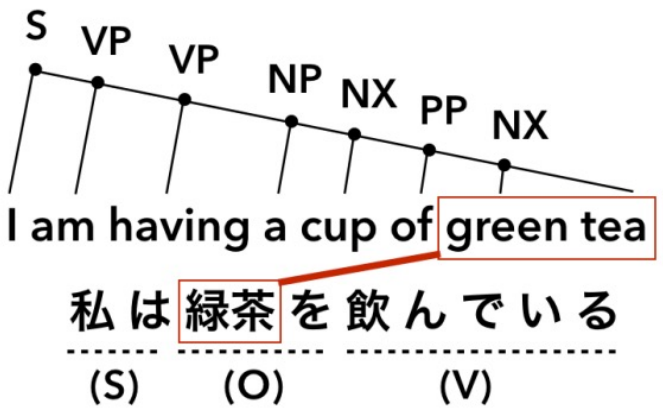
Tree Encoder

- Linguistic distance languages
  - Different syntax construction
  - Different lexical units such as words/phrase
  - E.g., if ‘緑茶’ only align with ‘green’ and ‘tea’,
  - Then ‘a cup of’ will align with ‘null’.



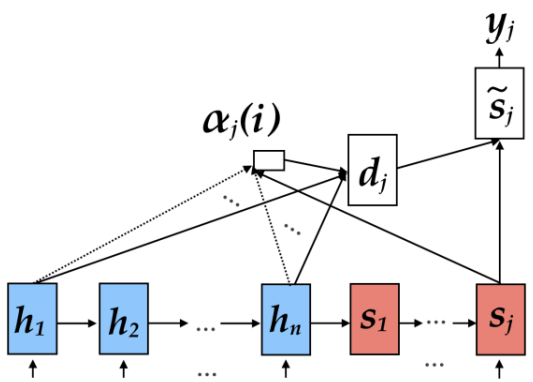


Figure 2: Attentional Encoder-Decoder model.

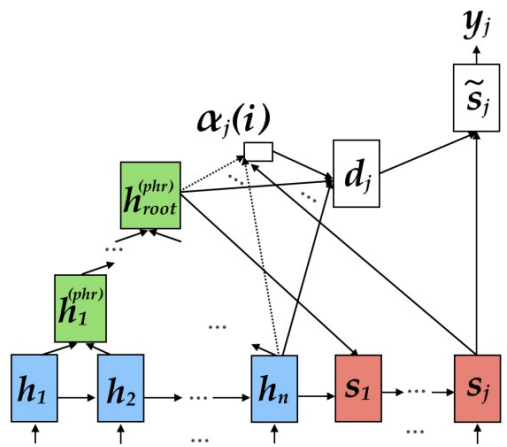


Figure 3: Proposed model: Tree-to-sequence attentional NMT model.

Tree-LSTM

$$\boldsymbol{h}_k^{(phr)} = f_{tree}(\boldsymbol{h}_k^l, \boldsymbol{h}_k^r),$$

$$\boldsymbol{s}_1 = g_{tree}(\boldsymbol{h}_n, \boldsymbol{h}_{root}^{(phr)}),$$

Context vector

$$\boldsymbol{d}_j = \sum_{i=1}^{n} \boldsymbol{\alpha}_j(i)\boldsymbol{h}_i + \sum_{i=n+1}^{2n-1} \boldsymbol{\alpha}_j(i)\boldsymbol{h}_i^{(phr)}.$$