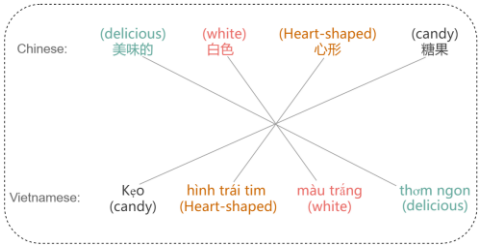
Linguistic Feature Template Integration for Chinese-Vietnamese Neural Machine Translation

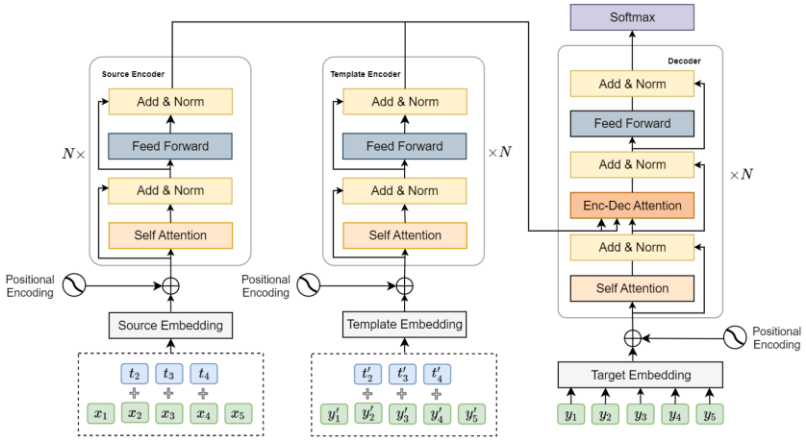
Zhiqiang YU, Yantuan XIAN, Zhengtao YU, Yuxin HUANG, Junjun GUO

Frontiers of Computer Science, DOI: 10.1007/s11704-022-1297-0

Problems & Ideas

- Problems of conventional stereo matching approaches:
  - previous template-based approaches only retrieve target sentence as template to generate translation, neglecting the utilization of linguistic feature in the low-resource language pair (e.g. Chinese-Vietnamese).
  - The existing Chinese-Vietnamese translation model can not integrate the Linguistic Feature Template.
- Ideas: 1) formalized represent the representative linguistic feature named modifier reverse in Chinese-Vietnamese as linguistic feature template and integrate it into well-designed template-based translation model.





Overview of our NMT framework. (t_2, t_3, t_4) and (t'_2, t'_3, t'_4) are corresponding modifier sequence for input sentence x and input template y'

Main Contributions

- Contributions:
  - We investigate a representative linguistic feature named modifier inverse in Chinese and Vietnamese in the perspective of linguistic difference and formalized represent it as modifier sequence.
  - We propose a novel translation framework which can integrate the modifier inverse into conventional template-based translation architecture.

Models	ALT	
	zh→vi	vi→zh
Moses	8.84	8.72
Transformer	8.51	8.44
SoftPrototype	9.72	9.37
Our approach (*noun*)	10.67	10.08
Our approach (*verb*)	**10.94**	**10.35**
Our approach (*adjective*)	10.78	10.23

Translation quality evaluation. noun,verb and adjective denote that noun, verb and adjective linguistic feature templates are adopted respectively.