DLHLP HW5-1

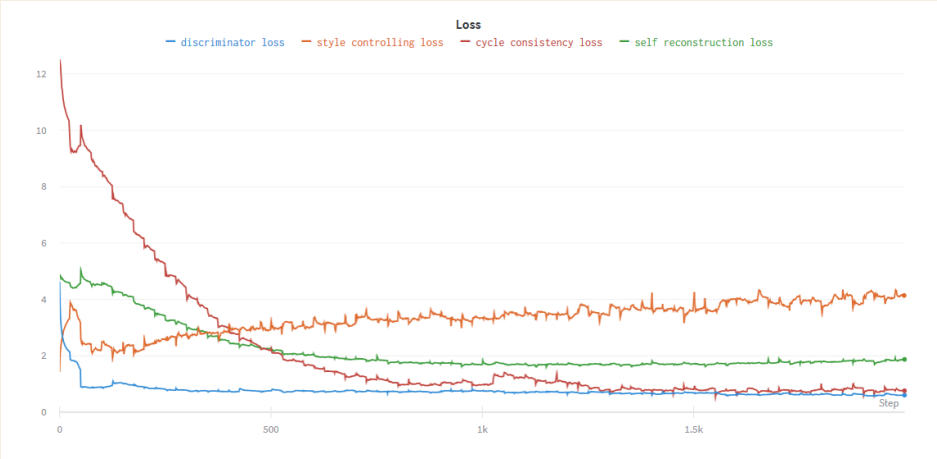
R08922067 鄭淵仁

Configuration

We changed embedding size to 512. Other settings follows the default.

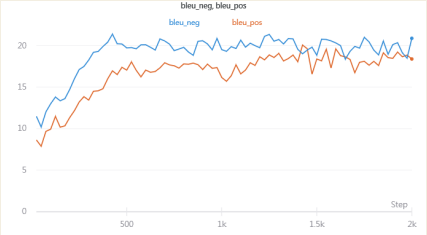
```
p main.py --do_train --use_wandb -embed_size 512
```

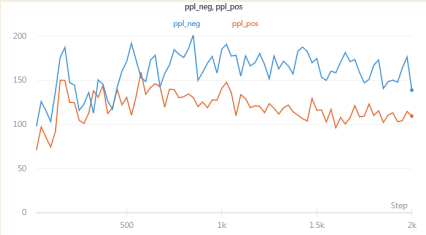
Loss


Loss
discriminator loss
style controlling loss
cycle consistency loss
self reconstruction loss
12
10
8
6
4
2
0
0
500
1k
1.5k
Step

Evaluation - Metrics

	pos	neg	avg
accuracy	0.92	0.788	0.854
ref-bleu	18.395	20.89	19.643
perplexity	109.569	138.972	124.27





Evaluation - Good Examples

```
******************** pos sample ********************
[gold] i was pleased with their service .
[raw ] i was pleased with their service .
[rev ] i was unhappy with their service .
```

“pleased” → “unhappy” is good.

```
******************** pos sample ********************
[gold] this was the best i have ever had !
[raw ] this was the best i have ever had !
[rev ] this was the worst i have ever had !
```

“the best” → “the worst” is good.

Evaluation - Problematic Examples

```
****************** pos sample ******************
[gold] well the food was great and the price of it was even better .
[raw ] well the food was great and the price of it was even better .
[rev ] well the food was horrible and the price of it was even better .
```

price should not be “better” after “even”, it should be some other adjectives, such as “even more expensive”.

```
****************** pos sample ******************
[gold] he is very thorough and genuinely cares for his customers .
[raw ] he is very thorough and genuinely cares for his customers .
[rev ] he is very unprofessional and black cares for his customers .
```

“black” can not be used to describe “care.”

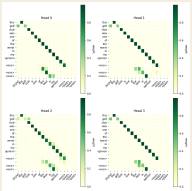
Observation

dimension 越大 performance 會越好，但是有時候也不會變好，要多跑幾次。

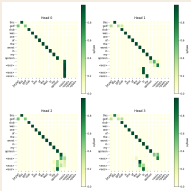
DLHLP HW5-2

R08922067 鄭淵仁

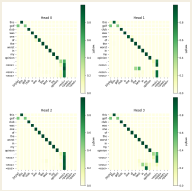
problem1



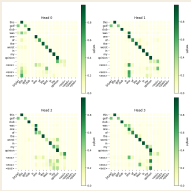
(a) layer 1



(b) layer 2



(c) layer 3



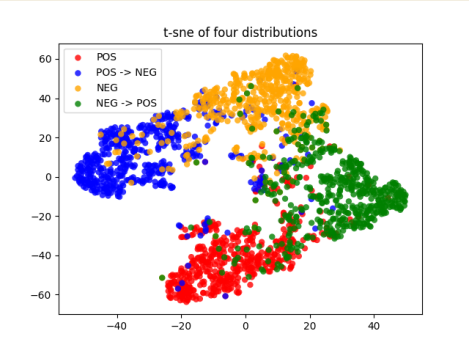
(d) layer 4

[gold] this golf club is one of the best in my opinion .
[rev] this golf club was one of the worst in my opinion .

在每一個 head 上，只有在輸出 golf 時會 attend 到 style。

我認為這可能是因為 model 會根據 “golf club” 這個主詞來找到相對應的形容詞，以便在後面輸出，所以才會在輸出 golf 時先 attend 在 style 上。

problem2



從圖中可以發現 POS、POS → NEG、NEG、NEG → POS 這四類句子在 t-sne 轉換之後，被大致分開到四個區域。

但是 POS → NEG 有一部分與 NEG 混在一起；NEG → POS 也有一部分與 POS 混在一起。這兩個部分很可能是 discriminator 無法分類清楚的部分。這也顯示出 generator 可以在一定程度上騙過 discriminator。

problem3

```
[ORG] the calzones are awesome and the lunch special they have is perfect .
[REV] the calzones are cold and the lunch special they have is flavorless .

[ORG] <unk> calzones are awesome and the lunch special they have is perfect .
[REV] the calzones are awesome and the lunch special they have is flavorless .

[ORG] the <unk> are awesome and the lunch special they have is perfect .
[REV] the food are cold and the lunch special they have is flavorless .

[ORG] the calzones <unk> awesome and the lunch special they have is perfect .
[REV] the calzones are cold and the lunch special they have is flavorless .

[ORG] the calzones are <unk> and the lunch special they have is perfect .
[REV] the calzones are rude and the lunch special they have is flavorless .

[ORG] the calzones are awesome <unk> the lunch special they have is perfect .
[REV] the calzones are awesome , the lunch special they have is flavorless .

[ORG] the calzones are awesome and <unk> lunch special they have is perfect .
[REV] the calzones are cold and the lunch special they have is flavorless .

[ORG] the calzones are awesome and the <unk> special they have is perfect .
[REV] the calzones are cold and the chicken special they have is flavorless .

[ORG] the calzones are awesome and the lunch <unk> they have is perfect .
[REV] the calzones are cold and the lunch and they have is flavorless .

[ORG] the calzones are awesome and the lunch special <unk> have is perfect .
[REV] the calzones are cold and the lunch special they have is flavorless .

[ORG] the calzones are awesome and the lunch special they <unk> is perfect .
[REV] the calzones are cold and the lunch special they are is flavorless .

[ORG] the calzones are awesome and the lunch special they have <unk> perfect .
[REV] the calzones are cold and the lunch special they have so flavorless .

[ORG] the calzones are awesome and the lunch special they have is <unk> .
[REV] the calzones are cold and the lunch special they have is terrible .
```

從圖中我發現當 mask 在名詞或形容詞周圍時，後面相關的形容詞會有變化。例如 “the calzones” 的 “the” 被 mask 掉後，“cold” 被變回 “awesome” 了；或是 “perfect” 被 mask 掉之後，“flavorless” 變成了 “terrible”。

我認為這是因為在 model 在 output 形容詞的時候，會參考前面的名詞跟現在的形容詞。因此當兩者其中一個有變化，形容詞就會跟著變動。

DLHLP HW5-3

Add More Styles

R08922067 鄭淵仁

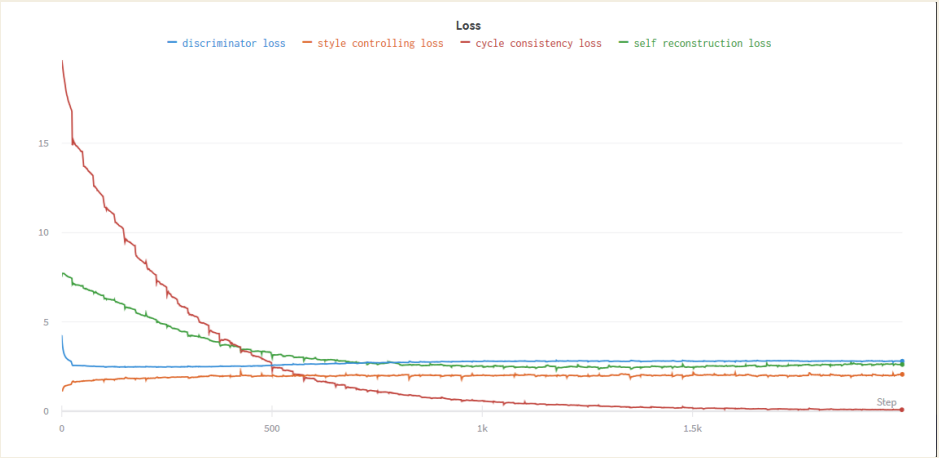
Configuration

We use the provided gender data. Other settings follows the default.

```
p main.py --do_train --use_wandb
```

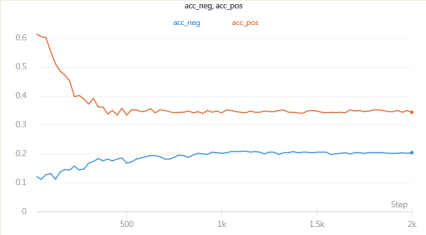
```
parser.add_argument("-data_path", default="./data/gender/"
```

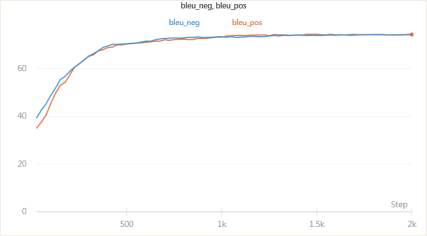
Loss

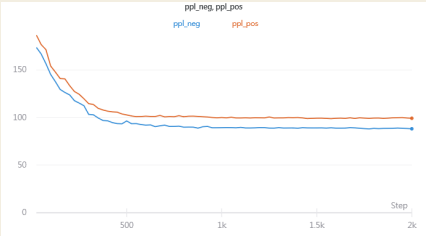

Loss
discriminator loss
style controlling loss
cycle consistency loss
self reconstruction loss
15
10
5
0
0
500
1k
1.5k
Step

Evaluation - Metric

	pos	neg	avg
accuracy	0.344	0.204	0.274
ref-bleu	74.425	74.208	74.3165
perplexity	99.224	88.175	93.695







Evaluation - Good Examples

```
******************** pos sample ********************
[gold] it ' s a disappointment that the service was subpar .
[raw ] it ' s a disappointment that the service was subpar .
[rev ] it ' s a disappointment that the service was subpar .
```

These sentences are exactly the same.

```
******************** pos sample ********************
[gold] as i said , my wife and i like to hang out towards the back
[raw ] as i said , my wife and i like to hang out towards the back
[rev ] as i said , my husband and i like to hang out towards the back
```

“wife” → “husband” means the gender of the speaker changes from male to female.

Evaluation - Problematic Examples

```
****************** pos sample ******************
[gold] i had the passion fruit sorbet cone , my wife went with her adventurous streak
[raw ] i had the passion fruit sorbet cone , my wife went with her adventurous streak
[rev ] i had the passion fruit sorbet cone , my husband went with her adventurous streak
```

“his” for husband, not “her”.

```
****************** pos sample ******************
[gold] my girlfriend ordered their special , a brisket sandwich , and i ordered 0 ribs
[raw ] my girlfriend ordered their special , a brisket sandwich , and i ordered 0 ribs
[rev ] my hubby ordered their special , a brisket sandwich , and i ordered 0 ribs
```

“grilfriend” vs. “boyfriend”. “wife” vs. “hubby”.

Observation

除了「丈夫」、「男朋友」等詞以外，其他詞幾乎都不會被變動。