

ARES-QINTEL と ARES-PRISMIX の性能比較

October 27, 2024

序論

ARES-QINTEL と ARES-PRISMIX は, それぞれ異なる探索戦略と学習手法に基づく強化学習アルゴリズムである. 両者の性能を厳密に比較するため, 以下の側面で理論的検討を行う.

1. 探索戦略の違い (QINTEL のベイズ最適化と PRISMIX のメタ学習)
2. 収束速度とメモリ効率
3. 汎化性能と適応力

探索戦略の違い

ARES-QINTEL は量子インスパイアードなベイズ最適化を用い, 行動価値関数の確率的最適解を求める. 一方, ARES-PRISMIX はメタ学習と因果推論を組み込んだ多層的な探索戦略を持つ.

QINTEL の行動選択

量子インスパイアードなベイズ最適化では, 行動価値 $Q(s, a)$ を確率分布と見立て, 選択する行動 a を次式で決定する.:

$$a = \arg \max_{a'} \Phi \left(\frac{Q(s, a') - \mu}{\sigma} \right) \quad (1)$$

ここで, Φ は標準正規分布の累積分布関数 (CDF), μ および σ は $Q(s, a)$ の平均と分散である.

PRISMIX の行動選択

PRISMIX はメタ学習によって環境変動に適応する探索率 ϵ を調整し, 以下のよう
に行動を決定する.:

$$\epsilon_{t+1} = f_{\text{meta}}(\epsilon_t, R_{t-1}, Q(s_t, a_t)) \quad (2)$$

ここで, f_{meta} はメタ学習モデルにより導かれる関数であり, R_{t-1} は前エピソード
での報酬である.

収束速度とメモリ効率

ARES-QINTEL はメモリ効率が低い設計であるが, PRISMIX は多層的なメタ学習
と因果推論を行うため, メモリ消費が大きくなる傾向がある.

QINTEL の収束速度

ARES-QINTEL の収束速度は, ベイズ最適化の性質上, 理論的には $O(\log N)$ のオー
ダーで早いと期待される.

PRISMIX の収束速度

PRISMIX の収束は以下のような二段階の収束過程が見られる.:

$$\lim_{t \rightarrow \infty} \mathbb{E}[R_t] = R^* - O\left(\frac{1}{\sqrt{t}}\right) \quad (3)$$

PRISMIX の収束は一見遅く見える場合があるが, メタ学習によって得られた情報
により, 長期的には環境変動にも安定して適応できるようになる.

3. 汎化性能と適応力

PRISMIX はメタ学習による高い適応能力があり, 環境の変動に対しても強力であ
る. 一方, QINTEL は局所的な最適解への収束が早いいため, 安定した環境において
強力である.

QINTEL の適応度

環境変動への適応度を示す関数 α_{QINTEL} は以下のように与えられる.:

$$\alpha_{\text{QINTEL}} = \frac{1}{1 + \exp(-\beta \cdot \Delta Q)} \quad (4)$$

ここで, ΔQ は行動価値関数の変動, β は探索率の減衰パラメータである.

PRISMIX の適応度

PRISMIX のメタ学習に基づく適応度 α_{PRISMIX} はメタ学習更新規則 $g(\cdot)$ を通じて次式で与えられる.:

$$\alpha_{\text{PRISMIX}} = g(\alpha_{t-1}, Q(s_t, a_t), R_t) \quad (5)$$

結論

- **ARES-QINTEL**: 早期収束と軽量なメモリ構成で, 特定の安定した環境で強力な性能を発揮する.
- **ARES-PRISMIX**: メタ学習と因果推論によって, 環境変動への適応力が高く, 長期的な汎化性能に優れる.