

## 科学研究費助成事業 研究成果報告書

平成 27 年 6 月 8 日現在

機関番号：17104

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24700025

研究課題名(和文)非正則なデータ構造上の非数値計算問題に対するスケルトン並列プログラミングの応用

研究課題名(英文)Applying Skeletal Parallel Programming to Non-numeric Computation on Unstructured Data

研究代表者

江本 健斗 (Emoto, Kento)

九州工業大学・大学院情報工学研究院・助教

研究者番号：00587470

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：道路ネットワークやソーシャルネットワークのようなデータから、特定の条件を満たす(最適な)部分を取り出すといった計算を高速に行うプログラムの構築を容易にするための、プログラミング手法の開発を行った。具体的には、既存の「生成・検査・集約プログラミング」のより一般的な構造への拡張、プログラムの変換の正しさを保証するための定理証明支援系による機械証明の枠組みの構築、条件付きの最短経路を例にした柔軟性と高速性に関する異なる既存手法の組合せ、を行った。

研究成果の概要(英文)：This research developed programming methods to ease making efficient programs dealing with large data such as road networks and social networks, in which, for example, programs will find (the best) portions that match with given conditions. The main results of this research are the extension of the Generate-Test-Aggregate programming paradigm to more general data structures, the development of framework for mechanical proof of correctness of the programming transformations via proof assistant software, and the combination of existing different methods for flexibility and performance improvements for shortest-path problems.

研究分野：計算機科学

キーワード：並列プログラミング スケルトン グラフ

## 1. 研究開始当初の背景

現在、計算機の多くがその計算性能の向上を並列処理に頼っている状況にある。さらに、そのアーキテクチャも、マルチコア CPU を持つ計算機、それらをネットワークで接続したクラスター、GPGPU (General Purpose computing on GPU) など、より複雑化している。一方、それらの環境を有効に利用するための並列プログラムの開発は、既存の逐次プログラミングの延長で行われており、一般のプログラマにとって複雑で困難なものである。また、個々の問題インスタンスごとに場当たり的に並列プログラムを開発するという状況は、開発者資源の浪費につながり社会の不利益となる。そのため、並列プログラムを容易にかつ系統的に開発するための手法が必要とされている。

非数値計算問題とは、演算の対象が単純な数値ではない計算問題である。特に、グラフやネットワークに代表される非正則なデータ構造上の組み合わせ最適化問題は、実用上重要な問題であり、近年の計算機性能の向上とともにその応用と並列化が重要視されつつある。しかしながら、既存の並列プログラミング手法の主な対象は、正則なデータ構造である配列や単純な数値を扱う数値計算である。よって、それらの手法は非正則なデータ構造上の非数値計算問題を解く並列プログラムの作成に効果的であるとは言えない。そのため、非正則なデータ構造上の非数値計算問題に対しても有効な、並列プログラムを容易にかつ系統的に開発するための手法の確立が強く望まれている。

## 2. 研究の目的

本研究では、非正則なデータ構造上の非数値計算問題に対する実用的なスケルトン並列プログラミングの構築を目的とし、特に次の2点について研究を進める。スケルトン並列プログラミング (Skeletal Parallel Programming) は、「並列スケルトン」と呼ばれる並列計算のための計算パターンを組み合わせてプログラムを作成する手法であり、ユーザは用意された並列スケルトンを組み合わせれば良く、並列性をそれほど意識することなく並列プログラムを作成できる枠組である。

(1) 非正則なデータ構造を扱う記述性の高い並列スケルトンの定式化: 既存のスケルトン並列プログラミングの研究では、グラフやネットワークに代表される非正則なデータ構造を操作するスケルトンの定式化がなされていない。これは、リストや配列等の正則なデータ構造と異なり、非正則なデータ構造に対してはどのような計算パターンをスケルトンの定式化に用いるべきであるかが非自明であることに起因する。一方、申請者は、先駆的なデータ並列言語 NESL や最新の並列言語 Fortress にある生成子

(generator) のアイデアを拡張し、生成子・検査子・集約子の3種類の計算要素に計算を分離することにより、正則なデータ構造上の非数値計算問題に対し柔軟な記述と効率的な実装とを与えられる並列プログラミングの枠組みを提案している。本研究では、この枠組みのアイデアを非正則なデータ構造上の非数値計算問題に応用する。

(2) 容易な記述と効率的な実装を提供する最適化機能を備えたプログラミング環境の実現: 非正則なデータ構造上の非数値計算問題に有効な、実用的スケルトン並列プログラミングに必要なとされる柔軟な記述性と高性能を達成する最適化能力を、プログラミング言語の観点から整理してライブラリとして実現する。近年、ライブラリレベルでの最適化機構実現に使用できる言語機能を備えたプログラミング言語が広まってきている。それらの能力を駆使することで、最適化能力を備えた並列プログラミングライブラリを実現する。

## 3. 研究の方法

理論に関する研究(1)と実現に関する研究(2)のそれぞれについて、複数の小項目に分けて研究を実施する。特に、理論に関する研究と実現に関する研究とをある程度並行して行うことで、研究が進まない状況できるだけ避けるようにする。

(1) 記述性の高い並列スケルトンとその最適化の定式化:(1-1)非正則データ構造用並列スケルトンの定式化、(1-2)最適化規則の開発、(1-3)より柔軟なプログラム構成手法の開発

(2) 最適化機能を備えたプログラミング環境の実現:(2-1)評価対象とする具体例の選定と比較対象の実装、(2-2)定式化に基づく並列スケルトン・最適化機構の実装、(2-3)記述性・最適化に関する評価実験

本研究は基本的には申請者一人で行う。ただし、特に理論的な定式化や設計の段階では他の研究者の意見が有益であるため、スケルトン並列プログラミングに関して活動してきた共同研究グループとの集中的な議論などを行う。また、実装したライブラリの評価実験などにはこの共同研究グループで所有するハードウェアも利用し、ライブラリの完成度を高める。

## 4. 研究成果

まず、理論的な観点から、最も正則なデータである列に於ける並列プログラミング手法「生成・検査・集約プログラミング」を、より一般的な構造である木に拡張した。この枠組では、(1)必要となる入力データの部分構造を生成子で全生成する、(2)興味のない構造を検査子で破棄する、(3)最終的に残った候補から集約子によって必要な情報を取り出す、という形で愚直なプログラム

の設計・記述を行い、言語システム側がその記述から効率的な並列プログラムを自動合成する。これにより、木構造上の高速な並列プログラムを愚直かつ柔軟なプログラム記述から自動合成できる枠組みが作られた。この枠組は、一般のグラフを木として捉える k 木や木分解の概念と合わせることで、グラフ上の高速な並列プログラムを柔軟かつ愚直な記述から自動合成できる枠組みの基礎となり得る。次に、一般のグラフ上の問題仕様記述からの木として捉える場合の問題仕様記述への自動変換手法の構築に取り組んだが、この正しさの証明を手作業で行うことが困難であることが明らかになり、変換の正しさを定理証明支援系による機械証明で保証するための研究を行った。この研究では、プログラムの構成要素となる準同型写像・効率的な実装と愚直な記述を繋ぐ半環の代数構造・柔軟な記述を与えるための半環構造の拡大・プログラムの変換の定式化に必要な各種データ構造について、その定理証明支援系における定式化・証明を行った。これらの定式化・証明は定理証明支援系の分野においても十分な貢献となった。また、定式化・証明を行ったプログラム変換を自動で行うライブラリを構築し、公開した。このライブラリは、既存の BSP モデル上の並列スケルトンを対象とした機械証明付き自動並列化ライブラリに、生成・検査・集約プログラミングによる自動効率化機構の証明付きライブラリを組み合わせたものになっている。前者は、定理証明支援系上に記述された逐次プログラムを定理証明支援系上で自動的に並列化し、さらにその証明付きのプログラムを既存のプログラミング言語上に抽出するというライブラリである。ただし、後者のライブラリとは異なり計算量の異なるプログラムの合成は行わない。両者のライブラリを合わせることで、定理証明支援系上の愚直な記述から既存のプログラミング言語上で動作する高速な並列プログラムの自動合成が達成された。この一連の変化が定理証明支援系により機械的に正しいと証明されていることが特徴である。最後に、グラフ上の最短路問題を対象に、柔軟な記述性を与えるプログラム変換技術と、高速な計算を行うためのデータ前処理技術との、異なる既存手法の組み合わせを試みた。これは、プログラム変換に基づく記述性の向上と、入力の前処理による並列スケルトン単体の性能の向上とを同時に達成し、より柔軟にかつより高速に問題を記述・計算出来るように枠組みを改善することを目的としたものである。結果として、これらの完全なる両立は無理であることが分かった。しかし、プログラム変換による記述性の一部を削ることはなるが既存の様々な問題記述を包含する十分な記述性を保ったまま、両者の恩恵に同時に与えることの出来る枠組みは構築できることが示された。

## 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 6 件)

L. Tung, K. Emoto and Z. Hu: Shortest Regular Category-Path Queries. GRACE-TR 2014-03, GRACE Center, National Institute of Informatics, August 2014, 12 pages. (査読無し)

[http://grace-center.jp/rsc\\_tr-html](http://grace-center.jp/rsc_tr-html)  
K. Emoto, K. Matsuzaki: An Automatic Fusion Mechanism for Variable-Length List Skeletons in SkeTo. International Journal of Parallel Programming (IJPP), Springer, 42(4): 546-563, 2014. (査読有り)

DOI: 10.1007/s10766-013-0263-8

Y. Liu, K. Emoto, Z. Hu: A Generate-Test-Aggregate Parallel Programming Library. Parallel Computing, 40(2): 116-135, 2014. (査読有り)

DOI: 10.1016/j.parco.2013.11.002

劉雨, 江本健斗, 松崎紀, 胡振江: Accumulative Computation on MapReduce. 情報処理学会論文誌: プログラミング, 7(1): 18-27, 2014. (査読有り)

<http://ci.nii.ac.jp/naid/110009658970>

K. Emoto, K. Matsuzaki: An Automatic Fusion Mechanism for Variable-Length List Skeletons in SkeTo. Technical Report METR 2013-04, 18 pages, Department of Mathematical Engineering and Information Physics, University of Tokyo, 2013. (査読無し)

<http://www.keisu.t.u-tokyo.ac.jp/research/techrep/2013.html>

K. Emoto, S. Fischer, Z. Hu: Filter-embedding Semiring Fusion for Programming with MapReduce. Formal Aspects of Computing, 24(4-6): 623-645, 2012. (査読有り)

DOI: 10.1007/s00165-012-0241-8

[学会発表](計 4 件)

K. Emoto, F. Loulergue, J. Tesson: A Verified Generate-Test-Aggregate Coq Library for Parallel Programs Extraction. In Proceedings of the 5th Conference on Interactive Theorem Proving (ITP 2014), Vienna, Austria, 14th-17th July, LNCS 8558, pp. 258-274, Springer, 2014. (査読有り)

DOI: 10.1007/978-3-319-08970-6\_17

Y. Liu, K. Emoto, Z. Hu: A Generate-Test-Aggregate Parallel Programming Library. In Proceedings of

the 2013 International Workshop on Programming Models and Applications for Multicores and Manycores (PMAM 2013), Shenzhen, China, 23rd February, pp. 71-81, ACM New York, 2013. (査読有り)

DOI: 10.1145/2442992.2443000

松崎公紀, 江本健斗: boost::proto を用いた融合変換機能付きライブラリの作成. 情報処理学会 第 93 回プログラミング研究発表会, 2 月国立情報学研究所, 2 月 28 日-3 月 1 日, 2013. (査読無)

K. Emoto, H. Imachi: Parallel Tree Reduction on Hadoop. In International Workshop on Practical Aspects of High-Level Parallel Programming, Omaha, USA, 4th-6th June, Procedia Computer Science 9, pp. 1827-1836, 2012. (査読有)

DOI: 10.1016/j.procs.2012.04.201

〔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

定理証明支援系 Coq 上の自動最適化機構付き

並列プログラミングライブラリ SyDPaCC

<http://traclifo.univ-orleans.fr/SyDPaCC/>

/

## 6. 研究組織

### (1) 研究代表者

江本 健斗 (EMOTO KENTO)

九州工業大学・大学院情報工学研究院・助教

研究者番号: 00587470