

TDGA における 2 遺伝子座ごとのエントロピー評価に関する一考察

1 概要

本稿では、森ら [1] が提案した GA における選択プロセスに熱力学的なエントロピーと温度の概念を導入した熱力学的遺伝アルゴリズム (Thermodynamical Genetic Algorithm: TDGA) の機能拡張に関する考察を述べる。具体的には、TDGA 特有の問題である TDGA 収束に対するアプローチとして森らが指摘した、エントロピーを 2 遺伝子座ごとに評価する手法について検討する。

2 実装上の変更点

実装上は、エントロピーを計算する際の個体表現を変更した。具体的には、対立遺伝子を 2 値 $\{0,1\}$ とした場合に隣接する 2 つの遺伝子座の遺伝子を 2 進数表現として捉え、10 進数表現に変換した。これによって、対立遺伝子は 4 値 $\{0,1,2,3\}$ となる。また、変換前の遺伝子長を M とすると、変換後の遺伝子長は $M-1$ となることに注意する。例えば以下のように、個体 x_1 は x'_1 に変換される。

$$x_1: \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0$$

$$x'_1: \quad 2 \quad 1 \quad 3 \quad 2 \quad 0 \quad 1 \quad 2$$

この変更に伴い、エントロピーの計算式は以下のようになる。

$$H^2 = \sum_{k=1}^{M-1} H_k^2, \quad H_k^2 = - \sum_{j \in \{0,1,2,3\}} P_j^k \log P_j^k \quad (1)$$

ここで、 H^2 は 2 遺伝子座ごとに計算したエントロピーを表し、 M は対立遺伝子を 2 値 $\{0,1\}$ とした場合の遺伝子長である。また、 H_k^2 は個体群の遺伝子座 k の遺伝子に関するエントロピーを、 P_j^k は遺伝子座 k における対立遺伝子 j の存在確率を表している。

3 評価実験

本章では、エントロピーを 2 遺伝子座ごとに評価した TDGA を、シミュレーションによって従来の TDGA および SGA と比較することにより、その有効性について検討する。

3.1 実験条件

シミュレーションの条件については森ら [1] に従い、表 1 に示す 30 荷物のナップサック問題を扱う。具体的には各アルゴリズムについて突然変異率および個体数を変え、30 回問題を解かせた場合に 100 世代以内に最適解を獲得した回数を調べる。異なる突然変異率を設定する場合、個体数として TDGA では 32, SGA では 64 を与えた。この実験においてそれぞれのアルゴリズムに対して最も多く最適解を得たときの突然変異率を、個体数を変化させた場合の突然変異率として設定する。また、いずれの場合も初期温度 $T = 10$ で設定した。

3.2 結果と考察

まず、図 1 に異なる突然変異率を設定した場合に各アルゴリズムが最適解を得た回数の比較を示す。図 1 から、2 遺伝子座ごとにエントロピーを計算する場合 (TDGA (bi-allele)) は、1 遺伝子座ごとにエントロピーを計算する場合 (TDGA) よりも最適解が得られる回数が少ないことがわかる。

次に、図 2 に異なる個体数を設定した場合の各アルゴリズムが最適解を得た回数の比較を示す。ここでそれぞれのアルゴリズムに対して使用した突然変異率は、TDGA では 0.005, TDGA (bi-allele) では 0.02, SGA では 0.02 である。図 2 からわかる通り、異なる個体数を設定した場合でも、TDGA (bi-allele) は TDGA の最適解獲得回数を下回った。

図 3～図 4 に各アルゴリズムを用いた場合の解の探索過程について示す。図 3 と図 4 を比較すると、2 遺伝子座ごとにエントロピーを評価する方が、1 遺伝子座ごとに評価するよりも目的関数値の最高値が最適解に到るまでに要する世代数が多いことがわかる。

また、図 5～図 6 に各アルゴリズムを用いた場合の各遺伝子座におけるエントロピーを示す。これらは個体群の多様性を維持する様子を表現した図である。図 5 と図 6 を比較すると、2 遺伝子座ごとにエントロピーを評価する方が、1 遺伝子座ごとに評価するよりも全体的にエントロピーが高いことが確認できる。

表 1: 30 荷物ナップサック問題詳細 ([1] より引用)

$\max_{x_i} z = \sum_i c_i x_i$ $\text{sub. to } \sum_i a_i x_i \leq b, \quad x_i \in \{0,1\}$															
i	1*	2*	3*	4*	5*	6*	7*	8*	9*	10*	11*	12*	13*	14*	15*
c_i	75	84	58	21	55	95	28	76	88	53	58	81	32	89	54
a_i	7	9	13	5	16	28	15	43	60	37	44	63	34	95	61
i	16+	17*	18*	19*	20	21+	22	23	24	25	26	27	28	29	30
c_i	23	42	52	58	53	30	26	40	40	26	39	25	23	16	12
a_i	29	57	72	83	84	48	45	74	78	52	79	64	64	55	74
$b = 744$ *:the optimal solution +:the solution by the greedy algorithm															

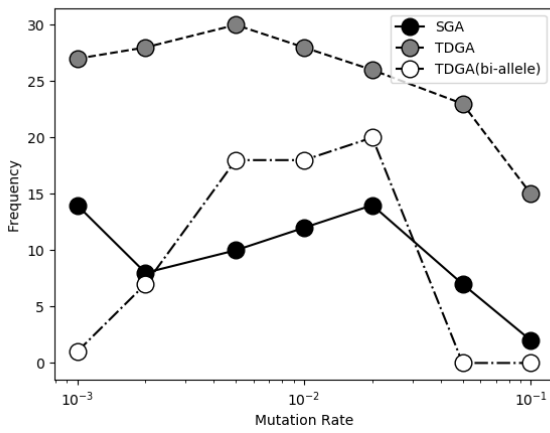


図 1: 突然変異率の比較

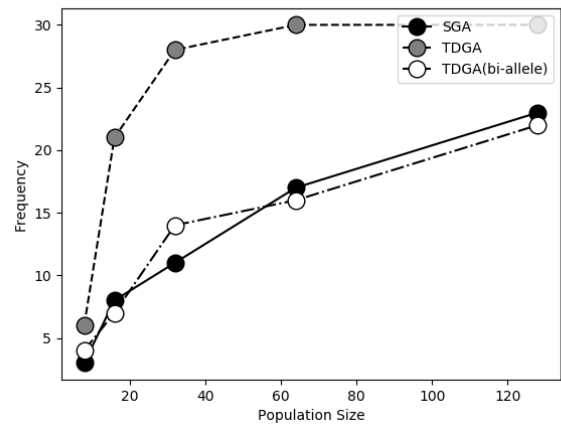


図 2: 個体数の比較

以上のことから、2 遺伝子座ごとにエントロピーを評価することで、より個体群の多様性を維持するようになることが考察される．そして今回対象とした問題では、その性質が解の探索を必要以上に複雑にし、結果として最適解を獲得する可能性を減少させたと考えられる．

4 おわりに

本稿では、TDGA 収束の対策として挙げられた 2 遺伝子座ごとにエントロピーを計算することの有効性について検討した．30 荷物ナップサック問題を対象としたシミュレーションの結果、2 遺伝子座ごとにエントロピーを計算する手法は、従来の 1 遺伝子座ごとにエントロピーを計算する手法よりも低い性能を示した．このことから、この問題に対しては 1 遺伝子座ごとにエントロピーを評価の方が適切であるといえる．実験において、従来の TDGA が

30 回すべての試行において最適解を獲得する場合も確認されたことから、今回扱った問題では TDGA 収束は起こりにくいことも考えられる．そのため、TDGA 収束が発生しやすい問題を対象として解かせた場合に、2 遺伝子座ごとにエントロピーを評価することの妥当性について、今後検討する必要がある．

参考文献

- [1] 森 直樹, 吉田 潤二, 喜多 一, 西川 禎一. 遺伝アルゴリズムにおける熱力学的選択ルールの提案. システム制御情報学会論文誌, 9(2):82-90, 1996.

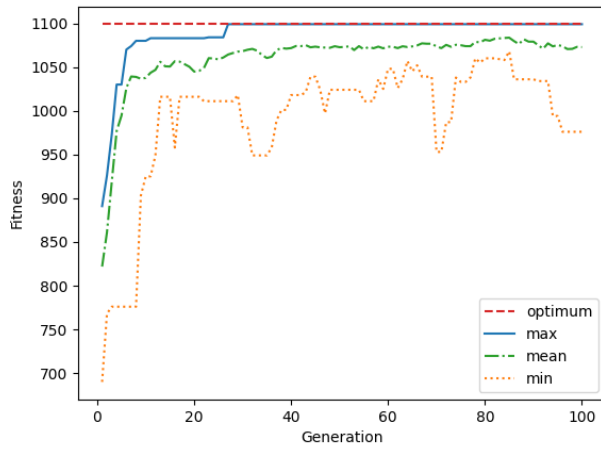


図 3: TDGA を用いた場合の解の探索過程

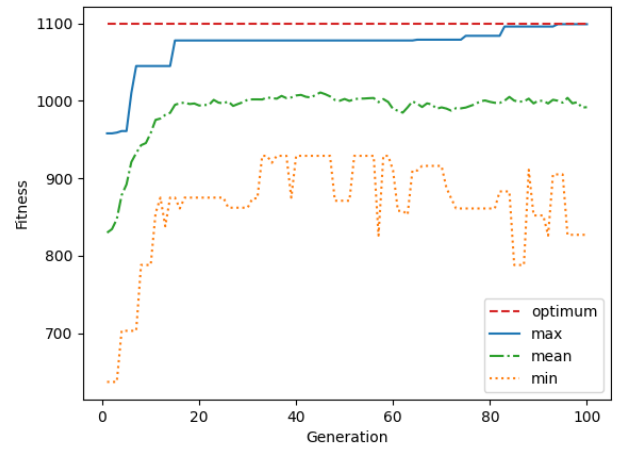


図 4: TDGA (bi-allele) を用いた場合の解の探索過程

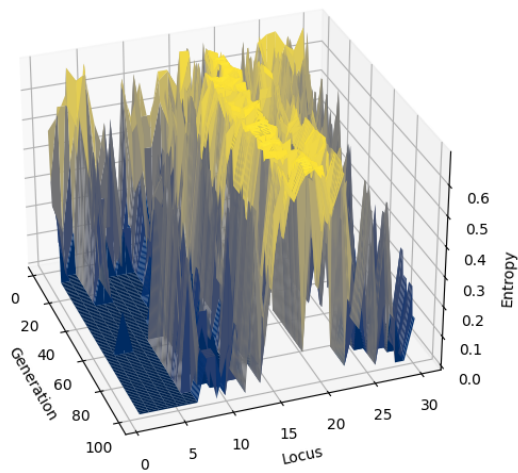


図 5: 各遺伝子座におけるエントロピー (TDGA)

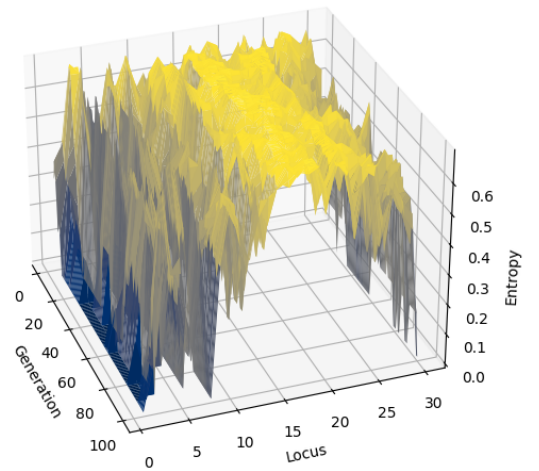


図 6: 各遺伝子座におけるエントロピー (TDGA (bi-allele))