episode_84_steps_42000 Total reward explorative Total steps explorative KL Div w.r.t. optimal allocation - generative 10^{4} 1.2 -1.0 8.0 10^{3} 0.6 0.4 10^{2} 0.2 0.0 Total reward greedy Total steps greedy KL Div w.r.t. optimal allocation - navig. constraints 10^{1} 1.0 8.0 10^{0} 0.6 0.4 10^{-1} 0.2 10^{-2} 0.0

Episode

Episode

Iteration