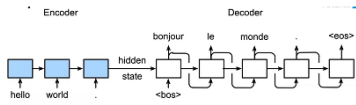


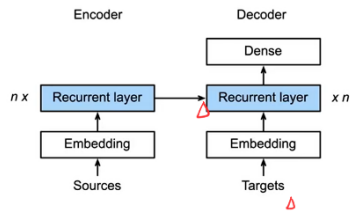
Seq2Seq

- 机器翻译

- 给定一个源句子，自动翻译成目标语言
- 这两个句子可以有不同长度
- 编码器是一个RNN（可以双向），最后一个时刻的隐藏状态传给解码器



训练时前一刻的输出不会放到下一时刻输入
推理时则需要



Encoder 是没有输出的 RNN

编码器最后时间步的隐状态用作解码器的初始隐状态
(和原始数据一起作解码器输入)

- 衡量生成序列好坏的 BLEU

- P_n 是预测中所有 n -gram 的精度（预测序列里长度为 n 的子串在标签序列的出现率）
 - 标签序列 A B C D E F 和预测序列 A B B C D 有: $P_1=4/5$, $P_2=3/4$, $P_3=1/3$, $P_4=0$
- BLEU 的定义:

$$BLEU = \exp \left[\min \left(0, 1 - \frac{\text{len}(\text{table})}{\text{len}(\text{pred})} \right) \right] \prod_{n=1}^k p_n^{\frac{1}{2n}}$$

控制序列长度，惩罚过短的预测

长匹配具有高权重