神经网络与记忆传递

物理科学与技术学院
2013301020115
李海波

摘要

本文依据物理学中的能量最低原理和计算机中的蒙特卡罗模拟，建立了神经网络中的记忆传递模型，并利用程序模拟了记忆的储存与回放过程，研究了该模型下记忆的稳健性以及记忆稳健性与记忆容量的关系，得出了记忆具有较强的稳健性，以及记忆稳健性与记忆容量成反相关关系的结论。

本文第一部分简单的介绍了一下大脑中的信号传递以及信息的记忆过程，将神经元的状态考虑成二值模型，将回忆的过程考虑成能量降低的过程，并依据能量最低原理和蒙特卡洛模拟建立了一个简单的记忆模型。

本文第二部分，利用N*N网格模拟神经网络，利用格点模拟神经元，将格点的坐标集合成矩阵，并利用第一部分给出的公式计算其相关关系，模拟网格中字母型图案的记忆与回忆过程，研究了该模型下记忆的稳健性以及稳健性与记忆容量的关系。

1 记忆的储存与回放模型

1.1 大脑中信号的传递

大脑中存在着数量达到10^{12}之高的神经元，神经元之间相互连接，互相影响，关系极为密切。由于生物化学作用，相互连接的神经元之间存在着微弱的电流，起着传递信号的作用。

神经元之间信号的传递，不管在形式上还是数量上，都极为复杂。现考虑最基本的模型，如图1。

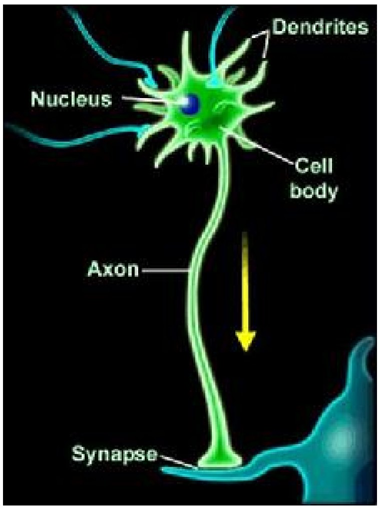


图 1: 信号传递的简单模型

简化模型，信号从轴突(axon)传出，通过连接两神经元的突触(synapse),传递到下一个神经元的树突(dendrite)，从而完成传递。

每个神经元，都有着大量的神经元与之相连，数量可达10^4。也就是说，每个神经元，都同时有大量的信号输入，所有的信号叠加起来，会得出一个电压差，普遍认为，信号是通过这个电压差来解析的。对信号进行解析的函数称为激励函数(firing rate), 可以写成$R = f(\sum V_i)$。

实验表明，该函数并非线性，而是拥有如图2所示的形式：

为了简单起见，本文将其考虑为符号函数$sign(x)$的形式。即二值化，当信号大于某个值时考虑成高电平，对应神经元的兴奋; 低于某个值时考虑为低电平，对应神经元的抑制。

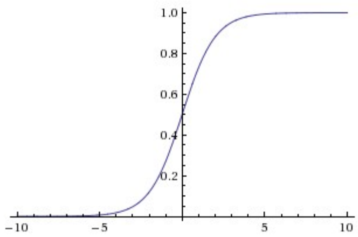


图 2: firing rate

如此，通过信号的传递，神经元的兴奋或者抑制状态不断变化，对应二进制，构成信息的表达。

神经元的兴奋与抑制，构成信息，保留下来则成了记忆。那么记忆是如何储存与回放的呢？

1.2 记忆的储存与回放

神经元的状态可以构成信息，形成想法，但大脑间信号的传递万千，人的思绪也是瞬息万变，神经元之间的状态不能保持一直不变，因而科学家认为记忆并不是由神经元的状态构成的。

第一部分已经提及，每个神经元都有数以万计的神经元与之相连，关系非常紧密，普遍认为记忆是储存在神经元之间的联系之中。每当需要回想某个信息的时候，对应的神经元开始传递号，通过神经元之间联系的不同，即激励函数的不同，造成不同的神经元的兴奋或者抑制，从而回想起某个信息。

如第一部分所示，为了简化模型，神经元的状态二值化，只考虑两种状态，$s = \pm 1$。

神经元之间的联系，可以记为$J_{i,j}$，表示第i个神经元与第j个神经元的联系。

如同物理学中的最低能量原理一样，记忆也遵循相似的原理，记某个神经网络的总能量为：

$$E = -\sum_{i,j} J_{i,j} * s_i * s_j \tag{1}$$

其中s_i和s_j为此时神经网络中对应的神经元的状态。大脑将会根据能量是否下降来决定神经元的状态是否变化，此过程可以考虑为蒙特卡洛过

程，即随机选一个神经元，改变状态，观察总能量的变化。如果总能量降低，则认可此变化，将该状态保持，否则变为原状态。经过充分多次的尝试之后，能量将会降到最低，即记忆的状态。

此模型下，能否达到理想的记忆状态，关键在于神经元之间的联系，$J_{i,j}$的选取。简单化的经验给出如下的公式：

$$J_{i,j} = \frac{1}{m} * \sum_{m} s_i(m) * s_j(m) \tag{2}$$

其中，m对应第m个记录的信息，即信息m状态下的s_i与s_j。

可以证明，给定一定数量的信息，确定$J_{i,j}$之后，从某个随机的初始状态出发，最终可以达到原来的某个信息状态，即形成了记忆。

此模型中，记忆在回放时，脑海中首先形成一个模糊的印象（忽略形成的过程），即对应部分的神经元，形成一个模糊的状态，与理想状态有一定程度的相似。在形成该状态后，大脑利用最低能量原理，逐渐改变神经元的状态，观察能量是否降低，决定是否改变，最终达到能量最低的状态，达到回忆的效果。

2 记忆储存与回放的程序模拟

2.1 初步模拟

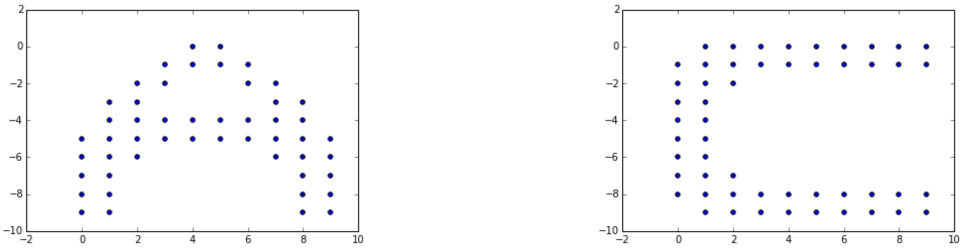
考虑一个10*10的网络，每个网格对应一个神经元，每个网格都与网络内的所有网格有联系，对应神经元的联系。每个网格点拥有两种状态，1或者-1。为了可视化，我们将利用此网格来记录字母A与C，表现如下：

首先利用公式(2)，算出每一个$J_{i,j}$，得出记忆的储存；在回忆模型的蒙特卡罗过程中，给出一个状态类似某个记忆的网格，随机抽取一个神经元，利用公式(1)，计算该神经元状态改变后的能量变化，如果：

1 $\delta E < 0$，则改变其状态

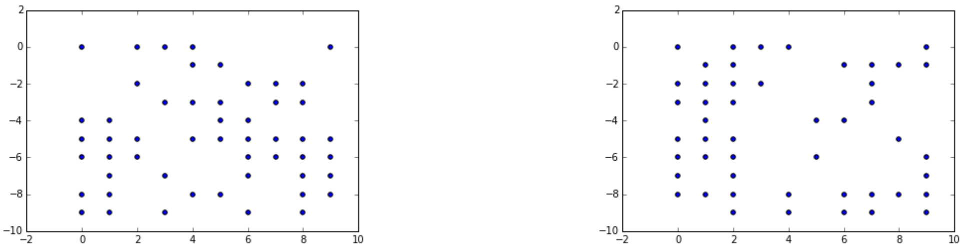
2 $\delta E \geq 0$，则保留原来的状态

将此过程进行网格总数的3倍次数后，得出最终的状态，再与理想的记忆对比。下面给出一些例子，如图4，随机选取30%A和C理想网格中的格



(a) A的网格 (b) C的网格

图 3: A与C的网格表达



(a) A的30%错误网格 (b) C的30%错误网格

图 4: A与C的30%错误网格

子，改变其状态，再进行模拟。这样的网格，模拟之后，都变成了对应的如图3的理想的网格，说明这样的记忆储存方式可行。

2.2 记忆受损的模拟

上述是完整的记忆状态，神经元连接完好，假如连接一定程度地受损，是否还能给出理想的回忆呢？下面给出80%受损的记忆连接（随机选80%的$J_{i,j}$取值为0），以及50%错误的初状态，观察其模拟之后的状态。

图5中可以看出，即使有80%的$J_{i,j}$受损，依然可以基本正确的回忆起信息。

此情形下，当受损的程度达到90%以后，记忆开始难以复原，也就是回忆不起来了。

事实上，记忆的稳健性，与记忆容量有很大关系。

在10*10的网格中，这里只记录了2个字母。按《Computational Physics》中的说法，N*N网格大约只能记$0.13 * N^2$个大图案，否则即使记忆连接完

(a) C的50%错误网格 (b) C模拟后的网格

图 5: 80%神经连接受损的模拟

(a) A的50%错误网络 (b) 模拟后的网络

图 6: 25%容量下的50%初印象和60%受损

整，记忆也可能出错，即记忆有一定的容量。

在达到最大容量前，记忆得越少，则记忆越稳健，即可接受的神经元连接的受损程度越大，可接受的初印象的模糊程度也越大。

当容量大约达到理论的25%时，10*10网络下即记忆3个字母A、C、T时，50%的初印象和60%的记忆连接受损，就已经很容易混淆了。如图6所示，A错认成了C。当初印象改为30%的时候，则不容易出错，仍然可以正确地回忆。

当然，这与模型的过于简单也有一定的关系。受笔者水平所限，无法对此进行进一步探究。

可以得出的结论是：

1 记忆越少，则可接受更模糊的初始状态，即更容易回忆

2 记忆越少，则可接受的神经连接受损越大，即更不容易忘记

上述两条结论都与日常经验相符合，说明模型虽然简单，但仍具有一

定的实用性。

3 参考文献

Computational Physics, Nicholas J. Giordano, Hisao Nakanishi, Tsinghua University Press