Relational Reasoning in Multi-Agent Learning

Wenhao Li 06/13/2019

Interaction Networks for Learning about Objects, Relations and Physics

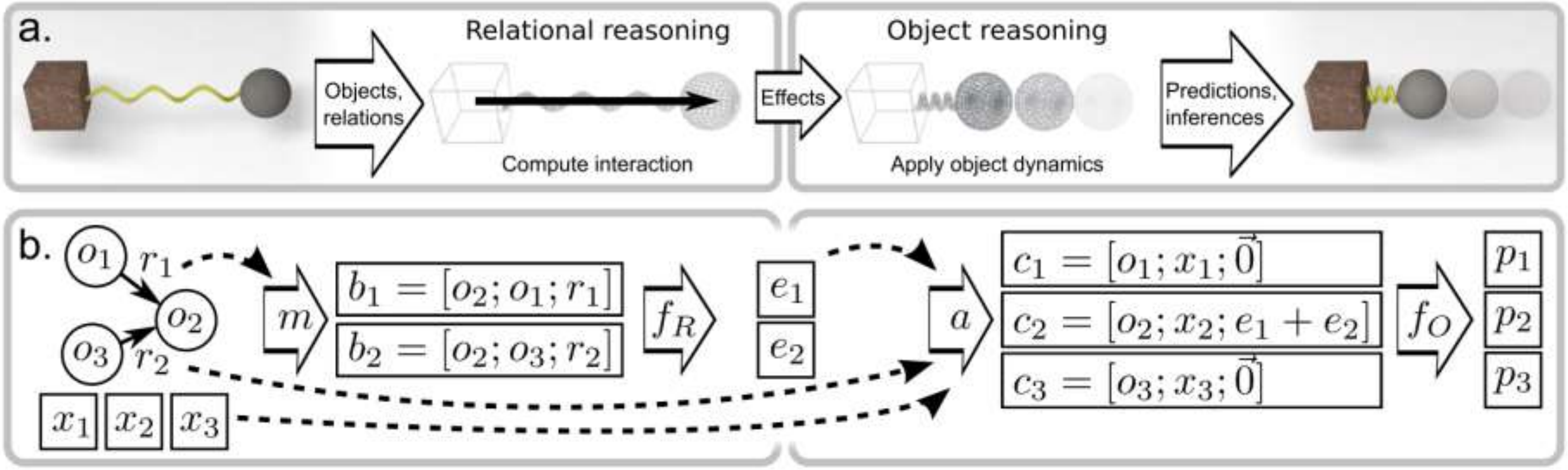


Figure 1: Schematic of an interaction network. *a.* For physical reasoning, the model takes objects and relations as input, reasons about their interactions, and applies the effects and physical dynamics to predict new states. *b.* For more complex systems, the model takes as input a graph that represents a system of objects, o_j, and relations, $\langle i, j, r_k \rangle_k$, instantiates the pairwise interaction terms, b_k, and computes their effects, e_k, via a relational model, $f_R(\cdot)$. The e_k are then aggregated and combined with the o_j and external effects, x_j, to generate input (as c_j), for an object model, $f_O(\cdot)$, which predicts how the interactions and dynamics influence the objects, p.

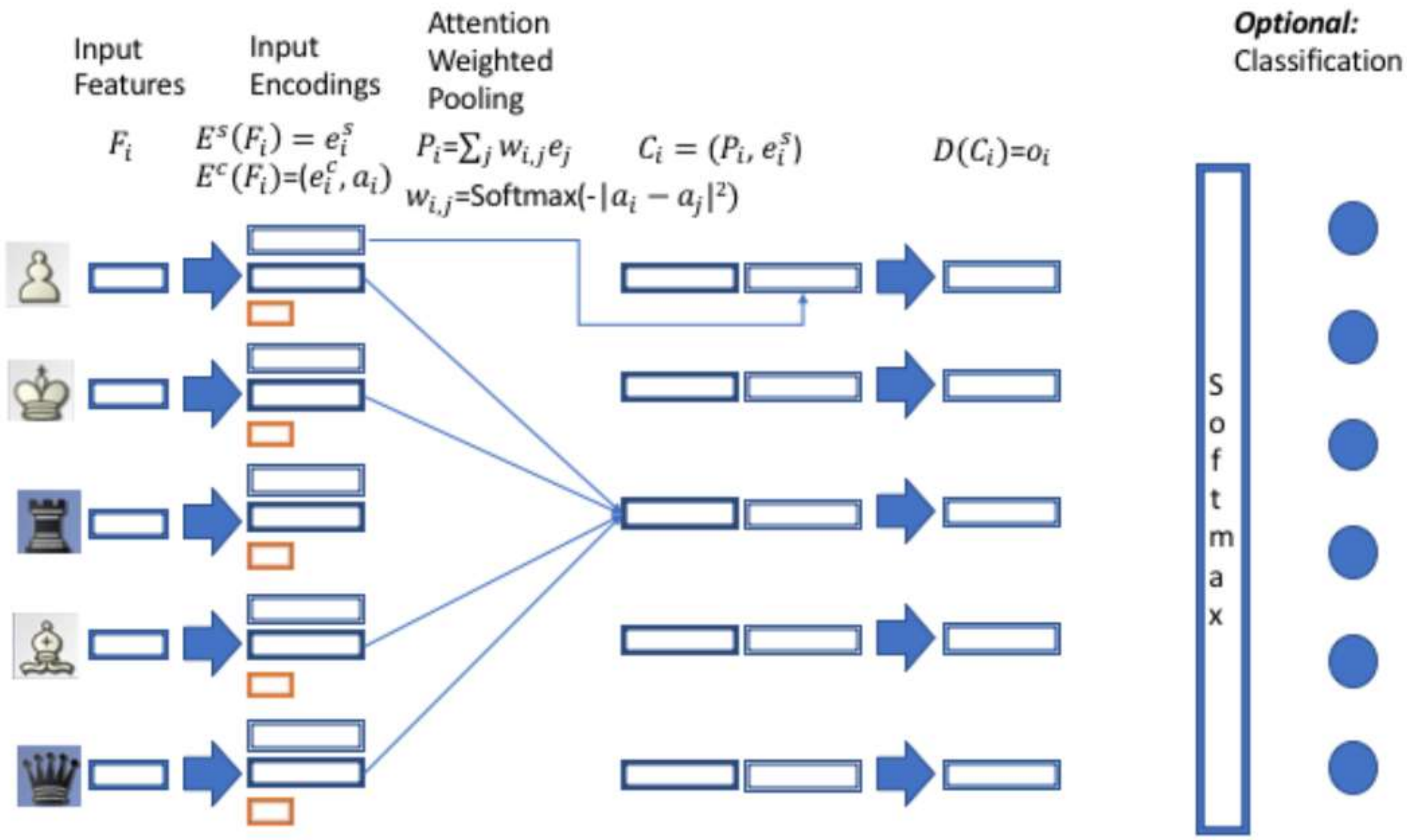
VAIN: Attentional Multi-agent Predictive Modeling

$$o_i = \theta(\sum_{j \neq i} \psi_{int}(x_i, x_j), \phi(x_i))$$

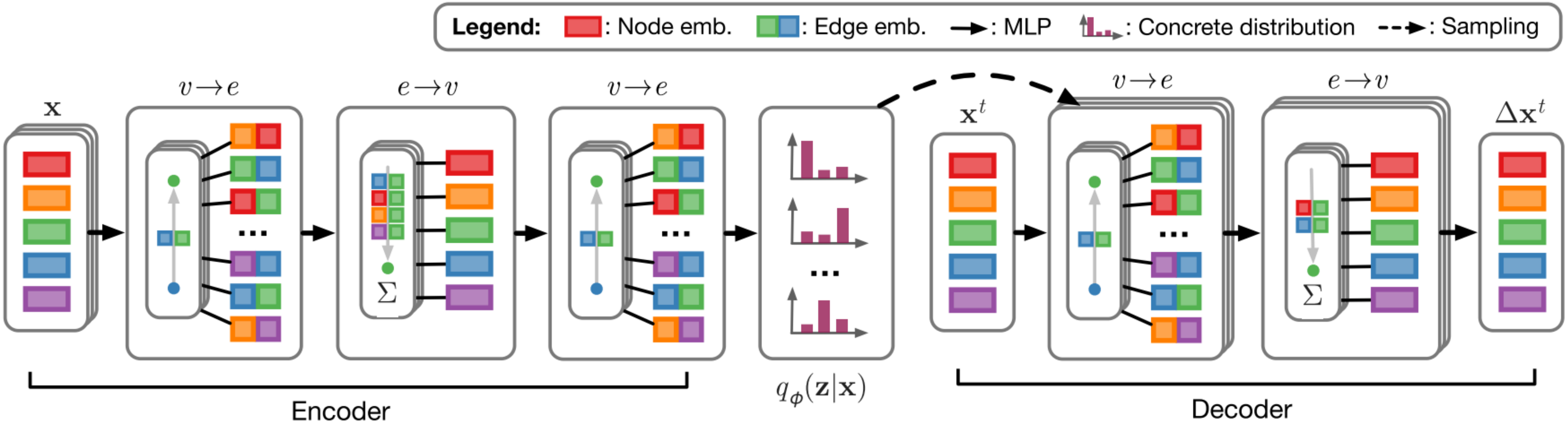
$$o_i = \theta(\sum_{j \neq i} e^{-\|a_i - a_j\|^2} \psi_{vain}(x_j), \phi(x_i))$$

$$K_{i,j} = e^{-\|a_i - a_j\|^2} / \sum_j e^{-\|a_i - a_j\|^2}$$

VAIN: Attentional Multi-agent Predictive Modeling



Neural relational inference for interacting systems



Neural relational inference for interacting systems

$$\begin{aligned}
\mathbf{h}_j^1 &= f_{\mathrm{emb}}\left(\mathbf{x}_j\right) \\
v \rightarrow e: \mathbf{h}_{(i,j)}^1 &= f_e^1\left(\left[\mathbf{h}_i^1, \mathbf{h}_j^1\right]\right) \\
e \rightarrow v: \mathbf{h}_j^2 &= f_v^1\left(\sum_{i \neq j} \mathbf{h}_{(i,j)}^1\right) \\
v \rightarrow e: \mathbf{h}_{(i,j)}^2 &= f_e^2\left(\left[\mathbf{h}_i^2, \mathbf{h}_j^2\right]\right)
\end{aligned}$$

$$\begin{aligned}
v \rightarrow e: \tilde{\mathbf{h}}_{(i,j)}^t &= \sum_k z_{ij,k} \tilde{f}_e^k\left(\left[\mathbf{x}_i^t, \mathbf{x}_j^t\right]\right) \\
e \rightarrow v: \boldsymbol{\mu}_j^{t+1} &= \mathbf{x}_j^t + \tilde{f}_v\left(\sum_{i \neq j} \tilde{\mathbf{h}}_{(i,j)}^t\right) \\
p\left(\mathbf{x}_j^{t+1} \mid \mathbf{x}^t, \mathbf{z}\right) &= \mathcal{N}\left(\boldsymbol{\mu}_j^{t+1}, \sigma^2 \mathbf{I}\right)
\end{aligned}$$

Neural relational inference for interacting systems

$$v \to e : \tilde{\mathbf{h}}_{(i,j)}^{t} = \sum_{k} z_{ij,k} \tilde{f}_{e}^{k} \left(\left[\mathbf{x}_{i}^{t}, \mathbf{x}_{j}^{t} \right] \right)$$

$$e \to v : \boldsymbol{\mu}_{j}^{t+1} = \mathbf{x}_{j}^{t} + \tilde{f}_{v} \left(\sum_{i \neq j} \tilde{\mathbf{h}}_{(i,j)}^{t} \right)$$

$$p \left(\mathbf{x}_{j}^{t+1} | \mathbf{x}^{t}, \mathbf{z} \right) = \mathcal{N} \left(\boldsymbol{\mu}_{j}^{t+1}, \sigma^{2} \mathbf{I} \right)$$

$$\begin{aligned}
\boldsymbol{\mu}_{j}^{2} &= f_{\text{dec}} \left(\mathbf{x}_{j}^{1} \right) \\
\boldsymbol{\mu}_{j}^{t+1} &= f_{\text{dec}} \left(\boldsymbol{\mu}_{j}^{t} \right) \quad t = 2, \ldots, M \\
\boldsymbol{\mu}_{j}^{M+2} &= f_{\text{dec}} \left(\mathbf{x}_{j}^{M+1} \right) \\
\boldsymbol{\mu}_{j}^{t+1} &= f_{\text{dec}} \left(\boldsymbol{\mu}_{j}^{t} \right) \quad t = M+2, \ldots, 2M \\
& \cdots
\end{aligned}$$

Neural relational inference for interacting systems

$$v \to e : \tilde{\mathbf{h}}_{(i,j)}^t = \sum_k z_{ij,k} \tilde{f}_e^k \left(\left[\tilde{\mathbf{h}}_i^t, \tilde{\mathbf{h}}_j^t \right] \right)$$

$$e \to v : \mathrm{MSG}_j^t = \sum_{i \neq j} \tilde{\mathbf{h}}_{(i,j)}^t$$

$$\tilde{\mathbf{h}}_j^{t+1} = \mathrm{GRU}\left(\left[\mathrm{MSG}_j^t, \mathbf{x}_j^t \right], \tilde{\mathbf{h}}_j^t \right)$$

$$\boldsymbol{\mu}_j^{t+1} = \mathbf{x}_j^t + f_{\mathrm{out}} \left(\tilde{\mathbf{h}}_j^{t+1} \right)$$

$$p\left(\mathbf{x}^{t+1} | \mathbf{x}^t, \mathbf{z} \right) = \mathcal{N}\left(\boldsymbol{\mu}^{t+1}, \sigma^2 \mathbf{I} \right)$$

$$\boldsymbol{\mu}_j^2 = f_{\mathrm{dec}} \left(\mathbf{x}_j^1 \right)$$

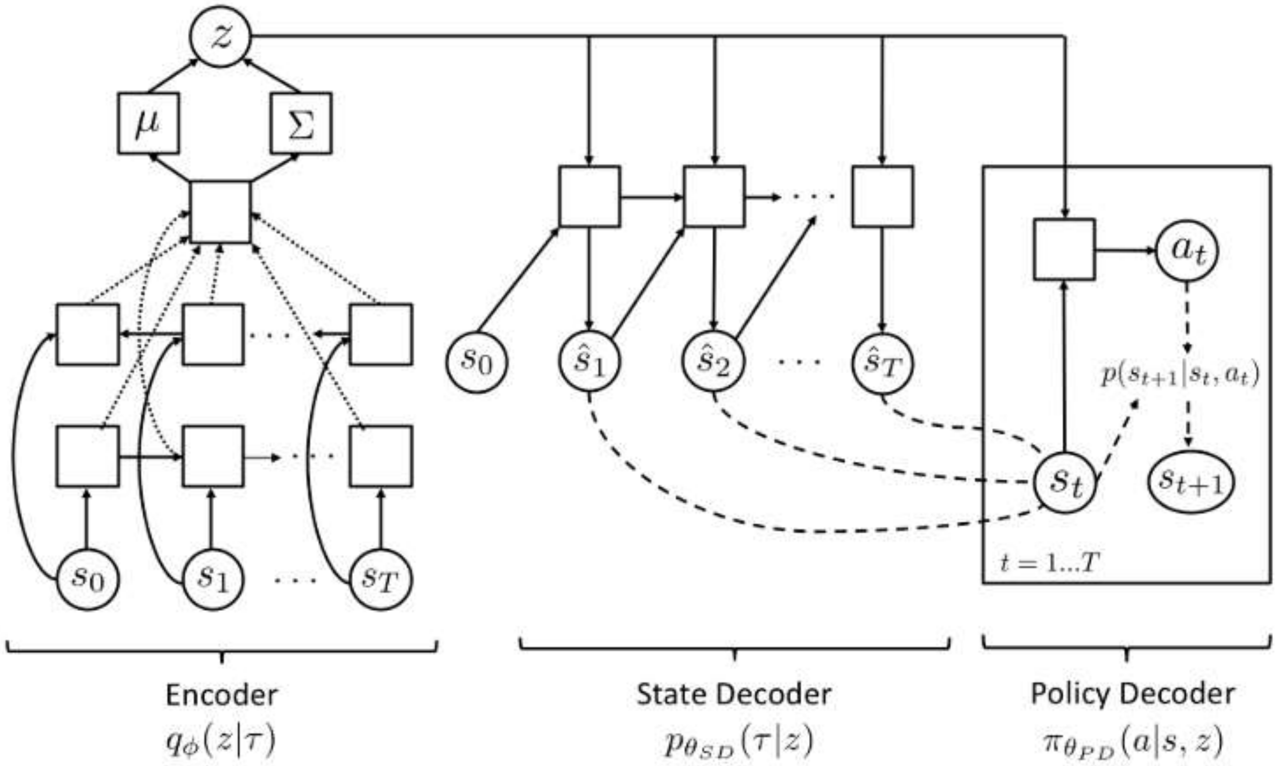
$$\boldsymbol{\mu}_j^{t+1} = f_{\mathrm{dec}} \left(\boldsymbol{\mu}_j^t \right) \quad t = 2, \ldots, M$$

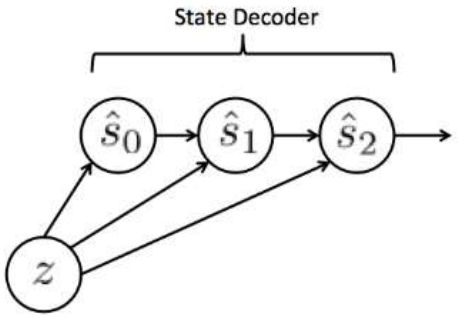
$$\boldsymbol{\mu}_j^{M+2} = f_{\mathrm{dec}} \left(\mathbf{x}_j^{M+1} \right)$$

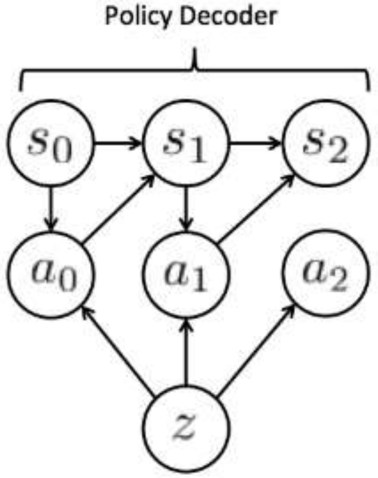
$$\boldsymbol{\mu}_j^{t+1} = f_{\mathrm{dec}} \left(\boldsymbol{\mu}_j^t \right) \quad t = M+2, \ldots, 2M$$

$$\cdots$$

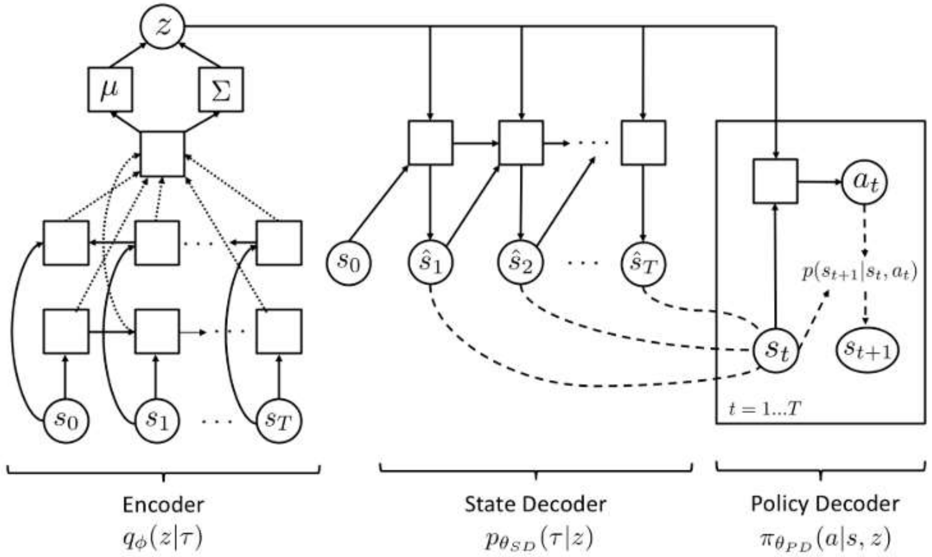
Self-consistent trajectory autoencoder: Hierarchical reinforcement learning with trajectory embeddings







Self-consistent trajectory autoencoder: Hierarchical reinforcement learning with trajectory embeddings



$$
\begin{aligned}
&\max && \log p(\tau) \\
&\text{subject to} && \mathbb{E}_{q_\phi}[D_{KL}(p_{\theta_{PD}}(\tau \mid z) \,\|\, p_{\theta_{SD}}(\tau \mid z))] = 0
\end{aligned}
$$

$$
\max_{\theta_{SD},\theta_{PD},\phi} \log p(\tau) - \lambda \mathbb{E}_{q_\phi}[D_{KL}(p_{\theta_{PD}}(\tau \mid z) \,\|\, p_{\theta_{SD}}(\tau \mid z))]
$$

$$
\begin{aligned}
&\log p(\tau) - \lambda \mathbb{E}_{q_\phi}[D_{KL}(p_{\theta_{PD}}(\tau \mid z) \,\|\, p_{\theta_{SD}}(\tau \mid z)] \\
&\quad \geq \mathbb{E}_{q_\phi}[\log p_{\theta_{SD}}(\tau \mid z))] - D_{KL}(q_\phi(z \mid \tau) \,\|\, p(z)) + \\
&\quad \lambda \left[\mathbb{E}_{q_\phi, p_{\theta_{PD}}(\tau|z)}[\log p_{\theta_{SD}}(\tau \mid z)] + \mathcal{H}(p_{\theta_{PD}}(\tau \mid z))\right]
\end{aligned}
$$

$$
\mathbb{E}_q[\log p_{\theta_{SD}}(\tau \mid z)] + \lambda \mathbb{E}_{q, p_{\theta_{PD}}(\tau'|z)}[\log p_{\theta_{SD}}(\tau' \mid z)].
$$

$$
\lambda \left[\mathbb{E}_{q, p_{\theta_{PD}}(\tau'|z)}[\log p_{\theta_{SD}}(\tau' \mid z)] + \mathcal{H}(p_{\theta_{PD}}(\tau \mid z))\right].
$$

Relational Reasoning in Multi-Agent Learning

Interaction Networks for Learning about Objects, Relations and Physics	IN
VAIN: Attentional Multi-agent Predictive Modeling	Attention IN
Neural relational inference for interacting systems	Graph IN
Self-consistent trajectory autoencoder: Hierarchical reinforcement learning with trajectory embeddings	Self-Consistent IN