

## 学習制御法

工学部 知能機械工学科 山田 訓, 浅野 翼(山田研究室)

Keywords: 強化学習, ロボット, 自律的学習制御, 適応制御, 入力空間表現

### 1. 研究目的

実世界で働くロボットは、様々な未知環境や変化する環境内で行動しなければならない。しかし、全ての場合に対し、プログラムしておくことはできないので、何らかの学習法が必要である。強化学習は制御した結果に対する評価(良いか、悪いか)のみを用いて、制御を自律的に学習するので、幅広い制御へ適用できる可能性がある。様々な制御課題を学習するためには、その課題に対して適切な方法で入力进行处理する必要がある。本研究では、センサ値から周囲の状況を判断するための効率的な入力空間表現法を検討した。

### 2. 環境と制御課題

本研究では、図1のような学習環境で、ロボットが障害物センサを用いて壁沿いを走行するというセンサ値の細かい違いを区別する必要がある制御課題で、空間表現法の性能を比較した。従来は、入力空間を一定の間隔で分割する表計算(CMAC)で行なっていたので、センサ値の細かい変化を識別するために、多数の要素(50000個)を必要とし、計算時間が長かかった(表1)。本研究(INGnet)では、その課題で必要な部分だけに要素を配置するので、要素を必要最小限(54個)にでき、計算時間を短くすることができた(約1/75)(表1)。

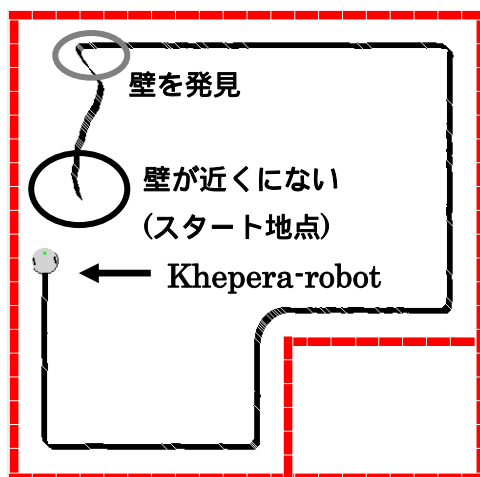


図1 学習環境と成功の軌跡

表1 各表現法で学習に必要な要素数

	要素数	計算時間 (1要素) [ $\mu$ sec]	計算時間 (全要素) [ $\mu$ sec]
従来法(CMAC)	50000	0.0986	4930
本研究(INGnet)	54	1.2111	65.3994

### 3. 応用の可能性

制御課題に適する必要最小限の要素で処理できるので、多数の入力や細かいセンサ値の区別が必要となる実世界の制御課題のリアルタイム制御へ適用できる可能性がある。なお、この研究は文科省高度化推進事業・社会連携研究推進事業の一環である。