

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 3 日現在

機関番号：13302

研究種目：基盤研究(B)

研究期間：2011～2014

課題番号：23300001

研究課題名(和文) 省メモリ計算モデル上でのアルゴリズム設計技法の開発

研究課題名(英文) Development of Algorithmic Paradigms on Memory-Constrained Computation

研究代表者

浅野 哲夫 (Asano, Tetsuo)

北陸先端科学技術大学院大学・学長

研究者番号：90113133

交付決定額(研究期間全体)：(直接経費) 14,000,000円

研究成果の概要(和文)：近年の計算機環境は著しい変化を遂げている。計算機の性能が上がるにつれ、より大きな規模の問題が解けるようになり、いわゆるビッグデータ処理が注目を集めているところである。このように問題のサイズは大きくなる一方であるが、データサイズが大きすぎて主記憶に収まりきらないという問題がある。そのために省メモリのアルゴリズムが要求されるが、この分野の研究は始まったばかりである。本研究では、計算幾何学に焦点をあてつつ、グラフ理論の諸問題に対する省メモリアルゴリズム、2値画像処理に対する効率の良いアルゴリズムを開発した。

研究成果の概要(英文)：With recent progress in computer environment, as the performance of computers is improved, people tend to be interested in larger-sized problems such as big-data processing. The problem size is continuously growing, we have a problem of too large memory size to be included in computer main memory. Because of this reason memory-constrained algorithms are requested in many ways, but research on small-memory algorithms has just begun. In this research we focus on computational geometry and develop memory-efficient algorithms, but we also developed efficient algorithms for graph problems and binary image processing problems.

研究分野：計算幾何学

キーワード：アルゴリズム 省メモリアルゴリズム 計算幾何学 グラフ理論 画像処理

1. 研究開始当初の背景

近年の計算機環境は著しい変化を遂げている。計算機の性能が上がるにつれ、より大きな規模の問題が解けるようになり、いわゆるビッグデータ処理が注目を集めているところである。このように問題のサイズは大きくなる一方であるが、記憶領域に関しては厳しい状況が続いている。つまり、データサイズが大きすぎて主記憶に収まりきらないのである。そのために省メモリのアルゴリズムが要求されるが、この分野の研究は始まったばかりであり、有効な方法が見出されていない場合が多い。本研究では、計算幾何学に焦点をあてつつ、それだけに拘らずに様々な問題に対して省メモリのアルゴリズムを開発した。大規模なグラフにサイクルが含まれるかどうかを判定する問題は非常に基本的な問題であるが、これを定数作業領域だけで解決する方法は古くから知られていた。その他にも定数作業領域で解ける問題については対数領域アルゴリズムの名前の下によく研究されてきた。本研究で目指すのは、もう少し大きな作業領域を用いて解ける問題である。すなわち、問題のサイズの平方根程度のメモリだけを用いて解くアルゴリズムの開発である。たとえば、1テラバイトのサイズの問題を1ギガ程度のメモリだけを用いて効率よく解こうとするものである。一般的には利用可能なメモリのサイズに応じてアルゴリズムの速度が変化するようなアルゴリズムを開発することが本研究の主テーマであった。

2. 研究の目的

本研究では省メモリのアルゴリズム開発を行うとともに、その理論的な限界についても考察する。従来から作業領域のサイズと計算時間の関係について様々な角度から研究が行われてきた。外部記憶装置とのデータ転送回数に注目した効率のよいアルゴリズム設計法や、キャッシュのサイズを知らなくてもキャッシュを効果的に利用できるアルゴリズムなどが研究されてきた。また、ネットワーク時代のオンラインデータ処理を対象としたストリーム・アルゴリズムの研究も盛んに行われている。このように、メモリと計算時間の関係は常にアルゴリズム研究の中心的な関心事であった。本研究では、現在の計算環境を反映した新たな省メモリ計算モデル上で、効率の良いアルゴリズムを開発するための一般的なアルゴリズム設計技法の開発を行う。特に、グラフに関する問題に対する省メモリ・アルゴリズムの開発に取り組む。

本研究の目標をまとめると次のようになる。

(1). 省メモリ・アルゴリズム設計技法の開発と個別問題への応用。

(2). 画像関連および計算幾何学の諸問題に対する省メモリ・アルゴリズムの開発。

(3). メモリ量に敏感な、すなわち、メモリ量に比例して高速になるアルゴリズムの設計原理の開発。

3. 研究の方法

研究目的で挙げた3つの視点から研究方法を述べる。

(1) 省メモリ・アルゴリズム設計技法の開発と個別問題への応用

様々な個別問題に対するアルゴリズムを開発することも重要であるが、それ以上に多くの問題に適用できる一般的なアルゴリズム開発技法を開発することの方がより重要であるとの視点にたって研究を行った。省メモリのアルゴリズムを開発するにあたって、入力をあらかじめ定められたサイズの部分集合に分割することが重要になる。最も簡単な方法はインデックスによる分割である。次はインスタンスの間の順序関係を利用した分割である。さらに、インスタンスの性質に基づく分割である。それぞれの分割を省メモリの環境で能率よく実行するための方法を必要なデータ構造とともに開発した。

(2) 画像関連および計算幾何学の諸問題に対する省メモリ・アルゴリズムの開発

画像に関しては入力データが2次元的に配置された画素であるのが最大の特長である。この特徴を最大限に活かしてアルゴリズムを構成することが必要になる。具体的には画像の3行程度の部分だけを管理することにより、接続関係をうまく保ちながら順にスキャンしていく方法を考案した。

(3) メモリ量に敏感な、すなわち、メモリ量に比例して高速になるアルゴリズムの設計原理の開発

使えるメモリを最大限に活かすという方針でもアルゴリズム開発を行った。そのために、 s というパラメータを想定し、 $O(s)$ のサイズのメモリだけを用いて最も効率よく問題を解く。そのためにはデータ構造の工夫が重要である。

4. 研究成果

(1) 省メモリ・アルゴリズム設計技法の開発と個別問題への応用

与えられた入力をパラメータ s に応じて小さな部分集合に分割しながら答えを導くという考え方を応用したアルゴリズムの例として、線分交差判定と列挙のアルゴリズムを挙げることができる。問題は、全部で n 本の線分が平面上に与えられたとき、その間の交差があるかどうかを判定する問題と、交点をすべて列挙するという問題である。それぞれの分割に応じて効率の良いアルゴリズムが構成できることを示した。より具体的には、 n 本の線分が平面上に与えられたとき、 $O(s)$

のメモリだけを用いて交差線分対が存在するかどうかを $O(n^2/s \log n)$ 時間で判定するアルゴリズムを得ることに成功した。

交差する線分対をすべて列挙することはより難しく、一般の線分を対象にした場合には問題サイズの2乗に近い時間がかかるが、2次元の点位置決定のアルゴリズムとデータ構造を上手につかうことにより2乗より少ない時間で交差を列挙するアルゴリズムを開発することができた。

これとは別に、単純な多角形の内部に指定された2点間の最短経路を効率よく求める問題についても結果を得た。具体的には、 n 頂点からなる多角形の場合、 $O(ns \log n)$ 時間の前処理で適当なデータ構造を構築することにより、 $O(s \log s)$ の時間で最短経路を求めるアルゴリズムを開発した。

さらに、計算幾何学の代表的な問題であるポロノイ図作成と平面上の点集合に対するユークリッド最小木の構成問題に対しても定数作業領域のアルゴリズムを与えることに成功した。これらのアルゴリズムは $O(s)$ のメモリを使ったアルゴリズムの構成にも役立つと思われる。

(2) 画像関連および計算幾何学の諸問題に対する省メモリ・アルゴリズムの開発

画像関連の問題に関しては、特に2値画像を対象とした効率の良いアルゴリズムを開発した。開発したのは、(A)連結成分の個数を数えるアルゴリズム、(B)最大の連結成分を求めるアルゴリズム、(C)連結成分ごとに異なるラベルを付与するアルゴリズムの3つである。連結成分の個数を数える方法はグラフのサイクルを求めるアルゴリズムの応用として捉えることができる。したがって、定数作業領域でも $O(n \log n)$ 時間、 $O(n)$ の作業領域を使えるなら $O(n)$ 時間で数えることができるという結果を得た。

さらに、最大の連結成分を求める問題については、連結成分の管理を動的に行うデータ構造を工夫することにより、 $O(n)$ の時間と $O(n)$ の領域で実現することに成功した。更に、最も難しいラベル付の問題についても、巧妙にラベルの生き死にを管理する方法により $O(n \log n)$ 時間で実行するアルゴリズムを得た。実は、 $O(n)$ の作業領域を用いても、 $O(n \log n)$ 時間でラベル付けを実行することは難しい中で、作業領域を $O(n)$ にまで下げたという点は評価されるべきものであろう。

以上の研究では入力画像を読み出し専用のメモリに入っていることを仮定し、それに作業領域を用いて作業を進めていた。一方、画像メモリに直接書き込みができる場合を検討した研究も行った。具体的には、入力画像の2値画像が与えられたとき、画像メモリとは別に $O(s)$ のメモリだけを用いて $O(s)$ のサイズの小さな連結領域を消す(画像の値を1から0に変える)という作業を $O(s \log s)$ の時間で実行するアルゴリズムを開発した。

(3) メモリ量に敏感な、すなわち、メモリ量に比例して高速になるアルゴリズムの設計原理の開発

深さ優先探索技法はグラフ理論における主要な技法の一つであり、この方法を応用することで様々な問題を解くことができる。しかし、深さ優先探索を少ないメモリだけで実現できるかどうかは分かっていなかった。本研究では、問題サイズと同じビット数を用いる方法を開発した。まずは、問題のサイズの2倍のビット数を用いて深さ優先探索を実行するアルゴリズムを開発した。これは4つの色を使うことに対応する。各頂点の色を変えることにより、既に探索したかどうかをうまく管理することでアルゴリズムを実現することに成功した。さらに、再帰を用いることによって色数を3色、さらに2色に減らすことで初期の目的を達成することができた。

別の研究では、グラフに関する最も基本的な問題の一つである到達可能性判定問題に取り組んだ。無向グラフに関しては定数作業領域だけで到達可能性を判定するアルゴリズムが既に知られているが、有向グラフに対しては効率の良いアルゴリズムが知られていなかった。本研究では、平面グラフに限定した場合、グラフのサイズの平方根程度のメモリだけを使って到達度判定を行うアルゴリズムの開発に成功した。平面グラフに対してアルゴリズムを再帰的に適用するために、平面グラフを、ほぼ同じサイズをもつ複数個の部分に分割する方法を確立しなければならないが、既に知られている平面分割定理を有向グラフ用に拡張することで再帰アルゴリズムを得ることに成功した。さらに、計算時間を問題サイズの多項式で抑えるために、普遍系列の概念を導入した。これも再帰の考え方に基づくものであるが、再帰を上手に使うことによって多項式で抑える工夫をしたものである。

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9件)

[1] T. Asano, K. Buchin, M. Buchin, M. Korman, W. Mulzer, G. Rote, A. Schulz: Memory-constrained algorithms for simple polygons. *Computational Geometry*, 46(8): 959-969 (2013) (査読あり)

[2] T. Asano, R. Kumar: A Small-Space Algorithm for Removing Small Connected Components from a Binary Image. *IEICE Transactions* 96-A(6): 1044-1050 (2013) (査読あり)

[3] M. Konagaya, T. Asano: Reporting All Segment Intersections Using an Arbitrary Sized Work Space. *IEICE Transactions*

96-A(6): 1066-1071 (2013) (査読あり)
[4] T. Asano, J. Jansson, K. Sadakane, R. Uehara, G. Valiente: Faster computation of the Robinson-Foulds distance between phylogenetic networks. Inf. Sci. 197: 77-90 (2012) (査読あり)
[5] T. Asano, E. D. Demaine, M. L. Demaine, R. Uehara: NP-completeness of generalized Kaboodle. JIP 20(3): 713-718 (2012) (査読あり)
[6] T. Asano: In-place Algorithm for Erasing a Connected Component in a Binary Image. Theory of Computing System, 50(1): 111-123 (2012) (査読あり)
[7] T. Asano, W. Mulzer, Y. Wang: Constant-Work-Space Algorithms for Shortest Paths in Trees and Simple Polygons. J. Graph Algorithms Appl. 15(5): 569-586 (2011) (査読あり)
[8] T. Asano, W. Mulzer, G. Rote, Y. Wang: Constant-Work-Space Algorithms for Geometric Problems. Journal of Computational Geometry, 2(1): 46-68 (2011) (査読あり)
[9] E. Chiba, T. Asano, T. Miura, N. Katoh, I. Mitsuka: Collision Probability in an In-Line Machines Model. Transactions on Computational Science 13: 1-12 (2011) (査読あり)

〔学会発表〕(計8件)

[1] T. Asano, T. Izumi, M. Kiyomi, M. Konagaya, H. Ono, Y. Otachi, P. Schweitzer, J. Tarui, R. Uehara: Depth-First Search Using $O(n)$ Bits. ISAAC 2014: 553-564, Jeonju, Korea, December 15, 2014. (査読あり)
http://link.springer.com/chapter/10.1007/978-3-319-13075-0_44
[2] T. Asano, D.G. Kirkpatrick, K. Nakagawa, O. Watanabe: $O(n)$ -Space and Polynomial-Time Algorithm for Planar Directed Graph Reachability. MFCS (2) 2014: 45-56, Budapest, Hungary, August 26, 2014. (査読あり)
<http://www.inf.u-szeged.hu/mfcs2014/>
[3] T. Asano, A. Elmasry, J. Katajainen: Priority Queues and Sorting for Read-Only Data. TAMC 2013: 32-41, May 20, 2013, Hong Kong (査読あり)
[4] T. Asano, D. G. Kirkpatrick: Time-Space Tradeoffs for All Nearest Larger Neighbors Problems. WADS 2013: 61-72, August 14, 2013, London, Canada. (査読あり)
[5] T. Asano, S. Bereg: A New Framework for Connected Components Labeling of Binary Images. IWCIA 2012: 90-102, November 28, 2012, Austin, USA. (査読あり)
[6] T. Asano, S. Bereg, L. Buzer: Small Work Space Algorithms for Some Basic

Problems on Binary Images. IWCIA 2012: 103-114, November 28, 2012, Austin, USA. (査読あり)

[7] T. Asano, K. Buchin, M. Buchin, M. Korman, W. Mulzer, G. Rote, A. Schulz: Memory-Constrained Algorithms for Simple Polygons. CoRR abs/1112.5904, March 19, 2012, Assisi, Italy (査読あり)

[8] T. Asano, B. Doerr: Memory-Constrained Algorithms for Shortest Path Problem. CCCG, August 10, 2011, Toronto, Canada. (査読あり)

〔図書〕(計0件)

〔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

6. 研究組織

(1) 研究代表者

浅野 哲夫 (ASANO TETSUO)
北陸先端科学技術大学院大学・学長
研究者番号: 90113133

(2) 研究分担者

上原 隆平 (UEHARA RYUHEI)
北陸先端科学技術大学院大学・情報科学研究科・教授
研究者番号: 00256471

大舘 陽太 (OTACHI YOTA)
北陸先端科学技術大学院大学・情報科学研究科・助教
研究者番号: 80610196