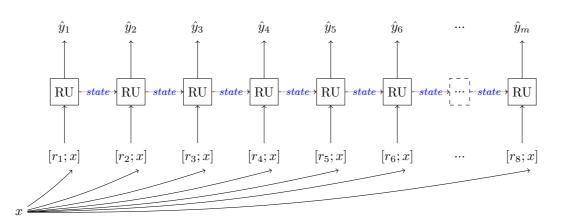


variable m outputs: $\hat{y_t}$ with prompts: r_t (often y_{t-1})

 x_1



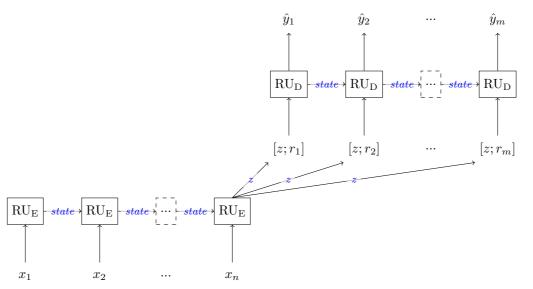
 x_5

 x_n

RU state RU state RU state RU state RU state

 x_3

 x_2



RNN Encoder-Decoder:

Variable n inputs: x_t Variable m outputs $\hat{y_t}$ Prompts: r_t (often y_{t-1})