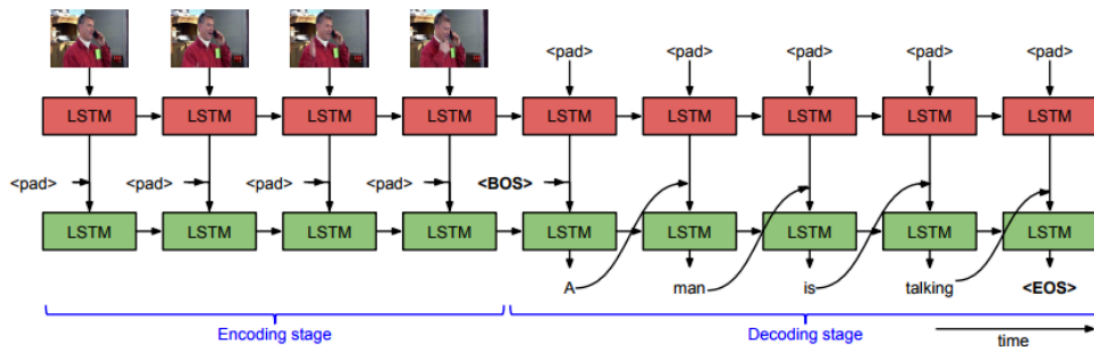


## 1. Model description (2%)

- Describe your seq2seq mod



- Training Epoch = 2000
- LSTM dimension = 1000
- Learning rate = 0.01
- vocab size = 6452
- AdamOptimizer

## 2. Attention mechanism(2%)

- How do you implement attention mechanism? (1%)

參考” Effective Approaches to Attention-based Neural Machine Translation” 修改原本的 model。match 的部分使用 dot 來算 alpha 值。

- Compare and analyze the results of models with and without attention mechanism. (1%)

從 blue 分數來看，沒有 attention 的 model 在 test 上拿到 0.679 分。

加了 attention 之後的 model 在 test 上拿到 0.716 分。但從產生的句子看不太出來 attention 有明顯的改善。

- How to improve your performance (1%)

加入 attention mechanism 和 schedule sampling 後 blue score 分數在 test 上有上升。有試著加入 attention regularization term 到 loss 中，但沒有獲得比較好的結果。

- Write down the method that makes you outstanding

Describe the model or technique (0.5%)

分數進步最多的部分是在加入 attention 這個部分，blue score 從原本的 0.679 提升到 0.716 分。

- Why do you use it (0.5%)

看完 paper 和聽老師上課的說明後覺得很厲害。

### 3.Experimental results and settings (1%)

- parameter tuning, schedule sampling ... etc

一開始使用的 LSTM dimension 為 1000，訓練的速度十分緩慢，後來降為 500 後訓練速度加快了訓練結果也差不多。

另外也有 使用 schedual sampling 的方式來做訓練，一開始選擇 ground truth 的機率為 1.0 之後每個 ephoch 乘以 0.99 得到新的機率，這樣的訓練方式可以解決 model 在 testing 的時候沒有 ground truth 的問題。從輸出結果來看，比起只拿 ground truth 來訓練 schedule sampling 訓練出來的 model 能夠產生更多樣化的句子而且在 test 上的表現也較好。