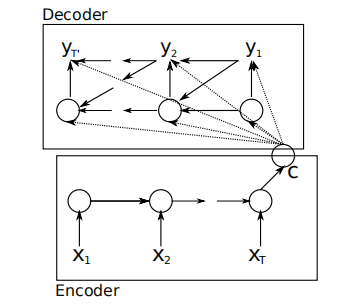
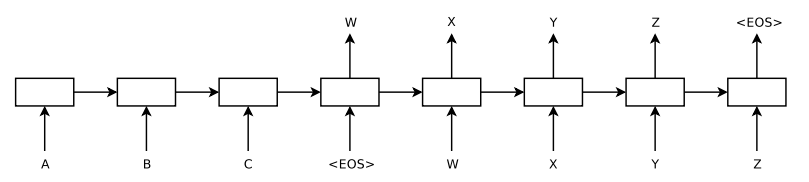
**Encoder-Decoder**

在原理上与seq2seq的想法是一致的，都是先编码再解码。只不过实现上稍有区别，



主要在解码阶段，每次都要接收3个输入，背景向量，目标输入和隐藏层输入。而seq2seq没有背景向量的输入。



训练效果目标如下图所示，首先将源序列通过编码器编码，将隐藏层信息（背景向量）传给解码器。解码阶段，首先第一个输入是开始标识<SOS>，然后将解码结果作为下一时刻的输入（选择输出层最大分量对应的字符）。以此类推直到输出出现结束符<EOS>为止。

**Decoder**

**A**

**B**

**C**

**<EOS>**

**X**

**Encoder**

**Y**

**Z**

**<EOS>**

**<SOS>**

训练阶段，一个有助于快速收敛的方法叫做Teacher Forcing，即如下所示，将目标输出直接作为解码器的输入。

**Decoder**

**A**

**B**

**C**

**<EOS>**

**X**

**Encoder**

**Y**

**Z**

**<EOS>**

**<SOS>**

**X**

**Y**

**Z**

**Teacher Forcing**

一般训练时要将上面这两种解码方式按一定比例混合后训练。