

遺伝的プログラミングを用いた自動ゲームプログラミング

Automatic Game Programming using Genetic Programming

09-03

河口 信夫†

Nobuo KAWAGUCHI†

長江 信隆†

Nobutaka NAGAE†

外山 勝彦†

Katsuhiko TOYAMA†

稲垣 康善†

Yasuyoshi INAGAKI†

†名古屋大学大学院工学研究科計算理工学専攻

†Dept. of Comp. Sci. and Eng., Nagoya Univ.

In this paper, we propose a learning method of position evaluation functions of games using genetic programming (GP). Each position evaluation function is composed of several operators and variables which have some information of a position. We examine two cases which differ in variable settings. One is using raw information for position variables. The other is using high level information obtained from experts knowledge. The method uses a set of pairs of positions, one of which is known to be superior to the other, as training signals. We applied the method to simple games under the several parameters, and obtained proper functions.

1 はじめに

良いゲームプログラムは、高い精度を持つ局面の静的評価関数と高速な探索手法の組み合わせによって成立する。チェスの世界チャンピオンに勝利したシステムには、高速な先読み探索を可能とする並列計算機と、長年の経験による評価関数のチューニングが不可欠であった[松原 97]。しかし、良い評価関数を構築するためには人の経験が必要であり、そのチューニングの方法論も明確でない。本研究はゲームの局面評価関数を遺伝的プログラミング (GP) により自動的に学習する手法を提案する。この手法により、困難であった評価関数のチューニングを自動化することができ、自動ゲームプログラミングが可能となる。本手法の特徴は、相対的な優劣のついた局面対 (事例) から、局面的情報をを用いて GP により評価関数を構成する点にある。我々はすでに局面情報としてゲームのエキスパートの知識を利用する場合について検討した [堀之内 97]。この場合、どのような局面情報を用いるかが問題となる。そこで本稿では、局面から直接得られる情報 (raw information) のみを用いる場合についても検討した。対象ゲームとして tic-tac-toe ゲームを用い、様々なバリエーションの下で実験を行い、直接情報でもある程度の学習が行えることを確認した。

2 遺伝的プログラミングによる局面評価関数の獲得

GP による評価関数の学習を簡単に解説する。詳しくは [堀之内 97] を参照されたい。GP は評価関数を個体として持ち、各評価関数は、局面により変わる変数と定数、演算子 (加算、乗算) によって構成される。相対的に優劣のついた局面対を事例とし、事例の集合にに対し、各評価関数が優劣を正しく判定できた事例の個数の割合が、GP における各個体の適応度となる。

連絡先: 河口 信夫 名古屋大学大学院工学研究科計算理工学専攻
〒464-01 名古屋市中区千種区不老町

Tel: 052-789-3630 Fax: 052-789-3800

E-mail: kawaguchi@mie.nagoya-u.ac.jp

3 局面から直接得られる情報による学習

エキスパートの知識を用いた局面情報 (例えばオセロの場合、角の駒に対して重みが大いこと) を用いる事により、適切な評価関数の獲得が可能なることはすでに確認されている [堀之内 97]。しかし、エキスパートの知識は完全ではなく、任意のゲームについてエキスパートが存在するとは限らないため、この手法には限界がある。そこで、我々は局面的情報をそのまま GP の変数として入力する手法を行い、評価関数の獲得が可能であるかを確かめた。この場合、変数をどのように与えるか、すなわちコーディングの方法が問題になる。我々は対象ゲームとして tic-tac-toe を用い、以下の2種類のコーディングについて実験を行った。

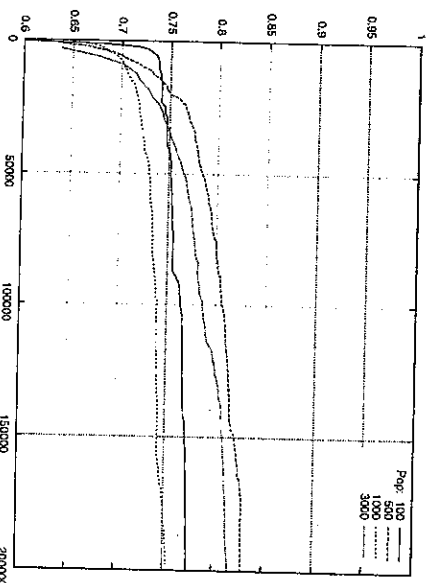


図 1: 直接情報による学習 (コーディング 1)

コーディング 1 各バスの {○, ×, (空白)} の状態に対し、それぞれ {1, -1, 0} を割り当てる変数 9 種類を用いる。

コーディング 2 各バスに対し、{○, ×, (空白)} の存在に対し、それぞれ {1, 0} を割り当てる変数 27 種類を用いる。

評価関数は加算・乗算からなる多項式である。コーディング 1 では変数の数は少ないが、あるバスの ○ と × が異なるだけで符号が反転し、各項に与える影響が大きい。コーディング 2 では各変数の影響は小さいが、逆にあるバスの状態を調べるために、3 つの変数を参照する必要がある。

表 1: 全解探索による勝率

	勝ち	引分け	負け
C1 先手番	76(0.73)	13(0.13)	15(0.14)
C1 後手番	336(0.60)	53(0.09)	176(0.31)
C2 先手番	68(0.90)	3(0.04)	5(0.06)
C2 後手番	338(0.61)	103(0.19)	110(0.20)
WK 先手番	72(0.95)	4(0.05)	0(0.00)
WK 後手番	402(0.69)	176(0.30)	7(0.01)

実験のパラメータは、交叉率 0.7, 突然変異率 0.3 に固定し、集団規模を 100 ~ 3000 と変化させ、定常状態再生を行った。5 回試行の平均の結果を図 1, 図 2 に示す。図の横軸は生成された個体の数、縦軸は適応度の値を示す。

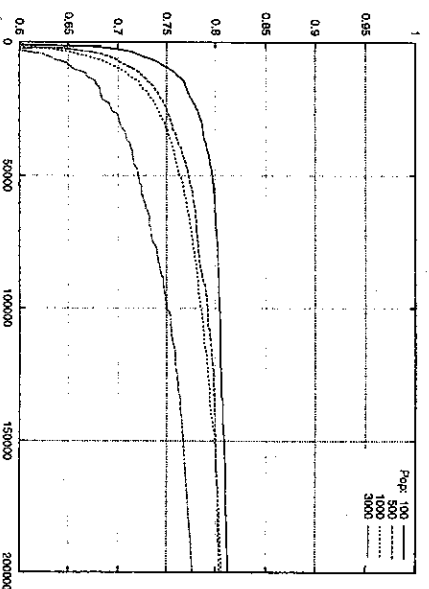


図 2: 直接情報による学習 (コーディネング 2)

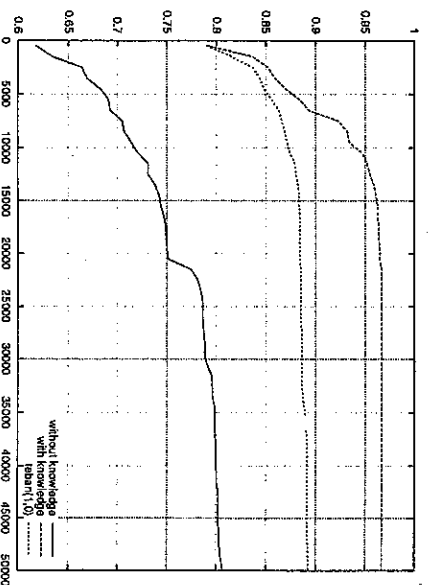


図 3: 知識を用いた学習効率の向上

コーディネング 1 では各試行において値のばらつきが大きく、学習の効果が集団規模に比例しない。一方、コーディネング 2 では学習の効果が集団規模と相関があり、集団規模が大きいと学習がゆっくりと進むことがわかる。これは、コーディネング 1 では、各変数の影響が大きいため、局所解

に陥る可能性が高く、コーディネング 2 では、変数の個数が多いため、学習効率が悪いことを示していると考えられる。

一方、エキスパートの知識を用いた局面情報を用いた学習では、学習効率は格段に向上する。図 3 では、直接得られる情報のみを用いた場合に比べ、知識を用いることにより、高い適応度を早い時点で達成していることが確認できる。また、teban(1.0) とかかれたグラフは、直接情報に手番の情報に加えて学習を行った場合である。手番の情報を付加するだけでも、学習効率が向上することが確認できた。

4 評価

最終的に得られた評価関数を、全解探索を行うプログラムと対戦させることにより評価を行った。評価結果を表 1 に示す。ここで、C1, C2 はそれぞれ直接得られる情報による学習のコーディネング 1, 2 を、WK は知識を用いた情報による学習によって得られた評価関数を表す。知識を用いた場合には、先手番では負けることのない強い評価関数が得られていることがわかるが、直接情報によっても、ある程度の強さの評価関数が得られることが確認できた。特にコーディネング 2 のような単純なコーディネングが有効であることが確かめられた。

5 まとめ

本稿では、遺伝的プログラミングを用いてゲームの静的局面評価関数を自動的に獲得する手法を提案した。特に、局面から直接得られる情報のみを用いても、効率は必ずしも良くないが適切な評価関数が獲得できることを確認した。今後の課題としては、コーディネングの工夫による探索の効率化、得られた評価関数から有用な局面情報の獲得、より複雑な他のゲームへの適用等が挙げられる。

参考文献

- [Koza 92] Koza, J.R.: Genetic programming : On the Programming of Computers by Means of Natural Selection, The MIT Press, 1992.
- [Koza 94] Koza, J.R.: Genetic Programming II: Automatic Discovery of Reusable Programs, The MIT Press, 1994.
- [堀之内 96] 堀之内, 河口, 稲垣: 遺伝的プログラミングを用いたゲームの局面評価関数の生成, 第 10 回人工知能学会全国大会, pp. 649-652, 1996.
- [堀之内 97] 堀之内, 河口, 稲垣: 遺伝的プログラミングを用いたゲームの局面評価関数の学習, 信学技報, KIBSE96-35, pp. 17-24, 1997.
- [松原 97] 松原: なぜチェス名人はコンピュータチェスに負けなければならなかったのか?, 情報処理, Vol 38, No.8, pp. 705-706, 1997.
- [松原 93] 松原: ゲームのアルゴリズム, コンピュータソフトウェア, Vol.10, No.6, pp.3-18, 1993.
- [Pell 92] Pell, B.: METAGAME: A new challenge for games and learning, *Heuristic Programming in Artificial Intelligence 3 - The Third Computer Olympiad*, Ellis Horwood, 1992.
- [Tesauro 95] Tesauro, G.: Temporal difference learning and TD-gammon, *CACM*, Vol.38, No.3, pp.58-68, 1995.