

人工神经网络

主要内容

- 人工神经网络概述

- 生物神经元

- 人工神经元

1. 人工神经网络概述

□ 什么是人工神经网络？

□ 人工神经网络是由具有适应性的简单单元组成的广泛并行互连的网络，它的组织能够模拟生物神经系统对真实世界物体所作出的交互反应-- T.Koholen

ANN研究的目的是意义

- (1) 通过揭示物理平面与认知平面之间的映射，了解它们相互联系和相互作用的机理，从而揭示思维的本质，探索智能本源。
- (2) 争取构造出尽可能与人脑具有相似功能的计算机，即ANN计算机。
- (3) 研究仿照脑神经系统的人工神经网络，将在模式识别、组合优化和决策判断等方面，取得传统计算机所难以达到的效果。

神经网络研究的发展

(1) 第一次热潮(40-60年代末)

1943年，美国心理学家McCulloch和数学家Pitts 在提出了**MP模型**

1961年，Rosenblatt等研制出了感知机

(2) 低潮(70-80年代初)

Minsky证明感知器不能实现对XOR逻辑函数，也不能实现高阶谓词结构，**错误宣判了神经网络的死刑。**

神经网络研究的发展

(3) 第二次热潮(80年代初-90年代中)

1982年，美国物理学家J.J.Hopfield提出**Hopfield 模型**：互联的非线性动力学网络，一种反复运算的动态过程，是符号逻辑处理方法所不具备的性质。

1987年首届国际ANN大会在圣地亚哥召开，国际ANN联合会成立，创办了多种ANN国际刊物。1990年12月，北京召开首届学术会议。

神经网络研究的发展

□ (4) 平稳期（90年代后期-新世纪初期）

■ 支持向量机(SVM)的兴起

□ (5) 深度神经网络(DNN)与深度学习(Deep Learning)的兴起（2006年-）

■ 2006年，加拿大多伦多大学教授、机器学习领域的泰斗Geoffrey Hinton等在《科学》上发表论文

神经网络研究的发展

- 2011年以来，微软研究院和Google先后采用DNN技术降低语音识别错误率20%~30%，是该领域十多年来最大的突破性进展。
- 2012年DNN技术在图像识别领域取得惊人的效果，在ImageNet评测上将错误率从26%降低到15%。
- 其他：微软同声翻译，2012；谷歌Google Brain，2012
- 百度深度学习研究院，2013年

神经网络的优点

- 一般而言，ANN与经典计算方法相比并非优越，只有当常规方法解决不了或效果不佳时ANN方法才能显示出其优越性。
- 对问题的机理不甚了解或不能用数学模型表示的系统，如故障诊断、特征提取和预测等问题，ANN往往是最有利的工具。
- 对处理大量原始数据而不能用规则或公式描述的问题，表现出极大的灵活性和自适应性。

神经网络的优点

- 人工神经网络以其具有自学习、自组织、较好的容错性和优良的非线性逼近能力，受到众多领域学者的关注。
- 在实际应用中，80%~90%的人工神经网络模型是采用BP网络，目前主要应用于函数逼近、模式识别、分类和数据压缩或数据挖掘。

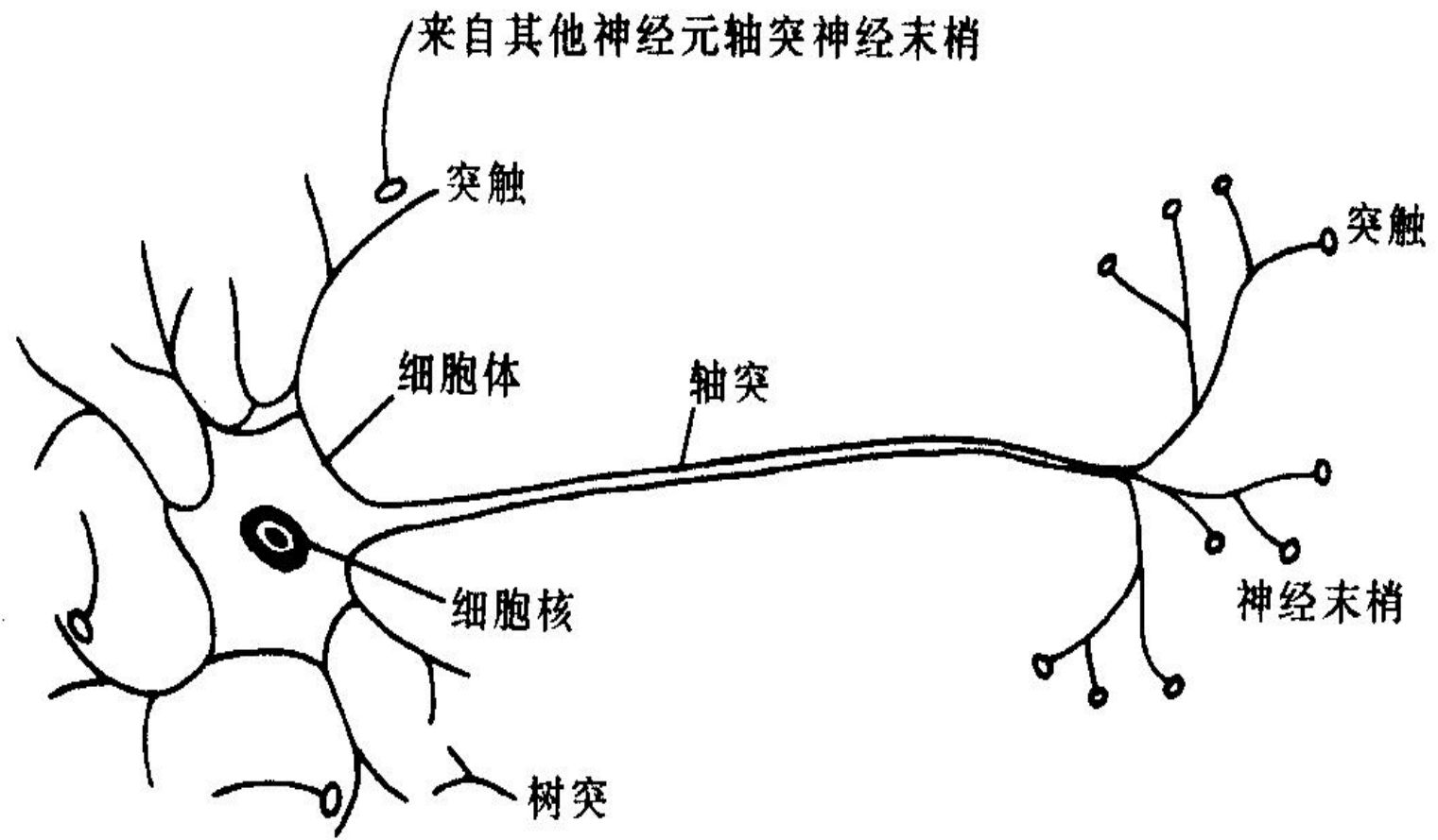
人工神经网络研究的局限性

- (1) ANN研究受到脑科学研究成果的限制
- (2) ANN缺少一个完整、成熟的理论体系
- (3) ANN研究带有浓厚的策略和经验色彩

2. 生物神经元

- 神经元是大脑处理信息的基本单元
- 人脑大约由 10^{11} 个神经元组成，神经元互相连接成神经网络
- 神经元以细胞体为主体，由许多向周围延伸的不规则树枝状纤维构成的神经细胞，其形状很像一棵枯树的枝干
- 神经元主要由细胞体、树突、轴突和突触组成

生物神经元示意图



生物神经元

□ 细胞体

- 用于处理从其他神经元传递过来的信息

□ 轴突

- 相当于神经元的输出电缆，利用尾部的神经末梢和梢端的突触输出

□ 树突

- 神经元的输入端

□ 突触

- 神经元之间相互连接的接口，即一个神经元的神经末梢与另一个神经元的树突相接触的交接面

突触的信息处理

- 神经元各组成部分的功能来看，信息的处理与传递主要发生在突触附近
- 神经元细胞体通过轴突将脉冲传递到突触前膜
- 当脉冲幅度达到一定强度，即超过其阈值电位后，突触前膜将向突触间隙释放神经传递的化学物质

突触的信息处理

- 生物神经元传递信息的过程为多输入、单输出
- 突触有两种类型：兴奋性突触和抑制性突触。前者产生正突触后电位，后者产生负突触后电位

突触的传递功能与特点

□ 不可逆性

- 脉冲只从突触前传到突触后，不逆向传递

□ 对信息传递具有时延和不应期，一般为0.3~1ms

- 相邻的两次输入之间，需要一定的间隔；在此间隔内，突触不传递信息

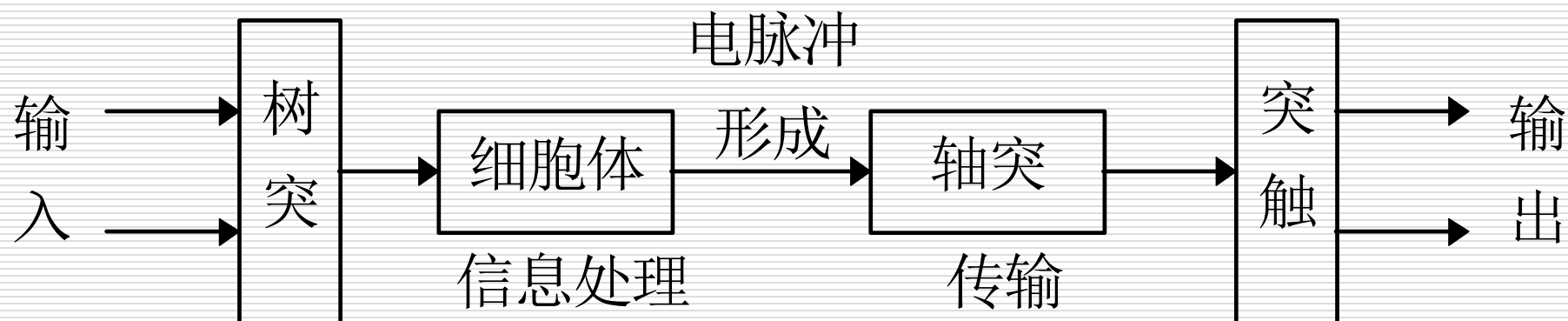
□ 可塑性

- 突触传递信息的强度是可变的，即具有学习功能

□ 学习、遗忘或疲劳（饱和）效应

- 对应突触传递作用增强、减弱和饱和

生物神经网络基本模型



3. 人工神经元

□ 人工神经网络

神经网络是一个并行和分布式的信息处理网络结构

它一般由大量神经元组成

- 每个神经元只有一个输出，可以连接到很多其他的神经元
- 每个神经元的输入有多个连接通道，每个连接通道对应于一个连接权系数

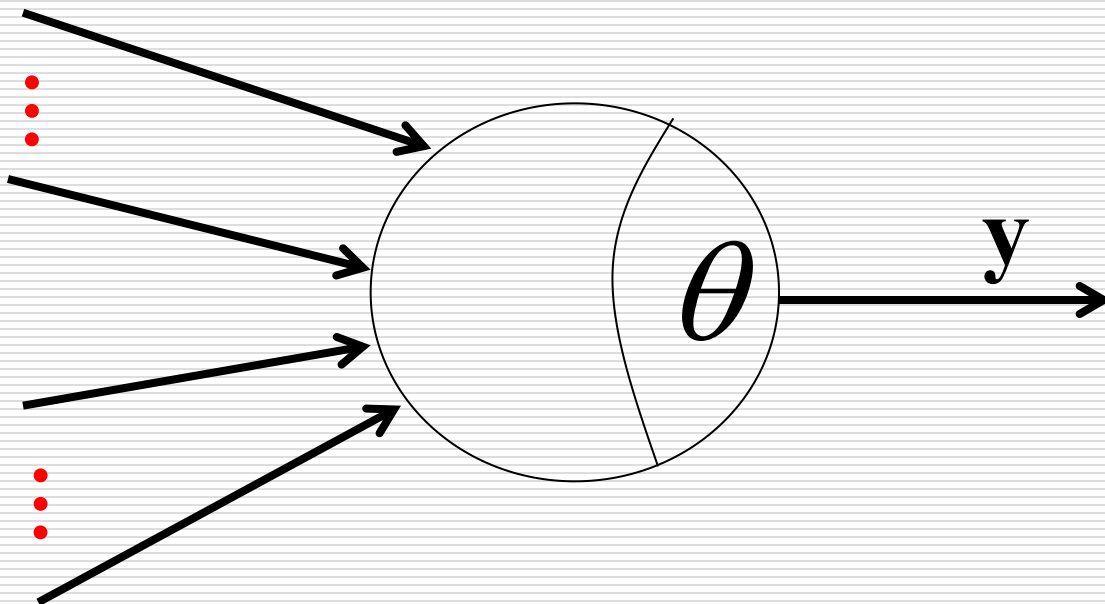
3. 人工神经元

□ M-P神经元模型

心理学家麦克洛奇(W.McCulloch)和数理逻辑学家皮兹(W.Pitts)于1943年首先提出了一个简化的神经元模型

兴奋性
输入 e

抑制性
输入 i



3. 人工神经元

□ 说明

- 圆表示神经元的细胞体；
- 外部输入包括兴奋性输入和抑制性输入
- 符号 θ 表示神经元兴奋的阈值

M-P模型的输入输出关系

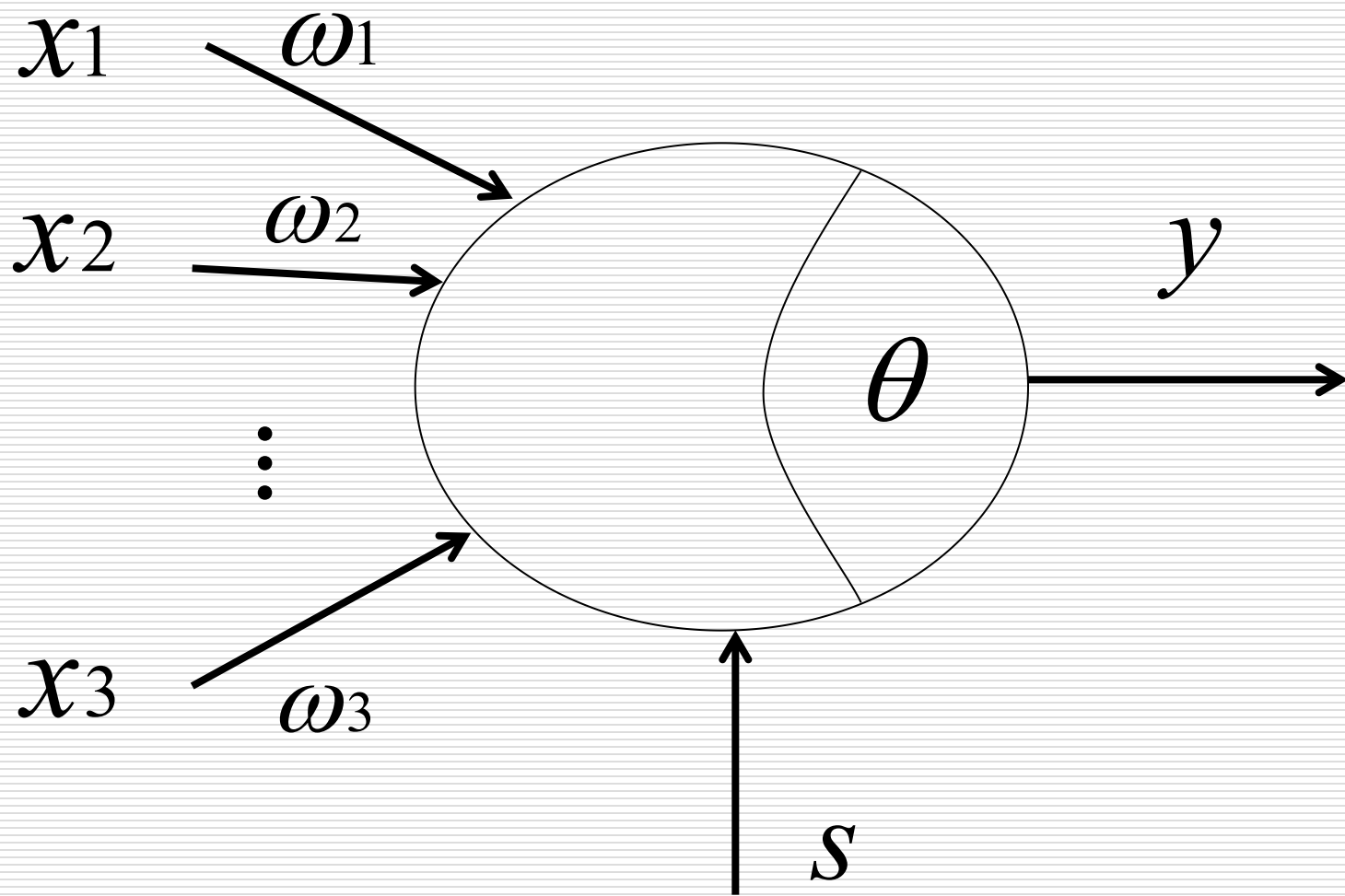
$$y = \begin{cases} 1, & \text{当 } \sum e \geq \theta, \sum i = 0 \\ 0, & \text{当 } \sum e \geq \theta, \sum i > 0 \\ 0, & \text{当 } \sum e < \theta, \sum i \leq 0 \end{cases}$$

其中， y 表示输出， θ 表示阈值；

e 表示兴奋性输入；

i 表示抑制性输入；

常用的神经元结构模型



模型说明

x_i 是神经元的输入； ω_i 为神经元与各输入的连接强度，称为连接权值；

θ 称为神经元的阈值； s 为外部输入的控制信号，可以用来调整神经元的连接权值；

y 是神经元的输出；

模型说明

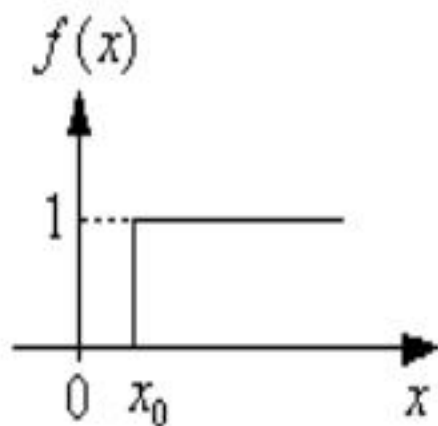
神经元的工作过程:

- (1) 从各输入端接收输入信号 z ;
- (2) 根据连接权值 ω_i , 求出所有输入的加权和 σ :

$$\sigma = \sum_{i=1}^n \omega_i x_i + s - \theta$$

- (3) 用某一特性函数 f 进行转换, 得到输出 y :

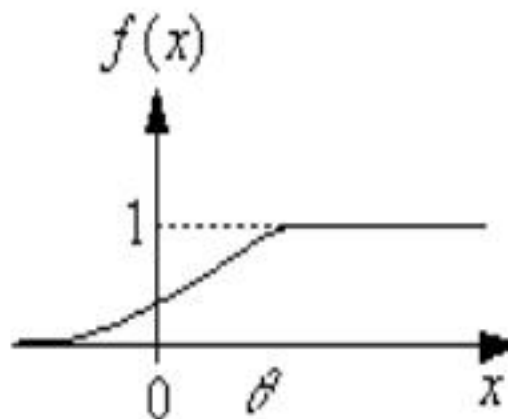
$$y = f(\sigma) = f\left(\sum_{i=1}^n \omega_i x_i + s - \theta\right)$$



(a)

$$f(x) = \begin{cases} 1, & x \geq x_0 \\ 0, & x < x_0 \end{cases}$$

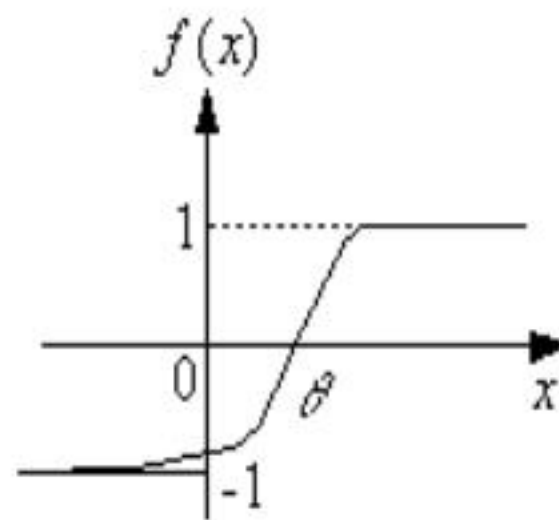
阈值函数



(b)

$$f(x) = \frac{1}{1 + e^{-ax}}, \quad 0 < f(x) < 1$$

S型函数



(c)

$$f(x) = \frac{1 - e^{-ax}}{1 + e^{-ax}}, \quad -1 < f(x) < 1$$

双曲正切函数

4. 人工神经网络的互联结构

□ ANN可分为两种

■ 相互连接网络

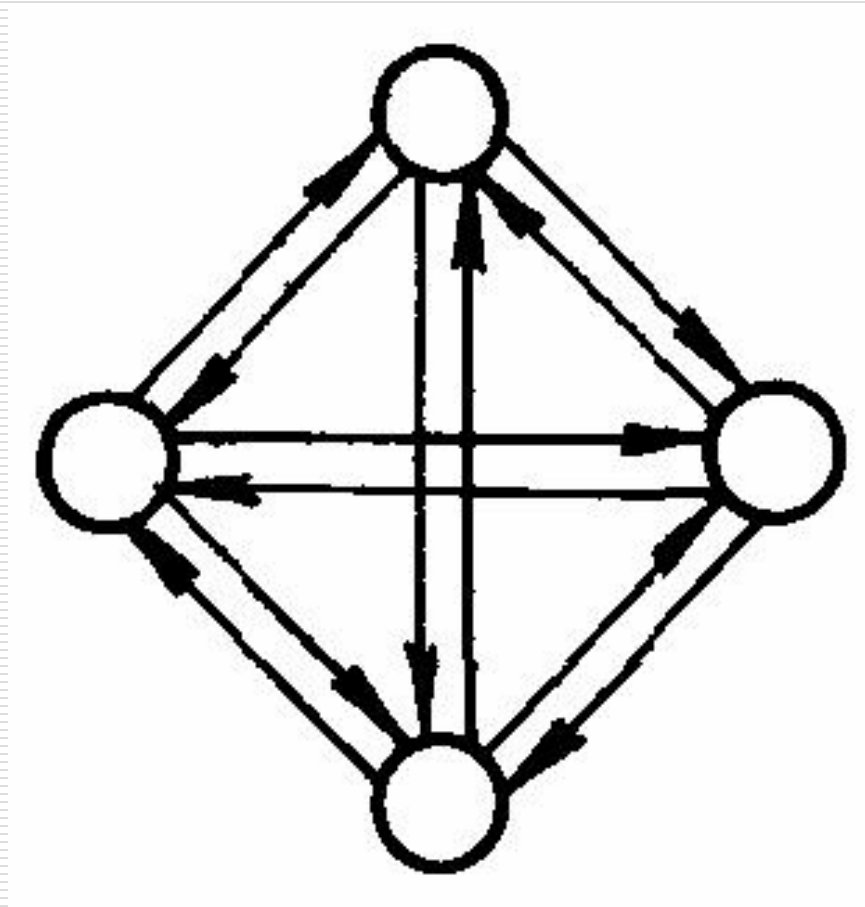
■ 分层网络

□ 单层

□ 两层网络结构

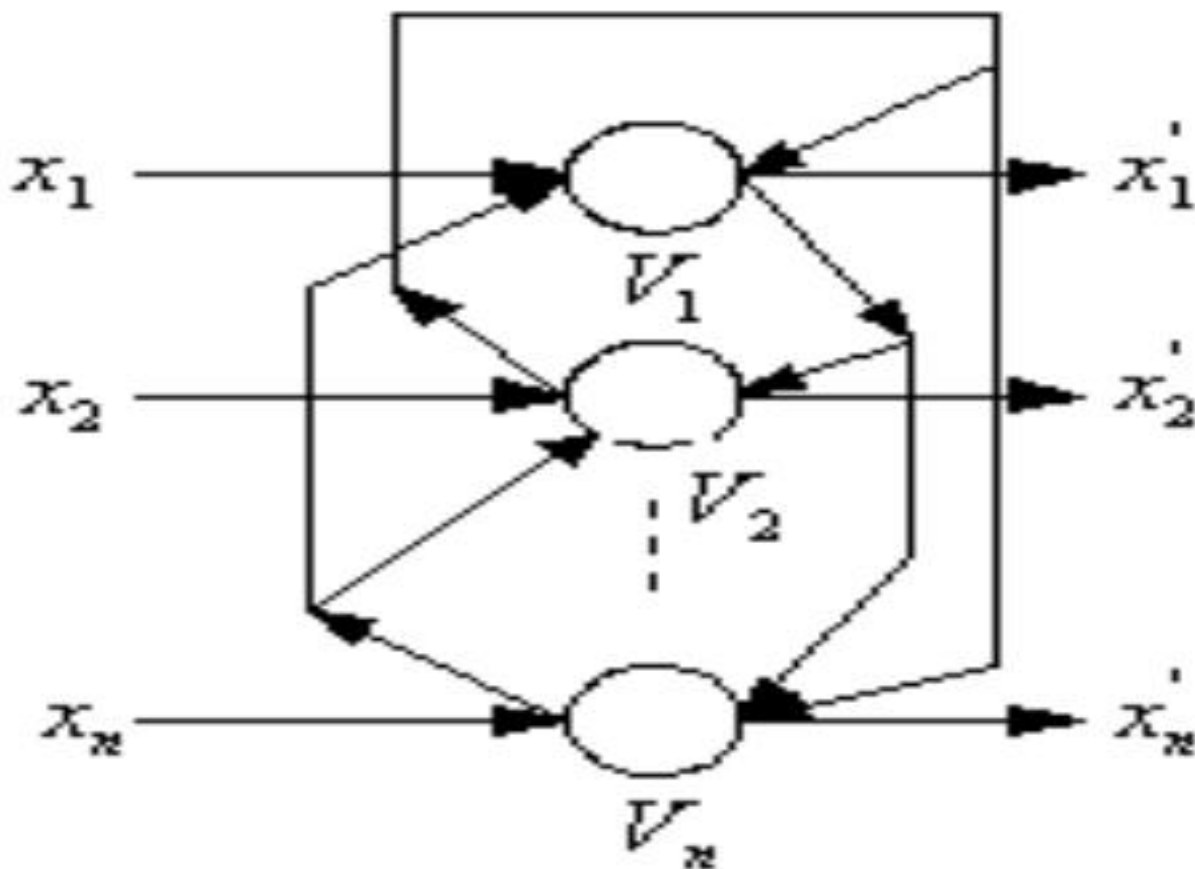
□ 多层网络结构

典型的网络结构



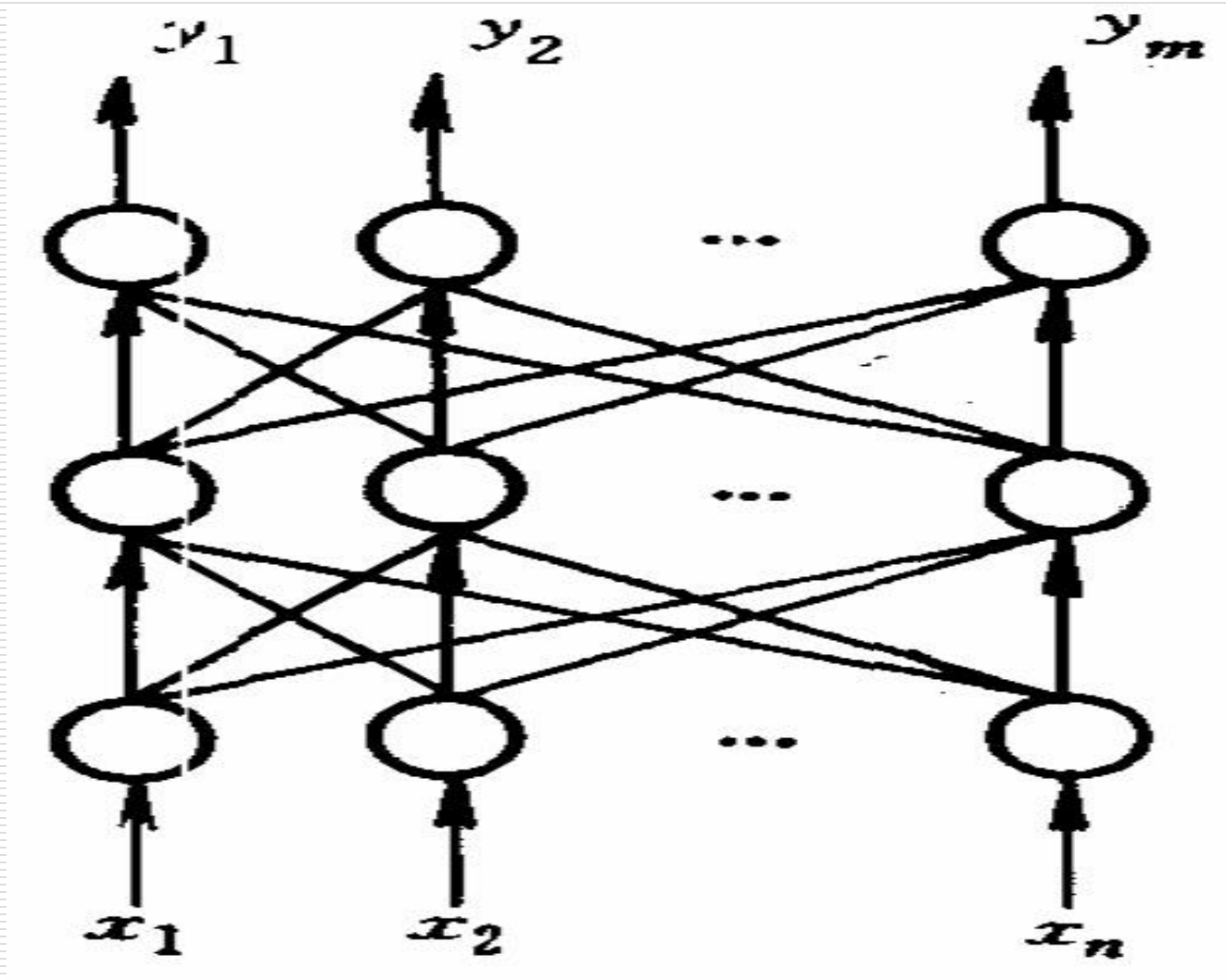
相互连接网络

典型的网络结构



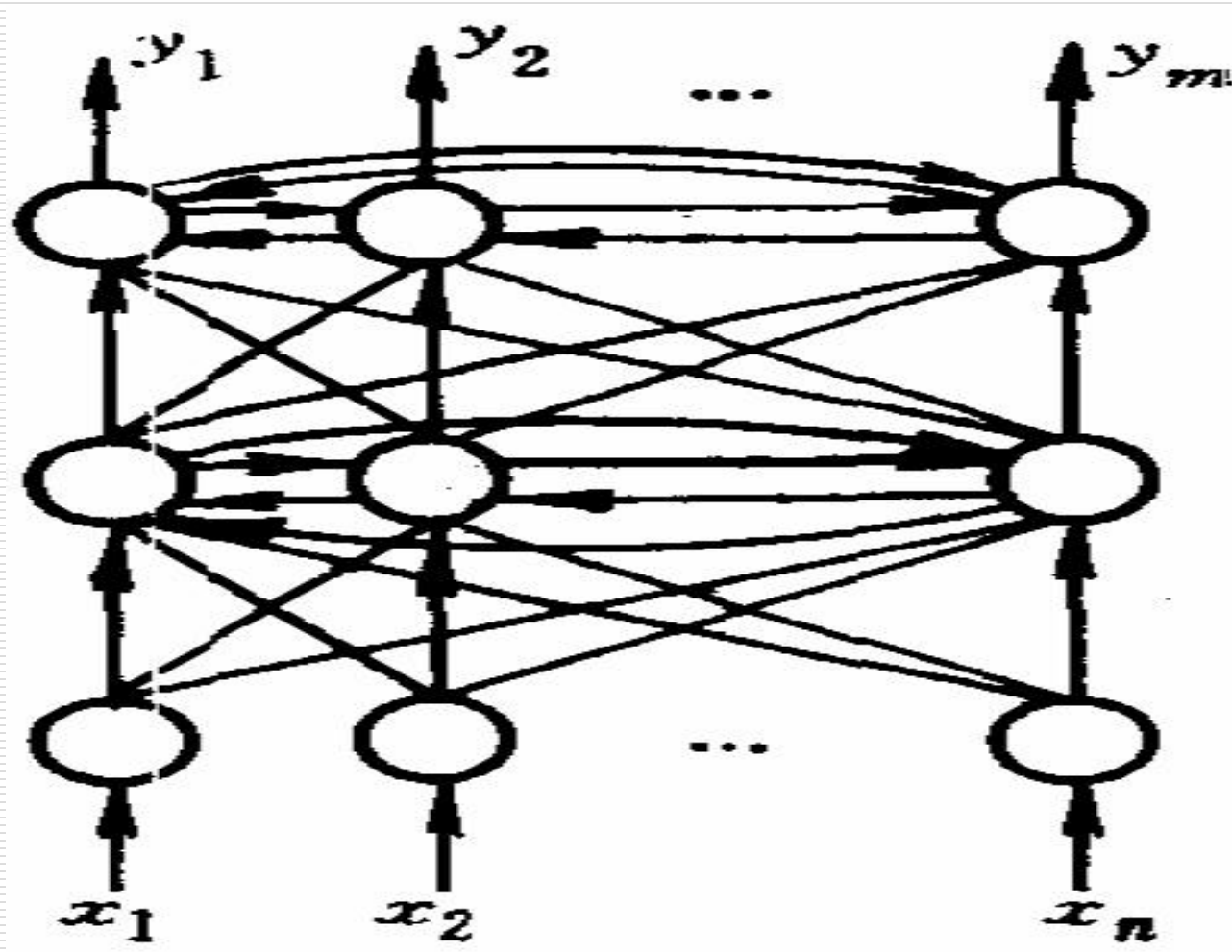
单层网络

典型的网络结构



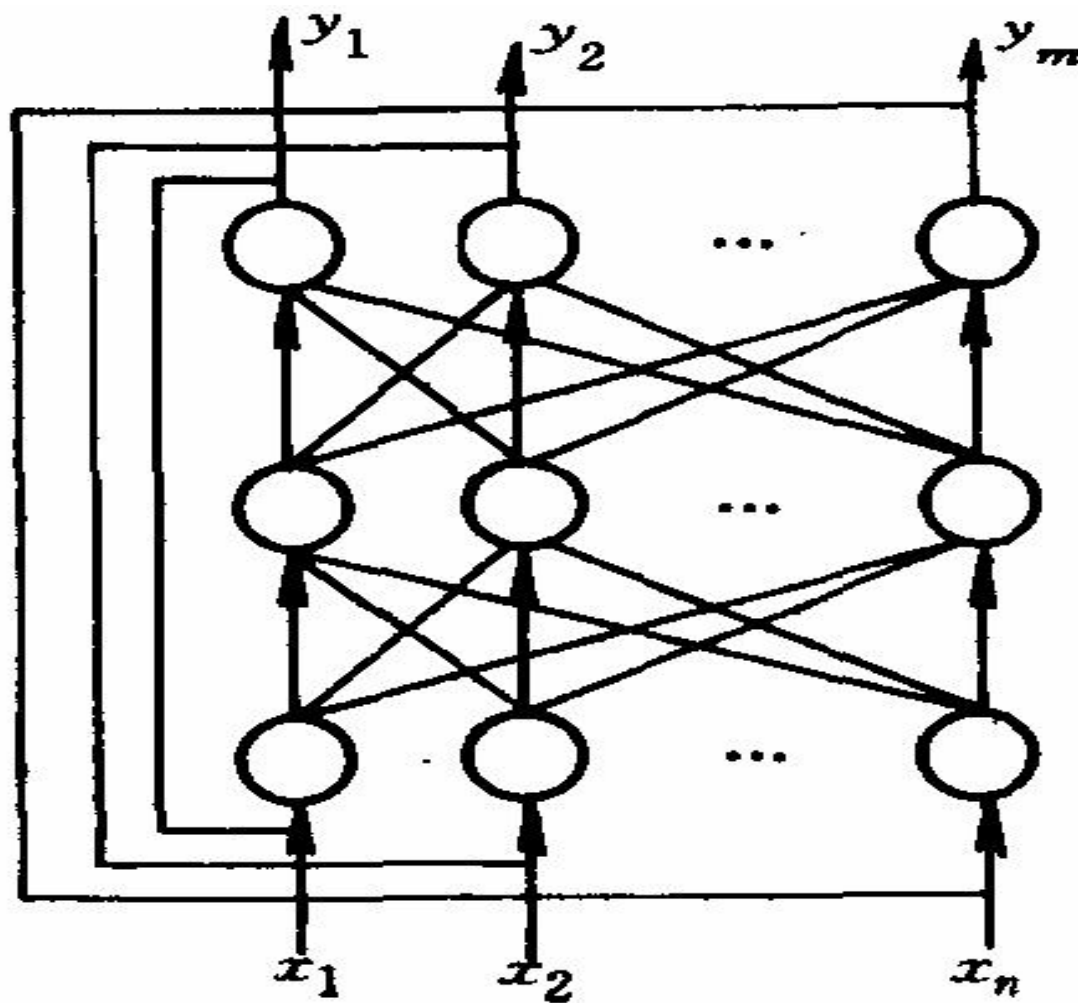
多层前向网络

典型的网络结构



带侧抑制的多层网络

典型的网络结构

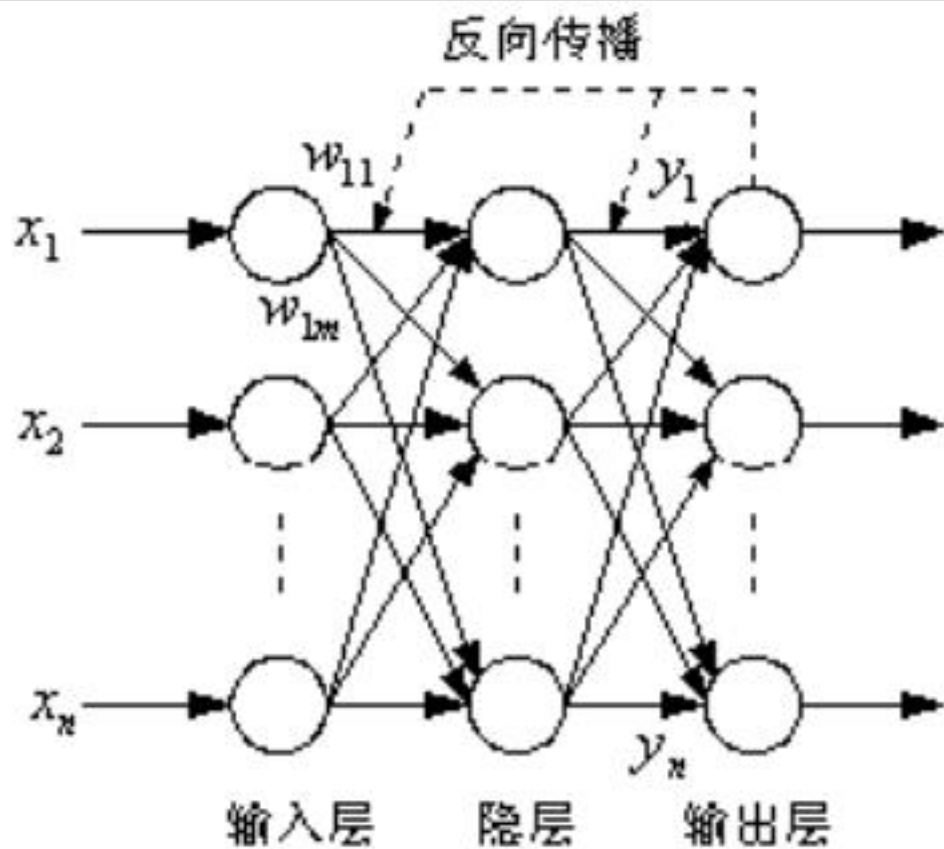
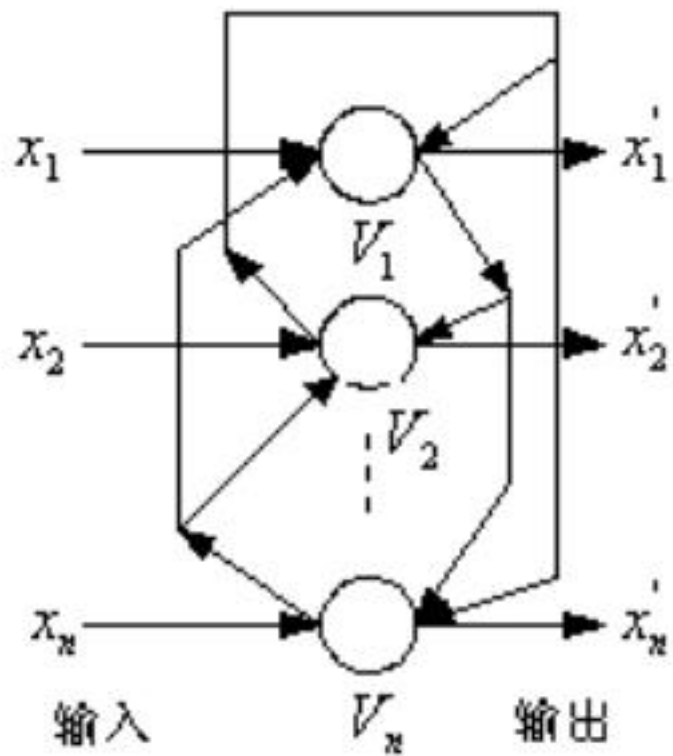


带反馈的多层网络

多层神经网络

- 3层及3层以上的神经网络，称多层神经网络
- 深度神经网络属于多层神经网络
- 按照各层功能可分为
 - 输入层
 - 中间层（隐层）：可包含多层
 - 输出层

多层神经网络



小 结

- 神经网络的定义
- 神经元模型
- 神经网络的典型结构及其分类