匹配．对于 n 维布尔空间有效结构样式建立的运算复杂度是 $O(n^2)$，有效模式建立的运算复杂度是 $O(2^{2n}\times 3^n)$，模式匹配过程的运算复杂度是 $O(2^n)$，所以整个学习算法的运算复杂度是 $O(2^{2n}\times 3^n)$，与文献 [2] 中介绍的几何学习算法的运算复杂度 $O(2^{2n})$ 比较，模式匹配学习算法的运算效率是一个需要改进的地方．

5 结 束 语

模式匹配学习算法的最大益处是得到了一个 "透明" 的学习结果，这对知识提取非常重要．从方法上看，模式匹配学习算法的核心是通过对采用的线性可分结构系进行结构分层，建立有效模式的集合并进行模式匹配；这种方法不仅对凹超立方体结构系适用，也对汉明球系和 PSP 函数系适用，在必要的时候，还可以同时使用几种线性可分结构系，这样虽然增加了学习时间，但最终得到的二进神经网络的表达效率和表达 "清晰" 性会更好．同一个输入样本集合，用不同的线性可分结构系进行模式匹配学习，学习结果也不一样，所以，如何选择线性可分结构系，最终得到最佳的学习结果，是一个值得进一步研究的相关问题．

参 考 文 献

[1] 陆阳，韩江洪，高隽，魏臻，二进神经网络中汉明球的逻辑意义及一般判别方法，计算机研究与发展，2002, 39(1), 79-86.

[2] J. H. Kim, S. Park, The geometrical learning of binary neural networks, IEEE Trans. on Neural Networks, 1995, 6(1), 237-247.

[3] 朱大铭，马绍汉，二进制神经网络分类问题的几何学习算法，软件学报， 1997, 8(8), 622-629.

[4] 马晓敏，杨义先，章照止，一种新的阈函数的分析框架及有关结论，计算机学报， 2000, 23(3), 225-230.

[5] 陆阳，韩江洪，张维勇，二进神经网络逻辑关系判据及等价性规则提取，模式识别与人工智能， 2001, 14(2), 171-176.

[6] 陆阳，韩江洪，高隽，二进神经网络隐元数目最小上界研究，模式识别与人工智能， 2000, 13(3), 254-257.

[7] 马晓敏，杨义先，章照止，二进神经网络学习算法研究，计算机学报，1999, 22(9), 931–935.

THE PATTERN MATCH LEARNING OF BINARY NEURAL NETWORKS

Lu Yang　Wei Zhen　Han Jianghong　Fan Yuqi

(*Computer and Information College, Hefei Univ. of Tech., Hefei* 230009, *China*)

Abstract It is necessary to know the logical meaning of every binary neuron when extracting knowledge from a binary neural network. Generally, it is difficult to analyze learning results of a learning algorithm for binary neural networks. In this paper, a new learning method is presented which is based on analyzing a set of linear separable structures. The most important benefit of this method is all binary neurons belong to one or more types of linear separable structure sets. If those linear separable structure sets have clear logical meaning, the whole knowledge of binary neural networks can be dug out.

Key words Binary neural networks, Linear separability, Pattern match

陆　阳：　男，1967 年生，副研究员，博士生，研究方向是智能控制，模式识别和信号处理．
魏　臻：　男，1965 年生，副研究员，博士生，研究方向是可靠性理论和智能控制．
韩江洪：　男，1954 年生，研究员，博士生导师，研究方向是智能控制，嵌入式系统和计算机信息系统．
樊玉琦：　男，1976 年生，硕士生，研究方向是智能控制和计算机信息系统．