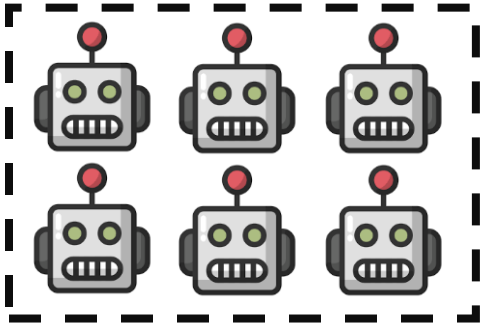
Energy-based Surprise Minimization for Multi-Agent Value Factorization

Karush Suri, Xiao Qi Shi, Konstantinos N. Plataniotis, Yuri A. Lawryshyn

Estimate surprise

$$log \sum_{a=1}^{N} e^{V^{a}_{surp}(s,u,\sigma)}$$



Update Beliefs

Minimize surprise

Novel surprise value function assigns values to surprising states

Energy operator encodes surprising configurations as intrinsic motivation

Agents jointly minimize surprise using free-energy minimization

Resulting joint policy executes surprise-agnostic exploration

Agents depict surprise-robust behaviors at execution time

$$L(\theta) = \mathbb{E}_{b\sim R}[\frac{1}{2}(r + \gamma \max_{u'} \min_{i} Q_i(u', s'; \theta^-) + \beta E - Q(u, s; \theta))^2]$$

Mutual energy acts as intrinsic motivation

Minimize surprise utilizing energy configurations across all agents

Optimal policy results in minimum surprise at thermal equilibrium

Energy-based scheme in conjunction with QMIX (EMIX) demonstrates improved performance on StarCraft II micromanagement scenarios