本周学习:循环神经网络(Recurrent Neural Network)

在之前的全连接神经网络和卷积神经网络都是单独处理输入，前一个输入和后一个输出没有关系。但是在某些任务中，我们需要处理序列的信息，前面的输入和后面的输出是有关系的。

那么我们就需要RNN来解决这个问题。

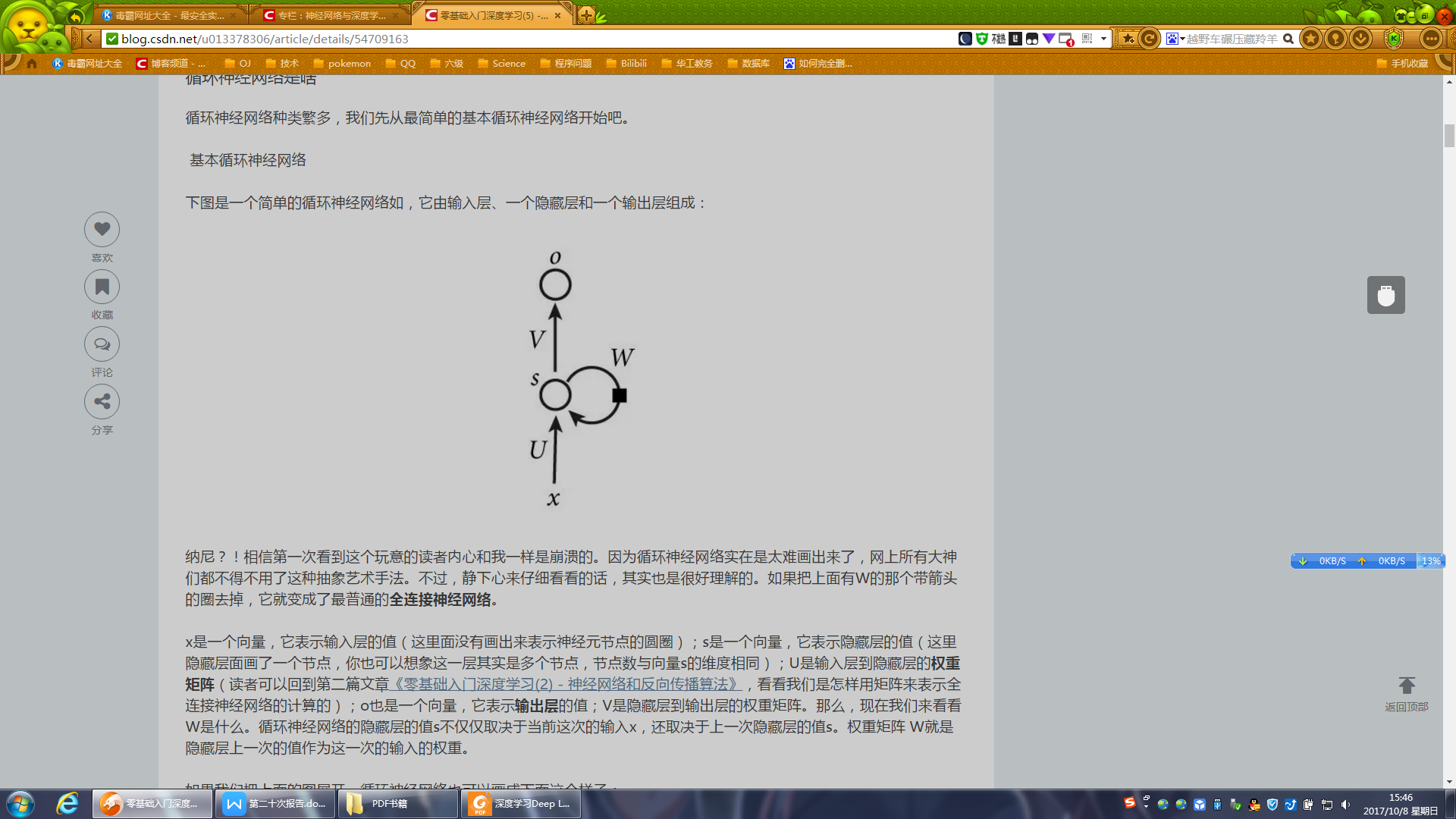
语言模型

RNN最早在自然语言处理（NLP）方面被应用，语言模型就是给定一个句子的前部分，预测下一部分是什么。

循环神经网络（RNN）

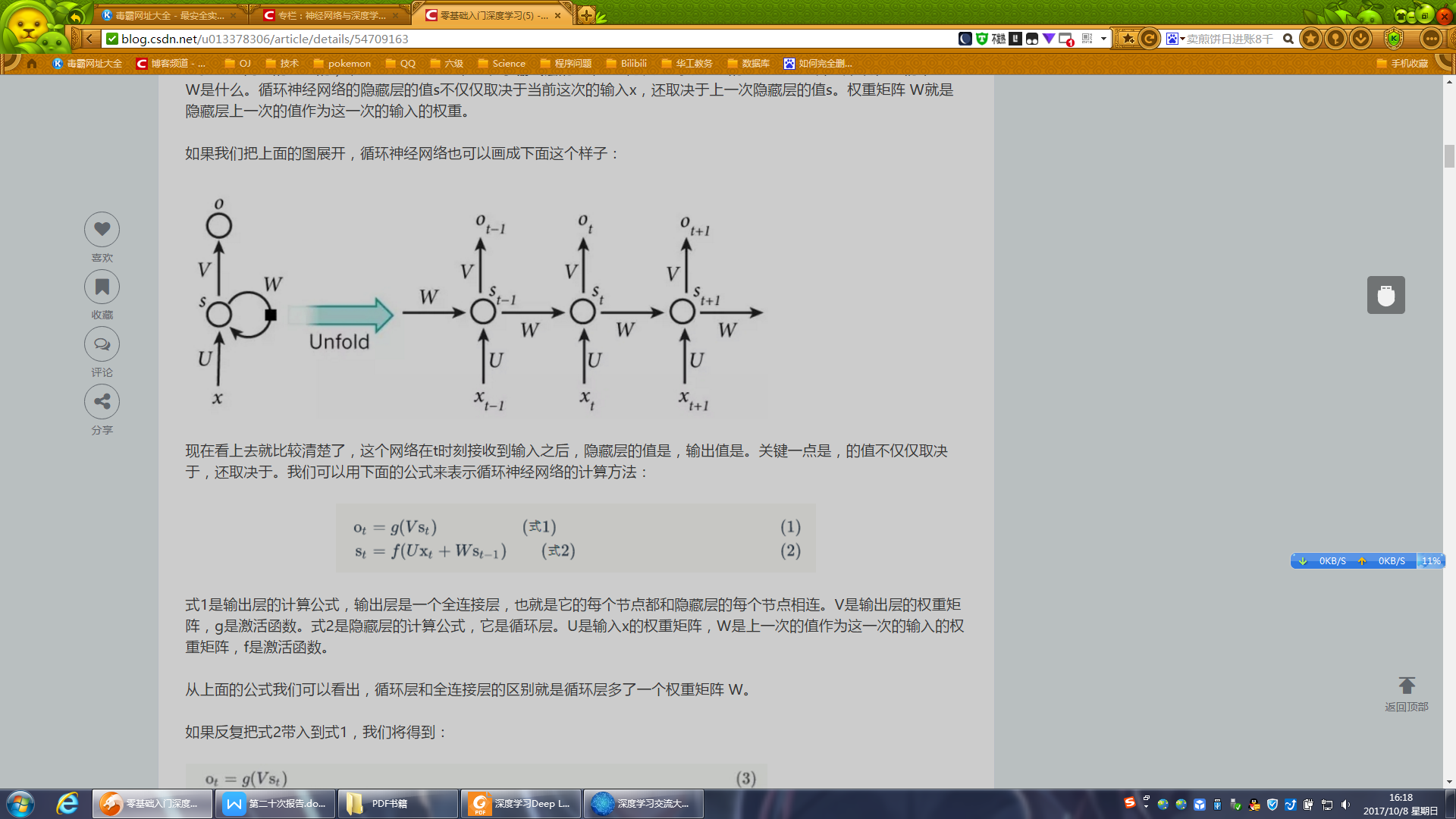
基本的循环神经网络

构成：输入层、隐藏层、输出层

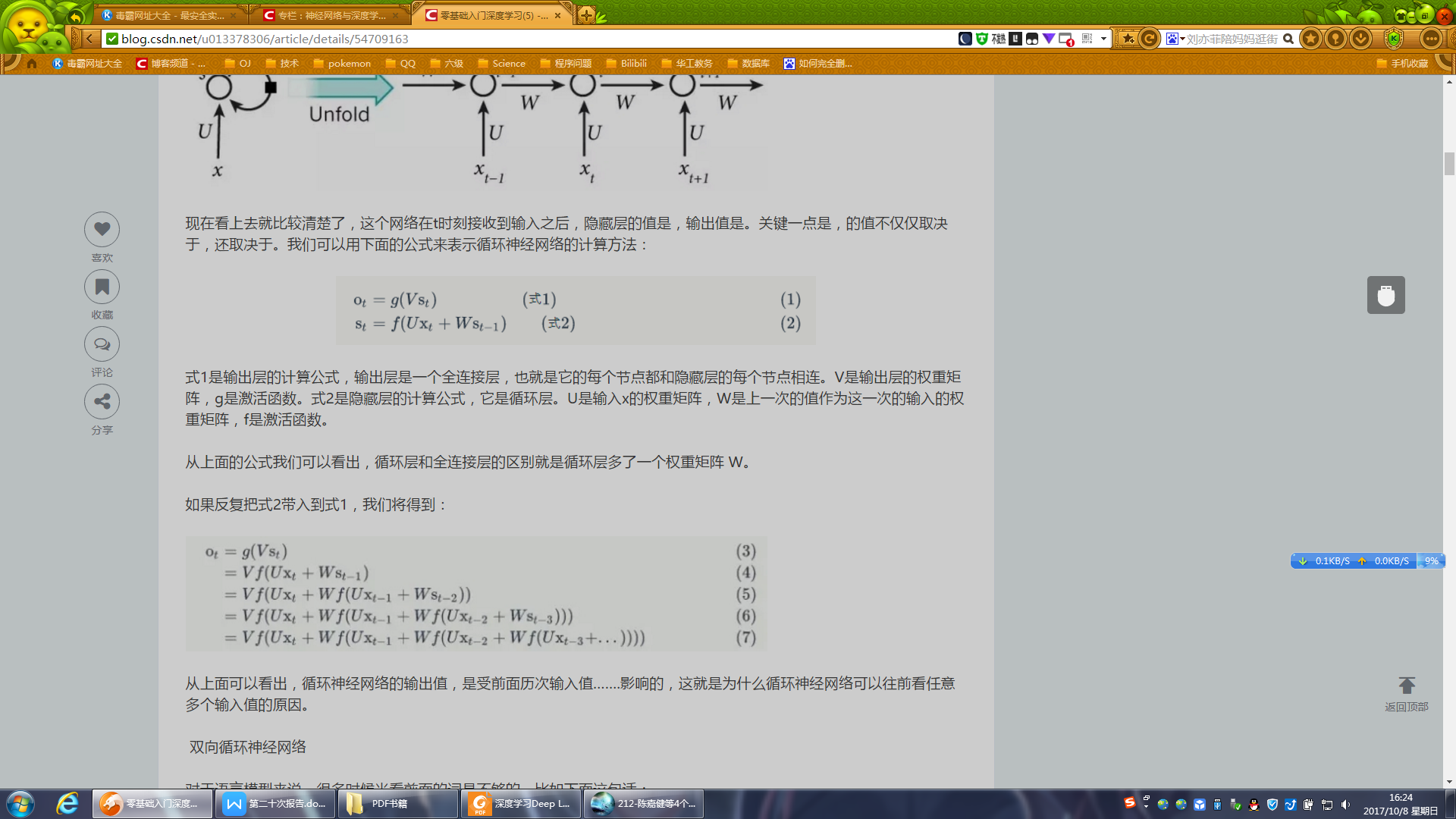


x是一个向量，它表示输入层的值；s是一个向量，它表示隐藏层的值；U是输入层到隐藏层的权重矩阵；o也是一个向量，它表示输出层的值；V是隐藏层到输出层的权重矩阵。那么，现在我们来看看W是什么。循环神经网络的隐藏层的值s不仅仅取决于当前这次的输入x，还取决于上一次隐藏层的值s。权重矩阵 W就是隐藏层上一次的值作为这一次的输入的权重。

展开的话可以描述为以下的形状：

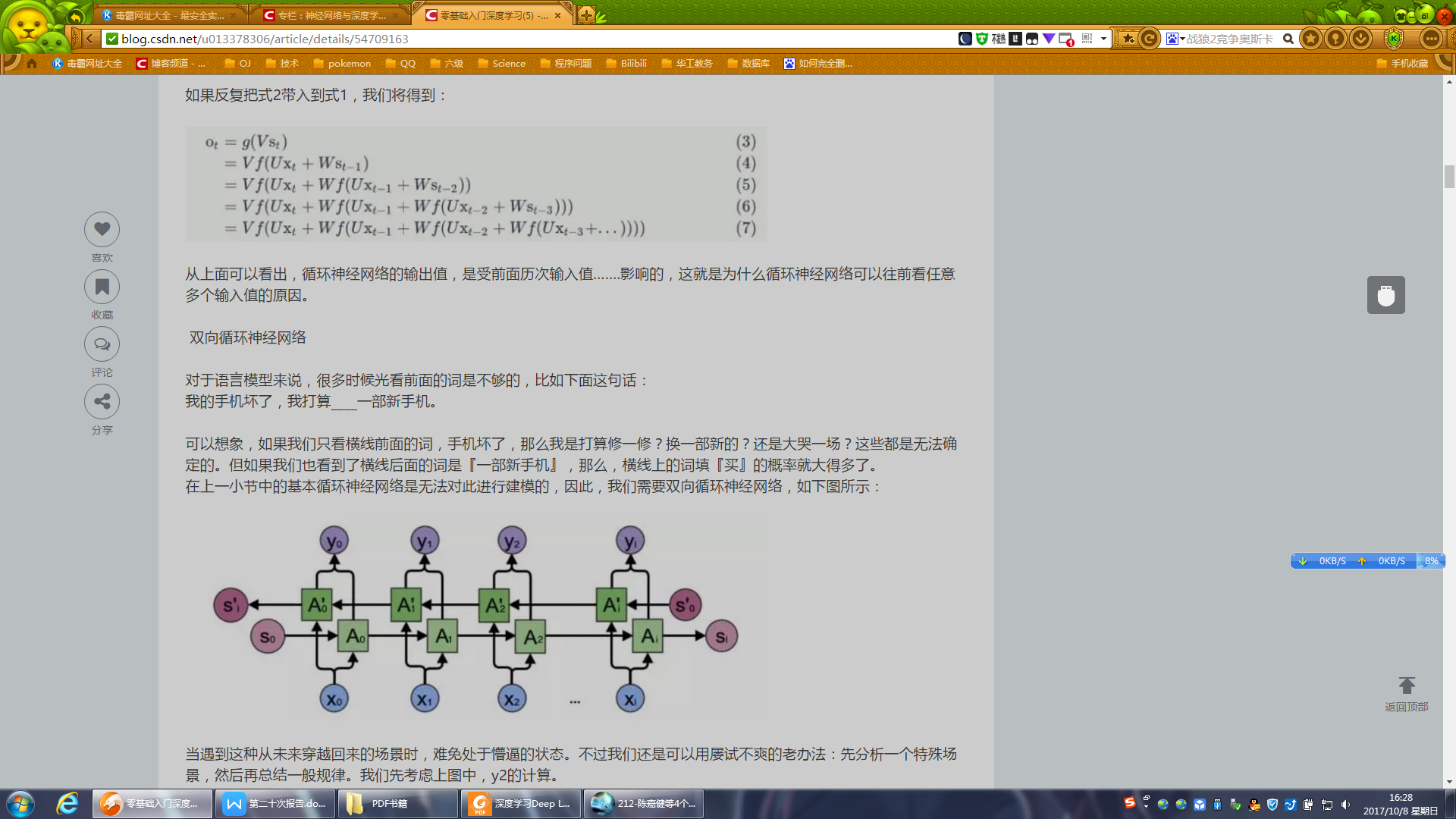


输出层和隐藏层的计算公式如下



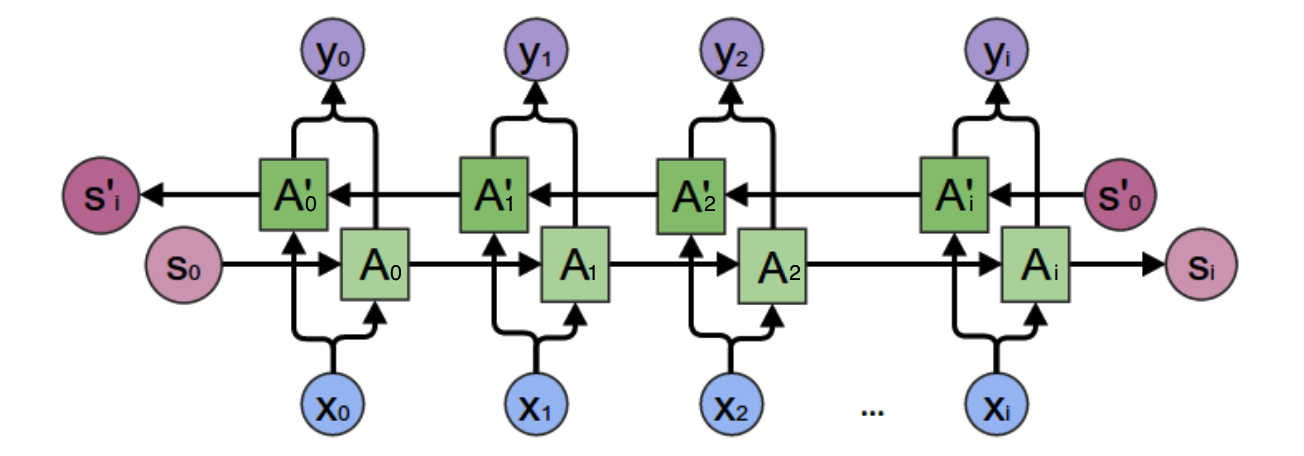
重点：循环层和全连接层的区别就是循环层多了一个权重矩阵 W

反复迭代公式可以得到：



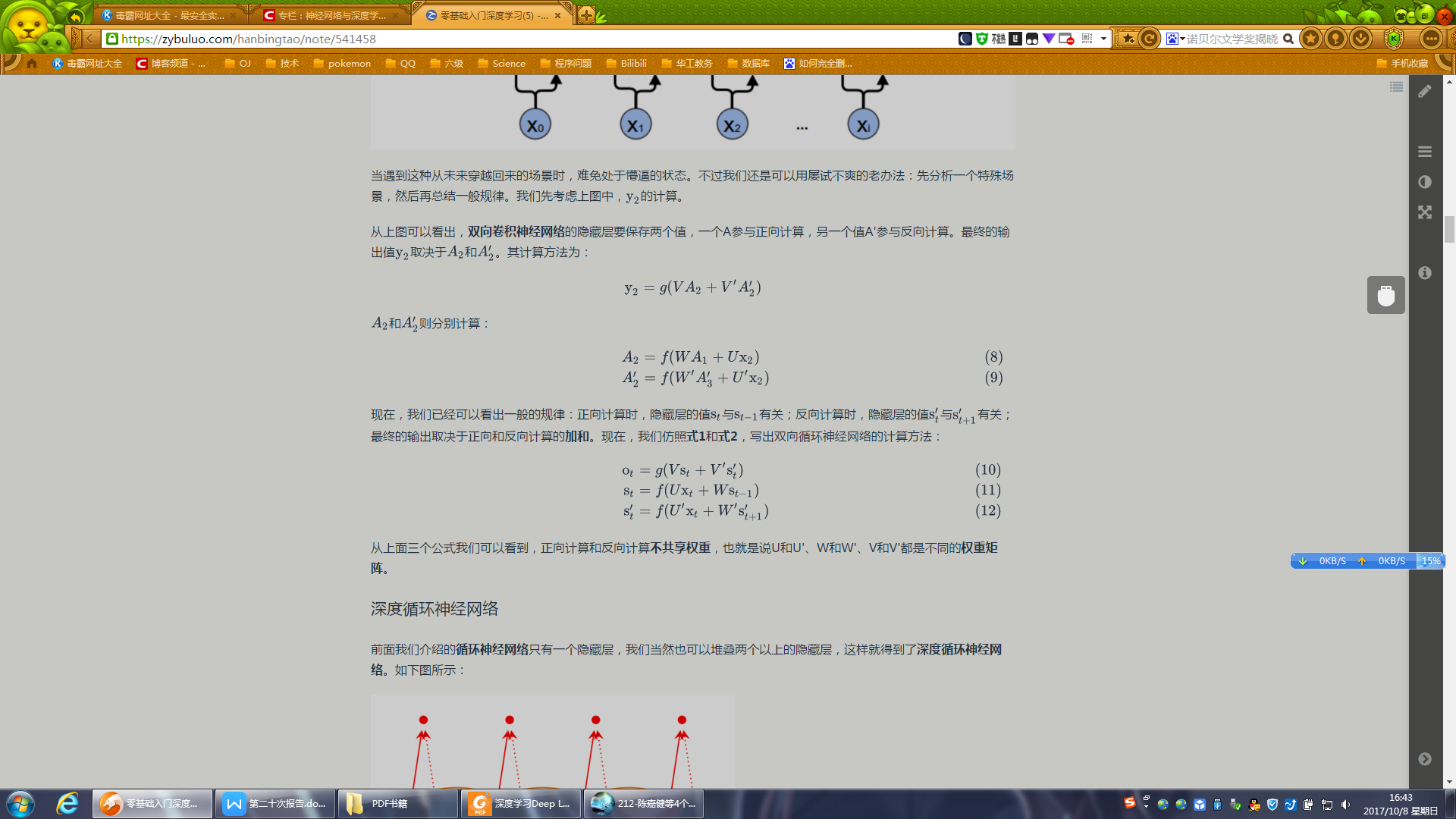
问题出现了：只看前面的信息就可以吗？当然不是，还要考虑后面的信息

双向循环神经网络



先看单独情况如何计算

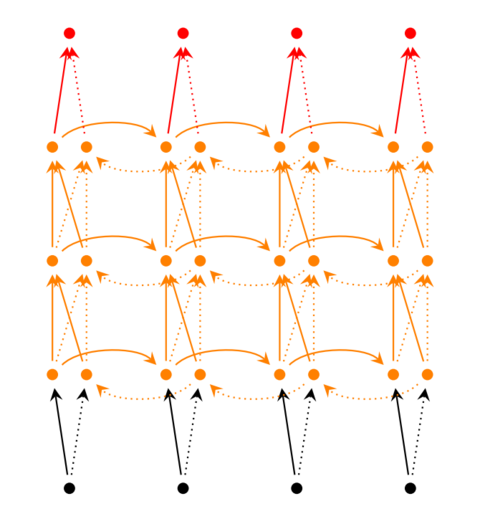
隐藏层要保存两个值，一个A参与正向计算，另一个值A'参与反向计算。最终的输出值y2取决于A2和A2’。其计算方法为：

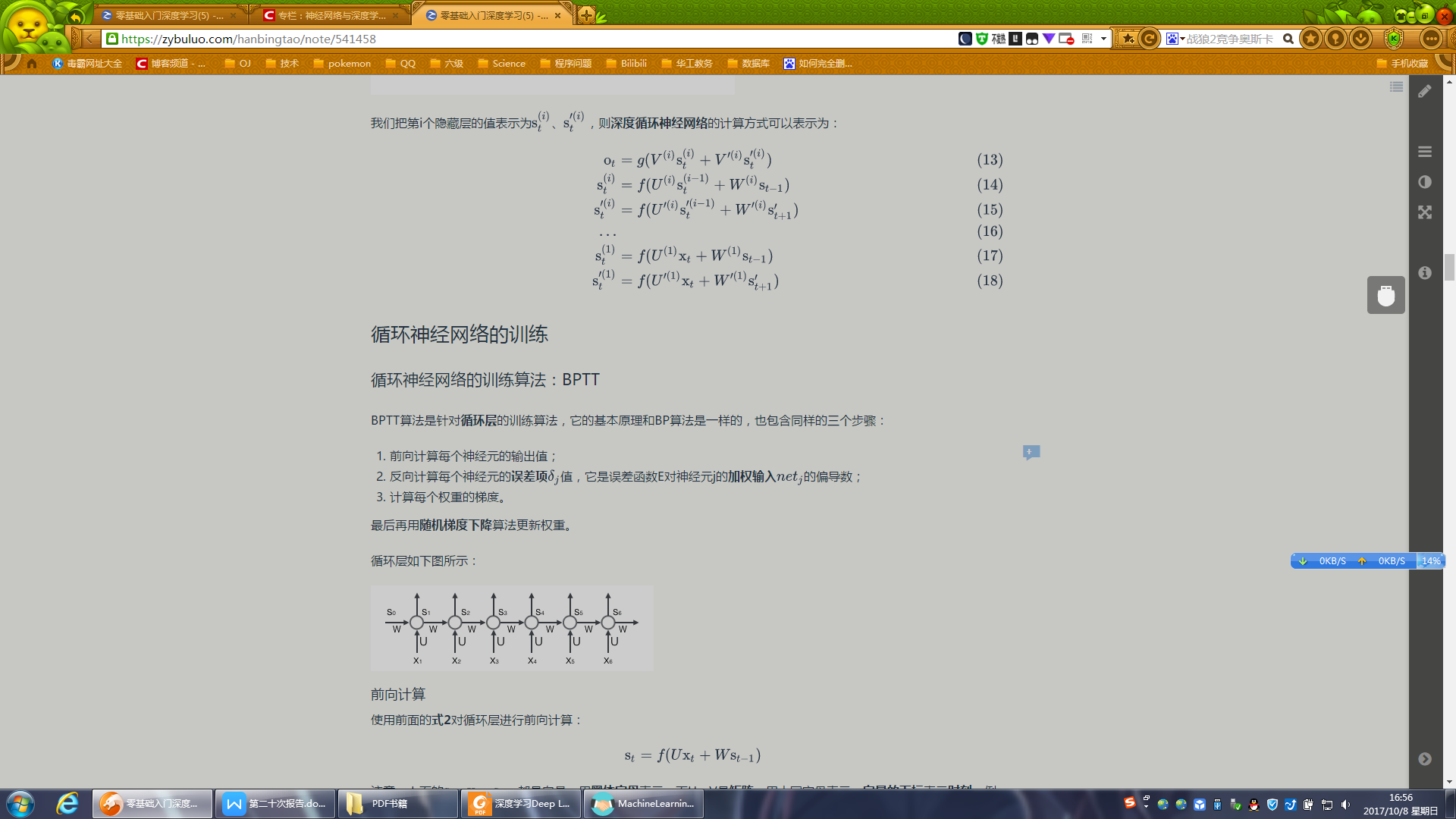


正向计算和反向计算不共享权重，U和U’，W和W’是不同的矩阵

深度循环神经网络

堆叠两个以上的隐藏层，这样就得到了深度循环神经网络。如下图所示：





参考资料：

[1].博客https://zybuluo.com/hanbingtao/note/541458

[2].博客http://blog.csdn.net/column/details/16408.html

[3].深度学习Deep Learning 中文版