

双向循环神经网络



未来很重要



I am _____
I am _____ very hungry,
I am _____ very hungry, I could eat half a pig.

未来很重要



I am **happy**.

I am **not** very hungry,

I am **very** very hungry, I could eat half a pig.

未来很重要



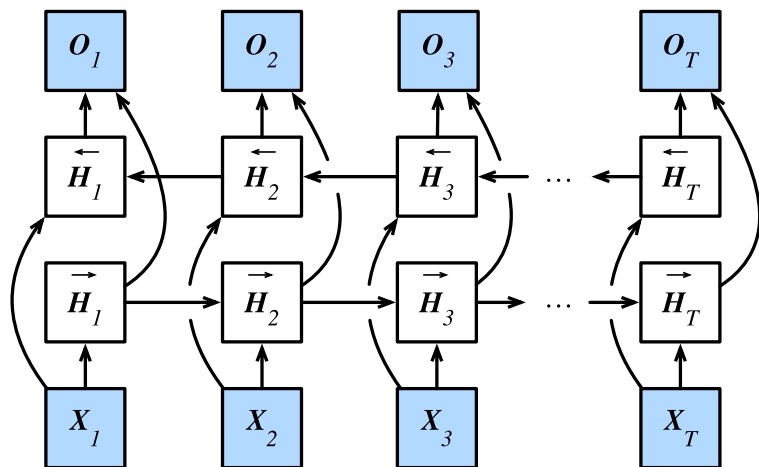
I am **happy** .

I am **not** very hungry,

I am **very** very hungry, I could eat half a pig.

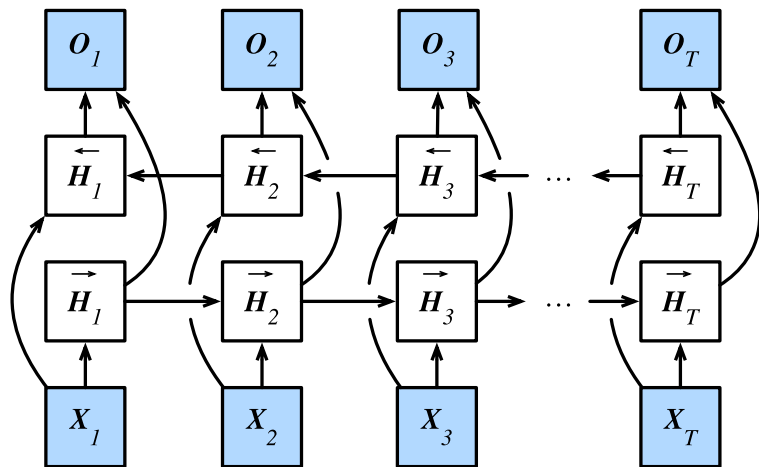
- 取决于过去和未来的上下文，可以填很不一样的词
- 目前为止RNN只看过去
- 在填空的时候，我们也可以看未来

双向 RNN



- 一个前向RNN隐层
- 一个反向RNN隐层
- 合并两个隐状态得到输出

双向 RNN



$$\vec{\mathbf{H}}_t = \phi(\mathbf{X}_t \mathbf{W}_{xh}^{(f)} + \vec{\mathbf{H}}_{t-1} \mathbf{W}_{hh}^{(f)} + \mathbf{b}_h^{(f)}),$$

$$\overleftarrow{\mathbf{H}}_t = \phi(\mathbf{X}_t \mathbf{W}_{xh}^{(b)} + \overleftarrow{\mathbf{H}}_{t+1} \mathbf{W}_{hh}^{(b)} + \mathbf{b}_h^{(b)}),$$

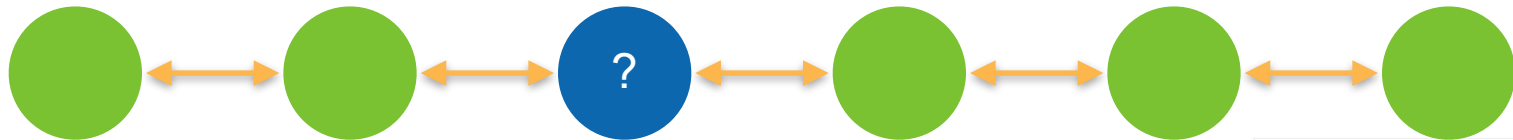
$$\mathbf{H}_t = [\vec{\mathbf{H}}_t, \overleftarrow{\mathbf{H}}_t]$$

$$\mathbf{O}_t = \mathbf{H}_t \mathbf{W}_{hq} + \mathbf{b}_q$$

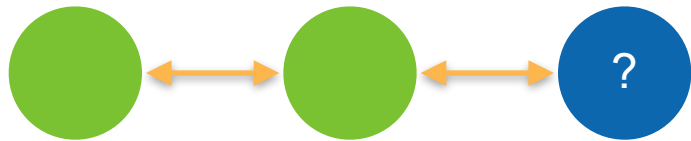
推理



• 训练：



• 推理：





总结

- 双向循环神经网络通过反向更新的隐藏层来利用方向时间信息
- 通常用来对序列抽取特征、填空，而不是预测未来