

Attention

□ Attention

- feed-input : 下一个单词知道上一个单词的 attention

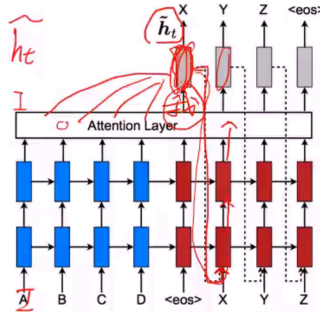


Figure from <https://arxiv.org/pdf/1508.04025.pdf>

Attention

□ Seq2Seq的技巧总结：

- dropout + reverse + attention + feed-input

System	Ppl	BLEU
Winning WMT'14 system – <i>phrase-based</i> + <i>large LM</i> (Buck et al., 2014)		20.7
Base	10.6	11.3
Base + <u>reverse</u>	9.9	12.6 (+1.3)
Base + reverse + <u>dropout</u>	8.1	14.0 (+1.4)
Base + reverse + dropout + global attention (<u>location</u>)	7.3	16.8 (+2.8)
Base + reverse + dropout + global attention (<u>location</u>) + <u>feed input</u>	6.4	18.1 (+1.3)

Ensemble

20.0

Figure from <https://arxiv.org/pdf/1508.04025.pdf>