本周学习：尝试理解LSTM网络（其一）

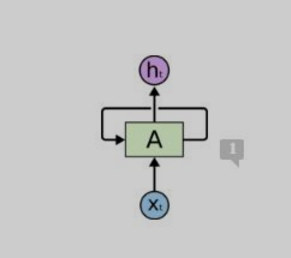
由Recurrent Neural Network（循环神经网络，RNN）演化而来

Long Short Term Memory Networks（长短时记忆网络，LSTM）与传统的前向神经网络不同，LSTM可以对之前的输入有选择地记忆，从而判断当前的输入，LSTM的这一个特点在处理有关时序的输入时就会显得有非常大的优势。

先看一看RNN，基于人的思维不会总是每时每刻都从当前的开始，特别是在阅读时，人类对词语的理解，和之前的阅读积累有关，换句话说，人类的思考具持续性。

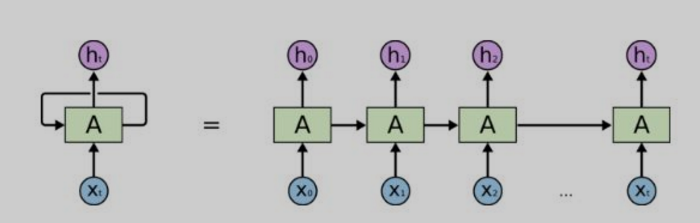
在解决时序相关的输入信号时，传统的前馈神经神经网络就无能为力，因为前向神经网络输出只和当前的输入有关，无法利用之前的信息。RNN的提出就很好地解决了这个问题。与传统的前馈神经网络不同，RNN是一个循环神经网络，可以很好地储存之前的信息。

最简单的RNN网络结构



其中，A表示神经网络的一块，Xt表示当前的输入，ht表示当前的输出，这个自循环的箭头表示输入信息可以从当前时刻传递给下一个时刻。

下面对上面的网络结构展开，以更加清晰地了解RNN

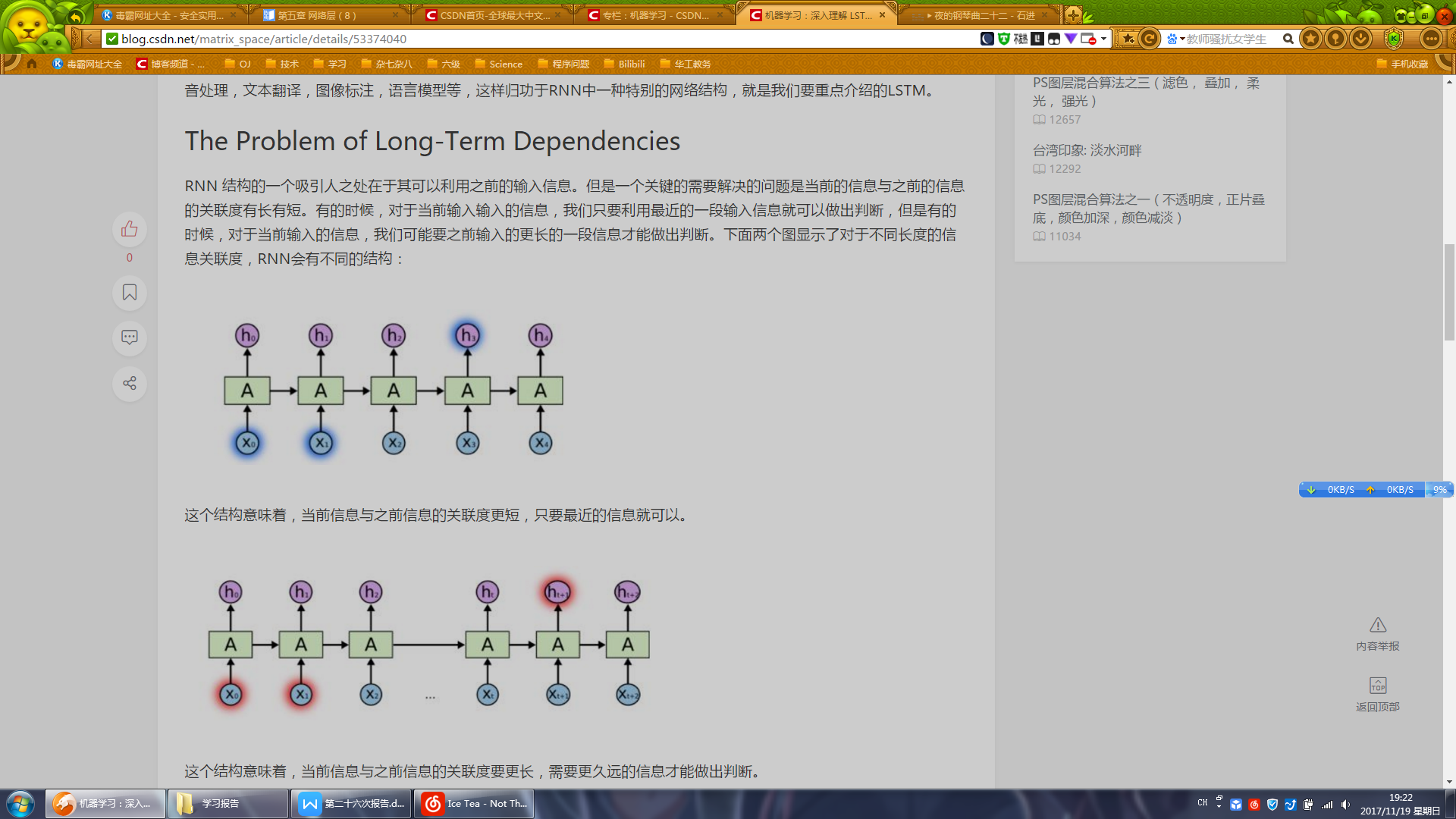


其实，RNN网络可以看做是同一个神经网络的多重复制，每一次都是把信息传递给下一个。形状上开来是一种链式结构，说明了RNN是与时序信号紧密相连的。因此，RNN看起来是对于处理时序信号是非常有效的。RNN已经成功地运用于语言处理、文本翻译、、图像标注、语言模型，这些都要归功于LSTM。

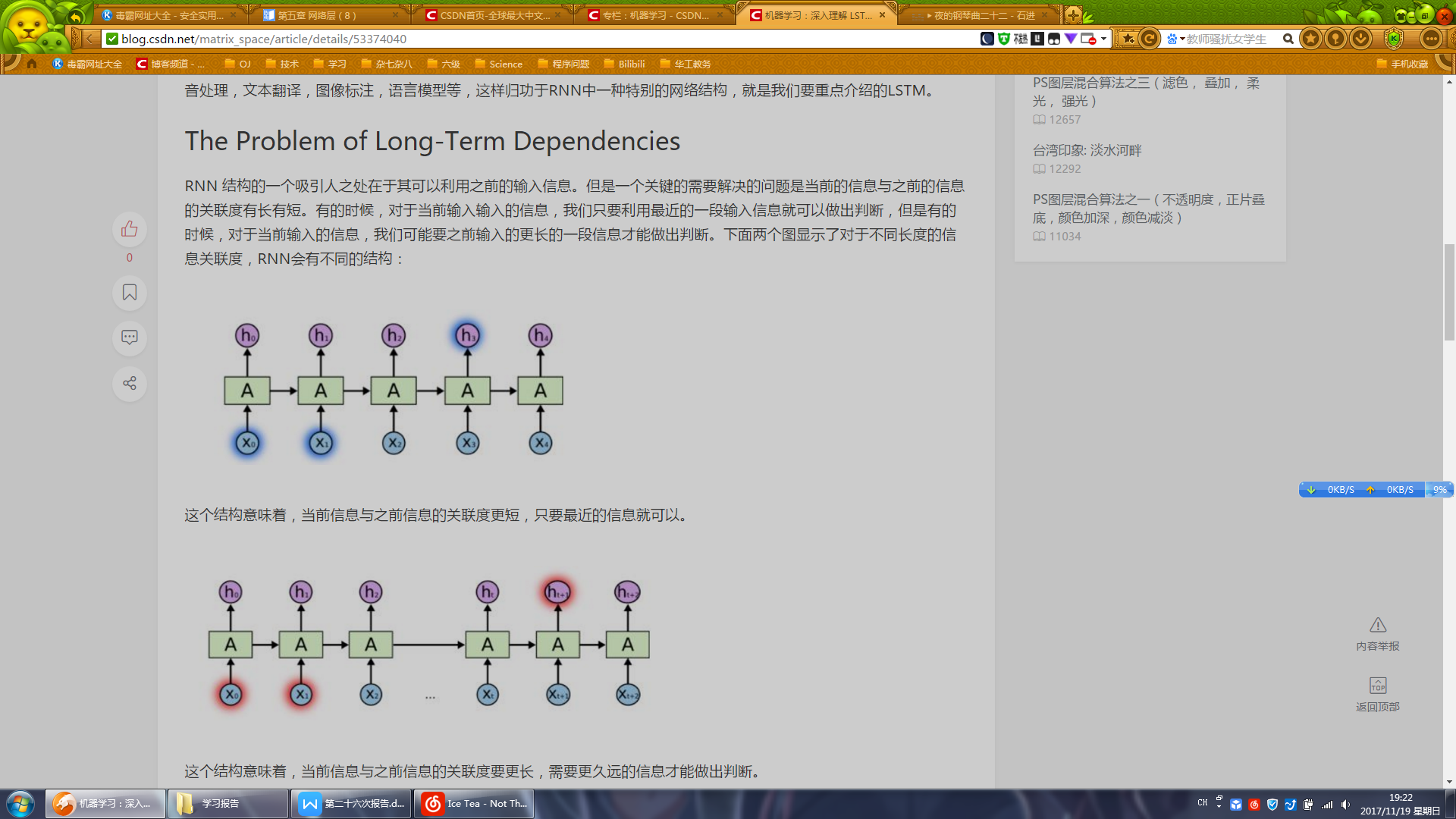
The Problem of Long-Term Dependencies（长期依赖问题）

RNN的特点是它可以利用之前输入的信息。但是一个关键的需要解决的问题是当前的信息与之前的信息的关联有一定的长度，即是长短不一。有时候，对于当前输入的信息，我们只需要利用最近的一次输入信息就可以判断。但是有时候，对于当前输入的信息，我们可能要依赖之前输入的更长的一段信息才能够判断。

对于不同长度的信息关联度，所构造的RNN的结构也是不一样的，



上图表示当前信息与之前信息的关联更短，只要最近的信息就可以



上图表示当前信息与之前信息的关联度要长，需要时间更加远的信息才能判断

随着，信息关联度越来越长，RNN 将变得无法去学习这些信息之间的联系，从而完全失去作用，为了解决这个问题，所以提出了 LSTM 的结构。

参考资料:

1. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
2. <http://blog.csdn.net/matrix_space/article/details/53374040>