Rethinking Translation Memory Augmented Neural Machine Translation

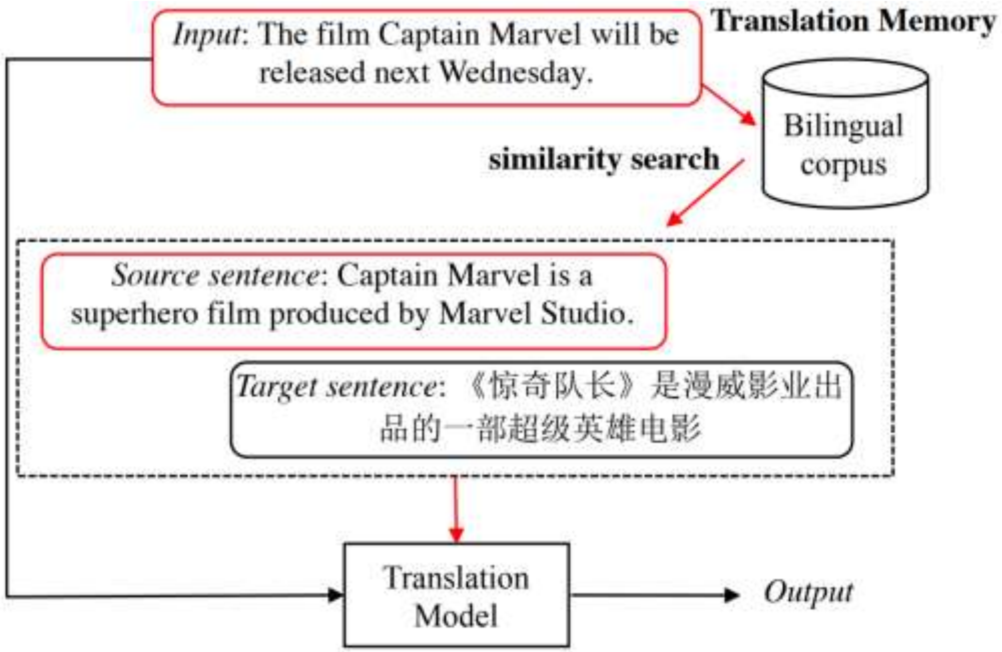
Presenter: Hongkun Hao

Date: 2023.6.24

Hongkun Hao[1*] **Guoping Huang**[2] **Lemao Liu**[2†]
Zhirui Zhang[2] **Shuming Shi**[2] **Rui Wang**[1†]
[1]Shanghai Jiao Tong University {haohongkun, wangrui12}@sjtu.edu.cn
[2]Tencent AI Lab {donkeyhuang, redmondliu, shumingshi}@tencent.com zrustc11@gmail.com

Background & Motivation

- Translation Memory (TM) augmented Neural Machine Translation (NMT) achieves gains over the vanilla NMT (without TM) under the conventional scenario.
- However, it fails to advance the vanilla NMT under a low-resource scenario.



Model	High-Resource	Low-Resource
w/o TM	60.83	54.54
w/ TM	63.76 ↑	53.92 ↓

Table 1: Testing BLEU comparison on JRC-Acquis German⇒English task. w/o TM and w/ TM denote the vanilla Transformer and TM-augmented Transformer, respectively; High-Resource and Low-Resource denote full and quarter train data are used for NMT training and TM retrieval.

Methods & Results

- TM-NMT can be viewed as a latent variable model. It has lower bias but higher variance.

Model	Single Encoder		Dual Encoder	
	Var	**Bias^2**	**Var**	**Bias^2**
w/o TM	**0.2088**	1.9519	**0.1573**	1.9992
w/ TM	0.2263	**1.7500**	0.2168	**1.8460**

Table 2: Estimated variance and bias on JRC-Acquis German⇒English task under single encoder and dual encoder backbone respectively.

Methods & Results

Model	Es⇒En	En⇒Es	De⇒En	En⇒De
w/o TM	58.44	56.11	54.54	49.97
TM-base	57.31	55.06	53.92	48.67
TM-single	57.56	55.24	54.03	48.82
TM-average	57.08	54.91	53.77	48.41
TM-weight	**59.14**	**56.53**	**55.36**	**50.51**

Table 5: Experimental results (test set BLEU scores) on four translation tasks of JRC-Acquis corpus under low-resource scenario.

- TM-single: only use top-1 TM.
- TM-average: average ensemble top-5 TMs.
- TM-weight: weighted ensemble top-5 TMs.