

確率モデルGAの順序問題への一適用法
 Probabilistic Model-Building Genetic Algorithms in Permutation Representation Domain

A-41

筒井茂義†
 Shigeyoshi Tsutsui

1. はじめに

近年、確率モデルGA (PMBGAs: probabilistic model-building genetic algorithms)の研究が活発に行われている¹⁾。確率モデルGAの研究は、バイナリコーディングGAでの研究から始まり、近年実数値問題への拡張の研究が行われている。しかし、順序問題への研究は始まったばかりである。本稿では、確率モデルGAの順序問題への一適用法としてエッジヒストグラムを用いるEHBSA (edge histogram based sampling algorithm)について述べる⁴⁾。

2. 確率モデルGAの概要

確率モデルGAは、従来のGAと異なり、集団における個体の分布に着目する。ある世代において、探索空間において個体が密となっている領域は確率的に有望な探索領域、粗となっている領域は確率的に有望でない探索領域と考える。有望と考えられる領域にはより多くの子個体を生成し、有望でないと考えられる領域にはより少ない子個体を生成する。これを実現するため、確率モデルGAでは、集団の個体分布の確率モデルを生成し、この確率モデルに基づいて子個体を生成する。生成された子個体は評価され、選択オペレータにより次世代の集団が形成される。

このように確率モデルGAでは、従来のGAにおける子個体の生成方法である交叉オペレータが「集団における個体の確率分布の推定とそれに基づくサンプリング」に置き換わる。この点以外は、従来のGAと全く同じであり、各種の選択オペレータや世代交代モデルを用いることが出来る。確率モデルGAは、バイナリコーディングGAでの研究が先行している。BOAなどがその代表例である²⁾。実数値コーディングへの確率モデルGAの研究も近年行われている³⁾。しかし、順序表現問題への確率モデルGAの研究は始まったばかりである。

3. EHBSAの構成

エッジは、順序表現ストリングにおけるノード間の連結関係のことであり、ストリングに関する重要な情報である。このエッジ情報を用いる交叉として、エッジ交叉(ER)や拡張エッジ交叉(eER)がよく知られている。しかしこれらは、2つの親のエッジしか考慮しない。本論文で提案するEHBSA (edge histogram based sampling algorithm)では、集団を構成する全個体のエッジの分布情報であるエッジヒストグラムマトリックス(EHM)を用いる。

EHBSAの構成は以下の通りである。

1. 個体をランダムに生成し、初期集団を作る。
2. 選択オペレータを適用し、有望集団を作る。
3. この集団からEHMを作成する。

4. このEHMを基に子個体をサンプリングする。

5. 終了条件が満たされるまで、ステップ2からステップ4までを繰り返す。

3.1 エッジヒストグラムマトリックス(EHM)

本稿では対称TSPのようなエッジに方向性を持たない場合を考える。世代*t*の集団*P(t)*における*k*番目の個体を*s_k* = ($\pi_k^t(0), \pi_k^t(1), \dots, \pi_k^t(L-1)$)で表すと、これは(0, 1, ..., L-1)の順列となる。集団*P(t)*のエッジヒストグラムマトリックスEHM^t (e_{ij}^t) ($i, j = 0, 1, \dots, L-1$)は、以下のようにL²の要素を持つ対称マトリックスで定義する。

$$e_{i,j}^t = \begin{cases} \sum_{k=1}^N (\delta_{i,j}(s_k^t) + \delta_{j,i}(s_k^t)) + \varepsilon & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (1)$$

ここで*N*は、集団サイズであり、 $\delta_{ij}(s_k^t)$ は、以下で定義されるデルタ関数である。また、 ε ($\varepsilon > 0$)は、式(3)の*B_{ratio}* (*B_{ratio}* > 0)で定義される定数である。

$$\delta_{i,j}(s_k^t) = \begin{cases} 1 & \text{if } \exists h [h \in \{0, 1, \dots, L-1\} \wedge \pi_k^t(h) = i \\ & \wedge \pi_k^t((h+1) \bmod L) = j] \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$\varepsilon = \frac{2N}{L-1} B_{\text{ratio}} \quad (3)$$

なお、ここでは対称EHMを定義したが、スケジューリング問題のようにエッジに方向性がある場合は、非対称EHMとして同様に定義することができる。EHMの一例を図1に示す。

$$\begin{matrix} s_1^t = (0, 1, 2, 3, 4) \\ s_2^t = (1, 3, 4, 2, 0) \\ s_3^t = (3, 4, 2, 1, 0) \\ s_4^t = (4, 0, 3, 1, 2) \\ s_5^t = (2, 1, 3, 4, 0) \end{matrix} \quad \begin{pmatrix} 0 & 3.1 & 2.1 & 2.1 & 3.1 \\ 3.1 & 0 & 4.1 & 3.1 & 0.1 \\ 2.1 & 4.1 & 0 & 1.1 & 3.1 \\ 2.1 & 3.1 & 1.1 & 0 & 4.1 \\ 3.1 & 0.1 & 3.1 & 4.1 & 0 \end{pmatrix}$$

(a) *P(t)*

(b) EHM

図1 EHMの一例 (*N* = 5, *L* = 5, *B_{ratio}* = 0.04)

3.2 サンプリング法

EHMに基づいたサンプリング法として、テンプレートをを用いない方法とテンプレートをを用いる方法の二つの方法を提案する。

3.2.1 EHBSAWO

テンプレートをを用いないEHBSA/WO (edge histogram based sampling algorithm without template)では、EHMに基づくルーレットにより、各ノードを順次サンプリングし、ストリングを生成する。

3.2.2 EHBSAWT

テンプレートをを用いるEHBSA/WT (edge histogram based sampling algorithm with template)では、まず、集団から一つ

† 阪南大学経営情報学部

の個体をランダムに選びテンプレートとして利用する。このテンプレートをランダムに決めた n 点 ($n > 1$) のカットポイントを適用して n 個のセグメントに分割する。この n 個のセグメントからランダムにひとつを選び、そのセグメントのノードのみを EHM に基づいてサンプリングを行う。残りの $n-1$ のセグメントに属するノードはそのまま新しい個体のノードとして用いる。このサンプリング法を EHBSA/WT n と記述する。図 2 に EHBSA/WT/3 の例を示す。

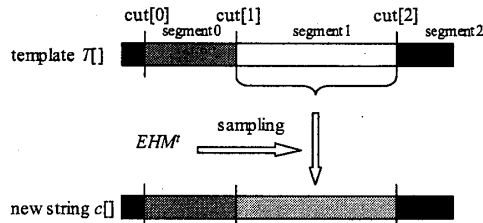


図 2 EHBSA/WT/3

4. 実験結果

4.1 世代交代モデル

EHBSA/WT に用いる世代交代モデルは以下の通りである。まず、個集団からテンプレート個体 $T[]$ をランダムに選ぶ。次に $T[]$ と EHM とから 3.2.2 項で述べた方法により子個体 $c[]$ を生成する。ここで、もし個体 $c[]$ の方が個体 $T[]$ よりも優れているときにのみ個体 $c[]$ を個体 $T[]$ と交換する。EHBSA/WO 用の世代交代モデルも基本的に同じである。本実験では比較のため、従来の順序表現向き交叉オペレータとして、PMX, OX, eER も取り上げた。これらの交叉オペレータ用の世代交代モデルも EHBSA/WT と類似なものにするため、GENITOR アルゴリズムと同様、2 親から一つの子個体を生成することとした。

4.2 テスト問題と結果

EHBSA を評価するためにここではよく知られた TSP を用いる。ここでは、gr24, gr48 および pr76 を用いた。なお、この実験では純粋に提案アルゴリズムの評価を行うため、突然変異オペレータやヒューリスティックスは適用しない。評価は、#OPT (10 回の実験中で最適ルートを発見した回数)、ANE (最適解を発見するのに要した評価回数の平均) および Aver (各実験の最良解の平均) で行う。

gr24 の結果を表 1 に示す。EHBSA/WO は、集団サイズ $N = 60, 120, 240$ に対して最適解をそれぞれ 7, 9, 6 回であるのに対して EHBSA/WT/ n はいずれの実験も最適解を 10 回とも見つけている。ANE の値も EHBSA/WO よりも優れている。交叉オペレータの中では、eER がいい結果を示しているが、EHBSA との比較では、EHBSA/WT/ n の方が優れている。EHBSA/WT と eER との大きな違いは、EHBSA/WT は、eER よりも小さな集団サイズで探索性能が得られることである。

gr48, pr76 の結果は、それぞれ表 2, 表 3 の通りである。この結果の傾向は、gr24 の結果とほぼ同じである。ただし、問題が大きくなるにしたがってカットポイントの数 n が大きいほどよい結果が得られる傾向にある。

表 1 gr24 の結果

Model	Population Size N															
	60			120			240			480			960			
	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	
EHBSA/WO	7	16328	1281	9	23637	1273	6	44853	1280	-	-	-	-	-	-	-
EHBSA/WT/2	10	9141	1272	10	17978	1272	10	35604	1272	-	-	-	-	-	-	-
EHBSA/WT/3	10	9523	1272	10	18251	1272	10	32956	1272	-	-	-	-	-	-	-
EHBSA/WT/4	10	10677	1272	10	17652	1272	10	33606	1272	-	-	-	-	-	-	-
EHBSA/WT/5	10	11170	1272	10	20489	1272	10	36078	1272	-	-	-	-	-	-	-
OX	0	-	1345	1	22449	1303	4	34140	1296	11	48674	1301	0	-	-	1484
eER	1	4738	1299	7	6237	1276	10	13394	1272	10	23785	1272	10	42767	1272	1272
PMX	0	-	1492	0	-	1414	2	23191	1341	11	49442	1316	0	-	-	1572
Other PMBGA*	#OPT=0, ANE is not available, best length = 1328 with MIMIC, best Aver = 1439 with EBNA															
Optimum: 1272																
$E_{max} = 50000, B_{ratio} = 0.03$																

表 2 gr48 の結果

Model	Population Size N															
	60			120			240			480			960			
	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	
EHBSA/WO	0	-	5212	0	-	5316	0	-	5773	-	-	-	-	-	-	-
EHBSA/WT/2	4	102691	5053	10	174125	5046	10	299391	5046	-	-	-	-	-	-	-
EHBSA/WT/3	10	85387	5046	10	134597	5046	10	240391	5046	-	-	-	-	-	-	-
EHBSA/WT/4	9	82701	5047	10	126444	5046	10	237260	5046	-	-	-	-	-	-	-
EHBSA/WT/5	10	89799	5046	10	157041	5046	10	257486	5046	-	-	-	-	-	-	-
OX	0	-	5527	0	-	5268	0	-	5200	11	162154	5099	2	287852	5082	5082
eER	0	-	5653	0	-	5233	0	-	5098	2	95075	5072	3	166286	5058	5058
PMX	0	-	8285	0	-	7374	0	-	6859	0	-	6116	0	-	-	5860
Other PMBGA*	#OPT=0, ANE is not available, best length = 6104 with MIMIC, best Aver = 6717 with MIMIC															
Optimum: 5046																
$E_{max} = 500000, B_{ratio} = 0.015$																

表 3 pr76 の結果

Model	Population Size N															
	60			120			240			480			960			
	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	#OPT	ANE	Aver	
EHBSA/WO	0	-	119136	0	-	128208	0	-	142206	-	-	-	-	-	-	-
EHBSA/WT/2	4	360128	108352	9	457147	108174	7	871319	108201	-	-	-	-	-	-	-
EHBSA/WT/3	4	248091	108385	9	472719	108171	8	853801	108201	-	-	-	-	-	-	-
EHBSA/WT/4	9	341482	108247	9	607544	108247	0	-	108496	-	-	-	-	-	-	-
EHBSA/WT/5	10	494674	108159	10	797963	108159	0	-	108807	-	-	-	-	-	-	-
OX	0	-	129603	0	-	121642	0	-	116591	0	-	113412	0	-	-	112259
eER	0	-	142003	0	-	122217	0	-	111839	11	109119	3	394887	108507	108507	108507
PMX	0	-	236827	0	-	213528	0	-	187601	0	-	164883	0	-	-	158515
Other PMBGA*	not available															
Optimum: 108159																
$E_{max} = 1000000, B_{ratio} = 0.005$																

5. むすび

以上本稿では、確率モデル GA の順序問題への適用の一手法として、エッジヒストグラムを用いる方法を提案し、TSP を用いて評価を行った。エッジヒストグラムから個体を生成するサンプリングアルゴリズムとして EHBSA/WO と EHBSA/WT の二つの方法を提案したが、EHBSA/WT が優れた性能を持っていることが明らかになった。また、EHBSA/WT 従来のよく知られた順序向きの交叉オペレータよりも少ない集団サイズで優れた性能を持つことも明らかになった。今後は、スケジューリング問題など一般的な問題で評価を進めていく予定である。

参考文献

- 1) Pelikan, M., Goldberg, D. E., and Lobo, F. G.: A survey of optimization by building and using probabilistic models, *Technical Report IlliGAL Report 99018*, University of Illinois at Urbana-Champaign (1999).
- 2) Pelikan, M., Goldberg, D. E., and Cantu-Paz, E.: BOA: The Bayesian optimization algorithm, *Proc. of the GECCO-99* (1999).
- 3) 筒井, M. Pelikan, D. E. Goldberg: 周辺分布ヒストグラムを用いた実数値確率モデル遺伝的アルゴリズムの一構成法, 情報処理学会論文誌「数理モデル化と応用」(掲載予定).
- 4) Tsutsui, S.: Probabilistic Model-Building Genetic Algorithms in Permutation Representation Domain Using Edge Histogram, *Proc. of the PPSN VII* (2002, to appear).