# 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 $h_{T_x}$

**Decoder**

initial state $s_0 = h_{T_x}$

$i$-th decoder state $s_i = \mathsf{RNN}_{\mathsf{dec}}(s_{i-1}, \hat{y}_{i-1}) = \tanh(U_d s_{i-1} + W_d E_d \hat{y}_{i-1} + b_d)$

$i$-th word score $t_i = \tanh(U_o s_i + W_o E_d \hat{y}_{i-1} + b_o)$ ("output projection")

output $\hat{y}_i = \arg\max V_o t_i$