第三章 Hopfield人工神经网络

XI'AN JIAOTONG UNIVERSITY

- 引言
- Hopfield人工神经网络的电路模型及其能量函数
- Hopfield人工神经网络A/D变换器
- Hopfield人工神经网络用于求解TSP
- Hopfield人工神经网络用于求解货流问题
- Hopfield人工神经网络用于联想记忆
- 双向联想记忆
- 小结

§3.1 引言

XI'AN JIAOTONG UNIVERSITY

1982年到1986年，美国物理学家Hopfield在美国科学院院刊上陆续发表了几篇论文，提出了反馈神经网络理论，报导了相应的研究成果，成为新一轮神经网络研究高潮的标志。因此，把这种单层反馈网络称为Hopfield网络。

下表是Hopfield网络和前向多层网络的简单比较。

	前向多层网络	Hopfield 网络
网络连接形式	不含反馈连接	包含反馈连接
输出输入关系	简单映射关系，不考虑滞后效应	要考虑输出输入间的延迟，要用差分或微分方程描述
学习算法	误差修正法（BP 算法），收敛慢	Hebb 规则，收敛快
应用（功能）	分类、联想	分类、联想、优化计算
稳定性理论	分析简单	分析复杂

§3.1 引言

XI'AN JIAOTONG UNIVERSITY

Hopfield的主要贡献有：

<1> 提出了利用能量函数研究反馈网络稳定状态的方法。

<2> 给出了利用模拟电子线路实现反馈型人工神经网络的电路模型。

<3> 成功求解了人工智能的典型难题——TSP问题。

以此为基础，人们对Hopfield网络进行了深入研究，主要有以下几个方面：寻找Hopfield网络的稳定性规律并进而研究其信息容量；提出各种改进的Hopfield网络模型；参照Hopfield电子线路模型研究人工神经网络的硬件实现方法；借助能量函数方法用Hopfield网络求解优化计算、组合数学、人工智能问题的多种实例。

本章主要讨论Hopfield人工神经网络的网络模型，以及其在优化计算、联想记忆方面的应用，并简单介绍Hopfield网络的稳定性问题。

§3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

一、电路模型

Hopfield用模拟电子线路构造了反馈型人工神经网络的电路模型，如图3-1所示。电路中的基本元件如下：

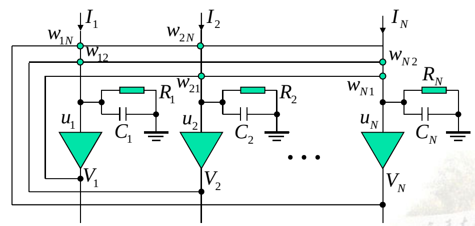


图3-1 Hopfield神经网络电路模型

§3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

1. 用具有正、反向输出的运算放大器模仿神经元的非线性函数关系。第i个神经元输出电压V_i与输入电压u_i之间的关系为S形函数，如图3-2所示。

$$V_i = f_i(u_i) = \frac{1}{2}\left[1+\tanh\left(\frac{u_i}{u_0}\right)\right] = \frac{1}{1+e^{-2u_i/u_0}}$$

u_0是归一化基准值。当 $u_0 \to 0$时，f_i成为硬限幅函数。

2. 输入电容C_i和电阻R_i，决定着电路的时间常数，它使放大器的输出相对于输入有一定的延时，这可以用来模仿神经元的动态特性。

3. 连接电导w_{ij}，是第j个神经元的输出和第i个神经元的输入之间的连接电导，

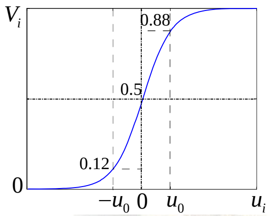


图3-2 神经元特性曲线

§3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

它模仿神经元之间的突触作用。也可以用电阻R_{ij}来表示，$R_{ij}=1/w_{ij}$。

4. 外加偏置电流I_i，模仿神经元的阈值作用。

图3-3是第i个神经元输入节点处元件的相互连接关系，显然，由KCL定律我们可得：

$$\sum_j w_{ij}(V_j - u_i) + I_i$$

$$= C_i\frac{du_i}{dt} + \frac{u_i}{R_i}$$

令

$$\frac{1}{R_i'} = \frac{1}{R_i} + \sum_j w_{ij}$$

得

$$C_i\frac{du_i}{dt} = \sum_j w_{ij}V_j + I_i - \frac{u_i}{R_i'}$$

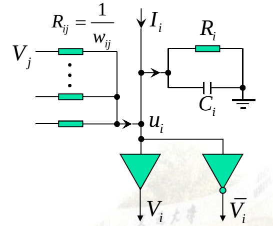


图3-3 神经元输入节点处的元件连接关系

§3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

如果各神经元参数相同，并取$R=R_i'$，$C=C_i=1$；令 $\tau=RC$。则有：

$$\begin{cases} \dfrac{du_i}{dt} = -\dfrac{u_i}{\tau} + \sum_j w_{ij}V_j + I_i \\ V_i = f_i(u_i) \end{cases} \quad i=1,2,\cdots,N \qquad \cdots\cdots\cdots \text{EQ3.5}$$

N个神经元的方程构成 N元联立非线性微分方程组。给定一个初值，可以求出此方程组的解；如果方程组的解为一组确定值，则表明系统演变的最终结果将得到一个稳定状态。所以，具体应用时都需要研究它的稳定性，稳定状态存在的规律及性质。

二、能量函数

Hopfield利用非线性动力学系统理论中的能量函数法（Liapunov第二方法）为网络建立了能量函数，借此研究反馈神经网络的稳定性，并利用此方法建立了求解优化问题的系统方程式。

§3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

$$E = -\frac{1}{2}\sum_{i\neq j}\sum_j w_{ij}V_iV_j + \sum_i \frac{1}{R_i'}\int_0^{V_i} f_i^{-1}(V)\,dV - \sum_i I_iV_i \qquad \cdots\cdots \text{EQ3.6}$$

E具有能量的量纲。

下面证明：E对时间的导数为负值，且达到稳定状态时，E取极小值。

使用复合微分 $\dfrac{dE}{dt} = \sum_i \dfrac{\partial E}{\partial V_i}\bullet\dfrac{dV_i}{dt}$

利用突触权重的对称性$w_{ij}=w_{ji}$，可得

$$\frac{\partial E}{\partial V_i} = -\sum_j w_{ij}V_j + \frac{u_i}{R_i'} - I_i \qquad \cdots\cdots\cdots\cdots \text{EQ3.8}$$

与EQ3.3 $C_i\dfrac{du_i}{dt} = \sum_j w_{ij}V_j + I_i - \dfrac{u_i}{R_i'}$ 比较，有 $\dfrac{\partial E}{\partial V_i} = -C_i\dfrac{du_i}{dt}$ 所以：

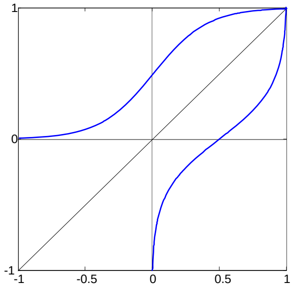
§ 3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

$$\frac{dE}{dt}=-\sum_i C_i\frac{du_i}{dt}\bullet\frac{dV_i}{dt}$$

$$=-\sum_i C_i\frac{du_i}{dV_i}\frac{dV_i}{dt}\bullet\frac{dV_i}{dt}$$

$$=-\sum_i C_i\frac{du_i}{dV_i}\left(\frac{dV_i}{dt}\right)^2$$



其中 $C_i>0$，$(dV_i/dt)^2\ge 0$

由右图，可知 $du_i/dV_i>0$ 。

所以 $dE/dt\le 0$，

且当 $dV_i/dt=0$ 时，$dE/dt=0$ 。

以上结果说明，随着时间的推移，在状态空间中，网络总是朝能量函数 E 减小的方向运动，网络达到稳态时 E 取极小值。

§ 3.2 Hopfield人工神经网络的电路模型及其能量函数

XI'AN JIAOTONG UNIVERSITY

如果神经元非线性函数 f 是硬限幅函数，则EQ3.6中的第二项（积分项）等于零。所以能量函数 E 的表达式可以简化为：

$$E=-\frac{1}{2}\sum_{i\ne j}\sum_j w_{ij}V_iV_j-\sum_i I_iV_i \quad \cdots\cdots\cdots\cdots\cdots \text{EQ3.13}$$

其实，在一般情况下，当 f 为sigmoid函数时，如果 V_i 的值接近于1或0，那么它的变动对积分结果的影响也不大，能量函数也可以近似为上式。

其矩阵形式为：

$$E=-\frac{1}{2}\mathbf{V}^T\mathbf{W}\mathbf{V}-\mathbf{V}^T\mathbf{I} \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.14}$$

设神经元数目为 N，那么式中 $\mathbf{V}$ 是 N 维列向量，$\mathbf{V}^T$ 是 $\mathbf{V}$ 的转置，$\mathbf{W}$ 是 $N\times N$ 的方阵，$\mathbf{I}$ 也是 N 维列向量。

§ 3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

一、Tank与Hopfield人工神经网络A/D变换器

模拟量 x	数字量 V_i
9	1001
8	1000
⋮	⋮

右图为Tank和Hopfield给出的神经网络A/D变换器的原理电路图。要求：当网络稳定以后，其输出 V_i 接近于0或1，且与模拟量 x 满足如下关系：

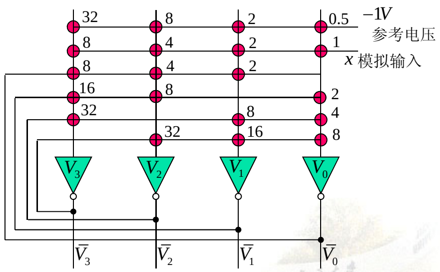


图3-7 人工神经网络A/D变换器

§ 3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

$$x\doteq\sum_{i=0}^{3}2^iV_i \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.25}$$

为了完成此A/D变换电路的设计，需要根据要求建立Hopfield网络的能量函数，并由此计算出网络的各突触电导值 w_{ij}。

A. 因为要求电路的转换功能完成以后，模拟量 x 与数字量 V_i 之间满足EQ3.25，所以能量函数中应该包含下式：

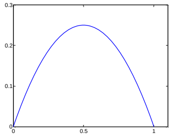
$$\frac{1}{2}\left(x-\sum_{i=0}^{3}2^iV_i\right)^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.26}$$

B. 要求输出 V_i 尽量靠近0或1。由于运放的输出是连续的，为保证输出数字量，能量函数中要附加下式：

$$-\frac{1}{2}\sum_{i=0}^{3}\left(2^i\right)^2V_i\left(V_i-1\right) \quad \cdots\cdots\cdots\cdots\cdots \text{EQ3.27}$$

§3.3 Hopfield神经网络A/D变换器

此式大于等于0，只有当 $V_i=0$ 或1时取最小值0。其中系数的选择是为了保证突触连接电导矩阵 $\mathbf{W}$ 的对角元素为0，即当 $j=i$ 时，$w_{ij}=0$。



将EQ3.26与EQ3.27相加，整理并去除与变量 V_i 无关的项，可得如下的能量函数：

$$E=-\frac{1}{2}\sum_{j=0}^{3}\sum_{\substack{i=0\\i\neq j}}^{3}\left(-2^{i+j}\right)V_iV_j-\sum_{i=0}^{3}\left(-2^{(2i-1)}+2^i x\right)V_i \qquad \cdots\cdots \text{EQ3.28}$$

More...

将此式与“标准的”Hopfield能量函数表达式

$$E=-\frac{1}{2}\sum_{i\neq j}\sum_{j}w_{ij}V_iV_j-\sum_{i}I_iV_i \qquad \cdots\cdots \text{EQ3.13}$$

进行比较，可得电路中的电导值为：$w_{ij}=-2^{(i+j)}$

$$I_i=-2^{(2i-1)}+2^i x$$

§3.3 Hopfield神经网络A/D变换器

其中，w_{ij} 表示各突触的电导值，利用运放的反相输出可以实现式中的负号。神经元偏置电流表达式由两项组成，第一项表示参考电压源和运放输入之间连接的电导值，若参考电压为 $-1V$，则电导值为 $2^{(2i-1)}$；第二项表示模拟输入到各运放所经过的电导值，即 2^i。显然这是一组归一化的电导值。

实际电路中还需要做以下修正：

(1) 当运放输出电压为0至 V_{BB} 伏时，有

$$w_{ij}=-2^{(i+j)}/V_{BB}$$

(2) 当参考电压源为 $-V_R$ 时，相应的一组电导值为：

$$2^{(2i-1)}/V_R$$

(3) 当输入信号 x 的最大值为 V_H 时，相应的一组电导值为：

$$\frac{2^i}{\frac{V_H}{16}}=\frac{2^{4+i}}{V_H}$$

§3.3 Hopfield神经网络A/D变换器

神经网络A/D变换器的优点包括：(a) 由于它是并行工作的，所以比一般的串行A/D变换器速度快；(b) 比一般的并行A/D变换器结构简单，需要的运放少。

这种电路的缺点是存在迟滞（hysteresis）现象。

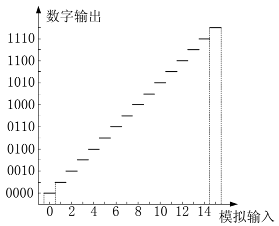


图3-8 正确工作情况下的A/D转换关系

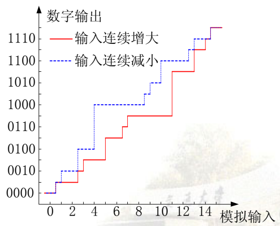


图3-9 迟滞现象

§3.3 Hopfield神经网络A/D变换器

出现迟滞现象的原因是：在给定输入模拟量的情况下，该电路的能量函数存在多个极小值，其中只有一个是全局最小值，其余为局部极小值，当系统收敛到局部极小解时，就出现错误映射。

解决办法：

(1) Tank与Hopfield指出，如果在每次变换之前先输入零，则可以避免（减小）迟滞现象。但电路复杂度增大，工作速度降低。

(2) 利用非对称Hopfield人工神经网络。

张宁、郑君里，“非对称人工神经网络A/D变换器”，电路与系统学报，1991，2。

§3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

二、非对称Hopfield神经网络及相应的A/D变换器

（一）非对称Hopfield网络结构

网络结构如右图所示。可以看到，网络结构有如下特点：高位输出只对低位输入有反馈，而低位输出对高位没有影响。网络的这个特点使得该网络总是稳定的，且不存在局部极小值。下面我们将证明这一点。

对神经元的要求：单调连续，权值 $T_{ij} \ge 0$。

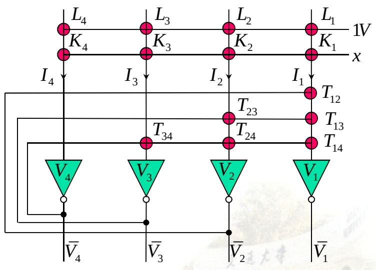


图3 非对称Hopfield网络A/D变换器

§3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

（二）关于网络稳定性和唯一性的证明

证明中要用到如下的定理。

【定理】 考虑微分方程

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y^{(1)} + a_n y = \phi(t)$$

其中 $a_1, a_2, \cdots, a_n$ 为实常数，$a_n \ne 0$，$\phi(t)$为连续函数，若方程满足：

(a) $\lim_{t\to\infty} \phi(t) = b$；

(b) 微分方程的特征方程的根具有负实部，

那么它的所有解当 $t \to \infty$ 时，趋于 b/a_n。

1. 稳定性

对于图示的网络来说，根据KCL，其运动方程为

$$C_i \frac{du_i}{dt} = -\sum_{j=i+1}^{N} T_{ij} V_j + I_i - \frac{u_i}{R_i} \qquad \cdots\cdots (1)$$

§3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

其中$T_{ij} \ge 0$，$i = 1, 2, \cdots, N$。

当$i = N$ 时，有

$$C_N \frac{du_N}{dt} = I_N - \frac{u_N}{R_N} \qquad \cdots\cdots (2)$$

由于$R_N > 0, C_N > 0$，根据上述定理（$a_n = \dfrac{1}{R_N}, b = I_N$）有 $t \to \infty$时，(2)式的稳态解为

$$u_N^e = R_N I_N \qquad \text{其中上标 } e \text{ 代表平衡状态。}$$

当$i = N-1$时，有

$$C_{N-1} \frac{du_{N-1}}{dt} = -T_{N-1,N} V_N + I_{N-1} - \frac{u_{N-1}}{R_{N-1}} \qquad \cdots\cdots (3)$$

由于 $V_N = g(u_N)$，其中 $g(\bullet)$是神经元功能函数，且$g(u)$是单调连续的。

§3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

所以：$\lim_{t\to\infty} V_N = g(u_N^e)$

根据上述定理，有

$$\lim_{t\to\infty} u_{N-1} = -R_{N-1} T_{N-1,N}\, g(u_N^e) + R_{N-1} I_{N-1}$$

同理可证，对 $i = 1, 2, \cdots, N$，方程(1)的稳态解为

$$u_i^e = -R_i \sum_{j=i+1}^{N} T_{ij}\, g(u_j^e) + R_i I_i \qquad \cdots\cdots (4)$$

可见方程(1)的解当 $t \to \infty$ 时是稳定的，且与初始条件无关。

2. 唯一性

当方程(1)达到稳定时，$C_i \dfrac{du_i}{dt} = 0$，那么从方程(1)显然可以得到其平衡状态的解为：

$$\frac{u_i^e}{R_i} = -\sum_{j=i+1}^{N} T_{ij}\, g(u_j^e) + I_i \qquad \cdots\cdots (5)$$

§ 3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

由于 I_i 对于某一具体输入是确定的，因此方程(1)的解是唯一的。比较(4)、(5)两式，可见两者是一致的，所以由方程(1)所描述的网络的解是唯一的，且解的吸引域为整个空间。

（三）非对称Hopfield网络的优点

1. 不存在局部极小点问题；2. 该系统总是稳定的。

（四）在A/D变换中的应用

为了用这个网络实现A/D变换，必须在网络输出与模拟输入之间建立一定的关系，使这些输出能代表模拟输入的值。设图中运放的输出高电平为 A 伏，低电平为 B 伏，模拟输入的动态范围为 C 伏，那么网络输出 V_j 应满足：

$$\frac{2^N-1}{C}x-0.5\le\sum_{j=1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}\le\frac{2^N-1}{C}x+0.5 \qquad \cdots\cdots\cdots (6)$$

其中 $K=A-B$，N 为A/D变换器的位数，展开(6)式

§ 3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

$$\frac{2^N-1}{C}x-0.5\le\sum_{j=1}^{i-1}\left(V_j-B\right)\frac{2^{j-1}}{K}+\left(V_i-B\right)\frac{2^{i-1}}{K}+\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}\le\frac{2^N-1}{C}x+0.5$$

$$\frac{2^N-1}{C}x-0.5-\sum_{j=1}^{i-1}\left(V_j-B\right)\frac{2^{j-1}}{K}\le\left(V_i-B\right)\frac{2^{i-1}}{K}+\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}\le\frac{2^N-1}{C}x+0.5$$

因为 $\sum_{j=1}^{i-1}\left(V_j-B\right)\frac{2^{j-1}}{K}\le\sum_{j=1}^{i-1}2^{j-1}=2^{i-1}-1$ ，所以

$$\frac{2^N-1}{C}x+0.5-2^{i-1}-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}\le\left(V_i-B\right)\frac{2^{i-1}}{K} \qquad \cdots\cdots (7)$$

$$\le\frac{2^N-1}{C}x+0.5-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}$$

定义如下决策函数：

§ 3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

$$u_{ii}=\frac{2^N-1}{C}x+0.5-2^{i-1}-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K} \qquad \cdots\cdots\cdots (8)$$

V_i 和决策函数有如下关系。当 $V_i=A$ 时，根据(7)式的右半不等式，有

$$\frac{2^N-1}{C}x+0.5-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}\ge\left(V_i-B\right)\frac{2^{i-1}}{K}=2^{i-1}$$

$$\Rightarrow\frac{2^N-1}{C}x+0.5-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}-2^{i-1}\ge0\Rightarrow u_{ii}\ge0$$

当 $V_i=B$ 时，根据(7)式的左半不等式，有

$$\frac{2^N-1}{C}x+0.5-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}-2^{i-1}\le0\Rightarrow u_{ii}\le0$$

所以有 $\begin{cases}\text{当}V_i=A\text{时},u_{ii}\ge0\\ \text{当}V_i=B\text{时},u_{ii}\le0\end{cases}$

§ 3.3 Hopfield神经网络A/D变换器

XI'AN JIAOTONG UNIVERSITY

反过来，当 $u_{ii}>0$ 时，有

$$0<u_{ii}=\frac{2^N-1}{C}x+0.5-2^{i-1}-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}$$

$$=\frac{2^N-1}{C}x-0.5+1-2^{i-1}-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}$$

$$\le\sum_{j=1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}+1-2^{i-1}-\sum_{j=i+1}^{N}\left(V_j-B\right)\frac{2^{j-1}}{K}$$

$$=\sum_{j=1}^{i-1}\left(V_j-B\right)\frac{2^{j-1}}{K}+1-2^{i-1}+\left(V_i-B\right)\frac{2^{i-1}}{K}$$

$$\le2^{i-1}-1+1-2^{i-1}+\left(V_i-B\right)\frac{2^{i-1}}{K}\Rightarrow V_i=A$$

§3.3 Hopfield神经网络A/D变换器

而当 $u_{ii}<0$ 时，有

$$0>u_{ii}=\frac{2^N-1}{C}x+0.5-2^{i-1}-\sum_{j=i+1}^{N}(V_j-B)\frac{2^{j-1}}{K}$$

$$\ge\sum_{j=1}^{N}(V_j-B)\frac{2^{j-1}}{K}-2^{i-1}-\sum_{j=i+1}^{N}(V_j-B)\frac{2^{j-1}}{K}$$

$$=\sum_{j=1}^{i-1}(V_j-B)\frac{2^{j-1}}{K}+(V_i-B)\frac{2^{i-1}}{K}-2^{i-1}$$

$$\ge(V_i-B)\frac{2^{i-1}}{K}-2^{i-1}\Rightarrow V_i=B$$

所以有 $\begin{cases}当u_{ii}>0时, V_i=A\\ 当u_{ii}<0时, V_i=B\end{cases}$

§3.3 Hopfield神经网络A/D变换器

由此可见，u_{ii}与 V_i的对应关系和u_i与V_i的对应关系是一致的。因此决策函数u_{ii}就决定了A/D变换器第 i 位取值所应遵循的规则，也就确定了网络方程的稳定解。根据这一点就可以导出相应的网络参数。

网络的稳态解为

$$u_i^e/R_i=-\sum_{j=i+1}^{N}T_{ij}g\left(u_j^e\right)+I_i$$

比较可得：

$$\begin{cases}T_{ij}=2^{j-1}/K & j=i+1,\cdots,N\\ I_i=\dfrac{2^N-1}{C}x+0.5-2^{i-1}+\dfrac{B}{K}\left(2^N-2^i\right) & i=1,2,\cdots,N\end{cases}$$

由式中的 I_i 可以确定：

$$K_i=\frac{2^N-1}{C}\qquad L_i=0.5-2^{i-1}+\frac{B}{K}\left(2^N-2^i\right)\qquad i=1,2,\cdots,N$$

§3.4 Hopfield网络用于求解TSP

一、TSP命题

TSP是Travelling Salesman Problem的缩写，中文称为“旅行商问题”，它是一个经典的人工智能难题。

TSP命题是：设有 N 个城市A、B、C…，它们之间的距离分别为 d_{AB}、d_{AC} 、… d_{BC} …。问题是寻找一条闭合路径，此路径经过每个城市且只经过一次，最后返回起始城市，要求路径的长度最短。

如图3-4是 $N=10$ 时的一种可能路径。其路径长度为

$$d=d_{AJ}+d_{JC}+d_{CB}+\cdots+d_{EA}$$

不难求得，对于给定的 N 可能存在的路径

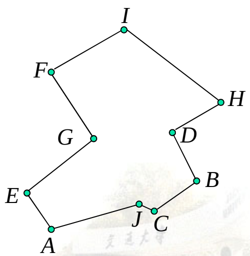


图3-4 10个城市的TSP路径示意

§3.4 Hopfield网络用于求解TSP

数为$\frac{1}{2}(N-1)!$。如果按照传统的方法，先计算出所有可能路径的长度然后再搜索最优路径，那么随着 N 的增大，计算量将急剧增大。

实际应用中，这类问题往往不需要得到严格的最优解，近似最优、准最优的解也是可以接受的。例如，图3-4中的路径也许就是准最优的。

二、求解TSP的网络模型

1. TSP状态与结果的表示

如果用行表示城市，用列表示步数，那么我们就可以用表3-1所示的矩阵来表示TSP的状态与结果。例如，表中的矩阵表示的路径顺序是CAEBD，路径总长度为

$$d=d_{CA}+d_{AE}+d_{EB}+d_{BD}+d_{DC}$$

表3-1 置换矩阵

	1	2	3	4	5
A	0	1	0	0	0
B	0	0	0	1	0
C	1	0	0	0	0
D	0	0	0	0	1
E	0	0	1	0	0

§3.4 Hopfield网络用于求解TSP

这种矩阵称为置换矩阵（permutation matrix），它具有以下特点：

(1) 每行只有一个“1”，其余元素为“0”；

(2) 每列只有一个“1”，其余元素为“0”；

(3) 全部元素中“1”的总和为 N。

2. 网络模型

这样，我们就可以用$N\times N$个神经元组成Hopfield神经网络，当网络达到稳定状态时，各神经元的状态对应置换矩阵的各元素值（“1”或“0”）。各城市间的距离 d_{AB}、d_{AC}、...等作为一组约束信息确定神经元之间的连接强度 w_{ij}。我们期望网络演变的最终结果给出最优解，即网络给出的置换矩阵所对应的路径是最短路径。

为了便于和置换矩阵对应，神经元的输出电压用双下标表示，如V_{Xi}，其中第一个下标X代表城市名（如A，B等），第二个下标i代表路径顺序（如1，2等）。相应地，放大器的输入电压也用双下标表示，记为 u_{Xi}。

§3.4 Hopfield网络用于求解TSP

为了完成网络电路的设计，需要求出其中的连接电导和偏置电流。因此，我们要根据优化目标构造能量函数，所建立的能量函数应该满足以下两方面的要求：(1) 网络稳定时，也即能量函数达到最小值时，神经元的输出符合置换矩阵的形式；(2) 网络稳定时，置换矩阵对应的路径最短。

为满足(1)，能量函数中应包括下式：

$$\frac{\alpha}{2}\sum_X\sum_i\sum_{j\neq i}V_{Xi}V_{Xj}+\frac{\beta}{2}\sum_i\sum_X\sum_{Y\neq X}V_{Xi}V_{Yi}+\frac{\gamma}{2}\left(\sum_X\sum_i V_{Xi}-N\right)^2 \quad \cdot\cdot\ \text{EQ3.17}$$

式中的α、β、γ 为实系数。其中第一项是置换矩阵所有各行每行内各元素两两交叉相乘（不包括自乘）的各种组合，此项是为了保证使置换矩阵的各行中1的个数不大于“1”；第二项是为了保证使置换矩阵的各列中“1”的个数不大于1；第三项是为了保证置换矩阵中共有N个“1”。当EQ3.17取最小值0时，网络的输出符合置换矩阵形式，其所代表的路径是一条“合法”路径。

§3.4 Hopfield网络用于求解TSP

为了满足(2)，能量函数应包含下式

$$\frac{\lambda}{2}\sum_X\sum_{Y\neq X}\sum_i d_{XY}V_{Xi}\left(V_{Y,i+1}+V_{Y,i-1}\right) \quad \cdots\cdots\cdots\ \text{EQ3.18}$$

上式表示的是某条“合法”路径的长度，我们希望它越小越好。要注意的是，上式中的下标按“模N”规律确定，即 $V_{Y,N+j}=V_{Y,j}$。

将上两式结合，得到能量函数表达式为：

$$E=\frac{\alpha}{2}\sum_X\sum_i\sum_{j\neq i}V_{Xi}V_{Xj}+\frac{\beta}{2}\sum_i\sum_X\sum_{Y\neq X}V_{Xi}V_{Yi}+\frac{\gamma}{2}\left(\sum_X\sum_i V_{Xi}-N\right)^2+\frac{\lambda}{2}\sum_X\sum_{Y\neq X}\sum_i d_{XY}V_{Xi}\left(V_{Y,i+1}+V_{Y,i-1}\right) \quad \cdot\cdot\cdot\ \text{EQ3.19}$$

把此式和“标准的”Hopfield神经网络能量函数（EQ3.13）相比较，就可以求得连接权值和阈值。

§3.4 Hopfield网络用于求解TSP

$$E=-\frac{1}{2}\sum_{i\neq j}\sum_j w_{ij}V_iV_j-\sum_i I_iV_i \quad \cdots\cdots\cdots\cdots\ \text{EQ3.13}$$

为了方便比较，我们把“标准”能量函数改写为双下标的形式：

$$E=-\frac{1}{2}\sum_X\sum_i\sum_Y\sum_j w_{Xi,Yj}V_{Xi}V_{Yj}-\sum_X\sum_i I_{Xi}V_{Xi}$$

More...

为了清楚地表示比较结果，引用如下符号：

$$\delta_{ij}=\begin{cases}1 & (i=j)\\ 0 & (i\neq j)\end{cases} \quad \cdots\cdots\cdots\cdots\ \text{EQ3.20}$$

通过比较，我们可以得到：

$$w_{Xi,Yj}=-\alpha\delta_{XY}\left(1-\delta_{ij}\right)-\beta\delta_{ij}\left(1-\delta_{XY}\right)-\gamma-\lambda d_{XY}\left(\delta_{j,i+1}+\delta_{j,i-1}\right) \quad \cdots\cdots\ \text{EQ3.21}$$

$$I_{Xi}=N\gamma \quad \cdots\cdots\cdots\cdots\ \text{EQ3.22}$$

§ 3.4 Hopfield网络用于求解TSP

XI'AN JIAOTONG UNIVERSITY

下面导出描述TSP网络的状态方程式，由EQ3.9，并令 $C = C_i$，有

$$-\frac{\partial E}{\partial V_i} = C\frac{du_i}{dt}$$

考虑到神经元非线性函数不一定是硬限幅函数，需要在能量函数表达式中增加 $f_i^{-1}(V)$ 函数积分项，最后求得：

$$\left.\begin{aligned}\frac{du_{Xi}}{dt} &= -\frac{u_{Xi}}{\tau} - \alpha\sum_{j\neq i} V_{Xj} - \beta\sum_{Y\neq X} V_{Yi} - \gamma\left(\sum_X\sum_i V_{Xi} - N\right) \\ &\quad - \lambda\sum_{Y\neq X} d_{XY}\left(V_{Y,i+1} + V_{Y,i-1}\right) \\ V_{Xi} &= f(u_{Xi}) = \frac{1}{2}\left[1 + \tanh\left(\frac{u_{Xi}}{u_0}\right)\right]\end{aligned}\right\} \quad \cdots\cdots \text{EQ3.23}$$

这是 $N\times N$ 个神经元状态方程式的通用表达式。为求TSP的优化结果，需要求解 $N\times N$ 个非线性一阶联立微分方程组。

§ 3.4 Hopfield网络用于求解TSP

XI'AN JIAOTONG UNIVERSITY

三、仿真实验

Hopfield给出了TSP的计算机仿真实例。实验中参数选择如下：

$$N = 10 \quad \tau = 1 \quad u_0 = 0.02$$

$$\alpha = \beta = 500 \quad \gamma = 200 \quad \lambda = 500$$

初始条件为：$u_{Xi} = u_{00} + \delta u_{Xi}$ ，

u_{00}满足 $t = 0$时，$\sum_X\sum_i V_{Xi} = 10$ 以利于收敛；

$-0.1u_0 \le \delta u_{Xi} \le 0.1u_0$ 为随机噪声。

利用数值计算方法对此微分方程组求解，经过若干次迭代即可求得各神经元的最终状态。给定不同的 δu_{Xi} ，就可能得到不同的最终状态，其中有一些可能是最优的，另一些可能是准最优的。图3-6是两种可能的解。

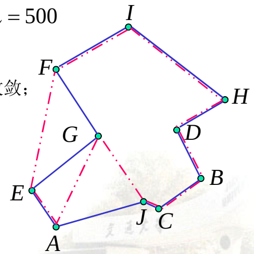


图3-6 两种可能的解

§ 3.4 Hopfield网络用于求解TSP

XI'AN JIAOTONG UNIVERSITY

关于仿真实验的几点讨论：

1. u_0的选择对计算结果有较大的影响。太小的 u_0将使优化路径过于粗略，所得结果与最优状态相差较远。而太大的 u_0则使最终的 V_{Xi} 难以接近1或0，即难以得到正确的置换矩阵形式。

2. 系数$\alpha, \beta, \gamma, \lambda$的选择也很重要，其中$\lambda$的选择更为重要。

3. 实验结果表明，选择20组不同的δu_{Xi}进行实验，其中大部分结果收敛于接近最优的路径，包括最优路径。对于 $N = 10$ ，共有181440条路径，所以这种算法改善计算效率为：$20/181440 \approx 10^{-5}$。

四、Hopfield网络求解TSP的意义、存在的问题及影响

利用Hopfield网络求解TSP的成功引起了人们极大的兴趣，但也存在着一些问题。

§ 3.4 Hopfield网络用于求解TSP

XI'AN JIAOTONG UNIVERSITY

1. 可以得到最优或接近最优的解，与人脑分析这种问题的特点相似。

2. 借助计算机进行仿真，只能证实这一方法的可行性，不一定能改善计算速度。

3. 如果用硬件实现，则解题时间仅是运放过渡特性时间常数的量级，非常快。而目前硬件实现的进展不顺利。

4. 按照Hopfield提出的方法计算TSP时，收敛效果并不好。已有很多文章从不同角度对此方法进行了改进。

5. Hopfield网络的构成原理既不同于数字计算机，也不同于传统的模拟计算机，它在结构原理和功能效果方面更接近人脑的特征。在解决TSP以外的许多优化问题中都取得了一定的进展，提供了一种解决优化问题的新思路，表现出了引人注目的潜在能力。

6. TSP属于NP难题，Hopfield网络求解TSP的方法对众多NP问题的研究有一定的启发。

§3.4 Hopfield网络用于求解TSP

五、NP理论简介

在研究算法时，需要研究算法的“时间复杂性”与“空间复杂性”。在研究算法时间复杂性的过程中逐步形成了NP理论，它已成为组合数学研究领域中的重要组成部分。

假定计算对象的规模为N，那么算法所需的时间将随N的增加而增大，其增长规律可能有以下几种类型：

(1) 对数型：$\log N$

(2) 多项式型：$N, N^2, N^3, \cdots$

(3) 指数型：$2^N, 3^N, 4^N, \cdots$

(4) 阶乘型：$N!$

如$T(N)=8N^3+5N+6$，算法的时间复杂性为$O(N^3)$，其中$O(\cdot)$表示取数量级。而FFT算法的时间复杂性为$O(N\log_2 N)$。

§3.4 Hopfield网络用于求解TSP

通常认为算法的时间复杂性存在多项式界时，计算时间是可以接受的，如果超出多项式界，那么计算时间将随N的增大而急剧增长，则该算法在实际应用中很难被接受。一般情况下有如下结论：

$$O(\log N)<O(N)<O(N\log N)<O(N^2)<\cdots<O(2^N)$$
$$<O(3^N)<\cdots<O(N!)$$

NP理论指出：如果算法的时间复杂性存在多项式界(Polynomial bounded)，则称为P类。如FFT算法。有许多问题，其算法的时间复杂度肯定超出多项式界，如求一个图的所有树。此外，还有许多问题未能证明它超出多项式界，但又未找到有效算法，称之为“非确定型多项式(Nondeterministic Polynomial)问题”，简称NP问题。当然，这不是NP问题的严格定义。

NP理论不期望寻找NP问题的具体算法，而是着重证明这类问题相互等价性，即证明它们的困难程度相当，如果其中一个问题获得多项式解法，则此类问题全部获得解决。

§3.5 Hopfield网络用于求解货流问题

一、以二值状态实现数字编码的几种方式

在某些优化计算中，输出结果需要连续取值。这可以利用连续时间连续幅度工作下的运放来实现，也可以用二值数据编码来实现。

1. 二进制

是一种一对一的映射，无冗余，虽然节省神经元的数目，但容错性较差。

$$N_d=\sum_{k=1}^{q}2^{k-1}V_k \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.47}$$

如 $5\Rightarrow 101 \qquad 7\Rightarrow 111$

2. 简单求和制

是一种一对多的映射，和二进制相比，表示相同的数字需要更多的神经元，但容错性好。

$$N_d=\sum_{i=1}^{q}V_i \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.48}$$

§3.5 Hopfield网络用于求解货流问题

如 $5\Rightarrow 1010111, 0011111, 1101011, \cdots \quad 7\Rightarrow 1111111$

3. 分组求和加权制

是上述两种方案的折衷。把q bit分成K组，每组M bit，$q=KM$。

每组内使用简单求和制，各组之间使用$M+1$进制的加权求和制。

$$N_d=\sum_{k=1}^{K}\left[(M+1)^{k-1}\sum_{i=1}^{M}V_{(k-1)M+i}\right] \quad \cdots\cdots\cdots\cdots \text{EQ3.46}$$

例如：$q=6, K=2, M=3$ 则

$$100100=4^1\times(1+0+0)+4^0\times(1+0+0)=5$$
$$5\Rightarrow 100001, 001100, 010010, \cdots$$

当$M=1$、$K=q$时，即每组只有一位，相当于不分组，可得二进制表达式EQ3.47；而当$M=q$、$K=1$时，即只分一组，此时可得简单求和制的表达式EQ3.48。

§3.5 Hopfield网络用于求解货流问题

XI'AN JIAOTONG UNIVERSITY

二、货流问题描述

货流问题（Hitchcock Problem）又称为运输问题。

设有m个货源 $x=1,2,\cdots,m$，每个源储存的（可供应的）物资量为S_x。有n个用户 $y=1,2,\cdots,n$，每个用户需要的物资量为D_y。从源 x 到用户y 运送单位物资量的消耗费用为C_{xy}。在以上条件下，确定一组从 x到y运送物资的流量 f_{xy}，使其满足供销平衡且总的消耗费用最小。

可以用数学语言描述如下：

$$\min\left[\sum_{x=1}^{m}\sum_{y=1}^{n}C_{xy}f_{xy}\right] \qquad \cdots\cdots\cdots\cdots \text{EQ3.49}$$

$$\sum_{y=1}^{n}f_{xy}=S_x \qquad x=1,2,\cdots,m \qquad \cdots\cdots\cdots \text{EQ3.50}$$

$$\sum_{x=1}^{m}f_{xy}=D_y \qquad y=1,2,\cdots,n \qquad \cdots\cdots\cdots \text{EQ3.51}$$

§3.5 Hopfield网络用于求解货流问题

XI'AN JIAOTONG UNIVERSITY

图3-11是$m=4$、$n=5$情况下的供销关系示意图。

假定已知消耗矩阵如下表：

x \ y	1	2	3	4	5
1	5	1	7	3	3
2	2	3	6	9	5
3	6	4	8	1	4
4	3	2	2	2	4

经分析计算可以得到如下的流量矩阵：

		D_1	D_2	D_3	D_4	D_5
		2	7	3	2	4
S_1	5	0	5	0	0	0
S_2	3	2	1	0	0	0
S_3	4	0	0	0	2	2
S_4	6	0	1	3	0	2

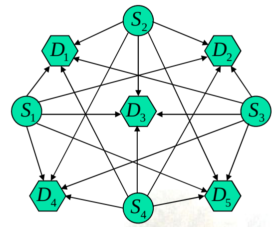


图3-11 货流问题供销关系

§3.5 Hopfield网络用于求解货流问题

XI'AN JIAOTONG UNIVERSITY

问题是：如何用人工神经网络来求解这一问题？

三、用Hopfield网络求解货流问题

流量矩阵中的每个元素值以 q个神经元的状态按数字编码（分组加权求和）的方式表示，因此求解货流问题需要的神经元个数$N=qmn$。我们用三个下标 x、y、j 确定神经元的位置，其中xy表示流量矩阵中的元素序号，而j表示神经元在该元素的 q个位置中的序号。如下表所示：

	$y=1$				$y=2$				$\cdots$				$y=n$			
	j				j				$\cdots$				j			
	q	$\cdots$	2	1	q	$\cdots$	2	1					q	$\cdots$	2	1
$x=1$																
$x=2$																
$\vdots$																
$x=m$																

$V_{2,1,2}$ $f_{1,2}$

§3.5 Hopfield网络用于求解货流问题

XI'AN JIAOTONG UNIVERSITY

引用EQ3.46，用分组求和加权法表示流量 f_{xy}：

$$f_{xy}=\sum_{k=1}^{K}\left[(M+1)^{k-1}\sum_{i=1}^{M}V_{xy,(k-1)M+i}\right] \qquad \cdots\cdots\cdots\cdots \text{EQ3.52}$$

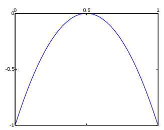
由于矩阵元素是用分组求和加权制表示的，所以下标 $j=(k-1)M+i$。

下面我们根据优化目标构造能量函数。

(1) 神经元的输出要尽量接近0或1，即二值化，能量函数应包括：

$$-\frac{A}{2}\sum_{x=1}^{m}\sum_{y=1}^{n}\sum_{k=1}^{K}\sum_{i=1}^{M}(M+1)^{k-1}\left[1-2V_{xy,(k-1)M+i}\right]^2$$

借用函数 $-(1-2V)^2 \quad (0\le V\le 1)$来满足这个要



(2) 供货量约束

所有用户对某一源的需求应为源的供货量。

§ 3.5 Hopfield网络用于求解货流问题

$$\frac{B}{2}\sum_{x=1}^{m}\left[S_x-\sum_{y=1}^{n}\sum_{k=1}^{K}\sum_{i=1}^{M}(M+1)^{k-1}V_{xy,(k-1)M+i}\right]^2$$

(3) 需货量约束，所有货源对某一用户的供货量等于这个用户的需货量。

$$\frac{C}{2}\sum_{y=1}^{n}\left[D_y-\sum_{x=1}^{m}\sum_{k=1}^{K}\sum_{i=1}^{M}(M+1)^{k-1}V_{xy,(k-1)M+i}\right]^2$$

(4) 消耗费用最小约束。

$$\frac{D}{2}\left[\sum_{x=1}^{m}\sum_{y=1}^{n}\sum_{k=1}^{K}\sum_{i=1}^{M}C_{xy}(M+1)^{k-1}V_{xy,(k-1)M+i}\right]^2$$

由于消耗费用恒为正值，所以上式中的平方可以去掉。以上四项相加，即为完整的能量函数，它与“标准”的Hopfield能量函数（EQ3.13）比较，就可以得到网络中的连接权值（电导）和阈值（偏置电流）。

§ 3.5 Hopfield网络用于求解货流问题

为了方便比较，将EQ3.13式改写为与本问题相同的形式：

$$E=-\frac{1}{2}\sum_{x=1}^{m}\sum_{y=1}^{n}\sum_{k=1}^{K}\sum_{i=1}^{M}\sum_{x'=1}^{m}\sum_{y'=1}^{n}\sum_{k'=1}^{K}\sum_{i'=1}^{M}w_{xy,(k-1)M+i;x'y',(k'-1)M+i'}V_{xy,(k-1)M+i}V_{x'y',(k'-1)M+i'}$$

$$-\sum_{x=1}^{m}\sum_{y=1}^{n}\sum_{k=1}^{K}\sum_{i=1}^{M}V_{xy,(k-1)M+i}\bullet I_{xy,(k-1)M+i}$$

借助上一节定义的δ_{ij} 符号，通过比较可得：

$$\begin{aligned}w_{xy,(k-1)M+i;x'y',(k'-1)M+i'}&=4A(M+1)^{k-1}\delta_{xx'}\delta_{yy'}\delta_{kk'}\delta_{ii'}\\&-B(M+1)^{k+k'-2}\delta_{xx'}-C(M+1)^{k+k'-2}\delta_{yy'}\\&-D(M+1)^{k+k'-2}C_{xy}C_{x'y'}\end{aligned} \quad \cdots\cdot \text{EQ3.55}$$

$$I_{xy,(k-1)M+i}=-2A(M+1)^{k-1}+B(M+1)^{k-1}S_x+C(M+1)^{k-1}D_y \quad \text{EQ3.56}$$

§ 3.5 Hopfield网络用于求解货流问题

如果令$M=1$、$K=q$，那么可以得到二进制编码方式的连接权和阈值：

$$\begin{aligned}w_{xyk;x'y'k'}&=4A2^{k-1}\delta_{xx'}\delta_{yy'}\delta_{kk'}\\&-B2^{k+k'-2}\delta_{xx'}-C2^{k+k'-2}\delta_{yy'}-D2^{k+k'-2}C_{xy}C_{x'y'}\end{aligned} \quad \cdots \text{EQ3.58}$$

$$I_{xyk}=-A2^{k}+B2^{k-1}S_x+C2^{k-1}D_y \quad \cdots\cdots\cdots\cdots \text{EQ3.59}$$

如果令$M=q$、$K=1$，那么就可以得到简单求和制编码方式的连接权和阈值：

$$w_{xyi;x'y'i'}=4A\delta_{xx'}\delta_{yy'}\delta_{ii'} \quad \cdots\cdots\cdots\cdots \text{EQ3.60}$$

$$-B\delta_{xx'}-C\delta_{yy'}-DC_{xy}C_{x'y'} \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.61}$$

$$I_{xyi}=-2A+BS_x+CD_y$$

§ 3.6 Hopfield网络用于联想记忆

Hopfield网络除用于优化计算外，也可以用于联想记忆或分类。在上一章，我们曾讨论过用前向多层神经网络实现异联想，和分类的情况类似，异联想可以用一种映射关系来描述。 $X^{(s)}\rightarrow Y^{(s)}$, $X^{(s)}+\Delta X^{(s)}\rightarrow Y^{(s)}$

本节我们以印刷体字符识别为例，讨论Hopfield网络用于联想记忆的原理、学习和识别的算法步骤。

待识别的8个标准字符如图3-12所示。

样本数$M=8$，每个字符用120个黑白象素（12×10）表示。我们可以用列矢量$\mathbf{X}$来描述字符。

$$\left.\begin{aligned}&\mathbf{X}=[x_0\ x_1\ \cdots\ x_{N-1}]^T\\&\mathbf{X}\in\{-1,1\}^N\end{aligned}\right\} \quad \cdots \text{EQ3.62}$$

其中，$N=120$为矢量维数。

图3-12 8个标准字符

§3.6 Hopfield网络用于联想记忆

网络模型如图3-13所示。利用标准样本对网络进行训练（学习），可以求得网络的突触权重连接矩阵 $\mathbf{W}$，从而使网络可以完成联想记忆的功能。

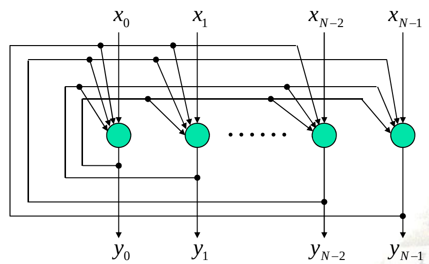


图3-13 Hopfield网络模型

§3.6 Hopfield网络用于联想记忆

算法步骤如下：

① 利用Hebb规则学习，求 w

$$w_{ij} = \begin{cases} \sum_{s=0}^{M-1} x_i^{(s)} x_j^{(s)}, & i \neq j \\ 0, & i = j \end{cases} \qquad 0 \le \begin{Bmatrix} i \\ j \end{Bmatrix} \le N-1 \quad \cdots\cdots \text{EQ3.63}$$

② 利用未知输入样本初始化

$$y_i(0) = x_i \qquad 0 \le i \le N-1 \quad \cdots\cdots\cdots\cdots \text{EQ3.64}$$

其中 $y_i(t)$ 表示节点 i 在 t 时刻的输出；$x_i \in \{-1,1\}$ 是输入信号的第 i 元素。

③ 运算、迭代至收敛

$$y_i(t+1) = f\left[\sum_{j=0}^{N-1} w_{ij} y_j(t)\right], \quad 0 \le i \le N-1 \quad \cdots\cdots \text{EQ3.65}$$

§3.6 Hopfield网络用于联想记忆

其中，$f[\bullet]$ 是神经元非线性函数。重复上述计算过程，直至 $y_i(t+k)$ 不再改变。最终，节点的输出描述了未知输入与样本的最佳匹配。

④ 返回②，可以再输入一个未知样本。

以上算法可以用**矩阵形式**表示如下：

$$\mathbf{W} = \sum_{s=0}^{M-1} \left(\mathbf{X}_s \mathbf{X}_s^T - I\right) \quad \cdots\cdots\cdots\cdots \text{EQ3.66}$$

$$\mathbf{Y}(0) = \mathbf{X} \quad \cdots\cdots\cdots\cdots \text{EQ3.67}$$

$$\mathbf{Y}(t+1) = f\left[\mathbf{W}\mathbf{Y}(t)\right] \quad \cdots\cdots\cdots\cdots \text{EQ3.68}$$

其中，I 为单位阵。 More...

如果输入样本没有噪声，即输入的是标准样本，运算将一次迭代成功，输出 $\mathbf{Y}$ 等于输入 $\mathbf{X}$。如果输入样本含有噪声，即输入 $\mathbf{X}+\Delta\mathbf{X}$，那么需要经过若干次迭代，最终收敛后使 $\mathbf{Y}$ 与某个样本 $\mathbf{X}$ 一致。

§3.6 Hopfield网络用于联想记忆

在仿真实验中，把样本“3”的25％的元素随机取反，经过6次迭代，最后输出收敛于标准样本“3”。

为实现联想记忆功能，我们希望网络最终收敛至预期的标准样本，即希望通过训练使样本成为网络的稳定状态。然而，稳定状态出现的规律十分复杂。

(1) 给定一组样本，经训练得到 w，网络将对应有一组稳定状态。除此之外，还可能有多余的稳定状态，称为伪稳定状态。

(2) 并非任何一组样本经训练后都可以构成一组稳定状态。

(3) 给定一个加噪的样本，最终不一定收敛到与其汉明距离最近的标准样本。

(4) 各样本之间的汉明距离分布对联想记忆功能的正确实现有重要影响。若样本之间相互正交（汉明距离是输入矢量维数的一半），效果最好。

(5) N 一定时，M 越小越好。

§ 3.7 双向联想记忆

双向联想记忆，Bidirectional Associative Memory(BAM)，是一种异联想，它可以存储两组矢量。

$\mathbf{A}=\left[a_0, a_1, \cdots, a_{N-1}\right]^T$　$\mathbf{A}\in\{-1,1\}^N$表示 N 维矢量，

$\mathbf{B}=\left[b_0, b_1, \cdots, b_{P-1}\right]^T$　$\mathbf{B}\in\{-1,1\}^P$表示 P 维矢量。

各M个矢量构成M 对矢量$(\mathbf{A}_s, \mathbf{B}_s)$，$s=0,1,\cdots,M-1$。

双向联想记忆可以

$$\left.\begin{matrix}\mathbf{A}\Rightarrow\mathbf{B}\\ \mathbf{B}\Rightarrow\mathbf{A}\end{matrix}\right\}\ \mathbf{A}\Leftrightarrow\mathbf{B}$$

当有噪声或缺损时，联想功能可使样本对复原。

利用人工神经网络实现BAM有多种结构形式，这里只介绍最基本的一种，Kosko型BAM网络，其网络模型如图3-16所示。

2004-8-8　《神经网络导论》——Hopfield神经网络　3-53

§ 3.7 双向联想记忆

$\mathbf{W}$为正向传输$\mathbf{B}\Rightarrow\mathbf{A}$ 的突触权重连接矩阵；$\mathbf{W}^T$是反向传输的权重连接矩阵。

当输入任意矢量时，网络要经过若干次迭代运算演变至稳定状态。下面通过与自联想记忆的对比来说明网络的演变过程。

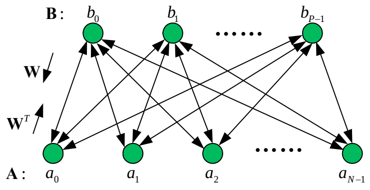


图3-16 双向联想记忆网络模型

自联想	双向联想
$\mathbf{WY}(t)\to\mathbf{Y}(t+1)$	$\mathbf{WB}(t)\to\mathbf{A}(t+1)$
$\mathbf{WY}(t+1)\to\mathbf{Y}(t+2)$	$\mathbf{W}^T\mathbf{A}(t+1)\to\mathbf{B}(t+2)$
$\vdots$	$\mathbf{WB}(t+2)\to\mathbf{A}(t+3)$
$\cdots\to\mathbf{Y}(t+k)$	$\vdots$
迭代直至收敛、稳定	直至**A**、**B**均为稳定状态

2004-8-8　《神经网络导论》——Hopfield神经网络　3-54

§ 3.7 双向联想记忆

网络的学习（训练）按Hebb规则。

假设给定 M个样本对：$(\mathbf{A}_0,\mathbf{B}_0),(\mathbf{A}_1,\mathbf{B}_1)\cdots(\mathbf{A}_{M-1},\mathbf{B}_{M-1})$。则

$$\mathbf{W}=\sum_{s=0}^{M-1}\mathbf{A}_s\mathbf{B}_s^T;\quad \mathbf{W}^T=\sum_{s=0}^{M-1}\mathbf{B}_s\mathbf{A}_s^T \quad \cdots\cdots\cdots\cdots \text{EQ3.95}$$

如果BAM网络神经元的阈值等于零，那么称为齐次BAM网络。定义齐次BAM网络的**能量函数**为

$$E(\mathbf{A},\mathbf{B})=-\frac{1}{2}\mathbf{A}^T\mathbf{WB}-\frac{1}{2}\mathbf{B}^T\mathbf{W}^T\mathbf{A} \quad \cdots\cdots\cdots\cdots \text{EQ3.97}$$

因为

$$\mathbf{B}^T\mathbf{W}^T\mathbf{A}=\mathbf{B}^T\left(\mathbf{A}^T\mathbf{W}\right)^T=\left(\mathbf{A}^T\mathbf{WB}\right)^T=\mathbf{A}^T\mathbf{WB}$$

所以

$$E(\mathbf{A},\mathbf{B})=-\mathbf{A}^T\mathbf{WB} \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{EQ3.99}$$

2004-8-8　《神经网络导论》——Hopfield神经网络　3-55

§ 3.7 双向联想记忆

对非齐次BAM网络，其**能量函数**定义为：

$$E(\mathbf{A},\mathbf{B})=-\mathbf{A}^T\mathbf{WB}+\mathbf{A}^T\boldsymbol{\theta}+\mathbf{B}^T\boldsymbol{\mu} \quad \cdots\cdots\cdots\cdots \text{EQ3.100}$$

其中$\boldsymbol{\theta}=\left(\theta_0,\theta_1,\cdots,\theta_{N-1}\right)^T$，$\theta_i$为$\mathbf{A}$的第$i$个元素$a_i$对应的阈值。

$\boldsymbol{\mu}=\left(\mu_0,\mu_1,\cdots,\mu_{P-1}\right)^T$，$\mu_j$为$\mathbf{B}$的第 j个元素b_j对应的阈值。

若神经元非线性函数f 为硬限幅特性，以 sgn 表示，则描述BAM动态特性的差分方程式如下：

正向传输（$\mathbf{B}\Rightarrow\mathbf{A}$）：

$$a_i(t+1)=\begin{cases}\operatorname{sgn}\left[\sum_{j=0}^{P-1}w_{ij}b_j(t)-\theta_i\right] & \text{当}\sum_{j=0}^{P-1}w_{ij}b_j(t)-\theta_i\neq 0\text{时}\\ a_i(t) & \text{当}\sum_{j=0}^{P-1}w_{ij}b_j(t)-\theta_i=0\text{时}\end{cases} \quad \text{EQ3.101}$$

2004-8-8　《神经网络导论》——Hopfield神经网络　3-56

§3.7 双向联想记忆

反向传输（$\mathbf{A} \Rightarrow \mathbf{B}$）：

$$b_j(t+2)=\begin{cases}\mathrm{sgn}\left[\sum_{i=0}^{N-1} w_{ij}a_i(t+1)-\mu_j\right] & \text{当}\sum_{i=0}^{N-1} w_{ij}a_i(t+1)-\mu_j \neq 0\text{时} \\ b_j(t+1) & \text{当}\sum_{i=0}^{N-1} w_{ij}a_i(t+1)-\mu_j = 0\text{时}\end{cases} \quad \text{EQ3.102}$$

下面我们证明，网络在演变过程中能量函数 E 递降，即 $\Delta E<0$。

设 a_i 发生变化，由于其取值为－1或1，所以 Δa_i 的取值为－2，2两种情况。为书写简便，下面的推导省去时序变量。

由EQ3.100可得：

$$\Delta E=-\left(\Delta\mathbf{A}^T\right)\mathbf{WB}+\left(\Delta\mathbf{A}^T\right)\boldsymbol{\theta}=-\left(\Delta\mathbf{A}^T\right)(\mathbf{WB}-\boldsymbol{\theta})$$

$$=-\sum_{i=0}^{N-1}(\Delta a_i)\left(\sum_{j=0}^{P-1} w_{ij}b_j-\theta_i\right) \quad \cdots \text{EQ3.103}$$

More...

§3.7 双向联想记忆

如果 $\Delta a_i=-2$，即 a_i 从1变为－1，由EQ3.101可知必有

$$\sum_{j=0}^{P-1} w_{ij}b_j-\theta_i<0 \quad \text{故有 } \Delta E<0\text{。}$$

如果 $\Delta a_i=2$，即 a_i 从－1变为1，由EQ3.101可知必有

$$\sum_{j=0}^{P-1} w_{ij}b_j-\theta_i>0 \quad \text{故有 } \Delta E<0\text{。}$$

联系

第二次实验

双向联想记忆BAM

地址：http://home.xjtu.edu.cn/teacher/zhangjg

§3.8 本章小结

本章我们讨论了Hopfield网络的电路模型及其能量函数。在此基础上介绍了Hopfield网络A/D变换器及非对称网络的A/D变换器；讨论了用Hopfield网络求解TSP问题及货流问题，这些都是Hopfield网络的典型应用。最后，介绍了Hopfield网络在联想记忆方面的应用。

Hopfield网络的几个主要特点：

(1) Hopfield网络是一个复杂的动力学系统，稳定性分析成为不可回避的困难问题，这不同于前向网络，它是一种输入到输出的映射模型。

(2) Hopfield网络的学习（训练）主要采用Hebb规则，收敛速度较前向网络快得多。

(3) Hopfield网络可以用于优化计算和联想记忆，两者实际上是对偶的。在优化计算方面，由于Hopfield网络较成功地求解了NPH问题，因此人们对它寄予厚望。

附录A. EQ3.28的推导

$$E=\frac{1}{2}\left(x-\sum_{i=0}^{3}2^iV_i\right)^2-\frac{1}{2}\sum_{i=0}^{3}2^{2i}V_i(V_i-1)$$

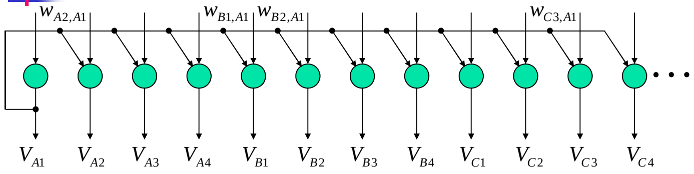
$$=\frac{1}{2}x^2-x\sum_{i=0}^{3}2^iV_i+\frac{1}{2}\left(\sum_{i=0}^{3}2^iV_i\right)^2-\frac{1}{2}\sum_{i=0}^{3}2^{2i}V_i^2+\frac{1}{2}\sum_{i=0}^{3}2^{2i}V_i$$

$$=-\sum_{i=0}^{3}\left(2^ix-2^{2i-1}\right)V_i+\frac{1}{2}\left(\sum_{i=0}^{3}2^iV_i\right)\left(\sum_{j=0}^{3}2^jV_j\right)-\frac{1}{2}\sum_{i=0}^{3}2^{2i}V_i^2$$

$$=\frac{1}{2}\sum_{j=0}^{3}\sum_{\substack{i=0\\i\neq j}}^{3}2^{(i+j)}V_iV_j-\sum_{i=0}^{3}\left(2^ix-2^{2i-1}\right)V_i$$

Return

附录B. 对TSP双下标的说明



$$w_{Xi,Yj}=-\alpha\delta_{XY}\left(1-\delta_{ij}\right)-\beta\delta_{ij}\left(1-\delta_{XY}\right)-\gamma-\lambda d_{XY}\left(\delta_{j,i+1}+\delta_{j,i-1}\right)$$

$$w_{A2,A1}=-\alpha-\gamma$$

$$w_{B1,A1}=-\beta-\gamma$$

$$w_{B2,A1}=-\gamma-\lambda d_{AB}$$

$$w_{C3,A1}=-\gamma$$

Return

附录C. 对自联想记忆的说明

假设现在输入的是标准样本 $\mathbf{X}_0$，那么

$$\begin{aligned}
\mathbf{Y}&=\operatorname{sgn}\left(\mathbf{W}\mathbf{X}_0\right)\\
&=\operatorname{sgn}\left[\left(\sum_{s=0}^{M-1}\mathbf{X}_s\mathbf{X}_s^T-\mathbf{I}\right)\mathbf{X}_0\right]\\
&=\operatorname{sgn}\left[\sum_{s=0}^{M-1}\mathbf{X}_s\mathbf{X}_s^T\mathbf{X}_0-M\mathbf{X}_0\right]\\
&=\operatorname{sgn}\left[\mathbf{X}_0\mathbf{X}_0^T\mathbf{X}_0+\sum_{s=1}^{M-1}\mathbf{X}_s\mathbf{X}_s^T\mathbf{X}_0-M\mathbf{X}_0\right]\\
&=\operatorname{sgn}\left[120\mathbf{X}_0+\sum_{s=1}^{7}\mathbf{X}_s\mathbf{X}_s^T\mathbf{X}_0-8\mathbf{X}_0\right]\\
&=\operatorname{sgn}\left[112\mathbf{X}_0+\sum_{s=1}^{7}\mathbf{X}_s\mathbf{X}_s^T\mathbf{X}_0\right]
\end{aligned}$$

Return

附录D. 对双向联想记忆能量变化的说明

设从 t 时刻到 $t+1$ 时刻，系统由 $\mathbf{B}$ 联想 $\mathbf{A}$，则

$$E(t)=-\mathbf{A}^T(t)\mathbf{W}\mathbf{B}(t)+\mathbf{A}^T(t)\boldsymbol{\theta}+\mathbf{B}^T(t)\boldsymbol{\mu}$$

$$E(t+1)=-\mathbf{A}^T(t+1)\mathbf{W}\mathbf{B}(t+1)+\mathbf{A}^T(t+1)\boldsymbol{\theta}+\mathbf{B}^T(t+1)\boldsymbol{\mu}$$

已知 $\mathbf{B}(t+1)=\mathbf{B}(t)$，令 $\Delta\mathbf{A}=\mathbf{A}(t+1)-\mathbf{A}(t)$ 可得：

$$\begin{aligned}
\Delta E&=E(t+1)-E(t)\\
&=-\left[\mathbf{A}^T(t+1)-\mathbf{A}^T(t)\right]\mathbf{W}\mathbf{B}+\left[\mathbf{A}^T(t+1)-\mathbf{A}^T(t)\right]\theta\\
&=-\Delta\mathbf{A}^T\mathbf{W}\mathbf{B}+\Delta\mathbf{A}^T\theta
\end{aligned}$$

Return