

# 神经网络

# Neural Networks

## 第一章

---

# 绪 论

史忠植

中国科学院计算技术研究所  
<http://www.intsci.ac.cn/>

---

# 教材

---

书名：《神经网络》

出版社：高等教育出版社

出版日期：2009年5月

定价：37元

作者：史忠植

# 内容提要

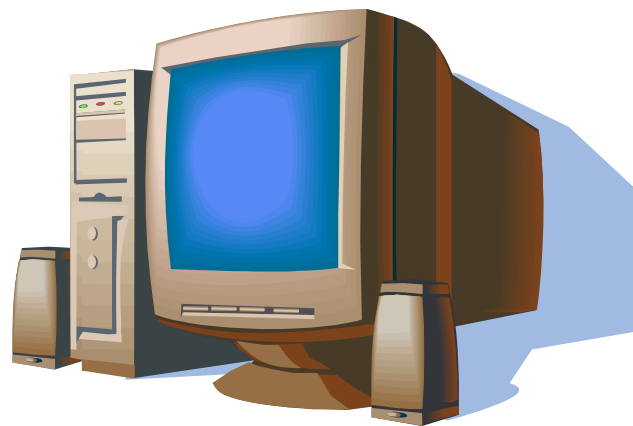
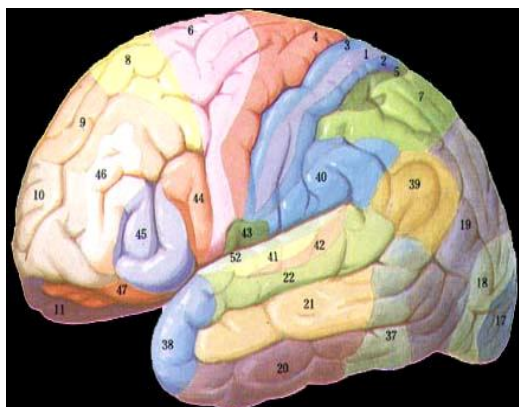
---

- 1.1 概述
- 1.2 神经网络的研究历史
- 1.3 人脑的神经系统
- 1.4 神经信息处理的基本原理
- 1.5 简单的神经网络模型
- 1.6 神经网络的研究内容
- 1.7 神经网络的分类
- 1.8 神经网络研究的发展方向

# 概述

## 人脑与计算机信息处理能力的比较

记忆与联想能力



# 概述

## 人脑与计算机信息处理能力的比较

信息加工能力



非逻辑加工

回忆  
联想  
想象

逻辑加工

模糊逻辑  
辩证逻辑  
二值逻辑

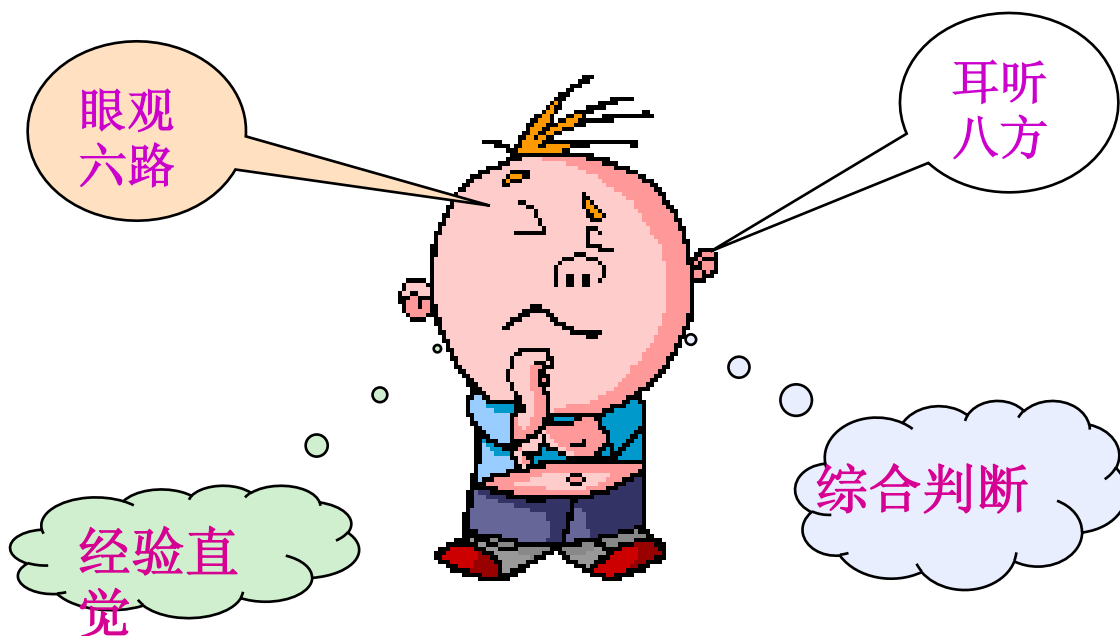


二值逻辑

# 概述

## 人脑与计算机信息处理能力的比较

信息综合能力



# 概述

## 人脑与计算机信息处理能力的比较

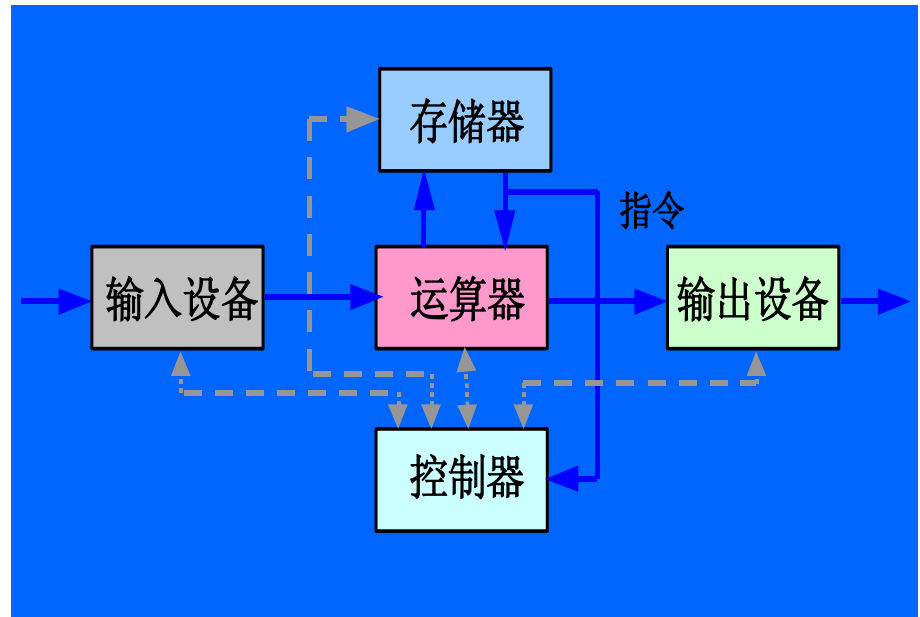
信息处理速度



# 概述

## 人脑与计算机信息处理机制的比较

### (一) 系统结构

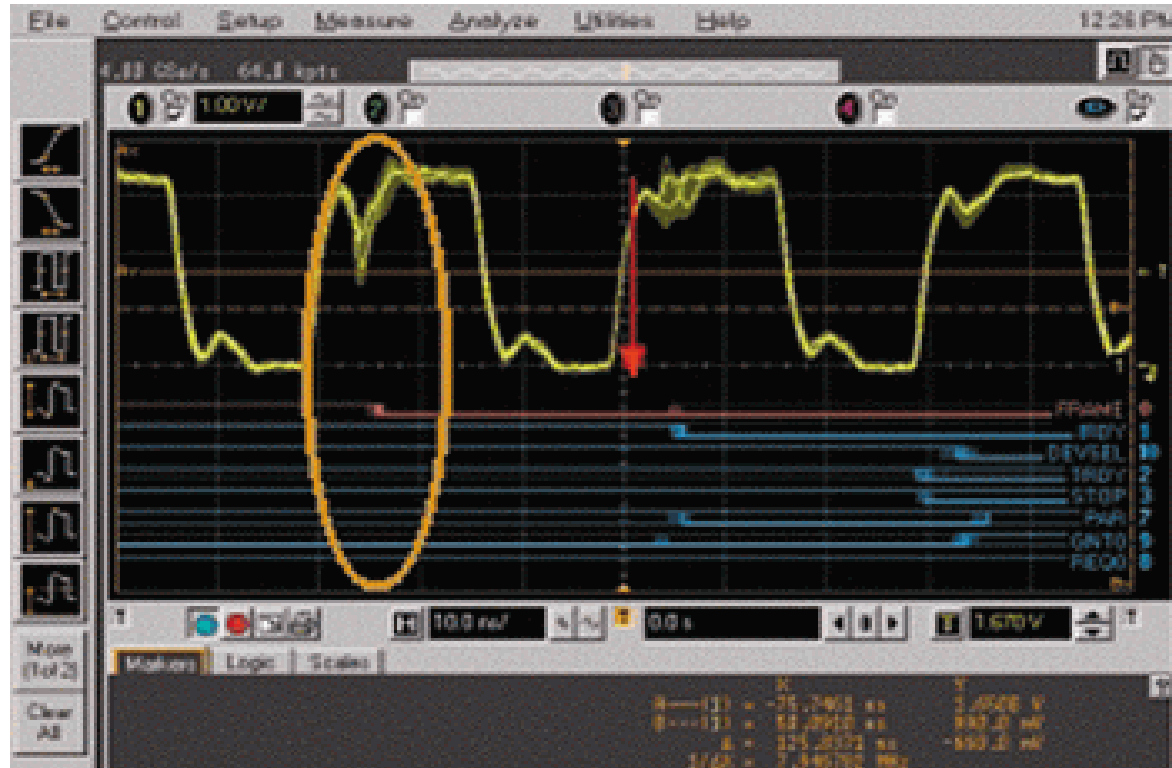




# 概述

## 人脑与计算机信息处理机制的比较

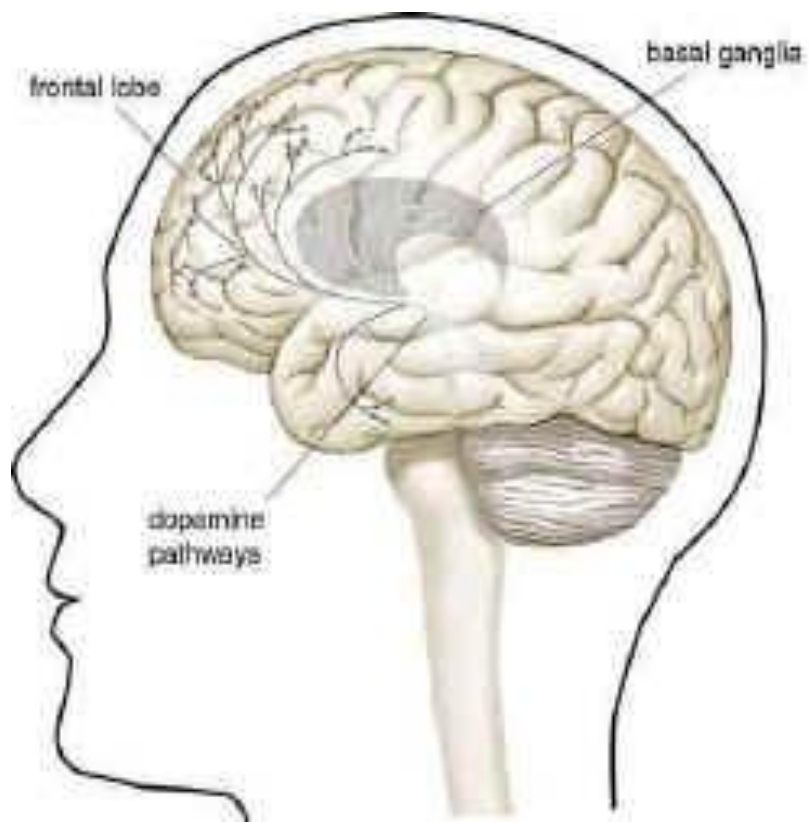
信号形式



# 概述

## 人脑与计算机信息处理机制的比较

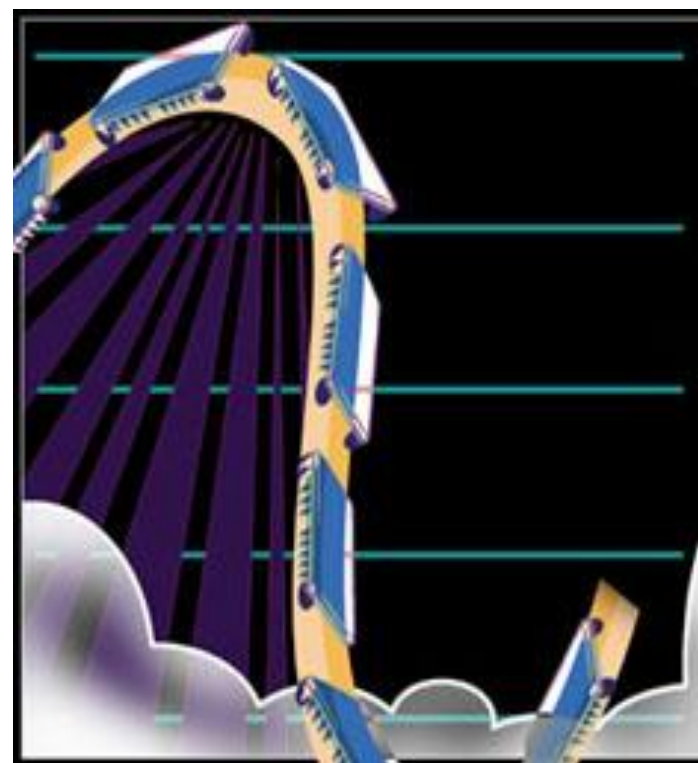
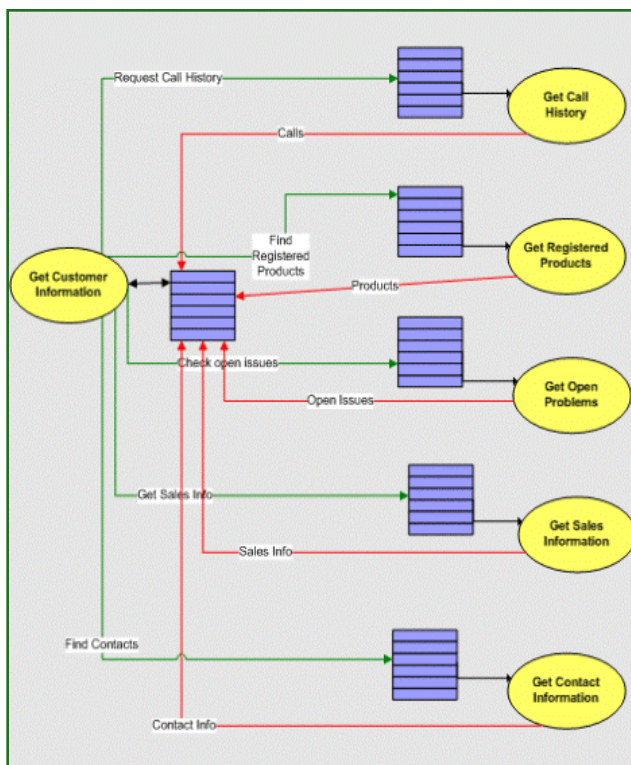
信息存储



# 概述

## 人脑与计算机信息处理机制的比较

信息处理方式



# 神经网络的提出

- 物理符号系统和人工神经网络系统的差别

项目	物理符号系统	人工神经网络
处理方式	逻辑运算	模拟运算
执行方式	串行	并行
动作	离散	连续
存储	局部集中	全局分布

# 神经网络的提出

## • 两种人工智能技术的比较

项目	传统的AI技术	ANN技术
基本实现方式	串行处理；由程序实现控制	并行处理；对样本数据进行多目标学习；通过人工神经元之间的相互作用实现控制
基本开发方法	设计规则、框架、程序；用样本数据进行调试（由人根据已知的环境去构造一个模型）	定义人工神经网络的结构原型，通过样本数据，依据基本的学习算法完成学习——自动从样本数据中抽取内涵（自动适应应用环境）
适应领域	精确计算：符号处理，数值计算	非精确计算：模拟处理，感觉，大规模数据并行处理
模拟对象	左脑（逻辑思维）	右脑（形象思维）

# 神经网络的特点

---

- 信息的分布表示
- 运算的全局并行和局部操作
- 处理的非线性

# 神经网络的概念

## •1、定义

### •1) Hecht—Nielsen (1988年)

人工神经网络是一个并行、分布处理结构，它由处理单元及其称为联接的无向讯号通道互连而成。这些处理单元（PE—Processing Element）具有局部内存，并可以完成局部操作。每个处理单元有一个单一的输出联接，这个输出可以根据需要被分枝成希望个数的许多并行联接，且这些并行联接都输出相同的信号，即相应处理单元的信号，信号的大小不因分支的多少而变化。

# 神经网络的概念

- (1) Hecht—Nielsen (1988年) (续)
- 处理单元的输出信号可以是任何需要的数学模型，每个处理单元中进行的操作必须是完全局部的。也就是说，它必须仅仅依赖于经过输入联接到达处理单元的所有输入信号的当前值和存储在处理单元局部内存中的值。



# 神经网络的概念

---

- **强调：**
  - ① 并行、分布处理结构；
  - ② 一个处理单元的输出可以被任意分枝，且大小不变；
  - ③ 输出信号可以是任意的数学模型；
  - ④ 处理单元完全的局部操作

# 神经网络的概念

- (2) Rumelhart, McClelland, Hinton的PDP
- 1) 一组处理单元 (PE或AN) ;
- 2) 处理单元的激活状态 ( $a_i$ ) ;
- 3) 每个处理单元的输出函数 ( $f_i$ ) ;
- 4) 处理单元之间的联接模式;
- 5) 传递规则 ( $\sum w_{ij} o_i$ ) ;
- 6) 把处理单元的输入及当前状态结合起来产生激活值的激活规则 ( $F_i$ ) ;
- 7) 通过经验修改联接强度的学习规则;
- 8) 系统运行的环境 (样本集合) 。

# 神经网络的概念

---

- **(3) Simpson (1987年)**
- 人工神经网络是一个非线性的有向图，图中含有可以通过改变权大小来存放模式的加权边，并且可以从不完整的或未知的输入找到模式。

# 神经网络的概念

- 2、关键点
  - (1) 信息的分布表示
  - (2) 运算的全局并行与局部操作
  - (3) 处理的非线性特征
- 3、对大脑基本特征的模拟
  - 1) 形式上：神经元及其联接；**BN**对**AN**
  - 2) 表现特征：信息的存储与处理

# 神经网络的概念

---

- 别名
- 人工神经系统 (ANS)
- 神经网络 (NN)
- 自适应系统 (Adaptive Systems)、自适应网 (Adaptive Networks)
- 联接模型 (Connectionism)
- 神经计算机 (Neurocomputer)

# 发展历史

- 萌芽期（20世纪40年代）

人工神经网络的研究最早可以追溯到人类开始研究自己的智能的时期，到1949年止。

- 1943年，心理学家McCulloch和数学家Pitts建立起了著名的阈值加权和模型，简称为M-P模型。发表于数学生物物理学会刊《Bulletin of Mathematical Biophysics》

- 1949年，心理学家D. O. Hebb提出神经元之间突触联系是可变的假说——Hebb学习律。

# 发展历史

- 第一高潮期（1950~1968）
  - 以Marvin Minsky, Frank Rosenblatt, Bernard Widrow等为代表人物，代表作是单级感知器（Perceptron）。
  - 可用电子线路模拟。
  - 人们乐观地认为几乎已经找到了智能的关键。许多部门都开始大批地投入此项研究，希望尽快占领制高点。

# 发展历史

---

- 反思期（1969~1982）
  - M. L. Minsky和S. Papert, 《Perceptron》，MIT Press, 1969年
  - “异或”运算不可表示
  - 二十世纪70年代和80年代早期的研究成果



# 发展历史

- 第二高潮期（1983~1990）
  - 1982年，J. Hopfield提出Hopfield网络
    - 用Lyapunov函数作为网络性能判定的能量函数，建立ANN稳定性的判别依据
    - 阐明了ANN与动力学的关系
    - 用非线性动力学的方法来研究ANN的特性
    - 指出信息被存放在网络中神经元的联接上

# 发展历史

- 第二高潮期（1983~1990）
  - 1984年，J. Hopfield设计研制了后来被人们称为**Hopfield网-Tank** 电路。较好地解决了著名的TSP问题，找到了最佳解的近似解，引起了较大的轰动。
  - 1985年，UCSD的Hinton、Sejnowsky、Rumelhart等人所在的并行分布处理（PDP）小组的研究者在Hopfield网络中引入了随机机制，提出所谓的**Boltzmann**

# 发展历史

- ☛ 1986年，并行分布处理小组的Rumelhart等研究者重新独立地提出多层网络的学习算法——**BP算法**，较好地解决了多层网络的学习问题。（Paker1982和Werbos1974年）
- ☛ 自适应共振理论（ART）
- ☛ 自组织特征映射理论

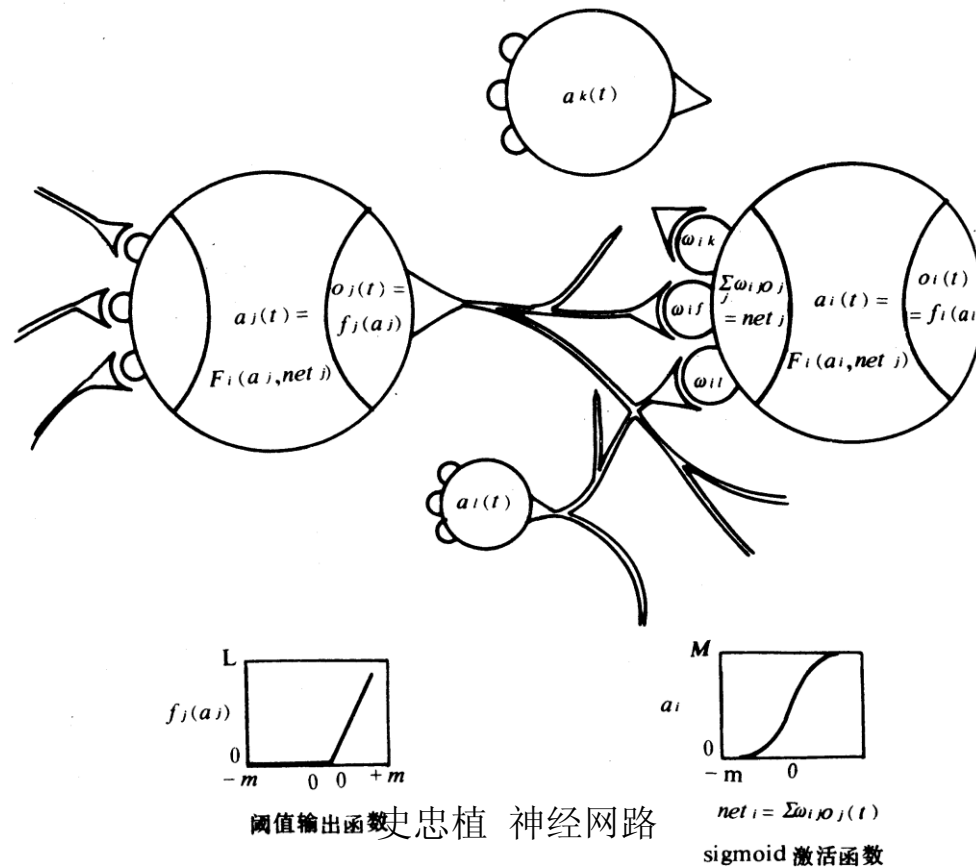
# 发展历史

---

- Hinton 等人最近提出了 Helmboltz 机
- 徐雷提出的 Ying-Yang 机理论模型
- 甘利俊一( S.Amari) 开创和发展的基于统计流形的方法应用于人工神经网络的研究,
- **国内首届神经网络大会**是1990年12月在北京举行的。

# 并行分布式理论框架

1986年，美国加州大学圣地亚哥分校（UCSD）Rumelhart, McClelland, Hinton: **Parallel and Distributed Processing**, MIT Press, Cambridge



# 并行分布式理论框架

## PDP模型

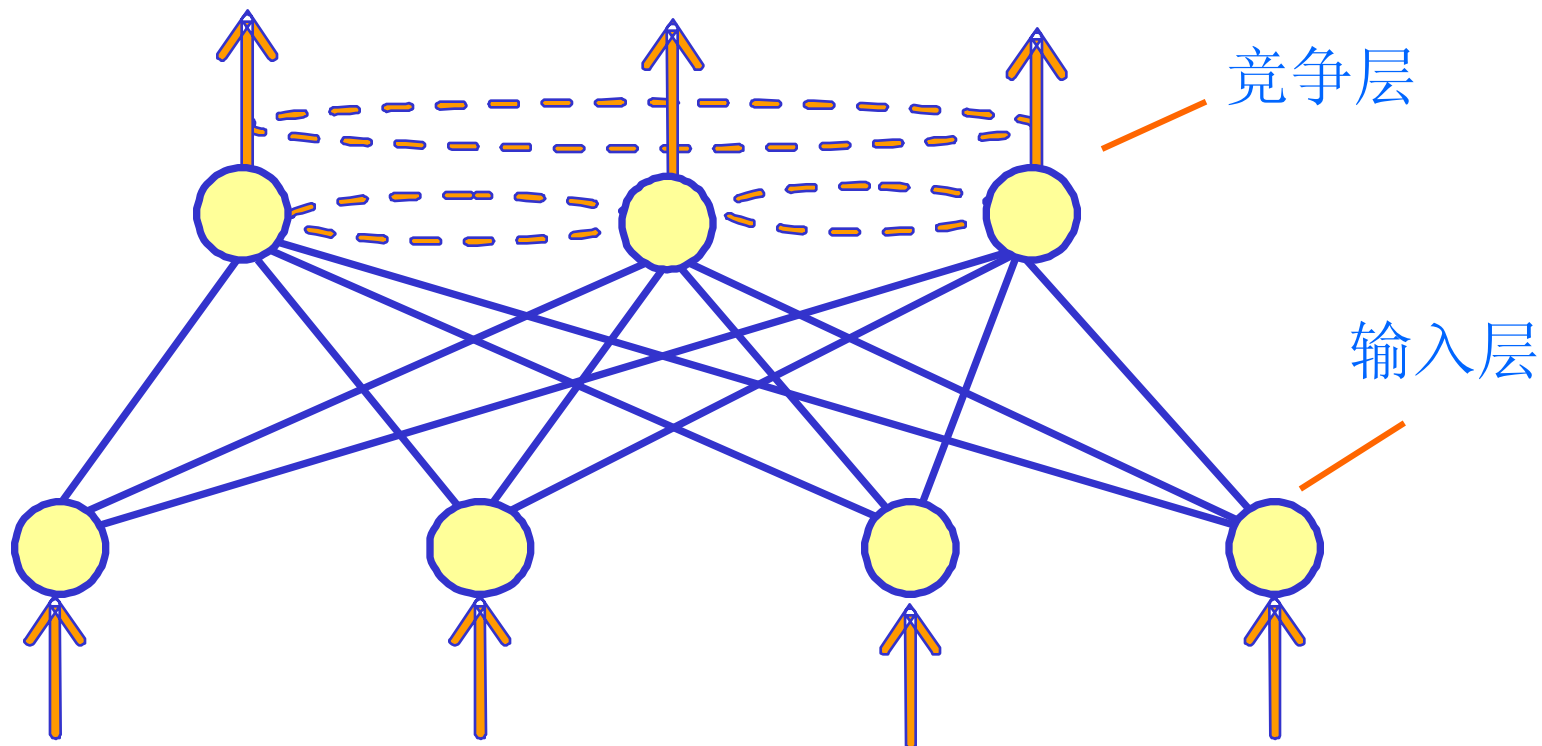
- 1) 一组处理单元 (PE或AN)
- 2) 处理单元的激活状态 ( $a_i$ )
- 3) 每个处理单元的输出函数 ( $f_i$ )
- 4) 处理单元之间的连接模式
- 5) 传递规则 ( $\sum w_{ij} o_i$ )
- 6) 把处理单元的输入及当前状态结合起来产生激活值的激活规则 ( $F_i$ )
- 7) 通过经验修改连接强度的学习规则
- 8) 系统运行的环境 (样本集合)

# 神经网络的维数

---

- Various types of **neurons**
- Various network **architectures**
- Various **learning algorithms**
- Various **applications**

# 交互与竞争 IAC神经网络



自组织神经网络的典型结构



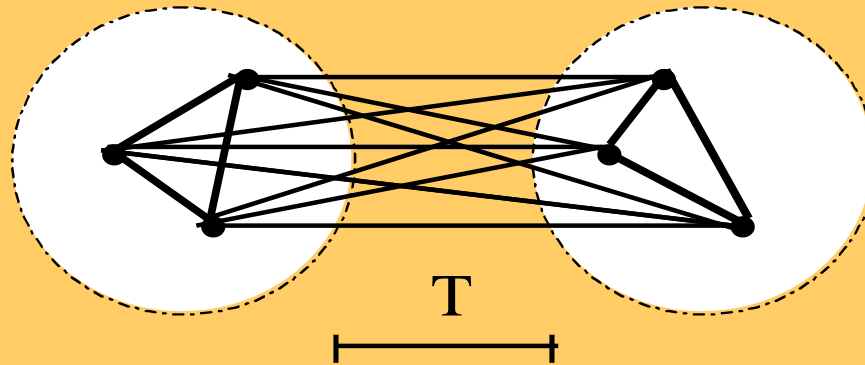
# 竞争学习

## 相似性测量\_欧式距离法

$$\|\mathbf{X} - \mathbf{X}_i\| = \sqrt{(\mathbf{X} - \mathbf{X}_i)^T (\mathbf{X} - \mathbf{X}_i)}$$

类 1

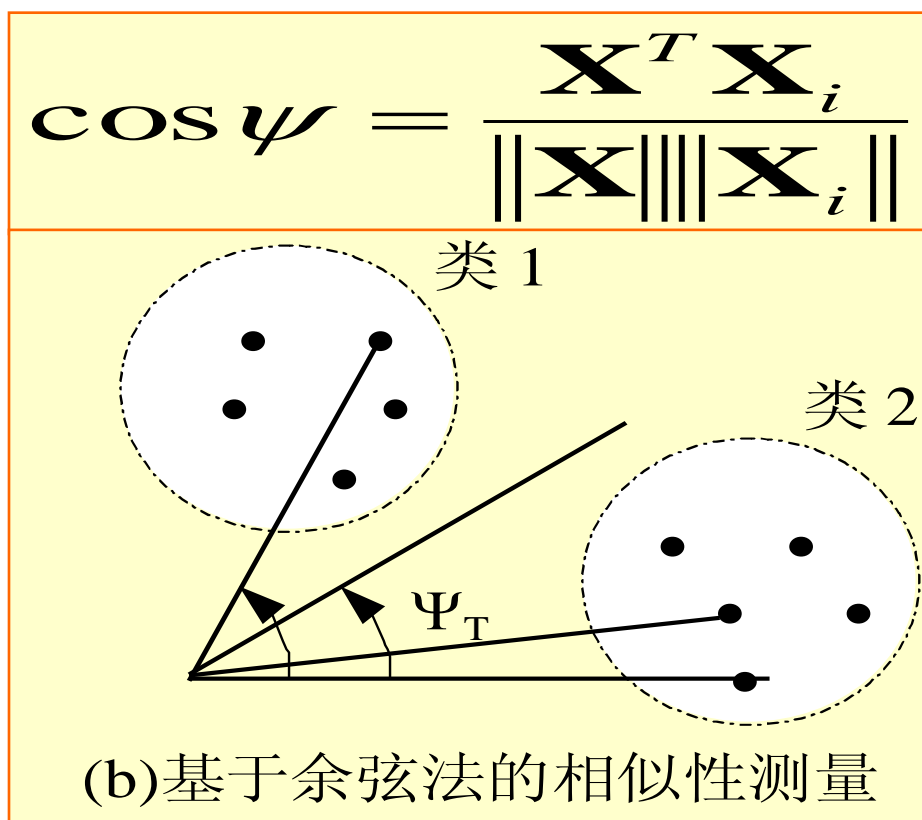
类 2



(a) 基于欧式距离的相似性测量

# 竞争学习

## 相似性测量\_余弦法



# 竞争学习原理

## 竞争学习规则——Winner-Take-All

网络的输出神经元之间相互竞争以求被激活，结果在每一时刻只有一个输出神经元被激活。这个被激活的神经元称为竞争获胜神经元，而其它神经元的状态被抑制，故称为**Winner Take All**。

# 竞争学习原理

**寻找获胜神经元** 当网络得到一个输入模式向量时，竞争层的所有神经元对应的内星权向量均与其进行相似性比较，并将最相似的内星权向量判为竞争获胜神经元。

欲使两单位向量最相似，须使其点积最大。即：

$$\hat{\mathbf{W}}_{j^*}^T \hat{\mathbf{X}} = \max_{j \in \{1, 2, \dots, m\}} (\hat{\mathbf{W}}_j^T \hat{\mathbf{X}})$$

# 竞争学习原理

$$\|\hat{\mathbf{X}} - \hat{\mathbf{W}}_{j^*}\| = \min_{j \in \{1, 2, \dots, m\}} \{\|\hat{\mathbf{X}} - \hat{\mathbf{W}}_j\|\}$$

$$\|\hat{\mathbf{X}} - \hat{\mathbf{W}}_{j^*}\| = \sqrt{(\hat{\mathbf{X}} - \hat{\mathbf{W}}_{j^*})^T (\hat{\mathbf{X}} - \hat{\mathbf{W}}_{j^*})}$$

$$= \sqrt{\hat{\mathbf{X}}^T \hat{\mathbf{X}} - 2\hat{\mathbf{W}}_{j^*}^T \hat{\mathbf{X}} + \hat{\mathbf{W}}_{j^*}^T \hat{\mathbf{W}}_{j^*}} = \sqrt{2(1 - \mathbf{w}_{j^*}^T \hat{\mathbf{X}})}$$

从上式可以看出，欲使两单位向量的欧式距离最小，须使两向量的点积最大。即：

$$\hat{\mathbf{W}}_{j^*}^T \hat{\mathbf{X}} = \max_{j \in \{1, 2, \dots, m\}} (\hat{\mathbf{W}}_j^T \hat{\mathbf{X}})$$

# 竞争学习原理

## 3. 网络输出与权值调整

$$o_j(t+1) = \begin{cases} 1 & j = j^* \\ 0 & j \neq j^* \end{cases}$$

$$\left. \begin{aligned} \mathbf{W}_{j^*}(t+1) &= \hat{\mathbf{W}}_{j^*}(t) + \Delta \mathbf{W}_{j^*} = \hat{\mathbf{W}}_{j^*}(t) + \mu(t)(\hat{\mathbf{X}} - \hat{\mathbf{W}}_{j^*}) \\ \mathbf{W}_j(t+1) &= \hat{\mathbf{W}}_j(t) & j \neq j^* \end{aligned} \right\}$$

步骤3完成后回到步骤1继续训练，直到学习率衰减到0。

# Thank You

---

Question!

Intelligence Science

<http://www.intsci.ac.cn/>

