Granger Causal Interaction Skill Chains

Core Takeaway: Goal-conditioned skills that induce interactions offer improved sample efficiency, overall performance and transferability in long-horizon factored environments.

Overview

Step 1
Learn Granger-Causal factored dynamics models: “passive” autoregressive model and “active” pairwise model.

Step 2
Identify desired interactions for skill goals: states with disagreement between the passive and active models.

Step 3
Learn Goal-conditioned policy with hindsight: reward desired interactions using previously learned policies as actions.

Granger Causal Interactions

Interaction Definition
\[ w^{\text{pass}}(s^{i-1}; \ldots, s^{i-
u}; s^{i+1}; \ldots) = \theta^i + \sum j \theta_j^{i-1} + \theta_i \]
\[ w^{\text{act}}(s^{i-1}; \ldots, s^{i-
u}; s^{i+1}; \ldots) = \theta^i + \sum j \theta_j^{i-1} + \theta_i \]
\[ c_{\text{pass}}(s_k, s_{k+1}; \theta) = \text{log} \text{m}^{\text{pass}}(s_k, s_{k+1}; \theta) \]
\[ c_{\text{act}}(s_k, s_{k+1}; \theta) = \text{log} \text{m}^{\text{act}}(s_k, s_{k+1}; \theta) \]
\[ h_{\text{act}}(s_k, s_{k+1}; \theta) = (c_{\text{act}}(s_k, s_{k+1}; \theta) > c_{\text{act}}) \land (c_{\text{pass}}(s_k, s_{k+1}; \theta) < c_{\text{pass}}) \]

Skill Termination Condition
\[ c_0(s_k, s_{k+1}, s_{k+2}, \theta) := \begin{cases} 1 & h_{\text{act}}(s_k, s_{k+1}, s_{k+2}; \theta) \land ||\Delta s_k - \epsilon \Delta s|| \leq 1 \land \epsilon > 0 \\ 0 & \text{otherwise.} \end{cases} \]

Evaluation

Sample Efficiency and Overall Performance on long-horizon tasks

Skill transfer to challenging in-domain tasks