Parallel Genetic Algorithms

片浦哲平，下坂久司

1　はじめに

進化的アルゴリズムは大量の数値計算 (評価計算) をする必要がある．そのため，コンピュータのパワーが必要となる．しかしながら，進化的アルゴリズムは，次のような理由から並列化を比較的簡単に行うことができる．

- 母集団ベースの探索手法
- 独立した適合度

その他にも並列化に関して，いくつかの利点がよく知られている．

- 高信頼性
- 高品質の解
- 高いスピードアップ

並列化には，次のようなシステム (ツール) を使用できる．

- ベオウルフ型クラスタシステム
- PVM や MPI などのパブリックドメインの通信ツール
- パブリックな GA のソースコード

2　概論

並列 GA のタイプとして，次の 4 つのようなものがある．

- 単一母集団マスタースレーブモデル
- 複数母集団モデル
- 細粒度並列化モデル
- 階層的な組合せモデル

2.1　単一母集団マスタースレーブモデルの概要

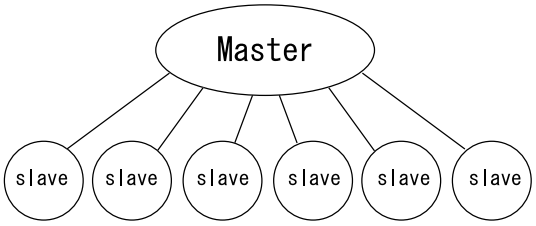

Fig. 1　Single-population master-slave

単一母集団マスタースレーブモデルの特徴は，単一の母集団であり，マスターは選択，交叉，突然変異を行う．スレーブでは評価を行う．

2.2　複数母集団モデルの概要

複数母集団モデルでは，地域交配集団同士が接続する．最も一般的な手法であるが，実装は複雑である．

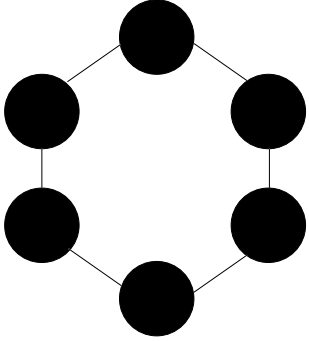

Fig. 2　Multiple population

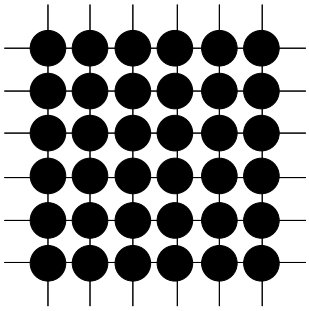

Fig. 3　Fine-grained

2.3　細粒度並列化モデルの概要

細粒度並列化モデルは，空間的な構造を持つ集団であり，近傍と選択や交叉を行う．

2.4　階層的な組合せモデルの概要

複数母集団はそれぞれが並列モデルである．例えば，Fig. 4 であれば，複数母集団モデルと単一母集団モデルの両方の利点を持つ．

3　単一母集団マスタースレーブモデル

3.1　同期モデル vs 非同期モデル

単一母集団マスタースレーブモデルには、マスターとスレーブが個体を移住させる方法に同期型と非同期型が存在する．

同期型の特徴は次の通りである．

- 全ての評価を待つ必要がある
- 遅いスレーブを待つ必要がある
- Simple GA と同じ

非同期型型には次のような特徴がある．

- 遅いスレーブを待つ必要がない
- 可能な限り速い
- Generation gap が生じる

3.2　同期型マスタースレーブモデル

同期型マスタースレーブモデルの一世代は，コンピュータ操作と通信である．

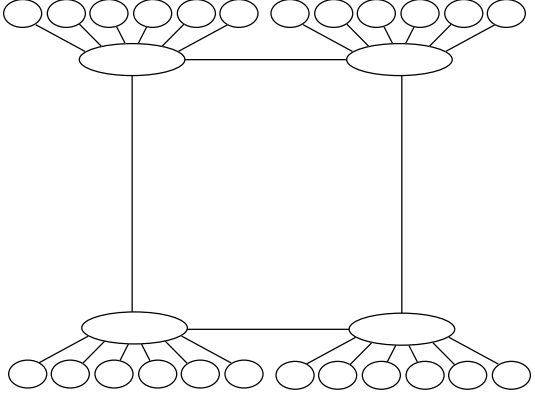

Fig. 4　Hierarchical

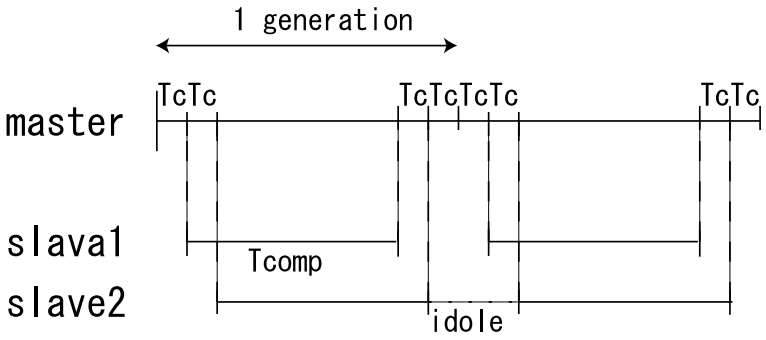

Fig. 5　Synchronous master-slave GAs

コンピュータ操作時間に関しては，次のような式で表せる．

$$Computation\ time : \frac{pop\ size \times eval\ time}{num\ proc} = \frac{nT_f}{P}$$

通信時間に関しては，次のような式で表せる．

$$Communication\ time : num\ proc \times comm\ time = PT_c$$

つまり，プロセス数が増加すれば，コンピュータの操作時間は減るが，通信時間は増加する．

$$T_p = \frac{nT_f}{P} + PT_c$$

よって，プロセス数の最適値は次のようになる．

$$P = \sqrt{\frac{nT_f}{T_c}}$$

また次のようなことを付加できるかもしれない．

- マスターでの評価計算
- 通信を隠す
- ロードバランスをとる

1 プロセスの時と比較したスピードアップに関しては次のような式で表せる．

$$S_p = \frac{T_s}{T_p} = \frac{nT_f}{\frac{nT_f}{P} + PT_c}$$

1000 個体で，$\frac{T_f}{T_c} = 1, 10, 100$ と変化させた場合の，スピードアップのグラフを Fig. 6 に示す．

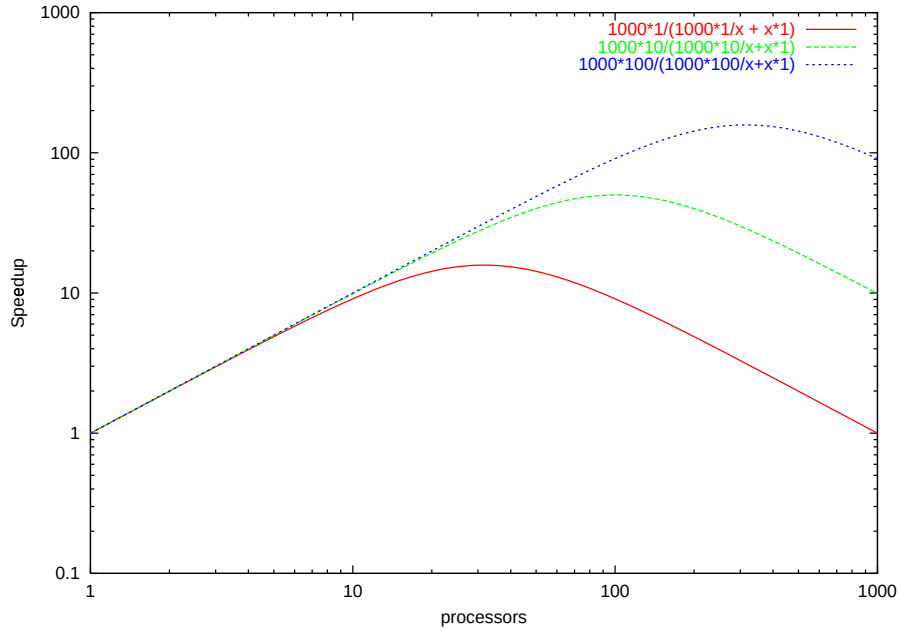

Fig. 6　speedup

3.3　非同期型マスタースレーブ モデル

非同期マスタースレーブモデルでは，マスターからスレーブに個体が送られる．スレーブの処理が終了した際，マスターは次のようなことを行う．

- 母集団に個体を挿入する (どうやって)
- 個体の生成 (どれくらい)
- 新しい個体をスレーブへ送る

非同期型マスタースレーブモデルでは次のような可能性がある．

- 遅いスレーブを待つ必要がない
- 評価時間が様々な場合に有効

一方で，GA オペレーションを並列化する方法もある．これは，アルゴリズムをより確実なものとするためのもので，スレーブノードでも選択や交叉を行うものである．このためには，更なる通信を付加する必要がある．

3.3.1　Parallel GA operations

GA オペレーションを並列化する際，母集団を分割するという方法もある．また，選択や交叉のためにマスターとスレーブは k 回の通信をする必要があるとすると，全体の処理時間は次のようにあらわせる．

$$T_p = \frac{nT_f}{P} + k(P-1)T_c$$

よって，最適なプロセス数は次のとおりである．

$$P = \sqrt{\frac{nT_f}{kT_c}}$$

3.3.2 1つか複数の母集団か？

母集団が一つの場合，全体の GA オペレーションを並列数で割った数に分割して処理できる．

複数の母集団を設定すると，母集団間で十分な通信を行う必要がある．その際，移住率や移住間隔を考慮する必要がある．

3.3.3 母集団のサイズ 1

マスタースレーブモデルのプロセス数は個体数に依存する．母集団サイズは解品質と持続期間を確定する．

多数の小さな母集団では，次のような特徴がある．

- 集中型よりも高速
- 解の品質は低い

3.3.4 Building Block(s)

Fig. 7 に Building Block(以下 BB) の例を示す．

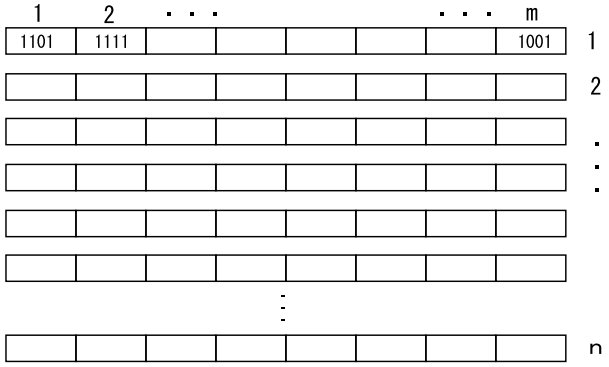

Fig. 7 Building block

BB は個体のビット列を m 個のパーティションに区切った一つを示す．Fig. 7 では，個体のビット列を 4 ビットずつ，m 個のパーティションに区切った，n 個の個体群である．

3.3.5 Deciding Well Between Two BBs

選択における 2 個体の競争において，良い個体が選ばれる必要があるが，たまに悪い個体が選ばれる．パーティション H_1 に，最適な BB(Building Block) を含む，個体 i_1 と，パーティション H_2 に，2 番目に最適な BB(Building Block) を含む，個体 i_2 の競争を考えた場合，選択操作で i_1 が選択されるが，誤って i_2 を選択する可能性がある．これは，他のパーティションが i_1 の適合度のアドバンテージを超えるほど大きな貢献をしたためである．

これらの 2 個体間の選択が正しく行われる可能性は，個体 i_1 の適合度 f_1 が個体 i_2 の適合度 f_2 が，等しいか大きい可能性である．つまり，$f_1 - f_2 > 0$ となる確率である．

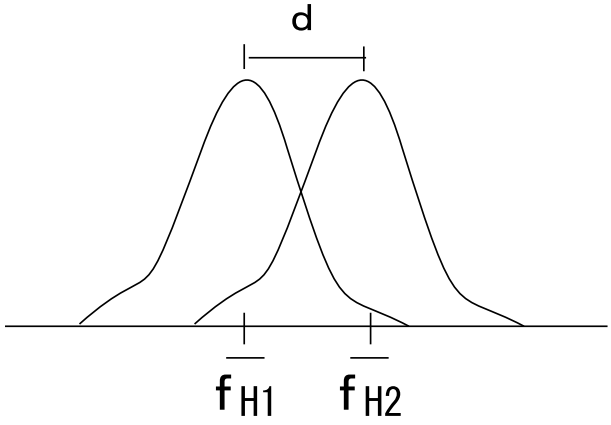

Fig. 8 Fitness distributions for two competing individuals

Fig. 8 は，H_1 と H_2 を含む個体群の分布である．H_1 を含む個体群の分布の平均値を，f_{H_1}，H_2 を f_{H_2} として，その距離を d とする．中心極限定理[1]により，f_1 と f_2 の分布は，正規分布となる．よって，$f_1 - f_2$ もまた，個体の平均値の差を平均値に，個体の分散の和を分散にとる正規形である．

$$f_1 - f_2 \sim N(\bar{f_{H_1}} - \bar{f_{H_1}}, \sigma^2_{H_1} + \sigma^2_{H_2})$$

$d = \bar{f_{H_1}} - \bar{f_{H_1}}$ より，一回試行で正しい選択を行う確立は次のように表される．

$$p = \phi(\frac{d}{\sqrt{\sigma^2_{H_1} + \sigma^2_{H_2}}})$$

ϕ は，平均値が 0 の標準偏差が 1 の標準正規分布の積分値である．

$\sigma^2_{H_1}$ と $\sigma^2_{H_2}$ の計算は，関数 F の適合度は m 個の独立したサブ関数 F_i(サイズ k) の和で計算できると仮定して，上の分散は次のように表現できる．

$$\sigma^2_F = \sum_{i=1}^{m} \sigma^2_{F_i}$$

外から受けるノイズは，$m' = m - 1$ であるので，$\sigma^2 = m'\sigma^2_{bb}$ となる．よって，一回試行で正しい選択が行われる確率は次のように表せる．

$$p = \phi(\frac{d}{\sqrt{2m'}\sigma_{bb}})$$

この確率は，解の精度を決定するために，母集団がどれくらい必要かを説明する最初のモデルとして作成された．これらのモデルは母集団のサイズを決定する問題に，外部からのノイズの効果をどのように取り入れるかを示す．そして，確率 p を見積もる方法を説明している．

3.3.6 Gambler's ruin model

Gambler's ruin 問題は，確立的なプロセスで結果を予測するための数学的ツールである，古典的な random walk の例題である．最も基本的な例は Fig. 9 のような，ある要素が確立的に左右に動く，一次元のものである．ステップサイズは一定である．要素の動作は，時々いくつかのポイントで制限を受ける．

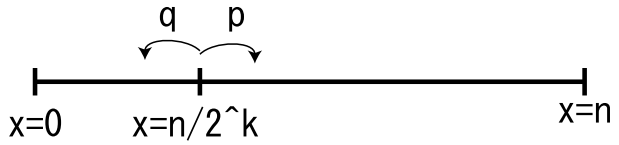

Fig. 9　The bounded one-dimensional space of gambler's ruin problem

Fig. 9 では，両端の点で要素が捕らえられる．$x = 0$ では破産，$x = n$ は勝利である．この場合，重要なのは，p と初期点の値である．

もし，パーティションが独立していて，一つのパーティションの列に注目したとき，Gambler's ruin problem は GA の選択操作に類似している．個体群のあるパーティション列において，x は正しい BBs の個数に置き換えられる．収束点は $x = 0$ と $x = n$ である．初期点 x_0 は，個体群をランダムに生成するので，BB が $k$bit の列であれば，$x_0 = \frac{n}{2^k}$ となる．

この Gambler's ruin model を用い，GA の解の精度を予測するためにいくつかの仮定を行う．まず一つ目に世代という明白な概念はなく，決定の結果は最適な BB のコピーを一つ増やすか減らすかで行う．一番良い BBs と二番目に良い BBs 間での競争しか行わない．この場合，最適な BBs を残す可能性は，$p = \phi(\frac{d}{\sqrt{2m'}\sigma_{bb}})$ となる．これは，暗黙のうちに 2 つの文字列の競争におけるトーナメント選択となっているが，他の選択法も調整することで使用可能である．また，BBs の初期化はランダムに行わなければならない．交叉や突然変異は重要な BBs の増やしたり減らしたりすることはない．よって，$x = 0, x = n$ の収束点に一度入ると抜け出ることはできない．

[1] 母集団分布が正規分布でなくても，標本が大きくなると標本平均値の分布は次第に正規分布に近づく

上のような条件から，$x = n$ で収束する可能性は，よく知られたランダムウォークの文献から，次のように表せる．

$$P_{bb} = \frac{1 - (\frac{q}{p})^{x_0}}{1 - (\frac{q}{p})^{n}}$$

$q = 1 - p$ は，BB のコピーが一つ減る可能性である．$p > 1 - p$(BestBB の適合度を持つ個体の分布の平均値は，Second BestBB のものより大きい) より，n が十分大きい場合は，$x_0 = \frac{n}{2^k}$ は次のように表せる．

$$P_{bb} = 1 - (\frac{1-p}{p})^{\frac{n}{2^k}}$$

各パーティションは独立しているので，全てが正しい BB になる確率は $\bar{Q} = mP_{bb}$ となる．
母集団のサイズを得るために，次の式が導ける．

$$n = \frac{2^k \log(\alpha)}{\log(\frac{1-p}{p})}$$

$\alpha = 1 - \frac{\bar{Q}}{m}$ は GA が失敗する可能性である．
また，先程導いた p を拡張し，次のように表す．

$$p = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} z$$

$z = \frac{d}{\sigma_{bb}\sqrt{2m'}}$ である．これを，代入すると，

$$n = \frac{2^k \log(\alpha)}{\log(\frac{1 - \frac{z\sqrt{2}}{\pi}}{1 + \frac{z\sqrt{2}}{\pi}})}$$

$$n = -2^{k-1} \log(\alpha) \frac{\sigma_{bb}\sqrt{\pi m'}}{d}$$

このことから，いくつかの直観的にわかるものがある．例えば，長い BBs(large k) は短い BBs(short k) よりも難しい．SN 比 (the-signal-to-noise ratio) に反比例して，母集団のサイズが必要である．分散の大きい問題は，良い解の影響が小さくなるので，難しい問題となる．ノイズが増えるので，Longer Problem (larger m) は，よりパーティションが少ないものより，難しくなる．問題サイズの平方根で母集団サイズが必要になる．

3.3.7 Experimental Verification

次のような条件で，SimpleGA の解品質を Gambler's ruin model を用いて，予測した結果を Fig. 10 に示す．

- The one-max problem
- 100-bit one max
- m=100(k=1)
- d=1
- $\sigma_{bb}^2 = 0.25$

選択において，正しい BBs を得られる確率は $p = 0.5565$ です．
また，4bit の substring の適合度を，1 の数によって，Fig. 11 のように変化させた trap madel を用い，次のような条件で，SimpleGA の解品質を Gambler's ruin model を用いて，予測した結果を Fig. 12 に示す．

- 適合度は，substring の適合度の和
- m=20 (k=4)
- d=1
- $\sigma_{bb}^2 = 1.215$

選択において，正しい BBs を得られる確率は $p = 0.5585$ です．

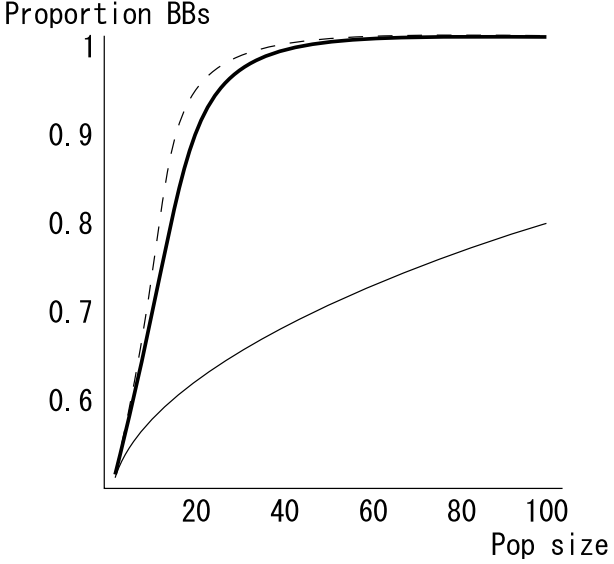

Fig. 10　pop size1：The predict of gambler's ruin model is in bold,the experimental results are dotted line,and previous decition-based model is the thin line

4　複数母集団モデル

複数母集団モデルとは, 別名,

- 島モデル
- 分割母集団モデル
- 粗粒度モデル

と呼ばれるモデルで, その名の通り母集団を複数に分割する方法である.
現在では最も一般的な GA のモデルであるが, 最も複雑なモデルであるとされる.
その理由として以下のようなパラメータの設定の複雑さが挙げられる.

- 母集団の分割数をどのように設定するか
- 母集団の個体数をどのように設定するか
- １つの母集団内で行う処理と通信する処理をどのように分けるか
- 移住率, 移住方法はどのように設定するか
- 通信トポロジーをどうするか

4.1　Isolated demes

isolated demes は言葉通り, 分割した複数の母集団内でのみ交配を行う手法である. これは,SGA を分割した母集団ごとに実行し, すべての母集団で計算が終わった段階で解を集めてその中で最も良かった解を最適解とする.
それぞれの母集団での解の平均品質は mP_{bb} で表現される．この手法で,Building blocks を横軸に取り，個体数を縦軸に取ると，Fig. 13 のように二項分布になる傾向が高い. (このパーティションをそれぞれ X_1,X_2,X_3,…,X_r とする．)

Fig. 13 は Building block が 10 個で 1 個体であるような個体が Building block 数中何個の最適解があるかを分布図で示したものである．図の通り，全てが最適解である可能性，全てが最適でない解の可能性は低く，ちょうど半分の 5 個前後になる確率が高いことは容易に理解できるだろう.

r 個の母集団で考える場合に，まず，図の正規化を行う．i 番目のパーティション $X(X_i)$ を正規化した値を z_i とすると，

$$z_i = \frac{X_i - mP_{bb}}{\sqrt{mP_{bb}(1-P_{bb})}}$$

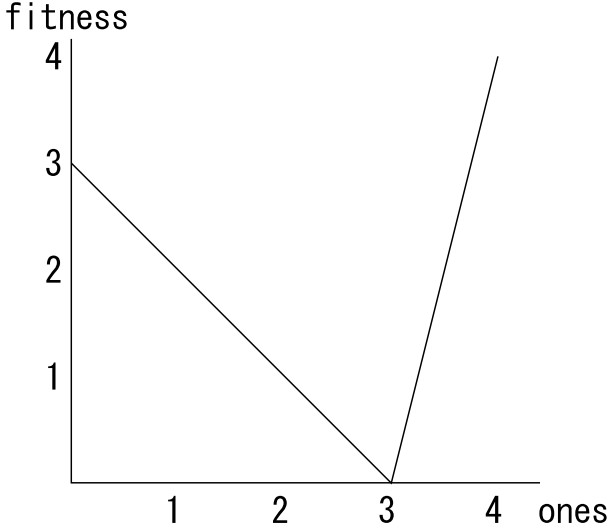

Fig. 11　trap function

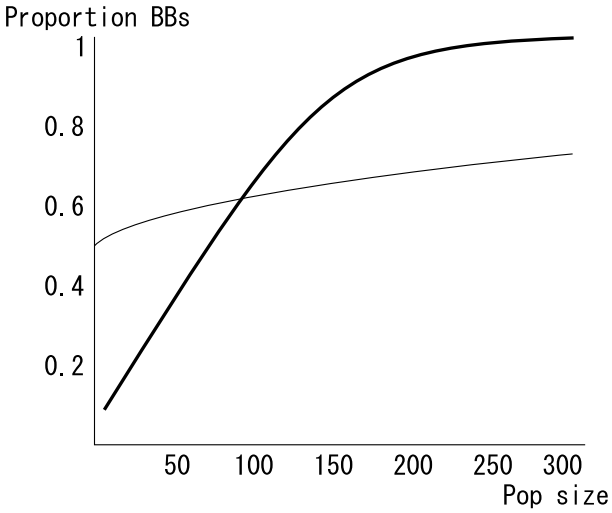

Fig. 12　pop size2:The predict of gambler's ruin model is in bold,and previous decition-making model is the thin line

となる．そして，Z_r の期待値を $\mu_{r:r}$ とすると，r 個の母集団全ての期待値 (平均)Q は，

$$\hat{Q}_{r:r} = mP_{bb} + \mu_{r:r}\sqrt{mP_{bb}(1-P_{bb})}$$

で表すことができる．この式に二項分布，および，正規分布から最も頻繁に起こるのは 5 の場合，すなわち P_{bb}=0.5 である．この値を上記の式に代入すると

$$\hat{Q} = E(X_r) \leq mP_{bb} + \frac{\mu_{r:r}}{2}\sqrt{m}$$

この式を P について解くと，

$$\hat{P} = \frac{\hat{Q}}{m} - \frac{\mu_{r:r}}{2\sqrt{m}}$$

となる．この式から，正しい building block が選ばれる確率は，分割母集団数 r が増えるにしたがって緩められると言うことが分かる．

また，上記の式は

$$\hat{P} = \frac{\hat{Q}}{m} - \sqrt{\frac{lnr}{2m}}$$

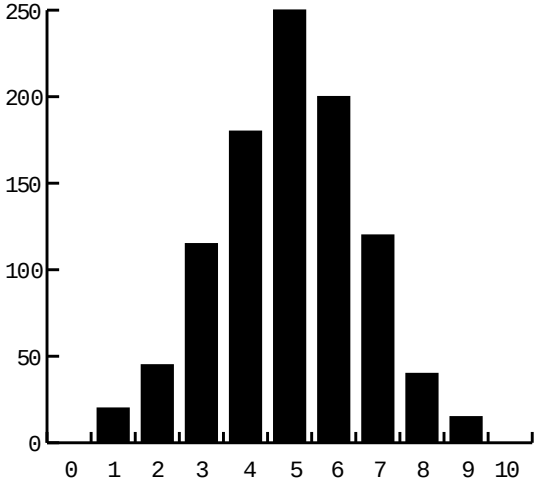

Fig. 13　二項分布

のようにかける．

4.1.1　並列化の効果

Gambler's ruin model で P_{bb} は

$$P_{bb} = 1 - (\frac{q}{p})^{\frac{n}{2^k}}$$

で求めることができた．同様にして，分割母集団ごとの確率は

$$\hat{P} = P_{bb} = 1 - (\frac{q}{p})^{\frac{n_d}{2^k}}$$

これから，

$$n_d = \frac{2^k ln(1-\hat{p})}{ln(\frac{q}{p})}$$

と表せる．これを近似すると，

$$n_d = 2^{k-1} ln(1-\hat{p}) \sqrt{\pi m'} \frac{\sigma_{bb}}{d}$$

となる．
並列化した場合の速度向上は単純に考えるならば，総個体数と分割母集団の個体数との比で表すことができる．式で表せば，

$$S_p = \frac{n}{n_d}$$

となる．上記の式では個体数でその比を表現しているが，これと同等の意味で単純な等式で表せる式が

$$S_p = \frac{T_s}{T_p} = \frac{ln(1-\frac{Q}{m})}{ln(1-P)}$$

である．この式から，スピードアップの比率も分割母集団数 r によって考えることができる．Fig. 14 にスピードアップの比率を示す．

この図は 4bit を 1BB として m=20，すなわち 1 個体が 80bit の個体に trap function がある場合のスピードアップの比率を表したものである．解の品質は BBs の 80%が正確であったものを選んでいる．

この結果は 100 試行の平均をとったものであり，実験の結果からも Isolated demes での並列化の効果は大きく，分割母集団 (1〜16) の個数を増やせば増やすほど速くなるということになった．しかし，分割母集団をどの程度まで増やすことができるのかを調べることは避けることのできない実験であるとしている．

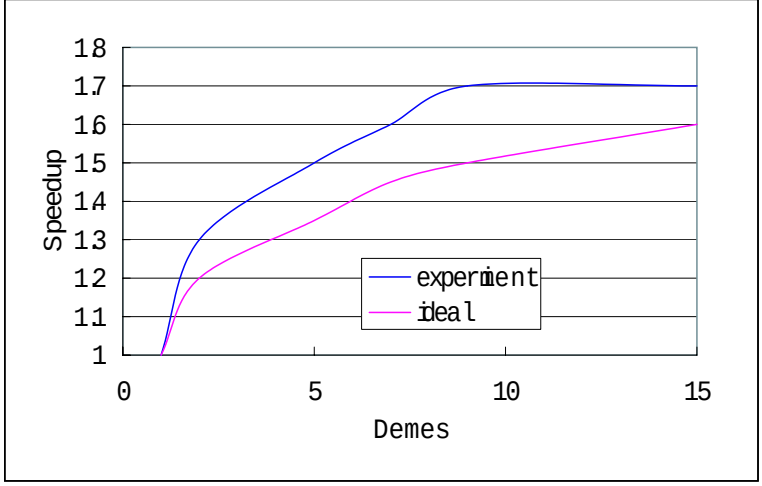

Fig. 14　Speedup

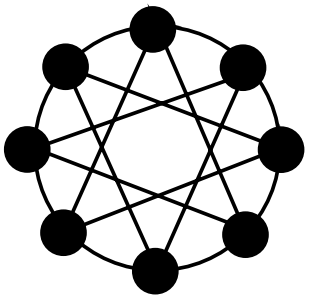

Fig. 15　Fully connected demes

4.2　Fully connected demes

先程の Isolated demes とは異なり Fig. 15 のように分割したすべての母集団の個体を設定した世代ごとに移住させる方法を Fully connected demes と呼ぶ.

この手法では, 最大移住 BB 数を

$$\frac{n_d}{r}$$

すなわち, 分割母集団の個体数／分割母集団数とし, 移住間隔が最も頻繁なものは 1 世代ごとに最も少ないものは Isolated demes と同じように計算が終わった後に解交換を行うものとしている. (Grosso,85,Braun,90;Munetomo etal,93)

4.2.1　アルゴリズム

Fully connected demes のアルゴリズムは以下のようになっている.

1. 各母集団ごとに収束するまで計算をする
2. 移住率にしたがい, すべての母集団で個体の交換を行う
3. 再び母集団ごとに計算をする

移住は, P_{bb} で選ばれた BB を他の母集団に与える方法をとる. 移住においては正しい BB もそうでない BB もそれぞれの確率によって選ばれ, 正しいもののみが選ばれるというわけではない. したがって, 移住が終わった後の BBs には正確な BB が nP_{bb} 個コピーされたことになる.

さらによい解を見つけるために, 先ほど説明した gambler's ruin model を用いる. 開始点は

$$P_{bb} = 1 - (\frac{q}{p})^{x_0}$$

で x_0=nP_{bb} となる点から探索を行う. したがって, 開始点が正確な遺伝子である確率は

$$P_{bb2} = 1 - (\frac{q}{p})^{n_d} P_{bb}$$

そして，同様に計算を進めて

$$E(X_r) = mP_{bb2} + \mu_{r:r}\sqrt{mP_{bb2}(1-P_{bb2})}$$

$$P_{bb2} = 1 - exp(\frac{-c^2 n_d^2}{2^k})$$

$$n_d = \sqrt{-2^k ln(1-\hat{p})\pi m' \frac{\Sigma_b b}{2d}})$$

Fig. 16 に並列化した場合の効果を示す.

4.2.2 並列化の効果

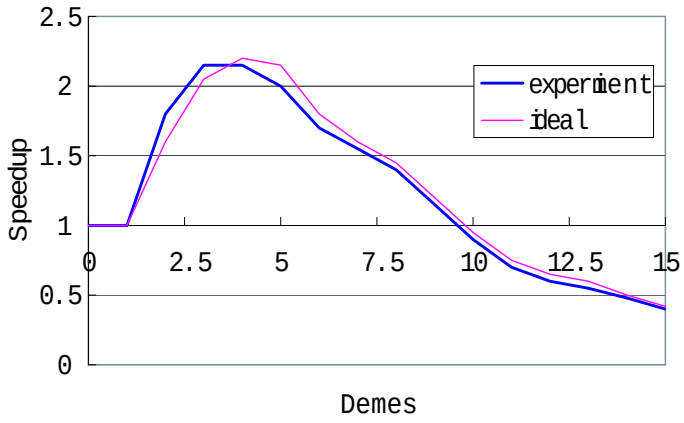

Fig. 16 Speedup1

4.2.3 並列化の効果 2

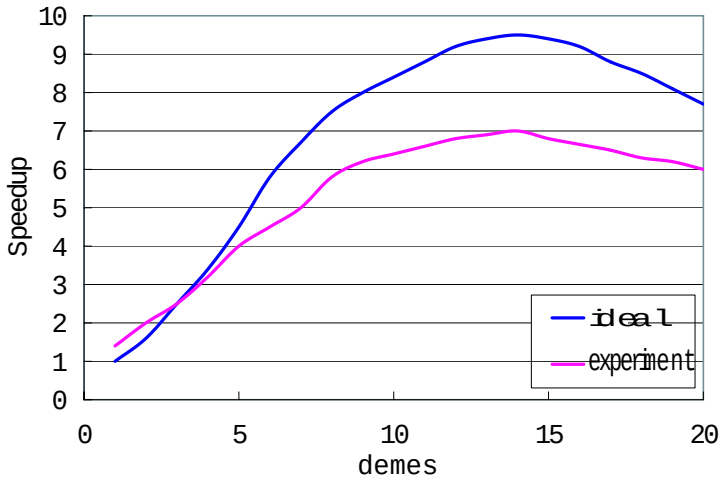

Fig. 17 Speedup2

Fig. 16 は 1BB が 4bit，1 個体が 80bit の trap function で実行した場合である．図を見れば分かる通り，Isolated demes の場合は母集団数を増やせば処理速度が向上していたのが，Fully connected demes の場合は 5 個あたりをピー

クに徐々に悪くなっているのが分かる．Fig. 17 は 1BB が 8bit，1 個体が 80bit の trap function で実行した場合の図である．この場合ならば，demes 数は 15 個あたりがピークになっているのが分かる．

Isolated demes 同様にスピードアップについても理論的な式で表すことができるのだが，自分の勉強不足で理解することができなかった．ただ，私の感想としては，いくら公式化したとしても，実際にその式がどのような傾向になっているのかを判断するには図示する以外方法が無いように思えた．

計算は各母集団での収束，移住，…の繰り返しとなるが，最も重要なのが各母集団で収束するまで計算をして移住を行う第 1 世代と，移住後に収束するまで計算を行う第 2 世代である．なぜならば，多くの難しい研究の対象は第 2 世代以降のもっと改善のされた解を解析することを目的としているが，逆に，その最初の 2 世代において高い精度を持つ解の探索法を考えることが重要であると考えるからだと述べられている．

今回述べた gambler's ruin model はその第 2 世代までに精度を高める方法として考えられたものである．そのため，どのくらいの母集団を集めればよいかなど初期のパラメータを正確に知っておく必要がある．

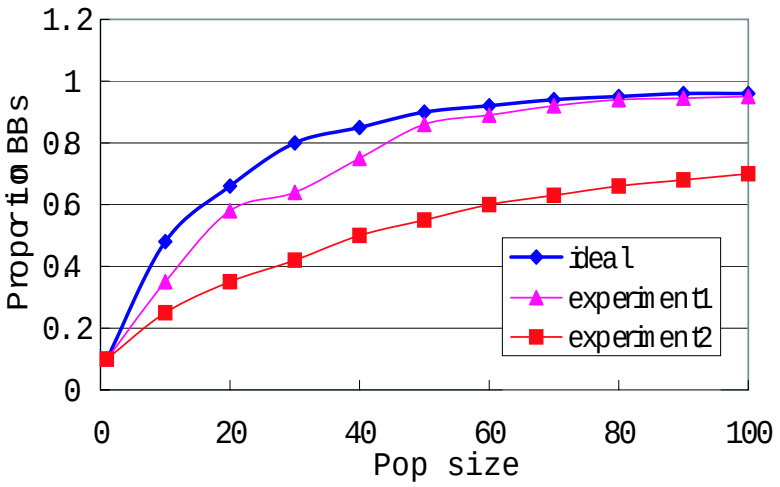

Fig. 18　BBs の 1bit あたりの fitness

4.2.4　移住率

移住率はより高くする方が効果的である．

Fig. 19 は，1BBs が 4bit で trap を用いている．また，分割母集団数は 4 で，各母集団の個体数は 50，最初の 2 世代の図である．

4.2.5　隣接数

Fig. 20 は，1BBs が 4bit で trap を用いている．また，10%の移住率である．これも，最初の 2 世代の図である．

5　細粒度並列化モデル

細粒度並列化モデルは，セルラー GA などとも呼ばれる．

パラメータには次のようなものがある．

- 母集団のサイズ
- 母集団の構造 (トポロジー)
- 交叉戦略
- 近傍構造の設定

5.1　交叉戦略

交叉する親を選択する際に，様々な方法が考えられる．

- 平均的な選択

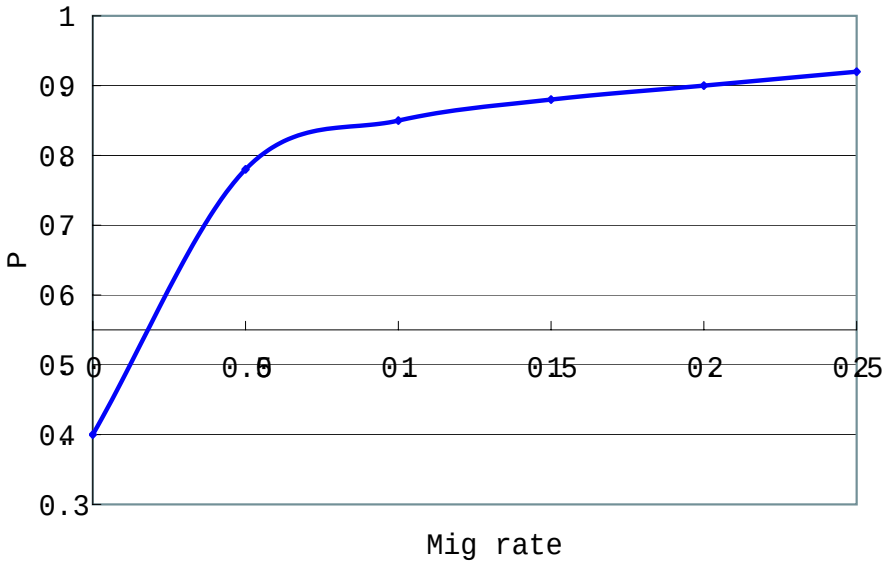

Fig. 19　移住率

- トーナメントによる選択
- ランダムウォークによる選択

子個体をどのように扱うかについて，ローカル集団の個体と置き換える方法がよくとられるが，これはあまり良くない方法である．

5.2　近傍構造の設定

近傍構造の設定は，選択に大きな影響を与える．重要なパラメータは次のようなものである．

$$Critical = \frac{radius\ neighborhood}{radius\ entire\ grid}$$

近傍を大きくとると，高速に収束する．これは，ある個体が他の個体と多く入れ替わるためである．よって，近傍のサイズで，個体が入れ替わる回数を分析する必要がある．

6　階層的な組合せモデル

階層的 GA には Fig. 21 や Fig. 22，Fig. 23 などのものが考えらえれる．

階層的組合せモデルでは，最適な地域交配集団のサイズやスレーブノードの数を，実験や計算等によって，探しだす必要がある．

参考文献

1) The Gambler's Ruin Problem,Genetic Algorithms,and the Sizing of Populations : George Harik 他 : IlliGAL Report No. 96004 July 1996

2) Parallel Genetic Algorithms : Erick Cantu-Paz :

文責：片浦哲平，下坂久司（hisashi@mikilab.doshisha.ac.jp）

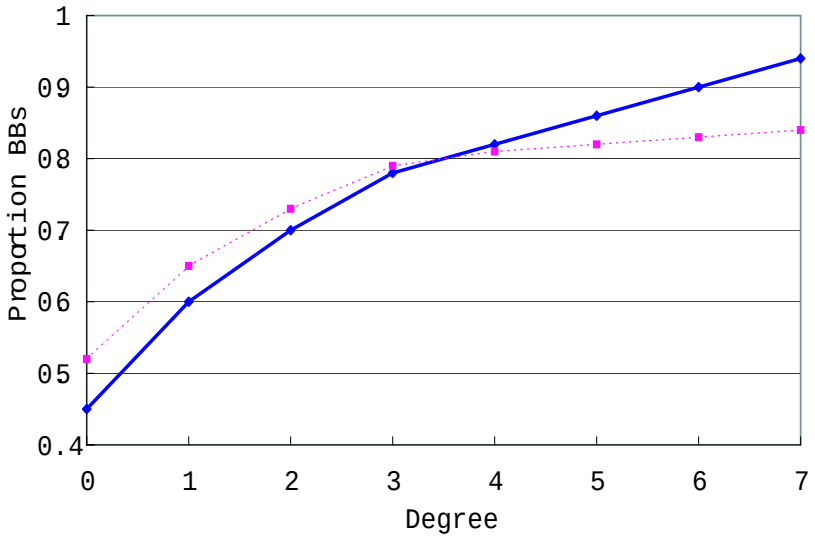

Fig. 20 Number of neighbors

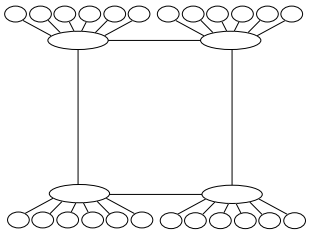

Fig. 21 Hierarchical1

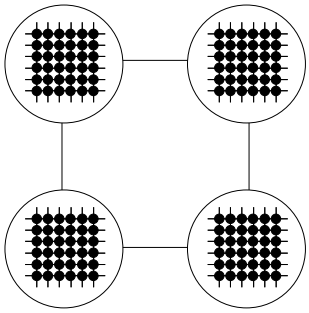

Fig. 22 Hierarchical2

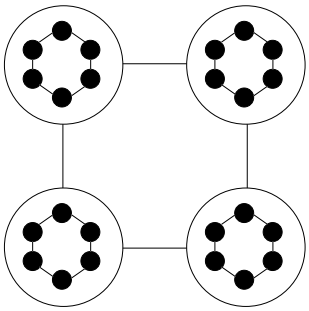

Fig. 23 Hierarchical3