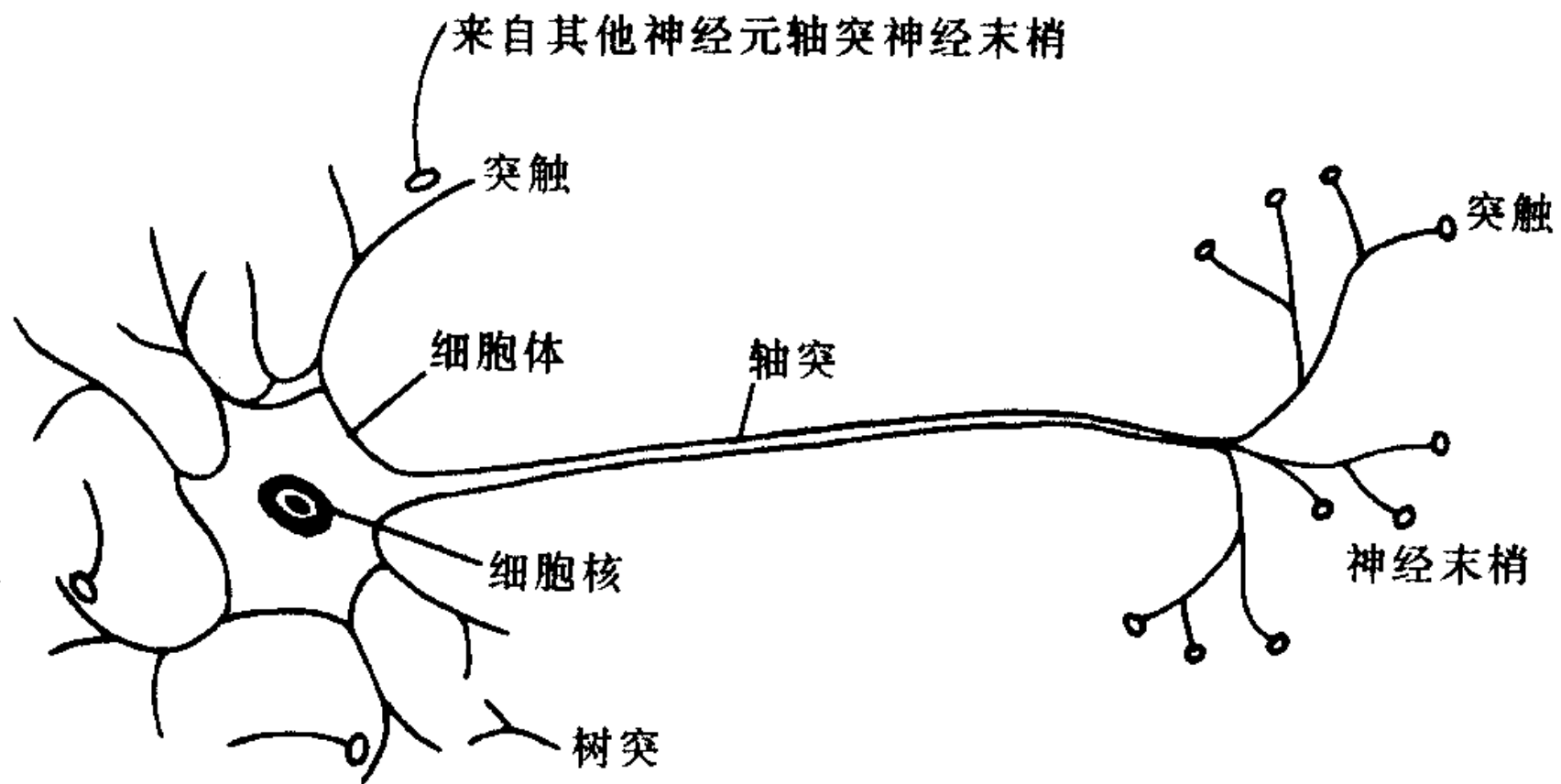


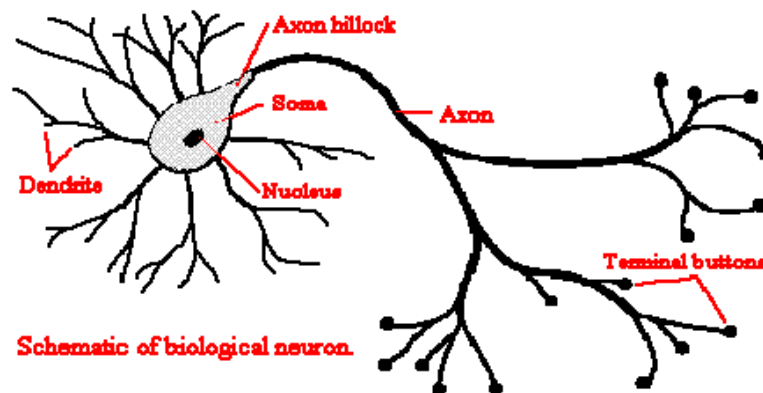
# 神经网络

Neural Network



# 神经网络模型的基本组成

## ☾★ 生物神经元的基本组成



☾★树突

☾★轴突

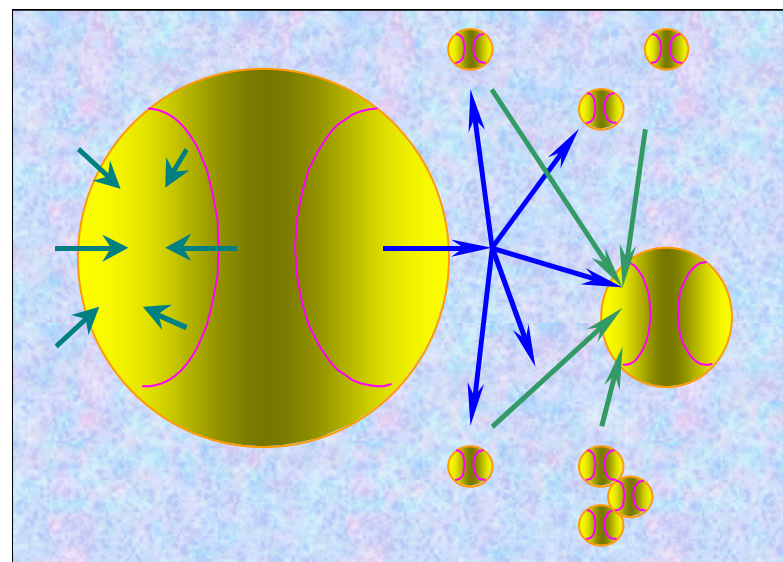
## ☾★ 人工神经元的基本结构

处理单元

连接

☾★输入

☾★输出



# 主要内容

- 智能及其实现
- **ANN**基础(Artificial Neural Network)
- **Perceptron**(感知器)
- **BP** (**Back Propagation**——向后传播) 网络
- **CPN** (**Counterpropagation Networks**) 对传网
- 统计方法
- **Hopfield** 网 与 **BAM** (**Bidirectional Associative Memory**) 双联存储器
- **ART** (**Adaptive Resonance Theory**) 自适应共振理论
- 深度学习理论 (**Deep Learning**)

# 1 NN介绍

1.1 人工神经网络的提出

1.2 人工神经网络的特点

1.3 历史回顾

# 1 NN介绍

- 人类对人工智能的研究可以分成两种方式对应着**两种不同的技术**:
  - 传统的人工智能技术——心理的角度模拟
  - 基于人工神经网络的技术——生理的角度模拟

# 1.1 人工神经网络的提出

- 人工神经网络（**Artificial Neural Networks**, 简记作**ANN**），是对人类大脑系统的一阶特性的一种描述。简单地讲，它是一个**数学模型**，可以用**电子线路**来实现，也可以用**计算机程序**来模拟，是人工智能研究的一种方法。

# 1.1 人工神经网络的提出

- 1.1.1 智能与人工智能
- 一、 智能的含义
- 智能是个体有目的的行为，合理的思维，以及有效的、适应环境的综合能力。
- 智能是个体认识客观事物和运用知识解决问题的能力。
- 人类个体的智能是一种综合能力。



# 1.1 人工神经网络的提出

- 智能可以包含8个方面
- **感知与认识客观事物、客观世界和自我的能力**
  - 感知是智能的基础——最基本的能力
- **通过学习取得经验与积累知识的能力**
  - 这是人类在世界中能够不断发展的最基本能力。
- **理解知识，运用知识和经验分析、解决问题的能力**
  - 这一能力可以算作是智能的高级形式。是人类对世界进行适当的改造，推动社会不断发展的基本能力。

# 1.1 人工神经网络的提出

- **联想、推理、判断、决策语言的能力**
  - 这是智能的高级形式的又一方面。
  - 预测和认识
  - “主动”和“被动”之分。联想、推理、判断、决策的能力是“主动”的基础。
- **运用语言进行抽象、概括的能力**
- **上述这5种能力，被认为是人类智能最为基本的能力**

# 1.1 人工神经网络的提出

- 作为5种能力综合表现形式的3种能力
  - 发现、发明、创造、创新的能力
  - 实时、迅速、合理地应付复杂环境的能力
  - 预测、洞察事物发展、变化的能力

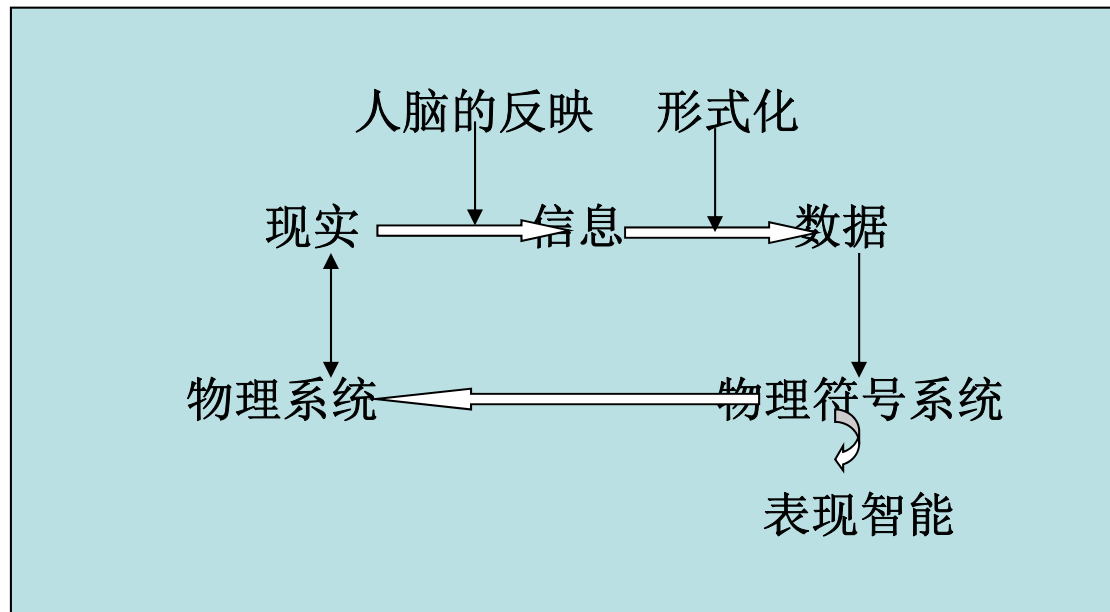
# 1.1 人工神经网络的提出

- 二、人工智能
- 人工智能：研究如何使类似计算机这样的设备去模拟人类的这些能力。
- 研究人工智能的目的
  - 增加人类探索世界，推动社会前进的能力
  - 进一步认识自己
- 三大学术流派
  - 符号主义学派（模拟人脑的逻辑思维）
  - 联接主义学派（模拟人脑的形象思维）
  - 进化主义学派（强调人的行为模拟）

# 1.1 人工神经网络的提出

- 1.1.2 物理符号系统

- 



# 1.1 人工神经网络的提出

- Newell和Simon假说：一个物理系统表现智能行为的充要条件是它有一个物理符号系统
- 概念：物理符号系统需要有一组称为符号的实体组成，它们都是物理模型，可以在另一类称为符号结构的实体中作为成分出现，以构成更高级别的系统。

# 1.1 人工神经网络的提出

- 困难：
  - 抽象——舍弃一些特性，同时保留一些特性
  - 形式化处理——用物理符号及相应规则表达物理系统的存在和运行。
- 局限：
  - 对全局性判断、模糊信息处理、多粒度的视觉信息处理等是非常困难的。

# 1.1 人工神经网络的提出

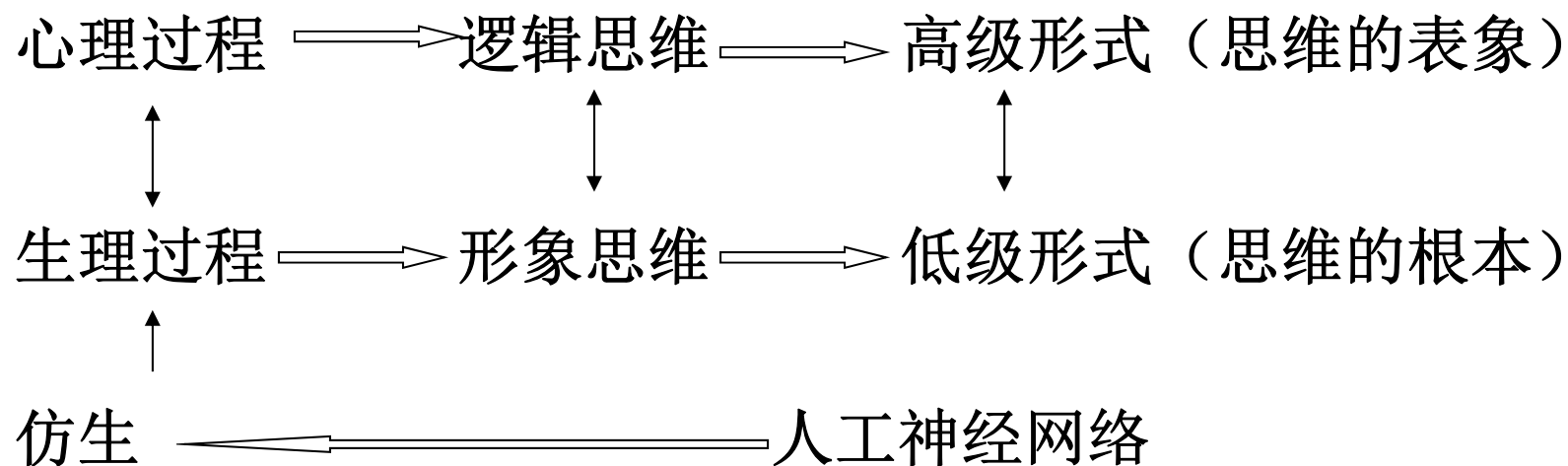
- 1.1.3 联接主义观点
- 核心：智能的本质是联接机制。
- 神经网络是一个由大量简单的处理单元组成的高度复杂的大规模非线性自适应系统
- ANN力求从四个方面去模拟人脑的智能行为
  - 物理结构
  - 计算模拟
  - 存储与操作
  - 训练



# 1.1 人工神经网络的提出

## • 1.1.4 两种模型的比较

### 物理符号系统



### 联结主义观点

# 1.1 人工神经网络的提出

- 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

# 1.1 人工神经网络的提出

- 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）

## 1.2 人工神经网络的特点

- 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

# 1.2.1 人工神经网络的概念

- 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。

## 1.2.1 人工神经网络的概念

- **(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。

## 1.2.1 人工神经网络的概念

- 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

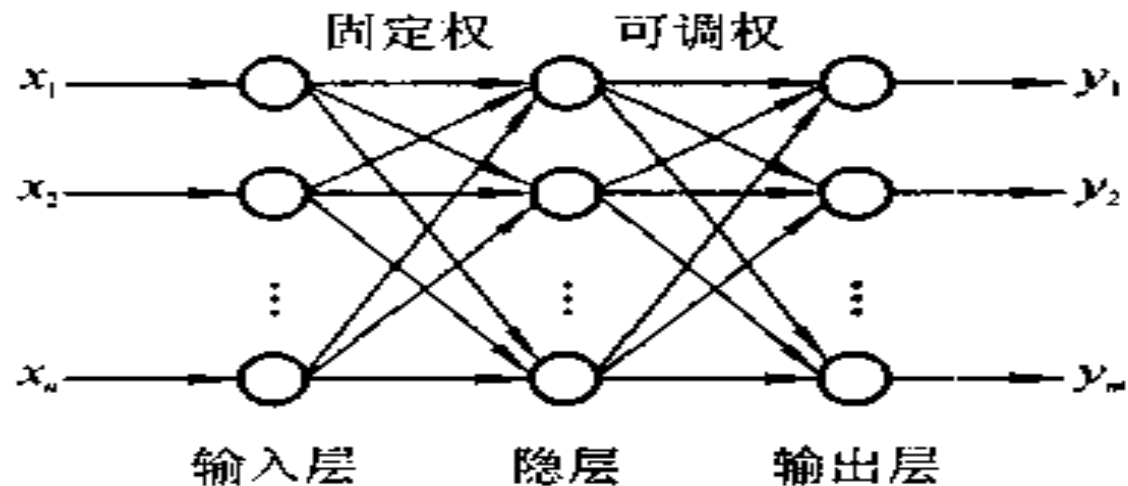
## 1.2.1 人工神经网络的概念

- **(2) Rumelhart, McClelland, Hinton的PDP (并行分布处理模型)**
  - **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) 。**



## 1.2.1 人工神经网络的概念

- (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。



## 1.2.1 人工神经网络的概念

- 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

## 1.2.1 人工神经网络的概念

- 4、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**)

## 1.2.2 学习 (Learning) 能力

- 人工神经网络可以根据所在的环境去改变它的行为

它在接受样本集合 $A$ 时，可以抽取集合 $A$ 中输入数据与输出数据之间的映射关系。——“抽象”功能。

- 不同的人工神经网络模型，有不同的学习/训练算法

## 1.2.3 基本特征的自动提取

- 由于其运算的不精确性，表现成“去噪音、容残缺”的能力，利用这种不精确性，比较自然地实现模式的自动分类。
- 普化（**Generalization**）能力与抽象能力

## 1.2.4 信息的分布存放

- 信息的分布存放提供容错功能
  - 由于信息被分布存放在几乎整个网络中，所以，当其中的某一个点或者某几个点被破坏时，信息仍然可以被存取。
- 系统在受到局部损伤时还可以正常工作。
- 并不是说可以任意地对完成学习的网络进行修改。也正是由于信息的分布存放，对一类网来说，当它完成学习后，如果再让它学习新的东西，这时就会破坏原来已学会的东西。

## 1.2.5适应性 (Applicability)问题

- 擅长两个方面：
  - 对大量的数据进行分类，并且只有较少的几种情况；
  - 必须学习一个复杂的非线性映射。
- 目前应用：
  - 人们主要将其用于语音、视觉、知识处理、辅助决策等方面。
  - 在数据压缩、模式匹配、系统建模、模糊控制、求组合优化问题的最佳解的近似解（不是最佳近似解）等方面也有较好的应用。

## 1.3 历史回顾

- **1.3.1 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。



## 1.3.2 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。

## 1.3.3 反思期（1969~1982）

- **M. L. Minsky和S. Papert, 《Perceptron》, MIT Press, 1969年**
- **二十世纪70年代和80年代早期的研究成果**
- **认识规律：认识——实践——再认识**

## 1.3.4 第二高潮期（1983~1990）

- 1982年，J. Hopfield提出循环网络
  - 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

## 1.3.4 第二高潮期（1983~1990）

- 2) 1984年， J. Hopfield设计研制了后来被人们称为Hopfield网的电路。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。
- 3) 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机。

## 1.3.4 第二高潮期（1983~1990）

- 4 ) 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）
- **国内首届神经网络大会**是**1990**年**12**月在**北**京举行的。

## 1.3.5 再认识与应用研究期 (1991~)

- 问题:
- 1) 应用面还不够宽
- 2) 结果不够精确
- 3) 存在可信度的问题

## 1.3.5 再认识与应用研究期 (1991~)

- 研究：
- 1) 开发现有模型的应用，并在应用中根据实际运行情况对模型、算法加以改造，以提高网络的训练速度和运行的准确度。
- 2) 充分发挥两种技术各自的优势是一个有效方法
- 3) 希望在理论上寻找新的突破，建立新的专用/通用模型和算法。
- 4) 进一步对生物神经网络进行研究，不断地丰富对人脑的认识。