

平成 22 年 6 月 14 日現在

研究種目：基盤研究（B）

研究期間：2007～2008

課題番号：19300075

研究課題名（和文）

分布推定型遺伝的プログラミングに基づく進化システムの構築に関する研究

研究課題名（英文）

Program Evolution by means of Estimation of Distribution Programming

研究代表者

伊庭 斉志（IBA HITOSHI）

東京大学・大学院工学系研究科・教授

研究者番号：40302773

研究成果の概要：

本研究では、確率モデルに基づくプログラム進化である「プログラムの分布推定：EDP(Estimation of Distribution Programming)」という考えを提案し、遺伝的プログラミングを拡張した自動プログラミング法を構築した。より詳細には、集団探索ではあるが交叉と突然変異を用いず、プログラムの確率分布を推定することにより進化を行い、頑強で効率のよい手法を実現した。提案手法では確率変数の依存関係を表す確率分布モデルを複数の候補で組み合わせ、プログラム集団の確率分布を推定することによりプログラム進化を実行する。この手法は以下で述べる従来の最適化アルゴリズム（EDA）を拡張したものである。さらに本研究では、提案する EDP に基づいて構築した進化システムの有効性を、ヒューマノイド・ロボットの動作生成、金融工学におけるデータマイニングと予測、バイオインフォマティクスにおける遺伝子因果関係の解析などの実際的な問題領域を対象にして検証した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	3,600,000	1,080,000	4,680,000
2008 年度	3,500,000	1,050,000	4,550,000
年度			
年度			
年度			
総計	7,100,000	2,130,000	9,230,000

研究分野：

科研費の分科・細目：

キーワード：

- (1) 遺伝的プログラミング (2) 遺伝的アルゴリズム (3) 進化計算
 (4) ヒューマノイドロボット (5) 自動作曲 (6) 創発デザイン
 (7) EDA アルゴリズム (8) 対話型進化計算

1. 研究開始当初の背景

遺伝的アルゴリズム (Genetic Algorithms, GA) などの進化論的手法は、生物の進化のメカニズムをまねてデータ構造を変形、合成、選択し、これにより最適化問題の解法や有益

な構造の生成を目指すアプローチである。その 1 つである遺伝的プログラミング (Genetic Programming, GP) は、GA の遺伝子型を構造的な表現が扱えるように拡張し、プログラム生成や学習、推論、概念形成などに

応用することを目標としている。GPは多くの分野、ロボティクス、回路設計、芸術、デザイン、金融工学など、さまざまな実的な領域でさかんに応用され、その有用性が確認されている。

しかしながら遺伝的プログラミングにおいては、(1)探索が必ずしも効率的でない、(2)探索過程でのプログラム(木構造)が膨大になる、(3)進化の結果として得られたプログラムの頑強性が高いとは限らない、などの点が指摘されている。これらは、GPの探索過程でプログラム(遺伝子型)の長さが不必要に増大することに原因がある。この現象はブloat(loat)、あるいは構造的複雑さの増大と呼ばれており、GPの効率的な探索を阻害する厄介な問題の1つである。ブloatは、プログラムのイントロン(全体のプログラムの動作に無関係なコード)や冗長部分が成長することに起因すると言われている。これらの背景をふまえて、本研究ではGPと統計的手法を統合し、GPとは異なるメカニズムでプログラム進化を行う手法を提案する。

2. 研究の目的

本研究では、GP(遺伝的プログラミング)とは異なるメカニズムでプログラム進化を行う手法を提案する。研究の目標は、「プログラムの分布推定: EDP(Estimation of Distribution Programming)」に基づいて遺伝的プログラミングを拡張し、より頑強な自動プログラミング手法を実現することである。提案手法は集団探索ではあるが交叉と突然変異は用いず、プログラムの確率分布を推定することにより進化を行う。この研究においては前述したブloat問題(プログラムの木構造の過大な成長減少)を効果的に解決することが究極の目標であるが、そのためにはプログラムの分布を効率的に推定して進化的探索を的確に導く必要がある。この問題を解決するためにさまざまなモデル推定のための手法を比較・考察し、より頑強で効率的な推定方法を提唱する。

さらに本研究では、提案するEDPに基づいて構築した進化システムの有効性を、ヒューマノイド・ロボットの動作生成、金融工学におけるデータマイニングと予測、バイオインフォマティクスにおける遺伝子因果関係の解析などの実的な問題領域を対象にして検証する。

3. 研究の方法

本研究では、提案した分布推定型遺伝的プログラミング(EDP)に基づいて進化型システムを構築する。この進化型システムでは、(1)探索の効率化、(2)頑強性、(3)実領域問題への適用可能性、の3点を重視して実装を行う。効率的な探索の実現のために、分布モデル

推定の高速化を各種モデルに対して検証する。一般にプログラムには数多くの変数が存在し、EDPにおいてはその間の因果関係を適切に推定する必要がある。そのため、スケラビリティを考慮して多くの変数の因果関係を的確に導出する手法の実現を試みる。また、頑強なシステム構築のために、集団的な学習手法をプログラム進化の枠組に取り入れる。つまり、複数の異種モデルを用いることを前提として、集団学習のアプローチによりモデルを統合してより頑強な推定手法を構築する。具体的には、異種モデルの推定結果を動的(競合的)に組み合わせる一方で、静的(協調的)に統合して新たな推定と仮説を生成する。バギングやブースティングなどの機械学習の手法により、与えられた訓練データからブootストラップ法(例題からの復元抽出)を繰り返して擬似的にデータ(仮説)を作り、複数の遺伝的プログラムを独立に学習し、多数決により頑強な学習手法を確立することを試みる。

この研究計画では提案するEDPに基づいて構築した進化システムの有効性を多岐にわたる分野で検証するが、それは進化型計算などの近似アルゴリズムの重要性は実領域で確認されることが多いとされるからである。本研究で扱う実験領域は、(1)金融データのマイニングと予測、(2)遺伝子因果関係解析、(3)ヒューマノイドロボット・プログラミングなどの分野である。これらはそれぞれ(1)金融工学、(2)バイオインフォマティクスおよび(3)工学的最適化設計に対応し、異なる実領域の代表と考えられる。

4. 研究成果

平成19年度には、以下の3点について研究を行った。

- (1)EDP(分布推定型遺伝的プログラミング)の基本部分の設計を行った。特に分布推定を効率的に導くためのプログラム内の変数に関する因果関係モデルの開発、及び分布モデル推定方法の効率的実装を実現した。
- (2)プログラム進化にはある種の拘束条件での準最適値の探索が不可欠である。そのために、構築したEDPの基本的な探索能力の検証をする。具体的には様々なベンチマーク問題に対して空間/時間計算量を従来の遺伝的プログラミングと比較し、探索能力を解析した。この結果に基づき可能な改良を行った。
- (3)EDPに基づいて進化システムを構築し、さまざまな工学的応用に適用可能なプログラム進化の枠組を実現した。

EDPの基本部分の設計のために、プログラム進化に寄与するさまざまなモデルを考察して推定手法を実装した。具体的には、ページ

アン・ネットワーク、強化学習、K2 アルゴリズム、Kullback-Leibler 情報量に基づく変数の因果関係推定手法を比較・考察し、より頑強で効率的な導出法を考案した。プログラム進化においては、可能なシンボルの数が増大し、依存関係の推定が困難になることがある。本研究では拡張構文木や情報量基準を導入しこの問題点の解決を試みた。より具体的には、確率変数が依存関係を持つ確率分布モデルを複数種類組み合わせ、プログラム集団の確率分布を推定した。これにより確かなプログラムを交叉や突然変異を利用することなく進化させることを目指した。

遺伝的プログラミングには空間計算量、プログラム実行時の負荷、及び進化の停滞現象などの課題があることも明らかになっている。本研究の中心であるプログラム進化の場合、遺伝子表現が進化の方向性（淘汰圧）を決定するので、言語の効率的な設計はより一層重要である。提案する「EDP：分布推定型遺伝的プログラミング」は進化する各個体（プログラム）を確率モデルで表現する。これは従来提案されている進化型プログラムより高速に実行可能であり、かつ空間計算量も少ないという利点がある。EDP を用いて空間的／時間的効率を改善することにより、遺伝的プログラミングにおける進化速度を数倍のオーダーで向上させることを目指した。さらに従来の標準的なシステムとの比較実験を行い、探索効率の解析と評価を行った。

(3)については、まず工学的な応用として提案した枠組をヒューマノイド・ロボットの動作設計に応用し、進化システムの有効性を検証した。想定するタスクは、与えられた条件を満たすヒューマノイド・ロボットの動作（例：ボールを蹴る、太極拳を踊るなど）をロボットの安定性（転ばないこと）と制御可能性（過度な負担がサーボにかからないこと）を考慮して設計するというものである。このタスクの遂行には複雑な力学的数値計算が必要であり、汎用的な解法は提案されていない。これに対してEDPを用いて進化的に動作をデザインすることを試みた。

この問題を最初に取り上げる理由は以下の通りである。

- ・安定して柔軟な動作を人手で設計するのはきわめて困難である
- ・生成結果を実環境で容易にテストできる
- ・得られる結果の有用性が極めて高い

つまり、適度に難しい問題であり、かつ成否の判断が容易なため、構築した枠組の有効性の検証に適している。なお、実際の検証実験には、実機ロボット（富士通製HOAP-1）とロボットシミュレータ（OpenHRP）を使用した。

平成20年度の研究では、プログラム進化のための分布推定モデルとしてさまざまなモデルを検証した。本研究では、複数の異種モデルを用いることを前提として、集団学習のアプローチに基づいてモデルを統合し、より頑強な推定手法を構築した。集団学習は、簡単な学習機械を組み合わせることで複雑な機械を構成する機械学習の手法であり、計算量の削減や過学習の回避などの長所を有する。具体的には、異種モデルの推定結果を動的（競合的）に組み合わせる一方で、静的（協調的）に統合して新たな推定と仮説を生成する。またバギングやブースティングなどの機械学習の手法により、ブートストラップ法（例題からの復元抽出）を繰り返して擬似的にデータ（仮説）を作り、複数の遺伝的プログラムを独立に学習し、多数決により頑強な学習手法を確立することを試みた。これにより目的とするプログラム分布推定を効率的に導出するEDPアプローチの拡張を提案し、さまざまな応用問題においてその有効性を実験的に明らかにした。

具体的には、前年度に行ったヒューマノイド・ロボットの動作設計についての研究を引き続き行うとともに、金融工学やバイオインフォマティクスの問題に適用して、以下に詳述するような内容で実証性を検証した。

バイオインフォマティクスにおいて遺伝子の因果関係の解析は重要な課題である。大規模な遺伝子ネットワークの推定問題を考えると、(1)ノイズの多さ、(2)データ量の少なさ、(3)得られる解（ネットワーク）候補の比較の難しさなどの観点から、工学的な意味での逆問題（Reverse Engineering）を超えた難しさを含んでいる。従来このためにさまざまな推定手法が提案されてきた。たとえば、S-systems、重み付けネット、微分方程式系、確率微分方程式、ベイジアンネットなど。しかしながらいずれの方法にも一長一短があり、どれがもっとも優れているかについての結論は出ていない。本研究では、EDPにより複数候補のモデルを遺伝的プログラムに基づいて生成し、集団学習のアプローチによりこれらのモデルを統合して推定を行う手法を提案した。これにより上述の遺伝子解析の問題点を合理的な計算資源で解決することが可能になった。

金融データの予測に関しては、これまでに各種の時系列予測を応用した方法が提案されている。しかしながら、汎化能力、リアルタイム性、頑強性などにおいて従来の手法は必ずしも望ましいものではなかった。そこで本研究では、EDPに基づく時系列予測のプログラムの構築を行った。さらに他の金融データのマイニング（オプション価格予測、ポートフォリオ構築、金融リスク分析・分類問題）などにEDPを応用し、その有効性を確認した。

5. 主な発表論文等

(研究代表者, 研究分担者及び連携研究者には下線)

[雑誌論文] (計 12 件)

- [1]長谷川禎彦, 伊庭斉志: "PCFG-LA 混合モデルに基づく分布推定アルゴリズム", 人工知能学会論文誌, vol. 24, no. 1SP-H, pp. 80-91, 2009.
- [2]Noman, N., Iba, H.: "Differential evolution for economic load dispatch problems", Electric Power Systems Research, vol. 78, pp. 1322-1331, 2008.
- [3]Noman, N., Iba, H.: "A Bayesian Network Approach to Program Generation", IEEE/ACM Transcriptions on Evolutionary Computation, vol. 12, no. 6, pp. 750-764, 2008.
- [4]Iba, H.: "Inference of differential equation models by genetic programming", Information Sciences, vol. 178, pp. 4453-4468, 2008.
- [5]Tohge, T., Iba, H.: "Evolutionary Morphology for Polycube Robots", Frontiers in Evolutionary Robotics,
- [6]Iba, H. (Ed.), I-Tech Education and Publishing, pp. 221-232, 2008.
- [7]Yanase, T., Iba, H.: "Evolutionary Motion Design for Humanoid Robots", Frontiers in Evolutionary Robotics, Iba, H. (Ed.), I-Tech Education and Publishing, pp. 567-586, 2008.
- [8]Iba, H.: "Hybrid Genetic Programming and GMDH System: STROGANOFF", Hybrid Self-Organizing Modeling systems (Studies in Computational Intelligence), Godfrey C. Onwubolu (ed.), pp. 27-98, Springer, 2009
- [9]Tohge, T., Iba, H.: "Search for the Structure of Polyube Robots using Memetic Algorithm", Emerging Tehnologies, Robotis and Control Systems, Third edition, Salvatore Pennacchio (ed.), pp. 154-159, Internationalsar, 2009
- [10]Aranha, C., Iba, H.: "The Memetic Tree-based Genetic Algorithm and its application to Portfolio Optimization", Memetic Computing, vol. 1, no. 2, DOI:10.1007/s12293-009-0010-2, pp. 139-151, 2009
- [11]Paul, T.-K., Iba, H.: "Predition of Caner Class with Majority Voting Genetic Programming Classifier Using Gene Expression Data", IEEE/ACM Transcriptions on Computational Biology and Bioinformatics, vol. 6, no. 2, pp. 353-367, 2009
- [12]Hasegawa, Y., Iba, H.: "Latent Variable

Model for Estimation of Distribution Algorithm Based on a Probabilistic Context-Free Grammar", IEEE/ACM Transcriptions on Evolutionary Computation, vol. 13, no. 4, pp. 858-878, 2009

[学会発表] (計 12 件)

- [1]Aranha, Claus C., Iba, H.: "A Tree-based GA Representation For The Portfolio Optimization Problem", Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2008), pp. 873-880, 2008.
- [2]Tanji, M., Ando, D., Iba, H.: "Musical Rhythm Parsing Using Mixture Probabilistic Conext-Tree Grammar", Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10), 2008.
- [3]Aranha, C., Iba, H.: "Application of a Memetic Algorithm to the Portfolio Optimization Problem", Advances in Artificial Intelligence, 21st Australasian Joint Conference on Artificial Intelligence (AI2008), LNCS 5360, pp. 512-521, 2008.
- [4]Tanji, M., Ando, D., Iba, H.: "Improving Metrical Grammar with Grammar Expansion", Advances in Artificial Intelligence, 21st Australasian Joint Conference on Artificial Intelligence (AI2008), LNCS 5360, pp. 180-191, 2008.
- [5]Yanase, T., Iba, H.: "Evolutionary Multi-Objective Optimization for Biped Walking", Simulated Evolution and Learning, 7th International Conference (SEAL2008), LNCS 5361, pp. 635-644, 2008.
- [6]Aranha, C., Iba, H.: "MVGPC+: An Integrated Approach to Machine-learning Classification", Proceedings of the IASTED International Conference Advanced in Computer Science and Engineering (ACSE2009), pp. 38-43, 2009.
- [7]Hirabayashi, A., Aranha, C., Iba, H.: "MVGPC+: An Integrated Approach to Machine-learning Classification", Proceedings of the IASTED International Conference Advanced in Computer Science and Engineering (ACSE2009), pp. 32-37, 2009.
- [8]Tanji, M., Iba, H.: "Program optimization by random tree sampling", Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2009), pp. 1131-1138, 2009
- [9]Yanase, T., Hasegawa, Y., Iba, H.:

"Binary encoding for prototype tree of probabilistic model building GP", Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2009), pp.1147-1154, 2009

[10]Aranha, C., Iba, H. : "Using memetic algorithms to improve portfolio performance in static and dynamic trading scenarios", Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2009), pp.1427-1434, 2009

[11]Hirabayashi, A., Aranha, C., Iba, H. : "Optimization of the trading rule in foreign exchange using genetic algorithm", Proceedings of the Genetic and Evolutionary Computation Conference (GECCO2009), pp.1529-1536, 2009

[12]. Watanabe, A., Tanji, M., Iba, H. : "Creating Singing Vocal Expressions by means of Interactive Evolutionary Computation", Proc. Of the 5th International Workshop on Computational Intelligence and Applications, IEEE SMC Hiroshima Chapter, Hiroshima University, pp.10-12, 2009

[図書] (計 3 件)

[1]Iba, H. (ed.) : Frontiers in Evolutionary Robotics, I-Tech Education and Publishing, Vienna, Austria, 1997.

[2]伊庭 齊志: Cによる探索プログラミング—基礎から遺伝的アルゴリズムまで, オーム社, 2008.

[3]Iba, H., Paul, T. -K, Hasegawa, Y. : "Applied Genetic Programming and Machine Learning", CRC Press, ISBN-10:1439803692, Taylor. Francis, 2009

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

6. 研究組織

(1) 研究代表者

伊庭 齊志 (IBA HITOSHI)

東京大学・大学院工学系研究科・教授

研究者番号 : 40302773

(2) 研究分担者

なし

(3) 連携研究者

なし