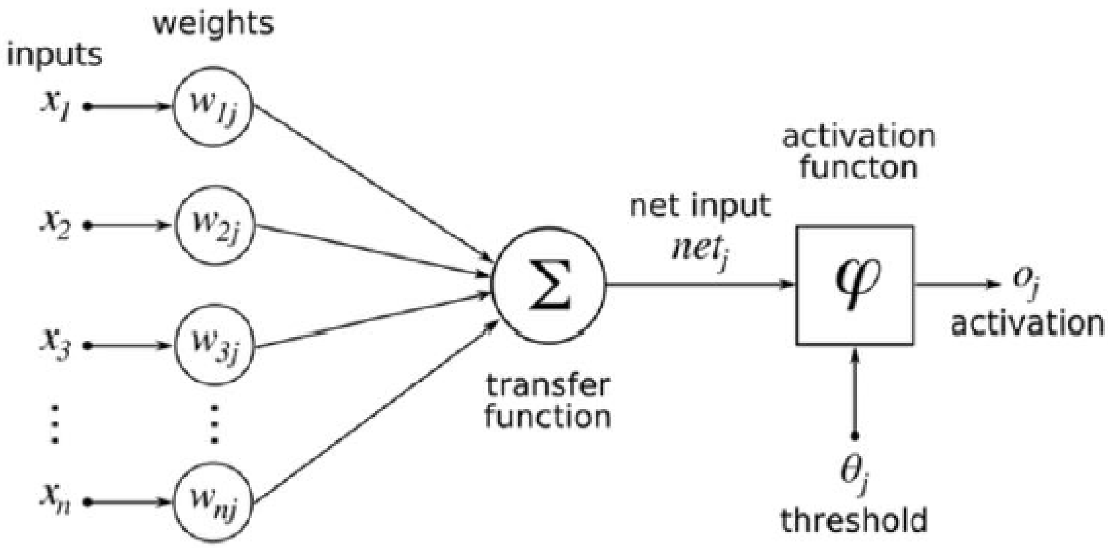
INT301 W1 Note

神经传输的原理



如果来自一个或多个神经元的加权输入信号的总和（求和）足够大（超过阈值）以导致消息传输，则单个神经元将通过此接口将消息传递给另一个神经元。

x_n 是输入信号，w_{nj} 是权重，则 x_n w_{nj} 是加权输入信号。所有 x_n w_{nj} 的总和会当作参数输入一个激活函数 φ，如果得到的结果大于某个阈值，则传递信号。

神经网络的阶段

神经网络分为两个阶段，即：

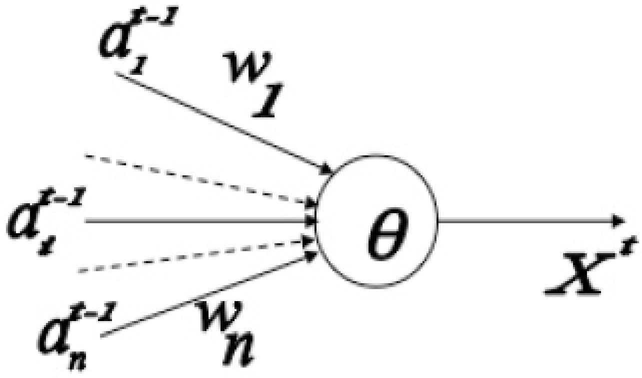
- learning phase，训练网络的阶段
- application phase，用训练好的网络验证或者推理的阶段。

The McCulloch-Pitts Neuron（MP

neuron)

作者将神经元建模为：

- A binary discrete-time element
- With ***excitatory*** and ***inhibitory*** inputs and an excitation threshold



在某一时刻 t，从第 i 个上一个突触来的 a^t_i 值为 **1** 或 **0**。

连接上的权重 w_i 有两个值：

- excitatory type connection 的 **+1**，
- inhibitory type connection 的 **-1**

以及 excitation threshold θ。

神经在下一个时刻 t+1 的输出 x^{t+1} 可以由下式定义：

$$x^{t+1}=1 \text{ if and only if } S^t=\sum_i w_i a_i^t \geq \theta$$

在 MP 神经元中，我们称某一个时刻的总输入 S^t 为：***instant state of the neuron***

注：上文说到的 binary discrete-time element，binary 指输入只有 0 和 1，而 discrete-time 则指对于每个时刻进行建模。

MP 神经元的状态 S^t 不依赖于神经元本身的先前状态，可以写为：

$$S^t=\sum_i w_{ij} a_i^t=f(t)$$

神经元输出 x^{t+1} 是其状态 S^t 的函数，因此输出也可以写成离散时间的函数：

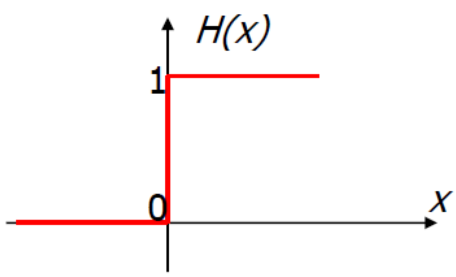
$$x(t)=g\left(S^t\right)=g(f(t))$$

其中 g 是 ***threshold activation function***：

$$g\left(S^t\right)=H\left(S^t-\theta\right)=\begin{cases}1, \text{ if } & S^t \geq \theta \\ 0, \text{ if } & S^t<\theta\end{cases}$$

这里 H 是 the Heaviside (unit step) function：

$$H(X)=\begin{cases}1, & x \geq 0; \\ 0, & x<0.\end{cases}$$



MP neuron 为机器（单元网络）奠定了基础，能够：1，存储信息，和 2，进行逻辑和算术运算。

ANN learning rules

接下来是大脑的另一个功能，learning 功能。

这个功能由改变权重 w_{ji} （又叫 free parameters）来实现。

Hebb’s Rule

Hebb's rule 的主要思想是 ***use-dependent modification***，即每次信息通过 connection（使用该连接），都会使 connection 产生变化。

简单来说，Hebb rule 是在下一个时刻增加 connection 的权重，方式如下：

$$w_{ji}^{k+1} = w_{ji}^{k} + \Delta w_{ji}^{k}$$

where

$$\Delta w_{ji}^{k} = C a_i^k x_j^k$$

其中，

$\mathbf{w_{ji}^{k}}$ 代表在当前时刻 k 时，编号 ji 的 connection 的权重；

$\mathbf{w_{ji}^{k+1}}$ 代表在下一时刻 k+1 时，编号 ji 的 connection 的权重；

$\mathbf{\Delta w_{ji}^{k}}$ 是增大 connection 权重的增量；

C 是影响增量 Δw_{ji}^{k} 大小的正系数，又叫 learning rate；

$\mathbf{a_i^k}$ 是 k 时刻前一个突触神经发送的信息，即当前神经元的输入；

$\mathbf{x_j^k}$ 是 k 时刻当前神经元的输出。

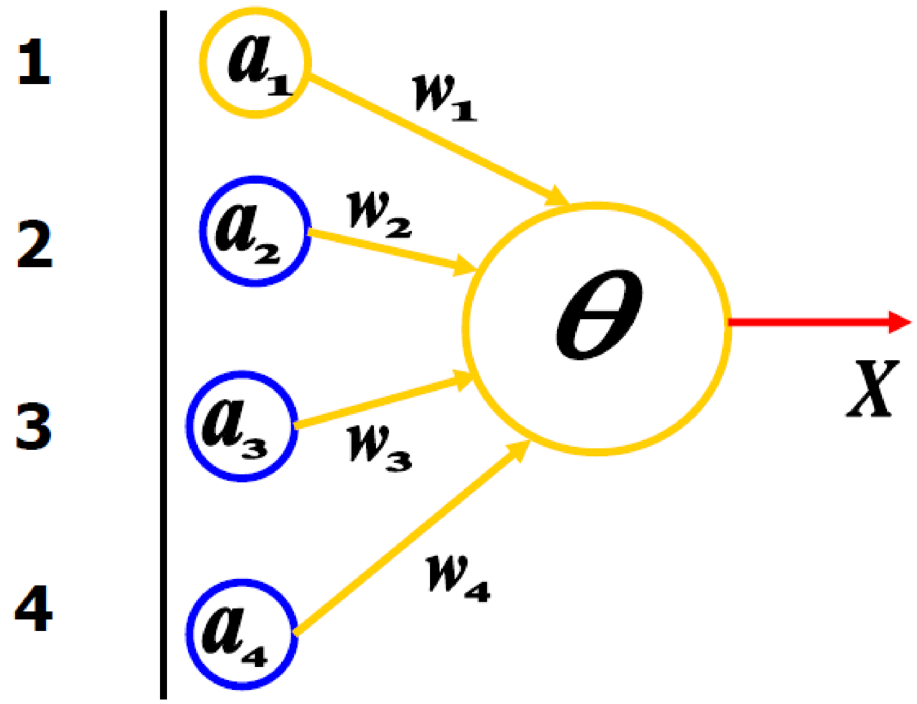
因此，在 Hebb's rule 中，下一时刻连接权重的变化与当前时刻的 **输入** 和 **结果** 有关。只有当两者都不为 0 时，下一时刻的连接权重才会变化。

这个公式强调了 Hebb 突触（Hebbian synapse）的 **相关性（correlation nature）**。

此外，由于 Hebb's rule 只能增加连接的权重，因此它会出现连接进入饱和状态，从而消除了网络中不同神经元的独特性能。

Hebb’s rule in practice.

给出一个 Hebb 网络模型如下：



设 learning rate C = 1，excitation threshold θ = 2，以及 k = 0, 1, 2 这 3 个时刻对该神经元的输入，和 k = 0 时各个权重的初始值（如下表）：

k = 0 时	w_1^0	w_2^0	w_3^0	w_4^0
	1	1	1	1

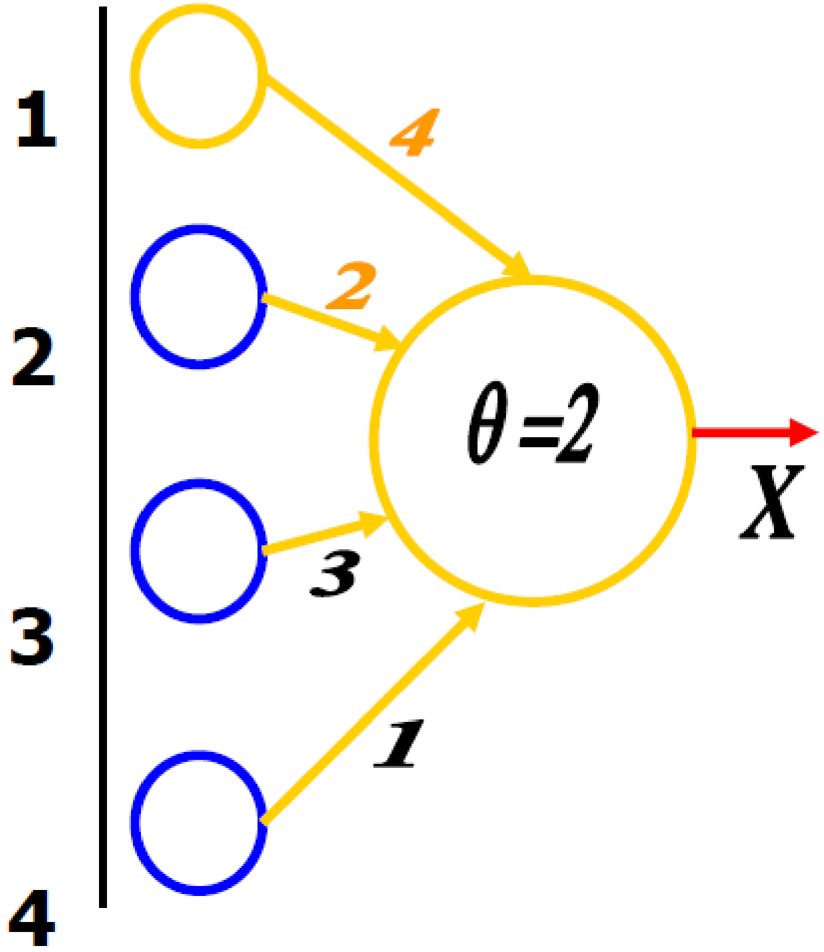
k = 0 时	a_1^0	a_2^0	a_3^0	a_4^0
	1	0	1	0

k = 1 时	a_1^1	a_2^1	a_3^1	a_4^1
	1	0	1	0

k = 2 时	a_1^2	a_2^2	a_3^2	a_4^2
	1	1	0	0

那么在 k = 3 的时候，该神经元的各个权重为多少?

答案如下:


1
2
3
4
4
2
3
1
θ=2
X