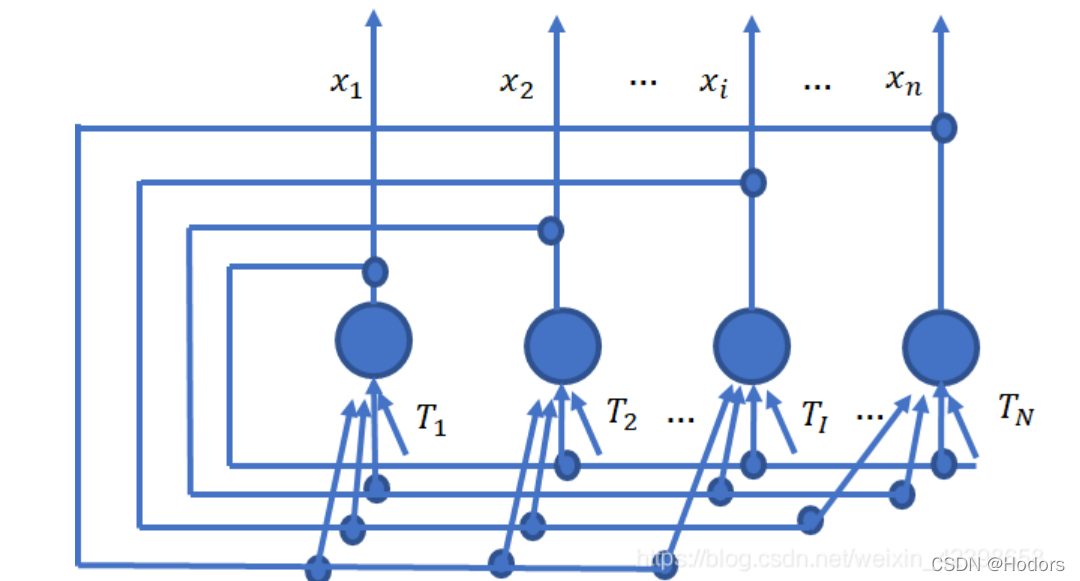
# RNN

## Hopfield网络

该网络的要点：**每个神经元除了接收输入向量中对一个的信号外，还接收除了自身之外其他所有神经元的输出作为自己的输入。**



神经元i的状态由以下关系决定：

W\_ij是神经元i和神经元j之间的连接权值，一般要求其权值对称，且对自身权值为0。对该系统的动态迭代相当于最小化李雅普诺夫函数。

对于自然神经系统神经元的研究发现，如果两个相连的神经元总是同时被激活或者抑制，他们之间的突触连接强度倾向于增强。

## 循环神经网络

### RNN的网络结构

将Hopfield网络的输入输出关系重写为下列格式：

最基本的RNN运算函数是非线性挤压的加权求和函数，比如：

这一函数中的加号可以看做是“串接”，即concat。之所以叫做挤压函数，是因为其将向量挤压到-1到1之间，当然也可以使用其他非线性挤压函数。最后得到的输出将为h\_t再经过一个权值矩阵从而变换到输出空间。

### RNN的训练

和多层感知机以及卷积神经网络是一致的，将误差梯度通过链式法则沿着反方向进行传播，并在每一个节点上计算误差对于权值的偏导数。在正向传播的过程中，每个时刻的权值矩阵是相同的。在反向传播的过程中，则是**在每个时刻的状态和输入下求误差对于权值的梯度**，而权值的更新是把所有时刻的梯度求和，**用梯度之和更新权值**。

在这一过程中会遇到与非常深层的神经网络类似的问题，即如果时间序列过长则**会产生梯度爆炸或者梯度消失**问题。

## LSTM长短时记忆模型

在RNN中，神经元h\_t承担着“传承记忆”的作用，这是一种短时记忆。如果我们把RNN中的循环迭代时间加长，让网络当前状态对后面很多时刻都有影响，则是一种长时记忆。但是这可能**导致网络需要传承大量并不重要的记忆**，使得网络学习负担过重。

### LSTM和RNN的对比

除了h\_t作为神经元之间的隐状态向量，LSTM还具有一套记忆状态向量C\_t。h\_t经过遗忘/记忆，和C\_t一起产生新的记忆状态向量C\_t+1。同时其也保留了RNN本身h\_t传递过程，不同的点在于h\_t经过权值矩阵后会再和C\_t+1运算，然后才产生h\_t+1。

