Encoder-Decoder Model – Formal Notation

Data

input tokens (source language) $\mathbf{x} = (x_1, \ldots, x_{T_x})$

output tokens (target language) $\mathbf{y} = (y_1, \ldots, y_{T_y})$

Encoder

initial state $h_0 \equiv \mathbf{0}$

j-th state $h_j = \mathsf{RNN}_{\mathsf{enc}}(h_{j-1}, x_j) = \tanh(U_e h_{j-1} + W_e E_e x_j + b_e)$

final state h_{T_x}