

対話的進化ロボティクスに基づく実環境ロボット学習

2B-02

片上 大輔 山田 誠二

東京工業大学大学院総合理工学研究科知能システム科学専攻

1 はじめに

従来のロボット学習では、評価関数やその他のパラメータ調整がうまくいかず、人間が意図するような行動を学習しない場合が多かった。また一方、実環境での強化学習は、学習の収束に時間がかかり、特に報酬を得るまでの初期学習の立ち上がりに大きな時間コストがかかる。しかも 1 回の行動に必要な大半の時間は、ロボットの感覚、行動系の処理時間に費やされるので、高速に学習するためには、学習試行数の削減自体が必要である。

そこで我々は、人間を評価系に組み込む対話型進化計算法 (IEC)[1] の評価能力を用いて効率のよい実環境ロボット学習を行うことを目的とし、少ない試行数で学習でき環境の多様性や動的状況の変化に適応可能なクラシファイアシステム [2] に基づく学習システム、対話型クラシファイアシステム (ICS)[3] を提案してきた。Fig.1 にシステムの概要を示す。これによりロボット学習分野における初期学習の効率化を図り、また人間が意図するような行動を学習することが可能である。

さらに、従来の対話型手法においてはその対話の主体における認知の重要性についてはあまり触れられてこなかった。本研究では対話をユーザからロボットへの教示であるとして、その教示の主体における認知について、教示する側の視点に基づく場合と学習する側の視点に基づく場合の 2 種類の場合を想定する。

2 教師の認知と学習者の認知

局所情報しか知らないロボットが、特定の大域目標の実現に対してどのように自己の行動を修正したら良いのかを知ることは一般に困難である。しかも、ロボットには目標が実現していることさえ明確に認識できないかもしれない。目標の実現を認識できるのは、システム内部のロボットではなく、むしろシステムの外部から全体を眺めることのできる観測者である。本研究では、この教師からの認知を **Teacher View**(Fig.2) と呼び、逆にシステム内部から状況を判断する学習者

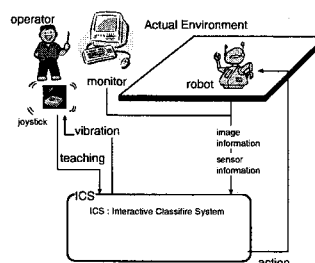


図 1: Teaching Environment

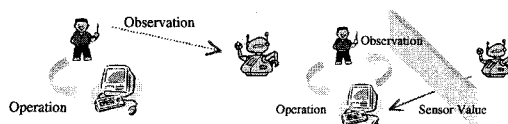


図 2: Teacher View

図 3: Learner View

の認知を **Learner View**(Fig.3) と呼ぶ。

ロボットの認知と外的な観測者による認知の間には一般に大きな差異がある。例えば実環境上でロボットを動かした場合、センサ情報の獲得によって形成されるロボットの内部モデルは、その状況および観測者の履歴に依存して大きく異なったものになってしまう。その結果、ロボットはあらかじめ設計者が設計した行動規則に従うものの、状況の変化に対応して設計者の意図する行動を生成しないことがある。これは、外的な観測者の想定するロボットの環境認識とロボットの実際の認識が異なっているために生じるものである。さらにその環境が複雑で、外的な観測が困難であればあるほどそのずれは大きくなることになり重要な問題となる。従来の評価関数の設定やパラメータの調整がうまくいかなかったのは、この認識のずれが問題であったためであると考えられる。

3 対話型クラシファイアシステム

対話型クラシファイアシステムはクラシファイアシステムの一つである XCS[4] に対して IEC による対話

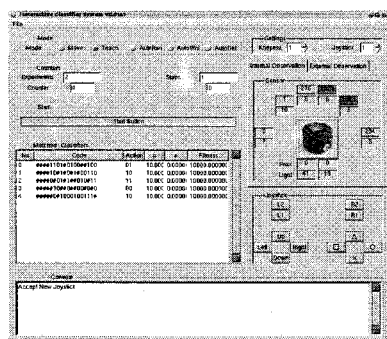


図 4: User Interface



図 5: A Mobile Robot: Khepera and Joystick

機能を組み込むことにより、自律的な学習に加え、教示による学習も行うことができるロボット学習システムである。開発は Linux 上で、GTK+, Video4Linux API を使って行った。開発したインタフェースを Fig.4 に示す。

4 実験

移動ロボットとして、Khepera を用いる。Khepera は、直径 55[mm]、高さ 32[mm]、重さ 70[g] で、モトローラ 68331, RAM 256[Kbyte], ROM 512[Kbyte] を搭載している。また、DC モータ (ロボットの移動 1[sec] あたり 8[mm]) を 2 つ、赤外線近接センサと光センサが一体になったものを Fig.5 の右図の位置に 8 つ装備している。また、教示に用いる入力装置には SONY 社製アナログコントローラ DUALSHOCK をパラレルポートに接続し使用する。それぞれ、Fig.5 に示す。

● Teacher View による教示実験

環境全体を見渡せるように外部に設置した CCD カメラからの情報 (Fig.6) を用いて Teacher View による教示を実現する。教師は Teacher View より環境全体の情報をみながら教示を行うことができる。これにより、教師はグローバルな情報をもとにタスクをうまくこなす教示ができるが、反面ロボットの状態に変化ない場合学習がうまく行か

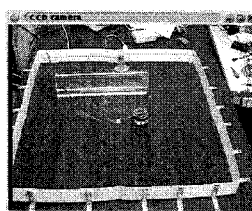


図 6: CCD Camera View

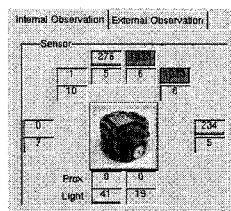


図 7: Sensor Panel

ないという問題が発生する。この教示は探索空間が広い場合や報酬がなかなか得られない場合に効果を発揮する。

● Leanear View による教示実験

Fig.7 に示した、ロボットの現在の状態を知らせるセンサパネルを用いて Learner View による教示を実現する。教師は Learner View より、ローカルな情報をもとにタスクをこなさなくてはならないため、タスクの遂行はより困難なものになるが、ロボットの学習の観点で考えるとより確実なルール作成が行え、教示による学習が効率的に進むと考える。この教示はタスクが複雑な場合や動的な環境において効果を発揮する。

この 2 つに教示法について比較実験を行い、現在タスクの難易度 (認識のずれの大きさ) によってこの教示法の結果に違いが出ることが確認されている。

5 おわりに

従来のロボット行動学習がとらわれて来た行動還元主義的なアプローチではなく、対話時の視点に焦点をあてた認知主義的な新しいアプローチにより教示を行う手法を提案した。現在教師の認知と学習者の認知の 2 つの場合においていくつかのタスクで実験中である。

参考文献

- [1] 高木 英行, 敵見 達夫, 寺野 隆雄. 対話型進化計算法の研究動向. 人工知能学会誌, 13(5):24-35, 1998.
- [2] John H. Holland and Judith S. Reitman. Cognitive systems based on adaptive algorithms. In Donald A. Waterman and Frederick Hayes-Roth, editors, *Pattern-Directed Inference Systems*, pages 313-329, Orlando, 1978. Academic Press.
- [3] D. Katagami and S. Yamada. Interactive classifier system for real robot learning. In *(International Workshop on Robot and Human Interaction ROMAN-2000)*, pages 258-263, 2000.
- [4] S.W. Wilson. Classifier fitness based on accuracy. *Evolutionary Computation*, 3(2):149-175, 1995.