FCAI fcai.fi

Background

Representation learning for RL

 $z_t = e_{\theta}(o_t)$

Encoder

Dynamics $\hat{z}_{t+1} = z_t + d_{\phi}(z_t, a_t)$

 $\hat{r}_{t+1} = r_{\phi}(z_t, a_t)$

Reward

 $q_t = Q_{\psi}(z_t, a_t)$

Critic

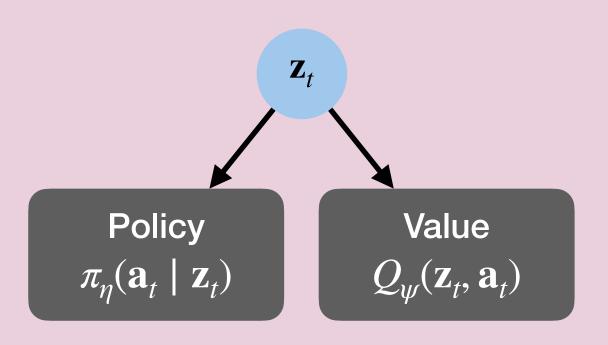
 $a_t \sim \pi_{\eta}(a_t \mid z_t)$

Policy

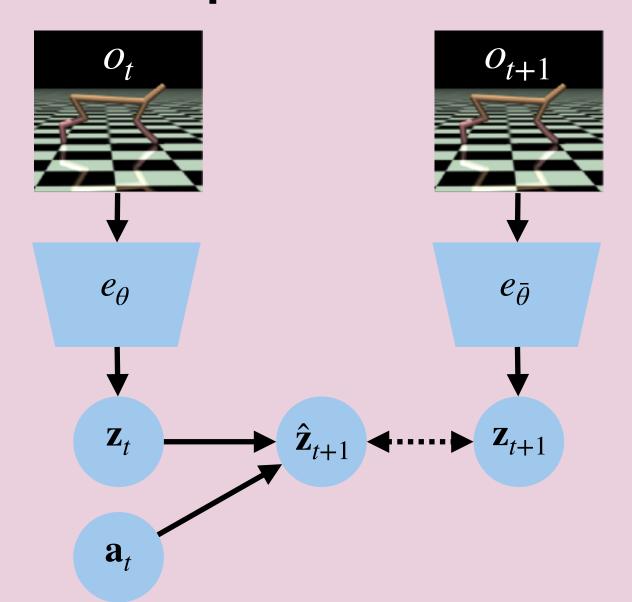
Latent-state consistency loss (representation learning)

 $\arg\min_{\theta,\phi} \sum_{h=t}^{t+H} \gamma^h \left(\frac{z_h + d_{\phi}(e_{\theta}(o_h), a_h)}{\|z_h + d_{\phi}(e_{\theta}(o_h), a_h)\|_2} \right)^{-1} \left(\frac{e_{\bar{\theta}}(o_{h+1})}{e_{\bar{\theta}}(o_{h+1})} \right) + \left\| r_{\phi}(e_{\theta}(o_h), a_h) - r_{h+1} \right\|_2^2$

2. Latent actor-critic



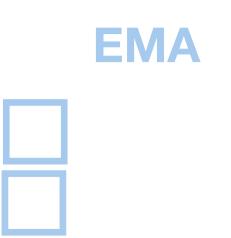
1. Learn representation



Zhao et al. (2023). Simplified Temporal Consistency Reinforcement Learning. ICML.

Representation is task specific!

Latent-state consistency with cosine similarity



Background

Representation learning for RL

Encoder
$$z_t = e_{\theta}(o_t)$$

Dynamics
$$\hat{z}_{t+1} = z_t + d_{\phi}(z_t, a_t)$$

Reward
$$\hat{r}_{t+1} = r_{\phi}(z_t, a_t)$$

Critic
$$q_t = Q_{\psi}(z_t, a_t)$$

Policy
$$a_t \sim \pi_{\eta}(a_t \mid z_t)$$

Latent-state consistency loss (representation learning)

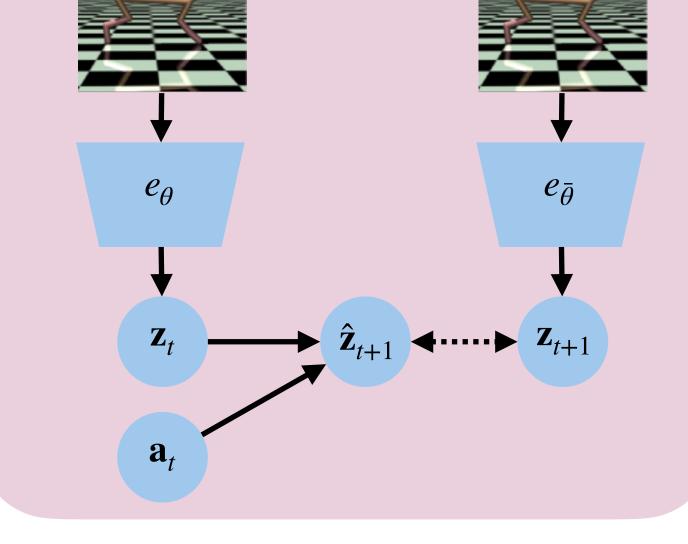
$$\arg\min_{\theta,\phi} \sum_{h=t}^{t+H} \gamma^{h} \left(\frac{z_{h} + d_{\phi}(e_{\theta}(o_{h}), a_{h})}{\|z_{h} + d_{\phi}(e_{\theta}(o_{h}), a_{h})\|_{2}} \right)^{\mathsf{T}} \left(\frac{e_{\bar{\theta}}(o_{h+1})}{\|e_{\bar{\theta}}(o_{h+1})\|} \right) + \left\| r_{\phi}(e_{\theta}(o_{h}), a_{h}) - r_{h+1} \right\|_{2}^{2}$$

.atent-state consistency with cosine similarity

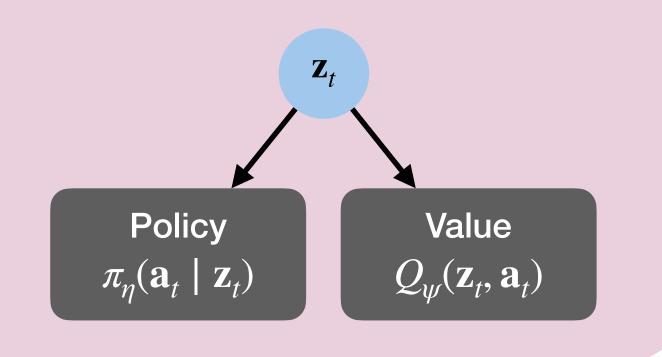
Representation is task specific!

$$\| r_{\phi}(e_{\theta}(o_h), a_h) - r_{h+1} \|_{2}^{2}$$

1. Learn representation



2. Latent actor-critic





Zhao et al. (2023). Simplified Temporal Consistency Reinforcement Learning. ICML.

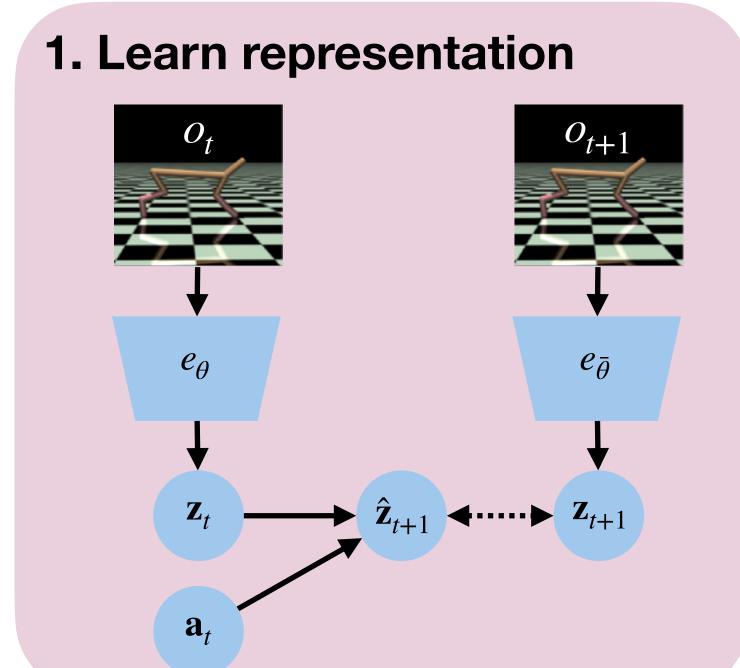
Background

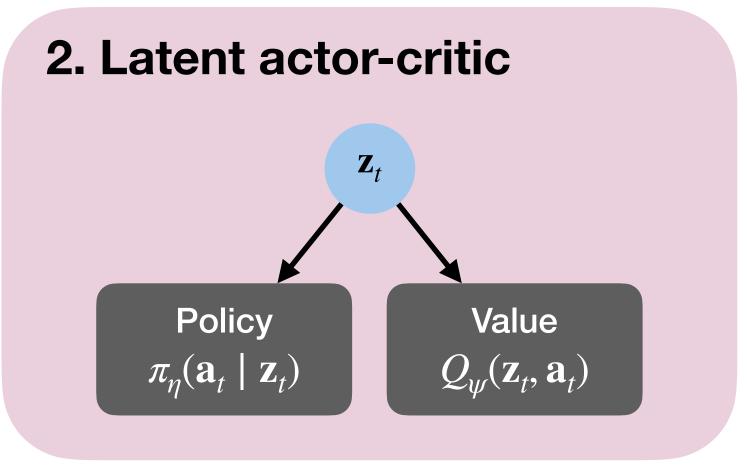
Task-agnostic representations for RL

Temporal Consistency RL (TCRL)

$$\arg\min_{\theta,\phi} \sum_{h=t}^{t+H} \gamma^h \left(\frac{z_h + d_{\phi}(e_{\theta}(o_h), a_h)}{\|z_h + d_{\phi}(e_{\theta}(o_h), a_h)\|_2} \right)^{\mathsf{T}} \left(\frac{e_{\bar{\theta}}(o_{h+1})}{\|e_{\bar{\theta}}(o_{h+1})\|_2} \right) + \left\| r_{\phi}(e_{\theta}(o_h), a_h) - r_{h+1} \right\|_2^2$$

Representation is task specific!





FCAI

fcai.fi