

Multi-way MT

Firat et al. 2016 (two papers)

- Assume only many bilingual parallel corpora.
- For N languages: learn N encoders and N decoders.
- But what about attention?