

組み合わせ問題に対する DNA 計算アプローチ

01704451	北海道大学	*	山本雅人	YAMAMOTO Masahito
	北海道大学		平山拓央	HIRAYAMA Takuo
1004631	北海道大学		大内東	OHUCHI Azuma
	北海道大学		柴肇一	SHIBA Toshikazu
	北海道大学		堤香織	TSUTSUMI Kaori
	北海道大学		棟方正信	MUNEKATA Masanobu
	北海道大学		滝谷重治	TAKIYA Shigeharu

1. はじめに

L. Adleman は、1994 年に組み合わせ問題である有向ハミルトン経路問題 (DHP) を DNA を用いた分子生物学的実験によって解く手法を提案した [1]. この計算パラダイムは、DNA 計算と呼ばれ、分子化学反応のもつ超並列性、高密度性、省エネルギー性などから最近我が国でも注目を集めてきた [2][3]. Adleman の実験後、Lipton によって他の NP-完全問題の一つである充足可能性問題が、また、最近、Ouyang らによって最大クリーク問題が解かれている [4][5]. しかし、DNA 計算に関する研究の多くは、理論的研究や将来性などの議論に関するものであり、実際の実験に基づいた議論が行なわれているものは多くない。

本論文では、DNA 計算を実際に行なう立場から、DNA 計算を実現する際の問題点を明確化し議論する。実際に行なった実験結果から DNA 計算の実現可能性について検討する。

2. Adleman の実験

有向ハミルトン経路問題 (DHP) は、 n 個の頂点をもつ有向グラフにハミルトン経路 (頂点 O_{in} から頂点 O_{out} までの経路で各頂点をちょうど一度だけ含む経路) が存在するかどうかを決定する問題であり、NP-完全問題の一つである。

Adleman は 1994 年に図 1 の有向グラフに対する有向ハミルトン経路問題を DNA 計算によって解いた。ただし、 $V_{in} = V_0$, $V_{out} = V_6$ であり、このグラフには $V_0 - V_1 - V_2 - V_3 - V_4 - V_5 - V_6$ というハミルトン経路が存在する。

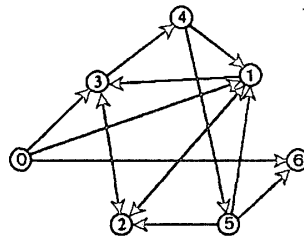
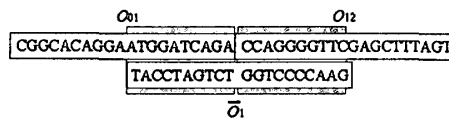


図 1: 有向グラフ

この DHP を Adleman は、Generation-and-Test 操作を基礎としたアルゴリズムを用いて解いている。すなわち、有向グラフじょうの経路を可能な限り生成し、ハミルトン経路が満たすべき条件によって徐々に選別していき、最後に残った経路がハミルトン経路であり、残った経路がなければハミルトン経路は存在しない。

DNA 計算を行なうためには、問題を DNA 上に表現する必要がある。まず、有向グラフの頂点 i を長さ 20 の (一本鎖) DNA O_i で表現する。

次に、頂点 i から頂点 j の有向辺は、図 2 のように O_i の 3' 末端側の 10 残基と O_j の 5' 末端側の 10 残基を合成した DNA $O_{i \rightarrow j}$ によって表現する。

図 2: 頂点 i から頂点 j への有向辺の DNA 表現

上記のように DNA で表現した有向ハミルトン経路問題を分子化学反応に基づく以下のような方法によって解いた。

Step 1: i, j, k を通るの経路を $O_{i \rightarrow j}$ と $O_{j \rightarrow k}$ とを連結反応によって生成する。その際、 O_j の Watson-Crick の塩基対 $\overline{O_j}$ を添え木的な役割として用いる。この結果、DNA は二本鎖になる。このように、ランダムな経路を多数生成する。

Step 2: O_{in} と $\overline{O_{out}}$ をプライマーとして用いて、PCR (複製連鎖反応) によって、 V_{in} で始まり、 V_{out} で終わる経路をエンコードする DNA のみを残す。

Step 3: agarose gel を用いた電気泳動によって、 n 個の頂点を含む DNA ($n = 7$ の時、140-base pair band) のみを分離し、PCR 増幅と gel purify を数回行い純度を上げる。

Step 4: Step 3 で残ったものを affinity purify する。さらに、結合 $\overline{O_i}$ に対してアニーリングを行うことで、頂点 i を含むもののみを残すことができる。これを、 $i = 0 \dots n$ まで繰り返す。

Step 5: PCR 増幅と gel purify する。

3. DNA 計算での問題点

現在、DNA 計算に関して以下のような問題点が指摘されている。

(1) 大規模な問題を扱うためには膨大な量の DNA が必要である。

(2) 化学反応を基礎としているための計算に誤りがある。

著者らは、(1) に関しては、現在の研究が DNA 計算がどのような問題に適用可能であるかを議論する段階にありこの問題点に対して結論を出すのは時期尚早であると考えている。実際、Adleman が行なった Generation-and-Test に基づく方法とは異なり、 μ -ブール式の評価を手続き (プログラム) を DNA 上に表現し DNA 計算によって行なう研究がなされている [6]。

(2) に関しては DNA 計算が避けては通れない問題の一つである。著者らは、DNA 計算の際の誤りが、問題の DNA 表現に大きな影響を受けていると考えている。すなわち、計算の誤りの一つは、DNA のミスアニーリング (相補対でない DNA 同士が結合して二重螺旋構造を形成する) が起こることであると考へた。

従って、一本鎖 DNA 間の干渉度という尺度を導入し、異なる DNA 配列間の干渉度が小さくなるように DNA 配列の生成を行なった。干渉度の計算方法は、以下の通りである。

定義 1 2 本の配列を並べて、相補塩基対が同じ位置に場所を調べ、 n 個の連続した相補対が並んでいれば 2^{n-1} のスコアを加算した合計が、2 本の配列をずらしながら最大となるものを 2 配列の相補性の値とする。□

定義 2 2 本の配列 A と B の干渉度を A, B の相補性と A, \overline{B} の相補性の合計とする。□

4. 実験

著者らは、新しく導入した干渉度という尺度を用いて DNA 配列を生成し、Adleman の実験と同じ問題を解いた。しかしながら、Adleman の実験では、もし、ハミルトン経路が見つければ問題ないが、ハミルトン経路が見つからなかったときには、実験過程でどのような経路が生成されていたかを知ることができない。DNA 計算に関して議論を行なう上で、実験過程でどのような経路が生成されているかを考察する必要がある。そこで、本実験では、Adleman の実験の [Step 4] の部分を以下のように変更した。

Step 4: Step 3 で残ったものを affinity purify する。さらに、大腸菌に導入し、各経路を種類ごとに分離する。その後、いくつかのコロニーから直接 PCR で DNA 断片を増幅し、シークエンサーにかけ塩基配列の決定を行った。

実験結果から、以下のことがわかった。

- ・ 20% 前後の割合でハミルトン経路を見つけている。
- ・ 長さが 140 でない経路が数多く発見されている (理論的には長さ 140 のものだけ取り出しているはず)

6. おわりに

本論文では、DNA 計算を実際に行なう立場から、DNA 計算における問題点を検討し、Adleman が行なった実験の追実験を行なった。今後は、他の問題への DNA 計算の適用可能性について検討する予定である。

参考文献

- [1] L. M. Adleman : Molecular Computation of Solutions to Combinatorial Problems. Science, Vol. 266, No. 11, pp. 1021-1024(1994)
- [2] 横森 貴 : Bio-Computing ってなんだろう? . Bit, Vol. 29, No. 8, pp. 34-38(1997)
- [3] L. Kari, 榎原 康文 : DNA コンピュータとは. 電子情報通信学会誌, Vol. 80, No. 9, pp. 935-939(1997)
- [4] R. J. Lipton : DNA Solution of Hard Computational Problems. Science, Vol. 268, No. 4. pp. 542-545(1994)
- [5] Q. Ouyang, P. D. Kaplan, S. Liu and A. Libchaber : DNA Solution of the Maximal Clique Problem. Science, Vol. 278, No. 10. pp. 446-449(1997)
- [6] M. Hagiya, M. Arita, D. Kiga, K. Sakamoto and S. Yokoyama : Towards Parallel Evaluation and Learning Boolean μ -Formula with Molecules. Proceedings of the 3rd DIMACS Workshop on DNA based Computers, pp. 105-114(1997)