

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 11 日現在

機関番号：32644

研究種目：挑戦的萌芽研究

研究期間：2011～2013

課題番号：23650009

研究課題名(和文) 統計に起因する諸問題への列挙算法の開発

研究課題名(英文) Development of the enumeration algorithm for the problems on statistics

研究代表者

松井 泰子 (MATSUI, YASUKO)

東海大学・理学部・准教授

研究者番号：10264582

交付決定額(研究期間全体)：(直接経費) 1,900,000円、(間接経費) 570,000円

研究成果の概要(和文)：列挙算法は、VLSI・ビッグデータ・遺伝子情報・合成化合物などの広範な分野で適用できる情報データ解析ツールの一つである。本研究では、統計に関する問題に対し、列挙算法を開発した。具体的には、ある種のグラフ上でのコスト付き頂点彩色問題の最適解を高速に列挙する算法である。結果は、国際会議で発表し、未発表の結果を含め、論文にまとめている。

研究成果の概要(英文)：Enumeration algorithm is one of tools for analyzing of VLSI, big data, gene information, synthetic compounds etc.

In this research, we developed some enumeration algorithms for the problems on statistics. In particular, we proposed speedy enumeration algorithm for the optimal cost vertex coloring on some graphs, and had talks at some international conferences. we are summarizing the results in the paper.

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：アルゴリズム 列挙

1. 研究開始当初の背景

(1)与えられた離散構造から、ある条件を満たす部分構造を重複無く全て列挙する問題を**列挙問題**と呼び、列挙問題を解く算法を**列挙算法**と呼ぶ。一般に、列挙される部分構造の総数は指数個で、列挙算法実行中では、それらを全てメモリに保持せず、部分構造が得られる都度出力し、メモリから削除している。そのため、部分構造を重複無く全て列挙する列挙解法構築の際には、様々な工夫が必要となる。一般に列挙算法開発は容易ではなく、部分構造同士の隣接関係の解明や、算法に適したデータ構造の提案が必須である。

(2)最近、スタンフォード大学のクヌース教授が著書にて列挙算法の研究を詳細に紹介したことも契機となり、列挙算法開発は話題性の高いアルゴリズム研究分野とみなされ、大勢の研究者が参入しつつある。

列挙算法は、VLSI・ビッグデータ・遺伝子情報・合成化合物などの解析にも適用され、応用面に対する有用性も注目を浴びている。また、大容量のメモリの低廉化により、列挙された部分構造を、特殊なデータ構造を用いて全て保持する研究も注目されている。

2. 研究の目的

(1)本研究の目的は、既存の列挙算法の枠組みとは異なるアプローチによる列挙算法の提案である。ただし、扱う列挙問題は統計に起因する諸問題に限定する。

既存の列挙算法の多くは、報告者の知る限り、理論的な興味によるものが主で、応用分野へのアプローチは、学習理論やデータマイニング、遺伝子情報解析等の、比較的新しい分野を扱う事が多い。

(2)統計分野では、実験データに対する、精度の高い検定のために、ある条件を満たす部分構造の“列挙・数え上げ・サンプリング”等が必須であるが、報告者の知る限り統計分野の列挙問題に対する算法開発は少ない。

もし、統計分野の列挙問題に対し、有用な列挙算法が開発されれば、医療統計を中心とする様々な統計データ解析に対する強力なツールとして、列挙算法が活用される。

以上のモチベーションから、報告者は統計分野の列挙問題に対する算法開発を試みた。

3. 研究の方法

(1)統計分野において、高速な列挙算法が望まれる列挙問題をリストアップし、リストアップした列挙問題の解空間がなす多面体における、解同士の隣接構造の特徴を解明する。  
 (2)リストアップした列挙問題に対し、列挙算法の適用の可能性を探る。一般に、列挙算法は、以下の項目を全て満たす必要がある。

列挙問題の最初の解を一つ、多項式時

間で求める。

解を得るたびに、他に解が存在するならば、それを求め、解が存在しないならば算法を終了する。

解を得るたびに出力し、メモリ上に保持しない。

上記 ~ を実現し、重複無く全て解を列挙出来ることが、列挙算法の要件である。これを満たす列挙算法構築の際には、既存の算法を前処理に組み込む事が多い。そこで、問題周辺分野の知識を広げつつ、理論的に列挙算法の正当性を確認しながら構築した。

(3)(2)で構築した列挙算法の高速化について内外の研究者と議論した。国内で年2回開催される、群馬大主催の列挙セミナーにて、他大学の研究者と列挙算法の高速化を実現するデータ構造について議論した。さらに国際会議で海外の研究者と意見交換し、研究の方向性を探った。

4. 研究成果

(1)統計に起因する列挙問題として、グラフ

木の最小コスト点彩色

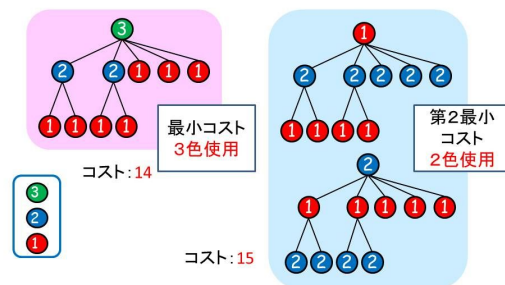


図1. 木の最小コスト点彩色の例。例の木では最適解は色数が多い方が、よりコストが小さくなる。

の木に対する最小コスト点彩色問題に着目した。グラフの木は、統計のグラフィカルモデルでは重要な意味を持つ離散構造であり、対象グラフを木に限定しても、列挙算法構築は意義がある。

一般に、グラフの頂点を、隣接する色が異なるように色数最小で塗り分ける「点彩色問題」は、NP-困難問題と呼ばれる、効率良く解けないと信じられる問題の一つである。さらに、色に重みが割り当てられ、頂点の重みの総和最小の頂点彩色を求める「最小コスト点彩色問題」は、「点彩色問題」同様、求解が容易ではないが、木では効率良く解けることが既知である。

しかし、木の形状によって、最適解に要する色数が異なるため、最適解を求めるには工夫が必要となる(図1)。そこで、木に対する「最小コスト点彩色問題」の最適解が一意的でない場合、それらを重複無く全て求める列

## 列挙アルゴリズム

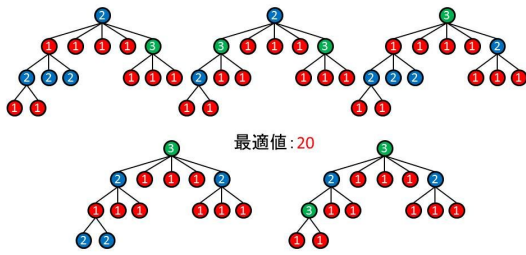


図2. 列挙算法で求めた全最適解. この例では, コスト 20 の相異なる点彩色が 5 つ存在する.

拳算法の構築を試み, 提案した.

(2) 報告者が提案した列挙算法は, 既存の算法を前処理の段階で用いて, 最初に一つ最適解を求めている. ただし, 既存の算法では最適解を求める際, 必要最低限の情報を保持し

## 一意性の特徴付け

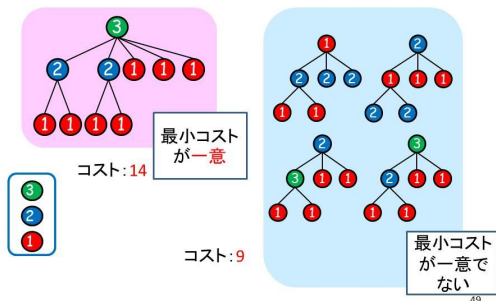


図3. 左の木は, 最適解が一意で, 右の木には最適解が一意でない例である. 右の木には最適解が 4 つ存在する.

つつ算法を実行するが, 報告者が提案する列挙算法に組み込むために, 実行中の情報を保持するよう変更している.

提案した列挙算法は, 複雑なデータ構造を用いることなく, 最適解を列挙する. そのため算法が要する記憶容量は多項式で抑えられる. また, 最適解を一つ当たり多項式時間で列挙する, 多項式時間遅延の高速算法である. 列挙された最適解の例を図2に示す.

(3) 報告者は, 列挙算法を構築する際, 木の形状によって, 「最小コスト点彩色問題」の最適解が要する色数が異なることに気付いた. 図3に最適解が一意な木と, 最適解が複数存在する木の例を示す.

本研究では, 最適解が一意である木の形状の特徴付けや, 最適解が与えられた色数を必ず用いる木の構成方法を提案した.

(4) 本研究の本研究の最終段階では, 研究初期に構築した列挙算法を拡張し, 対象を木のみならず, 1 - 木に対しても, 「最小コスト点彩色問題」の最適解の列挙を可能とした. この結果から, 対象とするグラフを木幅で分類し, 木幅によって提案した列挙解法が適用できるか否かが判別可能となった.

木幅でグラフを分類したアルゴリズムの提案は, アルゴリズムの理論分野では主流になっており, 本研究の成果は, 最新の理論的な結果に結び付いた. 得られた結果は, 未発表のものを含んでおり, 現在, 複数の論文にまとめて投稿準備をしている.

## 5. 主な発表論文等

[雑誌論文](計 1 件)

Masaki Yamamoto, Shuji Kijima, Yasuko Matsui, Polynomial-Time Perfect Q-Ising with a Vertex-Independent Noise, "Journal of Combinatorial Optimization, 査読有り, Volume 22, No. 3, 2011, 392-408.  
DOI: 10.1007/s10878-010-9309-7

[学会発表](計 8 件)

Yuma Tanaka, Atsuko Ikegami, Yasuko Matsui, Katsuki Fujisawa, Yuichiro Yasui, Fast counting of the number of primitive sorting network on  $n$  elements, The 25th Workshop on Topological Graph Theory in Yokohama, 2013 年 11 月 22 日, 横浜国立大学 (神奈川県)

Yasuko Matsui, Kento Kizaka, An enumeration algorithm for the optimal cost vertex colorings for trees, 2013 年 9 月 24 日, 日本数学会 2013 年度秋季総合分科会, 愛媛大学 (愛媛県)

Yasuko Matsui, Kento Kizaka, 木の最小コスト点彩色の列挙と一意性, 情報処理学会第 143 回アルゴリズム研究会, 2013 年 3 月 1 日, 飯坂温泉伊勢谷 (福島県)

Kento Kizaka, Yasuko Matsui, Enumeration and uniqueness for optimal cost vertex colorings of trees, 4th Pacific Workshop on Discrete Mathematics, 2012 年 11 月 29 日, 東海大学パシフィックセンター (Hawaii)

松井泰子, 選挙区割りに付随した列挙問題について, ERATO 湊離散構造処理系プロジェクト「2011 年度 秋のワークショップ」, 2011 年 11 月 1 日, 支笏湖丸駒温泉 (北海道)

Yasuko Matsui, Kento Kizaka, An enumeration algorithm for the optimal cost vertex-colorings for trees, 21th International Symposium on Mathematical Programming(ISMP 2012), 2012年8月21日, TU Berlin (Berlin)

川村一宏, 八森正泰, 松井泰子, Independence complexes of multipartite graphs, posets and regular cell complexes, 研究集会「変換群の幾何と組合せ論」, 2011年6月13日, 京都大学数理解析研究所(京都府)

松井泰子, 横棒数最小の円筒あみだくじの列挙について, ERATO 湊離散構造処理系プロジェクト「2011年度 初夏のワークショップ」, 2011年6月11日, 北海道大学(北海道)

〔図書〕(計 2件)

松井 泰子他, 朝倉出版, ものづくりに役立つ経営工学の事典-180 の知識-, 2014, 383 (310-311)  
安藤 清, 土屋 守正, 松井 泰子共著, 森北出版, 2013, 160 (1-160)

6. 研究組織

(1) 研究代表者

松井 泰子 (MATSUI, Yasuko)  
東海大学・理学部・准教授  
研究者番号: 10264582