

## 自ら学習し進化するロボットのための人工知能

**研究概要:** 人と協働するロボットの実現には、環境に応じて適切な動作ができる「自律化」が求められる。そして、自律化を達成するには、ロボットが状況に応じた行動を「知識」として保持していなければならない、これを形成するために「学習」が必要である。強化学習は、ロボットの行動学習手法として有望であるが、原理的に学習時間の長さが問題であり、容易に適用できない。本研究では、過去に獲得した知識を異なるタスクに転移することによって、学習を高速化する手法について研究している。

**キーワード:** 強化学習, 転移学習, 知識, 自律化, ロボット

### 研究背景

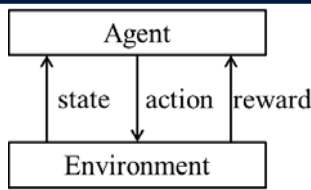


Fig.1 強化学習のモデル

強化学習は**試行錯誤**を伴う学習手法であるため、適切な政策を学習するまでに**多くの試行回数(時間)**を要する。

試行回数を削減する方法として知識の転移によって学習を効率化する**転移学習**がある。

### 学習アルゴリズム

- Actor-Criticをベースにした学習アルゴリズム。
- 状態空間はFuzzy ARTで自律的に構築。
- 知識の転移・消去・ランダムな政策を**交叉・淘汰・突然変異**の概念で再現。

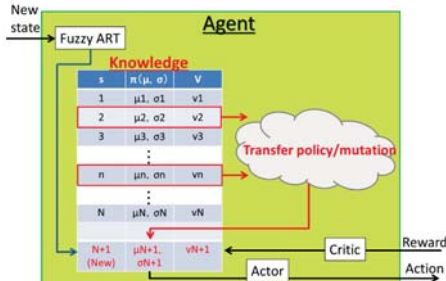
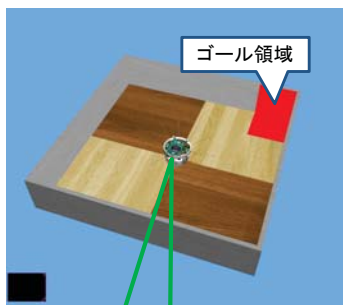


Fig.3 交叉・淘汰・突然変異の概念を用いた学習手法

### 応用例2: 小型2輪ロボットの行動学習



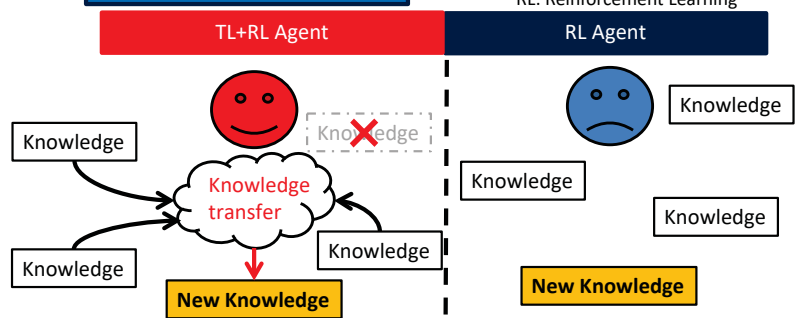
Robot radius	37 mm
Wheel radius	20.5 mm
Axle length	52 mm
Encoder resolution	159.23
Speed unit	0.00628 rad/s
Maximum angular speed	1000 units
Devises	Differential wheels, Proximity sensors, Light sensors, LEDs, Camera, Accelerometer, Ground sensors

	Time(h)
Without transfer	15.0
With transfer (policy)	13.7
With transfer (policy & value)	13.2

Fig.5 二輪ロボットのモータスキル獲得実験

成果: 学習時間を15.0(h)→13.2(h)と約10%削減(単一タスク)

### 転移学習を用いた強化学習



- ・過去の知識を元に学習する
- ・個別に知識を獲得するのみ

Fig.2 転移学習の有無によるエージェントの学習イメージの違い

TLを導入すると未知タスクに対して政策の獲得が早い

→ 遺伝的アルゴリズムの概念を用いた転移学習を提案

### 応用例1: 多関節冗長ロボットアームの経路学習

#### 学習条件

- アームの先端がゴール内に入れば報酬(r)+10。
- タスク失敗時の報酬はなし。
- 10個のタスクを学習。

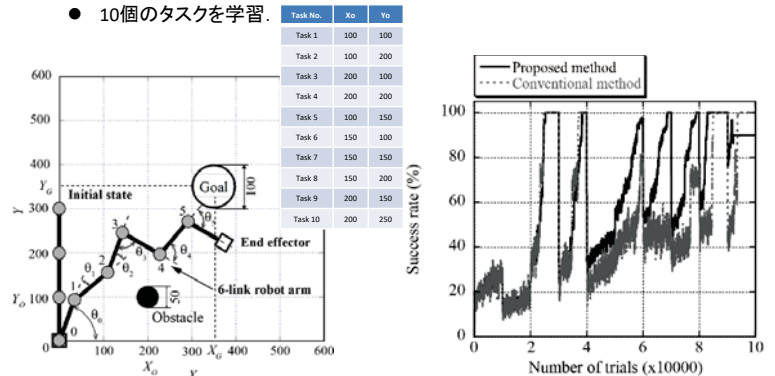


Fig.4 6関節ロボットアームによるリーチング動作獲得問題

成果: 学習時間を86.5(h)→10.2(h)と約90%削減

### 本研究の展開と関連研究の応用例

#### □ ロボットの自律化支援

- ヒューマノイドロボットのような冗長システムの動作生成を高速化。
- プログラムではなく、褒める、叱るなどのコミュニケーションを通してロボットの性能を向上させることで、専門知識を持たなくてもロボットを使用できる。
- ロボットを訓練し、技能を学習したロボットを現場へ投入する。
- 知識の転移により、ロボットの進化を加速。

#### □ 最適化手法としての利用

- アレイセンサの指向性設計への応用