

分子反応系の自動設計およびDNAのモデリング

川又 生吹*1 田中 文昭 萩谷 昌己*1
 Ibuki Kawamata Fumiaki Tanaka Masami Hagiya

*1 東京大学大学院 情報理工学系研究科 コンピュータ科学専攻

Department of Computer Science, Graduate School of Information Science and Technology, University of Tokyo

We propose an automatic design method to synthesize systems composed of biological molecules, and a modeling method of DNA molecule for efficient design. By regarding the design of molecular reaction system as a combinatorial optimization problem and applying heuristic algorithm, the purpose of the research is to automatically search for a system with a mechanism that has the function of interest. When the number of molecular species in a system increases, appropriate modeling is necessary to avoid a combinatorial explosion among molecules.

1. 導入

近年、生体由来の分子を素材として用い、人工的な機能を持つ系を合成する研究がされている。そのような研究はナノロボットや生体計測、生体内投薬制御等への応用が期待されており、DNAを素材とした回路 [Seelig 06, Zhang 07], RNAを入出力とするDNA回路 [Takinoue 08], 小分子を入力としてタンパクを出力するRNA回路 [Win 08] はその研究例である。これらの系を手により設計する場合、分子同士の相互作用の組み合わせを網羅して系の挙動を予測し、さらに予期しない反応が起きないことを保証する必要がある。しかしながら系中の分子種が増えるにつれ相互作用の組み合わせは爆発し、また分子同士は異なる速度で反応するため、系の挙動を予測することは人間にとって本質的に難しい問題である。この問題を解決する既存研究として、プログラミングモデルによる設計手法 [Yin 08, Phillips 09] が提案されているが、プログラミング自体は人手による試行錯誤が必要であり、系の設計が自動化されているとは言えない。

このような問題を踏まえ、我々はDNA論理回路の自動設計手法 [Kawamata 09] を提案した。既存研究において設計の対象とした系はDNA分子およびその水素結合のみを利用する単純な系であり、自動的に設計された系の一つについて実際の化学実験により挙動を検証した。自動設計手法では、分子反応系の設計を組み合わせ最適化問題としてとらえ、ヒューリスティックなアルゴリズムを適用することで、目的の機能を持つ系のメカニズムを自動的に発見する。

本研究では設計の対象となる分子および反応を拡張し、RNAを入出力とする論理回路の設計について説明する。特に系中の分子種の数が多い場合、分子同士の組み合わせ爆発を防ぐために適切なモデリングが必要になる。

2. DNAシステムの自動設計

我々の既存研究である自動設計手法 [Kawamata 09] の概略について説明する。DNA論理回路の設計を、評価値の最大化を行う組み合わせ最適化問題ととらえ、焼きなまし法 (SA) と呼ばれるヒューリスティックなアルゴリズムを実装した (図 1)。

連絡先: 川又生吹, 東京大学大学院情報理工学系研究科 コンピュータ科学専攻, 文京区 7-3-1 理学部 7 号館 217, ibuki@is.s.u-tokyo.ac.jp

図のフローチャートに示されているように、アルゴリズムでは初期のランダムな解から探索を始め、より評価値の高い解へ段々と遷移し、終了条件が満たされると解を出力する。

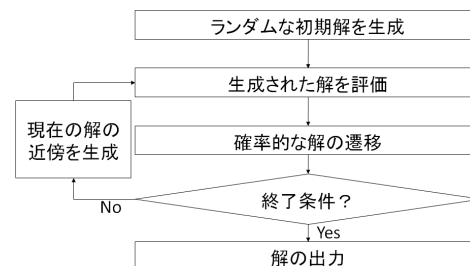


図 1: アルゴリズムのフローチャート



図 2: グラフによる DNA のモデル化

ランダムな初期解を定義するためにまず、DNAをグラフ構造によりモデル化した (図 2)。図はDNA論理回路 [Seelig 06] を左から右に向かってモデル化した例であり、反応の単位を頂点として、化学的な結合を有向辺と無向辺として表現した。さらに系を文字列により表現することで、ランダムな初期解はランダムな文字列により、近傍の生成は文字列の一部をランダムな文字で置き換える操作により定義した。解の評価の行う際は、反応速度論に基づいたシミュレーションを行い、目的の動作に近い程大きい値を返す評価関数を用いた。

分子のモデル化およびシミュレーションを拡張すればDNA以外の分子を素材とした系を設計の対象とすることができ、また評価関数を変更すれば論理回路以外の機能の設計に応用することができる。

3. RNAを入出力とした論理回路

酵素反応を利用する既存研究 [Takinoue 08] と同じ反応を扱えるようモデルを拡張することで、RNAを入出力としたDNAから成る論理回路の設計を行った。具体的には転写、逆転写、伸長の酵素反応をグラフの操作として追加し、系を長さ 24 の文字列により表現し、さらにDNA論理回路のAND演算の場合と同じ評価関数を用いてSAを適用した。

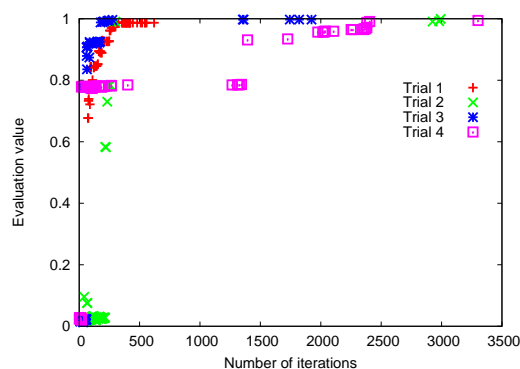


図 3: 評価値の遷移

4 回の実行例について、近傍に遷移した時に評価値をプロットしたものが図 3 である。横軸は探索した解の個数であり、縦軸は評価値の値を示している。いずれの実行例においても、評価値が増加しており局所最適解を発見したと言える。

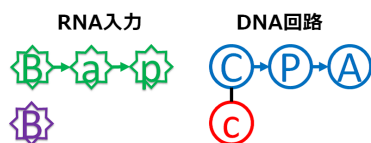


図 4: RNA を入出力とする DNA 回路のグラフ表現

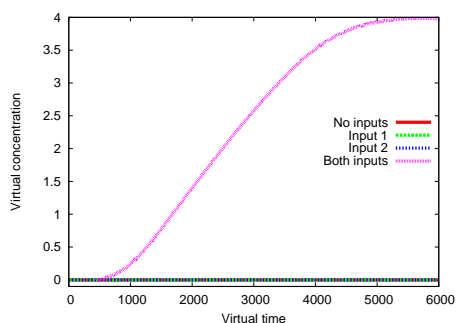


図 5: AND 回路出力の変化

図 4 は出力解の内の一つをグラフにより表現した図であり、図 5 は反応速度シミュレーションを行った結果である。横軸はシミュレーションの経過時間、縦軸は出力の濃度を示しており、単位はいずれも仮想的なものである。入力がない場合および一方だけの場合は出力がなく、両方の入力を与えられた場合のみ出力が増加しており、AND 回路の機能を確認できる。

4. DNA のモデリング

分子種の数が多い系を設計する場合、シミュレーションにより予測される構造数が爆発する問題がある。より実用的なシステムの設計に向けて、我々はグラフの部分構造に着目したモデルの抽象化を行っている [Kawamata 11]。同じ DNA からなる AND 回路の設計を対象として、抽象化を行っていない場合と行った場合で探索できる解の個数を比較する予備実験を行った。実験では短い 3 分間での探索を 5 回実行し、探索を開始してから終了するまでの間の解の個数からその平均を求めた。その結果、抽象化を行わなかった場合は平均 2248 個、行っ

た場合は平均 2492 個の解を探索しており、抽象化を行った方がより多くの解を探索することが分かった。

5. 考察

人手による分子反応系の設計は、系の挙動の予測やプログラミングの試行錯誤を伴うため本質的に難しいと考えられる。系の設計を計算機により自動的に行うため、我々はヒューリスティックなアルゴリズムを利用した組み合わせ最適化による設計手法を提案している。設計の対象となる系は、分子のモデル化およびグラフ操作を追加することで拡張できることを RNA および酵素反応を含む系を例に示した。より分子種の数が多い系を設計する場合は、単純なグラフによるモデル化では構造数が爆発してしまい適切ではない。グラフの部分構造に着目した抽象化を行い探索した場合、抽象化をしない場合に比べてより多くの解が探索できることが分かった。実験では短い時間による比較を行ったが、より長い時間探索を行えば、その差はさらに大きくなると考えられる。今後、探索手法とモデリングの関係を調べることで、人間には想像もできない程複雑な系を設計できることが期待される。

参考文献

- [Kawamata 09] Kawamata, I., Tanaka, F., and Hagiya, M.: Automatic Design of DNA Logic Gates Based on Kinetic Simulation, *DNA Computing and Molecular Programming*, Vol. 5877 of LNCS, pp. 88–96 (2009)
- [Kawamata 11] Kawamata, I., Tanaka, F., and Hagiya, M.: Abstraction of DNA Graph Structures for Efficient Enumeration and Simulation, in *Proceedings of The 2011 International Conference on Parallel and Distributed Processing Techniques and Applications*, to appear (2011)
- [Phillips 09] Phillips, A. and Cardelli, L.: A programming language for composable DNA circuits, *J. R. Soc. Interface*, Vol. 6, pp. 419–436 (2009)
- [Seelig 06] Seelig, G., Soloveichik, D., Zhang, D. Y., and Winfree, E.: Enzyme-Free Nucleic Acid Logic Circuits, *Science*, Vol. 314, No. 5805, pp. 1585–1588 (2006)
- [Takinoue 08] Takinoue, M., Kiga, D., Shohda, K.-I., and Suyama, A.: Experiments and simulation models of a basic computation element of an autonomous molecular computing system, *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.*, Vol. 78, No. 4 Pt 1, p. 041921 (2008)
- [Win 08] Win, M. N. and Smolke, C. D.: Higher-Order Cellular Information Processing with Synthetic RNA Devices, *Science*, Vol. 322, No. 5900, pp. 456–460 (2008)
- [Yin 08] Yin, P., Choi, H. M. T., Calvert, C. R., and Pierce, N. A.: Programming biomolecular self-assembly pathways, *Nature*, Vol. 451, No. 7176, pp. 318–322 (2008)
- [Zhang 07] Zhang, D. Y., Turberfield, A. J., Yurke, B., and Winfree, E.: Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA, *Science*, Vol. 318, No. 5853, pp. 1121–1125 (2007)