FCAI fcai.fi

DCWM: Components

 $\mathcal{L}(\theta, \phi, \xi; \mathcal{D}) = \mathbb{E}_{(\mathbf{o}, \mathbf{a}, \mathbf{o}', r)_{0:H} \sim \mathcal{D}} \left[\sum_{h=0}^{H-1} \gamma^h \left(\text{CE}(p_{\phi}(\hat{\mathbf{c}}_{h+1} \mid \hat{\mathbf{c}}_h, \mathbf{a}_h), \mathbf{c}_{h+1}) + \|R_{\xi}(\mathbf{c}_h, \mathbf{a}_h) - r_h\|_2^2 \right) \right]$

$\mathbf{x}_t = e_{\theta}(\mathbf{s}_t) \in \mathbb{R}^{d \times b}$

Encoder

```
Latent quantization \mathbf{c}_t = f(\mathbf{x}_t) \in \mathscr{C}
```

 $\hat{\mathbf{c}}_{t+1} \sim \text{Categorical}(p_1, ..., p_{|\mathcal{C}|})$ with $p_i = P_{\phi}(\mathbf{c}_{t+1} = \mathbf{c}^{(i)} \mid \mathbf{c}_t, \mathbf{a}_t)$

Dynamics

 $\hat{r}_{t+1} = R_{\xi}(\mathbf{c}_t, \mathbf{a}_t)$

Reward

World model loss

DCWM: Components

Encoder

$$\mathbf{x}_t = e_{\theta}(\mathbf{s}_t) \in \mathbb{R}^{d \times b}$$

Latent quantization $\mathbf{c}_t = f(\mathbf{x}_t) \in \mathscr{C}$

$$\mathbf{c}_t = f(\mathbf{x}_t) \in \mathscr{C}$$

Dynamics

$$\hat{\mathbf{c}}_{t+1} \sim \text{Categorical}(p_1, ..., p_{|\mathcal{C}|})$$
 with $p_i = P_{\phi}(\mathbf{c}_{t+1} = \mathbf{c}^{(i)} \mid \mathbf{c}_t, \mathbf{a}_t)$

Reward

$$\hat{r}_{t+1} = R_{\xi}(\mathbf{c}_t, \mathbf{a}_t)$$

World model loss

$$\mathcal{L}(\theta, \phi, \xi; \mathcal{D}) = \mathbb{E}_{(\mathbf{o}, \mathbf{a}, \mathbf{o}', r)_{0:H} \sim \mathcal{D}} \left[\sum_{h=0}^{H-1} \gamma^h \left(\text{CE}(p_{\phi}(\hat{\mathbf{c}}_{h+1} \mid \hat{\mathbf{c}}_h, \mathbf{a}_h), \mathbf{c}_{h+1}) + \|R_{\xi}(\mathbf{c}_h, \mathbf{a}_h) - r_h\|_2^2 \right) \right]$$

FCAI

fcai.fi

DCWM: Components

FCAI fcai.fi

16