

PHASE1

I. How to execute

1. Download the model from <https://drive.google.com/file/d/0B6z4WeJotpHgb2RjYWs1UHM1QUk/view?usp=sharing>
2. unzip and mkdir phase1/tmp_1
3. put the model and checkpoint files into tmp_1/ folder
4. run the script

II. How the code works

1. 修改tensorflow內建的rnn/translate/seq2seq_model.py裡的code。seq2seq model主要是先將原語言的句子input到encoder RNN變成fix length vector，再餵入decoder RNN部分並輸入翻譯語言的output。在這次作業我們使用multi-layers，所以下一層的RNN cell的input就是上一層的output。
2. 使用LSTM cell跟attention mechanism 確保整個句子的資訊可以被保留起來。
3. 結果：phase1 的training收斂比較慢，大約要train到兩萬多步perplexity才能降到2~3。若繼續train應該會再下降，但因為兩萬步就要train一天半，而且工作站很多人要用就沒有這麼嘗試了。

PHASE2

I. How to execute

1. Download the model from <https://drive.google.com/file/d/0B6z4WeJotpHgRmZZeVBGc0haTFE/view?usp=sharing>
2. unzip and mkdir phase2/tmp
3. put the model and checkpoint files into tmp/ folder
4. run the script

II. How the code works

1. 觀察：看了一些data發現，只要function是inform，回答的起手式都是XXX is a nice place，而request的開頭都是疑問詞，因此用排列feature vector的方式去train與翻譯的概念其實是相似的。
2. 讀入training data，parse每筆資料的function那行的parameter並存好對應的value，例如：inform(name='colibri mexican bistro';type=restaurant)，會分別被存成func=inform, para=["name", "type"], value=["colibri mexican bistro", "restaurant"]
3. 統計所有出現過的parameter，得到para_list.txt。並根據para_list，依序重新輸出training data成為以下格式：

func	para_list[0]=name	para_list[1]=type	para_list[2]=near	...
inform	'colibri mexican bistro'	restaurant	NULL	NULL
?confirm	NULL	NULL	west portal'	NULL
...

4. 將function那行得到的陣列當作train.en，下方的回答句子當作train.es，就可以視為翻譯的task，用phase1的方法去train model. train出來的正確率也有達到0.63。data如下圖所示。

```
1 inform 'colibri mexican bistro' NULL NULL NULL NULL NULL NULL NULL restaur
  NULL NULL NULL NULL
2 inform 'colibri mexican bistro' NULL NULL NULL NULL NULL NULL NULL restaur
  NULL NULL NULL NULL
3 ?request NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL food NULL NULL NULL
4 ?confirm NULL NULL NULL NULL NULL NULL 'west portal' NULL NULL NULL NULL N
  NULL
5 inform orexi NULL NULL NULL NULL NULL 'west portal' NULL NULL NULL NULL NU
  NULL
6 inform 'orexi' NULL 'hayes valley or west portal' '243 west portal avenue
  NULL NULL NULL NULL NULL NULL NULL NULL
7 inform orexi NULL NULL NULL NULL NULL NULL NULL NULL restaurant NULL NULL NULL
8 goodbye NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL
9 inform orexi NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL NULL
```

```
1 colibri mexican bistro is a nice restaurant
2 colibri mexican bistro is a nice restaurant
3 what kind of food would you like
4 did you say you are looking for a restaurant near west portal
5 orexi is a nice place it is near west portal
6 orexi is a nice place , it is in the area of west portal , their address is
  243 west portal avenue and it is in the area of hayes valley
7 orexi is a nice restaurant
8 goodbye
```

5. 結果：phase2 的training相較於phase1就收斂的快許多，可能是因為資料量比較小，還有training data相對單純許多。大約2000步就可以得到超過strong baseline的結果。