

# 間引きを用いたパス技術の自律学習

小林隼人<sup>†</sup> 畑埜晃平<sup>‡</sup> 石野明<sup>†</sup> 篠原歩<sup>†</sup>

<sup>†</sup> 東北大学大学院情報科学研究科      <sup>‡</sup> 九州大学大学院システム情報科学研究科

## 1 はじめに

ロボットが実世界で知的に活動するには、未知の環境に適応する能力、つまり学習能力が必要不可欠である。本研究では、自律ロボットによるサッカー競技である RoboCup サッカーを題材に、自律ロボットの学習に焦点をあてた研究を行う。RoboCup サッカーは人工知能とロボット工学の融合を目指して、日本の研究者らによって提唱された国際プロジェクトである。本研究では、RoboCup サッカーのスタンダードプラットフォームリーグで用いられる犬型ロボット、AIBO を用いて実験を行う。

AIBO のように肢体を持つロボットがサッカーをするには、歩く、走る、シュートする、パスするといった基本技術を習得しなければならない。これらの基本技術は、実環境との相互作用が重要な役割を担っているため、実用的な技術を獲得するには実環境で学習する必要がある。このような場合、ロボットに人手を介さずに技術を獲得させることで、実験の人的コストを抑えることができる。これは自律学習 (autonomous learning) として知られており、様々に研究が行われている。Fidelman と Stone [1] は、止まったボールを弾かずに捕捉する技術の自律学習法を示し、Kobayashi ら [3] は、向かってくるボールをトラップする技術の自律学習法を示した。

本研究では、サッカーの戦略に必要な基本技術であるパス技術の学習に取り組む。我々は、パス技術の自律学習手法を提案し、その学習過程の試行回数を削減するために、先行研究 [2] で提案した間引き (thinning-out) を利用する。先行研究では仮想環境上での学習問題に取り組んだが、本研究では実環境上での学習問題における間引きの効果を検証する。

## 2 パス技術の自律学習

パス技術は、ある固定された目的距離 (味方ロボットの位置など) までボールを正確に転がす精密なシュー

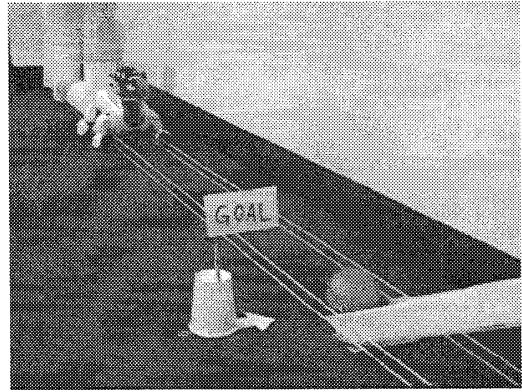


図 1: パス技術の自律学習のための実験装置

ト動作である。我々の開発したフレームワークでは、AIBO のシュート動作を複数のキーフレームから生成できる。キーフレームは、動作の骨組みを形成する特徴的な姿勢を表す 15 個 (=首 3 個+各足 3 個×4 本) の関節角度列である。例えば、ボールを蹴るためには足を後ろに引いた姿勢と、前に押し出した姿勢の 2 つのキーフレームが最低限必要である。

本研究では、シュート動作の方向を前方方向に制限することで自律学習を実現する。まず、図 1 に示す 2 本のロープとスロープからなる実験装置を用意した。GOAL の旗は目的距離を表しており、ボールが目的距離以上転がると手前に戻ってくるようになっている。ロボットは、自動的にボールが返ってくるゴルフパット練習機と同じように、ボールを取りにいく労力をほとんど使うこと無く自律的にパス技術を学習することができる。

シュート動作が左右対称であることから、キーフレームにおける首の左右方向の関節角度を 0 とし、左足の関節角度を右足と同じ値にすることができる。したがって、キーフレーム数を  $n_k$  とすると、探索空間は  $8n_k$  次元となる。スコア関数  $f: \mathbf{R}^{8n_k} \rightarrow \mathbf{R}$  は、 $x \in \mathbf{R}^{8n_k}$  から生成されるシュート動作を実行し、一定時間後ロボットからボールまでの距離  $d_{ball}$  を返す関数とする。このとき、目的距離を  $D_{goal}$  とすると、スロープにより常に  $d_{ball} \leq D_{goal}$  となることに注意する。ここでは、 $D_{goal} = 800$  とした。本研究における学習の目的は、このスコア関数を最大化することで、目的距離で正確にボールが止まるシュート動作を見つけることである。

**Autonomous Learning of Ball Passing Skills Using Thinning-out.** Hayato Kobayashi<sup>†</sup>, Kohei Hatano<sup>‡</sup>, Akira Ishino<sup>†</sup>, and Ayumi Shinohara<sup>†</sup>. <sup>†</sup>Graduate School of Information Sciences, Tohoku University <sup>‡</sup>Graduate School of Information Science and Electrical Engineering, Kyushu University. kobayashi@shino.ecei.tohoku.ac.jp, hatano@i.kyushu-u.ac.jp, ishino@ecei.tohoku.ac.jp, and ayumi@ecei.tohoku.ac.jp.

### 3 間引き

間引きは、探索木における枝刈り (pruning) と同様の概念であり、不必要な試行を省く手法である。我々はまず、スコア関数が探索空間上で  $g$ -Lipschitz 連続であると仮定する。 $g$ -Lipschitz 連続は、定数  $c$  に対する  $c$ -Lipschitz 連続を関数  $g$  で扱えるよう拡張したものである。このとき、関数  $g$  を正しく推定できれば、以下の間引き条件を満たす候補点の評価を省いても学習結果が悪化しないことを保証する。

**定義 3.1 (間引き条件)**  $x_c$  を探索空間から選んだ評価候補点とする。 $x_n$  を評価履歴中の  $x_c$  の最近傍点、 $x_b$  を評価履歴中の最高スコア  $f(x_b)$  を持つ点とする。 $d$  を探索空間上の距離関数とする。ある推定関数  $\hat{g}: \mathbf{R} \rightarrow \mathbf{R}$  について、 $f(x_n) + \hat{g}(d(x_c, x_n)) \leq f(x_b)$  を満たすとき、 $x_c$  は  $\hat{g}$  に関して間引き条件を満足するという。

本研究では、関数  $g$  の推定方法として差分収集 (GD) 法を利用する。GD 法は、近接する 2 点を結ぶ直線が、その近傍における関数の傾きを良く近似するというヒューリスティクスを使って関数  $g$  を推定する手法である。この手法は、先行研究での比較実験において、高次元の問題に対しても多くの候補を削減できることが示されている。

### 4 実験と結果

本研究では、遺伝的アルゴリズム (GA) に GD 法を適用したアルゴリズム GA+GD を用いて実験を行う。初期モーションは、胸でボールを押し出す強いシュート動作である。この動作は 6 個のキーフレームからなるため、探索空間は 48 次元である。ノイズ対策のため、各スコアはシュート動作を 5 回実行して中間値をとる。

図 2 は、評価回数 50 回における GA と GA+GD の学習過程を示している。それぞれの実験には約 3 時間、バッテリー交換 4 程度を要した。図より、GA は初めに良いスコアを得ているにもかかわらず、それを活かすことができずにいるのに対し、GA+GD は初期段階で良いスコアは得られなかったものの、徐々に最高スコアを更新していることが分かる。実際、GA+GD は平均的に GA よりも良い結果を示している。

表 1 は、図 2 における最高スコアと、学習に使用した評価候補数を示している。最高スコアの結果は、学習により得られたパス技術の精度が、GA は 66 mm、GA+GD が 30 mm であることを意味する。また、評価候補数の結果は、GA+GD の最高スコアが 2000 個以上の候補点を吟味した結果であることを示唆している。もし 2000 回の評価を行うならば、実験には約 80 時間 ( $\approx (2000 \text{ 回の評価}) \times (5 \text{ 回の中間値}) \times (30 \text{ 秒の}$

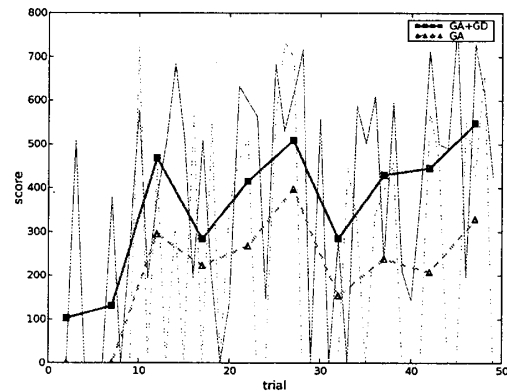


図 2: 評価回数 50 回における学習過程。細線は 1 評価毎のスコア、太線は 5 評価の平均スコアを表す。

表 1: 図 2 における最高スコアと評価候補数

学習法	GA	GA+GD
最高スコア	734	770
評価候補数	50	2192

動作時間)) が必要とされる。つまり、間引きにより 80 時間以上かかるはずの実験を約 3 時間で実現できたことを意味する。

### 5 まとめ

本研究では、犬型ロボット AIBO によるパス技術の自律学習に取り組んだ。ロボットはバッテリー交換以外は全く人手を介さずに、パス技術を獲得することができた。また、本研究により、実環境の問題に対しても、先行研究で提案した間引きが効果的に試行回数を削減できることを示した。実験の学習前、学習後の動画は、著者のサイト (<http://www.shino.ecei.tohoku.ac.jp/~kobayashi/movies.html>) で公開されている。

### 参考文献

- [1] P. Fiedelman and P. Stone. The Chin Pinch: A Case Study in Skill Learning on a Legged Robot. In *Proc. of RoboCup 2006*, pages 59–71. Springer-Verlag, 2007.
- [2] H. Kobayashi, K. Hatano, A. Ishino, and A. Shinohara. Reducing Trials Using Thinning-out in Skill Discovery. In *Proc. of DS2007*, pages 127–138. Springer-Verlag, 2007.
- [3] H. Kobayashi, T. Osaki, E. Williams, A. Ishino, and A. Shinohara. Autonomous Learning of Ball Trapping in the Four-legged Robot League. In *Proc. of RoboCup 2006*, pages 86–97. Springer-Verlag, 2007.