

タスク毎の DDPG を用いた行動価値空間の形成における行動選択

中村 悟 (Satoru Nakamura)

February 2021

- 離散行動における行動の細分化では，時間的制約があることに着目
 - － 我々は優先度を定義し複数タスク下で行動選択を行い，行動を細分化する研究をしている．行動選択の際に中間行動を生成し追加を行うと，学習に時間がかかることや行動が限定的なままであるという問題があった．
- 新たな行動価値空間の形成で，複数タスク下において連続行動用いた行動選択することを実現
 - － 本研究では，連続行動を扱う深層強化学習を用いて，新たに学習空間を形成し複数タスク下で行動選択が可能になる手法を提案する．
- 最短経路問題を用いて，複数タスク下での達成効率の向上を検証
 - － 最短経路問題を考慮した実験環境で実験を行い，連続化した行動によってタスクの達成が効率的になったことを確認した．

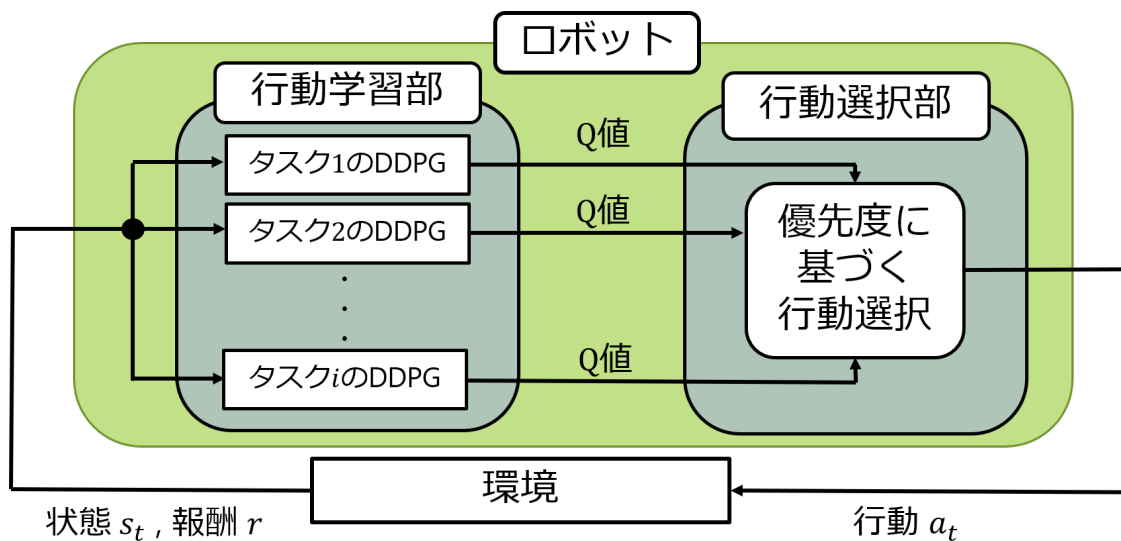


Fig.1 DDPG を用いたシステムの概要図