GAN在文本生成中的应用

姜衡军

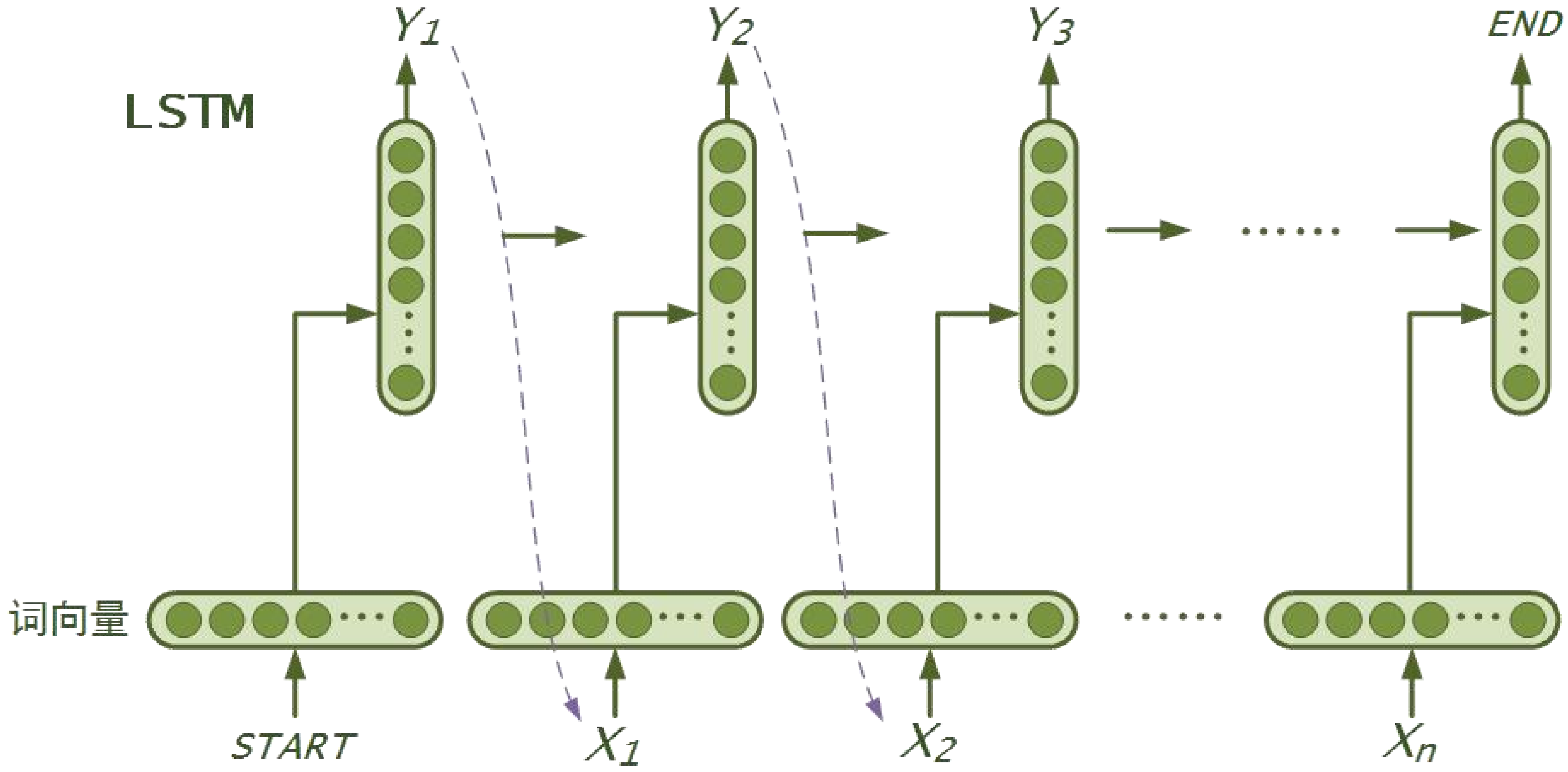
Outlines

- Neural Text Generation
  - methods/models
  - limitations
- LeakGAN
- MaskGAN
- LeakMaskGAN
- Experiments/Results

Neural Text Generation Models

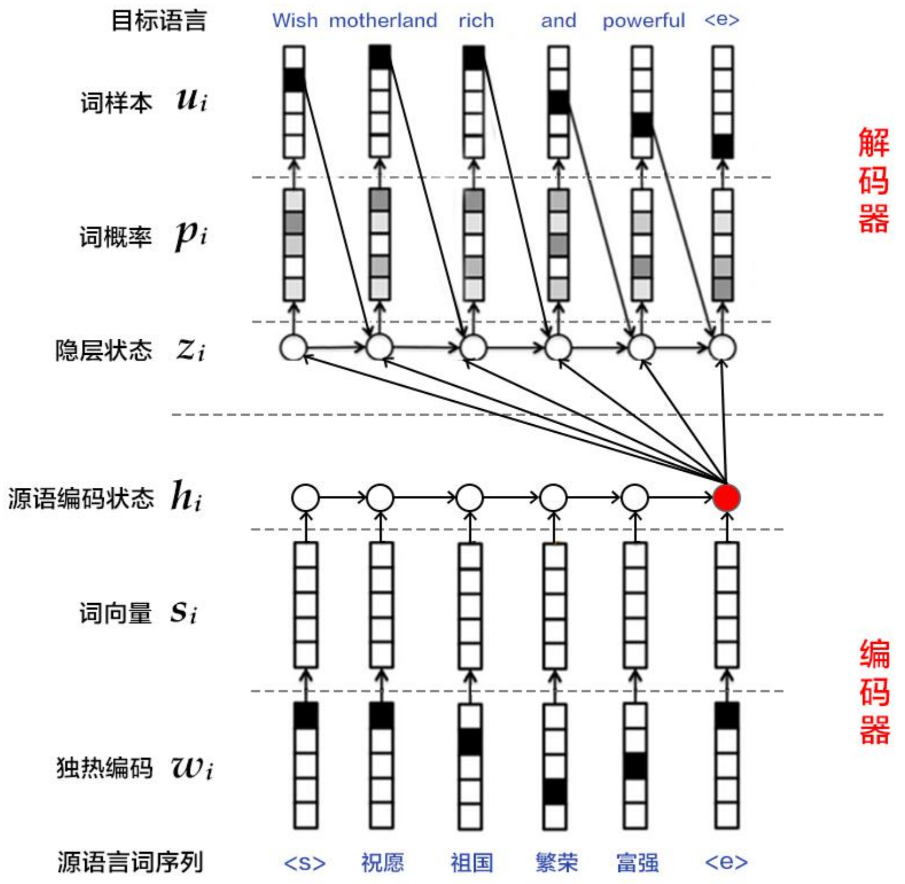
- Language Model(LM)
- Sequence to Sequence(seq2seq)
- Neural Turing Machine(NTM)
- Reinforcement Learning(RL)
- Generative Adversarial Nets(GAN)

Text Generation with language model



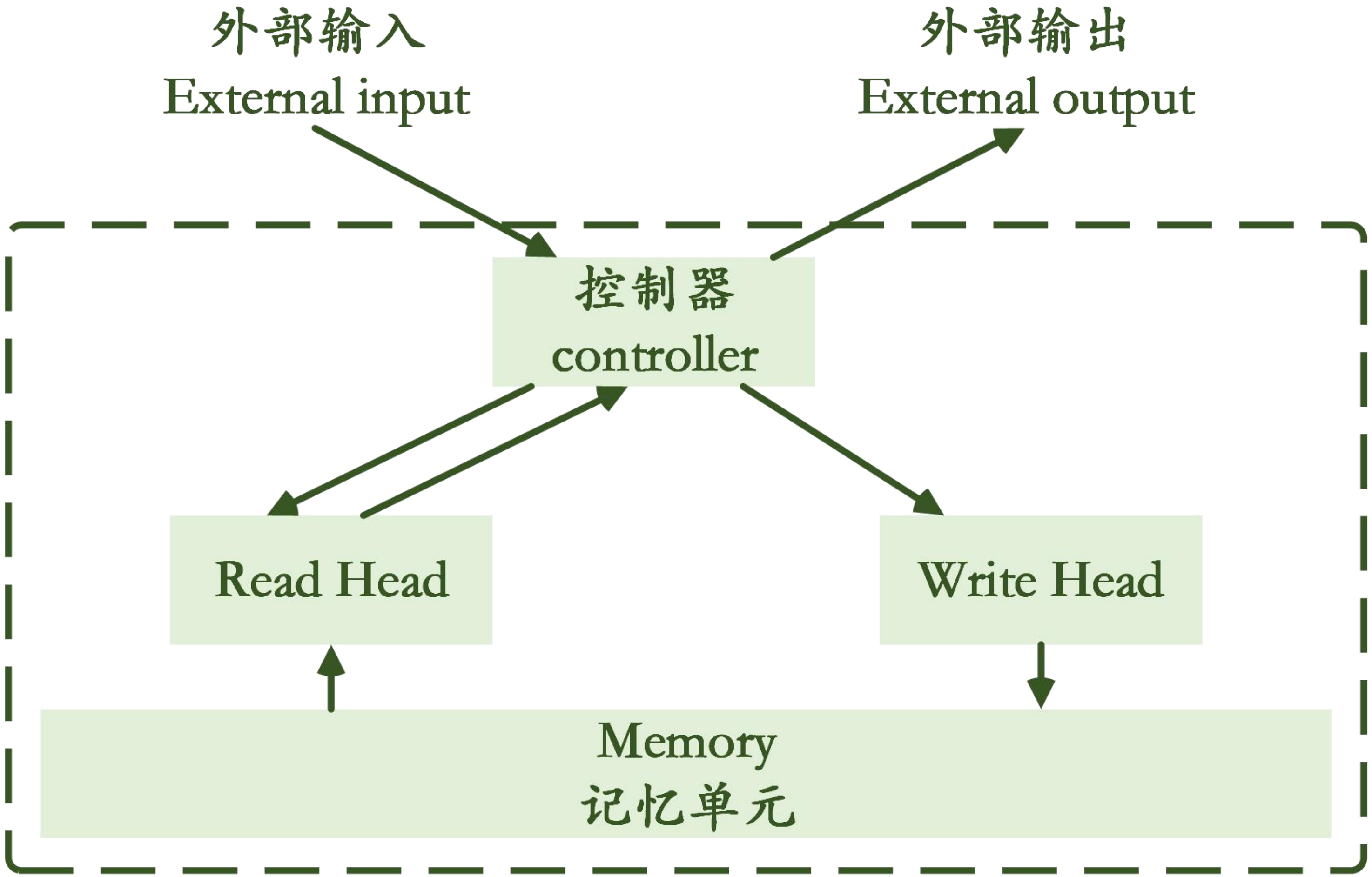
图片参考：https://gitbook.cn/books/58ec9b969741d1032f26c300/index.html

Text Generation with seq2seq



图片参考：https://gitbook.cn/books/58ec9b969741d1032f26c300/index.html

Text Generation with NTM



图片参考：https://gitbook.cn/books/58ec9b969741d1032f26c300/index.html

Text Generation with RL

$$J_\theta(\hat{s}_n) = -\sum_{t=0}^{n-1} R_t \log \hat{P}(x_t|s_t)$$

$$R_t = \mathbb{E}_{s\sim G_\theta(\cdot|s_{t+1})}[\mathbf{BLEU}(s)]$$

Text generation with RNNLMs can be viewed as a Markov decision process (MDP), the locally optimal policy of which can be found through reinforcement learning (RL) .

Limitations

Cons of MLE and PG-BLEU

- MLE suffers from so-called exposure bias, which is due to the inherent difference between the training stage and inference stage of language models trained via MLE. The effect of exposure bias becomes more obvious and serious as the sequence becomes longer.[2]
- First, BLEU is not a computationally cheap metric. Second, BLEU is not a perfect metric even not a strong one as it just counts the n-gram statistics similarity between the generated text and the reference text (corpus). Therefore, it introduces much unnecessary bias into the model.[2]

Scheduled Sampling

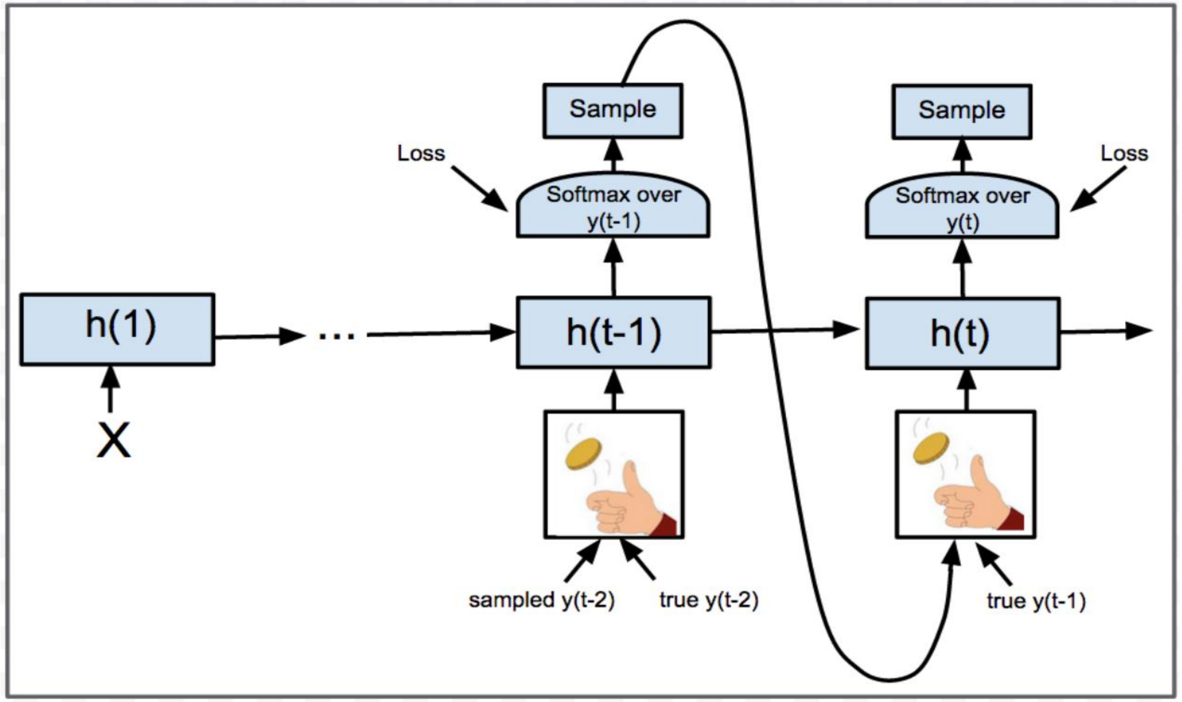
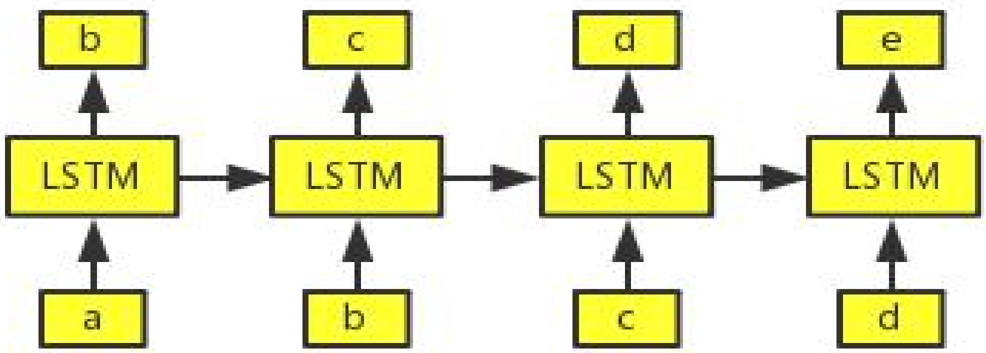
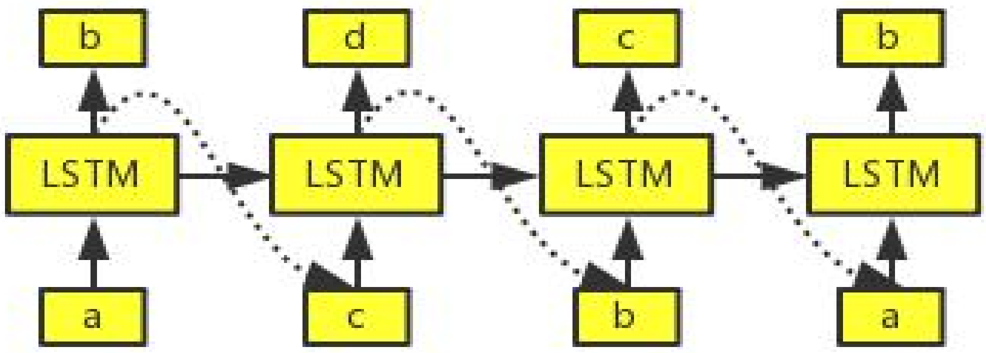


Figure 1: Illustration of the Scheduled Sampling approach, where one flips a coin at every time step to decide to use the true previous token or one sampled from the model itself.

Cons of Scheduled Sampling



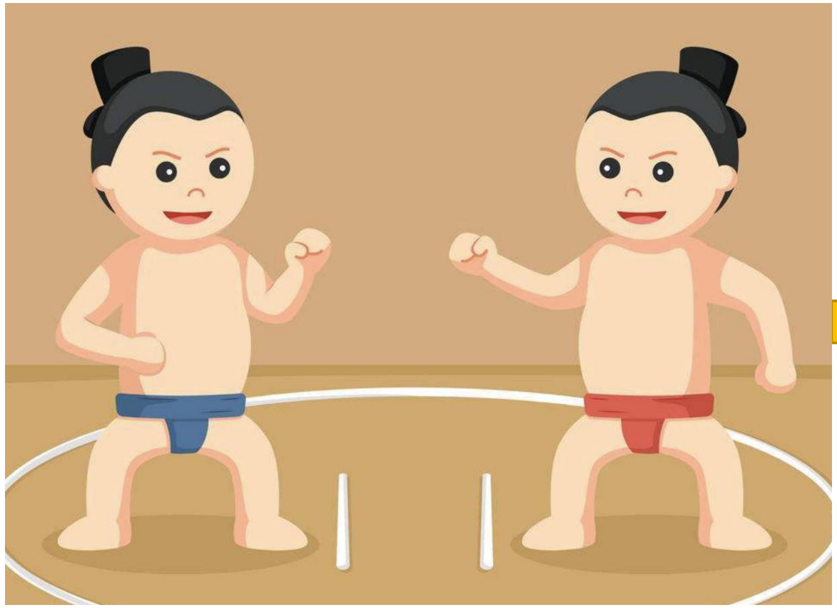
a→b
b→c
c→d
d→e



a→b
c→c(d)
b→d(c)
a→e(b)

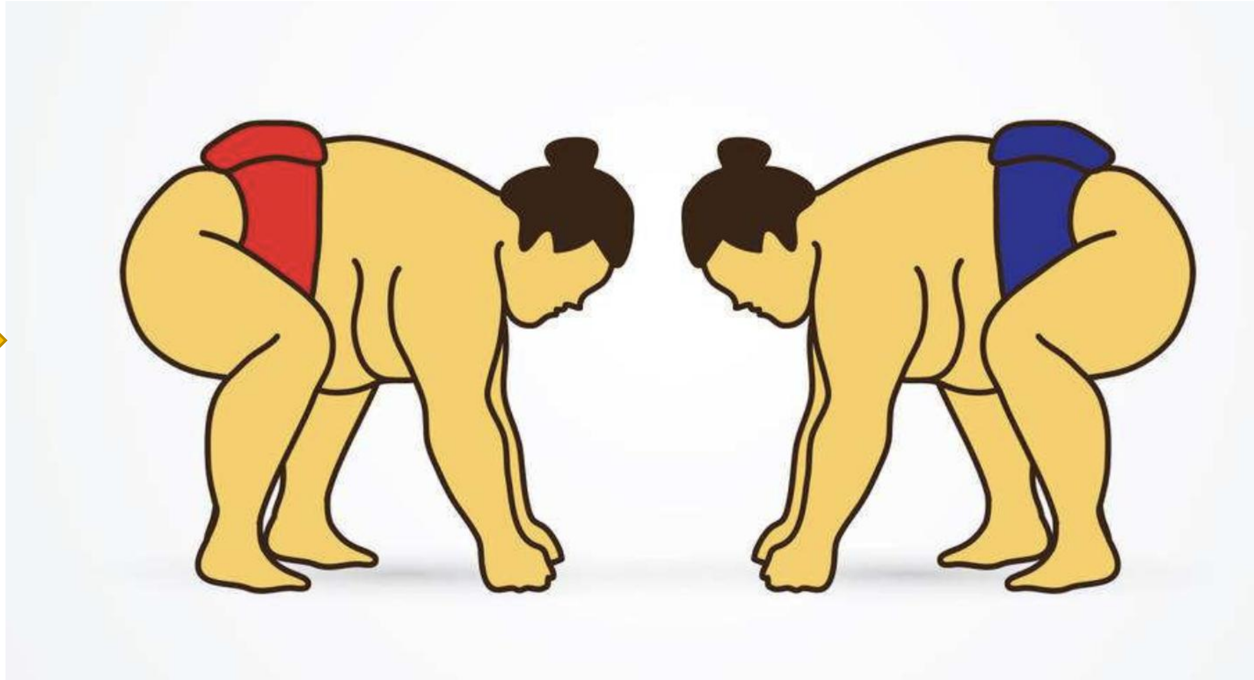
Text Generation with GAN

$$\min_{\theta}\max_{\phi}\mathbb{E}_{s\sim p_{\text{data}}}\left[\log(D_\phi(s))\right]+\mathbb{E}_{s\sim G_\theta(\cdot)}\left[\log(1-D_\phi(s))\right]$$



Adversarial

Training



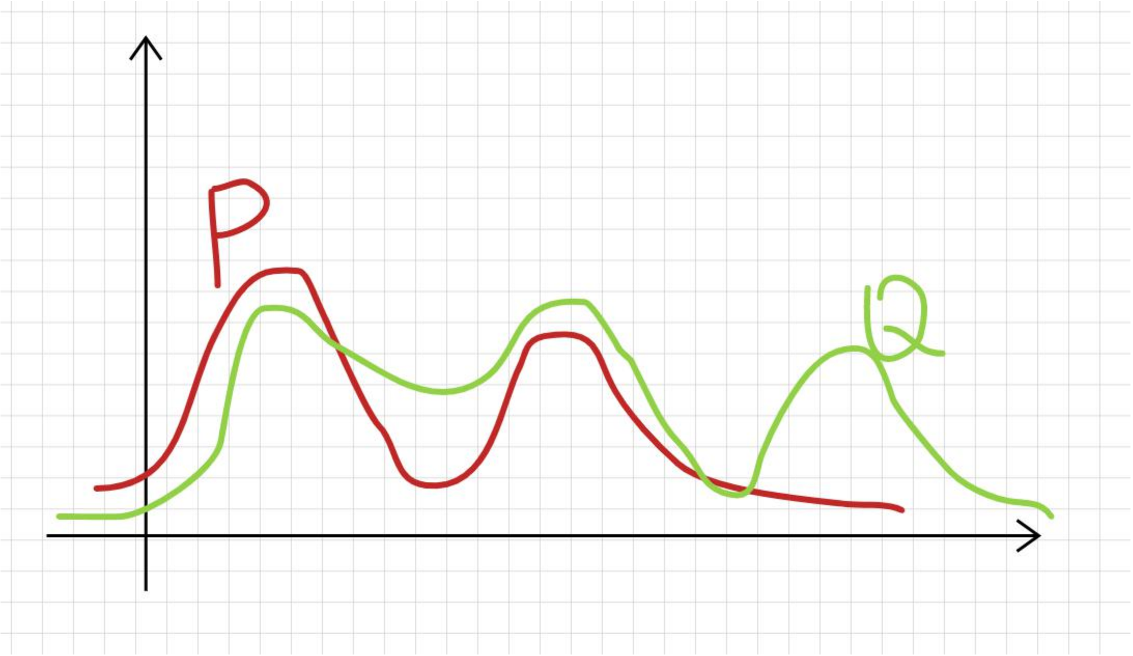
Text Generation with GAN

Why GAN?

Rethinking MLE

Maximise MLE = Minimise KL(p||q)

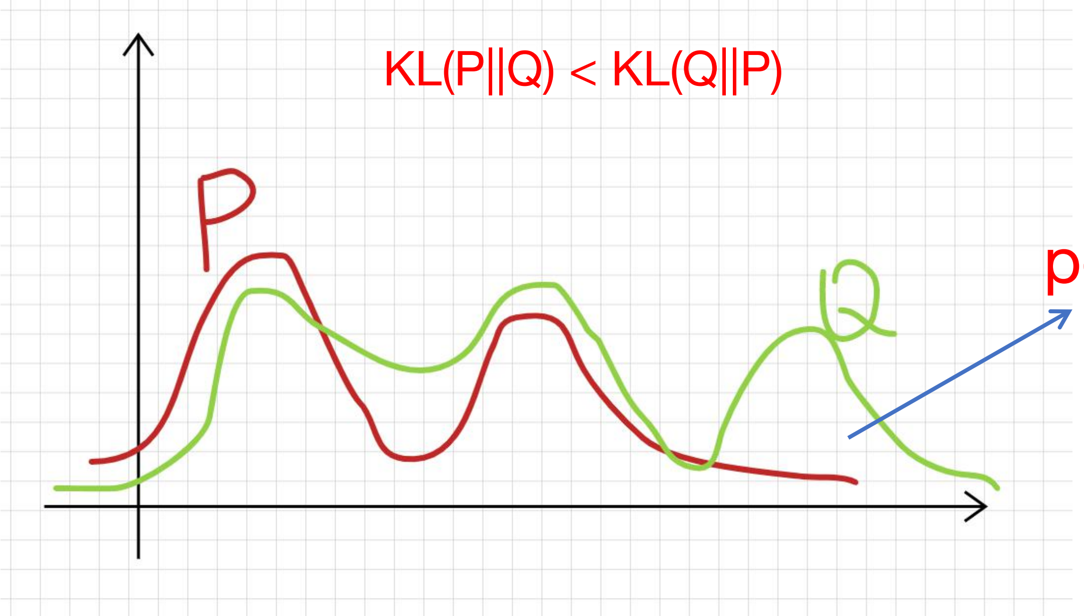
$$\mathrm{H}(p,q) = \mathrm{E}_p[-\log q] = \mathrm{H}(p) + D_{\mathrm{KL}}(p\|q)$$



Cons of MLE

Maximise MLE = Minimise KL(p||q)

$$\mathrm{H}(p,q) = \mathrm{E}_p[-\log q] = \mathrm{H}(p) + D_{\mathrm{KL}}(p\|q)$$



poor perceptual sample quality

Rethinking Scheduled Sampling

Consider a learning sequence of lengtl

MLE objective function:

$$
\begin{aligned}
D_{ML}[P\|Q] &= KL[P\|Q] \\
&= KL[P_{x_1}\|Q_{x_1}] +
\end{aligned}
$$

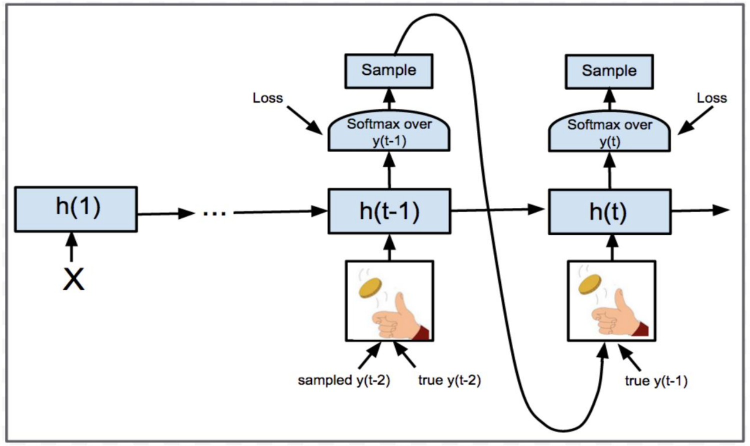


Figure 1: Illustration of the Scheduled Sampling approach, where one flips a coin at every time step to decide to use the true previous token or one sampled from the model itself.

Scheduled Sampling objective function:

$$
D_{SS}[P\|Q] = \underbrace{\boxed{KL[P_{x_1}\|Q_{x_1}] + \epsilon \mathbb{E}_{z\sim P_{x_1}} KL[P_{x_2|x_1=z}\|Q_{x_2|x_1=z}]}}_{\text{MLE objective}} + (1-\epsilon)\mathbb{E}_{z\sim Q_{x_1}} KL[P_{x_2}\|Q_{x_2|x_1=z}]
$$

Cons of Scheduled Samping

Consider a learning sequence of length 2 : (x_1, x_2)

Schduled Sampling objective function:

$$D_{SS}[P\|Q] = KL[P_{x_1}\|Q_{x_1}] + \epsilon \mathbb{E}_{z\sim P_{x_1}} KL[P_{x_2|x_1=z}\|Q_{x_2|x_1=z}] + (1-\epsilon)\mathbb{E}_{z\sim Q_{x_1}} KL[P_{x_2}\|Q_{x_2|x_1=z}]$$

$$\Downarrow$$

$$D_{SS}[P\|Q] = KL[P_{x_1}\|Q_{x_1}] + \mathbb{E}_{z\sim P_{x_1}} KL\left[\epsilon P_{x_1|x_1=z} + \frac{Q_{x_1}(z)}{Q_{x_1}(z)} P_{x_2} \middle\| Q_{x_2|x_1=z}\right] + C_{P,\epsilon}$$

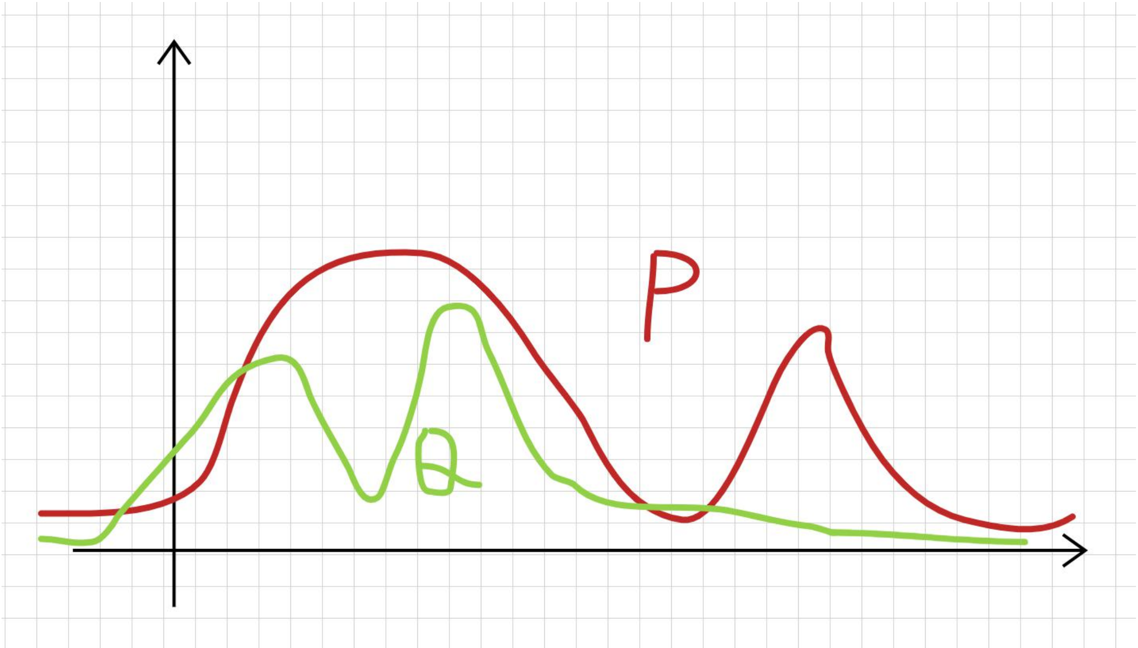
$$= KL\left[P_{x_1}\left(\epsilon P_{x_1|x_1} + (1-\epsilon)\frac{Q_{x_1}P_{x_2}}{P_{x_1}}\right) \middle\| Q_{x_1,x_2}\right] + C_{P,\epsilon}$$

As $\epsilon \to 0$, Minimise $D_{SS}[P\|Q] \longrightarrow P \neq Q$!

KL[Q||P]

Perceptual Quality Metric: $\mathbb{E}_{\mathbf{x}\sim Q} \log P(x)$

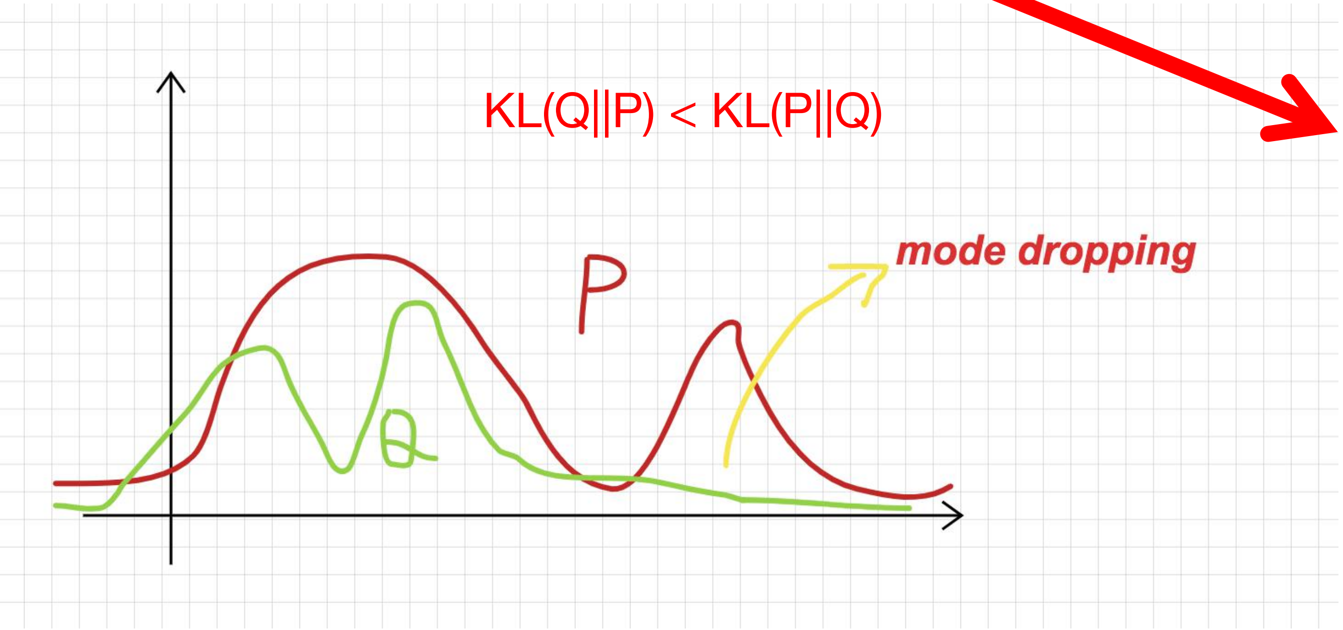
$$KL[Q||P] = \boxed{-\mathbb{E}_{\mathbf{x}\sim Q} \log P(x)} + \mathbb{E}_{\mathbf{x}\sim Q} \log Q(x)$$



Cons of KL[Q||P]

Perceptual Quality Metric: $\mathbb{E}_{\mathbf{x}\sim Q} \log P(x)$

$$KL[Q\|P] = \boxed{-\mathbb{E}_{\mathbf{x}\sim Q} \log P(x)} + \mathbb{E}_{\mathbf{x}\sim Q} \log Q(x)$$



intractable when x~Q is not in the support set of P

Jensen-Shannon Divergence

Jensen-Shannon divergence(symmetric):

$$JSD[P\|Q] = JSD[P\|Q] = \frac{1}{2}KL\left[P\middle\|\frac{P+Q}{2}\right] + \frac{1}{2}KL\left[Q\middle\|\frac{P+Q}{2}\right]$$

Adversarial training can be described as minimising an approximation to the Jensen-Shannon divergence between P and Q.[5]

Generalised Jensen-Shannon Divergence

Ger

JS_π

$(1-\pi)Q]$

$\lim_{\pi\to 0}$

$\pi)Q]\}$

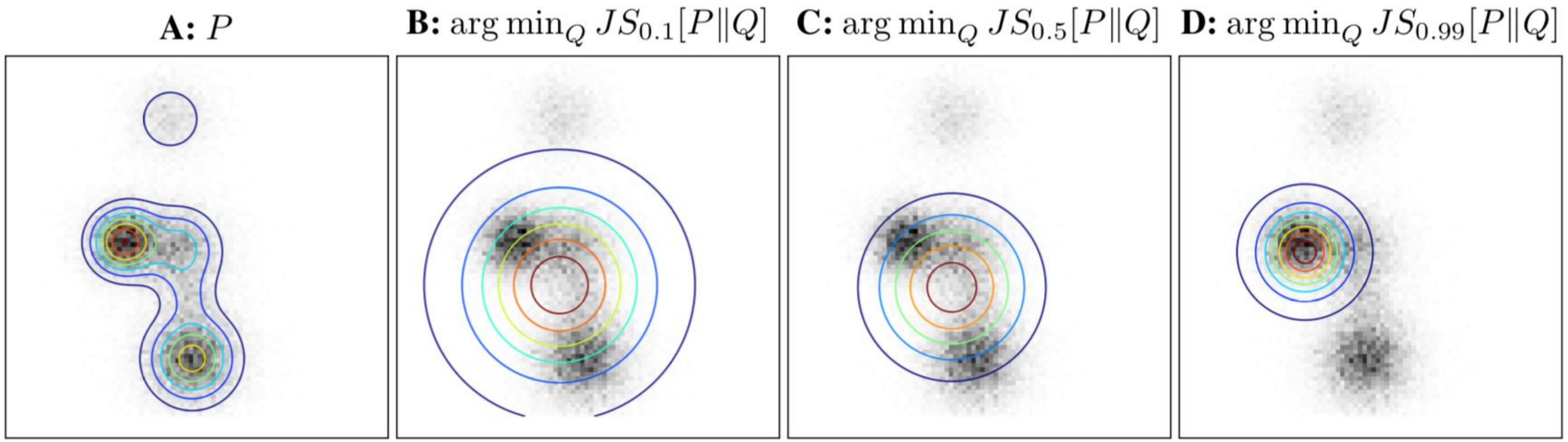


Figure 1: Illustrating the behaviour of the generalised JS divergence under model underspecification for a range of values of π. Data is drawn from a multivariate Gaussian distribution P **(A)** and we aim approximate it by a single isotropic Gaussian **(B-D)**. Contours show level sets the approximating distribution, overlaid on top of the 2D histogram of observed data. For $\pi = 0.1$, JS divergence minimisation behaves like maximum likelihood **(B)**, resulting in the characteristic moment matching behaviour. For $\pi = 0.99$ **(D)**, the behaviour becomes more akin to the mode-seeking behaviour of minimising $KL[Q\|P]$. For the intermediate value of $\pi = 0.5$ **(C)** we recover the standard JS divergence approximated by adversarial training. To produce this illustraiton we used software made available by Theis et al. (2015).

effectively

viour of KL[P||Q]

$JS_\pi[P\|($

Conclusion

Adversarial training is one of the best strategies for generative modelling, if the end goal is to draw the realistic samples from the model.

Text Generation with LeakGAN

Cons of GAN:

sparse reward(in NLP)

training instability

failed in long text generation

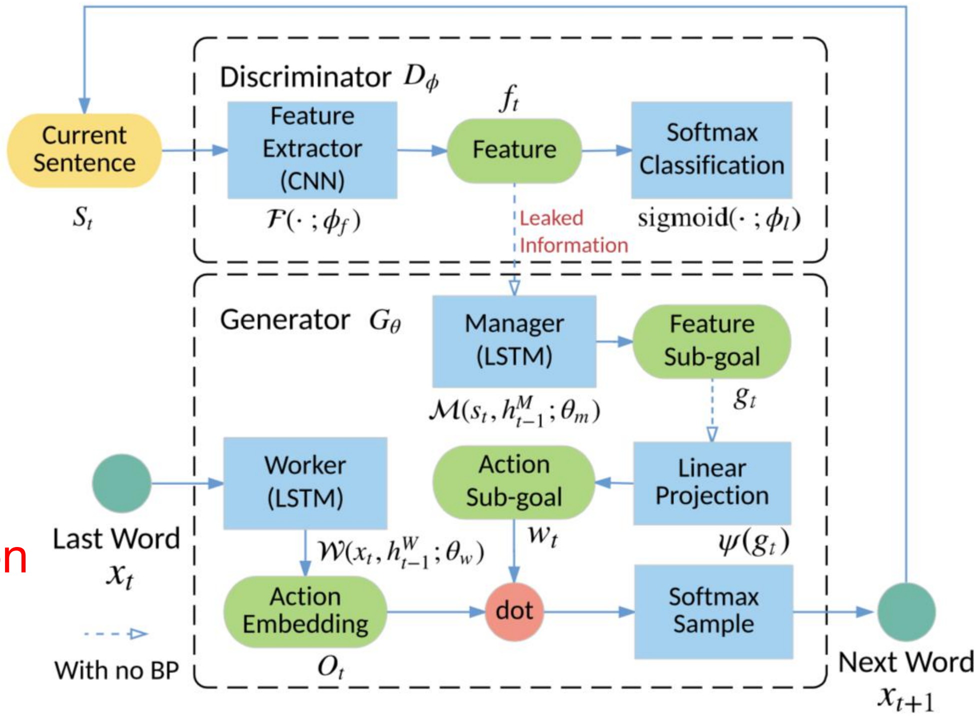


Figure 1: An overview of our LeakGAN text generation framework. While the generator is responsible to generate the next word, the discriminator adversarially judges the generated sentence once it is complete. The chief novelty lies in that, unlike conventional adversarial training, during the process, the discriminator reveals its internal state (feature f_t) in order to guide the generator more informatively and frequently. (See Methodology Section for more details.)

Text Generation with LeakGAN

$$\nabla^{\text{adv}}_{\theta_m} g_t = -Q_{\mathcal{F}}(s_t, g_t) \nabla_{\theta_m} d_{\cos}\Big(f_{t+c} - f_t, g_t(\theta_m)\Big)$$

where $Q_{\mathcal{F}}(s_t, g_t) = Q(\mathcal{F}(s_t), g_t) = Q(f_t, g_t) = \mathbb{E}[r_t]$

$$\nabla_{\theta_w} \mathbb{E}_{s_{t-1} \sim G}\Big[\sum_{x_t} r_t^I \mathcal{W}(x_t | s_{t-1}; \theta_w)\Big]$$

$$= \mathbb{E}_{s_{t-1} \sim G, x_t \sim \mathcal{W}(x_t | s_{t-1})}[r_t^I \nabla_{\theta_w} \log \mathcal{W}(x_t | s_{t-1}; \theta_w)]$$

where $r_t^I = \frac{1}{c} \sum_{i=1}^{c} d_{\cos}\Big(f_t - f_{t-i}, g_{t-i}\Big)$

Text Generation with LeakGAN

Table 2: BLEU scores performa

Method	SeqGAN	RankGA
BLEU-2	0.8590	0.778
BLEU-3	0.6015	0.478
BLEU-4	0.4541	0.411
BLEU-5	0.4498	0.463

Table 3: BLEU scores on C

Method	SeqGAN	RankGAN
BLEU-2	0.831	0.850
BLEU-3	0.642	0.672
BLEU-4	0.521	0.557
BLEU-5	0.427	0.544

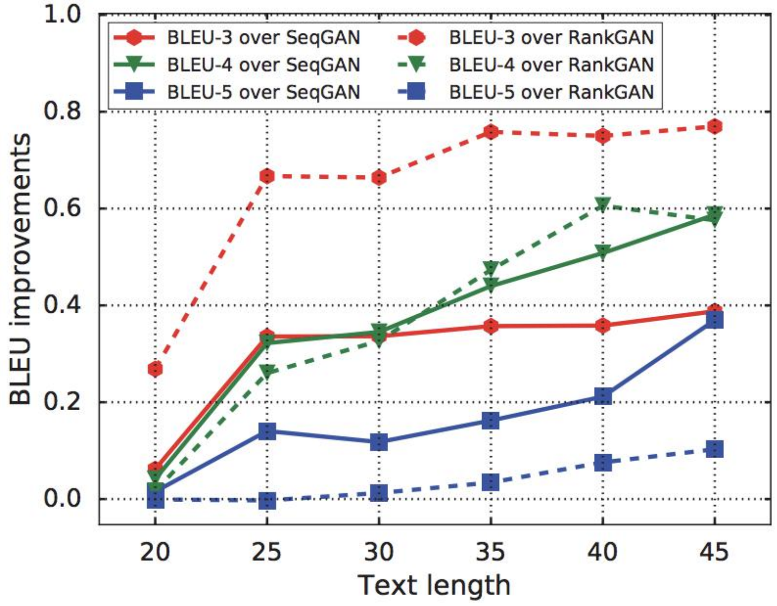


Figure 3: The illustration of BLEU improvement change along with the generated text length on WMT News.

Table 5: Turing test results for in real-world experiments.

Dataset	SeqGAN	LeakGAN	Ground Truth	p-value
WMT News	0.236	**0.554**	0.651	$< 10^{-6}$
COCO	0.405	**0.574**	0.675	$< 10^{-6}$

Text Generation with LeakGAN

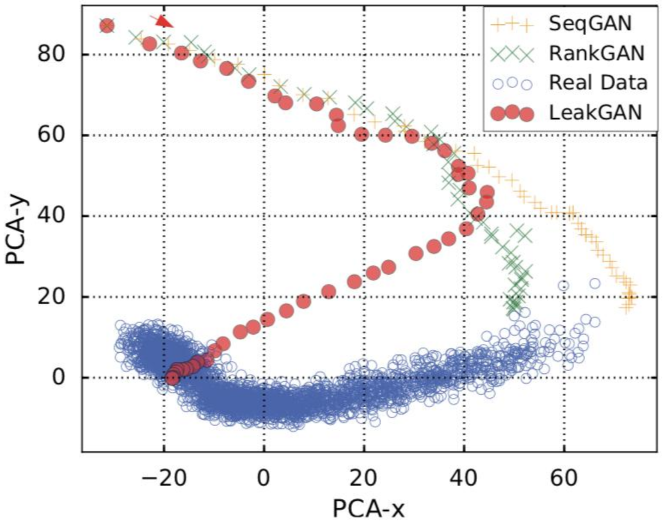
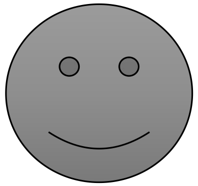


Figure 4: Feature traces during the generation (SeqGAN, RankGAN and LeakGAN) and features of completed real data (all compressed to 2-dim by PCA) on WMT News.



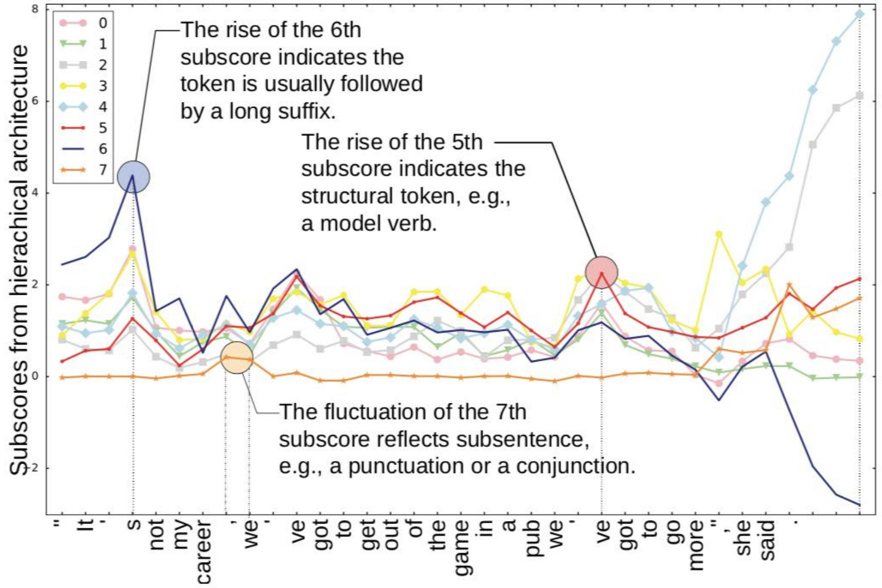
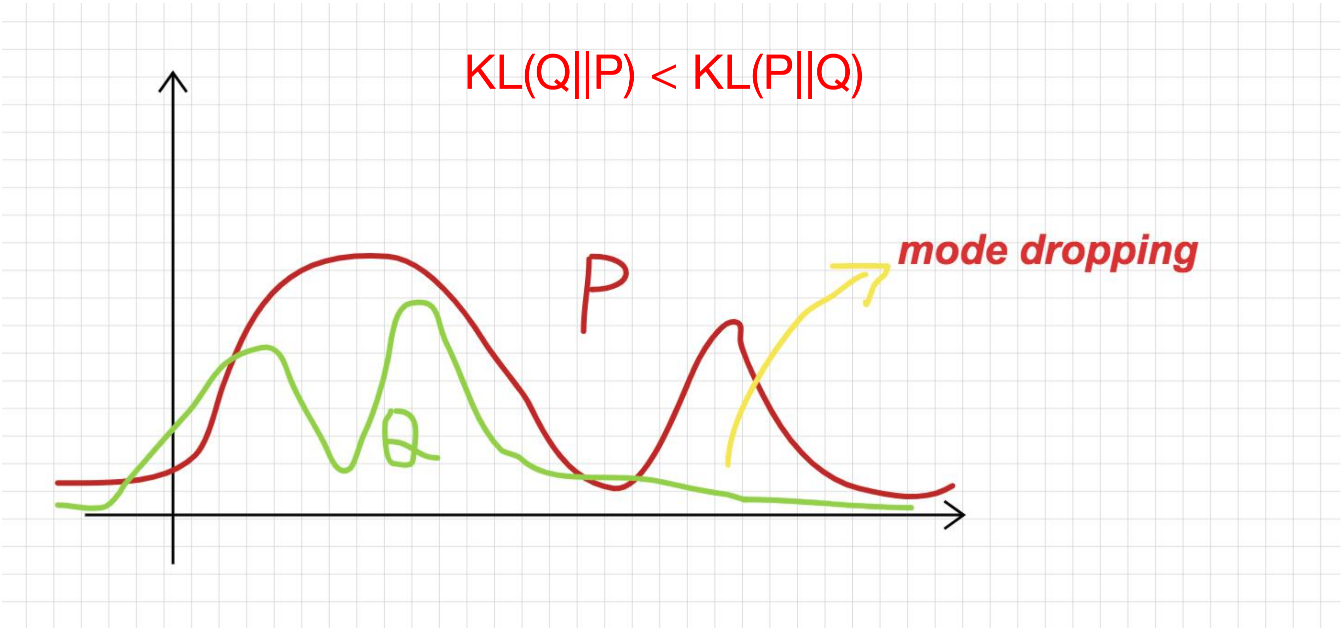


Figure 5: Illustration of WORKER and MANAGER's behaviors during a generation. (Dimension-wise Product of Worker and Manager)

Text Generation with MaskGAN

Cons of GAN:

- training instability(sparse reward)
- mode dropping(KL[Q||P])



Text Generation with MaskGAN

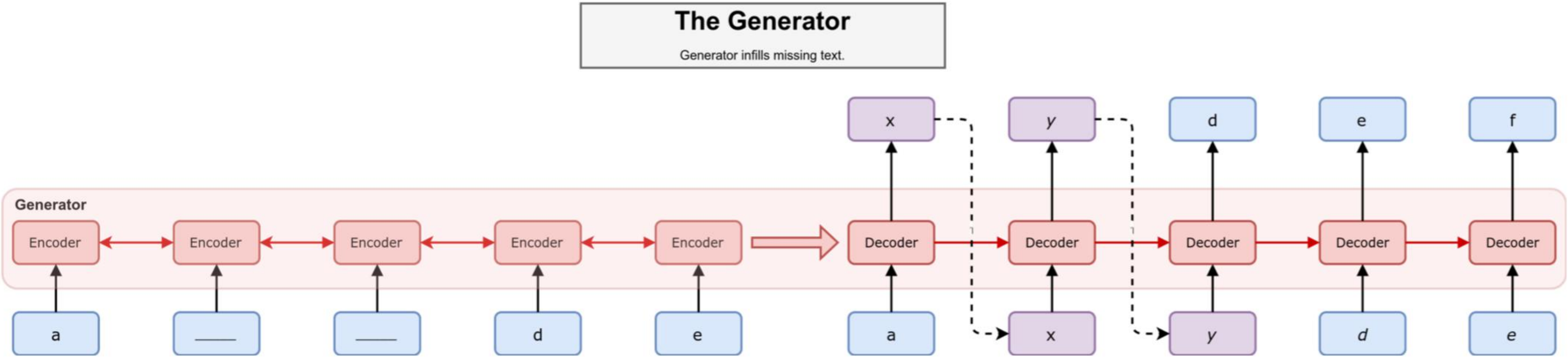


Figure 1: seq2seq generator architecture. Blue boxes represent known tokens and purple boxes are imputed tokens. We demonstrate a sampling operation via the dotted line. The encoder reads in a masked sequence, where masked tokens are denoted by an underscore, and then the decoder imputes the missing tokens by using the encoder hidden states. In this example, the generator should fill in the alphabetical ordering, (a,b,c,d,e).

Text Generation with MaskGAN

$$\nabla_\theta \mathbb{E}[R] = \mathbb{E}_{\hat{x}_t \sim G}\left[\sum_{t=1}^{T}(R_t - b_t)\nabla_\theta \log(G_\theta(\hat{x}_t))\right]$$
$$= \mathbb{E}_{\hat{x}_t \sim G}\left[\sum_{t=1}^{T}\left(\sum_{s=t}^{T}\gamma^s r_s - b_t\right)\nabla_\theta \log(G_\theta(\hat{x}_t))\right]$$

where $r_t \equiv \log D_\phi(\tilde{x}_t|\tilde{x}_{0:T}, \boldsymbol{m}(\boldsymbol{x}))$ $\quad b_t = V^G(x_{1:t})$

$$\nabla_\phi \frac{1}{m}\sum_{i=1}^{m}\left[\log D(x^{(i)})\right] + \log(1 - D(G(z^{(i)})\right]$$

Text Generation with MaskGAN

Preferred Model	Grammaticality %	Topicality %	Overall %
LM	15.3	19.7	15.7
MaskGAN	59.7	58.3	58.0
LM	20.0	28.3	21.7
MaskMLE	42.7	43.7	40.3
MaskGAN	49.7	43.7	44.3
MaskMLE	18.7	20.3	18.3
Real samples	78.3	72.0	73.3
LM	6.7	7.0	6.3
Real samples	65.7	59.3	62.3
MaskGAN	18.0	20.0	16.7

Table 7: A Mechanical Turk blind heads-up evaluation between pairs of models trained on IMDB reviews. 100 reviews (each 40 words long) from each model are unconditionally sampled and randomized. Raters are asked which sample is preferred between each pair. 300 ratings were obtained for each model pair comparison.

Preferred model	Grammaticality %	Topicality %	Overall %
LM	32.0	30.7	27.3
MaskGAN	41.0	39.0	35.3
LM	32.7	34.7	32.0
MaskMLE	37.3	33.3	31.3
MaskGAN	44.7	33.3	35.0
MaskMLE	28.0	28.3	26.3
SeqGAN	38.7	34.0	30.7
MaskMLE	33.3	28.3	27.3
SeqGAN	31.7	34.7	32.0
MaskGAN	43.3	37.3	37.0

Table 8: A Mechanical Turk blind heads-up evaluation between pairs of models trained on PTB. 100 news snippets (each 20 words long) from each model are unconditionally sampled and randomized. Raters are asked which sample is preferred between each pair. 300 ratings were obtained for each model pair comparison.

MaskGAN 108.3 ± 3.5

Table 5: The perplexity is calculated using a pre-trained language model that is equivalent to the decoder (in terms of architecture and size) used in the MaskMLE and MaskGAN models. This language model was used to initialize both models.

LeakGAN vs MaskGAN(COCO)

models\n-BLEU	2-gram BLEU	3-gram BLEU	4-gram BLEU	5-gram BLEU
LeakGAN	0.943	0.875	0.762	0.669
MaskGAN	0.946	0.851	0.717	0.616

LeakGAN vs MaskGAN(COCO)

LeakGAN samples(partial):

1. A woman is eating a white bowl and two people on it .
2. A woman in black shirt is holding a baseball bat .
3. A large kitchen with a stove , oven , sink and a refrigerator .
4. A black and white cat is parked in the ocean .
5. A woman is standing in the kitchen and a bunch of craft .
6. A bathroom with a toilet , sink , tub and a mirror .
7. A person holding a smile as she walking in a phone .
8. A very nice bathroom with a sink and a large sink .
9. A man is riding a bike with a bike .

MaskGAN samples(partial):

1. A cat stuck *up *on *a *couch *in *front *of *a *vase *.
2. Bicycles , cars and *a *stop *sign *in *front *of *an *old *truck *.
3. A lady talking *on *a *cell *phone *while *sitting *on *a *couch *.
4. A man standing in *front *of *a *TV *on *a *table *.
5. A person standing in *the *grass *with *a *surfboard *in *the *background *.
6. A man poses for a *picture *with *a *woman *in *a *yellow *shirt *.
7. A person with *a *surprised *standing *next *to *a *gate *that *is *in *the *desert *.
8. A crowd of skiers prepare *to *get *a *turn *on *the *beach *.
9. A woman takes *a *picture *of *a *man *in *a *cart *holding *glasses *.

Key Words

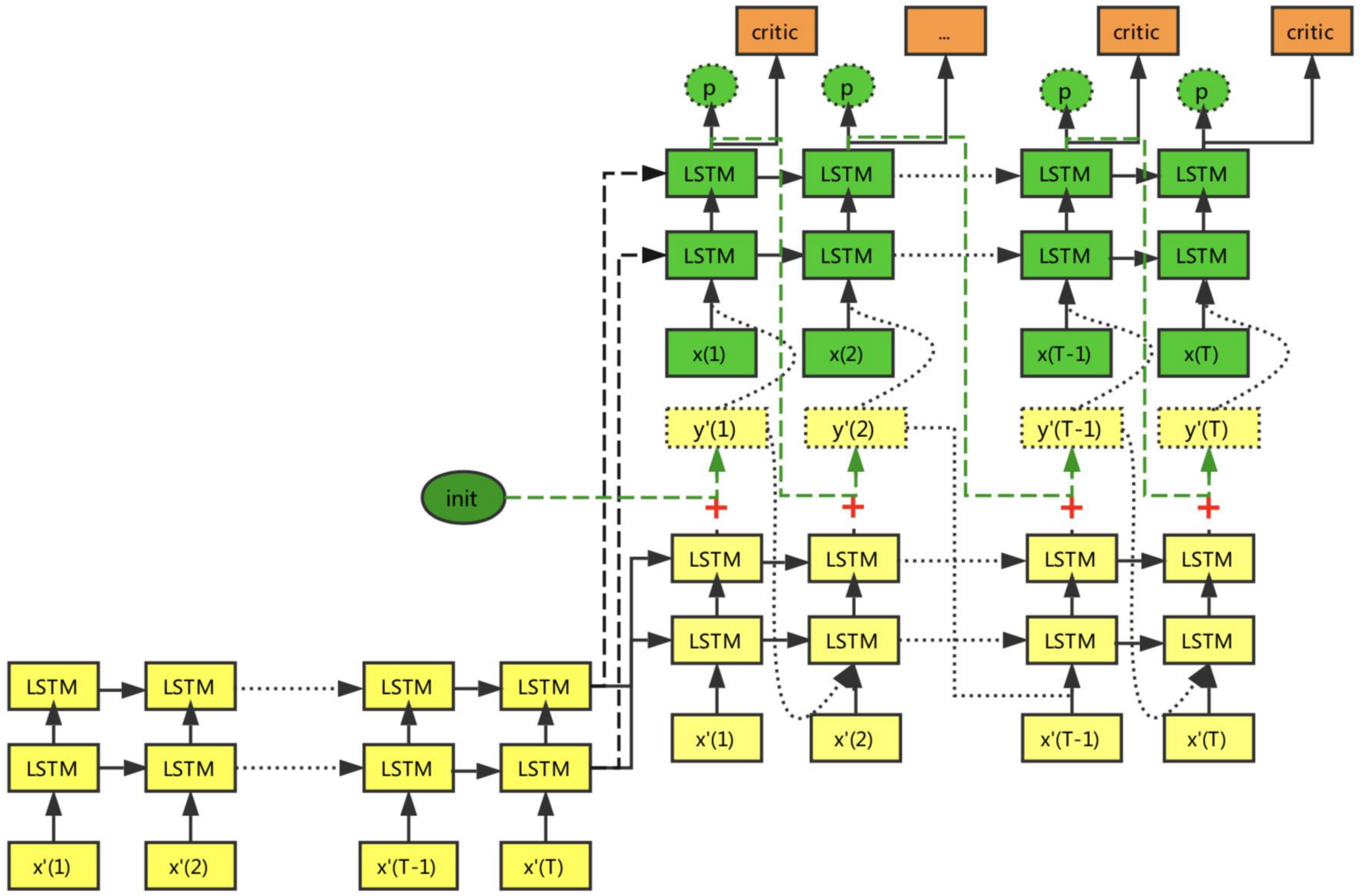
Pros of LeakGAN:

- Leaked information from Discriminator(more strong signal)
- Hierarchical reinforcement learning
- Interleaved Training

Pros of MaskGAN:

- In-filling task(masked LM)
- Actor-Critic architecture
- New evaluation metrics(unique n-grams)

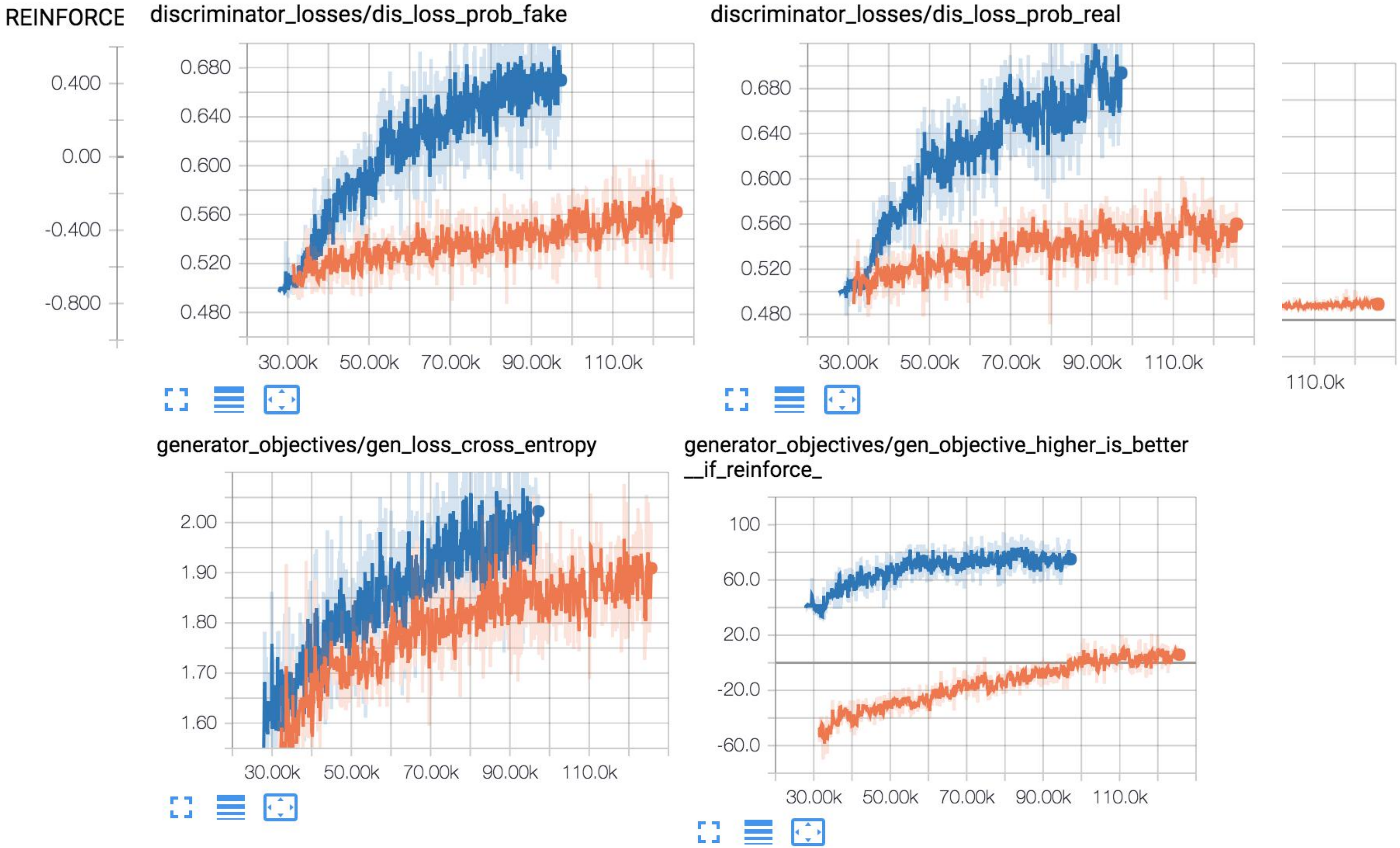
Text Generation with LeakMaskGAN



LeakMaskGAN vs MaskGAN(Car)

models\unique n-grams	% Unique bigrams	% Unique trigrams	% Unique quadgrams
MaskGAN	49.8	43.3	63.0
LeakMaskGAN	59.7	62.0	77.1

models\n-BLEU	2-gram BLEU	3-gram BLEU	4-gram BLEU	5-gram BLEU
MaskGAN	0.934	0.780	0.578	0.432
LeakMaskGAN	0.906	0.732	0.555	0.476


REINFORCE
0.400
0.00
-0.400
-0.800
discriminator_losses/dis_loss_prob_fake
0.680
0.640
0.600
0.560
0.520
0.480
30.00k 50.00k 70.00k 90.00k 110.0k
discriminator_losses/dis_loss_prob_real
0.680
0.640
0.600
0.560
0.520
0.480
30.00k 50.00k 70.00k 90.00k 110.0k
110.0k
generator_objectives/gen_loss_cross_entropy
2.00
1.90
1.80
1.70
1.60
30.00k 50.00k 70.00k 90.00k 110.0k
generator_objectives/gen_objective_higher_is_better
__if_reinforce_
100
60.0
20.0
-20.0
-60.0
30.00k 50.00k 70.00k 90.00k 110.0k

LeakMaskGAN vs MaskGAN(Car)

MaskGAN samples(partial):

1. 方向盘 很 轻 *, *底盘 *偏 *硬 *, *减震 *较 *硬 *, 硬 。 <eos>
2. 我 感觉 看 了 *这 *款 *车 *, *看 *着 *会 *挺 *顺眼 蛮 好 的 , 起码 里面 的 氛围 灯 我 很 喜欢 <eos>
3. 内饰 比较 普通 *, *很 *适合 *年轻人 *, *但是 *对 *整车 *的 , 这 款 车 的 内饰 用料 都 是 一些 硬 塑料 , 看上去 比较 简陋 。 <eos>
4. 一点 】 省 油 *, *觉得 *油耗 *低 *, *全景 *天窗 *! *关键 】 车 内 有 异响 <eos>
5. 遥控 钥匙 没有 用 *很多 *女儿 *就 *翻 *换 *到 *综合 *是 *这个 , 郁闷 死 了 ! <eos>
6. 内饰 做工 不错 , 德国车 *的 *<unk> *设计 *, *内饰 *做工 *也 *很 *精细 也 很 棒 , 比较 满意 <eos>
7. 方向盘 很 轻 *, *指 *哪 *打 *哪 *, *不 *有 *其他 *乘坐 *空间 *都 *可以 *了 *, *尤其 *是 *乘坐 *支撑 *不错 *, *特别 *是 *储物 *空间 *比较 *大 *, *后排 *空间 *很 *大 *, *后备箱 *也 *不 *小 *,
8. 我 感觉 看 了 *这 *款 *车 *, *有点 *好 *大家 *, *看 *着 *绝对 *适合 *厂家 *, *关键 *是 *led *好看 *, *能 *买 *红 *的 *原因 *, *, *就 *像 *前脸 *了 *, *没 *太 *多 *的 *、 *还有
9. 内饰 比较 普通 *, *设计 *也 *非常 *大气 *, *带 *起来 *很 *稳 *, *颜色 *就 *比较 *灵敏 *了 *, *座椅 *非常 *的 *舒服 *, *空间 *足够 *用 *, *空间 *大 *, *加速 *不错 *, *既 *也 *蛮 *轻松 *的 *,
10. 一点 】 省 油 *, *颜值 *高 *! *省 *油 *, *安全性 *和 *其他 *技术 *很 *可以 *, *百 *公里 *就 *1 *0 *个 *油 *。 *, *在 *磨合期 *才 *1 *0 *个 *左右 *, *有点 *高 *一些 *, *油耗 *较
11. 遥控 钥匙 没有 用 *那么 *舒适 *! *但是 *坐 *到 *2 *个 *人 *就 *上下班 *, *放 *一 *个 *打 *个 *问题 *, *就 *是 *后排 *空间 *较 *小 *, *副 *驾驶 *一般 *了 *, *乘坐 *空间 *还 *行 *, *前排

LeakMaskGAN samples(partial)：

1. 方向盘 很 轻 *, *操控性 *很 *好 *, *方向 *轻 *, *高速 硬 。 <eos>
2. 我 感觉 看 了 *很多 *车型 *, *还是 *经过 *丰田 *精彩 *的 *, 蛮 好 的 , 起码 里面 的 氛围 灯 我 很 喜欢 <eos>
3. 内饰 比较 普通 *, *这个 *价位 *的 *车 *购买 *, *大部分 *是 , 这 款 车 的 内饰 用料 都 是 一些 硬塑料 , 看上去 比较 简陋 。 <eos>
4. 一点 】 省 油 *, *省 *油 *, *【 *最 *不 *满意 *的 】 车 内 有 异响 <eos>
5. 遥控 钥匙 没有 用 *键 *启动 *, *提 *速 *快 *, *为什么 *三 , 郁闷 死 了 ! <eos>
6. 内饰 做工 不错 , 德国车 *的 *设计 *, *做工 *还 *可以 *, *中控台 *方面 也 很 棒 , 比较 满意 <eos>
7. 方向盘 很 轻 *, *转向 *挺 *准 *的 *, *转向 *精准 *, *配合 *得 *慢慢 *差点 *! *<eos>
8. 我 感觉 看 了 *福特 *车型 *, *总体 *来说 *年轻 *年轻 *运动 *, *越野 *性能 *相当 *好 *。 *<eos>
9. 内饰 比较 普通 *, *比较 *简单 *, *但是 *操控 *也 *该有 *的 *就 *是 *软包 *的 *, *硬质 *一点点 *件 *特别 *好 *, *行车 *电脑 *响 *维修 *, *跑 *高速 *就 *会 *有 *点 *疲软 *<eos>
10. 一点 】 省 油 *, *重要 *的 *事情 *说 *三 *遍 *! *【 *最 *不 *满意 *的 *一点 *】 *动力 *绝对 *够用 *! *1 *8 *0 *马力 *, *颜值 *高 *。 *6 *座 *内饰 *不 *协调 *, *<eos>
11. 遥控 钥匙 没有 用 *落 *合 *p *车窗 *, *在 *上坡 *上 *很 *顺畅 *。 *有点 *硬 *。 *全程 *状态 *少 *点 *。 *<eos>

Failure Modes

- Mode collapse
- Matching Syntax at boundaries
- Loss of global context

Trival Solutions

- Bi-directional RNN discriminator
- Bi-directional attention
- Global context information

Results(Car)

MaskGAN Samples(partial)

本人 对 油耗 *还是 *比较 *满意 *的 *， *这么 *大 *过程 *的， 一 公里 也 就 3 毛 多， 很 满意 油耗！ （备注 我 加 的 是 9 7 号 的 油） <eos>

一 档 怠速 1 *. *6 *t *的 *， *市区 *开车 *还 *可以 是 温柔 的 踏油 门， 没有 顿 挫感， <eos>

性价比 高， *配置 *高 *<eos>

我 都 是 走 超车道 *， *据说 *认可 *就 *算是 *我 *什么 *的 *讨厌 去 看 我 发 的 提车 作业 <eos>

的 一点 】 外观 *大气 *， *大 *灯 *流畅 *， *用料 *厚道 *， 晚上 的 白色 氛围 灯 好 浪漫， 高贵 <eos>

油耗 低， *座椅 *舒服 *。 *<eos>

一般 我 开 的 *1 *0 *0 *0 *公里 *， *早 *生 *9 6 0 的 速度， 正常 路面 是 感觉 很 好 的， 但是 <unk> 的 坡， 不 加 点 油 很 是 吃力 <eos>

外形 比较 符合 大众 这个 *价格 *。 *<eos>

比较 简约， 我 个人 *感觉 *不 *觉得 *很 *高档 *， *跟 *汽车 *的 车， 看 半天， 好多 按钮！ <eos>

起步 动力 偏 *肉 *， *起步 *慢 *一点 *<eos>

LeakMaskGAN Samples(partial)

本人 对 油耗 *不 *是 *很 *满意 *， *不过 *也 *不 *能， 一 公里 也 就 3 毛 多， 很 满意 油耗！ （备注 我 加 的 是 9 7 号 的 油） <eos>

一 档 怠速 1 *0 *0 *码 *时 *声音 *比较 *大 *， *音乐 是 温柔 的 踏 油门， 没有 顿 挫感， <eos>

性价比 高， *动力 *足 *<eos>

我 都 是 走 超车道 *的 *路面 *， *偶尔 *开 *的 *时候 *油耗 *有 去 看 我 发 的 提车 作业 <eos>

的 一点 】 外观 *漂亮 *， *外观 *漂亮 *， *内饰 *准确 *， *车漆 晚上 的 白色 氛围 灯 好 浪漫， 高贵 <eos>

油耗 低， *油耗 *低 *！ *<eos>

一般 我 开 的 *是 *比较 *温柔 *， *身材 *很 *瘦小 *， *我 6 0 的 速度， 正常 路面 是 感觉 很 好 的， 但是 <unk> 的 坡， 不 加 点 油 很 是 吃力 <eos>

外形 比较 符合 大众 这个 *级别 *的 *4 *骐达 *<eos>

比较 简约， 我 个人 *感觉 *不 *一样 *， *怕 *怎么 *说 *的 *， 车， 看 半天， 好多 按钮！ <eos>

起步 动力 偏 *弱 *， *但是 *家用 *够 *了 *。 *<eos>

NewMaskGAN Samples(partial)

本人 对 油耗 不怎么 关注， 但是 *省 *油 *， *油耗 *低 *， *省 *油 *， 就 3 毛 多， 很 满意 油耗！ （备注 我 加 的 是 9 7 号 的 油） <eos>

一 档 怠速 1 0 码， 一 档 换 *二档 *有 *一点点 *肉 *。 *没有 *顿 挫感 *， *没有 顿 挫感， <eos>

性价比 高， *配置 *高 *， *各 *方面 *都 *非常 *满意 *。 <eos>

我 都 是 走 超车道， 没 上牌 *过 *高速 *， *根本 *不 *敢 *买 *东 西 *， 发 的 提车 作业 <eos>

的 一点 】 外观， 中后排 的 舒适性 *， *大 *灯 *美观 *， *车灯 *设计 *不错 *， 灯 好 浪漫， 高贵 <eos>

油耗 低， 转向 底盘 *扎实 *， *油耗 *感人 *。 *<eos>

一般 我 开 *3 *1 *0 *0 *公里 *， *1 *0 *0 - 6 0 的 速度， 正常 路面 是 感觉 很 好 的， 但是 <unk> 的 坡， 不 加 点 油 很 是 吃力 <eos>

外形 比较 符合 大众 *口味 *， *质量 *也 *很 *出色 *<eos>

比较 简约， 我 个人 感觉 很快 就 熟悉 了 *， *本来 *买 *车 *的 *时候 *是 *导航 *， 好多 按钮！ <eos>

起步 动力 偏 弱， 对比 旧 途安 *的 *卡罗拉 *， *足够 *了 *<eos>

References

- 1. https://gitbook.cn/books/58ec9b969741d1032f26c300/index.html
- 2. Lu, Sidi, et al. "Neural Text Generation: Past, Present and Beyond." arXiv preprint arXiv:1803.07133 (2018).
- 3. Guo, Jiaxian, et al. "Long text generation via adversarial training with leaked information." arXiv preprint arXiv:1709.08624 (2017).
- 4. Fedus, William, Ian Goodfellow, and Andrew M. Dai. "Maskgan: Better text generation via filling in the _." arXiv preprint arXiv:1801.07736 (2018).
- 5. Huszár, Ferenc. How (not) to Train your Generative Model: Scheduled Sampling, Likelihood, Adversary?[J]. Computer Science, 2015.
- 6. Bengio, Samy, et al. "Scheduled sampling for sequence prediction with recurrent neural networks." Advances in Neural Information Processing Systems. 2015.

Q&A