

# 離散 2 次計画問題の 分子計算アルゴリズム

齋藤 誠慈<sup>\*</sup>, 藤原 祐介<sup>\*\*</sup>  
谷田 純<sup>\*</sup>, 小倉 裕介<sup>\*</sup>  
石井 博昭<sup>\*</sup>

本論文では、離散型、すなわち 0, 1 の 2 整数値をとりうる目的関数の制約条件付き最適化問題に対する分子コンピューティングのための解法手順を紹介する。特に、最大クリーク問題から派生する計画問題について述べる。

最大クリーク問題は組み合わせ最適化問題の一つであり、この問題を多項式時間で解くアルゴリズムは未だ発見されていない。また、最大クリーク問題は種々の 2 次計画問題に定式化することが可能であり、2 次計画問題を考えることにより解の上・下界を与える様々な条件が得られ、多項式時間で解の上・下界を計算することができる。一方、1994年に Adleman [1] が有向ハミルトン経路問題に対する DNA 分子計算を報告して以来、近年では DNA 分子計算により分子反応の並列性、自律性を用いて様々な組み合わせ最適化問題に対するアプローチが行われている。本研究では 2 次 0-1 計画問題で表現した最大クリーク問題に対する分子計算を考えることで、解の上・下界の適用により計算量の軽減が可能な自律計算アルゴリズムの開発を目的とする。

## 1. 諸 定 義

グラフを  $G = (V, E)$  とし、頂点の集合  $V$  とエッジの集合  $E$  で表現する。グラフの頂点はそれぞれ  $1, 2, \dots, n$  と表され、頂点の集合を  $V = \{1, 2, \dots, n\}$

---

\* 大阪大学 大学院情報科学研究科 情報数理学専攻

\*\* 大阪大学工学部応用自然科学科

離散 2 次計画問題の分子計算アルゴリズム

とする。また 2 頂点  $i, j$  間に存在するエッジは  $(i, j)$  と表され、その集合を  $E = \{(i, j) \mid i, j \in V, i \neq j\}$  とする。また  $n$  頂点のグラフ  $G$  に対し  $n$  次正方の実対称行列である隣接行列  $A_G = (a_{ij})_{n \times n}$  を導入する、その各成分  $a_{ij}$  は  $(i, j) \in E$  のとき  $a_{ij} = 1$ 、それ以外は 0 である。

与えられたグラフの部分グラフが完全グラフとなるとき、その部分グラフの頂点の集合をクリークといい、大きさが最大のクリークを最大クリークという。最大クリーク問題とは最大クリークの大きさ (クリーク数)  $\omega(G)$  を求める問題である。本論文では連結グラフのみを扱う。

## 2. 最大クリーク問題と 2 次 0-1 計画問題

最大クリーク問題は様々な 2 次計画問題に定式化が可能である [2]。本研究では次に示す 2 次 0-1 計画問題への定式化を考える。

$$\text{Minimize } f(\mathbf{x}) = -\sum_{i=1}^n x_i + 2 \sum_{i>j, (i,j) \in E} x_i x_j \quad (1)$$

$$\text{Subject to } \mathbf{x} \in \{0, 1\}^n \quad (2)$$

このとき最適解は  $x_i^* = 1 \Leftrightarrow i \in C$  ( $C$ : 最大クリーク) であり、クリーク数  $\omega(G) = -f(\mathbf{x}^*) = \sum_{i=1}^n x_i$  となる。最適解においては目的関数の 2 次項が全てゼロになることがわかる。

## 3. クリーク数の上・下界

最大クリーク問題を 2 次計画問題に定式化することで、次に挙げる条件などからクリーク数の上・下界を求めることができる [3, 4]。上界を与える条件には次の 3 つなどがある。

$$\omega(G) \leq \frac{3 + \sqrt{9 - 8(n-m)}}{2} \quad (n: \text{頂点の数}, m: \text{エッジの数}) \quad (3)$$

$$\omega(G) \leq \rho(A_G) + 1 \quad (\rho(A_G): A_G \text{ のスペクトル半径}) \quad (4)$$

$$\omega(G) \leq N_{-1} + 1 \quad (N_{-1}: A_G \text{ の固有値のうち、}-1 \text{ 以下のものの数}) \quad (5)$$

下界を与える条件には次の2つなどがある。

$$\omega(G) \geq \frac{1}{1-\delta} \quad (\delta: A_G \text{ における成分1の密度}) \quad (6)$$

$$\omega(G) \geq \frac{\lambda_p}{s^2 - \lambda_p} + 1 \geq \frac{\lambda_p}{n - \lambda_p} + 1 \quad (7)$$

( $\lambda_p$ ; Perron固有値,  $s = \mathbf{e}^T \mathbf{x}_p$ ,  $\mathbf{x}_p$ : Perron固有ベクトル,  $\mathbf{e} = (1, 1, \dots, 1)^T$ )

乱数を用いて作成した様々な大きさ、密度のランダムグラフに上記の条件などを適用し、それぞれの場合の上・下界の計算を行った。結果を図1に示す。ただし、得られた種々の上界の最小値を上限候補とし、下界の最大値を下限候補とする。また解候補の数とは下限候補から上限候補の間に含まれる自然数の数であり、クリーク数はそのうちのどれかの数である。

図1より、いずれの場合も頂点の数  $n$  に対して上限候補はおおよそ  $n/2$  程度に

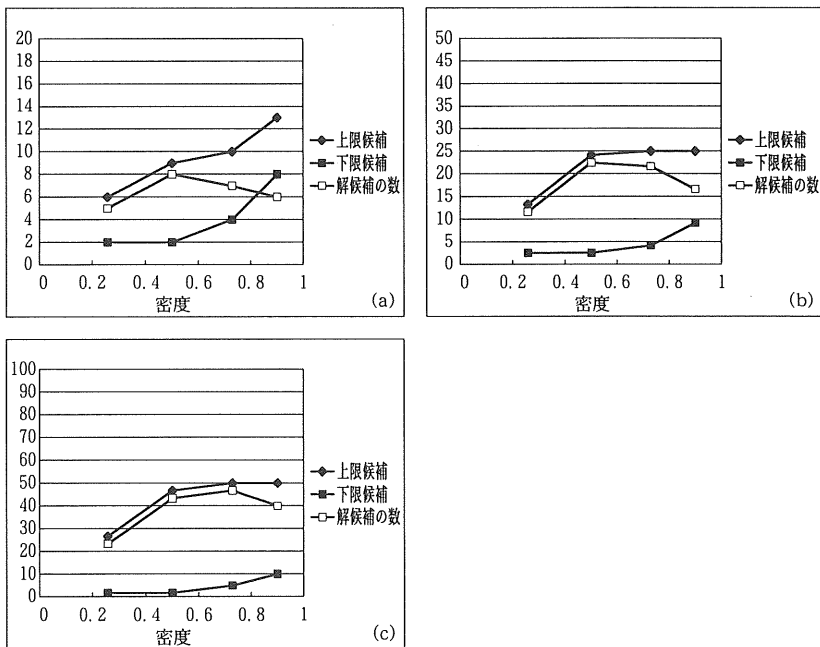


図1 グラフの密度と上・下界、解候補の数との関係  
(a) 頂点20 (b) 頂点50 (c) 頂点100

## 離散 2 次計画問題の分子計算アルゴリズム

抑えられていることがわかる。これはグラフの作成方法の都合上極端にエッジが局在した頂点がないためである。このようなグラフにおいては解候補の数は頂点の数  $n$  の約 2 ~ 4 割になっている。このことを考えると、次に述べる分子計算における計算を約 2 ~ 4 割に軽減することが可能である。

### 4. 分子計算アルゴリズム

本研究では式(1)の 2 次 0-1 計画問題の目的関数を DNA 分子に符号化することで、自律計算可能な分子計算方法を考案した。式(1)の目的関数は 1 次項と 2 次項からなっており、本計算法では 1 次項を符号化した DNA 配列と 2 次項を符号化した DNA 配列をそれぞれ別々に作成し、またそれぞれに異なる役割を持たせる。1 次項を符号化した配列（以下、1 次項配列）は解候補を表す役割を持ち、解候補の分類や上・下界の適用による計算量の軽減に用いられる。2 次項を符号化した配列（以下、2 次項配列）は解候補の評価を行う役割を持ち、不適な配列を自律的に除去するのに用いられる。アルゴリズムを以下に示す。

- Step 1
  - (1) 1 次項配列の作成
  - (2) 1 次項配列の分類
  - (3) 上・下界の適用による不適な配列の除去
- Step 2
  - (1) 2 次項配列の作成
  - (2) 1 次項配列と 2 次項配列の結合
  - (3) ヘアピン構造の形成による自律計算
  - (4) 解の評価

#### Step 1 (1) 1 次項配列の作成

1 次項配列の作成には次のような規則を用いて符号化を行う。

$$x_i = 0 \text{ のとき } \quad \begin{array}{|c|c|} \hline x_i & S_i \\ \hline \end{array}$$

$x_i=1$  のとき  $x_i$

ただし,  $x_i$  は変数  $x_i$  を表す配列であり,  $S_i$  は変数  $x_i$  が値 0 を持つことを表す配列である。

全ての変数に対して値 0 または 1 の割り当てを行い,  $x_1$  から  $x_n$  まで順に結合する。この操作によって各変数に 0 または 1 が割り当てられた  $2^n$  通りの 1 次項配列が作成される。概要を図 2 に示す。

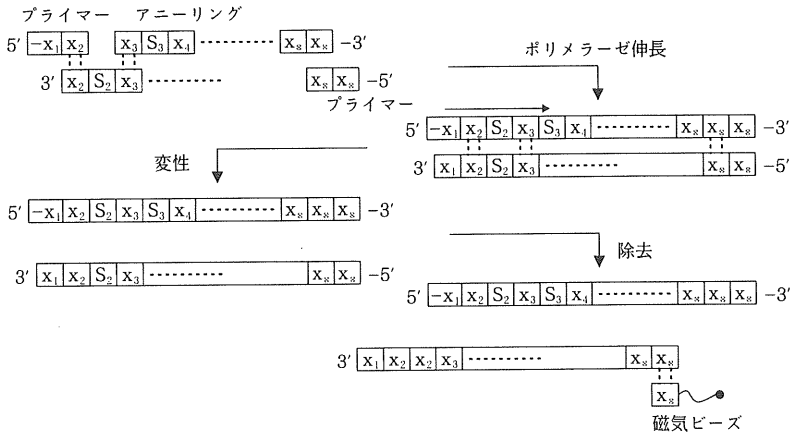


図 2 1 次項配列の作成方法

## (2) 1 次項配列の分類

以上のようにして作成した 1 次項配列は, 値 1 を持つ変数の数により長さが異なる。したがってゲル電気泳動を用いて 1 次項配列を長さにより分類することは, 値 1 を持つ変数の数により 1 次項配列を分類することを意味する。値 1 を持つ変数の数が  $k$  ( $k=0, 1, \dots, n$ ) 個の配列を試験管  $T_k$  に分類し,  $T_0$  から  $T_n$  までの  $n+1$  本の試験管を作成する。

## (3) 上・下界の適用による不適な配列の除去

計算により求めたクリーク数の上界を  $p$ , 下界を  $q$  とする。すると最適解において値 1 を持つ変数の数は  $q$  個以上  $p$  個以下であることが分かる。したがって作成した  $n+1$  本の試験管のうち, 最適解を符号化した配列を含む可能性が

離散 2 次計画問題の分子計算アルゴリズム

あるのは  $T_q \sim T_p$  のみであり、それ以外の試験管は不適であるので廃棄することができる。故に Step 2 で考えるべき試験管は  $T_q \sim T_p$  のみに限定される。このように上・下界を適用することで不適な配列を除去でき、Step 2 における計算量（操作回数）を軽減することができる。以上の概要を図 3 に示す。

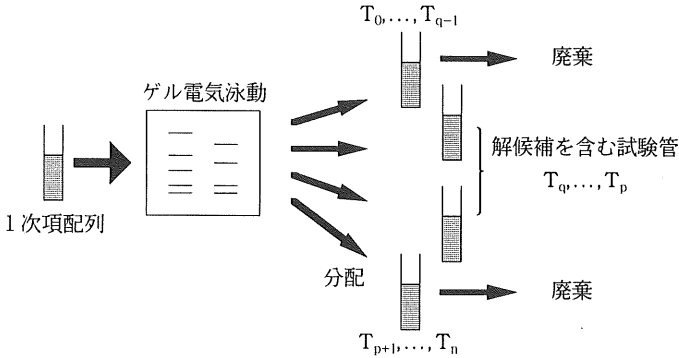


図 3 1 次項配列の分類と上・下界の適用による不適な配列の除去

Step 2 (1) 2 次項配列の作成

1 次項配列の作成には次のような規則を用いて符号化を行う。第  $t$  項  $x_i x_j$  に対して、

$$\begin{aligned}
 &x_i x_j = 0 \text{ のとき } \boxed{y_t} \\
 &x_i x_j = 1 \text{ のとき } \boxed{y_t} \quad \boxed{\bar{S}_i} \quad \boxed{\bar{S}_j}
 \end{aligned}$$

ただし、 $\boxed{y_t}$  は第  $t$  番目の項であることを示し、 $\boxed{\bar{S}_i}$  は変数  $x_i$  が値 1 を持つことを表す。また  $\boxed{\bar{S}_j}$  と 1 次項配列の作成の際に用いた  $\boxed{S_i}$  は互いに相補的である。

全ての 2 次項に対して値 0 または 1 の割り当てを行い、2 次項の初項から末項まで順に結合する。この操作によって各項に値 0 または 1 が割り当てられた 2 次項配列が作成される。作成の手順は 1 次項の場合と同様である。

(2) 1 次項配列と 2 次項配列の結合

Step 1 で作成した  $p-q+1$  本の試験管  $T_q \sim T_p$  に 2 次項配列を加え、酵素リ

ガーゼによるライゲーション反応を用いて1次項配列と2次項配列を結合する。この操作により1次項配列と2次項配列が連結した配列、つまり式(1)の目的関数を符号化した配列（以下、目的関数配列）が作成される。この操作の概要を図4に示す。

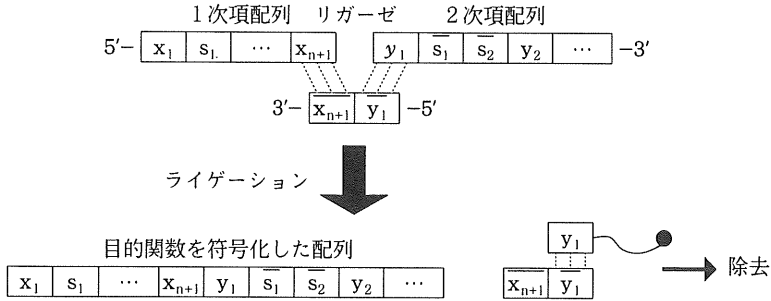


図4 ライゲーション反応を用いた1次項配列と2次項配列の結合

### (3) ヘアピン構造の形成による自律計算

前述(2)で作成した目的関数配列は1次項配列と2次項配列を単純に連結したものであり、1次項における変数  $x_i$ ,  $x_j$  と2次項における項  $x_i x_j$  との値が矛盾しているもの（例えば  $x_i=0$  に対し  $x_i x_j=1$  となっている配列など）も含まれる。このような不適な配列を除去するために、本研究では萩谷ら [5] が提案した分子のヘアピン構造形成による自律計算の手法を導入する。Step 1 (1) で作成した1次項配列において値0を表す配列  $\boxed{S_i}$  と、Step 2 (1) で作成した2次項配列において値1を表す配列  $\boxed{\bar{S}_i}$  が互いに相補的である考慮すると、1次項での変数の値と2次項での項の値が矛盾している場合は1本鎖DNA分子中に互いに相補的な部分が存在することが分かる。従って試験管中の溶液の濃度と温度を調節することにより分子内ハイブリダイゼーションが起こり、ヘアピン構造が形成される。試験管中に制限酵素を予め加えておくことにより、ヘアピン構造により2本鎖になった部分が特異的に切断される。これらの反応により、不適な配列が自律的に切断される。概要を図5に示す。

離散 2 次計画問題の分子計算アルゴリズム

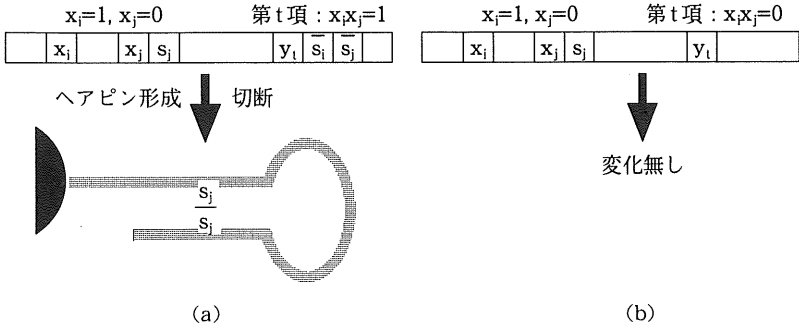


図5 ヘアピン構造の形成による自律計算  
(a) 値の割り当てが矛盾するとき (b) 矛盾しないとき

(4) 解の評価

今までの操作により、各試験管  $T_a \sim T_p$  には 1 次項における変数の値と 2 次項における項の値とが矛盾しない目的関数配列のみが含まれている。これらの試験管の中から最適解を符号化した配列を見つける必要がある。ここではゲル電気泳動法を用いて最適解を符号化した配列を検出する方法を提案する。まず各試験管中の目的関数配列の 1 次項部分と 2 次項部分とを切断する。次にこの 2 次項配列に対してゲル電気泳動法を用いて長さによる分離を行う。

ここで最大クリーク問題から導かれる 2 次 0-1 計画問題の最適解に着目すると、上述したように、最適解においてクリーク数は、

$$-f(\mathbf{x}^*) = \sum_{i=1}^n x_i^* \quad (8)$$

と表される。このことより、最適解を符号化する配列は次の 2 つの条件を満たす配列であることが分かる。

- (i) 2 次項配列が最短 (2 次項が全て 0 に対応)
- (ii) 1 次項配列が最長 (値 1 を持つ変数の数が最多)

ここで値 1 を持つ変数の数は試験管の番号に対応する (試験管  $T_k$  では値 1 を持つ変数の数が  $k$  個)。従って、最適解を符号化した配列の含まれる試験管は、

- (i) ゲル電気泳動において 2 次項が全て 0 に対応する長さのバンドが表れ



る

(ii) 試験管の番号が最大

であることがわかる。またクリーク数  $\omega(G)$  は

$$\omega(G) = -f(x^*) \tag{9}$$

と表されるので、クリーク数は最適解を符号化した配列の含まれる試験管の番号に等しい。以上の概要を図6に示す。

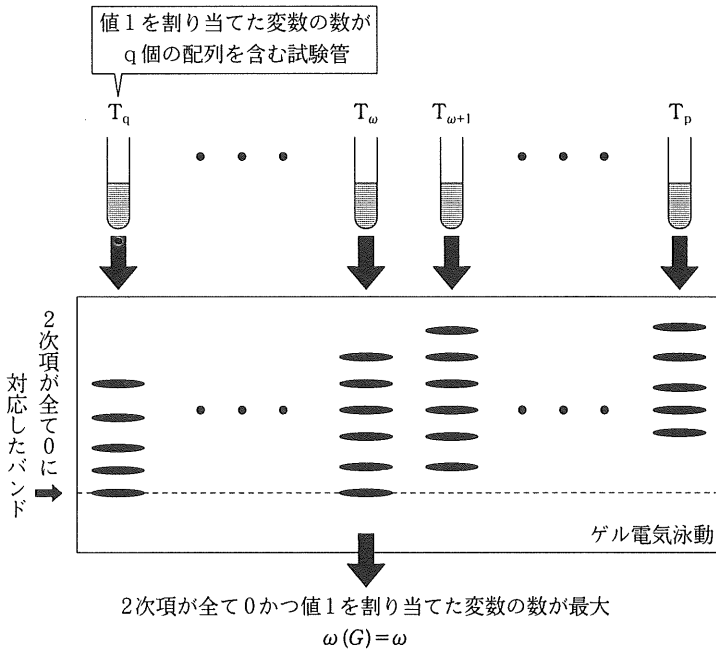


図6 ゲル電気泳動による解の評価

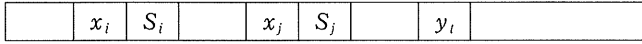
## 5. 考 察

本研究で考案した分子計算法では不適な解の一部が切断・除去されずに残るため、正しい解を導くことができない。これは Step 1 (1), Step 2 (1) において用いた符号化に起因し、これを改善するためには本論文で述べた符号化より優

離散2次計画問題の分子計算アルゴリズム

れた新しい符号化を用いる必要があると考えられる。不適な配列がヘアピン構造の形成により切断・除去されずに残る様子を以下に示す。

$x_i=0, x_j=0$ , 第  $t$  項  $x_i x_j=0$  のとき (矛盾していない)



→ヘアピン構造を作らない →適

$x_i=0, x_j=0$ , 第  $t$  項  $x_i x_j=1$  のとき (矛盾)



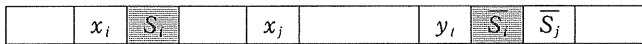
→ヘアピン構造形成切断 →適

$x_i=1, x_j=0$ , 第  $t$  項  $x_i x_j=0$  のとき (矛盾していない)



→ヘアピン構造を作らない →適

$x_i=1, x_j=0$ , 第  $t$  項  $x_i x_j=1$  のとき (矛盾)



→ヘアピン構造形成切断 →適

$x_i=1, x_j=1$ , 第  $t$  項  $x_i x_j=0$  のとき (矛盾)



→切断されずに残る →不適

$x_i=1, x_j=1$ , 第  $t$  項  $x_i x_j=1$  のとき (矛盾していない)



→ヘアピン構造を作らない →適

以上のように2つの変数  $x_i, x_j$  がともに1であるとき、その積が0であってもヘアピン構造は形成されず切断されない。これは2次項の値を評価する際に1次項の少なくとも一方が0であるかどうかで判断を行っているためである。これを改善するためには1分子での2入力1出力論理回路の評価を行うことができるような符号化を用いる必要がある。

## 結 論

本研究では最大クリーク問題を2次計画問題に置き換えて考えることにより、計算機的に求まる解の上・下界を分子計算に適用することによる計算量を軽減、2次計画問題の目的関数の特性を用いることで自律計算の可能性を示した。本論文で述べた符号化方法では一部不適な配列が残り正しい解を導くことができないが、日々刻々と提案されている新しい技術を導入することで本分子計算法の実行可能性も高まると考えられる。

## 参 考 文 献

- [1] L. M. Adleman, "Molecular Computation of Solution to Combinatorial Problems", *Science*, vol. 266, 1021-1024, 1994
- [2] I. M. Bomze, M. Budinich, P. M. Pardalos and M. Pellilo, "The Maximum Clique Problem", *Handbook of Combinatorial Optimization Supp. vol. A*, Kluwer Academic Publishers, 1999
- [3] A. T. Amin and S. L. Hakimi, "Upper Bounds on the Order of a Clique of a Graph", *SIAM Journal on Applied Mathematics*, vol. 22, No. 4, 569-573, 1972
- [4] M. Budinich, "Exact Bounds on the Order of the Maximum Clique of a Graph", *Discrete Applied Mathematics*, vol. 127, 535-543, 2003
- [5] K. Samamoto, H. Gouzu, K. Komiya, D. Kiga, S. Yokoyama, T. Yokomori and M. Hagiya "Molecular Computation by DNA Hairpin Formulation", *Science*, vol. 288, 1223-1226, 2000