

Sequence-to-sequence with attention

- Bahdanau et al. (2014)
- α_{ij} : weight of the j^{th} encoder state for the i^{th} decoder state
- no sentence embedding

