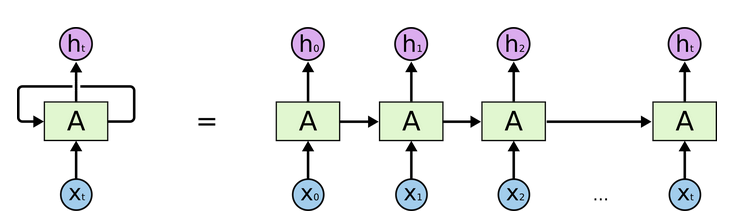
2017/7/20学习报告：

1.文献：基于LSTM型RNN的CAPTCHA识别方法 作者：张　亮　　黄曙光　　石昭祥　　胡荣贵(解放军电子工程学院网络系合肥　230037)；

2.索引：人工智能,脱机文字识别,全自动的区分计算机和人的图灵测试(CAPTCHA),长短时记忆(LSTM)；

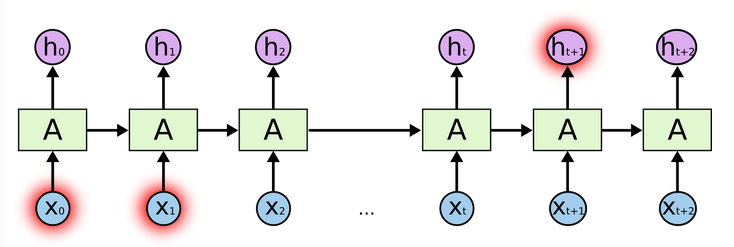
3.重点收获：

1. 了解**图灵测试**：测试者与被测试者（一个人和一台机器）隔开的情况下，通过一些装置（如键盘）向被测试者随意提问。进行多次测试后，如果有超过30%的测试者不能确定出被测试者是人还是机器，那么这台机器就通过了测试，并被认为具有[人类智能](https://baike.baidu.com/item/%E4%BA%BA%E7%B1%BB%E6%99%BA%E8%83%BD" \t "https://baike.baidu.com/item/%E5%9B%BE%E7%81%B5%E6%B5%8B%E8%AF%95/_blank)。
2. 了解RNN：循环（周期）神经网络



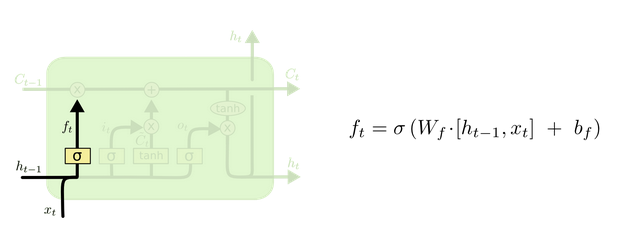
神经网络的模块A输入为Xi，输出为hi。模块A的循环结构使得信息从网络的上一步传到了下一步。

RNN存在问题：长期依赖问题（只输入X0、X1时，要跳过X2~Xt，随着间隔增长，RNNs变得难以学习连接之间的关系了，因此想要只根据X0、X1预测出ht+1的输出结果十分困难）



1. 了解LSTM：长短期记忆网络（一种时间递归神经网络，特殊的RNN），能够学习长时间依赖，从而解决RNN存在的长期依赖问题。因为记忆长期信息是LSTMs的默认行为

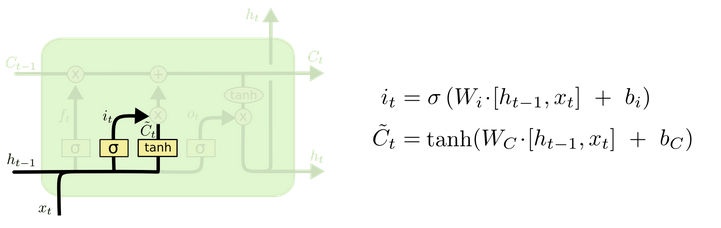
LSTM三种门（每个门都是一个sigmoid层）（以下为经典LSTM模型）：

1. 遗忘门：图示及公式：

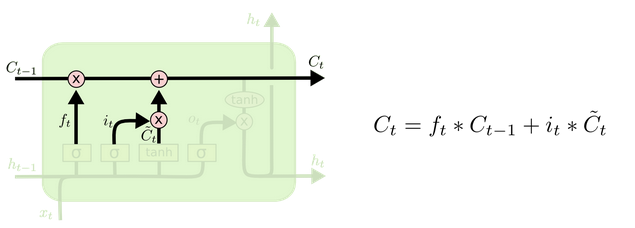
Sigmoid 层的输出在0到1之间，定义了各成分被放行通过的程度。0值意味着『不让任何东西过去』；1值意味着『让所有东西通过』。

1. 输入门：图示及公式：

此步决定细胞中要存储何种信息。它有2个组成部分。首先，由一个叫做『输入门层』的sigmoid层决定我们将要更新哪些值。其次，一个tanh层创建一个新的候选向量Ct，它可以加在状态之中。在下一步我们将结合两者来生成状态的更新。

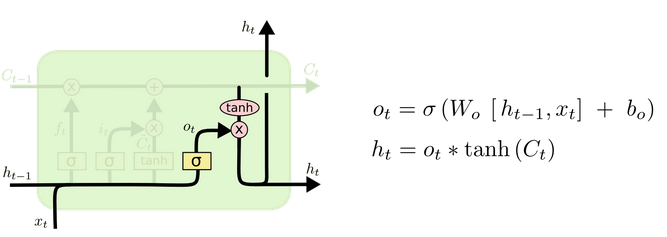


Ct-1--->Ct：



1. 输出门：使用sigmoid层来决定我们要输出细胞状态的哪些部分。

用tanh处理细胞状态Ct（将状态值映射到-1至1之间），将其与sigmoid门的输出值相乘，从而我们能够输出我们决定输出的值。如图所示：



1. 了解CAPTCHA：CAPTCHA属于图灵测试的一种,由机器人来判断测试者是自然人还是机器人 .自然人很容易通过 CAPTCHA,而机器人却很难通过 .

目前 CAPTCHA的实现形式有文字识别型 、邮件验证型 、手机验证型 、声音识别型等多种形式

早期CAPTCHA识别方法：图像分割

难题：粘着严重的字符难以分割，识别失败率很大

新的识别方法：基于LSTM型RNN的CAPTCHA 识别方法

优势：训练和识别时不需要对 CAPTCHA中的字符进行分割 .得到一个 CAPTCHA图片后 , 直接使用**滑动窗口在图片上提取特征值序列** , 然后与图片对应的字符序列构成一个训练样本, 即可使用这个样本对 RNN进行训练 .

1. 了解特征提取：

特征选择和提取决定着系统所能达到的识别精度。手写识别中 , 一般在滑动窗口上提取一组基于矩特征和结构特征形成的混合特征。

本文提取窗口的平均灰度值 、窗口重心 、窗口的二阶矩 、上下黑点位置及其相对变化率 、黑白转换次数 、上下黑点之间的黑像素比率等 9个参数作为 1组特征值 , 将这组特征称为混合特征 .

用混合特征作为输入训练RNN。

1. RNN解码:较为复杂，下周研究。