

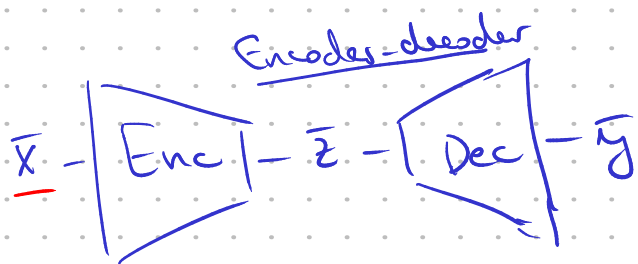
1) Policy gradient

$$\nabla_{\theta} J(\theta)$$

REINFORCE

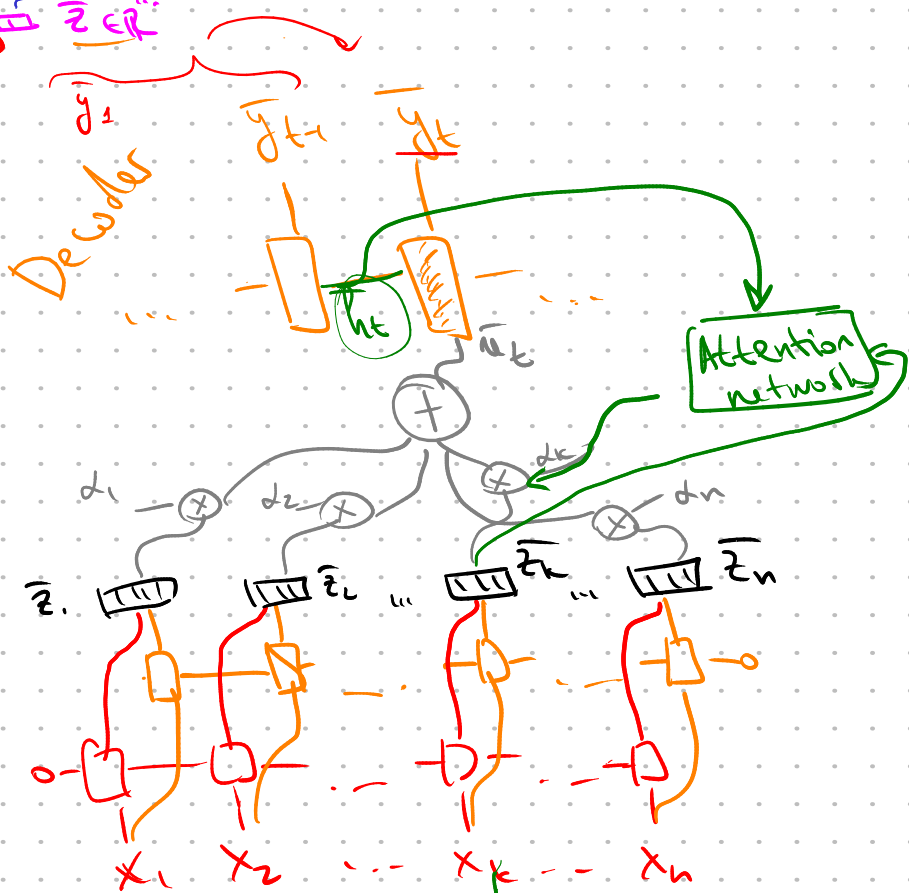
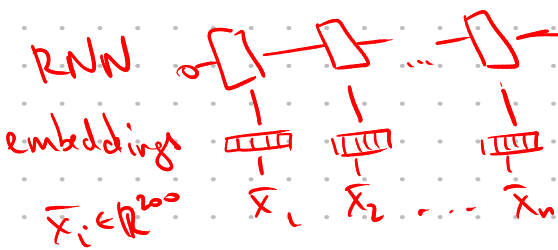
2) $p(\bar{x}) \approx q(\bar{x})$
 $p(\bar{x}) \approx q(\bar{x})$

Variational approx



- autoencoders
 $y \approx x$

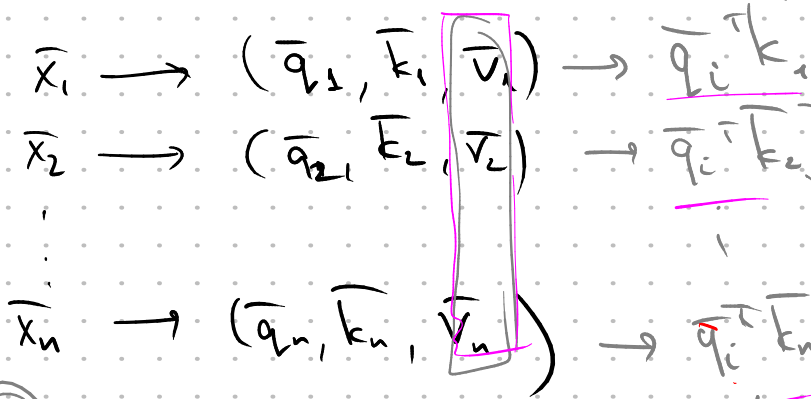
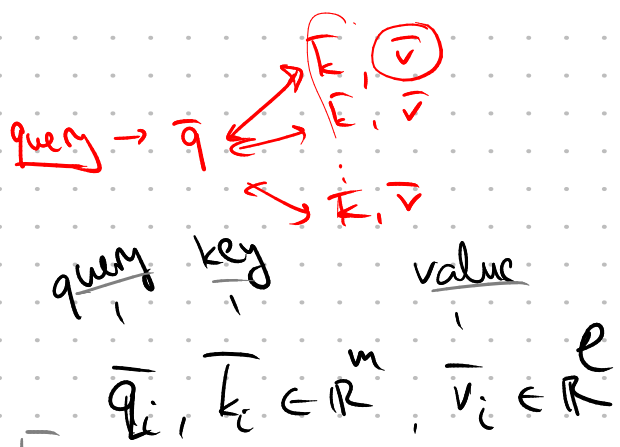
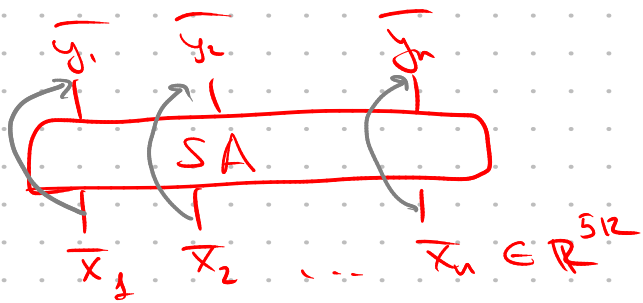
SMT:



Soft attention

$$u_t = \sum_{i=1}^n \alpha_i z_i$$

$$\alpha_k = \text{softmax}(\text{Att}(z_k, p_t))$$

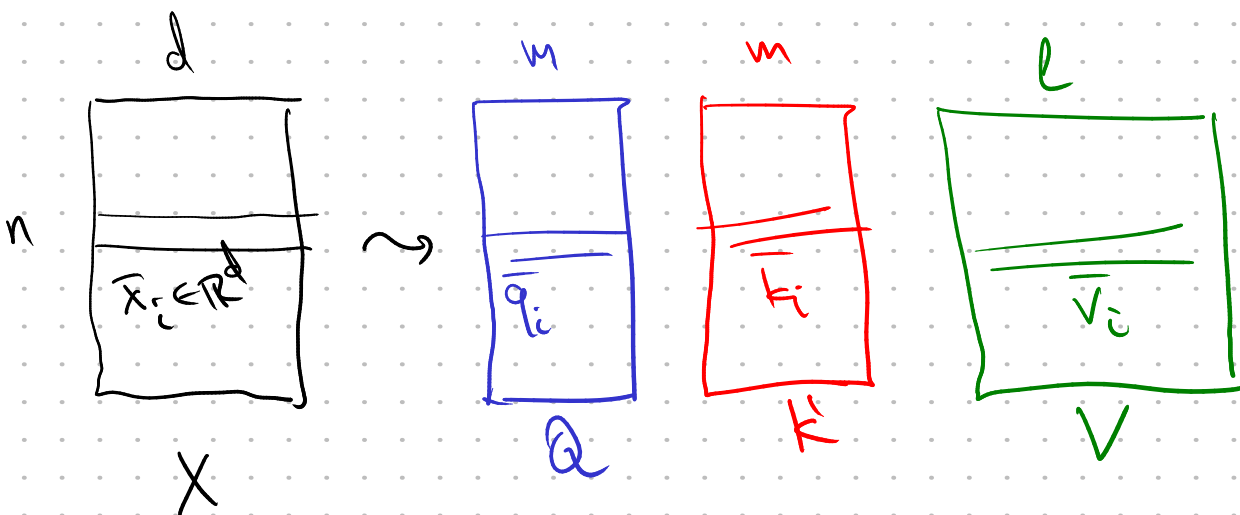


\bar{x}_i \rightarrow $(\bar{q}_i, \bar{k}_i, \bar{v}_i) \rightarrow \bar{q}_i^T \bar{k}_j$
 $\bar{x}_2 \rightarrow (\bar{q}_2, \bar{k}_2, \bar{v}_2) \rightarrow \bar{q}_i^T \bar{k}_2$
 \vdots
 $\bar{x}_n \rightarrow (\bar{q}_n, \bar{k}_n, \bar{v}_n) \rightarrow \bar{q}_i^T \bar{k}_n$

$\alpha_{ij} \approx \bar{q}_i^T \bar{k}_j$
 "how much attention \bar{x}_i pays to \bar{x}_j "

$$\alpha_{ij} = \text{softmax}\left(\frac{1}{\sqrt{m}} \bar{q}_i^T \bar{k}_j\right)$$

$$\bar{x}_i \rightarrow (\bar{q}_i, \bar{k}_i, \bar{v}_i) \rightarrow \bar{y}_i = \sum_{j=1}^n \alpha_{ij} \bar{v}_j$$



$$W_Q: \bar{q}_i = W_Q \bar{x}_i$$

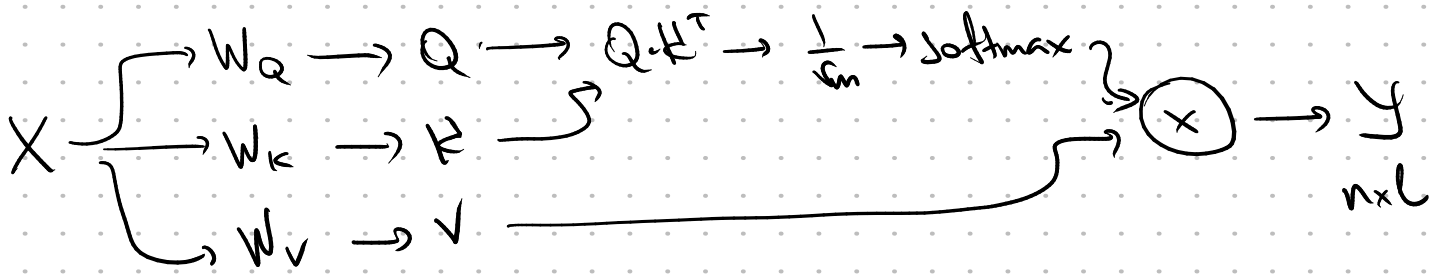
$$Q = W_Q X$$

$$W_K: \bar{k}_i = W_K \bar{x}_i$$

$$K = W_K X$$

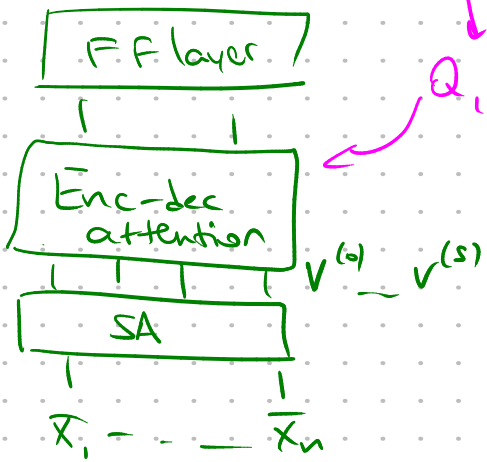
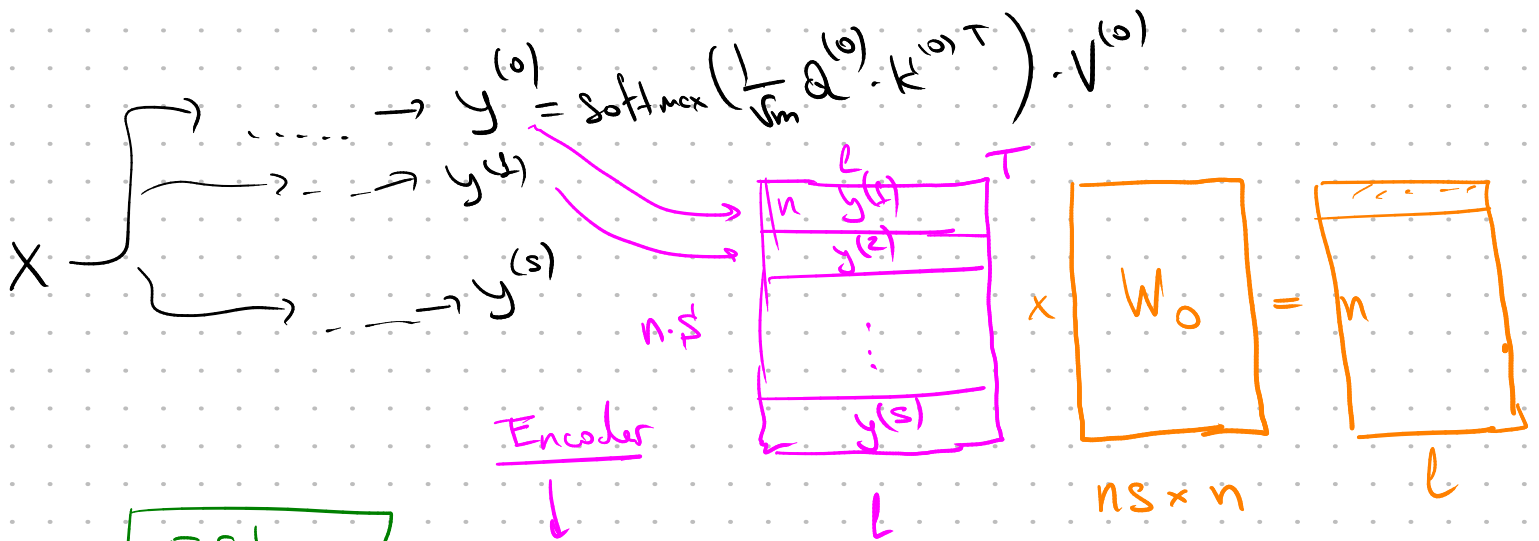
$$W_V: \bar{v}_i = W_V \bar{x}_i$$

$$V = W_V X$$



Multi-head attention

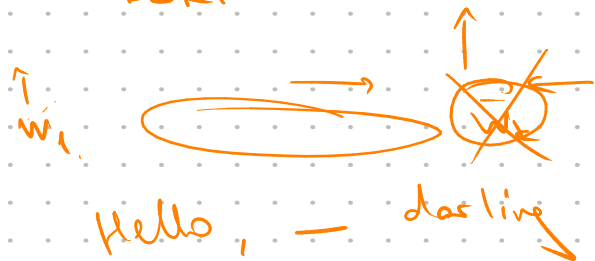
$(W_Q^{(1)}, W_K^{(1)}, W_V^{(1)}) \dots (W_Q^{(h)}, W_K^{(h)}, W_V^{(h)})$



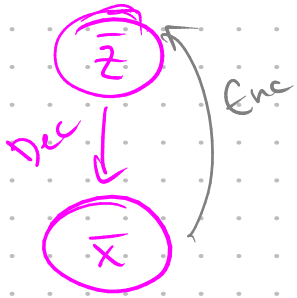
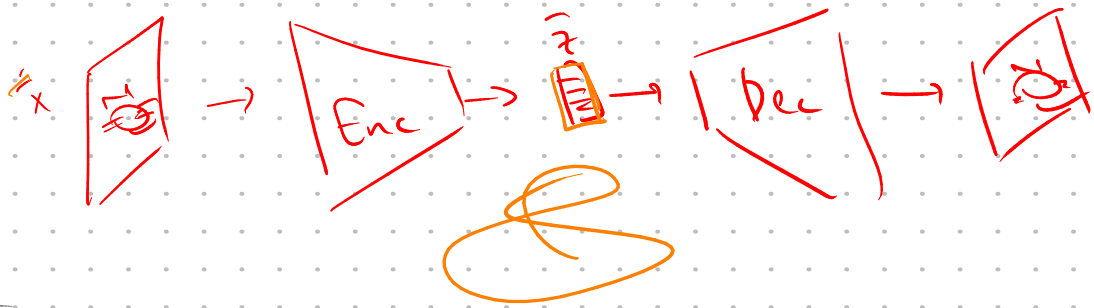
GPT



BERT



VAE



$$p(z|\bar{x}) = \frac{p(\bar{x}|z)p(z)}{p(\bar{x})}$$

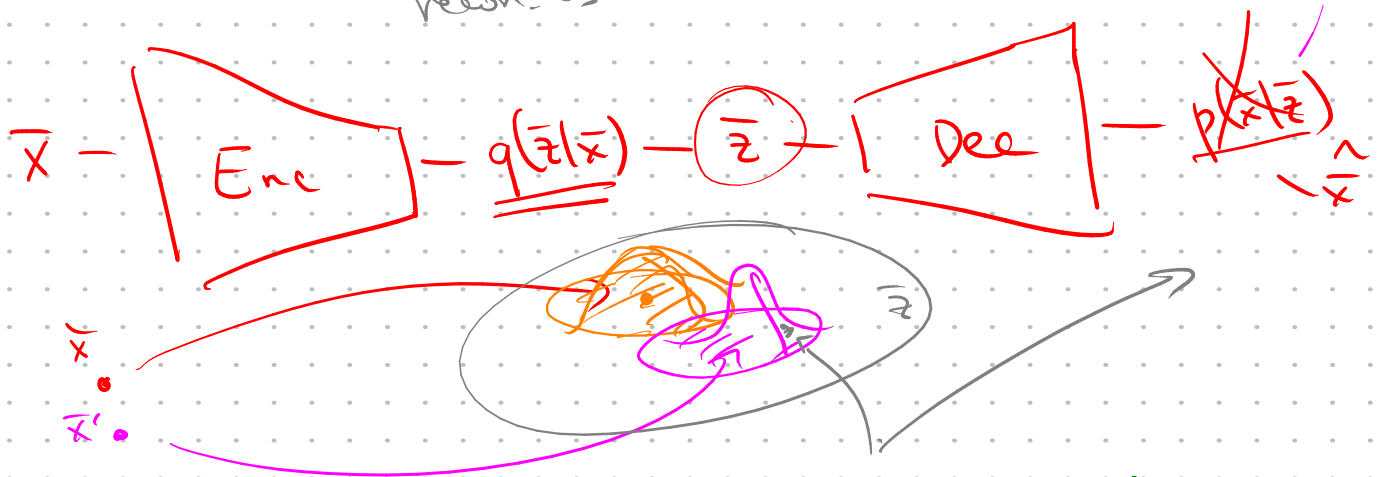
$$p(z|\bar{x}) \approx q(z|\bar{x})$$

$$\min_q KL(q||p)$$

$$\max_q \left[\mathbb{E}_{q(z|\bar{x})} [\log p(\bar{x}|z)] \right] - \text{KL}(q(z|\bar{x}) || p(z))$$

regularizer

recon. loss



$$q(z|\bar{x}) = \mathcal{N}(z | \bar{\mu}(\bar{x}), \bar{\sigma}(\bar{x}))$$

