

化学反応ネットワークにおける最適反応経路候補の列挙

中野 裕太^{1,a)} 瀧川 一学^{2,3,4,5,b)}

概要: 化学において近年, 人工力誘起反応法 (AFIR 法) と呼ばれる量子化学計算に基づいた方法による反応経路の探索が可能になった. しかし, この方法により生成される化学反応ネットワークは, 様々な遷移状態が含まれており複雑であるため, このネットワークにおける最適な反応経路の探索が課題となっている. そこで, 本論文ではまず, AFIR 法により生成されるネットワークを, 重み付きの有向グラフへ変換する方法を述べる. 続いて, そのグラフ上における経路に関して, 含まれる辺の重みの最大値と辺の個数に着目した優先順位を付け, K-Best での列挙へ帰着させることで, 化学反応ネットワークにおける最適な反応経路の候補を列挙する手法を提案する. また, この手法を実際の反応経路ネットワークへ適用することで, 大きなネットワークに対しても効率よく最適な反応経路の候補の列挙を行うことが可能であることを示す.

1. はじめに

化学反応は, ある物質を別の物質に変える物質変換の基礎であり, 電力などのエネルギー生成, 身の回りのあらゆる材料・製品の製造, 創薬・医療・食品や, 我々の生命活動現象そのものまで多様な形で生活の基盤を支えている. 新規の化学反応や効率良い化学反応を自由自在に設計・探索・理解することは化学研究の一つの究極的目標である.

近年の理論化学の発展とコンピュータの計算能力の向上により, 電子の運動に関するシュレーディンガー方程式を解く量子化学計算に基づいて化学反応の機構解明が可能となってきた. 化学反応は, エネルギー的に安定な平衡構造が一時的に不安定な遷移状態を経て別の安定平衡構造に至る過程である. 任意の多原子分子系でこの遷移状態構造を網羅的に探索することはこれまで不可能と考えられてきたが, 近年独自の理論に基づく画期的な反応経路自動探索法 [1,2] が考案され, 既存の反応機構の解明だけでなく, 実験に先立った生成物の予測や新しい反応設計, 新物質や材料の創出を担う新しい方法論として注目されている [3-6].

反応経路自動探索法により計算された系の安定平衡構造を頂点, その間の遷移状態を辺とし, 得られるグラフ構造を化学反応ネットワークと呼ぶ. 化学反応ネットワーク上の経路は理論的に可能な化学反応経路と対応するため, ある物質から他の物質への最適な反応経路を探索する合成経路設計, ある物質を生成するための最適な反応経路を探索する逆合成解析, ある物質から起こりうる最適な反応経路を探索する反応予測, など, 様々なシナリオで知識発見のための効率的な手法が期待されている.

着目する系に含まれる原子数が多くなるほど, 化学反応ネットワークは急激に複雑となる. C_6H_6 のような 12 原子の系でさえ, 化学反応ネットワークの頂点数は 2004 にもなる. 医薬品や有機材料に関する反応を考える場合, 少なくとも構成原子数が 100 程度の分子を扱う必要があり, より大規模なネットワークとなるため, 経路の全列挙や直接人手で経路の評価を行うことは現実的ではない. 一方, 実際の化学反応設計の際に考慮すべき点は複数あり, 反応に必要なエネルギーが少ない, 素反応のステップ数が少ない, 可能な限り不要な副生成物を生じない (反応収率が高い) といった定量的尺度の他に, 有害な中間生成物を生成しない, 実験が行いやすい, 反応物が手に入りやすい, など定量的ではない尺度もありうることから, 最適な反応経路を得る組合せ最適化としての定式化も困難である.

そこで本研究では, 複雑で非定量的な最適化基準を直接定式化する代わりに, 反応経路が備えるべき基本的な要件を満たすように, 最適となりうる反応経路候補を列挙する手法を提案する. 我々の先行研究 [7] では反応経路中のエネルギーの上限に制約を設けた反応経路列挙を可能とした

¹ 北海道大学 工学部
Sapporo, Hokkaido 060-8628, Japan
² 北海道大学 大学院情報科学研究科
Sapporo, Hokkaido 060-0814, Japan
³ 北海道大学 化学反応創成研究拠点 (WPI-ICReDD)
Sapporo, Hokkaido 001-0021, Japan
⁴ 理化学研究所 革新知能統合研究センター (AIP)
Chuo-ku, Tokyo 103-0027, Japan
⁵ JST さきがけ
Kawaguchi, Saitama 332-0012, Japan
a) nakano@art.ist.hokudai.ac.jp
b) takigawa@ist.hokudai.ac.jp

が、この手法では出力を決定グラフとして圧縮された構造で直接構築するとは言え、制約を満たす経路の全列挙となるため、適用できる化学反応ネットワークの規模に限界があった。提案手法では、反応ステップが短く、反応が経過する状態の最大最小エネルギー差が小さいという基本的な順位づけを考え、このランク付けに基づく K -Best 列挙を行う。最も有望そうな K 個の反応経路候補のみの列挙を行うことで、より大規模な化学反応ネットワークにも適用可能な効率的な手法を提案する。

以下の本文では、2章で準備として、化学反応ネットワークと反応経路探索における基準、および、グラフ上の K 最短路問題について述べる。3章では本稿で提案する反応経路へのランク付けについて説明し、そのランク付けに基づいた化学反応ネットワークの変換方法と、変換されたグラフ上におけるコスト関数について述べ、最適反応経路候補の列挙が K 最短路問題に帰着されることを示す。また最後に、実際の化学反応経路探索における、合成経路設計、反応予測、逆合成解析の3つのシナリオにおいて、先述の提案手法と既存の K 最短路問題に対するアルゴリズムを用い、最適反応経路候補の列挙を行う方法について述べ、4章でその性能評価のための計算機実験の結果を示す。5章で本稿をまとめる。

2. 準備

本章では準備として、先述した化学反応ネットワークと、反応経路探索において望ましい反応経路が満たすべき基本的な要件、および、 K 最短路問題について述べる。

2.1 化学反応ネットワーク

対象となる系としてある原子の集合を考える。例えば、炭素 C が 4 つ、水素 H が 6 つの C_4H_6 の系は、ブタジエンやシクロペンテンなどよく知られた分子を多数含み、これらの分子構造はエネルギー的に安定な平衡構造を取る。この系で起こりうる化学反応とは、こうした安定平衡構造間の変換であり、その過程ではエネルギー的に不安定な遷移状態を一時的に経過して、別の安定平衡構造へ至る。この PES 上の経路が化学反応経路に対応する。

化学反応経路を考えるためには、物質はエネルギーが小さいほど安定することから、まず与えられた構造（ある座標をとる原子集合）に対してエネルギー値を求める必要がある。量子化学計算では、対象とする分子の構造（系内の原子の座標）が与えられると、電子の運動に関するシュレーディンガー方程式を解くことにより、その構造のポテンシャルエネルギー値（位置エネルギー値）を計算する。エネルギー値は、構成する各々の原子の位置座標の値により変化するため、これを変数とした関数と見ることが可能であり、ポテンシャルエネルギー曲面（**PES, Potential Energy Surface**）と呼ばれる。

化学反応経路は、与えられた系の PES 上の極小点である安定平衡構造（**EQ, Equilibrium Structure**）と一次鞍点である遷移状態構造（**TS, Transition State**）で特徴付けられる。TS は二つの EQ を結ぶ曲面上の経路において峠にあたる中間状態に対応する構造であり、ある EQ からある TS を経て別の EQ へ変わる経路が実際の化学反応経路に対応する。化学反応系の原子数が増えるごとに構造変数は指数的に増えるため、EQ や TS の探索は非常に高次元で極小点および 1 次鞍点を求める問題となる。与えられた一つの構造のエネルギー値（PES の関数値）を計算するだけでもシュレーディンガー方程式を解く量子化学計算が必要なことから、特に、TS を網羅的に探索することはこれまで現実的には不可能と考えられてきた。

ところが、近年量子化学計算と独自の理論に基づいて、この反応経路の網羅的自動探索を行う画期的な計算法 [1, 2] が開発された。特に、前田・諸熊らにより 2010 年に開発された人工力誘起反応法（AFIR 法, Artificial Force Induced Reaction）[2] を用いることで、PES 上の全ての特徴点を網羅的、効率的に探索することが可能になった。これらの研究により、教科書的な $A+B \rightarrow C+D$ という単純に見える反応でも、実際はより複雑な多数の安定平衡構造と遷移状態構造からなる複雑なものであることが明らかになり、化学反応の設計・探索・理解のためには、この複雑な化学反応ネットワークを解析する汎用的な手法が必要と考えられる。

AFIR 法で得られる化学反応経路は、EQ を頂点、TS を無向辺としたグラフ構造を取る化学反応ネットワークを成し、ある化学反応経路がこのグラフ上の 1 つの経路に対応する。従って、化学反応経路の探索がこのネットワーク上の頂点間の経路を探索する問題に帰着される。ポテンシャルエネルギー曲面上の安定平衡構造の集合を EQ、遷移状態の集合を TS とする。本稿では、頂点集合を EQ、辺集合を TS とした重み付き無向グラフ、 (EQ, TS, w_{EQ}, w_{TS}) を化学反応ネットワークとする。 $w_{EQ} : EQ \rightarrow \mathbb{R}$, $w_{TS} : TS \rightarrow \mathbb{R}$ はそれぞれ頂点重み関数、辺重み関数で、対応する状態のポテンシャルエネルギー値を与える写像とする。例として、図 1 に 4 原子系 HCHO の化学反応ネットワークを示す。

2.2 反応経路探索で考慮すべき項目

本研究では、望ましい化学反応経路の基準が一般に複数かつ非定量的であることから、そのような複雑な基準を無理やり尺度化し単一の最適な反応経路を求めるのではなく、有望な化学反応経路が満たす基本的な要件に基づくランク付けにより、最適な反応経路の候補の列挙を行う。

この節では、佐藤ら [8] により述べられている、反応経路探索で考慮すべき項目について簡単に取り上げる。3.1 節では、こうした項目に基づき、提案する反応経路に対するランク付けの基準を述べる。

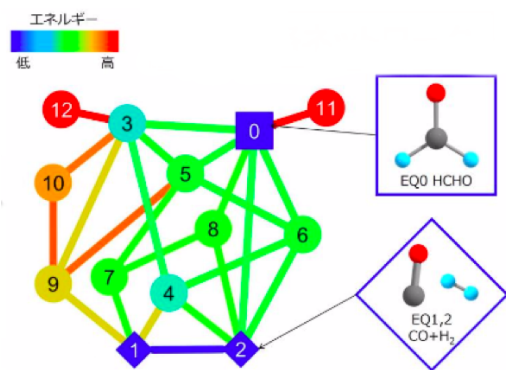


図 1 HCHO の化学反応ネットワーク

分子構造 分子を構成する原子がどのような配置にあるか、どのように結びついているかを表したものの。平面構造から立体構造などが考えられ、着目する物質を考える上で重要な基準となる。

EQ のポテンシャルエネルギー ある安定平衡構造に対応するエネルギーの値。EQ の安定性や反応性に対する基準となる。

TS のポテンシャルエネルギー ある遷移状態構造に対応するエネルギーの値。EQ のポテンシャルエネルギー値との差である **活性化エネルギー** が反応障壁となり、対応する反応の起こりやすさの基準となる。

反応回数 ある反応経路内に含まれる TS (辺) の個数で表され、対応する反応経路の効率に対する基準となる。

2.3 K 最短路問題

この節で述べる **K 最短路問題 (K Shortest Paths Problem)** とは、重み付きグラフ $G = (V, E, w)$ の頂点 $s, t \in V$ について、 s から t への経路集合 $P_{s,t} \in 2^{2^E}$ に対し、重みが小さい順に K 個の $s-t$ 経路を列挙するという問題である。

2.3.1 既存手法と K 最単純路問題

この問題は、幅優先探索 (**Breadth-First Search, BFS**) に基づく素朴な方法で解けることが古くから知られているが、これを改良した現時点で知られている最良の手法に Eppstein のアルゴリズム [9] がある。しかし、最短路に対し特に制約がないため、この手法において列挙される K 個の経路の中には、閉路を含む非単純路が含まれる可能性がある。そのため、経路候補を閉路を持たない単純路 (simple path) のみに制約した **K 最単純路問題** も考えられる。化学反応経路としては、閉路のある経路は意味がないため、単純路制約を考慮する必要がある。

K 最単純路問題 について、古くから知られている現時点で最良のアルゴリズムの一つに Yen のアルゴリズム [10] がある。このアルゴリズムの特徴として、Dijkstra 法などの標準的な最短路アルゴリズムを繰り返し用いることで、所望の K 個の単純路を列挙するという点がある。最短路

アルゴリズムを用いる部分以外は、グラフ上の頂点、辺を削除するなどの、簡単な操作で Yen のアルゴリズムが構成されている。

2.3.2 単一起点 (単一終点) の K 最短路問題

通常是最短路問題では始点 s から終点 t への $s-t$ 経路を考えるが、例えば始点 s から伸びる任意の経路、あるいは、終点 t を終端とする任意の経路について、その経路の重み順に列挙を行う単一起点 (単一終点) の **K 最短路問題** を考えることができる。つまり、

$$P_s = \bigcup_{v \neq s} P_{s,v}, \quad \text{もしくは} \quad P_t = \bigcup_{v \neq t} P_{v,t}$$

なる経路集合に対して重みが小さい K 個を列挙するという問題を考える。

P_s, P_t における探索や列挙に対する一般的な方法の一つとして、仮想頂点を導入して $s-t$ 経路列挙へ帰着する方法がある。図 2 に例として、 P_s に対する問題における、仮想頂点の導入を示す。左右のグラフにおいて、経路 $\{e_1, e_3\}$ と経路 $\{e_1, e_3, e'_3\}$ 、経路 $\{e_2\}$ と経路 $\{e_2, e'_2\}$ がそれぞれ対応している。この例からもわかるように、仮想頂点への辺重みを適切に定義することで、1 対 1 対応ができるため、2 頂点間の $s-t$ 経路列挙問題に帰着することができる。

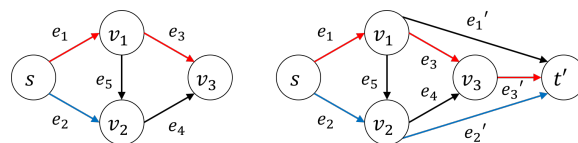


図 2 1 対 1 対応の例

また、前述の Eppstein の論文 [9] でも下記のように触れられている通り、この問題は素朴な幅優先探索により直接的に解けることもよく知られている。

Finding the k shortest paths between two terminals s and t has been a difficult enough problem to warrant much research. In contrast, the similar problem of finding paths with only one terminal s , ending anywhere in the graph, is much easier: one can simply use breadth first search.

3. 提案手法

ここではまず、先述した反応経路探索で考慮すべき項目に基づいて、本稿における反応経路のランク付けを説明する。次に、活性化エネルギー値に基づく化学反応ネットワークの変換方法と、変換を行ったグラフ上に対するコスト関数について述べる。最後に、本研究での最適反応経路候補列挙の問題が、変換を行ったグラフ上において、コスト関数をもとにした **K 最短路問題** と同様の問題になることを示す。

3.1 反応経路候補のランク付け

2.2節にて述べた佐藤らによる項目に基づいて、本稿における本稿における反応経路のランク付けを考える。

反応に対してまず挙げられる基準として、反応の起こりやすさがある。そのため第一基準として、起こりやすさを定量的に表した活性化エネルギーの値を用いる。EQは極小点であり各EQ周辺のTSのエネルギー値は必ずEQのエネルギー値より大きいことより、この値は正の実数値をとり、化学反応において、ある経路に対しこの最大値がより小さいものが起こりやすいとされている。そのため、起こりやすい反応経路である活性化エネルギーの最大値が、より小さい経路を高くランク付けする。

続いて、活性化エネルギーの最大値が同じ経路に対するランク付けを考える。ここで反応の効率に関わる反応回数(素反応のステップ数)が、より少ない反応が良いと考え、第二基準として反応回数を用い、同じ活性化エネルギーの最大値を持つ反応経路に対し、反応回数がより少ないものを高くランク付けする。

3.2 活性化エネルギーグラフへの変換

ランク付けの第一基準として活性化エネルギー値を用いるが、化学反応ネットワークのある辺 $e \in TS$ に対する重み $w_{TS}(e)$ は、遷移状態 (TS) に対するポテンシャルエネルギー値であるため、直接、辺の重みを考慮しランク付けを行うことは出来ない。ここで、直接辺の重みを利用できるようにするために、辺の重みが活性化エネルギー値である活性化エネルギーグラフというものを考える。

頂点对 $u, v \in EQ$ と、その間の辺 $e = \{u, v\} \in TS$ に対し、活性化エネルギー値の重みがついた u から v への有向辺、 v から u への有向辺を考え、 e をこの2つで置き換える。こうして得られた辺集合を E_{AE} 、頂点集合を $V_{AE} = EQ$ とした重み付き有向グラフ、 $G_{AE} = (V_{AE}, E_{AE}, w_{AE})$ を活性化エネルギーグラフと呼ぶ。 $w_{AE} : E_{AE} \rightarrow \mathbb{R}_{\geq 0}$ は辺重み関数で、対応する辺の活性化エネルギー値を与える写像とする。図3にこの変換の例を示す。左の化学反応ネットワークを変換したものが、右の活性化エネルギーグラフである。Algorithm 1に、この変換を行う擬似コードを示す。

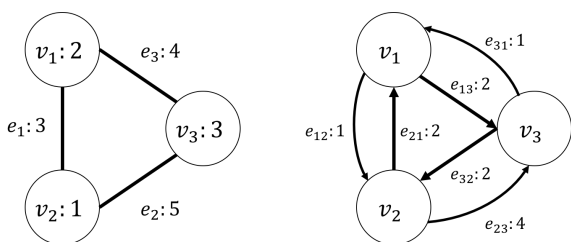


図3 化学反応ネットワークの変換例

Algorithm 1 ConvertCRNetwork(EQ, TS, w_{EQ} , w_{TS})

```

1:  $V_{AE} \leftarrow EQ$ 
2:  $E_{AE} \leftarrow \emptyset$ 
3: for all  $e = \{u, v\} \in TS$  do
4:    $w_{AE}((u, v)) \leftarrow w_{TS}(e) - w_{EQ}(u)$ 
5:    $w_{AE}((v, u)) \leftarrow w_{TS}(e) - w_{EQ}(v)$ 
6:    $E_{AE} \leftarrow E_{AE} \cup \{(u, v), (v, u)\}$ 
7: end for
8: return  $G_{AE} = (V_{AE}, E_{AE}, w_{AE})$ 

```

3.3 コスト関数の導入

2.2節の基準による経路の順位付けのために、活性化エネルギーグラフ上の経路に対するコスト関数 w を導入する。

活性化エネルギーグラフを $G_{AE} = (V_{AE}, E_{AE}, w_{AE})$ とし、 $E_{AE} = \{e_1, e_2, \dots, e_{|E_{AE}|}\}$ は重みの昇順にソートされているとする。このとき、 G_{AE} 上の経路 $P \in 2^{E_{AE}}$ を考える。 P に含まれる辺の中で最大の重みを持つ辺の添字の値 i_{\max} 、 P における辺の数を n_P とする。このとき、コスト関数 $w : 2^{E_{AE}} \rightarrow \mathbb{N}$ を式 (1) で定める。

$$w(P) = i_{\max} \cdot |E_{AE}| + n_P \quad (1)$$

このコスト関数は、 i_{\max} の値に対する変化量を大きくすることで、 $w(P)$ の小さい経路をより高いランクとすることが出来るため、 $w(P)$ を本稿で用いる基準を反映した経路に対する評価値とすることが可能である。つまり、この値が小さい経路が、より最適な反応経路と考えられる。

3.4 K 最短路問題への帰着

活性化エネルギーグラフへの変換とコスト関数の導入により、最適反応経路候補の列挙を、活性化エネルギーグラフ $G_{AE} = (V_{AE}, E_{AE}, w_{AE})$ における経路 $P \in 2^{E_{AE}}$ に対し、式 (1) の $w(P)$ が小さい順に列挙するという問題へ帰着することができる。

$w(P)$ の部分は、通常的最短路問題では、辺重み総和 $\sum_{e \in P} w_{AE}(e)$ であるが、式 (1) の $w(P)$ を経路長として考えても同様に求められる、 $w(P)$ というコスト関数を経路長としたときの、 G_{AE} 上の最短路問題を、Dijkstra 法などの既存の標準的な最短路アルゴリズム等を用い計算することが可能である。したがって、2.3節で述べた、 K 最短路問題へと帰着される。

そのため、標準的な最短路アルゴリズムを用い構成される Yen のアルゴリズムを、提案するコスト関数の値を基準とした K 最短路問題に用いることができる。

また w の定義である式 (1) より、ある経路の端点に対し隣接する辺を加えた経路のコストは、元の経路のコストよりも必ず大きくなるため、単一始点もしくは単一終点に対する K 最短路問題も、コスト関数の値をもとにした優先度付きキューを用いた素朴な幅優先探索により同様に列挙できる。

いずれの場合にも、本稿で考える経路は閉路のない単純

路とする。閉路がある経路は、化学反応経路の途中で生成された物質が、いくつかの反応を経た後に再度出現するということを意味するため、反応経路探索の観点からは意味を持たないためである。

3.5 典型的な反応経路探索シナリオに対する提案手法

この節では、実際の化学反応経路探索において考えられる合成経路設計・逆合成解析・反応予測の3つのシナリオに対し、望ましい反応経路の要件をもとに具体的な問題設定を与える。その上での最適反応経路候補の列挙方法を、先述した提案手法と2.3節で述べた K 最短経路問題に対する既存手法を用い具体的に述べる。

3.5.1 合成経路設計

合成経路設計において対象とする反応経路は、ある反応物からある生成物への反応経路である。つまり、この設定における最適反応経路列挙は以下の問題に帰着できる。

入力 (EQ, TS, w_{EQ} , w_{TS}), $s, t \in EQ, K \in \mathbb{N}$

出力 ある始点 s から終点 t への全ての単純路の中から最適反応経路の候補 K 個。

この問題は、先述した Yen のアルゴリズムを用いることで簡単に列挙ができる。まず、化学反応ネットワークを活性化エネルギーグラフへ変換する。続いて、得られた活性化エネルギーグラフに対し、ある候補集合サイズ K を定め Yen のアルゴリズムを用いることで、最適反応経路候補を列挙できる。

3.5.2 反応予測

反応予測において対象とする反応経路は、ある反応物から起こりうる反応経路である。つまり、この設定における最適反応経路列挙は以下の問題に帰着できる。

入力 (EQ, TS, w_{EQ} , w_{TS}), $s \in EQ, K \in \mathbb{N}$

出力 始点 s から s を除く任意の頂点への全ての単純路の中から最適反応経路の候補 K 個。

手法 1 (Yen)

2.3節で示したように、仮想頂点を導入することにより、この問題に対し、 s から仮想頂点への Yen のアルゴリズムを用いた列挙が可能である。ここで、列挙を行った各反応経路に対し、最後の辺（仮想頂点と共に導入された辺）を削除する必要がある。

この仮想頂点の導入に関し、対応する経路のコストを考える。例として、図2における、経路 $\{e_1, e_3\}$ と対応する経路 $\{e_1, e_3, e'_3\}$ のコストを考える。含まれる辺の重みの最大値と辺の個数について考えると、 $w_{AE}(e'_3) = 0$ より、

$$\max(e_1, e_2, e'_3) = \max(e_1, e_2) \quad (2)$$

$$|\{e_1, e_2, e'_3\}| = |\{e_1, e_2\}| + 1 \quad (3)$$

となることがわかる。この例からわかるように、 s から t への単純路のコストは、対応する経路のコストよりも1大きいものとなるため、仮想頂点による対応はコストの大小

を保存することがわかる。

手法 2 (BFS)

この設定では単一始点の K 最短単純経路問題となるため、素朴な幅優先探索に基づく直接的な列挙も可能である。ただし、経路が単純かどうかを判定しながら経路集合を管理する必要がある。

優先度付きキューにより経路を管理し、用いる優先度付きキューの優先度を、コスト関数によるランク付けと同様に定め、この優先度付きキューから取り出された順に K 個の経路を考え、幅優先的に探索していくことで、所望の最適反応経路の列挙を行うことが出来る。見つかった経路をキューに入れる際に、その経路が単純かどうかを判定することで、単純経路の条件を加味することができる。経路が単純かどうかの判定については、キューにおいて経路とその経路に含まれる頂点集合を同時に管理しておき、新たに加えられる頂点がすでに頂点集合の中に存在するかどうかで効率的に行うことができる。

3.5.3 逆合成解析

逆合成解析において対象とする反応経路は、ある生成物を生成するための反応経路である。つまり、この設定における最適反応経路列挙は以下の問題に帰着できる。

入力 (EQ, TS, w_{EQ} , w_{TS}), $t \in EQ, K \in \mathbb{N}$

出力 t を除く任意の頂点から終点 t への全ての単純路の中から最適反応経路の候補 K 個。

手法 1 (Yen)

前述する問題と逆のような問題を考えることとなり、単一終点の K 最短単純経路問題となるため、先程と同様に仮想頂点 s' を導入することにより、 $s'-t$ の単純経路を考えることで、Yen のアルゴリズムを用いた列挙が可能である。ここで、この列挙において最初の候補経路として得られる、仮想始点から仮想辺を用いた辺を無視し列挙する必要がある。

手法 2 (BFS)

また、先程と同様に優先度付きキューと幅優先探索を用いた直接的な列挙も可能である。しかし、単純に実行するだけでは先程と同様の列挙となり、得られる最適反応経路の候補は不適切なものとなる。ここで、活性化エネルギーグラフの辺の向きを反転させることにより、この手法を用いた列挙が可能になる。

4. 実際の化学反応系への応用

今回の各問題設定に対し、それぞれの手法を実装し、実際の化学反応ネットワークへ適用させた際の性能について述べる。実験環境は、OS が 64-bit Ubuntu 16.04 LTS, CPU が Intel Core i7-3930K 3.20GHz, RAM が 64GB である。また、実際に用いる化学反応ネットワークのデータとして、表1に示す2つを用いた。これらデータに対する、我々の既存手法である決定グラフに基づく全列挙手法 [7] の実行時間も同表に記載している。C₄H₆ のデータは、既

存手法 [7] で列挙に失敗したデータを用いた。

表 1 用いる化学反応ネットワーク

化学反応ネットワーク	EQ(頂点) 数	TS(辺) 数	既存手法 [7]
HCHO	13	24	20[ms]
C ₄ H ₆	65	225	メモリ不足で実行できず

表 2, 3, 4 に, ランダムに選んだ EQ に対し実際に最適反応経路の列挙を行った結果を示す。それぞれ, 候補集合サイズを $K = 100$ とし, 100 回動作させた際の平均値, 最大値, 最小値を記載している。表 3, 4 における, Yen, BFS はそれぞれ, Yen のアルゴリズムに基づく手法, 幅優先探索に基づく手法に対応する。

どの場合でも数 10 ミリ秒から数 100 ミリ秒単位で動作し, 高速であることがわかるため, 提案手法が有効であることが確認できる。また, 既存手法 [7] では列挙に失敗した, C₄H₆ のネットワークに対しての列挙に成功し, 同様に高速に動作していることがわかる。Yen のアルゴリズムに基づく手法, 幅優先探索に基づく手法を表 3, 4 から比較すると, 後者の手法の方がより高速で動作し, グラフサイズの増加に関わらず一定の時間で動作していることがわかる。

表 2 合成経路設計シナリオに対する結果 ($K = 100$)

化学反応ネットワーク	平均 [ms]	最大 [ms]	最小 [ms]
HCHO	10.59	35.15	0.07
C ₄ H ₆	167.80	312.54	74.84

表 3 反応予測シナリオに対する結果 ($K = 100$)

手法	化学反応ネットワーク	平均 [ms]	最大 [ms]	最小 [ms]
Yen	HCHO	10.48	43.56	2.68
	C ₄ H ₆	126.81	248.42	46.50
BFS	HCHO	0.44	1.22	0.20
	C ₄ H ₆	0.51	1.25	0.27

表 4 逆合成解析シナリオに対する結果 ($K = 100$)

手法	化学反応ネットワーク	平均 [ms]	最大 [ms]	最小 [ms]
Yen	HCHO	9.13	35.04	3.15
	C ₄ H ₆	147.25	261.21	74.53
BFS	HCHO	0.33	1.79	0.14
	C ₄ H ₆	0.47	1.22	0.18

5. おわりに

複雑な化学反応ネットワークに対する最適な反応経路の探索において, 本稿では, 反応経路に対しランク付けを行い, 上位一部を最適反応経路候補とする絞り込みを行い, その上で, 重み付き有向グラフへ変換し, コスト関数を導入することで, K 最短路問題へ帰着させる手法を提案した。

続いて, 実際の反応経路探索に関する問題をもとに, 実際に最適反応経路候補を列挙する手法を, 既存の K 最短路問題に対するアルゴリズムをベースに紹介した。

最後に計算機実験により, 既存の化学反応ネットワークに対し, 今回の手法を用いることで, 数 10 ミリ秒から数 100 ミリ秒と既存の手法と比較し, 高速に最適反応経路の候補を列挙できることを示した。また, 既存の手法では列挙に失敗したネットワークに対して列挙に成功した。

今後は, 最適な反応経路の候補を絞り込むためのよりよい基準の選定と, それに対応する候補列挙の手法を検討したい。また, K 最短路の列挙では, 上位経路が一つだけが異なる非常に似た経路となることが多い。従って, 何らかの化学的な基準により, 似ている反応経路をクラスタリングすることで, 異なる経路を絞り込みつつ, より幅広い種類の最適な反応経路の候補を列挙するという手法についても検討したい。

謝辞 今回, 実際の化学反応ネットワークのデータをいただいた, 北海道大学の前田理教授, 原淵祐助教, 住谷陽輔博士, 杉山佳奈美氏にお礼申し上げます。

参考文献

- [1] Ohno, K. and Maeda, S.: A scaled hypersphere search method for the topography of reaction pathways on the potential energy surface, *Chemical physics letters*, Vol. 384, No. 4-6, pp. 277-282 (2004).
- [2] Maeda, S. and Morokuma, K.: Communications: A systematic method for locating transition structures of A+B → X type reactions, *The Journal of Chemical Physics*, Vol. 132, No. 24, p. 241102 (2010).
- [3] North, S. W.: Atmospheric photochemistry: Roaming in the dark, *Nature Chemistry*, Vol. 3, pp. 504-505 (2011).
- [4] Grubb, M. P., Warter, M. L., Xiao, H., Maeda, S., Morokuma, K. and North, S. W.: No Straight Path: Roaming in Both Ground- and Excited-State Photolytic Channels of NO₃ → NO + O₂, *Science*, Vol. 335, No. 6072, pp. 1075-1078 (2012).
- [5] Maeda, S. and Ohno, K.: No activation barrier synthetic route of glycine from simple molecules (NH₃, CH₂, and CO₂) via carboxylation of ammonium ylide: a theoretical study by the scaled hypersphere search method, *Chemical physics letters*, Vol. 398, No. 1-3, pp. 240-244 (2004).
- [6] Mita, T., Sugawara, M. and Sato, Y.: One-Pot Synthesis of α -Amino Acids through Carboxylation of Ammonium Ylides with CO₂ Followed by Alkyl Migration, *The Journal of organic chemistry*, Vol. 81, No. 12, pp. 5236-5243 (2016).
- [7] 鈴木浩史, 中野裕太, 住谷陽輔, 湊真一, 前田理: 化学反応経路ネットワークにおける ZDD を用いたエネルギー制限付き経路列挙, 情報処理学会, 研究報告アルゴリズム (AL), Vol. 2018, No. 7, pp. 1-6 (2018).
- [8] 佐藤寛子, 小田朋宏, 中小路久美代, 宇野毅明, 田中宏明, 岩田寛, 大野公一: 「埋蔵分子」発掘プロジェクト—化学反応経路マップのインタラクティブ可視化に向けて, 情報処理学会, インタラクシオン (2014).
- [9] Eppstein, D.: Finding the k Shortest Paths, *SIAM Journal on Computing*, Vol. 28, No. 2, pp. 652-673 (1998).
- [10] Y. Yen, J.: Finding the K Shortest Loopless Paths in a Network, *Management Science*, Vol. 17, pp. 712-716 (1970).