

# 近傍解の評価値の確率分布推定アルゴリズムに基づく 近傍探索法

重弘 裕二<sup>1</sup> 増田 達也<sup>1</sup>

**概要：**著者らはこれまで、大規模な組合せ最適化問題に対して効率良く解探索を行う方法について、特に単位近傍操作の反復からなる近傍操作を用いて探索を行う方法について考察している。具体的には、単位近傍操作の反復回数の異なる多数の近傍操作を用意し、その中で最も良い操作を選んで解に適用する。最も良い操作を選ぶために、近傍操作により得られる解の評価値の分布を推定する。評価値の分布を推定するために、近傍操作により得られる解の評価値が従うであろう分布の形状を仮定し、探索の過程において収集した情報から分布を特徴づける統計量を推定する。本稿では特に、平均まわりの3次モーメントが0でない分布の形状を仮定し、その分布を推定する方法について述べる。また、計算機実験により手法の有効性を評価する。

**キーワード：**組合せ最適化, 確率分布, 最大充足可能性問題

## A Neighborhood Search Method Based on the Estimation of the Probability Density Function of the Objective Function Value of the Neighborhood Solution

YUJI SHIGEHIRO<sup>1</sup> TATSUYA MASUDA<sup>1</sup>

**Abstract:** We have considered effective search methods for large scale combinatorial optimization problems, in which many different neighborhood operations, which consist of iterations of the unit neighborhood operation, are employed. To select the best operation to apply, the probability distribution of the objective function values of the neighborhood solutions is estimated. To estimate the probability distribution, we assume the shape of the probability density function and estimate the statistics of the function using data collected through the search process. In this paper, we assume a probability distribution with non-zero third-order central moment, and consider how to estimate the statistics of the distribution. The experimental results are also shown to evaluate the proposed method.

**Keywords:** combinatorial optimization, probability distribution, maximum satisfiability problem

### 1. はじめに

著者らはこれまで、近傍操作のみを用いて、大規模な組合せ最適化問題 [1], [2] に対して効率良く解探索を行う方法を、確率論的観点から検討してきた。基本となる発想は、広域の探索を行うために適した近傍操作や、局所的な探索を行うために適した近傍操作等、多数の近傍操作をあらか

じめ定義しておき、広大な解空間を効率良く探索するためにそれらを使い分けるといったものである。ただ、ここで問題となるのが、探索の過程において近傍操作を行うたびに、どうやって最も適した近傍操作を選択すべきかということである。

この問題に対して、まず、解探索を行うエージェントが近傍操作の効率を計測しながら、適応的に近傍操作を切り替えるという手法 [3] について考察した。フィルタ回路合成問題に対する計算機実験を行い、進化型計算手法に匹敵

<sup>1</sup> 大阪工業大学  
Osaka Institute of Technology

する探索能力を持つことを確認しているが、処理が複雑になるのを避けることはできず、実際にソフトウェアを実装するのは面倒であった。

その後さらに、近傍操作により得られる解の評価値に対して統計的なモデルを仮定することにより、どのような探索手法が構築できるのかについて考察した [4], [5], [6]. 具体的には

- (1) 近傍操作により得られる解の評価値の統計量を、探索の過程において得られる情報から推定し、
  - (2) 個々の近傍操作に対して、それを適用したときの解探索の進み具合を表す量の期待値を求め、
  - (3) それを最大にするような近傍操作を選択する
- という考え方に沿って解探索手法を構成した。さらに、その手法を大規模な巡回セールスマン問題ならびに最大充足可能性問題 [2] に適用し、通常の近傍探索法や遺伝的アルゴリズムとの比較を行った。その結果、多くの問題に対して、提案する手法により理論的に期待された通りの結果が得られることを確認した。しかし一方で、ある種の最大充足可能性問題に対しては、期待するような近傍操作の選択が行われていないと考えざるを得ない結果を得るに至った [6].

そこで本稿では、より正しく近傍操作の選択が行われるように、近傍操作により得られる解の評価値に対する統計モデルの改善を試みる。その統計モデルによる確率分布を推定するためのアルゴリズムを導き、解探索手法に組み込んで計算機実験を行うことにより有効性を評価する。

## 2. 準備

関数  $f(x)$  の原点まわりの  $n$  次モーメント

$$\int_{-\infty}^{\infty} x^n f(x) dx \quad (1)$$

を  $m_n[f(x)]$  で表す。確率変数  $X$  の確率密度関数 [7] の原点まわりの  $n$  次モーメントを同様に、 $X$  の原点まわりの  $n$  次モーメントと呼び  $m_n[X]$  で表す。  $m_1[X]$  は特に**平均**、**期待値**と呼ばれる。これを  $E[X]$  で表す。原点まわりのモーメントは期待値を用いて

$$m_n[X] = E[X^n] \quad (2)$$

と表すことができる。確率変数  $X, Y$  に対して

$$E[aX + bY] = aE[X] + bE[Y] \quad (3)$$

である。同様に確率変数  $X_i$  に対して

$$E\left[\sum_i X_i\right] = \sum_i E[X_i] \quad (4)$$

である。また、独立な確率変数  $X, Y$  に対しては

$$E[XY] = E[X]E[Y] \quad (5)$$

である。

関数  $f(x)$  の**重心まわりの  $n$  次モーメント**

$$\int_{-\infty}^{\infty} (x - m_1[f(x)])^n f(x) dx \quad (6)$$

を  $M_n[f(x)]$  で表す。確率変数  $X$  の確率密度関数の重心まわりの  $n$  次モーメントを同様に、 $X$  の**平均まわりの  $n$  次モーメント**と呼び  $M_n[X]$  で表す。平均まわりのモーメントは期待値を用いて

$$M_n[X] = E[(X - E[X])^n] \quad (7)$$

と表すことができる。確率変数  $X$  に対して

$$M_1[X] = E[X - E[X]] = E[X] - E[X] = 0 \quad (8)$$

$$\begin{aligned} M_n[aX] &= E[(aX - E[aX])^n] \\ &= E[(aX - aE[X])^n] = E[a^n(X - E[X])^n] \\ &= a^n E[(X - E[X])^n] = a^n M_n[X] \end{aligned} \quad (9)$$

である。確率変数  $X, Y$  に対して

$$\begin{aligned} XY &= \{(X - E[X]) + E[X]\} \{(Y - E[Y]) \\ &\quad + E[Y]\} \\ &= (X - E[X])(Y - E[Y]) + (X - E[X])E[Y] \\ &\quad + E[X](Y - E[Y]) + E[X]E[Y] \end{aligned} \quad (10)$$

であるから

$$\begin{aligned} E[XY] &= E[(X - E[X])(Y - E[Y])] \\ &\quad + E[X - E[X]]E[Y] + E[X]E[Y - E[Y]] \\ &\quad + E[X]E[Y] \\ &= E[(X - E[X])(Y - E[Y])] + M_1[X]E[Y] \\ &\quad + E[X]M_1[Y] + E[X]E[Y] \\ &= E[(X - E[X])(Y - E[Y])] + E[X]E[Y] \end{aligned} \quad (11)$$

である。また、独立な確率変数  $X, Y$  に対して

$$\begin{aligned} M_2[X + Y] &= E[(X + Y - E[X + Y])^2] \\ &= E\left[\{(X - E[X]) + (Y - E[Y])\}^2\right] \\ &= E[(X - E[X])^2] \\ &\quad + 2E[(X - E[X])E[(Y - E[Y])]] \\ &\quad + E[(Y - E[Y])^2] \\ &= M_2[X] + 2M_1[X]M_1[Y] + M_2[Y] \\ &= M_2[X] + M_2[Y] \end{aligned} \quad (12)$$

$$\begin{aligned} M_3[X + Y] &= E[(X + Y - E[X + Y])^3] \\ &= M_3[X] + 3M_2[X]M_1[Y] + 3M_1[X]M_2[Y] \\ &\quad + M_3[Y] \\ &= M_3[X] + M_3[Y] \end{aligned} \quad (13)$$

である。従って、互いに独立な確率変数  $X_i$  に対して

$$M_2\left[\sum_i X_i\right] = \sum_i M_2[X_i] \quad (14)$$

$$M_3\left[\sum_i X_i\right] = \sum_i M_3[X_i] \quad (15)$$

である。

平均まわりのモーメントを原点まわりのモーメントで表すと

$$\begin{aligned} M_2[X] &= E[(X - E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= m_2[X] - (E[X])^2 \end{aligned} \quad (16)$$

$$\begin{aligned} M_3[X] &= E[(X - E[X])^3] \\ &= m_3[X] - 3m_2[X]E[X] + 3E[X](E[X])^2 \\ &\quad - (E[X])^3 \\ &= m_3[X] - 3m_2[X]E[X] + 2(E[X])^3 \end{aligned} \quad (17)$$

であり、これより

$$m_2[X] = M_2[X] + (E[X])^2 \quad (18)$$

$$\begin{aligned} m_3[X] &= M_3[X] + 3m_2[X]E[X] - 2(E[X])^3 \\ &= M_3[X] + 3\{M_2[X] + (E[X])^2\}E[X] \\ &\quad - 2(E[X])^3 \\ &= M_3[X] + 3M_2[X]E[X] + (E[X])^3 \end{aligned} \quad (19)$$

である。

### 3. 問題の定式化

解の集合（実行可能領域） $F$  と、解から実数値への写像（目的関数） $f: F \rightarrow R$  が与えられたとき（ここで  $R$  は実数全体の集合を表す）、 $F$  に含まれる解  $s \in F$  の中で、 $f$  により対応付けられる値  $f(s)$  を最小化するようなものを探索するという問題（最適化問題）について考える。 $f(s)$  を求めることを**評価**と呼び、 $f(s)$  を  $s$  の**評価値**と呼ぶ。ここでは特に、解が離散値を含む配列、順列、組合せ、グラフ、ネットワークのようなデータ構造、もしくはそれらの合成により表現されるような問題（組合せ最適化問題）について考える。なお、解の探索に要する時間計算量は、解の評価回数を基準に判断されることが多く、ここでもそれに従う。

**近傍操作**とは、ある解  $s$  を元にして別の解  $s'$  を生成する操作である。ここでは具体的に、**あらかじめ定められた手続きに従って**  $s$  を表現するデータ構造の構成要素の一部に変更を加え、その変更結果が表現する解を  $s'$  と考える。また、 $s$  に近傍操作を適用して生成された  $s'$  は、 $s$  の**近傍解**と呼ばれる。例えば、解が順列で表現されている場合、定められた確率に従って順列中から2つの要素を選び、それらを入れ替えるという近傍操作が考えられる。このように、以下では、ある解に近傍操作を適用すると、**近傍操作の手続きによって定まる確率に従って**、さまざまな近傍解が得られるものとする。

解  $s$  に近傍操作  $N$  を適用した結果、すなわち  $s$  を表現するデータ構造の構成要素に対して  $N$  の手続きにより変更を加えた結果、より評価値の小さな近傍解  $s'$  が得られ

たとする。探索の目標が解の評価値の最小化であることを考えると、 $N$  が  $s$  に変更を加えることによって  $s$  をより目標に近づけた、すなわち  $s$  を改善したと考えることができる。この改善の量は評価値の差  $f(s) - f(s')$  で表されるが、これを**近傍解改善量**と呼び、 $X(s, N)$  で表す。 $(f(s'))$  の方が  $f(s)$  より小さくない場合は値が非正となるが、この場合も  $f(s) - f(s')$  を近傍解改善量  $X(s, N)$  とする。）解  $s$  に近傍操作  $N$  を適用して得られる近傍解は確率的に定まるので、 $X(s, N)$  の値も確率的に定まる。すなわち  $X(s, N)$  は確率変数である。

以下では、未知の近傍解改善量の確率分布を推定するために、次の仮定を設ける。

**仮定 1** 任意の解を  $s$  とし、その近傍解（の一つ）を  $s'$  とする。 $s'$  の近傍解（の一つ）を  $s''$  とし、さらに同様に、 $s''$  の近傍解を  $s'''$ 、 $s'''$  の近傍解を  $s''''$ 、 $\dots$  とする。 $s$  に近傍操作  $N$  を適用したときの近傍解改善量  $X(s, N)$  の統計量（例えば平均や分散等）と、 $s'$  に  $N$  を適用したときの近傍解改善量  $X(s', N)$  の統計量は、（ほぼ）等しい。このとき帰納的に、 $X(s', N)$ 、 $X(s'', N)$ 、 $X(s''', N)$ 、 $X(s''', N)$ 、 $\dots$  の統計量と  $X(s, N)$  の統計量は（ほぼ）等しい。

また、解に関して可能な手続きを以下に限定する。

**ランダム解を生成** 解を表現するデータ構造の構成要素をランダムに設定することにより、新たな解を生成する。

**近傍操作を適用** 与えられた解に近傍操作を適用し、新たな解を生成する。

**評価** 与えられた解を評価し、評価値を得る。

この前提の中では、解の探索とは、これら3つの手続きを順序良く実行することである。

## 4. 提案手法

### 4.1 概要

本研究の目的は、前述した前提の下、定められた評価回数の中で、より評価値の小さな解を探索する方法を導き出すことである。その最初の段階として、本稿では図1に示すように、局所探索法（もしくは温度  $\rightarrow 0$  とした Simulated Annealing 法）[1], [2] の枠組みに沿って考察を進めるものとする。例えば複数の解を保持する [8] こと等に関する考察は今後の課題としたい。

探索の過程において、その時点までに評価した解の中で評価値が最も小さなものを**暫定最良解**と呼ぶことにする。局所探索法とは、最初にランダムに解を1つ生成し（図1の2行目）、以後、暫定最良解（同図の  $s_i$ ）に対して近傍操作を適用する（同図5行目）ことを繰り返すという手法である。もし近傍操作により得られた解（同図の  $s'_i$ ）の方が評価値が小さければ、その解が新たな暫定最良解となり（同図7行目）、そうでない場合は暫定最良解は変化しない（同図8行目）。

以下では局所探索法に倣って

考察する近傍探索法 ( $n$  は総評価回数)

- 1: 情報を蓄積するための変数を初期化
- 2:  $s_1 \leftarrow$  (ランダム解を生成);  $v_1 \leftarrow$  ( $s_1$  を評価)
- 3:  $i = 1, 2, \dots, n-1$  に対して
- 4: 近傍操作を選択
- 5:  $s'_i \leftarrow$  ( $s_i$  に近傍操作を適用);  $v'_i \leftarrow$  ( $s'_i$  を評価)
- 6: 近傍解改善量に関する情報を蓄積
- 7:  $v'_i \leq v_i$  のとき  $\{ s_{i+1} = s'_i; v_{i+1} = v'_i \}$
- 8: そうでないとき  $\{ s_{i+1} = s_i; v_{i+1} = v_i \}$
- 9:  $s_n$  を探索結果とする

図 1 本稿で考察する近傍探索法

Fig. 1 Neighborhood search method considered in this paper.

- ランダム解を生成するのは最初に 1 度だけとする
- 既に評価を行った解に対しては、暫定最良解を除き、近傍操作を適用しない

ものとする。すると、検討の余地があるのは、連続して近傍操作を適用する（近傍操作を適用した結果得られた解に対して、評価を行わずに、さらに近傍操作を適用する）ことであろう。

そこで本稿では、単位となる近傍操作が与えられるものとし、その近傍操作を複数回連続して繰り返す操作全体を近傍操作と考え、局所探索法の枠組みに沿って解の探索を行う方法について考察する [4], [5], [6]。以後、単位となる近傍操作を単位近傍操作と呼び、単位近傍操作を  $\alpha$  回繰り返すという操作を  $\alpha$ -近傍操作と呼ぶ。また、単に近傍操作と記述した場合は  $\alpha$ -近傍操作 ( $\alpha = 1, 2, \dots$ ) を指すものとする。

提案する近傍探索法では、局所探索法において近傍操作を適用するとき、毎回、 $\alpha$ -近傍操作 ( $\alpha = 1, 2, \dots$ ) の中から異なる近傍操作を選んで適用できるものとし（図 1 の 4, 5 行目）、そのために必要となる情報を近傍操作ごとに収集するものとする（同図 6 行目）。どのような情報をどうやって集めるかについては 4.4, 5.2 で述べる。その情報から 5.2 で述べるようにして、まず単位近傍操作の近傍解改善量の統計量を推定し、さらに 5.1 で述べるようにして  $\alpha$ -近傍操作の近傍解改善量の統計量を推定する。次に、 $\alpha$ -近傍操作の近傍解改善量の確率密度関数が 4.3 で述べる形状であると仮定し、5.3 で述べるようにして、その確率密度関数のパラメータを推定する。これにより  $\alpha$ -近傍操作の近傍解改善量の確率密度関数が定まるので、その結果を用いて、4.2 で述べるようにして近傍操作を選択する。

## 4.2 近傍操作の選択

図 1 の手順では、4~8 行目を繰り返し実行する。その  $i$  回目の実行を開始する時点の暫定最良解は  $s_i$  であり、その評価値は  $v_i$  である。また  $s'_i$  は  $s_i$  の近傍解であり、その評価値は  $v'_i$  である。 $v_i$  と  $v_{i+1}$  の差は、 $v_i$  と  $v'_i$  を用いて

$$v_i - v_{i+1} = \begin{cases} v_i - v'_i & (v'_i \leq v_i) \\ 0 & (v'_i > v_i) \end{cases} \quad (20)$$

と表すことができるが、これは  $i$  回目の実行により暫定最良解の評価値が改善された量を表している。これを最良解改善量と呼ぶことにする。

$v'_i$  の値は確率的に定まるので、最良解改善量も確率的に定まる。すなわち最良解改善量は確率変数である。もし近傍解改善量の確率分布を知ることができれば、最良解改善量の確率分布を求めることができる。その期待値は、近傍操作により暫定最良解の評価値がどの程度改善されると期待できるのかを表していると言える。そこで、 $\alpha$ -近傍操作 ( $\alpha = 1, 2, \dots$ ) が選択された場合の最良解改善量の期待値を（候補となる全ての  $\alpha = 1, 2, \dots$  に対して）求めて、その値が最も大きくなるような  $\alpha$ -近傍操作を選択するものとする。これにより、最良解改善量の期待値を最大化するという意味で最も良い近傍操作を選択することができる。

## 4.3 近傍解改善量の確率分布

実際の問題において、近傍解改善量の確率分布を知るとは難しい。そこで、近傍解改善量の確率密度関数の形状を仮定し、その統計量については探索の過程において収集した情報に基づき推定するものとする。

これまで、近傍解改善量の確率分布として正規分布を仮定することにより、多くの問題において良い結果を得ることができたが、ある種の最大充足可能性問題においては期待する結果を得ることができなかった [6]。そこで本稿では確率分布の形状として、ガウス関数の右半分と左半分

$$g_R(x) = \begin{cases} e^{-x^2} & (x > 0) \\ 0.5 & (x = 0) \\ 0 & (x < 0) \end{cases} \quad (21)$$

$$g_L(x) = \begin{cases} 0 & (x > 0) \\ 0.5 & (x = 0) \\ e^{-x^2} & (x < 0) \end{cases} \quad (22)$$

を組み合わせた関数

$$p_g(x) = K \left\{ \frac{a_2}{a_1} g_R \left( \frac{x-b}{a_1} \right) + \frac{a_1}{a_2} g_L \left( \frac{x-b}{a_2} \right) \right\}$$

ただし  $K = \frac{2}{(a_1 + a_2)\sqrt{\pi}}$ ,  $a_1 > 0$ ,  $a_2 > 0$  (23)

を考える。

**仮定 2**  $\alpha$ -近傍操作 ( $\alpha = 1, 2, \dots$ ) を適用したときの近傍解改善量の確率密度関数は (23) 式で与えられる。

## 4.4 近傍解改善量に関する情報の蓄積

近傍解改善量の確率密度関数として仮定する (23) 式は 3 つのパラメータ  $a_1$ ,  $a_2$ ,  $b$  を持つ。これらを決めるために、近傍解改善量の観測値 [7] に関する情報を収集する。なお

図 1 の手順で,  $i$  回目の実行において解  $s_i$  に近傍操作を適用して得られる近傍解改善量の観測値は  $y_i = v_i - v'_i$  である。

通常, 4~8 行目を繰り返す回数は多いため, 全ての  $y_i$  の値を保持しておくことは現実的ではない。ここでは (23) 式が 3 つのパラメータを持つことを考慮し,  $y_i, y_i^2, y_i^3$  の総和を求めるものとする。具体的には, 図 1 の 1 行目で

$$(z_k)_0 \leftarrow 0 \quad (k = 1, 2, 3) \quad (24)$$

とし, 6 行目で

$$y_i \leftarrow v_i - v'_i \quad (25)$$

$$(z_k)_i \leftarrow r \cdot (z_k)_{i-1} + y_i^k \quad (k = 1, 2, 3) \quad (26)$$

のように積算する。なおここでは, 探索の過程とともに近傍解改善量の統計量が徐々に変化していく可能性を考え, パラメータ  $r$  ( $0 < r \leq 1$ ) を用いて, 新しく  $y_i$  の値が得られるたびに, それ以前の値の影響を少しずつ減少させるものとしている。これにより,  $n'$  回目の実行が終わった時点で  $(z_k)_{n'}$  に  $y_i^k$  の総和 (重み付けを行って求めた総和)

$$(z_k)_{n'} = \sum_{i=1}^{n'} r^{n'-i} y_i^k \quad (k = 1, 2, 3) \quad (27)$$

の値を求めることができる。

## 5. 近傍解改善量の確率分布推定アルゴリズム

3 章では, 本稿でどのような問題を想定しているのかについて述べ, 4 章では, そのような問題に対してどのような方針で取り組むのかについて説明した。以上に基づき, 本章では近傍解改善量の確率分布の統計量, ならびに確率密度関数の推定法を導く。

### 5.1 近傍解改善量の統計量

図 1 の手順で  $i$  回目の実行において, 4 行目で  $\alpha_i$ -近傍操作が選択されたとする。すると, 5 行目で近傍操作が適用された際には  $\alpha_i$  回の単位近傍操作が繰り返されている。その  $j$  回目の単位近傍操作を適用したときに得られた近傍解を  $(s_i)_j$  とし, そのときの近傍解改善量を  $x_{ij} = f((s_i)_{j-1}) - f((s_i)_j)$  で表す。ここで,  $(s_i)_0 = s_i$  であり  $(s_i)_{\alpha_i} = s'_i$  である。また,  $\alpha_i$ -近傍操作 (全体) が適用されたときの近傍解改善量を 4.4 と同じく  $y_i = v_i - v'_i = f(s_i) - f(s'_i)$  で表す。このとき

$$\begin{aligned} y_i &= f(s_i) - f(s'_i) = \sum_{j=1}^{\alpha_i} f((s_i)_{j-1}) - f((s_i)_j) \\ &= \sum_{j=1}^{\alpha_i} x_{ij} \end{aligned} \quad (28)$$

である。

$x_{ij}, y_i$  の確率分布について考えるため, それらの確率

変数を  $X_{ij}, Y_i$  ( $i = 1, 2, \dots, n-1, j = 1, 2, \dots, \alpha_i$ ) で表す\*1。異なる  $X_{ij}, Y_i$  は互いに独立である。また (28) 式より

$$Y_i = \sum_{j=1}^{\alpha_i} X_{ij} \quad (29)$$

である。

仮定 1 より  $X_{ij}$  の期待値, 平均まわりの 2, 3 次モーメントは等しいとして, それらを  $\bar{E} = E[X_{ij}], \bar{M}_2 = M_2[X_{ij}], \bar{M}_3 = M_3[X_{ij}]$  で表す。  $Y_i$  の期待値は

$$\begin{aligned} E[Y_i] &= E\left[\sum_{j=1}^{\alpha_i} X_{ij}\right] = \sum_{j=1}^{\alpha_i} E[X_{ij}] = \sum_{j=1}^{\alpha_i} \bar{E} \\ &= \alpha_i \bar{E} \end{aligned} \quad (30)$$

であり, 同様に

$$M_2[Y_i] = \alpha_i \bar{M}_2 \quad (31)$$

$$M_3[Y_i] = \alpha_i \bar{M}_3 \quad (32)$$

である。このように, 単位近傍操作の近傍解改善量の統計量  $\bar{E}, \bar{M}_2, \bar{M}_3$  から,  $\alpha_i$ -近傍操作の近傍解改善量の統計量  $E[Y_i], M_2[Y_i], M_3[Y_i]$  を求めることができる。

さらに, 5.2 で必要となるので  $E[Y_i^2]$  等を求めると

$$\begin{aligned} E[Y_i^2] &= m_2[Y_i] = M_2[Y_i] + (E[Y_i])^2 \\ &= \alpha_i \bar{M}_2 + \alpha_i^2 \bar{E}^2 \end{aligned} \quad (33)$$

$$\begin{aligned} E[Y_i^3] &= m_3[Y_i] \\ &= M_3[Y_i] + 3M_2[Y_i]E[Y_i] + (E[Y_i])^3 \\ &= \alpha_i \bar{M}_3 + 3\alpha_i^2 \bar{M}_2 \bar{E} + \alpha_i^3 \bar{E}^3 \end{aligned} \quad (34)$$

$$\begin{aligned} &E[(Y_i^2 - E[Y_i^2])(Y_j - E[Y_j])] \\ &\quad (i \neq j \text{ のとき}) \\ &= E[Y_i^2 - E[Y_i^2]] E[Y_j - E[Y_j]] \\ &= M_1[Y_i^2] M_1[Y_j] = 0 \\ &\quad (i = j \text{ のとき}) \\ &= E[Y_i^2 Y_i] - E[Y_i^2] E[Y_i] - E[Y_i^2] E[Y_i] \\ &\quad + E[Y_i^2] E[Y_i] \\ &= E[Y_i^3] - E[Y_i^2] E[Y_i] \\ &= \alpha_i \bar{M}_3 + 3\alpha_i^2 \bar{M}_2 \bar{E} + \alpha_i^3 \bar{E}^3 \\ &\quad - (\alpha_i \bar{M}_2 + \alpha_i^2 \bar{E}^2) \alpha_i \bar{E} \\ &= \alpha_i \bar{M}_3 + 2\alpha_i^2 \bar{M}_2 \bar{E} \end{aligned} \quad (35)$$

となる。

### 5.2 蓄積した情報の統計量

次は 4.4 で述べた  $y_i^k$  の総和 (重み付けを行って求めた総和) について考える。図 1 の手順で  $n'$  回目の実行が終

\*1 確率変数  $X_{ij}, Y_i$  は確率的に様々な値を取り得る。その実際の値 (確率変数の観測値) が  $x_{ij}, y_i$  である。

わった時点で求めた  $(z_k)_{n'}$  の確率変数を  $Z_k$  で表すことにすると, (27) 式より

$$Z_k = \sum_{i=1}^{n'} r^{n'-i} Y_i^k \quad (k = 1, 2, 3) \quad (36)$$

である. 以下では

$$N_{kl} = \sum_{i=1}^{n'} r^{k(n'-i)} \alpha_i^l \quad (37)$$

として,  $E[Z_1]$  等を求めると

$$\begin{aligned} E[Z_1] &= E \left[ \sum_{i=1}^{n'} r^{n'-i} Y_i \right] = \sum_{i=1}^{n'} r^{n'-i} E[Y_i] \\ &= \sum_{i=1}^{n'} r^{n'-i} \alpha_i \bar{E} = \bar{E} \sum_{i=1}^{n'} r^{n'-i} \alpha_i = N_{11} \bar{E} \end{aligned} \quad (38)$$

$$\begin{aligned} M_2[Z_1] &= M_2 \left[ \sum_{i=1}^{n'} r^{n'-i} Y_i \right] = \sum_{i=1}^{n'} M_2 \left[ r^{n'-i} Y_i \right] \\ &= \sum_{i=1}^{n'} r^{2(n'-i)} M_2[Y_i] = \sum_{i=1}^{n'} r^{2(n'-i)} \alpha_i \bar{M}_2 \\ &= \bar{M}_2 \sum_{i=1}^{n'} r^{2(n'-i)} \alpha_i = N_{21} \bar{M}_2 \end{aligned} \quad (39)$$

$$M_3[Z_1] = \bar{M}_3 \sum_{i=1}^{n'} r^{3(n'-i)} \alpha_i = N_{31} \bar{M}_3 \quad (40)$$

$$\begin{aligned} E[Z_1^2] &= m_2[Z_1] = M_2[Z_1] + (E[Z_1])^2 \\ &= N_{21} \bar{M}_2 + N_{11}^2 \bar{E}^2 \end{aligned} \quad (41)$$

$$\begin{aligned} E[Z_2] &= E \left[ \sum_{i=1}^{n'} r^{n'-i} Y_i^2 \right] = \sum_{i=1}^{n'} r^{n'-i} E[Y_i^2] \\ &= \sum_{i=1}^{n'} r^{n'-i} (\alpha_i \bar{M}_2 + \alpha_i^2 \bar{E}^2) \\ &= \bar{M}_2 \sum_{i=1}^{n'} r^{n'-i} \alpha_i + \bar{E}^2 \sum_{i=1}^{n'} r^{n'-i} \alpha_i^2 \\ &= N_{11} \bar{M}_2 + N_{12} \bar{E}^2 \end{aligned} \quad (42)$$

$$\begin{aligned} E[Z_1^3] &= m_3[Z_1] \\ &= M_3[Z_1] + 3 M_2[Z_1] E[Z_1] + (E[Z_1])^3 \\ &= N_{31} \bar{M}_3 + 3 N_{21} N_{11} \bar{M}_2 \bar{E} + N_{11}^3 \bar{E}^3 \end{aligned} \quad (43)$$

$$\begin{aligned} E[Z_3] &= E \left[ \sum_{i=1}^{n'} r^{n'-i} Y_i^3 \right] = \sum_{i=1}^{n'} r^{n'-i} E[Y_i^3] \\ &= \sum_{i=1}^{n'} r^{n'-i} (\alpha_i \bar{M}_3 + 3 \alpha_i^2 \bar{M}_2 \bar{E} + \alpha_i^3 \bar{E}^3) \\ &= \bar{M}_3 \sum_{i=1}^{n'} r^{n'-i} \alpha_i + 3 \bar{M}_2 \bar{E} \sum_{i=1}^{n'} r^{n'-i} \alpha_i^2 \\ &\quad + \bar{E}^3 \sum_{i=1}^{n'} r^{n'-i} \alpha_i^3 \\ &= N_{11} \bar{M}_3 + 3 N_{12} \bar{M}_2 \bar{E} + N_{13} \bar{E}^3 \end{aligned} \quad (44)$$

となる. また,  $E[Z_2 Z_1]$  を求めると

$$\begin{aligned} E[Z_2 Z_1] &= E \left[ \left( \sum_{i=1}^{n'} r^{n'-i} Y_i^2 \right) \left( \sum_{j=1}^{n'} r^{n'-j} Y_j \right) \right] \\ &= E \left[ \sum_{i=1}^{n'} \sum_{j=1}^{n'} r^{n'-i} r^{n'-j} Y_i^2 Y_j \right] \\ &= \sum_{i=1}^{n'} \sum_{j=1}^{n'} r^{n'-i} r^{n'-j} E[Y_i^2 Y_j] \\ &= \sum_{i=1}^{n'} \sum_{j=1}^{n'} r^{n'-i} r^{n'-j} E[(Y_i^2 - E[Y_i^2])(Y_j - E[Y_j])] \\ &\quad + \sum_{i=1}^{n'} \sum_{j=1}^{n'} r^{n'-i} r^{n'-j} E[Y_i^2] E[Y_j] \end{aligned} \quad (45)$$

となるが, 第 1 項は

$$\begin{aligned} &\sum_{i=1}^{n'} \sum_{j=1}^{n'} r^{n'-i} r^{n'-j} E[(Y_i^2 - E[Y_i^2])(Y_j - E[Y_j])] \\ &= \sum_{i=1}^{n'} r^{n'-i} r^{n'-i} (\alpha_i \bar{M}_3 + 2 \alpha_i^2 \bar{M}_2 \bar{E}) \\ &= \bar{M}_3 \sum_{i=1}^{n'} r^{2(n'-i)} \alpha_i + 2 \bar{M}_2 \bar{E} \sum_{i=1}^{n'} r^{2(n'-i)} \alpha_i^2 \\ &= N_{21} \bar{M}_3 + 2 N_{22} \bar{M}_2 \bar{E} \end{aligned} \quad (46)$$

となり, 第 2 項は

$$\begin{aligned} &\sum_{i=1}^{n'} \sum_{j=1}^{n'} r^{n'-i} r^{n'-j} E[Y_i^2] E[Y_j] \\ &= \left( \sum_{i=1}^{n'} r^{n'-i} E[Y_i^2] \right) \left( \sum_{j=1}^{n'} r^{n'-j} E[Y_j] \right) \\ &= \left( \sum_{i=1}^{n'} r^{n'-i} (\alpha_i \bar{M}_2 + \alpha_i^2 \bar{E}^2) \right) \\ &\quad \times \left( \sum_{j=1}^{n'} r^{n'-j} \alpha_j \bar{E} \right) \\ &= \left( \bar{M}_2 \sum_{i=1}^{n'} r^{n'-i} \alpha_i + \bar{E}^2 \sum_{i=1}^{n'} r^{n'-i} \alpha_i^2 \right) \\ &\quad \times \left( \bar{E} \sum_{j=1}^{n'} r^{n'-j} \alpha_j \right) \\ &= (N_{11} \bar{M}_2 + N_{12} \bar{E}^2) N_{11} \bar{E} \\ &= N_{11}^2 \bar{M}_2 \bar{E} + N_{12} N_{11} \bar{E}^3 \end{aligned} \quad (47)$$

となるので

$$\begin{aligned} E[Z_2 Z_1] &= N_{21} \bar{M}_3 + 2 N_{22} \bar{M}_2 \bar{E} + N_{11}^2 \bar{M}_2 \bar{E} \\ &\quad + N_{12} N_{11} \bar{E}^3 \\ &= N_{21} \bar{M}_3 + (2 N_{22} + N_{11}^2) \bar{M}_2 \bar{E} + N_{12} N_{11} \bar{E}^3 \end{aligned} \quad (48)$$

となる.

$Z_k$  の観測値は  $(z_k)_{n'}$  であるので、 $(z_1)_{n'}$ ,  $\{(z_1)_{n'}\}^2$ ,  $\{(z_1)_{n'}\}^3$ ,  $(z_2)_{n'}$ ,  $(z_3)_{n'}$ ,  $\{(z_2)_{n'}\}\{(z_1)_{n'}\}$  をそれぞれ  $E[Z_1]$ ,  $E[Z_1^2]$ ,  $E[Z_1^3]$ ,  $E[Z_2]$ ,  $E[Z_3]$ ,  $E[Z_2Z_1]$  の推定値とすることができる。また  $N_{kl}$  については  $(z_k)_{n'}$  と同様にして求めることができる。具体的には、図 1 の 1 行目で

$$(N_{kl})_0 \leftarrow 0 \quad (49)$$

とし、4 行目で  $\alpha_i$ -近傍操作が選択されたとして、その  $\alpha_i$  の値を用いて 6 行目で

$$(N_{kl})_i \leftarrow r^k \cdot (N_{kl})_{i-1} + \alpha_i^l \quad (50)$$

のように積算すれば、 $N_{kl}$  の値を求めることができる。

これらの値が決定すれば、(38) 式は  $\bar{E}$  に関する 1 次方程式の、(41), (42) 式は  $\bar{M}_2$ ,  $(\bar{E}^2)$  に関する 2 元 1 次連立方程式の、(43), (44), (48) 式は  $\bar{M}_3$ ,  $(\bar{M}_2\bar{E})$ ,  $(\bar{E}^3)$  に関する 3 元 1 次連立方程式の形になっており、それぞれを解くことにより、順に  $\bar{E}$ ,  $\bar{M}_2$ ,  $\bar{M}_3$  の不変推定量を求めることができる。

### 5.3 仮定した確率密度関数の統計量

今度は、仮定 2 より近傍解改善量の確率密度関数が (23) 式の  $p_g(x)$  で与えられたとして、その統計量を求める方法について考える。図 1 の手順で  $i$  回目の実行において、4 行目で  $\alpha_i$ -近傍操作 ( $\alpha_i = 1, 2, \dots$ ) が選ばれたとする。その近傍解改善量  $Y_i$  の確率密度関数を  $p_g(x)$  とする。また  $E[Y_i]$ ,  $M_2[Y_i]$ ,  $M_3[Y_i]$  の値が 5.2, 5.1 のようにして求めたとする。このとき  $p_g(x)$  の 3 つのパラメータ  $a_1$ ,  $a_2$ ,  $b$  を求める方法について考える。

まず (21), (22) 式の  $g_R(x)$ ,  $g_L(x)$  の原点まわりの 0, 1, 2, 3 次モーメントと求めるために、ガウス積分ならびに部分積分法を用いて計算すると

$$m_0[g_R(x)] = m_0[g_L(x)] = \frac{\sqrt{\pi}}{2} \quad (51)$$

$$m_1[g_R(x)] = -m_1[g_L(x)] = \frac{1}{2} \quad (52)$$

$$m_2[g_R(x)] = m_2[g_L(x)] = \frac{\sqrt{\pi}}{4} \quad (53)$$

$$m_3[g_R(x)] = -m_3[g_L(x)] = \frac{1}{2} \quad (54)$$

となる。この結果と置換積分法を用いて  $p_g(x)$  の原点まわりのモーメント、すなわち  $Y_i$  の原点まわりのモーメントを計算すると

$$m_0[Y_i] = m_0[p_g(x)] = 1 \quad (55)$$

$$m_1[Y_i] = E[Y_i] = m_1[p_g(x)] = b \quad (56)$$

$$m_2[Y_i] = m_2[p_g(x)] = \frac{a_1 a_2}{2} + b^2 \quad (57)$$

$$\begin{aligned} m_3[Y_i] &= m_3[p_g(x)] \\ &= \frac{a_1 a_2}{\sqrt{\pi}}(a_1 - a_2) + \frac{3}{2} a_1 a_2 b + b^3 \end{aligned} \quad (58)$$

となり、さらに平均まわりのモーメントを求めると

$$M_2[Y_i] = m_2[Y_i] - (E[Y_i])^2 = \frac{a_1 a_2}{2} \quad (59)$$

$$\begin{aligned} M_3[Y_i] &= m_3[Y_i] - 3 m_2[Y_i] E[Y_i] + 2(E[Y_i])^3 \\ &= \frac{a_1 a_2}{\sqrt{\pi}}(a_1 - a_2) \end{aligned} \quad (60)$$

となる。 $b$  の値は (56) 式より直ちに求まる。また (59), (60) 式より  $a_1$ ,  $a_2$  に関する 2 次方程式が導かれるので、その方程式を (解の公式を用いて) 解くことにより、 $a_1$ ,  $a_2$  の値を求めることができる。すなわち

$$A = \frac{\sqrt{\pi} M_3[Y_i]}{2 M_2[Y_i]}, \quad B = 2 M_2[Y_i] \quad (61)$$

として

$$a_1 = \frac{A + \sqrt{A^2 + 4B}}{2}, \quad a_2 = \frac{-A + \sqrt{A^2 + 4B}}{2} \quad (62)$$

である\*2。

## 6. 評価

本稿で提案する手法を Haskell[9] により実現し、最大充足可能性問題に適用した。その際、乱数生成や定積分の計算等に GSL[10] と FFI[11] を利用した。最大充足可能性問題の解は、全ての変数の真偽値を要素とするベクトルにより表現し、ランダムに 1 つの変数を選んで真偽値を反転する操作を単位近傍操作とした。4.4 の  $r$  の値は 0.99 とした。簡単のため、暫定最良解に適用する近傍操作は、1, 2, 4, 8, 16, 32, 64, 128, 256-近傍操作に限定した。それら全てに対して最良解改善量の期待値を求めた上で、その値が最も大きなものを選択して暫定最良解に適用した。探索開始時等、5.2 の  $\bar{E}$ ,  $\bar{M}_2$ ,  $\bar{M}_3$  の値を計算することができない場合は、256-近傍操作を暫定最良解に適用した。

なお、提案手法の処理をそのまま計算機で実行したところ、探索の初期において  $\bar{M}_3$  を正しく求めることができなかった。 $\bar{M}_3$  は  $Y_i$  の 3 乗に影響を受けるため、 $Y_i$  のサンプル数が少ないと、 $Y_i$  の観測値のばらつきに大きく影響を受けてしまうのが原因と思われる。そこで今回は、暫定最良解の評価値が 26,000 未満であれば  $\bar{M}_3$  を 0 として確率密度関数を求めるものとした。

適用対象とする最大充足可能性問題として、文献 [6] の P5 を用いた。これは文献 [6] の手法では期待する結果を得ることができなかった問題である。この問題はランダムに生成され、変数の数は 10,000、節の数は 40,000、節当たりのリテラルの数は 10~15 である。節の重みは -100~100 の 0 を除く整数であり、節を構成するリテラルの組み

\*2 解の公式からは 2 組の解が得られるが、(23) 式より、 $a_2 \leq 0$  となる解は無効である。

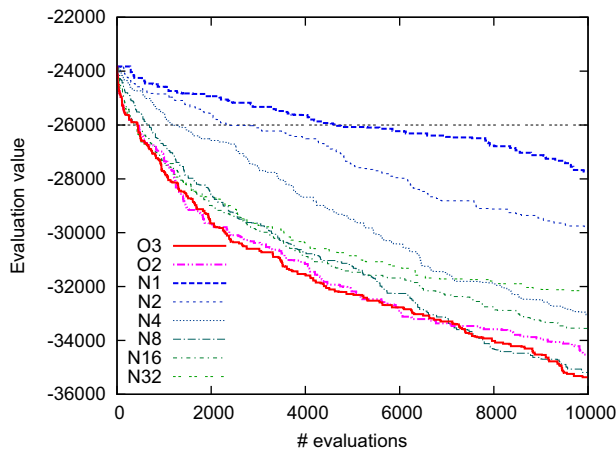


図 2 実験結果

Fig. 2 Experimental results.

合わせは全て異なる。

比較のために、前述の近傍操作のいずれか 1 つのみを用いる通常の局所探索法 (1 つの近傍操作しか選択することができない提案手法) と文献 [6] の手法を、提案手法と共に前述の最大充足可能性問題に適用した。文献 [6] の手法は、近傍改善量の確率分布として正規分布を仮定したもので、提案手法において  $\bar{M}_3$  を常に 0 とした場合に等しい。

結果を図 2 に示す。N1~N32 は 1~32-近傍操作を用いた局所探索法を表し、O2 は文献 [6] の手法 (2 次モーメントまで考慮)、O3 は提案手法 (3 次モーメントまで考慮) を表す。グラフの横軸と縦軸はそれぞれ、評価回数と暫定最良解の評価値を表す。なお、各手法の実装の都合上、最大充足可能性問題の評価値を正負反転して最小化問題として扱っており、その評価値をそのまま縦軸に示しているため、実際の評価値は、縦軸に示されている値を正負反転した値である。

提案手法 (O3) は通常の局所探索法 (N1) に対して明らかな優位性を示している。一方、8-近傍操作を用いた局所探索法 (N8) では、N1 よりもはるかに良い結果が得られている。このように、うまく近傍操作を選べば、単純な局所探索法でも良い結果を得ることができるとわかる。しかし、問題や最大評価回数が異なれば、適した近傍操作も異なる。(さらに、探索の進行に合わせて適した近傍操作が変化する。) 通常は、どの近傍操作が適しているのかわからないため、おそらく、このような問題に対して 8-近傍操作を用いた局所探索法が適用されることはないだろう。

グラフの傾きは、評価回数当たり平均してどの程度、暫定最良解の評価値が改善されているかを表しており、最良解改善量の期待値の大きさに対応している。O3 のグラフはムラはあるものの、特に探索の終盤にかけて、文献 [6] の手法 (O2) より傾きが大きくなっていると言えるだろう。これは、3 次モーメントまで考慮に入れることにより、最良解改善量の期待値の推定精度が向上したためと考えら

れる。

評価値が 33,000 程度において、O3 の傾きは N8 に及んでいない。想定している確率密度関数  $p_g(x)$  が実際の確率分布とは異なる以上、理想的な近傍操作選択の実現は難しいかもしれない。ただ、ここで取り上げた手法はすべて乱数に依存しているため、乱数系列が異なると結果 (すなわちグラフの傾き、さらには最適な近傍操作) も異なったものになる。より精密な評価を行うのであれば、評価の方法を再考する必要があるだろう。

## 7. まとめ

大規模な組合せ最適化問題に対し、近傍操作のみを用いて効率良く解探索を行う方法について、確率論的観点から考察した。また、最大充足可能性問題に対して、提案手法、文献 [6] の手法、局所探索法を適用して計算機実験を行い、提案手法が通常の局所探索法に対して明らかな優位性を示していること、ならびに、文献 [6] の手法を改善していることを確認した。

なお提案手法では、探索の初期において、平均まわりの 3 次モーメントをうまく求めることができなかった。統計量を正確に推定するためには多数のサンプルが必要となるが、探索の初期においては少数のサンプルしか得ることができないのが原因と思われる。これに対する対応法は今後の課題である。

## 参考文献

- [1] 久保幹雄：メタヒューリスティックス、離散構造とアルゴリズム IV, 近代科学社, pp. 171-230 (1995).
- [2] 柳浦睦憲, 茨木俊秀：組合せ最適化, 朝倉書店 (2001).
- [3] 重弘裕二, 増田達也：改善解探索速度に基づくエージェント探索法のフィルタ回路合成への適用, 電気学会論文誌 C, Vol. 130, No. 1, pp. 123-132 (2010).
- [4] 重弘裕二, 増田達也：近傍解の評価値の改善量の期待値に基づく近傍探索法, 情報処理学会研究報告, Vol. 2012-MPS-87, No. 17, pp. 1-7 (2012).
- [5] 重弘裕二, 増田達也：大規模組合せ最適化問題に対する複数の近傍を用いた探索手法, 電気学会電子・情報・システム部門大会講演論文集, pp. 1358-1363 (2012).
- [6] 重弘裕二, 増田達也：統計量推定に基づく組合せ最適化手法の最大充足可能性問題への適用, 電気学会電子・情報・システム部門大会講演論文集, pp. 1426-1431 (2013).
- [7] 森棟公夫, 照井伸彦, 中川 満, 西埜晴久, 黒住英司：統計学, 有斐閣 (2008).
- [8] 重弘裕二, 増田達也：近傍解コストの分布に基づく最適化手法の提案と評価, 情報処理学会研究報告, 2006-MPS-59, Vol. 2006, No. 56, pp. 81-84 (2006).
- [9] Jones, S. P.: *Haskell 98 Language and Libraries: The Revised Report*, Cambridge University Press (2003).
- [10] Gough, B.: *GNU Scientific Library Reference Manual*, Network Theory Ltd. (2009).
- [11] Chakravarty, M. M. T.: *The Haskell 98 Foreign Function Interface 1.0*, The University of New South Wales (online), available from <http://www.cse.unsw.edu.au/~chak/haskell/ffi/> (accessed 2012-01-28).