

## 動的環境問題における BOA への混合ベジアンネットワーク導入の効果について

堀伸哉<sup>†</sup> , 棟朝雅晴<sup>††</sup> 赤間清<sup>††</sup>

BOA(Bayesian Optimization Algorithm)はベジアンネットワークと条件付確率を用いて問題の構造を詳細に表現することで、広範囲の問題を解くことのできる最適化アルゴリズムである。BOAはその特性から広範囲の問題を解くことができるが、環境が動的に変化するような問題に対しては、ベジアンネットワークが一方の環境に対して収束し、結果として動的環境に対応できないという問題点がある。本論文は BOA に混合ベジアンネットワークの概念を導入した BOA with Mixture Distribution(BOA-MD) に新しい追加要素を導入することでそのような問題の解決を目指す。そして、その効果を考察することによって BOA の問題解決領域を広げることが目的とする。

### The effect of implementation of mixtures of Bayesian Network to dynamic environment problem

Shinya Hori<sup>†</sup>, Masaharu Munetomo<sup>††</sup>  
and Kiyoshi Akama<sup>††</sup>

Bayesian Optimization Algorithm (BOA) can solve wide-spectrum of optimization problems by modeling their probabilistic models with conditional probabilities based on Bayesian networks. BOA succeeds in solving a variety of difficult optimization problems, however, it has not been applied successfully to dynamic environment because it tend to converge at one specific network and cannot adapt to change of probabilistic distributions. In this paper, we propose a BOA that introduces mixture distributions and inheritance of probability distribution of previous generation to adapt to dynamic environment, and discuss effect of this introduction. Through these, this work motivates to broaden the optimization problem domain.

## 1. はじめに

BOA(Bayesian Optimization Algorithm)<sup>[1][2]</sup>は EDA(Estimation Distribution Algorithm)<sup>[2]</sup>の一つでベジアンネットワークを使用することで複雑な依存関係を有する問題の解を発見することができる最適化アルゴリズムである。BOA は GA(Genetic Algorithm)<sup>[3]</sup>などの他の進化計算アルゴリズムと比べて広い範囲の問題を解決することができ、現在最も注目されているメタヒューリスティックスの一つと言える。これまでの研究で BOA に混合ベジアンネットワークを導入した BOA-MD(BOA with Mixture Distribution)を提案した<sup>[11]</sup>。混合ベジアンネットワークは動的な環境での因果関係の記述や、因果関係を複数並列に学習させるために使用されるベジアンネットワークの混合分布であり、BOA に混合ベジアンネットワークを導入することで、複数の問題構造をもつような問題に対して BOA よりも良いパフォーマンスを見せることを確認した。本論文では動的環境問題に対して BOA-MD を適用する。動的環境問題は時間によって評価値を出す評価関数が増えるような問題で、BOA でこれを解くことは難しい。このような問題を解くために本論文では BOA-MD にさらに新たな確率分布継承の要素を追加することで動的環境問題に適応するようなアルゴリズムを提案し、BOA との比較実験を行った上でその効果を示すことで、BOA の拡張を行う。

## 2. BOA with Mixture Distribution (BOA-MD)

### 2.1 BOA

BOA はベジアンネットワークと条件付確率を利用して進化計算における遺伝子間の依存関係をモデル化することでビルディングブロックを破壊せずに複雑な問題の構造を記述、そしてその最適解を導出するような分布推定型の最適化アルゴリズムの一種である。進化計算において問題のビルディングブロックを破壊しないように探索を進めることが重要であるが、BOA は問題のビルディングブロックをベジアンネットワークと条件付確率で記述することで、問題の依存関係を表現、そして問題構造を LINC<sup>[3]</sup>や LIEM<sup>[3]</sup>などの他のリンケージ同定アルゴリズムや分布推定アルゴリズムよりも詳細に表現することが可能となる。PBIL(Population-Based Incremental Learning)<sup>[3]</sup>や UMDA(Univariate Marginal Distribution Algorithm)<sup>[3]</sup>などの他の分布推定アルゴリズムと同様に BOA はまず初期個体群を生成し、評価関数に従って初期個体群の中から優良個体の選択を行う。そして、選択された個体の情報をベジアンネットワークと条件付確率によって表現し、新たな個体を生

<sup>†</sup> 北海道大学大学院情報科学研究科  
Graduate school of Information science and technology Hokkaido University

<sup>††</sup> 北海道大学工学部情報基盤センター  
Hokkaido University information initiative center

成する。これらの動作を繰り返す中で問題構造を表現するネットワークと確率分布と解候補となる個体集合を更新していき、最適解を発見するのが BOA の簡単な動作となっている。これらのアルゴリズムの流れは Figure.1 で図示する。<sup>[4]</sup>

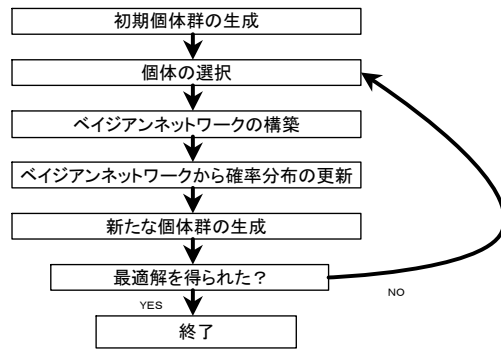


Figure.1 BOA のフローチャート

## 2.2 混合ベイジアンネットワーク<sup>[5][6][7]</sup>

ベイジアンネットワークは変数間の因果関係をネットワークと条件付確率で記述するグラフであり、ベイジアンネットワークを用いることで観測された情報から観測することのできない情報を推論することが可能となる。多くのデータを用いてそれらの環境に適合するベイジアンネットワークを構築する必要があるのだが、ベイジアンネットワークは状態の変化する環境への対応には向いていないと考えられる。また、問題が複数の問題構造をもっているような場合、正しいネットワークを構築するためには正しいデータを選択を求められ、それが失敗するとネットワーク構築の効率が落ちることが予想される。このように、ベイジアンネットワークは問題構造を表現する強力なグラフ構造であるが、対処すべき問題がいくつか存在していることがわかる。これらの問題に対処するためには単一ベイジアンネットワークではなく混合ベイジアンネットワークを使用する事が必要となる。混合ベイジアンネットワークは一つのベイジアンネットワークを一要素とした混合分布であり、それぞれの要素ベイジアンネットワークを独立に学習することによって、変化する環境への対応や、複数構造の記述が可能となる。

本論文は単一ベイジアンネットワークのみを使用している BOA に対してこの混合ベイジアンネットワークの考えを導入することで、ベイジアンネットワークの弱点となる領域を克服する。これによって BOA の解決することのできる問題領域は広がり、より有効なアルゴリズムへと進化することが予想される。

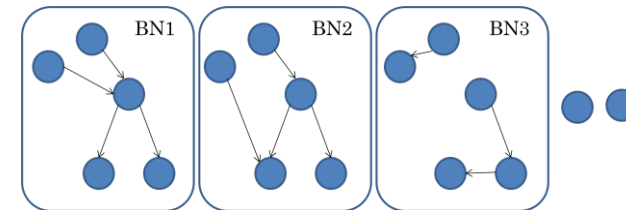


Figure.2 混合ベイジアンネットワーク

## 2.3 BOA with Mixture Distribution

本節では BOA に混合ベイジアンネットワークを導入したアルゴリズム、BOA with Mixture Distribution(BOA-MD)について説明をする。BOA に混合ベイジアンネットワーク<sup>[3][4][5]</sup>を導入するためには個体群からどのように混合ベイジアンネットワークを学習しなければいけないか考える必要がある。それを説明するために、まず混合ベイジアンネットワークの構築方法について考える。BOA-MD は EM アルゴリズム<sup>[8]</sup>を使用して各個体をそれぞれクラスターに振り分ける。そして、各クラスターに所属する個体集合のデータを集計して要素ベイジアンネットワークと条件付確率を構築する。

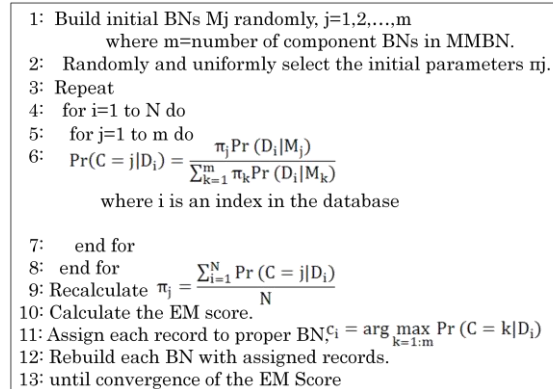


Figure.3 混合ベイジアンネットワークの学習

Figure.3 で示した混合ベイジアンネットワークの学習方法について説明をする。初めに生成した初期個体群を  $C$  個のクラスターへランダムに分割する。 $C$  は混合ベイジアンネットワークの要素数であり、アルゴリズムの実行の時に設定をする。本論文では  $C=2$  として

説明を続ける。2つのクラスターへと個体群を分割した後に、それぞれのクラスターで各クラスターに属する個体情報を表現するような要素ベイジアンネットワークを構築する。要素ベイジアンネットワークを構築するためのスコアとして(1)式の EM スコアを使用する。

$$EM_i = \sum_{j=1}^c \pi'_j (\log \Pr(D_i | M'_j) + \log \pi'_j) \quad (1)$$

各クラスターで構築された各要素ベイジアンネットワークから、各個体が各クラスターへと振り分けられる確率  $\Pr(C=j|D_i)$  の計算を行う。これは一つの個体がどのクラスターに属する確率が高いかを示す条件付確率である。 $\Pr(C=j|D_i)$  を全ての個体とクラスターに対して計算し、この  $\Pr(C=j|D_i)$  の値を用いて、全ての個体を C 個のクラスターに再配置する。再配置をした後で、それぞれのクラスターの個体情報を基に新しい要素ベイジアンネットワークを構築する。以上の処理をスコアが収束するまで繰り返すことで再帰的にパラメータを推定し、各個体のクラスタリングを行う。そして、各クラスターでの個体情報を基に要素ベイジアンネットワークを構築し、各要素ベイジアンネットワークから要素確率分布を決定する。それぞれの確率分布から新しい個体群を生成し、新個体群として次の世代に移る。以上が提案する新しいアルゴリズムとなる。この提案手法を BOA with Mixture Distribution(BOA-MD)と名付けた。Figure.4 で BOA-MD の詳しい処理を記述している。

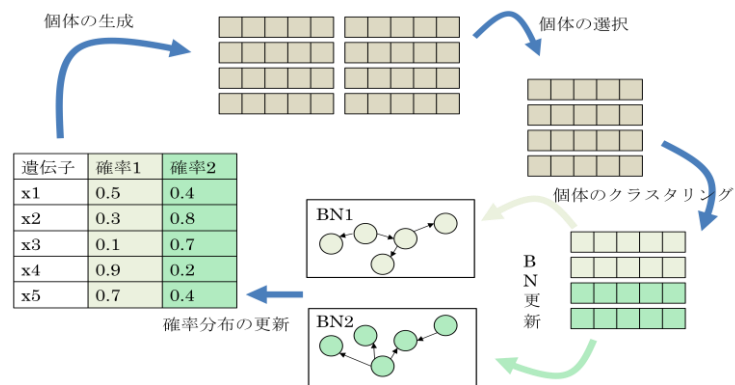


Figure.4 BOA-MD

- (1): Generate population
- (2): Select population
- (3): Clustering the population with EM algorithm
- (4): Learning the mixed Bayesian Network according to EM Score
- (5): Renew the probability for each component Bayesian Network
- (6): Generate new population using each probability
- (7): if (best solution is found) algorithm terminates  
 else go to (2)

Figure.5 BOA-MD のアルゴリズム

BOA-MD は複数の構造をもつような問題に対して BOA を上回る効果を見せる事が確認できた<sup>[11]</sup>。本論文は動的に変化するような環境に対して BOA-MD がどのような動作を見せ、BOA と比較し有効かどうかを説明していく。次章では動的環境問題についての説明をし、BOA での問題点、そして BOA-MD で実行した時の効果について説明する。

### 3. 動的環境問題に対応する BOA-MD

#### 3.1 動的環境問題

実世界には様々な問題が存在しておりそれらの解決がアルゴリズム作成の目的となっているが、その中でも環境が動的に変化することを考慮する必要のある問題は解決が難しい。進化計算においてこのような問題を解決するためには一つの環境のみに対応する個体のみを生き残らせるのではなく、個体群に多様性を維持する必要がある。通常の BOA はベイジアンネットワークを用いて問題構造を記述することで、複雑な問題の解決が可能であるが、ある程度個体群が収束すると、問題の環境が変化した場合への対応に問題が生じる。このような問題に対して、混合ベイジアンネットワークを利用した BOA-MD を用いることで、複数のベイジアンネットワークで複数環境のベイジアンネットワークをそれぞれ記述し、解を得ることができると考えた。動的環境問題に直面した時の BOA の問題点は環境の変化の対応の遅さである。例えば、環境 A と環境 B を 1 時間毎に交互に繰り返す問題について考える。BOA はまず環境 A に適応するようなベイジアンネットワークと確率分布を作り、環境 A で良いと推測される解を出す。環境 A にとっての優良解で個体群が埋められた後に環境 A から環境 B へとスイッチされる。この時、環境 A にとっての優良解の中から環境 B についても優良であると考えられる個体が選択され、そして環境 B を記述するようなベイジアン

ネットワークと確率分布の探索がスタートする。探索が進む中でようやく環境 B に関するネットワークと確率分布が作られたら、また環境 A へとスイッチされる。そうするとまた環境 A に適応するようなネットワークの構築が始まる。ここで問題になるのが、新しいネットワーク構築に必要な時間である。BOA は一度に一つのネットワークしか作ることができず、一つの環境のみを記述する。従って、環境が変化するたびにネットワーク構造を一から探索することになり、結果として非効率な動作をすることになる。一方で BOA-MD は複数のベイジアンネットワーク、混合ベイジアンネットワークを利用している。混合ベイジアンネットワークによって、BOA-MD が環境 A と環境 B の両方の環境の情報を常に保持することが可能となり、環境の変化に対して柔軟に対応することが可能となるのだ。しかし、予想とは裏腹に BOA-MD は動的環境問題に対してそれほどよい結果を見せることはできなかったため、ここで新たな概念を加える必要が生じた。

### 3.2 改良された BOA-MD

BOA-MD は混合ベイジアンネットワークによって複数環境の状態を記述することができるアルゴリズムだが、動的環境問題に対して高い効果を見せるわけではなかった。環境の変化によって個体群における優良個体が劇的に変化するような問題に直面した時、BOA-MD は二つの環境を表現する二つのネットワークを保持することは可能だったが、以前の環境の優良個体を保持することが困難だったため、正しい確率分布を生成できずに、結果として片方の環境のみに適応するような個体のみを解として生成することとなった。この問題点を改良するためにそれぞれの世代の確率分布を数世代後まで残しておく操作を行った。以前の環境のデータを十分に得るために、現在更新される確率分布だけではなく過去数世代の確率分布も使用することによって過去のデータ、つまり以前の環境のデータを得ることがきる。過去の環境の正しいデータを得ることができれば、それだけ正しい混合ベイジアンネットワークを生成することができ、それぞれの環境に対する正しい確率分布もまた得ることができるようになる。このような改良を加えた新しいアルゴリズムは以下の Figure.6 に記述する。

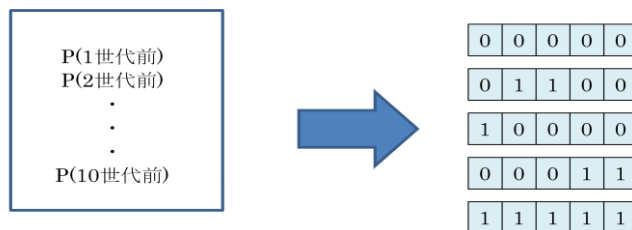


Figure.6 前世代の確率分布からの標本

Figure.6 では 1 世代前から 10 世代前までの確率分布を使用して個体の生成を行っている。10 世代前までの確率分布を保持しておくことで、個体群の多様性を維持し、環境の変化に対応しやすくさせる効果を見せている。本論文ではそれぞれの確率分布からは当確率で次の世代の個体を生成することとしている。次の章では改良した BOA-MD と通常の BOA を動的環境問題のベンチマークへと適用し、両アルゴリズムのパフォーマンスの比較を行い、新たに追加した確率分布継承要素の効果について考える。

## 4. 比較実験

### 4.1 ベンチマーク

ベンチマークとして Trap5 to Zeromax 問題を使用する。この関数は Trap5 関数<sup>[9]</sup>と Zeromax 関数を交互に繰り返すことで動的に変化する環境を再現している。以下の数式の通りに Trap5 to Zeromax 問題は評価値を返す。

$$\text{trap5}(u) = \begin{cases} 4 - u & \text{if } u < 5 \\ 5 & \text{if } u = 5 \end{cases} \quad (2)$$

$$\text{Zeromax}(S) = L - v \quad (3)$$

$$F(S) = \begin{cases} \frac{L}{5} \sum_{i=1}^5 \text{trap5}(u_i) & \text{if } (g \bmod 2t < t) \\ \text{Zeromax}(S) & \text{else} \end{cases} \quad (4)$$

S は問題を表すビット列を示している。u はビット列 S を 5 ビット毎に区切った部分列中の、v は S 中における「1」の値をもつビットの個数を表している。g は世代数を表す変数で、t は Trap5 関数と Zeromax 関数を何世代毎にスイッチするのかわを示している。もし t=5 なら 5 世代毎に Trap5 関数と Zeromax 関数が入れ替わることとなる。Trap5(y) は y に属する 5 ビット全ての値が 1 である時に最大の値 5 を返すが、それ以外の場合はビット中に 1 が多ければ多いほど高い値を返すだまし問題である。Trap5(y) の解を正しく得るためには、ビット列の中で Trap5(y) の引数となっている 5 つのビットをサブストリングとして考慮する必要がある。BOA においてはこれら 5 つのビットに対応するノードがエッジを張っているようなベイジアンネットワークを作る

ことによって、Trap5(x)の解を求める。一方で Zeromax(x)は全てのビット列を引数とする関数で、ビット列中の「0」の値をもつ変数の個数そのまま評価値となる関数である。Trap5(y)を解くために 5 ノードがつながっているベイジアンネットワークを作る必要がある一方で、Zeromax(x)はベイジアンネットワークを作る必要がなく、局所探索によって最適解を得ることができる。本論文ではこれら二つの関数を t 世代毎にスイッチすることで動的環境問題を再現するベンチマークとして使用している。今回の実験では t=5,10,15 の 3つの値で実験を行い、BOA-MD の効果を見る。

#### 4.2 実験結果

前述の Trap5 to Zeromax 関数を BOA と BOA-MD の両アルゴリズムで解き、パフォーマンスの比較を行う。実験条件は以下の通りである。

- ・問題長 L = 50
- ・個体群の大きさ N = 40000

以上の条件で、t=5,10,15 の 3つの動的環境において BOA と BOA-MD を実行させ、その効果を見た。実験結果は以下の通りとなった。

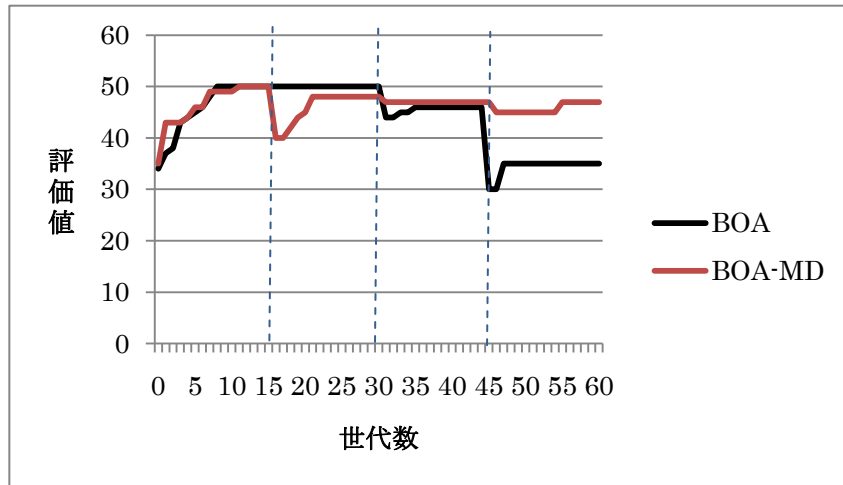


Figure.7 t=15

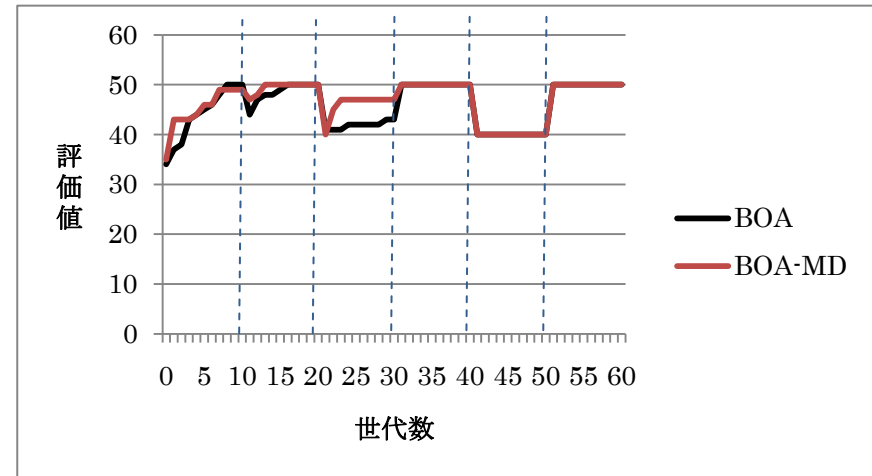


Figure.8 t=10

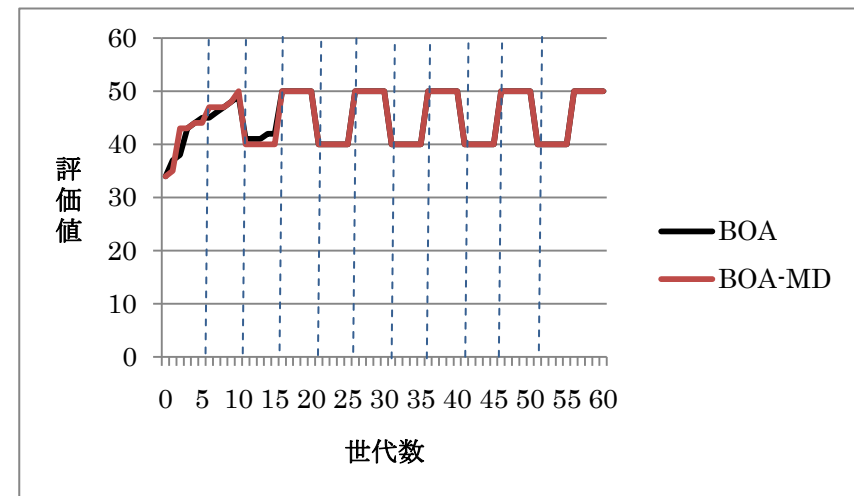


Figure.9 t=5

Figure.7 は T=15 の時の各アルゴリズムの評価値の変動の様子を示している。T=15 な

ので 15 世代毎に Trap5(x) と Zeromax(x) が入れ替わるのだが、BOA は最初の 15 世代は Trap5(x) をベイジアンネットワークで上手く表現して探索を行っているが、16 世代以降は Trap5(x) から Zeromax(x) へと環境が変化すると、Zeromax(x) 関数の環境のみに適応する個体のみが個体群の中に現れる。このような状況で 31 世代以降に Trap5(x) へと環境が移行しても、個体は既に Zeromax(x) の環境で適応するように進化を終えてしまい、環境の変化に対応しなくなってしまう。一方、BOA-MD は Trap5(x) から Zeromax(x) へと環境が移行してしまった後でも、Trap5(x) と Zeromax(x) の両方のベイジアンネットワークを混合ベイジアンネットワークによって保存することができるので、31 世代以降でも Trap5(x) の優良個体を探索することが可能となっている。次に Figure.8 は T=10 の時の両アルゴリズムの比較を行っている。T=10 では十分な探索を行う前に次の環境へと移行してしまうため、30 世代を超えたところで BOA も BOA-MD も ALL-0 へと収束してしまった。この実験においては BOA-MD の有効性を示すことができなかったが、20 世代目の BOA-MD の動作を見ると BOA よりも Trap5(x) に対してある程度探索を続けている事が言える。また、Figure.9 は T=5 の時の同様の実験結果だが、T=10 の時と同様に環境の変化が早すぎるので両アルゴリズムに大きな差は見られなかった。以上の実験から、環境の変化がある程度間隔を開けて起こるような問題にたいして BOA-MD は BOA の苦手とする動的環境の問題に対処することができたが、急激に変化する環境には対応することが難しいということがわかった。

## 5. おわりに

本論文は BOA に混合ベイジアンネットワークを導入した BOA-MD を提案し、動的環境問題での効果について説明した。我々の研究の最初に作成したアルゴリズム、BOA-MD は動的環境の問題では高い効果を見せることができなかったため、本論文では前世代の確率分布を引き継ぐという新しい要素を BOA-MD に追加した。前世代の確率分布を次世代まで保存しておくことによって、個体群に多様性を作ることを実現し、複数のベイジアンネットワークを構築することの手助けをすることにも成功した。前世代確率分布の継承を導入した BOA-MD は動的環境問題に対して、環境の変化がダイナミックに生じるような問題に対して BOA よりも適応できたことを確認した。ただし、環境の変化が急激に起こるような問題に対しては BOA とあまり変わらない結果を見せる事になったので、今後はこのような問題に対してどのように対応していくかを考える必要がある。

**謝辞** 本研究は科研費(22500196)の助成を受けたものである。

## 参考文献

- 1) Martin Pelikan, Kumara Sastry, Erick Cantu-Paz: Scalable optimization via Probabilistic Modeling, *Hierarchical Bayesian Optimization Algorithm*, pp63-90, Springer (2006)
- 2) Pedro Larranage and Jose A.Lozano: Estimation of Distribution Algorithms a New Tool for Evolutionary Computation, Kluwer Academic Publishers (2002)
- 3) 棟朝雅晴, 遺伝的アルゴリズム-その理論と先端的手法- 森北出版株式会社(2008)
- 4) Goldberg, David E. Genetic Algorithm in Search Optimization and Machine Learning. Addison Wesley, New York(1989)
- 5) Kyongryun Lee : A Mixture Model of Bayesian Networks, the degree of MASTER OF SCIENCE, Iowa State University (2005)
- 6) Bo Thiesson Christopher Meek ,David Maxwell Chickering , David Heckerman : Learning Mixture of Bayesian Networks, IN G.F.Cooper and S.Moral(Eds), *Proc. Fourteenth Conference on Uncertainty in Artificial Intelligence(UAI'98)*.San Francisco, Calif : Morgan Kaufmann(1998)
- 7) Bo Yhiesson, Christopher A.Meak. SPEECH RECOGNITION WITH MIXTURES OF BAYESIAN NETWORKS, Microsoft Corporation,WA(US), Patent No.: US 6,336,108 B1(1998)
- 8) Dempster,A, Larid, N, and Rubin, D : Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, B* 39:1-38(1977)
- 9) Deb, K and Goldberg, D.E. Sufficient conditions for deceptive and easy binary functions. *Annals of Mathematics and Artificial Intelligence*, 10:385-408(1994)
- 10) Andre R.Goncalves, Fernando J. Von Zuben : Online Learning in Estimation of Distribution Algorithms for Dynamic Environments, *Evolutionary Computation*,2011 IEEE congress on(2011)
- 11) 堀伸哉, 棟朝雅晴, 赤間清 : BOA における混合ベイジアンネットワークの導入の効果について 第 5 回進化計算フロンティア研究会(2011)