

人工智能概论-计算智能概述&神经网络

赵亚伟

zhaoyw@ucas.ac.cn

中国科学院大学 大数据分析技术实验室

2018.6.9

目录

- 计算智能概述
 - 神经计算
 - 小结
-

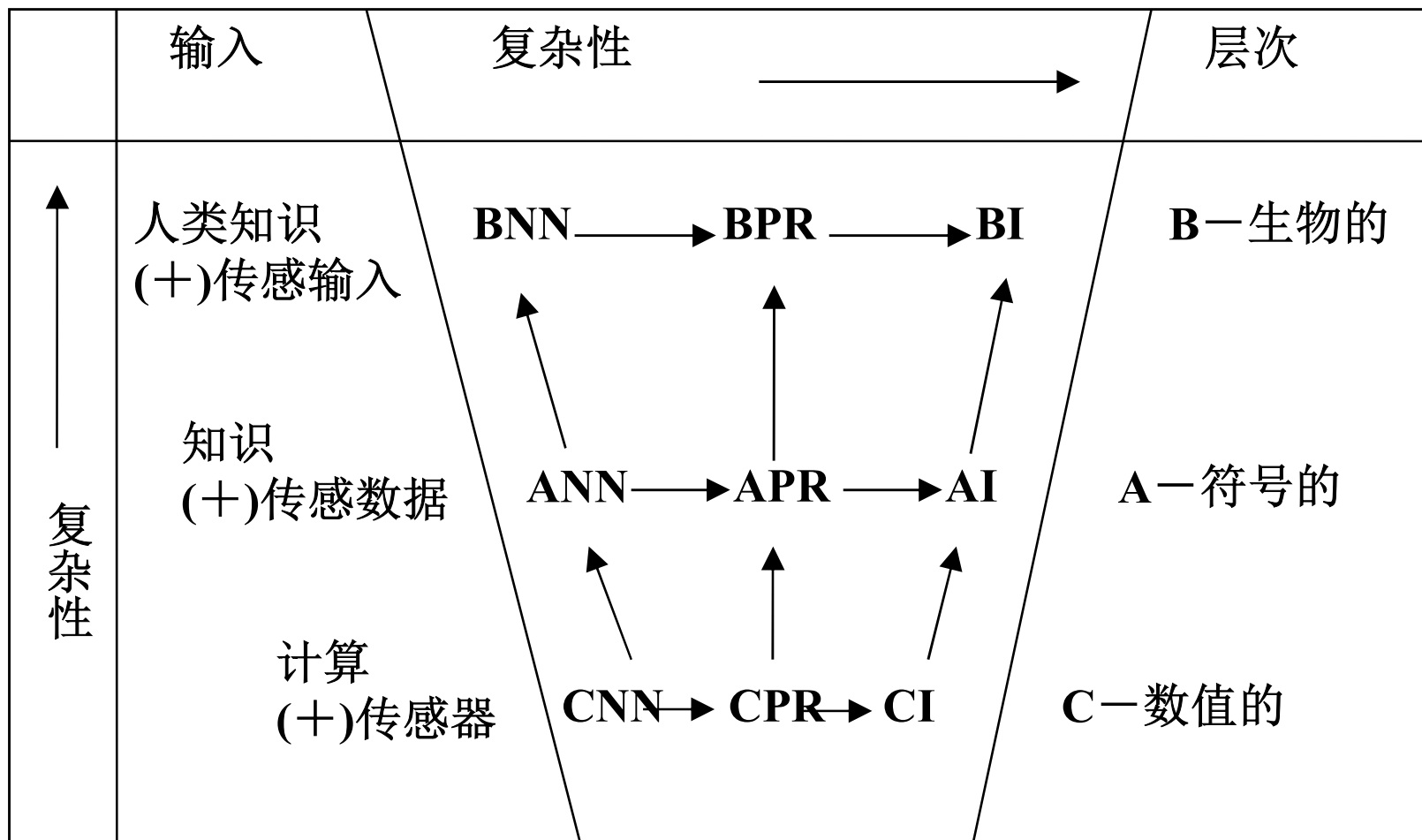
概述

- 信息科学与生命科学的相互交叉、相互渗透和相互促进是现代科学技术发展的一个显著特点。
- 计算智能涉及神经网络、模糊逻辑、进化计算和人工生命等领域，它的研究和发展正反映了当代科学技术多学科交叉与集成的重要发展趋势。

什么是计算智能

- 把神经网络（NN）归类于人工智能（AI）可能不大合适，而归类于计算智能（CI）更能说明问题实质。进化计算、人工生命和模糊逻辑系统的某些课题，也都归类于计算智能。
- 计算智能取决于制造者（manufacturers）提供的数值数据，不依赖于知识；另一方面，人工智能应用知识精品（knowledge tidbits）。人工神经网络应当称为计算神经网络。

计算智能与人工智能的区别和关系



-
- 上图由贝兹德克于1994年提出，表示ABC与神经网络（NN）、模式识别（PR）和智能（I）之间的关系
 - A—Artificial，表示人工的（非生物的）；
 - B—Biological，表示物理的+化学的+（？）=生物的；
 - C—Computational，表示数学+计算机
 - 计算智能是一种智力方式的低层认知，它与人工智能的区别只是认知层次从中层下降至低层而已。中层系统含有知识（精品），低层系统则没有。

ABC及其相关领域的定义

BNN	人类智能硬件：大脑	人的传感输入的处理
ANN	中层模型： CNN + 知识精品	以大脑方式的中层处理
CNN	低层，生物激励模型	以大脑方式的传感数据处理
BPR	对人的传感数据结构的搜索	对人的感知环境中结构的识别
APR	中层模型： CPR +知识精品	中层数值和语法处理
CPR	对传感数据结构的搜索	所有 CNN +模糊、统计和确定性模型
BI	人类智能软件：智力	人类的认知、记忆和作用
AI	中层模型： CI +知识精品	以大脑方式的中层认知
CI	计算推理的低层算法	以大脑方式的低层认知

目录

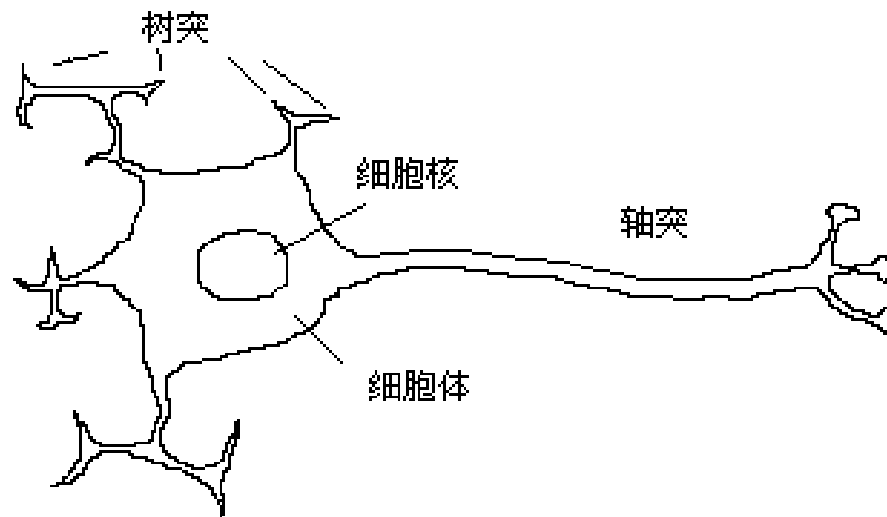
- 计算智能概述
 - 神经计算
 - 小结
-

神经计算-人工神经网络研究的进展

- **1943年**麦卡洛克和皮茨提出神经网络模型（称为MP模型）的概念。
- **20世纪60年代**威德罗和霍夫提出自适应线性元件。
- **60年代末期至80年代中期**，整个神经网络研究处于低潮。
- **80年代后期以来**，人工神经网络研究得到复苏和发展，在模式识别、图象处理、自动控制等领域得到广泛应用。

人工神经网络的特性

- 并行分布处理
- 非线性映射
- 通过训练进行学习
- 适应与集成
- 硬件实现

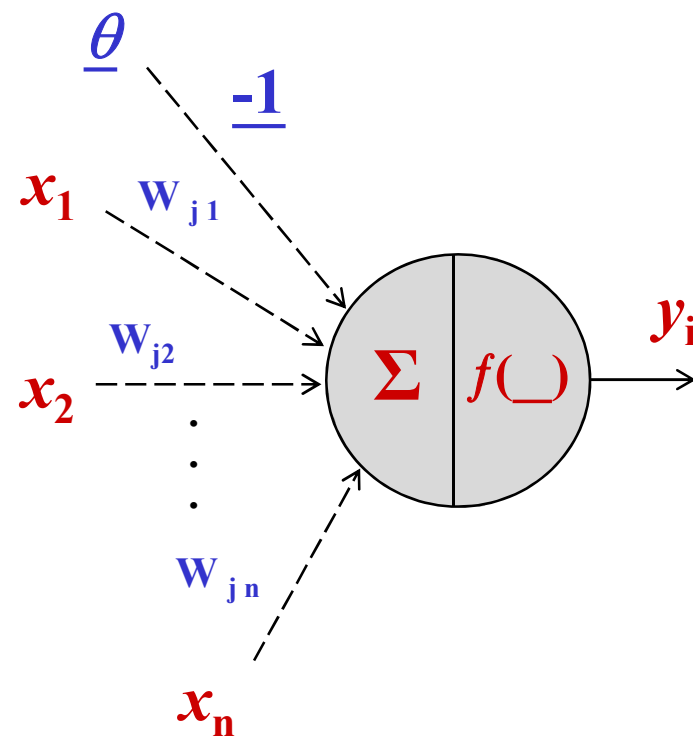


人工神经网络的结构

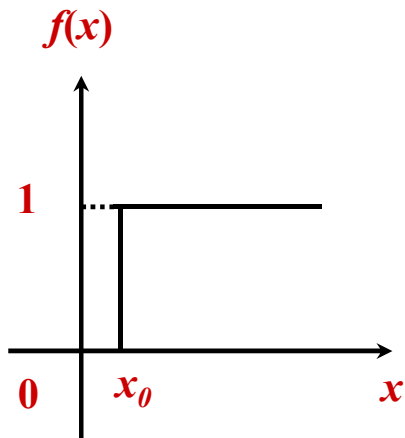
- 神经元单元由多个输入 x_i , $i=1, 2, \dots, n$ 和一个输出 y 组成。中间状态由输入信号的权和表示, 而输出为

$$y_j(t) = f\left(\sum_{i=1}^n w_{ji}x_i - \theta_j\right)$$

- 式中, θ_j 为神经元单元的偏置 (阈值), w_{ji} 为连接权系数。 n 为输入信号数目, y_j 为神经元输出, t 为时间, $f()$ 为输出变换函数

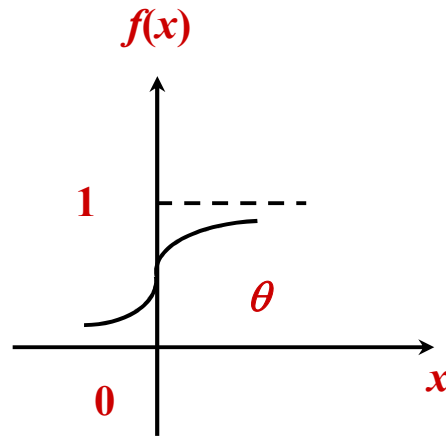


神经元中的激发函数



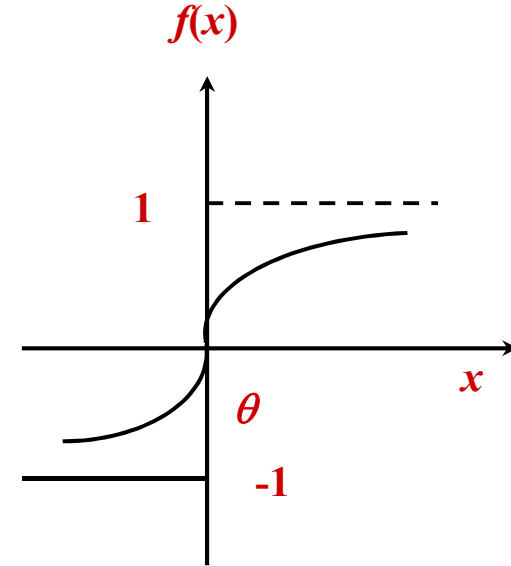
(a)

(a) 二值函数



(b)

(b) S形函数



(c)

(c) 双曲正切函数

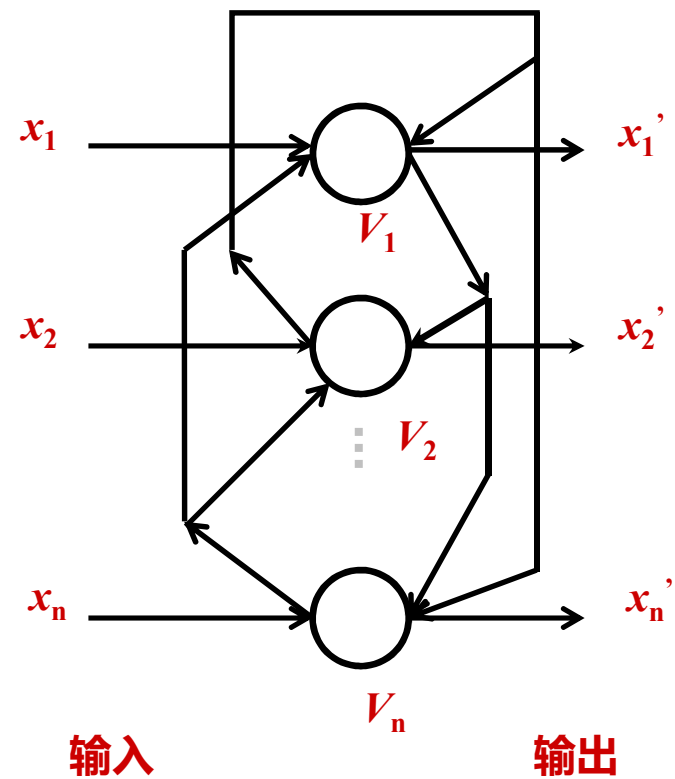
人工神经网络的基本特性和结构

□ 人工神经网络是具有下列特性的有向图：

- 对于每个节点 i 存在一个状态变量 x_i ；
- 从节点 j 至节点 i ，存在一个连接权系统数 w_{ji} ；
- 对于每个节点 i ，存在一个阈值 θ_i ；
- 对于每个节点 i ，定义一个变换函数 f_i ；对于最一般的情况，此函数取形式。

$$f_i(\sum_j w_{ij} x_j - \theta_i)$$

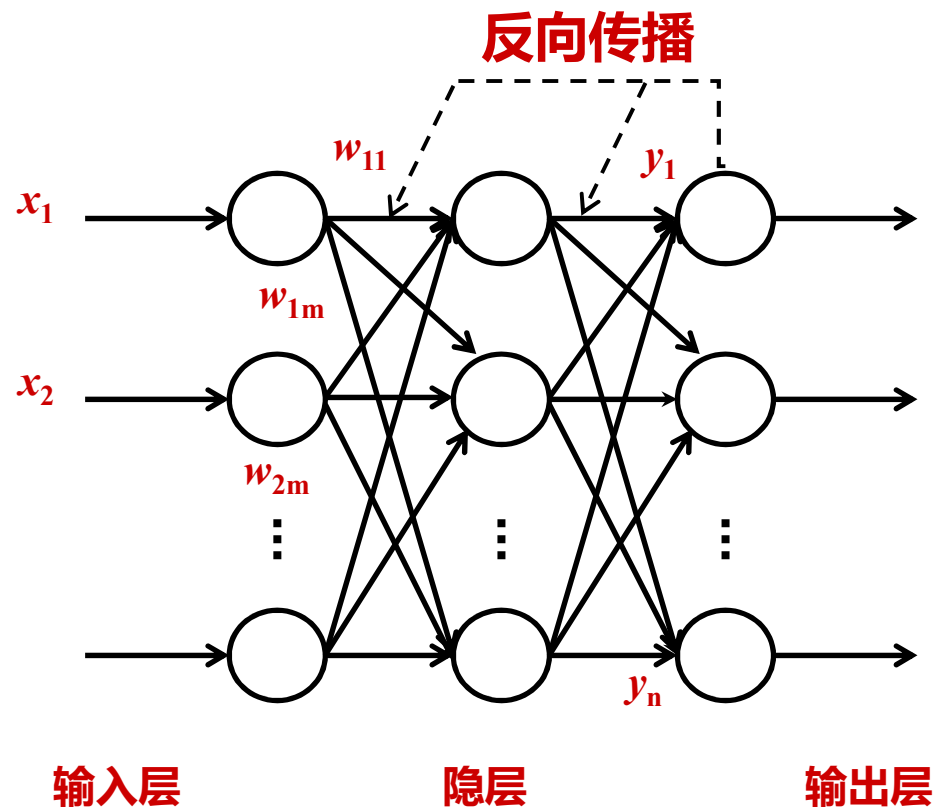
- **递归（反馈）网络：**在递归网络中，多个神经元互连以组织一个互连神经网络。



反馈网络

前馈网络

- **前馈（多层）网络**:前馈网络具有递阶分层结构，由同层神经元间不存在互连的层级组成。



人工神经网络的主要学习算法

- **有监督学习**：能够根据期望输出和实际网络输出（对应于给定输入）间的差别来调整神经元间连接的强度（权）。
- **无监督学习**：不需要知道期望输出。
- **强化学习**：采用一个“评论员”来评价与给定输入相对应的神经网络输出的优度（质量因数）。强化学习算法的一个例子是遗传算法（GA）。

基于神经网络的知识表示与推理

□ 1.基于神经网络的知识表示

- 神经网络中的知识表示是一种**隐式的表示方法**。
- 知识并不像在产生式系统中那样独立地表示为每一条规则，而是将某一问题的若干知识在同一网络中表示。
- 例如，在有些神经网络系统中，知识是用神经网络所对应的有向权图的邻接矩阵及阈值向量表示的。

例子：神经网络实现与逻辑

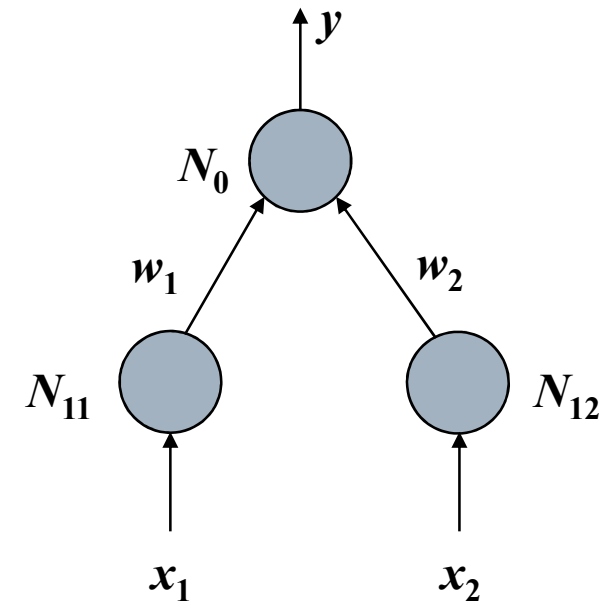
□ 使用一个二层神经网络实现与逻辑 (and)

- x_1 、 x_2 为网络输入
- w_1 、 w_2 为连接边的权值
- y 为网络的输出
- 定义一个输入输出关系函数：

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

其中， $\theta = 0.5$

- 根据网络的定义，输出 $y = f(x_1w_1 + x_2w_2)$ ，只要有一组合适的权值 w_1 、 w_2 ，就可以使输入数据 x_1 、 x_2 和输出 y 之间符合与逻辑



x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

□ 根据实验得到几组 w_1 、 w_2 的权值数据满足与关系

w_1	w_2		x_1	x_2	y
0.20	0.35	}	0	0	0
0.20	0.40		0	1	0
0.25	0.30		1	0	0
0.40	0.20		1	1	1

□ 验证:

■ $y = f(x_1w_1 + x_2w_2) = f(0 \times 0.20 + 0 \times 0.35) = 0$

■ $y = f(x_1w_1 + x_2w_2) = f(0 \times 0.20 + 1 \times 0.35) = 0$

■ $y = f(x_1w_1 + x_2w_2) = f(1 \times 0.20 + 0 \times 0.35) = 0$

■ $y = f(x_1w_1 + x_2w_2) = f(1 \times 0.20 + 1 \times 0.35) = 1$

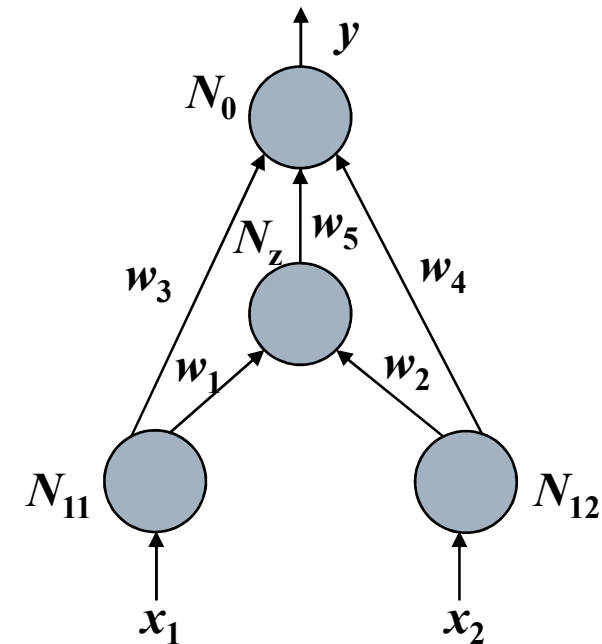
例子：神经网络实现异或逻辑

- ❑ 异或运算：相同为假
- ❑ 设计网络结构如图
- ❑ 训练后得到一组权值
- ❑ $(0.3, 0.3, 1, 1, -2)$
- ❑ 演算： $x_1=1, x_2=1$

$$\begin{aligned} z &= f(x_1 w_1 + x_2 w_2) \\ &= f(1 \times 0.3 + 1 \times 0.3) = 1 \end{aligned}$$

$$\begin{aligned} z &= f(x_1 w_3 + x_2 w_4 + z w_5) \\ &= f(1 \times 1 + 1 \times 1 + 1 \times (-2)) \\ &= f(0) = 0 \end{aligned}$$

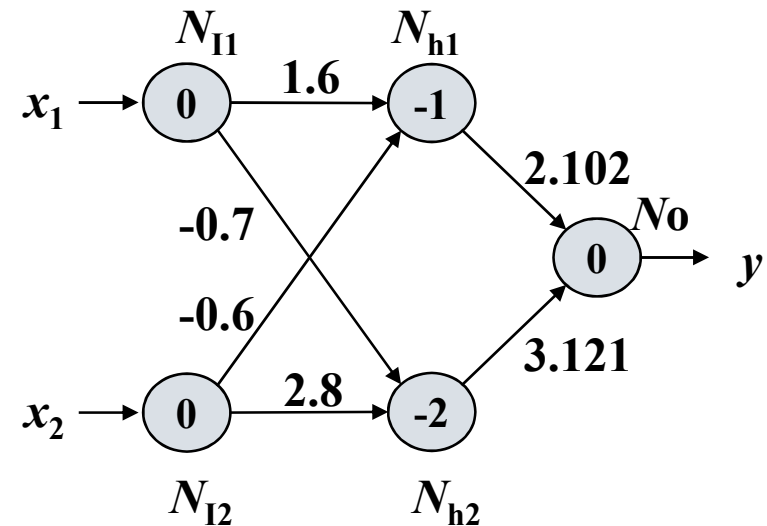
成立



x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

例子：异或逻辑的神经网络表示

- 实现异或逻辑的网络不是唯一的，下面的网络也可以实现
- 可采用神经网络对应的有向赋权图的邻接矩阵和阈值向量表示
- 节点阈值向量为：
(0, 0, -1, -2, 0)



表示“异或”问题的两层感知器模型

	N_{I1}	N_{I2}	N_{h1}	N_{h2}	N_o
N_{I1}	0	0	1.6	-0.7	0
N_{I2}	0	0	-0.6	2.8	0
N_{h1}	0	0	0	0	2.102
N_{h2}	0	0	0	0	3.121
N_o	0	0	0	0	0

基于神经网络的推理

- 基于神经网络的推理实质上是在一个已经训练成熟的网络上对未知样本进行预测或分类判断，其实是计算 $f(_)$ 的值
 - 网络的训练一般是根据已知输入 $\{(x_{11}, x_{12}, \dots, x_{1n}, y_1), (x_{21}, x_{22}, \dots, x_{2n}, y_2), \dots, (x_{m1}, x_{m2}, \dots, x_{mn}, y_m)\}$ 计算网络的一组权值 $\{w_1, w_2, \dots, w_k\}$ 的过程
 - 回顾：
 - 确定性推理：搜索 + $f(_)$ ；消解 $f(_)$ （谓词演算）
 - 不确定性推理：Bayes 求解后验概率 $P(H|E)$ ；可信度推理求解 $CF(H, E)$ ；证据理论求解 $Bel(A), Pl(A)$
 - 一切皆计算！推理就是计算
-

基于神经网络的推理：泛化能力

- 神经网络的训练是一个网络对训练样本内在规律的学习过程，而对网络进行训练的目的在于为了让网络模型模型对训练样本以外的数据具有正确的**映射能力**
 - 通常定义神经网络的**泛化能力**（也称**推广能力**），是指神经网络在训练完成之后输入训练样本以外的新数据时获得正确输出的能力，指**泛化性能**
 - **内插泛化**：也称弱泛化，在测试集上的泛化能力，训练集和测试集的分布是相同的
 - **外推泛化**：也称强泛化，在真实样本集上的泛化能力，分布与测试集或训练集可能不同
-

基于神经网络的推理：泛化能力Cont.

□ 影响泛化能力的三要素：

- 训练样本的质量和数量
- 网络结构
- 问题本身的复杂程度

□ 学习时间：网络的训练次数

- 学习时间增加可以提高网络的精度
 - 但也可能导致过拟合（对训练样本的误差逐渐减小到某一个很小的值），泛化能力降低，尤其针对外推测试
 - 并不是训练误差越小越好，而是要从实际出发，提高外推映射能力，提高泛化性能
-

基于神经网络的推理：泛化能力Cont.

- 泛化性能还体现在网络对噪声的抗干扰能力
 - 泛化过程分为三个阶段：
 - 第一阶段，泛化误差单调下降
 - 第二阶段，泛化动态较为复杂，泛化误差将达到最小值
 - 第三阶段，泛化误差又将单调上升
 - 最佳的泛化能力往往出现在训练误差的全局最小点出现之前，存在一个最优的训练停止时间，最优停止法设计神经网络的主要思想
-

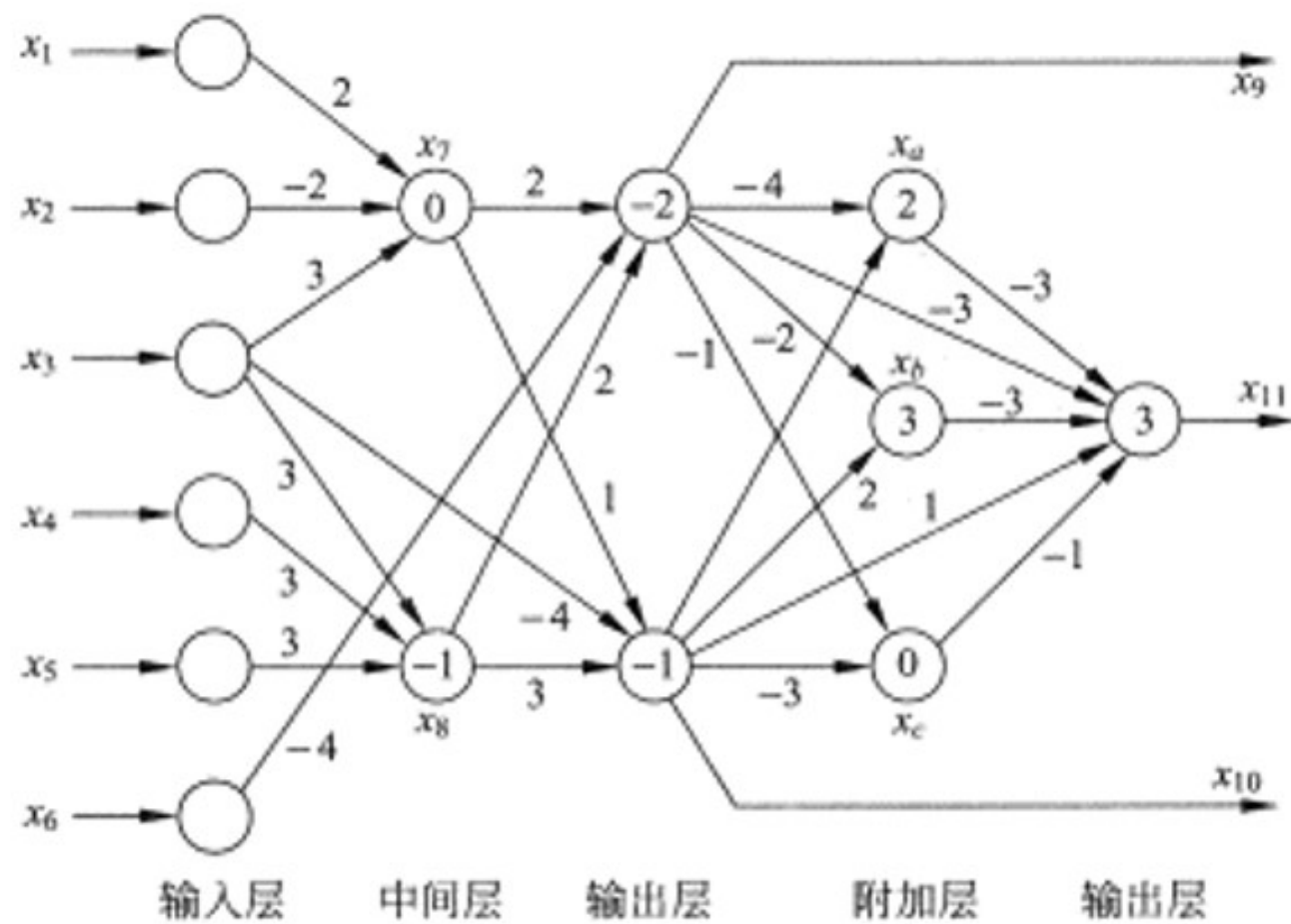
例子：基于B-P网络的医疗诊断系统

□ 该医疗诊断系统只考虑 6 种症状、2 种疾病、3种治疗方案

- 症状：对每一症状采集三种信息 — 有（1）、无（-1）、没有记录（0）
- 疾病：对每一疾病采集三种信息 — 有（1）、无（-1）、没有记录（0）
- 治疗方案：对每一治疗方案采集两种信息 — 是、否

□ 每个病人的信息构成一个训练样例，用一批训练样例对网络进行训练（B-P算法），假设得到的是如图所示的神经网络

- $x_1 - x_6$ 为症状（输入）
 - x_7, x_8 为疾病名
 - x_9, x_{10}, x_{11} 为治疗方案（输出）
-



正向网络推理的步骤

- 基于神经网络的推理是通过网络计算实现的。把用户提供的初始证据用作网络的输入，通过网络计算最终得到输出结果。
- 一般来说，正向网络推理的步骤如下：
 - 把已知数据输入网络输入层的各个节点。
 - 利用特性函数分别计算网络中各层的输出。
 - 用阈值函数对输出层的输出进行判定，从而得到输出结果。

ANN常见网络及其算法

□ 包括

- 自适应谐振理论（ART）网络
- 学习矢量量化（LVQ）网络
- Kohonen网络
- Hopfield网络

□ 参考教材或网上资源学习

目录

- 计算智能概述
 - 神经计算
 - 小结
-

小结

- **计算智能**理解为智力的低层认知，取决于数值数据而不依赖于知识。**人工智能**是在计算智能的基础上引入知识而产生的智力中层认知，**生物智能**尤其是人类智能，则是最高层次智能，即

$$CI \subset AI \subset BI$$

- 神经网络的基元是**神经元**，具有多个输入和一个输出，神经元之间是带权的有向连接，输入借助**激励函数**获得输出
 - 人工神经网络种类繁多，其中**反向传播**和**Hopfield网络**应用广泛，如目前NLP领域的RNN就是一种Hopfield网络
-