

階層型自然方策勾配法

福永修一

東京都立産業技術高等専門学校

強化学習

- 試行錯誤により報酬を最大化する方策を求める手法
 - 状態・行動空間が高次元になると学習が困難
- ➡ 方策を直接近似する方策勾配法が有効

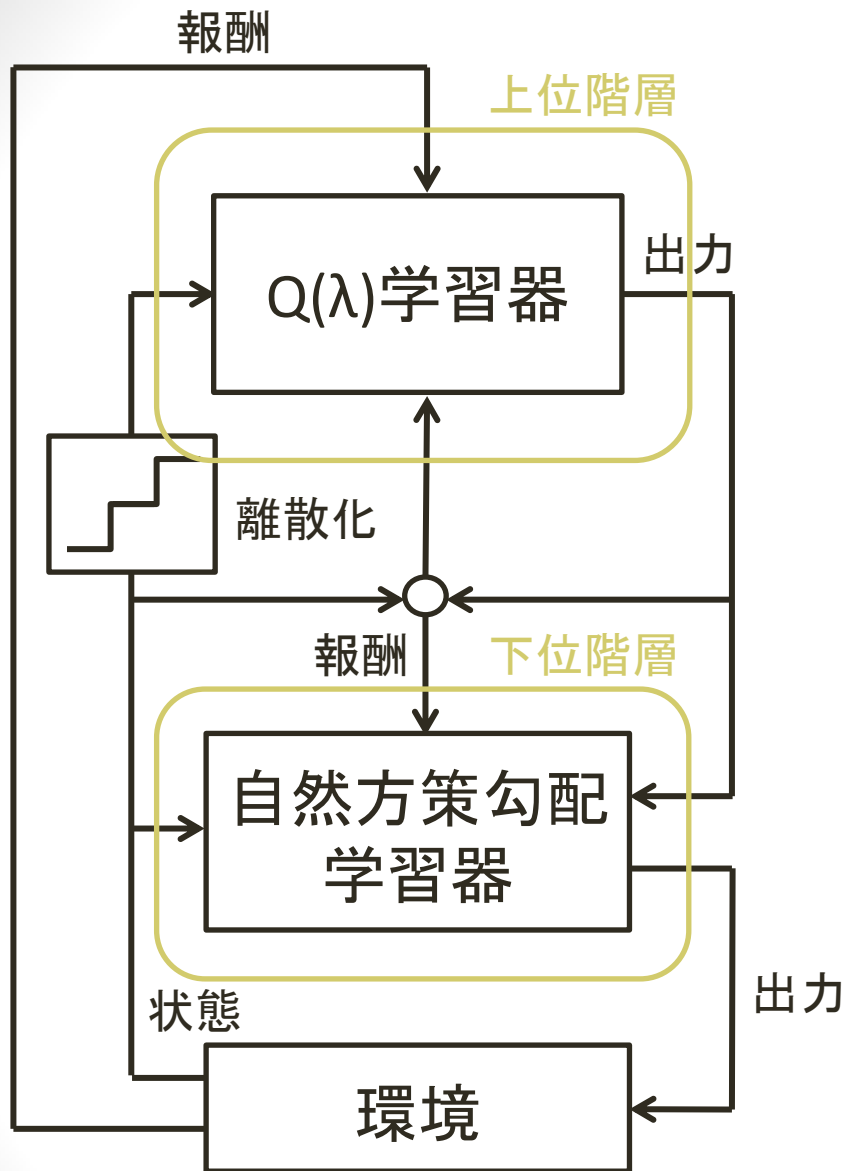
関連する手法として

- 階層型方策勾配法(Ghavamzadeh & Mahadevan, 2003)
- 自然方策勾配法(Kakade, 2002)



2つの手法を組み合わせること
により学習を高速化

階層型自然方策勾配法の提案



提案手法の模式図

階層型ハイブリッド型アルゴリズム (Morimoto & Doya, 2000)

Q(λ)学習+Actor Critic法



改良

提案手法

- 上位階層：Q(λ)学習
 - 離散化された状態を扱う
- 下位階層：自然方策勾配法
 - 連続状態を扱う
 - サブゴールに到達することを学習

数値例を用いて提案手法の有効性を検証