異なる学習空間を持つ複数エージェントを用いた単体ロボットの意思決定
-TD 誤差に基づいた行動選択確率の統合による環境適応能力の向上-

山田　広来

February 2025

- **環境変化後に過去の行動価値に基づいて行動選択を行う問題に着目**
  - 先行研究では環境変化が起きてもエージェントが受け取る状態が変わらない箇所では，エージェントは過去の知識に基づいて行動選択を行っている．しかし，環境変化によりタスク達成方法が変わったときに，学習が収束したグローバルエージェントの行動価値を統合することにより，過去の知識に基づいた行動が行われ，新たな環境への適応が遅れるという問題があった．
- **TD 誤差に基づいた行動選択確率の統合**
  - グローバルエージェントの学習の収束度合いを TD 誤差の重み付き平均から算出し，その値をグローバルエージェントの行動選択確率に重みとして累乗する．その後，グローバルエージェントの行動選択確率とローカルエージェントの行動選択確率を掛け算を用いて統合し，重みの値に応じて各エージェントの行動選択の影響度を調整する．
- **迷路問題を用いて提案手法による環境適応能力の向上を検証**
  - スタートとゴール，壁が存在する迷路問題を利用し，動的環境においてロボットの行動数を確かめ環境変化への適応能力が向上されているかの検証実験を行った．実験結果から TD 誤差に基づいたグローバルエージェントの行動選択確率を統合していることにより，環境変化後の行動数が削減された．このことから提案システムにより環境適応能力が向上したことを確認した．

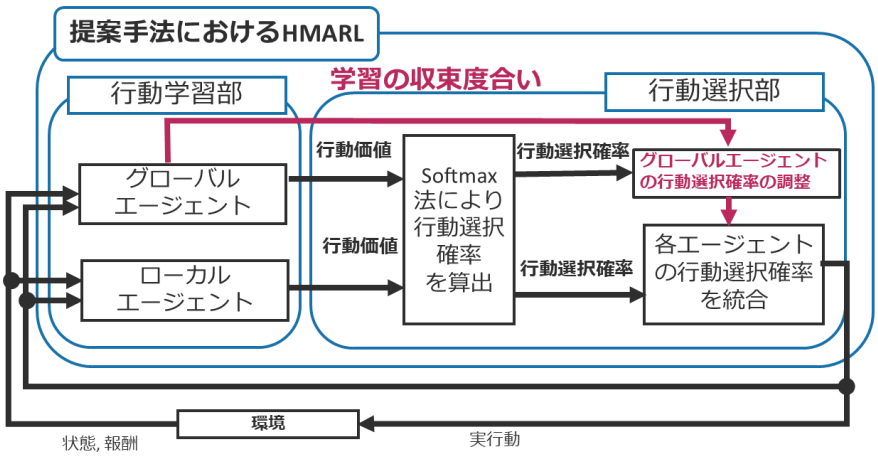

Fig.1　図