论文笔记《Style-transformer：unpaired text style transfer without disentangled latent representation》

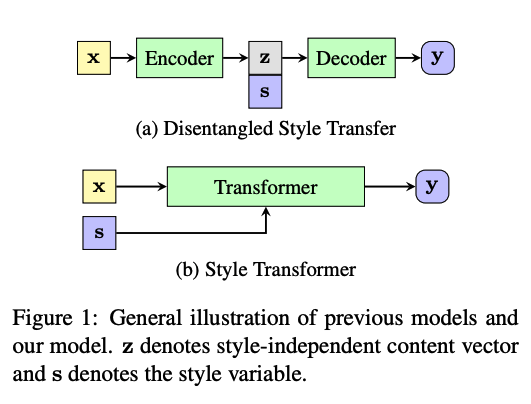
论文来源：2019 ACL

论文代码：<https://github.com/fastnlp/style-transformer>

论文主要内容：

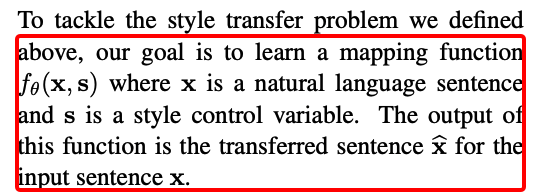
一般的style transfer，都是从source sentence中把content和style分隔开，这样的做法存在着一些问题。文中提出了一种做法，不采用disentangled latent representation的方式，并且作者采用了transformer结构，没有用RNN，并在一些指标上取得了state-of-art的效果。

下图是作者的整体流程与其他人的区别：



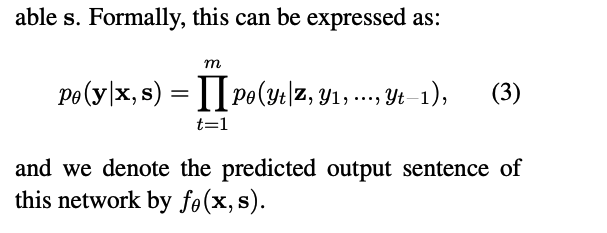
**Model Overview**

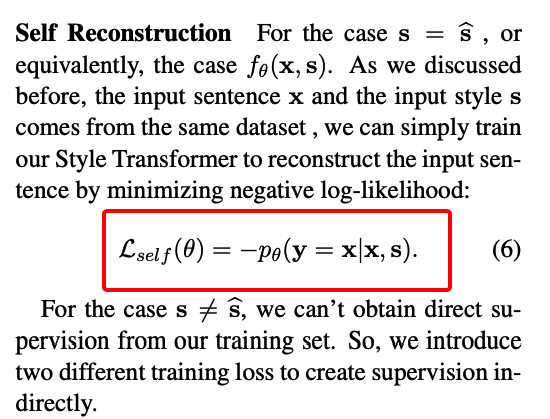
表示有k个数据集，每个数据集是一种style，是他的style symbol。

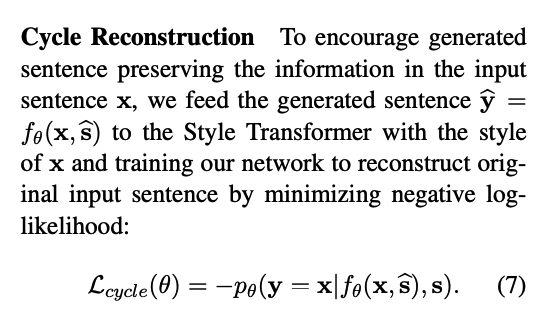


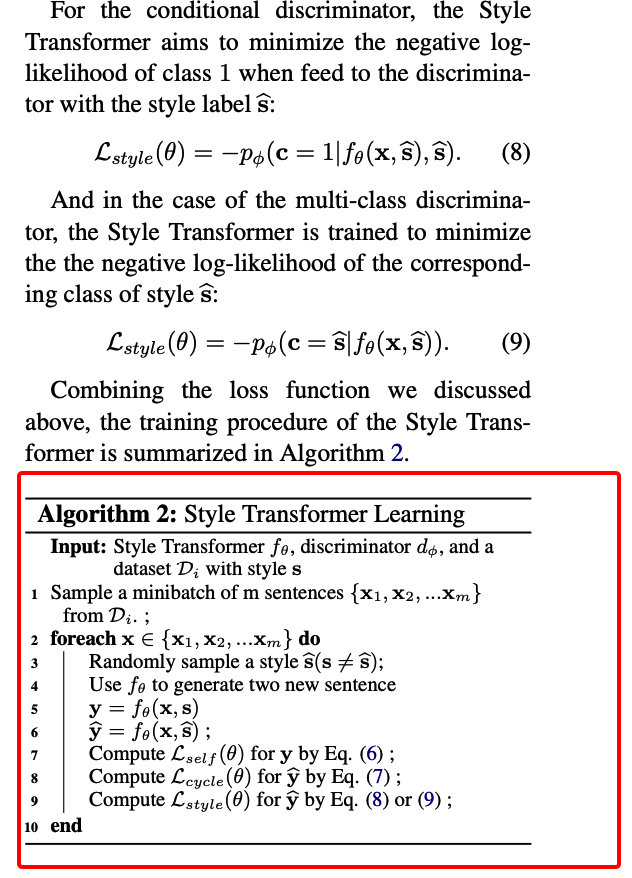
模型主要分为两部分，一个是style transformer network，将源语句转换为新style的句子，一个是discriminator network，鉴别当前语句是哪种style。

**Style transformer network**

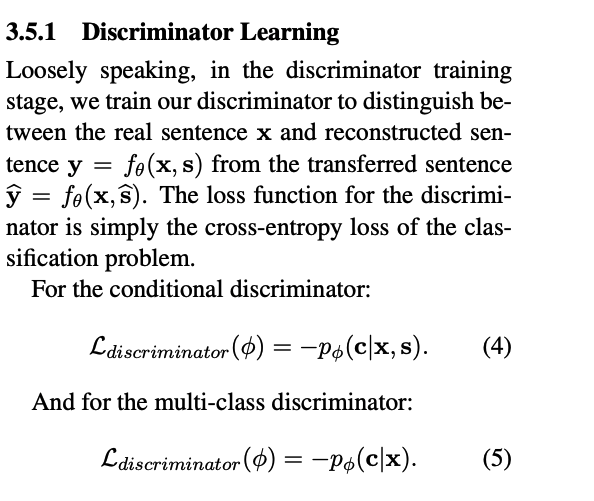


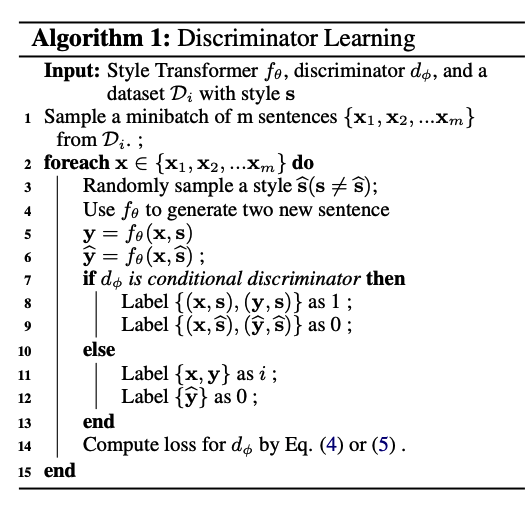
****

****

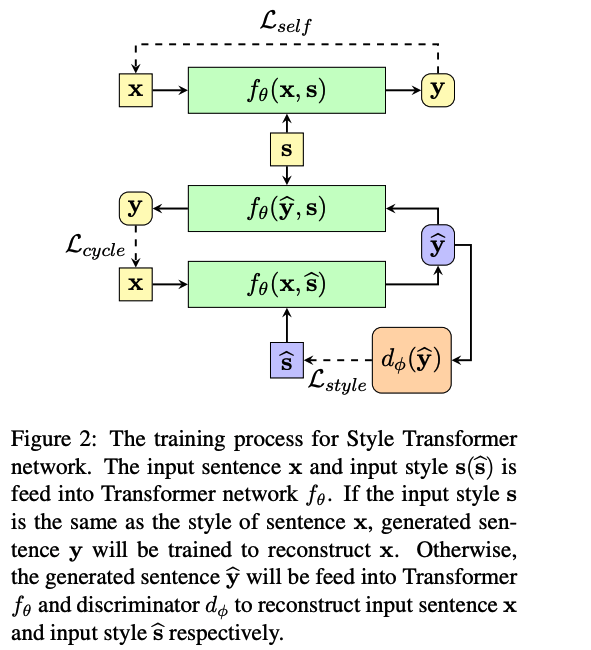
****

**Discriminator Network**

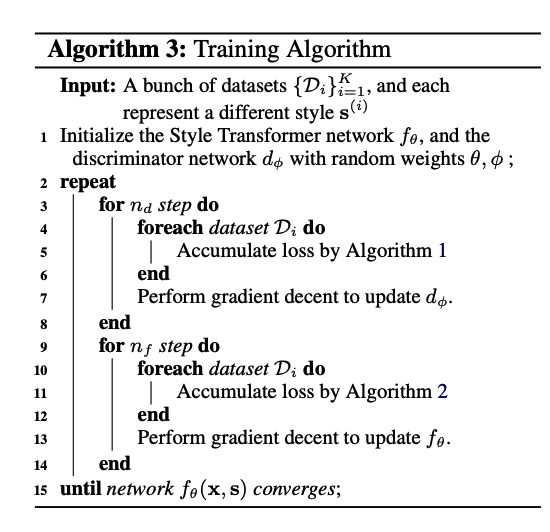




系统流程图：



**整体的训练过程：**



总而言之，先训练discriminator网络，在训练style网络，并且discriminator网络有两种不同的结构，主要区别在于conditional 判别器的输入有两个x和s，来判断句子x和style s是否一致，返回结果为0或1，而multi-task 判别器的输入只有句子x，判定句子x的风格类别，如果是生成的不同的句子则类别为0。

实验结果：

