

D-52

状態非依存-時刻依存方策による学習の提案

前田新一, 石井信 京都大学 情報学研究科

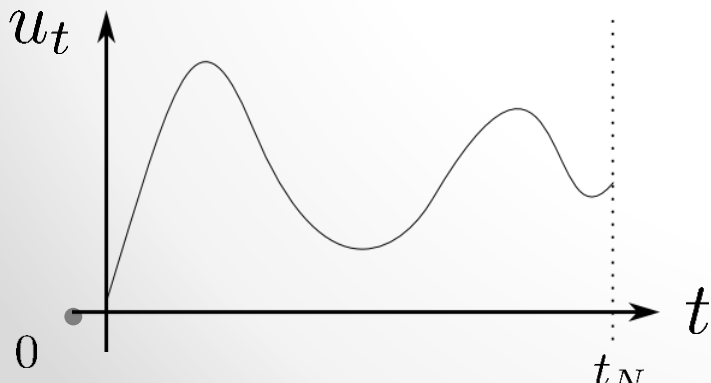
状態依存の関数は、状態空間が大きくなると学習が困難



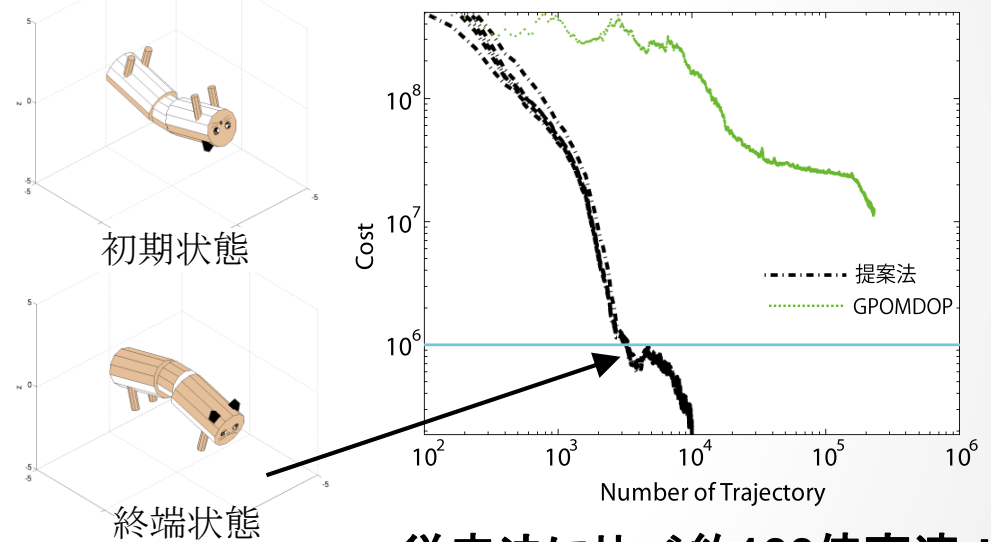
状態非依存-時刻依存方策

- 初期状態 x_0 は固定
- 各時刻でとるべき行動 u_t を時刻の関数 t で表現

$$u_t = f(t, x_0, \theta)$$



猫ひねり問題への適用



従来法に比べ約100倍高速!

想定される応用例

