论文笔记《A deep generative framework for paraphrase generation》

论文来源：2018 AAAI

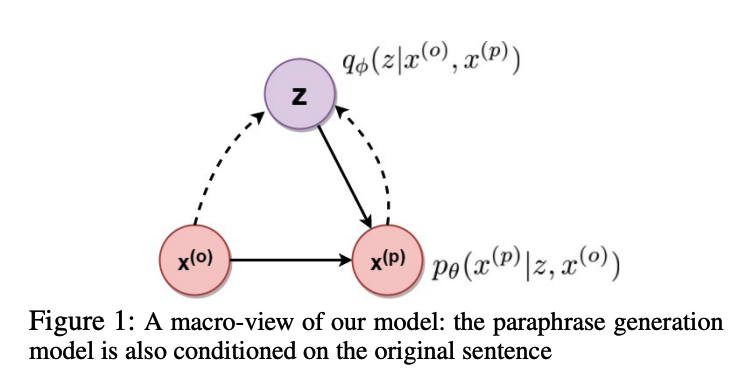
论文主要内容：

本文主要将VAE与LSTM结合起来生成复述。传统的AVE结合RNN时，生成的句子不符合复述要求，所以作者在模型的encoder和decoder部分都将原始的句子作为condition进行干涉，这样就达到了复述的效果。

模型结构：

在训练时作者使用了N对句子，其中原始句子用表示，复述句子用表示；原始句子的向量表示标记为，复述句子的向量表示为。

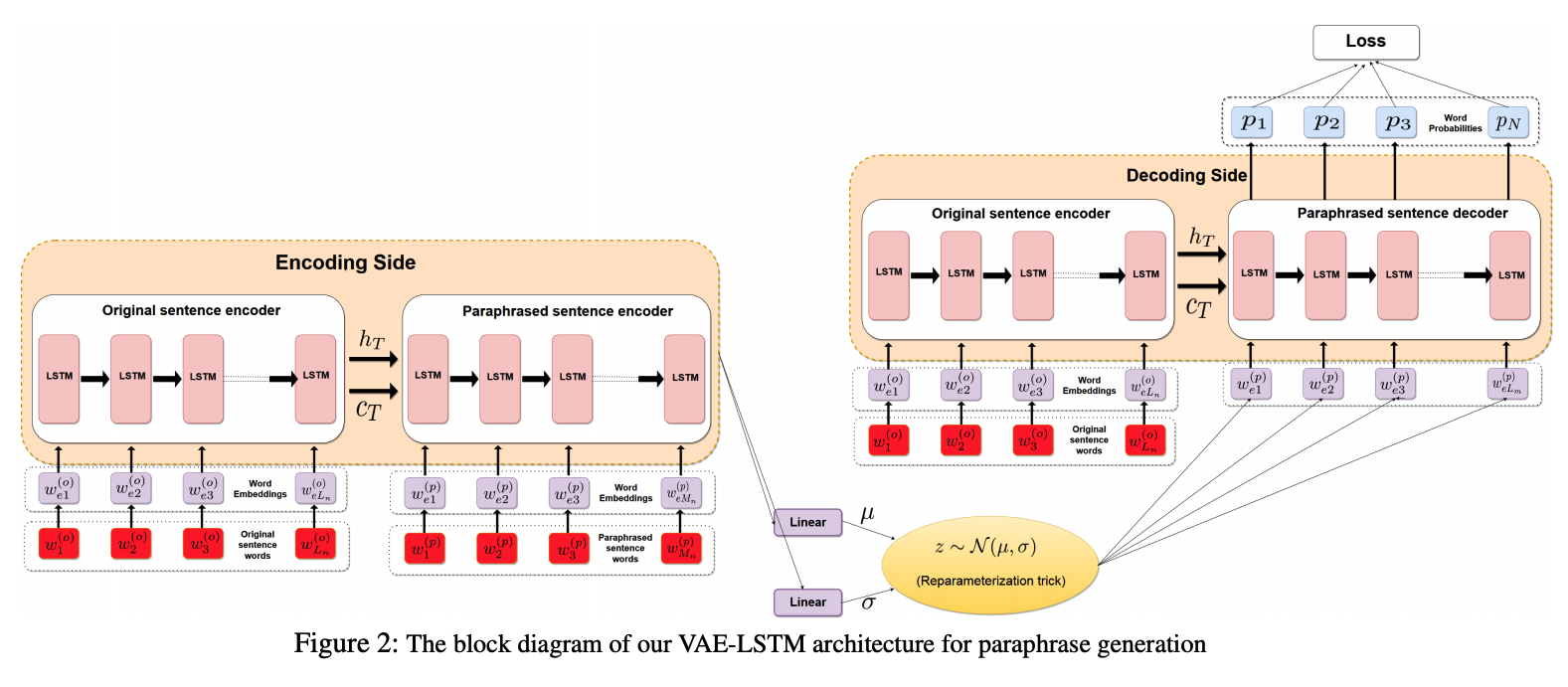
去掉LSTM后，本文的宏观模型如图所示：



从图中可以看出在生成隐向量z时，是基于和的；在生成复述表示时，是基于和隐向量z的。

区别传统的VAE，本文构建的模型将作为了condition加在了encoder和decoder之上。

具体的模型结构如下：

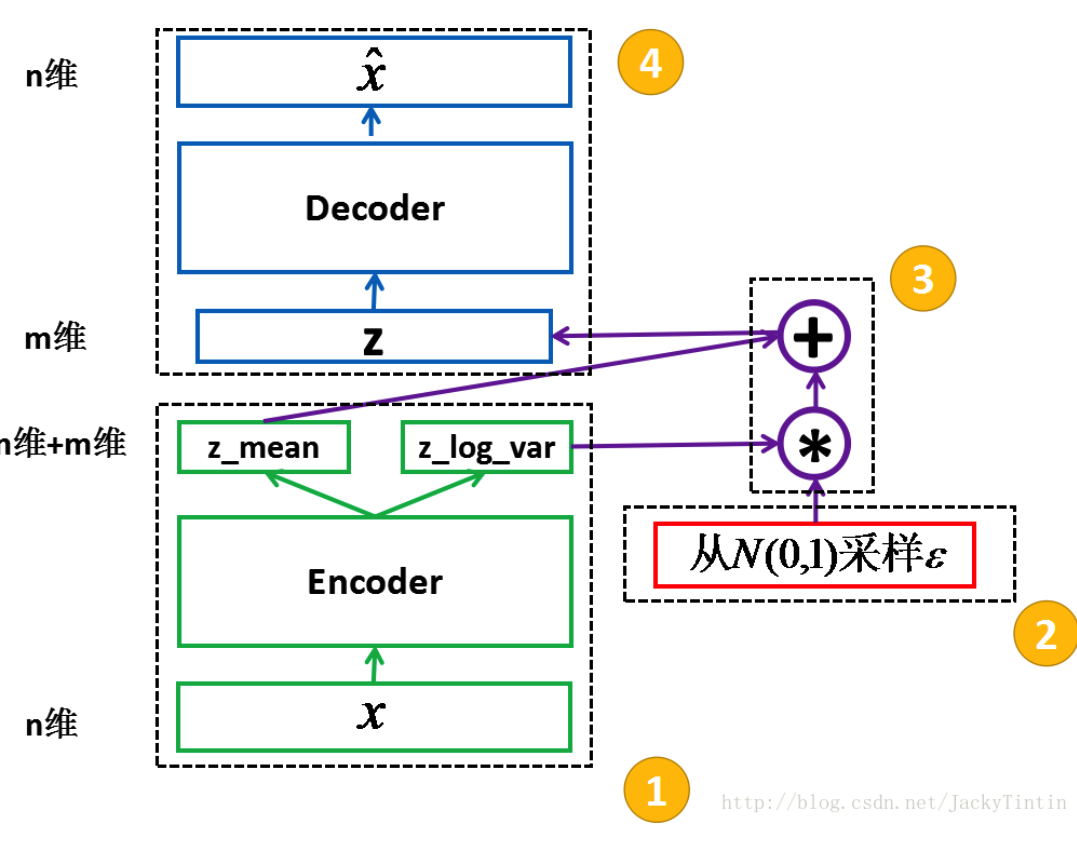


本文的模型主要包括四个LSTM，其中3个LSTM encoder，一个LSTM decoder。

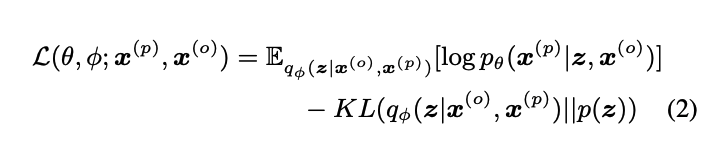
在VAE的encoder方面，两个LSTM encoder被使用，第一个转换原始句子到其向量表示，一个将复述句子转到其向量表示。将这两个句子都编码进了LSTM以后，使用一个前馈神经网络，获取向量表示的均值和方差，送入VAE。

在VAE的decoder方面，VAE输出隐变量z，第三个LSTM编码原始句子，然后将原始句子的向量表示与z一起送入第四个LSTM，产生复述句子。

将LSTM抽象化，可以得到整个模型的架构如下【也就是VAE的结构】：



训练的目标还是和VAE的目标一样，最大化下面的概率函数：



最后的实验结构如下【其中VAE-SVG与VAE-SVG-eq的区别就是，VAE-SVG-eq在encoder和decoder中对进行encoder的两个LSTM是参数共享的】：

