i. For i in number of episodes

i. Collect trajectory $\tau_i = \{o_t, a_t, o_{t+1}, r_t\}_{t=0}^{T}$

ii. Add trajectory to replay buffer $\mathscr{D} \leftarrow \mathscr{D} \cup \tau_i$

iii. For $T \times r_{\text{utd}}$ steps

i. Sample batch from replay buffer $\mathscr{D}$

ii. One encoder update

iii. One critic update

iv. One actor update


iQRL

Algorithm

iQRL

Algorithm

i. For i in number of episodes
   i. Collect trajectory $\tau_i = \{o_t, a_t, o_{t+1}, r_t\}_{t=0}^{T}$
   ii. Add trajectory to replay buffer $\mathscr{D} \leftarrow \mathscr{D} \cup \tau_i$
   iii. For $T \times r_{\text{utd}}$ steps
      i. Sample batch from replay buffer $\mathscr{D}$
      ii. One encoder update
      iii. One critic update
      iv. One actor update

Finite Scalar Quantization