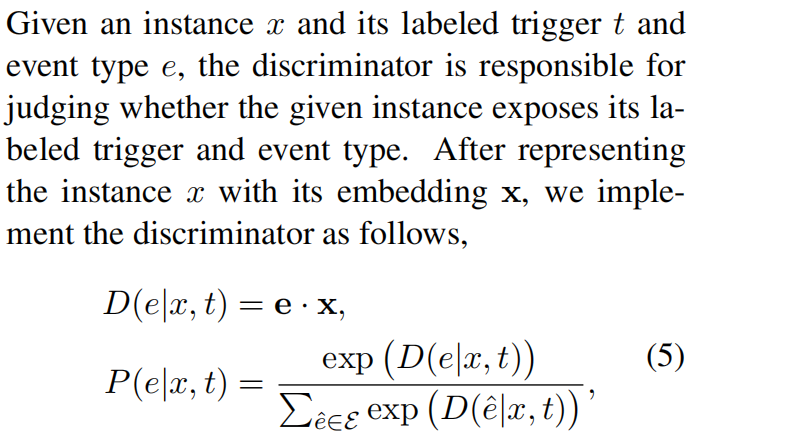
本质是想通过替换后生成尽可能真的事件的语义向量。

这篇的综述写的很好。。。

After representing

the instance *x* with its embedding **x**



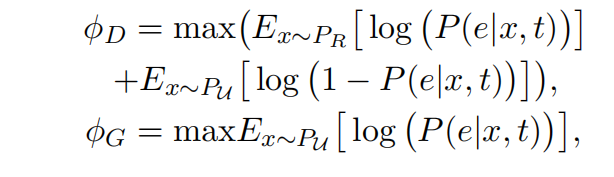
也许给生成样本器也有点启示

这篇其实就是个典型的gan思想，但是他是旨在**选择**好样本（远程监督。）（但他的判别器其实已经把任务做完了。我实际想的是只用一个判别器判断真假。而不应该像在语义对抗中把任务做完。其实整个这一套弄下来有好效果的概率非常高。在最夸张的时候可以采取一个语义平滑。）

我认为这样有个很大的问题，这样判别器的

Generator生成的被几乎认为是干扰。选出了迷惑性最强的（但如果是为生成而生呢）

Electra，weakly suprevised



我们只是为了替换对应参数，trigger判断真假

等于说是distant learning（但distant是规则样本的，但g如果是个选择器就不一样了）和 结合。

必要时可以迭代gan，原因主要在于gan的任务不应该太复杂，不然g完全可以生成噪音来让d完全没有策略。语义对抗学习中g的结构不应该太复杂。

Weakly-Supervised

Weakly-Supervised Dual Generative Adversarial Networks for Makeup-Removal

Distant argument

我想确认下文本类的gan不好做是（生成文本。）（相似度回归生成文本如何？）

分类本身用在gan前面是什么操作?来个玄学过头的如何？？如果我们后面是个很复杂的模型我是否实现了一个蒸馏？？？？蒸馏很多时候会直接去对应参数。

weakly supervisedweakly supervised到底有什么。

Gcn对抗学习。。

语义对抗

因为很多时候用的是 同质化的节点

图结构很多时候是样本决定的。‘

词向量每一维很难说是独立还是不独立的。

判别器的迁移。

Rnn的玄学，缺陷

如果真的判别任务有类别的话把判断真假和类别一起做时总有一种做宝搞的感觉，，，。

[1,]

虚数生成网络。

元技术核心：gan，gcn，迁移学习（zsl与预训练）。

Gan+ZSL的话会很难。

Gcn+gan。

辅技术：多语言，多任务，注意力。

不靠谱技术：复数网络。

它拿的高斯分布（拉普拉斯近似）近似的迪利克雷先验。gumbel分布可能更适合。。。或者二者混合。

如果zero shot用的隐藏vae如何

多个gan，但可能泛化的会太厉害。但如果成对可以gcn。

F（x）=0...

半监督gan那篇有些做宝搞的原因在于万一远程监督生成的是合理的样本怎么办。。。要生成特定的一类是有专门的应用的。特定的合理的一类呢？标签不在损失而是在输入那边，对子的格式。

Distill gan：太花了。。Distill没体现出来。。如果目标真的在kdd那么我们应该旨在如何在线学习。。

Self-regulation：缺点如果提取的不是特征而是噪音怎么办？这样的话另一头没起到任何作用，我们其实需要做个实验观察到底输出是否合理。。

反其道而行训练一张图？？（初始是完全链接的随机值，但会舍弃过小权重）（训练的图结构也是一个编码。）

对抗训练有两个作用，一是提高模型对恶意攻击的鲁棒性，二是提高模型的泛化能力。

在CV任务，根据经验性的结论，对抗训练往往会使得模型在非对抗样本上的表现变差，然而神奇的是，在NLP任务中，模型的泛化能力反而变强了，如[1]中所述：

people observe the opposite result for language models (Miyato et al., 2017; Cheng et al., 2019), showing that adversarial training can improve both generalization and robustness.

**在NLP任务中，对抗训练的角色不再是为了防御基于梯度的恶意攻击，反而更多的是作为一种regularization，提高模型的泛化能力**。

FreeLB-Roberta