

解説

RoboCup 中型リーグ研究事例紹介 - 阪大 Trackies -

Research Activities of Osaka Univ. Trackies in RoboCup Middle Size League

高橋 泰岳*1 浅田 稔*2*1大阪大学大学院工学研究科 *2大阪大学大学院工学研究科

Yasutake Takahashi*1 and Minoru Asada*2*1Graduate School of Engineering, Osaka University *2Graduate School of Engineering, Osaka University

1. はじめに

大阪大学 Trackies は1997年に行われた第一回ロボカップから参加し続けている。我々のチームはロボットの自律的行動獲得を目的として、特にロボットの学習機構について研究を行ってきた。ロボットに学習させる利点として、より少ない事前知識、環境の変化への順応性などがあげられるが、実際の環境で活動する実ロボットに適用するには様々な問題が表面化する。本稿では我々の研究室におけるロボカップを題材にした一連の研究を簡単に紹介する。

2. 移動ロボットを用いた行動獲得

2.1 強化学習の応用

視覚情報をもとにシュート行動やゴール守備行動を獲得するために強化学習を応用した [1] [11]。強化学習の手法として離散化された状態空間と行動空間を必要とする Q 学習を用いた [9]。後にこれを連続空間に拡張するための手法も提案した [4]。

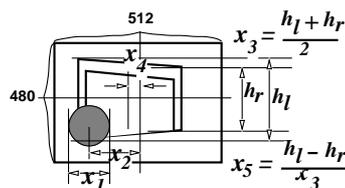


図1 Our robot picture and input vector composing state space

2.2 自律的な状態空間の分節化

強化学習を適用する際、必要とされる状態行動空間を設計者が与えていたのではロボットの自律性が大きく損なわれるので、ロボット自身が環境との相互作用を通してセン

原稿受付 2001年9月28日

キーワード: RoboCup MSL

*1 〒565-0871 大阪府吹田市山田丘 2-1

*2 〒565-0871 大阪府吹田市山田丘 2-1

*1 Yamadaoka 2-1, Suita-shi, Osaka 565-0871

*2 Yamadaoka 2-1, Suita-shi, Osaka 565-0871

サ情報を分割するのが望ましい。そこで成功事例に基づくオフライン法 [12] と関数近似に基づくオンライン法 [13] を開発した。

2.3 複数の学習器の階層的構築による行動獲得

同一構造の学習器を複数用い、これを階層的に自律的に構築することによる行動獲得手法を提案し、実機による実験を通してその有効性を確認した [14]。また、全状態行動空間を部分空間に分け高いレベルで全体の統合を行うことで、多種多様な論理センサ・アクチュエータを装備したロボットによる学習を行う手法を提案し、サッカーロボットによるシュート行動に適用することで有効性を確認した [3]。

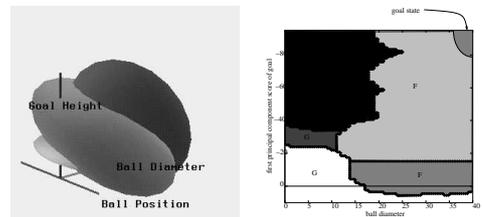


図2 Examples of state space construction : online method (left) and offline one(right)

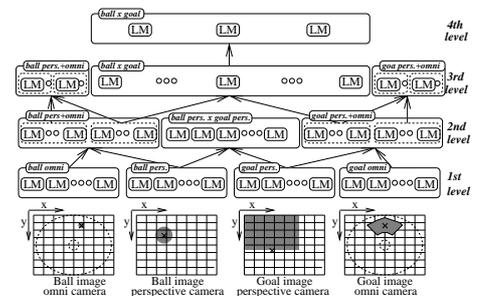


図3 Multi-layered learning system for vision based behaviors

2.4 適応的な評価関数を用いた進化的手法のための行動獲得

多目的な評価値に基づいて行動を評価する際、評価間のバランスが重要になる。本研究では個体集合の評価値の分布に応じて複数の評価値の重み付けを適応的に変更する枠組を提案した [8]。

3. 複数ロボットの協調行動の獲得

3.1 他者の状態推定

環境内に自律的に行動できるロボットが複数存在する場合、状態空間の構成そのものが困難となる。そこで、観測と行動のシーケンスから次元を含めた状態ベクトルを推定する手法を提案した [2]。

3.2 共進化による協調行動の創発

複数ロボットの同時学習として、進化的手法を用いて共進化による協調・競合行動の創発についての実験を行った。このときロボットに与えられるタスクの適切な複雑さが効果的な共進化を誘導することについて議論した [10]。

3.3 モジュール干渉検出機構に基づく複数タスクの動的割当

複数のタスクを動的にロボットに割り当てる手法を開発した。具体的には、各タスクの達成状況をロボット間で共有することにより各ロボットが同時に達成できるタスクのみを自律的に選択するための枠組を提案した [7]。

4. チーム戦略学習

マルチエージェントシステムで敵味方に分かれ非常に動的な環境、ロボカップでよく見られるこの特徴的な環境において、ロボット同士で明示的な通信を行わない場合でも協調的動作が創発されることを実際の競技の中で示した [5]。またこの知見をもとに、おのおの性格を持った選手の組合せを戦略と考え、監督の立場からゲーム進行状況から相手の戦略を推定し、相手に勝つ味方の戦略を推定する機構を提案している [6]。

5. ま と め

以上、大雑把に我々の研究室で行われて来た研究を紹介した。詳しくは以下の文献、もしくは我々のwebページ (<http://www.er.ams.eng.osaka-u.ac.jp>) を参照して頂ければ幸いである。

参 考 文 献

- [1] M. Asada, S. Noda, S. Tawaratumida, and K. Hosoda. Purposeful behavior acquisition for a real robot by vision-based reinforcement learning. *Machine Learning*, Vol. 23, pp. 279–303, 1996.
- [2] M. Asada, E. Uchibe, and K. Hosoda. Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development. *Artificial Intelligence*, Vol. 110, pp. 275–292, 1999.
- [3] Y. Takahashi and M. Asada. Multi-controller fusion in multi-layered reinforcement learning. In *International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI2001)*, pp. 7–12, 2001.
- [4] Yasutake Takahashi, Masanori Takeda, and Minoru Asada. Improvement continuous valued q-learning and its application to vision guided behavior acquisition. In *The Fourth International Workshop on RoboCup*, pp. 255–260, 2000.
- [5] Yasutake Takahashi, Takashi Tamura, and Minoru Asada. Cooperation via environmental dynamics caused by multi robots in a hostile environment. In *The Fourth IFAC Symposium on Intelligent Autonomous Vehicles*, pp. 413–418, September 2001.
- [6] Yasutake Takahashi, Takashi Tamura, and Minoru Asada. Strategy learning for a team in adversary environments. In *RoboCup 2001 book*, p. (to appear), 2001.
- [7] Eiji Uchibe, Tatsunori Kato, Minoru Asada, and Koh Hosoda. Dynamic task assignment in a multiagent/multitask environment based on module conflict resolution. In *Proc. of IEEE International Conference on Robotics and Automation*, pp. 3987–3992, 2001.
- [8] Eiji Uchibe, Masakazu Yanase, and Minoru Asada. Behavior generation for a mobile robot based on the adaptive fitness function. In *Proc. of Intelligent Autonomous Systems (IAS-6)*, pp. 3–10, 2000.
- [9] C. J. C. H. Watkins and P. Dayan. “Technical note: Q-learning”. *Machine Learning*, Vol. 8, pp. 279–292, 1992.
- [10] 内部英治, 浅田稔, 中村理輝. 共進化によるマルチ移動ロボット環境における協調行動の獲得. 遺伝的アルゴリズム 4, 第7章, pp. 193–220. 産業図書, 2000.
- [11] 加藤, 鈴木, 浅田. 強化学習によるゴール守備行動の獲得. 第3回 JSME ロボメカ・シンポジウム講演論文集, pp. 37–40, 1998.
- [12] 浅田, 野田, 細田. ロボットの行動獲得のための状態空間の自律的構成. 日本ロボット学会誌, Vol. 15, No. 6, pp. 886–892, 1997.
- [13] 高橋泰岳, 浅田稔. 実ロボットによる行動学習のための状態空間の漸次的構成. 日本ロボット学会誌, Vol. 17, No. 1, pp. 118–124, 1999.
- [14] 高橋泰岳, 浅田稔. 複数の学習器の階層的構築による行動獲得. 日本ロボット学会誌, Vol. 18, No. 7, pp. 1040–1046, 2000.

高橋 泰岳 (Yasutake Takahashi)

1972年12月13日生。1994年大阪大学大学院学研究科博士前期課程修了。2000年同大学博士後期課程中退、同年同大学大学院工学研究科助手。知能ロボットの行動獲得に関する研究に従事。(日本ロボット学会正会員)

浅田 稔 (Minoru Asada)

1982年大阪大学大学院基礎工学研究科後期課程修了。同年、大阪大学基礎工学部助手。1989年大阪大学工学部助教授。1995年同教授。1997年大阪大学大学院工学研究科知能・機能創成工学専攻教授となり現在に至る。この間、1986年から1年間米国メリーランド大学客員研究員。知能ロボットの研究に従事。1989年、情報処理学会研究賞、1992年、IEEE/RSJ IROS'92 Best Paper Award受賞。1996年日本ロボット学会論文賞受賞。博士(工学)。日本ロボット学会、電子情報通信学会、情報処理学会、人工知能学会、日本機械学会、計測自動制御学会、システム制御情報学会、IEEE R&A, CS, SMC societiesなどの会員。(日本ロボット学会正会員)