

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 3 日現在

機関番号：12612

研究種目：基盤研究（C）

研究期間：2010 年度 ～ 2012 年度

課題番号：22500010

研究課題名（和文） 分子ロボティクスのための反応系解析アルゴリズムの理論

研究課題名（英文） Theory of Algorithms to Analyse Chemical Reaction Systems for Molecular Robotics

研究代表者

小林 聡（KOBAYASHI SATOSHI）

電気通信大学・大学院情報理工学研究科・教授

研究者番号：50251707

研究成果の概要（和文）：

本研究では、分子がさまざまな形で干渉して多くのアセンブリを生成するような複雑な化学反応系の平衡状態を効率良く計算するための理論を確立した。この理論では、ハイパーグラフの理論と最適化理論を統合して利用することにより、化学平衡を計算するために必要な変数の個数を削減することに成功した。この理論を核酸配列の干渉系に応用した。また、この理論を拡張してRNAのフォールディングシミュレーションを近似的に高速に実行することに成功した。

研究成果の概要（英文）：

We established a theory of efficiently computing an equilibrium of a complex chemical reaction system, where molecules interact in various ways to produce a large set of assemblies. In this theory, we used hypergraph theory and optimization theory in a unified manner, and succeeded in the reduction of the number of variables which are needed for the computation of chemical equilibrium. The theory was applied to the computation of equilibrium of interacting nucleic acid strands. It was also extendedly applied to the approximate and efficient simulation of kinetic folding of an RNA molecule.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010 年度	1,400,000	420,000	1,820,000
2011 年度	1,300,000	390,000	1,690,000
2012 年度	600,000	180,000	780,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：総合領域

科研費の分科・細目：情報学・情報学基礎

キーワード：分子ロボット，化学平衡，平衡状態計算，シミュレーション

1. 研究開始当初の背景

生体高分子のように複雑な構造を形成し得

る分子が干渉し合う系は、極めて高次元の系となるため現実的な時間内に高精度で解析することが難しい。しかしながら、分子ロボ

ティクス, 分子計算, ナノテクノロジー, 生命科学などの分野で, このような系を解析する問題の重要性が急速に増している. 従って, 生成される構造体の集合が組合せ爆発を起こすような反応系を解析する問題は, さまざまな分野で重要な研究テーマとして認識されている. 物理化学の分野では, 「リガンド・インタラクション」の問題などにおいて, 古くから強い関心が持たれ, 最近ではバイオインフォマティクスの分野でも, 足場蛋白質の解析や核酸配列のインタラクションなど, 多くの研究が進められている. 高次元の化学反応系の平衡状態を解析するための他の有望なアプローチとして, マルコフ連鎖モンテカルロ (MCMC) 法がある. 高次元反応系の場合, ある特定の部分状態空間に長い時間マルコフ連鎖が留まってしまう「trapping problem」という難しい困難がつかまとう. しかしながら, 2003 年に Uncoupling-Coupling (UC) 法という非常に興味深い新しい手法が Fischer によって提案され注目を集めている.

2. 研究の目的

これらの従来研究との違いを鮮明に説明するため, 解の探索空間を「構造体の空間」と「濃度分布の空間」の2つの軸に分けて記述することにしたい. 図1において, 横軸には反応系で生成されるすべての構造体が概念的に記載されている. 縦軸は, 各構造の濃度である. このように軸をとると, 反応系の構造体の濃度分布は, 図1の実曲線で描かれたグラフとして記述できる. ここで, 平衡状態を求める問題は, 図1で実曲線のグラフとして描かれているさまざまな濃度分布の中で, 最小自由エネルギーをとる濃度分布を求める問題と理解できる.



図1 解の探索空間

このような問題を解くために従来の研究では, (I) 構造体の中で重要ではないと思われるものを除去して構造の爆発を回避する, (II) Macroscopic な状態を導入することにより, 対象とする系の解析を容易化するという手法がとられてきた. ここで, Macroscopic な状態とは, 構造体のある部分

集合を表す状態 (変数) のことであり, そのような変数を適切に導入することにより, 系を記述する式を単純化し, 系の巨視的な振舞いの解析を容易にしようとするものである.

これらの手法の問題点として, (I) の手法では, 指数爆発する構造の空間を構造を除去することによって対応するには限界があり, また, 構造の除去により平衡状態における濃度エラーが大きくなってしまおうという問題点が指摘されている. 一方, (II) の手法は極めて有効な手法であるが, 任意に与えられた構造の濃度を知りたい場合には用いることができない.

本研究では, 特に分子ロボティクスや分子計算の分野で用いられる核酸配列を中心とした生体高分子の反応系の解析の問題に焦点をあてて, 「グラフによって構造体集合を数え上げる」という新しい発想を用いて, 自己集合反応系の「平衡状態を計算する問題」と「シミュレーション問題」を高速に解くための一般的なアルゴリズム論を構築することとその限界を理論的に明らかにする. また, その限界を乗り越えるための糸口も理論的に考察する. 応用面では, この理論を核酸配列のフォールディングやそのインタラクション, 足場蛋白質の反応解析, DNA タイルの反応解析などの具体的な問題に適用し, 従来手法より劇的に高速な (数10倍以上高速な) システムを開発することを目指す.

3. 研究の方法

本研究では, 自己集合反応系の平衡状態計算問題とシミュレーション問題に対し, 組合せ爆発を起こさない効率の良い計算アルゴリズムを与える. そこで用いる仮定は,

(A) 構造体全体のギブス自由エネルギーがその部分構造の自由エネルギーの総和で求められる

という性質が反応系で成り立つことである. この仮定は, 核酸配列のインタラクションや DNA タイルアセンブリの反応系など, 多くの反応系の解析で用いられている仮定である. 申請者は, 本研究課題に関連し, 既に平衡状態計算問題に関して予備的な成果を得ている. そこでは, ハイパーグラフを用いて構造体集合を数え上げることにより, グラフの辺の個数だけの変数を用いて効率よく平衡状態を計算できることが示されている. しかしながら,

(P1) 応用システムが構築できておらず, 実用的な理論であることを実証できていない

(P2) シミュレーション問題に対するアルゴリズム理論が構築できていない

(P3) 木構造状の構造体集合が生成される反応系にしか適用できない

(P4) 計算量理論や最適化理論の立場から

みると、提案手法に対する解析に不十分な点がある

といった課題が未解決である。これらの問題を解決することに重点をおいて研究を進める。

4. 研究成果

(P4) の問題を解決するために、平衡状態を計算するための提案手法の解析を厳密に行い、平衡状態計算に対して数え上げの発想を適用する一般論を構築した ([8][9][10])。以下にその基本的な発想を述べる。

アサイクリックな有向グラフ $G=(V, E_g)$ を考える。ここで、 V は頂点の集合、 E_g は有向辺の集合である。入る辺を持たない頂点（初期頂点）の集合を V_0 、出る辺を持たない頂点（最終頂点）の集合を V_f で表す。グラフ G の初期頂点から最終頂点へのパスの集合を $PT(G)$ で表す。グラフによる構造体の列挙とは、 $PT(G)$ から構造体の集合 A への上への写像 ψ を考えることである。つまり、 G の初期頂点から最終頂点へのパスを数え上げることで構造体をすべて数え上げる。その場合、異なるパスにより同じ構造体を重複して数えることも許されるが、ここでは議論の最も本質的な部分のみを説明するため ψ は1対1であるとする。これに応じて、タイルアセンブリにおいても、タイルが x 軸まわりに回転して会合しないものとして単純化して説明する。図2は、そのような単純化したタイルアセンブリ反応系において、長さ6のタイルアセンブリを列挙するグラフである。ここで、 $S(1,p)$ および $S(6,p)$ といった形の頂点がそれぞれ初期頂点および最終頂点である。例えば、図5において $S(1,a) \rightarrow S(2,b) \rightarrow S(3,a) \rightarrow S(4,b) \rightarrow S(5,b) \rightarrow S(6,a)$ というパス α は、*ababba* というタイルアセンブリに対応する。このような対応 ψ を直感的な理解は関数呼び出しであり、有向辺は始点の関数呼び出しが終点の関数を呼び出すまでに行う処理を表す。図2のグラフでは、有向辺 $S(i,p) \rightarrow S(i+1,q)$ はタイル p と q の間の部分構造を生成していると考える。

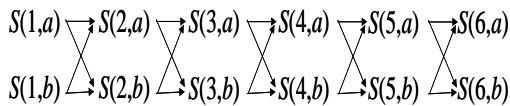


図2 グラフによる列挙

このように構成したグラフ G を用いて濃度分布間に同値関係を導入する。まず、任意の A の濃度分布 $[]$ を考える。このとき、辺 e に対して、次を定義する。

この定義を直感的に述べると、 $[e]$ は e が

生成する部分構造が反応系内で存在する濃度を表す。濃度分布 $[]_1$ と $[]_2$ は、すべての $e \in E_g$ に対して $[e]_1 = [e]_2$ となるとき局所的に等価であるという。この局所的等価関係は同値関係であり、 A の濃度分布をその同値類に分割して二段階最適化を適用することにより、平衡状態を計算するための変数の個数を劇的に削減できることが分かる。本研究では、この理論をさらに一般化し、ハイパーグラフの理論を用いて、木構造状の構造を生成する系に適用できる理論、さらには、ある種の条件下では、円環上の構造を生成する系に対しても適用可能な理論を構築した ([8])。この理論は、上記の写像 ψ が1対多の関係の場合においても適用可能な理論である。

(P1) の問題を解決するために、核酸配列のインタラクション反応に提案した理論を適用し、平衡状態を計算するアルゴリズムを開発した ([3][11][17])。これにより、従来指数時間の計算時間が必要な問題を、多項式個の変数からなる凸計画問題に帰着する非常に効率の良い計算手法を開発することができた。

(P2) の問題を解決するために、具体的にRNA分子のフォールディングのシミュレーションに提案手法を応用することを試みた ([1])。平衡状態を計算するために構成したグラフとそれに対応する凸計画問題の対象関数（凸関数）に対して、できるだけ実際の化学反応が進行する方向に凸関数の局面を降下していくことを目指すアプローチである。この手法によるシミュレーション例が以下の図3である。

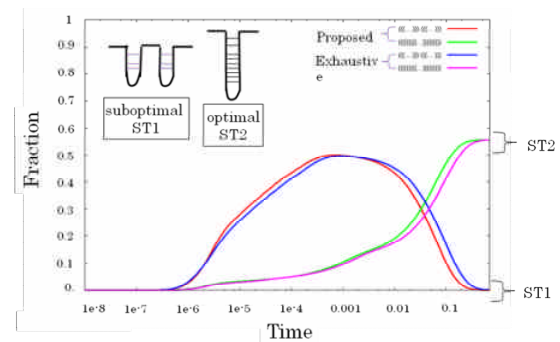


図3 RNAフォールディング

用いたRNA配列は、CCCUUUGGGACCCUAAAGGGである。この図において、Proposed とあるのが提案手法によるシミュレーション結果であり、Exhaustive とあるのがすべての構造を生成したシミュレーションした厳密なシミュレーション結果である。近似的なシミュレーションが提案手法によって実現できていることがわかる。さまざまな配列に対して適用してみたところ、配列が長くなるほど計算速度は提案手法が著しく速くなるこ

とが検証された。これは、提案手法が配列長さの多項式個の変数を用いたシミュレーションであるのに対し、厳密なシミュレーションは、配列長の指数関数個の変数を用いることから当然の帰結である。また、提案手法は厳密に平衡状態に正しく収束することを理論的に保証した([1])。

(P3)の問題に関しては、[8]において、ある種の条件下では、円環状の構造を生成する反応系に対しても効率良く平衡状態を計算できることを示した。しかしながら、平衡状態が効率良く計算できる場合とできない場合の境界をより理論的に明確にするところまでは至らなかった。この問題を解決することは今後の課題である。

本研究では、少数分子反応系の定常状態の計算問題に数え上げの発想を応用することも試みた。特に、一分子反応系で、分子がとる構造の個数が分子のサイズに関して指数関数的に組み合わせ爆発するような系に対して、分配関数が効率良く計算できることを示した([4])。

また本研究では、連携研究者らとの共同研究などにより、分子ロボティクスに利用できる化学反応回路の提案とそのシミュレーション問題に取り組んだ([2][16])。このような経験を通しながら、分子ロボティクスの分野で役立つ反応系解析の理論の今後の展開を見据えることができた([7][12][13][14][15])。

さらに、補助的な成果として、化学反応系の計算モデルの計算能力に関する理論的な成果を得ることに成功した([5][6])。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 8 件)

[1] Takumi Tanigawa, Satoshi Kobayashi, Efficient and Approximate Simulation Algorithm of Kinetic Folding of an RNA Molecule, Proc. of 2011 International Conference on Parallel and Distributed Processing Techniques and Applications, pp. 706-712, 2011. (査読有)
(<http://world-comp.org/p2011/PDP5089.pdf>)

[2] Ryo Hirose, Satoshi Kobayashi, Ken Komiya, DNA Logic Circuits with a DNA Polymerase and a Nicking Enzyme, Proc. of 2011 International Conference on Parallel and Distributed Processing Techniques and Applications, pp. 713-719, 2011. (査読有)

(<http://world-comp.org/p2011/PDP5090.pdf>)

[3] Takaya Kawakami, Satoshi Kobayashi, Enumeration Approach to the Analysis of Interacting Nucleic Acid Strands, Proc. of 17th International Conference on DNA Computing and Molecular Programming, Abstracts of Talks and Posters, p. 47, 2011. (査読有)

[4] Satoshi Kobayashi, Enumeration Approach to the Computational Analysis of High-Dimensional Monomolecular Chemical Master Equation, Proc. of 17th International Conference on DNA Computing and Molecular Programming, Abstracts of Talks and Posters, p. 52, 2011. (査読有)

[5] Fumiya Okubo, Satoshi Kobayashi, Takashi Yokomori, Reaction Automata, Theoretical Computer Science, 429, pp. 247-257, 2012. (査読有)
(DOI:10.1016/j.tcs.2011.12.045)

[6] Fumiya Okubo, Satoshi Kobayashi, Takashi Yokomori, On the properties of language classes defined by bounded reaction automata, Theoretical Computer Science, 454, pp. 206-221, 2013. (査読有)
(DOI:10.1016/j.tcs.2012.03.024)

[7] S. Murata, A. Konagaya, S. Kobayashi, H. Saito, M. Hagiya, Molecular Robotics: A New Paradigm for Artifacts, New Generation Computing, 31, 27-45, 2013. (査読有) (DOI: 10.1007/s00354-012-0121-z)

[8] Satoshi Kobayashi, Enumeration Approach to Computing Chemical Equilibria, Theoretical Computer Science, to appear. (査読有)

[学会発表] (計 8 件)

[1] Satoshi Kobayashi, Enumeration of Structures and Efficient Analysis of Complicated Chemical Reaction Systems, *9th International conference on Unconventional Computation*, Tokyo, June, 2010. (Invited)

[2] Satoshi Kobayashi, Enumeration Approach to the Analysis of Complex Chemical Reaction Systems, *16th Int. Conference on DNA Computing and Molecular*

Programming, Hong Kong, June,
2010. (Invited)

[3] 松本典子, 小林聡, 仮想核酸配列インタラクション反応系の平衡状態解析, 情報処理学会研究報告, Vol. 2011-MPS-086, 論文番号 25, pp.1-7, 東京, 12 月, 2011.

[4] 小林聡, 知能を実現する化学反応回路の構築を目指して, 生命医薬情報学連合大会 (招待講演), 東京, 10 月, 2012.

[5] 小林聡, 分子ロボットに知性は実現できるか - 計算論的立場から -, 人工知能学会合同研究会 第 51 回分子生物情報研究 (SIGMBI) (招待講演), 東京, 11 月, 2012.

[6] 小林聡, グラフによる分子種の数え上げと化学反応系の解析, 「細胞を創る」研究会 5.0 (招待講演), 神奈川, 11 月, 2012.

[7] 小林聡, 分子ロボティクス --- 感覚と知能を備えた分子ロボットの創成 ---, 計測自動制御学会 システム・情報部門 学術講演会 2012 (SSI2012) (招待講演), 名古屋, 11 月, 2012.

[8] 小林聡, DNA を用いた組み合わせ回路の高速化について, 第92回数理解モデル化と問題解決研究会, 武雄, 3 月, 2013.

[図書] (計 1 件)

[1] Satoshi Kobayashi, Takaya Kawakami, Enumeration Approach to the Analysis of Interacting Nucleic Acid Strands, In Biomolecular Information Processing - From Logic Systems to Smart Sensors and Actuators, E. Katz (Ed.), Chapter 12, p.225-244, Wiley-VCH, Weinheim, Germany, 2012. (査読有)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

<http://comp.cs.uec.ac.jp/projects.html>

6. 研究組織

(1) 研究代表者

小林 聡 (KOBAYASHI SATOSHI)
電気通信大学・大学院情報理工学研究科・教授
研究者番号: 50251707

(2) 研究分担者 なし

(3) 連携研究者

山下 雅史 (YAMASHITA MASAFUMI)
九州大学・大学院システム情報科学研究科・教授
研究者番号: 00135419

村田 智 (MURATA SATOSHI)
東北大学・大学院工学研究科・教授
研究者番号: 10334533

鈴木 泰博 (SUZUKI YASUHIRO)
名古屋大学・大学院情報科学研究科・准教授
研究者番号: 50292983

小宮 健 (KOMIYA KEN)
東京工業大学・大学院総合理工学研究科・助教
研究者番号: 20396790

葛谷 明紀 (KUZUYA AKINORI)
関西大学・工学部・准教授
研究者番号: 00456154

瀧ノ上 正浩 (TAKINOUE MASAHIRO)
東京工業大学・大学院総合理工学研究科・講師
研究者番号: 20511249