**Synthetic Data Experiments**

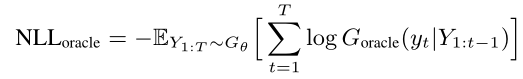
为了测试效果并增加我们对SeqGAN的理解，我们使用合成数据进行模拟测试

为了模拟真实世界的结构化序列，我们考虑一种语言模型来捕获token的依赖性

我们使用随机初始化的LSTM作为真实模型，也就是oracle，为以下实验生成实际数据分布p（x t | x 1，...，x t-1）

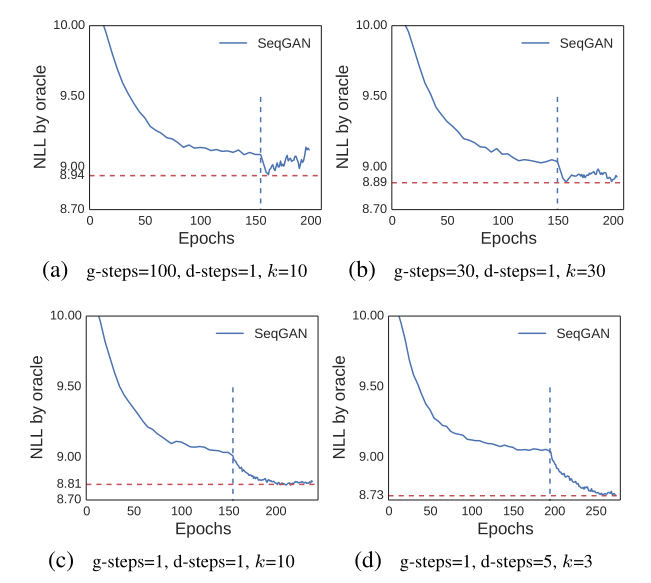
**Evaluation Metric（评估指标）**

我们知道MLE正试图最小化真实数据分布p和我们的近似q之间的交叉熵，即-E x〜p logq（x）。然而，评估生成模型的最准确方法是我们从中抽取一些样本，让人类观察者根据他们的先验知识对其进行评估。我们假设人类观察者已经学会了自然分布的精确模型p human（x）为了增加通过图灵测试的概率，我们实际上需要最小化完全相反的平均负对数似然-E x〜q logp human（x），其中p和q的角色被交换。在我们的合成数据实验中，我们可以认为oracle是现实世界问题的人类观察者，因此一个完美的评估指标应该是



其中Gθ和G oracle分别表示我们的生成模型和oracle模型

在测试阶段，我们使用Gθ生成100,000个序列样本，并通过G oracle计算每个样本的NLL oracle及其平均分数。还进行显着性测试以比较基线和SeqGAN之间的生成性能的统计特性



不同训练策略下SeqGAN的负对数似然（NLL）收敛性能，垂直虚线表示对抗训练的开始

将四种生成模型与SeqGAN进行比较，第一个模型是随机token生成模型，第二个模型是MLE训练的LSTM Gθ，第三个是scheduled sampling（Bengio et al.2015），第四个是BLEU的策略梯度模型（PG-BLEU）。

在scheduled sampling中，训练过程逐渐从完全引导的方案变为LSTM，进入较少的引导方案，该方案主要向LSTM提供其生成的标记。课程率ω用于控制用生成的token替换真实token的概率。为了获得良好且稳定的性能，我们将每个训练时期的ω降低0.002

在PG-BLEU算法中，我们使用BLEU，一个衡量生成序列和参考（训练数据）之间相似性的度量，从蒙特卡罗搜索中对完成的样本进行评分

由于评估指标具有根本性的指导意义，我们可以看到SeqGAN的影响，它明显优于其他基线。还对来自比较模型的生成序列的NLL oracle得分分布进行了显着性T检验，这证明了SeqGAN相对于所有比较模型的显着改进。

图4中显示的学习曲线明确地说明了SeqGAN的优越性，在大约150个训练时期之后，MLE和schedule sampling方法都收敛到相对较高的NLL oracle分数，而SeqGAN可以改善生成器的极限，其结构与基线相同。这表明将对抗性训练策略应用于离散序列生成模型以突破MLE局限性的前景。此外，SeqGAN优于PG-BLEU，这意味着GAN中的判别信号比预定义得分（例如BLEU）更通用和有效，以指导生成策略捕获序列数据的基础分布

**Discussion**

在我们的综合数据实验中，我们发现SeqGAN的稳定性取决于训练策略。这种策略导致快速收敛，但随着发生器快速改进，鉴别器无法完全训练，从而逐渐提供误导信号

上述所有三种训练策略中的d步都设置为1，这意味着我们只生成一组与给定数据集具有相同编号的否定示例，然后针对不同的k个时期训练其上的鉴别器

图3（d）显示这种技术可以提高整体性能并具有良好的稳定性，因为鉴别器显示的是更多的负面例子，每次都强调正面例子，这将导致更全面的培训生成器指导

在分析生成对抗网的收敛性时，一个重要的假设是允许鉴别器达到其最佳给定G，只有当鉴别器能够一致地区分真实数据和非自然数据时，来自它的监督信号才有意义，整个对抗训练过程才能稳定有效

对于文本生成场景，我们应用拟议的SeqGAN来生成中国诗歌和巴拉克奥巴马的政治演讲

在诗歌创作任务中，我们使用了16,394个中文绝句的语料库，每个绝句包含4行，共20个字符

在奥巴马的政治演讲生成任务中，我们使用了一个语料库，这是奥巴马政治演讲中的11,092段的集合

我们使用BLEU分数作为评估来度量生成的文本与人类创建的文本之间的相似度

我们混合了20首真正的诗歌和20首由SeqGAN和MLE生成的诗歌，然后邀请了70位中国诗歌专家来判断60首诗中的每首诗是由人还是机器创作的。一旦被认为是真实的，它获得+1分，否则为0，最后，计算每个算法的平均分数。结果表明，对于诗歌创作，SeqGAN的表现与真实的人类数据相当。