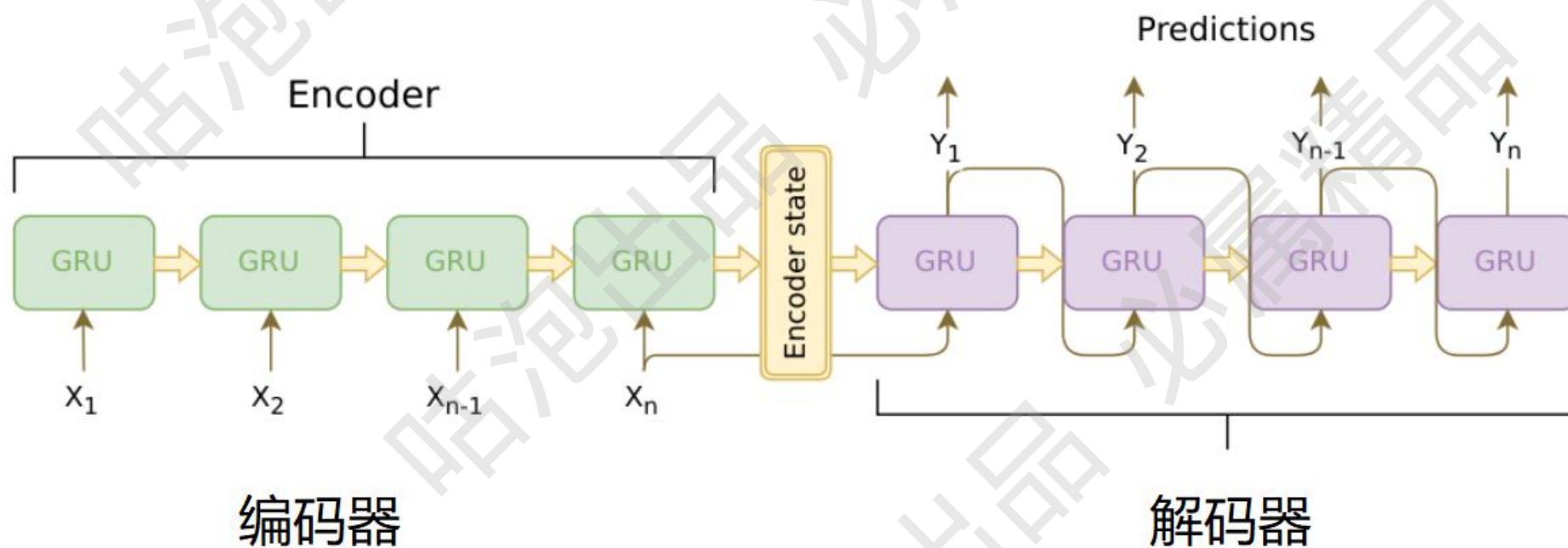


# seq2seq

✓ 模型长什么样子呢?

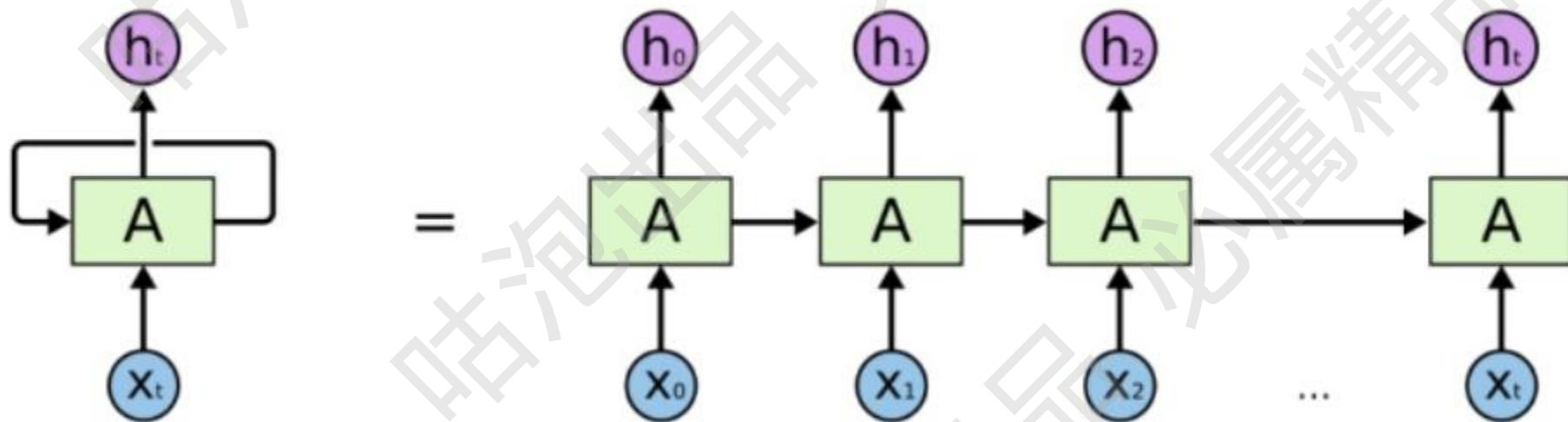
✎ 想想谍战片里，我方情报员与敌方情报员：



# seq2seq

✓ 先来回归下RNN

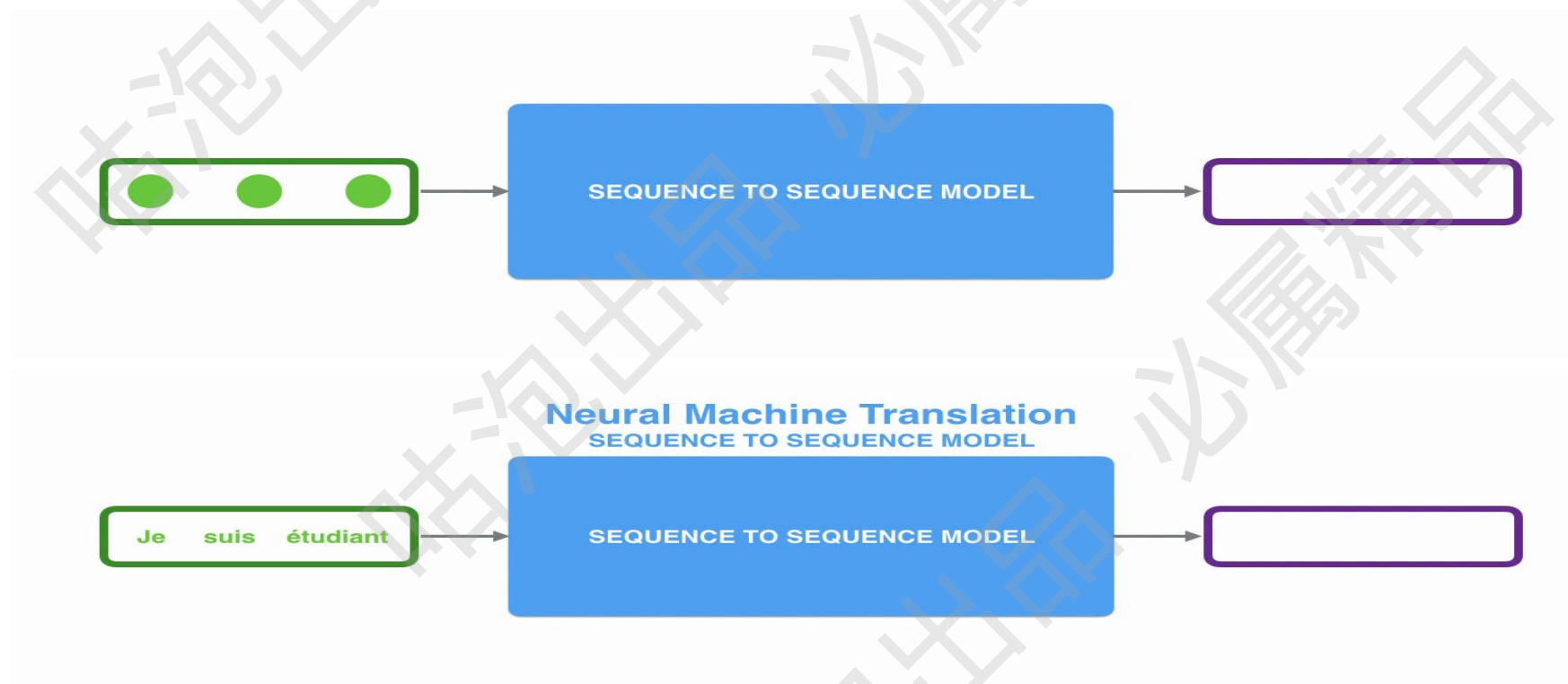
✎ 传统的递归神经网络，但是也有很多缺点！



# seq2seq

✓ 咋工作的呢?

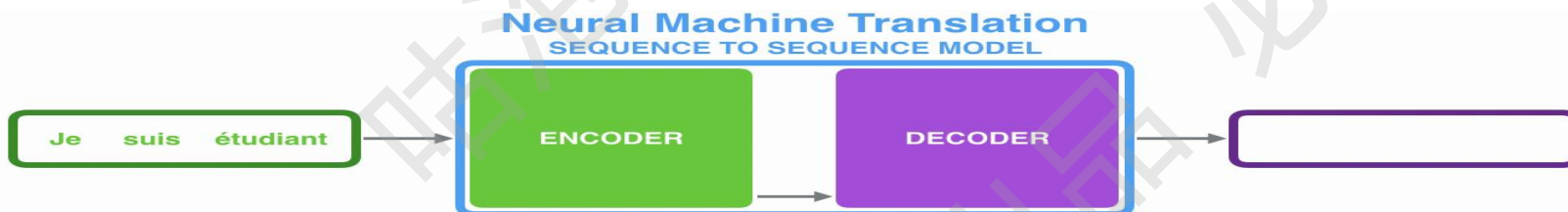
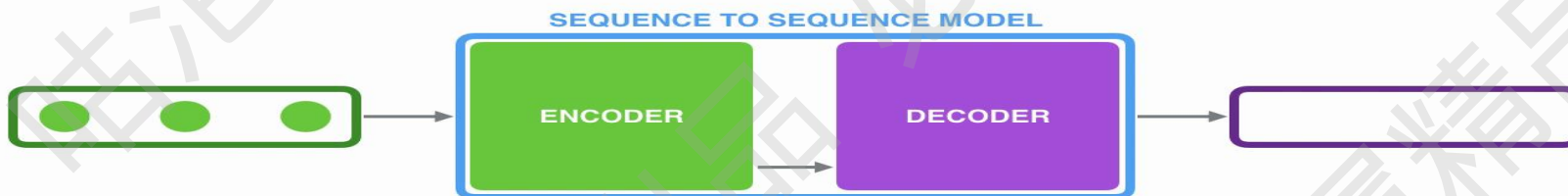
📎 整体模型看起来就像是先把所有输入一个个吃掉，然后再输出结果



# seq2seq

✓ 咋工作的呢?

✎ 这其中需要先用编码器完成所有编码任务，再有解码器输出



# seq2seq

✓ 编码特征长什么样子?

✎ 其实就是数值特征, 例如128, 256维

✎ 结合RNN一起来观察下:

CONTEXT

0.11
0.03
0.81
-0.62

0.11
0.03
0.81
-0.62

Time step:

Neural Machine Translation  
SEQUENCE TO SEQUENCE MODEL

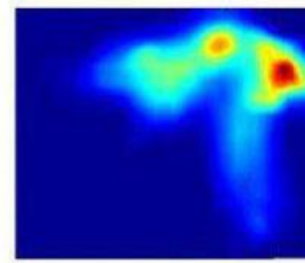
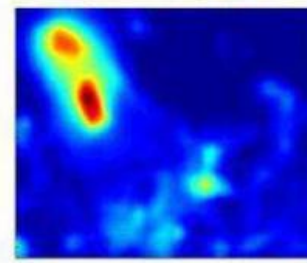
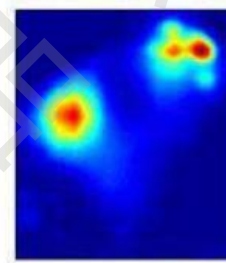
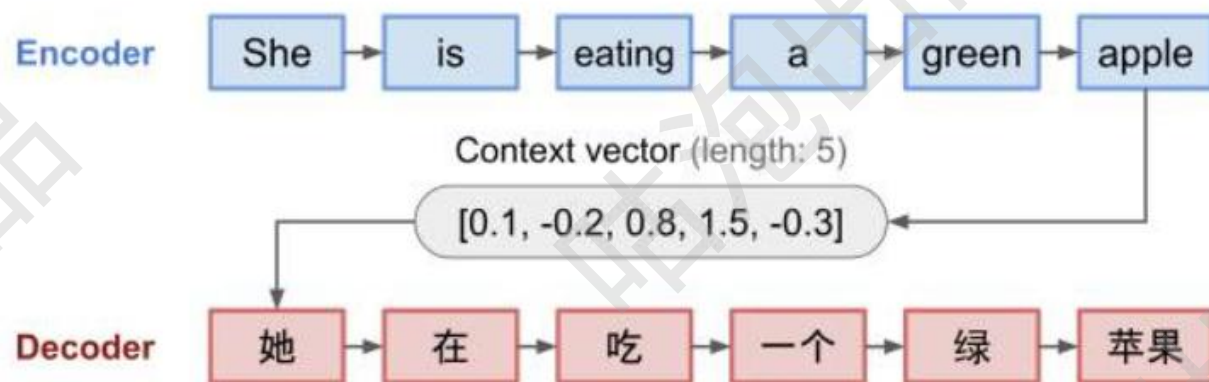


# Attention

✓ Attention是啥意思呢?

✎ 对于输入的数据, 你的关注点是什么?

✎ 如何才能让计算机关注到这些有价值的信息?



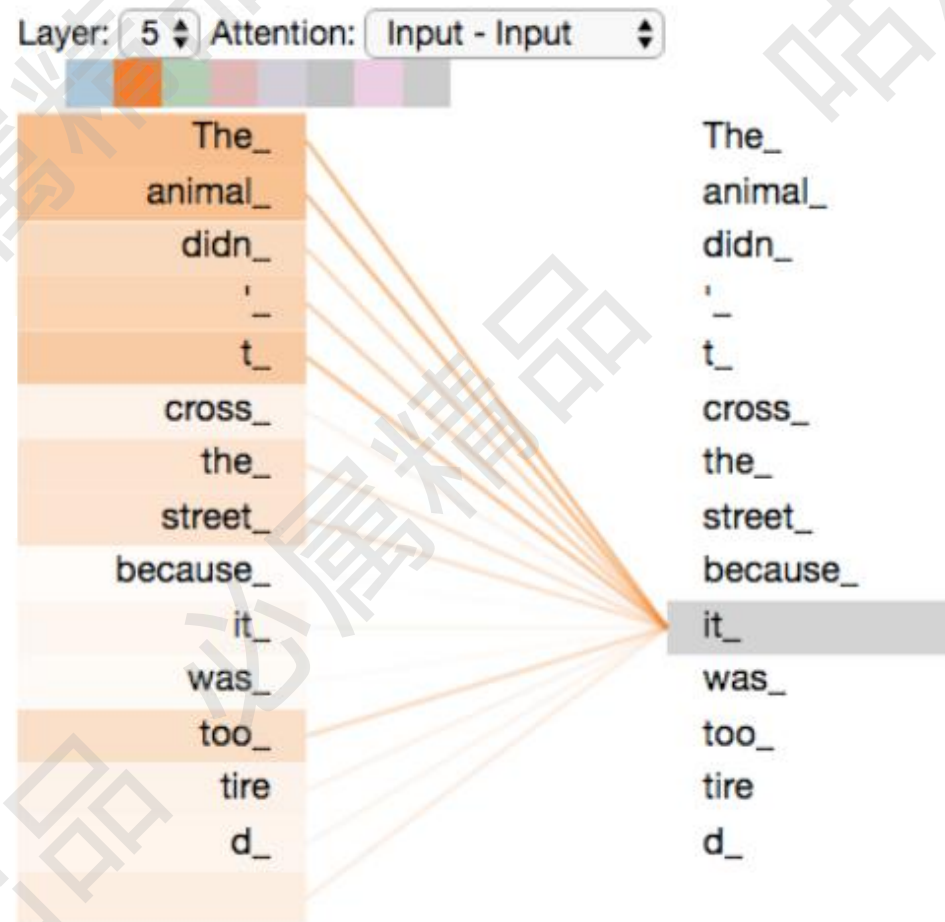


# Attention

✓ attention的效果:

The **animal** didn't cross the **street** because **it** was too **tired**.

The **animal** didn't cross the **street** because **it** was too **narrow**.



# seq2seq

✓ 加入Attention的Seq2Seq

Attention at time step 4

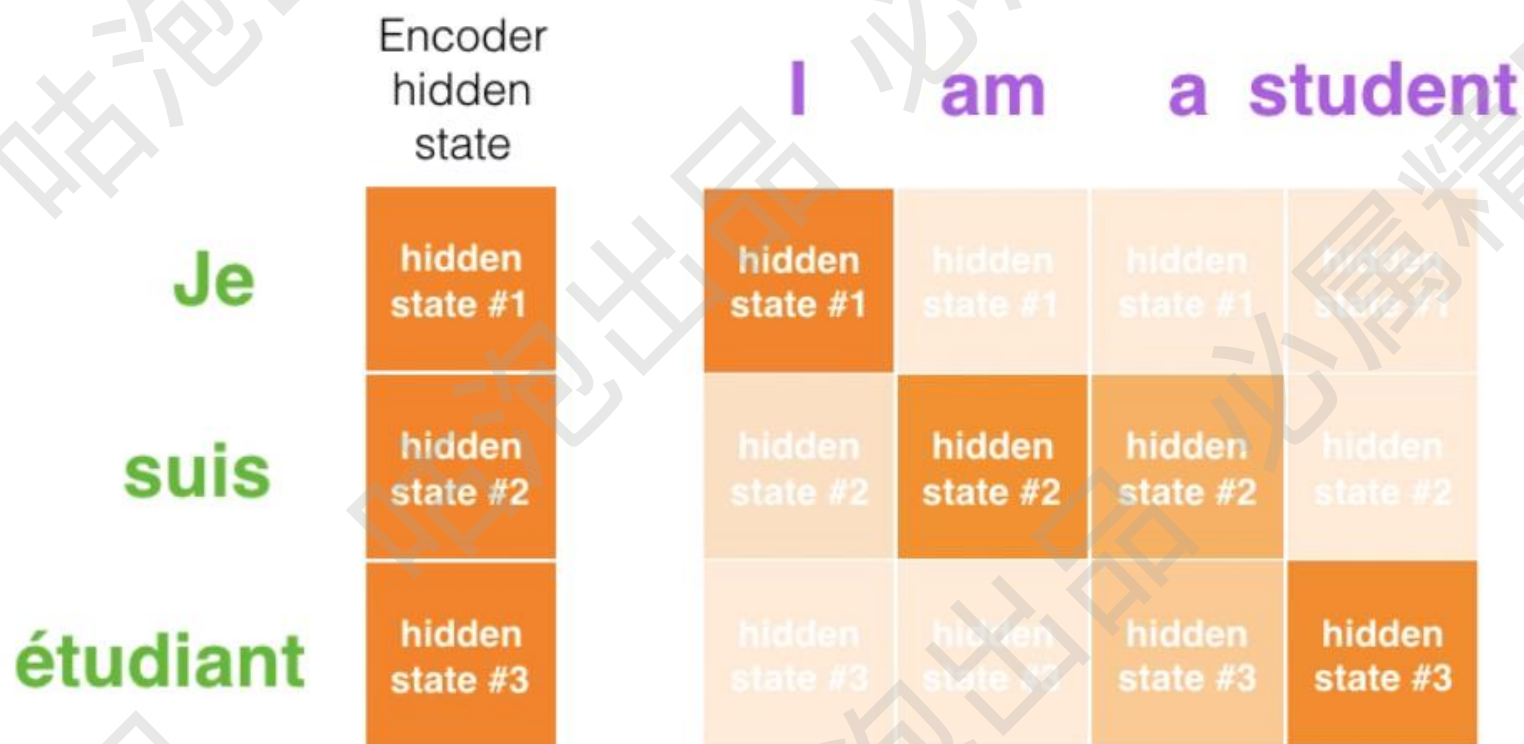




# seq2seq

✓ 加入Attention的Seq2Seq

✎ 其实就相当于每一个输出的关注点都是不同的！



# seq2seq

## ✓ Teacher Forcing

✎ 从小我的妈妈就跟我说，要跟好孩子在一块玩。。。

