

○吉井伸一郎（北大工） 鈴木恵二（北大工） 嘉数侑昇（北大工）

要旨

ラマルク進化による形質獲得は自然界においては確認されていないが、適応的な戦略の獲得を目的とした系において有効視されている。すなわち、ラマルク進化により、個体の環境への収束度を向上させることが可能となる。本研究では、ラマルク進化の方向性を遺伝的に決定することにより、適応的に個体の収束度の向上を図ることを目的とし、特にジョブショップスケジューリング問題における有効性を検証する。

1 はじめに

遺伝的アルゴリズム(GA)は、交叉、及び突然変異といった遺伝的操作によりランダムな方向へ情報を多様化させ、淘汰により、環境に対して有効な形質を獲得する。一方、適応的戦略獲得を目的とした系において、環境からの評価を直接各個体の戦略に反映させ、個体の環境への自己適応を目的としたラマルク進化の導入が図られている[Grefenstette].

そこで、本研究ではラマルク進化を導入したGAを最適化問題へと適用し、その効果について検証する。しかし、ラマルク進化を一種の強化学習として実現した場合、適用する問題空間に応じて収束の方向性を決定する必要がある。この問題を避けるため、本研究はラマルク進化の方向性を遺伝的に決定することにより、この解決を試みる。

2 最適化問題へのラマルク進化の適用

本研究においては、最適化問題への適用を目的として、行動選択テーブル(Lookup-Table;LUT)を遺伝子として持つ個体からなる集団を考える。LUTは適応的振る舞いを実現する手法の一つであり[Horn]、GAによる適応的な戦略の獲得も成されている[Iba]。ここではこれらをLUT-GAと呼び、最適化問題の一つであるジョブショップスケジューリング問題への適用を考える。 n 個のジョブ J_i ($i=1, \dots, n$) が m 台の機械 M_r ($r=1, \dots, m$) で処理される時、個体のLUTである S を次のように表す。

$$S = \{S_1, S_2, \dots, S_m\}. \quad (1)$$

$$S_r = \{S_r^1, S_r^2, \dots, S_r^n\}. \quad (2)$$

$$S_r^i = \{a_{ij}^1, a_{ij}^2, \dots, a_{ij}^n\}. \quad (3)$$

$$a_{ij}^r = [p, q] \in R. \quad (4)$$

S_r は各機械ごとのLUTであり、個体が機械 M_r 上でジョブ J_i の次に作業するジョブを決定するための S_r^i からなっている。 a_{ij}^r は個体が機械 M_r 上でジョブ J_i の次にジョブ J_j をスケジューリングする度合いを表す。

このLUTとGifferとThompsonによるアクティブスケジューリング生成アルゴリズム(GT法)を対応させ、各個体のスケジューリングを生成する。すなわち、GTアルゴリズムによって求められるコンフリクト集合において、どのジョブをスケジューリングするかを各個体のLUTにより決定する。こうして、各個体の適応度をそれぞれのLUTをもとに生成されたスケジューリングにより計算し、その後、適応度の高い個体の選択を行ない、各LUTに遺伝的操作を施す。

次に、このLUT-GAに強化学習的なラマルク進化を導入する。強化則として、現在のLUTに基づいて生成したスケジューリングが前回に生成したスケジューリングと比較して、良ければ今回生成したスケジューリングを生成する確率を上げ、悪ければ確率を下げるものとする。すなわち、機械 M_r 上での k 番目に処理されるジョブ J_k を $J_{i,k,r}$ とすると、今回のスケジューリングに対応するLUTの要素 $a_{\kappa, \kappa+1}^r$ ($\kappa, \kappa+1$ はジョブ $J_{i,k,r}, J_{i,k+1,r}$) について

$$a_{\kappa, \kappa+1}^r(t+1) = a_{\kappa, \kappa+1}^r(t) + \Delta. \quad (6)$$

ただし、

$$\Delta = \begin{cases} bs(q - a_{\kappa, \kappa+1}^r(t)) & (T_t < T_{\min}) \\ bt(q - a_{\kappa, \kappa+1}^r(t)) & (T_t < T_{t-1}) \\ wr(p - a_{\kappa, \kappa+1}^r(t)) & (T_t > T_{t-1}) \end{cases} \quad (7)$$

ここで、

T_t : 個体が t 回目に生成したスケジューリングの総所要時間

T_{\min} : 現在までの最小所要時間

bs, bt, wr : ラマルク進化の度合いを決定するパラメーターである。 bs と bt は、今回生成したスケジューリングの評価が高かった場合にそのスケジューリングを生成する確率を高くする、リワードに関するパラメーターであり、反対に、 wr は評価が低い場合にそのスケジューリングを生成する確率を低くするペナルティに関するパラメーターである。 bs, bt, wr の定義域は、

$$bs = [0.0, 1.0]. \quad (8)$$

$$bt = [0.0, 1.0]. \quad (9)$$

$$wr = [-1.0, 0.0]. \quad (10)$$

であり、絶対値が大きいほどラマルク進化の度合いは強くなる。

3 ラマルク進化の効果と問題点

まず、ラマルク進化の効果を検証するためにテストケースとして、最適解の知られている問題を用いて実験を行った。ここで、先に説明したラマルク進化の度合いを決定するパラメータを何回かの予備実験により、Table1のように設定した。Fig.1及びFig.2に 10×5 問題、 10×10 問題の実験結果を示す。

Fig.1, Fig.2からわかるように、 10×5 問題においてはcase1のパラメータセットを用いたLUT-GAの収束度は向上しているものの、 10×10 問題においては良い結果は得られなかった。一方、case2のパラメータセットについては、case1と反対の結果となった。

	<i>bs</i>	<i>bt</i>	<i>wr</i>
case1	0.8	0.3	0.1
case2	0.5	0.2	0.3

Table1 Parameters of Lamarckian Evolution.

異なるパラメータセットによるこのような収束度の違いは、問題空間に依存すると考えられる。すなわち、今回用いた10×5問題は、コンフリクトとなるジョブが少なく、スケジューリングの結果に対してクリティカルな作業が全ての機械に分散している。そのため、リワードに関するパラメータが大きいほど探索空間を素早く絞り込むことができるため、結果として収束度が向上している。一方、10×10問題は、その技術的順序から、ある機械上でコンフリクトとなるジョブが多いため、case1のようなパラメータセットでは、探索の初期段階において局所解に陥り、その後の探索はあまり進まない。逆に、case2のようにペナルティに関するパラメータを大きくすることで、多くのコンフリクト集合の中から誤った方向への探索を避け、10×5問題とは反対の形で探索を行なうことで探索空間を絞り込み、その結果、個体の収束度が向上したものと考えられる。

このように、多くの制約条件があり、複雑な探索空間を有するジョブショップスケジューリング問題において、効果的な適応を図るためには、問題空間に応じて個体の学習の方向を決定する必要があると考えられる。

4 適応的学習パラメータ機構の導入

これまでのラマルク進化を導入したLUT-GAは、予め対象とする問題空間において、望ましいと思われる方向へ学習を導いていたことから、一種の教師あり学習である。一方、予め適切な学習の方向が予測できない問題領域において、または、種としての進化と個体の学習という問題に換して、個体の学習の方向を遺伝的に監視しようという研究が報告されている[Ackley, Littman]。本研究では、この問題を解決するため、LUTを持つ集団(Pop_{LUT})とは別に、学習の方向を遺伝的に決定するもうひとつの集団(Pop_{para})を用意する。 Pop_{para} は Pop_{LUT} のラマルク進化のパラメータを遺伝子として持ち、 Pop_{LUT} のある個体は生まれながらに Pop_{para} のある個体の表現型に従って学習を行なうも

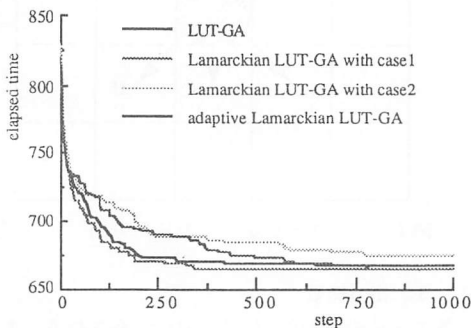


Fig.1 Experimental results of 10×5problem.

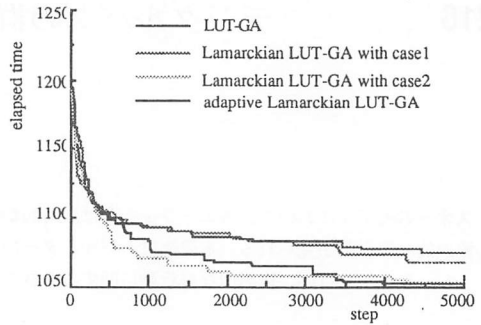


Fig.2 Experimental results of 10×10problem.

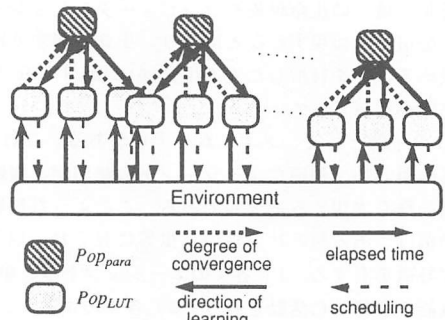


Fig.3 Adaptive learning mechanism.

のとす。また、 Pop_{para} は自分のパラメータを利用している Pop_{LUT} の収束度によって評価される。こうして、二つの集団は、別々の評価関数により評価され、相互に依存し合っている。Fig.3に適応的学習パラメータ機構の概念図を示す。

実験結果を、同じくFig.1及びFig.2に示す。これにより、予め予備実験によりパラメータを設定した場合に比較して優れた結果は得られなかったものの、異なる問題空間においても安定した結果を得られることが確認された。

5 おわりに

本研究では、ラマルク進化による適応的戦略獲得において、その学習の方向をも適応的に決定することにより、問題空間に依存せずに個体の収束度が向上することが確認できた。

参考文献

- [Grefenstette] John J. Grefenstette: Lamarckian Learning in Multi-agent Environments, *Proceeding of 3rd ICGA*, 1989.
- [Horn] Jeffrey Horn: Measuring the Evolving Complexity of Stimulus-Response Organisms, *Toward a Practice of Autonomous Systems*, The MIT Press, 1991.
- [Iba] Hitoshi Iba: Bugs: A Bug-Based Search Strategy using Genetic Algorithms, *Parallel Problem Solving from Nature 2*.
- [Ackley, Littman] David Ackley and Michael Littman: Interactions Between Learning and Evolution, *Artificial Life II*, 1992.