

移民世代数と深さ依存型交叉を用いた並列分散 GP

岩下 誠[†] 伊庭 斉 志[†]

遺伝的プログラミング (GP) の探索においては良好なスキーマを育てる局所的探索とスキーマを組み合わせる大域的な探索の両方が重要となる。島モデルの分散 GP においては各島ごとに独自のスキーマを育てることが可能となっている。よって、島それぞれでは局所的な探索を行い、移民によって異なる島のスキーマを組み合わせる大域的な探索を行うことによって探索効率が向上すると考えられる。本論文では分散 GP の性能を高める手法として、より大域的な探索を促進する深さ依存型交叉を導入し、島にどれだけ長くいるかの指標である移民世代数に応じてオペレータを使い分ける手法を提案し、評価を行った。結果、深さ依存型交叉の導入により成績を上げることに成功し、移民世代数を指標として用いることで低い深さ依存型交叉の選択率でも大きな効果をあげることが示した。

Parallel Distributed GP with Immigrants Aging and Depth-dependent Crossover

MAKOTO IWASHITA[†] and HITOSHI IBA[†]

This paper proposes a new method for parallel distributed Genetic Programming in the Island Model. The proposed method includes applying a depth-dependent crossover to the “deme-young” immigrants according to their immigrant ages. The immigrant ages show how long they survive in their original islands. This method can provide both local and global search strategies, both of which are important for the purpose of effective search. The experimental results have shown that induction of depth-dependent crossover and immigrant age works effectively.

1. 序 論

遺伝的プログラミング (Genetic Programming: GP) は遺伝的アルゴリズム (Genetic Algorithm: GA) の拡張で、構造的表現を可能にしたものである。GP は、自動プログラム生成やバイオインフォマティクス、芸術的分野など、様々な分野での応用が行われている¹⁾。

GP の探索が有効である理由として、解プログラムの一部に相当する小さな部品 (building blocks) が交叉によって組み合わせられてより大きな部品が作られ、これが繰り返されることによって最終的に解に到達するという積木仮説が成り立っていることが考えられる²⁾。よって、GP の探索においては良好なスキーマを育てる局所的探索とスキーマどうしを組み合わせる大域的探索の両方が重要となる。

GP を実行する際の問題点として、オペレーションによる良好なスキーマの破壊や局所解への早期収束

による集団の多様性の低下、木の爆発的成長であるブロートなどにより探索効率が低下することが知られている。特にブロートは GP において探索能力の低下をもたらす最大の問題というべきもので、構文的イントロンと呼ばれる遺伝子コードにおいて実際に到達されることのないコードが原因となり起こることが知られている⁴⁾。構文的イントロンの存在により、GP は本質的に GA よりも自己複製 (ここでいう複製とは遺伝子型の一致ではなく表現型の一致を意味する) を起こしやすく、ブロートとともに集団の多様性も失われることとなる。

一方、GP の基となっている GA の並列分散処理の手法である島モデル (Island Model) を用いると、移民の効果により大域的な探索を行うことが可能である。並列分散 GA の分野においては解探索メカニズムは単一母集団の GA とは異なり、並列独自の手法が有効に働くことが知られている。三木らは島単位での局所的な探索を行うオペレータである最良組合せ交叉の導入

[†] 東京大学大学院新領域創成科学研究科

Graduate School of Frontier Science, The University of Tokyo

イントロンには他に冗長表現である意味論的イントロンがある。詳細は Angeline の論文³⁾を参照されたい。

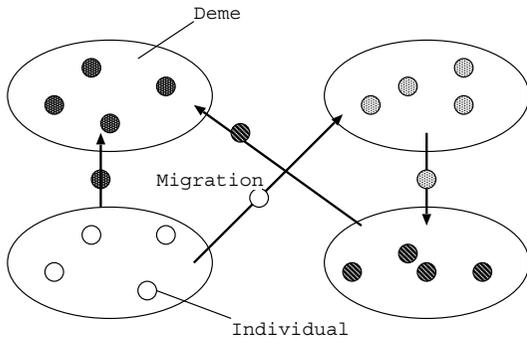


図 1 鳥モデル
Fig. 1 Island model.

と、移民と原住民を 100%の確率で交叉させるハイブリッド生成交叉との組合せによる並列分散 GA 独自の手法を用いることにより全体の処理効率が上がる例を報告している⁵⁾。

一方で並列分散 GP の分野においては、プロセス数に対し超線形の性能向上や移住率が探索性能に与える影響を報告した Andre らの研究⁶⁾や、Fernandez らの島数や個体数に関する研究⁷⁾、移住やトポロジーに関する研究⁸⁾などがあるが、並列分散独自の手法に関する研究は GA ほど活発には行われていない。そこで、本研究では並列分散 GP において探索効率を高める手法の提案を行う。

2. 鳥モデル

鳥モデルは、母集団を複数のサブ母集団(島)に分割し、島ごとに独立に進化を行い、定期的に島どうしで移住と呼ばれる個体の交換を行う粗粒度の並列モデルである(図 1)。この手法を用いることによって、高い並列性と少ない通信量での並列処理を行うことができる。

鳥モデルにおいては、島単位での多様性が失われても移住により個体を交換することで母集団全体として多様性を保つこと可能であり、より効率的な探索を行うことができる⁹⁾。よって、鳥モデルにおいては各島の中では独自のスキームを育てる局所的探索を重視し、移民に関してはスキームを保護する大域的探索を行うことによって探索を効率的に行うことができると考えられる。

2.1 鳥モデル GP

可変長の木構造を遺伝子として持つ GP は、その起源である GA とは異なる性質を持つ。過去の研究から鳥モデルにおいても GP は GA と異なる特性を持つことが知られている。鳥モデル GA においては、1 島

あたり 10 程度の個体数で島数を多くするのがよいとされている¹⁰⁾。一方で、GP においては島モデルにおいて探索能力を保つために必要なサブ母集団あたりの個体数は大きく、対象問題の性質だけではなく困難さに依存すると考えられている。Punch はサブ母集団あたり 700 個体(サブ母集団数 7)を用いて関数あたりは問題を解いた結果、木構造の深さ制限が 17 の場合鳥モデルは単一母集団 GP より好成績であったが、深さ制限を 25 に変え探索空間を広げることによって問題を困難にしただけで鳥モデルは単一母集団 GP より劣ったという結果を得ている¹¹⁾。また、移住率に関しても鳥モデル GA においては 10%以上の値で好成績を得ることが多い¹⁰⁾のに対し、GP においては移住率を 12%とすると単一母集団 GP より劣る結果となったという例が報告されている¹²⁾。

これらの現象の原因の 1 つとして GP 個体のプロット、およびそれにとまなう個体複製の増加があると考えられる。プロットは問題の探索空間の広さに対して集団個体数が小さいほど起こりやすいため、集団分割によって多様性を保つメリットとオーバヘッドの関係にある。つまり、サブ母集団の個体数を小さくするとサブ母集団内でのプロットが起こる可能性が上昇し、また移住率を大きくすると移民個体を原因としたプロットが起こりやすくなると考えられる。

3. 木の深さと遺伝的操作

3.1 深さとノード選択率

個体を木構造で表現する GP においては、深い地点ほど(木の葉に近いほど)ノード数が増加する傾向にある(図 2)。このため、遺伝的オペレータを施す際にランダムにノードを選択すると、より深いノードが高確率で選択されることになる。特に、木のサイズが大きくなると遺伝的オペレータが根に近いノードに施される確率は非常に低くなってしまい、ほとんどが深いノードでのオペレーションになってしまうという問題がある。

3.2 木の深さと遺伝的操作

GP の交叉と木構造の深さに関して次のような仮説が立てられている¹³⁾。

- (1) 交叉による遺伝子型の変化が大きいほど、表現型における変化は大きく、適合度も劇的に変わる。
- (2) 遺伝子型の変化が根のノードから遠いほど、表現型での変化は小さい。つまり、交叉点が根から遠いほど、その部分木が全体の評価に及ぼす影響は少なくなる。

Number of Nodes

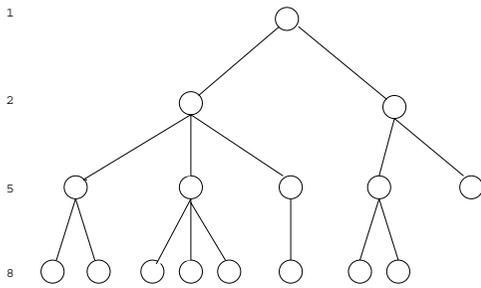


図2 木の深さとノード数

Fig. 2 Number of nodes and depth of tree.

Igel らは、人工蟻の探索問題、4 入力マルチプレクサ関数、太陽の黒点予測問題の 3 つの問題において仮説があてはまることを実験的に示した¹³⁾。また仮説 (2) に関してはその原因として根から遠いノードの変化が実際に小さい表現変化をもたらす可能性が高い場合と、根から遠いノードの変化が中立に働く可能性が高く結果として平均的な変化の度合いを小さくする場合の 2 通りがあることを示し、どちらの原因が支配的となるかは対象問題の性質に応じて異なることを示した。

また、積木仮説に基づいて考えると交叉による部分木の積み上げが進化にとって重要となるが、葉に近い点での交叉は根に近い点での交叉に比較して積み上げられた部品を破壊する可能性が大きくなる。これは根のノードから遠いほど表現型での変化は小さいとする仮説 (2) と矛盾するようにも感じられるが、論じる観点が異なるため矛盾するものとはならない。仮説 (2) が 1 個の個体だけに注目したものである一方で、この部品の破壊とは集団全体で持っている有用な部品の破壊を意味するからである。

3.3 深さ依存型交叉

Ito らは、このような問題を解決するために木の根に近いノードがより高確率で選択される深さ依存型交叉 (図 3) を提案した¹⁴⁾。この交叉法における交叉のノードの決定法は以下に示すとおりである。

- (1) 木構造に対して交叉を適用する深さ d を木構造の最大の深さによって決まる深さ選択率に応じて、確率的に選択する。
- (2) 選ばれた深さ d のあるノードをランダムに選び、このノードを交叉点とする。

深さ選択率は選ばれた木の最大の深さを $depth$ とし以下式によって決定する。

step1:select depth

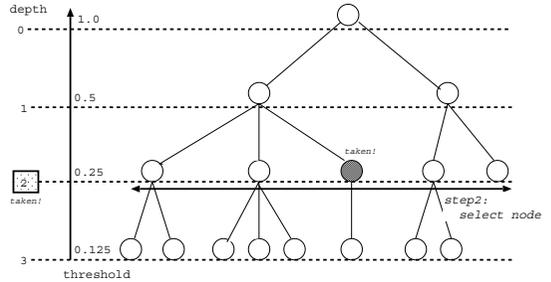


図3 深さ依存型交叉

Fig. 3 Depth-dependent crossover.

$$\text{深さ } d \text{ の選択率} = \frac{\text{threshold}_i}{\sum \text{threshold}_i}$$

$$\begin{cases} \text{threshold}_i = 1/2^i & \text{for } i = \text{depth} \\ \text{threshold}_i = \text{threshold}_{i+1} * 2 & \text{for } i = 0, 1, \dots, \text{depth} - 1 \end{cases}$$

深さ依存型交叉を用いることにより、破壊的な交叉の働きを抑制し、スキーマの蓄積を促進することが可能となっている。

Igel らは人工蟻の探索問題、4 ビット偶パリティ問題、8 入力マルチプレクサ問題にこの交叉法を適用した結果、4 ビット偶パリティ問題、8 入力マルチプレクサ問題においては優秀な成績を収めたが、人工蟻の探索問題においては通常の交叉よりも劣り、有効性は問題依存であると結論づけた。一方で深さ依存型交叉の問題点として個体の木のサイズが大きくなりやすく、プロートが起こりやすいことがこの交叉法の問題点としてあげられている。

4. 提案手法

4.1 移民世代数

並列分散 GP では、それぞれの母集団で異なるスキーマが育つ。よって新たに母集団に移住してきた移民個体、およびその子孫は母集団に土着した個体とは異なるスキーマを持つ可能性が高い。そこで、このような個体を区別するための指標が必要であると考え、各個体がどれだけ長くサブ母集団に属しているかを示す指標である移民世代数を新たに導入する。これは、三木らがハイブリッド生成交叉の際に行った移民と原住民の区別⁵⁾をその子孫にまで拡張したものであるといえる。移住率が高めに設定されているときはこの指標の意味は薄れるが、GP においては移住率が低めに設定されるため、この指標を用いて移民個体の区別を行うことは有用であると考えられる。移民個体の子孫

移民世代数の定義は以下のとおりである .

- 初期世代および他の母集団から移住してきた個体の移民世代数を 1 とする .
- 遺伝的オペレータが施された際に生成した子の移民世代数は親の移民世代数に 1 を加えたものとする .
- 交叉の際には両親のうち移民世代数の大きくない方の親の移民世代数に 1 を加えたものとする .

この定義は海外に移住した邦人の子供が日系 2 世 , その子供が日系 3 世と呼ばれることを例にとれば理解しやすいと思われる .

島での滞在が長くなり移民世代数が大きくなると , 他の島で育てられたスキーマは交叉や突然変異によって失われてしまう可能性があるが , 移民世代数の小さい個体は , 島にとって希少なスキーマ持っている可能性が高いと考えられる .

4.2 適用オペレータ

通常適用オペレータは決められた確率に基づきランダムに決定されるが , 本手法においては選択された親の移民世代数に応じて適用オペレータの決定方法を変更する . 移民世代数にしきい値を設け , 選択された個体の移民世代数がしきい値以下であった場合は確率的なオペレーションの決定を行わずにつねに深さ依存型交叉を行うものとする (図 4). これは , 島独自のスキーマと他の島のスキーマを組み合わせるより大きなスキーマを作成すること , および島に希少なスキーマが突然変異や葉に近い点での交叉によって破壊されることを防ぐことを目的としている . 移民世代数の大きい個体に対しては通常的手法で適用オペレータを決定し , 交叉は通常交叉を適用する . これは , 島ごとに局所的な探索を行いスキーマの最適化や , 新たなスキーマの生成を行うことを目的としている . 選択された個体の移民世代数がしきい値より大きい , あるいは同じ場合は従来手法と同様適用オペレータを確率的に決定する .

深さ依存型交叉のみを用いるとプロートが起りやすくなるというデメリットのほかに , 進化の序盤以外では比較的小さな部品の生成やスキーマの最適化が行われる可能性が低くなり世代が進むにつれて進化が行き詰まる可能性がある . そこで , 通常的手法を併用することによってこれらを防ぐ . この手法を用いることにより , 島モデルの特性を活かしたより効率的な探索を行うことが可能になると考えられる .

5. 評価実験

提案手法の評価を行うために , 計算ノード数 (島の

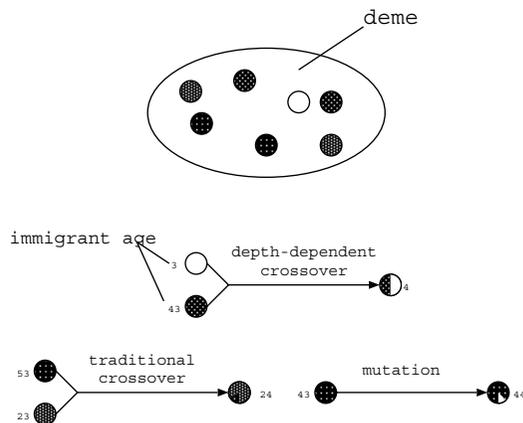


図 4 提案手法
Fig. 4 Proposed method.

数)が 8 の PC クラスタを用いて提案手法と通常の並列分散 GP の比較を行った . それぞれの問題について 20 回試行を行い , その結果の平均の比較を行った . 通常手法においては , 交叉が選択された際には深さ依存型交叉あるいは通常交叉を確率的に選択した . 交叉が選択された際に深さ依存型交叉が選択される確率を 0 , 0.2 , 0.4 , 0.6 , 0.8 , 1 と変えて適応オペレータ比率の変化が適合度に与える影響を調べた . この深さ依存型交叉の選択確率が 0.6 であり交叉率が 0.8 であれば深さ依存型交叉が施される確率は $0.8 \times 0.6 = 0.48$, 通常交叉が施される確率は $0.8 \times 0.4 = 0.32$ となることになる .

提案手法においては , 選択された個体の移民世代数がしきい値以下であればつねに深さ依存型交叉を適用し , それ以外の場合は通常手法と同様に確率的に適用オペレータを決定した . つまり , 交叉が選択された際の深さ依存型交叉適用率が 0.2 であるとする提案手法において深さ依存型交叉が適用されるのは , 選択された個体の世代数がしきい値以下であった場合およびしきい値以上の個体に対して確率的に深さ依存交叉の適用が決定された場合となる .

すべての実験において移住率を 5% とし , それぞれの島のエリート個体を移住個体として選択し , 移住のトポロジーとしては ring 型を採用した . 移住間隔を 10 世代ごと , 提案手法におけるオペレーション決定のしきい値を 5 としている . また , 交叉率 , 突然変異率はそれぞれ 0.8 , 0.2 に設定されている .

5.1 対象問題

評価の対象問題として , 以下の性質の異なる 3 つの対象問題による評価を行った .

- (1) 人工蟻の探索問題

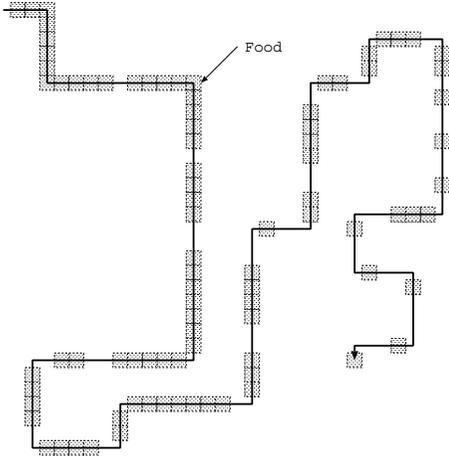


図5 サンタフェトレイル
Fig. 5 Santa Fe trail.

- (2) 8入力マルチプレクサ問題
- (3) からみ螺旋問題

5.1.1 人工蟻の探索問題

人工蟻の探索問題は GP のテスト問題としてよく知られたものである。この問題は人工蟻 (Ant) がサンタフェトレイル (図 5) と呼ばれる格子座標の地図上にある餌を限られたエネルギー内ですべて獲得するようなプログラムの生成を目的としている²⁾。

使用関数は、{if-food-ahead(A,B):Ant の一步前に餌があれば A を、さもなれば B を実行, prog2(A,B):A, B を順に実行, prog3(A,B,C):A, B, C を順に実行}, 終端記号として {left (左を向く), right (右を向く), move (前進)} を用いた。なお, Ant は終端記号を実行するごとにエネルギーを 1 消費し、餌を獲得することによってエネルギー 1 を得る。

この問題においては if-food-ahead の連鎖により構文的イントロンが創発する。また、この問題に関しては中立的なオペレーションを除くと適合度の変化はオペレーションの深さに異存しないことが知られており、深さ異存型交叉は他の対象問題とは異なり大域的探索を意味しないという特性がある。この問題においては総個体数は 2,048 とし、適合度に個体の遺伝子長 $\times 0.0001$ のペナルティを与えている。

5.1.2 8入力マルチプレクサ問題

人工蟻の探索問題と同様 GP のテスト問題としてよく知られたものであり、8 個の入力から制御信号に応じて選択的に 1 つを出力する回路を生成する問題である²⁾。使用関数は、GP のテスト問題としてよく用いられる関数セット {and, or, not, if} を用いて実験を行った。and, or はともに 2 入力である。

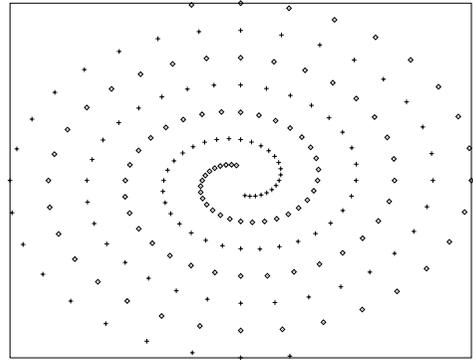


図6 からみ螺旋問題
Fig. 6 Inter-twined spirals problem.

この問題においては and, or などの組合せにより構文的イントロンが創発される。この問題においては総個体数は 1024 としている。適合度に個体の遺伝子長 $\times 0.001$ のペナルティを与えている。

5.1.3 からみ螺旋問題

からみ螺旋問題は、2 次元上からんだ螺旋を構成する 2 つのデータのクラス (図 6) を分類するという問題である¹⁵⁾。図に示したような螺旋上の点に対し、螺旋 1 上の点に対しては $f(x, y) > 0$ 、螺旋 2 上の点に対しては $f(x, y) < 0$ となるような関数を得るのが目的である。使用関数は {+, -, *, /, sin, cos, tan, log, ln}, 使用終端記号は {x, y, ランダム定数} である。

この問題においては演算関数のみが用いられるため、明確に定義された構文的イントロンは存在しない。ただし $\tan 1.6 + \sin A$ のような相対的な微小項を持つコードが創発されると、微小項の内部 (例では A の部分に相当する) が結果に与える影響は小さく、近似的に構文的イントロンと同様に振る舞うと考えられる。この問題においては総個体数は 4,096 としている。適合度に個体の遺伝子長 $\times 0.0001$ のペナルティを与えている。

5.2 結果

人工蟻の探索問題の実験結果を図 7, 図 8 に、8 入力マルチプレクサ問題の実験結果を図 9, 図 10 に、からみ螺旋問題の実験結果を図 11, 図 12 に示す。深さ依存交叉の割合を変えることにより、成績が変化することが分かる。なお、これらのグラフの適合度はペナルティ関数が加えられていない生の適合度であることに注意されたい。

まず、それぞれの図および図どうしの比較からこれらの結果に共通する特性として以下の結論が導き出される。

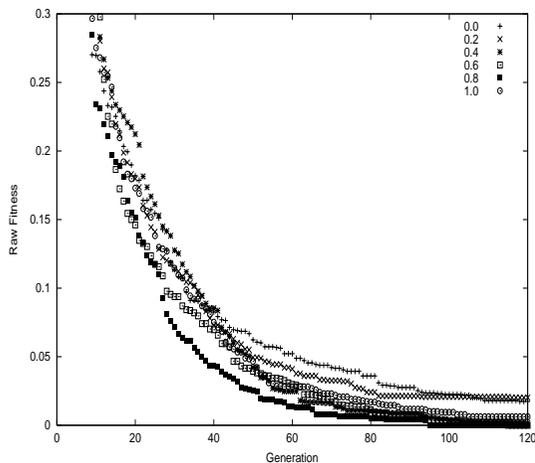


図7 人工蟻問題の世代-適合度グラフ (従来手法)
Fig. 7 Fitness vs. generation graph of ant problem (traditional method).

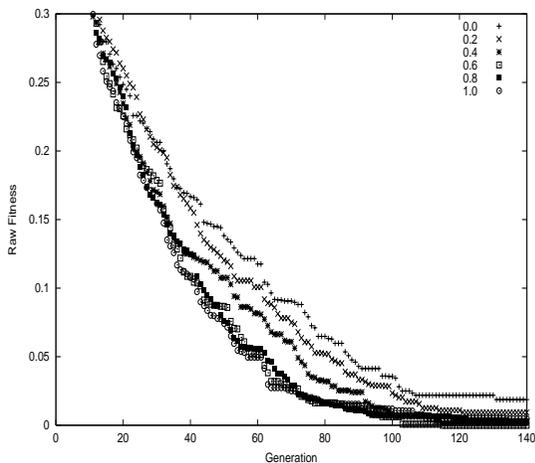


図9 8入力マルチプレクサ問題の世代-適合度グラフ (従来手法)
Fig. 9 Fitness vs. generation graph of 11-MX problem (traditional method).

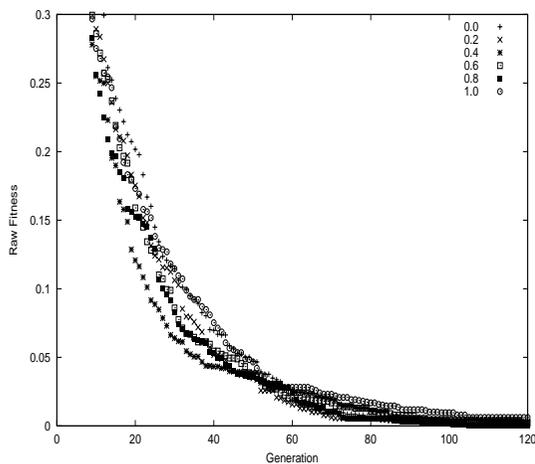


図8 人工蟻問題の世代-適合度グラフ (提案手法)
Fig. 8 Fitness vs. generation graph of ant problem (proposed method).

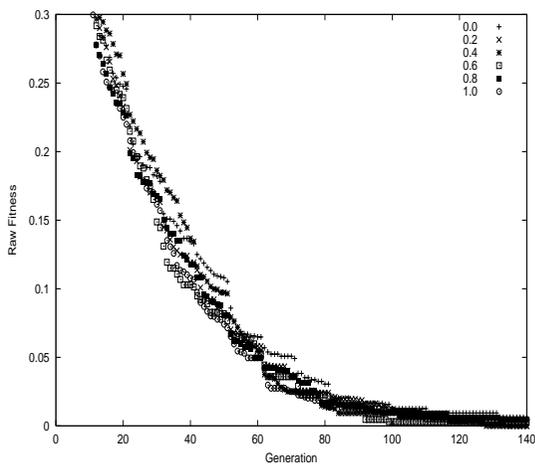


図10 8入力マルチプレクサ問題の世代-適合度グラフ (提案手法)
Fig. 10 Fitness vs. generation graph of 11-MX problem (proposed method).

- (1) 従来の確率的なオペレーション決定手法を用いると、通常の交叉のみを用いたものがすべての問題において最も悪い、あるいはそれに近い成績を残している。
- (2) 確率的な深さ依存型交叉選択率を0としたときは、移民世代数によるオペレータの決定を用いた提案手法の方が従来手法より大きく優れ、0.2に設定した場合でも提案手法の方が成績がよい。一方で深さ依存型交叉選択率が0.6以上に設定された場合は手法間に有意な差は見られない。
- (3) 深さ依存型交叉選択率を増やしていくと成績がよくなる傾向にあるが、深さ依存交叉のみを用

いるとからみ螺旋問題以外では成績が下がる。

まず、(1)より通常のGPの交叉のみを用いた手法は島モデルGPには適していないといえる。特に単一母集団モデルにおいて深さ依存型交叉が効果を発揮しなかった人工蟻の探索問題においても、明らかに深さ依存型交叉のみを用いた結果よりも劣っているのは特筆すべきである。これは、通常の交叉は大域的探索能力が乏しく、移民のスキームを破壊する可能性が高いため移民の効果を十分に活かしきれないためであると考えられる。

提案手法は、低い深さ依存交叉選択率において従来手法よりも大きく優れた結果を残している。実験の結

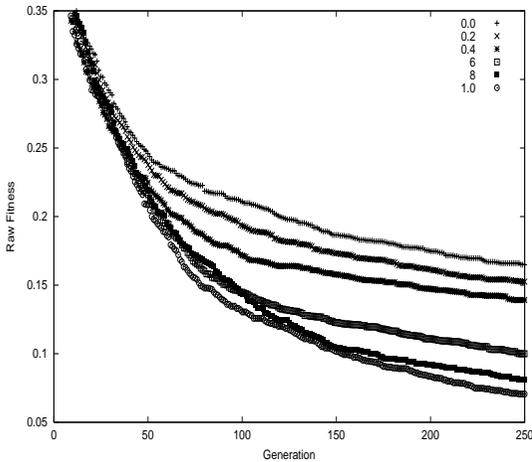


図 11 からみ螺旋問題の世代-適合度グラフ (従来手法)
Fig. 11 Fitness vs. generation graph of inter-twined spirals problem (traditional method).

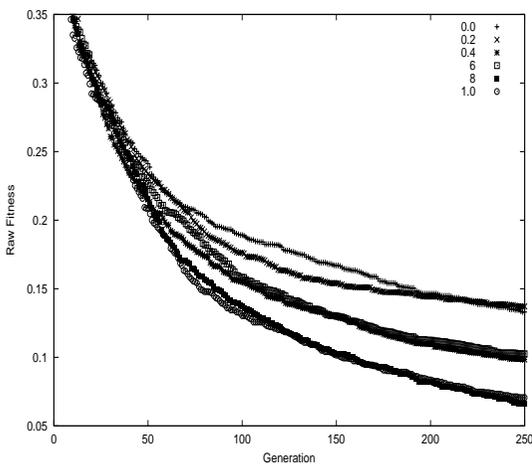


図 12 からみ螺旋問題の世代-適合度グラフ (提案手法)
Fig. 12 Fitness vs. generation graph of inter-twined spirals problem (proposed method).

果すべての場合において、確率的な深さ依存型交叉選択率を 0 とした提案手法は 2 割の確率で深さ依存型交叉を選ぶ従来手法に比べ同等以上の成績を残している。つまり、提案手法は単に深さ依存型交叉を上げることによって好成绩を得ているのではなく、島モデルの特性を活かして成績をあげていると考えられる。ただし、深さ依存型交叉の選択率が上がると本手法の意味は薄れ、結果的にも有意な差は得られていない。

また、からみ螺旋問題以外の 2 つの問題に対して提案手法を用いると低い深さ依存型交叉選択率で、深さ依存型交叉の選択率が高いものとほぼ同等の成績を収めているのは特筆すべき点である。深さ依存型交叉を

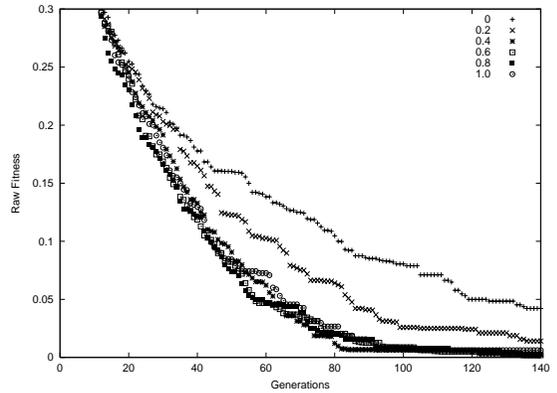


図 13 8 入力マルチプレクサ問題の世代-適合度グラフ：移住率 2% (従来手法)

Fig. 13 Fitness vs. generation graph of 11-MX problem: Immig. rate 2%.

用いる際には個体のサイズを適切に抑えないとプロットが起こり探索効率の低下や計算時間の増大をもたらすからである。本実験においては予備実験によりペナルティ関数の値を決めたため深さ依存型交叉選択率による個体サイズの差はそれほど大きなものではない。しかし個体サイズの抑制に失敗すると深さ依存型交叉の選択率の高いものほど個体サイズが育ちやすくなり、探索効率や評価計算のコストに悪影響を与えるためである。複雑な実問題などの応用の際には個体サイズの適切な抑制が行われない可能性があるため、確率的な深さ依存型交叉選択率を低めに設定し提案手法を用いることによって深さ依存型交叉に起因するプロットを抑制しつつ深さ依存型交叉導入による性能向上を実現できると考えられる。

特に従来手法において深さ依存型交叉の選択率が高いほど成績はよくなるが、深さ依存型交叉のみを用いると成績がやや低下する傾向にある。これは、深さ依存型交叉だけでは交叉を施すノードが根の近くに偏り、スキーマの最適化に必要な局所探索が効率良く行われていないためであると考えられる。

5.3 異なる移住率における実験

従来手法および提案手法が移住率の変化に対してどのような影響を受けるかを調べるために前節と同様の実験を移住率を 2%、10%として行った。8 入力マルチプレクサ問題の結果を図 13～図 16 に示す。

図 9 と図 13、図 15 の比較から分かるように通常の交叉のみを用いた従来手法においては移住率の変化にともない成績が低下している。一方で深さ依存型交叉を導入した際の効果はより大きく移住率を変えても、移住率 5% のときと同等の成績を保つことに成功して

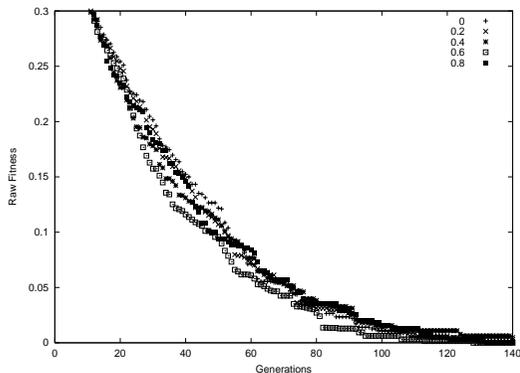


図 14 8 入力マルチプレクサ問題の世代-適合度グラフ：移住率 2% (提案手法)

Fig. 14 Fitness vs. generation graph of 11-MX problem: Immig. rate 2%.

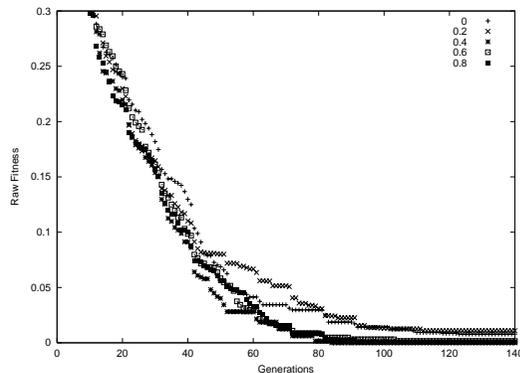


図 16 8 入力マルチプレクサ問題の世代-適合度グラフ：移住率 10% (提案手法)

Fig. 16 Fitness vs. generation graph of 11-MX problem: Immig. rate 10%.

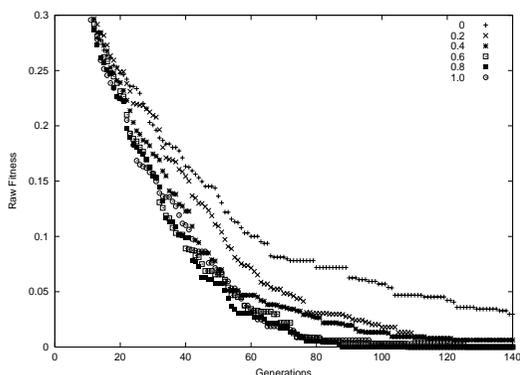


図 15 8 入力マルチプレクサ問題の世代-適合度グラフ：移住率 10% (従来手法)

Fig. 15 Fitness vs. generation graph of 11-MX problem: Immig. rate 10%.

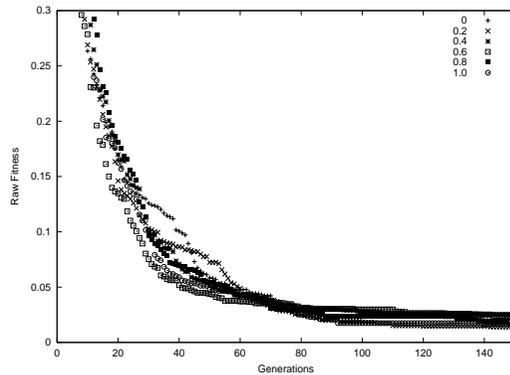


図 17 人工蟻問題の世代-適合度グラフ：単一母集団

Fig. 17 Fitness vs. generation graph of ant problem: single population (traditional method).

いる．深さ依存型交叉選択率が低い場合に提案手法は従来手法に比べて優位であり，特に移住率 2% の場合においてその差は顕著となっている．従来手法では移住率が低い場合に移民の効果を発揮できないまま滅ぼしてしまうことがしばしばあるのに対して，提案手法では確実に移民の利点を活用できていると考えられる．なお，他の問題に関しても提案手法は移住率の変化に対してよりロバストであるという結果を得ている．

5.4 単一母集団 GP における実験

島モデル GP において深さ依存型交叉と通常の手法の組合せは良好な成績をもたらした．単一母集団モデルにおいてもこの組合せは有効に働く可能性が大きいと考えられるため，対象問題それぞれに関して同様の条件における実験を行った．結果を図 17～図 19 に示す．

人工蟻の探索問題に関しては図に見られるように有

意な差は発見できなかった．Ito らは通常の交叉の方が優れていると結論づけていたがその差はさほど大きいものではなく，この問題に関しては初期値依存性が比較的大きいことが知られているためその影響が現れていたものと考えられる．8 入力マルチプレクサ問題およびからみ螺旋問題に関しては，従来手法を用いた島モデルの結果と同様に深さ依存型交叉選択率が高い方が好成績を収めている．つまり，個体サイズに対する選択圧が良好に働く場合においては単一母集団 GP においても深さ依存型交叉選択率を高めた方がよい成績となる．

6. 検証実験

6.1 移民世代数の有効性に関する検証

本論文において行われた実験に対して，確率的な深さ依存交叉選択率を低くしたときはすべての場合において提案手法は有効に働いた．特に，通常の交叉の確

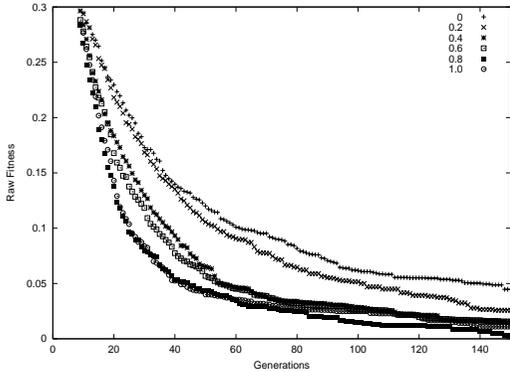


図 18 8 入力マルチプレクサ問題の世代-適合度グラフ：単一母集団

Fig. 18 Fitness vs. generation graph of 11-MX problem: single population (traditional method).

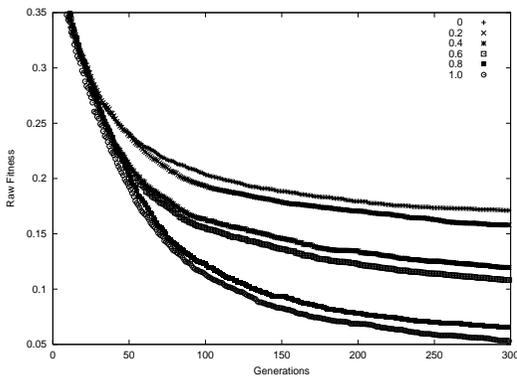


図 19 からみ螺旋問題の世代-適合度グラフ：単一母集団

Fig. 19 Fitness vs. generation graph of inter-twined spirals problem: single population (traditional method).

表 1 深さ依存型交叉の適用割合
Table 1 Rate of depth-dependent crossover.

確率的選択率	0	0.2	0.4	0.6	0.8
適用割合	0.09	0.26	0.45	0.64	0.82

率を 100%としたときにはその効果が非常に大きかった。ただし、提案手法は深さ依存型交叉が選択される確率を上げているため、それが探索効率の向上を生んでいるという可能性も否定はできない。そこで、今回の対象問題に対して提案手法を適用した際に実際に深さ依存型交叉が適用された割合をまず調べた。その結果得られた設定された深さ依存型交叉の選択率に対する実際の適用割合を表 1 にまとめる。

そこで移民世代数の指標としての有効性を確認するために、9%の割合で確率的に深さ依存型交叉を選択する手法と今回の提案手法において確率的な深さ依存型交叉の選択率を 0 とした手法の比較を行った。ここ

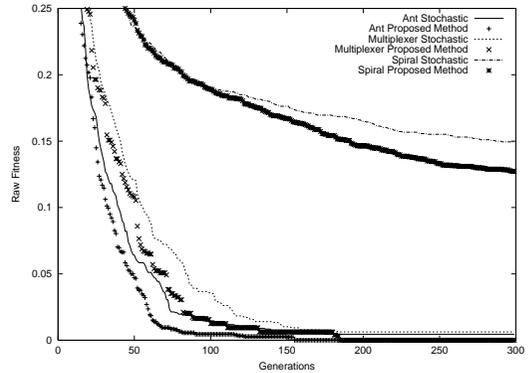


図 20 検証実験：移民世代数の有効性

Fig. 20 Verification: effectiveness of immigrant age.

で移住率は 5%としている。結果を図 20 に示す。

図 20 に示されるように、すべての場合において移民世代数を用いたものが探索効率に優れていることが分かる。これより、移民世代数は有効な指標として用いることができると結論づけることができる。

6.2 Island GP における有効なオペレーションに関する検証

深さ依存型交叉の選択率を低めに選択した際に、提案手法は今回の評価実験において単なる交叉法の選択率の変更以上の効果をもたらした。今回の対象問題について提案手法がうまく働く根拠を実験的に検証することを試みた。今回の対象問題のなかから論理回路素子を関数セットとして用いた 8-マルチプレクサ問題を選んで検証実験を行った。実験内容は、従来の手法でプログラムの進化を行い、優秀な成績を残した個体についてその個体がどのように生成されたかを調べるものである。優秀な個体の定義は、島それぞれで前世代の最良の個体よりもよい適合度を残した個体、すなわち島全体の最良適合度を上げたような個体を指す。ただし、移民によって島の最良適合度が上がっているケースは除外している。

この実験では交叉、突然変異のどちらが優秀な個体を生成しているかを調べるために、両者の生起率を等しくして実験を行った。10 回試行を行い、すべての優秀な個体について以下の情報の統計をとった。

- 個体生成の際のオペレーション
- オペレーションが施されたノードの深さと親の木全体の深さ
- 親の移民世代

進化が進むにつれ、優秀な個体を生み出すメカニズムがどのように変化するかを観測を行った。

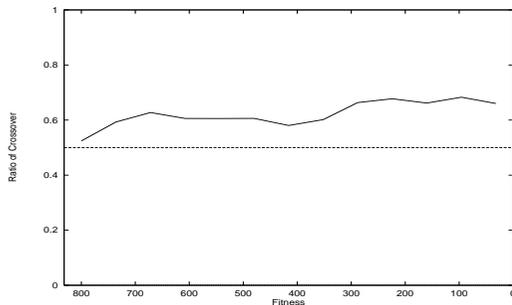


図 21 優秀個体の生成オペレーション比率

Fig. 21 Ratio of fitter individuals which is created by crossover.

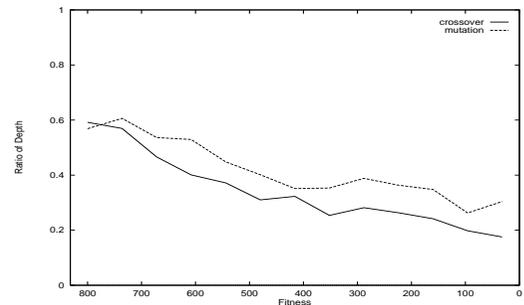


図 22 優秀個体生成オペレーションの深さ比率

Fig. 22 Depth ratio of operation by which fitter individuals are created.

6.2.1 結果

この実験の結果は移民世代との関連のグラフ以外は横軸を個体の最良適合度とした。進化の推移とメカニズムの関係という視点で考えると、世代よりも最良適合度の方が進化の推移を表現するのに適切であると考えたからである。たとえば、進化の序盤というのを表現する際に n 世代未満、というよりも最良適合度が n 以上という表現をした方がより適切だと考えたわけである。なお、エリート生き残り戦略を用いているため、最良個体の適合度が上がっているということは世代が進んでいるということになる。

まず、優秀な個体生成オペレーションのうち交叉の占める割合を図 21 に示す。横軸を一定の区間で区切り、その区間内での平均をとっている（以降の図に同じも同様である）。初期においては生起率同様 0.5 程度であるが、進化が進むにつれ交叉が直接的に優秀個体を生成していることが分かる。

次にオペレーションの深さについてのグラフを図 22 に示す。縦軸は木の深さとオペレーションが施されたノードの深さの比を表す。序盤においては深いノードでのオペレーションが有効になっている。積木仮説に基づくと、序盤は比較的小さい部品を生成するステージであると考えられ、部品生成という点では深い点でのオペレーションが適していると考えられる。進化が進むにつれて、より浅い点でのオペレーションの重要性が増していることが分かる。交叉だけではなく、突然変異においても浅い点でのオペレーションが有効であることは興味深い。

移民世代数についてのグラフを図 23 に示す。交叉によって生成される優秀な個体の移民世代数は、世代が進んでも小さい数字を保っている。島モデルの進化の過程においては移民が非常に重要な役割を占めることが分かる。

以上を総合して考えると、本手法では移民に深さ依

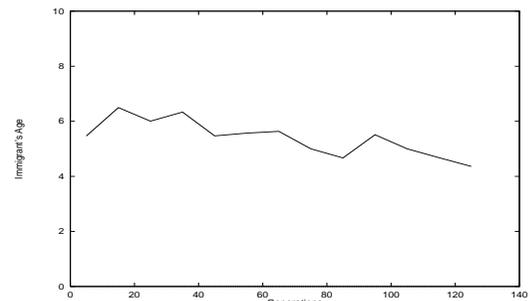


図 23 交叉による優秀個体の移民世代

Fig. 23 Immigrant age of fitter individuals which are created by crossover.

存型交叉を用いることによって優秀な個体の生成を促進し、中盤以降で新たなスキーマの生成やスキーマの最適化が起きにくいという深さ依存型交叉の欠点を通常のオペレーションによって補っていると考えられる。今後これに関してさらに定量的に解析を行っていく必要がある。

7. 結論

本論文では、並列分散 GP においては島ごとに独自のスキーマを育てることができていることに注目し、島に希少なスキーマを持つ個体の指標となる移民世代数の導入を行った。移民世代数に応じてより大域的な探索を行う深さ依存型交叉とより局所的な探索を行う従来の手法を組み合わせる手法を使い分けるという手法を新たに提案した。3つのテスト問題に関して評価を行った結果、深さ依存型交叉は島モデル GP の性能を向上させ、移住率の変化に対してもロバスト性が高くあまり成績を下げないことを示した。また移民世代数を指標として用いることで低い深さ依存型交叉率で大きな効果を得られることを示した。さらに今後の課題として、提案手法を用いる際に島数や移住間隔が与える影響についての検討、個体の選択手法および移住の

際の島のトポロジーが結果に与える影響についての検討や、さらに異なる交叉オペレータの導入などによる探索効率の向上などがあげられる。

参 考 文 献

- 1) 伊庭斉志：遺伝的プログラミング，秀和システムトレーディング (1996).
- 2) Koza, J.: *Genetic Programming: On the Programming of Computers by Natural Selection*, MIT Press, Cambridge, MA. (1992).
- 3) Angeline, P.: Subtree Crossover Causes Bloat, *Proc. 3rd Int'l Conf. on Genetic Programming*, pp.745-753 (1998).
- 4) Banzhaf, W., Nordi, P., Keller, R. and Francore, F.: *Genetic Programming: An Introduction*, Morgan Kaufmann Publishers Inc. (1998).
- 5) 三木光範，広安知之，吉田純一，大向一輝：並列分散遺伝的アルゴリズムにおける最適な交叉スキーム，第五回数理モデル化と問題解決シンポジウム論文集，pp.49-56，情報処理学会 (2000).
- 6) Andre, D. and Koza, J.: A Parallel Implementation of Genetic Programming That Achieves Super-Linear Performance, *Information Science Journal* (1997).
- 7) Fernandez, F., Tomassini, M., Punch, W. and Sanchez, J.: Experimental Study of Mutipopulation Parallel Genetic Programming, *EuroGP'00*, Banzhaf, W., Poli, R., Schoenauer, M. and Fogarty, T. (Eds.), Springer (2000).
- 8) Fernandez, F., Tomassini, M. and Vanneschi, L.: Studying the Influence of Communication Topology and Migration on Distributed Genetic Programming, *EuroGP'02*, Springer (2002).
- 9) Tanese, R.: Distributed Genetic Algorithms, *Proc. 3rd International Conference on Genetic Algorithms*, pp.434-439 (1989).
- 10) 廣安知之，三木光範，上浦二郎：分散遺伝的アルゴリズムにおけるパラメータの検討，第8回MPSシンポジウム，情報処理学会 (2001).
- 11) Punch, W.: How Effective Are Multiple Populations in Genetic Programming, *Proc. 3rd Annual Conference on Genetic Programming*, Koza, et al. (Eds.), pp.308-313 (1998).
- 12) Andre, D. and Koza, J.: A Parallel Implemen-

tation of Genetic Programming That Achieves Super-Linear Performance, *Information Science Journal*, pp.201-218 (1997).

- 13) Igel, C. and Chellapilla, K.: Investigating the Influence of Depth and Degree of Genotypic Change on Fitness in Genetic Programming, *Proc. Genetic and Evolutionary Computation Conference*, pp.1061-1068, Morgan Kaufman (1999).
- 14) Ito, T., Iba, H. and Sato, S.: Depth-Dependent Crossover for Genetic Programming, *Proc. IEEE International Conference on Evolutionary Computation*, pp.775-780, IEEE Press (1998).
- 15) Koza, J.: A Genetic Approach to the Truck Backer Upper Problem and the Inter-Twined Sparal Problem, *Proc. IJCNN International Joint Conference on Neural Networks*, Piscataway, NJ, pp.310-318, IEEE Press (1992).

(平成 14 年 2 月 5 日受付)

(平成 14 年 4 月 12 日再受付)

(平成 14 年 5 月 15 日採録)



岩下 誠

1975 年生．2000 年東京大学工学部電子情報工学科卒業．2002 年同大学大学院新領域創成科学研究科修士課程修了．現在同大学院博士後期課程在学中．並列計算，創発等に興味を持つ．



伊庭 斉志 (正会員)

1962 年生．1985 年東京大学理学部情報科学科卒業．1990 年同大学大学院工学系研究科情報工学専攻博士課程修了．工学博士．同年電子技術総合研究所入所．1998 年から東京大学大学院工学系研究科電子情報工学専攻助教授．1999 年から東京大学大学院新領域創成科学研究科基盤情報学専攻助教授．進化システムおよび人工知能基礎の研究に従事．特に遺伝的プログラミング，学習，推論，知能ロボットに興味を持つ．